Chapter 5

Modular Curves

5.1 Moduli of elliptic curves

Lemma 5.1.1. *$SL_2(\mathbb{R})$ acts transitively on the upper half plane $\mathbb{H} = \{\tau \in \mathbb{C} \mid \operatorname{Im}\tau > 0\}$ by fractional linear transformations. The stabilizer of i is $SO(2)$. Therefore, we can identify $\mathbb{H} = SL_2(\mathbb{R})/SO(2)$.*

I will omit the proof, which is not hard, and hopefully presented by one of you. We can view $\mathbb{H}$ as the upper hemisphere of the Riemann sphere $\mathbb{P}^1_{\mathbb{C}}$. The action of $SL_2(\mathbb{R})$ extends to the boundary $\partial\mathbb{H} = \mathbb{P}^1_{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$. In order to better visualize the action, it useful to note that $\mathbb{H}$ has a Riemannian metric, called the hyperbolic or Poincaré metric, where the geodesics are lines or circles meeting $\partial\mathbb{H}$ at right angles. The action of $SL_2(\mathbb{R})$ preserves this metric, so it takes a geodesic to another geodesic.

Recall that an elliptic curve can be written as a quotient $E_\tau = \mathbb{C}/L_\tau$ where $L_\tau = \mathbb{Z} + \mathbb{Z}\tau$ with $\tau \in \mathbb{H}$. The origin $0 \in E_\tau$ is a distinguished point, which is part of the structure. In particular, two elliptic curves E_τ and $E_{\tau'}$ are called isomorphic if there is there a holomorphic isomorphism $f : E_\tau \to E_{\tau'}$ taking the origin of the first curve to the origin of the second.

Theorem 5.1.2. *E_τ and $E_{\tau'}$ are isomorphic if and only of there exists a matrix*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

such that

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

Proof. The isomorphism $f : E_\tau \to E_{\tau'}$ is induced by a holomorphic map $F : \mathbb{C} \to \mathbb{C}$ satisfying

$$F(z + \lambda) = F(z) + \Lambda(z, \lambda)$$

for all $\lambda \in L_\tau$ and some function $\Lambda : \mathbb{C} \times L_\tau \to L_{\tau'}$. Since $\Lambda(-, \lambda)$ is necessarily continuous, it is constant. Differentiating the previous identify shows that F' is

doubly periodic, and therefore $F'(z) = \phi$ is constant. Therefore we can assume that $F(z) = \phi z$ since $f(0) = 0$. We must have $\phi L_\tau = L_{\tau'}$. Since $1, \tau'$ (resp. $1, \tau$) is a positively oriented basis of $L_{\tau'}$ (resp. ϕL_τ), we can find $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ such that

$$\begin{aligned} \tau' &= a\phi\tau + b\phi \\ 1 &= c\phi\tau + d\phi \end{aligned} \tag{5.1}$$

Therefore

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

Conversely, suppose the last equation is true. Set $\phi = (c\tau + d)^{-1}$. Then (5.1) holds. Therefore $\phi L_\tau = L_{\tau'}$. So that multiplication by ϕ gives an isomorphism $E_\tau \cong E_{\tau'}$. □

Corollary 5.1.3. *There is a natural bijection between the set of isomorphism classes of elliptic curves and the quotient space*

$$\mathcal{A}_1 := SL_2(\mathbb{Z})\backslash\mathbb{H} = SL_2(\mathbb{Z})\backslash SL_2(\mathbb{R})/SO(2)$$

At the moment, $\mathcal{A}_1$ is just a set. In order to give more structure, we need to analyze the action more carefully. First observe that $-I$ acts trivially on $\mathbb{H}$, so the action factors through $\Gamma = PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z})/\{\pm I\}$. Consider the closed region $F \subset \mathbb{H}$ lying above the unit circle and between the lines $\operatorname{Im} z = \pm 1/2$ depicted below.

