

# MODULAR CURVES

DAVID LOEFFLER

## CONTENTS

Waffle	1
0.1. Recap of modular forms	1
1. Modular curves as Riemann surfaces	2
1.1. Modular curves as topological spaces	2
1.2. Riemann surfaces: recap	2
1.3. Genus, ramification, Riemann-Hurwitz	3
1.4. Sheaves and Riemann-Roch	5
1.5. The Katz sheaf	6
2. Modular curves as algebraic curves	8
2.1. Modular curves over $\mathbf{C}$	8
2.2. Descending the base field	8
3. Modular curves as moduli spaces	11
3.1. Lattices and level structures	11
3.2. Moduli spaces and representable functors	11
Some comments about the questions on Question sheet 1	11
3.3. Elliptic curves over general base schemes	12
3.4. Smoothness	15
3.5. A complex-analytic digression	15
3.6. Quotients and $Y_0(N)$	17
3.7. General modular curves	18
3.8. General level structures	19
4. Leftovers	20
4.1. Katz modular forms	20
4.2. Cusps and the Tate curve	20

## WAFFLE

Lecture 1

Let  $\mathcal{H} = \{z \in \mathbf{C} \mid \Im(z) > 0\}$ , so  $\mathrm{SL}_2(\mathbf{R})$  is acting on  $\mathcal{H}$  by Möbius transformations. Let  $\Gamma$  be a subgroup of  $\mathrm{SL}_2(\mathbf{Z})$  of finite index, and let  $Y(\Gamma) = \Gamma \backslash \mathcal{H}$ . We will equip  $Y(\Gamma)$  with various interesting structures.

**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  $F : \mathcal{H} \rightarrow \mathbf{C}$  such that

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

and a growth condition on the boundary. 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  $q$ -expansion

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

where  $a_n \in \mathbf{C} \quad \forall n$  and  $q = e^{\frac{2\pi iz}{h}}$ . 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  $\pi : \mathcal{H} \rightarrow Y(\Gamma)$  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  $\tau_1$  and  $\tau_2$ , 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  $\tau_i$  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

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

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  $\Gamma$ .

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 neighbourhoods of  $\infty$  are the sets  $\{z \mid \Im(z) > R\}$  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  $\Gamma$  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  $\gamma_1, \dots, \gamma_n$  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  $\phi_i : X_i \rightarrow U_i$  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  $f : X \rightarrow \mathbf{C}$  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  $\Gamma$  in  $\mathrm{PSL}_2(\mathbf{Z}) = \Gamma/(\Gamma \cap \{\pm 1\})$ .

*Note.* If  $P$  is elliptic for  $\Gamma$ , 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  $\rho = e^{\frac{2\pi i}{3}}$ . So the set of elliptic points for any  $\Gamma$  is finite.

If  $P$  is not elliptic and not a cusp, then we can easily find a chart around  $P$ : take  $\tau$  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  $\tau$  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  $U$  containing  $\tau$  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  $\pi$  to  $U$  is a homeomorphism  $\mathrm{Stab}_{\bar{\Gamma}}(\tau) \backslash U \rightarrow V$ .

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  $\delta$  always exists.) This conjugates  $\mathrm{Stab}_{\bar{\Gamma}}(\tau)$  onto a finite group of Möbius transformations fixing 0 and  $\infty$ , 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  $U'' \rightarrow V$  gives a coordinate chart around  $P$ .

Lastly, if  $P$  is a cusp, we argue similarly: choose  $\delta$  mapping  $P$  to  $\infty$ . 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  $z \mapsto e^{2\pi iz/h}$  gives the local coordinate.

It's easy to see that all the above coordinate charts are compatible on overlaps; so we have proved the following: there exist Riemann surface structures (clearly unique) on  $Y(\Gamma)$  and  $X(\Gamma)$  such that  $\pi : \mathcal{H} \rightarrow Y(\Gamma)$  is holomorphic.

**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*

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

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  $\chi(M) = 2 - 2g$ .

**Proposition 1.3.2.** 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 holomorphic map

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

It is bijective (count zeros using contour integration), so (by the Open Mapping Theorem of complex analysis) it has a holomorphic inverse.  $\square$

Lecture 2

**Convention.** All Riemann surfaces are assumed to be connected.

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

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

**Definition 1.3.3.** (1) For  $f : X \rightarrow Y$  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. Note that points where  $e_P(f) > 1$  are isolated.

(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  $X(\Gamma) \rightarrow X(\mathrm{SL}_2(\mathbf{Z}))$  is  $[\mathrm{PSL}_2(\mathbf{Z}) : \bar{\Gamma}]$ .

**Theorem 1.3.4** (Riemann-Hurwitz). For  $f : X \rightarrow Y$  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 1.3.5.** For any  $\Gamma$ , 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  $\varepsilon_2$  (resp.  $\varepsilon_3$ ) is the number of elliptic points of order 2 (resp. order 3) and  $\varepsilon_\infty$  is the number of cusps.

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

- If  $P \in Y(\Gamma)$  is not in the  $\mathrm{SL}_2(\mathbf{Z})$ -orbit of  $i$  or  $\rho$ .

$$\begin{array}{ccc} \mathcal{H} & & \\ \downarrow & \searrow \alpha & \\ Y(\Gamma) & & \\ \downarrow & \swarrow & \\ Y(\mathrm{SL}_2(\mathbf{Z})) & & \end{array}$$

The map  $\alpha$  is unramified at any  $\tau$  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  $Y(\Gamma) \rightarrow Y(\mathrm{SL}_2(\mathbf{Z}))$  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  $\Gamma$ , so  $e_P(f) = 2$ . We use the definition of the degree to count the number of points above  $[i]$ . We have  $N = 1 \cdot \varepsilon_2 + 2 \cdot (\text{number of non-elliptic points of } Y[\Gamma] \text{ above } [i])$ . Hence, the number of non-elliptic points above  $[i]$  is  $(N - \varepsilon_2)/2$ . So

$$\sum_{P \in f^{-1}([i])} (e_P - 1) = \frac{N - \varepsilon_2}{2}.$$

- 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  $N = 12$ ,  $\epsilon_\infty = 2$  ( $[0]$  and  $[\infty]$ ),  $\epsilon_2 = \epsilon_3 = 0$  (exercise, c.f. Diamond and Shurman), so

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

*Exercise.* (1) Verify that  $\epsilon_2 = \epsilon_3 = 0$  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  $\Gamma$  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  $U \rightarrow \mathbf{C}$ . 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  $\mathcal{F} \otimes_{O_X} \{\text{sheaf of meromorphic functions}\}$ ).

**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

$$\begin{aligned} \deg(\mathcal{F} + \mathcal{G}) &= \deg(\mathcal{F}) + \deg(\mathcal{G}), \\ \deg(\mathcal{F}^{-1}) &= -\deg(\mathcal{F}). \end{aligned}$$

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  $\Omega$  is the sheaf of holomorphic differentials on  $X$ .

