ABELIAN VARIETIES AND SHIMURA VARIETIES

YIHANG ZHU

Abstract. These notes are based on a topics course on Shimura varieties delivered by Yihang Zhu at the University of Maryland in Spring of 2022. Notes were taken by Steven Jin and edited by Y.Z.. If you have questions or corrections please contact Y.Z..

Contents

A description of the course

Shimura varieties play a vital role in the arithmetic aspects of the Langlands Program. This course will focus on the point of view that they are moduli spaces of abelian varieties with additional structures. The course consists of four parts.

(1) **Modular curves.** In this part, we first discuss the classical complex analytic construction of modular curves. Then we introduce the idea that modular curves are moduli spaces of elliptic curves, and give the scheme theoretic construction of them over $\mathbb{Z}[1/N]$.
(2) **Abelian varieties.** We discuss the fundamentals of abelian varieties, including important theorems about line bundles and their cohomology. We mainly follow [Mum08], but we simplify the exposition by admitting Grothendieck's existence theorem for Picard schemes.
(3) **Siegel modular varieties.** We establish the representability of the moduli functor of polarized abelian schemes with level structure over $\mathbb{Z}[1/N]$ using Geometric Invariant Theory methods, following [MFK94].
(4) **The Mazur–Ribet Theorem.** We discuss how Ribet [Rib90] uses the geometry of modular curves and Shimura curves to prove that Taniyama–Shimura–Weil implies Fermat's Last Theorem.

Prerequisites. Basic understanding of algebraic geometry (such as Hartshorne's textbook), and some familiarity with algebraic number theory and modular forms.

References for the course.

- **Elliptic curves and modular curves:** The two volumes by Silverman [Sil09, Sil94] are the standard introductory textbooks on elliptic curves. A friendly introduction to modular curves is given in [DS05]. Our discussion of modular curves over $\mathbb{C}$ is similar in spirit to the first four chapters of [Mil11]. For the scheme-theoretic point of view, see the book by Katz–Mazur [KM85], and the handouts in Brian Conrad's course [Cone], especially [Cona, Conc] .
- **Abelian varieties:** The classic [Mum08]. Also useful is the online book draft [EvdGM], as well as various online lecture notes for instance [Cond].
- **Moduli of abelian schemes:** We will mainly follow [MFK94]. A brief summary of the relevant content in [MFK94] is found in [GN06]. The latter also contains material more towards the point of view of Shimura varieties, which we will not cover.
- **The Mazur–Ribet Theorem:** The original sources are Ribet's two papers [Rib90, Rib94]. For expositions, see [Pra95, Oes88, Edi97]. We also highly recommend Saito's two volumes [Sai13, Sai14] for excellent explanation of many important concepts in the subject of Fermat's Last Theorem. For more information on the proof of Fermat's Last Theorem, see [CSS97, DDT97].

References for further study of Shimura varieties. The original theory is obviously due to works of Shimura. The perspective which is nowadays more mainstream was initiated in Deligne's two ariticles [Del71b, Del79]. See also Milne's notes [Mil17b] for a self-contained overview. In [Lan], you can see the examples of many Shimura varieties. For a survey of more recent research on Shimura varieties, see [HH20].

1. Lecture 1

In this lecture we provide a brief motivational overview.

1.1. **Modular curves.** Let $\mathbb{H} = \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$. There is a left action of $SL_2(\mathbb{R})$ on $\mathbb{H}$, given as follows: for $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ and $z \in \mathbb{H}$, we define

$$g.z := \frac{az+b}{cz+d}.$$

Note that this yields a surjection

$$\mathrm{SL}_2(\mathbb{R}) \twoheadrightarrow \mathrm{Aut}_{\mathrm{hol}}(\mathbb{H})$$

where $\mathrm{Aut}_{\mathrm{hol}}(\mathbb{H})$ is the holomorphic automorphism group of $\mathbb{H}$. We are interested in certain discrete subgroups of $\mathrm{SL}_2(\mathbb{R})$ whose actions will produce interesting quotients of $\mathbb{H}$.

Definition 1.1.1. We say that a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ is a **congruence subgroup** if it is a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma(N) \subset \Gamma$ with finite index for some $N \geq 1$, where

$$\Gamma(N) = \left\{ g \in \mathrm{SL}_2(\mathbb{Z}) \mid g \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N\} \right\}.$$

We will also usually assume that Γ is **small enough**, i.e., that $\Gamma \subset \Gamma(N)$ for some $N \geq 3$.

The group Γ acts freely and properly discontinuously on $\mathbb{H}$, and this implies that $\Gamma\backslash\mathbb{H}$ has a canonical complex manifold structure given by the complex structure on $\mathbb{H}$. Furthermore, the quotient map $\mathbb{H} \to \Gamma\backslash\mathbb{H}$ is the universal covering.

Definition 1.1.2. The complex manifold $\Gamma\backslash\mathbb{H}$ is called a **modular curve**.

Proposition 1.1.3. *A modular curve $\Gamma\backslash\mathbb{H}$ enjoys the following properties:*

(1) $\Gamma\backslash\mathbb{H}$ *has the canonical structure of an algebraic variety over $\mathbb{C}$, which is compatible with the complex manifold structure.*
(2) $\Gamma\backslash\mathbb{H}$ *is the moduli space of elliptic curves over $\mathbb{C}$ with "Γ-level structure".*
(3) *The moduli interpretation in (2) also makes sense over some number field E (depending on Γ; e.g., $E = \mathbb{Q}(\zeta_N)$ if $\Gamma = \Gamma(N)$). This moduli problem over E is then represented by a quasi-projective E-scheme whose base change to $\mathbb{C}$ recovers $\Gamma\backslash\mathbb{H}$ as a $\mathbb{C}$-scheme. We say that $\Gamma\backslash\mathbb{H}$ has a* ***model*** *over E.*
(4) *The model over E in (3) can be canonically characterized by using "special points" in $\Gamma\backslash\mathbb{H}$ and the Hecke action, without reference to the moduli problem.*

Remark.

- The fact in (1) is not obvious, as the complex manifold $\Gamma\backslash\mathbb{H}$ is not compact. Hence, one cannot appeal to the usual equivalence between smooth projective curves over $\mathbb{C}$ and compact Riemann surfaces. The key is that there is a canonical compactification of $\Gamma\backslash\mathbb{H}$ which is a compact Riemann surface (the Baily–Borel compactification).
- In (2), the space $\Gamma\backslash\mathbb{H}$ is indeed a fine moduli space as a consequence of the assumption that Γ is small enough. Also, the moduli interpretation in (2) holds both algebraically and analytically. That is to say, one can formulate a moduli functor either over the category of (finite-type) $\mathbb{C}$-schemes or over the category of complex manifolds. Then $\Gamma\backslash\mathbb{H}$ as a $\mathbb{C}$-scheme or a complex manifold respectively represents the functor.

- In (3), if we consider a certain disconnected variant of this construction, we can descend the field of definition to $\mathbb{Q}$.
- The fact in (4) provides the uniqueness of the canonical model, but not its existence.

1.2. **Higher dimensional generalizations.** To obtain higher dimensional generalizations of the modular curves, we replace SL_2 by a reductive group G over $\mathbb{Q}$. One key example to keep in mind is the **symplectic similitude group** $G = \mathrm{GSp}_{2g}$. Recall that this is defined as follows. Fix a $2g$-dimensional symplectic vector space V over $\mathbb{Q}$, that is, a vector space V over $\mathbb{Q}$ equipped with a symplectic bilinear form $\langle -, - \rangle$. Then for any $\mathbb{Q}$-algebra R,

$$\mathrm{GSp}_{2g}(R) = \mathrm{GSp}(V)(R) = \{g \in \mathrm{GL}(V \otimes R) \mid \exists \nu(g) \in R^\times, \forall v, w \in V, \langle gv, gw \rangle = \nu(g)\langle v, w \rangle\}.$$

More informally, we write

$$\mathrm{GSp}_{2g} = \{g \in \mathrm{GL}(V) \mid \exists \nu(g) \in \mathbb{G}_m, \forall v, w \in V, \langle gv, gw \rangle = \nu(g)\langle v, w \rangle\}.$$

Remark. Setting $g = 1$ gives $\mathrm{GSp}_{2g} = \mathrm{GL}_2$, not SL_2.

At the same time, we wish to generalize $\mathbb{H}$ to a (possibly disconnected) homogeneous space X under a left action of the real Lie group $G(\mathbb{R})$ which admits a complex structure (invariant under the $G(\mathbb{R})$-action). We also require a Hermitian structure on each tangent space (invariant under the $G(\mathbb{R})$-action). This implies that there is constant curvature, which we insist must be negative. We require that X cannot be too small by requiring that up to connected components X is isomorphic to the **symmetric space** of $G^{\mathrm{ad}}(\mathbb{R})$, i.e., the quotient of $G^{\mathrm{ad}}(\mathbb{R})$ by a maximal compact subgroup.

Remark. For $G = \mathrm{GL}_n$ with $n \geq 3$, there is no such X.

Fix such G and X as above. We wish to generalize the notion of a congruence subgroup to this setting.

Definition 1.2.1. We say that a subgroup $\Gamma \subset G(\mathbb{Q})$ is a **congruence subgroup** if Γ contains $K \cap G(\mathbb{Q})$ with finite index for some compact open subgroup K of $G(\mathbb{A}_f)$ where $\mathbb{A}_f$ is the finite adeles of $\mathbb{Q}$. We also assume that Γ is "small enough", or **neat**.

Remark. The (canonical) topology on $G(\mathbb{A}_f)$ can be described in the following elementary manner. Fix an injective $\mathbb{Q}$-homomorphism $\rho : G \to \mathrm{GL}_n$ for some n. For each integer $N \geq 1$, define

$$K_{\rho,N} := \{g \in G(\mathbb{A}_f) \mid \rho(g) \in \mathrm{GL}_n(\widehat{\mathbb{Z}}) \subset \mathrm{GL}_n(\mathbb{A}_f), \rho(g) \mapsto 1 \in \mathrm{GL}_n(\mathbb{Z}/N\mathbb{Z})\}.$$

For the fixed ρ and varying N, the subgroups $K_{\rho,N}$ of $G(\mathbb{A}_f)$ are all open compact, and they form a neighborhood basis of 1. Thus a subgroup $\Gamma \subset G(\mathbb{Q})$ is a congruence subgroup if and only if it contains with finite index the inverse image under ρ of some congruence subgroup of $\mathrm{GL}_n(\mathbb{Q})$ (the latter defined in the same way as in the SL_2 case).

As in the modular curve case, for a neat congruence subgroup $\Gamma \subset G(\mathbb{Q})$, the set $\Gamma\backslash X$ can be canonically equipped with the structure of a complex manifold. This manifold enjoys similar properties as before.

Proposition 1.2.2.

- *The complex manifold $\Gamma\backslash X$ can be given the structure of a quasi-projective complex variety (again thanks to the Baily–Borel compactification which is a projective complex variety containing $\Gamma\backslash X$ as a Zariski open).*
- *For some specific choices of (G, X) (called **of PEL type**), we enjoy analogues of (2),(3), and (4) from Proposition 1.1.3, where in (2), the moduli of elliptic curves becomes the moduli of abelian varieties equipped with some additional structures.*

1.3. **Adelic reformulation.** Fix a compact open subgroup $K \subset G(\mathbb{A}_f)$, which we also assume to be small enough, or **neat**.

Remark. We are omitting the definitions of neatness for now, but we note that the notion of neatness for compact open subgroups of $G(\mathbb{A}_f)$ is closely related to the neatness for congruence subgroups of $G(\mathbb{Q})$.

We define the set

$$\mathrm{Sh}_K = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K.$$

Here $G(\mathbb{Q})$ acts diagonally on the left of $X \times G(\mathbb{A}_f)$ as follows:

- $G(\mathbb{Q})$ acts on X via the action of $G(\mathbb{R})$ on X.
- $G(\mathbb{Q})$ acts on $G(\mathbb{A}_f)$ by left multiplication.

Here K acts on the right of $X \times G(\mathbb{A}_f)$ by right multiplication on the factor $G(\mathbb{A}_f)$.

Notice that *a priori* Sh_K is merely a set, and a rather unmanageable one at that. The following finiteness result assuages us.

Proposition 1.3.1 (Borel finiteness)**.** *The set $G(\mathbb{Q})\backslash G(\mathbb{A}_f)/K$ is finite.*

We now equip Sh_K with the structure of a smooth quasi-projective complex variety. Fix $g_1, \ldots, g_n \in G(\mathbb{A}_f)$ to be representatives of $G(\mathbb{Q})\backslash G(\mathbb{A}_f)/K$. For each i, let

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

where the intersection is taken inside $G(\mathbb{A}_f)$. As K is neat, we know the Γ_i are neat congruence subgroups of $G(\mathbb{Q})$.

Proposition 1.3.2. *The map $\coprod_{i=1}^n \Gamma_i\backslash X \to \mathrm{Sh}_K$ given by sending $x \in \Gamma_i\backslash X$ to the double coset of (x, g_i) is a bijection.*

We leave the proof as an exercise.

We use the above bijection to equip Sh_K with the structure of a smooth quasi-projective complex variety, i.e., the disjoint union of the $\Gamma_i\backslash X$. The resulting structure is independent of the choices of the representatives g_i.

Remark. The variety Sh_K is disconnected in general, for two reasons:

- X might be disconnected in general (although sometimes becomes connected after taking the quotient by Γ_i)
- We typically have $n \geq 2$.

2. Lecture 2

2.1. **Continuation of the overview.** We have indicated that we are interested in compact open subgroups $K \subset G(\mathbb{A}_f)$ that are neat. We show that there is a rich source of such subgroups.

Suppose $\rho : G \to \mathrm{GL}_n$ is a faithful algebraic representation over $\mathbb{Q}$. This induces a morphism of topological groups

$$\rho : G(\mathbb{A}_f) \to \mathrm{GL}_n(\mathbb{A}_f).$$

We have a subgroup

$$H := \{g \in \mathrm{GL}_n(\widehat{\mathbb{Z}}) \mid g \mapsto 1 \text{ inside } \mathrm{GL}_n(\mathbb{Z}/N\mathbb{Z})\} \subset GL_n(\mathbb{A}_f)$$

where $N \geq 1$. Then the $K_{\rho,N} := \rho^{-1}(H) \subset G(\mathbb{A}_f)$ are compact open subgroups that form a neighborhood basis of 1, where ρ is fixed and we vary N.

Proposition 2.1.1 (Criterion for Neatness)**.** *If $K \subset G(\mathbb{A}_f)$ is a compact open subgroup such that $K \subset K_{\rho,N}$ for some ρ and $N \geq 3$ as above, then K is neat.*

Let $K \subset G(\mathbb{A}_f)$ be a neat compact open subgroup. As in the previous lecture, we define the set Sh_K as follows:

$$\mathrm{Sh}_K = G(\mathbb{Q})\backslash X \times G(\mathbb{A}_f)/K \cong \coprod_{i=1}^{n} \Gamma_i \backslash X$$

where $\Gamma_i \subset G(\mathbb{Q})$ are congruence subgroups. The fact that K is neat implies that the Γ_i are also neat. This equips the set Sh_K with the structure of a complex manifold.

Of course, we also desire the structure of an algebraic variety over $\mathbb{C}$ on Sh_K. *A priori*, there may be 0 or ≥ 2 possible ways to give a compatible variety structure on Sh_K whose analytification is the underlying complex manifold. (If Sh_K happens to be compact, then there is at most one way.) Nonetheless, there is in fact a canonical variety structure, coming from the Baily–Borel compatification: we are able to embed the complex manifold Sh_K into a normal projective variety such that Sh_K is a Zariski open set in said variety. In particular, Sh_K has the canonical structure of a quasi-projective smooth variety. Moreover, this variety structure can be characterized by a universal property (and is in fact absolutely unique), to be discussed in the next lecture.

Next, the gist of Shimura varieties is that they can be (canonically) defined over number fields. For this, it is important to upgrade the pair (G, X) to a **Shimura datum** in the sense of Deligne. This amounts to the extra datum of a $G(\mathbb{R})$-equivariant injective[1] map

$$X \hookrightarrow \mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, G_{\mathbb{R}})$$

satisfying some axioms which we omit for now. Here, $\mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_m)$ is the **Deligne torus**, and $\mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, G_{\mathbb{R}})$ is the set of $\mathbb{R}$-algebraic group homomorphisms $\mathbb{S} \to G_{\mathbb{R}}$, on which $G(\mathbb{R})$ acts on by conjugation on $G_{\mathbb{R}}$.

The following theorem is a deep result due to the effort of numerous mathematicians.

Theorem 2.1.2 (Shimura, Deligne, Milne, Borovoi, et al.)**.** *Fix a Shimura datum*

$$(G, X, X \hookrightarrow \mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, G_{\mathbb{R}})).$$

*For each neat compact open subgroup $K \subset G(\mathbb{A}_f)$, Sh_K has a canonical model over a canonical number field E. Here E, called the **reflex field**, depends only on the Shimura datum, not on K.*

Example. We give an example of a Shimura datum. Let $G = \mathrm{GL}_2$ and $X = \mathbb{H}^+ \coprod \mathbb{H}^- = \mathbb{C} \setminus \mathbb{R}$. Define an $\mathbb{R}$-homomorphism $h : \mathbb{S} \to G_{\mathbb{R}}$ as follows. For any $\mathbb{R}$-algebra R, we have

$$\mathbb{S}(R) = (R \otimes_{\mathbb{R}} \mathbb{C})^{\times} = \{a \otimes 1 + b \otimes i \mid a, b \in R,\ a^2 + b^2 \in R^{\times}\}.$$

We define $\mathbb{S}(R) \to G(R) = \mathrm{GL}_2(R)$ by

$$a \otimes 1 + b \otimes i \longmapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Now, note that $\mathrm{GL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cdot \mathbb{R}^{\times} \cong X$ under the map $g \mapsto g \cdot i$. (Here $\mathbb{R}^{\times}$ embeds into $\mathrm{GL}_2(\mathbb{R})$ as the scalar matrices.) Then, we define

$$X \cong \mathrm{GL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cdot \mathbb{R}^{\times} \hookrightarrow \mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, G_{\mathbb{R}})$$

[1] For some purposes injectivity can be loosen to finite-to-one.

via $g \mapsto \mathrm{Int}(g) \circ h, \forall g \in \mathrm{GL}_2(\mathbb{R})$, where $\mathrm{Int}(g)$ is the inner automorphism $x \mapsto gxg^{-1}$ of $\mathrm{GL}_{2,\mathbb{R}}$. In this case, the reflex field $E = \mathbb{Q}$.

In this course, we will mainly consider Shimura data such that the resulting Shimura varieties are closely related to certain moduli spaces of abelian varieties (with additional structures). More precisely, these Shimura varieties are isomorphic, up to disjoint unions, to such moduli spaces. Focusing on the moduli point of view will enable us to obtain good models over Zariski open sets of $\mathrm{Spec}\, \mathcal{O}_E$, as opposed to just $\mathrm{Spec}\, E$. Further, this will allow us to study the geometry and arithmetic of the reductions of the models modulo primes of E.

There are a number of applications along these lines.

(1) The construction of Galois representations attached to certain automorphic forms, as predicted in the Langlands program: This began with Deligne in the 1960s-1970s, who attached 2-dimensional Galois representations to cuspidal eigenforms on GL_2 of weight ≥ 2. This led to his famous resolution of the Ramanujan conjecture in this setting. A closely related problem is the computation of the Hasse–Weil zeta functions of these Shimura varieties. The prototype for this work is the Eichler–Shimura Theorem in the cases of modular curves and Shimura curves.
(2) Congruences for modular forms: one critical exhibition of such an application is Ribet's theorem, which played an important role in the resolution of Fermat's Last Theorem.

Our first goals in this course, to be addressed in the next few lectures, are the following:

- Classify elliptic curves / abelian varieties with additional structure over $\mathbb{C}$ via modular curves in the one dimensional case and Siegel modular varieties in higher dimensions.
- Prove representability of the moduli functor of principally polarized abelian varieties (ppav) over $\mathbb{Z}$, following Mumford.

3. Lecture 3

3.1. **Characterization of the algebraic variety structure.** Let (G, X) be a pair as in a Shimura datum. For this subsection, we do not require the datum of $X \hookrightarrow \mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, G_{\mathbb{R}})$. Here we will use the classical language as opposed to the adelic language. For every neat congruence subgroup $\Gamma \subset G(\mathbb{Q})$, we recall that $\Gamma\backslash X$ is a complex manifold, with its structure inherited from the universal covering $X \to \Gamma\backslash X$. As before, the Bailey–Borel compactification provides an algebraic variety structure on $\Gamma\backslash X$ that is compatible with its complex manifold structure.

The following theorem gives a characterization of the algebraic variety structure in terms of the complex manifold structure.

Theorem 3.1.1 (Borel)**.** *For every smooth variety $S/\mathbb{C}$, every holomorphic map $S^{\mathrm{an}} \to \Gamma\backslash X$ is algebraic, with respect to the algebraic structure on $\Gamma\backslash X$ given by Baily–Borel. In particular, there is in fact a unique algebraic variety structure on the complex manifold $\Gamma\backslash X$.*[2]

[2]For the last statement, use the following fact which is an easy consequence of Zariski's Main Theorem: If k is an algebraically closed field of characteristic zero, then every bijective algebraic morphism between two (irreducible) varieties over k, with the target normal, is an isomorphism. See for instance [Mil17a, Prop. 8.60].

3.2. **Naive classification of elliptic curves over $\mathbb{C}$.** In this subsection, all algebraic varieties are over $\mathbb{C}$. Recall that an **elliptic curve** E is a complete group variety of dimension one. We denote the identity element by $O = O_E \in E(\mathbb{C}) = E$.

We now list some facts about elliptic curves which we will admit.

(1) The group structure on an elliptic curve is automatically commutative.
(2) If we have an algebraic morphism $f : E \to E'$ between elliptic curves such that $f(O_E) = O_{E'}$, then f is automatically a group homomorphism.
(3) By a **lattice** in $\mathbb{C}$ we mean a $\mathbb{Z}$-submodule of $\mathbb{C}$ generated by an $\mathbb{R}$-basis of $\mathbb{C}$. For each lattice $\Lambda \subset \mathbb{C}$, we equip the compact Riemann surface $\mathbb{C}/\Lambda$ with the group structure coming from addition on $\mathbb{C}$. Since there is an equivalence of categories between compact Riemann surfaces and smooth projective algebraic curves, we see that $(\mathbb{C}/\Lambda, +)$ is an elliptic curve.

   Remark. We can in fact write down an explicit embedding $\mathbb{C}/\Lambda \hookrightarrow \mathbb{P}^2_{\mathbb{C}}$ using the Weierstrass $\wp$ function and its derivative $\wp'$.

(4) Every elliptic curve arises as $\mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$. More precisely, given an elliptic curve E, there is a holomorphic group homorphism

$$\exp : \operatorname{Lie} E \longrightarrow E$$

   coming from Lie group theory. Here $\operatorname{Lie} E$ is a 1-dimensional complex vector space. (Note that exp is not algebraic.) Then $\ker(\exp)$ is a lattice in $\operatorname{Lie} E$ and

$$(\operatorname{Lie} E)/\ker(\exp) \xrightarrow{\sim} E$$

   is an isomorphism of elliptic curves. Notice also that (non-canonically) the left hand side is isomorphic to $\mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$.

   Remark. The map $\exp : \operatorname{Lie} E \to E$ is a universal covering. Hence we have the following canonical isomorphisms:

$$\ker(\exp) \cong \pi_1(E, O) \cong \mathbf{H}_1(E, \mathbb{Z}).$$

(5) Suppose E and E' are elliptic curves. We have

$$\operatorname{Hom}(E, E') \xrightarrow{\sim} \{f : \operatorname{Lie} E \to \operatorname{Lie} E' \mid f \text{ is } \mathbb{C}\text{-linear and } f(\mathbf{H}_1(E, \mathbb{Z})) \subset \mathbf{H}_1(E', \mathbb{Z})\}$$

   where the assignment is given by

$$F \longmapsto \mathrm{d}F|_{\operatorname{Lie} E}.$$

Combining the above facts, we have an equivalence of categories

$$((V, \Lambda), V \text{ a 1-dim'l } \mathbb{C}\text{-vector space and } \Lambda \subset V \text{ a } \mathbb{Z}\text{-lattice}) \xrightarrow{\sim} (\text{Elliptic curves}) \tag{3.1}$$

given by

$$(V, \Lambda) \longmapsto V/\Lambda,$$

with the inverse functor

$$E \longmapsto (\operatorname{Lie} E, \mathbf{H}_1(E, \mathbb{Z})).$$

Now we say two lattices Λ and Λ' in $\mathbb{C}$ are **homothetic** if there exists $\lambda \in \mathbb{C}^\times$ such that $\Lambda = \lambda\Lambda'$. From the equivalence of categories (3.1), we obtain a bijection

$$\{\text{Lattices inside } \mathbb{C}\}/\text{homothety} \xrightarrow{\sim} \{\text{Elliptic curves}\}/\text{isomorphism}.$$

We now make the set on the left hand side more explicit. Set

$$\mathbb{H}^\pm := \mathbb{H}^+ \bigsqcup \mathbb{H}^- = \mathbb{C} \setminus \mathbb{R}.$$

For a lattice Λ in $\mathbb{C}$, we pick a $\mathbb{Z}$-basis $w_1, w_2 \in \mathbb{C}^\times$ of Λ. Then $w_1/w_2 \in \mathbb{H}^\pm$. If we choose another $\mathbb{Z}$-basis $\{w_1', w_2'\}$ of Λ, then there is a (unique) matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z})$ such that $w_1 = aw_1' + bw_2'$ and $w_2 = cw_1' + dw_2'$. Thus

$$\frac{w_1}{w_2} = \frac{aw_1' + bw_2'}{cw_1' + dw_2'} = \frac{a\left(\frac{w_1'}{w_2'}\right) + b}{c\left(\frac{w_1'}{w_2'}\right) + d}.$$

In other words, $\frac{w_1}{w_2}$ and $\frac{w_1'}{w_2'}$ are related by the $\mathrm{GL}_2(\mathbb{Z})$-action on $\mathbb{H}^\pm$. Thus every lattice $\Lambda \subset \mathbb{C}$ has an invariant in $\mathrm{GL}_2(\mathbb{Z})\backslash\mathbb{H}^\pm$. This construction induces a bijection from the set of homothety classes of lattices in $\mathbb{C}$ to the set $\mathrm{GL}_2(\mathbb{Z})\backslash\mathbb{H}^\pm$. Moreover, $\mathrm{GL}_2(\mathbb{Z})\backslash\mathbb{H}^\pm \cong \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+$. Thus we have proved the following theorem.

Theorem 3.2.1. *There is a natural bijection*

$$\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+ \xrightarrow{\sim} \{\textit{Elliptic curves}\}/\textit{isomorphism}$$

sending the $\mathrm{GL}_2(\mathbb{Z})$*-orbit of* $\tau \in \mathbb{H}^+$ *to the isomorphism class of* $\mathbb{C}/\mathbb{Z} + \mathbb{Z}\tau$.

3.3. **Alternative point of view: Hodge structures.** In the previous subsection, we classified elliptic curves V/Λ up to isomorphism by morally fixing the complex vector space V and varying Λ. As an alternative point of view, we may fix an abstract $\mathbb{Z}$-module Λ, finite free of rank 2, and ask how we could vary the $\mathbb{C}$-structure. We elaborate this idea below.

As before, an elliptic curve is given by $E \cong (\mathrm{Lie}\, E)/\mathbf{H}_1(E, \mathbb{Z})$. Also, we have a canonical isomorphism of 2-dimensional $\mathbb{R}$-vector spaces:

$$\mathrm{Lie}\, E \cong \mathbf{H}_1(E, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{R}.$$

Notice of course that $\mathrm{Lie}\, E$ also has a complex structure. Thus in order to reconstruct E, we need the abstract $\mathbb{Z}$-module $\mathbf{H}_1(E, \mathbb{Z})$ together with a complex structure on the $\mathbb{R}$-vector space $\mathbf{H}_1(E, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{R}$. In general, to define a complex structure on an $\mathbb{R}$-vector space V, it suffices to define multiplication by i such that $i^2 = -1$. In other words, a complex structure on V is exactly an element $J \in \mathrm{End}_{\mathbb{R}}(V)$ such that $J^2 = -1$, and this element corresponds to scalar multiplication by i.

Definition 3.3.1. An **integral Hodge structure of elliptic type** is a pair (Λ, J) where Λ is a finite free $\mathbb{Z}$-module of rank 2 and $J \in \mathrm{End}_{\mathbb{R}}(V)$ is a complex structure on $\Lambda \otimes_{\mathbb{Z}} \mathbb{R}$.

As such, we can rewrite the equivalence of categories (3.1) as the following equivalence of categories:

$$(\text{Elliptic curves}) \xrightarrow{\sim} (\text{integral Hodge structures of elliptic type}).$$

In particular, we obtain a bijection on the level of isomorphism classes of both categories respectively. It is worth noting that there is a more general notion of a Hodge structure:

Definition 3.3.2. Fix a subring $R \subset \mathbb{R}$, usually taken to be $\mathbb{Z}, \mathbb{Q}$, or $\mathbb{R}$. An ***R*-Hodge structure** is a finitely generated R-module Λ together with a direct sum decomposition

$$\Lambda \otimes_R \mathbb{C} = \bigoplus_{p,q \in \mathbb{Z}} F^{p,q}$$

as $\mathbb{C}$-vector spaces such that

$$\overline{F^{p,q}} \cong F^{q,p}.$$

Here the bar denotes the $\mathbb{R}$-linear automorphism of $\Lambda \otimes_R \mathbb{C}$ (viewed as an $\mathbb{R}$-vector space) given by $x \otimes y \mapsto x \otimes \bar{y}, \forall x \in \Lambda, y \in \mathbb{C}$. In general, the **type** of a Hodge structure refers to the set of (p, q) such that $F^{p,q} \neq 0$. We shall refer to this decomposition as the **Hodge decomposition**. The **Hodge filtration** refers to

$$\mathrm{Fil}^i = \bigoplus_{\substack{p \geq i \\ \text{all } q}} F^{p,q}.$$

This is a decreasing filtration by $\mathbb{C}$-subspaces on $\Lambda \otimes_R \mathbb{C}$, i.e., we have $\mathrm{Fil}^i \supset \mathrm{Fil}^{i+1}$.

In light of this new language, we observe that there is an equivalence of categories:

$$\begin{aligned}&\text{(Integral Hodge structures of elliptic type)}\\ &\qquad\qquad \xrightarrow{\sim} (\mathbb{Z}\text{-Hodge structures free of rank 2 of type } \{(-1,0), (0,-1)\}).\end{aligned}$$

The assignment is as follows. Given an integral Hodge structure of elliptic type (Λ, J), we define the decomposition $\Lambda \otimes_{\mathbb{Z}} \mathbb{C} = F^{-1,0} \oplus F^{0,-1}$ by letting $F^{-1,0}$ and $F^{0,-1}$ be the eigenspaces for $\tilde{J}$ of eigenvalues i and $-i$ respectively. Here $\tilde{J}$ denotes the $\mathbb{C}$-linear extension of $J \in \mathrm{End}_{\mathbb{R}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R})$ to $\mathrm{End}_{\mathbb{C}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{R} \otimes_{\mathbb{R}} \mathbb{C}) = \mathrm{End}_{\mathbb{C}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{C})$. We leave it as an exercise for the reader to work out the construction in the reverse direction.

4. Lecture 4

4.1. Classification of elliptic curves via Hodge structures.

Definition 4.1.1. Let R be a subring of $\mathbb{R}$, and Λ an R-Hodge structure. We say that Λ is **pure of weight** m if in the direct sum decomposition $\Lambda \otimes_R \mathbb{C} = \bigoplus_{p,q \in \mathbb{Z}} F^{p,q}$, all nonzero summands $F^{p,q}$ satisfy $p + q = m$.

Definition 4.1.2. Suppose Λ is a $\mathbb{Z}$-Hodge structure, with Hodge decomposition $\Lambda \otimes_{\mathbb{Z}} \mathbb{C} = \bigoplus_{p,q} F^{p,q}$. Let $\Lambda^\vee = \mathrm{Hom}_{\mathbb{Z}\textbf{-mod}}(\Lambda, \mathbb{Z})$. Then we have $\Lambda^\vee \otimes_{\mathbb{Z}} \mathbb{C} \cong \bigoplus_{p,q} (F^{p,q})^*$, where $*$ denotes the $\mathbb{C}$-linear dual. Define $G^{p,q} := F^{-p,-q}$. Then $\Lambda^\vee \otimes_{\mathbb{Z}} \mathbb{C} = \bigoplus_{p,q} G^{p,q}$ is a Hodge decomposition. We thus obtain a $\mathbb{Z}$-Hodge structure $(\Lambda^\vee, (G^{p,q})_{p,q})$, which we call the **dual** of Λ.

Remark. Suppose Λ is a Hodge structure pure of weight m. Then we can recover $F^{p,q}$ from $(\mathrm{Fil}^i)_i$ and m, as

$$F^{p,q} = \mathrm{Fil}^p \cap \overline{\mathrm{Fil}^q}.$$

Recall that there is an equivalence of categories

$$\text{(Elliptic curves)} \xrightarrow{\sim} \text{(Integral Hodge structures of elliptic type)},$$

sending E to $\Lambda = \mathbf{H}_1(E, \mathbb{Z})$ with complex structure J on $\mathbf{H}_1(E, \mathbb{Z}) \otimes \mathbb{R} \cong \mathrm{Lie}\, E$ given by that on $\mathrm{Lie}\, E$.

Remark. The usual Hodge decomposition

$$\mathbf{H}^1(E, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{C} \cong \mathbf{H}^1(E, \mathbb{C}) \cong \mathbf{H}^0(E, \Omega^1_{E/\mathbb{C}}) \oplus \mathbf{H}^1(E, \mathcal{O}_E)$$

endows the $\mathbb{Z}$-module $\mathbf{H}^1(E, \mathbb{Z})$ with the structure of a $\mathbb{Z}$-Hodge structure of type

$$\{(0,1), (1,0)\},$$

where we set $F^{0,1} = \mathbf{H}^1(E, \mathcal{O}_E)$ and $F^{1,0} = \mathbf{H}^0(E, \Omega^1_{E/\mathbb{C}})$. This Hodge structure is the dual of the integral Hodge structure of elliptic type that we attach to E.

In summary, the task of classifying elliptic curves is tantamount to classifying integral Hodge structures of elliptic type up to isomorphism. But this is not so difficult. Since we identify isomorphic Hodge structures, notice that for any integral Hodge structure of elliptic type (Λ, J), we can choose a $\mathbb{Z}$-module isomorphism $f : \Lambda \cong \mathbb{Z}^2$. Then f transports J to a complex structure on $\mathbb{Z}^2 \otimes \mathbb{R} = \mathbb{R}^2$. In other words, we have a bijection

$$\{\text{Integral Hodge structures of elliptic type}\}/\text{isom} \xrightarrow{\sim} \mathrm{GL}_2(\mathbb{Z})\backslash\{\text{Complex structures on } \mathbb{R}^2\}$$

where $\mathrm{GL}_2(\mathbb{Z})$ acts by conjugation.

We now explain how this bijection is related to our previous bijection

$$\mathrm{GL}_2(\mathbb{Z})\backslash\mathbb{H}^{\pm} \xrightarrow{\sim} \{\text{Elliptic curves}\}/\text{isomorphism}$$

given by

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

Firstly, we have the following construction.

Proposition 4.1.3. *There is a natural bijection*

$$\mathbb{H}^{\pm} \xrightarrow{\sim} \{\textit{Complex structures on } \mathbb{R}^2\}.$$

Proof. We first define the map. Recall that there is a $\mathrm{GL}_2(\mathbb{R})$-equivariant injection

$$\mathbb{H}^{\pm} \hookrightarrow \mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, \mathrm{GL}_{2,\mathbb{R}})$$

where $\mathbb{S} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathbb{G}_m)$ is the Deligne torus. Note that an arbitrary element of $\mathbb{H}^{\pm}$ can be described as $\tau = g \cdot i$ where $i \in \mathbb{H}^{\pm}$ is the base point and $g \in \mathrm{GL}_2(\mathbb{R})/\mathrm{SO}_2(\mathbb{R}) \cdot \mathbb{R}^{\times}$. The injection $\mathbb{H}^{\pm} \hookrightarrow \mathrm{Hom}_{\mathbb{R}}(\mathbb{S}, \mathrm{GL}_2(\mathbb{R}))$ is defined by sending $\tau = g \cdot i$ to the morphism $h_\tau : \mathbb{S} \to \mathrm{GL}_{2,\mathbb{R}}$ of algebraic groups over $\mathbb{R}$ defined as follows: for any $\mathbb{R}$-algebra T, we can write $\mathbb{S}(T) = (T \otimes_{\mathbb{R}} \mathbb{C})^{\times} = \{a \otimes 1 + b \otimes i \mid a^2 + b^2 \in T^{\times}\}$. The morphism h_τ is defined via $a \otimes 1 + b \otimes i \mapsto g \begin{pmatrix} a & b \\ -b & a \end{pmatrix} g^{-1}$. We then define

$$\mathbb{H}^{\pm} \longrightarrow \{\text{Complex structures on } \mathbb{R}^2\}$$

via $\tau \mapsto J_\tau := h_\tau(i) \in \mathrm{GL}_2(\mathbb{R})$ where $i \in \mathbb{C}^{\times} = \mathbb{S}(\mathbb{R})$. It is an exercise to see that J_τ is a complex structure and further that this assignment is a $\mathrm{GL}_2(\mathbb{R})$-equivariant bijection, where $\mathrm{GL}_2(\mathbb{R})$ acts on $\mathbb{H}^{\pm}$ by linear fractional transformations and $\mathrm{GL}_2(\mathbb{R})$ acts on $\{\text{Complex structures on } \mathbb{R}^2\}$ by conjugation. $\square$

Proposition 4.1.4. *We have a string of bijections*

$$\begin{array}{c} \{\textit{Elliptic curves}\}/\textit{isomorphism} \\ \downarrow \\ \{\textit{Integral Hodge structures of elliptic type}\}/\textit{isomorphism} \\ \downarrow \\ \mathrm{GL}_2(\mathbb{Z})\backslash\{\textit{Complex structures on } \mathbb{R}^2\} \\ \downarrow \\ \mathrm{GL}_2(\mathbb{Z})\backslash\mathbb{H}^{\pm} \\ \downarrow \\ \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^{+} \end{array}$$

where the third map is induced by the bijection $\tau \mapsto J_\tau$ in the previous proposition. The composition of these maps is the inverse to $\tau \mapsto \mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$.

Proof. Exercise. ∎

5. Lecture 5

5.1. **Some motivation for level structure.** Recall Proposition 4.1.4. In fact, we note that there is also an $\mathrm{SL}_2(\mathbb{Z})$-invariant holomorphic morphism $j : \mathbb{H}^+ \to \mathbb{C}$ inducing a bijection

$$\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+ \xrightarrow{\sim} \mathbb{C}.$$

Here j corresponds to evaluating the classical j-invariant of an elliptic curve. We may use this map to identify the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+$ with $\mathbb{C}$ in order to give the former a complex manifold structure.

Note that $\mathbb{H}^+ \to \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+$ is a holomorphic map, but not a local isomorphism. In other words, this is not a covering map; there is ramification over the images of i and $e^{\frac{2\pi i}{3}}$ with branching of order 2 and 3 respectively. This is related to the fact that the $\mathrm{SL}_2(\mathbb{Z})$-action on $\mathbb{H}^+$ is problematic in the following sense:

- $-I \in \mathrm{SL}_2(\mathbb{Z})$ acts trivially on $\mathbb{H}^+$. In particular, the $\mathrm{SL}_2(\mathbb{Z})$-action on $\mathbb{H}^+$ is not free.
- The naive solution is to now consider the action of $\mathrm{SL}_2(\mathbb{Z})/\{\pm I\}$ on $\mathbb{H}^+$. For this action, most points in $\mathbb{H}^+$ have trivial stabilizer, but points in the orbit of i and the orbit of $e^{\frac{2\pi i}{3}}$ have nontrivial stabilizers. So this is also not a solution.

This phenomenon exactly corresponds to the fact that for any elliptic curve E over $\mathbb{C}$ (or any algebraically closed field of characteristic away from 2 or 3), the automorphism group of E is either:

(1) $\mathbb{Z}/2\mathbb{Z}$, where the nontrivial automorphism is negation. This corresponds to the inclusion of $\{\pm I\}$ in all stabilizers.
(2) $\mathbb{Z}/4\mathbb{Z}$. This automorphism group applies to a unique isomorphism class of elliptic curves.
(3) $\mathbb{Z}/6\mathbb{Z}$. This automorphism group applies to a unique isomorphism class of elliptic curves.

Remark. The complex manifold structure we put on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+$ (using j) is the unique one such that the projection $\mathbb{H}^+ \to \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^+$ is holomorphic.

Remark. It is "more correct", in some sense, to define the orbifold (or Deligne–Mumford stack) quotient of $\mathbb{H}^+$ by $\mathrm{SL}_2(\mathbb{Z})$. This allows us to form a fine moduli space of elliptic curves that remembers the automorphisms (including the generic $\mathbb{Z}/2\mathbb{Z}$-automorphisms). As we will see in the future, the only obstruction to representability of the moduli of elliptic curves in the category of schemes is the presence of these automorphisms.

5.2. **Level structure.** Instead of seriously talking about orbifolds or stacks, we shall mainly focus on the following solution to the presence of automorphisms: we will rigidify the moduli problem by asking for some additional structures on the elliptic curves that will kill all automorphisms. In particular, no nontrivial automorphism of E will preserve this extra structure. Correspondinly we will need to shrink $\mathrm{SL}_2(\mathbb{Z})$ to some smaller congruence subgroup Γ such that the Γ-action on $\mathbb{H}^+$ is free.

Fix an integer $N \geq 3$ throughout. For any elliptic E over $\mathbb{C}$, consider the N-torsion subgroup

$$E[N] := \{z \in E \mid z + \cdots + z \ (N \text{ times}) \ = 0\}.$$

Recall that $E[N]$ is non-canonically isomorphic to $(\mathbb{Z}/\mathbb{Z})^2 = \mathbb{Z}/N\mathbb{Z} \oplus \mathbb{Z}/N\mathbb{Z}$ as $\mathbb{Z}/N\mathbb{Z}$-modules.

Definition 5.2.1. A choice of an isomorphism $\gamma : E[N] \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^2$ is called a **level-N structure** on E. Equivalently, this is a choice of an ordered basis (P, Q) of $E[N]$ as a free $\mathbb{Z}/N\mathbb{Z}$-module.

Recall that elliptic curves over $\mathbb{C}$ correspond to integral Hodge structures of elliptic type (Λ, J). For an elliptic curve $E/\mathbb{C}$, the corresponding integral Hodge structure of elliptic type was obtained by setting $\Lambda = \mathbf{H}_1(E, \mathbb{Z})$. We have the canonical isomorphisms

$$E[N] \cong \frac{1}{N}\Lambda/\Lambda \cong \Lambda/N\Lambda \cong \Lambda \otimes_{\mathbb{Z}} \mathbb{Z}/N\mathbb{Z}.$$

It thus makes sense to formulate the following definition.

Definition 5.2.2. A **level-N structure** on an integral Hodge structure of elliptic type (Λ, J) is a choice of an isomorphism $\gamma : \Lambda/N\Lambda \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^2$.

Now, define

$$\Gamma(N) := \{g \in \mathrm{SL}_2(\mathbb{Z}) \mid g \equiv 1 \mod N\} \trianglelefteq \mathrm{SL}_2(\mathbb{Z}).$$

The action of such groups on $\mathbb{H}^+$ behaves better than that of $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 5.2.3. *For $N \geq 3$, $\Gamma(N)$ acts freely and properly discontinuously on $\mathbb{H}^+$.*

Proof. We sketch the proof that the action is free. Suppose $\gamma \in \Gamma(N)$ has a fixed point in $\mathbb{H}^+$. Since the stabilizer of $i \in \mathbb{H}^+$ in $\mathrm{SL}_2(\mathbb{R})$ is $\mathrm{SO}_2(\mathbb{R})$ and since $\mathbb{H}^+$ is transitive under $\mathrm{SL}_2(\mathbb{R})$, we see that γ must lie in a $\mathrm{SL}_2(\mathbb{R})$-conjugate of $\mathrm{SO}_2(\mathbb{R})$. In particular γ must be semi-simple and its eigvenvalues in $\mathbb{C}$ have absolute value 1. On the other hand, the characteristic polynomial of γ is monic with integer coefficients, so the eigenvalues of γ are algebraic integers. Combined with the previous fact, we see that the eigenvalues of γ must be roots of unity. Pick a prime power p^e dividing N, and we can arrange that $p^e \geq 3$. Since $\gamma \equiv 1 \mod p^e$, each eigenvalue λ of γ in $\overline{\mathbb{Q}}_p$ must satisfy $v_p(\lambda - 1) \geq e$. We leave it as an exercise to the reader to show that any root of unity $\lambda \in \overline{\mathbb{Q}}_p$ satisfying $v_p(\lambda - 1) \geq e$ must be trivial provided that $p^e \geq 3$. Thus the eigenvalues of γ are trivial, so γ must be trivial.

We omit the proof that $\Gamma(N)$ acts properly discontinuously. See [DS05, §2.1]. $\square$

In particular, this implies that $\Gamma(N)\backslash\mathbb{H}^+$ has the natural structure of a Riemann surface and $\mathbb{H}^+ \to \Gamma(N)\backslash\mathbb{H}^+$ is a covering. Further, this is obviously the universal covering, since $\mathbb{H}^+$ is simply connected.

Definition 5.2.4. The **modular curve** $Y(N)$ is the complex manifold

$$Y(N) := \coprod_{j \in (\mathbb{Z}/N\mathbb{Z})^\times} \Gamma(N)\backslash\mathbb{H}_j^+$$

where $\mathbb{H}_j^+ := \mathbb{H}^+$.

Remark. Classically, the notation $Y(N)$ often refers to just a single connected component, namely $Y(N) = \Gamma(N)\backslash\mathbb{H}^+$.

For each $j \in (\mathbb{Z}/N\mathbb{Z})^\times$, fix once and for all an element $g_j \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ such that $\det(g_j) = j$. For instance, we may take $g_j = \begin{pmatrix} j & 0 \\ 0 & 1 \end{pmatrix}$.

For each $j \in (\mathbb{Z}/N\mathbb{Z})^\times$ and each $\tau \in \mathbb{H}_j^+$, we will define an integral Hodge structure of elliptic type together with a level-N structure:

$$\mathcal{V}_\tau = (\Lambda_\tau, J_\tau, \gamma_\tau : \Lambda_\tau/N\Lambda_\tau \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^2)$$

as follows:

- $\Lambda_\tau := \mathbb{Z}^2$.
- J_τ is the complex structure on $\mathbb{R}^2$ corresponding to τ, i.e., $J_\tau := h_\tau(i)$.
- $\gamma_\tau : \Lambda_\tau/N\Lambda_\tau \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^2$ is the isomorphism defined by $g_j : (\mathbb{Z}/N\mathbb{Z})^2 \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^2$. (By the definition of Λ_τ, $\Lambda_\tau/N\Lambda_\tau$ *is* $(\mathbb{Z}/N\mathbb{Z})^2$).

Observe that if $\tau, \tau' \in \mathbb{H}_j^+$ are related by the $\Gamma(N)$-action, then we have $(\Lambda_\tau, J_\tau, \gamma_\tau) \cong (\Lambda_{\tau'}, J_{\tau'}, \gamma_{\tau'})$, i.e., there is an isomorphism of integral Hodge structures compatible with the level structures. In turn, we get a map

(5.1)
$$Y(N) \to \{\text{Integral Hodge structures of elliptic type with level-}N\text{ structure}\}/\text{isomorphism}.$$

In the next lecture we will show that this map is a bijection.

