MODULAR CURVES

SARAH ZERBES

Contents

Waffle

Lecture 1

Let $\mathcal{H} = \{z \in \mathbf{C} | \Im(z) > 0\}$, so $\mathrm{SL}_2(\mathbf{R})$ is acting on $\mathcal{H}$ by Mobius transformations:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z \mapsto \frac{az+b}{cz+d}.$$

Note. This action factors through $\mathrm{PSL}_2(\mathbf{Z}) = \mathrm{SL}_2(\mathbf{Z})/\{\pm I\}$.

Let Γ be a subgroup of $\mathrm{SL}_2(\mathbf{Z})$ of finite index. Examples of such groups are

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\},$$

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : c \equiv 0 \pmod N \right\},$$

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : a, d \equiv 1 \pmod N, \quad c \equiv 0 \pmod N \right\}.$$

Let $Y(\Gamma) = \Gamma \backslash \mathcal{H}$. We will equip $Y(\Gamma)$ with various interesting structures.

Definition. A congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$ is a subgroup which contains $\Gamma(N)$ for some $N \geq 1$.

Remark. Every congruence subgroup is of finite index in $\mathrm{SL}_2(\mathbf{Z})$, but not every subgroup of $\mathrm{SL}_2(\mathbf{Z})$ of finite index is a congruence subgroup!

0.1. **Recap of modular forms.** Fix $\Gamma < \mathrm{SL}_2(\mathbf{Z})$ of finite index (level), and let $k \in \mathbf{Z}$. Then there exists a space $M_k(\Gamma)$, which is defined to be the set of holomorphic functions $F : \mathcal{H} \to \mathbf{C}$ such that

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$$

and a growth condition on the boudary. One can define the subspace $S_K(\Gamma) \subset M_k(\Gamma)$ of cusp forms.

Remark. Both $S_k(\Gamma)$ and $M_k(\Gamma)$ can be shown to be finite-dimensional $\mathbf{C}$-vector spaces.

Any modular form has a q-expension

$$f(z) = \sum_{n \geq 0} a_n q^n,$$

where $a_n \in \mathbf{C} \quad \forall n$ and $q = e^{\frac{2\pi i z}{h}}$. Here, h is the smallest positive integer such that $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \Gamma$.

1. Modular curves as Riemann surfaces

1.1. **Modular curves as topological spaces.** Clearly $\mathcal{H}$ has a topology, so $Y(\Gamma)$ inherits a quotient topology, which is the strongest topology such that the map $\pi : \mathcal{H} \to Y(\Gamma)$ is continuous.

Note. Quotient topologies can be pretty nasty, e.g. if we consider $\mathbf{Q}$ acting on $\mathbf{R}$ by translation, then the quotient $\mathbf{R}/\mathbf{Q}$ has the indiscrete topology.

Proposition 1.1.1. *For any $\tau_1, \tau_2 \in \mathcal{H}$, there exist neighbourhoods U_1 and U_2 of τ_1 and τ_2, respectively, such that if $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ satisfies $\gamma(U_1) \cap U_2 \neq \emptyset$, then $\gamma(\tau_1) = \tau_2$. We say that $\mathrm{SL}_2(\mathbf{Z})$* acts properly discontinuously *on $\mathcal{H}$.*

Proof. See Diamond + Shurman, Proposition 2.1.1. □

Corollary 1.1.2. *$Y(\Gamma)$ is a Hausdorff topological space.*

Proof. Let $P_1 \neq P_2$ be two points of $Y(\Gamma)$, and choose $\tau_1, \tau_2 \in \mathcal{H}$ lifting the P_i. Let U_1, U_2 be neighbourhoods of the τ_i as in Proposition 1.1.1. I claim that if we define $V_i = \pi(U_i)$, then V_1 and V_2 are open neighbourhoods of P_1 and P_2, respectively, such that $V_1 \cap V_2 = \emptyset$.

Suppose that $v_1 \cap V_2 \neq \emptyset$. Then

$$\pi^{-1}(V_1) \cap \pi^{-1}(V_2) \neq \emptyset \qquad \Leftrightarrow \qquad \bigcup_{\gamma \in \Gamma} \gamma U_1 \cap \bigcup_{\gamma' \in \Gamma} \gamma' U_2 \neq \emptyset.$$

Hence there exist $\gamma, \gamma' \in \Gamma$ such that $\gamma U_1 \cap \gamma' U_2 \neq \emptyset$, i.e.

$$(\gamma')^{-1} \gamma U_1 \cap U_2 \neq \emptyset.$$

Hence $(\gamma')^{-1}\gamma\tau_1 = \tau_2$ by our assumption on the U_i, which gives a contradiction as $P_1 \neq P_2$. □

We are also interested in the slightly larger space $X(\Gamma)$ which is a compactification of $Y(\Gamma)$.

Definition. Let $X(\Gamma) = Y(\Gamma) \cup C(\Gamma)$, where $C(\Gamma) = \Gamma \backslash \mathbf{P}^1(\mathbf{Q})$. We call $C(\Gamma)$ the *cusps of* Γ.

Example. $C(\mathrm{SL}_2(\mathbf{Z})) = \{\infty\}$.

Let $\mathcal{H}^* = \mathcal{H} \cup \mathbf{P}^1(\mathbf{Q})$, and give $\mathcal{H}^*$ a topology extending that of $\mathcal{H}$ as follows:

- the neighburhoods of ∞ are the sets $\{z | \Im(z) > R\}$ for some $R > 0$;
- the neighbourhoods of $x \in \mathbf{Q}$ are the circles tangent to $\mathbf{R}$ at x.

It is then easy to check that the action of Γ on $\mathcal{H}^*$ is still properly discontinuous, so $X(\Gamma)$ is Hausdorff.

Proposition 1.1.3. *$X(\Gamma)$ is compact.*

Proof. It suffices to find a compact subset of $\mathcal{H}^*$ mapping surjectively onto $X(\Gamma)$. Let

$$D^* = \{\infty\} \cup \left\{ z \in \mathcal{H} : -\frac{1}{2} \leq \Re(z) \leq \frac{1}{2}, \quad |z| > 1 \right\}.$$

It is a standard fact that D^* contains an element of every $\mathrm{SL}_2(\mathbf{Z})$-orbit on $\mathcal{H}$, and it is easy to check that D^* is compact. It follows that if $\gamma_1, \ldots, \gamma_n$ are coset representatives for $\Gamma \backslash \mathrm{SL}_2(\mathbf{Z})$, then $\bigcup_{i=1}^n \gamma_i D^*$ is compact and surjects onto $X(\Gamma)$. □

1.2. **Riemann surfaces: recap.**

Definition 1.2.1. A Riemann surface consists of the following data:

- a topological space X (Hausdorff and second-countable);
- a collection $(U_i, V_i, \phi_i)_{i\in I}$ where the $V_i \subset X$ are opens forming a cover of X, the U_i are opens in $\mathbf{C}$, and $\phi_i : X_i \to U_i$ are homeomorphisms,

such that if $V_i \cap V_j \neq \emptyset$, the map

$$U_i \cap \phi^{-1}(V_i \cap V_j) \xrightarrow{\phi_j^{-1}\circ\phi_i} U_j \cap \phi_j^{-1}(V_i \cap V_j)$$

is holomorphic.

Roughly, a Riemann surface is the least amount of structure on X needed to make sense of a function $f : X \to \mathbf{C}$ being holomorphic.

Definition 1.2.2. We say that $P \in Y(\Gamma)$ is an *elliptic point* if for some (hence any) $\tau \in \mathcal{H}$ lifting P, $\mathrm{Stab}_{\bar\Gamma}(\tau) \neq \{1\}$. Here, $\bar\Gamma$ denotes the image of Γ in $\mathrm{PSL}_2(\mathbf{Z}) = \Gamma/(\Gamma \cap \{\pm 1\})$.

Note. If P is elliptic for Γ, then it maps to an elliptic point of $Y(\mathrm{SL}_2(\mathbf{Z}))$, and there are only two of these: the orbits of i and of $\rho = e^{\frac{2\pi i}{3}}$; their stabilizers in $\mathrm{PSL}_2(\mathbf{Z})$ have orders 2 and 3, respectively. The set of elliptic points for any Γ is therefore finite.

Proposition 1.2.3. *There exist Riemann surface structures (clearly unique) on $Y(\Gamma)$ and $X(\Gamma)$ such that $\pi : \mathcal{H} \to Y(\Gamma)$ is holomorphic.*

Proof. If P is not elliptic and not a cusp, then we can easily find a chart around P: take τ to be a lifting of P to $\mathcal{H}$, and apply Proposition 1.1.1 with $\tau_1 = \tau_2 = \tau$. Let $U = U_1 \cap U_2$. Then U is a neighbourhood of τ such that $\gamma U \cap U = \emptyset$ for any $\gamma \neq 1 \in \bar\Gamma$. If V is the image of U in $Y(\Gamma)$, then $\phi = \pi|_U$ is a homeomorphism $U \cong V$.

Suppose now that P is elliptic. Proposition 1.1.1 gives us a U contining τ such that $U \cap \gamma U \neq \emptyset$ if and only if $\gamma \in \mathrm{Stab}_{\bar\Gamma}(\tau)$. The group $\mathrm{Stab}_{\bar\Gamma}(\tau)$ is finite, so (by replacing U with the open subset $\bigcap_{\gamma\in\mathrm{Stab}_{\bar\Gamma}(\tau)} \gamma U$) we may assume that U is fixed by this group. Then, if V is the image of U in $Y(\Gamma)$, the restriction of π to U is a homeomorphism $\mathrm{Stab}_{\bar\Gamma}(\tau)\backslash U \to V$.

Choose $\delta \in \mathrm{SL}_2(\mathbf{C})$ sending $\tau \mapsto 0$ and $\bar\tau \mapsto \infty$, and let U' be the image of U. (Such a δ always exists.) This conjugates $\mathrm{Stab}_{\bar\Gamma}(\tau)$ onto a finite group of Möbius transformations fixing 0 and ∞, which is therefore a cyclic group C of rotations by $e^{2\pi i/n}$ where $n \in \{2, 3\}$ is the order of the elliptic point. We therefore have a diagram (where all arrows are homeomorphisms)

$$\begin{array}{ccc} \mathrm{Stab}_{\bar\Gamma}(\tau)\backslash U & \longrightarrow & C\backslash U' \\ & \searrow & \downarrow \\ & & V \end{array}$$

However, two points in U' are in the same C-orbit if and only if they map to the same point under $z \mapsto z^n$. So we can extend this to a diagram

$$\begin{array}{ccccc} \mathrm{Stab}_{\bar\Gamma}(\tau)\backslash U & \longrightarrow & C\backslash U' & \xrightarrow{z\mapsto z^n} & U'' \\ & \searrow & \downarrow & \swarrow & \\ & & V & & \end{array}$$

and the right diagonal arrow $U'' \to V$ gives a coordinate chart around P.

Lastly, if P is a cusp, we argue similarly: choose δ mapping P to ∞. Then $\mathrm{Stab}_{\delta\bar\Gamma\delta^{-1}}(\infty)$ is the group of translations by $h\mathbf{Z}$, for some $h \in \mathbf{N}$, and $z \mapsto e^{2\pi i z/h}$ gives the local coordinate.

It's easy to see that all the above coordinate charts are compatible on overlaps. □

1.3. **Genus, ramification, Riemann-Hurwitz.** Fact: Riemann surfaces are connected orientable smooth 2-manifolds, and there are not very many of these. The compact ones "look like doughnuts": they are all homeomorphic to g-holed tori, for some integer $g \geq 0$.

Definition 1.3.1. Define the genus of a compact connected Riemann surface M as the unique integer $g = g(M)$ such that

$$H^1(M, \mathbf{Z}) \cong \mathbf{Z}^{2g}.$$

The genus is closely connected to the *Euler characteristic*

$$\chi(M) = \sum_{i \geq 0} (-1)^i \operatorname{rk} H^i(M, \mathbf{Z}) :$$

if M is as in the definition, then $H^0(M, \mathbf{Z}) \cong H^2(M, \mathbf{Z}) \cong \mathbf{Z}$ and $H^i(M, \mathbf{Z}) = 0$ for $i \geq 3$, so $\chi(M) = 2 - 2g$.

We need the following result from the theory of Riemann surfaces:

Proposition 1.3.2. *Let X be a compact Riemann surface, and let $f : X \to \mathbf{P}^1(\mathbf{C})$ be a non-constant meromorphic function. Then there exists an integer $n > 0$ (called the* valence *of f) such that f takes each value with multiplicity n.*

Proposition 1.3.3. *For $\Gamma = \mathrm{SL}_2(\mathbf{Z})$, the space $X(\Gamma)$ is isomorphic (as a Riemann surface, so in particular as a 2-manifold) to $\mathbf{P}^1(\mathbf{C}) \cong S^2$.*

Proof. The j-invariant

$$j(z) = q^{-1} + 744 + 196884q + \dots$$

is $\mathrm{SL}_2(\mathbf{Z})$-invariant and descends to a meromorphic map

$$X(\mathrm{SL}_2(\mathbf{Z})) \longrightarrow \mathbf{P}^1(\mathbf{C}).$$

Then j has valence 1 (one can show that it is holomorphic on $Y(\mathrm{SL}_2(\mathbf{Z}))$, and it clearly has a simple pole at ∞), so (by the Open Mapping Theorem of complex analysis) it has a holomorphic inverse. □

Convention. All Riemann surfaces are assumed to be connected.

Recap. We want to find $g(X(\Gamma))$ for all Γ. We know that $X(\mathrm{SL}_2(\mathbf{Z})) \cong \mathbf{P}^1(\mathbf{C})$ has genus 0. For all Γ, we have a map

$$X(\Gamma) \longrightarrow X(\mathrm{SL}_2(\mathbf{Z})).$$

Definition 1.3.4.

(1) For $f : X \to Y$ and $P \in X$, the ramification degree $e_P(f)$ is the unique integer $e \geq 1$ such that f looks like $z \mapsto z^e$ locally around P. Note that points where $e_P(f) > 1$ are isolated. Note also that in any neighbourhood of P, one can find $e(P)$ distinct points having the same image under f: f is locally $e(p)$-to-1.