*Note.* If  $\deg(\mathcal{F}) < 0$ , then  $\mathcal{F}$  has no non-zero global sections, so if  $\deg(\mathcal{F})$  is large, then  $H^0(X, \Omega \otimes \mathcal{F}^{-1}) = 0$ , and we get a formula for  $H^0(X, \mathcal{F})$ . Also, note that  $\dim H^0(X, \Omega) = g(X)$  by taking  $\mathcal{F} = \mathcal{O}_X$  and  $\deg(\Omega) = 2g - 2$  by taking  $\mathcal{F} = \Omega$ .

*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  $\Omega$  the *dualizing sheaf*.

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

**Definition 1.5.1.** Let  $\omega_k$  be the sheaf defined by

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

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

**Theorem 1.5.2.** (i)  $\omega_k$  is invertible;  
(ii)  $\omega_2 = \Omega_X(\text{cusps})$ .

The notation in (ii) is as follows:

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

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

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

*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  $\pi^{-1}(V) = \sqcup_{\gamma \in \bar{\Gamma}} \gamma U$  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

$$\omega_k(V) = \mathcal{O}_X(V) \cdot b.$$

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

$$\pi^{-1}(V) = \coprod_{\gamma \in \bar{\Gamma} / \text{Stab}_{\bar{\Gamma}}(\tau)} \gamma U,$$

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

$$\omega_k(V) = \{f : U \rightarrow \mathbf{C} \text{ holomorphic of wt. } k \text{ and invariant under } \text{Stab}_{\bar{\Gamma}}(\tau)\},$$

while

$$\mathcal{O}_X(V) \{f : U \rightarrow \mathbf{C} \text{ holomorphic of wt. } 0 \text{ and invariant under } \text{Stab}_{\bar{\Gamma}}(\tau)\}.$$

It follows that if  $\text{Stab}_{\bar{\Gamma}}(\tau) = 1$ , or more generally if the weight  $k$  and weight 0 actions of  $\text{Stab}_{\bar{\Gamma}}(\tau)$  coincide, then we can take  $b = 1$ .

If  $\tau$  is elliptic, conjugate onto  $\tau = 0$  as before. Then a function  $U \rightarrow \mathbf{C}$  is of weight  $k$  invariant under  $\{e^{\frac{2\pi i}{n}}\}$  if and only if it is of the form

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

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

If  $\tau$  is a cusp, we take without loss of generality  $\tau = \infty$ . If  $\infty$  is a regular cusp, or  $k$  is even, then the weight  $k$  and weight 0 actions of the stabilizer coincide and  $b = 1$  works. In the remaining case, we have

$$\mathcal{O}_X(V) = \{f : U \rightarrow \mathbf{C} \text{ holomorphic, } f(z+h) = f(z)\}$$

and

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

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  $\mathcal{O}_{\mathcal{H}} \cong \Omega_{\mathcal{H}}^1$  commuting with the action of  $\Gamma$ , if we give  $\mathcal{O}_{\mathcal{H}}$  the weight 2 action and  $\Omega_{\mathcal{H}}^1$  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  $\infty$  as usual.

In a sufficiently small neighbourhood  $V$  of  $\infty$ , the local coordinate is  $q = e^{\frac{2\pi iz}{h}}$ , so  $dq = \frac{2\pi i}{h}q dz$ . Thus  $dz = \frac{h}{2\pi i} \frac{dq}{q}$ , and so we have

$$\omega_2(V) = \mathcal{O}_X(V)dz = \mathcal{O}_X(V) \frac{dq}{q} = \frac{1}{q} \Omega^1(V)$$

and so holomorphic sections of  $\omega_2$  give differentials with simple poles. □

By construction,  $H^0(X(\Gamma), \omega_k) = M_k(\Gamma)$ , 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*

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

(Thus  $1 \leq r \leq 12$ .) 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  $\omega_r$  in the previous proof were 1, and the local bases of  $\omega_k$  depended only on  $k$  modulo  $r$ . So (local basis of  $\omega_r$ )  $\times$  (local basis of  $\omega_k$ ) = (local basis of  $\omega_{k+r}$ ) 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  $\Gamma$  is *neat*.

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

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

*Proof.* We have

$$\begin{aligned} \deg(\omega) &= \frac{1}{2} \deg(\omega^{\otimes 2}) \\ &= \frac{1}{2} (\deg(\Omega) + \varepsilon_{\infty}) \\ &= \frac{1}{2} (2g - 2 + \varepsilon_{\infty}), \end{aligned}$$

so if  $k \geq 2$ , then  $\deg(\omega^{\otimes k}) > 2g - 2$ , and Riemann-Roch implies that

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

□

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

*Example.*  $\Gamma = \Gamma_1(5)$  is neat (in fact  $\Gamma_1(N)$  is neat if and only if  $N \geq 5$ ), and  $g = 0$  and  $\varepsilon_{\infty} = 4$ , so  $\dim M_k(\Gamma) = k + 1$  for  $k \geq 2$ . For  $k = 1$ , we need to worry about  $H^0(X(\Gamma), \Omega^1 \otimes \omega^{-1})$ , but  $\deg \Omega^1 = 2g - 2 = -2$ , so  $\Omega^1 \otimes \omega^{-1}$  has degree  $-1$ .

If you try to do this for  $\Gamma_1(23)$ , it fails: the degree of  $\Omega^1 \otimes \omega^{-1}$  is zero. So the dimensions of the spaces of modular forms of weight 1 lie much deeper.

## 2. MODULAR CURVES AS ALGEBRAIC CURVES

2.1. Modular curves over  $\mathbf{C}$ .

**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  $X^{\text{an}}$  for a unique  $X$ .

(3) There exists an equivalence of categories

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

*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  $\Gamma$  there is an algebraic variety  $X(\Gamma)_{\mathbf{C}}$  and invertible sheaves  $\omega_k$  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.**  $X(\Gamma)_{\mathbf{C}} \cong \text{Proj}(\bigoplus_{k \geq 0} M_k(\Gamma))$ .

(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  $S_0 \cong \mathbf{C}$ ,

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

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.  $\square$

*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.  $\mathbf{P}_{\mathbf{Q}}^1$  and  $\{X^2 + Y^2 + 2Z^2\} \subset \mathbf{P}_{\mathbf{Q}}^1$  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

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

In particular, for  $X/\mathbf{Q}$  a curve,

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

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



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  $\omega$ .

Lecture 4

**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  $\Phi_N(j(z), j(Nz))$  for  $\Phi_N$  a symmetric polynomial.*
- (3) *If  $N = p$  is prime, then  $\Phi_p(X, Y) \equiv (Y^p - X)(Y - X^p) \pmod{p}$ .*

*Proof.* It is clear that  $j(Nz)$  is a meromorphic function on  $X_0(N)$ , so  $\mathbf{C}(j(z), j(Nz)) \subseteq \mathbf{C}(X_0(N))$ . Moreover,  $\mathbf{C}(X_0(N))$  has degree  $[\mathrm{PSL}_2(\mathbf{Z}) : \Gamma_0(N)] = [\mathrm{SL}_2(\mathbf{Z}) : \Gamma_0(N)] =: m$  over  $\mathbf{C}(j)$ .