6. Lecture 6

6.1. Points on the modular curve.

Proposition 6.1.1. *The map (5.1) is a bijection.*

Proof. We leave injectivity as an exercise. For surjectivity, let $\mathcal{V} = (\Lambda, J, \gamma)$ be an arbitrary integral Hodge structure of elliptic type together with a level N-structure. Pick a group isomorphism $u : \Lambda \xrightarrow{\sim} \mathbb{Z}^2$. Then u takes J to some complex structure on $\mathbb{Z}^2 \otimes_{\mathbb{Z}} \mathbb{R} = \mathbb{R}^2$, which must be of the form J_τ for some $\tau \in \mathbb{H}^\pm$. If $\tau \in \mathbb{H}^-$, we compose u with some element of $\mathrm{GL}_2(\mathbb{Z})$ of determinant -1, and as a result we can always assume that $\tau \in \mathbb{H}^+$. Now let γ' be the composition

$$(\mathbb{Z}/N\mathbb{Z})^2 \xrightarrow{u^{-1} \mod N} \Lambda/N\Lambda \xrightarrow{\gamma} (\mathbb{Z}/N\mathbb{Z})^2.$$

Then $\gamma' \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. We note the following fact:

Fact. (Strong approximation for SL_2.) The natural map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is surjective.

For a proof, see [DS05, Exercise 1.2.2]. Note that the statement is not true if we replace SL_2 by GL_2, since elements of $\mathrm{GL}_2(\mathbb{Z})$ all have determinants ± 1 .

Let $j = \det(\gamma')$, so $\gamma' g_j^{-1} \in \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. By the above fact, we can compose u with a suitable element of $\mathrm{SL}_2(\mathbb{Z})$ to arrange that $\gamma' = g_j$. When we do this the element τ we found in the above will be moved by the element of $\mathrm{SL}_2(\mathbb{Z})$, but the new τ still lies in $\mathbb{H}^+$. We think of it as an element of $\mathbb{H}_j^+$. It is then easy to check that u induces an isomorphism between $\mathcal{V}$ and $\mathcal{V}_\tau$. Thus the image of τ under (5.1) is the isomorphism class of $\mathcal{V}$. $\square$

Corollary 6.1.2. *Points on $Y(N)$ are in natural bijection with the isomorphism classes of elliptic curves with level N-structure.*

Remark. Classically, one considers only one connected component of $Y(N)$, and correspondingly imposes a stronger condition on the level N-structure. For this, recall that for each elliptic curve E, the **Weil pairing** is a canonical isomorphism of $\mathbb{Z}/N\mathbb{Z}$-modules

$\bigwedge^2 E[N] \xrightarrow{\sim} \mu_N$. (Here $\bigwedge^2$ is taken in the category of finite free $\mathbb{Z}/N\mathbb{Z}$-modules, and μ_N is the set of N-th roots of unity in $\mathbb{C}$.) Thus any level-N structure on E will induce an isomorphism between μ_N and $\bigwedge^2(\mathbb{Z}/N\mathbb{Z}) = \mathbb{Z}/N\mathbb{Z}$, or equivalently the choice of a generator of μ_N. If one insists that the level structure must induce a prescribed generator of μ_N, then under the correspondence one sees only one connected component of $Y(N)$. It turns out that the generator $e^{2\pi i/N}$ corresponds to the connected component $\Gamma(N)\backslash\mathbb{H}_1^+$. The verification of the last claim amounts to the explicit computation of the Weil pairing for $\mathbb{C}/\mathbb{Z}+\tau\mathbb{Z}$ as in [Sil94, Exercise 1.15].

Up to this point, we just described how points on $Y(N)$ correspond to isomorphism classes of elliptic curves or integral Hodge structures of elliptic type with level-N structure. We have not used the complex manifold structure or complex algebraic variety structure on $Y(N)$. We would now like to upgrade the pointwise correspondence to a moduli interpretation. Recall that the complex manifold $Y(N)$ has a unique compatible structure of an algebraic variety, by the theorems of Baily–Borel and Borel. The following is our target theorem.

Theorem 6.1.3. *The algebraic variety $Y(N)$ is the moduli space of elliptic curves with level-N structure. Namely, it represents the functor sending each finite-type $\mathbb{C}$-scheme V to the set of isomorphism classes of proper flat families of elliptic curves over V with level-N structure.*

We will explain the notions in the statement of the theorem in more detail later.

6.2. **Variation of Hodge structures.** One important ingredient towards the proof of the Theorem 6.1.3 is the correct notion of a "family of integral Hodge structures of elliptic type". Suppose S is a complex manifold and for each $s \in S$ we are given an integral Hodge structure of elliptic type (Λ_s, J_s). The question is how to formulate that these structures vary nicely as s varies in S. In particular, we need to take into account the complex manifold structure on S. The answer is provided in the following definition.

Definition 6.2.1. A **variation of integral Hodge structures of elliptic type on** S refers to a pair $(\underline{\Lambda}, \mathcal{L})$ consisting of a locally constant sheaf of abelian groups $\underline{\Lambda}$ on S that is locally free of rank 2 over $\mathbb{Z}$, and a holomorphic line subbundle $\mathcal{L}$ of the plane bundle $\underline{\Lambda} \otimes_{\mathbb{Z}} \mathcal{O}_S$. (Here $\mathcal{O}_S$ is the structure sheaf of holomorphic functions on S.) They should satisfy the following condition:

- For each $s \in S$, the 1-dimensional $\mathbb{C}$-subspace of $\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathbb{C}$ determined by the fiber of $\mathcal{L}$[3] is the Fil^0 associated with a Hodge structure of elliptic type on $\underline{\Lambda}_s$ (i.e., a complex structure on $\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathbb{R}$, or equivalently a Hodge decomposition $\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathbb{C} = F^{-1,0} \oplus F^{0,-1}$ such that $\overline{F^{-1,0}} = F^{0,-1}$).

Recall that for Hodge structures of elliptic type, or more generally pure Hodge structures, the Hodge decomposition and the Hodge filtration determine each other. Thus the above definition is equivalent to the datum of $\underline{\Lambda}$ together with a Hodge structure of elliptic type on $\underline{\Lambda}_s$ for each $s \in S$ such that the pointwise Fil^0's vary holomorphically in the sense that they come from some holomorphic line subbundle $\mathcal{L} \subset \underline{\Lambda} \otimes_{\mathbb{Z}} \mathcal{O}_S$.

The following is the more general definition.

Definition 6.2.2. Let R be a subring of $\mathbb{R}$, and let m be an integer. A **variation of R-Hodge structures of weight m on S** refers to a pair $(\underline{\Lambda}, \mathcal{L}^\bullet)$, where $\underline{\Lambda}$ is a locally

[3] Here $\underline{\Lambda}_s$ is the stalk of Λ at s, and we identify $\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathbb{C}$ with the fiber of the plane bundle $\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathcal{O}_S$ at s. Thus the fiber of $\mathcal{L}$ at s gives rise to a subspace.

constant sheaf of R-modules on S that is locally finite free, and $\mathcal{L}^\bullet$ is a decreasing filtration on $\underline{\Lambda} \otimes_R \mathcal{O}_S$ by holomorphic sub-vector bundles. They should satisfy the following two conditions:

- For each $s \in S$, the filtration on the $\mathbb{C}$-vector space $\underline{\Lambda}_s \otimes_R \mathbb{C}$ determined by the fibers of $\mathcal{L}^\bullet$ at s is the Hodge filtration associated with a (unique) Hodge structure of weight m on $\underline{\Lambda}_s$.
- (Griffiths transversality.) Let $\nabla : \underline{\Lambda} \otimes_R \mathcal{O}_S \to \underline{\Lambda} \otimes_R \Omega^1_S$ be the flat connection $\mathrm{id} \otimes d$. For each p, we have $\nabla(\mathcal{L}^p) \subset \mathcal{L}^{p-1} \otimes_{\mathcal{O}_S} \Omega^1_S$.

Note that for Hodge structures of elliptic type, Griffiths transversality is automatic.

If we have a smooth projective variety X over $\mathbb{C}$, then for each non-negative integer m we have the **Hodge decomposition** from Hodge theory:

$$\mathbf{H}^m(X, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{C} = \mathbf{H}^m(X, \mathbb{C}) = \bigoplus_{p+q=m} \mathbf{H}^q(X, \Omega^p_X).$$

This makes $\mathbf{H}^m(X, \mathbb{Z})$ a $\mathbb{Z}$-Hodge structure of weight m (with $F^{p,q} = \mathbf{H}^q(X, \Omega^p_X)$). The following result is the motivation for the definition of a variation of Hodge structures.

Theorem 6.2.3 (Griffiths)**.** *Let X and S be smooth algebraic varieties over $\mathbb{C}$, and let $X \to S$ be a smooth projective morphism. Then the Hodge structures on $\mathbf{H}^m(X_s, \mathbb{Z})$ for $s \in S$ come from a canonical variation of $\mathbb{Z}$-Hodge structures of weight m on S^{an}.*

As a special case, we obtain the following

Theorem 6.2.4. *Let S be a smooth algebraic variety over $\mathbb{C}$, and let $E \to S$ be a proper flat family of elliptic curves. Then the integral Hodge structures of elliptic type assigned to E_s for $s \in S$ come from a canonical variation of integral Hodge structures of elliptic type on S^{an}.*

This is indeed a special case of Griffiths' theorem, since the integral Hodge structure of elliptic type attached to an elliptic curve E over $\mathbb{C}$ is the dual of $\mathbf{H}^1(E, \mathbb{Z})$, and since $E \to S$ is automatically projective (i.e., after Zariski localization on S, it factors through $E \to \mathbb{P}^1_S$; a fact that is not true for higher dimensional abelian varieties). In fact, after Zariski localization on S, one can always find an S-embedding $E \hookrightarrow \mathbb{P}^1_S$ that is described by a Weierstrass equation. (More details in the future.)

7. Lecture 7

7.1. **Variation of Hodge structures, continued.**

We briefly recall some notions from before.

Definition 7.1.1. Let S be a complex manifold. A **variation of integral Hodge structures of elliptic type** over S is a pair $(\underline{\Lambda}, J)$ where:

- $\underline{\Lambda}$ is a locally constant sheaf of $\mathbb{Z}$-modules on S that is locally free of rank 2 over $\mathbb{Z}$.
- $J = (J_s)_{s \in S}$ is a family where that each J_s is a complex structure on the $\mathbb{R}$-vector space $\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathbb{R}$ such that $\mathrm{Fil}^0(\underline{\Lambda}_s \otimes_{\mathbb{Z}} \mathbb{C}) = (F^{0,-1})_s$ varies holomorphically in the sense that they all come from a holomorphic line sub-bundle $\mathcal{L}$ of the plane bundle $\underline{\Lambda} \otimes_{\mathbb{Z}} \mathcal{O}_S$.

Remark. Remembering $(\underline{\Lambda}, J)$ is the same as remembering $(\underline{\Lambda}, \mathcal{L})$.

Definition 7.1.2. By a **level-N structure** on a variation of integral Hodge structures of elliptic type $(\underline{\Lambda}, J)$, we mean a choice of sheaf isomorphism (as sheaves of groups)

$$\gamma : \underline{\Lambda} \otimes_{\mathbb{Z}} \mathbb{Z}/N\mathbb{Z} \xrightarrow{\sim} \underline{(\mathbb{Z}/N\mathbb{Z})^2}.$$

Remark. Such an isomorphism need not exist in general!

7.2. **Families of elliptic curves.**

Definition 7.2.1. Let S be a scheme. An **elliptic curve** over S is a proper and smooth morphism of schemes $\pi : E \to S$, with geometric fibers that are (smooth and) connected curves of genus 1, along with the datum of a section $e : S \to E$.

Theorem 7.2.2 ("Abel", see [KM85, §2.1])**.** *In the notation from the above definition, there is a canonical structure of an S-group scheme on E with $e : S \to E$ the identity section. Moreover, this group scheme is commutative.*

This group scheme structure is given similarly to the case where $S = \operatorname{Spec} k$, with k a field. For any S-scheme T, we will make $E(T)$ a group as follows. Write f for the structural morphism $E \to S$, and let f_T be the pullback of f over T:

$$\begin{array}{ccc} E_T & \longrightarrow & E \\ {\scriptstyle f_T}\downarrow & & \downarrow{\scriptstyle f} \\ T & \longrightarrow & S. \end{array}$$

For $P \in E(T)$, denote by D_P the Cartier divisor on E_T given by $\operatorname{im}(P)$. (Thus $\mathcal{O}_{E_T}(D_P)$ is the invertible $\mathcal{O}_{E_T}$-module that is inverse to the ideal sheaf of $\operatorname{im}(P)$.) Write e_T for the section of $E_T \to T$ induced by e. Then for any $P, Q, R \in E(T)$, we impose that

$$P + Q = R$$

if and only if

$$\mathcal{O}_{E_T}(D_P) \otimes \mathcal{O}_{E_T}(D_Q) \cong \mathcal{O}_{E_T}(D_{e_T}) \otimes \mathcal{O}_{E_T}(D_R) \otimes f_T^*(\mathcal{L})$$

for a line bundle $\mathcal{L}$ on T.

Remark. The attribution to Abel rests in his classical proof that for an elliptic curve E over $\mathbb{C}$ we have

$$E \xrightarrow{\sim} \operatorname{Jac}(E) = \{\deg 0 \text{ divisors on } E\}/\text{linear equivalence}$$

via $P \mapsto [D_P - D_e]$.

Proposition 7.2.3 ([KM85, §2.3])**.** *Suppose E/S is an elliptic curve. After localizing S, the map $E \to S$ is projective. More precisely, each point in S has an open neighborhood U such that $E_U \to U$ factors through an embedding $E_U \hookrightarrow \mathbb{P}^2_U$ whose image is described by a generalized Weierstrass equation*

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \qquad a_i \in \mathcal{O}_U(U).$$

Remark. Projectivity over the base (even after localization) is not true for general abelian schemes; this is a special phenomenon about elliptic curves.

Proposition 7.2.3, combined with Theorem 6.2.3, implies the following corollary:

Corollary 7.2.4. *Suppose S is a smooth variety over $\mathbb{C}$ and E/S is an elliptic curve. The pointwise integral Hodge structures of elliptic type associated to E_s for $s \in S(\mathbb{C})$ come from a variation of integral Hodge structures of elliptic type over S^{an}. (More precisely, the local system in the variation of integral Hodge structures is the $\mathbb{Z}$-linear dual of the first derived pushforward of $\underline{\mathbb{Z}}$ along $E^{\mathrm{an}} \to S^{\mathrm{an}}$.)*

Definition 7.2.5. A **(naive) level-N structure** on an elliptic curve E/S is a choice of isomorphism of S-group schemes

$$\gamma : E[N] \xrightarrow{\sim} \underline{(\mathbb{Z}/N\mathbb{Z})^2}.$$

Remark. In general, $E[N]$ is a finite locally free S-group of rank N^2. We will only consider the case when N is invertible on S, i.e., S is over $\mathbb{Z}[1/N]$. In this case $E[N]$ is finite étale over S, and after a finite étale base change $E[N]$ is isomorphic to the constant group scheme $\underline{(\mathbb{Z}/N\mathbb{Z})^2}$.

Let S be a $\mathbb{Z}[1/N]$-scheme, and E/S an elliptic curve. Using the above remark, it is an exercise to show that giving a level N-structure on E/S is the same as providing an ordered pair of sections P and Q of the S-scheme $E[N]$ such that for each geometric point $\bar{s}$ in S, the fibers of P and Q at $\bar{s}$ generate the group $E_{\bar{s}}[N]$.

Remark. There is an obvious version of Corollary 7.2.4 incorporating level-N structure.

7.3. **The moduli functor.**

Theorem 7.3.1. *Suppose $N \geq 3$. We define the contravariant functor*

$$S(N) : (\textit{finite-type schemes over } \mathbb{Z}[1/N]) \longrightarrow (\textit{sets})$$

by

$$S \longmapsto \{\textit{isomorphism classes of elliptic curves } E/S \textit{ with level-}N \textit{ structure}\}.$$

(On morphisms, this functor is defined by the obvious notion of pullback.) Then $S(N)$ is representable by a nice—in particular, finite type—scheme over $\mathbb{Z}[1/N]$, still denoted by $S(N)$.

Recall that by the Bailey–Borel compactification the complex manifold $Y(N)$ has a unique structure of an algebraic variety over $\mathbb{C}$.

Theorem 7.3.2. *$S(N)_{\mathbb{C}}$ is canonically isomorphic to $Y(N)$.*

We will indicate the proof of Theorem 7.3.1 using explicit manipulation with Weierstrass equations in the near future. Later we will prove it again by deducing it from Mumford's more general theorem for the moduli of abelian schemes. We now explain the proof of Theorem 7.3.2. For this we will first prove the following proposition.

Proposition 7.3.3. *The complex manifold $Y(N)$ represents the contravariant functor*

$$(\textit{complex manifolds}) \longrightarrow (\textit{sets})$$

sending S to the set of isomorphism classes of variations of $\mathbb{Z}$-Hodge structures of elliptic type with level-N structure on S. (On morphisms, this functor is defined by the obvious notion of pullback.)

Proof. Recall that for each $\tau \in Y(N)$, we have constructed an integral Hodge structure of elliptic type with level-N structure $\mathcal{V}_\tau$ which is well defined up to isomorphism; see the map (5.1) . We want to construct a variation of integral Hodge structures of elliptic type with level-N structure on $Y(N)$ that recovers this pointwise construction. In the rest of the proof, we abbreviate "variation of integral Hodge structures of elliptic type with level N-structure" simply as "VHSL".

Let $j \in (\mathbb{Z}/N\mathbb{Z})^\times$. Recall that we have fixed $g_j \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ such that $\det(g_j) = j$. First, we construct a VHSL on $\mathbb{H}_j^+$. Namely, we define

$$\widetilde{\mathcal{V}_j} = (\underline{\mathbb{Z}^2},\ (J_\tau)_{\tau \in \mathbb{H}_j^+},\ \gamma : \underline{\mathbb{Z}^2} \otimes \mathbb{Z}/N\mathbb{Z} = \underline{(\mathbb{Z}/N\mathbb{Z})^2} \xrightarrow{g_j} \underline{(\mathbb{Z}/N\mathbb{Z})^2}).$$

Here, as always, J_τ denotes $h_\tau(i) \in \mathrm{GL}_2(\mathbb{R})$. To show that $\widetilde{\mathcal{V}_j}$ is indeed a variation of Hodge structures, consider the map $\iota : \mathbb{H}_j^+ \to \mathbb{P}^1(\mathbb{C})$ sending $\tau \in \mathbb{H}_j^+$ to the 1-dimensional subspace $\mathrm{Fil}^0 \subset \mathbb{C}^2$ determined by J_τ (i.e., the $-i$-eigenspace of the complexification of J_τ). One checks that ι is the standard open embedding $\mathbb{H}_j^+ = \mathbb{C} - \mathbb{R} \hookrightarrow \mathbb{P}^1(\mathbb{C})$. This shows that the pointwise Fil^0's in $\widetilde{\mathcal{V}_j}$ indeed vary holomorphically; more precisely, they come from the tautological line bundle on $\mathbb{P}^1(\mathbb{C})$ restricted to $\mathbb{H}_j^+$.

Now, suppose we have two points $\tau, \tau' \in \mathbb{H}_j^+$ related by some (unique) $g \in \Gamma(N)$. Say $\tau' = g\tau$. Then for any open neighborhood U of τ in $\mathbb{H}_j^+$, we have an isomorphism

$$\widetilde{\mathcal{V}_j}|_U \xrightarrow{\sim} g^*(\widetilde{\mathcal{V}_j}|_{g(U)})$$

between VHSL's on U given by $g^{-1} : \underline{\mathbb{Z}^2} \xrightarrow{\sim} \underline{\mathbb{Z}^2}$. (Exercise: check that this indeed preserves the other structures.) These isomorphisms satsify the cocycle relation and give rise to a descent datum from $\mathbb{H}_j^+$ to $\Gamma(N)\backslash\mathbb{H}_j^+$, by which we obtain a VHSL on $\Gamma(N)\backslash\mathbb{H}_j^+$, denoted by $\mathcal{V}_j$. Note here that the local system in $\mathcal{V}_j$ is no longer constant, but rather its monodromy group is $\Gamma(N)$, which is the full fundamental group of $\Gamma(N)\backslash\mathbb{H}_j^+$. Taking disjoint union over the $j \in (\mathbb{Z}/N\mathbb{Z})^\times$, we obtain $\mathcal{V}^{\mathrm{univ}}$ on $Y(N)$.

In the next lecture, we will show that $\mathcal{V}^{\mathrm{univ}}$ is the universal VHSL. □

8. Lecture 8

8.1. Proof of Proposition 7.3.3, continued.

Proof. It remains to show that $\mathcal{V}^{\mathrm{univ}}$ is the universal VHSL. This amounts to showing that for any complex manifold S and any VHSL $\mathcal{V}$ on S, there exists a unique holomorphic map $f : S \to Y(N)$ such that $\mathcal{V} \cong f^*\mathcal{V}^{\mathrm{univ}}$. By Proposition 6.1.1, we have a bjiection

$$Y(N) \to \{\text{Integral Hodge structures of elliptic type with level } N \text{ structure}\}/\text{isomorphism}.$$

One checks that this bijection is exactly given by sending $\tau \in Y(N)$ to the isomorphism class of $(\mathcal{V}^{\mathrm{univ}})_\tau$. This implies that the desired map f must send each $s \in S$ to the unique $f(s) \in Y(N)$ such that $(\mathcal{V}^{\mathrm{univ}})_{f(s)}$ is isomorphic to $\mathcal{V}_s$. Thus we know that f is unique.

We must still check that f given by the above recipe is holomorphic and $f^*\mathcal{V}^{\mathrm{univ}} \cong \mathcal{V}$. If we wish to show that f is holomorphic at $s_0 \in S$, then we may shrink S to an open neighborhood of s_0, which we may assume is connected and simply connected. This is always possible, since S is just a complex manifold. In particular, after shrinking, we may assume that the local system in $\mathcal{V}$ is constant. We denote

$$\mathcal{V} = (\underline{\Lambda}, (J_s)_{s \in S}, \gamma : \underline{\Lambda} \otimes \mathbb{Z}/N\mathbb{Z} \xrightarrow{\sim} \underline{(\mathbb{Z}/N\mathbb{Z})^2}).$$

Pick an isomorphism $\underline{\Lambda} \cong \underline{\mathbb{Z}^2}$. Then γ becomes an element of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$. Here $(J_s)_{s \in S}$ becomes a family of elements $(\tau_s)_{s \in S}$ where $\tau_s \in \mathbb{H}^\pm$. Using the fact that Fil^0 varies holomorphically, we know that $S \to \mathbb{H}^\pm$ given by $s \mapsto \tau_s$ is holomorphic. As such, the image of $S \to \mathbb{H}^\pm$ is either in $\mathbb{H}^+$ or $\mathbb{H}^-$, and in particular we may compose the selected isomorphism $\underline{\Lambda} \cong \mathbb{Z}^2$ with a matrix in $\mathrm{GL}_2(\mathbb{Z})$ with determinant -1 if necessary in order to assume that $S \to \mathbb{H}^{\pm}$ lands in $\mathbb{H}^{+}$.

Now recall that strong approximation for SL_2 says that $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is surjective. In turn, we may further compose the isomorphism $\underline{\Lambda} \cong \mathbb{Z}^2$ by an element of $\mathrm{SL}_2(\mathbb{Z})$ if necessary to ensure that this isomorphism carries

$$\gamma : \underline{\Lambda} \otimes \mathbb{Z}/N\mathbb{Z} \to \underline{(\mathbb{Z}/N\mathbb{Z})^2}$$

to g_j for some j.

After these adjustments, we get a holomorphic map $\tilde{f} : S \to \mathbb{H}_j^{+}$ such that for all $s \in S$, we have $\mathcal{V}_s \cong (\widetilde{\mathcal{V}_j})_{\tilde{f}(s)}$. We have a commutative diagram

$$\begin{tikzcd} & \mathbb{H}_j^{+} \arrow[d] \\ S \arrow[ur, "\tilde{f}"] \arrow[r, "f"'] & Y(N). \end{tikzcd}$$

Note here that S has been shrunk to be connected and simply connected; of course such a lifting cannot be produced for general S. This implies that f is holomorphic since $\tilde{f}$ is holomorphic.

Now we need that $f^*\mathcal{V}^{\mathrm{univ}}$ is isomorphic to $\mathcal{V}$. First, we shrink S to an open neighborhood of any given point in S, and we construct $\tilde{f}$ as before. Then, on any such neighborhood, we have

$$\mathcal{V} \cong \tilde{f}^*(\widetilde{\mathcal{V}}_j) \cong f^*\mathcal{V}^{\mathrm{univ}}$$

where the first isomorphism can be checked using the construction of $\tilde{f}$ and the second isomorphism follows from the above commutative diagram. In other words, we have an open covering $(U_i)_{i \in I}$ of S such that for each i, there is an isomorphism

$$\varphi_i : \mathcal{V}|_{U_i} \xrightarrow{\sim} (f^*\mathcal{V}^{\mathrm{univ}})|_{U_i}.$$

This yields our claim locally. Now we note the following fact.

Fact 8.1.1. *For $N \geq 3$, on any complex manifold any VHSL has no automorphisms.*

Proof. Let S be a complex manifold and $\mathcal{V}$ a VHSL on S. We may assume that S is connected. Then any automorphism of $\mathcal{V}$ is uniquely determined by its behavior at one fiber (since it is after all an automorphism of a local system). Thus we reduce to the case where S is a point. In other words, we need to show that any integral Hodge structure of elliptic type with level-N structure (Λ, J, γ) has no automorphism. Suppose g is an automorphism. Choose an identification $\Lambda \cong \mathbb{Z}^2$. Then g becomes an element of $\mathrm{GL}_2(\mathbb{Z})$. Since g preserves γ, g lies in $\Gamma(N)$. Since g preserves J, g has a fixed point in $\mathbb{H}^{\pm}$. But $\Gamma(N)$ acts freely on $\mathbb{H}^{\pm}$ (Proposition 5.2.3), so $g = 1$. □

This fact implies that on nontrivial intersections $U_i \cap U_j$, the isomorphisms φ_i and φ_j must agree; else $\varphi_i^{-1} \circ \varphi_j$ is a nontrivial automorphism of $\mathcal{V}|_{U_i \cap U_j}$. Thus we can glue together the isomorphisms φ_i to get a global isomorphism $\varphi : \mathcal{V} \xrightarrow{\sim} f^*\mathcal{V}^{\mathrm{univ}}$. □

9. Lecture 9

9.1. **Isomorphism between the algebraic and analytic moduli spaces.** Recall the following two theorems from before. Let N be an integer ≥ 3.

Theorem 9.1.1. *We define a contravariant functor*

$$S(N) : (\textit{finite type } \mathbb{Z}[1/N]\textit{-schemes}) \longrightarrow (\textit{Sets})$$

given by

$$T \longmapsto \{\textit{iso. cl. of elliptic curves over } T \textit{ with level-}N \textit{ structure}\}.$$

Then $S(N)$ is representable by a "nice" $\mathbb{Z}[1/N]$-scheme, also denoted by $S(N)$. In particular, $S(N)_{\mathbb{C}}$ is a smooth $\mathbb{C}$-variety.[4]

Theorem 9.1.2. *We have a natural isomorphism of $\mathbb{C}$-varieties*

$$S(N)_{\mathbb{C}} \cong Y(N).$$

We will prove Theorem 9.1.2 assuming Theorem 9.1.1.

Proof. We write "VHSL" for "variation of integral Hodge structures of elliptic type with level-N structure". For any smooth $\mathbb{C}$-variety T, by the version of Corollary 7.2.4 incorporating level-N structures, we have a functor

$$(\text{elliptic curves over } T \text{ with level-}N \text{ structure}) \to (\text{VHSL on } T^{\mathrm{an}}).$$

In particular, we have a natural map between the sets of isomorphism classes. By Proposition 7.3.3 and Theorem 9.1.1, this is tantamount to a map

$$\mathrm{Hom}_{\mathbb{C}\text{-sch}}(T, S(N)_{\mathbb{C}}) \longrightarrow \mathrm{Hom}_{\mathrm{hol}}(T^{\mathrm{an}}, Y(N)).$$

Consider the universal case, i.e., $T = S(N)_{\mathbb{C}}$. Then the distinguished element of the left hand side—namely, the identity—gives rise to a distinguished holomorphic map $f : S(N)_{\mathbb{C}}^{\mathrm{an}} \to Y(N))$. By Borel's Theorem, f is algebraic. Moreover, f is a bijection on $\mathbb{C}$-points, since the induced map on $\mathbb{C}$-points is the familiar bijection

$$\begin{aligned}&\{\text{isom. cl. of elliptic curves with level-}N \text{ structure}\}\\ &\quad\xrightarrow{\sim} \{\text{isom. cl. of integral Hodge structures of elliptic type with level-}N \text{ structure}\}.\end{aligned}$$

Fact 9.1.3. *Suppose k is an algebraically closed field of characteristic 0 and $f : X \to Y$ is a morphism of k-varieties. If Y is normal and f is a bijection on k-points, then f is an isomorphism.*

This fact implies that $f : S(N)_{\mathbb{C}} \to Y(N)$ is an isomorphism of $\mathbb{C}$-varieties, since the target is smooth and *a fortiori* normal. □

9.2. **The representability of $S(N)$.** We will discuss the strategy first. Suppose that S is a scheme and E/S is an elliptic curve. By the Riemann–Roch Theorem, we will be able to construct "meromorphic functions with controlled poles". From this we will attain Weierstrass coordinates on E locally on S, i.e., after shrinking S, we will find a closed S-embedding

$$E \longrightarrow \mathbb{P}^2_S$$

given by a Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

(The indexing is such that every term $a_ix^jy^k$ satisfies $i+2j+3k = 6$.) The level-N structure on E, which is used to rigidify the abstract elliptic curve, will serve to rigidify the concrete

[4] Here and below, by a variety over an algebraically closed field k we mean a reduced finite-type k-scheme such that each connected component is irreducible. Thus we allow a finite disjoint union of what are usually called k-varieties.

Weierstrass equation, in the sense that ambiguous choices of a_i for the same elliptic curve will be rigidified. Then, we will construct $S(N)$ as the spectrum of $\mathbb{Z}[1/N][a_i]$ modulo some relations.

In fact, we will only be able to do this explicitly for small values of N. We will do this for $N = 3$ and $N = 4$, and then bootstrap the general case from $S(3)$ and $S(4)$. Namely, in the general case, this is done after passing through the open cover $\{\operatorname{Spec}\mathbb{Z}[1/N]\cap \operatorname{Spec}\mathbb{Z}[1/3],\ \operatorname{Spec}\mathbb{Z}[1/N]\cap\operatorname{Spec}\mathbb{Z}[1/4]\}$ of $\operatorname{Spec}\mathbb{Z}[1/N]$.

Remark. In [KM85], instead of a construction of $S(3)$ and $S(4)$, they construct $S(3)$ and a rigidified version of $S(2)$. (Note that elliptic curves with level-2 structure always have automorphism group $\mathbb{Z}/2\mathbb{Z}$, the non-trivial element being negation.) The construction for the rigidified version of $S(2)$ in fact contains a mistake, as pointed out in [Conc, §2]. As a correction, one should just construct $S(3)$ and $S(4)$ instead. We will only explain the construction of $S(3)$. **The construction of $S(4)$ as well as the explanation of the mistake in [KM85] are left as a presentation topic.** The main reference is [Conc].

9.3. **Generalities on relative curves.** We collect some general facts about relative curves. Suppose S is a locally noetherian scheme. Suppose $f : X \to S$ is a proper, smooth morphism of schemes whose geometric fibers are all connected of dimension 1. We also assume that f admits a section $P : S \to X$, which we fix. We note that P is always a closed immersion. We write $\mathcal{I}_P$ for the ideal sheaf of $\operatorname{im}(P)$ in X.

Facts:

(1) The $\mathcal{O}_X$-module $\mathcal{I}_P$ is invertible. More precisely for all $s \in S$, after shrinking S near s, we can find an open neighborhood U of $\operatorname{im}(P)$ in X and $t \in \mathcal{O}_U(U)$ such that $\mathcal{I}_P|_U = t\cdot\mathcal{O}_U$ and $t : \mathcal{O}_U \to \mathcal{O}_U$ is injective. This implies that $\mathcal{I}_P$ is invertible. We shall call t a **local coordinate near** P. When U is fixed, the choice of t is unique up to multiplication by $\mathcal{O}_U(U)^\times$.

Let n be a positive integer. We will write $\mathcal{O}(nP)$ for $\mathcal{I}_P^{\otimes -n}$, a line bundle on X. This is the "sheaf of functions that are allowed to have a pole of order n along P". After choosing a local coordinate as above and shrinking U around $\operatorname{im}(P)$ in X, we get an element $t^{-1} \in \mathcal{O}(P)(U)$ such that $\mathcal{O}(P) = t^{-1}\cdot\mathcal{O}_U$ and $\mathcal{O}(nP) = t^{-n}\mathcal{O}_U$ for all $n \in \mathbb{Z}$.

(2) We have a natural map of $\mathcal{O}_S$-modules "taking the $-n$-th coefficient of the Laurent expansion at P"

$$\mathrm{Lead}_n : f_*\mathcal{O}(nP) \longrightarrow P^*(\Omega_{X/S})^{\otimes -n}.$$

Here we note that $\Omega_{X/S}$ is a line bundle on X thanks to our assumptions, so $P^*\Omega_{X/S}$ is a line bundle on S and its negative tensor powers are defined.

10. Lecture 10

10.1. **Generalities on relative curves, continued.** We maintain the same assumptions as the previous section. Namely, suppose S is a locally noetherian scheme. Fix $f : X \to S$ to be a proper smooth morphism of schemes with all geometric fibers connected of dimension 1. Fix a section $P \in X(S)$. We will describe the morphism of $\mathcal{O}_S$-modules

$$\mathrm{Lead}_n : f_*\mathcal{O}(nP) \longrightarrow P^*(\Omega_{X/S})^{\otimes -n}.$$

First, we will provide an abstract description that does not rely on a choice of a local coordinate. Note that by adjunction, there is a natural map

$$f_*\mathcal{O}(nP) \longrightarrow f_*P_*P^*\mathcal{O}(nP) = P^*\mathcal{O}(nP).$$

So it suffices to define $P^*\mathcal{O}(nP) \to P^*(\Omega_{X/S})^{\otimes -n}$. Observe that we have the following string of isomorphisms of $\mathcal{O}_X/\mathcal{I}_P = \mathcal{O}_S$-modules:

$$P^*\mathcal{O}(nP) = P^*(\mathcal{I}_P^{-n}) = \mathcal{I}_P^{-n} \otimes_{\mathcal{O}_X} \mathcal{O}_X/\mathcal{I}_P = \mathcal{I}_P^{-n}/\mathcal{I}_P^{1-n} = (\mathcal{I}_P^{-1}/\mathcal{O}_X)^{\otimes n}.$$

(The last n-th tensor power is over $\mathcal{O}_S$.)

It suffices to define an $\mathcal{O}_S$-module map

$$\mathcal{I}_P^{-1}/\mathcal{O}_X \longrightarrow P^*(\Omega_{X/S})^{\otimes -1},$$

or equivalently, an $\mathcal{O}_S$-bilinear pairing

$$\mathcal{I}_P^{-1}/\mathcal{O}_X \otimes_{\mathcal{O}_S} P^*\Omega_{X/S} \longrightarrow \mathcal{O}_S.$$

Now as a general fact (see [Sta18, Tag 0474]), $P^*\Omega_{X/S}$ is canonically identified with the **conormal sheaf** of $\mathrm{im}(P)$ in X, namely $\mathcal{I}_P/\mathcal{I}_P^2$ viewed as an $\mathcal{O}_X/\mathcal{I}_P = \mathcal{O}_S$-module. The identification $\mathcal{I}_P/\mathcal{I}_P^2 \xrightarrow{\sim} P^*\Omega_{X/S}$ is given by $a \mapsto da$. Then, we define the pairing

$$\mathcal{I}_P^{-1}/\mathcal{O}_X \otimes_{\mathcal{O}_S} \mathcal{I}_P/\mathcal{I}_P^2 \longrightarrow \mathcal{O}_S$$

by multiplication followed by reduction modulo $\mathcal{I}_P$.

For concreteness, we now also describe our morphism Lead_n explicitly after choosing a local coordinate. Thus assume we have an open neighborhood U of $\mathrm{im}(P)$ in X and a local coordinate $t \in \mathcal{O}_U(U)$ such that $t : \mathcal{O}_U \to \mathcal{O}_U$ is injective and $\mathcal{I}_P|_U = t \cdot \mathcal{O}_U$ as before. Then, we have

$$f_*\mathcal{O}(nP) = (\mathcal{I}_P^{-1}/\mathcal{O}_X)^{\otimes n} = (t^{-1}\mathcal{O}_U/\mathcal{O}_U)^{\otimes n}$$

viewed as $\mathcal{O}_U/t\mathcal{O}_U = \mathcal{O}_S$-modules. Also, we have

$$P^*\Omega_{X/S} = \mathcal{I}_P/\mathcal{I}_P^2 = t\mathcal{O}_U/t^2\mathcal{O}_U = (\mathcal{O}_U/t\mathcal{O}_U) \cdot dt|_{t=0} = \mathcal{O}_S \cdot dt|_{t=0}.$$

Here the symbol $dt|_{t=0}$ denotes the image of t in $t\mathcal{O}_U/t^2\mathcal{O}_U$, and $\mathcal{O}_S \cdot dt|_{t=0}$ is a rank 1 free $\mathcal{O}_S$-module generated by $dt|_{t=0}$. Finally, our pairing for $n = 1$ can be described as

$$(t^{-1}\mathcal{O}_U/\mathcal{O}_U) \otimes_{\mathcal{O}_S} \mathcal{O}_S \cdot dt|_{t=0} \longrightarrow \mathcal{O}_S$$

via

$$(t^{-1}\zeta, \epsilon\ dt) \longmapsto \zeta\epsilon \mod t.$$

Remark. For more clarity, we can ask what the picture looks like in the classical case. Suppose $S = \operatorname{Spec} k$ where k is an algebraically closed field. Then $f_*\mathcal{O}(nP) = H^0(X, \mathcal{O}(nP))$ is the k-vector space consisting of $\zeta \in k(X)$ allowed to have a pole of at worst order n at P and regular at all other points. On the other side, we have $(P^*\Omega_{X/S})^{\otimes -n} = (T_P^*X)^{\otimes_k -n}$, where T_P^*X is the cotangent space of X at P. After choosing a local coordinate t near P, we have $\widehat{\mathcal{O}}_{X,P} = k[[t]]$. We have the composition

$$f_*\mathcal{O}(nP) \hookrightarrow k(X) \to \operatorname{Frac} \widehat{\mathcal{O}}_{X,P} = k((t))$$

which we think of as taking the Laurent expansion of a function $\zeta \in f_*\mathcal{O}(nP)$ at P. Also, T_P^*X is a 1-dimensional k-vector space with a basis $dt|_{t=0}$ arising from the choice of t. Then the pairing is defined by sending $(\zeta, dt^{\otimes n})$ to the $-n$-th coefficient of the Laurent expansion of ζ at P.

10.2. **More generalities: cohomology and base change.** Suppose $f : X \to S$ is an arbitrary proper morphism of schemes, with S locally noetherian. Suppose $\mathcal{F}$ is a coherent sheaf of $\mathcal{O}_X$-modules flat over $\mathcal{O}_S$, i.e., for all $x \in X$, the stalk $\mathcal{F}_x$ (which is an $\mathcal{O}_{X,x}$-module) is flat over $\mathcal{O}_{S,f(x)}$. For each $s \in S$, we write X_s for the fiber of X_s over s, namely the fiber product

$$\begin{tikzcd} X_s \arrow[r] \arrow[d] & X \arrow[d] \\ \operatorname{Spec} k(s) \arrow[r, "s"'] & S, \end{tikzcd}$$

Write $\mathcal{F}|_{X_s}$ for the pullback of $\mathcal{F}$ along $X_s \to X$.

Theorem 10.2.1. *The following statements hold.*

(1) *(See* [MFK94, §0.5] *or* [Conb, Cor. 1.2, Prop. 2.1]*.) Suppose* $\mathbf{H}^1(X_s, \mathcal{F}|_{X_s}) = 0$ *for all* $s \in S$*. Then* $f_*\mathcal{F}$ *is a vector bundle over* S*, and* $R^1 f_* \mathcal{F} = 0$*. Moreover, the formation of* $f_*\mathcal{F}$ *commutes with arbitrary base change in the following sense. Suppose* S' *is locally noetherian and* $g : S' \to S$ *is an arbitrary morphism. Define* $f' : X' \to S'$ *to be the pullback of* f *along* g*, i.e., we have the cartesian diagram*

$$\begin{tikzcd} X' \arrow[r] \arrow[d, "f'"'] & X \arrow[d, "f"] \\ S' \arrow[r, "g"'] & S. \end{tikzcd}$$

Let $\mathcal{F}'$ *be the pullback of* $\mathcal{F}$ *to* X'*. Then the natural map (the "base change map") of* $\mathcal{O}_{S'}$*-modules*

$$g^* f_* \mathcal{F} \longrightarrow f'_* \mathcal{F}'$$

is an isomorphism.

As a special case, we can take g *to be the map* $\operatorname{Spec}(k(s)) \to S$ *defined by a point* $s \in S$*. Then the natural map of* $k(s)$*-vector spaces*

$$(f_*\mathcal{F}) \otimes_{\mathcal{O}_S} k(s) \longrightarrow \mathbf{H}^0(X_s, \mathcal{F}|_{X_s})$$

is an isomorphism. In particular, this means that any $k(s)$*-basis of* $\mathbf{H}^0(X_s, \mathcal{F}|_{X_s})$ *can be lifted to a trivialization of the vector bundle* $f_*\mathcal{F}$ *near* $s \in S$*.*

(2) *(See* [Conb, Thm. 1.1, Prop. 2.1]*.) Suppose for each* $s \in S$*, the natural map*

$$(f_*\mathcal{F}) \otimes_{\mathcal{O}_S} k(s) \longrightarrow \mathbf{H}^0(X_s, \mathcal{F}|_{X_s})$$

is surjective. Then the map is an isomorphism for each $s \in S$*. Moreover,* $f_*\mathcal{F}$ *is a vector bundle, and the formation of* $f_*\mathcal{F}$ *commutes with arbitrary base change.*

11. Lecture 11

11.1. **Application: pushforward of the structure sheaf.** As an application of Theorem 10.2.1 (2) above, we have the following useful result (cf. [Conb, Cor. 1.3]).

Proposition 11.1.1. *Let* S *be a locally noetherian scheme. Let* $f : X \to S$ *be a proper, flat, surjective morphism such that all its geometric fibers are connected and reduced. Then the natual map* $\mathcal{O}_S \to f_*\mathcal{O}_X$ *is an isomorphism.*

Proof. Since X is flat over S, $\mathcal{O}_X$ is flat over S. Now for each $s \in S$ the natural map

$$(f_*\mathcal{O}_X) \otimes_{\mathcal{O}_S} k(s) \longrightarrow \mathbf{H}^0(X_s, \mathcal{O}_{X_s})$$

is surjective since it is non-zero (as $1 \mapsto 1$) and the right hand side is 1-dimensional. Thus $\mathcal{O}_X$ satisfies the assumptions in Theorem 10.2.1 (2) . We conclude that $f_*\mathcal{O}_X$ is a vector bundle and its formation commutes with base change. Now since $\mathcal{O}_S \to f_*\mathcal{O}_X$ is a map between vector bundles, in order to check that it is an isomorphism it suffices to check that the induced maps $\mathcal{O}_S \otimes_{\mathcal{O}_S} k(s) \to (f_*\mathcal{O}_X) \otimes_{\mathcal{O}_S} k(s)$ are isomorphisms, for all $s \in S$. But the two sides are 1-dimensional $k(s)$-vector spaces and the map is non-zero, so it is an isomorphism. □

11.2. **Riemann–Roch for a relative elliptic curve.** Let S be a locally noetherian scheme, and $f : E \to S$ an elliptic curve with identity section e. Recall that this means that f is a proper smooth morphism with all geometric fibers being connected smooth projective curves of genus 1, and $e : S \to E$ is a distinguished section of f. Recall from before that for every positive integer n we have a line bundle $\mathcal{O}(ne)$ on E, as well as a map of $\mathcal{O}_S$-modules $\mathrm{Lead}_n : f_*\mathcal{O}(ne) \to e^*(\Omega_{E/S})^{\otimes -n}$.

Theorem 11.2.1. *Let n be a positive integer. The following statements hold.*

(1) *For each $s \in S$, $\mathbf{H}^1(E_s, \mathcal{O}(ne)|_{E_s}) = 0$, and $\mathbf{H}^0(E_s, \mathcal{O}(ne)|_{E_s})$ has dimension n over $k(s)$.*
(2) *The $\mathcal{O}_S$-module $f_*\mathcal{O}(ne)$ is a vector bundle of rank n, and its formation commutes with arbitrary base change.*
(3) *The composition of the natural maps of $\mathcal{O}_S$-modules $\mathcal{O}_S \to f_*\mathcal{O}_X \to f_*\mathcal{O}(e)$ is an isomorphism.*
(4) *The natural complex of $\mathcal{O}_S$-modules*

$$0 \to f_*\mathcal{O}(ne) \to f_*\mathcal{O}((n+1)e) \xrightarrow{\mathrm{Lead}_{n+1}} e^*(\Omega_{E/S})^{\otimes -(n+1)} \to 0$$

is exact.

Proof. **(1)** Note that $\mathcal{O}(ne)|_{E_s} = \mathcal{O}_{E_s}(ne_s)$, where e_s is the identity section of the elliptic curve E_s over $k(s)$. Hence the statement follows from the following special case of Riemann–Roch: Let X be a smooth projective curve over a field k (not necessarily algebraically closed) that is geometrically connected and has genus g. Let $P \in X(k)$. Then for all $n > 2g-2$, we have $\mathbf{H}^1(X, \mathcal{O}_X(nP)) = 0$ and $\dim_k \mathbf{H}^0(X, \mathcal{O}_X(nP)) = 1 - g + n$.

(2) This follows from part (1) and Theorem 10.2.1 (1). (Here, since $\mathcal{O}(ne)$ is a line bundle on E and since E is flat over S, we know that $\mathcal{O}(ne)$ is indeed flat over S.)

(3) Both $\mathcal{O}_S$ and $f_*\mathcal{O}(e)$ are vector bundles whose formation commutes with base change. Thus we reduce to the case where S is the spectrum of a field. Then the statement is classical.

(4) Note the following (easy) fact: Suppose R is a local ring with residue field k, and $0 \to A \to B \to C \to 0$ is a complex of finite free R-modules. Then this complex is exact if and only if $0 \to A \otimes k \to B \otimes k \to C \otimes k \to 0$ is exact.

Now all three terms in the complex in question are vector bundles over S, by part (2) and by the fact that $\Omega_{E/S}$ is a line bundle on E. Using the above paragraph, we see that we only need to check that the complex in question becomes exact after $\otimes_{\mathcal{O}_S} k(s)$ for each $s \in S$. But the formation of the three terms commutes with base change (by part (2), and by the functoriality of $\Omega_{E/S}$). Hence we reduce to the case where S is the spectrum of a field. Then the assertion follows from the description of Lead_{n+1} as taking the $-(n+1)$-th coefficient of the Laurent expansion after choosing a local coordinate around e.