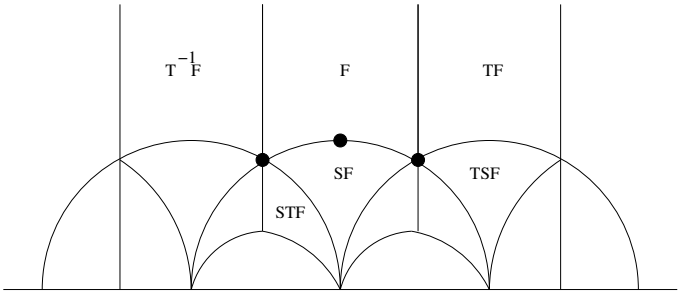

Figure 5.1: Fundamental domain

Let $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. These act by $z \mapsto -1/z$ and $z \mapsto z + 1$ respectively. S is a reflection about i which interchanges the regions $|z| \geq 1$ and $|z| \leq 1$. They generate a subgroup $G \subseteq \Gamma$.

Theorem 5.1.4.

(a) The union of translates gF, $g \in G$, covers $\mathbb{H}$.

(b) An interior point of F does not lie in any other translate of F under Γ.

(c) The isotropy group of $z \in F$ is trivial unless it is one of the points $\{i, e^{\pi i/3}, e^{2\pi i/3}\}$ marked in the diagram. The isotropy group is $\langle S\rangle, \langle ST\rangle, \langle TS\rangle$ respectively.

Proof. The intuition behind this can be understood from the picture. Repeatedly applying S and $T^{\pm 1}$ to F gives a *tiling* of $\mathbb{H}$ by hyperbolic triangles. Choose $\tau \in \mathbb{H}$, we want to find $A' \in SL_2(\mathbb{Z})$ and $\tau' \in F$ such that $A' \cdot \tau' = \tau$. Using (**??**), we can see that $\{\operatorname{Im} A \cdot \tau \mid A \in SL_2(\mathbb{Z})\}$ has a maximum M. Choose an A which realizes this maximum. Choose an integer n so that $\tau' = T^n A\tau$ has real part in $[-1/2, 1/2]$. Observe that $\operatorname{Im} \tau' = M$. If $|\tau'| < 1$ then $-1/\tau'$ would have imaginary bigger than M which is impossible. It follows that $\tau' \in F$, and τ lies in its orbit. This proves (a). For the remaining parts, see page 79 of [Serre, A Course in Arithmetic]

□

The set F is called a *fundamental domain* for the action of G. We can draw a number of useful conclusions.

Corollary 5.1.5. *$G = PSL_2(\mathbb{Z})$, i.e. S and T generate $PSL_2(\mathbb{Z})$.*

Proof. Let $z \in F$ be an interior point, and $h \in \Gamma$. Then $hz = gz$ for some $g \in G$. Since $z \in h^{-1}gF$, we must have $h^{-1}g = I$. □

Corollary 5.1.6. *The nontrivial elements of finite order in $PSL_2(\mathbb{Z})$ (resp. $SL_2(\mathbb{Z})$) are conjugate to S or $(ST)^{\pm 1}$ (resp. $-I, \pm S, \pm(ST)^{\pm 1}$).*

Proof. It is enough to prove the statement for $PSL_2(\mathbb{Z})$. A nontrivial element of $PSL_2(\mathbb{Z})$ of finite order must lie in the isotropy group of some point in $\mathbb{H}$. The points in the plane with nontrivial isotropy groups must be a translate of i or $e^{2\pi i/3}$. Their isotropy groups must be conjugate to the isotropy groups of one these two points. □

Corollary 5.1.7. *The action of $PSL_2(\mathbb{Z})$ is properly discontinuous, which means that for every point $p \in \mathbb{H}$, there is a neighbourhood U such that $gU \cap U = \emptyset$ for all but finitely many g.*

We can give $\mathcal{A}_1$ the quotient topology where $U \subseteq \mathcal{A}_1$ is open if and only its pullback to $\mathbb{H}$, under the projection $\pi : \mathbb{H} \to \mathcal{A}_1$ is open.

Proposition 5.1.8. *The topology on $\mathcal{A}_1$ is Hausdorff. In fact, it is homeomorphic to $\mathbb{C}$*

Proof. The first statement follows immediately from the last corollary. Using the above results, one can see that A_1 is obtained by gluing the two bounding lines of F and folding the circlular boundary in half. This is easily seen to be homeomorphic to the sphere minus the north pole. □

