Modular forms and Galois representations

by

Tracey Chen

Supervised by
Dr. Alex Ghitza

Submitted in partial fulfilment of the requirements
for the degree Master of Science in the
School of Mathematics and Statistics

The University of Melbourne
2022

Abstract

Modular forms are a central object of study in modern Arithmetic Geometry. We review the classical theory of modular forms and their associated Galois representations, with the aim of understanding the Modularity Theorem. To this end we will also study elliptic curves, modular curves, and moduli spaces. Finally, we will provide a brief overview of the generalisation to modular forms modulo a prime power.

Preface

In the latter half of the 20th century, surprising links were found between modular forms and elliptic curves [30]. Versions of the Modularity Theorem, also known as the Taniyama-Shimura-Weil conjecture, were first conjectured in the 1950's [7]. The theorem posits that "all rational elliptic curves arise from modular forms" [7]. In 1995, the result was proved for a large class of elliptic curves by Wiles and Taylor [31] [27], providing a proof of Fermat's Last Theorem. A proof of the full conjecture was given in 2001 [3].

We will explain a version of the Modularity Theorem: For any elliptic curve E over $\mathbb{Q}$ there exists a newform f such that the Fourier coefficients $a_p(f)$ are equal to

$$a_p(E) := \left(\begin{array}{c} \text{the number of solutions of the Weierstrass} \\ \text{equation E working modulo } p \end{array} \right) - 1.$$

The focus will be to first develop enough machinery to understand the theorem precisely. To do so, we will need to study modular forms, Hecke operators, modular curves, elliptic curves, and representations of the absolute Galois group. The relationships between these objects give rise to various forms of the statement of the Modularity Theorem.

We will then briefly discuss the extension of the notion of modularity to Galois representations modulo a prime power p^m, providing a summary of recent developments in this area. In particular, we will look at generalisations to the mod p^m case of the level raising and level lowering theorems of Ribet [20] from the mod p case.

Contents

1 Introduction and basic definitions

Modular forms are particular holomorphic functions on the complex upper half-plane that satisfy certain transformation properties under the action of the modular group. They arise in connection to various areas of mathematics such as number theory, geometry, and mathematical physics [17]. The transformation property makes the function a section of a line bundle on a modular curve.

This chapter provides a summary of the background material concerning classical modular forms and their closely related objects, namely elliptic curves and modular curves, loosely based on the first chapter of [7]. We will begin with a definition of modular forms with respect to $\mathrm{SL}_2(\mathbb{Z})$, then generalise this definition to congruence subgroups $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ in section 1.2.

Following this, in Section 1.3 we will define modular forms of character χ. Specifically, χ will be a Dirichlet character modulo some positive integer N, a group homomorphism $(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$.

In Section 1.4, we will introduce complex elliptic curves as the quotient of the complex plane by a lattice. Geometrically, this quotient is a compact Riemann surface of genus 1, i.e. a complex torus, and algebraically, it forms an abelian group. By embedding the torus in the complex projective plane $\mathbb{C}P^2$, we will see that it is in bijective correspondence with the set of ordered pairs satisfying an equation of the form

$$y^2 = 4x^2 - g_2 x - g_3, \quad g_2, g_3 \in \mathbb{Z},\ g_2^2 - 27g_3^2 \neq 0.$$

The map is an isomorphism of Riemann surfaces, and by equipping the curve with an appropriate group structure, an isomorphism of groups.

In section 1.5, we will introduce the moduli curves as the quotient of the upper half-plane by the action of a congruence subgroup. Such curves are in bijection with moduli spaces, equivalence classes of complex elliptic curves equipped with appropriate torsion data.

1.1 Modular forms

The set of transformations of interest are the elements of the *modular group*,

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,\ b,\ c,\ d \in \mathbb{Z},\ ad - bc = 1 \right\}$$

which act on $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ via

$$z \mapsto \frac{az+b}{cz+d}.$$

It can be easily seen that for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, the corresponding transformation maps the upper half-plane $\mathbb{H} = \{\tau \in \mathbb{C} \ : \ \mathrm{Im}(\tau) > 0\}$ to itself as

$$\mathrm{Im}(\gamma(\tau)) = \frac{\mathrm{Im}((a\tau+b)(c\bar{\tau}+d))}{|c\tau+d|^2} = \frac{(ad-bc)\mathrm{Im}(\tau)}{|c\tau+d|^2} = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Thus, this action is well-defined on $\mathbb{H}$. We will often work instead with $\mathrm{PSL}_2(\mathbb{R})/\{\pm 1\}$ as the matrices γ and $-\gamma$ act equivalently on $\mathbb{H}$.

The modular group is generated by the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Analogously, the group of transformations is generated by the maps

$$\tau \mapsto \tau + 1 \quad \text{and} \quad \tau \mapsto -1/\tau.$$

Definition 1.1. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. The **factor of automorphy** is

$$j(\gamma, \tau) = c\tau + d.$$

Definition 1.2. Let k be an integer and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. The **weight-$k$ operator** $[\gamma]_k$ on functions $f : \mathbb{H} \to \mathbb{C}$ is defined by

$$(f[\gamma]_k)(\tau) = j(\gamma, \tau)^{-k} f(\gamma(\tau))$$

Definition 1.3. Let $k \in \mathbb{Z}$. A meromorphic function $f : \mathbb{H} \to \mathbb{C}$ is **weakly modular of weight k** if for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, we have that

$$f[\gamma]_k = f.$$

That is, for all $\tau \in \mathbb{H}$,

$$f(\gamma(\tau)) = (c\tau + d)^k f(\tau).$$

We will refer to this property as the *modular transformation property.*

A modular form is a weakly modular form that is holomorphic on $\mathbb{H}$ and at ∞. To make precise what it means to be holomorphic at ∞, recall that the matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

is contained in $\mathrm{SL}_2(\mathbb{Z})$, corresponding to the translation $\tau \mapsto \tau + 1$. If f is weakly modular, then the rule for weak modularity requires that $f(\tau + 1) = f(\tau)$, so every weakly modular function is $\mathbb{Z}$-periodic.

Let $D = \{q \in \mathbb{C} : |z| < 1\}$ be the open unit disc in $\mathbb{C}$. The map $\tau \mapsto q := e^{2\pi i z}$ is holomorphic, $\mathbb{Z}$-periodic, and maps $\mathbb{H}$ onto the punctured disc, $D - \{0\}$. For any weakly modular form f, there is a corresponding map $g : D - \{0\} \to \mathbb{C}$ where $g(q) = f(\log(q)/(2\pi i))$. If f is holomorphic on $\mathbb{H}$, then g is holomorphic on the punctured disc since the logarithm can be defined holomorphically about each point. Thus, g has a Laurent expansion

$$g(q) = \sum_{n \in \mathbb{Z}} a_n q^n$$

for $q \in D - \{0\}$.

Notice that $|q| = |e^{2\pi i(\mathrm{Re}(\tau) + \mathrm{Im}(\tau))}| = e^{-2\pi \mathrm{Im}(\tau)} \to 0$ as $\mathrm{Im}(\tau) \to \infty$. Thinking of the point at ∞ as lying in the imaginary direction of $\mathbb{H}$, we provide the following definition.

Definition 1.4. f is **holomorphic at** ∞ if g extends holomorphically to the point $q = 0$.

If f is holomorphic on $\mathbb{H}$ and holomorphic at ∞, then f has a Fourier expansion (also known as a **q-expansion**)

$$f(\tau) = \sum_{n\in\mathbb{Z}} a_n(f)q^n, \quad q = e^{2\pi i\tau}.$$

Definition 1.5. Let f be a weakly modular form of weight k. f is a **modular form of weight** k if

i. f is holomorphic on $\mathbb{H}$, and

ii. f is holomorphic at ∞.

The set of modular forms of weight k is denoted $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

$\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ forms a vector space over $\mathbb{C}$, and the sum

$$\mathcal{M}(\mathrm{SL}_2(\mathbb{Z})) = \bigoplus_{k\in\mathbb{Z}} \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$$

forms a graded ring.

Example. For even integers $k > 2$, the *Eisenstein series* of weight k is a 2-dimensional analogue of the Riemann zeta function and is given by

$$G_k(\tau) = \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(c\tau + d)^k}, \quad \tau \in \mathbb{H}.$$

The series absolutely converges to a holomorphic function on $\mathbb{H}$ and it can be shown with some work that its Fourier expansion is

$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n, \quad q = e^{2\pi i\tau}$$

where the coefficients are

$$\sigma_{k-1}(n) = \sum_{\substack{m|n \\ m>0}} m^{k-1}.$$

It follows that G_k is holomorphic at ∞.

For any $\gamma = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, we check the modular transformation property, computing

$$\begin{aligned} G_k(\gamma(\tau)) &= \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{\left(c\left(\frac{a'\tau+b'}{c'\tau+d'}\right) + d\right)^k} \\ &= \left(c'\tau + d'\right)^k \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{\left(c\left(a'\tau + b'\right) + d(c'\tau + d)\right)^k} \end{aligned}$$

$$= \left(c'\tau + d'\right)^k \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{\left(\left(ca'+dc'\right)\tau + \left(cb'+dd'\right)\right)^k}$$

Notice that $\begin{pmatrix} ca'+dc' & cb'+dd' \end{pmatrix} = \begin{pmatrix} c & d \end{pmatrix}\gamma$. Since $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, γ is invertible, and so as (c,d) runs through $\mathbb{Z}^2 - \{(0,0)\}$, so does $(ca'+dc', cb'+dd')$. Therefore, $G_k(\gamma(\tau)) = (c\tau+d)^k G_k(\tau)$ and G_k is weakly modular of weight k. Thus, G_k is a modular form of weight k for $k > 2$.

Definition 1.6. If f satisfies the additional condition that

iii. $a_0(f) = 0$,

then we say f is a **cusp form**.

The set of cusp forms of weight k is denoted $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$.

$\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ forms a vector subspace of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, and the graded ring

$$\mathcal{S}(\mathrm{SL}_2(\mathbb{Z})) = \bigoplus_{k\in\mathbb{Z}} \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})$$

is an ideal in $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$.

Example. Let $g_2(\tau) = 60G_4(\tau)$ and let $g_3(\tau) = 140G_6(\tau)$. Define the discriminant function

$$\begin{aligned} \Delta : \mathbb{H} &\to \mathbb{C} \\ \tau &\mapsto g_2(\tau)^3 - 27g_3(\tau)^2. \end{aligned}$$

The discriminant function is holomorphic on $\mathbb{H}$ and weakly modular of weight 12 with q-expansion

$$\Delta(\tau) = (2\pi)^1 2 \sum_{n=1}^{\infty} \tau(n) q^n, \quad q = e^{2\pi i \tau}$$

where τ is Ramanujan's tau function for which $\tau(1) = 1$. Therefore Δ is a nonzero element in $\mathcal{S}_{12}(\mathrm{SL}_2(\mathbb{Z}))$. One can also show that Δ is nonvanishing on $\mathbb{H}$. This shows that the modular invariant to be defined below is holomorphic on $\mathbb{H}$.

Definition 1.7. The **modular invariant** or ***j*-invariant** is the modular function

$$j : \mathbb{H} \to \overline{\mathbb{C}}, \quad j = 1728\frac{g_2^3}{\Delta}$$

The numerator and denominator of j are both of weight 12 so j must be $\mathrm{SL}_2(\mathbb{Z})$-invariant, as its name suggests. j has a simple pole at ∞ so it is not a modular form. Nonetheless, the j-invariant will play an important role in our study of elliptic curves where we will revisit both the discriminant and the j-invariant in a more general context.

Remark. The theory of modular forms can be extended to automorphic forms by relaxing the holomorphicity condition, and instead only requiring the function be meromorphic on both the upper half-plane and at the cusps. In this setting, j is an automoprhic form of weight 0 with respect to $\mathrm{SL}_2(\mathbb{Z})$.

Remark. The leading term 1728 normalises the Laurent series of j to be

$$j(\tau) = \frac{1}{q} + \sum_{n=0}^{\infty} a_n q^n, \quad q =^{2\pi i\tau}$$

where the coefficients a_n are integers.

1.2 Congruence subgroups and their associated modular forms

The notion of weak modularity can be generalised to certain subgroups $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$, allowing us to define modular forms with respect to Γ.

Definition 1.8. Let N be a positive integer. The **principal congruence subgroup of level** N is

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

where the matrix congruence is entry-wise.

$\Gamma(N)$ is the kernel of the natural homomorphism $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ where we take congruences entry-wise. As this map is surjective, it induces an isomorphism $\mathrm{SL}_2(\mathbb{Z})/\Gamma(N) \cong \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Thus, the index $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)]$ is finite and it can be shown that

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = |\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})| = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right)$$

where the product is over all prime divisors of N (see Section 1.2 in [7]).

Definition 1.9. A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a **congruence subgroup of level** N if $\Gamma(N) \subset \Gamma$.

Example. Aside from the principal congruence subgroups, the congruence subgroups of interest here are

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

where $\star$ means the value is unspecified, and

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

Lemma 1.10. *For any positive integer N, we have*

$$\Gamma_1(N)/\Gamma(N) \cong \mathbb{Z}/N\mathbb{Z}$$

and

$$\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/N\mathbb{Z})^{\times}.$$

Proof. The map $\Gamma_1(N) \to \mathbb{Z}/N\mathbb{Z}$ defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto b \pmod N$$

is surjective with kernel $\Gamma(N)$, so the third isomorphism theorem gives $\Gamma_1(N)/\Gamma(N) \cong \mathbb{Z}/N\mathbb{Z}$. Similarly the map $\Gamma_0(N) \to (\mathbb{Z}/N\mathbb{Z})^\times$ defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \pmod N$$

is surjective with $\Gamma_1(N)$ lying in the kernel. To see that the kernel is in fact equal to $\Gamma_1(N)$, notice that if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$, then $1 = \det(\gamma) \equiv ad \pmod N$. If γ is in the kernel, then $d \equiv 1 \pmod N$, so we also must have $a \equiv 1 \pmod N$. □

Definition 1.11. A meromorphic function $f : \mathbb{H} \to \mathbb{C}$ is **weakly modular of weight k with respect to** Γ if it is weight-k invariant. That is, if

$$f[\gamma]_k = f$$

for all $\gamma \in \Gamma$.

Since $c\tau + d$ is always finite and never zero (since the determinant of γ is non-zero), $f[\gamma]_k$ always has the same zeros and poles as f. The following are some basic properties.

Lemma 1.12. *For all* $\gamma, \gamma' \in \mathrm{SL}_2(\mathbb{Z})$ *and* $\tau \in \mathbb{H}$,

i. $j(\gamma\gamma', \tau) = j(\gamma, \gamma'(\tau))j(\gamma', \tau)$,

ii. $(\gamma\gamma')(\tau) = \gamma(\gamma'(\tau))$,

iii. $[\gamma\gamma']_k = [\gamma]_k[\gamma']_k$,

iv. $\mathrm{Im}(\gamma(\tau)) = \frac{\mathrm{Im}(\tau)}{|j(\gamma,\tau)|^2}$,

v. $\frac{d\gamma(\tau)}{d\tau} = \frac{1}{j(\gamma,\tau)^2}$.

Finally, to define modular forms with respect to a congruence subgroup we want to extend our definition of holomorphicity at ∞. If Γ is a congruence subgroup of level N, meaning it contains $\Gamma(N)$, and hence contains the matrix

$$\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}.$$

Due to the modular transformation property, it follows that every weakly modular form f with respect to Γ must satisfy $f(\tau + N) = f(\tau)$. We follow steps identical to those in Section 1.1 except that we consider the holomorphic and $N\mathbb{Z}$-periodic map $\tau \mapsto e^{2\pi i\tau/N}$ and $q = q_N = e^{2\pi i\tau/N}$. It follows that f has a Laurent expansion of the form

$$f(\tau) = \sum_{n=0}^{\infty} a_n q_N^n, \quad q_N = e^{2\pi i\tau/N}.$$

Instead of adjoining only a single point at infinity, for a congruence subgroup Γ, we adjoin $\mathbb{Q} \cup \{\infty\} = \mathbf{P}^1(\mathbb{Q})$ forming the extended upper half-plane

$$\mathbb{H}^* := \mathbb{H} \cup \mathbf{P}^1(\mathbb{Q}),$$

then identify points that are Γ-equivalent where Γ acts on $\mathbf{P}^1(\mathbb{Q})$ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} [x : y] = [ax + by : cx + dy].$$

The elements of $\mathbb{H}^* \setminus \mathbb{H} = \mathbf{P}^1(\mathbb{Q})$ are called **cusps**. In the case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, all rational numbers are Γ-equivalent, and so we had only one cusp represented by ∞. When Γ is a proper subgroup of $\mathrm{SL}_2(\mathbb{Z})$, this is no longer the case, and Γ will have other cusps represented by equivalence classes of rational numbers.

Definition 1.13. Let $k \in \mathbb{Z}$ and let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. A function $f : \mathbb{H} \to \mathbb{C}$ is a **modular form of weight k with respect to Γ** if

i. f is weakly modular of weight k with respect to Γ,

ii. f is holomorphic on $\mathbb{H}$,

iii. $f[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

The set of modular forms of weight k with respect to Γ are denoted $\mathcal{M}_k(\Gamma)$.

The third condition needs only be checked on a set of cusp representatives. Since $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$ is finite, we need only check on a finite set of transformations.

Definition 1.14. If f satisfies the additional condition that

iv. $a_0 = 0$ in the Fourier expansion of $f[\alpha]_k$ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$,

then we say f is a **cusp form** of weight k with respect to Γ. $\mathcal{S}_k(\Gamma)$ denotes the set of such cusp forms.

1.3 Dirichlet characters and their associated modular forms

Definition 1.15. For any positive integer N, a **Dirichlet character modulo N** is a group homomorphism

$$\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times.$$

Remark. A Dirichlet character is sometimes extended to $\mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ (or $\mathbb{Z} \to C$) by setting $\chi(a) = 0$ if $\gcd(a, N) > 1$.

If χ and ψ are any two Dirichlet characters modulo N, then their product is defined by $(\chi\psi)(n) = \chi(n)\psi(n)$ for all $n \in (\mathbb{Z}/N\mathbb{Z})^\times$.

Lemma 1.16. *The set of Dirichlet characters modulo N equipped with the above multiplication rule forms a group.*

Proof. The identity in the group is the trivial character modulo N that maps every element of $(\mathbb{Z}/N\mathbb{Z})^\times$ to 1. If χ and ψ are any two Dirichlet characters modulo N and $n, m \in (\mathbb{Z}/N\mathbb{Z})^\times$, then

$$\begin{aligned}(\chi\psi)(nm) &= \chi(nm)\psi(nm)\\ &= \chi(n)\chi(m)\psi(n)\psi(m)\\ &= \chi(n)\psi(n)\chi(m)\psi(m)\\ &= (\chi\psi)(n)(\chi\psi)(m)\end{aligned}$$

so the product is again a Dirichlet character modulo N. The multiplicative group $(\mathbb{Z}/N\mathbb{Z})^\times$ has order $\varphi(N)$ where φ is the Euler totient. In particular, $(\mathbb{Z}/N\mathbb{Z})^\times$ is finite, and so the values taken by any Dirichlet character modulo N, say χ, are complex roots of unity, so the inverse is given by complex conjugation. That is, $\chi^{-1}(n) = \overline{\chi(n)}$ for all $n \in (\mathbb{Z}/N\mathbb{Z})^\times$. □

The group of Dirichlet characters modulo N is called the *dual group* $\widehat{(\mathbb{Z}/N\mathbb{Z})}^\times$ of $(\mathbb{Z}/N\mathbb{Z})^\times$.

If p is prime, then $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p-1$. Let g be a generator of $(\mathbb{Z}/p\mathbb{Z})^\times$ and $\zeta \in \mu_{p-1}$ be a primitive $(p-1)$st complex root of unity. Then the group of Dirichlet characters modulo p is generated by the homomorphism taking g to ζ and is also cyclic of order $p-1$. Thus, it is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^\times$.

In fact, this holds in the general case by the following proposition due to Serre [22].

Proposition 1.17. $\widehat{(\mathbb{Z}/N\mathbb{Z})}^\times$ *is (noncanonically) isomorphic to* $(\mathbb{Z}/N\mathbb{Z})^\times$.

If d is a positive divisor of N and let $\pi_{N,d} : (\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times$ denote the natural projection. Then every Dirichlet character χ modulo d lifts to a Dirichlet character $\chi_N = \chi \circ \pi_{N,d}$ modulo N. In the other direction, a Dirichlet character modulo N is not necessarily the lift of some Dirichlet character modulo d.

Definition 1.18. Let χ be a Dirichlet character modulo N. The **conductor** of χ is the smallest positive divisor d of N such that there exists a Dirichlet character χ_d modulo d such that $\chi = \chi_d \circ \pi_{N,d}$.

Equivalently, d is the smallest positive divisor of N such that χ is trivial on the normal subgroup

$$K_{N,d} := \ker(\pi_{N,d}) = \{n \in (\mathbb{Z}/N\mathbb{Z})^\times : n \equiv 1 \pmod{d}\}.$$

Definition 1.19. A Dirichlet character modulo N is **primitive** if its conductor is N.

Definition 1.20. Let $k, N \in \mathbb{Z}$ and let χ be a Dirichlet character modulo N. Then $f : \mathbb{H} \to \mathbb{C}$ is a **modular form of weight k, character χ and level** N if

i. $f[\gamma]_k = \chi(d)f$ for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$

ii. f is holomorphic on $\mathbb{H}$,

iii. $f[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

The set of all such modular forms are denoted $\mathcal{M}_k(N, \chi)$.

Definition 1.21. If f satisfies the additional condition that

iv $a_0 = 0$ in the Fourier expansion of $f[\alpha]_k$ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$,

then we say f is a **cusp form of weight k, character χ and level N**. $\mathcal{S}_k(N, \chi)$ denotes the set of such cusp forms.

Proposition 1.22. *For any positive integer N, we have the decompositions*

$$\mathcal{M}_k(\Gamma_1(N)) = \bigoplus_{\chi:(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}} \mathcal{M}_k(N, \chi)$$
$$\mathcal{S}_k(\Gamma_1(N)) = \bigoplus_{\chi:(\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}} \mathcal{S}_k(N, \chi)$$

where the summations run over all the Dirichlet character modulo N.

Proof. We prove the statement for the set of modular forms – the argument for cusp forms is similar. Recall from Lemma 1.10 that $\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$, so view $\mathcal{M}(\Gamma_1(N))$ as a module over $\Gamma_0(N)/\Gamma_1(N)$. The action of $\Gamma_0(N)$ on $\mathcal{M}_k(\Gamma_1(N))$ defined by by $\gamma \cdot f = f[\gamma]_k$ induces a representation of $\Gamma_0(N)/\Gamma_1(N)$ on $\mathcal{M}_k(\Gamma_1(N))$. Since $\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/N\mathbb{Z})^\times$ is finite and abelian, this representation completely reduces as the sum of irreducible one-dimensional representations of $\Gamma_0(N)/\Gamma_1(N)$. The one-dimensional representations are precisely those induced by the Dirichlet characters modulo N. □

1.4 Complex tori and elliptic curves

A complex torus is the quotient of $\mathbb{C}$ by a lattice $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ where $\{\omega_1, \omega_2\}$ is a basis for $\mathbb{C}$ over $\mathbb{R}$. Algebraically, $\mathbb{C}/\Lambda$ forms an abelian group under the addition inherited from $\mathbb{C}$. Geometrically, it is a compact and connected Riemann surface of genus 1. By convention, we normalise so that $w_1/w_2 \in \mathbb{H}$. However, this does not uniquely specify a lattice.

Lemma 1.23. *Let $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ and $\Lambda' = \omega_1'\mathbb{Z} \oplus \omega_2'\mathbb{Z}$ be lattices. Then $\Lambda = \Lambda'$ if and only if*

$$\begin{pmatrix} \omega_1' \\ \omega_2' \end{pmatrix} = A \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$$

for some $A \in \mathrm{SL}_2(\mathbb{Z})$.

Proof. If $\Lambda = \Lambda'$, then $\omega_1', \omega_2' \in \Lambda$, so there exist integers a, b, c, d such that $\omega_1' = a\omega_1 + b\omega_2$ and $\omega_2' = c\omega_1 + d\omega_2$. Then

$$\begin{pmatrix} \omega_1' \\ \omega_2' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}.$$

Conversely, suppose $\omega_1' = a\omega_1 + b\omega_2$ and $\omega_2' = c\omega_1 + d\omega_2$. For $z \in \Lambda'$, we can write $z = f\omega_2' + g\omega_2'$ for some $f, g \in \mathbb{Z}$. Then $z = (fa + gc)\omega_1 + (fb + gd)\omega_2$ and hence $z \in \Lambda$. Similarly for the other direction except we have $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \in \mathrm{SL}_2(\mathbb{Z})$. □

Lemma 1.24. *Let $\varphi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ be a holomorphic map. Then there exist $m, b \in \mathbb{C}$ with $m\Lambda \subset \Lambda'$ such that $\varphi(z + \Lambda) = mz + b + \Lambda'$. φ is invertible if and only if $m\Lambda = \Lambda'$.*

Corollary 1.25. *Suppose $\varphi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ is holomorphic with $\varphi(z + \Lambda) = mz + b + \Lambda'$. Then the following are equivalent:*

i *φ is a group homomorphism,*

ii *$b \in \Lambda'$,*

iii *$\varphi(0) = 0$.*

In particular, $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ are isomorphic if and only if there exists $m \in \mathbb{C}$ such that $m\Lambda = \Lambda'$.