□

11.3. **Local Weierstrass coordinates.** Keep the above notation. After replacing S by a Zariski open neighborhood of an arbitrary point, we may assume that $S = \operatorname{Spec} R$ and that the line bundle $e^*\Omega_{E/S}$ on S is trivial. Fix a trivialization $e^*(\Omega_{E/S})^{\otimes -1} \xrightarrow{\sim} \mathcal{O}_S$, or in other words a basis ω of the rank 1 free R-module $e^*(\Omega_{E/S})^{\otimes -1}(S)$. Then for each positive integer n, we have a basis ω^n for the rank 1 free R-module $e^*(\Omega_{E/S})^{\otimes -n}(S)$. Write $H(n)$ for the R-module $(f_*\mathcal{O}(ne))(S)$. We have a short exact sequence of R-modules

$$0 \to H(n) \to H(n+1) \xrightarrow{\mathrm{Lead}_{n+1}} R\omega^{n+1} \to 0.$$

Now $H(1)$ is canonically identified with $\mathcal{O}_S(S) = R$ (see Theorem 11.2.1 (3)), and $R\omega^{n+1}$ is always free of rank 1. Thus by induction we know that $H(n)$ is a free R-module of rank n for each $n \geq 1$. Moreover, the basis $\{1\}$ of $H(1) = R$ can be extended to a basis $\{1, x\}$ of $H(2)$ with

$$\mathrm{Lead}_2(x) = \omega^2,$$

and this can be further extended to a basis $\{1, x, y\}$ of $H(3)$ with

$$\mathrm{Lead}_3(y) = \omega^3.$$

Now using that $\mathrm{Lead}_4(x^2) = \omega^4$, we see that $\{1, x, y, x^2\}$ is a basis of $H(4)$. Similarly, $\{1, x, y, x^2, xy\}$ is a basis of $H(5)$ since $\mathrm{Lead}_5(xy) = \omega^5$; and $\{1, x, y, x^2, xy, x^3\}$ is a basis of $H(6)$ since $\mathrm{Lead}_6(x^3) = \omega^6$. Now $y^2 \in H(6)$, so

$$y^2 = Ax^3 - a_1xy + a_2x^2 - a_3y + a_4x + a_6$$

for unique $A, a_1, a_2, a_3, a_4, a_6 \in R$. Comparing Lead_6 of both sides we see that $A = 1$. Thus we conclude that x and y satisfy the Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

for unique $a_i \in R$.

12. Lecture 12

12.1. **Local Weierstrass coordinates, continued.** We have found a basis $\{1, x\}$ of $H(2) = f_*(\mathcal{O}(2e))(S)$ and a basis $\{1, x, y\}$ of $H(3) = f_*(\mathcal{O}(3e))(S)$ such that $\mathrm{Lead}_2(x) = \omega^2$ and $\mathrm{Lead}_3(y) = \omega^3$. We showed that x and y satisfy the Weierstrass equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

for unique $a_i \in R$. We shall call such x and y a choice of **Weierstrass coordinates**.

Now let $\mathcal{W}$ be the closed subscheme of $\mathbb{P}^2_S = \mathbb{P}^2_R$ defined by the (homogenization of the) above equation (with point at infinity $(0 : 1 : 0)$). Then the S-morphism $E \to \mathbb{P}^2_S$ defined by the basis $(x, y, 1)$ of $H(3)$ factors through an S-morphism

$$\phi : E \longrightarrow \mathcal{W}.$$

(Here the identity section e goes to the constant section $(0 : 1 : 0)$.) We claim that ϕ is an isomorphism. For this, we use the fibral criterion for isomorphism:

Fact: Let X, Y be two S-schemes that are locally of finite type and flat over S. Let $\phi : X \to Y$ be an S-morphism that is of finite type and separated. Then ϕ is an isomorphism if and only if $\phi_s : X_s \to Y_s$ is an isomorphism for each $s \in S$.

In order to apply this criterion to $\phi : E \to \mathcal{W}$, we need to know that $\mathcal{W}$ is flat over S, but this follows from the general fact that any hypersurface in $\mathbb{P}^n_R$ defined by a single homogeneous equation whose coefficients generate the unit ideal is flat over $\operatorname{Spec} R$. It is also not hard to check that $\phi : E \to \mathcal{W}$ is of finite type and separated. Thus we reduce to checking that $\phi_s : E_s \to \mathcal{W}_s$ is an isomorphism for all $s \in S$. But the formation of ϕ

commutes with base change, so we reduce to the case where S is the spectrum of a field. Then the claim is classical; see [Sil09, §III, Prop. 3.1].

Remark. It follows from our claim that $\mathcal{W}$ is smooth over S. In particular, the discriminant Δ of the Weierstrass equation (which is a universal polynomial in the variables a_i with integer coefficients, and homogeneous of degree 12 if a_i is assigned weight i; see [Sil09, §III.1]) is non-zero in $k(s)$ for all $s \in S$[5], and therefore $\Delta \in R^\times$.

Now let us analyze the uniqueness of the Weierstrass coordinates x, y. If we fix the basis ω of $e^*(\Omega_{E/S})^{\otimes -1}(S)$ fixed, then x and y are unique up to the transformation

$$\begin{cases} x \longmapsto x + a \\ y \longmapsto y + bx + c. \end{cases}$$

by the defining properties of x and y. On the other hand ω is unique up to $\omega \mapsto u\omega$ for $u \in R^\times$, and this can be matched by the transformation

$$\begin{cases} x \longmapsto u^2 x \\ y \longmapsto u^3 y. \end{cases}$$

Thus the group of all admissible transformations is generated by the above two types of transformations. One sees that a general transformation is of the form

$$\begin{cases} x \mapsto u^2 x + a \\ y \mapsto u^3 y + u^2 bx + c. \end{cases}$$

for $a, b, c \in R$ and $u \in R^\times$.

Thus we have shown that given any elliptic curve E/S, locally on S we can identify E with the planar curve defined by a Weierstrass equation, and moreover the identification is unique up to the effect of the above transformation group. We might imagine that the "moduli space of elliptic curves" should be given by "Spec $\mathbb{Z}[a_1, a_2, a_3, a_4, a_6]$ modulo the action of the transformation group". Unfortunately this does not work in the category of schemes. In the next few lectures, we will show that a level-3 structure can help to rigidify the Weierstrass equation to the extent that none of the above transformations are allowed. Then we will be able to prove the representability of $S(3)$.

13. Lecture 13

13.1. **Construction of $S(3)$.** As before, let $f : E \to S$ be an elliptic curve, with identity section $e : S \to E$. Recall that if $S = \operatorname{Spec} R$ and $e^*\Omega_{E/S} \cong \mathcal{O}_S$, then we can find Weierstrass coordinates x and y on E such that E is the relative curve in $\mathbb{P}^2_S$ defined by

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6$$

for unique $a_i \in R$. Here e corresponds to the homogeneous coordinate $(0 : 1 : 0)$. Moreover, the x and y are unique up to transformation

$$\begin{cases} x' = u^2 x + a \\ y' = u^3 y + bu^2 x + c \end{cases}, \qquad a, b, c \in R, \qquad u \in R^\times.$$

[5]This is because over a field, a Weierstrass equation defines a non-singular curve if and only if $\Delta \neq 0$. In fact, in the classical proof ([Sil09, §III, Prop. 3.1]) that our $\phi : E \to \mathcal{W}$ is an isomorphism when S is the spectrum of a field, it is shown *a priori* that the Weierstrass equation must be non-singular.

We will explain how to use a full level-3 structure to rigidify the situation, which will eventually lead to the construction of the moduli space $S(3)$. Our explanation will be informal to begin with.

Suppose $S = \operatorname{Spec} k$ where k is a field with $\operatorname{char} k \neq 3$. Let (P, Q) be a level 3 structure, namely, $P, Q \in E[3](k)$ such that they form an $\mathbb{F}_3$-basis of

$$E[3](\bar{k}) \cong \mathbb{F}_3^2.$$

Note in particular that this implies that $E[3](k) = E[3](\bar{k})$. Then we take Weierstrass coordinates x and y in the classical sense such that E is defined by $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ where $a_i \in k$.

Now, the fact that $3P = e$ implies that $3[P] - 3[e]$ is a principal divisor, i.e., there exists $\varphi \in k(E)$ such that $\operatorname{div}(\varphi) = 3[P] - 3[e]$. Note that we have $\varphi \in H(3) \setminus H(2)$ where $H(n) := (f_*\mathcal{O}(ne))(k) = H^0(E, \mathcal{O}(ne))$. We may assume that $\operatorname{Lead}_3(\varphi) = \omega^3$ where ω is the fixed basis of $e^*(\Omega_{E/S})^{\otimes -1}(k)$. Hence we could have chosen y to be φ. Namely, we may assume $\operatorname{div}(y) = 3[P] - 3[e]$. So we have $y(P) = 0$. Now $x(P)$ may not be zero, but we can always replace x by $x - x(P)$. In summary, we may assume $P = (0, 0)$.

Next we discuss negation on E. Take $T \in E(k)$. By the definition of the group law, the three points $T, -T$, and the "point at infinity" $e = (0 : 1 : 0) \in \mathbb{P}^2(k)$ are colinear. Thus the x-coordinate of $-T$ is the same as that of T. Now if y_1, y_2 are the two roots of the equation $y^2 + a_1x_0y + a_3y = x_0^3 + a_2x_0^2 + a_4x_0 + a_6$ for some fixed x_0, then $y_1 + y_2 = -(a_1x_0 + a_3)$ by Vieta theorem. Hence if T has coordinate (x, y) then $-T$ has coordinate $(x, -y - a_1x - a_3)$. Now applying this to $P = (0, 0)$, we get $-P = (0, -a_3)$. Since $-P \neq P$ (as P is a non-trivial 3-torsion point), we have

$$a_3 \neq 0.$$

Now recall that $x \in H(2) \setminus H(1)$, which means that the only pole that x has is a pole of order 2 at e. We have seen that both P and $-P$ has zero x-coordinate, so x has at least a zero at P and at least a zero at $-P$. Therefore $\operatorname{div}(x) = [P] + [-P] - 2[e]$. Now recall that $\operatorname{div}(y) = 3[P] - 3[e]$. From here, we see that

$$\operatorname{ord}_P(y^2 + a_1xy + a_3y) \geq 3.$$

We also see that $\operatorname{ord}_P(x^3) = 3$. Hence $\operatorname{ord}_P(a_2x^2 + a_4x + a_6) \geq 3$. But we know that $\operatorname{ord}_P(x) = 1$. This implies that $a_2 = a_4 = a_6 = 0$. Therefore the Weierstrass equation is forced to be of the form

$$y^2 + a_1xy + a_3y = x^3$$

where $a_3 \neq 0$ and $\Delta \neq 0$. Here Δ is the discriminant $\Delta = \Delta(a_1, a_3) = a_1^3a_3^3 - 27a_3^4$ (see [Sil09, §III.1]). Note that the only allowed transformations of coordinates are

$$\begin{cases} y' = u^3y \\ x' = u^2x \end{cases}, \qquad u \in R^\times$$

by easy considerations on the vanishing orders of x and y at P.

Now, we consider Q. There exists unique $A, B \in k$ such that

$$\operatorname{div}(y - Ax - B) = 3[Q] - 3[e]$$

because the right hand side is a principal divisor and has exactly a pole of order 3 at e. Now, we consider the system of equations

$$\begin{cases} y = Ax + B \\ y^2 + a_1xy + a_3y = x^3 \end{cases}.$$

After substituting the first equation into the second equation, we obtain

$$x^3 - (Ax+B)^2 - a_1x(Ax+B) - a_3(Ax+B) = 0.$$

The fact that $y - Ax - B$ has a triple zero at Q precisely means that the above equation has a triple zero at $x = x(Q)$. Hence we have

$$x^3 - (Ax+B)^2 - a_1x(Ax+B) - a_3(Ax+B) = (x - x(Q))^3.$$

By comparing coefficients, we will see that a_1, a_3 are related to A and B.

14. Lecture 14

14.1. **Construction of $S(3)$, continued.** Recall our setup from before. Fix an elliptic curve E over a field k with $\operatorname{char} k \neq 3$. We also fix a level-3 structure (P, Q). By choosing Weierstrass coordinates x and y such that $P = (0,0)$ and $\operatorname{div}(y) = 3[P] - 3[e]$, we can deduce that the Weierstrass equation is of the form

$$y^2 + a_1xy + a_3y = x^3.$$

We also know that the discriminant $\Delta = \Delta(a_1, a_3) := a_1^3a_3^3 - 27a_3^4$ is non-zero, which implicitly implies $a_3 \neq 0$.

Now we consider Q. Since Q is also 3-torsion and $Q \neq e$, we know that $3[Q] - 3[e]$ is principal divisor. Hence we can find unique $A, B \in k$ such that $\operatorname{div}(y - Ax - B) = 3[Q] - 3[e]$. The fact that $y - Ax - B$ has a triple zero at Q precisely means the following identity

$$x^3 - (Ax+B)^2 - a_1x(Ax+B) - a_3(Ax+B) = (x - x(Q))^3. \tag{14.1}$$

Note that different powers of x are linearly independent over k, so we can compare coefficients in (14.1) to get relations among a_1, a_3, A, B. In particular, for the quadratic coefficient, we see $-A^2 - a_1A = -3x(Q)$. Note also that $Q \neq \pm P$ since (P, Q) is a level 3 structure. This implies $x(Q) \neq 0$. So we have $A \neq 0$. Then by the change of coordinates

$$\begin{cases} x \mapsto A^2x \\ y \mapsto A^3y \end{cases}$$

we may assume that $A = 1$.

Now set $x(Q) =: C \neq 0$. Since $y - x - B$ vanishes at Q, we have $y(Q) = B + C$, i.e., $Q = (C, B + C)$. Comparing coefficients in (14.1) we get

$$\begin{cases} a_1 = 3C - 1, \\ a_3 = -3C^2 - B - 3BC, \\ B^3 = (B + C)^3. \end{cases}$$

Using the above, we think of $\Delta = \Delta(a_1, a_3)$ as a polynomial in B, C. In order to avoid conflict of notation we denote this polynomial by $\Delta_{B,C} \in \mathbb{Z}[B, C]$.

In conclusion, starting with (E, P, Q), we can choose Weierstrass coordinates x and y on E such that $P = (0,0)$ and $Q = (C, B + C)$, and such that the Weierstrass equation (uniquely determined after choosing x and y) is of the form

$$y^2 + a_1xy + a_3y = x^3$$

where

$$\begin{cases} a_1 = 3C - 1, \\ a_3 = -3C^2 - B - 3BC. \end{cases}$$

Moreover, $B, C \in k$ satisfy

$$\begin{cases} C \neq 0, \\ \Delta_{B,C} \neq 0, \\ B^3 = (B+C)^3. \end{cases} \tag{14.2}$$

(Recall that $C \neq 0$ comes from the condition $Q \notin \{\pm P, e\}$, and from $\Delta_{B,C} \neq 0$ we have $a_3 \neq 0$ which corresponds to $P \notin \{-P, e\}$.)

Conversely, suppose we start with $B, C \in k$ satisfying the conditions (14.2). We set $a_1 := 3C - 1$ and $a_3 := -3C^2 - B - 3BC$, and then define

$$\begin{aligned} E_{B,C} &:= \{y^2 + a_1xy + a_3y = x^3\} \subset \mathbb{P}^2_k, \\ P_{B,C} &:= (0,0), \\ Q_{B,C} &:= (C, B+C). \end{aligned}$$

Then $E_{B,C}$ is an elliptic curve over k and $(P_{B,C}, Q_{B,C})$ is a level-3 structure for $E_{B,C}$.

Also, for any given (E, P, Q), there is a unique choice of Weierstrass coordinates x and y rendering (E, P, Q) in the above standard form for unspecified B, C. Indeed, if x and y render (E, P, Q) in the standard form, then we can show that

$$\begin{aligned} \operatorname{div}(y) &= 3[P] - 3[e], \\ \operatorname{div}(x) &= [P] + [-P] - 2[e], \\ \operatorname{div}(y - x - B) &= 3[Q] - 3[e]. \end{aligned}$$

It is then easy to see that there is no non-trivial coordinate change

$$\begin{cases} x' = u^2x + a \\ y' = u^3y + bu^2x + c \end{cases}, \qquad a, b, c \in k, \qquad u \in k^\times$$

preserving these three conditions.

This uniqueness means that for any (E, P, Q), there exist *unique* B, C and a *unique* isomrophism $(E, P, Q) \xrightarrow{\sim} (E_{B,C}, P_{B,C}, Q_{B,C})$.

Now imagine that we may perform this construction in the relatively setting, i.e., we can show that for all elliptic curves E/S where S is defined over $\mathbb{Z}[1/3]$, after localizing S (i.e., passing to a Zariski open covering), there exist unique Weierstrass coordinates x and y rendering (E, P, Q) in the standard form as above. In particular, by uniqueness, we see that passing to a Zariski open covering of S is not necessary, because the local Weierstrass coordinates and the local sections B, C must be compatible on the overlaps of the open covering. Thus, for all locally noetherian $S/\mathbb{Z}[1/3]$ (not necessarily affine) and any E/S with level-3 structure (P, Q), there exist unique $B, C \in \mathcal{O}_S(S)$ and a unique isomorphism from (E, P, Q) to the standard $(E_{B,C}, P_{B,C}, Q_{B,C})$ inside $\mathbb{P}^2_S$. We may state this even more precisely as follows.

Theorem 14.1.1. *The functor*

$$S(3) : (\textit{locally noetherian schemes over } \mathbb{Z}[1/3]) \longrightarrow (\textit{Sets})$$

sending S to the set of isomorphism classes of elliptic curves E/S with level-3 structure (P, Q), is represented by the $\mathbb{Z}[1/3]$-scheme

$$\operatorname{Spec} \mathbb{Z}[1/3, B, C, C^{-1}, \Delta_{B,C}^{-1}]/(B^3 - (B+C)^3).$$

The universal object is given by $(E_{B,C}, P_{B,C}, Q_{B,C})$.

The proof roughly follows the same ideas as our discussion over a field. For the complete rigorous proof, see [Conc, §4]. There are two interesting points in the proof which we intend to elaborate on:

(1) Assume $S = \operatorname{Spec} R$ and $e^*\Omega_{E/S}$ is trivial, so Weierstrass coordinates exist. Up to further shrinking S, for any fixed Weierstrass coordinates x and y, there are unique $a, b \in R$ such that the suitable analog of the statement "$\operatorname{div}(y+ax+b) = 3[P]-3[e]$" holds. Thus we can replace y by $y + ax + b$.

(2) In the same setting as (1), the morphism $[-1] : E \to E$ is still given by $(x : y : z) \mapsto (x : -y - a_1x - a_3 : z)$, as in the case over a field.

15. Lecture 15

15.1. **Construction of $S(3)$, technical details.** Recall that last lecture, we informally showed that the functor

$$S(3) : (\text{locally noetherian schemes over } \mathbb{Z}[1/3]) \longrightarrow (\text{Sets})$$

sending S to the set of isomorphism classes of elliptic curves E/S with level-3 structure (P, Q), is represented by the $\mathbb{Z}[1/3]$-scheme

$$\operatorname{Spec} \mathbb{Z}[1/3, B, C, C^{-1}, \Delta_{B,C}^{-1}]/(B^3 - (B + C)^3).$$

The universal object is given by $(E_{B,C}, P_{B,C} = (0, 0), Q_{B,C} = (C, B + C))$.

Remark. The scheme $S(3)$ is affine and smooth over $\mathbb{Z}[1/3]$, with fibers of pure dimension 1.

We now elaborate on the two technical details stated at the end of last lecture. In this lecture we discuss (1). Suppose we have Weierstrass coordinates x, y, adapted to the choice of an R-basis ω of $e^*(\Omega_{E/S})^{\otimes -1}(S)$ (so $\operatorname{Lead}_2(x) = \omega^2$ and $\operatorname{Lead}_3(y) = \omega^3$). Recall that for each $p \in E(S)$, we have the ideal sheaf for the closed subscheme $p(S) \subset E$, denoted $\mathcal{I}_p \subset \mathcal{O}_E$. This is an invertible $\mathcal{O}_E$-module, and we have the notation $\mathcal{O}(np) := \mathcal{I}_p^{\otimes -n}$. Recall that $\{1, x\}$ is an R-basis of $(f_*\mathcal{O}(2e))(S) = \mathcal{O}(2e)(E) = \mathcal{I}_e^{\otimes -2}(E)$ such that $\operatorname{Lead}_2(x) = \omega^2$ where

$$\operatorname{Lead}_2 : (f_*\mathcal{O}(2e))(S) \longrightarrow (e^*\Omega_{E/S})^{\otimes -2}(S) \cong R \cdot \omega^2.$$

Also, $\{1, x, y\}$ is an R-basis of $\mathcal{I}_e^{\otimes -3}(E)$ such that $\operatorname{Lead}_3(y) = \omega^3$ where

$$\operatorname{Lead}_3 : \mathcal{I}_e^{\otimes -3}(E) \longrightarrow R \cdot \omega^3.$$

Note that there is a natural injective map of $\mathcal{O}_E$-modules

$$\mathcal{I}_p^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3} \hookrightarrow \mathcal{I}_e^{\otimes -3}$$

as $\mathcal{I}_p^{\otimes 3} = \mathcal{I}_p^3$ is an ideal sheaf of $\mathcal{O}_E$. So there is a natural $\mathcal{O}_S$-module embedding

$$f_*(\mathcal{I}_p^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3}) \hookrightarrow f_*(\mathcal{I}_e^{\otimes -3}).$$

The desired generalization of the statement "$\exists!\ a, b$ such that $\operatorname{div}(y + ax + b) = 3[P] - 3[e]$" which makes sense over a field is the following statement:

- (After further shrinking S if necessary), $\exists!\ a, b \in R$ such that the global section $y + ax + b$ of $f_*(\mathcal{I}_e^{\otimes -3})$ generates the $\mathcal{O}_S$-submodule $f_*(\mathcal{I}_p^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3})$. (Since S is affine, this is equivalent to requiring that $y+ax+b$ is an R-basis of the R-submodule $f_*(\mathcal{I}_p^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3})(S) \subset f_*(\mathcal{I}_e^{\otimes -3})(S)$.)

We now prove this statement. Recall that the group law on E is given such that for three points $p, q, r \in E(S)$ we have $p + q = r$ if and only if

$$\mathcal{I}_p \otimes \mathcal{I}_q \otimes \mathcal{I}_e^{\otimes -2} \cong \mathcal{I}_r \otimes \mathcal{I}_e^{\otimes -1} \otimes f^*\mathcal{L}$$

for some line bundle $\mathcal{L}$ on S. Since $P \in E[3]$, we have $\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3} \cong f^*\mathcal{L}$ for some line bundle $\mathcal{L}$ on S. After shrinking S, we may assume that $\mathcal{L}$ is trivial. In particular, we may ensure that $\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3} \cong \mathcal{O}_E$ (non-canonically).

Now recall that since E/S is proper and flat, with reduced and connected geometric fibers, we know that $f_*\mathcal{O}_E \cong \mathcal{O}_S$ canonically. So we know that non-canonically, we have

$$f_*(\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3}) \cong \mathcal{O}_S,$$

i.e., $f_*(\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3})$ is a (trivial) line bundle on S.

Now consider the composition

$$\varphi_{E,P} : f_*(\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3}) \hookrightarrow f_*(\mathcal{I}_e^{\otimes -3}) \xrightarrow{\mathrm{Lead}_3} e^*(\Omega_{E/S})^{\otimes -3}.$$

We claim that $\varphi_{E,P}$ is an isomorphism. Since both the source and target are line bundles, it suffices to check that $\varphi_{E,P}$ induces isomorphisms on all geometric fibers. For this, note that for each geometric point $\bar{s} : \operatorname{Spec} k \to S$, the map

$$f_*(\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3}) \underset{\mathcal{O}_{S,\bar{s}}}{\otimes} k \longrightarrow e^*(\Omega_{E/S})^{\otimes -3} \underset{\mathcal{O}_{S,\bar{s}}}{\otimes} k$$

induced by $\varphi_{E,P}$ is identified with $\varphi_{E_{\bar{s}},P_{\bar{s}}}$, i.e., the same construction but applied to the elliptic curve $E_{\bar{s}}$ over k and the point $P_{\bar{s}} \in E_{\bar{s}}[3](k)$ induced by P. Thus in order to check that $\varphi_{E,P}$ is an isomorphism, we may assume that $S = \operatorname{Spec} k$ for k an algebraically closed field. Then $\varphi_{E,P}$ becomes, for a choice of a local coordinate z around e, the map

$$\varphi_{E,P} : \mathbf{H}^0(E, \mathcal{O}(3[P] - 3[e])) \xrightarrow{\mathrm{Lead}_3} k,$$

sending each function to the coefficient of z^{-3} in its Laurent expansion near e. Since $P \in E[3]$, we see that there is a function in the left hand side whose divisor is precisely $3[P] - 3[e]$. Hence φ is surjective. Also, we see that $\varphi^{-1}(0)$ is the set of $h \in k(E)$ that have at most 2 poles at e, no other poles, and at least 3 zeros at P. But this must just be 0. Hence φ is injective as well. The claim is proved.

By the claim, there is a unique global section of $f_*(\mathcal{I}_P^{\otimes 3} \otimes \mathcal{I}_e^{\otimes -3})$ that generates this line bundle and which has image ω^3 under Lead_3. Then this global section must be of the desired form $y + ax + b$ for unique $a, b \in R$.

16. Lecture 16

16.1. **The inversion formula and the Rigidity Lemma.** Next, we will prove the following fact, which we recall is another key ingredient in the proof of the representability of $S(3)$.

Proposition 16.1.1. *Let $S = \operatorname{Spec} R$ be an affine scheme, and suppose E/S is an elliptic cuve given by the Weierstrass equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ in $\mathbb{P}^2_R$ (with identity section $(0 : 1 : 0)$). Then $[-1] : E \to E$ is given by $(x : y : z) \mapsto (x : -y - a_1x - a_3 : z)$.*

Proof. Note that the claimed formula gives an endomorphism of E, in the sense that this is an S-scheme morphism $E \to E$ preserving the identity section; (this indeed implies that this is a morphism of group schemes, but we will not need this fact). This formula and the abstract morphism $[-1] : E \to E$ are two endomorphisms of E that agree on geometric fibers. We conclude by the following fact. $\square$

Fact 16.1.2. *If E/S is an elliptic curve and $\varphi, \varphi' : E \to E$ are endomorphisms of elliptic curves (in the above sense) such that they agree on $E_{\overline{s}}$ for each geometric point $\overline{s}$ of S, then $\varphi = \varphi'$.*

Remark. In fact the analogous statement is true for general abelian schemes.

This fact following from the following Rigidity Lemma.

Theorem 16.1.3 (Rigidity Lemma)**.** *Suppose G is a group scheme over an arbitrary base scheme S. Suppose $f : X \to S$ is a proper (or just closed, which is enough) morphism such that $f_*\mathcal{O}_X \cong \mathcal{O}_S$. Suppose $\varphi, \varphi' : X \to G$ are two S-morphisms that agree on each geometric fiber. Then φ and φ' differ by multiplication (with respect to the group structure of G) by a section $\gamma \in G(S)$.*

For a proof, see [Conb, §4].

Remark. Recall one sufficient set of conditions for the hypothesis $f_*\mathcal{O}_X \cong \mathcal{O}_S$: when f is proper, flat, and surjective with connected and reduced geometric fibers.

Remark. To see that we need not have $\varphi = \varphi'$, consider the case when $S = \operatorname{Spec} k[\epsilon]/(\epsilon^2)$ for some field k, and $G = \mathbb{G}_a/S = \mathbb{A}^1_S$. Set $f : X = S \to S$ to be the identity. Then S-maps from X to G form the additive group $G(S) = (k[\epsilon]/(\epsilon^2), +)$. We may set $\varphi = 0$ and $\varphi' = \epsilon$. They indeed agree on geometric fibers.

Proof of Fact 16.1.2. By the Rigidity Lemma, φ and φ' differ by some $\gamma \in E(S)$. But φ and φ' both preserve the identity section, and hence γ must be trivial. $\square$

16.2. **Relative representability of level structures.** We have already shown the representability of $S(3)$ over $\mathbb{Z}[1/3]$ by an affine, smooth scheme over $\mathbb{Z}[1/3]$. We accept the same for $S(4)$ (over $\mathbb{Z}[1/2] = \mathbb{Z}[1/4]$) without proof. Our next objective is to prove the following theorem.

Theorem 16.2.1. *For $N \geq 3$, the functor*

$$S(N) : (\textit{Locally noetherian schemes over } \mathbb{Z}[1/N]) \longrightarrow (\textit{Sets})$$

given by

$$T \longmapsto \{\textit{iso. cl. of elliptic curves } E/T \textit{ with level-}N \textit{ structure}\}$$

is representable by a smooth affine scheme over $\mathbb{Z}[1/N]$.

The idea is to consider the forgetful map $S(N) \to S(1)$, where $S(1)$ is the fibered groupoid of elliptic curves (i.e., for each test scheme S, $S(1)(S)$ is the **groupoid** of all elliptic curves over S, namely the category of all elliptic curves over S where the only allowed morphisms are isomorphisms; if we have a morphism $S \to S'$, then we have a functor $S(1)(S') \to S(1)(S)$ given by pullback). We want to show that for each $\mathbb{Z}[1/N]$-scheme S and each object of $S(1)(S)$, i.e., an elliptic curve E over S, the pullback of $S(N) \to S(1)$ over S along $E \in S(1)(S)$ should be representable by a scheme over S. Let us state this more concretely as follows.

Theorem 16.2.2 (Relative representability of level structures)**.** *Let $N \geq 1$. Let S be a scheme over $\mathbb{Z}[1/N]$, and let E/S be an elliptic curve. The functor*

$$(\textit{schemes over } S) \longrightarrow (\textit{Sets})$$

given by

$$T \longmapsto \{\textit{level-}N \textit{ structures on } E_T = E \times_S T\}$$

is representable by an S-scheme $I_{E/S,N}$. Moreover, $I_{E/S,N}$ is an étale $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$-torsor over S under the natural $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$-action. (See the beginning of the next lecture for étale torsors.)

17. Lecture 17

17.1. **Relative representability of level structures, continued.** We first briefly recall some terminology. Suppose G is a finite group and S is a scheme. An **étale G-torsor** over S is a scheme X over S which is finite étale and surjective over S, together with an action of G on X via S-automorphisms such that the map

$$G_S \times_S X \longrightarrow X \times_S X$$

given for all S-schemes T via $G \times X(T) \to X(T) \times X(T), (g, x) \mapsto (gx, x)$ is an isomorphisms of S-schemes. (Here G_S denotes the constant group scheme over S given by G.) Equivalently, since X is finite étale and surjective over S, the condition on the action of G amounts to imposing that for all geometric points $\operatorname{Spec} \overline{k} \to S$, the action of G on $X(\overline{k})$ is free and transitive.

Example. Suppose k'/k is a finite Galois extension of fields. Take $X = \operatorname{Spec} k'$ and $S = \operatorname{Spec} k$. Let $G = \mathrm{Gal}(k'/k)$. Then X is an étale G-torsor over S. Note that the action of G on $X(k)$ is not free; $X(k)$ has only one element.

We now state and prove the relative representability of $S(N)$ over $S(1)$.

Theorem 17.1.1 (relative representability of level structures)**.** *Let N be a positive integer and S a scheme over $\mathbb{Z}[1/N]$. Fix an elliptic curve E/S. The functor*

$$(S\textit{-schemes}) \longrightarrow (\textit{Sets})$$

given by

$$T \mapsto \{\textit{level-}N \textit{ structures on } E_T = E \times_S T, \textit{ i.e., isomorphisms } \gamma : \underline{(\mathbb{Z}/N\mathbb{Z})}^2_T \xrightarrow{\sim} E_T[N]\}$$

is representable by a scheme $I_{E/S,N}$ which is an étale $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$-torsor over S.

Proof. We will use the following two facts, which we admit.

1. Since S is over $\mathbb{Z}[1/N]$, the group scheme $E[N]$ is finite étale over S.
2. For any S-scheme T, we have a canonical identification $E_T[N] \cong E[N] \times_S T$.

We define an S-scheme

$$M := E[N] \times_S E[N] = \text{“}\{(P, Q) \mid P, Q \in E[N]\}\text{”}.$$

We have a universal homomorphism

$$h : \underline{(\mathbb{Z}/N\mathbb{Z})}^2_M \to E_M[N] = E[N] \times_S M$$

given by

$$h(i, j) = iP + jQ, \quad \forall i, j \in \mathbb{Z}/N\mathbb{Z},$$

where (P, Q) is the “universal element of $E[N] \times E[N]$”, i.e., the section of

$$E_M[N] \times_M E_M[N] = E[N] \times_S E[N] \times_S M = M \times_S M \longrightarrow M$$

given by the diagonal $M \to M \times_S M$. Abstractly, the functor

$$(S\text{-schemes}) \longrightarrow (\text{Sets}), \qquad T \longmapsto \{\text{homomorphisms } \underline{(\mathbb{Z}/N\mathbb{Z})}^2_T \to E_T[N]\}$$

is represented by the S-scheme M, and h is the universal object.

We want to construct $I_{E/S,N}$ as a suitable locus in M over which h is an isomorphism. This can indeed be done by throwing away certain connected components of M. More precisely, notice that for each connected component M_i of M, exactly one of the following must be true.

(1) $h|_{M_i} : \underline{(\mathbb{Z}/N\mathbb{Z})}^2_{M_i} \to E_{M_i}[N]$ is an isomorphism.
(2) For each geometric point $x : \operatorname{Spec} \bar{k} \to M_i$, the homomorphism $h_x : \underline{(\mathbb{Z}/N\mathbb{Z})}^2_x \to E_x[N]$ is not an isomorphism.

Indeed, for each geometric point $x : \operatorname{Spec} \bar{k} \to M_i$, we have the following equivalence of categories:

$$(\text{finite étale } M_i\text{-schemes}) \longrightarrow (\text{finite } \pi_1^{\text{ét}}(M_i, x)\text{-sets}), \qquad Y \longmapsto Y_x(\bar{k}).$$

This follows from Grothendieck's Galois theory for schemes and the fact that M_i is connected. Now both the source and target of $h|_{M_i}$ are finite étale M_i-schemes. If h_x is an isomorphism, then the image of $h|_{M_i}$ under the above equivalence of categories is an isomorphism between $\pi_1^{\text{ét}}(M_i, x)$-sets (since it is a bijection of sets and equivariant under $\pi_1^{\text{ét}}(M_i, x)$). This implies that h_{M_i} is an isomorphism.

Now notice that for any S-scheme T and any isomorphism $\gamma : \underline{(\mathbb{Z}/N\mathbb{Z})}^2_T \to E_T[N]$, the resulting S-scheme morphism $T \to M$ arising from γ will factor through the union U of those connected components M_i satisfying (1) above, because otherwise there will be a geometric point of T over which γ is not an isomorphism. Conversely, if we have an S-scheme morphism $T \to U \subset M$, then the pullback of h to T is an isomorphism. In particular, this indicates that the functor in the theorem is represented by $I_{E/S,N} := U$.

Since M is finite étale over S, we know that $I_{E/S,N}$ is finite étale as well. Also, using the moduli interpresentation of $I_{E/S,N}$, we can easily see that $I_{E/S,N} \to S$ is surjective and a $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$-torsor. □

17.2. **Back to the construction of the modular curve.** We now have enough to start proving the representability of $S(N)$.

Theorem 17.2.1. *Let $N \geq 3$. The functor*

$$S(N) : (\textit{locally noetherian schemes over } \mathbb{Z}[1/N]) \longrightarrow (\textit{Sets})$$

given by

$$S \longmapsto \{\textit{iso. cl. of elliptic curves over } S \textit{ with level-}N \textit{ structure}\}$$

is representable by a smooth affine scheme over $\mathbb{Z}[1/N]$ with all fibers pure of dimension 1.

Proof. We assume this for $N = 3$ and $N = 4$. Suppose that $N \geq 5$ is general. We proceed along cases.

Case (a): Suppose $3|N$. We have an open immersion $\operatorname{Spec} \mathbb{Z}[1/N] \subset \operatorname{Spec} \mathbb{Z}[1/3]$. We have already constructed $S(3)$ as a moduli scheme over $\mathbb{Z}[1/3]$. We denote the universal object over $S(3)$ as (E_3, γ_3). We write $S(3)[\frac{1}{N}]$ for the base change of $S(3)$ over $\mathbb{Z}[1/N]$. Then we obtain $S(N)$ by the following fiber product:

$$\begin{tikzcd} S(N) \arrow[r] \arrow[d] & I_{E_3/S(3)[\frac{1}{N}],N} \arrow[d, "\text{forget}"] \\ S(3)[\frac{1}{N}] \arrow[r, "\gamma_3"] & I_{E_3/S(3)[\frac{1}{N}],3} \end{tikzcd}.$$

Here, the right vertical map is defined by the canonical process of obtaining a level 3-structure from a level N-structure, namely by identifying $\underline{(\mathbb{Z}/3\mathbb{Z})}^2$ (resp. $E[3]$) with the 3-torsion inside $\underline{(\mathbb{Z}/N\mathbb{Z})}^2$ (resp. $E[N]$). This map is finite étale. Hence $S(N)$ if finite étale over $S(3)[1/N]$, and therefore it is smooth, affine, of pure relative dimension 1 over $\mathbb{Z}[1/N]$ as $S(3)[1/N]$ has the same properties.

Case (b): Suppose $4|N$. Then we proceed as in Case (a), making use of $S(4)$ instead of $S(3)$.

In the next lecture, we will treat the case when neither 3 or 4 divides N. □

18. Lecture 18

18.1. The construction of the modular curve, continued.

Proof of Theorem 17.2.1, continued. In the last lecture we constructed $S(N)$, in the case where 3 or 4 divides N, by taking a fiber product. That the fiber product indeed represents the correct functor follows from statement (1) in Lemma 18.1.1 below.

Case (c): Suppose N is coprime to 6. We have an open covering

$$\{\operatorname{Spec}\mathbb{Z}[1/2N], \operatorname{Spec}\mathbb{Z}[1/3N]\}$$

of $\operatorname{Spec}\mathbb{Z}[1/N]$. It suffices to construct $S(N)$ over $\mathbb{Z}[1/2N]$ and over $\mathbb{Z}[1/3N]$ separately, because over $\operatorname{Spec}\mathbb{Z}[1/2N] \cap \operatorname{Spec}\mathbb{Z}[1/3N]$ the two constructions must be canonically isomorphic by Yoneda's lemma, which allows us to glue the two constructions together to obtain $S(N)$ over $\mathbb{Z}[1/N]$.

We construct $S(N)$ over $\mathbb{Z}[1/3N]$. We already have $S(3N)$ over $\mathbb{Z}[1/3N]$ by Case (a). Let $K = \ker(\mathrm{GL}_2(\mathbb{Z}/3N\mathbb{Z}) \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}))$. The group $\mathrm{GL}_2(\mathbb{Z}/3N\mathbb{Z})$, and therefore K, acts on $S(3N)$ via the moduli interpretation of $S(3N)$ by permuting the level-$3N$ structure, i.e.,

$$g\cdot(E,\gamma) = (E, \gamma\circ g^{-1}), \quad \forall (E,\gamma)\in S(N)(S), \forall \text{ loc. noeth. } \mathbb{Z}[1/3N]\text{-scheme } S.$$

Now for any locally noetherian $\mathbb{Z}[1/3N]$-scheme S and any elliptic curve E/S, after étale localization on S, to give a level N-structure on E/S is the same as to give a K-orbit of level-$3N$ structures. (The étale localization on S is needed just to guarantee the existence of one level $3N$-structure on E; for instance we can use the finite étale cover $I_{E/S,3N} \to S$ to achieve this.) This suggests that the K-action should make $S(3N)$ an étale K-torsor over $S(N)|_{\mathbb{Z}[1/3N]}$ (which is yet to be constructed). In turn, we should construct $S(N)|_{\mathbb{Z}[1/3N]}$ as the **quotient** of $S(3N)$ by K. This quotient is of the simplest type in algebraic geometry as we now explain.

Suppose $X \to Y$ is a morphism of finite type between affine schemes $X = \operatorname{Spec} B$ and $Y = \operatorname{Spec} A$. Assume that Y is noetherian. Let G be a finite group acting on X via Y-scheme automorphisms. Suppose for any geometric point $\operatorname{Spec} k \to Y$, the action of G on $X(k) = \{Y\text{-morphisms } \operatorname{Spec} k \to X\}$ is free. Then $X/G := \operatorname{Spec}(B^G)$ is again a Y-scheme of finite type, and the natural map $X \to X/G$ is an étale G-torsor. Moreover, X/G is the categorical quotient of X by G, in the sense that for every Y-scheme Z, every G-invariant Y-scheme map $X \to Z$ factors uniquely through $X \to X/G$. Furthermore, X/G is the geometric quotient of X by G in the sense that the topological space $|X/G|$ is the quotient space of $|X|$ by G. The reference for these statements is [Gro03, V, §1, §2]. (See also [Mum08, §6] when the base is an algebraically closed field.)

Now in our case, the K-action on the $\mathbb{Z}[1/3N]$-scheme $S(3N)$ satisfies the freeness hypothesis in the above paragraph, by statement (1) in Lemma 18.1.1 below. (For each geometric point $\operatorname{Spec} k \to \operatorname{Spec} \mathbb{Z}[1/3N]$, each K-orbit in $S(3N)(k)$ is equal to the set on the right hand side of the bijection in Lemma 18.1.1 (1) for some choice of (E, γ) (with $N' = 3N$). But clearly the K-action on the left hand side is free.) Hence we can form the quotient $S(N)|_{\mathbb{Z}/[1/3N]} := S(3N)/K$, and $S(3N) \to S(N)|_{\mathbb{Z}/[1/3N]}$ is an étale K-torsor.

We now give a rigorous argument justifying that $S(N)|_{\mathbb{Z}[1/3N]}$ constructed above indeed represents the correct functor. To simplify notation we write $\mathscr{S} = S(N)|_{\mathbb{Z}[1/3N]}$. We will use the assumption that N is coprime to 3 in order to simplify the argument, although this could be avoided without too much difficulty. Firstly, we need to construct a universal object over $\mathscr{S}$. Over $S(3N)$ we have the universal object (E_{3N}, γ_{3N}), and we let γ'_N be the level-N structure on E_{3N} induced by γ_{3N}. Then (E_{3N}, γ'_N) descends to $\mathscr{S}$ by finite étale descent since it is "invariant" under the action of K.[6] The resulting elliptic curve with level-N structure on $\mathscr{S}$ will serve as the universal object, and we denote it by (E_N, γ_N). Now suppose we have a locally noetherian $\mathbb{Z}[1/3N]$-scheme S, and an elliptic curve with level N-structure (E, γ) over S. Since $(3, N) = 1$, we have $E[3N] \cong E[3] \times E[N]$, and similarly $\mathbb{Z}/3N\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$. Therefore, for any S-scheme T, to give a level-$3N$ structure on E_T compatibly with the prescribed level-N structure γ is the same as to give simply a level-3 structure on E_T. Thus we have a map $\varphi : I_{E/S,3} \to S(3N)$, where for any level-3 structure we "combine" it with γ to produce a level $3N$-structure. Now φ is K-equivariant (with $K \cong \mathrm{GL}_2(\mathbb{Z}/3\mathbb{Z})$ acting non-trivially on both sides), and therefore it descends to a map $\psi : S \to \mathscr{S}$ (by finite étale descent, using that $I_{E/S,3}$ is an étale K-torsor on S). One then checks that ψ is the unique map that pulls (E_N, γ_N) back to (E, γ) up to isomorphism.

Similarly, we construct $S(N)|_{\mathbb{Z}[1/2N]}$ as the quotient of $S(4N)$ (which is already constructed in Case (b)) by $\ker(\mathrm{GL}_2(\mathbb{Z}/4N\mathbb{Z}) \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}))$. Then we glue the two constructions to obtain $S(N)$ over $\mathbb{Z}[1/N]$ as we have already explained. By construction and finite étale descent, $S(N)$ is affine, smooth, of pure relative dimension 1 over $\mathbb{Z}[1/N]$.

Case (d): We are left with the case where $N = 2d$, with d coprime to 6. By Case (c) we already have $S(d)$ over $\mathbb{Z}[1/d]$. Since $d|N$, we can construct $S(N)$ from $S(d)$ in the same way as how we constructed $S(N)$ from $S(3)$ in Case (a). (Note that since $(2, d) = 1$, we in fact have $S(N) \cong I_{E_d/S(d)[\frac{1}{2}],2}$ by the decomposition $E[N] = E[d] \times E[2]$ for any elliptic curve E over any locally noetherian $\mathbb{Z}[1/N]$-scheme S.) $\square$

Statement (1) in the following lemma is used for several times in the above proof of Theorem 17.2.1.

Lemma 18.1.1. *Let $N \geq 3$. Let E be an elliptic curve over S, and γ be a level-N structure on E. The following statements hold.*

(1) *Let N' be a positive multiple of N. For any level N'-structure γ' on E, denote by $\gamma'|_N$ the level-N structure induced by γ' (as explained in Case (a) in the proof of Theorem 17.2.1). Then the natural map*

$$\begin{aligned}&\{\gamma' \mid \gamma' \text{ is a level-}N' \text{ str. on } E, \gamma'|_N = \gamma\}\\ &\qquad\longrightarrow \{(E', \gamma') \text{ ell. cv. with level } N'\text{-str. on } S \mid (E', \gamma'|_N) \cong (E, \gamma)\}/\mathit{isom}\end{aligned}$$

is a bijection. Here on the right hand side we quotient out by the isomorphisms between elliptic curves with level-N' structures.

[6] More precisely, for any $k \in K$, writing f_k for the automorphism of $S(3N)$ given by k, we have a canonical isomorphism $f_k^*(E_{3N}, \gamma'_N) \xrightarrow{\sim} (E_{3N}, \gamma'_N \circ k^{-1}) = (E_{3N}, \gamma'_N)$. These isomorphisms satisfy the cocycle relation and therefore give rise to a descent datum.

(2) *The pair (E, γ) has no non-trivial automorphism, i.e., the only automorphism of E (as an elliptic curve over S) preserving γ is the identity.*

Proof. We first show that (2) implies (1). Clearly the two sets in (1) are simultaneously empty or non-empty, and when they are non-empty the map in question is surjective. To show injectivity, suppose γ' and γ'' are two elements of the left hand side such that $(E, \gamma') \cong (E, \gamma'')$. Thus there is an automorphism τ of E carrying γ' to γ''. Since $\gamma'|_N = \gamma''|_N = \gamma$, we see that τ preserves γ, and therefore must be the identity by (2). It follows that $\gamma' = \gamma''$.

We now prove (2). By the rigidity of the endomorphisms of E (see Fact 16.1.2), it suffices to treat the case where S is the spectrum of an algebraically closed field k. In this case the statement is classical, but we give a proof.

Let $H = \mathrm{End}(E)$. This is a $\mathbb{Z}$-algebra, and is a free $\mathbb{Z}$-module of rank 2 or 4. Moreover, the $\mathbb{Q}$-algebra $H \otimes_{\mathbb{Z}} \mathbb{Q}$ is either a quadratic imaginary field or a quaternion algebra over $\mathbb{Q}$ ramified at ∞ and p, with the latter happening only when $\mathrm{char}\, k = p$. Now for any $h \in H$, either $h \in \mathbb{Z}$ or the minimal polynomial $P_h(T)$ of h over $\mathbb{Q}$ is a degree 2 monic polynomial in $\mathbb{Z}[T]$. (Here h is integral over $\mathbb{Z}$ since H is a finite $\mathbb{Z}$-module.)