$\mathcal{A}_1$ has a natural compactification $\bar{\mathcal{A}}_1$ given by adding single point at infinity to make it a sphere. We will follow the convention of the automorphic form literature and call it a *cusp*. It is important to keep in mind that this clashes with the usual terminology in algebraic geometry, that a cusp is a singularity of the form $y^2 = x^3$. We will refer the last thing as cuspidal singularity in order to avoid confusion. We can construct this a quotient as follows. Let $\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\} \subset \mathbb{P}^1$. The action of Γ on $\mathbb{P}^1$ stabilizes $\mathbb{H}^*$. On $\mathbb{H}$ it coincides with the standard action, and on $\mathbb{Q} \cup \{\infty\}$ it consists of a single orbit. Thus $\Gamma\backslash\mathbb{H}^* = \bar{\mathcal{A}}_1$ as a set. In order to get the correct topology on the quotient, one needs a somewhat exotic topology of $\mathbb{H}^*$. On $\mathbb{H}$ it's the usual one, but on $\partial\mathbb{H}^*$ a fundamental system of punctured neighbourhoods of (a translate of) ∞ are (translates of) strips $\operatorname{Im} z > n$, $n \in \mathbb{N}$. These can be visualized as interiors of circles tangent to the boundary circle $\partial\mathbb{H}$.

5.2 Modular forms

Since $\mathcal{A}_1$ has a topology, we can talk about continuous functions on it. We can see that $f : A_1 \to \mathbb{C}$ is continuous if and only if it's pullback $\pi^* f := f \circ \pi$ is continuous. Let us also declare that a function on an open subset of $\mathcal{A}_1$ is holomorphic or meromorphic if its pullback to $\mathbb{H}$ has the same property. This means that such functions correspond to Γ-invariant functions on $\mathbb{H}$. Before constructing nontrivial examples, we want to relax the condition. We say that f is automorphic, with automorphy factor $\phi_\gamma(z)$, if it satisfies the functional equation

$$f(\gamma z) = \phi_\gamma(z) f(z)$$

This is very similar to what we did with theta functions. If we have two such functions with the same factor, their ratio would be invariant. Note that for this to work, we need to impose a consistency condition

$$\phi_{\gamma\xi}(z) f(z) = f(\gamma\xi z)$$

$$= \phi_\gamma(\xi z) f(\xi z) = \phi_\gamma(\xi z)\phi_\xi(z) f(z)$$

Cancelling f, leads to a so called cocycle condition on the automorphy factor

$$\phi_{\gamma\xi}(z) = \phi_\gamma(\xi z)\phi_\xi(z)$$

As the terminology suggests, ϕ_γ does give an element of a certain cohomology group. Rather than pursuing this direction, let us look for natural automorphic forms/factors in nature. Given a meromorphic differential form $\omega = f(z)dz$ on $\mathbb{H}$, let us see how it transforms under $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. We can see that

$$\omega \mapsto f(\gamma \cdot z) d\left(\frac{az+b}{cz+d}\right) = (cz+d)^{-2} f(\gamma \cdot z) dz$$

We say that $f(z)$ is a weakly modular form of weight 2, with respect to Γ, if $f(z)dz$ is invariant. We say that f is *weakly modular of weight* $2k$ if it

$$f(z) = (cz+d)^{-2k} f\left(\frac{az+b}{cz+d}\right) \tag{5.2}$$

This means that the tensor $f(z)dz^{\otimes k}$ is invariant. More generally, it makes sense to consider weakly modular forms of arbitrary integer weight ℓ, satisfying

$$f(z) = (cz+d)^{-\ell} f\left(\frac{az+b}{cz+d}\right)$$

However, when ℓ is odd, taking $\gamma = -I$, shows that $f = -f$, so it's zero! Natural nonzero examples do exist for other groups however, as we shall see shortly.

To drop the "weakly", we impose holomorphy conditions on $\mathbb{H}$ but also at infinity. To understand what the last part means, we first note that by using S and T, (5.2) is equivalent to

$$\begin{aligned} f(z+1) &= f(z) \\ f(-1/z) &= z^{2k} f(z) \end{aligned} \tag{5.3}$$

The first condition means that we have a Fourier expansion

$$f(z) = \sum_{-\infty}^{\infty} a_n e^{2\pi i n z} = \sum_{-\infty}^{\infty} a_n q^n$$

where $q = e^{2\pi i z}$. Note that as $z \to i\infty$, $q \to 0$. So we want to think of q as the local parameter at infinity. Then the Fourier series becomes the Laurent series in q. f is a *modular form of weight* $2k$ if it is holomorphic in $\mathbb{H}$, (5.2) holds, and the Fourier coefficients $a_n = 0$ for $n < 0$. It is called a *cusp form* of weight $2k$ if in addition $a_0 = 0$.