Example. For any $\tau \in \mathbb{H}$, we set

$$\Lambda_\tau = \mathbb{C}/(\mathbb{Z} \oplus \tau\mathbb{Z}).$$

Every complex torus is isomorphic to some Λ_τ. To see this, let $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ be a lattice. Up to swapping ω_1 and ω_2, we can assume that $\tau = \omega_2/\omega_1 \in \mathbb{H}$, so

$$\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z} = \omega_1(\mathbb{Z} \oplus \tau\mathbb{Z})$$

and hence $\mathbb{C}/\Lambda \cong \Lambda_\tau$.

It follows from Lemma 1.23 that Λ_τ and $\Lambda_{\tau'}$ are isomorphic if and only if there exists $m \in \mathbb{C}$ satisfying

$$m\begin{pmatrix} \tau' \\ 1 \end{pmatrix} = \gamma\begin{pmatrix} \tau \\ 1 \end{pmatrix}$$

for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Therefore Λ_τ and $\Lambda_{\tau'}$ are equivalent if and only if τ and τ' are equivalent under the action $(\gamma, \tau) \mapsto \gamma(\tau)$ of the modular group on the upper half plane.

Definition 1.26. An **isogeny** is a morphism of algebraic groups that is surjective and has finite kernel.

In particular, a nonzero holomorphic group homomorphism $f : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ is an isogeny, since any non-constant holomorphic map between compact Riemann surfaces is surjective, and the preimage of any point is discrete; hence the kernel is discrete, and therefore finite.

Example. Let N be a positive integer. Then the multiply-by-integer map

$$[N] : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda, \quad z + \Lambda \mapsto Nz + \Lambda$$

is an isogeny. Letting E denote the torus, the kernel of this map is denoted $E[N]$ and consists of the N-torsion points of E.

Example. Let $E = \mathbb{C}/\Lambda$ be a complex torus and let N be a positive integer. Let H be a cyclic subgroup of $E[N]$ of order N. The cyclic quotient map

$$\pi : E \to E/H, \quad z + \Lambda \to z + H$$

is an isogeny.

Lemma 1.27. *Every isogeny can be expressed as the composition of the previous two examples. Moreover, isogeny is an equivalence relation.*

An elliptic curve over a field k is a smooth, projective curve of genus 1 with a distinguished k-rational point (often a point at ∞). By a projective curve, we mean that it is the set of solutions to a non-constant homogeneous polynomial equation $F(X, Y, Z) = 0$ in the projective plane $\mathbb{P}^2(k)$ over k. In affine coordinates, every elliptic curve is given a cubic in two variables such that the discriminant (to be defined below) is nonzero. The constraint on the discriminant ensures that the curve is nonsingular. One takes the elliptic curve to be the set of all points (x, y) in the algebraic closure of k which satisfy the above polynomial.

Here, we will first focus on the case $k = \mathbb{C}$ before generalising to an arbitrary field k. Elliptic curves over $\mathbb{C}$ can be viewed as the embedding of a complex torus in the complex projective plane $\mathbb{C}P^2$. To see this explicitly, we look at the meromorphic functions on the complex torus. In particular, given a lattice Λ, the Weierstrass $\wp$-function is a meromorphic function on $\mathbb{C} \setminus \Lambda$ defined by

$$\wp_\Lambda(z) = \wp(z, \Lambda) = \frac{1}{z^2} + \sum_{w \in \Lambda - \{0\}} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right), \quad z \in \mathbb{C} - \{\Lambda\}$$

and its derivative is

$$\wp'_\Lambda(z) = \wp'(z, \Lambda) = -2 \sum_{w \in \Lambda} \frac{1}{(z-w)^3}.$$

It is clear from the formulae that $\wp$ is even and $\wp'$ is odd. Both $\wp$ and $\wp'$ are periodic in Λ, and therefore descend to meromorphic functions on $\mathbb{C}/\Lambda$. In fact, it follows from the Riemann-Roch theorem that the field of meromorphic functions on $\mathbb{C}/\Lambda$ is precisely $\mathbb{C}(\wp, \wp')$, the rational expressions in these two functions, so these are the only two examples we need.

Example. For lattices Λ, the Eisenstein series can be generalised such that we sum over the lattice points,

$$G_k(\Lambda) = \sum_{w \in \Lambda \setminus \{(0,0\}} \frac{1}{w^k}, \quad k > 2 \text{ even}.$$

Definition 1.28. Let k be a field of characteristic 0. A **Weierstrass equation over** k is a cubic equation of the form

$$E : y^2 = 4x^3 - g_2 x - g_3, \quad g_2, g_3 \in k.$$

The **discriminant** of the equation is

$$\Delta = g_2^3 - 27 g_3^2 \in k$$

and if $\Delta \neq 0$ define the **invariant** of the equation to be

$$j = \frac{1728 g_2^3}{\Delta} \in k.$$

Note that we work with this form of the Weierstrass equation only when $\mathrm{char}(k) = 0$. The form of the Weierstrass equation in arbitrary characteristic will be introduced in the following chapter. The reason for the slight discrepancy is for the sake of simplicity; since this form of the Weierstrass equation is closely related to the Weierstrass $\wp$-function, it will simplify our discussion of function fields of moduli spaces in Section 5.1.

Definition 1.29. Let $\overline{k}$ be an algebraic closure of k. If a Weierstrass equation E has nonzero discriminant, it is called **nonsingular** and the set

$$\mathcal{E} = \{(x, y) \in \overline{k}^2 \text{ satisfying } E(x, y)\} \cup \{\infty\}$$

is called an **elliptic curve over** k.

Lemma 1.30. *The Laurent expansion of* $\wp$ *is*

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{n=2 \\ n \text{ even}}} (n+1)G_{n+2}(\Lambda)z^n, \quad 0 < |z| < \inf\{|\omega| : \omega \in \Lambda - \{0\}\}$$

Lemma 1.31. *Let* Λ *be a lattice in* $\mathbb{C}$. *Then the functions* $\wp_\Lambda$ *and* $\wp'_\Lambda$ *satisfy*

$$(\wp'_\Lambda(z))^2 = 4(\wp_\Lambda(z))^3 - g_2(\Lambda)\wp_\Lambda(z) - g_3(\Lambda)$$

where $g_2(\Lambda) = 60G_4(\Lambda)$ *and* $g_3(\Lambda) = 140G_6(\Lambda)$.

Lemma 1.32. *Let* $\Lambda = w_1\mathbb{Z} \oplus w_2\mathbb{Z}$ *and let* $w_3 = w_1 + w_2$. *Then the cubic equation* $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$ *satisfied by* $(x, y) = (\wp_\Lambda, \wp'_\Lambda)$ *is*

$$y^2 = 4(x - e_1)(x - e_2)(x - e_3), \quad e_i = \wp(w_i/2).$$

This equation is nonsingular.

Hence the map $z \mapsto (\wp_\Lambda(z), \wp'_\Lambda(z))$ takes points on $\mathbb{C} - \Lambda$ to points $(x, y) \in \mathbb{C}^2$ such that $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$. This map is bijective and extends to all $z \in \mathbb{C}$ by sending the lattice points to a suitably defined point at infinity. Thus, every lattice is associated to an elliptic curve via the corresponding Weierstrass-$\wp$ function and its derivative $\wp'$,

$$(\wp, \wp') : \mathbb{C}/\Lambda \to \{(x, y) \ : \ y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)\} \cup \{\infty\}.$$

If z_1 and z_2 are two nonzero points in $\mathbb{C}/\Lambda$, then their image points $P = (\wp(z_1), \wp'(z_1))$ and $Q = (\wp(z_2), \wp'(z_2))$ determine a secant or tangent line $ax + by + c = 0$ of the elliptic curve in $\mathbb{C}^2$. Thinking of the curve as lying in the complex projective plane, it turns out that there is exactly one more point R lying on this line. Thus the elliptic curve inherits a group structure from $\mathbb{C}/\Lambda$ with the addition law given by

$$P + Q + R = 0 \iff P, Q, R \text{ are collinear.}$$

1.5 Modular curves and moduli spaces

A moduli space can be seen as a geometric space that gives a solution to a geometric classification problem. The points in a moduli space correspond to geometric objects that are in some sense "the same" (that is, up to a suitable notion of equivalence).

In the section 1.4, we saw that two complex tori are holomorphically group-isomorphic if and only if there exists $m \in \mathbb{C}$ such that $m\Lambda = \Lambda'$. Thus, the left quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ is in bijection with the set of isomorphism classes of complex tori. This quotient can be equipped with a natural complex structure which we can show to be isomorphic to $\mathbb{C}$. In this way, we can endow the set of equivalence classes of elliptic curves with a geometry.

As we will see in this section, given a congruence subgroup Γ, there is a similar bijection between the equivalence classes of points in $\mathbb{H}$ under the action of Γ and equivalence classes of elliptic curves enhanced by corresponding torsion data.

Definition 1.33. An **enhanced elliptic curve for** $\Gamma_0(N)$ is an ordered pair (E, C) where E is a complex elliptic curve and C is a cyclic subgroup of E of order N.

We say that two such pairs (E, C) and (E', C') are equivalent if there exists an isomorphism $\varphi : E \to E'$ such that $\varphi(C) = C'$. The set of equivalence classes is denoted $S_0(N)$ and an element of $S_0(N)$ is denoted $[E, C]$.

We can make analogous definitions for $\Gamma_1(N)$ and $\Gamma(N)$ with slight modifications to the torsion data we must specify.

Definition 1.34. An **enhanced elliptic curve for** $\Gamma_1(N)$ is an ordered pair (E, Q) where E is a complex elliptic curve and Q is point of E of order N.

We say that two such pairs (E, Q) and (E', Q') are equivalent, if there exists an isomorphism $E \to E'$ such that Q to Q'. The set of equivalence classes is denoted $S_1(N)$.

Definition 1.35. An **enhanced elliptic curve for** $\Gamma(N)$ is an ordered pair $(E, (P, Q))$ where E is a complex elliptic curve and (P, Q) is a pair of points of E that generate $E[N]$ with Weil pairing $e_N(P, Q) = e^{2\pi i/N}$.

Two such pairs $(E, (P, Q))$ and $(E', (P', Q'))$ are equivalent, if there exists an isomorphism $E \to E'$ that sends P to P' and Q to Q'. The set of equivalence classes is denoted $S(N)$.

The sets of equivalence classes $S_0(N)$, $S_1(N)$, and $S(N)$ are examples of moduli spaces of isomorphism classes of complex elliptic curves and N-torsion data.

Definition 1.36. Given a congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$, acting on the left on $\mathbb{H}$, the **modular curve** $Y(\Gamma)$ is defined as

$$Y(\Gamma) = \Gamma\backslash\mathbb{H} = \{\Gamma\tau \ : \ \tau \in \mathbb{H}\}.$$

$Y(\Gamma)$ is a Riemann surface that can be compactified by adding finitely many points, namely the cusps of Γ. The corresponding compactified modular curve is denoted $X(\Gamma)$. We will not prove this statement here and instead refer the reader to Chapter 2 of [7].

For any positive integer N, the modular curves

$$X_0(N) = \Gamma_0(N)\backslash\mathbb{H}^*, \quad X_1(N) = \Gamma_1(N)\backslash\mathbb{H}^*, \quad X(N) = \Gamma(N)\backslash\mathbb{H}^*$$

can be described as algebraic curves – that is, they are the solution sets of systems of polynomial equations. The existence of such polynomials comes from a general theorem of Riemann surface theory that tells us that every compact Riemann surface is an algebraic curve via an embedding in some complex projective space $\mathbb{C}P^n$.

What is less obvious is that $X_0(N)$ and $X_1(N)$ are curves over $\mathbb{Q}$, that is, the polynomials can be taken to have rational coefficients. Further, $X(N)$ can be defined by polynomials over $\mathbb{Q}(\mu_N)$. We will see in Chapter 2 how this arises from studying elliptic curves.

Theorem 1.37. *Let N be a positive integer.*

The moduli space $S_0(N)$ of $\Gamma_0(N)$. *Define an equivalence relation by letting $[E_\tau, \langle 1/N + \Lambda_\tau\rangle]$ and $[E_{\tau'}, \langle 1/N + \Lambda_{\tau'}\rangle]$ be equal if and only if $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$. The moduli space for $\Gamma_0(N)$ is*

$$S_0(N) = \{[E_\tau, \langle 1/N + \Lambda_\tau\rangle] : \tau \in \mathbb{H}\}.$$

There is a bijection

$$\begin{aligned} \psi_0 : S_0(N) &\to Y_0(N) \\ [\mathbb{C}/\Lambda_\tau, \langle 1/N + \Lambda_\tau\rangle] &\mapsto \Gamma_0(N)\tau. \end{aligned}$$

The moduli space $S_1(N)$ of $\Gamma_1(N)$. *Define an equivalence relation by letting $[E_\tau, 1/N + \Lambda_\tau]$ and $[E_{\tau'}, 1/N + \Lambda_{\tau'}]$ be equal if and only if $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$. The moduli space for $\Gamma_1(N)$ is*

$$S_1(N) = \{[E_\tau, 1/N + \Lambda_\tau] : \tau \in \mathbb{H}\}.$$

There is a bijection

$$\begin{aligned} \psi_1 : S_1(N) &\to Y_1(N) \\ [\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau] &\mapsto \Gamma_1(N)\tau. \end{aligned}$$

The moduli space $S(N)$ of $\Gamma(N)$. *Define an equivalence relation by letting $[\mathbb{C}/\Lambda_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)]$ and $[\mathbb{C}/\Lambda_{\tau'}, (\tau'/N + \Lambda_{\tau'}, 1/N + \Lambda_{\tau'})]$ be equal if and only if $\Gamma(N)\tau = \Gamma(N)\tau'$. The moduli space for $\Gamma(N)$ is*

$$S(N) = \{[\mathbb{C}/\Lambda_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)] : \tau \in \mathbb{H}\}.$$

There is a bijection

$$\begin{aligned} \psi : S(N) &\to Y(N) \\ [\mathbb{C}/\Lambda_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)] &\mapsto \Gamma(N)\tau. \end{aligned}$$

Proof. See Theorem 1.5.1. in [7]. □

Thus, maps of the modular curves $Y_0(N)$, $Y_1(N)$, and $Y(N)$ induce maps of moduli spaces $S_0(N)$, $S_1(N)$, and $S(N)$.

Example. Consider the map $Y_1(N) \to Y_0(N)$ given by $\Gamma_1(N)\tau \mapsto \Gamma_0(N)\tau$. This corresponds to a map $S_1(N) \to S_0(N)$ sending $[E, Q]$ to $[E, \langle Q\rangle]$. That is, we keep the subgroup generated by Q but we forget the generator.

Example. Consider the action of the quotient group $\Gamma_0(N)/\Gamma_1(N)$ on $Y_1(N)$. This induces an action on $S_1(N)$ given by

$$\Gamma_1(N)\gamma : [E, Q] \to [E, dQ],$$

where $\gamma \equiv \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \pmod N$.

Definition 1.38. Let $k \in \mathbb{Z}$ and let Γ be one of $\Gamma_0(N), \Gamma_1(N)$, or $\Gamma(N)$. A complex-valued function F of the enhanced elliptic curves for Γ is **degree-k homogeneous with respect to** Γ if given any $m \in \mathbb{C}^\times$, we have

$$\left.\begin{matrix} F(\mathbb{C}/m\Lambda, mC) \\ F(\mathbb{C}/m\Lambda, mQ) \\ F(\mathbb{C}/m\Lambda, (mP, mQ)) \end{matrix}\right\} = \begin{cases} m^{-k}F(\mathbb{C}/\Lambda, C), & \text{if } \Gamma = \Gamma_0(N) \\ m^{-k}F(\mathbb{C}/\Lambda, Q), & \text{if } \Gamma = \Gamma_1(N) \\ m^{-k}F(\mathbb{C}/\Lambda, (P, Q), & \text{if } \Gamma = \Gamma(N). \end{cases} \tag{1.1}$$

Given F satisfying the preceding definition, define the corresponding **dehomogenized** function $f : \mathbb{H} \to \mathbb{C}$ defined by

$$f(\tau) = \begin{cases} F(\mathbb{C}/\Lambda_\tau, \langle 1/N + \Lambda_\tau\rangle), & \text{if } \Gamma = \Gamma_0(N) \\ F(\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau), & \text{if } \Gamma = \Gamma_1(N) \\ F(\mathbb{C}/\Lambda_\tau, (\tau/N + \Lambda_\tau, 1/N + \Lambda_\tau)), & \text{if } \Gamma = \Gamma_0(N). \end{cases} \tag{1.2}$$

Then for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, we have $f(\gamma(\tau)) = (c\tau + d)^k f(\tau)$. In the other direction, if $f : \mathbb{H} \to \mathbb{C}$ is weight-k invariant, then equation (1.2) defines a function F on enhanced elliptic curves of the form $(\mathbb{C}/\Lambda_\tau$, torsion data) satisfying (1.1) and extending to a degree-k homogeneous function of enhanced elliptic curves for Γ.

Example. Let $N > 1$ be an integer and let $v = (c_v, d_v) \in \mathbb{Z}^2$ be such that the reduction $\overline{v}$ of v modulo N is nonzero. Define functions of enhanced elliptic curves for $\Gamma(N)$, $\Gamma_1(N)$, and $\Gamma_0(N)$ respectively by

$$\begin{aligned} F_0^{\overline{v}}(\mathbb{C}/\Lambda, (P, Q)) &= \frac{g_2(\Lambda)}{g_3(\Lambda)}\wp_\Lambda(c_vP + d_vQ) \\ F_0^{\overline{d}}(\mathbb{C}/\Lambda, Q) &= \frac{g_2(\Lambda)}{g_3(\Lambda)}\wp_\Lambda(dQ), \quad d \in \mathbb{Z},\ d \not\equiv 0 \pmod N \\ F_0(\mathbb{C}/\Lambda, C) &= \frac{g_2(\Lambda)}{g_3(\Lambda)}\sum_{Q \in C - \{0\}} \wp_\Lambda(Q). \end{aligned}$$

We will use without proof that these functions are degree-0 homogeneous with respect to its congruence subgroup. We obtain the corresponding weight-0 invariant functions

$$\begin{aligned} f_0^{\overline{v}}(\tau) &= \frac{g_2(\tau)}{g_3(\tau)}\wp_\tau\left(\frac{c_v\tau + d_v}{N}\right) \\ f_0^{\overline{d}}(\tau) &= \frac{g_2(\tau)}{g_3(\tau)}\wp_\tau\left(\frac{d}{N}\right) = f_0^{\overline{(0,d)}}(\tau) \end{aligned}$$

$$f_0(\tau) = \frac{g_2(\tau)}{g_3(\tau)} \sum_{d=1}^{N-1} \wp_\tau\left(\frac{d}{N}\right) = \sum_{d=1}^{N-1} f_0^{\overline{d}}(\tau).$$

We will revisit these functions in Section 5.1.

2 Galois representations

One of the ways of studying the finite extensions of $\mathbb{Q}$ is by looking at the representations of the absolute Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. In this section, we will begin with a review of some classical Galois theory, and in section 2.3 we will see how this can be extended to infinite field extensions.

Following this, we will relate Galois representations to the central objects of this paper, i.e. modular forms. For a more detailed discussion of Galois theory related to modular forms, one should see [19].

2.1 Galois extensions

Let F be a number field and fix an algebraic closure $\overline{F}$. A Galois representation is a continuous representation of the group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on a finite-dimensional vector space V. Typically, V will be a vector space over $\mathbb{Q}_p$, the field of p-adic numbers for a prime number p.

An important property of a Galois representation is whether or not it is *unramified* at a prime p. In the case that it is, we can formulate a well-defined action of the Frobenius endomorphism $x \to x^p$ in $\mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p)$ on V up to conjugation.

In this section we will only cover briefly some basic definitions from algebraic number theory, commutative algebra, and representation theory. One may wish to see [21], [12] or [2] for details.

Throughout, F is an algebraic number field. Recall that for a number field F, the index $[F : \mathbb{Q}]$ is finite. Associated with the number field is its ring of algebraic integers, denoted $\mathcal{O}_F$.

Definition 2.1. Let E/F be an algebraic field extension.

E is said to be **normal over** F if every irreducible polynomial $f \in F[x]$ with at least one root in E splits completely into linear factors over E.

E is said to be **separable over** F if for every $\alpha \in E$, the minimal polynomial of α over F is a separable polynomial. That is, its roots are distinct in $\overline{F}$, an algebraic closure of F.

A field extension E/F is **Galois** if E is both normal and separable over F. If E/F is Galois, then $\mathrm{Aut}(E/F)$ is called the **Galois group** of E/F, also denoted $\mathrm{Gal}(E/F)$.

In particular, the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ is Galois over $\mathbb{Q}$, and similarly, $\overline{\mathbb{F}}_p$ is Galois over $\mathbb{F}_p$.

Definition 2.2. The group $G_{\mathbb{Q}} := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is called the **absolute Galois group of** $\mathbb{Q}$.

Let $\mathcal{K}$ be the set of all Galois number fields partially ordered by $E \leq E' \iff E \subseteq E'$. We have transition morphisms given by restriction

$$\begin{gathered}\mathrm{Gal}(E/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(E'/\mathbb{Q}) \\ \sigma \mapsto \sigma|_{E'}\end{gathered}$$

for $E \subseteq E'$. Then we can view the absolute Galois group as the (inverse) limit

$$G_{\mathbb{Q}} = \varprojlim_{E \in \mathcal{K}} \mathrm{Gal}(E/\mathbb{Q}).$$

Thus, $G_{\mathbb{Q}}$ is an example of a profinite group, i.e. a group that is the inverse limit of a system of discrete finite groups. As such, $G_{\mathbb{Q}}$ is Hausdorff, compact, and totally disconnected.

Each finite Galois group $\mathrm{Gal}(E/F)$ can be viewed as a topological space with the discrete topology. By endowing the inverse limit with the subspace topology with respect to the inclusion $G_{\mathbb{Q}} \subset \prod_E \mathrm{Gal}(E/\mathbb{Q})$, we can define a topology on $G_{\mathbb{Q}}$. This topology is not the discrete topology, and we call it the *profinite topology* or the *Krull topology.*

A basis of the Krull topology on $G_{\mathbb{Q}}$ is

$$\{U_\sigma(E) \,:\, \sigma \in G_{\mathbb{Q}},\ E \text{ is a Galois number field}\}$$

where $U_\sigma(E) = \sigma \cdot \ker(G_{\mathbb{Q}} \to \mathrm{Gal}(E/\mathbb{Q}))$.

Remark. Galois extensions $\mathrm{Gal}(E/F)$ are examples of topological groups. That is, they are topological spaces with multiplication and inversion maps that are continuous and satisfy the group axioms.

Theorem 2.3. *(The Fundamental Theorem of Galois Theory). Let L/K be a finite Galois extension with Galois group $G = \mathrm{Gal}(L/K)$. Then there is an inclusion-reversing bijection*

$$\begin{aligned} \{\textit{closed subgroups } H \leq G\} &\longleftrightarrow \{\textit{subextensions } L/M/K\} \\ H &\longmapsto L^H \\ \mathrm{Aut}(L/M) &\longleftarrow\!\shortmid M \end{aligned}$$

which restricts to bijections

$$\begin{aligned} \{\textit{normal closed subgroups } H \leq G\} &\longleftrightarrow \{\textit{Galois subextensions } L/M/K\} \\ \{\textit{open subgroups } H \leq G\} &\longleftrightarrow \{\textit{finite subextensions } L/M/K\}. \end{aligned}$$

2.2 p-adic numbers

For the remainder of this chapter, let p denote a prime number.

For any prime p, we can introduce a metric on the field of rational numbers $\mathbb{Q}$ known as the p-adic metric. The p-adic numbers are the completion of $\mathbb{Q}$ with respect to this metric, forming a field $\mathbb{Q}_p$.

Given any two positive integers $n \geq m$ there is an inclusion

$$p^n\mathbb{Z} \subseteq p^m\mathbb{Z}$$

which induces a ring homomorphism

$$\mathbb{Z}/p^n\mathbb{Z} \to \mathbb{Z}/p^m\mathbb{Z}$$

given by reduction mod p. The homomorphism for all pairs of integers $n \geq m$ forms a directed system of rings.

Definition 2.4. The **ring of p-adic integers** $\mathbb{Z}_p$ is the (inverse) limit

$$\mathbb{Z}_p = \varprojlim_n (\mathbb{Z}/p^n\mathbb{Z})$$

under the maps $\mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ given by reduction mod p^n.

An p-adic integer z can we written as a formal base-p power series

$$z = \sum_{n=0}^{\infty} a_n p^n, \quad a_n \in \mathbb{Z}/p\mathbb{Z}$$

as shorthand for the sequence of partial sums

$$z = (a_0, a_0 + a_1 p, a_0 + a_1 p + a_2 p^2, a_0 + a_1 p + a_2 p^2 + a_3 p^3, \ldots). \tag{2.1}$$

Definition 2.5. The **p-adic valuation** is the discrete valuation $v_p : \mathbb{Z}_p \to \mathbb{Z} \cup \{\infty\}$ given by setting $v_p((a_i)_{i \geq 0})$ equal to the index i of the first nonzero term a_i and $v_p(0) = \infty$.

The p-adic valuation induces a norm given by

$$|z|_p = p^{-v_p(z)}$$

and thus a metric

$$d(z, w) = |z - w|_p.$$

Under this metric, the sequence (2.1) converges to

$$z = \sum_{n=0}^{\infty} a_n p^n.$$

Remark. There is a natural ring injection $\mathbb{Z} \hookrightarrow \mathbb{Z}_p$ given by reducing $a \in \mathbb{Z}$ modulo each p^n. We view $\mathbb{Z}$ as a subring of $\mathbb{Z}_p$ in this way. This map also gives a natural isomorphism

$$\begin{aligned} \mathbb{Z}/p\mathbb{Z} &\xrightarrow{\sim} \mathbb{Z}_p/p\mathbb{Z}_p \\ a + p\mathbb{Z} &\mapsto a + p\mathbb{Z}_p. \end{aligned}$$

Equipped with the metric introduced above, $\mathbb{Z}_p$ is a completion of $\mathbb{Z}$ – that is, d is a complete metric and $\mathbb{Z}$ is dense in $\mathbb{Z}_p$.