We have $\mathrm{Aut}(E) = H^\times$, and by the previous description of H we know that $\mathrm{Aut}(E)$ is finite. Let $g \in \mathrm{Aut}(E)$ and suppose that g preserves γ. Then g acts trivially on $E[N]$. If $g \in \mathbb{Z}$, then $g = \pm 1$, and then $g = 1$ since -1 does not act trivially on $E[N]$ (as $N \geq 3$). We may thus assume that $g \notin \mathbb{Z}$. Then $P_h(T)$ is a quadratic monic irreducible polynomial in $\mathbb{Z}[T]$ whose roots are roots of unity (since g is of finite order), and therefore it must be one of the following:

$$T^2 + 1, T^2 - T + 1, T^2 + T + 1.$$

On the other hand the fact that g acts trivially on $E[N]$ implies that $g - 1 \in N \cdot H$, i.e., $\frac{g-1}{N} \in H$. It is easy to see that

$$P_g(T) = N^2 P_{\frac{g-1}{N}}(\frac{T-1}{N}).$$

Since $P_{\frac{g-1}{N}}(T)$ is monic integral, we see that

$$P_g(T) \equiv T^2 - 2T + 1 \mod N.$$

But this is not true for the three candidates of $P_g(T)$, contradiction. □

19. Lecture 19

We reviewed some key points from last lecture, including Case (c) in the proof of Theorem 17.2.1, and Lemma 18.1.1 (1).

19.1. **Abelian schemes.** We would like to generalize the modular curves to the higher-dimensional **Siegel modular varieties**. These are moduli spaces of polarized abelian schemes with level structure. Our next goal is to prove that such a moduli functor is indeed representable over a base like $\mathbb{Z}[1/N]$, generalizing Theorem 17.2.1.

From now on, all schemes are assumed to be locally noetherian.

Definition 19.1.1. Let S be a scheme. An **abelian scheme** over S is a smooth proper group scheme over S all of whose geometric fibers are connected.

One can prove several useful facts about abelian schemes using the following Rigidity Lemma, which is a (more powerful) variant of Theorem 16.1.3.

Theorem 19.1.2 (Rigidity Lemma)**.** *Let S be a scheme, and G be a group scheme over S and separated over S. Let $f : X \to S$ be a scheme morphism such that*

(1) *f is flat.*
(2) *either f is proper, or f is closed and admits a section.*
(3) *For each $s \in S$, the $k(s)$-vector space $\mathbf{H}^0(X_s, \mathcal{O}_{X_s})$ is 1-dimensional.*

Then for any two S-morphisms $\phi, \phi' : X \to G$, if ϕ and ϕ' agree on one geometric fiber (or equivalently, on one fiber) for each connected component of S, then ϕ and ϕ' differ by multiplication by a section in $G(S)$.

Proof. See [MFK94, Prop. 6.1]. (Note that in *loc. cit.* all schemes are assumed to be separated over $\operatorname{Spec} \mathbb{Z}$, so all scheme maps are automatically separated.) $\square$

Remark. In Theorem 19.1.2, if f is flat and proper, then assumption (3) implies that the natural map $\mathcal{O}_S \to f_* \mathcal{O}_X$ is an isomorphism. Thus in this case the set of assumptions on f is strictly stronger than Theorem 16.1.3 (where f is not assumed to be flat whatsoever). Also, in Theorem 16.1.3 the group scheme G is not assumed to be separated over S. In any case, the assumptions on f in Theorem 19.1.2 are satisfied if f is flat proper with connected and reduced geometric fibers.

We have the following interesting consequence, which tells us that we can "separate variables" for G-valued functions in two variables under suitable assumptions.

Corollary 19.1.3. *Let X and G over S be as in Theorem 19.1.2. Assume either that $X \to S$ is proper, or that it is universally closed and admits a section. Let Y be a connected scheme over S and assume that $Y \to S$ admits a section ϵ. Then for any S-scheme morphism $\varphi : X \times_S Y \to G$, there are S-scheme morphisms $g : X \to G$ and $h : Y \to G$ such that φ is given by $(x, y) \mapsto g(x) \cdot h(y)$.*

20. Lecture 20

20.1. **Abelian schemes, continued.** We continue to assume all schemes are locally noetherian.

Proof of Corollary 19.1.3. Let $f : X \to S$ be the structure map. We consider the following commutative diagram

$$\begin{tikzcd} X \times_S Y \arrow[rr, shift left, "\Phi"] \arrow[rr, shift right, "\Phi'"'] \arrow[dr] & & G \times_S Y \arrow[dl] \\ & Y & \end{tikzcd}$$

where we define

$$\Phi(x, y) := (\varphi(x, y), y), \qquad \Phi'(x, y) = (\varphi(x, \epsilon(f(x))), y).$$

The Y-scheme $X \times_S Y$ and the Y-group scheme $G \times_S Y$ satisfy the hypotheses of Theorem 19.1.2. Now Φ, Φ' are Y-morphisms and for any $y_0 \in \operatorname{im}(\epsilon)$, the morphisms Φ and Φ' agree on the fiber of $X \times_S Y$ over y_0. Hence by Theorem 19.1.2 we know that Φ, Φ' differ by multiplication by a section of $G \times_S Y \to Y$, which is of the form $y \mapsto (h(y), y)$ for some S-map $h : Y \to G$. Then we have $\varphi(x, y) = \varphi(x, \epsilon(f(x))) \cdot h(y)$. Setting $g(x) := \varphi(x, \epsilon(f(x)))$ we can conclude the proof. $\square$

Recall that an abelian scheme is a proper smooth group scheme with connected geometric fibers.

Corollary 20.1.1. *Suppose X/S is an abelian scheme and G/S is a separated group scheme. Any S-map $\varphi : X \to G$ preserving the neutral section is a group homomorphism. In particular, the group structure on X is determined by the neutral section.*

Proof. We may assume that S is connected. Then X is connected, since $X \to S$ is closed, surjective, and has connected fibers. Consider the composition

$$\Phi : X \times_S X \xrightarrow{\mu} X \xrightarrow{\varphi} G$$

where μ is the multiplication map, i.e., $\Phi(x, y) = \varphi(x \cdot y)$. Then by Corollary 19.1.3, we have $\varphi(x \cdot y) = g(x) \cdot h(y)$ for some $g : X \to G$ and $h : X \to G$. Now observe that

$$e = \varphi(e \cdot e) = g(e)h(e).$$

This implies that $h(e) = g(e)^{-1}$. Then we have

$$\varphi(x) = \varphi(x \cdot e) = g(x)h(e) = g(x)g(e)^{-1}.$$

Also,

$$\varphi(x) = \varphi(e \cdot x) = g(e)h(x).$$

So we have

$$g(x) = \varphi(x)g(e), \quad h(x) = g(e)^{-1}\varphi(x).$$

Hence we have

$$\varphi(x \cdot y) = g(x)h(y) = \varphi(x)g(e)g(e)^{-1}\varphi(y) = \varphi(x)\varphi(y).$$

This concludes. □

Corollary 20.1.2. *Suppose X/S is an abelian scheme. Then the group structure is commutative.*

Proof. Apply Corollary 20.1.1 to the inversion $X \to X, x \mapsto x^{-1}$. □

20.2. **Picard schemes.** Again we demand all schemes to be locally noetherian.

For a scheme X, we define $\mathrm{Pic}(X)$ to be the abelian group of isomorphism classes of invertible $\mathcal{O}_X$-modules. A morphism $f : X \to Y$ of schemes gives a group homomorphism $\mathrm{Pic}(Y) \to \mathrm{Pic}(X)$ via $\mathcal{L} \mapsto f^*\mathcal{L}$. This yields the **absolute Picard functor** $\mathrm{Pic} : (\mathrm{Sch})^{\mathrm{op}} \to (\mathrm{Ab})$.

Suppose $f : X \to S$ is a scheme morphism. It will be easier to work with the **relative Picard functor**, defined as

$$\mathrm{Pic}_{X/S} : (S\text{-schemes})^{\mathrm{op}} \longrightarrow (\mathrm{Ab}), \qquad T \longmapsto \mathrm{Pic}(X_T)/f_T^* \mathrm{Pic}(T),$$

where $X_T := X \times_S T$ and $f_T : X_T \to T$ is the base change of f.

Theorem 20.2.1 (Grothendieck)**.** *Suppose $X \to S$ is a flat projective morphism with all geometric fibers integral (irreducible and reduced). Also assume that X/S has a section. Then $\mathrm{Pic}_{X/S}$ is representable by a commutative group scheme over S which is locally of finite type and separated over S.*

Proof. See [Kle05, Thm. (9.)4.8]. □

Remark. (1) If $e \in X(S)$ is a section, then $\mathrm{Pic}_{X/S} \cong \mathrm{Pic}_{X/S,e}$ where $\mathrm{Pic}_{X/S,e}$ is the **rigidified Picard functor** sending each S-scheme T to the group of isomorphism classes of pairs $(\mathcal{L}, \rho)$, where $\mathcal{L}$ is a line bundle on X_T and ρ is an isomorphism $e_T^*\mathcal{L} \xrightarrow{\sim} \mathcal{O}_T$ (called a **rigidification of $\mathcal{L}$ along** e_T). Here, e_T denotes the section of $X_T \to T$ induced by the section e of $X \to S$. More explicitly, we have inverse bijections

$$\begin{aligned} \operatorname{Pic}(X_T)/f_T^* \operatorname{Pic}(T) &\longleftrightarrow \operatorname{Pic}_{X/S,e}(T) \\ \mathcal{L} &\longleftarrow (\mathcal{L}, \rho) \\ \mathcal{L} &\longmapsto (\mathcal{L} \otimes f_T^* e_T^* \mathcal{L}^{-1}, \text{ canonical } \rho). \end{aligned}$$

Here the canonical ρ is defined by noting that $e_T^*(\mathcal{L} \otimes f_T^* e_T^* \mathcal{L}^{-1})$ is canonically isomorphic to $e_T^*\mathcal{L} \otimes e_T^*\mathcal{L}^{-1}$.

(2) Without assuming the existence of a section but with the other assumptions in force, the theorem still holds for the fppf-sheafification (or just the étale sheafification) of $\operatorname{Pic}_{X/S}$. See the discussion in [Kle05, §§3–4].

21. Lecture 21

21.1. **Projective morphisms.** Recall that a morphism $f : X \to S$ is called **projective**, if the S-scheme X is isomorphic to a closed subscheme of the projective bundle $\mathbb{P}(\mathcal{E})$ over S attached to some coherent $\mathcal{O}_S$-module $\mathcal{E}$ on S. We caution the reader that in general this is not the same as requiring that X is isomorphic to a closed subscheme of $\mathbb{P}^n_S$ for some n. However, when S admits an ample inveritble sheaf (e.g., when S is affine), the two definitions are the same; see for instance [Sta18, Tag 0B45].

If $f : X \to S$ is projective, then it is proper, and there is an open covering (U_i) of S such that $X|_{U_i}$ is U_i-isomorphic to a closed subscheme of $\mathbb{P}^{n_i}_{U_i}$ for each i (see [Sta18, Tag 01WB]). The converse is not true. Thus a locally projective morphism (i.e., one that becomes projective after passing to an open covering of the target) need not be projective.

21.2. **The torsion component of the Picard scheme.** The following result enhances Theorem 20.2.1.

Theorem 21.2.1 (Grothendieck). *Let $f : X \to S$ be a flat projective morphism whose geometric fibers are integral. Assume that f admits a section, and that f is smooth. Assume that S is noetherian. Then there is a closed and open subgroup scheme $\operatorname{Pic}^\tau_{X/S}$ of $\operatorname{Pic}_{X/S}$ (over S), called the* ***torsion component****, satisfying the following conditions:*

(1) *For each $s \in S$, the fiber of $\operatorname{Pic}^\tau_{X/S}$ over s consists of the torsion connected components of $(\operatorname{Pic}_{X/S})_s$. Here we say that a connected component is torsion if its image under the multiplication-by-n map $[n] : (\operatorname{Pic}_{X/S})_s \to (\operatorname{Pic}_{X/S})_s$ lies in the identity connected component for some $n \geq 1$.*
(2) *$\operatorname{Pic}^\tau_{X/S}$ is projective over S.*

Proof. See [Kle05, Thm. 9.6.16, Exc. 9.6.18]. (It seems that the necessity of the assumption that S is noetherian is overlooked in Grothendieck's original article [Gro62, Cor. 4.2].) □

21.3. **Dual abelian schemes.** Now if X/S is an abelian scheme, then all the assumptions in Theorem 20.2.1 are satisfied. If we assume that X/S is projective and that S is noetherian, then the assumptions in Theorem 21.2.1 are satisfied as well. Furthermore we have the following result:

Theorem 21.3.1. *Let X/S be a projective abelian scheme, and assume that S is noetherian. Then $\operatorname{Pic}^\tau_{X/S}$ is smooth and has connected geometric fibers. Hence in view of Theorem 21.2.1 we know that $\operatorname{Pic}^\tau_{X/S}$ is a projective abelian scheme.*

Proof. To show that $\mathrm{Pic}^{\tau}_{X/S}$ has connected geometric fibers, we reduce to the case where S is the spectrum of an algebraically closed field (since the formation of $\mathrm{Pic}^{\tau}_{X/S}$ commutes with base change). Then this is a fundamental result in the theory of abelian varieties over a field; see [Mum08, §13]. The proof that $\mathrm{Pic}^{\tau}_{X/S}$ is smooth is found in [MFK94, Prop. 6.7]. □

Definition 21.3.2. In the setting of Theorem 21.3.1, we call $\mathrm{Pic}^{\tau}_{X/S}$ the **dual abelian scheme of** X, and denote it by $X^{\vee}$.

Remark. For an abelian variety X over a field k, it is a classical result that $\mathrm{Pic}^{\tau}_{X/k}$ is (geometrically) connected. Hence for an abelian scheme X/S one could equivalently define $X^{\vee}$ by requring that fiberwise it is the identity connected component of $(\mathrm{Pic}_{X/S})_s$. (Note that the identity connected component of $(\mathrm{Pic}_{X/S})_s$ is automatically geometrically connected since it has a rational point.) One can interpret the definition of $X^{\vee}$ as a *subfunctor* of $\mathrm{Pic}_{X/S}$, even without knowing that $\mathrm{Pic}_{X/S}$ is representable. Namely, for each S-scheme T we declare that an element $\xi \in \mathrm{Pic}_{X/S}(T)$ belongs to $X^{\vee}(T)$ if for every geometric point $\mathrm{Spec}\, k \to T$ there exist a connected k-scheme V and an element of $\mathrm{Pic}_{X/S}(V)$ which specializes to ξ and to 0 at two k-points of V. Hence for an arbitrary abelian scheme X/S one could ask whether the functor $X^{\vee}$ is represented by an abelian scheme, even without requiring that $\mathrm{Pic}_{X/S}$ is representable.

To answer this question, first we have a relatively easy generalization of Theorem 21.3.1: For any locally projective abelian scheme X over any (locally noetherian) S, $\mathrm{Pic}^{\tau}_{X/S}$ is a locally projective abelian scheme. This would immediately follow from Theorem 21.3.1 once we check that $X^{\vee}$ is a Zariski sheaf. A much deeper generalization is to drop all the assumptions whatsoever: For any abelian scheme X over any S, $X^{\vee}$ is an abelian scheme. (This holds even without assuming that S is locally noetherian.) This result is due to Raynaud and Deligne on top of Artin's general result on the representability of Pic by algebraic spaces; see [FC90, §I.1].

21.4. **Isogenies.**

Definition 21.4.1. Let A, B be two abelian schemes over an arbitrary (locally noetherian) S. By an **isogeny**, we mean an S-group scheme homomorphism $A \to B$ that is surjective and quasi-finite.

Lemma 21.4.2. *Any isogeny $\phi : A \to B$ is finite and flat.*

Proof. Since both A and B are proper over S, we know that ϕ is proper. But a proper and quasi-finite map is finite ([Sta18, Tag 02LS]), so ϕ is finite.

Since both A and B are flat and finite-type over S, we have the *fiberwise criterion for flatness*. Namely, in order to check that $\phi : A \to B$ is flat, we only need to check that $\phi_s : A_s \to B_s$ is flat for each $s \in S$. (See [Gro66, (11.3.11)] or [Sta18, Tag 039E].) Hence we reduce to the case where S is the spectrum of a field k. We may also assume that k is algebraically closed since flatness satisfies fpqc descent (see [Gro65, (2.2.11) (iv)] or [Sta18, Tag 02L2]). Now ϕ is a surjective map between two finite-type schemes over a field, and the target is integral. Hence we have the *generic flatness* (see [Gro65, §6.9]): There exists a non-empty open subscheme $U \subset B$ over which ϕ is flat. Since ϕ is a group homomorphism, we can use the group structure on B to translate U, in order to obtain an open covering of B such that ϕ is flat over each member of the covering. (For this step we need to use that k is algebraic closed.) It follows that ϕ is flat, as desired. □

22. Lecture 22

22.1. **The Mumford Λ-construction.** Let S be a noetherian scheme, and $f : A \to S$ a projective abelian scheme over S. For any line bundle L on A, we define the **Mumford line bundle** $\mathfrak{M}(L)$ on $A \times_S A$ by

$$\mathfrak{M}(L) := \mu^* L \otimes p_1^* L^{-1} \otimes p_2^* L^{-1},$$

where p_1, p_2 are the two projections $A \times_S A \to A$, and μ is the group law $A \times_S A \to A$.

Recall that for any S-scheme T to give an S-map $T \to \mathrm{Pic}_{A/S}$ is the same as to specify an element of $\mathrm{Pic}(A_T)/f_T^* \mathrm{Pic}(T)$, where $A_T = A \times_S T$. Thus for $T = A$, the Mumford line bundle $\mathfrak{M}(L)$ on $A \times_S A = A_A$ gives rise to an S-map

$$\Lambda(L) : A \longrightarrow \mathrm{Pic}_{A/S}.$$

Lemma 22.1.1. *The S-map $\Lambda(L)$ takes the neutral section of A to the neutral section of $\mathrm{Pic}_{A/S}$.*

Proof. Let $e \in A(S)$ be the neutral section. To compute $\Lambda(L) \circ e$, we need to compute the pullback of $\mathfrak{M}(L)$ under

$$A = A \times_S S \xrightarrow{(\mathrm{id}, e)} A \times_S A, \quad x \longmapsto (x, e(f(x))).$$

Note that the compositions of the above map followed by $\mu, p_1, p_2 : A \times_S A \to A$ respectively are $\mathrm{id}, \mathrm{id}, e \circ f$. Hence the pullback of $\mathfrak{M}(L)$ under the above map is isomorphic to $f^* e^* L^{-1}$. This line bundle on A represents the zero element of $\mathrm{Pic}_{A/S}(S) = \mathrm{Pic}(A)/f^* \mathrm{Pic}(S)$. Hence $\Lambda(L) \circ e$ is the neutral section of $\mathrm{Pic}_{A/S}$. $\square$

As a consequence, we know that $\Lambda(L) : A \to \mathrm{Pic}_{A/S}$ is a group homomorphism by the Rigidity Lemma (see Corollary 20.1.1). Moreover, by the fiberwise connectedness of A, we know that $\Lambda(L)$ is a homomorphism $A \to A^\vee$. Similarly to the proof of the above lemma, one shows that for two line bundles L, M on A we have

$$\Lambda(L \otimes M^{\pm 1}) = \Lambda(L) \pm \Lambda(M).$$

22.2. **The case over an algebraically closed field.** We now assume that A is an abelian variety over an algebraically closed field k. In this case A is automatically projective. (Strictly speaking for our course we do not need this information, because we will be exclusively working with abelian schemes A/S which are *assumed* to be projective; in any case in the following we assume that A/k is projective.)

Note that $\mathrm{Pic}_{A/k}(k) = \mathrm{Pic}(A)$, since $\mathrm{Pic}(k)$ is trivial.

Let L be a line bundle on A. It is easy to see that at the level of k-points the map $\Lambda(L)$ is given by

$$A(k) \longrightarrow A^\vee \subset \mathrm{Pic}_{A/k}(k) = \mathrm{Pic}(A), \quad x \longmapsto t_x^* L \otimes L^{-1},$$

where $t_x : A \to A$ is translation by x. The fact that $\Lambda(L)$ is a group homomorphism thus entails the following:

Theorem 22.2.1 (Theorem of Square)**.** *For any line bundle L on A and any $x, y \in A(k)$, we have an isomorphism of line bundles*

$$t_{x+y}^* L \otimes L \cong t_x^* L \otimes t_y^* L.$$

We make two further observations.

Lemma 22.2.2. *Let L be a line bundle on A. Then $\Lambda(L) = 0$ if and only if the Mumford line bundle $\mathfrak{M}(L)$ is trivial.*

Proof. The "if" direction is clear from the very definition of $\Lambda(L)$ in terms of $\mathfrak{M}(L)$. Suppose that $\Lambda(L) = 0$. Then we know, again by the definition of $\Lambda(L)$, that $\mathfrak{M}(L)$ dies in $\operatorname{Pic}(A \times_k A)/p_2^* \operatorname{Pic}(A)$. In general, suppose M is a line bundle on $A \times_k T$ which isomorphic to $p_2^*(N)$ for some line bundle N on T. Write e_T for the map $T \to A \times_k T, t \mapsto (e, t)$. Then

$$p_2^* e_T^* M \cong p_2^* e_T^* p_2^* N \cong p_2^* N \cong M,$$

where the second isomorphism is because $p_2 \circ e_T = \mathrm{id}_T$. In particular,

$$M \otimes p_2^* e_T^* M^{-1} \cong \mathcal{O}_{A \times T}.$$

Applying this to $M = \mathfrak{M}(L)$, we know that

$$\mathfrak{M}(L) \otimes p_2^*(e, \mathrm{id})^* \mathfrak{M}(L)^{-1} \cong \mathcal{O}_{A \times A}.$$

A computation shows that the left hand side is isomorphic to $\mathfrak{M}(L) \otimes e_2^* L$, where e_2 is the map $A \times_k A \to A, (x, y) \mapsto e$. But $e_2^* L$ is trivial since e_2 factors through $\operatorname{Spec} k$ and $\operatorname{Pic}(k) = 0$. Hence $\mathfrak{M}(L)$ is trivial. $\square$

Since A/k is projective, it has ample line bundles.

Lemma 22.2.3. *Let L be an ample line bundle on A. Then $\ker(\Lambda(L))$ is a finite subgroup scheme of A.*

Proof. Suppose not. Then one can find a positive-dimensional abelian subvariety $B \subset A$ contained in $\ker(\Lambda(L))$. Note that $\Lambda(L)|_B = \Lambda(L|_B)$, and $L|_B$ is ample on B. Thus for the sake of deducing a contradiction we may assume that $B = A$, i.e., $\Lambda(L) = 0$. By Lemma 22.2.2, $\mathfrak{M}(L)$ is trivial. The pullback of $\mathfrak{M}(L)^{-1}$ along the "anti-diagonal"

$$(\mathrm{id}, [-1]) : A \longrightarrow A \times_k A, \quad x \longmapsto (x, -x)$$

is $L \otimes [-1]^* L$, and it must be trivial on A. Since L is ample and $[-1]$ is an automorphism of A, $L \otimes [-1]^* L$ is ample. Thus the trivial line bundle on A is ample, a contradiction with the fact that A is projective and positive-dimensional. $\square$

The following is the "main theorem" for line bundles on an abelian variety.

Theorem 22.2.4 (Main Theorem)**.** *Fix an ample line bundle L on A. For any line bundle M on A, we have $\Lambda(M) = 0$ if and only if $M \cong \Lambda(L)(x) = t_x^* L \otimes L^{-1}$ for some $x \in A(k)$.*

The proof of the above theorem will be given in the future few lectures. Let us now deduce some important consequences.

Definition 22.2.5. Let M, M' be two line bundles on A. We say that M is **algebraically equivalent** to M', if there exists a connected k-scheme T and a line bundle on $A \times_k T$ specializing to M and M' at two k-points of T.

Lemma 22.2.6. *Two line bundles M and M' are algebraically equivalent if and only if the isomorphism class of $M \otimes M'^{-1}$ lies in $A^\vee(k) \subset \operatorname{Pic}_{A/k}(k) = \operatorname{Pic}(A)$.*

Proof. Exercise. (Use the moduli interpretation of $\operatorname{Pic}_{A/k}$.) $\square$

The following corollary of Theorem 22.2.4 says that one can detect algebraic equivalence of line bundles by looking at $\Lambda(\cdot)$.

Corollary 22.2.7. *Let M be a line bundle on A. Then M is algebraically equivalent to zero if and only if $\Lambda(M) = 0$.*

We will prove this next time.

23. Lecture 23

23.1. **Criterion for algebraic equivalence.** Let A be an abelian variety over an algebraically closed field k. The following two corollaries are consequences of Theorem 22.2.4.

Corollary 23.1.1. *Let M be a line bundle on A. Then M is algebraically equivalent to zero if and only if $\Lambda(M) = 0$.*

Proof. Suppose M is algebraically equivalent to zero. Then there is a connected k-scheme T and a line bundle $\tilde{M}$ on $A \times_k T$ which specializes to M and to $\mathcal{O}_A$ at two points $t_1, t_2 \in T(k)$. Consider the T-group scheme homomorphism

$$\Lambda(\tilde{M}) : A_T = A \times_k T \longrightarrow \mathrm{Pic}_{A_T/T}.$$

On the fiber of A_T over t_2, the map induced by $\Lambda(\tilde{M})$ is $\Lambda(\mathcal{O}_A) = 0$. Thus $\Lambda(\tilde{M})$ and the zero homomorphism agree on one fiber over T. Since they both preserve the neutral section, they must be equal by the Rigidity Lemma (see Theorem 19.1.2). Thus $\Lambda(\tilde{M}) = 0$. But on the fiber over t_1, the map induced by $\Lambda(\tilde{M})$ is $\Lambda(M)$. Hence $\Lambda(M) = 0$.

Conversely, suppose that $\Lambda(M) = 0$. Then by Theorem 22.2.4, there exists an (ample) line bundle L on A and a point $x \in A(k)$ such that $M = \Lambda(L)(x)$. But $\Lambda(L)(A(k)) \subset A^\vee(k)$, so $M \in A^\vee(k)$. Thus M is algebraically equivalent to zero by Lemma 22.2.6. $\square$

Corollary 23.1.2. *Let L be an ample line bundle on A. Then $\Lambda(L) : A \to A^\vee$ is an isogeny.*

Proof. By Lemma 22.2.3, $\Lambda(L)$ is quasi-finite. To see that it is surjective, let $M \in A^\vee(k)$. Then $\Lambda(M) = 0$ by Corollary 23.1.1. Hence $M \in \mathrm{im}(\Lambda(L))$ by Theorem 22.2.4 . $\square$

The above two corollaries will be the essential tools needed to study "polarizations". Note that these they together imply the "only if" direction of Theorem 22.2.4, which is the more difficult direction.

23.2. **Proof of Theorem 22.2.4.** We first show the "if" direction. Suppose $M = t_x^* L \otimes L^{-1}$. Then for any $y \in A(k)$ we compute

$$\Lambda(M)(y) = t_y^* M \otimes M^{-1} \cong t_{x+y}^* L \otimes t_y^* L^{-1} \otimes t_x^* L^{-1} \otimes L,$$

which is trivial by the Theorem of Square (Theorem 22.2.1). Thus $\Lambda(M) = 0$. This proves the "if" direction.

For the "only if" direction we need some preparations.

Lemma 23.2.1. *Let M be a line bundle on A with $\Lambda(M) = 0$. Then for any k-scheme T and any two k-scheme maps $f, g : T \to A$, we have $(f+g)^* M \cong f^* M \otimes g^* M$. (Here the addition in $f+g$ is the group law.) For any $n \in \mathbb{Z}$, we have $[n]^* M \cong M^{\otimes n}$, where $[n]$ is the multiplication-by-n map $A \to A, x \mapsto x + \cdots + x$ (n times).*

Proof. By Lemma 22.2.2, $\mathfrak{M}(M)$ is trivial. The pullback of $\mathfrak{M}(M)$ along

$$(f, g) : T \longrightarrow A \times_k A$$

is $(f+g)^* M \otimes f^* M^{-1} \otimes g^* M^{-1}$, and this is trivial. This proves the first statement. For the second statement, note that it is obviously true for $n = 0$ or 1. (For $n = 0$, note that $[0]^*$ of any line bundle is trivial since $[0] : A \to A$ factors through $\mathrm{Spec}\, k$.) Applying the first statement to $f = [m]$ (with $m \geq 1$) and $g = [1]$, we prove by induction that the statement holds for all $n \geq 1$. For negative n, use that $\mathcal{O}_A \cong [0]^* M \cong [n-n]^* M \cong [n]^* M \otimes [-n]^* M$. $\square$

Lemma 23.2.2. *Let M be a line bundle on A with $\Lambda(M) = 0$. If M is non-trivial, then $\mathbf{H}^j(A, M) = 0$ for all $j \geq 0$.*

Proof. We induct on j. For $j = 0$, suppose $\mathbf{H}^0(A, M) \neq 0$. Then $M \cong \mathcal{O}_A(-D)$ for some effective divisor D. By Lemma 23.2.1, $\mathcal{O}_A(D) \cong M^{-1} \cong [-1]^*M \cong \mathcal{O}_A([-1]^*(-D))$, i.e.,

$$\mathcal{O}_A \cong \mathcal{O}_A(D + [-1]^*D).$$

But both D and $[-1]^*D$ are effective, so $D = 0$, a contradiction with the non-triviality of M.

For the induction step, assume $j \geq 1$. We factorize id_A as

$$A \xrightarrow{(\mathrm{id}, e)} A \times_k A \xrightarrow{\mu} A.$$

Thus the identity map on $\mathbf{H}^j(A, M)$ factors through $\mathbf{H}^j(A \times_k A, \mu^*M)$, and it suffices to prove that $\mathbf{H}^j(A \times_k A, \mu^*M) = 0$. Since $\mathfrak{M}(M)$ is trivial (by Lemma 22.2.2), we have

$$\mathbf{H}^j(A \times_k A, \mu^*M) \cong \mathbf{H}^j(A \times_k A, p_1^*M \otimes p_2^*M).$$

By the Künneth formula the above is isomorphic to

$$\bigoplus_{u+v=j} \mathbf{H}^u(A, M) \otimes_k \mathbf{H}^v(A, M),$$

and this is zero by the induction hypothesis since for every pair (u, v) with $u + v = j$ at least one of u, v is strictly less than j. □

We are now ready to prove the "only if" direction of Theorem 22.2.4. (The reference for this proof is [Mum08, §8, Thm. 1].) For the sake of contradiction suppose that $\forall x \in A(k)$, M is not isomorphic to $t_x^*L \otimes L^{-1}$. Define a line bundle on $A \times_k A$:

$$K := \mathfrak{M}(L) \otimes p_2^*M^{-1}.$$

For any $x \in A(k)$, we write $K|_{\{x\}\times A}$ for the pullback of K along $A \to A \times_k A, y \mapsto (x, y)$. We have

$$K|_{\{x\}\times A} \cong t_x^*L \otimes L^{-1} \otimes M^{-1},$$

and this is non-trivial by our assumption. Now

$$\Lambda(K|_{\{x\}\times A}) = \Lambda(t_x^*L \otimes L^{-1} \otimes M^{-1}) = \Lambda(t_x^*L \otimes L^{-1}) - \Lambda(M).$$

The first term is zero by the "if" direction of the theorem, and the second term is zero by assumption. Hence

$$\mathbf{H}^j(A, K|_{\{x\}\times A}) = 0, \qquad \forall j \geq 0$$

by Lemma 23.2.2. Since $x \in A(k)$ is arbitrary, this implies that

$$R^j p_{1,*}K = 0, \qquad \forall j \geq 0$$

by "cohomology and base change"; see [Mum08, §5, Cor. 3]. (To use this result one needs to check that $\mathbf{H}^j(\operatorname{Spec} k(x) \times_{\operatorname{Spec} k} A, K|_{\{x\}\times A}) = 0$ for *all* points $x \in A$, not just the closed points. However by semi-continuity [Mum08, §5, Cor. 1] knowing the vanishing for all closed points $x \in A(k)$ is already enough.) By the Leray spectral sequence

$$E_2^{p,q} = \mathbf{H}^p(A, R^q p_{1,*}K) \Longrightarrow \mathbf{H}^{p+q}(A \times_k A, K),$$

we conclude that

$$\mathbf{H}^i(A \times_k A, K) = 0, \qquad \forall i \geq 0.$$

We now look at the other Leray spectral sequence

$$E_2^{p,q} = \mathbf{H}^p(A, R^q p_{2,*}K) \Longrightarrow \mathbf{H}^{p+q}(A \times_k A, K). \tag{23.1}$$

For $x \in A(k)$, we have

$$K|_{A\times\{x\}} \cong t_x^*L \otimes L^{-1},$$

and so $\Lambda(K|_{A\times\{x\}}) = 0$ again by the "if" direction of the theorem. If $t_x^*L\otimes L^{-1}$ is non-trivial, i.e., $x \notin \ker(\Lambda(L))(k)$, then again by Lemma 23.2.2 we have

$$\mathbf{H}^j(K|_{A\times\{x\}}) = 0, \qquad \forall j \geq 0.$$

Hence for each j, $R^jp_{2,*}K$ is supported on the closed subscheme $\ker(\Lambda(L)) \subset A$. But L is ample, so $\ker(\Lambda(L))$ is a finite (maybe non-reduced) k-scheme by Lemma 22.2.3. Thus

$$\mathbf{H}^i(A, R^jp_{2,*}K) = 0, \qquad \forall i \geq 1, \forall j \geq 0.$$

Thus the only non-zero terms in the E_2-page of the spectral sequence (23.1) are those in the 0-th row. It follows that

$$\mathbf{H}^0(A, R^jp_{2,*}K) = \mathbf{H}^j(A\times_k A, K), \qquad \forall j \geq 0.$$

We have already seen that this is zero. Since $R^jp_{2,*}K$ is finitely supported, we conclude that

$$R^jp_{2,*}K = 0, \qquad \forall j \geq 0.$$

Again by "cohomology and base change" (see [Mum08, §5, Cor. 4]), this implies that for any $x \in A(k)$, we have

$$\mathbf{H}^j(A, K|_{A\times\{x\}}) = 0, \qquad \forall j \geq 0.$$

Taking $j = 0$ and $x = e$, we get $\mathbf{H}^0(A, \mathcal{O}_A) = 0$, which is absurd since this should be k. This finishes the proof of Theorem 22.2.4.

24. Lecture 24

24.1. Theorem of Cube and consequences.

Theorem 24.1.1 (Theorem of Cube)**.** *Let X, Y, Z be three abelian varieties over a field k (not necessarily algebraically closed). Let L be a line bundle on $X\times Y\times Z$ (all the products are over k) such that its restrictions to $\{e\}\times Y\times Z \cong Y\times Z, X\times\{e\}\times Z \cong X\times Z, X\times Y\times\{e\} \cong X\times Y$ are all trivial. Then L is trivial.*

Proof. Fix a trivialization ρ of $L|_{X\times Y\times\{e\}}$. Then (L,ρ) defines an element of $\mathrm{Pic}_{Z/k,e}(X\times Y)$, or in other words a k-map $F : X\times Y \to Z$. By Corollary 19.1.3, F is of the form $F(x, y) = g(x) + h(y)$ for k-maps $g : X \to Z$ and $h : Y \to Z$. By our assumptions, we have $g(e) + h(y) = g(x) + h(e) = 0$ for all $x \in X, y \in Y$. Thus g and h are constant (i.e., they factor through the structure maps $X \to \mathrm{Spec}\, k, Y \to \mathrm{Spec}\, k$ respectively). Then $F = 0$, and so (L,ρ) represents the trivial element of $\mathrm{Pic}_{Z/k,e}(X\times Y)$, which in particular implies that L is trivial. □

Corollary 24.1.2. *Let A be an abelian variety over a field k, and let L be a line bundle on A. Let T be a k-scheme, and $f, g, h : T \to A$ be k-maps. Then*

$$(f + g + h)^*L \cong (f + g)^*L \otimes (g + h)^*L \otimes (f + h)^*K \otimes f^*L^{-1} \otimes g^*L^{-1} \otimes h^*L^{-1}.$$

Proof. Let p_i be the i-th projection $A\times A\times A \to A$, and $p_{ij} := p_i + p_j : A\times A\times A \to A$. Let $m := p_1 + p_2 + p_3 : A\times A\times A \to A$. Define a line bundle on $A\times A\times A$:

$$M := m^*L \otimes p_{12}^*L^{-1} \otimes p_{13}^*L^{-1} \otimes p_{23}^*L^{-1} \otimes p_1^*L \otimes p_2^*L \otimes p_3^*L.$$

Then the difference of the two sides of the desired isomorphism is the pullback of M along

$$(f, g, h) : T \to A\times A\times A.$$

Hence it suffices to check that M is trivial. But one directly checks that M satisfies the hypothesis in the Theorem of Cube. □

Corollary 24.1.3. *Let A and L be as in Corollary 24.1.2. Then for each $n \in \mathbb{Z}$ we have*

$$[n]^*L \cong L^{n^2} \otimes (L \otimes [-1]^*L^{-1})^{(n-n^2)/2}.$$

Proof. Write L_n for $[n]^*L$. Applying Corollary 24.1.2 to $f = [n+1], g = [1], h = [-1]$, we get

$$L_{n+1} \cong L_{n+2} \otimes L_0 \otimes L_n \otimes L_{n+1}^{-1} \otimes L_1^{-1} \otimes L_{-1}^{-1},$$

i.e.,

$$L_{n+2} \otimes L_{n+1}^{-2} \otimes L_n \cong L \otimes L_{-1}.$$

Note that the desired isomorphism obviously holds for $n = 0$ and $n = 1$. The proof is then done by induction in the two directions, i.e, knowing the desired formula for L_n and L_{n+1} we can compute L_{n+2}, and knowing the desired formula for L_{n+2} and L_{n+1} we can compute L_n. □

We say that a line bundle L on A is **symmetric** if $L \cong [-1]^*L$. If L is symmetric, then $[n]^*L \cong L^{n^2}$. Now pick an ample line bundle L_0 and let $L = L_0 \otimes [-1]^*L_0$. Then L is both ample and symmetric. In particular, on A we have an ample line bundle L such that $[n]^*L \cong L^{\otimes n^2}$. From this, it is easy to see that $[n] : A \to A$ is an isogeny. Moreover, for any ample line bundle L on A and any isogeny $f : A \to A$, we have $\deg(L) \neq 0, \deg(L^m) = m^{\dim A} \deg(L)$, $\forall m \geq 1$, and $\deg(f^*L) = \deg(f) \deg(L)$. It follows that $\deg([n]) = n^{2 \dim A}$. Here the degree of a line bundle is defined using the Hilbert polynomial; see [Mum08, Appendix to §6] for details.

The following result will be proved in the next lecture.

Corollary 24.1.4. *Let A be an abelian variety over an algebraically closed field k. For any ample line bundle L on A, we have $\mathbf{H}^i(A, L) = 0$ for all $i > 0$.*

25. Lecture 25

25.1. **Cohomology of an ample line bundle.** In the following, let k be an algebraically closed field, and A/k an abelian variety.

Corollary 25.1.1. *For any line bundle L on A and any integer n, the line bundle $[n]^*L$ is algebraically equivalent to $L^{\otimes n^2}$.*

Remark. Recall from Lemma 23.2.1 that if L is a line bundle satisfying $\Lambda(L) = 0$, then $[n]^*L \cong L^{\otimes n}$. In this case L is algebraically equivalent to zero, so this does not contradict with the current corollary.

Proof. By Corollary 24.1.3, we only need to show that $\Delta := L \otimes [-1]^*L^{-1}$ is algebraically equivalent to 0. By Corollary 23.1.1, we only need to show that $\Lambda(\Delta) = 0$. For any $x \in A(k)$, we compute

$$\begin{aligned}\Lambda(\Delta)(x) \cong t_x^*L \otimes L^{-1} \otimes t_x^*[-1]^*L^{-1} \otimes [-1]^*L &\cong t_x^*L \otimes L^{-1} \otimes [-1]^*(t_{-x}^*L^{-1} \otimes L) \\ &\cong t_x^*L \otimes L^{-1} \otimes (t_{-x}^*L^{-1} \otimes L)^{-1},\end{aligned}$$

where the last isomorphism is because $[-1]^*M \cong M^{-1}$ for all M such that $\Lambda(M) = 0$ (see Lemma 23.2.1). The above is isomorphic to

$$t_x^*L \otimes t_{-x}^*L \otimes L^{-2},$$

and this is isomorphic to $t_0^*L \otimes L^{-1} \cong \mathcal{O}_A$ by the Theorem of Square. Hence $\Lambda(\Delta) = 0$. □

Corollary 25.1.2. *For any ample line bundle L on A, we have $\mathbf{H}^i(A, L) = 0$ for all $i > 0$.*

Proof. Since L is ample, there is an integer n_0 such that

$$\mathbf{H}^i(A, L^n) = 0, \qquad \forall i \geq 1, n \geq n_0.$$

Suppose L' is a line bundle algebraically equivalent to L^n. By Theorem 22.2.4, since $L' \otimes L^{-n}$ is algebraically equivalent to zero and since L^n is ample, we know that $L' \otimes L^{-n} \cong t_x^* L^n \otimes L^{-n}$ for some $x \in A(k)$. Thus $L' \cong t_x^* L^n$, and so $\mathbf{H}^i(A, L') = 0$ for all $i \geq 1$ since t_x is an automorphism of A.

Now by Corollary 25.1.1, $[n]^* L$ is algebraically equivalent to L^{n^2}, so we have

$$\mathbf{H}^i(A, [n]^* L) = 0, \qquad \forall i \geq 1, n \geq \sqrt{n_0}.$$

But the above is also isomorphic to

$$\mathbf{H}^i(A, [n]_* [n]^* L)$$

since $[n] : A \to A$ is finite. We claim that if $\operatorname{char} k$ does not divide n, then L is a direct summand of $[n]_* [n]^* L$. This would finish the proof since we can pick n sufficiently large and not divisible by $\operatorname{char} k$.

To prove the claim, first note that by the projection formula we have $[n]_* [n]^* L \cong L \otimes [n]_* \mathcal{O}_A$. Hence it suffices to show that $\mathcal{O}_A$ is a direct summand of $[n]_* \mathcal{O}_A$. Since $[n]$ is finite flat, there is a canonical trace map $[n]_* \mathcal{O}_A \to \mathcal{O}_A$ (see [Sta18, Tag 0BVH]) which when composed with the natural map $\mathcal{O}_A \to [n]_* \mathcal{O}_A$ on the left is the map $\mathcal{O}_A \to \mathcal{O}_A$ that multiplies each section by $\deg[n] \in \mathbb{Z}$. But $\deg[n] = n^{2 \dim A}$ is invertible over A, so the claim follows. □

26. Lecture 26

26.1. **Global sections of ample line bundles.** Let A be an abelian variety over an algebraically closed field k and take L to be an ample line bundle on A. We showed last lecture that $\mathbf{H}^i(A, L) = 0$ for all $i > 0$. We want now to show that for such L, we have the following result.

Theorem 26.1.1.

$$\dim \mathbf{H}^0(A, L) = \sqrt{\deg \Lambda(L)}.$$

We first briefly examine the case of elliptic curves. If $A = E$ is an elliptic curve, we note that $L_1 = \mathcal{O}(e)$ is a canonical choice of an ample line bundle, as $L_1^{\otimes 3}$ is very ample. In this case, we know that $\dim \mathbf{H}^0(E, L_1) = 1$. Hence by the theorem we have $\deg \Lambda(L_1) = 1$, which in particular says that $\Lambda(L_1) : E \to E^\vee$ is an isomorphism. As such, we may canonically identify E with $E^\vee$ via $\Lambda(L_1)$.

Remark. We know that $\Lambda(L_1)$ is a group homomorphism, and on k-points it is given by

$$E(k) \longrightarrow E^\vee(k), \qquad P \longmapsto t_P^*(\mathcal{O}(e)) \otimes \mathcal{O}(e)^{-1} = \mathcal{O}([P] - [e]).$$

Hence we once again see that the group structure on $E(k)$ is given by the following rule: We have $P + Q = R$ if and only if $\mathcal{O}([P] - [e]) \otimes \mathcal{O}([Q] - [e]) \cong \mathcal{O}([R] - [e])$, i.e., $[P] + [Q] - 2[e]$ is linearly equivalent to $[R] - [e]$.

On an elliptic curve E, consider more generally the line bundles $L_n := \mathcal{O}(n[e]) = L_1^{\otimes n}$. Then $\Lambda(L_n) = n \cdot \Lambda(L_1) = [n] \circ \Lambda(L_1)$. If we use $\Lambda(L_1)$ to canonically identify E and $E^\vee$, then $\Lambda(L_n)$ is identified with $[n]$. Recall that by Riemann–Roch, we have $\dim \mathbf{H}^0(E, L_n) = n$. We have also seen that $\deg[n] = n^2$. This verifies the theorem for L_n.

To prove Theorem 26.1.1 in general we need some preparations.

26.2. **The Poincaré line bundle.** Suppose A/S is a projective ablian scheme where S is noetherian. Recall that $\mathrm{Pic}_{A/S}$ represents the rigidified Picard functor

$$\mathrm{Pic}_{A/S,e} : (\text{locally noetherian } S\text{-schemes}) \longrightarrow (\text{Abelian groups})$$

given by

$$T \longmapsto \{(L,\rho) \mid L \text{ a line bundle on } A_T = A \times_S T,\ \rho : e_T^* L \xrightarrow{\sim} \mathcal{O}_T\}/\cong.$$

Over $A_{\mathrm{Pic}_{A/S}} = A \times_S \mathrm{Pic}_{A/S}$, we have a universal pair (L,ρ) where L is a line bundle on $A_{\mathrm{Pic}_{A/S}}$ and $\rho : (e,\mathrm{id})^* L \xrightarrow{\sim} \mathcal{O}_{\mathrm{Pic}_{A/S}}$. This pair is unique up to isomorphism. We may restrict (L,ρ) to $A \times_S A^\vee$. In particular, we acquire the so-called Poincaré line bundle $\mathcal{P}$ on $A \times_S A^\vee$, which comes equipped with a trivialization along (e,id).

Remark. Let $e^\vee$ be the neutral section of $A^\vee \to S$. Then $(\mathrm{id}, e^\vee)^*\mathcal{P}$ on A is also equipped with a trivialization and this trivialization is compatible with the previous trivialization in the sense that these two induce the same isomorphism

$$(e,e^\vee)^*\mathcal{P} \xrightarrow{\sim} \mathcal{O}_S.$$

Remark. We have the "flipping" identification $f : A^\vee \times A \xrightarrow{\sim} A \times A^\vee, (x,y) \mapsto (y,x)$. Notice that $f^*\mathcal{P}$ is a line bundle on $A^\vee \times A$ again equipped with two compatible trivializations along $(e^\vee, \mathrm{id})$ and (id, e). We can use this structure to obtain a canonical map $A \to A^{\vee\vee}$, which turns out to be an isomorphism.

Remark. The role played by the Poincaré line bundle in the duality theory for abelian varieties is analogous to the evaluation morphism $\mathrm{ev} : V \otimes_k V^\vee \to k$ for a vector space V over k.

Remark. Given an abelian variety A/k where $k = \overline{k}$, we can pick (non-canonically) an ample line bundle L to acquire an isogeny $\Lambda(L) : A \to A^\vee$. This in particular implies that

$$\dim A^\vee = \dim A.$$

We will study the cohomology of $A \times A^\vee$ with coefficients in the sheaf $\mathcal{P}$ as an intermediate step in attacking Theorem 26.1.1.