Let  $\gamma_1, \dots, \gamma_m$  be such that  $\mathrm{SL}_2(\mathbf{Z}) = \Pi_i \Gamma_0(N) \gamma_i$ . 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  $[\mathbf{C}(j(z), j(Nz)) : \mathbf{C}(j(z))] \geq 1$ , so (1) would follow.

Suppose  $j(N\gamma_i z) = j(N\gamma_j z)$  for all  $z \in \mathcal{H}$ . Then

$$j \left( \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \gamma_j \gamma_i^{-1} \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}^{-1} z \right) = j(z) \quad \forall z \in \mathcal{H}.$$

By a question from the problem sheet we must have  $\begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix} \gamma_j \gamma_i^{-1} \begin{pmatrix} N & 0 \\ 0 & 1 \end{pmatrix}^{-1} \in \mathrm{SL}_2(\mathbf{Z})$ , which forces  $\gamma_j \in \Gamma_0(N) \gamma_i$ , so  $j = i$ . So the  $j(N\gamma_i z)$  are distinct and we are done.

Now (2). We have from (1) that the minimal polynomial is

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

The coefficients are symmetric polynomials in the functions  $j(N\gamma_i z)$ ; in particular they are holomorphic on  $\mathcal{H}$ . Thus they are polynomials in  $j(z)$  (not just rational functions). So  $\Phi_N(X, Y) \in \mathbf{C}[X, Y]$ .

Now 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  $\gamma_i$  explicitly, we see that all  $j(N\gamma_i z)$  are of the form  $j(\frac{az+b}{d})$  with  $ad = N$ . Hence  $j(N\gamma_i z)$  has  $q$ -expansion in  $\mathbf{Z}[\zeta_N][[q^{\frac{1}{N}}]]$  for all  $i$ .

Hence the coefficients of  $\Phi_N(j, Y)$  are polynomials in  $j$  whose  $q$ -expansion lie in  $\mathbf{Z}[\zeta_N][[q]]$ . We claim that this implies that they lie in  $\mathbf{Z}[\zeta_N][j]$ . Let  $P = \sum_{n=1}^k b_n j^n$ , 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. 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  $\mathbf{C}$ , but it must be Galois invariant, so it lands 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

$$\Phi_N \left( j \left( -\frac{1}{Nz} \right), j \left( -\frac{N}{Nz} \right) \right) = 0 \quad \Leftrightarrow \quad \Phi_N(j(Nz), j(z)) = 0$$

for all  $z$ . Thus  $\Phi_N(Y, X)$  is a constant multiple of  $\Phi_N(X, Y)$ , necessarily  $\pm 1$ . But if  $\Phi_N(Y, X) = -\Phi_N(X, Y)$ , then  $\Phi_N(X, X) = 0$ , contradicting 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  $\gamma_i$  explicitly, and we have

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

Hence  $\Phi_N(j(z), Y)$  has  $q$ -expansion

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

As elements of  $\mathbf{Z}[\zeta_N](Y)$ , the  $j\left(\frac{z+i}{p}\right)$  are all congruent modulo the unique prime ideal  $\wp$  above  $p$  in  $\mathbf{Z}[\zeta_N]$ , so this is equal to

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

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

$$\Phi_N(X, Y) \equiv (Y - X^p)(Y^p - X) \pmod{p}.$$

□

*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  $\Phi_N$  are *huge*. For instance, for  $N = 2$  we have

$$\begin{aligned} \Phi_N(X, Y) &= X^3 + Y^3 - X^2Y^2 + 1488XY(X + Y) - 162000(X^2 + Y^2) \\ &\quad + 40773375XY + 8748000000(X + Y) - 15746400000000. \end{aligned}$$

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

**Definition.** We define  $X_0(N)$  as the unique smooth projective curve  $/\mathbf{Q}$  with function field  $\mathbf{Q}(X)[Y]/\Phi_N(X, Y)$ .

This is obviously a rather natural model, and the map  $X_0(N)_{\mathbf{C}} \rightarrow X_0(1)_{\mathbf{C}} = \mathbf{P}^1$  descends to  $\mathbf{Q}$ , which is a good sign!

**Theorem 2.2.2.** *There is a sheaf  $\omega_{k, \mathbf{Q}}$  on  $X_0(N)_{\mathbf{Q}}$  whose base-extension to  $\mathbf{C}$  is  $\omega_k$ .*

*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  $\omega_2 \cong \Omega_{X_0(N)_{\mathbf{C}}}^1(\text{cusps})$ . More generally, we can always write

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

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  $\infty$  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  $X_0(N) \rightarrow X_0(1)$  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

$$\Omega_{X_0(N)_{\mathbf{C}}}^1 = \mathbf{C} \otimes \Omega_{X_0(N)_{\mathbf{Q}}}^1$$

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

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

□

**Corollary 2.2.3.** *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  $q$ -expansions have coefficients in  $\mathbf{Q}$ .*

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

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

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

Now any section of  $\omega_{k, \mathbf{Q}}$  is an element of  $\mathbf{Q}(X_0(N)_{\mathbf{Q}})$  multiplied by  $(dj)^{\otimes \frac{k}{2}}$ , which is a meromorphic section of  $(\Omega_{X_0(N)_{\mathbf{Q}}}^1)^{\otimes \frac{k}{2}}$ . 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.  $\square$

*Remark.* Unfortunately this method does not extend to other  $\Gamma$ : it is hard to write down  $\mathbf{C}(X(\Gamma))$  explicitly, and even harder to get one's hands on  $\omega_k$  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.** Recall that if  $\Lambda$  is a lattice in  $\mathbf{C}$  (i.e. a discrete subgroup isomorphic to  $\mathbf{Z}^2$ ) then  $\Lambda$  is homothetic to a lattice of the form

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

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

So

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

*Exercise.* (i) For any  $N \geq 1$ , there is a bijection between  $Y_0(N)$  and equivalence classes of pairs  $(\Lambda, C)$ , where  $\Lambda$  is a lattice and  $C$  is a cyclic subgroup of  $\mathbf{C}/\Lambda$  of order  $N$ , and we say  $(\Lambda, C) \sim (\Lambda', C')$  if there is an isomorphism  $\mathbf{C}/\Lambda \cong \mathbf{C}'/\Lambda'$  mapping  $C$  to  $C'$ .

(ii) Similarly, there is a bijection between  $Y_1(N)$  and equivalence classes of pairs  $(\Lambda, P)$ , where  $P$  is a point of exact order  $N$  in  $\mathbf{C}/\Lambda$ , and  $(\Lambda, P) \sim (\Lambda', P')$  if there is an isomorphism  $\mathbf{C}/\Lambda \cong \mathbf{C}'/\Lambda'$  mapping  $P$  to  $P'$ .