(2) If X, Y are compact, then the sum

$$\sum_{P \in f^{-1}(Q)} e_P(f)$$

is independent of $Q \in Y$; it is called the *degree of f*.

Remark. The degree of the map $X(\Gamma) \to X(\mathrm{SL}_2(\mathbf{Z}))$ is $[\mathrm{PSL}_2(\mathbf{Z}) : \bar{\Gamma}]$.

Theorem 1.3.5 (Riemann-Hurwitz)**.** *For $f : X \to Y$ non-constant of degree N, X, Y compact, we have*

$$2g(X) - 2 = N \cdot (2g(Y) - 2) + \sum_{P \in X} (e_P(f) - 1).$$

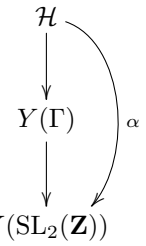
Corollary 1.3.6. *For any Γ, we have*

$$g(X(\Gamma)) = 1 + \frac{[\mathrm{PSL}_2(\mathbf{Z}) : \bar{\Gamma}]}{12} - \frac{\varepsilon_2}{4} - \frac{\varepsilon_3}{3} - \frac{\varepsilon_\infty}{2},$$

where ε_2 (resp. ε_3) is the number of elliptic points of order 2 (resp. order 3) and ε_∞ is the number of cusps.

Proof. We need to analyse ramfication of $f : X(\Gamma) \to X(\mathrm{SL}_2(\mathbf{Z}))$ at each $P \in X(\Gamma)$.

- If $P \in Y(\Gamma)$ is not in the $\mathrm{SL}_2(\mathbf{Z})$-orbit of i or ρ. Let $\tau \in \mathcal{H}$ be any lift of P, and let U be a neighbourhood of τ as in Proposition ... As $\mathrm{Stab}_{\bar{\Gamma}}(\tau) = \mathrm{Stab}_{\mathrm{PSL}_2(\mathbf{Z})}(\tau) = \{1\}$, U maps isomorphically to a neighbourhood of P in $X(\Gamma)$, resp. to a neighbourhood of $f(P)$ in $X(\mathrm{SL}_2(\mathbf{Z}))$, so $e_P(f) = 1$.



  The map α is unramified at any τ lifting a non elliptic point of $Y(\mathrm{SL}_2(\mathbf{Z}))$, so $e_P(f) = 1$.
- If P maps to $[i]$: all such P are either non-elliptic or elliptic of order 2. If P is elliptic of order 2, then $Y(\Gamma) \to Y(\mathrm{SL}_2(\mathbf{Z})$ is locally an isomorphism at P so $e_P(f) = 1$. If P is non-elliptic, then local coordinate for $\mathrm{SL}_2(\mathbf{Z})$ is square of that for Γ, so $e_P(f) = 2$. We use the definition of the degree to count the number of points above $[i]$. We have $N = 1 \cdot \epsilon_2 + 2 \cdot$ (number of non-elliptic points of $Y[\Gamma]$ above $[i]$). Hence, the number of non-elliptic points above $[i]$ is $(N - \epsilon_2)/2$. So

$$\sum_{P \in f^{-1}([i])} (e_P - 1) = \frac{N - \epsilon_2}{2}.$$

- If P maps to $[\rho]$, (where $\rho = e^{2\pi i/3}$). Then

$$e_P(f) = \begin{cases} 1 & \text{if } P \text{ is elliptic} \\ 3 & \text{if } P \text{ is not elliptic} \end{cases}.$$

  We use the same argument as before, using the definition of degree, to get that the number of non-elliptic points above $[\rho]$ is $(N - \epsilon_3)/3$. Hence

$$\sum_{P \in f^{-1}([\rho])} (e_P - 1) = \frac{2(N - \epsilon_3)}{3}$$

- If P is a cusp: let h be the width of the cusp P (that is the integer such that $e^{2\pi iz/h}$ is a local coordinate for $X(\Gamma)$ at P). A local coordinate for $X(\mathrm{SL}_2(\mathbf{Z})$ at $[\infty]$ is $(e^{2\pi iz/h})^h$, so $e_p(f) = h$. Thus

$$\sum_{P \in f^{-1}([\infty])} (e_P - 1) = \left(\sum_{P \in f^{-1}([\infty])} e_P \right) - e_\infty = N - e_\infty$$

Putting all of this together, we get

$$\begin{aligned} 2g(X(\Gamma)) - 2 &= (-2)N + \frac{N - \epsilon_2}{2} + \frac{2(N - \epsilon_3)}{3} + (N - \epsilon_\infty) \\ g(X(\Gamma)) &= 1 + \frac{N}{12} - \frac{\epsilon_2}{4} - \frac{\epsilon_3}{3} - \frac{\epsilon_\infty}{2} \end{aligned}$$

□

Example. Consider $\Gamma = \Gamma_0(11)$. Then $N = 12$, $\varepsilon_\infty = 2$ ([0] and $[\infty]$), $\varepsilon_2 = \varepsilon_3 = 0$ (exercise, c.f. Diamond and Shurman), so

$$g = 1 + \frac{12}{12} - 0 - 0 - \frac{2}{2} = 1.$$

Exercise. (1) Verify that $\varepsilon_2 = \varepsilon_3 = 0$ for $X_0(11)$.
(2) Show that the only primes p such that $g(X(\Gamma_0(p))) = 0$ are $\{2, 3, 5, 7, 13\}$.

Remark. For any g, there exists finitely many congruence subgroups Γ of $\mathrm{PSL}_2(\mathbf{Z})$ of genus g. (J.G. Thompson)

1.4. **Sheaves and Riemann-Roch.** Now let X be a Riemann surface, and let O_X be its structure sheaf, so $O_X(U)$ are holomorphic functions $U \to \mathbf{C}$. This is a sheaf of rings, so we can make sense of a sheaf of O_X-modules.

Definition 1.4.1. An invertible sheaf on X is a sheaf of O_X-modules that is locally free of rank 1. This is equivalent to it having an inverse with respect to the tensor product of O_X-modules.

Note. Invertible sheaves $\leftrightarrow$ line bundles with holomorphic structure.

We now specialise to the case when X is compact. We then have the notion of *meromorphic sections* of an invertible sheaf $\mathcal{F}$ (= sections of $\mathcal{F} \otimes_{O_X}$ {sheaf of meromorphic functions}).

Theorem 1.4.2 (Riemann existence theorem)**.** *An invertible sheaf on a compact Riemann surface has a non-zero global meromorphic section.*

Remark. This implies that there is the notion of the *degree* of an invertible sheaf, which is the sum of orders of vanishing of any non-zero meromorphic section. Note that this is well-defined, as the sum of the orders of the zeros and poles of a meromorphic function is 0.

We have

$$\deg(\mathcal{F} \otimes \mathcal{G}) = \deg(\mathcal{F}) + \deg(\mathcal{G}),$$
$$\deg(\mathcal{F}^{-1}) = -\deg(\mathcal{F}).$$

Moreover, invertible sheaves are a group under $\otimes$, with O_X as identity, and deg is a homomorphism to $\mathbf{Z}$.

Theorem 1.4.3. *[Riemann-Roch] Let X be a compact Riemann surface and $\mathcal{F}$ an invertible sheaf on X. Then*

(1) $H^0(X, \mathcal{F}) := \mathcal{F}(X)$ is finite-dimensional over $\mathbf{C}$;
(2) $\dim H^0(X, \mathcal{F}) - \dim H^0(X, \Omega \otimes \mathcal{F}^{-1}) = 1 - g + \deg(\mathcal{F})$, where Ω is the sheaf of holomorphic differentials on X.

Lecture 2

Corollary 1.4.4. *(1) If $\mathcal{F}$ is an invertible sheaf and $\deg(\mathcal{F}) < 0$, then $\mathcal{F}$ does not have any non-zero global sections.*
(2) If $\deg(\mathcal{F}) \gg 0$, then $H^0(X, \Omega \otimes \mathcal{F}^{-1}) = 0$, and we get a formula for $H^0(X, \mathcal{F})$.
(3) We have $\dim H^0(X, \Omega) = g(X)$.
(4) We have $\deg(\Omega) = 2g - 2$.

Proof. (1) Any global section would have to have a pole and hence can't be holomorphic.
(2) If $\deg(\mathcal{F}) \gg 0$, then

$$\deg(\Omega \otimes \mathcal{F}^{-1}) = \deg(\Omega) - \deg(\mathcal{F}) < 0,$$

so we conclude by (1).
(3) Take $\mathcal{F} = O_X$ and observe that the only functions on a compact Riemann surface which are evrywhere holomorphic are the constants, so $H^0(X, O_X) = 1$.
(4) Take $\mathcal{F} = \Omega$ □

Remark. There exists a cohomology theory for sheaves for which $H^0(X, \mathcal{F})$ is global sections. Then Riemann-Roch is a combination of two things:

- a formula for
$$\chi(\mathcal{F}) = \sum_{i \geq 0} (-1)^i \dim H^i(X, \mathcal{F});$$
- Serre duality:
$$H^i(X, \mathcal{F}) = H^{1-i}(X, \Omega \otimes \mathcal{F}^{-1})^*.$$

We call Ω the *dualizing sheaf.*

1.5. **The Katz sheaf.** Let $X = X(\Gamma)$ for some Γ, and let $k \in \mathbf{Z}$.

Definition 1.5.1. Let ω_k be the sheaf defined by

$$\omega_k(V) = \{\text{holomorphic functions } f \text{ on } \pi^{-1}(V) \subset \mathcal{H} \text{ satisfying } f(\gamma z) = j(\gamma, z)^k f(z)\ \forall \gamma \in \Gamma\}.$$

Here, if $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $j(\gamma, z) = cz + d$.

This is a sheaf of O_X-modules. If k is odd and $-1 \in \Gamma$, then it is the zero sheaf. Assume that this is not the case.

Definition. Let $\mathcal{L}$ be a sheaf of O_X-modules, and let $D = \sum_i n_i D_i$ be a divisor. Define

$$\mathcal{L}(D)(V) = \{\text{meromorphic sections } x \text{ of } \mathcal{L} \text{ over } V \text{ with } \operatorname{div}(x) + D \geq 0\}.$$

We think of $\mathcal{L}(P)$ as 'allowing a simple pole at P' and $\mathcal{L}(-P)$ as 'sections vanishing at P'.

Theorem 1.5.2. *(i) ω_k is invertible;*
(ii) $\omega_2 = \Omega_X(cusps)$.

Proof. (i) This is a case by case check. We just need to show it on a open neighbourhood of every $P \in X(\Gamma)$.

For P non-elliptic, not cusp, we can find $V \ni P$ open such that $\pi^{-1}(V) = \sqcup_{\gamma \in \overline{\Gamma}} \gamma U$ and $\omega_k(V) \cong \mathcal{O}_{\mathcal{H}}(U) \cong \mathcal{O}_X(V)$.

Other case: we want to show that for any $P \in X(\Gamma)$ there exists a neighbourhood V of P and $b \in \omega_K(V)$ such that

$$\omega_k(V) = O_X(V) \cdot b.$$

Choose $\tau \in \mathcal{H}^*$ lifting P and $U \subset \mathcal{H}^*$ open such that U is fixed by $\operatorname{Stab}_{\overline{\Gamma}}(\tau)$ and

$$\pi^{-1}(V) = \amalg_{\gamma \in \overline{\Gamma}/\operatorname{Stab}_{\overline{\Gamma}}(\tau)} \gamma U,$$

where $V = \pi(U)$. So

$$\omega_k(V) = \{f : U \to \mathbf{C} \quad \text{holomorphic and wt. } k \text{ invariant under } \operatorname{Stab}_{\overline{\Gamma}}(\tau)\},$$

while

$$O_X(V)\{f : U \to \mathbf{C} \quad \text{holomorphic and wt. } 0 \text{ invariant under } \operatorname{Stab}_{\overline{\Gamma}}(\tau)\}.$$

It follows that if $\operatorname{Stab}_{\overline{\Gamma}}(\tau) = 1$,then we can take $b = 1$.

If τ is elliptic, conjugate onto $\tau = 0$ as before. We have seen that $\operatorname{Stab}_{\overline{\Gamma}}(\tau)$ is isomorphic to the cyclic group $\langle e^{\frac{2\pi i}{n}} \rangle$, where $n \in \{2, 3\}$. This element $e^{\frac{2\pi i}{n}}$ corresponds to the matrix $\begin{pmatrix} e^{\frac{i\pi}{n}} & 0 \\ 0 & e^{-\frac{i\pi}{n}} \end{pmatrix}$ (check), so a function $U \to \mathbf{C}$ is weight k invariant under $\{e^{\frac{2\pi i}{n}}\}$ if and only if it satisfies

$$f\left(e^{\frac{2\pi i}{n}} z\right) = e^{-\frac{\pi i k}{n}} f(z),$$

which happens if and only if f is of the form

$$z \mapsto z^a f(z^n),$$

where a is the least non-negative integer such that $a \equiv \frac{k}{2} \pmod{n}$, and $b = z^a$ works.

Note:
- if $2n|k$, then $b = 1$ is a local basis;
- if there exists an elliptic point of order 2, then there are no non-trivial modular forms of odd weight.

If τ is a cusp, we take without loss of generality $\tau = \infty$. If ∞ is a regular cusp (i.e. if its stabilizer is generated by $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ for some $h \in \mathbf{Z}$), or k is even, then the weight k and weight 0 actions of the stabilizer coincide and $b = 1$ works. In the remaining case, the stabilizer is generated by $-\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ for some $h \in \mathbf{Z}$, and we have

$$O_X(V) = \{f : U \to \mathbf{C} \text{ holomorphic}, f(z + h) = f(z)\}$$

and

$$\omega_k(V) = \{f : U \to \mathbf{C} \text{ holomorphic}, f(z+h) = -f(z)\},$$

so we can take $b(z) = e^{i\pi z/h}$ as our local basis.

(ii) Clearly $f \mapsto f(z)dz$ gives a bijection $\mathcal{O}_{\mathcal{H}} \cong \Omega^1_{\mathcal{H}}$ commuting with the action of Γ, if we give $\mathcal{O}_{\mathcal{H}}$ the weight 2 action and $\Omega^1_{\mathcal{H}}$ the natural action; so passing to invariants we have $\omega_2|_{Y(\Gamma)} \cong \Omega_{Y(\Gamma)}$. Thus we only need to worry about cusps, and it suffices to treat the cusp ∞ as usual.