Theorem 26.2.1. *Suppose A is an abelian variety over an algebraically closed field k with dimension g. We have*

$$\mathbf{H}^n(A \times_k A^\vee, \mathcal{P}) = \begin{cases} 0, & n \neq g \\ k, & n = g. \end{cases}$$

Proof. In the following we write $A \times A^\vee$ for $A \times_k A^\vee$. We consider $R^n p_{2,*}\mathcal{P}$ where $p_2 : A \times A^\vee \to A^\vee$ is the second projection. Note that for $x \in A^\vee(k) \setminus \{e^\vee\}$, the isomorphism class of the line bundle $\mathcal{P}|_{A\times\{x\}}$ on A tautologically corresponds to $x \in \mathrm{Pic}(A)$. Since we are assuming that x is nontrivial, this line bundle must also be nontrivial and algebraically equivalent to 0. Hence, for such x, by Corollary 23.1.1 and Lemma 23.2.2, we have

$$\mathbf{H}^n(A \times \{x\}, \mathcal{P}|_{A\times\{x\}}) = 0, \qquad \forall n \geq 0.$$

In particular, for all $n \geq 0$, we know that $R^n p_{2,*}\mathcal{P}$ is supported at the neutral section $e^\vee$. By the Leray spectral sequence, we have

$$E_2^{p,q} = \mathbf{H}^p(A^\vee, R^q p_{2,*}\mathcal{P}) \implies \mathbf{H}^{p+q}(A \times_k A^\vee, \mathcal{P}).$$

Now $\mathbf{H}^p(A^\vee, R^q p_{2,*}\mathcal{P}) = 0$ except when $p = 0$, since $R^q p_{2,*}\mathcal{P}$ is supported on a finite scheme. This yields

$$\mathbf{H}^0(A^\vee, R^n p_{2,*}\mathcal{P}) \cong \mathbf{H}^n(A \times A^\vee, \mathcal{P})$$

for all n. If $n > g$, then $R^n p_{2,*}\mathcal{P} = 0$ as p_2 has relative dimension g. Hence $\mathbf{H}^n(A\times A^\vee, \mathcal{P}) = 0$ for $n > g$. Now notice that

$$\omega_{A\times A^\vee} = \wedge^{2g}\Omega_{A\times A^\vee/k}$$

is a trivial line; this is because $\Omega_{A\times A^\vee/k}$ is a trivial vector bundle already by the fact that $A\times A^\vee$ is an abelian variety (see [Mum08, §4 (iii)]). Hence by Serre duality applied to the $2g$-dimensional $A\times A^\vee$, we have

$$\mathbf{H}^n(A\times A^\vee, \mathcal{P}) \cong \mathbf{H}^{2g-n}(A\times A^\vee, \mathcal{P}^{-1})^\vee.$$

In particular, we know that $H^i(A\times A^\vee, \mathcal{P}^{-1})^\vee = 0$ for $i < g$. It is an easy exercise to show that

$$\mathcal{P}^{-1} \cong ([1],[-1])^*\mathcal{P}.$$

Since $([1],[-1]) : A\times A^\vee \to A\times A^\vee$ is an automorphism, this implies that $\mathcal{P}$ and $\mathcal{P}^{-1}$ have the same cohomology. Hence $\mathbf{H}^n(A\times A^\vee, \mathcal{P}) = 0$ for all $n < g$ as well. It remains to show that $\mathbf{H}^g(A\times A^\vee, \mathcal{P}) = k$. □

27. Lecture 27

27.1. The Poincaré line bundle, continued.

Proof of Theorem 26.2.1, continued. Let $p_2 : A\times A^\vee \to A^\vee$ be the second projection. By examining the cohomology of the fibers, we showed that $R^n p_{2,*}\mathcal{P}$ is supported at $e^\vee$. This implies that for $i \geq 1$, we have $\mathbf{H}^i(A^\vee, R^n p_{2,*}\mathcal{P}) = 0$. As such, the Leray spectral sequence

$$E_2^{i,j} = \mathbf{H}^i(A^\vee, R^j p_{2,*}\mathcal{P}) \implies \mathbf{H}^{i+j}(A\times A^\vee, \mathcal{P})$$

degenerates, and we have

$$\mathbf{H}^0(A^\vee, R^n p_{2,*}\mathcal{P}) \cong \mathbf{H}^n(A\times A^\vee, \mathcal{P}). \tag{27.1}$$

These vanish for $n > g$, as p_2 has relative dimension g. Now notice that for any group variety B over k, we have $\omega_{B/k} = \wedge^{\dim B}\Omega_{B/k} \cong \mathcal{O}_B$, as $\Omega_{B/k}$ is a trivial vector bundle. As such, for any line bundle L on any abelian variety B, we have $\mathbf{H}^i(B,L) \cong \mathbf{H}^{\dim B - i}(B, L^{-1})^\vee$ by Serre duality. Setting $B = A\times A^\vee$, we have

$$\mathbf{H}^n(A\times A^\vee, \mathcal{P}^{-1}) \cong \mathbf{H}^{2g-n}(A\times A^\vee, \mathcal{P})^\vee,$$

and we have already seen that the right hand side vanishes $n < g$. But $\mathcal{P}$ and $\mathcal{P}^{-1}$ have the same cohomology, as $\mathcal{P}^{-1} \cong ([1],[-1])^*\mathcal{P}$, where $([1],[-1]) : A\times A^\vee \to A\times A^\vee$ is the automorphism given by $(x,y)\mapsto (x,-y)$. In particular, this implies that for $n < g$, we have

$$\mathbf{H}^n(A\times A^\vee, \mathcal{P}) = 0.$$

Now since $R^n p_{2,*}\mathcal{P}$ is supported on a finite scheme, it is determined by its global sections. Thus by (27.1) and the vanishing of $\mathbf{H}^n(A\times A^\vee, \mathcal{P})$ for $n \neq g$, we know that $R^n p_{2,*}\mathcal{P} = 0$ for $n \neq g$.

It remains to show that $\mathbf{H}^g(A\times A^\vee, \mathcal{P}) = k$. For this, let $R = \mathcal{O}_{A^\vee, e^\vee}$ be the local ring of $A^\vee$ at the closed point $e^\vee$, and denote by $\mathfrak{m}$ the maximal ideal of R. Let M be the (finitely generated) R-module corresponding to the pullback of $R^g p_{2,*}\mathcal{P}$ to $\operatorname{Spec} R$. By (27.1), it is enough to show that $M \cong R/\mathfrak{m} = k$ as R-modules.

Now, the vanishing of $R^n p_{2,*}\mathcal{P}$ for all $n > g$ implies by cohomology and base change that $R^g p_{2,*}\mathcal{P}$ is compatible with fiber cohomology in the sense that for all $x \in A^\vee$, the natural base change map

$$(R^g p_{2,*}\mathcal{P}) \otimes_{\mathcal{O}_{A^\vee}} k(x) \longrightarrow \mathbf{H}^g(A\times \{x\}, \mathcal{P}|_{A\times\{x\}})$$

is an isomorphism of $k(x)$-vector spaces. (For this, use [Conb, Thm. 1.1] and reverse induction to show that all the fiber cohomology vanish at degrees $> g$. Then feed the vanishing of fiber cohomology at degree $g+1$ into [Conb, Cor. 1.2].) For $x = e^\vee$, the right hand side is just $\mathbf{H}^g(A, \mathcal{O}_A)$, and by Serre duality this is dual to $\mathbf{H}^0(A, \mathcal{O}_A^{-1}) = k$. Hence $M \otimes_R R/\mathfrak{m} = k$. It follows that M is generated by one element as an R-module by Nakayama's lemma. Write $M \cong R/J$. On the other hand, M is finite-length as an R-module, i.e., M is a finite dimensional k-vector space; this follows from the finite-dimensionality of $\mathbf{H}^0(A^\vee, R^g p_{2,*}\mathcal{P})$. Hence we have (exercise)

$$\sqrt{J} = \mathfrak{m}.$$

We will show that $J = \mathfrak{m}$ using Grothendieck duality, following [EvdGM, §9]. In view of the fact that p_2 is proper smooth of relative dimension g and the fact that $\Omega^g_{A\times A^\vee/A^\vee} \cong \mathcal{O}_{A\times A^\vee}$ (again because $\Omega_{A\times A^\vee/A^\vee}$ is a trivial vector bundle), Grothendieck duality says that for any bounded complexes $F^\bullet$ and $G^\bullet$ of quasi-coherent $\mathcal{O}_{A\times A^\vee}$-modules and $\mathcal{O}_{A^\vee}$-modules respectively, there is a natural bijection

$$\mathrm{Hom}_{D(A\times A^\vee)}(F^\bullet, p_2^* G^\bullet[g]) \xrightarrow{\sim} \mathrm{Hom}_{D(A^\vee)}(Rp_{2,*}F^\bullet, G^\bullet).$$

Here for a variety X we write $D(X)$ for the derived category of the abelian category of quasi-coherent $\mathcal{O}_X$-modules. Applying the proposition to the case when $F^\bullet = \mathcal{P}[0]$ and $G^\bullet = G[-g]$ where G is an arbitrary quasi-coherent $\mathcal{O}_{A^\vee}$-module, we have

$$\mathrm{Hom}_{\mathcal{O}_{A\times A^\vee}}(\mathcal{P}, p_2^* G) \xrightarrow{\sim} \mathrm{Hom}_{\mathcal{O}_{A^\vee}}(R^g p_{2,*}\mathcal{P}, G). \tag{27.2}$$

In particular, we will take G to be the quasicoherent $\mathcal{O}_{A^\vee}$-module corresponding to the R-modules R/J and $R/\mathfrak{m}$ respectively. Then the right hand side of (27.2) is R/J and $R/\mathfrak{m}$ respectively in the two cases. We thus obtain a commutative diagram:

$$\begin{tikzcd}
\mathrm{Hom}_{\mathcal{O}_{A\times A^\vee}}(\mathcal{P}, \mathcal{O}_{A\times A^\vee} \otimes_{\mathcal{O}_{A^\vee}} R/J) \arrow[r, "\cong"] \arrow[d] & R/J \arrow[d] \\
\mathrm{Hom}_{\mathcal{O}_{A\times A^\vee}}(\mathcal{P}, \mathcal{O}_{A\times A^\vee} \otimes_{\mathcal{O}_{A^\vee}} R/\mathfrak{m}) \arrow[r, "\cong"'] & R/\mathfrak{m},
\end{tikzcd}$$

where the horizontal arrows are bijections, and the right vertical arrow is induced by id_R and hence surjective. Hence the left vertical arrow is also surjective. Now, in the lower left set, there is a canonical element ϕ coming from the trivialization of $\mathcal{P}$ along $(\mathrm{id}_A, e^\vee)$. More precisely, we have a canonical isomorphism $\bar{\phi} : (\mathrm{id}_A, e^\vee)^*\mathcal{P} \xrightarrow{\sim} \mathcal{O}_A$ as part of the structure of the Poincaré line bundle, and this is precisely the datum of a morphism $\phi : \mathcal{P} \longrightarrow \mathcal{O}_{A\times A^\vee} \otimes_{\mathcal{O}_{A^\vee}} R/\mathfrak{m}$. By the surjectivity of the left vertical arrow, there exists a morphism $\tilde{\phi} : \mathcal{P} \longrightarrow \mathcal{O}_{A\times A^\vee} \otimes_{\mathcal{O}_{A^\vee}} R/J$ lifting ϕ . Now $\tilde{\phi}$ corresponds to the datum of a morphism

$$\bar{\tilde{\phi}} : \mathcal{P}|_{(A\times A^\vee)\times_{A^\vee}\mathrm{Spec}\, R/J} \to \mathcal{O}_{(A\times A^\vee)\times_{A^\vee}\mathrm{Spec}\, R/J}.$$

Note that $\bar{\tilde{\phi}}$ lifts $\bar{\phi}$ (i.e., it induces $\bar{\phi}$ if we base change along $\mathrm{Spec}\, R/\mathfrak{m} \to \mathrm{Spec}\, R/J$). Since $J = \sqrt{\mathfrak{m}}$, the natural map

$$(A \times A^\vee) \times_{A^\vee} \mathrm{Spec}\, R/\mathfrak{m} = A \times \mathrm{Spec}\, R/\mathfrak{m} \to A \times \mathrm{Spec}\, R/J = (A \times A^\vee) \times_{A^\vee} \mathrm{Spec}\, R/J$$

induces a homeomorphism of underlying topological spaces. Hence, since $\bar{\phi}$ is an isomorphism, we know $\bar{\tilde{\phi}}$ is an isomorphism as well. By the moduli interpretation of $(A^\vee, \mathcal{P})$, the fact that there exists a trivialization of $\mathcal{P}|_{A\times \mathrm{Spec}\, R/J}$ lifting $\bar{\phi}$ implies that $\mathrm{Spec}\, R/J \to A^\vee$

factors through $e^\vee : \operatorname{Spec} k \to A^\vee$. As such, $\mathfrak{m} = J$. This concludes the proof of Theorem 26.2.1. □

28. Lecture 28

28.1. **Generalities on G-torsors.** We now use Theorem 26.2.1 to prove Theorem 26.1.1. We will first discuss some general facts about torsors.

Let k ve a field and G be a finite group scheme over k (not necessarily commutative). Let X be a finite-type k-scheme. (In particular, X is noetherian.)

Definition 28.1.1. A **G-torsor over X** is a scheme Y over X with a G-action on Y (i.e., there is a k-scheme map $G \times_k Y \to Y$ such that for all k-schemes T, we have a group action $G(T) \times Y(T) \to Y(T)$) such that

(1) The structure map $Y \to X$ is finite flat and surjective, as well as G-equivariant where G acts trivially on X.
(2) The map $G \times_k Y \to Y \times_X Y$ defined via $(g, y) \mapsto (gy, y)$ is an isomorphism. (The fact that this map is well defined follows from the G-equivariance of $Y \to X$.)

Proposition 28.1.2. *Suppose $\phi : Y \to X$ is a G-torsor in the language above. For any coherent sheaf $\mathcal{F}$ on X, we have*

$$\chi(\phi^*\mathcal{F}) = \deg \phi \cdot \chi(\mathcal{F}).$$

Proof. The proof proceeds via dévissage and noetherian induction, and in particular needs that X is noetherian. See [Cond, Lem. 7.3.2] or [Mum08, §12, Thm. 2] for a proof. □

Lemma 28.1.3. *Let $\phi : A \to B$ be an isogeny between abelian varieties over a field k. For all coherent sheaves $\mathcal{F}$ on B, we have $\chi(\phi^*\mathcal{F}) = \deg \phi \cdot \chi(\mathcal{F})$.*

Proof. The morphism $\phi : A \to B$ is a G-torsor, where $G = \ker \phi$. Then we apply Proposition 28.1.2. □

28.2. **Global sections of ample line bundles, continued.** Now we prove Theorem 26.1.1. To recall our setup, let A be an abelian variety over an algebraically closed field k and let L be an ample line bundle on A.

Proof of Theorem 26.1.1. Recall from Corollary 25.1.2 that $\mathbf{H}^i(A, L) = 0$ for all $i > 0$. Hence it suffices to show that

$$\chi(L) = \sqrt{\deg \Lambda(L)}.$$

Recall that we have three morphisms $\mu, p_1, p_2 : A \times A \to A$; here μ is the group law and p_i is the projection to the i-th factor. The idea is to relate the Mumford line bundle $\mathfrak{M}(L) = \mu^*L \otimes p_1^*L^{-1} \otimes p_2^*L^{-1}$ on $A \times A$ with $\mathcal{P}$ on $A \times A^\vee$. It is an exercise to see that for the morphism $(\mathrm{id}_A, \Lambda(L)) : A \times A \to A \times A^\vee$, we have $\mathfrak{M}(L) \cong (\mathrm{id}_A, \Lambda(L))^*\mathcal{P}$. Note that $\deg(\mathrm{id}_A, \Lambda(A)) = \deg \Lambda(L)$. By Theorem 26.2.1, we have $\chi(\mathcal{P}) = (-1)^g$, where $g = \dim A$. By Lemma 28.1.3, we have

$$\chi(\mathfrak{M}(L)) = \deg(\Lambda(L))\chi(\mathcal{P}) = (-1)^g \deg(\Lambda(L)).$$

We now need to relate $\chi(L)$ with $\chi(\mathfrak{M}(L))$. We will examine the higher direct images of $\mathfrak{M}(L)$ along $p_1 : A \times A \to A$. If $x \in A(k) \setminus \ker(\Lambda(L))$, then $\mathfrak{M}(L)|_{\{x\} \times A}$ is nontrivial, but $\Lambda(\mathfrak{M}(L)|_{\{x\} \times A}) = 0$. Hence, we know that

$$\mathbf{H}^i(A, \mathfrak{M}(L)|_{\{x\} \times A}) = 0$$

for all i, by Lemma 23.2.2. As such, we know that for each i, the higher direct image $R^i p_{1,*}\mathfrak{M}(L)$ is supported on $\ker \Lambda(L)$, which is finite since L is ample.

By the projection formula ([Sta18, Tag 01E8]), we have

$$R^i p_{1,*}\mathfrak{M}(L) = R^i p_{1,*}(\mu^* L \otimes p_2^* L^{-1} \otimes p_1^* L^{-1}) \cong R^i p_{1,*}(\mu^* L \otimes p_2^* L^{-1}) \otimes_{\mathcal{O}_A} L^{-1}.$$

Now, since L^{-1} is a line bundle and the left hand side is finitely supported, we know that $R^i p_{1,*}(\mu^* L \otimes p_2^* L^{-1})$ must be finitely supported and isomorphic to the left hand side. Thus

$$R^i p_{1,*}\mathfrak{M}(L) \cong R^i p_{1,*}(\mu^* L \otimes p_2^* L^{-1}).$$

By the Leray spectral sequence, we have

$$\chi(\mathfrak{M}(L)) = \chi(\mu^* L \otimes p_2^* L^{-1}).$$

Now, we have an automorphism $(\mu, p_2) : A \times A \to A \times A$ given by $(x, y) \mapsto (x + y, y)$, and it is easy to see that

$$(\mu, p_2)^*(p_1^* L \otimes p_2^* L^{-1}) \cong \mu^* L \otimes p_2^* L^{-1}.$$

This yields

$$\chi(\mathfrak{M}(L)) = \chi(p_1^* L \otimes p_2^* L^{-1}).$$

By the Künneth formula, the right hand side is $\chi(L) \cdot \chi(L^{-1})$. Finally, by Serre duality, we have $\mathbf{H}^i(A, L) \cong \mathbf{H}^{g-i}(A, L^{-1})^\vee$, and so

$$\chi(L^{-1}) = (-1)^g \cdot \chi(L).$$

Combining the results we have

$$\chi(\mathfrak{M}(L)) = (-1)^g \deg \Lambda(L) = (-1)^g \chi(L)^2,$$

which concludes the proof. □

29. Lecture 29

29.1. **Very ample line bundles.** Let A be an abelian variety over an algebraically closed field k. Let L be an ample line bundle on A. We have shown that

$$h^i(A, L) = \begin{cases} 0 & , i > 0, \\ \sqrt{\deg \Lambda(L)} & , i = 0. \end{cases}$$

This recovers our knowledge of $h^i(E, \mathcal{O}(ne))$ for E and elliptic curve and $n \geq 1$. (Recall that if we identify E with $E^\vee$ via the isomorphism $\Lambda(\mathcal{O}(e))$ then $\Lambda(\mathcal{O}(ne))$ is identified with $[n]$.) In the case of an elliptic curve, we also know that $\mathcal{O}(3e)$ is very ample, i.e., the third power of the ample line bundle $\mathcal{O}(e)$ is very ample. This phenomenon has the following generalization.

Theorem 29.1.1. *Let A be an abelian variety over an algebraically closed field k, and let L be an ample line bundle on A. Then L^3 is very ample.*

Sketch of proof. For the detailed proof see [Mum08, §17].

(1) Every effective divisor on A is linearly equivalent to a multiplicity-free effective divisor. To see this, consider the simple situation with a divisor of the form $n\mathfrak{p}$, where $n \geq 2$ and $\mathfrak{p}$ is a prime divisor. By the Theorem of Square, for any points $x_1, \cdots x_n \in A(k)$ such that $\sum x_i = 0$, we have that $\sum t_{x_i}^* \mathfrak{p}$ is linearly equivalent to $n\mathfrak{p}$, and is multiplicity free as long as the x_i's are "generic enough". The general situation is treated similarly.

(2) Since $h^0(A, L) > 0$ (as L is ample), we have $L \cong \mathcal{O}(D)$ for some effective divisor D. By (1), we may assume that D is multiplicity-free.

(3) We need to show that the linear system $|3D|$ separates points. Suppose not. By translation, we may assume that $|3D|$ does not separate $0 \in A(k)-$ and some other point $x_1 \in A(k), x_1 \neq 0$. This means that for each $D' \in |3D|$, we have $0 \in \mathrm{Supp}(D') \Rightarrow x_1 \in \mathrm{Supp}(D')$. Now observe that for any $x, y \in A(k)$, we have

$$t_x^* D + t_y^* D + t_{-x-y}^* D \in |3D|$$

by the Theorem of Square. Thus for any x, y we have

$$0 \in \mathrm{Supp}(t_x^* D) \Longrightarrow x_1 \in \mathrm{Supp}(t_x^* D) \cup \mathrm{Supp}(t_y^* D) \cup \mathrm{Supp}(t_{-x-y}^* D).$$

Now for each given x, we can choose y such that

$$x_1 \notin \mathrm{Supp}(t_y^* D) \cup \mathrm{Supp}(t_{-x-y}^* D).$$

Hence we conclude that

$$\forall x \in A(k), 0 \in \mathrm{Supp}(t_x^* D) \Longrightarrow x_1 \in \mathrm{Supp}(t_x^* D).$$

From this it is easy to see that $\mathrm{Supp}(D)$ is invariant under t_{x_1}. Since D is multiplicity free, we conclude that D (as a divisor) is invariant under t_{x_1}.

In particular, we have $x_1 \in \ker(\Lambda(L))$, which is a finite k-group scheme as L is ample. Let F be the k-subgroup scheme generated by x_1. Thus F is a finite étale (and hence constant) group scheme over k; the constant value of F is just the subgroup of $A(k)$ generated by x_1. We can form the quotient A/F, which is still an abelian variety, and it is equipped with an isogeny $A \to A/F$ which is an étale F-torsor. (See [Mum08, §7, Thm. 4], and cf. [Mum08, §12].[7]) The fact that D is invariant under t_{x_1} and multiplicity-free implies that D is the pullback of a divisor D_1 on A/F. Moreover, every $s \in \mathbf{H}^0(A, L)$ whose zero divisor $(s)_0$ equals D comes from a global section of $\mathcal{O}(D_1)$ on A/F.

Now as a general fact about finite étale torsors, there are only finitely many isomorphism classes of line bundles $L_1, \cdots, L_m$ on A/F whose pullback to A are isomorphic to L. (The set of L_i's is in bijection with the set of all characters $F \to k^\times$; see [Mum08, §7, Prop. 3].) Let $s \in \mathbf{H}^0(A, L)$. If $(s)_0$ is multiplicity-free, then we can take D to be $(s)_0$ in the previous discussion, and conclude that s comes from a global section of some L_i. Now each L_i is still ample on A/F since L is ample on A, and therefore we have

$$h^0(A/F, L_i) = \chi(L_i) = \frac{1}{\deg(A \to A/F)}\chi(L) = \frac{1}{|F|}\chi(L) < \chi(L) = h^0(A, L).$$

Here $|F| > 1$ because $x_1 \neq 0$. Thus we have

$$\mathbf{H}^0(A, L) = V_1 \cup \cdots \cup V_m \cup W,$$

where each V_i is a proper vector subspace of $\mathbf{H}^0(A, L)$, and W is the set of $s \in \mathbf{H}^0(A, L)$ such that $(s)_0$ is not multiplicity-free. In the next lecture we will show that such a decomposition is impossible unless $W = \mathbf{H}^0(A, L)$. Since we have already seen that $W \neq \mathbf{H}^0(A, L)$, the proof is finished.

(4) We need to show that $|3D|$ separates tangent vectors. The argument is similar to (3) but more technical. We refer the reader to [Mum08, §17]. □

[7] Note that in the language of [Mum08, §7], a finite subgroup of A means a finite subgroup of $A(k)$, which is identified with the corresponding constant k-group scheme.

29.2. **Globalization.** Let S be a noetherian scheme, and $\pi : A \to S$ a projective abelian scheme. Let L be a line bundle on A that is relatively ample with respect to π. Recall [Gro61a, §4.6] that this means there exists a Zariski open covering (U_i) of S such that $L|_{\pi^{-1}(U_i)}$ is ample for each i. In the current setting, since S is noetherian and π is proper, this condition is also equivalent [Gro61b, Thm. 4.7.1] to the condition that the pullback of L to each fiber of π is ample.

Theorem 29.2.1 (cf. [MFK94, Prop. 6.13])**.** *The following statements hold.*

(1) $R^i\pi_*L = 0$ *for all* $i > 0$ *.*
(2) π_*L *is a vector bundle on* S*, and its formation commutes with arbitrary base change (cf. Theorem 10.2.1 (1)). Let* r *be its rank.*
(3) *The* S*-homomorphism* $\Lambda(L) : A \to A^\vee$ *is an isogeny, and its degree is* r^2*.*
(4) L^3 *is relatively very ample, i.e., the canonical* S*-map* $A \to \mathbb{P}_S(\pi_*L^3)$ *is a closed immersion.*

Proof. By cohomology and base change, the first two statements follow from Corollary 25.1.2 and the fact that L is fiberwise ample. In particular, we know that r is equal to the dimension of $\mathbf{H}^0(A_s, L|_{A_s})$ for any geometric point s of S. In view of this, the third statement is easily reduced to the corresponding statement for the fibers, in which case it is Theorem 26.1.1. The fourth statement is reduced to the corresponding statement for the fibers by [Gro61b, Prop. 4.6.7] and the fact that the formation of $\pi_*(L^3)$ commutes with arbitrary base change (by (2) applied to L^3). □

Remark. By (2) applied to L^3, we may Zariski localize S and fix a trivialization $\pi_*L^3 \xrightarrow{\sim} \mathcal{O}_S^{\oplus n}$. Then by (4) we obtain a closed immersion of S-schemes $A \to \mathbb{P}_S^{n-1}$.

Definition 29.2.2. By a **polarization on** A, we mean an S-group homomorphism $\lambda : A \to A^\vee$ such that for each geometric point $x : \operatorname{Spec} \bar{k} \to S$, the induced homomorphism

$$\lambda_x : A_x \longrightarrow (A^\vee)_x \cong (A_x)^\vee$$

(from the abelian variety A_x over $\bar{k}$ to its dual) is of the form $\Lambda(L_x)$ for some ample line bundle L_x on A_x. Note that a polarization is necessarily an isogeny, since each $\lambda_x = \Lambda(L_x)$ is surjective and quasi-finite by the ampleness of L_x.

30. Lecture 30

30.1. **Multiplicities of divisors in a linear system.** The following result is needed in the proof of Theorem 29.1.1. In [Mum08, §17] this result is taken for granted, but we find it not so straightforward.

Lemma 30.1.1. *Let* X *be a smooth projective variety over an algebraically closed field* k*. Let* L *be a line bundle on* X*. Let* $V = \mathbf{H}^0(X, L)$*, and* $W = \{s \in V \mid (s)_0$ *is not multiplicity-free*$\}$*. Suppose* $V \neq 0$ *and* $V \neq W$*. Then it is impossible to have*

$$V = V_1 \cup \cdots \cup V_m \cup W$$

where each V_i *is a proper vector subspace of* V*.*

Sketch of proof. It suffices to show that the set of $s \in V - \{0\}$ such that $(s)_0$ is multiplicity-free contains a non-empty Zariski open subset of V (viewed as the maximal spectrum of an affine space).

Write M for the k-scheme such that $M(k) = V - \{0\}$ (i.e., $M = \mathbb{A}_k^n - \{0\}$ for some n). Then the set-theoretical map sending every $s \in M(k)$ to the divisor $(s)_0$ in X can be upgraded to a geometric family over M. More precisely, inside $X \times_k M$ we have a closed subscheme D, which is the zero scheme of the tautological global section of $p_1^* L$ on $X \times_k M$. For $s \in M(k)$, the fiber of D over s is identified with the zero scheme of s inside $X \cong X \times \{s\}$.

Given this construction, for $s \in M(k)$, the condition that $(s)_0$ is multiplicity-free is equivalent to the condition that the fiber D_s is reduced. One can show that D is flat over M, in the same way as the argument showing that $q^{-1}(U) \to U$ is flat in the proof of [Sta18, Tag 0FD6]. It now suffices to cite the general fact [Sta18, Tag 0C0E] that for a proper flat morphism $Y \to S$ of finite presentation, the set of $s \in S$ such that Y_s is geometrically reduced is a Zariski open subset of S (which may be empty). □

30.2. **The line bundle attached to a polarization.** Let S be a noetherian scheme, and $\pi : A \to S$ a projective abelian scheme. In the last lecture we defined the notion of a polarization of A, which is a special kind of S-homomorphism $A \to A^\vee$. We now define a general construction that takes a homomorphism $A \to A^\vee$ to a line bundle on A.

Suppose $\lambda : A \to A^\vee$ is an S-homomorphism. Consider the map $(\mathrm{id}_A, \lambda) : A \to A \times_S A^\vee$. We set

$$L^\Delta(\lambda) := (\mathrm{id}_A, \lambda)^* \mathcal{P}$$

where $\mathcal{P}$ is the Poincaré line bundle. Equivalently, the datum of λ gives rise to a line bundle on $A_A = A \times_S A$ trivialized along $(e \circ \pi, \mathrm{id}) : A \to A \times_S A$ by the moduli interpretation of $\mathrm{Pic}_{A/S} \cong \mathrm{Pic}_{A/S,e}$, and $L^\Delta(\lambda)$ is the pullback of this line bundle under the diagonal $\Delta : A \to A \times_S A$. Note that according to either definition, $L^\Delta(\lambda)$ is equipped with a canonical trivialization along e, i.e., an isomorphism $\rho_{\mathrm{can}} : e^* L^\Delta(\lambda) \xrightarrow{\sim} \mathcal{O}_S$.

Proposition 30.2.1 ([MFK94, Prop. 6.10])**.** *Let λ be a polarization on A. Then*

$$\Lambda(L^\Delta(\lambda)) = 2\lambda.$$

Proof. Since the construction $\lambda \mapsto \Lambda(L^\Delta(\lambda))$ commutes with base change, by the Rigidity Lemma we reduce to the case where S is the spectrum of an algebraically closed field k. Since λ is a polarization and $S = \operatorname{Spec} k$, we have $\lambda = \Lambda(L)$ for some line bundle L on A, and we know that λ is an isogeny. We need to show that $\Lambda(L^\Delta(\lambda)) = 2\Lambda(L)$. Since the right hand side is $\Lambda(L^2)$, it suffices to show that $L^\Delta(\lambda)$ is algebraically equivalent to L^2 in view of Corollary 23.1.1. Since $\lambda = \Lambda(L)$, unraveling the definition one sees that $L^\Delta(\lambda)$ is the pullback of the Mumford line bundle $\mathfrak{M}(L)$ on $A \times_k A$ along $\Delta : A \to A \times_k A$. This can be explicitly computed to be $[2]^* L \otimes L^{-2}$. Hence it suffices to note that $[2]^* L$ is algebraically equivalent to L^4, by Corollary 25.1.1. □

Definition 30.2.2. We say that an isogeny $\lambda : A \to A^\vee$ is **symmetric** if it satisfies $\Lambda(L^\Delta(\lambda)) = 2\lambda$. Thus Proposition 30.2.1 says that polarizations are special examples of symmetric isogenies. (Recall that a polarization is automatically an isogeny.)

Lemma 30.2.3. *Let $\lambda : A \to A^\vee$ be an isogeny, and $n \geq 1$. Then λ is symmetric if and only if $n\lambda$ is symmetric.*

We postpone the proof to the next lecture.

Proposition 30.2.4 ([MFK94, Prop. 6.11][8])**.** *Fix a line bundle L on A together with a trivialization $\rho : e^* L \xrightarrow{\sim} \mathcal{O}_S$. Fix a positive integer n. Assume that $\Lambda(L) : A \to A^\vee$ is an isogeny (which is true, e.g., when L is relatively ample with respect to $A \to S$). The following statements hold.*

[8] In this reference, the condition of symmetric isogeny in statement (1) is missing, and the proof is invalid without this condition.

(1) *For any S-scheme T (always assumed to be noetherian), there exists at most one symmetric isogeny $\lambda : A_T \to A_T^\vee$ such that the base change to T of the pair (L, ρ) is isomorphic to the pair $(L^\Delta(n\lambda), \rho_{\mathrm{can}})$.*
(2) *There exists a closed (necessarily unique) subscheme $S_0 \subset S$ such that for every S-scheme T, the existence of λ in (1) holds if and only if $T \to S$ factors through S_0.*

Sketch of proof. **(1)** We show that the uniqueness of λ holds even if we just assume that the naked line bundles $L|_{A_T}$ and $L^\Delta(n\lambda)$ are isomorphic. Thus suppose λ and λ' are two symmetric isogenies $A_T \to A_T^\vee$ such that $L^\Delta(n\lambda) \cong L^\Delta(n\lambda')$. Since $n\lambda$ and $n\lambda'$ are symmetric by Lemma 30.2.3, we have

$$2n\lambda = \Lambda(L^\Delta(n\lambda)) = \Lambda(L^\Delta(n\lambda')) = 2n\lambda'.$$

To show that $\lambda = \lambda'$, we may pass to the case where S is the spectrum of an algebraically closed field by Rigidity Lemma. Then the homomorphism $\lambda - \lambda' : A \to A^\vee$ sends the connected A to the finite group scheme $A^\vee[2n]$, and hence is constantly 0. Thus $\lambda = \lambda'$.

(2) Firstly, a necessary condition for the existence of λ is that $2n\lambda = \Lambda(L)$ (since $n\lambda$ is symmetric by Lemma 30.2.3). Note that $2n\lambda = [2n]_{A^\vee} \circ \lambda = \lambda \circ [2n]_A$. Thus we need $\Lambda(L) : A \to A^\vee$ to factor through $[2n]_A$. Now $[2n]_A : A \to A$ is a K-torsor where K is the finite flat S-group scheme $\ker([2n]_A)$. By flat descent, $\Lambda(L)$ factors through $[2n]_A$ if and only if $\Lambda(L)|_K$ is "constantly zero", i.e., equal to $e^\vee \circ \pi : K \xrightarrow{\pi} S \xrightarrow{e^\vee} A^\vee$. Of course without a suitable base change there is no reason for this to hold. Let K' be the kernel of $\Lambda(L)|_K$. Using the fact that K is finite flat over S, one can show that there exists a maximal closed subscheme $S_1 \subset S$ such that $K|_{S_1} = K'|_{S_1}$.

The desired S_0 is certainly contained in S_1, so we may replace S by S_1. Thus we may assume that $\Lambda(L)$ factors through $[2n]_A$, and so we can find a (necessarily unique) isogeny $\lambda : A \to A^\vee$ such that $\Lambda(L) = 2n\lambda$. Since we have assumed that $\Lambda(L)$ is an isogeny, the same proof as Proposition 30.2.1 shows that $\Lambda(L)$ is a symmetric isogeny. Therefore λ is a symmetric isogeny by Lemma 30.2.3. Now (L, ρ) and $(L^\Delta(\lambda), \rho_{\mathrm{can}})$ correspond to two sections of $\mathrm{Pic}_{A/S} \to S$. The desired locus S_0 is the maximal closed subscheme of S over which the two sections agree, which exists (maybe empty) since $\mathrm{Pic}_{A/S} \to S$ is separated. □

31. Lecture 31

31.1. More on symmetric isogenies.

Proof of Lemma 30.2.3. Recall that $L^\Delta(\lambda)$ is the pullback of $\mathcal{P}$ along

$$A \xrightarrow{\Delta} A \times A \xrightarrow{(\mathrm{id},\lambda)} A \times A^\vee,$$

and $L^\Delta(n\lambda)$ is the pullback of $\mathcal{P}$ along

$$A \xrightarrow{\Delta} A \times A \xrightarrow{(\mathrm{id},\lambda)} A \times A^\vee \xrightarrow{(\mathrm{id},[n]_{A^\vee})} A \times A^\vee.$$

By the moduli interpretation of the group law on $A^\vee$, we have $(\mathrm{id}, [n]_{A^\vee})^*\mathcal{P} \cong \mathcal{P}^n$. Hence $L^\Delta(n\lambda) \cong L^\Delta(\lambda)^n$, and $\Lambda(L^\Delta(n\lambda)) = n\Lambda(L^\Delta(\lambda))$. Thus the lemma follows from the fact that two isogenies $f, g : A \to B$ between abelian varieties are equal if $mf = mg$ for some $m \geq 1$. □

Proposition 31.1.1. *Let $\lambda : A \to A^\vee$ be a homomorphism. Then the following are equivalent. (In the literature one often finds (2) as the definition of a polarization.)*

(1) *λ is a polarization.*

(2) *λ is a symmetric isogeny, and $L^{\Delta}(\lambda)$ is relatively ample with respect to $A \to S$.*

Proof. All the three notions "polarization", "symmetric isogeny", "relatively ample" can be checked fiberwise, so we may assume that S is the spectrum of an algebraically closed field k. (Then relatively ample just means ample.)

(1) $\Rightarrow$ (2). By Proposition 30.2.1, λ is a symmetric isogeny. By definition $\lambda = \Lambda(L)$ for some ample L. Then by being symmetric we have $2\lambda = \Lambda(L^{\Delta}(\lambda))$. Now $2\lambda = \Lambda(L^2)$ and L^2 is ample. In order to show that $L^{\Delta}(\lambda)$ is ample, it suffices to show the following claim: A line bundle M on A is ample if $\Lambda(M) = \Lambda(M_0)$ for some ample M_0. To show the claim, note that $\Lambda(M \otimes M_0^{-1}) = 0$ and so we can find $x \in A(k)$ such that $M \otimes M_0^{-1} \cong \Lambda(M_0)(x) = t_x^* M_0 \otimes M_0^{-1}$. Then $M \cong t_x^* M_0$ is ample. (Exactly the same argument appeared in the proof of Corollary 25.1.2).

(2) $\Rightarrow$ (1). We admit the following fact ([Mum08, §23, Thm. 3]): For any line bundle M on A, if $A[m] \subset \ker(\Lambda(M))$, then M admits an m-th root, that is, there exists a line bundle N on A such that $N^m = M$. Now since $\Lambda(L^{\Delta}(\lambda)) = 2\lambda$, we have $L^{\Delta}(\lambda) = L^2$ for some line bundle L. Then $2\Lambda(L) = 2\lambda$, and so $\Lambda(L) = \lambda$. To finish the proof it suffices to show that L is ample. Since $\Lambda(L^2) = \Lambda(L^{\Delta}(\lambda))$ and $L^{\Delta}(\lambda)$ is ample, by the claim in the previous paragraph we know that L^2 is ample. Hence L is ample. $\square$

Remark. Our definition of a symmetric isogeny is not the standard definition, which involves the notion of dual isogenies and the identification $A \cong (A^\vee)^\vee$. According to the standard definition, λ is symmetric if $\lambda^\vee : (A^\vee)^\vee \cong A \to A^\vee$ is equal to λ. The two definitions are equivalent, because in general we have $\Lambda(L^{\Delta}(\lambda)) = \lambda + \lambda^\vee$; see the proof of [Lan13, Prop. 1.3.2.14].

Remark. There is an analogy between the duality of abelian schemes and the duality of vector spaces. Thus the notions of a homomorphism $A \to A^\vee$, an isogeny $A \to A^\vee$, a symmetric isogeny $A \to A^\vee$, and a polarization on A respectively correspond to, loosely speaking, a bilinear pairing on a vector space (say over $\mathbb{Q}$), a non-degenerate bilinear pairing, a symmetric non-degenerate bilinear pairing, and a positive definite symmetric bilinear pairing.

32. Lecture 32

32.1. **Automatic deformation of abelian group structure.**

Remark. In the above proof of Proposition 31.1.1, we used the fact that for any abelian variety A over an algebraically closed field k and any line bundle M on A, if $A[m] \subset \ker(\Lambda(M))$, then M admits an m-th root. The set of all m-th roots of M is a torsor under $\mathrm{Pic}_{A/k}[m](k)$, which is exactly $A^\vee[m](k)$, as $\pi_0(\mathrm{Pic}_{A/k})$ is torsion-free. Since $A^\vee$ is an abelian variety of the same dimension as A, one can study the possible number of k-points of $A[m]$. e.g., if $\mathrm{char}(k) = 0$, then we have $m^{2\dim A}$ points.

We briefly summarize some of the most important facts we have proven thus far.

1. ("Riemann–Roch") For L an ample line bundle on an abelian variety A over an algebraically closed field k, we have that $\mathbf{H}^0(A, L)$ has rank equal to $\sqrt{\deg \Lambda(L)}$ and $\mathbf{H}^i(A, L) = 0$ for $i > 0$.
2. For such L, we know that L^3 is very ample.
3. Suppose A/S is a projective abelian scheme and S is noetherian. Let (L, ρ) be a pair where L is a line bundle on A such that $\Lambda(L) : A \to A^\vee$ is an isogeny and $\rho : e^*L \xrightarrow{\sim} \mathcal{O}_S$ is an isomorphism. Given $n \geq 1$, there exists a (unique) closed

subscheme $S_0 \subset S$ such that for all S-schemes T, the morphism $T \to S$ factors through S_0 if and only if on T, we have $(L, \rho)_T \cong (L^{\Delta}(n\lambda), \rho_{\mathrm{can}})$ for a unique symmetric isogeny $\lambda : A_T \to A_T^{\vee}$.

We wish to add another important fact to this list, a theorem of Grothendieck.

Theorem 32.1.1 (Automatics deformation of abelian group structure)**.** *Suppose S is a connected, locally noetherian scheme, $\pi : A \to S$ is a projective and smooth morphism, and $e : S \to A$ is a section of π. Suppose further that there is a geometric point $s : \operatorname{Spec} k \to S$ of S such that $\pi_s : A_s \to \operatorname{Spec} k$ is an abelian variety with zero section e_s (the base change of e). Then $\pi : A \to S$ is an abelian scheme with zero section e.*

Remark. In the following, we simply say that (π, e) is an abelian scheme when we mean that $\pi : A \to S$ is an abelian scheme with zero section e. Recall that the group structure on an abelian scheme is uniquely determined by the zero section. Hence whether (π_s, e_s) or (π, e) is an abelian scheme is a question of the *existence* of a group scheme structure.

The proof of this theorem proceeds in three steps. We leave the complete proof of the theorem as a presentation topic; see [MFK94, §6.3]. In the course we shall only discuss the first step, which we should consider as an infinitesimal version of the theorem.

Suppose R is a local Artin ring with maximal ideal $\mathfrak{m}$ and residue field k. Consider a quotient ring $R_0 = R/I$ where I satisfies $\mathfrak{m}I = 0$ (which in particular implies that $I^2 = 0$). Such a quotient map $R \to R_0$ is often called a small extension of local Artin rings, and in general any surjection of local Artin rings can be factored into finitely many small extensions. We write $S = \operatorname{Spec} R$ and $S_0 = \operatorname{Spec} R_0$.

Proposition 32.1.2. *Let $\pi : A \to S$ be a proper and smooth morphism and $e : S \to A$ a section such that $(\pi_0, e_0) := (\pi, e) \times_S S_0$ is an abelian scheme. Then (π, e) is an abelian scheme.*

Remark. The proof uses obstruction theory. Similar considerations also imply that any proper smooth morphism $A_0 \to S_0$ together with section e_0 automatically deforms to proper smooth $A \to S$ together with section e. Combining this fact with the proposition, we see that any abelian scheme over S_0 always deforms to an abelian scheme over S.

33. Lecture 33

33.1. Automatic deformation of abelian group structure, continued.

Proof of Proposition 32.1.2. Let $\overline{A} := A \times_S \operatorname{Spec} k$. We can encode the group structure of (A_0, e_0) into the difference map $\mu_0 : A_0 \times_{S_0} A_0 \to A_0, (x, y) \mapsto x - y$. The goal is to deform μ_0 to a map $\mu : A \times_S A \to A$ satisfying the group axioms (that is, we define the group law $+ : A \times_S A \to A$ using the zero section e and the difference map μ by $x + y := x - (0 - y)$, and require that $+$ satisfies the group axioms). The possibility of lifting μ_0 to μ (without worrying about the group axioms) is controlled by the vanishing of a certain

$$\beta \in \mathbf{H}^1(\overline{A} \times_k \overline{A}, \overline{\mu}^*(\mathcal{T} \otimes_k I)),$$

where $\mathcal{T}$ is the tangent bundle of $\overline{A}$, I is as in $R_0 = R/I$ (thus a finite -dimensional k-vector space), and $\overline{\mu} : \overline{A} \times_k \overline{A} \to \overline{A}$ is induced by μ_0. We would like to show that β_0.

Define $g_1 : A_0 \to A_0 \times_{S_0} A_0$ via $x \mapsto (x, 0)$ and $g_2 : A_0 \to A_0 \times_{S_0} A_0$ via $x \mapsto (x, x)$. We have $\mu_0 \circ g_1 = id : A_0 \to A_0$ and $\mu_0 \circ g_2 = e_0 \circ \pi : A_0 \to A_0$.

Briefly, suppose that $g : B \times_S S_0 \to A_0 \times_{S_0} A_0$ is an arbitrary morphism where B is any proper smooth S-scheme. Then one acquires a morphism $\mu_0 \circ g : B \times_S S_0 \to A_0$. The obstruction to deform $\mu_0 \circ g$ to a morphism $B \to A$ is controlled by

$$\overline{g}^* \beta \in \mathbf{H}^1(\overline{B}, \overline{g}^* \overline{\mu}^* (\mathcal{T} \otimes_k I)),$$

where $\overline{g} : B \times_S \operatorname{Spec} k \to \overline{A}$ is induced by g. As such, in our case, the obstruction to deform $\mu_0 \circ g_i$ to a morphism $A \to A$ is

$$\overline{g_i}^* \beta \in \mathbf{H}^1(\overline{A}, \overline{g_i}^* \overline{\mu}^* (\mathcal{T} \otimes_k I)).$$

Notice that $\mu_0 \circ g_1$ deforms to id_A and $\mu_0 \circ g_2$ deforms to $e \circ \pi$. So $\overline{g_i}^* \beta = 0$ for $i = 1, 2$. We will use this information to show that $\beta = 0$.

We have $\mathcal{T} \cong \mathcal{O}_{\overline{A}} \otimes_k T$ where T is the tangent space at zero $T_{\overline{e}} \overline{A}$ (which is a finite dimensional k-vector space); this description of the tangent bundle is a general fact that holds for any group variety over a field.

To simplify notation, for any variety Y over k we write $\mathbf{H}^i(Y)$ for $\mathbf{H}^i(Y, \mathcal{O}_Y)$. Defining the k-vector space $W := T \otimes_k I$, we have $\beta \in \mathbf{H}^1(\overline{A} \times_k \overline{A}) \otimes_k W$ and $\overline{g_i}^* \beta \in \mathbf{H}^1(\overline{A}) \otimes_k W$. Pick an arbitrary k-linear map $q : W \to k$, and let

$$\beta_q \in \mathbf{H}^1(\overline{A} \times_k \overline{A}) \otimes_k k = \mathbf{H}^1(\overline{A} \times_k \overline{A})$$

be the functorial image of β under q. It suffices to prove that $\beta_q = 0$ since q is arbitrary. Now from $\overline{g_i}^* \beta = 0$ we have

$$\overline{g_i}^* \beta_q = 0 \in \mathbf{H}^1(\overline{A}).$$

By the Künneth formula, we have

$$\mathbf{H}^1(\overline{A} \times_k \overline{A}) = p_1^* \mathbf{H}^1(\overline{A}) \oplus p_2^* \mathbf{H}^1(\overline{A}).$$

Write

$$\beta_q = p_1^* \beta_1 + p_2^* \beta_2$$

where $\beta_i \in \mathbf{H}^1(\overline{A})$. We have

$$0 = \overline{g_1}^* \beta_q = \overline{g_1}^* p_1^* \beta_1 + \overline{g_1}^* p_2^* \beta_2 = \beta_1 + 0$$

and

$$0 = \overline{g_2}^* \beta = \beta_1 + \beta_2.$$

Hence $\beta_1 = \beta_2 = 0$, and so $\beta_q = 0$ as desired.