Note that in (ii),  $(\Lambda, P)$  is always equivalent to  $(\Lambda, -P)$ . In (i) there are lots more exceptional cases coming from elliptic points.

*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 Ring (all rings are assumed commutative and unital here), R - Alg (algebras over a fixed ring  $R$ , a "slice category" of Ring) and Set. (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 Ring  $\rightarrow$  Set, or R - Alg  $\rightarrow$  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.

SOME COMMENTS ABOUT THE QUESTIONS ON QUESTION SHEET 1

Lecture 5

*Question 4.* Some people got misled by a typo in the notes: it should read

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

for  $\Gamma$  neat, and not  $(k - 1)(g - 2)$ . Thanks to Robin Bartlett for pointing this out.

*Question 5.* Can be done without using the explicit formula for  $\dim M_k(\Gamma_0(5))$  given in Diamond + Shurman; but most people used this formula, without saying anything about its proof, for which I took off 1 mark.

*Question 10.* See recent arxiv preprints by Nadim Rustom.

*Question 11.* There is a unique index 2 subgroup of  $\mathrm{SL}_2(\mathbf{Z})$  (coming from the index 2 subgroup of  $\mathrm{SL}_2(\mathbf{F}_2) \cong S_3$ ), and it works.

*Question 12.* Very nice argument from Jan Vonk: can deduce from Q9 that the quotient of  $\overline{\Gamma_0(11)}$  by the normal subgroup generated by the stabilizers of the cusps 0 and  $\infty$  is isomorphic to  $\mathbf{Z}^2$  (because it's the fundamental group of the compactification  $X_0(11)$ , which has genus 1.) Hence there are infinitely

many finite-index normal subgroups  $\Gamma$  of  $\Gamma_0(11)$  such that the corresponding map  $X(\Gamma) \rightarrow X_0(11)$  is unramified at the cusps, and hence everywhere, since there are no elliptic points. By Riemann-Hurwitz, an unramified cover of a genus 1 curve is a genus 1 curve, so all the modular curves  $X(\Gamma)$  have genus 1.

We now discuss some properties of functors and representable functors. 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 functor  $\mathcal{F} : \mathcal{C} \rightarrow \underline{\mathrm{Set}}$  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:  $\mathrm{id}_X$ . Need to know a corresponding element of  $\mathcal{F}(X)$ , as  $h^X(X) \cong \mathcal{F}(X)$ . Having specified a  $\varphi \in \mathcal{F}(X)$  corresponding to  $\mathrm{id}_X$ , this determines an element of  $\mathcal{F}(Y)$  for every homomorphism  $\alpha : X \rightarrow Y$ : take  $\mathcal{F}(\alpha)(\varphi)$ .

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

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

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

(We have discussed *covariant* functors, but we get the same for *contravariant* functors by replacing  $\mathcal{C}$  with  $\mathcal{C}^{\mathrm{opp}}$ .)

It follows that if  $\mathcal{F}$  is representable, the bijection  $\mathcal{F}(Y) \cong h^X(Y)$  is determined by a  $\varphi \in \mathcal{F}(X)$ ; we are saying that for every  $Y \in \mathrm{Ob}(\mathcal{C})$  and  $\psi \in \mathcal{F}(Y)$ , there exists a unique homomorphism  $\alpha : X \rightarrow Y$  such that  $\mathcal{F}(\alpha)(\varphi) = \psi$ . We say that  $(X, \varphi)$  *represents*  $\mathcal{F}$ . Note that  $\varphi$  is an essential part of the data!

*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  $\alpha : \mathbf{Z}[T] \rightarrow R$  such that  $\alpha(T) = r$ ;
- $\mathcal{F}(R) = R^\times$  is represented by  $(\mathbf{Z}[T, T^{-1}], T)$ ;
- $\mathcal{F}(R) = \{\text{nth roots of unity in } R\}$  is represented by  $\mathbf{Z}[T]/(T^n - 1), T$ . (Caveat: 'primitive nth roots of 1' is not a functor on  $R$ .)

*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  $b \in A$ . Then for any ring  $S$  and an element  $s \in S$  which is a square, there exists a unique homomorphism  $\alpha : A \rightarrow S$  such that  $\alpha(a) = s$ . But: take  $S = \mathbf{Z}[T]$  and  $s = T^2$ . Then there exists a unique  $\alpha : A \rightarrow \mathbf{Z}[T]$  with  $\alpha(a) = T^2$ , so  $\alpha(b) \in \{\pm T\}$ . Let  $\sigma : S \rightarrow S$  be  $T \rightarrow -T$ , 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  $\sigma \circ \alpha(b) \neq \alpha(b)$ . This contradicts the uniqueness of  $\alpha$ .

The moral of this non-example: *automorphisms are bad for representability!*

**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  $\pi : \mathcal{E} \rightarrow S$  (an  $S$ -scheme) such that  $\pi$  is proper and flat and all fibres are smooth genus 1 curves, given with a section "0" :  $S \rightarrow \mathcal{E}$ .

*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’. For  $\mathcal{E}/S$  as above,

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

are the sections of  $\pi : \mathcal{E} \rightarrow S$  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  $\mathcal{E}/\text{Spec}(\mathbf{Z}_p)$ , can have points of order  $p$  reducing (mod  $p$ ) to 0 (at closed points of  $\text{Spec}(\mathbf{Z}_p)$ ).

**Proposition 3.3.2.** *If  $\mathcal{E}/S$  is an elliptic curve, then  $\mathcal{E}$  has a Weierstrass equation locally on  $S$ , i.e. there exists a covering  $\coprod U_i \rightarrow S$  in the Zariski topology such that  $\mathcal{E}|_{U_i}$  has a Weierstrass equation for all  $i$ . More precisely, we will show the following:  $\omega_{\mathcal{E}/S} = \pi_*(\Omega_{\mathcal{E}/S}^1)$  is an invertible sheaf on  $S$ , and any local basis  $\omega$  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  $\omega = \frac{-dx}{2y+a_1x+a_3}$ .*

*Sketch.* The invertibility of  $\pi_*(\Omega_{\mathcal{E}/S}^1)$  comes from a calculation in sheaf cohomology, c.f. p.53 of Mumford’s ‘Abelian varieties’. Now, given  $U$  and  $\omega$  a basis of  $\pi_*(\Omega_{\mathcal{E}/U}^1)$ , this gives a local parameter on  $\mathcal{E}$  at 0 such that

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

(a local parameter adapted to  $\omega$ ).  $\pi_*\mathcal{O}_{\mathcal{E}}(2(0))$  is locally free of rank 2 over  $U$ . If  $U = \text{Spec}(A)$  is affine, then  $\pi_*\mathcal{O}_{\mathcal{E}}(2(0))$  is  $A \cdot (1, x)$ , where  $x = \frac{1}{T^2}(1 + \dots)$ . Similarly,

$$\pi_*\mathcal{O}_{\mathcal{E}}(3(0)) = A \cdot (1, x, y), \quad y = \frac{1}{T^3} + \dots,$$

$$\pi_*\mathcal{O}_{\mathcal{E}}(4(0)) = A \cdot (1, x, y, x^2),$$

$$\pi_*\mathcal{O}_{\mathcal{E}}(5(0)) = A \cdot (1, x, y, x^2, xy).$$

Now  $y^2 - x^3 \in \pi_*\mathcal{O}_{\mathcal{E}}(5(0))$ , 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

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

which implies  $\omega = \frac{-dx}{2y+a_1x+a_3}$  as the global differentials on  $\mathcal{E}$  are a rank 1  $A$ -module.