In a sufficiently small neighbourhood V of ∞, the local coordinate is $q = e^{\frac{2\pi i z}{h}}$, so $\mathrm{d}q = \frac{2\pi i}{h} q\,\mathrm{d}z$. Thus $\mathrm{d}z = \frac{h}{2\pi i}\frac{\mathrm{d}q}{q}$, and so we have

$$\omega_2(V) = \mathcal{O}_X(V)\mathrm{d}z = \mathcal{O}_X(V)\frac{\mathrm{d}q}{q} = \frac{1}{q}\Omega^1(V)$$

and so holomorphic sections of ω_2 give differentials with simple poles.

□

Note. By construction, we have

$$H^0(X(\Gamma), \omega_k) = M_k(\Gamma) \qquad \text{and} \qquad H^0\left(X(\Gamma), \omega_k(-\text{cusps})\right) = S_k(\Gamma),$$

so we've "brought modular forms into the world of sheaves".

Proposition 1.5.3. *Let r be the least common multiple of the set*

$$\{\#\operatorname{Stab}_\Gamma(P) : P \in \mathcal{H}\} \cup \{2 \text{ if } \exists \text{ irregular cusps}, 1 \text{ otherwise}\}.$$

(Thus $1 \le r \le 12$.) Then $\omega_{k+r} = \omega_k \otimes \omega_r$ for all $k \in \mathbf{Z}$. In particular, if $r = 1$ then $\omega_k = (\omega_1)^{\otimes r}$.

Proof. Our definition of r implies that all the local bases of ω_r in the previous proof were 1, and the local bases of ω_k depended only on k modulo r. So (local basis of ω_r) $\times$ (local basis of ω_k) = (local basis of ω_{k+r}) at every point. □

Definition 1.5.4. If $r = 1$ above (i.e. $-1 \notin \Gamma$, all cusps are regular and there are no elliptic points) then we say Γ is *neat*.

Corollary 1.5.5. *If Γ is neat, then for any $k \ge 2$ we have*

$$\dim M_k(\Gamma) = (k-1)(g-1) + \frac{k}{2}\varepsilon_\infty.$$

Proof. We have

$$\begin{aligned}\deg(\omega) &= \frac{1}{2}\deg(\omega^{\otimes 2}) \\ &= \frac{1}{2}(\deg(\Omega) + \varepsilon_\infty) \qquad \text{by Theorem1.4.3 (ii)} \\ &= \frac{1}{2}(2g - 2 + \varepsilon_\infty),\end{aligned}$$

so if $k \ge 2$, then $\deg(\omega^{\otimes k}) > 2g - 2$, so

$$\deg\left(\Omega \otimes (\omega^{\otimes k})^{-1}\right) = \deg(\Omega) - \deg(\omega^{\otimes k}) < 0$$

and hence $\Omega \otimes (\omega^{\otimes k})^{-1}$ has no global holomorphic sections. Therefore Riemann-Roch implies that

$$\begin{aligned}\dim H^0(X(\Gamma), \omega^{\otimes k}) &= k\big(g - 1 + \frac{1}{2}\varepsilon_\infty\big) - g + 1 \\ &= (k-1)(g-1) + \frac{k}{2}\varepsilon_\infty.\end{aligned}$$

□

There are similar (but messier) formulae for non-neat Γ, cf. chapter 3 of Diamond+Shurman.

1.6. **An extended example.** We start with the following result:

Lemma 1.6.1. *Let Γ be a level and let $\gamma_1, \ldots, \gamma_r$ be such that $\mathrm{SL}_2 \mathbf{Z} = \bigsqcup_j \Gamma\gamma_j$. Then $[\gamma_j i] \in Y(\Gamma)$ is elliptic if and only if $\gamma_j \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \gamma_j^{-1} \in \Gamma$.*

Proof. The stabiliser of $\gamma_j i$ in $\mathrm{SL}_2 \mathbf{Z}$ is just $\gamma_j \mathrm{Stab}_{\mathrm{SL}_2 \mathbf{Z}}(i)\gamma_j^{-1}$, which is the cyclic group generated by $\gamma_j \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \gamma_j^{-1}$. The orbit of $\gamma_j i$ is an elliptic point of $Y(\Gamma)$ if and only if this subgroup is contained in Γ, i.e. if and only if $\gamma_j \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \gamma_j^{-1} \in \Gamma$. □

Corollary 1.6.2. *Let p be an odd prime. Then $Y_0(p)$ has two elliptic points of order 2 if $p \equiv 1 \pmod 4$, and none if $p \equiv 3 \pmod 4$.*

Proof. In the case $\Gamma = \Gamma_0(p)$ we may take a set of coset representatives to be $\gamma_j = \begin{pmatrix} 1 & 0 \\ j & 1 \end{pmatrix}$ for $0 \leq j < p$ and $\gamma_p = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then $\gamma_j \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \gamma_j^{-1}$ is $\begin{pmatrix} j & -1 \\ 1+j^2 & -j \end{pmatrix}$, which is in Γ if and only if $j < p$ and $1 + j^2 = 0 \bmod p$. So if $p = 3 \bmod 4$ there are no elliptic points of order 2, while if $p = 1 \bmod 4$ there is at least one.

To see that there are two when $p = 1 \bmod 4$, we must check that the elliptic points $\gamma_j i$ and $\gamma_{-j} i$, where j is the square root of -1 modulo p, are distinct; but if $\gamma_j i$ and $\gamma_{-j} i$ (where $-j$ is taken modulo p) are in the same orbit, then we must have $\gamma_{-j} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \gamma_j \in \Gamma$. But this matrix has bottom left corner $1 - j^2 = 2 \neq 0$. □

We can similarly prove the analogous result for elliptic points of order 3 (exercise):

Corollary 1.6.3. *Let p be an odd prime. Then $Y_0(p)$ has two elliptic points of order 3 if $p \equiv 1 \pmod 3$, and none if $p \equiv 2 \pmod 3$.*

We now consider $\Gamma = \Gamma_0(5)$.

Lemma 1.6.4. *$X_0(5)$ has two cusps, so its genus is 0.*

Proof. We claim that the cusps are $[0]$ and $[\infty]$. To see that the cusps are distinct, assume that there exists $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(5)$ such that $\gamma.0 = \infty$. Then $d = 0$, but this gives a contradiction as $c \equiv 0 \pmod 5$.

Now let $\frac{m}{n} \in \mathbf{Q}^*$; assume that $(m, n) = 1$. If $5 \nmid n$, then we can find $a, c \in \mathbf{Z}$ such that $an - 5cm = 1$. Then $\gamma = \begin{pmatrix} a & m \\ 5c & n \end{pmatrix}$ is in $\Gamma_0(5)$ and satisfies $\gamma.0 = \frac{m}{n}$. If $d | n$, then we can find $b, d \in \mathbf{Z}$ such that $md - nb = 1$, and then $\gamma = \begin{pmatrix} m & b \\ n & d \end{pmatrix}$ is in $\Gamma_0(5)$ and satisfies $\gamma.\infty = \frac{m}{n}$.

We deduce from the previous two corollaries that $X_0(5)$ has two elliptic points of order 2 and no elliptic points of order 3. Moreover, we know from last lecture that the degree of the map $X(\mathrm{SL}_2(\mathbf{Z})) \to X_0(5)$ is equal to $[\mathrm{PSL}_2(\mathbf{Z}) : \bar{\Gamma}_0(5)] = [\mathrm{SL}_2(\mathbf{Z}) : \Gamma_0(5)]$, which is equal to 6. We therefore deduce from the Riemann-Hurwitz formula that $g(X_0(5)) = 0$. □

Proposition 1.6.5. *The space $S_4(\Gamma)$ is one-dimensional, and if F is a basis vector of $S_4(\Gamma)$, then multiplication by F is an isomorphism $M_k(\Gamma) \to S_{k+4}(\Gamma)$ for all $k \in \mathbf{Z}$.*

Proof. Since there are no elliptic points of order 3, we deduce from Proposition 1.5.3 that the sheaves ω_k satisfy $\omega_{k+4} = \omega_k \otimes \omega_4$, and hence

$$\omega_{k+4}(-\mathrm{cusps}) = \omega_k \otimes \omega_4(-\mathrm{cusps})$$

for all $k \in \mathbf{Z}$. So it suffices to show that $\dim H^0(X, \omega_4(-\mathrm{cusps})) = 1$. This will follow if $\omega_4(-\mathrm{cusps})$ has degree 0, as then $\deg\left(\Omega \otimes (\omega_4(-\mathrm{cusps}))^{-1}\right) = -2$, so $\Omega \otimes (\omega_4(-\mathrm{cusps}))^{-1}$ has no global holomorphic sections by Corollary 1.4.4, and we deduce the result from Riemann-Roch (Theorem 1.4.3).

However, we know $\omega_2(-\text{cusps}) \cong \Omega^1$ has degree $2g-2 = -2$, so ω_2 has degree 0 and thus (as the degree is a group homomorphism) ω_2^2 also degree 0. Let $\tau \in \mathcal{H}$ be a lift of one of the elliptic points, and conjugate τ to 0. Then as shown in the proof of Theorem 1.5.2, a section of ω_4 around P looks like $f(z^2)$ in a neighbourhood of τ. Similarly, a section of ω_2^2 around P looks like $zf(z^2)$ in a neighbourhood of τ, and it is easy to check that the two sheaves are locally isomorphic everywhere else. Hence $\omega_2^2 = \omega_4(-\text{ell.pts})$, and so $\deg \omega_4(-\text{ell.pts}) = 0$.

Finally, there are 2 elliptic points and 2 cusps, so we have $\deg \omega_4(-\text{cusps}) = \deg \omega_4(-\text{ell.pts}) = 0$ as required. □

2. Modular curves as algebraic curves

2.1. Modular curves over C.

Theorem 2.1.1. *(1) The $\mathbf{C}$-points of a smooth connected projective algebraic curve over $\mathbf{C}$ are canonically a Riemann surface; $X \mapsto X^{\text{an}}$.*
(2) Every compact Riemann surface is X^{an} for a unique X.
(3) There exists an equaivalence of categories

$$\{\text{loc. free sheaves of } O_X\text{-modules}\} \quad \Leftrightarrow \quad \{\text{loc. free sheaves of } O_{X^{\text{an}}}\text{-modules}\}.$$

Lecture 3

Remark. (1) is basically the implicit function theorem. We will later see a bit about the proof of (2). (3) is Serre's GAGA theorem. The functors are on the one hand

$$\mathcal{F} \mapsto O_{X^{\text{an}}} \otimes_{O_X} \mathcal{F}$$

and on the other hand

$$\mathcal{F} \mapsto \{\text{subsheaf of } \mathcal{F} \text{ whose sections over } U \text{ are elts. of } \mathcal{F}(U) \text{ extending to merom. sections on } X\}.$$

We deduce that for every Γ there is an algebraic variety $X(\Gamma)_{\mathbf{C}}$ and invertible sheaves ω_k on it such that

$$M_k(\Gamma) = H^0(X(\Gamma)_{\mathbf{C}}, \omega_k).$$

Here is an alternative (nicer) construction:

Theorem 2.1.2. $X(\Gamma)_{\mathbf{C}} \cong \text{Proj}\left(\bigoplus_{k\geq 0} M_k(\Gamma)\right)$.

(Cf. Hartshorne §II.2 for the definition of Proj of a graded ring.)

Proof. One knows that for any Noetherian graded $\mathbf{C}$-algebra S with $S_0 \cong \mathbf{C}$,

$$\text{Proj}(S) = \text{Proj}(S_{n\bullet}) \qquad \text{for any } n \geq 1,$$

where $S_{n\bullet} = \bigoplus_{k\geq 0} S_{nk}$.

Choose n to be the r from the last section, so

$$S_{n\bullet} = \bigoplus_{k\geq 0} H^0(X(\Gamma), \omega_n^{\otimes k}).$$

We now quote a standard fact in algebraic geometry: invertible sheaves of positive degree on curves are *ample*, so their sections give an embedding into projective space. □

Remark. In fact, the same argument can be used to prove Theorem 2.1.1 (ii): take any ample invertible sheaf on a Riemann surface, then its sections give an embedding into $\mathbf{P}^N$ for $N \gg 0$.

2.2. Descending the base field.

Question. Does there exist an algebraic curve over some number field K such that we get $X(\Gamma)_{\mathbf{C}}$ by base extension? (a "descent" of $X(\Gamma)_{\mathbf{C}}$ to K?)

Let's think a bit about what this means.

- Clearly not all varieties over $\mathbf{C}$ are definable over number fields: consider the elliptic curve $y^2 = X^3 + X + \pi$. This is not defined over any number field, as its j-invariant is $\frac{6192}{27\pi^2+4}$.
- Sometimes descents exist for nonobvious reasons: e.g. $\pi Y^2 = X^3 + X$ has a descent to $\mathbf{Q}$, even though its defining equations aren't rational, because it's isomorphic over $\mathbf{C}$ to $Y^2 = X^3 + X$.
- Even if a descent exists, it may not be unique: e.g. $\mathbf{P}^1_{\mathbf{Q}}$ and $\{X^2 + Y^2 + 2Z^2\} \subset \mathbf{P}^2_{\mathbf{Q}}$ become isomorphic over $\mathbf{C}$.

So we need to ask: is there a descent to a number field that *means something*?

The curves – fields correspondence. For any field k, there is a bijection

$$\text{smooth connected curves}/k \quad \leftrightarrow \quad \text{field extensions } K/k \text{ of transcendence deg. 1 with } \bar{k} \cap K = k.$$

In particular, for $X/\mathbf{C}$ a curve and $k \subseteq \mathbf{C}$,

$$\text{models of } X/k \quad \leftrightarrow \quad \text{subfields } K \text{ of } \mathrm{Rat}(X) \text{ generating it } /\mathbf{C} \text{ and st. } K \cap \bar{k} = k.$$

This gives a promising candidate: we can try to find a good subfield of $\mathrm{Rat}(X(\Gamma))$ (=meromorphic wt. 0 modular functions) and use this to descend to smaller fields.

Note. We have seen that the j-invariant gives an isomorphism $j : X(\mathrm{SL}_2(\mathbf{Z})) \cong \mathbf{P}^1(\mathbf{C})$. It follows that $\mathbf{C}(j(z))$ is the function field of $X(\mathrm{SL}_2(\mathbf{Z}))$.