Lemma 2.6. *The set of units of $\mathbb{Z}_p$ is given by*

$$\mathbb{Z}_p^\times = \{z \ : \ |z|_p = 1\}.$$

Proof. The condition that $|z|_p = 1$ is equivalent to the condition that $v_p(z) = 0$, or equivalently $a_0 \in \mathbb{Z}/p\mathbb{Z}$. If $a_0 = 0$, then for any $(b_i) \in \mathbb{Z}_p$ we have $a_0 b_0 = 0$, and so z is not invertible. On the other hand, if $a_0 \in \mathbb{Z}/p\mathbb{Z}$, then for all i we have that a_i is relatively prime to p since a_i projects to $a_0 \neq 0 \in \mathbb{Z}/p\mathbb{Z}$. Hence a_i is relatively prime to p^i and there exists $b_i = a_i^{-1} \in \mathbb{Z}/p^i\mathbb{Z}$. Since this holds for all i, we have $(a_i)(b_i) = (a_i b_i) = (1)$, so $z \in \mathbb{Z}_p^\times$. □

Definition 2.7. The p**-adic field** $\mathbb{Q}_p$ is the field of fractions of $\mathbb{Z}_p$.

We extend the p-adic valuation to a discrete valuation, also called the p-adic valuation, on $\mathbb{Q}_p$ by

$$\begin{aligned} v_p : \mathbb{Q}_p &\to \mathbb{Z} \\ \frac{a}{b} &\mapsto v_p(a) - v_p(b). \end{aligned}$$

We define the metric on $\mathbb{Q}_p$ by the same formula as for $\mathbb{Z}_p$. Similarly, an element $z \in \mathbb{Q}_p$ can also be written as a sum

$$\sum_{n \geq k} a_n p^n, \quad a_n \in \mathbb{Z}/p\mathbb{Z}$$

except now we must allow for finitely many negative exponents. This is justified by the following lemma.

Lemma 2.8. *Every element $z \in \mathbb{Q}_p$ can be written uniquely as $z = up^n$ with $u \in \mathbb{Z}_p^\times$ and $n = v_p(z)$.*

Proof. For existence, we can write z in the form $\sum_{n=k}^{\infty} a_n p^n$ with $a_n \in \mathbb{Z}/p\mathbb{Z}$ where $k = v_p(z)$ and $a_k \neq 0$. Then $z = (\sum n = 0^\infty a_{n+k} p^n) p^k$ is of the desired form. For uniqueness, suppose $up^n = u'p^n$ where $u, u' \in \mathbb{Z}_p^\times$. Then $u^{-1}u'p^n = p^n$, so $u^{-1}u' = 1$. Therefore $u = u'$. □

Remark. Some authors use this to define the p-adic valuation.

Definition 2.9. The p**-adic topology** is the topology on $\mathbb{Q}_p$ induced by the p-adic valuation.

Lemma 2.10. *$\mathbb{Z}_p$ is compact in the p-adic topology.*

Proof. Let

$$P = \prod_{n=1}^{\infty} \mathbb{Z}/p^n\mathbb{Z}.$$

Since each $\mathbb{Z}/p^n\mathbb{Z}$ is compact, P is compact by the Tychonoff Theorem. An element of P is of the form $(a_n + p^n\mathbb{Z})_{n=1}^{\infty}$. For each positive integer n let $\pi_{n+1,n} : \mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z}$ be the projection $(a + p^{n+1}\mathbb{Z} \to a + p^n\mathbb{Z}$. The graph of $\pi_{n+1,n}$ is

$$G_n := \{a_{n+1} + p^{n+1}\mathbb{Z}, a_n + p^n\mathbb{Z} : a_n \equiv a_{n+1} \pmod{p^n}\} \subset \mathbb{Z}/p^{n+1}\mathbb{Z} \times \mathbb{Z}/p^n\mathbb{Z}.$$

$\mathbb{Z}/p^{n+1}\mathbb{Z} \times \mathbb{Z}/p^n\mathbb{Z}$ is finite and hence has the discrete topology, so G_n is a closed subset of $\mathbb{Z}/p^{n+1}\mathbb{Z} \times \mathbb{Z}/p^n\mathbb{Z}$. The subset of P

$$G'_n = G_n \times \prod_{m \neq n, n+1} \mathbb{Z}/p^m\mathbb{Z} \subset P$$

is also closed since the complement $G_n^c \times \prod_{m \neq n,n+1} \mathbb{Z}/p^m\mathbb{Z} \subset P$ is open, where G_n^c is the complement of G_n in $\mathbb{Z}/p^{n+1}\mathbb{Z} \times \mathbb{Z}/p^n\mathbb{Z}$.

$\mathbb{Z}_p$ is the subspace of P consisting of elements $(a_n + p^n\mathbb{Z})_{n=1}^{\infty}$ such that $a_n \equiv a_{n+1} \pmod{p}$ for $1 \leq n \leq \infty$. That is,

$$\mathbb{Z}_p = \bigcap_{n=1}^{\infty} G'_n,$$

which is closed in the compact space P. Thus, $\mathbb{Z}_p$ is compact.

□

2.3 Representations of the absolute Galois group

In the study of representations of $G_{\mathbb{Q}}$, we are only interested in the maps $G_{\mathbb{Q}} \to \mathrm{GL}_n(F)$ that are continuous. The topology on $G_{\mathbb{Q}}$ was introduced in the previous section, but we are yet to specify a topology on $\mathrm{GL}_n(F)$. For this, we must assume that F is a *topological field*.

Definition 2.11. A **topological ring** is a triple $(R, +, \cdot)$ where R is a topological space equipped with the structure of a ring on its underlying set, such that addition $+$ and multiplication $\cdot$ are continuous functions.

A **topological field** is a topological ring whose underlying ring R is a field, and such that multiplicative inversion

$$\begin{gathered} i : R - \{0\} \to R - \{0\} \\ u \to u^{-1} \end{gathered}$$

is continuous with respect to the subspace topology inherited from R.

Given that F is a topological field, with F^{n^2} equipped with the product topology, $\mathrm{GL}_n(F)$ inherits its topology as a subset of F^{n^2}.

Definition 2.12. Let F be a topological field. A **Galois representation of dimension n over F** is a continuous group homomorphism

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(F).$$

ρ and ρ' are **equivalent** if $\rho' : G_{\mathbb{Q}} \to \mathrm{GL}_n(F)$ is another such representation and there exists $m \in \mathrm{GL}_n(F)$ such that $\rho'(\sigma) = m^{-1}\rho(\sigma)m$ for all $\sigma \in G_{\mathbb{Q}}$. Equivalence is denoted $\rho \sim \rho'$.

If F is a field extension over $\mathbb{Q}_p$, we call ρ an **p-adic Galois representation**.

Example. (The p-adic cyclotomic character). Let

$$\mu_n = \left\{ \zeta \in \overline{\mathbb{Q}}^{\times} : \zeta^n = 1 \right\}$$

be the set of nth roots of unity in $\overline{\mathbb{Q}}^{\times}$. Then there is an ismomorphism

$$\varphi_n : \mathrm{Gal}(\mathbb{Q}(\mu_n)/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/n\mathbb{Z})^{\times} \tag{2.2}$$

$$\sigma \mapsto g \quad \text{such that } \sigma(\zeta) = \zeta^g \text{ for all } \zeta \in \mu_n. \tag{2.3}$$

Consider the set of p^nth roots of unity for positive integers n, writing $\mu_{p^\infty} = \cup_{n \geq 1} \mu_{p^n}$. Then, taking the inverse limit with respect to the natural restriction maps, we have

$$\mathrm{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q}) \cong \varprojlim_{n \geq 1} \mathrm{Gal}(\mathbb{Q}(\mu_{p^n})/\mathbb{Q}).$$

For any pair of positive integers $n \leq m$, we have the following commutative diagram with the restriction map $\mathrm{Gal}(\mathbb{Q}(\mu_{p^m})/\mathbb{Q}) \to \mathrm{Gal}(\mathbb{Q}(\mu_{p^n})/\mathbb{Q})$ given by reduction modulo p^n from $(\mathbb{Z}/p^m\mathbb{Z})^{\times}$ to $(\mathbb{Z}/p^n\mathbb{Z})^{\times}$.

$$\begin{tikzcd}
\mathrm{Gal}(\mathbb{Q}(\mu_{p^m}/\mathbb{Q})) \arrow[r] \arrow[d, "\varphi_m", "\wr"'] & \mathrm{Gal}(\mathbb{Q}(\mu_{p^n}/\mathbb{Q})) \arrow[d, "\varphi_n", "\wr"'] \\
(\mathbb{Z}/p^m\mathbb{Z})^{\times} \arrow[r, "\text{mod } p^n"] & (\mathbb{Z}/p^n\mathbb{Z})^{\times}
\end{tikzcd}$$

Therefore, for each n we can construct a representation:

$$\chi_{p,n} : G_{\mathbb{Q}} \twoheadrightarrow \mathrm{Gal}(\mathbb{Q}(\mu_{p^n})/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/p^n\mathbb{Z})^{\times}$$

where the first map is restriction. The *p-adic cyclotomic character* is given by

$$\chi_p : \ G_{\mathbb{Q}} \longrightarrow\!\!\!\!\!\rightarrow \mathrm{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q}) \xrightarrow{\ \sim\ } \varprojlim(\mathbb{Z}/p^n\mathbb{Z})^{\times} \xrightarrow{\ \sim\ } \mathbb{Z}_p^{\times}.$$

To view this as an p-adic Galois representation, we simply consider the inclusion $\mathbb{Z}_p^{\times} \hookrightarrow \mathbb{Q}_p^{\times} = \mathrm{GL}_1(\mathbb{Q}_\mathrm{p})$.

Proposition 2.13. *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(E)$ be a Galois representation. Then ρ is isomorphic to a Galois representation $\rho' : G_{\mathbb{Q}} \to \mathrm{GL}_n(\mathcal{O}_E)$.*

To prove this proposition, we will need the following definition.

Definition 2.14. Let d be a positive integer. A **d-dimensional p-adic Galois representation** is a topological vector space V of dimension d over L where L is a finite field extension of $\mathbb{Q}_p$, and moreover V is a $G_{\mathbb{Q}}$-module such that the action

$$\begin{aligned} V \times G_{\mathbb{Q}} &\to V \\ (v, \sigma) &\mapsto \sigma(v) \end{aligned}$$

is continuous.

Any two such representation V and V' are **equivalent** if there exists a continuous $G_{\mathbb{Q}}$-module isomorphism of L-vector spaces $V \to V'$.

This definition is compatible with Definition 2.12. To see this, let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_d(L)$ be a Galois representation as in Definition 2.12. Any choice of ordered basis for L^d identifies $\mathrm{GL}_d(L)$ with $\mathrm{Aut}(L^d)$. Since ρ is continuous,

$$\begin{aligned} L^d \times G_{\mathbb{Q}} &\to L^d \\ (v, \sigma) &\mapsto \rho(\sigma)(v) \end{aligned}$$

is also continuous, making L^d a $G_{\mathbb{Q}}$-module satisfying Definition 2.14.

In the other direction, let V be a vector space of dimension d over L as in Definition 2.14. Choosing an ordered basis $\mathcal{B} = (b_1, \ldots, b_d)$ for V identifies $\mathrm{Aut}(V)$ with $\mathrm{GL}_d(L)$. Therefore the map

$$\begin{aligned} G_{\mathbb{Q}} &\to \mathrm{Aut}(V) \\ \sigma &\mapsto \sigma|_V \end{aligned}$$

gives a map $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_d(L)$ with matrix entries given by $\rho(\sigma)_{ij} = x_j(\sigma(b_i))$ where x_i is the projection

$$\begin{aligned} x_i : V &\to L \\ \sum_k a_k b_k &\mapsto a_i. \end{aligned}$$

Since each matrix entry ρ_{ij} is a continuous function, ρ is also continuous, making ρ a Galois representation in the sense of Definition 2.12.

Let us now return to the proof of Proposition 2.13.

Proof. We will use without proof that for any finite extension L over $\mathbb{Q}_p$ is of the form F_λ for some number field F and maximal ideal $\lambda|p$ of $\mathcal{O}_F$. The ring $\mathcal{O}_L = \mathcal{O}_{F,\lambda}$ is independent of choice of F and λ.

Let $V = E^n$ and let $\Lambda = \mathcal{O}_E^d$. Since Λ is a lattice of V it is finitely generated as a $\mathbb{Z}_p$-module. Since $\mathbb{Z}_p$ is compact, V is also compact. Since $G_\mathbb{Q}$ is also compact, the image Λ' of $\Lambda \times G_\mathbb{Q}$ under the action $V \times G_\mathbb{Q} \to V$ is compact. Therefore $\Lambda' \subset \lambda^{-r}\Lambda$ for some positive integer r. Λ' is finitely generated and contains Λ so its rank is greater than or equal to d. Since $\mathcal{O}_E$ is a principal ideal domain, its rank is exactly d. Moreover, $G_\mathbb{Q}$ takes elements of Λ' to points in Λ', so any basis of $\mathcal{O}_E$ gives ρ' as desired. □

2.4 Ramification

Let E/F be a Galois extension of number fields and let $\mathfrak{p}$ be a nonzero prime ideal in $\mathcal{O}_F$. Algebraic number theory tells us that $\mathcal{O}_E$ is a Dedekind domain. As such, lifting $\mathfrak{p}$ to $\mathcal{O}_E$, the ideal $\mathfrak{p}\mathcal{O}_E$ of $\mathcal{O}_E$ factors uniquely into a finite product of distinct prime ideals $\mathfrak{P}_i \supset p\mathcal{O}_E$ of $\mathcal{O}_E$ with positive multiplicities,

$$\mathfrak{p}\mathcal{O}_E = \prod_{i=1}^{g} \mathfrak{P}_i^{e_i}. \tag{2.4}$$

The **decomposition index** g is the number of distinct prime ideals over $\mathfrak{p}$. We call e_i the **ramification index** of $\mathfrak{P}_i$ over $\mathfrak{p}$. If $e_i > 1$ for any i, then we say that $\mathfrak{p}$ **ramifies in** E. Similarly, suppose $\mathfrak{P}$ is a nonzero prime ideal in $\mathcal{O}_E$ and $\mathfrak{p} := \mathfrak{P} \cap \mathcal{O}_F$. In this case, we say that $\mathfrak{P}$ lies above $\mathfrak{p}$. If the ramification index of $\mathfrak{P}$ in the factorisation of $\mathfrak{p}\mathcal{O}_E$ is greater than 1, then we say that $\mathfrak{P}$ **ramifies over** F.

The inclusion $\iota : \mathcal{O}_F \to \mathcal{O}_E$ induces an $\mathcal{O}_F/\mathfrak{p}$ structure on the ring $\mathbf{f}_{\mathfrak{P}_j} := \mathcal{O}_E/\mathfrak{P}_j$ giving a natural embedding of the residue field $\mathcal{O}_F/\mathfrak{p}$ into $\mathcal{O}_E/\mathfrak{P}_j$ for each j. The degree

$$f_j = [\mathcal{O}_E/\mathfrak{P}_j : \mathcal{O}_F/\mathfrak{p}]$$

is called the **inertial degree** of $\mathfrak{P}_j$ over $\mathcal{O}_F$.

The Galois group $\mathrm{Gal}(E/F)$ acts from the left on the set of prime ideals $S := \{\mathfrak{P}_i : i = 1, \dots, g\}$ via $\sigma \cdot \mathfrak{P}_i = \sigma(\mathfrak{P}_i)$. This action is transitive.

Proposition 2.15. *Let $\mathfrak{P}_i$ and $\mathfrak{P}_j$ be primes of $\mathcal{O}_E$ lying above $\mathfrak{p}$. Then there exists $\sigma \in \mathrm{Gal}(E/F)$ such that $\sigma(\mathfrak{P}) = \mathfrak{P}'$.*

Proof. Suppose that $\sigma(\mathfrak{P}_i) \neq \mathfrak{P}_j$ for all $\sigma \in \mathrm{Gal}(E/F)$. By the Chinese Remainder Theorem, there exists $x \in \mathfrak{P}_i$ such that $x \notin \mathfrak{P}_j$ for all $j \neq i$. Set $X = \prod_{\sigma\in\mathrm{Gal}(E/F)} \sigma(x)$. Then $X \in \mathfrak{P}_i \cap \mathcal{O}_F = \mathfrak{p}$. Since $\mathfrak{p} \subset \mathfrak{P}_j$, we have $X \in \mathfrak{P}_j$, but X is a product of the $\sigma(x)$, so this contradicts the primality of $\mathfrak{P}_j$. Thus, there exsits $\sigma \in \mathrm{Gal}(E/F)$ such that $\sigma(\mathfrak{P}) = \mathfrak{P}'$ □

In particular, an immediate consequence is that the ramification indices of all the primes dividing $\mathfrak{p}\mathcal{O}_E$ are equal, and it follows that the inertial degrees must also be equal. Therefore we can write

$$\mathfrak{p}\mathcal{O}_E = (\mathfrak{P}_1 \dots \mathfrak{P}_g)^e, \quad efg = [E : F] = |\mathrm{Gal}(E/F)|.$$

Definition 2.16. The **decomposition group** $D_{\mathfrak{P}}$ of $\mathfrak{P}$ is the subgroup of $\mathrm{Gal}(E/F)$ consisting of elements that fix $\mathfrak{P}$.

The order of the decomposition group is ef. It acts on the residue field $\mathcal{O}_E/\mathfrak{P}$ via

$$\sigma(x+\mathfrak{P}) = \sigma(x) + \mathfrak{P}, \quad x \in \mathcal{O}_E,\ \sigma \in D_{\mathfrak{P}}.$$

Definition 2.17. The **inertia group** $I_{\mathfrak{P}}$ of $\mathfrak{P}$ is the kernel of the above action,

$$I_{\mathfrak{P}} = \{\sigma \in D_{\mathfrak{P}} : x^\sigma \equiv x \pmod{\mathfrak{P}} \text{ for all } x \in \mathcal{O}_E\}.$$

The order of the inertia group is e, so it is trivial if $\mathfrak{P}$ is unramified over $\mathfrak{p}$.

In the case $F = \mathbb{Q}$ and $\mathfrak{p} = (p)$ where $p \in \mathbb{Z}$ is prime, $\mathbf{f}_{\mathfrak{P}} = \mathcal{O}_E/\mathfrak{P}$ is a degree-f vector space over $\mathcal{O}_F/\mathfrak{p} = \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ and so we have $\mathcal{O}_E/\mathfrak{P} \cong \mathbb{F}_{p^f}$. Viewing $\mathbb{F}_p$ as a subfield of $\mathbf{f}_{\mathfrak{P}}$, there is an injection

$$D_{\mathfrak{P}}/I_{\mathfrak{P}} \to \mathrm{Gal}(\mathbf{f}_{\mathfrak{P}}/\mathbb{F}_p).$$

Since both groups have order f, this is in fact an isomorphism, so we have a short exact sequence

$$\begin{tikzcd} 1 \arrow[r] & I_{\mathfrak{P}} \arrow[r] & D_{\mathfrak{P}} \arrow[r] & \mathrm{Gal}(\mathbf{f}_{\mathfrak{P}}/\mathbb{F}_p) \arrow[r] & 1. \end{tikzcd}$$

Since the Frobenius automorphism $\sigma_p : x \mapsto x^p$ generates $\mathrm{Gal}(\mathbf{f}_{\mathfrak{P}}/\mathbb{F}_p)$, we get a corresponding generator of $D_{\mathfrak{P}}/I_{\mathfrak{P}}$ that maps to σ_p.

Any representative of this generator in $D_{\mathfrak{P}}$ is called a *Frobenius element* of $\mathrm{Gal}(E/\mathbb{Q})$ and we denote it $\mathrm{Frob}_{\mathfrak{P}}$. The action of $\mathrm{Frob}_{\mathfrak{P}}$ on $\mathcal{O}_E$ descends to $\mathbf{f}_{\mathfrak{P}}$ where it is the action of the Frobenius automorphism σ_p.

Definition 2.18. Let $E/\mathbb{Q}$ be a Galois extension. Let $p \in \mathbb{Z}$ be prime and let $\mathfrak{P}$ be a prime ideal of $\mathcal{O}_E$ lying over p. A **Frobenius element** of $\mathrm{Gal}(E/\mathbb{Q})$ is an element $\mathrm{Frob}_{\mathfrak{P}}$ satisfying

$$\mathrm{Frob}_{\mathfrak{P}}(x) \equiv x^p \pmod{\mathfrak{P}}$$

for all $x \in \mathcal{O}_E$.

The extension $E/\mathbb{Q}$ is unramified at (p) if and only if the inertial group $I_{\mathfrak{P}}$ is trivial so that $\mathrm{Frob}_{\mathfrak{P}}$ is unique.

Proposition 2.19. *Let $\mathfrak{P}$ and $\mathfrak{P}'$ be prime ideals of $\mathcal{O}_E$ lying above (p). Then the Frobenius elements $\mathrm{Frob}_{\mathfrak{P}}$ and $\mathrm{Frob}_{\mathfrak{P}'}$ are conjugate in $\mathrm{Gal}(E/\mathbb{Q})$.*

Proof. Since $\mathrm{Gal}(E/\mathbb{Q})$ acts transitively on the prime ideals lying above (p), let $\sigma \in \mathrm{Gal}(E/\mathbb{Q})$ be such that $\mathfrak{P}' = \sigma(\mathfrak{P})$. For any $x \in \mathcal{O}_E$ we have $\mathrm{Frob}_{\mathfrak{P}}(\sigma^{-1}(x)) \equiv \sigma^{-1}(x)^p \pmod{\mathfrak{P}}$. Applying σ to both sides, we have

$$\sigma\mathrm{Frob}_{\mathfrak{P}}\sigma^{-1}(x) \equiv \sigma(\sigma^{-1}(x)^p) \pmod{\mathfrak{P}'} \equiv \sigma(\sigma^{-1}(x))^p \pmod{\mathfrak{P}'} \equiv x^p \pmod{\mathfrak{P}'}.$$

Thus $\sigma\mathrm{Frob}_{\mathfrak{P}}\sigma^{-1} = \mathrm{Frob}_{\sigma(\mathfrak{P})} = \mathrm{Frob}_{\mathfrak{P}'}$. □

Returning to the case of representations of the absolute Galois group let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(F)$ be an n-dimensional Galois representation over a field F. We want to extend the notion of ramification to such representations. First, we will need provide analogous definitions for the *absolute* decomposition and inertial groups in $G_{\mathbb{Q}}$. The approach will be similar to the case of finite extensions.

Given $p \in \mathbb{Z}$ prime and $\mathfrak{P}$ a prime ideal in $\overline{\mathbb{Z}}$ lying over (p), we would like to define $I_{\mathfrak{P}}$ and $D_{\mathfrak{P}}$ such that they fit into a short exact sequence

$$\begin{tikzcd} 1 \arrow[r] & I_{\mathfrak{P}} \arrow[r] & D_{\mathfrak{P}} \arrow[r] & \mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p) \arrow[r] & 1. \end{tikzcd}$$

Let $\mathfrak{P} \subset \overline{\mathbb{Z}}$ be a prime ideal lying over (p). Then as before, we have a reduction map $\overline{\mathbb{Z}} \to \overline{\mathbb{F}_p}$ defined by setting $\mathfrak{P}$ to be the kernel.

Definition 2.20. The **decomposition group** of $\mathfrak{P}$ is

$$D_{\mathfrak{P}} = \{\sigma \in G_{\mathbb{Q}} : \sigma(\mathfrak{P}) = \mathfrak{P}\}.$$

Observe that an element $\sigma \in D_{\mathfrak{P}}$ acts on $\overline{\mathbb{Z}}/\mathfrak{P}$ by $\sigma(x + \mathfrak{P}) = \sigma(x) + \mathfrak{P}$. The reduction map $D_{\mathfrak{P}} \to \mathrm{Gal}(\overline{\mathbb{F}}_p/\mathbb{F}_p)$ is surjective, and we define $I_{\mathfrak{P}}$ to be its kernel.

Definition 2.21. An **absolute Frobenius element over** p is any preimage $\mathrm{Frob}_{\mathfrak{P}} \in D_{\mathfrak{P}}$ of the Frobenius automorphism $\sigma_p \in G_{\mathbb{F}_p}$. Frob_p is thus defined only up to $I_{\mathfrak{P}}$.

As in the case of finite extension, the primes of $\overline{\mathbb{Z}}$ over (p) are conjugate, and thus

$$\mathrm{Frob}_{\sigma(\mathfrak{P})} = \sigma^{-1}\mathrm{Frob}_{\mathfrak{P}}\sigma, \quad \sigma \in G_{\mathbb{Q}}.$$

Theorem 2.22. *(Tchebotarov Density Theorem, weak version) Let $E/\mathbb{Q}$ be a Galois number field. Then every element of* $\mathrm{Gal}(E/\mathbb{Q})$ *is of the form* $\mathrm{Frob}_{\mathfrak{P}}$ *for finitely many maximal ideals $\mathfrak{P}$ of $\mathcal{O}_E$.*

A proof, as well as the strong statement of this theorem, can be found in [25]. Combining the above theorem with the definition of the Krull topology, we have the following.