We have shown that μ_0 lifts to $\mu : A \times_k A \to A$. The choice of lift is not unique, and need not *a priori* provide a group structure on A.

In fact, the possible choices of lift μ form a torsor under

$$\mathbf{H}^0(\overline{A} \times_k \overline{A}, \overline{\mu}^* (\mathcal{T} \otimes_k I)) = \mathbf{H}^0(\overline{A} \times_k \overline{A}) \otimes_k W = W.$$

Likewise, deformations of the morphism

$$e_0 = \mu_0 \circ (e_0, e_0) : S_0 \xrightarrow{(e_0, e_0)} A_0 \times_{S_0} A_0 \xrightarrow{\mu_0} A_0$$

to $S \to A$ either do not exist, or form a torsor under

$$\mathbf{H}^0(\operatorname{Spec} k, (\overline{e}, \overline{e})^* \overline{\mu}^* (\mathcal{T} \otimes_k I)) = W.$$

(Of course we know that we are in the latter case, since e is a deformation of e_0.) Therefore, there is a unique choice of μ such that

$$\mu \circ (e, e) : S \xrightarrow{(e,e)} A \times_k A \xrightarrow{\mu} A$$

is equal to e.

Now it suffices to check that this choice of μ satisfies the group axioms. In [MFK94, Prop. 6.15] Mumford suggests that one employs the Rigidity Lemma for this verification; this approach might be problematic[9]. As an alternative solution, we notice that the previous argument showing the uniqueness of μ deforming μ_0 and taking (e, e) to e can be generalized as follows. Consider

$$f_0 : \underbrace{A_0 \times_{S_0} \cdots \times_{S_0} A_0}_{m \text{ times}} \to A_0$$

such that $f_0 \circ (e_0, \dots, e_0) = e_0$. Then there exists at most one deformation $f : A \times_S \cdots \times_S A \to A$ of f_0 such that $f \circ (e, \dots, e) = e$. The proof is the same as before: The set of f deforming f_0 is either empty or a torsor under W.

Using this uniqueness, one can check all the group axioms. For instance, suppose we want to check $0 - (0 - x) = x$. Define $f : A \to A, x \mapsto 0 - (0 - x)$ and $f' : A \to A, x \mapsto x$. Then f and f' are deformations of the same map $A_0 \to A_0$ (because A_0 is a group) and they both take e to e. Here, the fact that f takes e to e follows from the fact that μ takes (e, e) to e. Thus by the above paragraph we have $f = f'$. □

34. Lecture 34

34.1. **The moduli problem.** We fix integers $g \geq 1, d \geq 1, N \geq 3$.

Definition 34.1.1. Let $\mathcal{A}_{g,d,N}$ be the functor from the category of noetherian schemes over $\mathbb{Z}[1/N]$ to the category of sets, sending S to the set of isomorphism classes of triples (A, λ, γ), where

- $\pi : A \to S$ is a projective abelian scheme of relative dimension g.
- $\lambda : A \to A^\vee$ is a polarization on A whose degree (as an isogeny) is d^2.
- γ is a level-N structure on A, namely an isomorphism of S-group schemes $(\mathbb{Z}/N\mathbb{Z})_S^{2g} \xrightarrow{\sim} A[N]$.

Here on morphisms the functor $\mathcal{A}_{g,d,N}$ is defined using the obvious notion of pullback.

Theorem 34.1.2. *The functor $\mathcal{A}_{g,d,N}$ is representable.*

Remark.
(1) For $\pi : A \to S$ as above, $A[N]$ is a finite étale group scheme over S, and étale locally on S it is isomorphic to $(\mathbb{Z}/N\mathbb{Z})_S^{2g}$. From this, similar to the modular curve case, it is easy to show that the functor from S-schemes to sets sending T to the set of level-N structures on $A \times_S T$ is representable by an S-scheme, and that this S-scheme is a finite étale $\mathrm{GL}_{2g}(\mathbb{Z}/N\mathbb{Z})$-torsor on S.
(2) For $N \geq 3$, if A is a projective scheme over S and γ is a level-N structure on A, then there is no non-trivial S-automorphism of A preserving γ.
(3) In view of (1) and (2), by the same argument as in the modular curve case we know that to prove the theorem for general $N \geq 3$ we only need to prove it for two choices of N that are coprime to each other. In particular, it is enough to prove the theorem for all sufficiently large N.
(4) The so-called **Siegel modular variety**, which is the most fundamental example of a Shimura variety, is the subfunctor of $\mathcal{A}_{g,d,N}$ where a certain compatibility condition between λ and γ is imposed. More precisely, for (A, λ, γ) over S, there is a canonical

[9] It seems that the best one could prove using the Rigidity Lemma [MFK94, Prop. 6.1] (note that [MFK94, Cor. 6.2] is not directly applicable since the target is not known to be a group) is that certain maps $A \times_S \cdots \times_S A \to A$ factor through $e : S \to A$. Thus one would be able to prove statements about μ and e such as $x - x = 0$ or $(x - (x - y)) - y = 0$, but not the statements about the equality of two morphisms, such as the axiom $0 + x = x$, meaning $0 - (0 - x) = x$.

S-scheme morphism $A[N] \times_S A^\vee[N] \to \mu_{N,S}$ which is bilinear with respect to the group structures, and from this we obtain a bilinear map

$$(\mathbb{Z}/N\mathbb{Z})_S^{2g} \times_S (\mathbb{Z}/N\mathbb{Z})_S^{2g} \xrightarrow{(\gamma,\gamma)} A[N] \times_S A[N] \xrightarrow{(\mathrm{id},\lambda)} A[N] \times_S A^\vee[N] \to \mu_{N,S}.$$

The Siegel modular variety is defined by asking that the last map should belong to a certain prescribed list. Once we know the representability of $\mathcal{A}_{g,d,N}$, we immediately deduce that the Siegel modular variety is represented by a union of some connected components of $\mathcal{A}_{g,d,N}$.

We will prove the theorem in two steps. Firstly, we consider a "framed" version of the moduli problem, namely a functor $\mathcal{H}_{g,d,N}$ which classifies (A, λ, γ) together with an extra structure of a projective embedding into some fixed $\mathbb{P}^m$. (The precise meaning of this will be made clear soon.) We will show that $\mathcal{H}_{g,d,N}$ is representable by relating it to Hilbert schemes. Secondly, we show that the natural PGL_{m+1}-action on $\mathcal{H}_{g,d,N}$ (by acting on the projective space $\mathbb{P}^m$) admits a "nice" quotient, and this quotient is the desired scheme representing $\mathcal{A}_{g,d,N}$.

34.2. **The framed moduli problem.** We make precise the definition of $\mathcal{H}_{g,d,N}$. First we make some observations. Let S be a noetherian $\mathbb{Z}[1/N]$-scheme, $\pi : A \to S$ a projective abelian scheme of relative dimension g, and $\lambda : A \to A^\vee$ a polarization of degree d^2. Then $L^\Delta(\lambda)$ is relatively ample with respect to π by Proposition 31.1.1, and then by Theorem 29.2.1 (4) $L^\Delta(\lambda)^3 = L^\Delta(3\lambda)$ is relatively very ample. Thus we have canonical closed immersion $A \to \mathbb{P}(\pi_* L^\Delta(3\lambda))$. Moreover, by Theorem 29.2.1 (1) (2), we have $R^i\pi_* L^\Delta(3\lambda) = 0$ for $i \geq 1$, and $\pi_* L^\Delta(3\lambda)$ is a vector bundle on S of rank $r = \sqrt{\deg \Lambda(L^\Delta(3\lambda))}$. Since 3λ is a symmetric isogeny, we have $\Lambda(L^\Delta(3\lambda)) = 6\lambda$, and hence $r = 6^g d$. (Recall that $\deg[n]_A = n^{2g}$.) Set

$$m = 6^g d - 1$$

once and for all. Thus if we localize S and trivialize the vector bundle $\pi_* L^\Delta(3\lambda)$, then $\mathbb{P}(\pi_* L^\Delta(3\lambda))$ is isomorphic to $\mathbb{P}_S^m$. This motivates the following definition.

Definition 34.2.1. By a **linear rigidification** of (A, λ), we mean the choice of an S-isomorphism

$$\phi : \mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\sim} \mathbb{P}_S^m.$$

We explain that this notion behaves well with base change. Suppose we have $T \to S$, then a linear rigidification ϕ of (A, λ) will determine a linear rigidification of (A_T, λ_T) (the base change of (A, λ) from S to T) as follows. Write L for $L^\Delta(3\lambda)$. First note that the line bundle $L^\Delta(3\lambda_T)$ on A_T is canonically identified with $L|_{A_T}$, the base change of L on A to A_T. Second, since $R^i\pi_* L = 0$ for all $i \geq 1$, we know that the formation of $\pi_* L$ commutes with base change, i.e., that $\pi_{T,*}(L|_{A_T})$ is canonically isomorphic to $\pi_* L \otimes_{\mathcal{O}_S} \mathcal{O}_T$. Combining these two facts, we see that $\mathbb{P}(\pi_{T,*} L^\Delta(3\lambda_T))$ is canonically identified with $\mathbb{P}(\pi_* L^\Delta(3\lambda)) \times_S T$. Hence a linear rigidification of (A, λ) canonically induces one for (A_T, λ_T).

By the above discussion, it makes sense to define the following functor.

Definition 34.2.2. Let $\mathcal{H}_{g,d,N}$ be the functor from the category of noetherian schemes over $\mathbb{Z}[1/N]$ to the category of sets, sending S to the set of isomorphism classes of quadruples $(A, \lambda, \gamma, \phi)$, where (A, λ, γ) is a triple over S as in the definition of $\mathcal{A}_{g,d,N}$, and ϕ is a linear rigidification of (A, λ).

The strategy to show that $\mathcal{H}_{g,d,N}$ is representable is to embed it into a larger functor which is closely related to Hilbert schemes.

Definition 34.2.3. Let $\tilde{\mathcal{H}}$ be the functor from the category of noetherian schemes over $\mathbb{Z}[1/N]$ to the category of sets, sending S to the set of closed subschemes $Z \subset \mathbb{P}^m_S$ that are flat over S together with $2g+1$ sections $\sigma_0, \cdots, \sigma_{2g} : S \to Z$.

We construct a natural transformation $\mathcal{H}_{g,d,N} \to \tilde{\mathcal{H}}$ as follows. Let S be a noetherian $\mathbb{Z}[1/N]$-scheme, and let $(A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(S)$. We define an element $(Z, \sigma_0, \cdots, \sigma_{2g}) \in \tilde{\mathcal{H}}(S)$ as follows. From (A, λ, ϕ), we obtain a closed immersion $A \hookrightarrow \mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\phi} \mathbb{P}^m_S$; this is our definition of Z. Define σ_0 to be the neutral section e of A. The S-scheme $(\mathbb{Z}/N\mathbb{Z})^{2g}_S$ has $2g$ canonical sections, corresponding to the basis vectors

$$(0, \ldots, 0, 1, 0, \cdots, 0) \in (\mathbb{Z}/N\mathbb{Z})^{2g}.$$

Using γ, we view these $2g$ sections as sections $\sigma_1, \cdots, \sigma_{2g}$ of A. This completes the definition of the map $\mathcal{H}_{g,d,N}(S) \to \tilde{\mathcal{H}}(S)$.

In the next lecture, we will show that $\mathcal{H}_{g,d,N} \to \tilde{\mathcal{H}}$ is a subfunctor, and is locally closed, i.e., roughly, for any $\xi \in \tilde{\mathcal{H}}(S)$, we have a locally closed locus in S over which ξ comes from $\mathcal{H}_{g,d,N}$.

The point of this is that by standard theory of Hilbert schemes, $\tilde{\mathcal{H}}$ is representable. It then follows that the locally closed subfunctor $\mathcal{H}_{g,d,N}$ is also representable (by a locally closed subscheme of the scheme representing $\tilde{\mathcal{H}}$).

35. Lecture 35

35.1. **The framed moduli problem, continued.** Recall that we are interested in proving the representability of the functor $\mathcal{A}_{g,d,N}$. For this, we are first studying an auxiliary functor, the framed moduli functor $\mathcal{H}_{g,d,N}$ given by sending a noetherian scheme S over $\mathbb{Z}[1/N]$ to the set of isomorphism classes of quadruples $(A, \lambda, \gamma, \phi)$ where:

- $\pi : A \to S$ is a projective abelian scheme of relative dimension g
- $\lambda : A \to A^\vee$ is a polarization of degree d^2
- $\gamma : (\mathbb{Z}/N\mathbb{Z})^{2g}_S \xrightarrow{\sim} A[N]$ is an isomorphism of S-group schemes
- $\phi : \mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\sim} \mathbb{P}^m_S$ is an S-isomorphism, called linear rigidification, where $m = 6^g d - 1$.

From the data of an element of $\mathcal{H}_{g,d,N}(S)$, we canonically acquire a closed embedding $i_{\text{can}} : A \hookrightarrow \mathbb{P}^m_S$ defined as the composition

$$A \hookrightarrow \mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\sim} \mathbb{P}^m_S$$

where the first arrow is a closed embedding given by the relative very ampleness of $L^\Delta(3\lambda)$.

We define another auxiliary functor $\tilde{\mathcal{H}}$ sending a noetherian scheme over $\mathbb{Z}[1/N]$ to the set of closed subschemes $Z \subset \mathbb{P}^m_S$ flat over S together with $2g+1$ sections $\sigma_0, \ldots, \sigma_{2g} : S \to Z$. We obtain a natural transformation $\mathcal{H}_{g,d,N} \to \tilde{\mathcal{H}}$ given by sending $(A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(S)$ to an element of $\tilde{\mathcal{H}}(S)$ in the following way:

- $i_{\text{can}} : A \hookrightarrow \mathbb{P}^m_S$ is a closed subscheme where A is clearly flat over S, as $\pi : A \to S$ is smooth since it is an abelian scheme
- $\sigma_0 = e : S \to A$ is the identity section
- For each $1 \le i \le 2g$, define $f_i := (0, \ldots, 1, \ldots, 0) \in (\mathbb{Z}/N\mathbb{Z})^{2g}$ with only identity in the ith entry and 0 elsewhere. Then we obtain $2g$ sections

$$\sigma_i : S \xrightarrow{f_i} (\mathbb{Z}/N\mathbb{Z})^{2g}_S \xrightarrow{\gamma} A[N] \subset A.$$

Proposition 35.1.1 ([MFK94, Chap. 7, §2])**.** *The map $\mathcal{H}_{g,d,N}(S) \to \tilde{\mathcal{H}}(S)$ is injective.*

Proof. From $A \subset \mathbb{P}^m_S$ and $\sigma_0 = e$, we recover the abstract abelian scheme structure on A, that is, the group structure. From $\sigma_1, \ldots, \sigma_{2g}$ we obtain the level structure. To obtain λ, we need the fact that the pullback of $\mathcal{O}_{\mathbb{P}^m_S}(1)$ under ϕ differs from $\mathcal{O}(1)$ on $\mathbb{P}(\pi_* L^\Delta(3\lambda))$ by something in $\mathrm{Pic}(S)$ (by which we mean a line bundle that is the pullback of a line bundle on S). Similarly, pullback of $\mathcal{O}(1)$ on $\mathbb{P}(\pi_* L^\Delta(3\lambda))$ to A differs from $L^\Delta(3\lambda)$ by something in $\mathrm{Pic}(S)$. In particular, we must have $i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1) \equiv L^\Delta(3\lambda) \mod \pi^* \mathrm{Pic}(S)$. Now recall that $L^\Delta(3\lambda)$ is canonically trivialized along e. As such, we have

$$L^\Delta(3\lambda) \cong i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1) \otimes \pi^* e^* (i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1))^{-1}.$$

Hence we have recovered $L^\Delta(3\lambda)$. Applying the Λ-construction, we obtain 6λ, and in particular we also obtain λ. It remains to recover ϕ. From $i_{\mathrm{can}} : A \to \mathbb{P}^m_S$ and λ, we must acquire an isomorphism $\mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\sim} \mathbb{P}^m_S$. Define $M := e^* i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(-1)$. This is a line bundle on S. Let $f : \mathbb{P}^m_S \to S$ be the structure map. By adjunction, we acquire a canonical morphism:

$$\Phi : f_* \mathcal{O}_{\mathbb{P}^m_S}(1) \otimes_{\mathcal{O}_S} M \longrightarrow \pi_*(i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1) \otimes \pi^* M).$$

We claim that this is an isomorphism. Notice first that both sides are vector bundles by cohomology and base change, since for each $s \in S$ we have $\mathbf{H}^1(\mathbb{P}^m_s, \mathcal{O}(1)) = 0$ and $\mathbf{H}^1(A_s, i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1)|_{A_s}) = 0$ (by the relative ampleness of $i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1)$). So it suffices to check that Φ induces isomorphisms on all fibers. Again by the fiberwise vanishing of $\mathbf{H}^1$, formation of pushforwards on each side commutes with base change. Hence we may assume that $S = \mathrm{Spec}\, k$ where k is an algebraically closed field. In this case, the claim is clear.

By taking the projectivization of both sides of Φ, we then acquire an isomorphism

$$\psi : \mathbb{P}(f_* \mathcal{O}_{\mathbb{P}^m_S}(1) \otimes_{\mathcal{O}_S} M) \to \mathbb{P}(\pi_*(i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1) \otimes \pi^* M)).$$

The left hand side is isomorphic to $\mathbb{P}^m_S$. Also, by the previous step, we know that

$$i^*_{\mathrm{can}} \mathcal{O}_{\mathbb{P}^m_S}(1) \otimes \pi^* M \cong L^\Delta(3\lambda).$$

Hence, ψ can be viewed as an isomorphism $\mathbb{P}^m_S \xrightarrow{\sim} \mathbb{P}(\pi_* L^\Delta(3\lambda))$. Once checks that ϕ must be the inverse of ψ, and as such we have recovered ϕ. ☐

Proposition 35.1.2 ([MFK94, Prop. 7.3])**.** *The natural transformation $\mathcal{H}_{g,d,N} \hookrightarrow \tilde{\mathcal{H}}$ is locally closed, i.e., for all S and all $\xi \in \tilde{\mathcal{H}}(S)$, there exists a unique locally closed subscheme $S_0 \subset S$ such that for any $T \to S$, the pullback $\xi|_T \in \tilde{\mathcal{H}}(T)$ lies in $\mathcal{H}_{g,d,N}(T)$ if and only if $T \to S$ factors through S_0.*

Proof. Write $\xi = (i : A \hookrightarrow \mathbb{P}^m_S, e, \sigma_1, \ldots, \sigma_{2g})$. The proof proceeds in a few steps.

(1) Note that there is a unique open subscheme $S' \subset S$ such that for all geometric points s of S, the base change A_s is smooth if and only if s belongs to S'. Hence we may replace S by S', and in particular we may assume that $A \to S$ is smooth.
(2) Since $A \to S$ is smooth and proper, we can apply Theorem 32.1.1, which states that for each connected component S^+ of S, if (A, e) becomes an abelian scheme when base changed to one geometric point of S^+, then (A, e) is an abelian scheme on S^+. So we can and must throw away all connected components of S not satisfying this condition. Then for the new S, the pair (A, e) is an abelian scheme over S.
(3) Asking that $\sigma_1, \ldots, \sigma_{2g}$ are sections of $A[N]$ (instead of just sections of A) is a closed condition on S. For them to define a level structure γ is an open condition. Thus we may assume that $\sigma_1, \ldots, \sigma_{2g}$ indeed come from a (unique) level structure γ.

(4) To conclude, it suffices to show that the condition that i comes from a polarization λ together with a linear rigidification ϕ is also locally closed on S. We will do this next time.

$\square$

36. Lecture 36

Proof of Proposition 35.1.2, continued. Write $\xi = (i : A \hookrightarrow \mathbb{P}^m_S, \sigma_0, \dots, \sigma_{2g})$ where $m = 6^g d - 1$. We showed last time that we may assume that (A, σ_0) is an abelian scheme and $(\sigma_1, \dots, \sigma_{2g})$ corresponds to a level-N structure on A. Now we need to show that there exists $S_0 \subset S$ such that over S_0, the inclusion $i : A \hookrightarrow \mathbb{P}^m_S$ comes from $A \xrightarrow{\text{can}} \mathbb{P}(\pi_* L^\Delta(3\lambda))$ for a polarization λ of degree d^2 and a linear rigidification $\phi : \mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\sim} \mathbb{P}^m_{S_0}$ and similarly after base change.

Now recall Proposition 30.2.4: fix (L, ρ) on A where L is a line bundle, $\rho : e^* L \xrightarrow{\sim} \mathcal{O}_S$ is an isomorphism, and such that $\Lambda(L) : A \to A^\vee$ is an isogeny (e.g., if L is relatively ample). Then there exists a locally closed $S_1 \subset S$ such that for any S-scheme T, the morphism $T \to S$ factors through S_1 if and only if $(L, \rho)_T \cong (L^\Delta(3\lambda), \rho_{\text{can}})$ for some (unique) symmetric isogeny $\lambda : A_T \to A_T^\vee$. Now in our current setting we define

$$L = i^* \mathcal{O}(1) \otimes \pi^* e^* i^* \mathcal{O}(1)^{-1}$$

and take $\rho : e^* L \xrightarrow{\sim} \mathcal{O}_S$ to be the canonical isomorphism. Here, L is relatively ample, so we can apply Proposition 30.2.4 to (L, ρ) and find a locally closed $S_1 \subset S$ over which (L, ρ) comes from a unique symmetric isogeny λ via $L^\Delta(3\lambda)$. Here the idea is that if i indeed comes from some polarization λ' and linear rigidification ϕ, then we must have $L \cong L^\Delta(3\lambda')$ by the same argument as the proof of Proposition 35.1.1. Thus the sought-after S_0 must be contained in S_1, and λ' must be λ (by the uniqueness in Proposition 30.2.4).

Since the third power of $L^\Delta(\lambda)$ is L, which is relatively ample, we know that $L^\Delta(\lambda)$ is relatively ample. Further, λ is symmetric. This implies that λ is a polarization by Proposition 31.1.1.

We may replace S by S_1. So we have reduced to the case where we have a unique polarization λ on A such that

$$i^* \mathcal{O}(1) \otimes \pi^* e^* i^* \mathcal{O}(1)^{-1} \cong L^\Delta(3\lambda).$$

Finally, we just need to find $S_0 \subset S$ over which $i : A \hookrightarrow \mathbb{P}^m_{S_0}$ factors as $(\phi \circ \text{can})$ where $\phi : \mathbb{P}(\pi_* L^\Delta(3\lambda)) \xrightarrow{\sim} \mathbb{P}^m_{S_0}$ is an isomorphism and $\text{can} : A \hookrightarrow \mathbb{P}(\pi_* L^\Delta(3\lambda))$ is the canonical embedding. To be precise, as in the proof of Proposition 35.1.1, in our current situation we already have a map of coherent $\mathcal{O}_S$-modules

$$\Phi : f_* \mathcal{O}(1) \otimes_{\mathcal{O}_S} e^* i^* \mathcal{O}(-1) \longrightarrow \pi_*(L^\Delta(3\lambda)),$$

where $f : \mathbb{P}^m_S \to S$ is the structure morphism. If i factors as $\phi \circ \text{can}$, then Φ must be an isomorphism and ϕ must be the inverse of $\mathbb{P}(\Phi)$, and vice versa. Hence the sought-after S_0 is just the open subscheme of S over which Φ is an isomorphism. $\square$

36.1. **Hilbert schemes.** Fix an integer $n \geq 1$. We define a functor $\text{Hilb}_{\mathbb{P}^n}$ from locally noetherian schemes to sets sending each S to the set of all closed subschemes $Z \subset \mathbb{P}^n_S$ such that Z is flat over S.

Theorem 36.1.1 (Grothendieck)**.** *The functor* $\text{Hilb}_{\mathbb{P}^n}$ *is representable by a locally noetherian scheme, still denoted by* $\text{Hilb}_{\mathbb{P}^n}$*.*

Write H for $\mathrm{Hilb}_{\mathbb{P}^n}$. We have the universal closed subscheme $Z_{\mathrm{univ}} \subset \mathbb{P}^n_H$ which is flat over H. Using this, we can easily represent certain variations of the functor $\mathrm{Hilb}_{\mathbb{P}^n}$. For instance, for $r \geq 1$, we define a functor $\mathrm{Hilb}_{\mathbb{P}^n,r}$ by sending a locally noetherian scheme S to the set of all closed subschemes $Z \subset \mathbb{P}^n_S$ with Z flat over S, along with the data of r (ordered) sections $S \to Z$. The functor $\mathrm{Hilb}_{\mathbb{P}^n,r}$ is representable by the r-fold product

$$Z_{\mathrm{univ}} \times_H \cdots \times_H Z_{\mathrm{univ}}.$$

In particular, we have that $\tilde{\mathcal{H}} = \mathrm{Hilb}_{\mathbb{P}^m,2g+1} \times_{\mathrm{Spec}\,\mathbb{Z}} \mathrm{Spec}\,\mathbb{Z}[1/N]$ is representable. One issue is that $\mathrm{Hilb}_{\mathbb{P}^n}$ and $\mathrm{Hilb}_{\mathbb{P}^n,r}$ are too large. We want to shrink them to more manageable subschemes.

To do this, we will need the notion of Hilbert polynomials. For a closed subscheme $Z \subset \mathbb{P}^n_k$ where k is algebraically closed, we obtain a polynomial $P_Z(T) \in \mathbb{Z}[T]$ such that for all $n \in \mathbb{Z}$, we have $P_Z(n) = \chi(\mathcal{O}_Z(n))$, where $\mathcal{O}_Z(n)$ is the nth power of $\mathcal{O}(1)|_Z$.

A basic fact is that for S locally noetherian *connected* and $Z \subset \mathbb{P}^n_S$ a closed subscheme flat over S, for geometric points s of S the polynomial $P_{Z_s}(T)$ is independent of s. Thus we obtain a natural decomposition

$$\mathrm{Hilb}_{\mathbb{P}^n} = \coprod_{P(T) \in \mathbb{Z}[T]} \mathrm{Hilb}^{P(T)}_{\mathbb{P}^n}$$

where each $\mathrm{Hilb}^{P(T)}_{\mathbb{P}^n}$ is open and closed (i.e., a union of connected components) in $\mathrm{Hilb}_{\mathbb{P}^n}$, determined by the condition that over its geometric points the Hilbert polynomial of Z_{univ} is $P(T)$.

Theorem 36.1.2 (Grothendieck)**.** *Each* $\mathrm{Hilb}^{P(T)}_{\mathbb{P}^n}$ *is a quasi-projective scheme over* $\mathbb{Z}$.

Similarly, we define $\mathrm{Hilb}^{P(T)}_{\mathbb{P}^n,r} := \mathrm{Hilb}_{\mathbb{P}^n,r} \times_{\mathrm{Hilb}_{\mathbb{P}^n}} \mathrm{Hilb}^{P(T)}_{\mathbb{P}^n}$; the moduli problem it represents is clear. This is again a quasi-projective scheme over $\mathbb{Z}$, since $\mathrm{Hilb}_{\mathbb{P}^n,r}$ is obviously projective over $\mathrm{Hilb}_{\mathbb{P}^n}$.

Lemma 36.1.3. *The subscheme* $\mathcal{H}_{g,d,N} \subset \tilde{\mathcal{H}} = \mathrm{Hilb}_{\mathbb{P}^m,2g+1}[1/N]$ *is contained in*

$$\mathrm{Hilb}^{P(T)}_{\mathbb{P}^m,2g+1}$$

for an explicit $P(T)$*. In particular,* $\mathcal{H}_{g,d,N}$ *is quasi-projective over* $\mathbb{Z}[1/N]$.

Proof. We will compute the Hilbert polynomial $P(T)$. Take an arbitrary some $(A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(k)$ where k is algebraically closed. As usual we obtain

$$i = i_{\mathrm{can}} : A \hookrightarrow \mathbb{P}(\pi_* L^{\Delta}(3\lambda)) \xrightarrow{\phi} \mathbb{P}^m_k.$$

We need to compute $P_{i(A)}(T)$.

Take any integer $n \geq 1$. We have

$$P_{i(A)}(n) = \chi(i^*\mathcal{O}(n)) = \chi(L^{\Delta}(3\lambda)^{\otimes n}) = \chi(L^{\Delta}(3n\lambda)).$$

Note that $L^{\Delta}(3n\lambda)$ is ample. Hence by Corollary 25.1.2 and Theorem 26.1.1 we have

$$\chi(L^{\Delta}(3n\lambda)) = h^0(L^{\Delta}(3n\lambda)) = \sqrt{\deg \Lambda(L^{\Delta}(3n\lambda))}$$
$$= \sqrt{\deg 6n\lambda} = \sqrt{\deg[6n] \cdot \deg \lambda} = (6n)^g d = 6^g d \cdot n^g.$$

Since $n \geq 1$ is arbitrary, $P_{i(A)}(T)$ must be $6^g d T^g$. We have thus shown that $\mathcal{H}_{g,d,N}$ is contained in $\mathrm{Hilb}^{6^g d T^g}_{\mathbb{P}^m,2g+1}$. ∎

36.2. **The projective linear group.** The idea is to roughly construct $\mathcal{A}_{g,d,N}$ as a quotient $\mathcal{H}_{g,d,N}/\operatorname{PGL}_{m+1}$, where PGL_{m+1} acts by modifying the linear rigidification

$$\phi : \mathbb{P}(\pi_*(L^{\Delta}(3\lambda))) \xrightarrow{\sim} \mathbb{P}^m_S$$

via its natural action on $\mathbb{P}^m_S$.

Recall that PGL_n is roughly speaking the group GL_n modulo the scalars. As a group scheme, we define PGL_n as $\operatorname{Proj}\mathbb{Z}[a_{11}, a_{12}, \ldots, a_{nn}]$ minus the locus where $\det(a_{ij}) = 0$. Then PGL_n is an open subscheme of $\mathbb{P}^{n^2-1}_{\mathbb{Z}}$. The group structure is simply given by matrix multiplication. We know that PGL_n is a smooth group scheme over $\mathbb{Z}$. For each scheme S, the group $\operatorname{PGL}_n(S)$ acts on $\mathbb{P}^n_S$ via linear transformations, which are S-scheme automorphisms of $\mathbb{P}^n_S$.

We have the following important non-trivial fact:

Fact 36.2.1 (See [MFK94, Chap. 0, §5, (b)])**.** *For each noetherian scheme S, the homomorphism $\operatorname{PGL}_n(S) \to \operatorname{Aut}_{S\text{-}sch}\mathbb{P}^{n-1}_S$ is an isomorphism.*

37. Lecture 37

37.1. **The "most ideal" quotient.** For $g, d \geq 1$ and $N \geq 3$, we have shown that $\mathcal{H}_{g,d,N}$ is representable by a locally closed subscheme of the quasi-projective $\mathbb{Z}$-scheme $\operatorname{Hilb}^{P(T)}_{\mathbb{P}^m,2g+1}$, where $m = 6^g d - 1$ and $P(T) = 6^g dT^g$.

Recall that for a group scheme G over $\mathbb{Z}$ and a scheme X over $\mathbb{Z}$, a G-action on X means a functorial collection of $G(S)$-actions on $X(S)$ for all test schemes S. In particular, the action determines and is determined by a morphism $G \times X \to X$, which we call the **action morphism**.

We have a natural PGL_{m+1}-action on $\mathcal{H}_{g,d,N}$ as follows: For $g \in \operatorname{PGL}_{m+1}(S)$ and $(A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(S)$, we define $g(A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(S)$ to be $(A, \lambda, \gamma, g \circ \phi)$, where in writing $g \circ \phi$ we think of g as an S-scheme automorphism of $\mathbb{P}^m_S$. Our idea of showing the representability of $\mathcal{A}_{g,d,N}$ is to construct it as a suitable quotient of $\mathcal{H}_{g,d,N}$ by PGL_{m+1}. Of course we need to make precise the notion of a quotient. In general, there are many possible versions of the notion of a quotient, but we will only need the version which is in some sense the most ideal. That is, we ask that the quotient map should be a PGL_{m+1}-torsor in the following sense. We keep the convention that all schemes are locally noetherian.

Definition 37.1.1. Let $G = \operatorname{PGL}_n$, or more generally any finite-type flat group scheme over $\mathbb{Z}$. Let X be a scheme equipped with a G-action. We say that a scheme map $\Phi : X \to Y$ is a **G-torsor**, if it satisfies the following conditions:

1. It is of finite type, and flat.
2. It is G-invariant, where G acts trivially on Y. In other words, for any scheme S, the map $\Phi : X(S) \to Y(S)$ is $G(S)$-invariant, where $G(S)$ acts trivially on $Y(S)$.
3. By (2), we have a map $G \times_{\operatorname{Spec}\mathbb{Z}} X \to X \times_Y X$ which on S-points is given by $(g, x) \mapsto (g \cdot x, x)$. (Here the fiber product $X \times_Y X$ is formed with respect to $\Phi : X \to Y$.) We ask that this map is an isomorphism.

Remark. If $\Phi : X \to Y$ is a G-torsor, then one should think of Y as an "ideal quotient" of X by G. When the scheme X with G-action is fixed, there can exist at most one pair (Φ, Y) up to unique isomorphism. That is, if (Φ, Y) and (Φ', Y') are two pairs, then there exists a unique isomorphism $u : Y \xrightarrow{\sim} Y'$ such that $u \circ \Phi = \Phi'$. On the other hand, the existence of a pair (Φ, Y) is a very subtle issue; it depends on how "good" the G-action on X is.

Theorem 37.1.2. *Fix $g, d \geq 1$. For all N sufficiently large with respect to (g, d), the scheme $\mathcal{H}_{g,d,N}$ with PGL_{m+1}-action (where $m = 6^g d - 1$ as always) admits a PGL_{m+1}-torsor $\mathcal{H}_{g,d,N} \to Y$. Moreover, Y is quasi-projective over $\mathbb{Z}$.*

Proposition 37.1.3. *Let Y be as in Theorem 37.1.2. There is an isomorphism between (the functor of points of) the scheme Y and the functor $\mathcal{A}_{g,d,N}$, under which the map $\mathcal{H}_{g,d,N} \to Y$ corresponds to the natural forgetful map $\mathcal{H}_{g,d,N} \to \mathcal{A}_{g,d,N}$. In particular, $\mathcal{A}_{g,d,N}$ is representable by a quasi-projective $\mathbb{Z}$-scheme as long as N is large enough with respect to (g, d).*

Proof. The proof is formal in the sense that it doesn't require the knowledge of how Y is constructed, but just the fact that we have a PGL_{m+1}-torsor $\mathcal{H}_{g,d,N} \to Y$. See [MFK94, Prop. 7.6]. The basic idea is that supposing $(A, \lambda, \gamma, \phi)$ is the universal object over $\mathcal{H}_{g,d,N}$, we want to show that the polarized abelian scheme with level-N structure (A, λ, γ) descends to Y, and that the descended object over Y is the universal object for the moduli functor $\mathcal{A}_{g,d,N}$. □

We have already observed that for fixed (g, d), the representability of $\mathcal{A}_{g,d,N}$ for all $N \geq 3$ follows from the representability of $\mathcal{A}_{g,d,N}$ for all sufficiently large N, or even just two coprime choices of N. Thus we have completed the proof of Theorem 34.1.2 modulo Theorem 37.1.2.

37.2. **A blackbox from Geometric Invariant Theory.** In order to prove Theorem 37.1.2, we introduce a huge blackbox which gives a sufficient condition for a scheme X with a PGL_{n+1}-action to admit a PGL_{n+1}-torsor $X \to Y$. In the following we fix n, and let G denote PGL_{n+1}.

Definition 37.2.1. Let k be an algebraically closed field, and l a positive integer. We say a collection C of l points in $\mathbb{P}^n_k(k)$ is **stable**, if for every proper linear subspace $H \subsetneq \mathbb{P}^n_k$, we have

$$\frac{|C \cap H|}{l} < \frac{\dim H + 1}{n + 1}.$$

For example, in $\mathbb{P}^2_k(k)$, a collection of 4 points is stable if and only if no 3 of the points are colinear. In general, by considering 0-dimensional H, we see that a necessary condition for any collection of l points in $\mathbb{P}^n_k(k)$ to be stable is that $l > n + 1$.

Fact 37.2.2. *Let $l \geq 1$. There is a unique maximal open subscheme $(\mathbb{P}^n)^l_{\text{stable}}$ of $(\mathbb{P}^n)^l$ (where $\mathbb{P}^n$ is over $\mathbb{Z}$) such that for every algebraically closed field k and every*

$$\xi = (\xi_1, \cdots, \xi_l) \in (\mathbb{P}^n)^l_{\text{stable}}(k),$$

the l components ξ_i of ξ form a stable collection of l points in $\mathbb{P}^n_k(k)$.

Note that the natural diagonal action of G on $(\mathbb{P}^n)^l$ stabilizes the open subscheme $(\mathbb{P}^n)^l_{\text{stable}}$. Here is the blackbox.

Theorem 37.2.3. *Suppose X is a finite-type $\mathbb{Z}$-scheme with a G-action. Assume that there exists a G-equivariant morphism $f : X \to (\mathbb{P}^n)^l_{\text{stable}}$ for some $l \geq 1$. Assume the following technical condition:*

- (*) *There exists a G-equivariant line bundle L on X such that L is relatively ample with respect to f.*

Then there exists a G-torsor $X \to Y$ and Y is quasi-projective over $\mathbb{Z}$.

Proof. Combine [MFK94, Def. 3.7/Prop. 3.4, Thm. 3.8, Prop. 7.1]. □

Here, a G-equivariant line bundle on X means a line bundle L on X together with a G-equivariance structure, meaning an isomorphism $\sigma^* L \xrightarrow{\sim} p_2^* L$ satisfying certain "cocycle conditions", where σ is the action morphism $G \times X \to X$ and p_2 is the second projection $G \times X \to X$. See [MFK94, Chap. 1, §3] for more details, where a G-equivariance structure is called a G-linearization.

For our application to $X = \mathcal{H}_{g,d,N}$ with the PGL_{m+1}-action, the technical condition (*) is always automatically satisfied. The real challenge is to construct a G-equivariant $f : \mathcal{H}_{g,d,N} \to (\mathbb{P}^m)^l_{\mathrm{stable}}$ for some l.

37.3. **Mapping into the stable locus.** We will prove the following lemma in the next lecture, using some intersection theory.

Lemma 37.3.1. *Let k be an algebraically closed field. Let A be an abelian variety over k embedded in some $\mathbb{P}^n_k$ such that it is not contained in any hyperplane. Let g be the dimension of A, and let r be the degree of A as a closed subvariety of $\mathbb{P}^n_k$. Let N be any integer larger than $\sqrt{(n+1)r}$ and coprime to $\mathrm{char}(k)$. Then the set $A[N](k)$, viewed as a collection of N^{2g} points in $\mathbb{P}^n_k(k)$, is stable.*

Proof of Theorem 37.1.2. We prove the theorem for $N > 6^g d\sqrt{g!}$. Let $G = \mathrm{PGL}_{m+1}$. We construct a G-equivariant map $f : \mathcal{H}_{g,d,N} \to (\mathbb{P}^m)^{N^{2g}}$ as follows. Fix an enumeration $y_1, \cdots, y_{N^{2g}}$ of $(\mathbb{Z}/N\mathbb{Z})^{2g}$. Let S be a noetherian $\mathbb{Z}[1/N]$-scheme, and let $\xi = (A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(S)$. As usual, from ξ we obtain a canonical closed embedding $i_{\mathrm{can}} : A \hookrightarrow \mathbb{P}^m_S$ which is the composition of the canonical closed embedding $A \hookrightarrow \mathbb{P}(\pi_* L^{\Delta}(3\lambda))$ and the isomorphism $\phi : \mathbb{P}(\pi_* L^{\Delta}(3\lambda)) \xrightarrow{\sim} \mathbb{P}^m_S$. We can view each y_i as a section in $(\mathbb{Z}/N\mathbb{Z})_S(S)$, and thus we get a section $i_{\mathrm{can}}(\gamma(y_i)) \in \mathbb{P}^m_S(S)$. We define $f(\xi) = (i_{\mathrm{can}}(\gamma(y_1)), \cdots, i_{\mathrm{can}}(\gamma(y_{N^{2g}})))$. The above construction $\xi \mapsto f(\xi)$ is functorial in S, so we obtain morphism $f : \mathcal{H}_{g,d,N} \to (\mathbb{P}^m)^{N^{2g}}$, which is clearly G-equivariant.

By Theorem 37.2.3, to complete the proof we only need to check that f lands in the stable locus, and that (*) is satisfied. For the first, it suffices to show that for any algebraically closed field k whose characteristic is coprime to N and any $\xi = (A, \lambda, \gamma, \phi) \in \mathcal{H}_{g,d,N}(k)$, the collection $i_{\mathrm{can}}(A[N](k))$ of N^{2g} points in $\mathbb{P}^m_k(k)$ is stable. By Lemma 37.3.1, it suffices to show that $N > \sqrt{(m+1)r}$ where r is the degree of $i_{\mathrm{can}}(A)$. Recall that the degree of a projective variety is equal to the factorial of the dimension times the leading coefficient of the Hilbert polynomial. We have shown before that the Hilbert polynomial of $i_{\mathrm{can}}(A)$ is $P(T) = 6^g d T^g$. Hence $r = g! 6^g d$. Also $m + 1 = 6^g d$, so the desired bound is $N > 6^g d \sqrt{g!}$, which is what we initially assumed.

To check (*), it suffices to show that $\mathcal{H}_{g,d,N}$ admits a G-equivariant line bundle which is absolutely ample. Now $\mathcal{H}_{g,d,N}$ is a G-equivariant locally closed subscheme of $\mathrm{Hilb}^{P(T)}_{\mathbb{P}^m, 2g+1}$, and it is a general fact that the latter admits a G-equivariant line bundle which is absolutely ample. Pulling back to $\mathcal{H}_{g,d,N}$ we obtain an absolutely ample G-equivariant line bundle. This concludes the proof. □

At this point, we have proved the representability of $\mathcal{A}_{g,d,N}$ for all $N \geq 3$ modulo Lemma 37.3.1.

38. Lecture 38

38.1. **Proving the key lemma using intersection theory.**

Proof of Lemma 37.3.1. We follow [MFK94, Prop. 7.7]. The proof uses intersection theory, which we will assume as a blackbox. First note that it suffices to show that for any given hyperplane $H \subset \mathbb{P}^n_k$ (of codimension 1), we have

$$y := \frac{\#A[N] \cap H(k)}{N^{2g}} < \frac{1}{n+1}.$$

We recall the formalism of Chow ring. The standard reference is [Ful98], also cf. [EH16]. Let X be a smooth projective variety over k of dimension g. The Chow ring of a X is a graded ring $\mathrm{CH}(X) = \bigoplus_{i=0}^{g} \mathrm{CH}^i(X)$, where $\mathrm{CH}^i(X)$ is the quotient of the free abelian group generated by codimension i subvarieties of X (elements of this group are called codimension i cycles on X) modulo rational equivalence. Here rational equivalence is the equivalence relation generated additively by the following rule: We set a codimension i cycle $\sum_j n_j Z_j$, where each $n_j \in \mathbb{Z}$ and Z_j is a codimension i subvariety of X, to be equivalent to 0, if there is a codimension $(i-1)$ subvariety W of X, and a rational function f on W, such that $\sum_j n_j Z_j$ is the divisor of f. Note that we do not require W to be smooth, so the notion of the divisor of a rational function on W is more technical, but we omit the details. The multiplication on $\mathrm{CH}(X)$ is graded-commutative, meaning that for $u \in \mathrm{CH}^i(X)$ and $v \in \mathrm{CH}^j(X)$ we have $u \cdot v = (-1)^{ij} v \cdot u \in \mathrm{CH}^{i+j}(X)$. Intuitively, if U is a codimension i subvariety and V is a codimension j subvariety such that U and V intersect transversally, then we define $[U] \cdot [V]$ to be $[U \cap V]$, where $[\cdot]$ denotes the image in $\mathrm{CH}(X)$. To really get this idea to work and obtain a ring structure on $\mathrm{CH}(X)$ some non-trivial work is needed. Finally, we have a homomorphism $\deg : \mathrm{CH}^g(A) \to \mathbb{Z}$ sending the class of every closed point to 1. We introduce the following notation: For $u \in \mathrm{CH}^i(X)$ and $v \in \mathrm{CH}^{g-i}(X)$, we write $\langle u, v\rangle$ for $|\deg(u \cdot v)| \in \mathbb{Z}_{\geq 0}$. This is regarded as the intersection number of u and v.

Back to our situation, let $h = A \cap H$, a codimension 1 subvariety of A. Pick a codimension $(g-1)$ linear subspace U of $\mathbb{P}^n_k$ and let $\gamma = A \cap U$. Let ψ denote $[N]_A : A \to A$. When U is "general enough", γ is a codimension $(g-1)$ subvariety of A, and both $\psi^*\gamma \cap h$ and $\gamma \cap h$ are codimension g subvarieties of Ain $\mathbb{P}^n_K$. By pre-composing the embedding $A \to \mathbb{P}^n_k$ with some translation on A (which does not change the previous conditions), we may assume that $e_A \in \gamma$. In this case, the intersection number $\langle \psi^*\gamma, h\rangle$ is larger or equal to $\#((\psi^*\gamma) \cap h)(k)$, which is a finite number larger or equal to $N^{2g}y$. We claim that for any $x \in \mathrm{CH}^{g-1}(A)$, we have $\langle x, \psi^* h\rangle = N^2 \langle x, h\rangle$. By the claim, we have

$$y \leq N^{-2g} \langle \psi^*\gamma, h\rangle = N^{-2g-2} \langle \psi^*\gamma, \psi^* h\rangle = N^{-2g-2} \deg(\psi) \cdot \langle \gamma, h\rangle = N^{-2} \langle \gamma, h\rangle.$$

Now we have $\langle \gamma, h\rangle = r$, which implies the desired bound for y.

To prove the claim, we have a natural isomorphism $\mathrm{CH}^1(A) \cong \mathrm{Pic}(A)$. Recall that for $L \in \mathrm{Pic}(A)$, $\psi^* L$ is algebraically equivalent to $L^{\otimes N^2}$. Translating the fact to the Chow group side, we have that $\psi^* h$ is algebraically equivalent to $N^2 \cdot h$ for a suitable definition of algebraic equivalence. Under this equivalence the intersection number does not change, and the claim follows. □

38.2. **Galois representations associated to modular forms.** We now start a brand new topic, the *Mazur–Ribet Theorem* on level-lowering for modular representations over finite fields. This work shows that the Taniyama–Shimura–Weil Conjecture, which states that every elliptic curve over $\mathbb{Q}$ is modular, implies Fermat's Last Theorem, which states that the equation $a^n + b^n = c^n$ for $n \geq 3$ has no integer solution with $abc \neq 0$.

For this part of the course, the original sources are Ribet's two papers [Rib90, Rib94]. For expositions, see [Pra95, Oes88, Edi97]. Many important concepts that we will encounter are excellently explained in the two-volume [Sai13, Sai14].

To start let us recall the basics of modular forms. We will exclusively work with weight-2 cusp forms for $\Gamma_0(N)$ with trivial nebentypus, so in the following we simply call them **cusp forms of level** N**.** We first recall their definition.