The following result hands us a $\mathbf{Q}$-model of $X_0(N)$ on a plate: the model with function field $\mathbf{Q}(j(z), j(Nz))$. One can handle $X_1(N)$ and $X(N)$ similarly, but it won't be very illuminating, and we have lost sight of a vital ingredient: the sheaf ω.

Theorem 2.2.1. *Let $N \geq 2$. Then*

1. $\mathbf{C}(X_0(N)) = \mathbf{C}(j(z), j(Nz))$;
2. *The minimal polynomial of $j(Nz)$ over $\mathbf{C}(j(z))$ has coefficients in $\mathbf{Z}[j(z)]$, and it is equal to $\Phi_N(j(z), j(Nz))$ for Φ_N a symmetric polynomial.*
3. *If $N = p$ is prime, then $\Phi_p(X, Y) \equiv (Y^p - X)(Y - X^p) \pmod{p}$.*

Proof. Note that $j(Nz)$ is a meromorphic function on $X_0(N)$ if and only if $j(N\gamma.z) = j(Nz)$ for all $\gamma \in \Gamma_0(N)$. Write $\gamma = \begin{pmatrix} a & b \\ Nc & d \end{pmatrix}$. Then

$$j(N\gamma.z) = j\left(\frac{Naz + Nb}{Ncz + d}\right) = j\left(\frac{a(Nz) + Nb}{c(Nz) + d}\right) = j(Nz)$$

as $\begin{pmatrix} a & Nb \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z})$. Hence $\mathbf{C}(j(z), j(Nz)) \subseteq \mathbf{C}(X_0(N))$. Moreover, $\mathbf{C}(X_0(N))$ has degree

$$\left[\mathrm{PSL}_2(\mathbf{Z}) : \overline{\Gamma_0(N)}\right] = [\mathrm{SL}_2(\mathbf{Z}) : \Gamma_0(N)] =: m$$

over $\mathbf{C}(j)$.

Let $\gamma_1, \ldots, \gamma_m \in \mathrm{SL}_2(\mathbf{Z})$ be such that

$$\mathrm{SL}_2(\mathbf{Z}) = \amalg_i \Gamma_0(N)\gamma_i.$$

Then the functions $z \mapsto j(N\gamma_i z)$ are all meromorphic functions on $\mathcal{H}$, and they are conjugate to $j(Nz)$ under an automorphism leaving $j(z)$ fixed (namely $z \mapsto \gamma_i z$). If we can show that they are distinct, then Galois theory implies that $[\mathbf{C}(j(z), j(Nz)) : \mathbf{C}(j(z))] = m$, so (1) would follow.

Suppose $j(N\gamma_i z) = j(N\gamma_j z)$ for all $z \in \mathcal{H}$ and for some $i \neq j$. Since j defines an isomorphism $j : \mathrm{SL}_2(\mathbf{Z})\mathcal{H}^* \cong \mathbf{P}^1(\mathbf{C})$, this implies that there exists $g \in \mathrm{SL}_2(\mathbf{Z})$ such that $N\gamma_i.z = gN\gamma_j.z$ for all z, so

$$\begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \gamma_i = \pm g \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \gamma_j$$

$$\Rightarrow \qquad \gamma_i \gamma_j^{-1} \in \mathrm{SL}_2(\mathbf{Z}) \cap \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}^{-1} \mathrm{SL}_2(\mathbf{Z}) \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}.$$

But

$$\mathrm{SL}_2(\mathbf{Z}) \cap \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}^{-1} \mathrm{SL}_2(\mathbf{Z}) \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} = \Gamma_0(N),$$

which contradicts the assumption that γ_i and γ_j lie in different cosets. Hence the $j(N\gamma_i z)$ are distinct as required.

We now prove (2). We know from (1) that the minimal polynomial of $j(Nz)$ over $\mathbf{C}(j)$ is

$$\Phi_N(j(z), Y) = \prod_i (Y - j(N\gamma_i z)).$$

The coefficients are symmetric polynomials in the functions $j(N\gamma_i z)$ (and hence invariant under $z \mapsto \gamma_i.z$); in particular they are holomorphic on $\mathcal{H}$. As they are rational runctions in $j(z)$, they must in fact be polynomials. So $\Phi_N(X,Y) \in \mathbf{C}[X,Y]$. It remains to show that the polynomial has coefficients in $\mathbf{Z}$.

We know that $j(z) = q^{-1} + \sum_{n\geq 0} a_n q^n$ with $a_i \in \mathbf{Z}$. Moreover, by writing down coset representatives γ_i explicitly, one can show that all $j(N\gamma_i z)$ are of the form $j\left(\frac{az+b}{d}\right)$ with $ad = N$ (exercise). As

$$e^{\frac{2\pi i(az+b)}{d}} = e^{2\pi i z a/d} e^{2\pi i b/d},$$

we see that $j(N\gamma_i z)$ has q-expansion in $\mathbf{Z}[\zeta_N]((q^{\frac{1}{N}}))$ for all i.

Hence the coefficients of $\Phi_N(j,Y)$ are elements of $\mathbf{C}[j]$ (polynomials in j) whose q-expension lie in $\mathbf{Z}[\zeta_N]((q))$. We claim that this implies that they lie in $\mathbf{Z}[\zeta_N][j]$. Let $P(j) = \sum_{n=1}^k b_n j^n$ be one of these coefficients, and suppose that the q-expansion of $P(j(z))$ is in $\mathbf{Z}[\zeta_N]((q))$. Looking at the lowest-order term, we see that $b_k \in \mathbf{Z}[\zeta_N]$. Induction now gives that all $b_n \in \mathbf{Z}[\zeta_N]$.

Hence $\Phi_N(X,Y) \in \mathbf{Z}[\zeta_N][X,Y]$, say $\Phi_N(X,Y) = \sum c_{r,s} X^r Y^s$. By construction, $c_{0,m} = 1$ and $c_{r,m} = 0$ for $r > 0$. Substitute the q-expansion of $j(z)$ (which has coefficients in $\mathbf{Q}$). We get a grotesque mess of infinitely many *linear* equations in the $c_{r,s}$ with coefficients in $\mathbf{Q}$. This has a unique solution in $\mathbf{C}$ (because of the uniqueness of the monic minimal polynomial), but it must be Galois invariant, so it takes values in $\mathbf{Q}$. But $\mathbf{Q} \cap \mathbf{Z}[\zeta_N] = \mathbf{Z}$.

Let us finally show that $\Phi_N(X,Y)$ is symmetric. We have $\Phi_N(j(z), j(Nz)) = 0$ for all z, so

$$\Phi_N\Big(j\big(-\frac{1}{Nz}\big), j\big(-\frac{N}{Nz}\big)\Big) = 0 \qquad \Leftrightarrow \qquad \Phi_N(j(Nz), j(z)) = 0$$

for all z. Thus $\Phi_N(Y,X)$ is a constant multiple of $\Phi_N(X,Y)$, necessarily ± 1. But if $\Phi_N(Y,X) = -\Phi_N(X,Y)$, then $\Phi_N(X,X) = 0$, so $(Y-X)$ is a factor of $\Phi_N(X,Y)$. This contradicts the irreducibility of $\Phi_N(j,Y) \in \mathbf{C}(j)[Y]$. This proves (2).

Finally, let's do (3), which is included since (a) it is the key to understanding $\pmod p$ reductions of modular curves, and (b) it's very cool.

For $N = p$, we can write down the γ_i explicitly (see last lecture), and we have

$$\{j(p\gamma_i z) : 1 \leq i \leq m\} = \{j(pz)\} \cup \left\{ j\left(\frac{z+i}{p}\right) : 0 \leq i < p \right\}.$$

Hence

$$\Phi_p(j(z), Y) = (Y - j(pz)) \prod_{i=0}^{p-1} \left(Y - j\left(\frac{z+i}{p}\right)\right).$$

As elements of $\mathbf{Z}[\zeta_p]((q))$, the $j\big(\frac{z+i}{p}\big)$ for $0 \leq i < p$ are all congruent modulo the unique prime ideal $\wp$ above p in $\mathbf{Z}[\zeta_p]$, so

$$\begin{aligned}\Phi_p(j(z), Y) \equiv& (Y - j(pz))(Y - j(z/p))^p \pmod{\wp \mathbf{Z}[\zeta_p]((q))} \\ \equiv& (Y - j(z)^p)(Y^p - j(z)) \pmod{\wp \mathbf{Z}[\zeta_N]((q))}.\end{aligned}$$

But the coefficients of $(Y - j(z)^p)(Y^p - j(z))$ are in $\mathbf{Z}((q))$, so the congruence is $\pmod{p\mathbf{Z}((z))}$. Since we can read off coefficients of $\Phi_p(X,Y)$ from q-expansions, this implies

$$\Phi_p(X,Y) \equiv (Y - X^p)(Y^p - X) \pmod p.$$

□

Remark. (1) This does not mean that $X_0(N)_{\mathbf{C}}$ is the projective curve defined by $\Phi_N(X,Y) = 0$. This curve is highly singular in general; the statement is that the smooth curve $X_0(N)_{\mathbf{C}}$ is *birational* to this curve.

(2) The coefficients of Φ_N are *huge*. For instance, for $N = 2$ we have

$$\begin{aligned}\Phi_N(X,Y) = X^3 + Y^3 - X^2Y^2 + 1488XY(X^2Y + Y^2X) - 162000(X^2 + Y^2) \\ + 40773375XY + 8748000000(X + Y) - 157464000000000.\end{aligned}$$

Other examples can be found at *https://math.mit.edu/∼drew/ClassicalModPolys.html*.

(3) Note that the mod$p$ curve defined by Φ_p is reducible: it is two intersecting copies of $\mathbf{P}^1$.

Definition. Define $X_0(N)$ to be the unique smooth projective curve $/\mathbf{Q}$ which has function field $\mathbf{Q}(X)[Y]/\Phi_N(X,Y)$.

Remark. The map $X_0(N)_{\mathbf{C}} \to X_0(1)_{\mathbf{C}} = \mathbf{P}^1$ descends to $\mathbf{Q}$, which is a good sign.

Lemma 2.2.2. *The cusp and the elliptic points of $X_0(1)$ are defined over $\mathbf{Q}$.*

Proof. This is clear for the cusp, as it is just the $\mathbf{Q}$-point ∞ of $X_0(1)$. To see that the elliptic points are defined over $\mathbf{Q}$, note that the isomorphism $j : \mathcal{H}/\operatorname{SL}_2(\mathbf{Z}) \to X_0(1)_{\mathbf{C}}$ send $[i]$ to 1 and $[\rho]$ to 0. □

Theorem 2.2.3. *There is a sheaf $\omega_{k,\mathbf{Q}}$ on $X_0(N)_{\mathbf{Q}}$ whose base-extension to $\mathbf{C}$ is ω_k.*

Proof. Since $-1 \in \Gamma_0(N)$, $\omega_{k,\mathbf{C}}$ is the zero sheaf for odd k, so we may assume k is even.

We saw above that $\omega_2 \cong \Omega^1_{X_0(N)_{\mathbf{C}}}(\text{cusps})$. More generally, by comparing local bases using Theorem 1.5.2, we see that

$$\omega_{2k} = \left(\Omega^1_{X_0(N)_{\mathbf{C}}}\right)^{\otimes k}(D_k),$$

where D_k is a $\mathbf{Z}$-linear combination of the divisors (cusps), (elliptic points of order 2) and (elliptic points of order 3).

Claim: These three divisors are defined over $\mathbf{Q}$.

For (cusps) this is easy, since the cusps of $X_0(N)$ are just the preimages of the $\mathbf{Q}$-point ∞ of $X_0(1)_{\mathbf{Q}}$. For (elliptic points of order 2) this is slightly harder, since not all preimages of the order 2 elliptic point $[i] \in X_0(1)_{\mathbf{Q}}$ are elliptic. But the elliptic points of $X_0(N)$ are exactly the preimages of $[i]$ at which the projection map $X_0(N) \to X_0(1)$ is *ramified*, and the ramification degree of a map defined over $\mathbf{Q}$ depends only on the Galois orbit, so we're fine. Similarly for elliptic points of order 3.

Now, we obviously have

$$\Omega^1_{X_0(N)_{\mathbf{C}}} = \mathbf{C} \otimes_{\mathbf{Q}} \Omega^1_{X_0(N)_{\mathbf{Q}}}$$

and we've seen that the D_k are defined over $\mathbf{Q}$, so we can take

$$\omega_{2k,\mathbf{Q}} = \left(\Omega^1_{X_0(N)_{\mathbf{Q}}}\right)^{\otimes k}(D_k). \tag{1}$$

□

Corollary 2.2.4. *For any $k \geq 2$ even and any $N \geq 1$, the spaces $S_k(\Gamma_0(N))$, $M_k(\Gamma_0(N))$ have a basis consisting of forms whose q-expensions have coefficients in $\mathbf{Q}$.*

Proof. We give the argument for M_k; the case of S_k is similar.

$$\begin{aligned} M_k(\Gamma_0(N)) &= H^0(X_0(N)_{\mathbf{C}}, \omega_k) \\ &= \mathbf{C} \otimes H^0(X_0(N)_{\mathbf{Q}}, \omega_{k,\mathbf{Q}}) \end{aligned}$$

Claim. The image of $H^0(X_0(N)_{\mathbf{Q}}, \omega_{k,\mathbf{Q}})$ consists of functions with q-expansions in $(2\pi i)^{\frac{k}{2}}\mathbf{Q}[[q]]$.

By (1), any section of $\omega_{k,\mathbf{Q}}$ is a meromorphic section of $\left(\Omega^1_{X_0(N)_{\mathbf{Q}}}\right)^{\otimes \frac{k}{2}}$, so it can be written as an element of $\mathbf{Q}(X_0(N)_{\mathbf{Q}})$ multiplied by $(dj)^{\otimes \frac{k}{2}}$. As the q-expansions of $j(z)$ and $j(Nz)$ have coefficients in $\mathbf{Q}$, the same is true for an element of $\mathbf{Q}(X_0(N)_{\mathbf{Q}})$. It therefore suffices to calculate

$$dj = j'(z)dz = J'(q) \cdot 2\pi i q \cdot dz$$

where $J(q) \in \mathbf{Q}((q))$. Hence $dj \in 2\pi i \cdot \mathbf{Q}((q))dz$, as required. □

Remark. Unfortunately this method does not extend to other Γ: it is hard to write down $\mathbf{C}(X(\Gamma))$ explicitly, and even harder to get one's hands on ω_k for odd k when $-1 \notin \Gamma$. Hence we need a different method.