(We don’t have such a nice characterisation of the Weierstrass equation if 2 is not invertible on  $S$ .)  $\square$

**Definition 3.3.3.** For  $S$  a scheme,  $\alpha, \beta \in \Gamma(S, \mathcal{O}_S)$ , let  $E(\alpha, \beta)$  be the subscheme of  $\mathbf{P}_S^2$  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  $\Delta(\alpha, \beta) \in \Gamma(S, \mathcal{O}_S)^\times$ , this is an elliptic curve over  $S$ .

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

$$-P = (0 : -\beta : 1),$$

$$2P = (-\beta : \beta(\alpha - 1) : 1),$$

$$-2P = (-\beta : 0 : 1),$$

$$3P = (1 - \alpha : \alpha - \beta - 1 : 1),$$

$$-3P = (1 - \alpha : (\alpha - 1)^2 : 1),$$

...

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

**Proposition 3.3.4.** *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  $\alpha, \beta \in \Gamma(S, \mathcal{O}_S)$  such that  $\Delta(\alpha, \beta) \in \Gamma(S, \mathcal{O}_S)^\times$  and a unique isomorphism  $E(\alpha, \beta) \cong E$  mapping  $(0 : 0 : 1)$  to  $P$ .*

We now prove Proposition 3.3.4.

*Proof.* First, assume that  $E$  has a Weierstrass equation over  $S$ . By a translation  $s \mapsto x + s$ ,  $y \mapsto y + t$ , we can assume that  $P = (0, 0)$ . Since  $P$  does not have order 2 in any fibre, gradients of tangent line at  $P$  is in  $\Gamma(S, \mathcal{O}_S)$ , so by replacing  $y$  with  $y + rx$  for some  $r$  we can put the equation into the form

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2$$

for some  $a_i \in \Gamma(S, \mathcal{O}_S)$ . Since  $P$  does not have order 3 in any fibre,  $(0, 0)$  is not an inflexion point, so  $a_2 \in \Gamma(S, \mathcal{O}_S)^\times$ . So by scaling  $x \mapsto u^2x$ ,  $y \mapsto u^3y$ , 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  $\Gamma(U_i, \mathcal{O}_S)$ , so we get  $\alpha_i, \beta_i \in \Gamma(U_i, \mathcal{O}_S)$  such that  $E|_{U_i}, P_{U_i} \cong (E(\alpha_i, \beta_i), (0, 0))$ . Since  $\alpha_i, \beta_i$  are unique, they must agree on  $U_i \cap U_j$ . The sheaf property of  $\mathcal{O}_S$  then implies that there exist  $\alpha, \beta \in \Gamma(S, \mathcal{O}_S)$  such that  $\text{res}_{U_i}(\alpha) = \alpha_i$  and  $\text{res}_{U_i}(\beta) = \beta_i$ . Then  $(E, P) \cong (E(\alpha, \beta), (0, 0))$ .  $\square$

*Remark.* The last step used in an essential way the *uniqueness of  $(\alpha, \beta)$* ; ‘‘local uniqueness gives global existence’’.

**Corollary 3.3.5.**      • *The pair*

$$\left( \text{Spec } \mathbf{Z}[A, B, \Delta(A, B)^{-1}], (E(A, B), (0 : 0 : 1)) \right)$$

represents the functor  $\text{Sch}^{\text{opp}} \longrightarrow \text{Set}$ ,

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

• *The pair*

$$\left( \text{Spec } \mathbf{Z}[B, \Delta[1 + B, B]^{-1}], (E(1 + B), (0, 0)) \right)$$

represents the functor

$S \mapsto \{ (E, P), E/S \text{ elliptic curve, } P \text{ a point of exact order } 5 \text{ in every fibre} \}$ .

*Proof.* (i) is a restatement of Proposition 3.3.4. For (ii), just equate  $3P = -2P$ :

$$\begin{aligned} 3P &= (1 - A, A - B - 1), \\ -2P &= (-B, 0). \end{aligned}$$

$\square$

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

Is it reasonable to define  $Y_1(5)_{\mathbf{Z}}$  to be  $\text{Spec } \mathbf{Z}[B, \Delta(1 + B, B)^{-1}]$ ? Sadly *no*; our definition of ‘point of exact order 5’ is too naive in characteristic 5: if  $E/\mathbf{F}_5$  is supersingular,  $E[5]$  is a single point with multiplicity 25, so there are no points of order 5, even over  $\overline{\mathbf{F}}_5$ . Hence this scheme  $\text{Spec } \mathbf{Z}[B, \Delta(1 + B, B)^{-1}]$  has empty fibre over supersingular  $j$ -invariants.

**Definition 3.3.6.** We set

$$Y_1(5)_{\mathbf{Z}[\frac{1}{5}]} = \text{Spec } \mathbf{Z}[\frac{1}{5}, B, \Delta(1 + B, B)^{-1}].$$

This represents the same functor as before in the category of  $\mathbf{Z}[\frac{1}{5}]$ -schemes.

More generally, we have the following definition:

**Definition.** (continued) For  $N \geq 4$ , let  $\mathcal{Y}_N$  be the closed subscheme of  $\mathcal{Y} = \text{Spec } \mathbf{Z}[A, B, \Delta(A, B)^{-1}]$ , where  $N \cdot (0 : 0 : 1) = (0 : 1 : 0)$ , and let

$$Y_1(N)_{\mathbf{Z}[\frac{1}{N}]} = \left( \mathcal{Y}_N - \bigcup_{d|N, 4 \leq d < N} \mathcal{Y}_d \right) \times_{\text{Spec } \mathbf{Z}} \text{Spec } \mathbf{Z}[\frac{1}{N}].$$

By construction, this represents the functor

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

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

*Remark.* More precisely,  $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}$  has a universal elliptic curve over it by restricting  $E(\alpha, \beta)/\mathcal{Y}$ , and this has a point  $(0, 0)$ . The triple  $(Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}, \text{this curve, this point})$  represents the above functor.

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  $\mathbf{C}$ ?

### 3.4. Smoothness.

**Definition 3.4.1.** A morphism of schemes  $\phi : X \rightarrow Y$  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  $\mathcal{S} \rightarrow S$  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  $N \geq 1$  is invertible on  $S$ , then  $[N] : E \rightarrow E$  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.  $\square$

**Proposition 3.4.3.** (*Functorial criterion for smoothness*) *Let  $X \rightarrow \text{Spec}(R)$  be a scheme of finite type over  $R$ , where  $R$  is noetherian. Then the map  $X \rightarrow \text{Spec}(R)$  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*

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

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

*Proof.* See Stacks Project §36.9.  $\square$

*Remark.* If we replace surjective with bijective, then we get the notion of ‘formally étale’.