Let N be a positive integer and $\Gamma_0(N)$ be the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ consisting of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $c \equiv 0 \pmod N$. By definition, a cusp form of level N is a holomorphic function $f : \mathbb{H}^+ \to \mathbb{C}$ (where $\mathbb{H}^+$ is the upper half plane) satisfying:

(1) For all $\tau \in \mathbb{H}^+$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, we have

$$f(\frac{a\tau + b}{c\tau + d}) = (c\tau + d)^2 f(\tau).$$

(2) f vanishes at the cusps for $\Gamma_0(N)$. Here recall that $Y_0(N) = \Gamma_0(N)\backslash\mathbb{H}^+$ is a non-compact Riemann surface, and it can be completed to a compact Riemann surface $X_0(N)$ by adding finitely many points called **cusps**.[10] The vanishing condition is precisely formulated by expressing f in terms of a local chart of $X_0(N)$ near each cusp.

In fact, the two conditions precisely mean that the holomorphic differential $f(\tau)d\tau$ on $\mathbb{H}^+$ descends to $Y_0(N)$ and then extends to $X_0(N)$. Denote the space of all cusp forms of level N by $\mathscr{C}(N)$. Thus $\mathscr{C}(N)$ is naturally identified with $\mathbf{H}^0(X_0(N), \Omega^1)$, the space of (global) holomorphic differentials on $X_0(N)$. In particular, this is a finite-dimensional $\mathbb{C}$-vector space, of dimension equal to the genus of $X_0(N)$.

For each $f \in \mathscr{C}(N)$, we have a unique q-expansion $f(\tau) = \sum_{n=1}^{\infty} a_n q^n$, where $a_n \in \mathbb{C}$ and $q = e^{2\pi i \tau}$. Conversely, two elements having the same q-expansion are equal. For each prime number $p \nmid N$, we have the **Hecke operator** $T_p : \mathscr{C}(N) \to \mathscr{C}(N)$, which in terms of q-expansions sends $f = \sum_{n=1}^{\infty} a_n q^n$ to

$$\sum_{n=1}^{\infty} a_{np} q^n + \sum_{n=1}^{\infty} p a_n q^{np}.$$

Of course this definition seems ad hoc, and it is not even clear that the above formula defines an element of $\mathscr{C}(N)$. Later we will discuss the geometric origin of T_p.

Now let $f = \sum a_n q^n \in \mathscr{C}(N)$ be a normalized eigenform, meaning that $a_1 = 1$ and f is an eigenvector for all T_p with $p \nmid N$. In this case, a_p is in fact the eigenvalue of f under T_p, and the set $\{a_p\}_{p\nmid N} \subset \mathbb{C}$ is contained in $\mathcal{O}_K$ for a number field K. In particular, the ring

$$R(f) := \mathbb{Z}[a_p, p \nmid N]$$

is a finite free $\mathbb{Z}$-module. These facts follow from the so-called Hecke theory, and we omit all the proofs. We now fix a subfield $\mathbb{F}$ of $\overline{\mathbb{F}}_\ell$ for some prime ℓ, and fix a ring homomorphism $\pi : R(f) \to \mathbb{F}$.

By the work of Shimura and others, we have the following result. (In fact, we have a similar result for cusp forms of weights other than 2 and non-trivial nebentypus as well. For weights ≥ 2, this is due to Deligne [Del71a]; for weight 1, Deligne–Serre [DS74].) We shall refer to it as the Shimura–Deligne Theorem.

[10] Here is the reason why the "missing points" on $Y_0(N)$ are called cusps. On $Y_0(N)$ we have a natural Hermitian metric of constant curvature -1 inherited from that on $\mathbb{H}^+$. According to this metric, the surface becomes infinitely narrow near any cusp and yet the cusp is at infinite distance from any point on the surface.

Theorem 38.2.1 (See for instance [Oes88, §II.2])**.** *In the above setting, associated to $f = \sum a_n q^n$ and $\pi : R(f) \to \mathbb{F}$) there is a unique (up to isomorphism) semi-simple continuous representation $\rho = \rho_{f,\pi} : G_{\mathbb{Q}} := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F})$ (where $\mathrm{GL}_2(\mathbb{F})$ has discrete topology) with the following properties:*

(1) *For each prime $p \nmid N\ell$, ρ is unramified at p. This means that for any choice of decomposition group $D_p \subset G_{\mathbb{Q}}$ at p (well defined up to conjugacy), the inertia subgroup I_p of D_p is mapped to $\{1\}$ under ρ.*

(2) *By (1), for each $p \nmid N\ell$, we have a well-defined conjugacy class $\rho(\mathrm{Frob}_p)$ in $\mathrm{GL}_2(\mathbb{F})$. We require that this conjugacy class has trace $\pi(a_p)$ and determinant p.*

In the theorem, since $G_{\mathbb{Q}}$ is compact, by Galois theory we know that the continuity of ρ just amounts to the existence of a finite Galois extension $F/\mathbb{Q}$ such that ρ factors through $\mathrm{Gal}(F/\mathbb{Q})$. Thus ρ is essentially just a 2-dimensional semi-simple representation of the finite group $\mathrm{Gal}(F/\mathbb{Q})$ over $\mathbb{F}$. (Since $\mathbb{F}$ has positive characteristic, semi-simplicity is not automatic.) The uniqueness of ρ is easy to show, so the essential part of the theorem is the existence.

We shall call the representation ρ as in the theorem a **modular representation over** $\mathbb{F}$, where the word "modular" means "coming from a modular form".[11] More precisely, we say that ρ is **modular of level** N. The purpose of this language is that we want to emphasize ρ more than remembering f; we just want to remember the representation ρ itself, and remember the fact that it is associated to some unspecified f in $\mathscr{C}(N)$. It is totally possible that a given ρ can be modular of level N and modular of level N' simultaneously, and that phenomenon is in fact what the Mazur–Ribet Theorem is trying to understand.

39. Lecture 39

39.1. Statement of the Mazur–Ribet Theorem.

Let us recall the concept of a modular representation from the last lecture.[12]

Definition 39.1.1. Let $\mathbb{F}$ be a subfield of $\overline{\mathbb{F}}_\ell$ for some prime ℓ. Let N be a positive integer. A representation $\rho : G_{\mathbb{Q}} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F})$ is called **modular of level** N, if it satisfies the following conditions:

(1) ρ is continuous and semi-simple, i.e., there exists a finite Galois extension $F/\mathbb{Q}$ such that ρ comes from a semi-simple representation $\mathrm{Gal}(F/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F})$.

(2) ρ is unramified at each prime $p \nmid N\ell$.

(3) There exists a normalized eigenform $f = \sum_n a_n q^n \in \mathscr{C}(N)$ and a ring homomorphism $\pi : \mathbb{Z}[a_p, p \nmid N] \to \mathbb{F}$ such that for each prime $p \nmid N\ell$ the element $\rho(\mathrm{Frob}_p) \in \mathrm{GL}_2(\mathbb{F})$ (well defined up to conjugacy) has trace $\pi(a_p)$ and determinant p.

Remark. (1) If ρ is modular, then the character $\det \rho : G_{\mathbb{Q}} \to \mathbb{F}^\times$ sends $\mathrm{Frob_p}$ to p for almost all p. Using class field theory (for $\mathbb{Q}$) we see that $\det \rho$ must be given by the ℓ-th cyclotomic character $\chi_\ell : G_{\mathbb{Q}} \to \mathbb{F}_\ell^\times$. The latter is defined by the rule $\sigma(\zeta) = \zeta^{\chi_\ell(\sigma)}$ for all $\sigma \in G_{\mathbb{Q}}$ and any primitive ℓ-th root of unity $\zeta \in \overline{\mathbb{Q}}$. In particular, if $c \in G_{\mathbb{Q}}$ is any choice of complex conjugation (well defined up to

[11] In the context of representation theory, the word "modular" usually just means representations over positive characteristic, as in "mod" p.

[12] In this course we only consider Galois representations associated to weight-2 eigenforms, and we omit the adjective "weight-2" in the definition of a modular representation. In general the notion of a modular representation should allow for representations associated to eigenforms of other weights.

conjugacy), then $\rho(c)$ is of order 2 and has determinant -1, and hence conjugate to the diagonal matrix $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$.

(2) Suppose ρ is modular and irreducible, and $\ell \geq 3$. Then ρ is absolutely irreducible, i.e., irreducible when base changed to $\overline{\mathbb{F}}$. In fact, if over $\overline{\mathbb{F}}$ there is a 1-dimensional subrepresentation, then that must be equal to the eigenspace of $\rho(c)$ over $\overline{\mathbb{F}}$ of eigenvalue either 1 or -1. But $\rho(c)$ is already diagonalizable over $\mathbb{F}$, so that 1-dimensional subrepresentation is already defined over $\mathbb{F}$, a contradiction.

(3) Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$ be an arbitrary semi-simple continuous representation, where $\mathbb{F} \subset \overline{\mathbb{F}}_\ell$. We say that ρ is defined over a subfield $\mathbb{F}' \subset \mathbb{F}$, if there is a semi-simple continuous representation of $G_{\mathbb{Q}}$ over $\mathbb{F}'$ whose base change to $\mathbb{F}$ is isomorphic to ρ. Let $\mathbb{F}_{\min}$ be the subfield of $\mathbb{F}$ generated by the coefficients of the characteristic polynomials of $\rho(g)$ for all $g \in G_{\mathbb{Q}}$. Clearly $\mathbb{F}_{\min}$ is a finite field, and is contained in any field of definition of ρ. Then ρ is in fact defined over $\mathbb{F}_{\min}$. To see this, we immediately reduce to the case where $\mathbb{F}$ is finite, and we may replace $G_{\mathbb{Q}}$ by its finite image under ρ. Then we use the following general fact (see [DS74, Lem. 6.13]): Any semi-simple representation ρ of a finite group G over a finite field $\mathbb{F}$ is defined over the minimal subfield of $\mathbb{F}$ containing the coefficients of the characteristic polynomials of $\rho(g)$ for all $g \in G$.

We can now state the Mazur–Ribet Theorem.

Theorem 39.1.2 (Mazur–Ribet). *Let $\mathbb{F} \subset \overline{\mathbb{F}}_\ell$, and let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$ be a modular representation of level N. Assume that $\mathbb{F} = \mathbb{F}_{\min}$ is the minimal field of definition of ρ, so in particular $\mathbb{F}$ is finite. Assume that ρ is irreducible. Let p be an odd prime dividing N precisely once (we write $p \| N$), satisfying the following conditions:*

(1) *ρ is finite at p (see below).*
(2) *Either $p \not\equiv 1 \pmod{\ell}$, or $\ell \nmid N$.*

Then ρ is modular of level N/p.

We explain the condition "finite at p". Let us agree that "$\mathbb{F}$-vector space" always means "finite-dimensional $\mathbb{F}$-vector space". We have an equivalence of categories between the category of $\mathbb{F}$-vector space schemes over $\mathbb{Q}$ (i.e., finite commutative group schemes V over $\mathbb{Q}$ together with endomorphisms $\lambda : V \to V$ for all $\lambda \in \mathbb{F}$ satisfying the axioms for scalar multiplication in a vector space; or equivalently, $\mathbb{Q}$-schemes V together with $\mathbb{F}$-vector space structures on $V(S)$ for all test $\mathbb{Q}$-schemes S which are functorial in S) and the category of continuous representations of $G_{\mathbb{Q}}$ on $\mathbb{F}$-vector spaces. The equivalence is given by $V \mapsto V(\overline{\mathbb{Q}})$. Now suppose V is the $\mathbb{F}$-vector space scheme over $\mathbb{Q}$ corresponding to ρ. We say that ρ is **finite at** p, if the $\mathbb{F}$-vector space scheme $V \times_{\operatorname{Spec}\mathbb{Q}} \operatorname{Spec}\mathbb{Q}_p$ over $\mathbb{Q}_p$ extends to an $\mathbb{F}$-vector space scheme over $\mathbb{Z}_p$ that is *flat* over $\mathbb{Z}_p$.

We comment that if $p \neq \ell$, then ρ is finite at p if and only if it is unramified at p.

39.2. **Taniyama–Shimura–Weil implies Fermat.** Suppose Fermat's Last Theorem is false. Then we can find non-zero integers a, b, c, and a prime $\ell \geq 5$ such that $a^\ell + b^\ell = c^\ell$ and $a \equiv 3 \pmod 4$, $b \equiv 0 \pmod 2$. (Here we can assume $\ell \geq 5$ because FLT is known for exponent 4.) We then construct the "Frey curve" E, an elliptic curve over $\mathbb{Q}$ given by $y^2 = x(x - a^\ell)(x + b^\ell)$. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_\ell)$ be the continuous representation arising from the natural $G_{\mathbb{Q}}$-action on the 2-dimensional $\mathbb{F}_\ell$-vector space $E[\ell](\overline{\mathbb{Q}})$. The curve E enjoys the following "incredibly good" properties (cf. [DDT97, Prop. 2.15]):

(1) $E[2](\overline{\mathbb{Q}}) = E[2](\mathbb{Q})$. (This alone is not uncommon.) In particular, by results of Mazur, ρ is irreducible (see [Ser87, §4.1, Prop. 6]).
(2) The conductor N_E of E is square-free, or equivalently, E has semi-stable reduction everywhere; see [Ser87, §4.1]. (Again, this alone is not uncommon.) In fact N_E is the product of all the distinct prime divisors of abc.
(3) ρ is finite at every odd prime p. Equivalently, $E[\ell]_{\mathbb{Q}_p}$ extends to a flat group scheme over $\mathbb{Z}_p$. If $p \nmid N_E$, then $E_{\mathbb{Q}_p}$ extends to an elliptic curve over $\mathbb{Z}_p$, so this is clear. Suppose $3 \leq p | N_E$. Then since we already know E has multiplicative reduction at p, the criterion for ρ to be finite at p is that $\ell | v_p(j_E)$ (see [DDT97, Prop. 2.12]). As computed in [Ser87, (4.1.9)], we have
$$v_p(j_E) = -2v_p(a^\ell b^\ell c^\ell) = -2\ell v_p(abc),$$
which is indeed divisible by ℓ.

Now by Taniyama–Shimura–Weil, E is modular. Whatever that means, that implies that ρ is modular of level N_E. Since N_E is square-free and even, and since ρ is irreducible and finite at every odd prime p, we can repeatedly apply Theorem 39.1.2 to conclude that ρ is modular of level 2. In particular this implies that $\mathscr{C}(2) = \mathbf{H}^0(X_0(2), \Omega)$ is non-zero. This is absurd since $X_0(2)$ has genus 0.

Now a few words on TSW. As we observed in (2) above, the Frey curve E, if exists, has semi-stable reduction everywhere, or simply *semi-stable over* $\mathbb{Q}$. Thus for the previous argument to work one only needs TSW for semi-stable elliptic curves over $\mathbb{Q}$; this is what Wiles and Taylor–Wiles proved around 1995. Later, around 2000, the full TSW was proved by Breuil–Conrad–Diamond–Taylor, known as the Modularity Theorem.

40. Lecture 40

40.1. **Mazur–Ribet and Fermat's Last Theorem.** Recall the statement of the Mazur–Ribet Theorem:

Theorem 40.1.1 (Mazur–Ribet)**.** *Let $\mathbb{F}$ be a subfield of $\overline{\mathbb{F}}_\ell$ for some odd prime ℓ. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$ be an irreducible modular representation of level N. Assume that $\mathbb{F} = \mathbb{F}_{\min}$ is the minimal field of definition and in particular $\mathbb{F}$ is finite. Let p be an odd prime such that $p \| N$ and such that the following two conditions hold:*

(1) *ρ is finite at p, that is, there exists an $\mathbb{F}$-vector space (of finite $\mathbb{F}$-dimension) scheme $\mathcal{V}/\mathbb{Z}_p$ flat over $\mathbb{Z}_p$ such that the action of $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$ on $\mathcal{V}(\overline{\mathbb{Q}}_p)$ is isomorphic (as a continuous $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-representation on a two-dimensional $\mathbb{F}$-vector space) to $\rho|_{D_p}$, where D_p is any choice of a decomposition group at p in $G_{\mathbb{Q}}$.*
(2) *Either of the following holds:*
   (a) $p \neq 1 \mod \ell$
   (b) $\ell \nmid N$.

Then ρ is modular of level N/p.

Note here that the case of (2)(a) is due to Mazur and (2)(b) is due to Ribet.

We now recall how the Taniyama–Shimura–Weil Conjecture implies Fermat's Last Theorem. Assume that Fermat's Last Theorem is falst. Then we may find $a^\ell + b^\ell = c^\ell$ with $abc \neq 0$ where ℓ is a prime such that $\ell \geq 5$. Further, we may always arrange that $a \equiv 3 \mod 4$ and b is even. Then we obtain an elliptic curve $E/\mathbb{Q}$ given by
$$E : y^2 = x(x - a^\ell)(x + b^\ell).$$

Further, define a continuous representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_\ell)$ by the action of $G_{\mathbb{Q}}$ on $E[\ell](\overline{\mathbb{Q}}) \cong \mathbb{F}_\ell^2$. Using properties of E, we can prove that ρ is irreducible and finite at every odd prime p, and further that the conductor N_E of E is square free. If the Taniyama–Shimura–Weil Conjecture is true for E, then ρ is modular of level N_E. Now there are two cases:

Case 1: Suppose $\ell | N_E$. Then take $p = \ell$. This obviously gives $p \not\equiv 1 \mod \ell$. By Mazur's part of Theorem 40.1.1, we see that p is modular of level N_E/p. Hence we have reduced to the following case.

Case 2: ρ is modular of some level N which is square free and coprime to ℓ. By Ribet's part of Theorem 40.1.1, we see that ρ is modular of level given by N divided by all odd prime factors of N. This leaves the level as being either 1 or 2. (In fact, N_E is even, so here can assume the level is 2.) But $\mathscr{C}(1) = \mathscr{C}(2) = 0$. This is a contradiction.

40.2. **Reformulation of modular representations.** Suppose $N \geq 1$ is an integer. For each prime $p \nmid N$, recall the Hecke operator

$$T_p : \mathscr{C}(N) \longrightarrow \mathscr{C}(N)$$

which in terms of q-expansions sends $f = \sum_{n=1}^{\infty} a_n q^n$ to

$$\sum_{n=1}^{\infty} a_{np} q^n + \sum_{n=1}^{\infty} p a_n q^{np}.$$

For $p | N$, we also have a Hecke operator

$$T_p : \mathscr{C}(N) \longrightarrow \mathscr{C}(N)$$

via

$$f = \sum_{n=1}^{\infty} a_n q^n \longmapsto \sum_{n=1}^{\infty} a_{np} q^n.$$

The T_p's all commute with each other. We define T_n for all integers $n \geq 1$ by equating the two formal series

$$\sum_{n=1}^{\infty} \frac{T_n}{n^s} = \prod_{p \nmid N} (1 - T_p p^{-s} + p \cdot p^{-2s})^{-1} \prod_{p | N} (1 - T_p p^{-s})^{-1}.$$

(This defines each T_n as an integral-coefficient polynomial in the T_p's with $p|n$.) The **Hecke algebra** $\mathbb{T}(N)$ is the subring of $\mathrm{End}_{\mathbb{C}}(\mathscr{C}(N))$ generated by T_p for all primes p, or equivalently, generated by T_n for all $n \geq 1$.

Fact 40.2.1. *$\mathbb{T}(N)$ is a commutative ring, and is a finite free $\mathbb{Z}$-module.*

Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}(N)$. Let $k_{\mathfrak{m}} := \mathbb{T}(N)/\mathfrak{m}$ be the residue field. Since $\mathbb{T}(N)$ is finite over $\mathbb{Z}$, $k_{\mathfrak{m}}$ is a finite field. Let $\ell = \operatorname{char} k_{\mathfrak{m}}$.

Theorem 40.2.2 (Reformulation of the Shimura–Deligne Theorem)**.** *For the fixed $\mathfrak{m}$, there exists a unique (up to isomorphism) continuous semi-simple representation $\rho_{\mathfrak{m}} : G_{\mathbb{Q}} \to \mathrm{GL}_2(k_{\mathfrak{m}})$ such that for all primes $p \nmid N\ell$, we have*

1. *$\rho_{\mathfrak{m}}$ is unramified at p.*
2. *$\mathrm{Tr}(\rho_{\mathfrak{m}}(\mathrm{Frob}_p)) = T_p \mod \mathfrak{m}$.*
3. *$\det(\rho_{\mathfrak{m}}(\mathrm{Frob}_p)) = p$.*

For a proof of this result admitting the Shimura–Deligne theorem, see [Rib90, Prop. 5.1].

Remark. (1) $\rho_{\mathfrak{m}}$ is a modular representation of level N in the previous sense.

(2) For an arbitrary continuous semi-simple $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$ where $\mathbb{F} \subset \overline{\mathbb{F}}_\ell$, we have that ρ is modular of level N if and only if there exists a maximal ideal $\mathfrak{m} \subset \mathbb{T}(N)$ together with $k_{\mathfrak{m}} \hookrightarrow \overline{\mathbb{F}}_\ell$ such that $\rho \otimes_{\mathbb{F}} \overline{\mathbb{F}}_\ell$ is isomorphic to $\rho_{\mathfrak{m}} \otimes_{k_{\mathfrak{m}}} \overline{\mathbb{F}}_\ell$.

40.3. **Canonical models of modular curves.** Recall that $Y_0(N) := \Gamma_0(N)\backslash\mathbb{H}^+$ is a non-compact Riemann surface, and it sits inside a compact Riemann surface $X_0(N)$ such that $X_0(N) - Y_0(N)$ consists of finitely many cusps. Both $Y_0(N)$ and $X_0(N)$ are algebraic curves over $\mathbb{C}$ and have canonical models over $\mathbb{Q}$. To construct the $\mathbb{Q}$-models, take $M \geq 3$ divisible by N. By our previous work, we have the smooth affine curve $S(M)$ over $\mathbb{Z}[1/M]$, and $Y(M) = S(M)(\mathbb{C})$ is identified with a disjoint union of $\phi(M)$ copies of $\Gamma(M)\backslash\mathbb{H}^+$. There is a natural map $Y(M) \to Y_0(N)$ which is a (possibly ramified) quotient (as Riemann surfaces, or as $\mathbb{C}$-varieties) by the action of some finite group $\Delta_{M,N}$ on $Y(M)$. This action actually comes from an action of $\Delta_{M,N}$ on $S(M)_{\mathbb{Q}}$ as a $\mathbb{Q}$-variety. Since $S(M)_{\mathbb{Q}}$ is quasi-projective, we can take the quotient $Y_0(N)_{\mathbb{Q}} := \Delta_{M,N}\backslash S(M)_{\mathbb{Q}}$ in the category of $\mathbb{Q}$-schemes (see for instance [Sai13, §A.3]). This gives the model of $Y_0(N)$ over $\mathbb{Q}$.

We caution that when N is small, the action of $\Delta_{M,N}$ on $Y(M)$ may not be free (correspondingly $Y(M) \to Y_0(M)$ may be a ramified covering). Hence the finite $\mathbb{Q}$-morphism $S(M)_{\mathbb{Q}} \to Y_0(N)_{\mathbb{Q}}$ may not be étale.

We know that $Y_0(1)_{\mathbb{Q}}$ is canonically isomorphic to $\mathbb{A}^1_{\mathbb{Q}}$ (induced by the j-invariant $j : S(M)_{\mathbb{Q}} \to \mathbb{A}^1_{\mathbb{Q}}$). The canonical embedding $\mathbb{A}^1_{\mathbb{Q}} \hookrightarrow \mathbb{P}^1_{\mathbb{Q}}$ is a $\mathbb{Q}$-model for $Y_0(1) \hookrightarrow X_0(1)$. We thus take $X_0(1)_{\mathbb{Q}}$ to be $\mathbb{P}^1_{\mathbb{Q}}$. We have a commutative diagram of $\mathbb{C}$-varieties

$$\begin{tikzcd} Y_0(N) \arrow[r, hook] \arrow[d] & X_0(N) \arrow[d] \\ Y_0(1) = \mathbb{A}^1_{\mathbb{C}} \arrow[r, hook] & X_0(1) = \mathbb{P}^1_{\mathbb{C}} \end{tikzcd}$$

such that $X_0(N)$ is the integral closure of $X_0(1)$ in $Y_0(N)$. Moreover, the vertical map on the left and the bottom map are both defined over $\mathbb{Q}$. We thus construct $X_0(N)_{\mathbb{Q}}$ by taking the integral closure of $X_0(1)_{\mathbb{Q}}$ in $Y_0(N)_{\mathbb{Q}}$. Thus we have a commutative diagram of $\mathbb{Q}$-varieties whose base change to $\mathbb{C}$ recovers the previous diagram:

$$\begin{tikzcd} Y_0(N)_{\mathbb{Q}} \arrow[r, hook] \arrow[d] & X_0(N)_{\mathbb{Q}} \arrow[d] \\ Y_0(1)_{\mathbb{Q}} = \mathbb{A}^1_{\mathbb{Q}} \arrow[r, hook] & X_0(1)_{\mathbb{Q}} = \mathbb{P}^1_{\mathbb{Q}} \end{tikzcd}$$

Fact 40.3.1. *The variety $X_0(N)_{\mathbb{Q}}$ is a smooth projective curve over $\mathbb{Q}$, and $Y_0(N)_{\mathbb{Q}}$ is a dense open subvariety.*

Fact 40.3.2. *The $\mathbb{Q}$-variety $Y_0(N)_{\mathbb{Q}}$ is a coarse moduli space for the moduli functor over $\mathbb{Q}$ of elliptic curves with $\Gamma_0(N)$-level structure. Here, a $\Gamma_0(N)$-level structure on an elliptic curve $E \to S$ (for a $\mathbb{Q}$-scheme S) refers to the datum of a closed S-subgroup scheme $C \subset E$ such that C is finite étale over S and all geometric fibers of C over S are isomorphic to $\mathbb{Z}/N\mathbb{Z}$; often we refer to such a datum as "a cyclic subgroup of order N". Similarly, $X_0(N)_{\mathbb{Q}}$ is also a coarse moduli space, where we allow certain degenerate variants of elliptic curves.*

Here, for any scheme S, to say that an S-scheme Y is a **coarse moduli space** for a functor $\mathcal{Y}$ from S-schemes to sets, we mean that there is a morphism $\psi : \mathcal{Y} \to Y$ which is initial among all morphisms from $\mathcal{Y}$ to S-schemes, and such that for any geometric point

$s : \operatorname{Spec} \bar{k} \to S$ the map $\mathcal{Y}(\bar{k}) \to Y(\bar{k})$ is a bijection. Thus we allow $\mathcal{Y}$ to be not representable, but we ask that Y should be the "best approximation" of $\mathcal{Y}$ and "has the expected geometric points". It is easily seen that for fixed $\mathcal{Y}$, there can exist at most one coarse moduli space up to canonical isomorphism. Moreover if $\mathcal{Y}$ is representable then the coarse moduli space is the scheme representing $\mathcal{Y}$ (in which case we say it is a fine moduli space). Thus the above fact gives an abstract characterization of the $\mathbb{Q}$-varieties $Y_0(N)_{\mathbb{Q}}$ and $X_0(N)_{\mathbb{Q}}$. In particular, we know that they are actually independent of the choice of M in the construction.

For more details on $Y_0(N)_{\mathbb{Q}}$ and $X_0(N)_{\mathbb{Q}}$, the reader should consult [Sai13, §§2.1–2.4]

41. Lecture 41

41.1. **Jacobians of modular curves.** From now on, we simply use the notations $Y_0(N)$ and $X_0(N)$ to denote the canonical models of modular curves over $\mathbb{Q}$.

Recall that $X_0(N)$ is a smooth projective curve over $\mathbb{Q}$, and $Y_0(N)$ is a dense open subvariety. Consider the Jacobian of $X_0(N)$, denoted by $J_0(N) := \operatorname{Jac}(X_0(N)) = \operatorname{Pic}^0_{X_0(N)/\mathbb{Q}}$. This is the identity component of $\operatorname{Pic}_{X_0(N)/\mathbb{Q}}$, which represents the fppf sheafification of the usual relative Picard functor for $X_0(N)/\mathbb{Q}$. It is a fact that $J_0(N)$ is an abelian variety over $\mathbb{Q}$.

Now let n be a positive integer. We can define a smooth (perhaps disconnected) curve $Y_0(N,n)$ over $\mathbb{Q}$ which is a coarse moduli space for tuples $(E_1, C_1, E_2, C_2, \phi)$ where $(E_i, C_i) \in Y_0(N)$ and $\phi : E_1 \to E_2$ is a degree n isogeny such that $\phi(C_1) = C_2$. In turn, we obtain an algebraic correspondence $Y_0(N) \dashrightarrow Y_0(N)$ by the following diagram, where the solid arrows send $(E_1, C_1, E_2, C_2, \phi)$ to (E_1, C_1) and to (E_2, C_2) respectively, and the dashed arrow should be thought of as a "one-to-many map".

$$\begin{tikzcd} & Y_0(N,n) \arrow[dl] \arrow[dr] & \\ Y_0(N) \arrow[rr, dashed] & & Y_0(N). \end{tikzcd}$$

This correspondence naturally extends to an algebraic correspondence $X_0(N) \dashrightarrow X_0(N)$, which we call the n-th Hecke correspondence. By the functoriality[13] of the Jacobian, for each n the n-th Hecke correspondence $X_0(N) \dashrightarrow X_0(N)$ induces a homomorphism

$$T_n^{\text{geom}} : J_0(N) \longrightarrow J_0(N).$$

Note that the cotangent space $\operatorname{Hom}_{\mathbb{Q}}(\operatorname{Lie} J_0(N)^\vee, \mathbb{Q})$ of $J_0(N)^\vee$ at identity is isomorphic to $\mathbf{H}^0(X_0(N), \Omega)$. As such, we have

$$\operatorname{Hom}_{\mathbb{Q}}(\operatorname{Lie} J_0(N)^\vee, \mathbb{C}) \cong \mathscr{C}(N).$$

By functoriality, T_n^{geom} induces an endomorphism of $\mathscr{C}(N)$. It is a fact that this endomorphism is T_n defined before in terms of q-expansions. In view of this, and since the functor sending an abelian variety to its cotangent space at identity is faithful, we obtain a well-defined faithful action of $\mathbb{T}(N)$ on $J_0(N)$ such that $T_n \in \mathbb{T}(N)$ acts by T_n^{geom}. In other words, we obtain an injective ring homomorphism

$$\mathbb{T}(N) \longrightarrow \operatorname{End}_{\mathbb{Q}}(J_0(N)), \quad T_n \longmapsto T_n^{\text{geom}}, \ \forall n \geq 1$$

[13] In fact, the Jacobian of a smooth projective curve admits a canonical degree 1 polarization, and therefore it is both a contravariant functor and a covariant functor in the curve. These two functorialities are respectively referred to as "Picard functoriality" and "Albanese functoriality" in the literature. Here we need to use the contravariant, Picard functoriality.

where $\mathrm{End}_{\mathbb{Q}}(J_0(N))$ indicates endomorphisms of $J_0(N)$ in the category of abelian varieties over $\mathbb{Q}$. From now on we do not distinguish between T_n and T_n^{geom}.

Now for any maximal ideal $\mathfrak{m} \subset \mathbb{T}(N)$, we define

$$\rho_{\mathfrak{m}}^{\mathrm{Jac}} := J_0(N)(\overline{\mathbb{Q}})[\mathfrak{m}] = \{x \in J_0(N)(\overline{\mathbb{Q}}) \mid \forall f \in \mathfrak{m},\ f(x) = 0\}.$$

(Similarly we shall apply the notation $M[\mathfrak{m}]$ to any other $\mathbb{T}(N)$-module M to mean $\mathfrak{m}$-torsion.) Note that this is a finite set: If ℓ denotes the characteristic of $k_{\mathfrak{m}}$, then our set is contained in $J_0(N)(\overline{\mathbb{Q}})[\ell]$, which is of order ℓ^{2g} with $g = \dim J_0(N) = g(X_0(N))$. In particular, $\rho_{\mathfrak{m}}^{\mathrm{Jac}}$ is a finite-dimensional $k_{\mathfrak{m}}$-vector space (induced by the $\mathbb{T}(N)$-module structure) with a compatible $G_{\mathbb{Q}}$-action. We have the following result:

Theorem 41.1.1. *Let $\mathfrak{m} \subset \mathbb{T}(N)$ be a maximal ideal such that $\rho_{\mathfrak{m}}$ is irreducible. Then $\rho_{\mathfrak{m}}^{\mathrm{Jac}}$ is isomorphic to $\rho_{\mathfrak{m}}^{\oplus d}$ as a $k_{\mathfrak{m}}[G_{\mathbb{Q}}]$-module for some positive integer d. (Necessarily $d \leq g$.) If $\operatorname{char} k_{\mathfrak{m}} \nmid 2N$, then $d = 1$.*

This result tells us that all the information about irreducible modular representations is encoded in the $\mathbb{T}(N)$-action on $J_0(N)$.

41.2. **Reduction of $J_0(N)$ modulo a prime.** For an arbitrary abelian variety A over $\mathbb{Q}_p$, there exists a unique (up to canonical isomorphism) smooth commutative group scheme $\mathcal{A}$ over $\mathbb{Z}_p$ with generic fiber identified with A, called the **Néron model**, characterized by the following "Néron extension property": for all smooth $\mathbb{Z}_p$-schemes T, the natural map $\mathrm{Mor}_{\mathbb{Z}_p}(T, \mathcal{A}) \to \mathrm{Mor}_{\mathbb{Q}_p}(T_{\mathbb{Q}_p}, A)$ is a bijiection. In general, $\mathcal{A}$ is not necessarily proper, so it is not an abelian scheme.

We write $A_{\mathbb{F}_p}$ for $\mathcal{A}_{\mathbb{F}_p}$. With this notation, we are interested in $J_0(N)_{\mathbb{F}_p}$, which is a smooth commutative group scheme over $\mathbb{F}_p$ with a $\mathbb{T}(N)$-action (inherited functorially from the $\mathbb{T}(N)$-action on $J_0(N)$). We study two important cases.

Case 1: good reduction. Suppose $p \nmid N$. Then $X_0(N)_{\mathbb{Q}_p}$ has a (canonical) integral model $\mathfrak{X}_0(N)$ over $\mathbb{Z}_p$ which is smooth projective over $\mathbb{Z}_p$ has a similar description as a coarse moduli space over $\mathbb{Z}_p$ of elliptic curves (with some permitted degenerations) with level structure. This can be constructed from $S(M) \otimes_{\mathbb{Z}[1/M]} \mathbb{Z}_p$ for some suitable M just as how we constructed $X_0(N)$ over $\mathbb{Q}$ from $S(M)_{\mathbb{Q}}$. The Néron model of $J_0(N)_{\mathbb{Q}_p}$ over $\mathbb{Z}_p$ is an abelian scheme, and in fact it is identified with $\mathrm{Pic}^0_{\mathfrak{X}_0(N)/\mathbb{Z}_p}$. The special fiber $J_0(N)_{\mathbb{F}_p}$ is also an abelian variety, and it is identified with the Jacobian of the curve $\mathfrak{X}_0(N)_{\mathbb{F}_p}$. In this case, we have the following injections of rings:

$$\mathbb{T}(N) \hookrightarrow \mathrm{End}_{\mathbb{Q}_p}(J_0(N)_{\mathbb{Q}_p}) \hookrightarrow \mathrm{End}_{\mathbb{F}_p}(J_0(N)_{\mathbb{F}_p}).$$

Case 2: semi-stable reduction. Suppose $p \| N$. In this case, we have the Deligne–Rapoport integral model $\mathfrak{X} = \mathfrak{X}_0(N)$ of $X_0(N)_{\mathbb{Q}_p}$. This is projective and flat over $\mathbb{Z}_p$. Moreover, $\mathfrak{X}$ is **weakly semi-stable**, meaning that the special fiber $\mathfrak{X}_{\mathbb{F}_p}$ is a curve with at worst nodal singularities, and $\mathfrak{X} \to \operatorname{Spec} \mathbb{Z}_p$ is smooth away from the nodes in $\mathfrak{X}_{\mathbb{F}_p}$. In fact, $\mathfrak{X}_{\mathbb{F}_p}$ is the union of two curves C_1 and C_2, each canonically isomorphic to $\mathfrak{X}_0(N/p)_{\mathbb{F}_p}$ (hence projective and smooth), intersecting transversally at finitely many points. In particular, $C_1 \cap C_2$ is a finite scheme over $\mathbb{F}_p$. Moreover, $C_1 \cap C_2$ is the supersingular locus $\mathfrak{X}_0(N/p)^{\mathrm{ss}}_{\mathbb{F}_p}$ (i.e., the coarse moduli space of supersingular elliptic curves plus level structure) inside each of $C_1 \cong \mathfrak{X}_0(N/p)_{\mathbb{F}_p}$ and $C_2 \cong \mathfrak{X}_0(N/p)_{\mathbb{F}_p}$. However, the way that C_1 and C_2 are glued together is not via the identity map from the supersingular locus in $C_1 \cong \mathfrak{X}_0(N/p)_{\mathbb{F}_p}$ to the supersingular locus in $C_2 \cong \mathfrak{X}_0(N/p)_{\mathbb{F}_p}$; rather, it is via the absolute p-Frobenius. We caution the reader that the two maps $\mathfrak{X}_0(N/p)_{\mathbb{F}_p} \xrightarrow{\sim} C_i \hookrightarrow \mathfrak{X}_{\mathbb{F}_p}$ for $i = 1, 2$ do not lift to characteristic zero.

As we have mentioned, the structure map $\mathfrak{X} \to \operatorname{Spec} \mathbb{Z}_p$ is smooth away from $C_1 \cap C_2$. The singularities at $C_1 \cap C_2$ have the following description: Set $W = W(\overline{\mathbb{F}}_p)$ to be the Witt vectors over $\overline{\mathbb{F}}_p$. For each $x \in (C_1 \cap C_2)(\overline{\mathbb{F}}_p)$, the complete local ring of $\mathfrak{X}_W$ at x is isomorphic to

$$W[[S,T]]/(ST - p^{e(x)})$$

as a W-algebra (as is true in general for any weakly semi-stable curve over $\mathbb{Z}_p$; see [Sai14, Lem. B.9]), where $e(x)$ is some positive integer determined by the following recipe. We view x as a point of $C_1(\overline{\mathbb{F}}_p) \cong \mathfrak{X}_0(N/p)(\overline{\mathbb{F}}_p)$, and so it corresponds to a pair (E, C) where E is a (supersingular) elliptic curve over $\overline{\mathbb{F}}_p$ and C is a cyclic subgroup of order N/p. Then we have

$$e(x) = \frac{1}{2} \# \operatorname{Aut}(E, C).$$

(Note that $\{\pm 1\}$ is always a subgroup of $\operatorname{Aut}(E, C)$, so the order of the latter is even.)

Using the above description of $\mathfrak{X}$ and general work of Raynaud on Néron models of Jacobians, Ribet showed that $J = J_0(N)_{\mathbb{F}_p}$ can be described as follows:

(1) There is a Hecke-equivariant short exact sequence

$$1 \to J^\circ \to J \to \Phi \to 1$$

where Φ is a constant finite group scheme.

(2) There is a Hecke equivariant short exact sequence

$$1 \to T \to J^\circ \to A \to 1$$

where T is a torus over $\mathbb{F}_p$ and A is an abelian scheme over $\mathbb{F}_p$.

(3) We have an isomorphism

$$A \cong J_0(N/p)_{\mathbb{F}_p} \times J_0(N/p)_{\mathbb{F}_p}$$

equivariant for T_n with n coprime to N. We may say that "A sees C_1 and C_2".

(4) Write $\widehat{T}$ for $X^*(T)$. As a $\operatorname{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$-module, $\widehat{T}$ is isomorphic to the group of formal integral linear combinations of $\mathfrak{X}_0(N/p)^{\mathrm{ss}}_{\mathbb{F}_p}(\overline{\mathbb{F}}_p)$ with the coefficients adding to 0. In particular, T splits over $\mathbb{F}_{p^2}$ since every point of $\mathfrak{X}_0(N/p)^{\mathrm{ss}}_{\mathbb{F}_p}(\overline{\mathbb{F}}_p)$ is defined over $\mathbb{F}_p$. Moreover, T_p acts on $\widehat{T}$ in the same way as Frob_p. In particular, this action is an involution (i.e., squaring to the identity). We may say that "T sees $C_1 \cap C_2$".

(5) The $\mathbb{T}(N)$-module Φ is such that for every prime $q \nmid N$, T_q acts in the same way as $1 + q$. One can describe Φ more precisely using the function $(C_1 \cap C_2)(\overline{\mathbb{F}}_p) \to \mathbb{Z}_{\geq 1}, x \mapsto e(x)$, but we omit it. We may say that "Φ sees $C_1 \cap C_2$ as well as the local equations of $\mathfrak{X}_W$ at these points".

42. Lecture 42

42.1. Proof of Mazur's theorem.

Lemma 42.1.1. *Let ρ_1, ρ_2 be two continuous semi-simple representations $G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$, where $\operatorname{char} \mathbb{F} = \ell > 2$. Note that for almost all primes p, ρ_1 and ρ_2 are unramified at p, so $\operatorname{Tr}(\rho_1(\operatorname{Frob}_p))$ and $\operatorname{Tr}(\rho_2(\operatorname{Frob}_p))$ are well defined. Assume that these two traces are equal for almost all p. Then ρ_1 and ρ_2 are isomorphic.*

Proof. Recall that the Brauer–Nesbitt theorem says that two n-dimensional semi-simple representations of a finite group over a field k are isomorphic if they have the same character and either $\operatorname{char} k = 0$ or $\operatorname{char} k > n$. We can view ρ_1 and ρ_2 as semi-simple representations of $\operatorname{Gal}(F/\mathbb{Q})$ for some finite Galois extension $F/\mathbb{Q}$. By the Chebotarev density theorem, every conjugacy class in $\mathrm{Gal}(F/\mathbb{Q})$ contains Frob_p for infinitely many primes p (among those which are unramified in F). The lemma follows from these two theorems. $\square$

Definition 42.1.2. Let $N \geq 1$ and let M be a $\mathbb{T}(N)$-module. We say that M is **Eisenstein**, if for almost all primes q we have $T_q = q+1$ in $\mathrm{End}(M)$.

Example. Suppose $p \| N$. The component group Φ of $J_0(N)_{\mathbb{F}_p}$ from last lecture is an Eisenstein $\mathbb{T}(N)$-module.

Lemma 42.1.3. *Let M be an Eisenstein $\mathbb{T}(N)$-module. Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}(N)$ with residue characteristic $\ell > 2$. If $\rho_{\mathfrak{m}}$ is irreducible, then $M[\mathfrak{m}] = 0$, and $M/\mathfrak{m}M = 0$.*

Proof. First, we observe that $M/\mathfrak{m}M$ is also Eisenstein, and it is equal to $(M/\mathfrak{m}M)[\mathfrak{m}]$. Hence it suffices to prove the vanishing of $M[\mathfrak{m}]$ for all Eisenstein M.

Suppose $M[\mathfrak{m}]$ contains a non-zero element f. Then the annihilator of f in $\mathbb{T}(N)$ must be $\mathfrak{m}$. On the other hand, for almost all primes q, $T_q - q - 1$ lies in this annihilator. Hence $T_q = q+1$ in $k_{\mathfrak{m}}$. Therefore for almost all q, $\mathrm{Tr}(\rho_{\mathfrak{m}}(\mathrm{Frob}_q)) = q+1$ by the characterizing property of $\rho_{\mathfrak{m}}$. On the other hand, the representation $\rho = \chi_\ell \oplus 1$ of $G_{\mathbb{Q}}$, where χ_ℓ is the ℓ-th cyclotomoic character $G_{\mathbb{Q}} \to \mathbb{F}_\ell^\times$ and 1 is the trivial character $G_{\mathbb{Q}} \to \{1\} \subset \mathbb{F}_\ell^\times$, also satisfies that $\mathrm{Tr}(\rho(\mathrm{Frob}_q)) = q+1$ for almost all q. Hence $\rho_{\mathfrak{m}}$ is isomorphic to (the base change to $k_{\mathfrak{m}}$ of) $\chi_\ell \oplus 1$ by Lemma 42.1.1, and is therefore reducible. $\square$

The following lemma will be the "source" of all level lowering phenomena.

Lemma 42.1.4. *Let N_1, N_2 be two positive integers. Assume that there is an abelian group Y and ring maps $\alpha_i : \mathbb{T}(N_i) \to \mathrm{End}(Y)$ for $i = 1, 2$, and assume that $\alpha_1(T_q) = \alpha_2(T_q)$ for almost all primes q. Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}(N_2)$ of residue characteristic $\ell > 2$ such that either $Y/\alpha_2(\mathfrak{m}) \cdot Y \neq 0$ or $Y[\alpha_2(\mathfrak{m})] \neq 0$. Then $\rho_{\mathfrak{m}}$ is modular of level N_1.*

Proof. Let S be the set of primes q such that $\alpha_1(T_q) = \alpha_2(T_q)$. Let $\mathbb{T}'$ be the subring of $\alpha_2(\mathbb{T}(N_2))$ generated by $\alpha_2(T_q)$ for $q \in S$. Our hypothesis implies that $\alpha_2(\mathfrak{m})$ is a maximal ideal of $\alpha_2(\mathbb{T}(N_2))$. Let $\mathfrak{m}' = \alpha_2(\mathfrak{m}) \cap \mathbb{T}'$. Then $\mathfrak{m}'$ is a maximal ideal of $\mathbb{T}'$ since $\mathbb{T}' \subset \alpha_2(\mathbb{T}(N_2))$ is a finite ring extension. Also, we have $\mathbb{T}' \subset \alpha_1(\mathbb{T}(N_1))$, and this is again a finite ring extension. By going-up, we can find a maximal ideal $\mathfrak{n}'$ of $\alpha_1(\mathbb{T}(N_1))$ containing $\mathfrak{m}'$. Let $\mathfrak{n}$ be the inverse image of $\mathfrak{n}'$ in $\mathbb{T}(N_1)$. Let k denote the field $\mathbb{T}'/\mathfrak{m}'$. Then we have constructed field maps $f : k \to k_{\mathfrak{m}}$ and $g : k \to k_{\mathfrak{n}}$ such that for almost all primes q we have elements $U_q \in k$ with $f(U_q) = T_q \in k_{\mathfrak{m}}$ and $g(U_q) = T_q \in k_{\mathfrak{n}}$. Since $k, k_{\mathfrak{m}}, k_{\mathfrak{n}}$ are finite fields of characteristic ℓ, we can obviously embed all of them into $\overline{\mathbb{F}}_\ell$ in such a way that f and g become inclusion maps. Then if we base change both $\rho_{\mathfrak{m}}$ and $\rho_{\mathfrak{n}}$ to $\overline{\mathbb{F}}_\ell$, their characters agree on T_q for almost all primes q, and so they are isomorphic by Lemma 42.1.1. Thus $\mathfrak{m}$ is modular of level N_1. $\square$

Remark. Typically we will apply Lemma 42.1.4 to situations where $N_1 | N_2$, but this condition itself is not necessary for the lemma.