3. Modular curves as moduli spaces

3.1. **Lattices and level structures.** Let Λ be a lattice in $\mathbf{C}$ (i.e. a discrete subgroup isomorphic to $\mathbf{Z}^2$).

Definition. Two lattices Λ_1, Λ_2 are homothetic if there exists $\alpha \in \mathbf{C}^\times$ such that $\Lambda_2 = \alpha\Lambda_1$.

Lemma 3.1.1. *Every lattice Λ is homothetic to a lattice of the form*

$$\Lambda_\tau = \mathbf{Z} + \mathbf{Z}\tau, \tau \in \mathcal{H}.$$

Moreover, τ is unique modulo $\operatorname{SL}_2(\mathbf{Z})$.

Proof. Exercise. □

We also need the following result:

Lemma 3.1.2.
- *Every elliptic curve over* $\mathbf{C}$ *is isomorphic to* $E_\tau := \mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau)$*, for some* $\tau \in \mathcal{H}$
- *Two elliptic curves* E_τ *and* $E_{\tau'}$ *are isomorphic if and only if* τ *and* τ' *are in the same* $\mathrm{SL}_2(\mathbf{Z})$*-orbit, in which case any* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *mapping* τ *to* τ' *gives an isomorphism* $\mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau') \to \mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau)$ *via* $z \mapsto (c\tau+d)z$ *on* $\mathbf{C}$*.*

Corollary 3.1.3. *We have*

$$Y(\mathrm{SL}_2(\mathbf{Z})) = \{\textit{homothety classes of lattices}\} = \{\textit{iso. classes of elliptic curves/}\mathbf{C}\}.$$

Proposition 3.1.4.
(1) The map $\tau \mapsto (\mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau), \frac{1}{N}\mathbf{Z})$ *gives a bijection between* $\Gamma_0(N)\backslash\mathcal{H}$ *and the set of equivalence classes of pairs* (E, C)*, where* E *is an elliptic curve over* $\mathbf{C}$ *and* C *is a cyclic subgroup of* E *of order* N*.*

(2) The map $\tau \mapsto (\mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau), \frac{1}{N})$ *gives a bijection between* $\Gamma_1(N)\backslash\mathcal{H}$ *and the set of equivalence classes of pairs* (E, P)*, where* E *is an elliptic curve over* $\mathbf{C}$ *and* P *is a point of* E *of exact order* N*.*

Proof. Exercise. □

Note. In (2), (Λ, P) is always equivalent to $(\Lambda, -P)$. In (1) there are lots more exceptional cases coming from elliptic points.

We have a similar description for $X(N)$: Let E be an elliptic curve over $\mathbf{C}$, and $N > 1$. The *Weil pairing* is a perfect pairing

$$E[N] \times E[N] \to \mu_N.$$

Fact: if $E = \mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau)$, then $\langle \tau/N, 1/N\rangle_{E[N]} = e^{2\pi i/N}$.

Proposition 3.1.5. *The map* $\tau \mapsto (\mathbf{C}/(\mathbf{Z}+\mathbf{Z}\tau), \frac{\tau}{N}, \frac{1}{N})$ *gives a bijection between* $\Gamma(N)\backslash\mathcal{H}$ *and the set of equivalence classes of triples* (E, P, Q) *with* E *an elliptic curve over* $\mathbf{C}$ *and* P, Q *two points of order* N *on* E *with* $\langle P, Q\rangle_{E[N]} = e^{2\pi i/N}$*.*

Lecture 4

Proof. Using a similar argument to the one above, we see that the natural map $E_{\gamma\tau} \to E_\tau$ sends $1/N$ to $\frac{c\tau+d}{N}$ and τ/N to $\frac{a\tau+b}{N}$ modulo $\mathbf{Z}+\mathbf{Z}\tau$. This shows that the map is injective (and well-defined).

Given a triple (E, P, Q), we may assume without loss of generality that $E = E_\tau$ for some τ. We have $P, Q = \frac{a\tau+b}{N}, \frac{c\tau+d}{N}$ for some $a, b, c, d \in \mathbf{Z}/N\mathbf{Z}$, and since $\langle P, Q\rangle = e^{2\pi i/N}$, we know that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$. It is a known result that $\mathrm{SL}_2(\mathbf{Z})$ surjects onto $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$, so so we can choose some $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ such that $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mod N. Then the map $E_{\gamma\tau} \to E_\tau$ sends $(\tau/N, 1/N)$ to P, Q, and hence $(E, P, Q) \sim (E_{\tau'}, \tau'/N, 1/N)$ where $\tau' = \gamma\tau$. □

Natural question: If $x \in Y_0(N)_{\mathbf{Q}}(\mathbf{Q})$, then does (E, C), E the elliptic curve $\mathbf{C}/\Lambda_x$, descend to $\mathbf{Q}$?
This is the right sort of question to ask to understand the arithmetic of modular curves.

Remark. This actual question is vacuous for $N \gg 0$ as $Y_0(N)(\mathbf{Q})$ is actually empty by a theorem of Mazur, but we don't know this yet!

3.2. **Moduli spaces and representable functors.** We have categories $\underline{Ring}$ (all rings are assumed commutative and unital here), $\underline{R-Alg}$ (algebras over a fixed ring R, a "slice category" of $\underline{Ring}$) and $\underline{Set}$.

Remark. These categories are large – the collections of their objects are not sets – but we won't worry too much about foundational issues here.

"Most sets that naturally arise in algebraic geometry" (whatever that means) are actually functors $\underline{Ring} \to \underline{Set}$, or $\underline{R-Alg} \to \underline{Set}$ for some R. E.g.

(1) Sets of points of varieties or schemes
(2) Sets of varieties of a certain kind, or structures on such varieties.

A lot of the fun and power of algebraic geometry comes from the fact that many instances of example (2) are actually instances of example (1) in disguise! These are the so-called *moduli spaces*: geometric objects that parametrize other geometric objects.

We now discuss some properties of functors and representable functors.

Definition. Let $\mathcal{C}$ be a *locally small* category (homomorphisms between any two objects are a set). For an object X in $\mathcal{C}$, denote by h^X the functor (*covariant Hom-functor*)

$$\mathrm{Hom}(X, \sim) : \mathcal{C} \longrightarrow \underline{\mathrm{Set}}.$$

A covariant functor $\mathcal{F} : \mathcal{C} \to \underline{\mathrm{Set}}$ is *representable* if there exists an isomorphism of functors $\mathcal{F} \cong h^X$ for some $X \in \mathrm{Ob}(\mathcal{C})$.

How do we specify the isomorphism $\mathcal{F} \cong h^X$? Note that $h^X(X)$ has a canonical element: id_X. Let x be the element of $\mathcal{F}(X)$ corresponding to id_X under the isomorphism $h^X(X) \cong \mathcal{F}(X)$.

Note. The choice of x determines an element of $\mathcal{F}(Y)$ for every homomorphism $\alpha : X \to Y$: take $\mathcal{F}(\alpha)(x)$.

Proposition 3.2.1 (Yoneda's lemma)**.** *This construction gives a bijection*

$$\{\textit{nat. transformations } h^X \to \mathcal{F}\} \xrightarrow{\sim} \mathcal{F}(X)$$

for any $\mathcal{F} : \mathcal{C} \to \underline{\mathrm{Set}}$ *and* $\mathcal{X} \in \mathrm{Ob}(\mathcal{C})$.

Corollary 3.2.2. *If* $\mathcal{F}$ *is representable, then the bijection* $\mathcal{F}(Y) \cong h^X(Y)$ *is determined by an object* $x \in \mathcal{F}(X)$.

In other words, for every $Y \in \mathrm{Ob}(\mathcal{C})$ and $y \in \mathcal{F}(Y)$, there exists a unique homomorphism $\alpha : X \to Y$ such that $\mathcal{F}(\alpha)(x) = y$. We say that (X, x) *represents* $\mathcal{F}$. Note that x is an essential part of the data!

Remark. We have assume that $\mathcal{F}$ is covariant, but we get the same for contravariant functors by replacing $\mathcal{C}$ with $\mathcal{C}^{\mathrm{opp}}$. In other words, if $\mathcal{G} : \mathcal{C} \to \underline{\mathrm{Set}}$ is a contravariant functor, then (X, x) represents $\mathcal{G}$ if for every $Y \in \mathrm{Ob}(\mathcal{C})$ and $y \in \mathcal{G}(Y)$, there exists a unique homomorphism $\alpha : Y \to X$ such that $\mathcal{G}(\alpha)(x) = y$.

Examples. Let $\mathcal{C} = \underline{\mathrm{Ring}}$.

- $\mathcal{F}(R) = R$ (the 'forgetful functor') is represented by $(\mathbf{Z}[T], T)$: for any ring R, $r \in R$, there exists a unique $\alpha : \mathbf{Z}[T] \to R$ such that $\alpha(T) = r$;
- $\mathcal{F}(R) = R^\times$ is represented by $(\mathbf{Z}[T, T^{-1}], T)$;
- $\mathcal{F}(R) = \{n\text{th roots of unity in } R\}$ is represented by $\mathbf{Z}[T]/(T^n - 1), T)$. (Caveat: 'primitive nth roots of 1' is not a functor on R.)
- Consider the contravariant functor $\mathcal{G} : \underline{Top} \to \underline{Sets}$ which maps a topological space T to the set of open subsets of T. Then $\mathcal{G}$ is represented by $(\{x, y\}, x)$, where the two-point space $\{x, y\}$ has the topology for which the open sets are $\{\varnothing, \{x\}, \{x, y\}\}$. To check this, note that if $f : A \to B$ is a continuous map, then $\mathcal{G}(f)$ sends an open subset of B to its preimage under f. Now let A be a topological space, and let $C \subseteq A$ be open. Define $f : A \to \{x, y\}$ by

$$f(z) = \begin{cases} x & \text{if } z \in C \\ y & \text{otherwise} \end{cases}.$$

  Then f is continuous (so it is a morphism in $\underline{\mathrm{Top}}$), and $\mathcal{G}(f)(x) = f^{-1}(\{x\}) = C$, as required.

Non-example. (taken from some online notes by Zach Norwood) The functor $\mathcal{F}(R) = \{\text{squares in } R\}$ is not representable: suppose that $\mathcal{F}$ is represented by (A, a) for some ring A and $a \in A$ with $a = b^2$ for some $b \in A$. The for any ring S and an element $s \in S$ which is a square, there exists a unique homomorphism $\alpha : A \to S$ such that $\alpha(a) = s$. But: take $S = \mathbf{Z}[T]$ and $s = T^2$. Then there exists a unique $\alpha : A \to \mathbf{Z}[T]$ with $\alpha(a) = T^2$, so $\alpha(b) \in \{\pm T\}$. Let $\sigma : S \to S$ be $T \to -T$, so $\sigma(s) = s$. Then $\sigma \circ \alpha \in \mathrm{Hom}(A, S)$ also sends a to s, but $\sigma \circ \alpha \neq \alpha$ as $\sigma \circ \alpha(b) \neq \alpha(b)$. The contradicts the uniqueness of α.

The moral of this non-example: *automorphisms are bad for representability!* Here is a second non-example.

Lemma 3.2.3. *Let $\mathcal{F}$ be a representable functor $\underline{Ring} \to \underline{Set}$, and let $(R_i)_{i\geq 1}$ is a projective system of rings (i.e. a collection of rings R_i and morphisms $R_{i+1} \to R_i$). Set $R = \varprojlim R_n$. Then $\mathcal{F}(R) = \varprojlim_n \mathcal{F}(R_n)$.*

Proof. It suffices to show that if S is any ring then there is a bijection $\hom(S, \varprojlim_i R_i) \to \varprojlim_i \hom(S, R_i)$. But this is just the definition of the inverse limit. □

Proposition 3.2.4. *The functor $\mathcal{F} : \underline{Ring} \to \underline{Set}$ mapping a ring R to the set of roots of unity in R is not representable.*

Proof. Fix a prime p and consider the rings $R_i = \mathbf{Z}/p^i$. Since R_i is finite, every invertible element of R_i is a root of unity, so $\varprojlim_i \mathcal{F}(R_i) = \varprojlim_i R_i^\times = \mathbf{Z}_p^\times$. But $1+p$ is an element of $\mathbf{Z}_p^\times$ which is not a root of unity in $\mathbf{Z}_p$, so $\mathcal{F}(\mathbf{Z}_p) \neq \mathbf{Z}_p^\times$, and hence $\mathcal{F}$ can't be representable by Lemma 3.2.3. □

3.3. **Elliptic curves over general base schemes.** We will define the notion of *elliptic curves over S*, where S is a scheme.

Definition 3.3.1. Let S be a scheme. An elliptic curve over S is a scheme $\mathcal{E}$ with a morphism $\pi : \mathcal{E} \to S$ (an S-scheme) such that π is proper and flat and all fibres are smooth genus 1 curves, given with a section "0" : $S \to \mathcal{E}$.

Note. If $\mathcal{E}$ is an elliptic curve over S and $R \to S$ is a morphism of schemes, then $\mathcal{E}' = R \times_S \mathcal{E}$ together with the natural section $R \to \mathcal{E}'$ is an elliptic curve over R. In other words, the functor

$$\begin{aligned} \mathrm{Sch} &\longrightarrow \underline{\mathrm{Set}} \\ S &\mapsto \{\text{elliptic curves } \mathcal{E}/S\} \end{aligned}$$

is contravariant.

Example. In Silverman's book, there is the equation

$$Y^2 + XY = X^3 - \frac{36}{j-1728}X - \frac{1}{j-1728}.$$

The associated homogeneous cubic

$$Y^2Z + XYZ = X^3 - \frac{36}{j-1728}XZ^2 - \frac{1}{j-1728}Z^3$$

is a subscheme of $\mathbf{P}^2/R$, where R is the ring $\mathbf{Z}[j, j^{-1}, (j-1728)^{-1}]$. This is an elliptic curve over $\mathrm{Spec}(R)$, with discriminant $\Delta = \frac{j^2}{(j-1728)^3}$.