Theorem 5.2.1. *The Eisenstein series*

$$G_{2k}(z) = \sum_{\mathbb{Z}^2 - 0} \frac{1}{(mz+n)^{2k}}$$

is a modular form of weight $2k$*, when* $k \geq 2$*.*

$$\Delta(z) = (60G_4(z))^3 - 27(140G_6(z))^2$$

is a cusp form of weight 12*.*

Proof. The sum can be seen to converge uniformly on compact sets, so it must converge to a holomorphic function on $\mathbb{H}$. One has

$$G_{2k}(\frac{az+b}{cz+d}) = (cz+d)^{2k} \sum \frac{1}{(ma+ndc)z + (mb+nd))^{2k}}$$

The vectors $(ma + ndc, mb + nd)$ can be seen to run over $\mathbb{Z}^2 - 0$. So the right side can be rewritten as

$$(cz + d)^{2k} G_{2k}(z)$$

as required.

We have to check holomorphicity at infinity. By uniform convergence, we can evaluate the limit as $z \to \infty$ term by term. When $m \neq 0$, have $(mz+n)^{-2k} \to 0$ as $z \to \infty$. Therefore

$$\lim_{z\to\infty} G_{2k}(z) = 2\sum_{n=1}^{\infty} \frac{1}{n^{2k}} = 2\zeta(2k)$$

where ζ is the Riemann zeta function. Euler gave explicit formulas for the values

$$\zeta(4) = \frac{\pi^4}{90}$$

$$\zeta(6) = \frac{\pi^6}{945}$$

This allows us to evaluate $\lim_{z\to\infty} \Delta(z)$ and check that it's zero. □

Corollary 5.2.2.

$$j(z) = 1728\frac{(60G_4(z))^3}{\Delta}$$

is weakly modular of weight 0.

The strange normalization is explained by the next result.

Proposition 5.2.3. *The function $j(z)$ is holomorphic on $\mathbb{H}$, and it has q-expansion*

$$j(z) = \frac{1}{q} + 744 + \dots$$

where the series has integer coefficients.

See Serre [A Course in Arithmetic] for the proof.

5.3 Modular curves

We want to make $\bar{\mathcal{A}}_1$ into a Riemann surface in a natural way. We already have a topology on it. We just have to say what holomorphic functions are. Let $\Gamma(1) = SL_2(\mathbb{Z})$. Given $U \subset \mathcal{A}_1$, let $\tilde{U} \subset \mathbb{H}$ denote its preimage with projection $\pi : \tilde{U} \to U$. Let us say that $f : U \to \mathbb{C}$ is holomorphic if and only if $f \circ \pi : U \to \mathbb{C}$ is holomorphic in the usual sense. We define $q = e^{2\pi i}$ to be the coordinate at ∞. So function is holomorphic at ∞ if it can be expanded as a power series in q, or equivalently, the Fourier expansion of the pullback to (an open subset of) $\mathbb{H}$, has no negative coefficients.

Proposition 5.3.1. *$\bar{\mathcal{A}}_1$ is a Riemann surface.*

Sketch. The key point is to show that any point $x \in \bar{\mathcal{A}}_1$ has a neighbourhood D with a homeomorphism z, called a local coordinate or parameter, to a disk in $\mathbb{C}$, such that holomorphic functions on both disks coincide. There are three cases: $x = \infty$, x is an image of one of the fixed points $i, e^{2\pi i/3}$, or x is any other point. The first case was done above. The third case is straight forward. The map $\pi : \mathbb{H} \to X(1)$ is unramified over x A local coordinate z at a point $y \in \mathbb{H}$ lying over x will give a local coordinate at x. The map π is ramified at i and $e^{2\pi i/3}$ with ramification index $r = 2$ and 3 respectively. z^r will give a local coordinate at the image. □

The importance of $j(z)$ stems from the following.

Corollary 5.3.2. *Two elliptic curves E_τ and $E_{\tau'}$ are isomorphic if and only if $j(\tau) = j(\tau')$.*

Proof. By proposition 5.2.3 j factors through a holomorphic map $\mathcal{A}_1 \to \mathbb{C}$, with pole of order 1 at ∞. Therefore j induces a holomorphic map $\bar{\mathcal{A}}_1 \to \mathbb{P}^1$ of degree 1, which is necessarily bijective. Since $\mathcal{A}_1$ is the set equivalence classes of elliptic curves, the corollary follows.