Theorem 2.23. *For each maximal ideal $\mathfrak{P}$ of $\overline{\mathbb{Z}}$ lying over a finite set of primes p of $\mathbb{Z}$, fix an absolute Frobenius element* $\mathrm{Frob}_{\mathfrak{P}}$. *Then the set of all such Frobenius elements is dense in $G_{\mathbb{Q}}$.*

Proof. Let $U = U_\sigma(E)$ in G_Q. We need find an element of the form $\mathrm{Frob}_{\mathfrak{P}} \in U$. This exists if there is some ideal $\mathfrak{P}$ of $\overline{Z}$ such that $\mathrm{Frob}_{\mathfrak{P}}|_E = \sigma|_E$. Since $\sigma|_E$ is of the form $\mathrm{Frob}_{\mathfrak{P}_E}$ for some prime ideal $\mathfrak{P}_E$ of $\mathcal{O}_E$, by the Tchebotarov Density Theorem we can lift $\mathfrak{P}_E$ to a prime ideal $\mathfrak{P}$ of $\overline{\mathbb{Z}}$. □

Given a Galois representation ρ, we want to evaluate ρ at absolute Frobenius elements. That is, we want to make sense of $\rho(\sigma)$ for $\sigma \in G_{\mathbb{Q}}$, but since each absolute Frobenius element is defined only up to $I_{\mathfrak{P}}$, $\rho(\mathrm{Frob}_{\mathfrak{P}})$ is well defined only if $I_{\mathfrak{P}} \subset \ker(\rho)$. This is what it means for ρ to be unramified at p.

If $\mathfrak{P}$ and $\mathfrak{P}'$ lie over (p), then their inertia groups $I_{\mathfrak{P}}$ and $I_{\mathfrak{P}'}$ are conjugate in $G_{\mathbb{Q}}$. Since $\ker(\rho)$ is a normal subgroup of $G_{\mathbb{Q}}$, the condition $I_{\mathfrak{P}} \subset \ker(\rho)$ depends only on p.

Definition 2.24. Let ρ be a Galois representation and let $p \in \mathbb{Z}$ be prime. Then we say that ρ is **unramified at** p if $I_{\mathfrak{P}} \subset \ker(\rho)$ for any maximal ideal $\mathfrak{P} \subset \overline{\mathbb{Z}}$ lying over (p).

Example. Let $\chi : (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^\times$ primitive Dirichlet character. Recall that this means that the conductor of χ is n. Let $\pi_n : G_{\mathbb{Q}} \to \mathrm{Gal}(\mathbb{Q}(\mu_n)/\mathbb{Q})$ be the restriction. Recall the isomorphism $\varphi_n : \mathrm{Gal}(\mathbb{Q}(\mu_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ (equation 2.2) and let

$$\rho_\chi = \chi \circ \varphi_n \circ \pi_n : G_{\mathbb{Q}} \to \mathbb{C}^\times.$$

Let p be a prime not dividing N, and let $\mathfrak{P}$ be a maximal ideal lying over (p). The map π_n sends $I_{\mathfrak{P}}$ to $I_{\mathfrak{P}_N}$ where $I_{\mathfrak{P}_N} = \mathfrak{P} \cap \mathbb{Q}(\mu_N)$. Since p does not divide N, (p) is unramified in $\mathbb{Q}(\mu_N)$ so $I_{\mathfrak{P}} \subset \ker(\rho_\chi)$. That is, ρ_χ is unramified at p.

Example. Similarly, in the example of the p-adic cyclotomic character χ_p, suppose $p \neq p$ is prime and let $\mathfrak{P}$ lie over (p). Since p is unramified in $\mathbb{Q}(\mu_{p^n})$, every element in $I_{\mathfrak{P}}$ acts trivially on $\mathbb{Q}(\mu_{p^n})$ for all positive n. Thus $I_{\mathfrak{P}} \subset \ker(\chi_p)$ so χ_p is unramified at p.

Remark. If a Galois representation is unramified at all but finitely many primes p, the values $\rho(\mathrm{Frob}_{\mathfrak{P}})$ for $\mathfrak{P}$ over the unramified primes p will determine the representation ρ everywhere by continuity.

3 Elliptic curves and modular curves as algebraic curves

3.1 Elliptic curves over k

Definition 3.1. Let k be a field (of any characteristic). A **Weierstrass equation over** k is a cubic equation of the form

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \quad a_1, \ldots, a_6 \in k. \tag{3.1}$$

For simplification, define the following:

$$\begin{gathered} b_2 = a_1^2 + 4a_2, \qquad b_4 = a_1a_3 + 2a_4, \qquad b_6 = a_3^2 + 4a_6, \\ b_8 = a_1^2a_6 - a_1a_3a_4 + a_2a_3^2 + 4a_2a_6 - a_4^2, \\ c_4 = b_2^2 - 24b_4, \qquad c_6 = -b_2^3 + 36b_2b_4 - 216b_6. \end{gathered}$$

Define the **discriminant** of the equation to be

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$

and if the discriminant is nonzero, then define the **invariant** of the equation to be

$$j = c_4^3/\Delta.$$

To each Weierstrass equation E we associate a corresponding **Weierstrass polynomial**

$$E(x, y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 \in k[x, y]. \tag{3.2}$$

This definition is rather cumbersome and we will often prefer to work with a simpler equation via a change of variables. We will want to preserve the geometric group structure that three collinear points sum to zero when we make these change of variables. Such change of variables are precisely those given by affine transformations, and so we give the following definition.

Definition 3.2. A **admissible change of variables** in the equation of an elliptic curve is

$$x = u^2x' + r, \quad y = u^3y' + su^2x' + t, \qquad u, r, s, t \in k,\ u \neq 0.$$

If $\mathrm{char}(k) \neq 2$, then by an admissible change of variables with $u = r = 0$, $s = -a_1/2$, and $t = -a_3/2$, the Weierstrass equation can be reduced to the form

$$E : y^2 = x^3 + (b_2x^2 + 2b_4x + b_6)/4, \quad b_2, b_4, b_6 \in k. \tag{3.3}$$

If $\mathrm{char}(k) \neq 2, 3$ then by another admissible change of variables, this time with $u = 1/6$, $r = b_2/12$, and $s = t = 0$, the Weierstrass equation can be further reduced to the form

$$E : y^2 = x^3 - 27c_4x - 54c_6, \quad c_4, c_6 \in k. \tag{3.4}$$

Remark. Recall that in the case of elliptic curves over $\mathbb{C}$ we obtained cubic equations of the form $y^2 = 4x^3 - g_2x - g_3$. This cubic cannot be put in the form 3.1 via admissible changes of variables, and so is not a Weierstrass equation under our definition. There are more general definitions of Weierstrass equations and admissible changes of variables that one can give that includes cubics of this form. However, normalising the coefficients of y^2 and x^3 to 1 will simplify the formulae in the discussion to follow, so we work with these definitions. See the discussion following Definition 8.1.1. of [7].

Definition 3.3. Let $\bar{k}$ be an algebraic closure of k and let E be a Weierstrass equation over k. If E has nonzero discriminant Δ, then it is called **nonsingular** and the set

$$\mathcal{E} = \{(x, y) \in \bar{k}^2 \ : \ E(x, y) = 0\} \cup \{\infty\}$$

is called an **elliptic curve over** k.

A Weierstrass curve E is nonsingular if and only if the corresponding elliptic curve $\mathcal{E}$ is geometrically nonsingular. That is, at each point on the curve, at least one of the partial derivatives of the Weierstrass polynomial is nonzero.

Henceforth, the point at ∞ will be denoted $0_{\mathcal{E}}$. We view the elliptic curve $\mathcal{E}$ as sitting inside $\mathbf{P}^2(\bar{k})$ with the projective point $0_{\mathcal{E}}$ infinitely far in the y-direction, i.e. $0_{\mathcal{E}} = [0 : 1 : 0]$. For the affine curve $E(x, y) = y^2 - 4x^3 + g_2x + g_3$ we work with the corresponding homogeneous equation

$$E'([x : y : z]) = y^2z + a_1xyz + a_3yz^2 - x^3 - a_2x^2z - a_4xz^2 - a_6z^3.$$

If $z = 0$ for some point $[x : y : z]$ on the curve, then $x = 0$, so we may assume $y = 1$ by scaling the point $[0 : y : 0]$ by a factor of $1/y$. Thus, $0_{\mathcal{E}}$ is the unique point on the curve with $z = 0$. Since every other point $P \neq 0_{\mathcal{E}}$ on the curve has nonzero z-coordinate, by scaling we can write $P' = [x_0 : y_0 : 1]$ corresponding to the solution $P = (x_0, y_0)$ of the affine equation E.

To formalise the addition law on $\mathcal{E}$, we will need the following theorem.

Theorem 3.4. *(Bézout's Theorem) Let C and D be two projective plane curves over k (of any characteristic) whose defining polynomials have degrees d_1 and d_2 and are relatively prime in $\bar{k}[x, y]$. Then, counting with multiplicity, their intersection in $\mathbf{P}^2(\bar{k})$ consists of d_1d_2 points.*

Proof. The most natural setting to prove this theorem is in the language of schemes which we will not discuss here. See (I, Corollary 7.8) of [11] for a proof. □

Bézout's Theorem tells us that cubic curves, and only cubic curves, produce triples of collinear points, i.e. points satisfying an equation of the form $ax + by + cz = d$ where $a, b, c \in k$ are all not all zero. Now we can restate the addition law on elliptic curves $\mathcal{E}$ that collinear triples sum to $0_{\mathcal{E}}$

$$P + Q + R = 0_{\mathcal{E}} \iff P, Q, R \text{ are collinear.}$$

Remark. If $\text{char}(k) \neq 2, 3$, the choice $0_{\mathcal{E}} = [0 : 1 : 0]$ gives the addition law a particularly pleasing geometry. If $P = (x_0, y_0) \in \mathcal{E} \setminus \{0_{\mathcal{E}}\}$, then the point $Q = (x_0, -y_0)$ is also a solution to the Weierstrass polynomial, and so Q also lies on $\mathcal{E}$. The projective line through P and Q is given by $x = x_0z$ which also has solution $0_{\mathcal{E}}$. Thus, the three points $P, Q, 0_{\mathcal{E}}$ are collinear and the group law gives $P + Q = 0_{\mathcal{E}}$. Therefore $Q = -P$ – that is, the inverse of a point is given by reflection in the x-axis.

In general, for any fixed value of x, the y-values satisfying the two Weierstrass equation sum to $-a_1x - a_3$. The third collinear point is $0_{\mathcal{E}}$, so the additive inverse of a point $P = (x_p, y_p)$ is

$$-P = (x_p, -y_p - a_1x_p - a_3).$$

Given any points P and Q of $\mathcal{E}$, let $R = (x_R, y_R)$ be the third collinear point. The addition law $P + Q = -R$ says that in terms of affine coordinates, we have $P + Q = (x_R, -y_R - a_1x_R - a_3)$.

Commutativity of the addition law follows immediately from the addition law. We can also easily check that $0_{\mathcal{E}}$ acts as the identity element: the line between $0_{\mathcal{E}}$ and any point P intersects the curve at $-P$, noting that if $P = -P$, then this is a double intersection. Therefore $0_{\mathcal{E}} + P + (-P) = 0_{\mathcal{E}}$, so $0_{\mathcal{E}} + P = P$.

Associativity is much less obvious and will not be proved here. Chapter 3 of [15] presents a geometric proof of this result as a direct consequence of the Riemann Roch theorem.

Theorem 3.5. *Every elliptic curve $\mathcal{E}$ forms an abelian group with the point $0_{\mathcal{E}}$ as its additive identity.*

Algebraically, the group law is defined by rational functions over k. Let k_{prime} denote the prime subfield of k. Then the group law is defined by rational functions r, s over $k_{\text{prime}}(\{a_i\})$ where a_i are the coefficients of the Weierstrass equation 3.1.

Let $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ be points of $\mathcal{E}\backslash\{0_{\mathcal{E}}\}$. If $x_P = x_Q$ and $y_Q = -y_P - a_1x_P - a_3$, then $P = -Q$ so $P+Q = 0_{\mathcal{E}}$. Otherwise, $P+Q$ must lie in the affine part of $\mathcal{E}$. Let $P+Q = (x_{P+Q}, y_{P+Q})$ denote their sum. Then

$$\begin{aligned} x_{P+Q} = r(x_P, x_Q, y_P, y_Q) &= \lambda^2 + a_1\lambda - a_2 - x_P - x_Q \quad \text{and} \\ y_{P+Q} &= -(\lambda + a_1)r(x_P, x_Q, y_P, y_Q) - \mu - a_3 \end{aligned}$$

where

$$\lambda = \begin{cases} \frac{y_Q - y_P}{x_Q - x_P}, & x_P \neq x_Q \\ \frac{3x_P^2 + 2a_2x_P + a_4 - a_1y_P}{a_1x_P + a_3 + 2y_P}, & x_P = x_Q \end{cases} \quad \text{and} \quad \mu = \begin{cases} \frac{x_Qy_P - x_Py_Q}{x_Q - x_P}, & x_P \neq x_Q \\ \frac{-x_P^3 + a_4x_P + 2a_6 - a_3y_P}{a_1x_P + a_3 + 2y_P}, & x_P = x_Q. \end{cases}$$

The line $y = \lambda x + \mu$ passes through P and Q when $P \neq Q$ and it is tangent to $\mathcal{E}$ at P when $P = Q$.

Definition 3.6. Let K/k be an algebraic extension over k, i.e. $k \subset K \subset \overline{k}$. Then the **set of K-points of $\mathcal{E}$** is

$$\mathcal{E}(K) = \{P \in \mathcal{E} \setminus \{0_\epsilon\} : (x_P, y_P) \in K^2\} \cup \{0_{\mathcal{E}}\}.$$

Since $0_{\mathcal{E}} = [0 : 1 : 0]$ has coordinates in k and because the group law is rational over k, $\mathcal{E}(K)$ is a subgroup of $\mathcal{E}$. In particular, if $\mathcal{E}$ is an elliptic curve over $\mathbb{Q}$, then the set of points with rational affine coordinates along with $0_{\mathcal{E}}$ is a subgroup.

As is often the case in mathematics, to better understand a class of objects we will want to study the structure-preserving maps between them. In this case, we want to understand the isogenies of elliptic curves. In particular, isogenies give us a useful way to classify curves due to the following theorem.

Theorem 3.7. *Let k be a field. Let $\varphi : E_1 \to E_2$ be an isogeny between elliptic curves over k. Then there exists a* ***dual isogeny*** *$\psi : E' \to E$. In particular, the existence of an isogeny between elliptic curves is an equivalence relation.*

Proof. See Theorem 7.8.1 in [7]. □

Example. We can generalise the multiplication-by-N maps $[N]$ that we defined for complex tori to elliptic curves. Let

$$[N] : \mathcal{E} \to \mathcal{E}$$

denote N-fold addition, i.e. for all $P \in \mathcal{E}$, $[N]P$ is the sum of N copies of P. This is an isogeny, and to any elliptic curve $\mathcal{E}$, for every positive integer N we can associate a set of N-torsion points.

Definition 3.8. The N-torsion points of $\mathcal{E}$ is group given by the kernel of $[N]$,

$$\mathcal{E}[N] = \{P \in \mathcal{E} : [N]P = 0_{\mathcal{E}}\}.$$

Theorem 3.9. *(Structure theorem for $\mathcal{E}[N]$). Let $\mathcal{E}$ be an elliptic curve over a field k and let $N \in \mathbb{Z}^+$. Then*

$$\mathcal{E}[N] \cong \prod \mathcal{E}[p^{e_p}], \quad \textit{where } N = \prod p^{e_p}.$$

If $p \neq \operatorname{char}(k)$, then

$$\mathcal{E}[p^e] \cong (\mathbb{Z}/p^e\mathbb{Z})^2.$$

If $p = \operatorname{char}(k)$, then either

$$\mathcal{E}[p^e] \cong \mathbb{Z}/p^e\mathbb{Z} \quad \textit{or} \quad \mathcal{E}[p^e] = \{0\}$$

for all $e \geq 1$.

Proof. See Theorem 7.1.3. in [7]. □

In particular, if $\operatorname{char}(k) = 0$, then for any positive integer N we have

$$\mathcal{E}[N] \cong (\mathbb{Z}/N\mathbb{Z})^2.$$

We will revisit this result in section 5.3.

Remark. The isomorphism in the theorem above is non-canonical – each choice of ordered basis (P, Q) of $\mathcal{E}[N]$ determines such an isomorphism.

Definition 3.10. Let k be a field of characteristic p. Then if $\mathcal{E}[p] \cong \mathbb{Z}/p\mathbb{Z}$, then we say that $\mathcal{E}$ is **ordinary**. Otherwise if $\mathcal{E}[p] = \{0\}$, we say that $\mathcal{E}$ is **supersingular**.

Lemma 3.11. *The projective point $[0 : 1 : 0]$ is always nonsingular.*

Proof. Let $E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ be a Weierstrass equation over k. Homogenising the polynomial with z, we obtain the corresponding projective equation

$$E'[x : y : z] = y^2z + a_1xyz + a_3yz^2 - x^3 - a_2x^2z - a_4x - a_6.$$

Dehomogenising E' by setting $y = 1$, we obtain

$$\tilde{E}(\tilde{x}, \tilde{z}) = \tilde{z} + a_1\tilde{x}\tilde{z} + a_3\tilde{z}^2 - \tilde{x}^3 - a_2\tilde{x}^2\tilde{z} - a_4\tilde{x}\tilde{z}^2 - a_6\tilde{z}^3$$

with the point $[0 : 1 : 0]$ corresponding to $(\tilde{z}, \tilde{w}) = (0, 0)$. The Jacobian matrix of $\tilde{E}$ is

$$D\tilde{E} = \begin{bmatrix} a_1\tilde{z} - 3\tilde{x}^2 - 2a_2\tilde{z} - a_4\tilde{z}^2 & 1 + a_1\tilde{x} + 2a_3\tilde{z} - a_2\tilde{x}^2 - 2a_4\tilde{x} - 3a_6\tilde{z}^2 \end{bmatrix}$$

so $D\tilde{E}_{(0,0)} = \begin{bmatrix} 0 & 1 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \end{bmatrix}$. □

It follows that a singular point P of a Weierstrass equation is affine. Furthermore, the coordinates of P lie in k. To see this, first suppose $\text{char}(k) \neq 2$. In this case, the Weierstrasss equation E can be given in the form (3.3) so we may take $a_1 = a_3 = 0$. The partial derivatives are $\partial E/\partial x = -3x^2 - 2a_2x - a_4$ and $\partial E/\partial y = 2y$. The partial derivatives vanish at P so $P = (x_0, 0)$ where x_0 is a repeated root of the cubic polynomial $f(x) = x^3 + a_2x^3 + a_4x + a_6$ over k, and so $g = \gcd(f, f')$ is a nonconstant polynomial over k. Since $\deg(g) \leq \deg(f') = 2$, the degree of g is either 1 or 2. If $\deg(g) = 1$ then $g(x) = x - x_0 \in k[x]$ so $x_0 \in k$. Otherwise if $\deg(g) = 2$ then g is a multiple of f'. Thus f' divides f, and so f' cannot have distinct roots, otherwise they would both be double roots of the cubic f. Hence $f' = \alpha(x - x_0)^2 = \alpha(x^2 + 2x_0 + x_0^2) \in k[x]$ for some $\alpha \in k$ so $x_0 \in k$.

If $\text{char}(k) = 2$ then the partial derivatives of the Weierstrass equation are $\partial E/\partial x = -x^2 - a_4$ and $\partial E/\partial y = a_1x + a_3$. If $a_1 \neq 0$ then for any singular point $P = (x_0, y_0)$ the vanishing of the partial derivative with respect to y gives $x_0 = -a_3/a_1 \in k$. Substituting this back into the Weierstrass equation we see that $y_0^2 = -x_0^3 + a_2x_0^2 - a_4x_0$. If $a_1 = 0$, then $x_0^2 = -a_4$ and $y_0^2 = x_0^3 + a_2x_0^2 + a_4x_0 + a_6$. In either case, since every element of k is square, the coordinates lie in k.

Thus, the admissible change of variables $x = x' - x_0$ and $y = y' - y_0$ takes any a singular point to $(0, 0)$. In the new coordinates, the condition for both partial derivatives to vanish at zero for the Weierstrass polynomial to be of the form

$$E(x, y) = y^2 + a_1xy - x^3 - a_2x^2 = (y - m_1x)(y - m_2x)x^3, \quad c_4 = (m_1 - m_2)^4$$

and the point $P = (0, 0)$ is the only singular point of E – that is, if a Weierstrass equation has a singular point, then it is its unique singular point.

Definition 3.12. A singular point P is called a **node** if two distinct tangent lines pass through the curve at P, and it is called a **cusp** is there is only one tangent line.

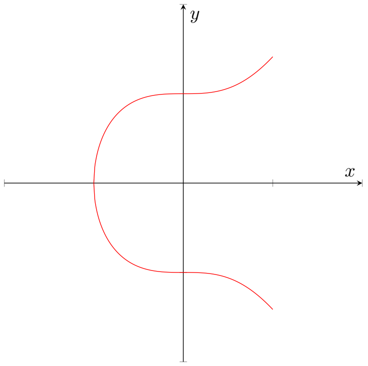
In particular, if $m_1 \neq m_2$, then P is a node, and if $m_1 = m_2$ the P is a cusp. This leads to the following proposition.

Proposition 3.13. *Suppose E is a Weierstrass equation over k. Then*

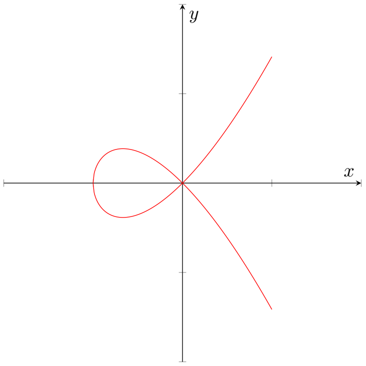
i. E is the Weierstrass equation of an elliptic curve $\iff \Delta \neq 0$,

ii. E describes a curve with a node $\iff \Delta = 0$ and $c_4 \neq 0$,

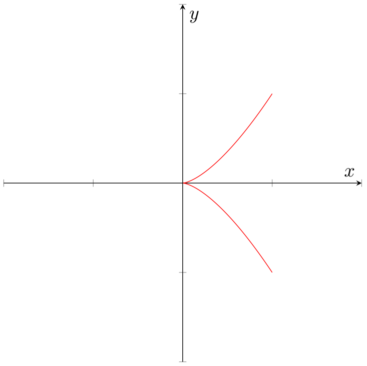
iii. E describes a curve with a cusp $\iff \Delta = 0$ and $c_4 = 0$.



(i) An elliptic curve



(ii) A curve with a node



(iii) A curve with a cusp

3.2 Function fields of algebraic curves

In the previous section, we saw that elliptic curves have both an algebraic structure (as an abelian group) and a geometric structure (as a smooth, projective curve of genus 1). In this section, we will focus on the latter and first familiarise ourselves with the mathematical language used to describe algebraic curves. This is motivated by an important result from algebraic geometry that curves are determined by their function fields – the precise version of the statement, which will be presented at the end of this section, will require us to impose appropriate equivalence relations on both curves and on fields.

Let k be a field and let $n, m \in \mathbb{Z}^+$. Consider a set of m polynomials $\varphi_1, \ldots, \varphi_m \in k[x_1, \ldots, x_n]$. Let

$$I = \langle \varphi_1, \ldots, \varphi_m \rangle \subset \overline{k}[x_1, \ldots, x_n]$$

denote the ideal generated by the φ_i in the ring of polynomials over the algebraic closure $\overline{k}$ of k. Let

$$C = \{P \in \overline{k}^n : \varphi(P) = 0 \forall \varphi \in I\}$$

be the simultaneous solutions of all polynomials in I.

Definition 3.14. Suppose that I is a prime ideal. Then the **coordinate ring $\overline{k}[C]$ of C over $\overline{k}$** is

$$\overline{k}[C] = \overline{k}[x_1, \ldots, x_n]/I.$$

The **function field $\overline{k}(C)$ of C over $\overline{k}$** is the field of fractions of the coordinate ring.

Definition 3.15. Let t be transcendental over $\overline{k}$. If $\overline{k}(C)$ is a finite extension of $\overline{k}(t)$ the we say C is an **affine algebraic curve over k**.

If for each point $P \in C$ the derivative matrix $[D_j \varphi_i(P)]$ has rank $n - 1$, then C is **nonsingular**.

For convenience, we will often work in the affine coordinate system. The projective version of the curve is obtained as follows.

Homogenise the polynomials $\varphi_1, \ldots, \varphi_n$ with another variable x_0 (i.e. multiply each term by a power of x_0 such that the resulting polynomial is homogeneous). Denote the homogenised polynomials

$\varphi_{1,\text{hom}}, \dots, \varphi_{n,\text{hom}}$ and let

$$I_{\text{hom}} = \langle \varphi_{1,\text{hom}}, \dots, \varphi_{n,\text{hom}} \rangle$$

be the ideal generated by the homogenised polynomials in $\overline{k}[x_0, \dots, x_n]$.Then the **projective algebraic curve over** k is given by

$$C_{\text{hom}} = \{P \in \mathbf{P}^n(k) : \varphi(P) = 0 \, \forall \text{ homogeneous } \varphi \in I_{\text{hom}}\}.$$

Similarly, the definition of nonsingularity can be extended where for the points outside $\overline{k}^2$ we dehomogenise in appropriate coordinate systems.

Remark. Elliptic curves correspond to the projective case where $m = 1$, $n = 2$, and C is generated by a Weierstrass polynomial $E(x, y)$. The function field is $\overline{k}(x)[y]/\langle E(x, y)\rangle$. Since the constraint on the discriminant of the corresponding Weierstrass equation ensures that at least one of the partial derivatives, they are projective algebraic curves over k that are nonsingular in the sense of Definition 3.15.