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

*Proof.* Let  $A$  be a local  $\mathbf{Z}[\frac{1}{N}]$ -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  $\text{Spec}(A_0)$ . Lift coefficients arbitrarily to  $A$  to get  $E/A$  lifting  $E_0$ ; 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  $[N] : E \rightarrow E$  is smooth, and a composition of smooth morphisms is smooth. (We apply this to  $[N]$  composed with the structure map  $E \rightarrow \text{Spec } A$ .) Hence  $(E_0, P_0)$  lifts to  $(E, P)$ , and we are done.  $\square$

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

### 3.5. A complex-analytic digression.

Let  $\Lambda \subset \mathbf{C}$  be a lattice.

**Definition 3.5.1.** The Weierstrass  $\wp$ -function  $\wp_\Lambda(z)$  is the unique holomorphic function  $\mathbf{C}/\Lambda \rightarrow \mathbf{P}^1(\mathbf{C})$  such that  $\wp_\Lambda(z) = \frac{1}{z^2} + O(1)$  at  $z = 0$ , and  $\wp_\Lambda(s)$  is holomorphic away from  $\Lambda$ .

(This is the  $x$ -coordinate on  $\mathbf{C}/\Lambda$  corresponding to the differential  $dz$  in the construction of Proposition 3.3.2.)

The machinery of Proposition 3.3.2 shows that  $\mathbf{C}/\Lambda$  is isomorphic to

$$Y^2 = X^3 - g_4X - g_6$$

via  $z \mapsto (\wp(z), -\frac{1}{2}\wp'(z))$ , where  $g_4, g_6$  are constants depending on  $\Lambda$ . If  $\Lambda = \Lambda\tau = \mathbf{Z} + \tau\mathbf{Z}$ , then  $g_4$  and  $g_6$  are constant multiples of the Eisenstein series  $E_4$  and  $E_6$  (holomorphic as functions of  $\tau$ ).

**Proposition 3.5.2.** *Let*

$$\mathcal{U} = \{\tau, z \in \mathcal{H} \times \mathbf{C} : z, 2z, 3z \notin \Lambda_\tau\}.$$

*Then there exists a holomorphic map*

$$\mathcal{U} \xrightarrow{(\alpha, \beta)} \mathbf{C}^2 - \{\alpha, \beta : \Delta(\alpha, \beta) = 0\}$$

such that

$$E(\alpha(\tau, z), \beta(\tau, z)) \cong \mathbf{C}/\Lambda_\tau, \\ (0, 0) \leftrightarrow z \pmod{\Lambda_\tau}.$$

*Proof.* Start from the pair

$$\left( Y^2 = X^3 - g_4 X - g_6, \quad (\wp(z), -\frac{1}{2}\wp'(z)) \right),$$

where  $g_4, g_6$  correspond to  $\Lambda_\tau$ . Manipulate this as in Proposition 3.3.4 to put this in Tate normal form. All coefficients of rescalings and translations are holomorphic on  $U$  as functions of  $(\tau, z)$ , and hence so are the resulting  $\alpha, \beta$ .  $\square$

**Corollary 3.5.3.**  $Y_1(N)_{\mathbf{Z}[\frac{1}{N}]}(\mathbf{C})$  is isomorphic to  $\Gamma_1(N)\backslash\mathcal{H}$  as a Riemann surface.

Recall that we characterised  $Y_0(N)/\mathbf{Q}$  using  $q$ -expansions. We'll now explain a construction which (among other things) shows that our model of  $Y_1(N)$  does *not* have such a nice characterization.

**Proposition 3.5.4.** (Siegel) Let  $E/S$  be an elliptic curve, and let  $c > 1$  be an integer not divisible by 2 or 3. Then there exists a unique element  ${}_c\Theta_E \in O(E \setminus E[c])^\times$  with the following properties:

- (1)  $\operatorname{div}({}_c\Theta_E) = c^2(0) - E[c]$ ;
- (2)  $N_a({}_c\Theta_E) = {}_c\Theta_E$  for  $a$  coprime to  $c$ , where  $N_a$  is the norm map

$$O(E \setminus E[ac])^\times \longrightarrow O(E \setminus E[c])$$

attached to the  $a$ -multiplication on  $E$ .

Moreover, if  $E = \mathbf{C}/(\mathbf{Z} + \tau\mathbf{Z})$  and  $S = \mathbf{C}$ , then we have

$${}_c\Theta_E = q^{\frac{c^2-1}{12}} (-t)^{-\frac{c(c-1)}{2}} \gamma_q(t)^{c^2} \gamma_q(t^c)^{-1},$$

where  $t = e^{2\pi iz}$  (for  $z \in \mathbf{C}/(\mathbf{Z} + \tau\mathbf{Z})$ ),  $q = e^{2\pi i\tau}$  and

$$\gamma_q(t) = \prod_{n \geq 0} (1 - q^n t) \prod_{n \geq 1} (1 - q^n t^{-1}).$$

*Proof.* First, note that this unit is unique if it exists: assume that  $f$  satisfying (i), (ii) exists. Then any other  $g$  satisfying these conditions is  $g = uf$  for some  $u \in O(S)^\times$ . Then

$$\begin{aligned} N_3(g) = g &\quad \Rightarrow \quad N_3(uf) = uf \\ &\quad \Rightarrow \quad u^3 f = uf \\ &\quad \Rightarrow \quad u^8 = 1. \end{aligned}$$

Similarly,  $N_2(g) = g$  implies that  $u^3 = 1$ , and hence

$$u = (u^3)^2 (u^8)^{-1} = 1.$$

Hence we get uniqueness, and we can prove existence locally on  $S$ . It suffices to show that  $c^2(0) - E[c]$  is locally on  $S$  a principal divisor. There exists a theory of *relative Cartier divisors* and a map

$$\frac{\text{degree 0 divisors on } E}{\text{pullbacks of ones on } S} \longrightarrow E(S).$$

Since  $c^2(0) - E[c] = 0$  in  $E(S)$ , it follows that  $c^2(0) - E[c]$  is the pullback of a divisor on  $S$  and hence locally-on- $S$  principal.

Let  $f$  be such that  $\operatorname{div}(f) = c^2(0) - E[c]$ . Since  $\operatorname{div}(f)$  is invariant under  $N_a$ , we must have  $N_a(f) = u_a f$  for some  $u_a \in O(S)^\times$ . Since  $N_a N_b = N_b N_a$ , we have

$$u_a (b^2 - 1) = a_b^{a^2-1}$$



for all  $a, b$  coprime to  $c$ . It follows that if we put  $g = u_2^{-3} u_3 u_a f$ , we have

$$\begin{aligned} N_a(g) &= u_2^{-3a^2} u_3^{a^2} u_a f \\ &= u_2^{3(a^2-1)} u_3^{a^2-1} u_a g \\ &= u_a^{3(2^2-1)} u_a^{3^2-1} u_a g \\ &= g. \end{aligned}$$

Hence  ${}_c\Theta_E$  exists locally and by uniqueness it exists globally.