□

While it's intuitively clear what it means that $\mathcal{A}_1$ parameterizes elliptic curves, the actual statement requires a bit more precision. Let us define an analytic family of (compact) complex manifolds to be a (proper) holomorphic submersion of complex manifolds $f : E \to B$. We recall that a submersion is map such that derivative is surjective on tangent spaces. This implies that fibres $E_b = f^{-1}(b)$ are complex submanifolds. By an analytic family of elliptic curves we mean an analytic family of compact complex manifolds $f : E \to B$ with a holomorphic section $s : B \to E$ such that each fibre E_b is a compact Riemann surface of genus one. We can regard E_b as an elliptic curve with origin $s(b)$. Given an elliptic curve $E = E_\tau$, set $j(E) = j(\tau)$.

Theorem 5.3.3. *$\mathcal{A}_1$ has the following properties:*

(a) The map $E \mapsto j(E)$ gives a bijection between the set of isomorphism classes of elliptic curves over $\mathbb{C}$ and points of $\mathcal{A}_1$.

(b) Given an analytic family elliptic curves $E \to B$, the map $B \to \mathcal{A}_1$, called the classifying map, given by $b \mapsto j(E_b)$ is holomorphic.

One might hope for stronger property:

Question 5.3.4. *Does there exists a universal family of elliptic curves over $\mathcal{A}_1$, i.e. a family of elliptic curves such that any other family is obtained by pulling it back with respect to j?*

It would be very desirable, for a number of reasons, to have an affirmative answer (in which case $\mathcal{A}_1$ would be called a *fine moduli space*). Unfortunately, the answer is no, as shown by the following example:

Example 5.3.5. *If $\mathcal{A}_1$ were a fine moduli space, then any family of elliptic curves with constant j-invariant would be trivial. However, let E be either E_i or $E_{\exp(2\pi i/3)}$ Either curve has a nontrivial automorphism group G, which is cyclic in both cases. Choose a manifold $\tilde{B}$ on which G acts freely, e.g. $\mathbb{C}^*$. The quotient $(E \times \tilde{B})/G \to \tilde{B}/G$ is a nontrivial family with constant j-invariant.*

There are a number of things we could do at this point:

1. State the precise universal property that $\mathcal{A}_1$ satisfies. This is weaker than the existence of a universal family, but stronger than what the last theorem says.

2. Explain what *stack* is. This addresses the lack of fineness in a different way.

3. Show how to add additional structure to get a universal family.

We will follow the 3rd option, which the easiest to explain. The reason why example 5.3.5 works is because the elliptic curves E_i or $E_{\exp(2\pi i/3)}$ have extra automorphisms. This is closely related to the fact that $SL_2(\mathbb{Z})$ has torsion and that it does not act freely on $\mathbb{H}$. The solution to both problems is to pass to subgroup. Given an integer $N > 0$, the *principal congruence subgroup of level* N of $\Gamma(1) = SL_2(\mathbb{Z})$ is

$$\Gamma(N) = \ker[\Gamma(1) \to SL_2(\mathbb{Z}/N)] = \{M \in \Gamma(1) \mid M \equiv I \mod N\}$$

Lemma 5.3.6. *If $N \geq 3$, $\Gamma(N)$ is torsion free and it acts freely on $\mathbb{H}$.*

Proof. This follows easily from corollary 5.1.6 □

Given such a group, it will act on $\mathbb{H}^*$, let $Y(N) = \Gamma(N)\backslash\mathbb{H}$ and let $X(N) = \Gamma'\backslash\mathbb{H}^*$. $X(N)$ can be made into a compact Riemann surface the same way as we did for $\bar{\mathcal{A}}_1$, and $Y(N) \subset X(N)$ is an open subset, and the points of $X(N) - Y(N)$ are called cusps. Given an elliptic curve E, the group of N-torsion points is isomorphic to $(\mathbb{Z}/N)^2$. A *level* N-structure for E is a basis for this group, i.e. a pair of N-torsion points which generates it.