Let C be a nonsingular algebraic curve over k and let $f + I \in \overline{k}[C]$ be any element in the coordinate ring. Since $\varphi(P) = 0$ for all $\varphi \in I$, $f + I$ takes a well-defined point in $\overline{k}$ at each point $P \in C$. That is, if $g \in f + I$ is any other representative, $f(P) = g(P)$ for all $P \in C$.

Thus, a polynomial on C defined a map $C \to \overline{k}$. We would also like to know if a rational function can similarly define a map from $C \to \overline{k} \cup \{\infty\} = \mathbb{P}^1(\overline{k})$. It is not clear whether such a map is well defined since an element $f/g \in \overline{k}(C)$ might be such that $f(P) = g(P) = 0$. We will see that there is, in fact, such a mapping on all of C_{hom}. To get there, we will first need to discuss localisations of the coordinate ring.

Let $P = (p_1, \dots, p_n) \in C$. By Hilbert's Nullstellensatz, the functions in the coordinate ring $\overline{k}[C]$ that vanish at P,

$$m_P := \{f \in \overline{k}[C] : f(P) = 0\} = \langle \{x_i - p_i + I : 1 \leq i \leq n\} \rangle$$

is a maximal ideal of $\overline{k}[C]$. In particular, it is prime so we can consider the localisation

$$\overline{k}[C]_P = (\overline{k}[C] - m_P)^{-1}\overline{k}[C] = \{f/g \in \overline{k}(C) : g(P) \neq 0\}$$

which we call the **local ring of** C **over** $\overline{k}$ **at** P. The unique maximal ideal

$$M_P = m_P \overline{k}[C]_P = \{f/g \in \overline{k}[C]_P : f(P) = 0\}$$

of $\overline{k}[C]_P$ is called the **local maximal ideal at** P.

Lemma 3.16. *Let C be a nonsingular algebraic curve over k. For any $P \in C$, the quotient m_P/m_P^2 is a 1-dimensional vector space over $\overline{k}$.*

Proof. See Lemma 7.2.2. in [7]. □

Proposition 3.17. *Let C be a nonsingular curve over k and let $P \in C$ be a nonsingular point. The ideal M_P is principal.*

Proof. Define the following function:

$$\begin{aligned} i : m_P &\to M_P/M_P^2 \\ f &\mapsto f + M_P^2. \end{aligned}$$

The kernel of this map is

$$\ker(i) = m_P \cap M_P^2.$$

We claim that $m_P \cap M_P^2 = m_P^2$. The inclusion $m_P^2 \subset m_P \cap M_P^2$ is clear. In the other direction, suppose $r \in m_P \cap M_P^2$. Then $s = r/t$ for some $r \in m_P^2$ and $t \in \overline{k}[C] - m_P$. Since m_P is a maximal ideal in $\overline{k}[C]$, $t + m_P$ is invertible in the field $\overline{k}[C]/m_P$, so there exists $v \in \overline{k}[C]$ such that $tv - 1 \in m_P$. Since $r = st$, we have $0 = v(r - st) = s - (stv - rv - s) = s - (s(tv-1) - rv)$, and so $s = s(tv-1) - rv$. Since $s(tv-1)$, $rv \in m_P^2$, it follows that $s = m_P^2$.

The map i is also surjective since for any $f/g \in M_P$, we have $i(f/g(P)) = f/g + M_P^2$ since $f/g(P) - f/g = f(g - g(P))/(g(P)g) \in M_P^2$. Hence i induces an isomorphism

$$\tilde{i} : m_P/m_P^2 \xrightarrow{\sim} M_P/M_P^2.$$

Combining this with the previous lemma, M_P/M_P^2 is a 1-dimensional vector space over $\overline{k}$. Let $a \in \overline{k}[C]$ be such that $i(a)$ is nonzero. Let $N = a\overline{k}[C]_p$. Since M_P is a finitely generated $\overline{k}[C]_p$ module, so is the quotient M_P/N. Since $M_P(M_P/N) = (N + M_P^2/N)$ and $N + M_P^2 = (a\overline{k} + M_P^2)\overline{k}[C]_P$, we have that $M_P(M_P/N) = M_P/N$. By Nakayama's Lemma, $M_P/N = 0$. Thus a also generates M_P as an ideal of $\overline{k}[C]_P$. □

Definition 3.18. A generator a of M_P is called a **uniformiser at** P.

Suppose $F \in \overline{k}[C]_P$ is nonzero and let a be a uniformiser at P. Either F is a unit or there is a nonzero element $F_1 \in \overline{k}[C]_P$ such that $F = tF_1$. Similarly, F_1 is either a unit or it is of the form $F_1 = tF_2$ for some nonzero element $F_2 \in \overline{k}[C]_P$. This process continues, leading to an ascending chain of ideals

$$\langle F \rangle \subsetneq \langle F_1 \rangle \subsetneq \langle F_2 \rangle \subsetneq \dots.$$

However, this chain must be finite – $\overline{k}$ is Noetherian since it is a field, so by the Hilbert Basis Theorem, any polynomial ring $\overline{k}[x_1, \dots, x_n]$ over $\overline{k}$ is also Noetherian. Any (two-sided) ideal I in a Noetherian ring R is also Noetherian since the ideals in R/I correspond precisely to ideals in R containing I. Thus, $\overline{k}[C]$ is also Noetherian. Finally, any localisation of a Noetherian ring is Noetherian (see Proposition 3.5 in [16]), and so $\overline{k}[C]_P$ is Noetherian, and it follows that the chain of ideal is finite. Therefore F must be of the form $F = t^e u$ where $e \in \mathbb{N}$ and u is a unit in $\overline{k}[C]_P$.

Furthermore, this representation is unique – if $F = t^e u = t^{e'} u'$ with $e \geq e'$ then $t^{e-e'} u = u'$, so $e = e'$ and $u = u'$.

Definition 3.19. Let a be a uniformiser at P. The **valuation at** P on the coordinate ring $\overline{k}[C]$ is the function

$$\nu_P : \overline{k}[C] \to \mathbb{Z}_{\geq 0} \cup \{+\infty\}$$

$$\nu_P(f) = \begin{cases} +\infty, & f = 0, \\ e, & f = a^e u. \end{cases}$$

This extends to the function field $\overline{k}(C)$ via $\nu_P(f/g) = \nu_P(f) - \nu_P(g)$.

We check that the extension of ν_P to the function field $\overline{k}(C)$ is well-defined. If $F = f/g = f'/g' \in \overline{k}(C)$ with $f = a^e u$, $g = a^d v$, $f' = a^{e'}u'$, and $g = a^{d'}v'$, then we have $a^{e-d}uv^{-1} = a^{e'-d'}u'v'^{-1}$. Therefore $a^{(e-d)-(e'-d')} = u'v(uv')^{-1}$ where the right-hand side is a unit. Therefore $\nu_P(f) - \nu_P(g) - (\nu_P(f') - \nu_P(g')) = (e-d) - (e'-d') = 0$.

Proposition 3.20. *Let C be a nonsingular algebraic curve and let $P \in C$ be any point on the curve. The valuation $\nu_P : \overline{k}(C) \to \mathbb{Z} \cup \{+\infty\}$ is surjective. Additionally, for all nonzero $F, G \in \overline{k}(C)$,*

i. $\nu_P(FG) = \nu_P(F) + \nu_P(G)$,

ii. $\nu_P(F+G) \geq \min\{\nu_P(F), \nu_P(G)\}$ *with equality if* $\nu_P(F) \neq \nu_P(G)$.

Proof. For surjectivity, let a be a uniformiser at P. Then $\nu_P(0) = +\infty$ and for any $n \in \mathbb{Z}$ we have $\nu_P(a^n) = n$, so ν_P is surjective. For (i), let $F = f/f'$ and $G = g/g'$. By multiplying both the numerators and the denominators by a unit, we can assume $f = a^e u$, $g = a^d v$, $f' = a^{e'}$, and $g = a^{d'}$ where $u, v \in \left(\overline{k}[C]_P\right)^\times$. Then $FG = a^{e+d}uv/(a^{e'+d'})$, so $\nu_P(FG) = \nu_P(a^{e+d}uv) + \nu_P(a^{e'+d'}) = e + d + e' + d' = (e+e') + (d+d') = \nu_P(F) + \nu_P(G)$.

Finally for (ii), let F and G be as above. Without loss of generality, we may assume $\nu_P(F) \leq \nu_P(G)$ and let $e + d' + c = e' + d$ for some $c \geq 0$ with equality if $\nu_P(F) = \nu_P(G)$. Then

$$F + G = \frac{a^e u}{a^{e'}} + \frac{a^d v}{a^{d'}} = \frac{a^{e+d'}u + a^{d+e'}v}{a^{e'+d'}} = \frac{a^{e+d'}(u + a^c v)}{a^{e'+d'}}.$$

Due to (i) we have

$$\begin{aligned}\nu_P(F+G) &= \nu_P(a^{e+d'}) - \nu_P(a^{e'+d'}) + \nu_P(u + a^c v) \\ &= \min\{\nu_P(F), \nu_P(G)\} + \nu_P(u + a^c v).\end{aligned}$$

If $\nu_P(F) \neq \nu_P(G)$, then $c > 0$ so $(u + a^c v)(P) = u(P) \neq 0$ and therefore $u + a^c v$ is a unit. Hence $\nu_P(u + a^c v) = 0$ and we have equality $\nu_P(F+G) = \min\{\nu_P(F), \nu_P(G)\}$. On the other hand if $\nu_P(F) = \nu_P(G)$ then $c = 0$ and $u + a^c v = u + v$. If $(u+v)(P) \neq 0$ then we have equality. Otherwise if $(u+v)(P) = 0$ then $u + v \in M_p$ so $\nu_P(u + a^c v) \in \mathbb{Z}_{\geq 0} \cup \{\infty\}$. □

We are now ready to describe a mapping $C \to \mathbb{P}^1(\overline{k})$ defined by an element of the function field $\overline{k}(C)$. If $F \in \overline{k}(C)$ is the zero function, then it defines the zero mapping. Otherwise, at each point $P \in C$, any nonzero F can be written in the form

$$F = a^{\nu_P(F)}\frac{f}{g}, \quad \nu_P(F) \in \mathbb{Z},\ f, g \in \overline{k}[C],\ f(P), g(P) \neq 0$$

defining a mapping

$$F(P) = \begin{cases} 0, & \nu_P(F) > 0, \\ \infty, & \nu_P(F) < 0, \\ f(P)/g(P), & \nu_P(F) = 0. \end{cases}$$

Remark. Since the mapping depends only on local information, it can be extended to all of C_{hom} by using an appropriate coordinate system at each point. The nonempty affine pieces obtained from dehomogenising the polynomials defining a nonsingular projective curve C_{hom} lead to the same function field up to isomorphism irrespective of which variable we set to 1. Moreover, the isomorphisms between any two such affine function fields preserves valuations on the overlap of the two corresponding affine pieces of C_{hom}. We will revisit this idea when we discuss reduction of algebraic curves in Section 3.5.

As we indicated in the introduction to this section, our current goal is to establish a correspondence between appropriate equivalence classes of algebraic curves and function fields. Now, we will describe what these equivalence relations are.

We will begin with curves. Let C be a nonsingular projective algebraic curve over k and let $r \in \mathbb{Z}^+$. Consider an element

$$h = [F_0 : \ldots : F_r] \in \mathbb{P}^r(\overline{k}(C)).$$

For each point $P \in C$, let a_P be a uniformiser at P and define the valuation of h to be

$$\nu_P(h) = \min\{\nu_P(F_0), \ldots, \nu_P(F_r)\}.$$

Note that this is well-defined since the valuation at a point is invariant under multiplication by a unit. Then h defines a map

$$\begin{aligned} h : C &\to \mathbb{P}^r(\overline{k}) \\ P &\mapsto [(a_P^{-\nu_P(h)} F_0)(P), \ldots, (a_P^{-\nu_P(h)} F_r)(P)]. \end{aligned}$$

If $C' \subset \mathbb{P}^r(\overline{k})$ is another nonsingular algebraic curve over k, then h is a morphism from C to C' if $h(P) \in C'$ for all $P \in C$. We also define isomorphisms in the usual way, that is, a morphism h from C to C' is an isomorphism if there exists a morphism h' from C' to C such that the compositions $h' \circ h$ and $h \circ h'$ are the identity maps of C and C' respectively, in which case we say C and C' are isomorphic.

This is clearly an equivalence relation, but we will want a finer one for our purposes.

Definition 3.21. Let C be a nonsingular projective algebraic curve over k characterised by the zero set of the ideal $I \subset \overline{k}[x_1, \ldots, x_n]$. An element $F \in \overline{k}(C)$ is **defined over** k if it can be written in the form $F = f/g$ where $f, g \in k[C] = k[x_1, \ldots, x_n]/I_k$, $I_k = I \cap k[x_1, \ldots, x_n]$.

The set of elements in $\overline{k}(C)$ defined over k form a subfield denoted

$$k(C) = \{F \in \overline{k}(C) : F \text{ is defined over } k\}.$$

A morphism $h = (F_1, \ldots, F_r)$ of curves over k is **defined over** k if each component F_i is defined over k.

The gives us the equivalence relation we need: two nonsingular projective curves C and C' are isomorphic over k if they are isomorphic via morphisms h and h' over k.

Remark. In the case of elliptic curves over k, the isomorphisms turn out to be precisely the admissible change of variables. The distinction between the two isomorphism classes is that the former allows the terms u, r, s, t in the change of variable formulae to lie in $\overline{k}$, whereas in the latter they must lie in k.

Now we define the equivalence relation on function fields.

Definition 3.22. Let k be a field. Then a field K is a **function field over** k if $K \cap \overline{k} = k$ and K is a finite extension of some field $k(t)$ where t is transcendental over k.

Remark. This definition is consistent with our definition for function fields $k(C)$ and $\overline{k}(C)$ of an algebraic curve C over k, i.e., $k(C)$ is a function field over k and $\overline{k}(C)$ is a function field over $\overline{k}$.

Definition 3.23. Let K and K' be function fields over k. Then K and K' are **conjugate over** k if there exists an isomorphism of fields $K \to K'$ that fixes k pointwise.

This gives us a the notion of equivalence on function fields that we need.

Theorem 3.24. *(Curves-Fields correspondence). The map*

$$C \mapsto k(C)$$

induces a bijection

$$\left\{ \begin{array}{c} \textit{isomorphism classes over } k \textit{ of} \\ \textit{nonsingular projective algebraic curves} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{conjugacy classes over } k \textit{ of function} \\ \textit{fields over } k \end{array} \right\}.$$

For any nonconstant morphism $h : C \to C'$ over k of nonsingular projective curves over k, let

$$\begin{aligned} h^* &: k(C') \to k(C) \\ h^*(G) &= G \circ h \end{aligned}$$

be the pullback. Then the map

$$h \mapsto h^*$$

is a bijection from the set of nonconstant morphisms $C \to C'$ over k to the set of k-injections of $k(C')$ and $k(C)$.

This statement of the correspondence has been taken from section 7.2 of [7] where it presented without proof. A proof of a more general statement (in terms of varieties) can be found in section 7.4 of [29].

The significance of the Curves-Field correspondence is that rather than constructing algebraic curves directly, we can instead construct suitable function fields.

In particular, suppose $\text{char} k = 0$ and that K is a function field over over $k(t)$ where t is transcendental. The corresponding curve is obtained as follows. Since the extension $K/k(t)$ is finite, the Primitive Element Theorem tells us that $K = k(t, u)$ for some generator u satisfying an irreducible polynomial over $k(t)$. By clearing denominators we obtain a polynomial relation $\varphi(t, u) = 0$ for some $\varphi \in k([x, y]$. φ is also irreducible over the algebraic closure $\overline{k}$ since $K \cap \overline{k} = k$. The zero set of φ in $\overline{k}$ defines a plane curve C, possibly with finitely many singular points. The curve C can be desingularised via blowup to produce a nonsingular curve C' with function field K. However, the map (t, u) taking C' to C is not necessarily an isomorphism. Instead, we can only guarantee that the map is a *birational equivalence.*

The natural setting for defining and studying birational morphisms is that of algebraic geometry. For simplicity, we will only provide an ad-hoc definition here. Roughly, a rational map from C' to C is an element $h \in \mathbf{P}^r(\overline{k}(C))$ that is only required to define a map from all but finitely many points of C'. Such a map is a birational equivalence if there exists a rational map C to C' such that both composites are the respective identity maps on C' and C.

Remark. The condition that h be a nonconstant morphism comes from a generalisation of the result that holomorphic maps from a compact Riemann surface to any Riemann surface is either constant or surjective. The statement for curves is that any morphism from a complete (in the sense of an algebraic variety) nonsingular curve over k to any curve over k is either constant or surjective (see Section II Proposition 6.8 in [11]).

Definition 3.25. If $h : C \to C'$ is a nonconstant morphism of nonsingular projective algebraic curves, then the **degree of** h is defined as the degree of the field extension

$$\deg(h) = [\overline{k}(C) : h^*(\overline{k}(C))].$$

We extend this definition to constant morphisms by giving them degree 0.

The degree of a composition of nonconstant morphisms is the product of the degrees. Thus, a morphism is an isomorphism if and only if its degree is 1. The degree counts the number of inverse images under h of each point in C', counted with a suitably defined algebraic notion of ramification.

Definition 3.26. Let $h : C \to C'$ be a morphism of nonsingular projective algebraic curves. The **ramification degree of** h at a point $P \in C$ is

$$e_P(h) = \nu_P(t' \circ h)$$

where t' is a uniformiser at $h(P)$.

3.3 Divisors on algebraic curves

Suppose C is a nonsingular projective algebraic curve over an algebraically closed field k. As in the complex case, we can associate to X a group $\mathrm{Div}(C) = \mathrm{Div}(C/\overline{k})$, the set of formal finite $\mathbb{Z}$-linear combinations $\sum_{x \in C} n_x x$ of closed (under the Zariski topology) points in X. Let $\mathrm{Div}^0(C)$ denote the subgroup of degree-0 divisors on C.

If we relax the condition on k and only require that k is a perfect field, let the group $\mathrm{Div}(C/k)$ be the subgroup

$$\mathrm{Div}(C/k) = H^0(\mathrm{Gal}(\overline{k}/k), \mathrm{Div}(C/\overline{k})$$

consisting of elements of $\mathrm{Div}(C/\overline{k})$ that are fixed by all the automorphisms of $\overline{k}/k$. Similarly, let $\mathrm{Div}^0(C/k)$ be the degree-0 divisors in $\mathrm{Div}(C/k)$.

There is a natural homomorphism $\overline{k}(C)^\times \to \mathrm{Div}(X)$ associating to a rational function $F \in \overline{k}(C)$ its divisor

$$(F) = \sum_{P \in C} \mathrm{ord}_P(f)P.$$

Example. Suppose $\mathcal{E}$ is the elliptic curve over $\mathbb{C}$ defined by $y^2 = x^3 + ax + b$. Consider the rational functions x and y. The divisor of x is

$$(x) = (0, \sqrt{b}) + (0, -\sqrt{b}) - 20_{\mathcal{E}}$$

and the divisor of y is

$$(y) = (x_1, 0) + (x_2, 0) + (x_3, 0) - 0_{\mathcal{E}}$$

where x_1, x_2, and x_3 are the roots of the cubic $x^3 + ax + b = 0$. The divisor of x/y is therefore

$$\left(\frac{x}{y}\right) = (0, \sqrt{b}) + (0, -\sqrt{b}) - (x_1, 0) - (x_2, 0) - (x_3, 0) + 0_{\mathcal{E}}.$$

It is a result in Riemann surface theory that a nonzero meromorphic function on a compact Riemann surface always has the same number of zeros as the number of poles, counting with multiplicity (see Corollary 4.25 in [9]). This extends to projective nonsingular algebraic curves, that is, every principal divisor on such a curve as degree 0. This result can be found as Corollary 6.10. in section II of [11]. Consequently, if $f \in \overline{k}(C)$ is nonzero, then $(f) \in \mathrm{Div}^0(X)$. As such, it makes sense to define the Picard group in an analogous way to the complex case.

Definition 3.27. Let C be an algebraic curve. A divisor of the form $(F) \in \mathrm{Div}(C)$ for some $F \in \overline{k}(C)$ is called **principal**. The set of principal divisors is denoted $\mathrm{Div}^P(C)$. Any two divisors of C are **linearly equivalent** if their difference is a principal divisor.

Definition 3.28. The **Picard group** $\mathrm{Pic}(C)$ of a nonsingular projective algebraic curve is the group $\mathrm{Div}(C)$ of divisors on C modulo linear equivalence.

The **degree-**0 **Picard group of** C is the subgroup of degree-0 divisors on C modulo linear equivalence.

There is an exact sequence of abelian groups

$$\begin{tikzcd} 0 \arrow[r] & \overline{k}(C)^{\times} \arrow[r] & \mathrm{Div}^0(C) \arrow[r] & \mathrm{Pic}^0(X) \arrow[r] & 0. \end{tikzcd}$$

The of these groups, the group of divisor classes of degree zero $\mathrm{Pic}^0(C)$ will be of particular interest to us. In algebraic geometry it perhaps more well-known as the *it* Jacobian of the curve C; more precisely, $\mathrm{Pic}^0(C)$ can be shown to be canonically isomorphic to $\mathrm{Jac}(C)$ via the Abel-Jacobi map. The general construction of the Jacobian as an abelian variety is rather complicated and will not be discussed here. This construction will note be of interest and we will instead continue to work with $\mathrm{Pic}^0(C)$.

Definition 3.29. Let C be a smooth projective curve over k with a rational point $0_C \in C(k)$. The **Abel-Jacobi map** is defined as

$$\begin{aligned} C(k) &\to \mathrm{Pic}^0(C) \\ P &\mapsto P - 0_C. \end{aligned}$$

Now suppose $\pi : C \to C'$ is a morphism of nonsingular projective algebraic curves over k. We can consider both the pullback and the push forward of the divisor groups $\mathrm{Div}(C)$ and $\mathrm{Div}(C')$.

If $D \in \mathrm{Div}(C')$, then the pullback $\pi^*(D)$ is a divisor on C defined as follows. For any point $P \in \mathrm{Div}(C'/\overline{k})$, let

$$\pi^*(P) = \sum_{Q \in \pi^{-1}(P)} e_{Q/P} Q$$

where $e_{Q/P}$ is the ramification degree of Q over P. Then $\pi^*(D)$ is obtained by extending $\mathbb{Z}$-linearly.

Lemma 3.30. *Suppose $\pi : C \to C'$ is a morphism of nonsingular projective algebraic curves over k. The pullback*

$$\pi^* : \mathrm{Div}(C') \to \mathrm{Div}(C)$$

induces a homomorphism

$$\tilde{\pi}^* : \mathrm{Pic}^0(C') \to \mathrm{Pic}^0(C).$$

In fact, this describes a contravariant functor, known as the **Picard functor**, from the category of nonsingular projective algebraic curves over k to the category of abelian groups. That is, the Picard functor sends a curve C to $\mathrm{Pic}^0(C)$ and sends a morphism $\pi : C \to C'$ to $\tilde{\pi}^* : \mathrm{Pic}^0(C') \to \mathrm{Pic}^0(C)$.

On the other hand, we define the push forward as follows. For a point $P \in \mathrm{Div}(C/\overline{k})$, let $\pi_*(P) = \pi(P)$. Then $\pi_* : \mathrm{Div}(C) \to \mathrm{Div}(C')$ induces a homomorphism

$$\tilde{\pi}_* : \mathrm{Pic}^0(C) \to \mathrm{Pic}^0(C').$$

This describes a covariant functor sending a curve C to $\mathrm{Pic}^0(C)$ and a morphism $\pi : C \to C'$ to $\tilde{\pi}_* : \mathrm{Pic}^0(C) \to \mathrm{Pic}^0(C')$. This is known as the **Albanese functor**.

Proposition 3.31. *Let C be a nonsingular projective algebraic curve over k. Let $S \subset C$ be any finite subset. Then every divisor class in $\mathrm{Pic}^0(C)$ contains a representative $\sum_{P \in C} n_P P$ such that $n_P = 0$ for all $P \in S$.*

Proof. See Proposition 7.3.1. in [7].

□

Lemma 3.32. *Let $\mathcal{E}$ be an elliptic curve over k and let $P, Q \in \mathcal{E}$. Then the divisor $P - Q$ is principal if and only if $P = Q$.*

Proof. If $P = Q$ the the divisor $P - Q$ is the divisor of the rational function 0. In the other direction, suppose $F \in \overline{k}(\mathcal{E})$ is a rational function with divisor $(F) = P - Q$. If $P \neq Q$ then F is an isomorphism $\mathcal{E} \to \mathbb{P}^1(\overline{k})$. However, $\mathcal{E}$ has genus 1 and $\mathbb{P}^1(\overline{k})$ has genus 0. By the Riemann-Hurwitz formula the genus is an algebraic invariant so such an isomorphism cannot exist. □

Proposition 3.33. *Let C be a smooth projective curve over $\overline{k}$ of genus $g \geq 1$. Then the Abel-Jacobi map is bijective.*

Proof. The surjectivity of the Abel-Jacobi map is the result known as Jacobi's inversion theorem. For injectivity, suppose for a contradiction that we have two distinct points P, Q on C such that $(P) - (P_0)$ is linearly equivalent to $(Q) - (P_0)$. Then $P - Q$ is a principal divisor, but this contradicts the previous lemma. □

Lemma 3.34. *Let $P, Q \in \mathcal{E}$. Then there is a congruence of divisors*

$$P + Q - 20_{\mathcal{E}} \cong P + Q - 0_{\mathcal{E}} \,(\operatorname{mod} \operatorname{Div}^P(\mathcal{E})).$$

Proof. Let $ax + by + c = 0$ be the equation of the line through P and Q and let $R = -P - Q$ be the third collinear point of the elliptic curve $\mathcal{E}$ and let $F(x, y) = ax + by + c$. If $R = 0_{\mathcal{E}}$ then on the left we have $(F) = P + Q + R - 30_{\mathcal{E}} = P + Q - 20_{\mathcal{E}} \in \operatorname{Div}^P(\mathcal{E})$, and on the right we have the empty divisor which also principal, so the congruence holds.