Think of this as a family of elliptic curves: one for every $j \neq 0, 1728$, varying in an 'algebraic way'.

Definition. For $\mathcal{E}/S$ an elliptic curve,

$$\mathcal{E}(S) = \mathrm{Hom}_{S-\mathrm{Sch}}(S, \mathcal{E})$$

are the sections of $\pi : \mathcal{E} \to S$ picking out a point on each fibre.

Warning. If $P \in \mathcal{E}(S)$ has order N, i.e. $N \cdot P = 0$ and $M \cdot P \neq 0$ for all $1 \leq M < N$, it is not necessarily true that P_x has order N on $\mathcal{E}_x$ for every $x \in S$! For example, if $\mathcal{E}/\mathrm{Spec}(\mathbf{Z}_p)$, can have points of order p reducing $\pmod p$ to 0 (at closed point of $\mathrm{Spec}(\mathbf{Z}_p)$).

Lemma 3.3.2. *Let $\omega_{\mathcal{E}/S} = \pi_*\left(\Omega^1_{\mathcal{E}/S}\right)$. Then $\omega_{\mathcal{E}/S}$ is an invertible sheaf on S.*

Proof. The invertibility of $\pi_*\left(\Omega^1_{\mathcal{E}/S}\right)$ comes from a calculation in sheaf cohomology, c.f. p.53 of Mumford's 'Abelian varieties'. □

Remark. If $S = \mathrm{Spec}(K)$ where K is a field, then a basis of $\omega_{\mathcal{E}/S}$ is the invariant differential on $\mathcal{E}$.

Proposition 3.3.3. *If $\mathcal{E}/S$ is an elliptic curve, then $\mathcal{E}$ has a Weierstrass equation locally on S, i.e. there exists a covering $\amalg U_i \to S$ in the Zariski topology such that $\mathcal{E}|_{U_i}$ has a Weierstrass equation for all i. More precisely, any local basis ω of $\omega_{\mathcal{E}/S}$ (over some $U \subset S$ open) determines a Weierstrass equation over U. If 2 is invertible on S, we can do this in such a way that $\omega = \frac{-dx}{2y}$.*

Sketch. Given U and ω a basis of $\pi_*(\Omega^1_{\mathcal{E}/U})$, we can choose a local parameter on $\mathcal{E}$ at 0 such that

$$\omega = dT \cdot (1 + \text{higher order terms});$$

T is a local parameter *adapted to* ω.

Also, it follows from Riemann-Roch that $\pi_*O_{\mathcal{E}}(n(0))$ is locally free of rank n over U for all $n > 0$. Hence, if $U = \operatorname{Spec}(A)$ is affine, then $\pi_*O_{\mathcal{E}}((0)) \cong A{\cdot}1$, and $\pi_*O_{\mathcal{E}}(2(0)) \cong A{\cdot}(1, x)$, where $x = \frac{1}{T^2}(1+\dots)$.

Similarly,

$$\begin{aligned}
\pi_*O_{\mathcal{E}}(3(0)) &= A \cdot (1, x, y), \qquad \text{where } y = \frac{1}{T^3} + \dots, \\
\pi_*O_{\mathcal{E}}(4(0)) &= A \cdot (1, x, y, x^2), \\
\pi_*O_{\mathcal{E}}(5(0)) &= A \cdot (1, x, y, x^2, xy).
\end{aligned}$$

Now $y^2 - x^3 \in \pi_*O_{\mathcal{E}}(5(0))$, so $y^2 - x^3 \in A \cdot (1, x, y, x^2, xy)$, and that is a Weierstrass equation over $A[x, y]$. Moreover, $dx = -\frac{2dT}{T^3} + \dots$ and $y = \frac{1}{T^3} + \dots$, so if 2 is invertible on S, then

$$\omega = \frac{-dx}{2y} \pmod{TdT},$$

which implies $\omega = \frac{-dx}{2y}$ as $\pi_*(\Omega^1_{\mathcal{E}/U})$ is a rank 1 A-module.

(We don't have such a nice characterisation of the Weierstrass equation if 2 is not invertible on S.) □

Definition 3.3.4. For S a scheme, $\alpha, \beta \in \Gamma(S, O_S)$, let $E(\alpha, \beta)$ be the subscheme of $\mathbf{P}^2_S$ defined by

$$Y^2Z + \alpha XYZ + \beta YZ^2 = X^3 + \beta X^2Z,$$

and let

$$\Delta(\alpha, \beta) = \beta^3(\alpha^4 - \alpha^3 + 8\alpha^2\beta - 36\alpha\beta + 16\beta^2 + 27\beta)$$

be its discriminant. If $\Delta(\alpha, \beta) \in \Gamma(S, O_S)^\times$, this is an elliptic curve over S.

Remark. If $\Delta(\alpha, \beta) \in \Gamma(s, O_S)^\times$, then the elliptic curve $E(\alpha, \beta)$ has j-invariant

$$j(\alpha, \beta) = \frac{\big((\alpha + 4\beta)^2 - 24\alpha\beta\big)^3}{\Delta}.$$

Note that $P = (0 : 0 : 1) \in E(S)$, and we calculate

$$\begin{aligned}
-P &= (0 : -\beta : 1), \\
2P &= (-\beta : \beta(\alpha - 1) : 1), \\
-2P &= (-\beta : 0 : 1), \\
3P &= (1 - \alpha : \alpha - \beta - 1 : 1), \\
-3P &= (1 - \alpha : (\alpha - 1)^2 : 1), \\
&\dots
\end{aligned}$$

so P does not have order 1, 2 or 3 in any fibre.

Proposition 3.3.5. *For any scheme S, E/S an elliptic curve and $P \in E(S)$ such that $P, 2P, 3P \neq 0$ in any fibre, there exist unique $\alpha, \beta \in \Gamma(S, O_S)$ such that $\Delta(\alpha, \beta) \in \Gamma(S, O_S)^\times$ and a unique isomorphism $E(\alpha, \beta) \cong E$ mapping $(0 : 0 : 1)$ to P.*

Proof. First, assume that E has a Weierstrass equation over S. By a translation $x \mapsto x + s$, $y \mapsto y + t$, we can assume that $P = (0, 0)$, so the Weierstrass equation for E is of the form

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

Since P does not have order 2 in any fibre, the gradient r of tangent line at P is in $\Gamma(S, O_S)$ (exercise), so by replacing y with $y + rx$ we can put the equation into the form

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

for some $a_i \in \Gamma(S, O_S)$. Since P does not have order 3 in any fibre, $(0, 0)$ is not an inflexion point, which implies that $a_2 \in \Gamma(S, O_S)^\times$. (Check that the tangent at $(0, 0)$ intersects the curve in two points which are distinct in every fibre.) So by scaling $x \mapsto u^2x$, $y \mapsto u^3y$ with $u = a_3/a_2$, we can arrange that $a_2 = a_3$. Then $E = E(a_1, a_2)$. This gives an isomorphism to a curve in Tate normal form.

Now consider a general E/S. We know that there exists an affine covering $S = \bigcup_i U_i$, such that $E|_{U_i}$ has a Weierstrass equation over $\Gamma(U_i, O_{U_i})$, so we get $\alpha_i, \beta_i \in \Gamma(U_i, O_S)$ such that $E|_{U_i}, P_{U_i}) \cong (E(\alpha_i, \beta_i), (0,0))$. Since α_i, β_i are unique, the must agree on $U_i \cap U_j$. The sheaf property of O_S then implies that there exist $\alpha, \beta \in \Gamma(S, O_S)$ such that $\mathrm{res}_{U_i}(\alpha) = \alpha_i$ and $\mathrm{res}_{U_i}(\beta) = \beta_i$. Then $(E, P) \cong (E(\alpha, \beta), (0,0))$. □

Remark. The last step used in an essential way the *uniqueness of* (α, β); "local uniqueness gives global existence".

Corollary 3.3.6. *Denote by* Sch *the category of schemes.*

- *The pair*
$$\Big(\operatorname{Spec} \mathbf{Z}[A, B, \Delta(A,B)^{-1}], \big(E(A,B), (0:0:1)\big)\Big)$$
represents the functor $\mathcal{F} : \mathrm{Sch}^{\mathrm{opp}} \longrightarrow \underline{\mathrm{Set}}$,

$S \mapsto \big\{$ *eq. classes of pairs* (E, P), $\quad E/S$ *elliptic curve,* $P \in E(S)$ *a point not of order* $1, 2, 3$ *in any fibre* $\big\}$.

- *The pair*
$$\Big(\operatorname{Spec} \mathbf{Z}[B, \Delta(1+B,B)^{-1}], \big(E(1+B,B), (0:0:1)\big)\Big)$$
represents the functor
$$S \mapsto \big\{(E, P), \quad E/S \text{ elliptic curve, } P \text{ a point of exact order } 5 \text{ in every fibre}\big\}.$$

Proof. For (i), we need to check that if $\mathcal{E} \to S$ is an elliptic curve and $P \in \mathcal{E}(S)$ a point not of order $1, 2, 3$ in any fibre, then there exists a unique homomorphism
$$S \to \operatorname{Spec} \mathbf{Z}[A, B, \Delta(A,B)^{-1}]$$
such that $\mathcal{E} = S \times_{\operatorname{Spec} \mathbf{Z}[A,B,\Delta(A,B)^{-1}]} E(A,B)$ and P is the pullback of $(0:0:1)$. But this is a restatement of Proposition 3.3.5. For (ii), just equate $3P = -2P$:
$$3P = (1 - A, A - B - 1),$$
$$-2P = (-B, 0).$$
□

Note. Note that $\Delta(1+B, B) = B^5(B^2 + 11B - 1)$, and the dicriminant of the quadratic is 5^3.

1st attempt: define $Y_1(5)_{\mathbf{Z}}$ to be $\operatorname{Spec} \mathbf{Z}[B, \Delta(1+B,B)^{-1}]$.

Problem. There is no $\beta \in \overline{\mathbf{F}}_5$ such that $\Delta(1+\beta, \beta) \neq 0$ and $j(1+\beta, \beta) = 0$. Hence the map
$$j : Y_1(5)_{\mathbf{Z}} \longrightarrow \mathbb{A}^1_{\mathbf{Z}}$$
is not surjective, since it fails to hit 0 in the fibre above $\mathbf{F}_5$. (The reason is that there exist elliptic curves over $\mathbf{F}_5$ with j-invariant 0, but these do not have a point of exact order 5; they are *supersingular*.) However, we want the j-invariant be surjective, so the answer seems to be to restrict the domain and codomain of j.

Definition 3.3.7. We set
$$Y_1(5)_{\mathbf{Z}[\frac{1}{5}]} = \operatorname{Spec} \mathbf{Z}\big[\tfrac{1}{5}, B, \Delta(1+B,B)^{-1}\big].$$
This represents the same functor as before in the category of $\mathbf{Z}\big[\frac{1}{5}\big]$-schemes.

More generally, we have the following definition:

Definition. (continued) For $N \geq 4$, let $\mathcal{Y}_N$ be the closed subscheme of $\mathcal{Y} = \operatorname{Spec} \mathbf{Z}[A, B, \Delta(A,B)^{-1}]$ cut out by the equation $N \cdot (0:0:1) = (0:1:0)$, and let
$$Y_1(N)_{\mathbf{Z}[\frac{1}{N}]} = \left(\mathcal{Y}_N - \bigcup_{d|N, 4 \leq d < N} \mathcal{Y}_d\right) \times_{\operatorname{Spec} \mathbf{Z}} \operatorname{Spec} \mathbf{Z}\big[\frac{1}{N}\big].$$

Remark. Note that $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}$ has a universal elliptic curve $\mathcal{E}$ over it by restricting $E(\alpha,\beta)/\mathcal{Y}$, and this has a point $(0:0:1)$ which is of precise order N. The triple $\Big(Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}, \mathcal{E}, (0:0:1)\Big)$ represents the functor

$$S \mapsto \big\{\text{elliptic curves } E/S \text{ with point of exact order } N\big\}$$

on the category of $\mathbf{Z}\big[\frac{1}{N}\big]$-schemes.

Remark. $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}$ is equipped with an action of the group $(\mathbf{Z}/N\mathbf{Z})^\times$. More precisely, the group acts on the scheme by $\mathrm{Spec}\,\big(\mathbf{Z}[\frac{1}{5}]\big)$-automorphisms. This observation will be important later.

There are two natural questions:

(1) What does $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}$ look like— Is it non-singular?
(2) There exists a bijection of *sets* between $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}(\mathbf{C})$ and $\Gamma_1(N)\backslash\mathcal{H}$. Is it a map of algebraic varieties over $\mathbf{C}$?

Lecture 5

3.4. **Smoothness.**

Definition 3.4.1. A morphism of schemes $\phi : X \to Y$ is *smooth* if it is locally of finite presentation, flat, and for every point $y \in Y$, the fibre $\phi^{-1}(y)$ is a smooth variety over $k(y)$.

Remark. Our definition of elliptic curves over S requires that $\mathcal{S} \to S$ be a smooth morphism.

Lemma 3.4.2. *(1) The composition of smooth morphisms is smooth.*
(2) If E/S is an elliptic curve and $N \geq$ is invertible on S, then $[N] : E \to E$ is smooth.

Proof. (i) is standard (see EGA – follow the trail of references from Wikipedia).
(ii) The morphism $[N]$ multiplies a global differential by N, so it induces an isomorphism of tangent space. In other words, it is an étale morphism, and étale morphisms are smooth. □

Proposition 3.4.3. *(Functorial criterion for smoothness) Let $X \to \mathrm{Spec}(R)$ be a scheme of finite type over R, where R is noetherian. Then the map $X \to \mathrm{Spec}(R)$ is a smooth morphism if and only if it is formally smooth, i.e. for any local R-algebra A and a nilpotent ideal $I \subset A$, the map*

$$\mathrm{Hom}_{\mathrm{Sch}/R}(\mathrm{Spec}\, A, X) \to \mathrm{Hom}_{\mathrm{Sch}/R}(\mathrm{Spec}\, A_0, X)$$

is surjective, where $A_0 = A/I$.