For the case  $S = \mathbf{C}$ ,  $E = E_\tau$ , we just check that the given function has properties (i) and (ii) (c.f. 3rd problem sheet).  $\square$

**Definition 3.5.5.** For  $N \geq 4$  and  $c = 1$  with  $(c, 6N) = 1$ , the Siegel unit  ${}_c g_N$  is the pullback of  ${}_c\Theta_{\mathcal{E}}$  along the order  $N$  section  $Y_1(N) \rightarrow \mathcal{E}$ , where  $\mathcal{E}$  is the universal elliptic curve over  $Y_1(N)$ .

*Remark.* These units are the building blocks of *Euler systems* – c.f. Kato’s paper (Astérisque 295, 2004) and my 2013 paper with Lei and Zerbes.

Important corollary:  $Y_1(N)$  is not characterised over  $\mathbf{Q}$  by having  $q$ -expansions of elements of  $Q(Y_1(N))$  in  $\mathbf{Q}((q))$ . To see this, calculate the  $q$ -expansion of  ${}_c g_5 \in \mathbf{Q}(Y_1(5))^\times$ : The order  $N$  section is  $z = \frac{1}{5} \pmod{\mathbf{Z} + \tau\mathbf{Z}}$ , so

$$\begin{aligned} t &= e^{2\pi iz} = e^{\frac{2\pi i}{5}} \notin \mathbf{Q}, \\ {}_c g_5 &= q^{\frac{c^2-1}{12}} \left( -e^{\frac{2\pi i}{5}} \right)^{-\frac{c(c-1)}{2}} \prod \dots \end{aligned}$$

has  $\zeta_5$ ’s everywhere.

In fact, one can show that  $f \in \mathbf{Q}(Y_1(N))$  if and only if  $f \in \mathbf{C}(Y_1(N))$  and its  $q$ -expansion lands in  $\mathbf{Q}(\zeta_N)((q))$  and satisfies  $a_n(f)^\sigma = a_n(\langle \sigma \rangle f)$  for all  $\sigma \in \text{Gal}(\mathbf{Q}(\mu_N)/\mathbf{Q}) \cong (\mathbf{Z}/N)^\times$ .

### 3.6. Quotients and $Y_0(N)$ .

**Proposition 3.6.1.** *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  $X \rightarrow X/G$  representing the functor*

$$Y \mapsto (\text{homs. of } S\text{-schemes } X \rightarrow Y \text{ commuting the } G\text{-action}).$$

*Proof.* Uniqueness is obvious (representing a functor). Existence: for  $X = \text{Spec}(A)$  affine,  $\text{Spec}(A^G)$  works, and one can show that these patch nicely. (One needs quasiprojectiveness and finiteness of  $G$  here.)  $\square$

**Definition 3.6.2.** For  $N \geq 4$ , let  $Y_0(N) = Y_1(N)/(\mathbf{Z}/N)^\times$  (as a  $\mathbf{Z}[\frac{1}{N}]$ -scheme).

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

Construction: Let  $S$  be a  $\mathbf{Z}[\frac{1}{N}]$ -scheme. There is a map

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

defined as follows. let  $(E, C)$  be an element of LHS. Then there exists  $S' \rightarrow S$  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

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

In general, this is neither injective nor surjective, but if  $S = \text{Spec}(\bar{k})$  for  $\bar{k}$  algebraically closed, then it is bijective:

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

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

is not injective. (There exist obstructions coming from quadratic twists, etc.) For  $k$  a field, can check that the image  $(\iota_k)$  is the set of pairs  $(E, C)$  defined over  $K$  modulo isomorphisms over  $\bar{k}$ .

*Surjectivity.* One can show that for  $k$  a field,  $\iota_k$  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  $S \rightarrow Y_0(N)$ . (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!)

Fact:  $Y_0(N)$  is smooth over  $\mathbf{Z}[\frac{1}{N}]$  and agrees with our earlier construction over  $\mathbf{Q}$ . (Sketch of the last point: it suffices to show that  $j(z)$  and  $j(Nz)$  lie in  $\mathbf{Q}(Y_1(N))^G$  – just take  $j(E)$  and  $E(E/\langle P \rangle)$ ).

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

**Definition 3.7.1.** Let  $R$  be a ring.

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

- objects are diagrams  $\begin{array}{c} E \\ \downarrow \\ S \end{array}$ , 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  $\mathcal{P} : \underline{\text{Ell}}/R \rightarrow \underline{\text{Set}}$ .

(3) We say that  $\mathcal{P}$  is representable if it is representable; it is *relatively representable* if, for every  $E/S \in \text{Ob}(\underline{\text{Ell}}/R)$ , the functor  $\underline{\text{Sch}}/S \rightarrow \underline{\text{Set}}, T \mapsto \mathcal{P}(E \times_S T/T)$  is representable.

*Remark.* The category  $\underline{\text{Ell}}/R$  is  $\underline{\text{Sch}}/Y$  for a  $Y$  that does not exist. If the functor  $S \mapsto \{\text{ell. curves}/S\}$  were representable by some  $(Y, E/Y)$ , then the objects of  $\underline{\text{Ell}}/R$  would be just maps  $S \rightarrow Y$ . This is the idea of *stacks*.

**Proposition 3.7.2.** For  $\mathcal{P}$  a moduli problem, let  $\tilde{\mathcal{P}} : \underline{\text{Sch}}/R \rightarrow \underline{\text{Set}}$  be the functor

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

If  $\mathcal{P}$  is representable on  $\underline{\text{Ell}}/R$ , then  $\tilde{\mathcal{P}}$  is representable on  $\underline{\text{Sch}}/R$ .

*Remark.* The converse is not quite true.

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

**Definition 3.7.3.**  $\mathcal{P}$  is *rigid* if for all  $E/S \in \text{Ob}(\underline{\text{Ell}}/R)$ ,  $\text{Aut}(E/S)$  acts on  $\mathcal{P}(E/S)$  without fixed points.

- Exercise.*
- (1) Any representable functor is rigid;
  - (2) if  $\mathcal{P}$  is rigid and  $\tilde{\mathcal{P}}$  is representable, then  $\mathcal{P}$  is representable.

**Theorem 3.7.4.** (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  $\mathbf{Z}[\frac{1}{3}]$ ;
- ‘Legendre moduli problem’ ( $\Gamma(2)$  with choice of differential) over  $\mathbf{Z}[\frac{1}{2}]$ .

Both have group actions ( $\text{GL}_2(\mathbf{F}_3)$  and  $\text{GL}_2(\mathbf{F}_2) \times \{\pm 1\}$ ). 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  $\text{GL}_2(\mathbf{F}_3)$ -action. Take invariants (this is OK since  $\mathcal{P}$  is rigid), so we get an object  $\mathcal{E}/S$  representing  $\mathcal{P}$  on  $\underline{\text{Ell}}/R[\frac{1}{3}]$ .