If $R \neq 0_{\mathcal{E}}$ then consider the line through R and $-R = P + Q$ which is defined by an affine equation of the form $G(x, y) = x + c$. The divisor of G is $(G) = R + (P + Q) - 20_{\mathcal{E}} \in \operatorname{Div}^P(\mathcal{E})$. Then $(F/G) = P + Q - 0_{\mathcal{E}} - 20_{\mathcal{E}} - ((P + Q) - 0_{\mathcal{E}}) \in \operatorname{Div}^P(\mathcal{E})$ gives the desired congruence. □

Theorem 3.35. *Let $\mathcal{E}$ be an elliptic curve. Then the map*

$$\begin{aligned} f : \operatorname{Div}(\mathcal{E}) &\to \mathcal{E} \\ \sum n_P(P) &\mapsto \sum [n_P]P \end{aligned}$$

induces an isomorphism

$$\operatorname{Pic}^0(\mathcal{E}) \xrightarrow{\sim} \mathcal{E}.$$

Proof. That the map $f : \operatorname{Div}(\mathcal{E}) \to \mathcal{E}$ is a homomorphism is clear from its definition. The restriction $f|_{\operatorname{Div}^0(\mathcal{E}}$ is surjective since for any point $P \in \mathcal{E}$ we have $f(P - 0_{\mathcal{E}}) = P$. The kernel of the restriction is precisely the principal divisors $\operatorname{Div}^P(\mathcal{E})$ because if $\sum n_P(P) \in \operatorname{Div}^0(\mathcal{E})$ then by Lemmas 3.32 and 3.34,

$$\begin{aligned} \sum [n_P]P = 0_{\mathcal{E}} &\iff \left(\sum [n_P]P\right) - 0_{\mathcal{E}} \text{ is principal} \\ &\iff \sum n_P(P - 0_{\mathcal{E}}) \text{ is principal.} \end{aligned}$$

Since $\sum n_P(P)$ is a 0-divisor, $\sum n_P = 0$, and so $\sum n_P(P - 0_{\mathcal{E}}) = \sum n_P(P)$. Thus $\sum [n_P]P = 0_{\mathcal{E}}$ if and only if $\sum n_P(P) \in \operatorname{Div}^P(\mathcal{E})$. □

Corollary 3.36. *If $\mathcal{E}$ is an elliptic curve over $\overline{k}$ then the Abel-Jacobi map is a group isomorphism.*

3.4 The Weil pairing

Definition 3.37. Let $\mathcal{E}$ be an elliptic curve over k and let N be coprime to $\operatorname{char}(k)$. The **Weil pairing** is a map

$$e_N : \mathcal{E}[N] \times \mathcal{E}[N] \to \mu_N$$

defined as follows.

Let μ_N denote the group of Nth roots of unity in $\overline{k}$. Let $P, Q \in \mathcal{E}[N]$. By Theorem 3.35 there exists $F_Q \in \overline{k}(\mathcal{E})$ such that

$$(F) = NQ - N0_{\mathcal{E}}.$$

Since the map $[N] : \mathcal{E} \to \mathcal{E}$ is unramified, the composition $F \circ [N]$ has divisor

$$(F \circ [N]) = \sum_{R:[N]R=Q} NR - \sum_{S:[N]S=0_{\mathcal{E}}} NS.$$

Let $Q' \in \mathcal{E}[N^2]$ be such that $[N]Q' = Q$. Then we can rewrite the divisor as

$$(F \circ [N]) = N \sum_{S \in \mathcal{E}[N]} (Q' + S) - S.$$

By Theorem 3.35 there exists $g \in \overline{k}(\mathcal{E})$ such that $(F \circ [N]) = N(g)$ and hence $f \circ [N] = (\alpha g)^N$ for some nonzero constant α. Let $g_Q = \alpha g$. Then for any point $T \in \mathcal{E}$,

$$g_Q(X + P)^N = f([N]X + N[P]) = f([N]X) = g_Q(X)^N.$$

Now consider the rational function $g_Q(X + P)/g_Q(X)$. Its Nth power is 1 so it must be constant with its image lying in μ_N. We define the image point to be the Weil pairing of the points P and Q,

$$e_N(P, Q) = \frac{g_Q(X + P)}{g_Q(X)}.$$

Proposition 3.38. *Let $\mathcal{E}$ be an elliptic curve over k and let M and N be positive integers coprime to* $\mathrm{char}(k)$. *The Weil pairing satisfies the following properties:*

i. Bilinear: $e_N(aP + bP', cQ + dQ') = e_N(P, Q)^{ac} e_N(P, Q')^{ad} e_N(P', Q)^{bc} e_N(P', Q')^{bd}$;

ii. Alternating: $e_N(P, P) = 1$ and $e_N(P, Q) = e_N(Q, P)^{-1}$;

iii. Non-degenerate: if $P \neq 0_{\mathcal{E}}$ then there exists $Q \in \mathcal{E}[N]$ such that $e_N(P, Q) \neq 1$;

iv. Compatibility: $e_{MN}(P, Q) = e_N(MP, Q)$ for all $P \in \mathcal{E}[MN]$ and $Q \in \mathcal{E}[N]$;

v. Galois-equivariant: $e_N(P, Q)^\sigma = e_N(P^\sigma, Q^\sigma)$ for all $\sigma \in \mathrm{Gal}(\overline{k}/k)$.

Proof. See Proposition 7.4.1. in [7]. □

Corollary 3.39. *Suppose $P, Q, P', Q' \in \mathcal{E}[N]$ are such that*

$$\begin{pmatrix} P' \\ Q' \end{pmatrix} = \gamma \begin{pmatrix} P \\ Q \end{pmatrix}$$

for some $\gamma \in \mathrm{M}_{2\times 2}(\mathbb{Z}/N\mathbb{Z})$. Then

$$e_N(P', Q') = e_N(P, Q)^{\det(\gamma)}.$$

Proof. See Corollary 7.4.2. in [7]. □

In particular, if (P, Q) is an ordered basis of $\mathcal{E}[N] \cong (\mathbb{Z}/N\mathbb{Z})^2$ over $\mathbb{Z}/N\mathbb{Z}$ then $e_N(P, Q)$ is a primitive Nth root of unity.

Remark. In the case of complex tori we have the following. Let $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ be a lattice with $\omega_1/\omega_2 \in \mathbb{H}$. Let P and Q be points in the N-torsion group $E[N]$. Then there exists $\gamma \in \mathrm{M}_{2\times 2}(\mathbb{Z}/N\mathbb{Z})$ such that

$$\begin{pmatrix} P \\ Q \end{pmatrix} = \gamma \begin{pmatrix} \omega_1/N + \Lambda \\ \omega_2/N + \Lambda \end{pmatrix}.$$

The Weil pairing $e_N : E[N] \times E[N] \to \mu_N$ of P and Q is given by

$$e_N(P, Q) = e^{2\pi i \det(\gamma)/N}.$$

This is well-defined since $\det(\gamma)$ is defined up to modulo N and it is independent of choice of basis $\{\omega_1, \omega_2\}$.

3.5 Reductions of elliptic curves over $\mathbb{Q}$

Given a Weierstrass equation with coefficients in k, we can clear denominators to obtain a Weierstrass equation with coefficients in the ring of integers $\mathcal{O}_k$. By then reducing the coefficients modulo a prime ideal, we obtain a Weierstrass equation with coefficients in a finite field $\mathbb{F}_q$.

This lead us to ask the following question: *Given an elliptic curve $\mathcal{E}$ over $\mathbb{Q}$, do we obtain an elliptic curve over $\mathbb{F}_p$ by reducing it modulo a prime p?* That is, under what conditions is the reduced Weierstrass equation nonsingular?

To understand the reduction of an elliptic curve, will first need to look at some properties of algebraic curves in characteristic p.

We denote the field of q elements by $\mathbb{F}_q$, and denote any algebraic closure $\overline{\mathbb{F}}_q$. Let p be prime and let e be a positive integer. The splitting field of the polynomial $f(x) = x^{p^e} - x$ over $\mathbb{F}_p$ gives the a unique subfield $\mathbb{F}_{p^e} \subset \overline{\mathbb{F}}_p$. Each such finite field extension $\mathbb{F}_{p^e}/\mathbb{F}_p$ is Galois, and its Galois group is cyclic of order e with generator σ_p, the Frobenius automorphism $x \mapsto x^p$.

Assume that $\mathrm{char}(k) = p > 0$ and let $q = p^e$. Consider the Frobenius endomorphism

$$\begin{aligned} \sigma_p : \overline{\mathbb{F}}_p &\to \overline{\mathbb{F}}_p \\ x &\mapsto x^p. \end{aligned}$$

The fixed points of σ_p are precisely $\mathbb{F}_p \subset \overline{\mathbb{F}}_p$, the set of roots of the polynomial $x^p - x$. More generally, the set $\mathbb{F}_p$ is the fixed field of the e-fold composition σ_p^e.

Now we would like to introduce a Frobenius map on a curve C. Firstly, for any positive integer n, the Frobenius map on $\overline{\mathbb{F}}_p^n$ is the bijective map

$$\begin{aligned} \sigma_p : \overline{\mathbb{F}}_p^n &\to \overline{\mathbb{F}}_p^n \\ (x_1, \ldots, x_n) &\mapsto (x_1^p, \ldots, x_n^p). \end{aligned}$$

This induced a well-defined bijection between the projective classes

$$\sigma_p : \mathbf{P}^n(\overline{\mathbb{F}}_p) \to \mathbf{P}^n(\overline{\mathbb{F}}_p)$$

$$[x_0 : x_1 : \cdots : x_n] \mapsto [x_0^p : x_1^p : \cdots : x_n^p].$$

Suppose $\varphi \in \overline{\mathbb{F}}_p[x_0, \ldots, x_n]$ is a homogeneous polynomial. Let us write the function in the form $\varphi(x) = \sum_d a_d x^d$ where $x = (x_0, \ldots, x_n)$, $d = (d_1, \ldots, d_n)$, and we write x^d as shorthand for $x_0^{d_0} \cdots x_n^{d_n}$. By applying the Frobenius map to its coefficients we obtain a map

$$\varphi^{\sigma_p}(x) := \sum_d a_d^{\sigma_p} x^d.$$

Since $(a+b)^p = a^p + b^p$ in characteristic p (consider the binomial expansion), we have

$$\varphi^{\sigma_p}(x^{\sigma_p}) = \varphi(x)^{\sigma_p}.$$

Thus, if $P \in \mathbf{P}^n(\overline{\mathbb{F}}_p)$ is a zero of φ, then $\varphi^{\sigma_p}(P^{\sigma_p}) = 0$ also. Let C be a projective curve defined over $\overline{\mathbb{F}}_p$ by polynomials $\varphi_1, \ldots, \varphi_m$ and let C^{σ_p} be the corresponding curve defined by the polynomials $\varphi_1^{\sigma_p}, \ldots, \varphi_m^{\sigma_p}$. Then σ_p restricted to C gives a well-defined morphism $C \to C^{\sigma_p}$.

Definition 3.40. Let C be a projective curve over $\overline{\mathbb{F}}_p$. The **Frobenius map on C** is

$$\begin{aligned} \sigma_p : C &\to C^{\sigma_p} \\ [x_0 : x_1 : \cdots : x_n] &\mapsto [x_0^p : x_1^p : \cdots : x_n^p]. \end{aligned}$$

Example. The Frobenius map on an elliptic curve over $\mathbb{F}_p$ is given by $\sigma_p(x, y) = (x^p, y^p)$ on the affine part. The pullback of function fields

$$\begin{aligned} \sigma_p^* : k(E^{\sigma_p}) &\to k(E) \\ G &\mapsto G \circ \sigma_p \end{aligned}$$

gives an extension of function fields $k(E^{\sigma_p})/k(E)$

$$\begin{aligned} k(E) &= \mathbb{F}_p(x)[y]/\langle E(x, y) \rangle \\ k(E^{\sigma_p}) &= \mathbb{F}_| \end{aligned}$$

Consider a Weierstrass equation E defined over $\mathbb{Q}$,

$$E(x, y) = y^2 + a_1 xy + a_3 y - x^3 - a_2 x^2 - a_4 x - a_6 \in \mathbb{Q}[x, y].$$

By an admissible change of variable of the form $(x, y) = (u^2 x', u^3 y')$ with $u \in \mathbb{Q}$ we obtain a Weierstrass equation $E'(x', y')$ with coefficients $a_i' = a_i/u^i$. With a suitable choice of u the coefficients a_i' can be made to be integral over $\mathbb{Q}$, so let us assume the original coefficients are integral. We view any two integral Weierstrass equations as *equivalent* if they are related by an admissible change of variable over $\mathbb{Q}$.

Recall the p-adic valuation v_p from section 2.2. For any prime p and integral Weierstrass equation E, define v_p to be the smallest value of the p-adic valuation of the discriminant $\Delta(E')$ over all integral Weierstrass equations E' equivalent to E,

$$v_p(E) = \min\{v_p(\Delta(E')) : E' \text{ integral over } \mathbb{Q} \text{ , } E' \text{ equivalent to } E\}.$$

Definition 3.41. The **global minimal discriminant of** E is

$$\Delta_{\min}(E) = \prod_{p \text{ prime}} p^{v_p(E)}.$$

The product is finite since $v_p(E) = 0$ for all $p \nmid \Delta(E)$.

Definition 3.42. Let E be a Weierstrass equation with integer coefficients. If $\Delta(E) = \Delta_{\min}(E)$ then E is called a **global minimal Weierstrass equation**.

Proposition 3.43. *Let E be a Weierstrass equation with integer coefficients. Then E is equivalent to a global minimal Weierstrass equation E' such that $\Delta(E') = \Delta_{\min}(E)$.*

Moreover, the global minimal Weierstrass equation is unique up to change of coordinates

$$(x', y') = (u^2x'' + r, u^3y'' + u^2sx' + t), \quad u \in \mathbb{Z}^\times,\ r, s, t \in \mathbb{Z}.$$

Proof. See Section VII Proposition 1.3 in [24]. □

Having chosen a global minimal Weierstrass equation E, we can reduce its coefficients modulo p. Viewing $\mathbb{F}_p$ as $\mathbb{Z}/p\mathbb{Z}$, the reduction map

$$\begin{aligned} \tilde{}: \mathbb{Z} &\to \\ n &\mapsto \tilde{n} = n = p\mathbb{Z} \end{aligned}$$

reduces a global minimal Weierstrass equation

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \quad a_1, a_2, a_3, a_4, a_6 \in \mathbb{Z}$$

to a Weierstrass equation $\tilde{E}$ over $\mathbb{F}_p$

$$\tilde{E} : y^2 + \tilde{a_1}xy + \tilde{a_3}y = x^3 + \tilde{a_2}x^2 + \tilde{a_4}x + \tilde{a_6}, \quad a_1, a_2, a_3, a_4, a_6 \in \mathbb{Z}/p\mathbb{Z}.$$

We call $\tilde{E}$ the **reduction of** E **modulo** p. Since we started with any global minimal Weierstrass equation for E, the reduction $\tilde{E}$ is unique up to admissible change of variables in $\mathbb{F}_p$.

$\tilde{E}$ defines an elliptic curve over $\mathbb{F}_p$ if and only if $p \nmid \Delta_{\min}(E)$.

Definition 3.44. Let $\tilde{E}$ be a reduction of E modulo p. We say that the reduction is **good** if $\tilde{E}$ again an elliptic curve. In this case, we also say that the reduction is **ordinary** if $\tilde{E}[p] \cong \mathbb{Z}/p\mathbb{Z}$ and **supersingular** if $\tilde{E}[p] = \{0\}$.

We say that the reduction is **bad** if $\tilde{E}$ is not an elliptic curve. In this case $\tilde{E}$ has one singular point (as we saw in section 1.4). We additionally say that the reduction is **semistable** (or multiplicative) if $\tilde{E}$ has a node, and **unstable** (or additive) if $\tilde{E}$ has a cusp.

Since an admissible change of variables between two global minimal Weierstrass equations preserves Δ and c_4, the reduction type is independent of choice of global minimal Weierstrass equations. When $\tilde{E}$ has good reduction at p, the two Weierstrass equations differ by a reduced admissible change of variable so they are isomorphic as elliptic curves and have identical p-torsion, preserving the ordinary and supersingular properties.

If the reduction is bad, then $v_p(\Delta_{\min}(E)) >)$, in which case $v_p(c_4) \geq 0$. If $v_p(c_4) = 0$ the reduction is multiplicative, meaning that the nonsingular points of $\tilde{E}$ form a multiplicative group isomorphic to $\overline{\mathbb{F}}_p^\times$. On the other hand, if $v_p(c_4) > 0$ then the reduction is additive and the nonsingular points form an additive group isomorphic to $\overline{\mathbb{F}}_p$.

4 Hecke operators, newforms, and eigenforms

We now turn to the problem of constructing a canonical basis for the space of cusp forms $\mathcal{S}_k(\Gamma_1(N))$. For this purpose, we introduce linear operators on the spaces of modular forms $\mathcal{M}_k(\Gamma)$ and cusp forms $\mathcal{S}_k(\Gamma)$ to modular forms and cusp forms of the same weight.

A Hecke operator is a certain kind of averaging operator similar to a trace. For details on the subject, see Chapter 5 of [7] and for an alternative formulation in terms of lattices, see Part I Chapter 2 of [13].

4.1 Hecke operators

4.1.1 The weight-k double coset operator

Definition 4.1. Let Γ_1 and Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. The set

$$\Gamma_1\alpha\Gamma_2 = \{\gamma_1\alpha\gamma_2 \ : \ \gamma_1 \in \Gamma_1, \ \gamma_2 \in \Gamma_2\}$$

is a **double coset** in $\mathrm{GL}_2^+(\mathbb{Q})$.

The action of Γ_1 on $\Gamma_1\alpha\Gamma_2$ by left multiplication partitions the double coset into orbits of the form $\Gamma_1\beta$ with representative $\beta = \gamma_1\alpha\gamma_2$.

Lemma 4.2. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. Then $\alpha^{-1}\Gamma\alpha \cap \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.*

Lemma 4.3. *Set $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2 \subset \Gamma_2$. Then left multiplication by α induces a natural bijection from $\Gamma_3\backslash\Gamma_2$ to $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$. In particular, $\{\gamma_{2,j}\}$ is a set of orbit representatives for $\Gamma_3\backslash\Gamma_2$ if and only if $\{\beta_j\} = \{\alpha\gamma_{2,j}\}$ is a set of orbit representatives for $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$.*

Corollary 4.4. *The orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$ is finite.*

Definition 4.5. The **weight-k double coset operator** is

$$\begin{aligned}[\Gamma_1\alpha\Gamma_2]_k : \mathcal{M}_k(\Gamma_1) &\to \mathcal{M}_k(\Gamma_2)\\ f[\Gamma_1\alpha\Gamma_2]_k &:= \sum_j f[\beta_j]_k\end{aligned}$$

where $\{\beta_j\}$ are orbit representatives.

We check that this operator is well-defined. To see that the double coset operator is independent of choice of representatives β_j, suppose $\{\tau_j\}$ is another choice of orbit representatives. Then for each j, there exists $\gamma_j \in \Gamma_1$ such that $\beta_j = \gamma_j\tau_j$. Hence, $f[\beta_j]_k = f[\gamma_j\tau_j]_k = f[\tau_j]_k$.

To verify that the image of the map lies in $\mathcal{M}_k(\Gamma_2)$, we need to check that $f[\Gamma_1\alpha\Gamma_2]_k$ is Γ_2-invariant and holomorphic at the cusps. Let $\gamma_2 \in \Gamma_2$ and let $\{\beta_j\}$ be a set of orbit representatives. The map given by right multiplication by γ_2, i.e.$\Gamma_1\beta \mapsto \Gamma_1\beta\gamma_2$, permutes the orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$. Thus, the map is bijective onto the orbit space. Hence, $\{\beta_j\gamma_2\}$ is also a set of orbit representatives and we have

$$(f[\Gamma_1\alpha\Gamma_2]_k)[\gamma_2]_k = \sum_j f[\beta_j\gamma_2]_k = f[\Gamma_1\alpha\Gamma_2]_k.$$

Finally, we show that $f[\Gamma_1\alpha\Gamma_2]_k$ is holomorphic at the cusps. For any $\delta \in \mathrm{SL}_2(\mathbb{Z})$, the function $(f[\Gamma_1\alpha\Gamma_2]_k)[\delta]_k$ is a finite sum of functions of the form $g_j = f[\gamma_j]_k$ where $\gamma_j = \beta_j\delta \in \mathrm{GL}_2^+(\mathbb{Q})$. For all j, $f[\gamma_j]_k$ is holomorphic at infinity, so it has a Fourier expansion. Hence, their sum also has a Fourier expansion, so $f[\Gamma_1\alpha\Gamma_2]_k$ is holomorphic at the cusps.

Consider the following special cases.

(i) $\Gamma_1 \supset \Gamma_2$ and $\alpha = I$. In this case, the double coset operator $f[\Gamma_1\alpha\Gamma_2]_k = f$ is the natural inclusion $\mathcal{M}_k(\Gamma_1) \overset{i}{\hookrightarrow} \mathcal{M}_k(\Gamma_2)$.

(ii) $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$. The double coset operator $f[\Gamma_1\alpha\Gamma_2]_k = f[\alpha]_k$ is the natural translation $\mathcal{M}_k(\Gamma_1) \xrightarrow{\sim} \mathcal{M}_k(\Gamma_2)$.

(iii) $\Gamma_1 \subset \Gamma_2$ and $\alpha = I$. Let $\{\gamma_{2,j}\}$ be the set of coset representatives of $\Gamma_1\backslash\Gamma_2$. The double coset operator is $f[\Gamma_1\alpha\Gamma_2]_k = \sum_j f[\gamma_{2,j}]_k$. This is the surjective trace map projecting $\mathcal{M}_k(\Gamma_1) \overset{s}{\twoheadrightarrow} \mathcal{M}_k(\Gamma_2)$ by symmetrising over the quotient.

Remark. The double coset operator has a geometric interpretation in terms of transferring points between corresponding modular curves, giving rise to an algebraic interpretation of the double coset as a homomorphism of divisor groups. Given Γ_1, Γ_2, and α, set $\Gamma_3 = \alpha^{-1}\Gamma_2\alpha \cap \Gamma_2$ as in the setup of Lemma 4.3. Set $\Gamma_3' = \alpha\Gamma_3\alpha^{-1} = \Gamma_1 \cap \alpha\Gamma_2\alpha^{-1}$. Then $\Gamma_1 \supset \Gamma_3')$ as in case (i), $\alpha^{-1}\Gamma_3'\alpha = \gamma_3$ as in case (ii), and $\Gamma_3 \subset \Gamma_2$ as in case (*iii*). Thus every double coset operator can be written as a composition

$$\mathcal{M}_k(\Gamma_1) \overset{i}{\hookrightarrow} \mathcal{M}_k(\Gamma_3') \xrightarrow{\sim} \mathcal{M}_k(\Gamma_3) \overset{s}{\twoheadrightarrow} \mathcal{M}_k(\Gamma_2)$$

$$f \mapsto f \mapsto f[\alpha]_k \mapsto \sum_j f[\alpha\gamma_{2,j}]_k.$$

Recall from section 1.5 that associated to every congruence subgroup Γ is a modular curve $X(\Gamma)$ consisting of orbits $\Gamma\tau$. Then we have a corresponding configuration for modular curves. Diagrammatically,

$$\begin{tikzcd} \Gamma_3 \arrow[r, "\sim"] \arrow[d, hook] & \Gamma_3' \arrow[d, hook] \\ \Gamma_2 & \Gamma_1 \end{tikzcd} \qquad \begin{tikzcd} X_3 \arrow[r, "\sim", "\alpha"'] \arrow[d, "\pi_2"'] & X_3' \arrow[d, "\pi_1"] \\ X_2 & X_1 \end{tikzcd}$$

The isomorphism on the left is $\gamma \to \alpha\gamma\alpha^{-1}$ and the isomorphism of modular curves $\alpha : X_3 \to X_3'$ is $\Gamma_3\tau \mapsto \Gamma_3'\alpha(\tau)$. This is well-defined since for any element $\gamma \in \Gamma_3$ we have $\gamma' = \alpha\gamma\alpha^{-1} \in \Gamma_3'$ satisfying $\gamma'\alpha = \alpha\gamma\alpha^1\alpha = \alpha\gamma$, so $\alpha(\gamma(\tau))$ and $\alpha(\tau)$ must lie in the same equivalence class in X_3'. The maps π_1 and π_2 are finite index covers.