Lemma 42.1.5. *Let $N \geq 1$ and $p \| N$. Let A be the abelian variety part of $J_0(N)_{\mathbb{F}_p}$. Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}(N)$ of residue characteristic $\ell > 2$. If $A(\overline{\mathbb{F}}_p)[\mathfrak{m}] \neq 0$, then $\rho_{\mathfrak{m}}$ is modular of level N/p.*

Proof. Apply Lemma 42.1.4 to the natural action of $\mathbb{T}(N)$ on $A(\overline{\mathbb{F}}_p)$ and the action of $\mathbb{T}(N/p)$ on $A(\overline{\mathbb{F}}_p)$ via the identification $A \cong J_0(N/p)^2_{\mathbb{F}_p}$. $\square$

Theorem 42.1.6 (Mazur)**.** *Let $N \geq 1$, and $p \| N$. Let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}(N)$ of residue characteristic $\ell > 2$, and assume that $\rho_{\mathfrak{m}}$ is irreducible and finite at p. Assume that $p \not\equiv 1 \mod \ell$. Then $\rho_{\mathfrak{m}}$ is modular of level N/p.*

Proof. For simplicity, we only prove the theorem assuming that $p \neq \ell$. (However, recall that for the application to Fermat's Last theorem, we precisely need the case $p = \ell$; the idea of proof in this case is nevertheless similar to the case $p \neq \ell$, except that it involves more technical discussion about group schemes.) Let V be the vector space of $\rho_{\mathfrak{m}}$. By Theorem 41.1.1, we can choose a $k_{\mathfrak{m}}[G_{\mathbb{Q}}]$-module embedding $\iota' : V \hookrightarrow J_0(N)(\overline{\mathbb{Q}})[\mathfrak{m}]$. Fix an embedding $\overline{\mathbb{Q}}_p \hookrightarrow \overline{\mathbb{Q}}$ and use it to define the decomposition group $D_p \cong \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p) \subset G_{\mathbb{Q}}$. Since $\rho_{\mathfrak{m}}$ is finite at p, we can find a $k_{\mathfrak{m}}$-vector space scheme $\mathcal{V}$ over $\mathbb{Z}_p$ which is flat over $\mathbb{Z}_p$ and such that $\mathcal{V}(\overline{\mathbb{Q}}_p)$ gives rise to $\rho_{\mathfrak{m}}|_{D_p}$. Now ι' determines a $\mathbb{T}(N)$-equivariant morphism of $\mathbb{Q}_p$-schemes $\mathcal{V}_{\mathbb{Q}_p} \to J_0(N)_{\mathbb{Q}_p}$. Since $p \neq \ell$, $\mathcal{V}$ is in fact étale over $\mathbb{Z}_p$, and in particular smooth over $\mathbb{Z}_p$. Hence by the Néron extension property, we get a $\mathbb{T}(N)$-equivariant morphism of $\mathbb{F}_p$-schemes $\mathcal{V}_{\mathbb{F}_p} \to J_0(N)_{\mathbb{F}_p}$, and in particular a $k_{\mathfrak{m}}[D_p]$-module morphism

$$\iota : V \longrightarrow J_0(N)_{\mathbb{F}_p}(\overline{\mathbb{F}}_p)[\mathfrak{m}],$$

where D_p acts on the right via the quotient $\mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$. It is easy to check that ι is injective, by the injectivity of ι'.

We now assume that $\rho_{\mathfrak{m}}$ is not modular of level N/p. Let Φ, A, T be the component group, the abelian variety part, and the torus part of $J_0(N)_{\mathbb{F}_p}$. Then by the fact that Φ is Eisenstein and by Lemmas 42.1.3 and 42.1.5, we know know that $\mathrm{im}(\iota)$ is contained in $T(\overline{\mathbb{F}}_p)[\mathfrak{m}]$. Now we identify $T(\overline{\mathbb{F}}_p)[\mathfrak{m}]$ with

$$\mathrm{Hom}(\widehat{T}/\mathfrak{m}\widehat{T}, (\overline{\mathbb{F}}_p^{\times})[\ell]) = \mathrm{Hom}(\widehat{T}/\mathfrak{m}\widehat{T}, \mu_\ell(\overline{\mathbb{F}}_p)).$$

On $\widehat{T}/\mathfrak{m}\widehat{T}$, Frob_p acts by order at most 2, and its action also agrees with T_p, which acts via scalar multiplication by $(T_p \mod \mathfrak{m}) \in k_{\mathfrak{m}}$. Since $k_{\mathfrak{m}}$ is a field, we conclude that Frob_p acts on $\widehat{T}/\mathfrak{m}\widehat{T}$ by scalar multiplication by 1 or -1. On the other hand, Frob_p acts on $\mu_\ell(\overline{\mathbb{F}}_p)$ via multiplication (if the abelian group $\mu_\ell(\overline{\mathbb{F}}_p)$ is written additively) by p. Since ι is injective and D_p-equivariant, we conclude that $\rho_{\mathfrak{m}}(\mathrm{Frob}_p)$ is conjugate in $\mathrm{GL}_2(k_{\mathfrak{m}})$ to either one of $\pm \begin{pmatrix} p & \\ & p \end{pmatrix}$. In particular, $\det(\rho_{\mathfrak{m}}(\mathrm{Frob}_p)) = p^2$. As we have observed before, $\det(\rho_{\mathfrak{m}})$ is in fact the ℓ-th cyclotomic character on $G_{\mathbb{Q}}$. Since $p \neq \ell$, we have $\det(\rho_{\mathfrak{m}}(\mathrm{Frob}_p)) = p$. Thus $p = p^2$ in $k_{\mathfrak{m}}$, i.e., $p \equiv 1 \mod \ell$. □

Combining the analysis of $T(\overline{\mathbb{F}}_p)[\mathfrak{m}]$ in the above proof with the last statement of Theorem 41.1.1, one can prove the following result.

Lemma 42.1.7. *Let $p \| N$, and let $\mathfrak{m}$ be a maximal ideal of $\mathbb{T}(N)$ of residue characteristic $\ell \nmid 2N$. Suppose $\rho_{\mathfrak{m}}$ is irreducible. Let T be the torus part of $J_0(N)_{\mathbb{F}_p}$. If $\widehat{T}/\mathfrak{m}\widehat{T}$ has $k_{\mathfrak{m}}$-dimension at least 2, then $p \equiv 1 \mod \ell$.*

Proof. By a general fact about Néron models, there is an injective map $T(\overline{\mathbb{F}}_p) \to J_0(N)(\overline{\mathbb{Q}}_p)$ which is $\mathbb{T}(N)$-equivariant and D_p-equivariant. (It is a section of the natural reduction map $J_0(N)(\overline{\mathbb{Q}}_p) \to J_0(N)_{\mathbb{F}_p}(\overline{\mathbb{F}}_p)$ coming from the Néron extension property, and hence $\mathbb{T}(N)$-equivariant.) It follows that we have an injective map of $k_{\mathfrak{m}}(D_p)$-modules $T(\overline{\mathbb{F}}_p)[\mathfrak{m}] \to J(\overline{\mathbb{Q}}_p)[\mathfrak{m}] \cong J(\overline{\mathbb{Q}})[\mathfrak{m}]|_{D_p}$. By the last statement of Theorem 41.1.1, the right hand side is isomorphic to $\rho_{\mathfrak{m}}|_{D_p}$ (with multiplicity 1). Using $T(\overline{\mathbb{F}}_p)[\mathfrak{m}] \cong \mathrm{Hom}(\widehat{T}/\mathfrak{m}\widehat{T}, \mu_\ell(\overline{\mathbb{F}}_p))$, our hypothesis implies that $T(\overline{\mathbb{F}}_p)[\mathfrak{m}]$ has $k_{\mathfrak{m}}$-dimension at least 2. Since $\rho_{\mathfrak{m}}$ has $k_{\mathfrak{m}}$-dimension

2, we conclude that $T(\overline{\mathbb{F}}_p)[\mathfrak{m}] \cong \rho_{\mathfrak{m}}|_{D_p}$ as $k_{\mathfrak{m}}[D_p]$-modules. Now as showed in the proof of Theorem 42.1.6, the determinant of Frob_p acting on $T(\overline{\mathbb{F}}_p)[\mathfrak{m}]$ is $p^2 \in k_{\mathfrak{m}}$. On the other hand the determinant of Frob_p acting on $\rho_{\mathfrak{m}}$ is $p \in k_{\mathfrak{m}}$ (since $\det \rho_{\mathfrak{m}} = \chi_\ell$). This concludes the proof. □

42.2. **Shimura curves.** To prove Ribet's theorem, we follow Ribet to study certain reductions of Shimura curves. These curves are certain one-dimensional Shimura varieties different from modular curves and attached to a quaternion algebra over $\mathbb{Q}$ ramified at two finite places p and q. We will study the reduction modulo q of the Jacobian of the Shimura curve, and in a miraculous way relate it to the reduction of the Jacobian of modular curves modulo the *different prime* p!

To start, we always fix two distinct primes p, q, and fix a positive integer M coprime to pq. Let D be the unique (up to isomorphism) quaternion algebra over $\mathbb{Q}$ ramified exactly at p and q. Thus for each place v of $\mathbb{Q}$, $D_v := D \otimes_{\mathbb{Q}} \mathbb{Q}_v$ is isomorphic to $M_2(\mathbb{Q}_v)$ as a $\mathbb{Q}_v$-algebra if $v \notin \{p, q\}$, and D_v is a division algebra over $\mathbb{Q}_v$ for $v \in \{p, q\}$. Fix an order $\mathcal{O}_D \subset D$ (i.e., $\mathcal{O}_D$ is a subring of D and a finite free $\mathbb{Z}$-module such that it contains a $\mathbb{Q}$-basis of D) such that for every finite place v of $\mathbb{Q}$, $\mathcal{O}_D \otimes_{\mathbb{Z}} \mathbb{Z}_v$ is a maximal $\mathbb{Z}_v$-order in D_v (i.e., maximal among subrings of D_v that are finite free over $\mathbb{Z}_v$ and contain a $\mathbb{Q}_v$-basis of D_v) if $v \nmid M$, and is the order

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z}_v, c \equiv 0 \pmod{v} \right\}$$

in $D_v \cong M_2(\mathbb{Q}_v)$ if $v|M$. We can then form the complex Shimura curve:

$$X = X_{D,M} := D^\times \backslash \mathbb{H}^\pm \times (D \otimes_{\mathbb{Q}} \mathbb{A}_f)^\times / \prod_{v<\infty} (\mathcal{O}_D \otimes_{\mathbb{Z}} \mathbb{Z}_v)^\times.$$

Here $D^\times$ acts diagonally on the two factors, and $\prod_{v<\infty} (\mathcal{O}_D \otimes_{\mathbb{Z}} \mathbb{Z}_v)^\times$ acts only on the second factor. The action of $D^\times$ on $\mathbb{H}^\pm$ is inherited from the usual $\mathrm{GL}_2(\mathbb{R})$-action via $D^\times \hookrightarrow D_\infty^\times$ and a fixed isomorphism $D_\infty \cong M_2(\mathbb{R})$. One can show that X is actually connected (which is not the case in general for the double quotients in the general definition of Shimura varieties), and is actually isomorphic to

$$\Gamma \backslash \mathbb{H}^+,$$

where $\Gamma = \mathcal{O}_D^\times \cap \mathrm{SL}_2(\mathbb{R})$, a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$; the last intersection is inside $D_\infty^\times$ which has been identified with $\mathrm{GL}_2(\mathbb{R})$.[14] Unlike $Y_0(N) = \Gamma_0(N) \backslash \mathbb{H}^+$ which is a non-compact Riemann surface, X is a compact Riemann surface. Moreover, X has a canonical model over $\mathbb{Q}$, which is a smooth projective curve over $\mathbb{Q}$ and is the coarse moduli space of

[14] Similarly, we can present $Y_0(N) = \Gamma_0(N) \backslash \mathbb{H}^+$ (at the level of complex points) as a double quotient as follows. Take B to be the $\mathbb{Q}$-algebra $M_2(\mathbb{Q})$, and take $\mathcal{O}_B \subset B$ to be the order consisting of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $a, b, c, d \in \mathbb{Z}$ and $N|c$. Then $Y_0(N)$ is isomorphic to $B^\times \backslash \mathbb{H}^\pm \times (B \otimes_{\mathbb{Q}} \mathbb{A}_f)^\times / \prod_{v<\infty} (\mathcal{O}_B \otimes_{\mathbb{Z}} \mathbb{Z}_v)^\times$. (Here the product group on the right is also $K_0(N) := \{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \mid c \equiv 0 \mod N \}$.) To prove that both double quotients with D and B are actually connected, one uses strong approximation, the fact that the derived subgroups of $D^\times$ and $B^\times$ (as reductive groups over $\mathbb{Q}$) are simply connected, the fact that the abelianization maps $D^\times \to \mathbb{G}_m$ and $B^\times \to \mathbb{G}_m$ map $(\mathcal{O}_D \otimes \mathbb{Z}_v)^\times$ and $(\mathcal{O}_B \otimes \mathbb{Z}_v)^\times$ respectively onto $\mathbb{Z}_v^\times$ for all finite v, and the fact that $\mathbb{Q}_{>0}^\times \backslash \mathbb{A}_f^\times / \widehat{\mathbb{Z}}^\times = \{1\}$, i.e., $\mathbb{Z}$ has trivial narrow class group. See [Mil17b, Thm. 5.17] for how to compute connected components of such double quotients in general when the derived subgroup is simply connected.

abelian surfaces A together with an injective ring map $\mathcal{O}_D \to \mathrm{End}(A)$ and a certain "level-M structure". This coarse moduli space can be constructed from $(\mathscr{A}_{g,d,N})_{\mathbb{Q}}$ for $g = 2, d = 1$, and large enough N by taking a finite quotient of a suitable subscheme. For more details, see [BC91].

From now on, we write X for the canonical model of the Shimura curve over $\mathbb{Q}$. Just as for modular curves, for each integer $n \geq 1$, we have a Hecke correspondence $X \dashrightarrow X$, and by functoriality this gives rise to an endomorphism T_n of the Jacobian $J(X) := \mathrm{Jac}(X)$ (which is again an abelian variety over $\mathbb{Q}$). We define the **Hecke algebra for** X to be the subring $\mathbb{T}(X)$ of $\mathrm{End}_{\mathbb{Q}}(J(X))$ generated by the T_n's. We will be mainly interested in the reduction modulo q of (the Néron model over $\mathbb{Z}_q$ of) $J(X)$, denoted by $J(X)_{\mathbb{F}_q}$. This is a smooth commutative group scheme over $\mathbb{F}_q$ equipped with a $\mathbb{T}(X)$-action. We will see that this is actually related to $J_0(Mpq)_{\mathbb{F}_p}$ and $J_0(Mp)_{\mathbb{F}_p}$ in an extremely interesting way.

43. Lecture 43

43.1. **Reduction of the Shimura curve.** Let p, q be distinct primes, M be a positive integer coprime to pq, and $X = X_{D_{p,q},M}$ be the the (canonical model of) Shimura curve attached to (M, p, q). Thus X is a smooth projective curve over $\mathbb{Q}$. Recall that we have the Hecke algebra $\mathbb{T}(X)$ acting on $J(X) = \mathrm{Jac}(X)$. We are interested in describing $J(X)_{\mathbb{F}_q}$ (i.e., the reduction of the Néron model of $J(X)_{\mathbb{Q}_q}$ over $\mathbb{Z}_q$ modulo q) together with the $\mathbb{T}(X)$-action.

Similar to the case of the reduction of the modular curve at a prime dividing the level precisely once, we have a weakly semi-stable integral model $\mathfrak{X}$ over $\mathbb{Z}_q$ of the Shimura curve $X_{\mathbb{Q}_q}$, constructed and studied by Čerednik and Drinfeld. Recall that weakly semi-stable means that $\mathfrak{X} \to \mathrm{Spec}\, \mathbb{Z}_q$ is smooth away from finitely singularities which are nodes of $\mathfrak{X}_{\mathbb{F}_q}$. More specifically, $\mathfrak{X}_{\mathbb{F}_q}$ is the union of two families of parallel $\mathbb{P}^1$'s such that whenever a $\mathbb{P}^1$ in the first family intersects with a $\mathbb{P}^1$ in the second family the intersection is transverse. Let $\mathcal{V}_1$ (resp. $\mathcal{V}_2$) be the set of $\mathbb{P}^1$'s in the first (resp. second) family, and let $\mathcal{E}$ be the set of intersection points. Then the intersection configuration is described by the maps $\gamma_1 : \mathcal{E} \to \mathcal{V}_1, \gamma_2 : \mathcal{E} \to \mathcal{V}_2$ defined in the obvious way. (Equivalently, the dual graph of $\mathfrak{X}_{\mathbb{F}_q}$ has $\mathcal{E}$ as the set of edges, $\mathcal{V}_1 \sqcup \mathcal{V}_2$ as the set of vertices, and (γ_1, γ_2) as the boundary map from edges to vertices.) The datum $(\mathcal{E}, \mathcal{V}_1, \mathcal{V}_2, \gamma_1, \gamma_2)$ has the following concrete description. We introduce a notation. For any integer $K \geq 1$ and any prime $v \nmid K$, we write $\Sigma(K, v)$ for the set of $\overline{\mathbb{F}}_v$-points of the supersingular locus of $\mathfrak{X}_0(K)_{\mathbb{Z}_v}$ (the "good" canonical integral model of $X_0(K)_{\mathbb{Q}_v}$ over $\mathbb{Z}_v$).

Then we have natural identifications $\mathcal{E} \cong \Sigma(Mq, p)$ and $\mathcal{V}_1 \cong \mathcal{V}_2 \cong \Sigma(M, p)$. Moreover, the maps $\gamma_1, \gamma_2 : \Sigma(Mq, p) \to \Sigma(M, p)$ come from two maps $X_0(Mq) \to X_0(M)$ (definable over characteristic zero) that can be explicitly described both in terms of the moduli interpretation and in terms of the complex coordinates in $\mathbb{H}^+$. If τ is the coordinate in $\mathbb{H}^+$, then the two maps are induced by $\tau \mapsto \tau$ and $\tau \mapsto q\tau$.

Moreover, the exponent $e(x)$ in the local equation $ST - q^{e(x)}$ for $\mathfrak{X}$ near each $x \in \mathcal{E} \cong \Sigma(Mq, p)$ is the same as the exponent in the local equation $ST - p^{e(x)}$ for $\mathfrak{X}_0(Mpq)_{\mathbb{Z}_p}$ near x. (Here recall that the nodes in $\mathfrak{X}_0(Mpq)_{\mathbb{F}_p}$ are also parametrized by $\Sigma(Mq, p)$.)

Using the above information about $\mathfrak{X}$ over $\mathbb{Z}_q$, Ribet proves the following statements.

(1) The group scheme $J(X)_{\mathbb{F}_q}$ is an extension of a finite abelian constant group scheme $\Phi_{X,q}$ by a torus $T_{X,q}$ over $\mathbb{F}_q$. In general, if A is an abelian variety over $\mathbb{Q}_q$ with Néron model $\mathcal{A}$ over $\mathbb{Z}_q$ such that the identity connected component $\mathcal{A}^{\circ}_{\mathbb{F}_q}$ of $\mathcal{A}_{\mathbb{F}_q}$ is an extension of an abelian variety by a maximal torus (i.e., A has semi-stable reduction modulo q), then $\mathrm{End}_{\mathbb{Q}_q}(A)$ injects into $\mathrm{End}_{\mathbb{F}_q}(\mathcal{A}^\circ_{\mathbb{F}_q})$. Thus we know that $\mathbb{T}(X)$ acts faithfully on $T_{X,q}$.

(2) We have a short exact sequence

$$0 \to \widehat{T}_{X,q} \xrightarrow{\alpha} \widehat{T}_{Mpq,p} \xrightarrow{\beta} \widehat{T}^2_{Mp,p} = \widehat{T}_{Mp,p} \oplus \widehat{T}_{Mp,p} \to 0 \tag{43.1}$$

coming from the description of the dual graph of $\mathfrak{X}_{\mathbb{F}_q}$. Here $\widehat{T}_{Mpq,p}$ denotes the character group of the torus part of $J_0(Mpq)_{\mathbb{F}_p}$, and similarly for $\widehat{T}_{Mp,p}$. Recall that $\widehat{T}_{Mpq,p}$ (resp. $\widehat{T}_{Mp,p}$) is the abelian group of formal integer linear combinations of elements of $\Sigma(Mq,p)$ (resp. $\Sigma(M,p)$) with the coefficients summing to zero. The map β is induced by the two maps $\gamma_1, \gamma_2 : \Sigma(Mq,p) \to \Sigma(M,p)$.

(3) The map α is equivariant for each T_n. (Here T_n acts on $\widehat{T}_{X,q}$ via $T_n \in \mathbb{T}(X)$ and acts on $\widehat{T}_{Mpq,p}$ via $T_n \in \mathbb{T}(Mpq)$; these two rings are a priori unrelated, but they both have elements indexed by an integer n denoted by T_n.) In particular, remembering that $\mathbb{T}(X)$ acts faithfully on $T_{X,q}$, we see that there is a well-defined surjective ring map $\mathbb{T}(Mpq) \twoheadrightarrow \mathbb{T}(X)$ sending each T_n to T_n. We shall use this map to view every $\mathbb{T}(X)$-module as a $\mathbb{T}(Mpq)$-module. Then α is $\mathbb{T}(Mpq)$-linear.

(4) Since α is a map of $\mathbb{T}(Mpq)$-modules, there is a unique structure of $\mathbb{T}(Mpq)$-module on $\widehat{T}^2_{Mp,p}$ such that β is $\mathbb{T}(Mpq)$-linear. This structure is given as follows. For each prime v, define $U_v \in M_2(\mathbb{T}(Mp))$ by

$$U_v = \begin{cases} \begin{pmatrix} T_v & \\ & T_v \end{pmatrix}, & v \neq q \\ \\ \begin{pmatrix} T_q & -1 \\ q & 0 \end{pmatrix}, & v = q. \end{cases}$$

Then $T_v \in \mathbb{T}(Mpq)$ acts on $\widehat{T}^2_{Mp,p}$ via the natural action of U_v on $\widehat{T}^2_{Mp,p}$. In the sequel, we will always view $\widehat{T}^2_{Mp,p}$ as a $\mathbb{T}(Mpq)$-module in this way.

(5) Let Ψ be the cokernel of the endomorphism of $\widehat{T}^2_{Mp,p}$ induced by $T_q^2 - 1 \in \mathbb{T}(Mpq)$ (i.e., the endomorphism $U_q^2 - 1$). Then we have an exact sequence

$$0 \to A \to \Psi \to \Phi_{X,q} \to B \to 0 \tag{43.2}$$

of $\mathbb{T}(Mpq)$-modules where A and B are Eisenstein.

43.2. **Lemmas for proving Ribet's theorem.** From now on, we fix an odd prime ℓ, and all maximal ideals of all Hecke algebras are assumed to be of residue characteristic ℓ. Let M, p, q be as in §43.1.

Lemma 43.2.1. *Let $\mathfrak{m} \subset \mathbb{T}(Mpq)$ be a maximal ideal. Assume that $\widehat{T}^2_{Mp,p}/\mathfrak{m}\widehat{T}^2_{Mp,p} \neq 0$. Then $\rho_{\mathfrak{m}}$ is modular of level Mp.*

Proof. Apply Lemma 42.1.4 to the actions of $\mathbb{T}(Mpq)$ and $\mathbb{T}(Mp)$ on $\widehat{T}^2_{Mp,p}$. Here for each prime v, $T_v \in \mathbb{T}(Mpq)$ acts by U_v (as always), and $T_v \in \mathbb{T}(Mp)$ acts by $\begin{pmatrix} T_v & \\ & T_v \end{pmatrix}$. These two actions satisfy the hypothesis of Lemma 42.1.4 by the discussion in **(4)** in §43.1. $\square$

Lemma 43.2.2. *Let $\mathfrak{m} \subset \mathbb{T}(Mpq)$ be a maximal ideal such that $\rho_{\mathfrak{m}}$ is not modular of level Mp. Then $\widehat{T}_{X,q}/\mathfrak{m}\widehat{T}_{X,q}$ and $\widehat{T}_{Mpq,p}/\mathfrak{m}\widehat{T}_{Mpq,p}$ have the same $k_{\mathfrak{m}}$-dimension.*

Proof. By Lemma 43.2.1, we have $\widehat{T}^2_{Mp,p}/\mathfrak{m}\widehat{T}^2_{Mp,p} = 0$. Since $\widehat{T}^2_{Mp,p}$ is a finite $\mathbb{Z}$-module and hence a finite $\mathbb{T}(Mpq)$-module, by Nakayama we have $(\widehat{T}^2_{Mp,p})_{\mathfrak{m}} = 0$. Localizing the short exact sequence (43.1) of $\mathbb{T}(Mpq)$-modules at $\mathfrak{m}$ and tensoring to $k_{\mathfrak{m}}$ gives the result. □

Lemma 43.2.3. *Let X be the Shimura curve attached to (M, p, q). Let $\mathfrak{m} \subset \mathbb{T}(Mpq)$ be a maximal ideal such that $\rho_{\mathfrak{m}}$ is irreducible. Then the following are equivalent.*

(1) $\Phi_{X,q}/\mathfrak{m}\Phi_{X,q} \neq 0$.
(2) $\widehat{T}^2_{Mp,p}/\mathfrak{m}\widehat{T}^2_{Mp,p} \neq 0$ *and* $T_q^2 - 1 \in \mathfrak{m}$.

Proof. Let the notation be as in (43.2). Since A is Eisenstein and $\rho_{\mathfrak{m}}$ is irreducible, by Lemma 42.1.3 we have $A/\mathfrak{m}A = 0$. Since A is finite over $\mathbb{Z}$, by Nakayama we have $A_{\mathfrak{m}} = 0$. Similarly, we have $B_{\mathfrak{m}} = 0$. Thus the exact sequence (43.2) implies that $\Psi_{\mathfrak{m}} \cong (\Phi_{X,q})_{\mathfrak{m}}$ as $\mathbb{T}(Mpq)_{\mathfrak{m}}$-modules. In particular,

$$\Psi/\mathfrak{m}\Psi \cong \Phi_{X,q}/\mathfrak{m}\Phi_{X,q}$$

as $k_{\mathfrak{m}}$-vector spaces. Now if (2) holds, then it immediately follows from the definition of Ψ that $\Psi/\mathfrak{m}\Psi = \widehat{T}^2_{Mp,p}/\mathfrak{m}\widehat{T}^2_{Mp,p} \neq 0$. Hence (1) holds. Conversely, if (1) holds, then $\Psi/\mathfrak{m}\Psi \neq 0$. But the annihilator of $\Psi/\mathfrak{m}\Psi$ in $\mathbb{T}(Mpq)$ contains $\mathfrak{m}$ and contains $T_q^2 - 1$. Hence $T_q^2 - 1 \in \mathfrak{m}$. Given this, we also have $\Psi/\mathfrak{m}\Psi = \widehat{T}^2_{Mp,p}/\mathfrak{m}\widehat{T}^2_{Mp,p}$. Hence (2) holds. □

Finally, a general lemma on abelian varieties.

Lemma 43.2.4. *Let A be an abelian variety over an algebraically closed field k. Let S be a commutative subring of* $\mathrm{End}(A)$ *and let I be a maximal ideal of S of residue characteristic ℓ. (Thus $\ell > 0$ since S is of finite rank over $\mathbb{Z}$.) Assume that $\ell \neq \mathrm{char}\, k$. Then $A(k)[I] \neq 0$.*

Proof. Consider the ℓ-adic Tate module $T = T_\ell(A)$, which is a finite free $\mathbb{Z}_\ell$-module with a natural action by $S \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. We shall use two general facts: First, $T/\ell T \cong A(k)[\ell]$. Second, the action of $S \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$ on T is faithful.

Write R for $S \otimes_{\mathbb{Z}} \mathbb{Z}_\ell$. We can identify R with the ℓ-adic completion of S, and as such the ℓ-adic completion J of I is a maximal ideal of R (with residue field S/I) containing the image of I under $S \to R$. It suffices to show that the R-module $T/\ell T$ satisfies $(T/\ell T)[J] \neq 0$. We now give two proofs of this.

First proof. Since R is of finite rank over the complete local ring $\mathbb{Z}_\ell$, we know that $R \cong \prod_{i=1}^n R_i$, where each R_i is a complete local ring. (See http://www.math.lsa.umich.edu/~hochster/615W14/ModFinComp.pdf) Now using the idempotents in R, we can decompose the R-module T into $\bigoplus_{i=1}^n T_i$ such that R acts on each T_i via the projection $R \to R_i$ and a certain R_i-action. Since the R-action on T is faithful, we have $T_i \neq 0$ for each i. Without loss of generality we may assume that $J = \mathfrak{m}_1 \times R_2 \times \cdots \times R_n$, where $\mathfrak{m}_1$ is the maximal ideal of R_1. Then it suffices to show that the R_1-module $M := T_1/\ell T_1$ satisfies $M[\mathfrak{m}_1] \neq 0$. Since T_1 is a non-zero finite free $\mathbb{Z}_\ell$-module, we have $M \neq 0$. Also M is a finite abelian group. Hence by Nakayama we have $M \supsetneq \mathfrak{m}_1 M \supsetneq \cdots \supsetneq \mathfrak{m}_1^r M = 0$ for some $r \geq 1$. But then $0 \neq \mathfrak{m}_1^{r-1} M \subset M[\mathfrak{m}_1]$.

Second proof. We claim that $J \in \mathrm{supp}_R(T/\ell T)$ (i.e., the localization $(T/\ell T)_J \neq 0$). Indeed, if not, then $T/\ell T = J(T/\ell T)$, and so $T = JT$. Since T is finite over $\mathbb{Z}_\ell$ and *a fortiori* finite over R, we have, by Nakayama, an element $\alpha \in 1 + J \subset R - \{0\}$ such that $\alpha T = 0$. This contradicts with the fact that R acts faithfully on T. Now since $T/\ell T$ is a finite-length module (being a finite abelian group) over the noetherian ring R, $\mathrm{supp}_R(T/\ell T)$ is equal to the set of associated primes, i.e., the prime ideals of R of the form $\mathrm{Ann}(x)$ for some $x \in T/\ell T$. (See Cor. 1.6.10 of https://faculty.math.illinois.edu/~r-ash/ComAlg/ComAlg1.pdf, or see [Eis95, Cor. 2.17, Thm. 3.1a].) Hence $J = \mathrm{Ann}(x)$ for some $x \in T/\ell T$, which shows that $(T/\ell T)[J] \neq 0$. □

43.3. **Proof of Ribet's theorem.**

Theorem 43.3.1 (Ribet)**.** *Let $N \geq 1$, and let p be a prime with $p \| N$. Let $\mathfrak{m}_0 \subset \mathbb{T}(N)$ be a maximal ideal of residue characteristic $\ell \nmid 2N$. Suppose $\rho_{\mathfrak{m}_0}$ is irreducible and finite at p (equivalently, unramified at p). Then $\rho_{\mathfrak{m}_0}$ is modular of level N/p.*

Proof. We write M for N/p. Suppose $\rho_{\mathfrak{m}_0}$ is not modular of level M. Then by the same argument as in the proof of Theorem 42.1.6, we have an injection of $k_{\mathfrak{m}_0}[D_p]$-modules $\rho_{\mathfrak{m}_0}|_{D_p} \hookrightarrow T_{Mp,p}(\overline{\mathbb{F}}_p)[\mathfrak{m}_0]$. It follows that

$$\widehat{T}_{Mp,p}/\mathfrak{m}_0\widehat{T}_{Mp,p} \neq 0. \tag{43.3}$$

Recall that $\rho_{\mathfrak{m}_0}(c)$, where c is any complex conjugation in $G_{\mathbb{Q}}$, is conjugate in $\mathrm{GL}_2(k_{\mathfrak{m}_0})$ to $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$. By Chebotarev density we can find a prime $q \nmid N\ell$ such that $\rho_{\mathfrak{m}_0}$ is unramified at q and $\rho_{\mathfrak{m}_0}(\mathrm{Frob}_q)$ is conjugate to $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$. In particular, we have

$$T_q \in \mathfrak{m}_0, \tag{43.4}$$

$$q \equiv -1 \mod \ell \tag{43.5}$$

by looking at the trace and determinant of $\rho_{\mathfrak{m}_0}(\mathrm{Frob}_q)$. We now form the Shimura curve X using (M, p, q). By Cayley–Hamilton, we know that $U_q \in M_2(\mathbb{T}(Mp))$ (notation as in **(4)** in §43.1) satisfies that $U_q^2 - T_qU_q + q = 0$ (where $T_q \in \mathbb{T}(Mp)$). Reducing this modulo $\mathfrak{m}_0$ and using (43.4) and (43.5), we get

$$U_q^2 - 1 \equiv \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \mod \mathfrak{m}_0. \tag{43.6}$$

Let R be the image of $\mathbb{T}(Mpq)$ in $\mathrm{End}(\widehat{T}^2_{Mp,p})$, let R_1 be the subring of $\mathrm{End}(\widehat{T}^2_{Mp,p})$ generated by the images of $T_v \in \mathbb{T}(Mp)$ for primes $v \neq q$, and let R_2 be the image of $\mathbb{T}(Mp)$ in $\mathrm{End}(\widehat{T}^2_{Mp,p})$. Then we have finite ring extensions $R_1 \subset R$ and $R_1 \subset R_2$. By (43.3), $\mathrm{im}(\mathfrak{m}_0 \to R_2)$ is a maximal ideal of R_2. Then $\mathfrak{m}_1 = R_1 \cap \mathrm{im}(\mathfrak{m}_0 \to R_2)$ is a maximal ideal of R_1. By going-up, we can find a maximal ideal $\mathfrak{m}'$ of R containing $\mathfrak{m}_1$. Let $\mathfrak{m}$ be the inverse image of $\mathfrak{m}'$ in $\mathbb{T}(Mpq)$. By (43.6), the image of $T_q^2 - 1 \in \mathbb{T}(Mpq)$ in R is zero. Hence $T_q^2 - 1 \in \mathfrak{m}$. Clearly we have field maps $R_1/\mathfrak{m}_1 \to k_{\mathfrak{m}}$ and $R_1/\mathfrak{m}_1 \to k_{\mathfrak{m}_0}$ such that for almost all primes v, $T_v \in k_{\mathfrak{m}}$ and $T_v \in k_{\mathfrak{m}_0}$ come from a common element of $R_1/\mathfrak{m}_1$. By the same argument as in Lemma 42.1.4, we conclude that $\rho_{\mathfrak{m}} \cong \rho_{\mathfrak{m}_0}$ (after some common extension of scalars). It remains to show that $\rho_{\mathfrak{m}}$ is modular of level M.

For this, it suffices to show that $\rho_{\mathfrak{m}}$ is modular of level Mq, since then we can apply Mazur's theorem to $q \| Mq$ (in view of (43.5)) to conclude. In the following we assume $\rho_{\mathfrak{m}}$ is not modular of level Mq, and deduce a contradiction.

By (43.3) and the definition of $\mathfrak{m}_1$, we have $\widehat{T}^2_{Mp,p}/\mathfrak{m}_1\widehat{T}^2_{Mp,p} \neq 0$. But $\widehat{T}^2_{Mp,p}/\mathfrak{m}\widehat{T}^2_{Mp,p} = (\widehat{T}^2_{Mp,p}/\mathfrak{m}_1\widehat{T}^2_{Mp,p}) \otimes_{R_1/\mathfrak{m}_1} k_{\mathfrak{m}}$, so this is again non-zero. Recall that $T_q^2 - 1 \in \mathfrak{m}$. Hence by the (2) ⇒ (1) direction of Lemma 43.2.3, we conclude that $\Phi_{X,q}/\mathfrak{m}\Phi_{X,q} \neq 0$. In particular, the image of $\mathfrak{m}$ under $\mathbb{T}(Mpq) \twoheadrightarrow \mathbb{T}(X)$ is a maximal ideal. Thus $J(X)(\overline{\mathbb{Q}})[\mathfrak{m}] \neq 0$ by Lemma 43.2.4.

We claim that the semi-simplification of the non-zero $k_{\mathfrak{m}}[G_{\mathbb{Q}}]$-module $J(X)(\overline{\mathbb{Q}})[\mathfrak{m}]$ is $\rho_{\mathfrak{m}}^d$ for some $d > 0$. Indeed, we know from the work of Eichler and Shimura that for almost all primes v, the action of Frob_v on $J(X)_{\mathbb{F}_v}$ satisfies $\mathrm{Frob}_v^2 - T_v \mathrm{Frob}_v + v = 0$, where $T_v \in \mathbb{T}(X)$. Writing V for $J(X)(\overline{\mathbb{Q}})[\mathfrak{m}]$, we immediately deduce, using Chebotarev and Brauer–Nesbitt, that the semi-simplification of the $k_{\mathfrak{m}}[G_{\mathbb{Q}}]$-module $V \oplus \mathrm{Hom}_{k_{\mathfrak{m}}}(V, \mu_\ell)$ is isomorphic to $\rho_{\mathfrak{m}}^{2d}$ for some positive integer d. (The point of forming this direct sum is to make sure that the eigenvalues of Frob_v, for almost all v, come in pairs λ and v/λ.) Since $\rho_{\mathfrak{m}}$ is irreducible our claim follows.[15]

Hence we can choose an injective map of $k_{\mathfrak{m}}[G_{\mathbb{Q}}]$-modules $\iota' : \rho_{\mathfrak{m}} \hookrightarrow J(X)(\overline{\mathbb{Q}})[\mathfrak{m}]$. We now consider the reduction of $J(X)$ modulo p. (Up to now we have only considered the reduction of $J(X)$ modulo q!) Since $\rho_{\mathfrak{m}}$ is not modular of level Mq, by Lemma 43.2.1 (with p and q reversed), we have $\widehat{T}^2_{Mq,q}/\mathfrak{m}\widehat{T}^2_{Mq,q} = 0$. Then by Lemma 43.2.3 (with p and q reversed), we have $\Phi_{X,p}/\mathfrak{m}\Phi_{X,p} = 0$. Since $\Phi_{X,p}$ is a finite $\mathbb{Z}$-module (in fact a finite abelian group), by Nakayama we have $\alpha\Phi_{X,p} = 0$ for some $\alpha \in 1 + \mathfrak{m} \subset \mathbb{T}(Mpq)$. This implies immediately that $\Phi_{X,p}[\mathfrak{m}] = 0$. Now since $\rho_{\mathfrak{m}}$ is finite at p, as in the proof of Theorem 42.1.6 we know that ι' induces an injective map of $k_{\mathfrak{m}}[D_p]$-modules $\iota : \rho_{\mathfrak{m}} \hookrightarrow J(X)_{\mathbb{F}_p}(\overline{\mathbb{F}}_p)[\mathfrak{m}]$. Since $\Phi_{X,p}[\mathfrak{m}] = 0$, the right hand side is equal to $T_{X,p}(\overline{\mathbb{F}}_p)[\mathfrak{m}]$. Hence $\dim_{k_{\mathfrak{m}}}(T_{X,p}(\overline{\mathbb{F}}_p)[\mathfrak{m}]) = \dim_{k_{\mathfrak{m}}}(\widehat{T}_{X,p}/\mathfrak{m}\widehat{T}_{X,p}) \geq 2$. Since $\rho_{\mathfrak{m}}$ is not modular of level Mq, by Lemma 43.2.2 (with p and q reversed) we get $\dim_{k_{\mathfrak{m}}}(\widehat{T}_{Mpq,q}/\mathfrak{m}\widehat{T}_{Mpq,q}) \geq 2$. Now by Lemma 42.1.7 applied to $q \| Mpq$, we get $q \equiv 1 \mod \ell$. This contradicts with (43.5). □

References

[BC91] J.-F. Boutot and H. Carayol. Uniformisation p-adique des courbes de Shimura: les théorèmes de Cerednik et de Drinfeld. Number 196-197, pages 7, 45–158 (1992). 1991. Courbes modulaires et courbes de Shimura (Orsay, 1987/1988). 87

[Cona] Brian Conrad. Course notes for Math 248B Modular Curves. http://virtualmath1.stanford.edu/~conrad/248BPage/handouts/modularcurves.pdf. 5

[Conb] Brian Conrad. Handout on cohomology and base change. http://virtualmath1.stanford.edu/~conrad/248BPage/handouts/cohom.pdf. 27, 36, 55

[Conc] Brian Conrad. Handout on isogenies and level structures. http://virtualmath1.stanford.edu/~conrad/248BPage/handouts/level.pdf. 5, 25, 34

[Cond] Brian Conrad. Lecture notes on abelian varieties. http://virtualmath1.stanford.edu/~conrad/249CS15Page/handouts/abvarnotes.pdf. 5, 56

[Cone] Brian Conrad. Math 248B. Modular Curves. course website http://virtualmath1.stanford.edu/~conrad/248BPage/handouts.html. 5

[CSS97] Gary Cornell, Joseph H. Silverman, and Glenn Stevens, editors. *Modular forms and Fermat's last theorem.* Springer-Verlag, New York, 1997. Papers from the Instructional Conference on Number Theory and Arithmetic Geometry held at Boston University, Boston, MA, August 9–18, 1995. 5

[DDT97] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. In *Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993)*, pages 2–140. Int. Press, Cambridge, MA, 1997. 5, 77, 78

[Del71a] Pierre Deligne. Formes modulaires et représentations l-adiques. In *Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363*, volume 175 of *Lecture Notes in Math.*, pages Exp. No. 355, 139–172. Springer, Berlin, 1971. 75

[Del71b] Pierre Deligne. Travaux de Shimura. In *Séminaire Bourbaki, 23ème année (1970/71), Exp. No. 389*, pages 123–165. Lecture Notes in Math., Vol. 244. Springer, Berlin, 1971. 5

[Del79] Pierre Deligne. Variétés de Shimura: interprétation modulaire, et techniques de construction de modèles canoniques. In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, pages 247–289. Amer. Math. Soc., Providence, R.I., 1979. 5

[DS74] Pierre Deligne and Jean-Pierre Serre. Formes modulaires de poids 1. *Ann. Sci. École Norm. Sup. (4)*, 7:507–530 (1975), 1974. 75, 77

[15] Using the same argument and based on the Eichler–Shimura relation for modular curves, one can prove the first statement in Theorem 41.1.1 at least up to semi-simplification.

[DS05] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005. 5, 16, 17

[Edi97] Bas Edixhoven. Serre's conjecture. In *Modular forms and Fermat's last theorem (Boston, MA, 1995)*, pages 209–242. Springer, New York, 1997. 5, 74

[EH16] David Eisenbud and Joe Harris. *3264 and all that—a second course in algebraic geometry*. Cambridge University Press, Cambridge, 2016. 74

[Eis95] David Eisenbud. *Commutative algebra*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. With a view toward algebraic geometry. 90

[EvdGM] Bas Edixhoven, Gerard van der Geer, and Ben Moonen. Abelian varieties. preliminary version, available at `http://van-der-geer.nl/~gerard/AV.pdf`. 5, 55

[FC90] Gerd Faltings and Ching-Li Chai. *Degeneration of abelian varieties*, volume 22 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1990. With an appendix by David Mumford. 45

[Ful98] William Fulton. *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, second edition, 1998. 74

[GN06] A Genestier and BC Ngô. Lectures on shimura varieties, 2006. available at `https://www.math.uchicago.edu/~ngo/Shimura.pdf`. 5

[Gro61a] Alexander Grothendieck. éléments de géométrie algébrique : II. étude globale élémentaire de quelques classes de morphismes. *Publications Mathématiques de l'IHÉS*, 8:5–222, 1961. 59

[Gro61b] Alexander Grothendieck. éléments de géométrie algébrique : III. étude cohomologique des faisceaux cohérents, Première partie. *Publications Mathématiques de l'IHÉS*, 11:5–167, 1961. 59

[Gro62] Alexander Grothendieck. Technique de descente et théorèmes d'existence en géométrie algébrique. VI. Les schémas de Picard : propriétés générales. In *Séminaire Bourbaki : année 1961/62, exposés 223-240*, number 7 in Séminaire Bourbaki. Société mathématique de France, 1962. talk:236. 44

[Gro65] Alexander Grothendieck. éléments de géométrie algébrique : IV. étude locale des schémas et des morphismes de schémas, Seconde partie. *Publications Mathématiques de l'IHÉS*, 24:5–231, 1965. 45

[Gro66] Alexander Grothendieck. éléments de géométrie algébrique : IV. étude locale des schémas et des morphismes de schémas, Troisième partie. *Publications Mathématiques de l'IHÉS*, 28:5–255, 1966. 45

[Gro03] Alexander Grothendieck. *Revêtements étales et groupe fondamental (SGA 1)*, volume 3 of *Documents Mathématiques (Paris)*. Société Mathématique de France, Paris, 2003. Séminaire de géométrie algébrique du Bois Marie 1960–61. , Directed by A. Grothendieck, With two papers by M. Raynaud, Updated and annotated reprint of the 1971 original [Lecture Notes in Math., 224, Springer]. 39

[HH20] Thomas Haines and Michael Harris. *Shimura varieties*, volume 457. Cambridge University Press, 2020. 5

[Kle05] Steven L. Kleiman. The Picard scheme. In *Fundamental algebraic geometry*, volume 123 of *Math. Surveys Monogr.*, pages 235–321. Amer. Math. Soc., Providence, RI, 2005. 43, 44

[KM85] Nicholas M. Katz and Barry Mazur. *Arithmetic moduli of elliptic curves*, volume 108 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1985. 5, 20, 25

[Lan] Kai-Wen Lan. An example-based introduction to shimura varieties. Available at `https://www-users.cse.umn.edu/~kwlan/articles/intro-sh-ex.pdf`. 5

[Lan13] Kai-Wen Lan. *Arithmetic compactifications of PEL-type Shimura varieties*, volume 36 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2013. 62

[MFK94] D. Mumford, J. Fogarty, and F. Kirwan. *Geometric invariant theory*, volume 34 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)]*. Springer-Verlag, Berlin, third edition, 1994. 5, 27, 42, 45, 59, 60, 63, 65, 67, 68, 71, 72, 73, 74

[Mil11] James S. Milne. Shimura varieties and moduli, 2011. Available at `https://www.jmilne.org/math/xnotes/svh.html`. 5

[Mil17a] James S. Milne. Algebraic geometry (v6.02), 2017. Available at `https://www.jmilne.org/math/CourseNotes/ag.html`. 10

[Mil17b] James S. Milne. Introduction to shimura varieties, 2017. Available at `https://www.jmilne.org/math/xnotes/svi.html`. 5, 86

[Mum08] David Mumford. *Abelian varieties*, volume 5 of *Tata Institute of Fundamental Research Studies in Mathematics*. Published for the Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition. 5, 39, 45, 49, 50, 51, 54, 56, 57, 58, 59, 62

[Oes88] Joseph Oesterlé. Nouvelles approches du "théorème" de Fermat. Number 161-162, pages Exp. No. 694, 4, 165–186 (1989). 1988. Séminaire Bourbaki, Vol. 1987/88. 5, 74, 76

[Pra95] Dipendra Prasad. Ribet's theorem: Shimura-Taniyama-Weil implies Fermat. In *Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994)*, volume 17 of *CMS Conf. Proc.*, pages 155–177. Amer. Math. Soc., Providence, RI, 1995. 5, 74

[Rib90] K. A. Ribet. On modular representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms. *Invent. Math.*, 100(2):431–476, 1990. 5, 74, 79

[Rib94] Kenneth A. Ribet. Report on mod l representations of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. In *Motives (Seattle, WA, 1991)*, volume 55 of *Proc. Sympos. Pure Math.*, pages 639–676. Amer. Math. Soc., Providence, RI, 1994. 5, 74

[Sai13] Takeshi Saito. *Fermat's last theorem*, volume 243 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2013. Basic tools, Translated from the Japanese original by Masato Kuwata. 5, 74, 80, 81

[Sai14] Takeshi Saito. *Fermat's last theorem*, volume 245 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2014. The proof, Translated from the 2009 Japanese original by Masato Kuwata, Iwanami Series in Modern Mathematics. 5, 74, 83

[Ser87] Jean-Pierre Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. *Duke Math. J.*, 54(1):179–230, 1987. 78

[Sil94] Joseph H. Silverman. *Advanced topics in the arithmetic of elliptic curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994. 5, 18

[Sil09] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009. 5, 30, 31

[Sta18] The Stacks Project Authors. *Stacks Project*. https://stacks.math.columbia.edu, 2018. 26, 44, 45, 52, 57, 60