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

**3.8. General level structures.** Fix  $N$  and a subgroup  $H \subset \mathrm{GL}_2(\mathbf{Z}/N)$ .

**Fact 3.8.1.** *There exists a moduli problem  $\mathcal{P}_H$  on  $\underline{\mathrm{Ell}}/\mathbf{Z}[\frac{1}{N}]$  such that if  $\bar{k}$  is algebraically closed,  $E/\bar{k} \in \mathrm{Ob}(\underline{\mathrm{Ell}}/\mathbf{Z}[\frac{1}{N}])$ , 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  $\Gamma(N)$ ,

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

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

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

**Proposition 3.8.2.**  $\mathcal{P}_H$  is relatively representable and étale over  $\underline{\mathrm{Ell}}/\mathbf{Z}[\frac{1}{N}]$  (i.e. for all  $E/S \in \mathrm{Ob}(\underline{\mathrm{Ell}}/\mathbf{Z}[\frac{1}{N}])$ , the functor  $T \mapsto \mathcal{P}_H(E \times_S T)$  is represented by an étale  $S$ -scheme).

*Proof.* For  $H = \{1\}$ , for  $E/S \in \mathrm{Ob}(\underline{\mathrm{Ell}}/\mathbf{Z}[\frac{1}{N}])$ , we can find an explicit  $S$ -scheme representing  $\mathcal{P}_H$  on  $\mathrm{Sch}/S$ ; it is an open subscheme of  $E[N] \times_S E[N]$  given by non-vanishing of Weil pairings. For general  $H$  just take the quotient of this by  $H$ .  $\square$

*Remark.* We conclude that it is easier to relatively represent  $\mathcal{P}_H$  than to define it.

**Proposition 3.8.3.**  $\mathcal{P}_H$  is rigid on  $\underline{\mathrm{Ell}}/R[\frac{1}{6}]$  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  $\geq 5$ ,  $E/k$  is an elliptic curve,  $\phi \in \mathrm{Aut}(E)$ , then the pair  $(E, \phi)$  lifts to characteristic 0. (This is shown somewhere in chapter VI of Deligne-Rapoport.)  $\square$

So if  $H$  satisfies the hypotheses of Proposition 3.8.3, 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 3.8.3, then the functors  $\mathcal{P}_H$  and  $\tilde{\mathcal{P}}_H$  are not representable, but the proof of Theorem 3.7.4 gives a scheme  $Y_{\mathcal{P}_H}$  over  $\mathbf{Z}[1/N]$  that is “the best possible approximation” to representing the functor  $\tilde{\mathcal{P}}_H$ : there are maps

$$\tilde{\mathcal{P}}_H(S) \rightarrow Y_{\mathcal{P}_H}(S)$$

which are surjective for  $S$  a field, and bijective if  $S$  is algebraically closed, as in the special case of  $Y_0(N)$  above. 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  $\Gamma$  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 intrinsically as

$$Y_{\mathcal{P}_H}(\mathbf{C}) = \mathrm{GL}_2(\mathbf{Q}) \backslash \mathrm{GL}_2(\mathbf{A}) / (\mathbf{R}_{>0} \cdot \mathrm{SO}_2(\mathbf{R}) \cdot U)$$

where  $U$  is the preimage of  $H$  in  $\mathrm{GL}_2(\hat{\mathbf{Z}})$  (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,  $\underline{\omega}_{E/S} = \pi_*(\Omega_{E/S}^1)$ .

**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 from Chapter 2.*

*Proof.* Just unravel the definitions. We will show that both have the same pullbacks to  $\mathcal{H}$  and the actions of  $\Gamma$  agree. By definition, the pullback of  $\omega_{\text{Katz}}$  is  $\underline{\mathbf{C}}$ . The pullback of  $\underline{\omega}_{E/S}$  is

$$\pi_*(\text{rel. differentials of } \mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau)) = \underline{\mathbf{C}} \cdot 2\pi idz.$$

But the isomorphism  $\mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau) \cong \mathbf{C}/(\mathbf{Z} + \mathbf{Z}\gamma\tau)$  for  $\gamma \in \text{SL}_2(\mathbf{Z})$  is multiplication by  $(c\tau + d)^{-1}$  on  $\mathbf{C}$ , so it multiplies  $dz$  by this constant. Hence the action agrees with the one defined in the construction of  $\omega_{\text{Katz}}$ .  $\square$

**Definition 4.1.2.** For  $\Gamma$  a torsion free congruence subgroup of level  $N$ ,  $R$  a  $\mathbf{Z}[\frac{1}{N}]$ -algebra, define

$$\text{KM}_k(\Gamma, R) = H^0(Y(\Gamma) \times R, \underline{\omega}_{E/\sim}^k)$$

(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,  $\alpha \in \mathcal{P}_H(E/S)$ ,  $\omega$  a basis of  $\Gamma(E, \Omega_{E/S}^1)$ ) an element of  $\Gamma(S, \mathcal{O}_S)$  such that

- it is compatible with base change in  $S$ ,
- it is homogeneous of weight  $k$  in  $\omega$ .

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

Fun thing: over  $R = \mathbf{Z}[\frac{1}{6}]$ , 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  $(E, \omega)$  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

$$y^2 + xy = x^3 + a_4x + a_6,$$

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 discriminant of Tate( $q$ ) is exactly the  $q$ -expansion of  $\Delta$  (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 check 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*

$$(1) \quad (X(u, q), Y(u, q)) \oplus (X(v, q), Y(v, q)) = (X(uv, q), Y(uv, q)).$$

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

*Proof.* Take

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

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

Sneaky part: there exists a straightforward change of coordinates from  $\text{Tate}(q)$  to

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

which is  $\mathbf{C}/(\mathbf{Z} + \mathbf{Z}\tau)$  via  $(\wp(z, \tau), \wp'(z, \tau))$ .  $X$  and  $Y$  are just  $\wp$  and  $\wp'$  as power series in  $u = e^{2\pi iz}$  and  $q = e^{2\pi i\tau}$ . So the identity (1) holds for all  $u, q$  in an open subset of  $\mathbf{C} \times \mathbf{C}$ , so it holds as an identity of power series.  $\square$

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

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

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

*Example.*  $Y_1(5)$ : the points of order 5 on  $\text{Tate}(q)$  over  $\mathbf{Z}[\frac{1}{5}, \zeta_5][[q^{\frac{1}{5}}]]$  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  $\text{Gal}(\mathbf{Q}(\mu_5)/\mathbf{Q})$  on the cusps is now easy to see: it factors through  $\text{Gal}(\mathbf{Q}(\mu_5)^+/\mathbf{Q}) \cong C_2$ , and the nontrivial element swaps  $\{\zeta_5, \zeta_5^4\}$  with  $\{\zeta_5^2, \zeta_5^3\}$  and fixes the other two cusps.