Let $\{\gamma_{2,j}\}$ be a set of orbit representatives of $\Gamma_3\backslash\Gamma_2$ and $\beta_j = \alpha\gamma_{2,j}$ for each j. Each point $x = \Gamma_2\tau$ of X_2 is taken back via $\pi_1 \circ \alpha \circ \pi_2^{-1}$ to a set of points in X_1. To be precise, we have the following picture.

$$\begin{tikzcd}
\{\Gamma_3\gamma_{2,j}(\tau)\} \arrow[r, mapsto, "\alpha"'] & \{\Gamma_3'\beta_j(\tau)\} \arrow[d, mapsto, "\pi_1"] \\
\Gamma_2\tau \arrow[u, mapsto, "\pi_2^{-1}"] & \{\Gamma_1\beta_j(\tau)\}
\end{tikzcd}$$

To count with multiplicity, we actually want to view π_2^{-1} as taking each point in $x \in X_2$ to the *multiset* of overlying points $y \in X_3$ each with multiplicity according to its degree of ramification e_y. We represent this multiset as the divisor

$$\sum_{y \in \pi_2^{-1}(x)} e_y y \in \mathrm{Div}(X_3).$$

Any map $\varphi : X \to X'$ between two spaces induces a $\mathbb{Z}$-linear map on divisors $\mathrm{Div}(X) \to \mathrm{Div}(X')$ via $\sum_{x \in X} n_x x \mapsto \sum_{x \in X} n_x \varphi(x)$. Thus, in terms of divisors, the weight-k double coset operator $[\Gamma_1\alpha\Gamma_2]_k$ is the composition $(\pi_1)_* \circ \alpha_* \circ (\pi_2)^*$ and we have

$$\begin{aligned} [\Gamma_1\alpha\Gamma_2]_k : X_2 &\to \mathrm{Div}(X_1) \\ \Gamma_2\tau &\mapsto \sum_j \Gamma_1\beta_j(\tau) \end{aligned}$$

which uniquely extends $\mathbb{Z}$-linearly to a divisor group homomorphism

$$[\Gamma_1\alpha\Gamma_2]_k : \mathrm{Div}(X_2) \to \mathrm{Div}(X_1).$$

This construction will define Hecke operators at the level of divisors.

4.1.2 The $\langle n\rangle$ and T_n operators

The first type of Hecke operator we will introduce is the *diamond operator* $\langle n\rangle$. At level N, we will first define $\langle d\rangle$ for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$ and then extend the definition to any positive integer n.

Let $\alpha \in \Gamma_0(N)$ and let $\Gamma_1 = \Gamma_1(N) = \Gamma_2$. We are in the special case of (ii) from the previous section so the double coset operator $[\Gamma_1(N)\alpha\Gamma_1(N)]_k$ translates each function $f \in \mathcal{M}_k(\Gamma_1(N))$ by

$$f[\Gamma_1(N)\alpha\Gamma_1(N)]_k = f[\alpha]_k.$$

This defines a group action of $\Gamma_0(N)$ on $\mathcal{M}_k(\Gamma_1(N))$ with the subgroup $\Gamma_1(N) \subset \Gamma_0(N)$ acting trivially, so we should really think of this as the action of the quotient $\Gamma_0(N)/\Gamma_1(N)$. Recall from section 1.2 that we identified $(\mathbb{Z}/N\mathbb{Z})^\times \cong \Gamma_0(N)/\Gamma_1(N)$ as the kernel of the map

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \,(\mathrm{mod}\ N)$$

so the action of $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$ on $\mathcal{M}_k(\Gamma_1(N))$ is determined by d (mod N).

Definition 4.6. (Diamond operator) The diamond operator for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$ is

$$\begin{aligned} \langle d \rangle : \mathcal{M}_k(\Gamma_1(N)) &\to \mathcal{M}_k(\Gamma_1(N)) \\ \langle d \rangle f &= f[\alpha]_k \end{aligned}$$

for any $\alpha = \left(\begin{smallmatrix} a & b \\ c & d' \end{smallmatrix}\right) \in \Gamma_0(N)$ with $d' \equiv d \pmod{N}$.

For $n \in \mathbb{Z}^+$ such that $\gcd(n, N) = 1$, $\langle n \rangle$ is determined by $n \pmod{N}$. Otherwise if $\gcd(n, N) > 1$, define $\langle n \rangle = 0$.

Remark. If χ is any Dirichlet character modulo N, the diamond operator $\langle d \rangle$ acts on $\mathcal{M}_k(N, \chi)$ by multiplication by $\chi(d)$. Thus the subspaces $\mathcal{M}_k(N, \chi) \subset \mathcal{M}_k(\Gamma_1(N))$ are precisely the χ-eigenspaces of the diamond operators and the diamond operator $\langle d \rangle$ repects the decomposition

$$\mathcal{M}_k(\Gamma_1(N)) = \bigoplus_{\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}} \mathcal{M}_k(N, \chi)$$

from Proposition 1.22.

The second type of Hecke operator we will introduce is also a weight-k double coset operator.

Definition 4.7. (T_p operator) Let $p \in \mathbb{Z}^+$ be prime and let $\alpha_p = \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)$. As before, let $\Gamma_1 = \Gamma_1(N) = \Gamma_2$. The T_p operator is

$$\begin{aligned} T_p : \mathcal{M}_k(\Gamma_1(N)) &\to \mathcal{M}_k(\Gamma_1(N)) \\ T_p f &= f[\Gamma_1(N)\alpha_p\Gamma_1(N)]_k. \end{aligned}$$

For prime powers, define T_{p^r} inductively by

$$T_{p^r} = T_p T_{p^{r-1}} - p^{k-1}\langle p \rangle T_{p^{r-2}}, \quad r \geq 2.$$

Finally extend the definition multiplicative for all $n \in \mathbb{Z}^+$ by

$$T_n = \prod T_{p_i^{e_i}}, \quad \text{where } n = \prod p_i^{e_i}.$$

For p prime, the double coset is

$$\Gamma_1(N) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma_1(N) = \left\{ \gamma \in \mathrm{M}_2(\mathbb{Z}) \,:\, \gamma \equiv \begin{pmatrix} 1 & \star \\ 0 & p \end{pmatrix} \pmod{N},\ \det(\gamma) = p \right\}$$

For details, see the discussion following Proposition 3.8.5. of Diamond and Shurman [7]. We may replace the matrix $\left(\begin{smallmatrix} 1 & \star \\ 0 & p \end{smallmatrix}\right)$ by any matrix in this double coset.

To see that the Hecke operators commute, the following proposition suffices.

Proposition 4.8. *Let $d, e \in (\mathbb{Z}/N\mathbb{Z})^\times$ and let p and q be prime. Then*

(i) $\langle d \rangle T_p = T_p \langle d \rangle$,

(ii) $\langle d \rangle \langle e \rangle = \langle e \rangle \langle d \rangle$,

(iii) $T_pT_q = T_{pq}$ *if* $\gcd(p,q) > 1$.

Proof. See Proposition 5.2.3. in [7]. □

Proposition 4.9. *Let p be prime. The operator T_p on $\mathcal{M}_k(\Gamma_1(N))$ is given by*

$$T_p(f) = \begin{cases} \sum_{j=0}^{p-1} f\left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right]_k, & p|N \\ \sum_{j=0}^{p-1} f\left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right]_k + f\left[\begin{pmatrix} m & n \\ N & p \end{pmatrix}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\right]_k, & p\nmid N,\ mp - nN = 1. \end{cases}$$

Proof. Let $\alpha_p = \left(\begin{smallmatrix} 1 & 0 \\ 0 & p \end{smallmatrix}\right)$. The set $\Gamma_3 = \alpha_p^{-1}\Gamma_1(N)\alpha_p \cap \Gamma_1(N)$ acts on $\Gamma_1(N)$ by left multiplication by α_p so the orbits of $\Gamma_1(N)\backslash\Gamma_1(N)\alpha\Gamma_1(N)$ are the cosets of $\Gamma_3\backslash\Gamma_1(N)$. If $\{\beta_j\}$ is a set of orbit representatives of $\Gamma_3\backslash\Gamma_1(N)$ then $\{\alpha_p\beta_j\}$ is a set of orbit representatives of $\Gamma_1(N)\backslash\Gamma_1(N)\alpha_p\Gamma_1(N)$.

Let

$$\Gamma^0(p) = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} \star & 0 \\ \star & \star \end{pmatrix} \pmod{p}\right\}.$$

We claim that

$$\Gamma_3 = \Gamma_1^0(N,p) := \Gamma^0(p) \cap \Gamma_1(N).$$

Note $\alpha_p^{-1} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & p^{-1} \end{smallmatrix}\right)$ and let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(N)$. Every element in Γ_3 is of the form

$$\alpha_p^{-1}\gamma\alpha_p = \begin{pmatrix} a & pb \\ p^{-1}c & d \end{pmatrix} \equiv \begin{pmatrix} \star & 0 \\ \star & \star \end{pmatrix} \pmod{p}$$

so $\Gamma_3 \subset \Gamma^0(p)$. Since $\Gamma_3 \subset \Gamma_1(N)$, we have $\Gamma_3 \subset \Gamma_1(N) \cap \Gamma^0(p) = \Gamma_1^0(N,p)$.

For the reverse inclusion, let $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1^0(N,p)$. Then p divides b, so let us write $b = pb'$ where $b' \in \mathbb{Z}$. Let $c' = pc$. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & p^{-1} \end{pmatrix}\begin{pmatrix} a & b' \\ c' & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} = \alpha_p^{-1}\begin{pmatrix} a & b' \\ c' & d \end{pmatrix}\alpha_p.$$

Since $\Gamma_1^0(p) \subset \Gamma_1(N)$, $a, d \equiv 1 \pmod{N}$ and $c \equiv 0 \pmod{N}$. Thus $c' = pc \equiv 0 \pmod{N}$ and $\left(\begin{smallmatrix} a & b' \\ c' & d \end{smallmatrix}\right) \in \Gamma_1(N)$. Thus, $\Gamma_3 = \Gamma_1^0(N,p)$. In fact, Γ_3 is $\Gamma_1(N)$ subject to the additional requirement that $b \equiv 0 \pmod{p}$, so we have obvious candidates, namely

$$\gamma_{2,j} = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}, \quad 0 \le j < p,$$

as coset representatives for $\Gamma_3\backslash\Gamma_1(N)$.

If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(N)$ then $\gamma \in \Gamma_3\gamma_{2,j}$ if $\gamma\gamma_{2,j}^{-1} \in \Gamma_3$ for some j. We have $\gamma\gamma_{2,j}^{-1} \in \Gamma_1(N)$ for any j since $\gamma, \gamma_{2,j}^{-1} \in \Gamma_1(N)$. Since

$$\gamma\gamma_{2,j}^{-1} = \begin{pmatrix} a & b - aj \\ c & d - cj \end{pmatrix},$$

we also require $b - ja \equiv 0 \pmod{p}$.

We split this into two cases. Firstly, if $p \nmid a$ then setting $j = ba^{-1} \pmod{p}$ does the job. Otherwise if $p|a$ then $b - ja$ cannot be congruent to 0 modulo p for any j since we would have $p|b$ and so $p|ad - bc = 1$, a contradiction since p is prime. The case $\gamma \in \Gamma_1(N)$ with $p|a$ occurs if and only if $p \nmid N$ – in this case, $\gamma_{2,0}, \dots, \gamma_{2,p-1}$ fail to be a set of orbit representatives of $\Gamma_3\backslash\Gamma_2$.

Instead, define

$$\Gamma_{2,\infty} = \begin{pmatrix} mp & n \\ N & 1 \end{pmatrix}$$

where $mp - nN = 1$. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $p|a$, we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \gamma_{2,\infty}^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -n \\ -N & mp \end{pmatrix} = \begin{pmatrix} a - Nb & bmp - na \\ c - dN & dmp - nc \end{pmatrix}$$

with $dmp - na \equiv 0 \pmod{p}$ since $p|a$. Therefore $\gamma\gamma_{2,\infty}^{-1} \in \Gamma_3$.

We verify that $\gamma_{2,\infty}$ and $\gamma_{2,0}, \dots, \gamma_{2,p-1}$ represent distinct cosets. Let $0 \le j, j' < p$ with $j \neq j'$. Since

$$\gamma_{2,j}\gamma_{2,j'}^{-1} = \begin{pmatrix} 1 & j - j' \\ 0 & 1 \end{pmatrix},$$

$\gamma_{2,j}\gamma_{2,j'}^{-1} \in \Gamma_3$ if and only if $j - j' \equiv 0 \pmod{p}$, but this is impossible. Thus $\gamma_{2,j}$ and $\gamma_{2,j'}$ represent distinct cosets. It remains to show that the coset represented by $\gamma_{2,\infty}$ is distinct from the $\gamma_{2,j}$. We similarly compute

$$\gamma_{2,j}\gamma_{2,\infty}^{-1} = \begin{pmatrix} 1 - Nj & jmp - n \\ -N & mp \end{pmatrix}.$$

Since $mp - nN = 1$, we have $p \nmid nN$, and so $p \nmid n$ as p is prime. Therefore $jmp - n \not\equiv 0 \pmod{p}$, and so $\gamma_{2,j}\gamma_{2,\infty}^{-1} \notin \Gamma_3$ and $\gamma_{2,\infty}$ and $\gamma_{2,0}, \dots, \gamma_{2,p-1}$ represent distinct cosets.

Thus $\beta_j = \alpha_p\gamma_{2,j}$ and $\beta_\infty = \alpha_p\gamma_{2,\infty}$ give distinct orbit representatives for $\Gamma_1(N)\backslash\Gamma_1(N)\alpha_p\Gamma_1(N)$. □

It is often more convenient to work with the Fourier expansion of a modular form, so we will also want to understand Hecke operators in this context.

Proposition 4.10. *Let $f \in \mathcal{M}_k(\Gamma_1(N))$ have Fourier expansion*

$$f(\tau) = \sum_{m=0}^{\infty} a_m(f)q^m, \quad q = e^{2\pi i\tau},$$

then for all $n \in \mathbb{Z}^+$, the Fourier expansion of $T_n f$ is

$$(T_n f)(\tau) = \sum_{m=0}^{\infty} \left(\sum_{d \,|\, \gcd(m,n)} d^{k-1} a_{mn/d^2}(\langle d \rangle f) \right) q^m.$$

In particular, if $f \in \mathcal{M}_k(N, \chi)$, then

$$a_m(T_n f) = \sum_{d|\gcd(m,n)} \chi(d) d^{k-1} a_{mn/d^2}(f).$$

Proof. See Proposition 5.3.1. of [7]. □

4.1.3 The Petterson inner product

Let us introduce an inner product on the space of cusp forms $\mathcal{S}_k(\Gamma_1(N))$, namely the Petersson inner product which will be defined as an integral. This will prove to be useful in studying this space. However, the integral does not converge on the larger space $\mathcal{M}_k(\Gamma_1(N))$, so this structure is restricted to the cusp forms.

Let

$$d\mu(\tau) = \frac{dx\ dy}{y^2}, \quad \tau = x + iy \in \mathbb{H}$$

be the hyperbolic measure on the upper half-plane. $d\mu$ is $\mathrm{SL}_2(\mathbb{Z})$-invariant, meaning that $d\mu(\gamma(\tau)) = d\mu(\tau)$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Since $\mathbb{Q} \cup \{\infty\}$ is countable, it has measure zero, so we can integrate over the extended upper half-plane using $d\mu$.

Lemma 4.11. *The hyperbolic measure is invariant under the group action of* $\mathrm{SL}_2(\mathbb{R})$*, and hence also invariant under* $\mathrm{SL}_2(\mathbb{Z})$*.*

Proof. Let us write $dz = dx + idy$ and $\overline{dz} = dx - idy$. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$ and write $\tau' = x' + iy' = \gamma(\tau)$. Since $ad - bc = 1$, we find that

$$x' = \frac{ac(x^2 + y^2) + x(ad + bc) + bd}{|c\tau + d|^2}$$

and

$$y' = \frac{y}{|c\tau + d|^2}.$$

Therefore

$$\begin{aligned} dx' + idy' = \frac{a(c\tau + d) - c(a\tau + b)}{(c\tau + d)^2}(dx + idy) &= \frac{ac\tau + ad - ca\tau - cb}{(c\tau + d)^2}(dx + idy) \\ &= \frac{1}{(c\tau + d)^2}(dx + idy) \end{aligned}$$

and so

$$d\mu(\gamma(\tau)) = \frac{d\tau'\overline{d\tau'}}{y'^2} = \frac{\frac{d\tau\overline{d\tau}}{|c\tau+d|^4}}{\frac{y^2}{|c\tau+d|^4}} = \frac{d\tau\overline{d\tau}}{y^2} = d\mu(\tau).$$

□

Rather than integrating over the entire extended upper half plane, for a congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$, we instead want to define an integral over the modular curve $X(\Gamma)$. It will first be useful to describe a fundamental domain of $\mathbb{H}^*$ under the action of $\mathrm{SL}_2(\mathbb{Z})$.

Lemma 4.12. *A fundamental domain of* $\mathbb{H}^*$ *under the action of* $\mathrm{SL}_2(\mathbb{Z})$ *is given by*

$$\mathcal{D}^* = \{\tau \in \mathbb{H} : |\mathrm{Re}(\tau)| \leq 1/2, |\tau| \geq 1\} \cup \{\infty\}$$

with the appropriate edges identified.

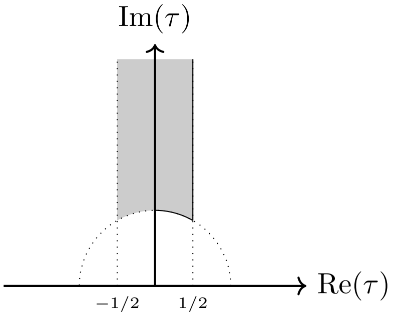


Figure 1: $\mathcal{D}$, a fundamental domain of $\mathbb{H}$ under the action of the modular group.

Proof. Let $\tau = x + iy \in \mathbb{H}$. By repeatedly applying either $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ or $\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)$ we can translate τ into the region $\{\tau' \in \mathbb{H} : |\text{Re}(\tau)| \leq 1/2\}$. Replace τ by this transform. If $\tau \notin \mathcal{D} = \{\tau \in \mathbb{H} : |\text{Re}(\tau)| \leq 1/2, |\tau| \geq 1\}$, then $|\tau| < 1$, so $\text{Im}\left(\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)(\tau)\right) = \text{Im}(-1/\tau) > \text{Im}(-\bar{\tau}/|\tau|^2) = \text{Im}(\tau/|\tau|^2) > \text{Im}(\tau)$. Since there are only finitely many pairs of integers (c, d) such that $|c\tau + d| < 1$, repeatedly replacing τ with $-1/\tau$ results in a point $\tau \in \mathcal{D}$. This process terminates with some $\tau \in \mathcal{D}$ since $\text{Im}(\gamma\tau) = \text{Im}(\tau)/|c\tau + d|$ for any $\gamma = \left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right) \in \text{SL}_2(\mathbb{Z})$.

Now suppose τ, τ' are distinct points in $\mathcal{D}$ with $\gamma\tau = \tau'$ for some $\gamma \in \text{SL}_2(\mathbb{Z})$. We claim that τ and τ' must lie on the boundary of $\mathcal{D}$. By symmetry, we may assume that $\text{Im}(\tau') > \text{Im}(\tau)$. Let $\gamma = \left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$. Then $|c\tau+d|^2 \leq 1$ and $\text{Im}(\tau) \geq \sqrt{3}/2$ since $\tau \in \mathcal{D}$. Therefore $|c|\sqrt{3}/2 \leq |\text{Im}(c\tau+d)| \leq |c\tau+d| \leq 1$. Since $c \in \mathbb{Z}$, either $c = 0$ or $c = \pm 1$.

If $c = 0$ then $\gamma = \pm\left(\begin{smallmatrix}1&b\\0&1\end{smallmatrix}\right)$ so $\text{Re}(\tau') = \text{Re}(\tau) + b$. Since $|\text{Re}(\tau)|, |\text{Re}(\tau')| \leq 1/2$, we must have $|b| = 1$, and so $\text{Re}(\tau) = \pm 1/2$ and $\tau' = \tau \mp 1$. We identify $1/2 + yi \sim -1/2 + yi$ for all $y \geq \sqrt{3}/2$.

Otherwise if $c = \pm 1$, then $(\text{Re}(\tau) \pm d)^2 + (\text{Im}(\tau))^2 = |\tau \pm d|^2 \leq 1$. Therefore $(\text{Re}(\tau) \pm d)^2 \leq 1 - (\text{Im}(\tau))^2 \leq 1 - (\sqrt{3}/2)^2 = 1/4$, so $|\text{Re}(\tau) \pm d| \leq 1/2$. Since $|\text{Re}(\tau)| \leq 1/2$, we must have $|d| \leq 1$. Since $d \in \mathbb{Z}$, $d = 0$ or $|d| = 1$. If $|d| = 1$ then $\text{Im}(\tau) = \sqrt{3}/2$ and $\text{Re}(\tau) = \pm 1/2$ since $|\text{Re}(\tau) \pm 1| = 1/2$. On the other hand if $d = 0$ then $|\tau| \leq 1$. Since $\tau \in \mathcal{D}$, $|\tau| = 1$ and $\text{Im}(\tau') = \text{Im}(\tau)$. By symmetry, we also have $|\tau'| = 1$, so $\text{Re}(\tau') = -\text{Re}(\tau)$. We identify $\tau = x + iy \sim -x + iy$ if $|\tau| = 1$.

Finally, every point $x \in \mathbb{Q} \cup \{\infty\}$ transforms under $\text{SL}_2(\mathbb{Z})$ to ∞. To see this, write $x = p/q$ where p and q are relatively prime integers. By Bézout's identity, there exist $a, b \in \mathbb{Z}$ such that $ap + bq = 1$. Letting $\gamma = \left(\begin{smallmatrix}a&b\\-q&p\end{smallmatrix}\right)$, we have $\gamma \in \text{SL}_2(\mathbb{Z})$ and the action of γ on $x = p/q$ is $\gamma(p/q) = \frac{a(p/q)+b}{-q(p/q)+p} = \infty$. □

Let $\Gamma \subset \text{SL}_2(\mathbb{Z})$ be a congruence subgroup and let $\{\alpha_j\} \subset \text{S}_2(\mathbb{Z})$ represent the coset space $\{\pm I\}\Gamma\backslash\text{SL}_2(\mathbb{Z})$. Given any continuous, bounded function $\varphi : \mathbb{H} \to \mathbb{C}$ and any $\alpha \in \text{SL}_2(\mathbb{Z})$, the integral $\int_{\mathcal{D}^*} \varphi(\alpha(\tau))d\mu(\tau)$ converges. If φ is also Γ-invariant, then the sum

$$\sum_j \int_{\mathcal{D}^*} \varphi(\alpha_j(\tau))d\mu(\tau)$$

is independent of choice of coset representatives. Since the hyperbolic measure is $\mathrm{SL}_2(\mathbb{Z})$-invariant,

$$\sum_j \int_{\mathcal{D}^*} \varphi(\alpha_j(\tau))d\mu(\tau) = \int_{\bigcup_j \alpha_j(\mathcal{D}^*)} \varphi(\tau)d\mu(\tau).$$

Since the set $\bigcup_j \alpha_j(\mathcal{D}^*)$ represents $X(\Gamma)$ up to identification of boundary, we define identify integration of the modular curve $X(\Gamma)$ by

$$\int_{X(\Gamma)} \varphi(\tau)d\mu(\tau) = \int_{\bigcup_j \alpha_j(\mathcal{D}^*)} \varphi(\tau)d\mu(\tau).$$

Definition 4.13. Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. The **Petersson inner product** is the mapping $\langle\ ,\ \rangle_\Gamma : \mathcal{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \to \mathbb{C}$ given by,

$$\langle f, g\rangle_\Gamma = \frac{1}{V_\Gamma} \int_{X(\Gamma)} f(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k d\mu(\tau)$$

where

$$V_\Gamma = \int_{X(\Gamma)} d\mu(\tau)$$

is the volume of $X(\Gamma)$.

Lemma 4.14. *Let $f, g \in \mathcal{S}_k(\Gamma)$ for any congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$. The function $f(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k$ is $\mathrm{SL}_2(\mathbb{Z})$-invariant.*

Proof. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. The modular transformation property tells us that $f(\gamma\tau) = (c\tau + d)^k f(\tau)$ and similarly $g(\gamma\tau) = (c\tau + d)^k g(\tau)$. Recalling that $\mathrm{Im}(\gamma(\tau)) = \mathrm{Im}(\tau)/|c\tau + d|^2$, we have

$$\begin{aligned} f(\gamma\tau)\overline{g(\gamma\tau)}(\mathrm{Im}(\gamma\tau))^k &= (c\tau + d)^k f(\tau)\overline{(c\tau + d)^k g(\tau)} \frac{\mathrm{Im}(\tau)^k}{|c\tau + d|^{2k}} \\ &= |c\tau + d|^{2k} f(\tau)\overline{g(\tau)} \frac{\mathrm{Im}(\tau)^k}{|c\tau + d|^{2k}} \\ &= f(\tau)\overline{g(\tau)}\mathrm{Im}(z)^k \end{aligned}$$

as desired. □

Lemma 4.15. *The Petersson inner product is well-defined and Hermitian.*

Proof. Let $f, g, h \in \mathcal{S}_k(\Gamma)$, and let $c \in \mathbb{C}$. Since f is a cusp form, the integrand decays exponentially as $\mathrm{Im}(\tau) \to \infty$ and hence the integral converges. The integral is independent of choice of fundamental domain by the previous lemma, so the integral is well-defined.

Now we check that it is a Hermitian inner product. That $\langle f + g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$ and $\langle cf, g\rangle = c\langle f, g\rangle$ follows immediately from the linearity of integration.

$$\overline{\langle g, f\rangle} = \overline{\frac{1}{V_\Gamma} \int_{X(\Gamma)} g(\tau)\overline{f(\tau)}(\mathrm{Im}(\tau))^k\, d\mu(\tau)}$$

$$\begin{aligned}
&= \frac{1}{V_\Gamma}\int_{X(\Gamma)} \overline{g(\tau)} f(\tau) (\mathrm{Im}(\tau))^k \, \overline{d\mu(\tau)} \\
&= \frac{1}{V_\Gamma}\int_{X(\Gamma)} f(\tau)\overline{g(\tau)} (\mathrm{Im}(\tau))^k \, d\mu(\tau).
\end{aligned}$$

If $f = 0$ then $\langle f, f\rangle = 0$ since the integrand is zero. Assuming $f \neq 0$,

$$\langle f, f\rangle = \frac{1}{V_\Gamma}\int_{X(\Gamma)} |f(\tau)|^2 (\mathrm{Im}(\tau))^k \, d\mu(\tau) > 0.$$

□

The normalising factor $1/V_\Gamma$ ensures that if $\Gamma' \subset \Gamma$ then $\langle\,,\rangle_{\Gamma'} = \langle\,,\rangle_\Gamma$ when we restrict $\langle\,,\rangle_{\Gamma'}$ to $\mathcal{S}_k(\Gamma)$.