Proof. See Stacks Project §36.9. □

Remark. If we replace surjective with bijective, then we get th enotion of 'formally étale'.

Theorem 3.4.4. *$Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}$ is smooth over $\mathbf{Z}\big[\frac{1}{N}\big]$.*

Proof. Let A be a local $\mathbf{Z}\big[\frac{1}{N}\big]$–algebra, and let $I \subset A$ be nilpotent. Let $(E_0, P_0) \in Y_1(N)(A_0)$. The ring A_0 is local, so E_0 has a Weierstrass equation over $\mathrm{Spec}(A_0)$. Lift coefficients arbitrarily to A to get E/A lifiting E_0; note that $\Delta(E) \in A^\times$ since its image in A_0 is in $A_0^\times$.

Can we lift P_0 to an N-torsion point of E, i.e. is $E[N]$ smooth? Yes, since $[N] : E \to E$ is smooth, and a composition of smooth mrophisms is smooth. (We apply this to $[N]$ composed with the structure map $E \to \mathrm{Spec}\, A$.) Hence (E_0, P_0) lifts to (E, P), and we are done. □

Note. The schemes $\mathcal{Y}_N/\mathbf{Z}$ are very rarely smooth; it was true for $N = 5$ essentially by accident.

3.5. **Quotients and $Y_0(N)$.**

Proposition 3.5.1. *Let X be a quasiprojective S-scheme (for some base scheme S), and let G be a finite group acting on X by S-automorphisms. Then there exists a unique S-scheme X/G and a unique morphism $X \to X/G$ representing the functor*

$$Y \mapsto \big(\textit{morphisms of } S\textit{-schemes } X \to Y \textit{ commuting the the } G\textit{-action}\big).$$

Here, we consider Y as a G-module with the trivial action.

Remark. Explicitly, this means that given any scheme Y and $\alpha : X \to Y$ a morphism of S-schemes such that $\alpha(g.x) = \alpha(x)$ for all $x \in X$ and $g \in G$, there exists a unique S-morphism $f : X/G \to Y$ such that α factors as

$$X \longrightarrow X/G \xrightarrow{f} Y.$$

Proof. Uniqueness is obvious (representing a functor). Existence: for $X = \mathrm{Spec}(A)$ affine, we can take $\mathrm{Spec}(A^G)$ and the map $\mathrm{Spec}(A) \to \mathrm{Spec}(A^G)$ induced by the inclusion $A^G \hookrightarrow A$, and one can show that these patch nicely. (One needs quasiprojectiveness and finiteness of G here.) □

Remark. It is clear from the construction that points on X in the same G-orbit map to the same point in X/G ,which justifies the notation.

Remember from last lecture that $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}$ is equipped with an action of the group $(\mathbf{Z}/N)^\times$.

Definition 3.5.2. For $N \geq 4$, let $Y_0(N) = Y_1(N)/(\mathbf{Z}/N)^\times$ (as a $\mathbf{Z}\left[\frac{1}{N}\right]$-scheme).

Note. The $\mathbf{C}$-points of this are $\Gamma_0(N)\backslash\mathcal{H}$.

Proposition 3.5.3. *$Y_0(N)$ is smooth over $\mathbf{Z}\left[\frac{1}{N}\right]$*

Does this definition agree with the explicit model of $X_0(N)$ that we constructed in Section 2.2?

Proposition 3.5.4. *$Y_0(N)$ agrees with our earlier construction as the unique smooth projective curve over $\mathbf{Q}$ with function field $\mathbf{Q}(X)[Y]/\Phi_N(X,Y)$, where $\Phi_N(j(z),Y)$ is the minimal polynomial of $j(Nz)$ over $\mathbf{Q}(j(z))$.*

Proof. (Sketch) Let $G = (\mathbf{Z}/N)^\times$. It suffices to show that $\mathbf{Q}(j(z), j(Nz)) \subseteq \mathbf{Q}(Y_1(N))^G$, i.e. that the functions $j(\tau)$ and $\ell(\tau) := j(N\tau)$ lie in $\mathbf{Q}(Y_1(N))^G$. First note that by definition, $j(\tau) = j(E_\tau)$, where $E_\tau \cong \mathbf{C}/(\mathbf{Z} + \tau\mathbf{Z})$, and

$$j(N\tau) = j(E_{N\tau}) = j(E/\langle 1/N\rangle),$$

since $\mathbf{C}/(\mathbf{Z} + \tau N\mathbf{Z}) \cong E_\tau/\langle 1/N\rangle$.

Now j and ℓ clearly define complex-valued functions on the complex points of $Y_1(N)$, and we know from Proposition 3.1.4 that the complex points on $Y_1(N)$ are in bijection with pairs $(E_\tau, \frac{1}{N})$. We deduce from the previous discussion that

$$j\left((E_\tau, 1/N)\right) = j(\tau) \qquad \text{and} \qquad \ell\left((E_\tau, 1/N\right) = j(E_\tau/\langle 1/N\rangle).$$

Now an element $g \in G = (\mathbf{Z}/N\mathbf{Z})^\times$ acts on $Y_1(N)$ by $(E, P) \mapsto (E, g.P)$. As $\langle P\rangle = \langle g.P\rangle$, it is clear that $j(\tau)$ and $j(N\tau)$ are invariant under the action of G. □

Question. Is $Y_0(N)$ a moduli space for elliptic curves with a subgroup of order N?

Proposition 3.5.5. *Let S be a $\mathbf{Z}\left[\frac{1}{N}\right]$-scheme. There is a natural map*

$$\left\{\begin{array}{c} \text{iso. classes of pairs } (E,C)\text{:} \\ E/S \text{ ell. curve, } C \subset E \\ \text{subgroup-scheme, étale loc.} \\ \text{isom. to } \mathbf{Z}/N \end{array}\right\} \longrightarrow Y_0(N)(S)$$

Proof. We define the map as follows: let (E, C) be an element of LHS. Then there exists $S' \to S$ étale and $P \in E(S')$ such that $C = \langle P\rangle$, and this gives a point of $Y_1(N)(S')$. Changing P changes this by an elements of $G = (\mathbf{Z}/N)^\times$, so we get a G-orbit of elements of $Y_1(N)(S')$. By a scary lemma (étale descent of morphisms) this gives an S-point of $Y_0(N)$. Thus we have a well-defined map

$$\iota_S : \{(E, C)/S\} \longrightarrow Y_0(N)(S).$$

□

Problem. The map ι_S is in general neither injective nor surjective.

Injectivity. If L/K is a finite field extension, then $Y_0(N)(K) \to Y_0(N)(L)$ is obviously injective, but

$$\big((E, C)/K\big) \longrightarrow \big((E, C)/L\big)$$

is not injective, as the following example shows.

Example. Let E be an elliptic curve over $\mathbf{Q}$. Let K be a quadratic extension of $\mathbf{Q}$, and denote by E' the twist of E be the corresponding quadratic character, so E and E' are not isomorphic over $\mathbf{Q}$, but they become isomorphic over K. Let $\varphi : E \to E'$ be an isomorphism over K. It is then easy to check that if $\sigma \in \mathrm{Gal}(K/\mathbf{Q})$ is non-trivial, then $\sigma \circ \varphi = -\varphi$.

Let $C \subset E(\overline{\mathbf{Q}})$ be a subgroup of order N, and assume that C is defined over $\mathbf{Q}$ (which means that any $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ permutes the elements of C). Then $C' = \varphi(C)$ is also defined over $\mathbf{Q}$: if $x \in C$ and $\tau \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, then

$$\tau(\varphi.x) = \varphi^\tau.\tau(x) = -\varphi(\tau.x) = \varphi(-\tau.x) \in \varphi(C).$$

Hence (E, C) and (E', C') are isomorphic over K but not over $\mathbf{Q}$.

However, if k is a field, then one can check that the image $(\iota_{\mathrm{Spec}(k)})$ is the set of pairs (E, C) defined over k modulo isomorphisms over $\bar{k}$.

Surjectivity. One can show that for k a field, ι_k is surjective. (This is fairly hard, c.f. Proposition VI.3.2 of Deligne-Rapoport.) However, for a non-field S, surjectivity can also fail. For example, if $S = Y_0(N)$, then in general there is no elliptic curve over S corresponding to the identity homomorphism

$$[S \to Y_0(N)] \in Y_0(N)(S) = \mathrm{Hom}_{S-\mathrm{Sch}}(S, Y_0(N)).$$

(One can try to use $E/(\mathbf{Z}/N)^\times$ where E is the universal elliptic curve over $Y_1(N)$, but fibres over points with non-trivial stabilizers might not be elliptic curves!)

Is there a conceptual way of understanding the failure of $Y_0(N)$ to classify elliptic curves with level structure? We will investigate this question in the next section.

3.6. **General modular curves.** (following Katz-Mazur)

Definition 3.6.1. Let R be a ring.

(1) Let $\underline{\mathrm{Ell}/R}$ be the following category:

- objects are diagrams $\begin{array}{c} \mathrm{E} \\ \downarrow \\ \mathrm{S} \end{array}$, where S is some R-scheme and E is an elliptic curve over S;
- morphisms are squares

$$\begin{array}{ccc} E & \longrightarrow & E' \\ \downarrow & & \downarrow \\ S & \longrightarrow & T \end{array}$$

where $E \cong E' \times_T S$.

(2) A *moduli problem for elliptic curves over* R is a contravariant functor $\mathcal{P} : \underline{\mathrm{Ell}/R} \to \underline{\mathrm{Set}}$.

(3) If $\mathcal{P}$ is a moduli problem, let $\tilde{\mathcal{P}} : \underline{\mathrm{Sch}/R} \to \underline{\mathrm{Set}}$ be the functor

$$\tilde{\mathcal{P}} : S \mapsto \big(\text{pairs } (E, \alpha),\ E/S \text{ elliptic curve},\ \alpha \in \mathcal{P}(E/S)\big).$$

Example. Here are some examples of moduli problems for elliptic curves over R:

- $(E \to S) \mapsto \{\text{points on } E(S) \text{ of exact order } N\}$;
- $(E \to S) \mapsto \{\text{subgroups of } E(S) \text{ of order } N \text{ in every fibre}\}$.

Note that these functors really are contravariant: for example, if T is an S-scheme and $P = (S \to E) \in E(S)[N]$, then we can basechange P to $T \to (E \times_S T) \in (E \times_S T)(T)[N]$.

Then the associated functors $\underline{\mathrm{Sch}/S} \to \underline{\mathrm{Set}}$ are precisely the ones that we studied in the previous chapters.

Fact 3.6.2. *Fix N and a subgroup $H \subset \mathrm{GL}_2(\mathbf{Z}/N)$. Then there exists a moduli problem $\mathcal{P}_H$ on $\underline{\mathrm{Ell}/\mathbf{Z}\left[\frac{1}{N}\right]}$ such that if $\bar{k}$ is algebraically closed, $E/\bar{k} \in \mathrm{Ob}(\underline{\mathrm{Ell}/\mathbf{Z}\left[\frac{1}{N}\right]})$, then*

$$\mathcal{P}_H(E/\bar{k}) = \{H\text{-orbits of isomorphisms } (\mathbf{Z}/N)^2 \xrightarrow{\cong} E[N]\}.$$

For $H = \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\rangle$, *this is* $\Gamma(N)$,

$$E/S \mapsto \big(\text{pairs of sections } P, Q \in E[S] \text{ generating } E[N] \text{ in every fibre}\big).$$

For $H = \left\{ \begin{pmatrix} \star & \star \\ 0 & 1 \end{pmatrix} \right\}$, *this is* $\Gamma_1(N)$; *for* $H = \left\{ \begin{pmatrix} \star & \star \\ 0 & \star \end{pmatrix} \right\}$, *this is* $\Gamma_0(N)$.

Remark. If k is a field, then E/k is the image of $\mathcal{P}_H(E/k)$ in $\mathcal{P}_H(E/\bar{k})$ is

$$\{H\text{-orbits of bases of } E[N](\bar{k}) \text{ in which image of } \mathrm{Gal}(\bar{k}/k) \text{ lands in } H\}.$$

Definition. A moduli problem $\mathcal{P}$ is *representable* if $\mathcal{P}$ is representable as a functor. In other words, $\mathcal{P}$ is a representable moduli problem if there exists an elliptic curve $\mathcal{E}$ over some scheme S and $\alpha \in \mathcal{P}(\mathcal{E}/S)$ such that for any $(E \to T)$ and any $\beta \in \mathcal{P}(E/T)$ there exists a unique morphism $T \to S$ such that $E \cong \mathcal{E} \times_S T$ and $\beta = \mathcal{P}(\pi)(\alpha)$, where π is the stucture map $\mathcal{E} \times_S T \to \mathcal{E}$.

Example. We saw in the proof of Corollary 3.3.6 that the functor

$$(E \to S) \mapsto \{\text{points on } E(S) \text{ which are of exact order 5 in every fibre}\}$$

is representable by $(E(1+B, B), (0:0:1))$. This was the crucial step in proving that the functor

$$S \mapsto \{\text{iso. classes } (E, P)\text{: } E \text{ an elliptic curve } /S,\ P \text{ of exact order 5 in every fibre}\}$$

is representable. As the following result shows, this is a general principle.

Proposition 3.6.3. *If $\mathcal{P}$ is representable on $\underline{\mathrm{Ell}/R}$, then $\tilde{\mathcal{P}}$ is representable on $\underline{\mathrm{Sch}/R}$.*

Proof. If $(E/S, \alpha)$ represents $\mathcal{P}$, then one can check that $(S, (E, \alpha))$ represents $\tilde{\mathcal{P}}$. □

We would like to quantify when a moduli problem is representable. For this, we introduce a much weaker concept, relative representability.

Definition. A moduli problem $\mathcal{P}$ is *relatively representable* if, for every $E/S \in \mathrm{Ob}(\underline{\mathrm{Ell}/R})$, the functor $\underline{\mathrm{Sch}/S} \to \underline{\mathrm{Set}}$, $T \mapsto \mathcal{P}(E \times_S T/T)$ is representable.