Theorem 5.3.7. *If $N \geq 3$, $Y(N)$ is a fine moduli space parameterizing pairs (E, L), where E is an elliptic curve with a level N-structure.*

Sketch. Let G be the group of holomorphic automorphisms of $\mathbb{H} \times \mathbb{C}$ generated by $\Gamma(N)$ acting on the first factor and the translations

$$(\tau, z) \mapsto (\tau, z + 1)$$

$$(\tau, z) \mapsto (\tau, z + \tau)$$

As an abstract group G is the semidirect product $\Gamma(N) \rtimes \mathbb{Z}^2$. Using lemma 5.3.6, we can see that G acts freely on $\mathbb{H} \times \mathbb{C}$. Therefore $\mathcal{E}(N) = G\backslash(\mathbb{H} \times \mathbb{C})$ is a complex manifold. The projection on the first factor gives a holomorphic

map $\mathcal{E}(N) \to Y(N)$ which can be seen to a family of elliptic curves. The maps $\tau \mapsto 1/N$ and $\tau \mapsto \tau/N$ give sections $\sigma_i : Y(N) \to \mathcal{E}(N)$ which give level N-structures on each fibre. The family $(\mathcal{E}(N), \sigma_1, \sigma_2)$ can be checked to be the universal family of elliptic curves with level N-structure. □

Theorem 5.3.8. *When $N \geq 3$, the genus of $X(N)$ is*

$$g = 1 + \frac{d(N-6)}{12N}$$

where

$$d = \frac{1}{2}|SL_2(\mathbb{Z}/N\mathbb{Z})|$$

(There are standard formulas for computing the order, e.g. if $N = p$ is prime, then $|SL_2(\mathbb{Z}/p\mathbb{Z})| = p(p^2-1)$.)

Proof. We have a holomorphic map

$$\pi : X(N) \to X(1) = \mathbb{P}^1$$

induced by inclusion $\Gamma(N) \subset \Gamma(1) = SL_2(\mathbb{Z})$. This is a branched covering. So we can compute the genus using the Riemann-Hurwitz formula, which says that if $Y \to X$ is a degree d branched covering of compact Riemann surfaces of genus $g(Y)$ and $g(X)$, then

$$2g(Y) - 2 = (2g(X) - 2)d + \underbrace{\sum_{y \in Y}(e_y - 1)}_{R}$$

where e_y is the ramification index which counts the number of sheets which "come together" at y.

The covering $\pi : X(N) \to X(1)$ is Galois with group $G = PSL_2(\mathbb{Z})/\operatorname{im}\Gamma(N)$. Let $\bar{S}, \bar{T} \in G$ denote the images of S and T. The degree of this covering $|G| = d$, when $N \geq 3$. Let p_2 and p_3 represent the images of i and $e^{2\pi i/3}$ in $X(1)$. Then p_2, p_3, ∞ are the ramification points. Given one of these points p, and $q \in \pi^{-1}(p)$, e_q is the order of the isotropy group $G_q = \{g \in G \ gq = q\}$. This is independent of q, because all the isotropy groups are conjugate. For $p = p_2$ and suitable q, $G_q = \langle \bar{S} \rangle$ so that $e_q = 2$. We can calculate the other values in a similar way to get

$$p = p_2, G_q = \langle \bar{S} \rangle, e_q = 2$$
$$p = p_3, G_q = \langle \bar{S}\bar{T} \rangle, e_q = 3$$
$$p = \infty, G_q = \langle \bar{T} \rangle, e_q = N$$

We also have $|\pi^{-1}(p)| = d/|G_q|$. This allows to calculate the ramification term above to get

$$R = \frac{d}{2} + \frac{2d}{3} + \frac{d(N-1)}{N}$$

Putting this into Riemann-Hurwitz and simplifying proves the theorem.

□

Corollary 5.3.9. *There are nonzero modular forms of weight* 2 *for* $\Gamma(N)$ *as soon as* $N \geq 6$.

If X is a compact Riemann surface of genus $g \geq 2$, then Hurwitz showed that the automorphism group satisfies

$$|Aut(X)| \leq 84(g-1)$$

The next example shows that this bound is sharp.

Example 5.3.10. *The group* $PSL_2(\mathbb{Z}/7\mathbb{Z})$ *has cardinality* 168 *and it acts on* $X(7)$. *By the above formula, it has genus* 3.

It is useful to consider a generalization as follows. Let $\Gamma \subseteq SL_2(\mathbb{Z})$ be a finite index subgroup. Then Γ will act on $\mathbb{H}$ and $\mathbb{H}^*$ as before, and the quotient $\Gamma\backslash\mathbb{H}$ (resp. $\Gamma\backslash\mathbb{H}^*$) can be made into a (compact) Riemann surface. This construction is particularly interesting to number theorists, when Γ is a *congruence* group, which means that it contains some $\Gamma(N)$. In this case, the quotients are called modular curves. These curves $\Gamma\backslash\mathbb{H}$ will parameterize elliptic curves with extra structure. For example, for

$$\Gamma_1(N) = \{M \in \Gamma(1) \mid M \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \mod N\}$$

the extra structure is the choice of a point of order N. For more information about these curves, and their applications to number theory, see [Diamond, Shurman, A First Course in Modular Forms]. One of the more dramatic applications was the proof of Fermat's Last Theorem by Wiles about 30 years ago. Some, but by no means all, of the ingredients are explained in that book.

5.4 Belyi's theorem

Given a subfield $K \subset \mathbb{C}$, we say that complex algebraic variety is definable over K, if we can choose coefficients of the defining equations to lie in K. This becomes particularly interesting when $K = \mathbb{Q}$ or $\bar{\mathbb{Q}}$.

Example 5.4.1. *There exist elliptic curves which are not defined over* $\bar{\mathbb{Q}}$, *and in fact most are not. This comes down to the simple observation that* $\mathcal{A}_1$ *has uncountably many points but only countably many could correspond to elliptic curves defined over* $\bar{\mathbb{Q}}$.

There is a surprisingly simple criterion for when a curve can be defined over $\bar{\mathbb{Q}}$, which was discovered not that long ago.

Theorem 5.4.2 (Belyi, 1979)**.** *Let* X *be a smooth projective curve over* $\mathbb{C}$. *The following are equivalent.*

(1) X *is definable over* $\bar{\mathbb{Q}}$.

(2) There exists a finite index subgroup $\Gamma \subseteq SL_2(\mathbb{Z})$ *such that* $X \cong \Gamma\backslash\mathbb{H}^*$.

(3) There exists a finite holomorphic map $\pi : X \to \mathbb{P}^1$ *unramified outside of* $\{0, 1, \infty\}$.

Proof. We note that (2) $\Rightarrow$ (1) is not at all obvious, and it follows from a deep result of Grothendieck[1] that if Y is a smooth curve over an algebraically closed field K of characteristic 0, then the category of étale covers of Y is equivalent to the category of étale covers of $Y \times_{Spec\, K} Spec\, L$ for any algebraically closed field $L \supset K$. For the implication, one applies this with $Y = \mathbb{P}^1 - \{0, 1, \infty\}$, $K = \bar{\mathbb{Q}}$ and $L = \mathbb{C}$. We just outline the equivalence of (2) and (3); for the remaining step , (1) $\Rightarrow$ (3), see [Serre, Lectures on the Mordell-Weil theorem].

For (2) $\Rightarrow$ (3), the inclusion $\Gamma \subseteq SL_2(\mathbb{Z})$ induces a map $X = \Gamma\backslash\mathbb{H}^* \to SL_2(\mathbb{Z})\backslash\mathbb{H}^* = \mathbb{P}^1$. This is unramified outside of 3 points corresponding to the images $i, e^{2\pi i/3}, \infty$. After composing with an automorphism of the line, we can move these points to $0, 1, \infty$

For (3) $\Rightarrow$ (2), we need to know that we can find subgroup $\Gamma_2 \subset SL_2(\mathbb{Z})$ of finite index, which acts freely on $\mathbb{H}$, such that $\Gamma_2\backslash\mathbb{H} \cong \mathbb{P}^1 - \{0, 1, \infty\}$. Suppose that $\pi : X \to \mathbb{P}^1$ unramified outside of $\{0, 1, \infty\}$ Then it follows that $Y = \pi^{-1}(\mathbb{P}^1 - \{0, 1, \infty\})$ is of the form $Y = \Gamma\backslash\mathbb{H}$ for some $\Gamma \subseteq \Gamma_2$ of finite index. Since $\Gamma\backslash\mathbb{H}^*$ is the unique Riemann surface compactification of Y, $X = \Gamma\backslash\mathbb{H}^*$. □

[1] SGA1: Revêtement étales et groupe fondamental