Proposition 3.6.4. *Let $N \geq 1$, and let $H \subset \mathrm{GL}_2(\mathbf{Z}/N)$. Then the moduli problem $\mathcal{P}_H$ is relatively representable and étale over $\underline{\mathrm{Ell}/\mathbf{Z}\left[\frac{1}{N}\right]}$. In other words, for all $E/S \in \mathrm{Ob}(\underline{\mathrm{Ell}/\mathbf{Z}\left[\frac{1}{N}\right]})$, the functor $T \mapsto \mathcal{P}_H(E \times_S T)$ is represented by an étale S-scheme.*

Proof. For $H = \{1\}$ and E and elliptic curve over S, we can find an explicit S-scheme representing $\mathcal{P}_H$ on $\underline{\mathrm{Sch}/S}$; it is an open subscheme Z of $E[N] \times_S E[N]$ given by a condition on the Weil pairing in every fibre: the condition is that the Weil pairing of two sections is a primitive Nth root of unity in every fibre. To see that Z has the required properties, let T be an S-scheme, and let $P, Q \in E \times_S T$ be generating $(E \times_S T)[N]$ in every fibre. Define $\alpha : T \to Z$ as the composition

$$\alpha = (\pi \times \pi) \circ (P, Q),$$

where

$$(P, Q) : T \to (E \times_S T)[N] \times (E \times_S T)[N]$$

and $\pi : E \times_S T \to E$ is the natural projection. Then it is easily seen that α has the required property.

For general H just take the quotient of Z by H. □

So what prevents a relatively representable moduli problem from being representable?

Definition 3.6.5. $\mathcal{P}$ is *rigid* if for all $E/S \in \mathrm{Ob}(\underline{\mathrm{Ell}/R})$, $\mathrm{Aut}(E/S)$ acts on $\mathcal{P}(E/S)$ without fixed points.

Example. The moduli problem

$$(E \to S) \mapsto \{\text{subgroups } C \text{ of } E(S) \text{ which have exact order } N \text{ in every fibre}\}$$

is not rigid: the element $[-1] \in \mathrm{Aut}(E/S)$ sends such a subgroup to itself.

We also know from the previous section together with Proposition 3.6.3 that this moduli problem is not representable. The following result of Katz-Mazur that in the case of a relatively representable moduli problem, these kinds obstructions are the only ones:

Theorem 3.6.6. *(Katz-Mazur) $\mathcal{P}$ is representable if and only if it is relatively representable and rigid.*

Proof. (Sketch) Start from two basic moduli problems:

- 'naive level $\Gamma(3)$' over $\mathbf{Z}\left[\frac{1}{3}\right]$;
- 'Legendre moduli problem' ($\Gamma(2)$ with choice of differential) over $\mathbf{Z}\left[\frac{1}{2}\right]$.

Both have group actions ($\mathrm{GL}_2(\mathbf{F}_3)$ and $\mathrm{GL}_2(\mathbf{F}_2)\times\{\pm1\}$). Given $\mathcal{P}$ relatively representable and rigid, construct one object by taking $\mathcal{E}/Y(3)$ – relative representability gives us a scheme over $Y(3)$ – and this has a $\mathrm{GL}_2(\mathbf{F}_3)$-action. Take invariants (this is OK since $\mathcal{P}$ is rigid), so we get an object $\mathcal{E}/S$ represnting $\mathcal{P}$ on $\underline{\mathrm{Ell}/R\left[\frac{1}{3}\right]}$.

Legendre gives an object over $R\left[\frac{1}{2}\right]$ similarly. By rigidity these agree over $R\left[\frac{1}{6}\right]$, so we get a representing object over R. □

It turns out that one can determine precisely for which subgroups of $\mathrm{SL}_2(\mathbf{Z}/N\mathbf{Z})$ the moduli problem $\mathcal{P}_H$ is rigid:

Proposition 3.6.7. *$\mathcal{P}_H$ is rigid on $\underline{\mathrm{Ell}/R\left[\frac{1}{6}\right]}$ if and only if the preimage in $\mathrm{SL}_2(\mathbf{Z})$ of $H\cap\mathrm{SL}_2(\mathbf{Z}/N)$ contains no elements of finite order (i.e. has no elliptic points and does not contain -1).*

Proof. (Sketch) Over $\mathbf{C}$ this is routine. To prove the statement in general it suffices to check it on objects $E/\bar{k}$, where $\bar{k}$ is algebraically closed and not too large. If $\bar{k}$ has characteristic 0, we can embed it into $\mathbf{C}$.

One can show: if k has finite characteristic ≥ 5, E/k is an elliptic curve, $\phi\in\mathrm{Aut}(E)$, then the pair (E,ϕ) lifts to characteristic 0. (This is shown somewhere in chapter VI of Deligne-Rapoport.) □

Corollary 3.6.8. *If H satisfies the hypotheses of Proposition 3.6.7, there is a scheme $Y=Y_{\mathcal{P}_H}$, an elliptic curve $\mathcal{E}/Y$ and an $\alpha\in\mathcal{P}_H(\mathcal{E}/Y)$ representing the functor $\mathcal{P}_H$. One can check that $Y_{\mathcal{P}_H}$ is smooth over $\mathbf{Z}[1/N]$.*

Remarks.

(1) If H does not satisfy the hypotheses of Proposition 3.6.7, then the functors $\mathcal{P}_H$ and $\widetilde{\mathcal{P}}_H$ are not representable, but one can show that there is a scheme $Y_{\mathcal{P}_H}$ over $\mathbf{Z}[1/N]$ that is "the best possible approximation" to representing the functor $\widetilde{\mathcal{P}}_H$: there are maps
$$\widetilde{\mathcal{P}}_H(S)\to Y_{\mathcal{P}_H}(S)$$
which are surjective for S a field, and bijective if S is algebraically closed, as in the special case of $Y_0(N)$ (Proposition 3.5.5). These schemes $Y_{\mathcal{P}_H}$ for non-rigid $\mathcal{P}_H$ are sometimes called *coarse moduli spaces* (while the $Y_{\mathcal{P}_H}$ for rigid H, which do represent functors, are sometimes called *fine moduli spaces*).

(2) The complex points of $Y_{\mathcal{P}_H}$ are closely related to $\Gamma\backslash\mathcal{H}$, where Γ is the preimage of H in $\mathrm{SL}_2(\mathbf{Z})$, but they are not always equal. The correct statement is that $Y_{\mathcal{P}_H}(\mathbf{C})$ is a possibly disconnected Riemann surface, whose components biject with the quotient $(\mathbf{Z}/N)^\times/\det(H)$, and the component corresponding to the coset of $1\in(\mathbf{Z}/N)^\times$ is $\Gamma\backslash\mathcal{H}$.

If you are happy with adèles you can write this more intrisically as
$$Y_{\mathcal{P}_H}(\mathbf{C})=\mathrm{GL}_2(\mathbf{Q})\backslash\mathrm{GL}_2(\mathbf{A})/\left(\mathbf{R}_{>0}\cdot\mathrm{SO}_2(\mathbf{R})\cdot U\right)$$
where U is the preimage of H in $\mathrm{GL}_2(\widehat{\mathbf{Z}})$ (an open subgroup of $\mathrm{GL}_2(\mathbf{A}_{\mathrm{fin}})$).

4. Leftovers

4.1. **Katz modular forms.** Recall that we defined, for E/S an elliptic curve, $\underline{\omega}_{E/S}=\pi_*(\Omega^1_{E/S})$. Write ω_{Katz} for the line bundle from Definition 1.5.1.

Proposition 4.1.1. *If $S=Y_{\mathcal{P}_H}$ for some H as before, E/S universal elliptic curve, then $\underline{\omega}_{E/S}$ is the Katz sheaf ω_1 from Chapter 2.*

Proof. Exercise (You need to show that both line bundles have the same pullbacks to $\mathcal{H}$ and the actions of Γ agree.) □

Definition 4.1.2. For Γ a torsion free congruence subgroup of level N, R a $\mathbf{Z}\left[\frac{1}{N}\right]$-algebra, define
$$\mathrm{KM}_k(\Gamma,R)=H^0\left(Y(\Gamma)\times R,\underline{\omega}^k_{E/Y(\Gamma)}\right)$$
(an R-module).

Concretely: a Katz modular form of weight k over R is a rule attaching to each triple $(E/S,\alpha,\omega)$ (S an R-scheme, E/S elliptic curve, $\alpha\in\mathcal{P}_H(E/S)$, ω a basis of $\Gamma(E,\underline{\omega}_{E/S})$) an element of $\Gamma(S,O_S)$ such that

- it is compatible with base change in S,
- it is homogeneous of weight k in ω.

(Compare with Katz "p-adic properties of modular schemes and modular forms, Springer LNM 330.)

Fun thing: over $R = \mathbf{Z}\left[\frac{1}{6}\right]$, for an elliptic curve E/R and $\omega \in \Omega^1$, there exists a unique Weierstrass equation such that $\omega = \frac{dX}{y}$, and E_4 (resp. E_6) are the maps which send (E, ω) to the a_4- (resp. a_6-)coefficient of this equation.

4.2. **Cusps and the Tate curve.** Consider the ring

$$\mathbf{Z}((q)) = \left\{ \sum_{n=-N}^{\infty} a_n q^n : a_n \in \mathbf{Z} \right\}.$$

We will define an elliptic curve over this ring together with a differential such that evaluating at this pair gives the q-expansion of a Katz modular form.

Definition 4.2.1. Tate(q) is the elliptic curve

$$y^2 + xy = x^3 + a_4 x + a_6,$$

where

$$a_4 = -\sum_{n \geq 1} \frac{5n^3 q^n}{1 - q^n},$$

$$a_6 = -\sum_{n \geq 1} \frac{1}{12} \cdot \frac{(7n^5 + 5n^3) q^n}{1 - q^n}.$$

Note that $a_4, a_6 \in \mathbf{Z}[[q]]$.

We find that the dicriminant of Tate(q) is exactly the q-expansion of Δ (weight 12 cusp form) in $q + q^2\mathbf{Z}[[q]] \subset \mathbf{Z}((q))^\times$. Hence Tate$(q)$ is an elliptic curve:

$$\begin{aligned} \text{Tate}(q) &= \text{“}q\text{-expansion of } \mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau)\text{”} \\ &= \mathbf{C}^\times / q^{\mathbf{Z}}. \end{aligned}$$

Proposition 4.2.2. *If $\tau \in \mathcal{H}$, then the series defining* Tate(1) *converges at $q = e^{2\pi i \tau}$ and defines a curve $\cong \mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau)$.*

Remark. Convergence is easy, and we chack that $j(\text{Tate}(q))$ is the q-expansion of $j(\tau)$.

Proposition 4.2.3. *There exist series $X(u,q)$, $Y(u,q)$ in $\mathbf{Z}[u, u^{-1}(1-u)^{-1}]$ such that*

$$\big(X(u,q), Y(u,q)\big) \oplus \big(X(v,q), Y(v,q)\big) = \big(X(uv,q), Y(uv,q)\big). \tag{2}$$

Here, $\oplus$ denotes the group law on Tate(q). *(We interpret $\big(X(u,q), Y(u,q)\big)$ as ∞ if $u = 1$.)*

Proof. Take

$$X(u,q) = \frac{u}{(1-u)^2} + \sum_{d \geq 1} \left(\sum_{m \mid d} m(u^m + u^{-m} - 2) \right) q^d,$$

$$Y(u,q) = \frac{u^2}{(1-u)^3} + \sum_{d \geq 1} \left(\sum_{m \mid d} \frac{m(m-1)}{2} u^m - \frac{m(m+1)}{2} u^{-m} + m \right) q^d.$$

Sneaky part: there exists a straightforward change of coordinates from Tate(q) to

$$y^2 = x^3 - g_4(\tau)x - g_6(\tau)$$

which is $\mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau)$ via $\big(\wp(z,\tau), \wp'(z,\tau)\big)$. X and Y are just $\wp$ and $\wp'$ as power series in $u = e^{2\pi i z}$ and $q = e^{2\pi i \tau}$. So the identity (2) holds for all u, q in an open subset of $\mathbf{C} \times \mathbf{C}$, so it holds as an identity of power series. □

Remark. In other words, the map

$$\mathbb{G}_{m/\mathbf{Z}((q))} \longrightarrow \mathrm{Tate}(q)$$

given by $u \mapsto (X(u,q), Y(u,q))$ is a group homomorphism, and it can be shown to factor through $q^{\mathbf{Z}}$. (This map does not actually exist, since the power series $X(u,q)$ and $Y(u,q)$ do not converge for general $u \in \mathbf{Z}(q)$), but it explains why the Tate curve is often written as $\mathbb{G}_m/q^{\mathbf{Z}}$.)

Proposition 4.2.4. *Cusps of $Y_{\mathcal{P}_H}$ correspond to*

$$\big\{\mathcal{P}_H\text{-level structures on } \mathrm{Tate}(q) \text{ over } \mathbf{Z}[[q^{\frac{1}{N}}, \zeta_N]] \text{ mod. automorphisms } q^{\frac{1}{N}} \to \zeta_N^a q^{\pm\frac{1}{N}}\big\}.$$

Note. We thus get an action of $\mathrm{Gal}(\mathbf{Q}(\mu_N)/\mathbf{Q})$ on the set of cusps.

Example. $Y_1(5)$: the points of order 5 on $\mathrm{Tate}(q)$ over $\mathbf{Z}\big[\frac{1}{5}, \zeta_5\big][[q^{\frac{1}{5}}]]$ are the images of $q^{\frac{a}{5}}\zeta_5^b$, with $a, b \in (\mathbf{Z}/5)^2 - \{(0,0)\}$.

These do not all give distinct cusps, because we need to keep track of the automorphisms. We find that there are 4 cusps

$$\{\zeta_5, \zeta_5^4\}, \quad \{\zeta_5^2, \zeta_5^3\}, \quad \{q^{\pm 1/5}\zeta_5^a : a \in \mathbf{Z}/5\mathbf{Z}\}, \quad \{q^{\pm 2/5}\zeta_5^a : a \in \mathbf{Z}/5\mathbf{Z}\}.$$

The action of $\mathrm{Gal}(\mathbf{Q}(\mu_5)/\mathbf{Q})$ on the cusps is now easy to see: it factors through $\mathrm{Gal}(\mathbf{Q}(\mu_5)^+/\mathbf{Q}) \cong C_2$, and the nontrivial element swaps $\{\zeta_5, \zeta_5^4\}$ with $\{\zeta_5^2, \zeta_5^3\}$ and fixes the other two cusps.