Remark. Since the convergence of Petersson inner product requires only the product fg vanish at each cusp, Definition 4.13 can be generalised to to a bilinear form $\mathcal{M}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \to \mathbb{C}$.

Lemma 4.16. *In the inner product space $\mathcal{S}_k(\Gamma_1(N))$ equipped with the Petterson inner product, the adjoints of the Hecke operators $\langle p\rangle$ and T_p for $p \nmid N$ are given by*

$$\langle p\rangle^* = \langle p\rangle^{-1} \quad \textit{and} \quad T_p^* = \langle p\rangle^{-1} T_p.$$

It follows that the Hecke operators $\langle n\rangle$ and T_n are normal if $\gcd(n, N) = 1$.

Theorem 4.17. *The space $\mathcal{S}_k(\Gamma_1(N))$ has an orthogonal basis of eigenforms for the Hecke operators*

$$\{T_n, \langle n\rangle : \gcd(n, N) = 1\}.$$

Proof. See Theorem 5.5.4. in [7]. □

4.2 Oldforms and newforms

So far, the Hecke operators have given us maps between vector spaces of modular forms and cusps forms at one generic level N. Now, we move our attention to comparing modular forms at different levels. In particular, suppose M divides N and d divides N/M. If $f \in \mathcal{M}_k(\Gamma_1(M))$, define $\tilde{f}$ by $\tilde{f}(\tau) := f(d\tau)$. Then $\tilde{f} \in \mathcal{M}_k(\Gamma_1(N))$ and so we have an embedding $\mathcal{S}_k(\Gamma_1(M)) \subset \mathcal{S}_k(\Gamma_1(N))$. While the cusp forms in the image of the embedding have level N, they arise in this manner from a smaller level M. In this sense, we consider such forms "old" and thus want to distinguish these from the truly "new" forms. To this end, we introduce the theory of oldforms and newforms due to Atkin and Lehner [14].

For each positive divisor of d of N, we define i_d to be the map

$$\begin{aligned}
i_d : (\mathcal{S}(\Gamma_1(Nd^{-1})))^2 &\to S_k(\Gamma_1(N)) \\
(f, g) &\to f + g[\alpha_d]_k
\end{aligned}$$

where $\alpha_d = \left(\begin{smallmatrix} d & 0 \\ 0 & 1\end{smallmatrix}\right)$ so that $f[\alpha_d]_k(\tau) = d^{k-1} f(d\tau)$.

Definition 4.18. The subspace of **oldforms at level** N is

$$\mathcal{S}_k(\Gamma_1(N))^{\text{old}} = \sum_{\substack{p|N \\ p \text{ prime}}} i_p((\mathcal{S}(\Gamma_1(Np^{-1})))^2).$$

The subspace of **newforms at level** N is the orthogonal complement of the set of oldforms at level N with respect to the Petersson inner product,

$$\mathcal{S}_k(\Gamma_1(N))^{\text{new}} = (\mathcal{S}_k(\Gamma_1(N))^{\text{old}})^{\perp}.$$

Lemma 4.19. *$\mathcal{S}_k(\Gamma_1(N))^{old}$ and $\mathcal{S}_k(\Gamma_1(N))^{new}$ are stable under the Hecke operators $\langle n\rangle$ and T_n for all positive integers n.*

Proof. See Proposition 5.6.2. in [7].

□

Corollary 4.20. *The spaces of oldforms $\mathcal{S}_k(\Gamma_1(N))^{old}$ and newforms $\mathcal{S}_k(\Gamma_1(N))^{new}$ have orthogonal bases of eigenforms for the Hecke operators away from the level. That is, they are eigenforms with respect to*

$$\{T_n, \langle n\rangle : \gcd(n, N) = 1\}.$$

The condition that $\gcd(n, N) = 1$ can be removed for newforms. To prove this, we will need the Main Lemma which is due to Atkin and Lehner [14].

Define the map $\iota_d : \mathcal{S}_k(\Gamma_1(M)) \to \mathcal{S}_k(\Gamma_1(N))$ by $(\iota_d f)(\tau) = f(d\tau)$. On q-expansions ι_d acts by

$$\sum_{n=1}^{\infty} a_n q^n \mapsto \sum_{n=1}^{\infty} a_n q^{dn}, \quad q = e^{2\pi i\tau}.$$

Thus if $f \in \mathcal{S}_k(\Gamma_1(N))$ is of the form $f = \sum_{p|N} \iota_p f_p$ with $f_p \in \mathcal{S}_k(\Gamma_1(N/p))$, then the q-expansion $f(\tau) = \sum a_n(f)q^n$ has $a_n(f) = 0$ for all n coprime to N. The result of the Main Lemma is that the converse is true as well.

Theorem 4.21. *(Main Lemma). Suppose $f \in \mathcal{S}_k(\Gamma_1(N))$ has Fourier expansion*

$$f(\tau) = \sum a_n(f)q^n, \quad q = e^{2\pi i\tau}$$

with $a_n(f) = 0$ whenever $\gcd(n, N) = 1$. Then f can be written in the form $f = \sum_{p|N} \iota_p f_p$ with $f_p \in \mathcal{S}_k(\Gamma_1(N/p))$.

Definition 4.22. An **eigenform** is a modular form which is an eigenvector for the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$.

An eigenform that is also a cusp form is called a **cuspidal eigenform**.

Definition 4.23. Let $f = \sum_{n=0}^{\infty} a_n(f)q^n$ be a Hecke eigenform. We say that f is **normalised** if $a_1(f) = 1$.

Definition 4.24. A **newform** is a normalised eigenform in $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$.

Corollary 4.25. *If $f(\tau) = \sum_{n=0}^{\infty} a_n(f)q^n$ is a Hecke eigenform, then for all $n \geq 1$ we have $a_1(T_n f) = a_n(f)$. That is, the eigenvalue of T_n for f is precisely the nth Fourier coefficient of $T_n f$.*

Proof. This follows from Proposition 4.10 □

Theorem 4.26. *Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be a nonzero eigenform for the Hecke operators away from the level. Then f is an eigenform. Moreover, if f' also satisfies the conditions for f above and has the same eigenvalues for the Hecke operators T_n, then $f' = cf$ for some constant c.*

The set of newforms in $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ is an orthogonal basis of the space. Each newform lies in an eigenspace $\mathcal{S}_k(N,\chi)$ and $T_n f = a_n(f)f$ for all positive integers n.

Proof. Let $f \in \mathcal{S}_k(\Gamma_1(N))$ be an eigenform for the Hecke operators for all T_n and $\langle n \rangle$ with $\gcd(n,N) = 1$. Let $c_n, d_n \in \mathbb{C}$ denote the eigenvalues $T_n f = c_n f$ and $\langle n \rangle f = d_n f$. The map $n \mapsto d_n$ defines a Dirichlet character χ modulo N, so $f \in \mathcal{S}_k(N,\chi)$. By Proposition 4.10, we have $a_1(T_n f) = a_n(f)$ for all positive integers n. Therefore, when $\gcd(n,N) = 1$ we have $a_n(f) = c_n a_1(f)$. Thus if if $a_1(f) = 0$, then $a_n(f) = 0$ whenever $\gcd(n,N) = 1$ and by the Main Lemma, $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{old}}$.

Now suppose that $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$, $f \neq 0$. Then the preceding argument shows that $a_1(f) \neq 0$, so we may assume that f is normalised so that $a_1(f) = 1$. Let m be a positive integer and let $g_m = T_m f - a_m(f)f$. Since $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ is stable under the Hecke operators, $g_m \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$. Moreover, if $\gcd(n,N) = 1$ then $T_n g_m = T_n T_m f - a_m(f) T_n f = (T_m - a_m(f))c_n f = c_n g_m$, so g_m is also an eigenform for the Hecke operators T_n away from the level. Similarly, $\langle n \rangle g_m = \langle n \rangle T_m f - a_m(f)\langle n \rangle f = (T_m - a_m(f))d_n f = d_n g_m$, so g_m is an eigenform for the Hecke operators $\langle n \rangle$ away from the level. The first coefficient of g_m is $a_1(g_m) = a_1(T_m f) - a_1(a_m(f)f) = a_m(f) - a_m(f) = 0$ where we have used that f is normalised so $a_1(f) = 1$. Thus $g_m \in \mathcal{S}_k(\Gamma_1(N))^{\text{old}}$, but now we have $g_m \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}} \cap \mathcal{S}_k(\Gamma_1(N))^{\text{old}} = \{0\}$, so $g_m = 0$ and $T_m f = a_m(f)f$.

We check that the set of newforms in $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ is linearly independent. Suppose

$$\sum_{i=1}^{n} c_i f_i = 0, \quad c_i \in \mathbb{C}$$

is a nontrivial linear relation with the minimal number of terms. Let p be prime and apply the linear operator $T_p - a_p(f_1)$ to obtain the relation

$$\sum_{i=1}^{n} c_i(T_p - a_p(f_1))f_i = c_1(a_p(f_1) - a_p(f_1)) + \sum_{i=2}^{n} c_i(a_p(f_i) - a_p(f_1))f_i = \sum_{i=2}^{n} c_i(a_p(f_i) - a_p(f_1))f_i = 0.$$

This has fewer terms than linear relation we started with, so it must be trivial, and so $a_p(f_i) = a_p(f_1)$ for all i. Thus $f_i = f_1$ for all i, but this is a contradiction since we assumed the original linear relation was nontrivial. □

4.3 Modular Jacobians and Hecke operators

The Jacobian of a curve is an abelian variety that can be seen as a parameter space for divisor classes of degree 0. Over the complex numbers, the Jacobian has an analytical construction as detailed in Appendix A. Of particular importance is Abel's theorem which identifies the Jacobian $\mathrm{Jac}(X)$ with the degree-0 Picard group $\mathrm{Pic}^0(X)$.

Let

$$J_1(N) = \mathrm{Jac}(X_1(N))$$

denote the Jacobian associated to $\Gamma_1(N)$ and let

$$J_0(N) = \mathrm{Jac}(X_0(N))$$

denote the Jacobian associated to $X_0(N)$.

Recall the description of the double coset operator in terms of modular curves viewing $[\Gamma_1\alpha\Gamma_2]_k$ as the composition $(\pi_1)_* \circ \alpha_* \circ (\pi_2)^*$, the composition of a pullback and push forwards as described in Section 3.3. This induces a map $(\tilde{\pi_1})_* \circ \tilde{\alpha}_* \circ (\tilde{\pi_2})^*$ on degree-0 Picard groups which we will denote by the same symbol,

$$[\Gamma_1\alpha\Gamma_2]_k : \mathrm{Pic}^0(X_2) \to \mathrm{Pic}^0(X_1),$$

and which via Abel's Theorem gives a map of Jacobians,

$$[\Gamma_1\alpha\Gamma_2]_k : \mathrm{Jac}(X_2) \to \mathrm{Jac}(X_1).$$

To rephrase this in terms of modular forms we will need the following result.

Lemma 4.27. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. Let $f \in \mathcal{S}_2(\Gamma)$. Then $f(z)dz$ defines a holomorphic differential on $X_0(N)$. Moreover, the map $f \mapsto f(z)dz$ is an isomorphism of complex vector spaces.*

Consequently, the dual spaces $\mathcal{S}_2(\Gamma)^\vee$ and $(\Omega^1_{\mathrm{hol}}(X(\Gamma))^\vee$ are also identified, and there is an isomorphism

$$\mathrm{Jac}(X(\Gamma)) \cong \mathcal{S}_2(\Gamma)^\vee/\mathrm{H}_1(X(\Gamma), \mathbb{Z}).$$

To simplify notation, for the remainder of this section we will identify the Jacobian of the modular curve $X(\Gamma)$ as

$$\mathrm{Jac}(X(\Gamma)) = \mathcal{S}_2(\Gamma) \vee /\mathrm{H}_1(X(\Gamma), \mathbb{Z})$$

and work with this new definition.

Let X_1 and X_2 be the modular curves associated to the congruence subgroups Γ_1 and Γ_2. Let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$ be such that $\alpha\Gamma_1\alpha^{-1} \subset \Gamma_2$ and consider the map

$$\begin{aligned} h : X_1 &\to X_2 \\ \Gamma_1\tau &\mapsto \Gamma_2\alpha(\tau). \end{aligned}$$

The pullback $h^\star$ in the differential-geometric sense is compatible with the weight-2 operator in the sense that

$$\begin{tikzcd}
\mathcal{S}_2(\Gamma_2) \arrow[r, "{[\alpha_2]}"] \arrow[d] & \mathcal{S}_2(\Gamma_1) \arrow[d] \\
\Omega^1_{\text{hol}}(\Gamma_2) \arrow[r, "h^\star"] & \Omega^1_{\text{hol}}(\Gamma_1)
\end{tikzcd}$$

commutes. The induced push forward on dual spaces is

$$\begin{aligned} h_\star : \mathcal{S}_2(\Gamma_1)^\vee &\to \mathcal{S}_2(\Gamma_2)^\vee \\ \varphi &\mapsto \varphi \circ [\alpha]_2 \end{aligned}$$

and we have a commutative diagram

$$\begin{tikzcd}
\mathcal{S}_2(\Gamma_1) \arrow[r, "\sum_j [\gamma_{2,j}]_2"] \arrow[d] & \mathcal{S}_2(\Gamma_2) \arrow[d] \\
\Omega^1_{\text{hol}}(\Gamma_2) \arrow[r, "\text{tr}_h"] & \Omega^1_{\text{hol}}(\Gamma_1)
\end{tikzcd}$$

where $\alpha\Gamma_1\alpha^{-1}\backslash\Gamma_2 = \cup_j \alpha\Gamma_1\alpha^{-1}\gamma_{2,j}$. The map $\sum_j [\gamma_{2,j}]_2$ induces the map

$$\begin{aligned} \text{tr}_h^\vee : \mathcal{S}_2(\Gamma_2)^\vee &\to \mathcal{S}_2(\Gamma_1)^\vee \\ \psi &\mapsto \psi \circ \sum_j [\gamma_{2,j}]_2. \end{aligned}$$

The commutativity of the two diagrams ensures that h_* and $\text{tr}_h^\vee$ descend to the maps h_* and h^* on Jacobians as before. The pullback of the weight-2 double coset operator

$$[\Gamma_1\alpha\Gamma_2]_2 : \mathcal{S}_2(\Gamma_1) \to \mathcal{S}_2(\Gamma_2), \quad f[\Gamma_1\alpha\Gamma_2]_2 = \sum_j f[\beta_j]_2$$

is

$$[\Gamma_1\alpha\Gamma_2]_2 : \mathcal{S}_2(\Gamma_1)^\vee \to \mathcal{S}_2(\Gamma_2)^\vee, \quad [\Gamma_1\alpha\Gamma_2]_2 f = f[\Gamma_1\alpha\Gamma_2]_2.$$

Proposition 4.28. *The weight-2 Hecke operators $T = T_p$ and $T = \langle d \rangle$ act by composition on the right on $J_1(N)$,*

$$\begin{aligned} T : J_1(N) &\to J_1(N) \\ [\varphi] &\mapsto [\varphi \circ T], \quad \varphi \in \mathcal{S}_2(\Gamma_1(N))^\vee. \end{aligned}$$

The Hecke operator T_p acts by composition on the right on $J_0(N)$.

Proof. See Proposition 6.3.2. in [7]. □

Definition 4.29. The **Hecke algebra over** $\mathbb{Z}$ for is the $\mathbb{Z}$-algebra of endomorphisms of $\mathcal{S}_2(\Gamma_1(N))$ generated by the Hecke operators,

$$\mathbb{T}_{\mathbb{Z}} = \mathbb{Z}[\{T_n, \langle n \rangle : n \in \mathbb{Z}^+\}].$$

The **Hecke algebra over** C, denoted $\mathbb{T}_{\mathbb{C}}$ is defined similarly except with $\mathbb{C}$ in place of $\mathbb{Z}$.

Remark. Each level, N, has its own Hecke algebra, but we have omitted N from the notation as it will not play a role in the discussion to follow.

Let $f(\tau) = \sum_{n=1}^{\infty} a_n(f)q^n$ be a normalised eigenform. Define the homomorphism $\lambda_f : \mathbb{T}_\mathbb{Z} \to \mathbb{C}$ by $Tf = \lambda_f(T)f$. Viewing $\mathbb{T}_\mathbb{Z}$ as the ring of endomorphisms of $\mathrm{H}_1(X_1(N), \mathbb{Z})$ which is a finitely generated $\mathbb{Z}$-module, $\mathbb{T}_\mathbb{Z}$ must also be finitely generated over $\mathbb{Z}$. Thus the image $\lambda_f(\mathbb{T}_\mathbb{Z}) = \mathbb{Z}[\{a_n(f) : n \in \mathbb{Z}^+\}]$ is a finitely generated $\mathbb{Z}$-module. That is, ring generated by the eigenvalues $a_n(f)$ has finite rank as a $\mathbb{Z}$-module. Let

$$I_f = \ker(\lambda_f).$$

Then by the first isomorphism theorem be have an isomorphism

$$\mathbb{T}_\mathbb{Z}/I_f \overset{\sim}{\to} \mathbb{Z}[\{a_n(f)\}].$$

Definition 4.30. Let $f \in \mathcal{S}_2(\Gamma_1(N))$ be a normalised eigenform with q-expansion $f(\tau) = \sum_{n=1}^{\infty} a_n(f)q^n$. The **number field of** f is the field $K_f = \mathbb{Q}(\{a_n(f)\})$ generated by the Fourier coefficients of f over $\mathbb{Q}$.

Theorem 4.31. *Let f be a normalised eigenform of weight 2 so that $f \in \mathcal{S}_2(N, \chi$ for some Dirichlet character χ modulo N. Let K_f be its number field. Let $\sigma : K_f \hookrightarrow \mathbb{C}$ be any embedding. Define f^σ by*

$$f^\sigma(\tau) = \sum_{n=1}^{\infty} \sigma(a_n)q^n.$$

Then f^σ is also a normalised eigenform and $f \in \mathcal{S}_2(N, \chi^\sigma)$ where $\chi^\sigma(n) = \sigma(\chi(n))$. If f is a newform then so is f^σ.

Proof. See Theorem 6.5.4. in [7]. □

Lemma 4.32. *The space $\mathcal{S}_k(\Gamma_1(N))$ has a basis of newforms with coefficients in $\mathbb{Z}$.*

Proof. See Corollary 6.5.6. in [7]. □

Now let $f \in \mathcal{S}_{\in}(-_{\infty}(\mathcal{M}_{\{}))$ be a newform at some level M_f. The action of the Hecke algebra $\mathbb{T}_\mathbb{Z}$ on the Jacobian $J_1(M_f)$ is well-defined, so the subgroup $I_f J_1(M_f)$ is also well-defined.

Definition 4.33. The **abelian variety associated to** f is the quotient

$$A_f = J_1(M_f)/I_f J_1(M_f).$$

5 The Modularity Theorem and Galois representations

5.1 Function fields over $\mathbb{C}$ and $\mathbb{Q}$

Recall the modular curves $X_0(N) = \Gamma_0(N)\backslash\mathbb{H}^*$, $X_1(N) = \Gamma_1(N)\backslash\mathbb{H}^*$, and $X(N) = \Gamma(N)\mathbb{H}^*$. In this section, we look briefly at their function fields with the aim of defining universal elliptic curves and introducing the concept of analytic conductors of elliptic curves. Proofs in this section will largely be omitted; see Sections 7.5 and 7.6 of [7] for details.

We will first discuss the case over fields of characteristic zero, so let k be a field with $\mathrm{char}(k) = 0$ and consider Weierstrass equations of the form

$$E : y^2 = 4x^3 - g_2x - g_3, \quad g_2, g_3 \in k$$

from Section 1.4. In this context, the **admissible change of variable** is

$$x = u^2x', \quad y = u^3y', \quad u \in k^\times$$

which transforms a Weierstrass equations of this form to another Weierstrass equation of this form. It also sends g_2 to g_2/u^4, g_3 to g_3/u^6, and Δ to Δ/u^{12}. The invariant j is preserved and the admissible change of variable induces an isomorphism between the corresponding elliptic curves.

Recall the function $f_0^{\overline{v}}$ from the end of Section 1.5. We will use without proof that $f_0^{\overline{v}} \in \mathbb{C}(X(N))$. Verifying this statement amounts to checking that $f_0^{\overline{v}}$ is meromorphic on the upper half-plane and at the cusps. Assuming this holds, since the functions $f_0^{\overline{d}}$ and f_0 are special cases of $f_0^{\overline{v}}$ that are invariant under $\Gamma_1(N)$ and $\Gamma_0(N)$ respectively, we conclude that $f_0^{\overline{d}} \in \mathbb{C}(X_1(N))$ and $f_0 \in \mathbb{C}(X_0(N))$.

Also recall the j-invariant from Section 1.1 and define

$$j_N(\tau) = j(N\tau).$$

We will also use without proof that $j_N \in \mathbb{C}(X_0(N))$.

Proposition 5.1. *The fields of meromorphic functions on $X(N)$, $X_1(N)$, and $X_0(N)$ are*

$$\begin{aligned}
\mathbb{C}(X(N)) &= \mathbb{C}(j, \{f_0^{\pm\overline{v}} : \pm v \in ((\mathbb{Z}/N\mathbb{Z})^2 - \{(0,0)\})/\pm\}) \\
&= \mathbb{C}(j, f_{1,0}, f_{0,1}) \\
\mathbb{C}(X_1(N)) &= \mathbb{C}(j, \{f_0^{\pm\overline{d}} : \pm\overline{d} \in (\mathbb{Z}/N\mathbb{Z} - \{0\})/\pm\}) \\
&= \mathbb{C}(j, f_1) \\
\mathbb{C}(X_0(N)) &= \mathbb{C}(j, f_0) \\
&= \mathbb{C}(j, j_N)
\end{aligned}$$

where to simplify notation we have written

$$f_{1,0} = f_0^{\pm\overline{(1,0)}} \quad \textit{and} \quad f_{0,1} = f_1 = f_0^{\pm\overline{(0,1)}}.$$

Proof. See Proposition 7.5.1. in [7]. □

Recall that the map $(\wp_\tau, \wp'_\tau)$ sends the complex elliptic curve $\mathbb{C}/\Lambda_\tau$ to the algebraic curve given by the zero set of

$$E_\tau : y^2 = 4x^3 - g_2(\tau)x - g_3(\tau).$$

Combining Proposition 5.1 with the equation

$$f_0^{\overline{v}}(\tau) = \frac{g_2(\tau)}{g_3(\tau)}\wp_\tau\left(\frac{c_v\tau + d_v}{N}\right)$$

from Section 1.5 shows that the field of meromorphic functions on $X(N)$ is generated by the j-invariant and by functions of τ that are related to the x-coordinates of the nonzero N-torsion points on the elliptic curve defined by E_τ. We can scale E_τ to a new curve that is generated exactly by j and the N-torsion x-coordinates.

To do so, fix $\tau \in \mathbb{H}$ such that $j(\tau) \neq 0, 1728$. Since $j = 1728g_2^3/(g_2^3 - 27g_3^2)$, this means that $g_2(\tau), g_3(\tau) \neq 0$. Let fix $(g_2(\tau)/g_3(\tau))^{1/2}$ to be either of the complex square roots. The map

$$\left(\frac{g_2(\tau)}{g_3(\tau)}\wp_\tau, \left(\frac{g_2(\tau)}{g_3(\tau)}\right)^{3/2}\wp'_\tau\right) : \mathbb{C}/\Lambda_\tau \to \mathbb{C}^2 \cup \{\infty\}$$

differs from $(\wp_\tau, \wp'_\tau)$ by the admissible change of variable

$$(x, y) = (u^2x', u^3y'), \quad u = (g_3(\tau)/g_2(\tau))^{1/2}.$$

This sends both $g_2(\tau)$ and $g_3(\tau)$ to $g_2(\tau)^3/g_3(\tau)^2$, and maps nonzero points $z + \Lambda_\tau$ to points (x, y) satisfying

$$E_{j(\tau)} : y^2 = 4x^3 - \frac{g_2(\tau)^3}{g_3(\tau)^2}x - \frac{g_2(\tau)^3}{g_3(\tau)^2}.$$

Since $g_2^3/g_3^2 = 27j/(j - 1728)$ we have equivalently

$$E_{j(\tau)} : y^2 = 4x^3 - \frac{27j(\tau)}{j(\tau) - 1728}x - \frac{27j(\tau)}{j(\tau) - 1728}.$$

The map $\mathbb{C}/\Lambda_\tau \to E_{j(\tau)}$ restricts to an isomorphism of N-torsion subgroups, taking the generators $\tau/N + \Lambda_\tau$ and $1/N + \Gamma_\tau$ of $(\mathbb{C}/\Lambda_\tau)[N]$ to

$$P_\tau = \left(\frac{g_2(\tau)}{g_3(\tau)}\wp_\tau(\tau/N), \left(\frac{g_2(\tau)}{g_3(\tau)}\right)^{3/2}\wp'_\tau(\tau/N)\right)$$

$$Q_\tau = \left(\frac{g_2(\tau)}{g_3(\tau)}\wp_\tau(1/N), \left(\frac{g_2(\tau)}{g_3(\tau)}\right)^{3/2}\wp'_\tau(1/N)\right)$$

respectively. Choosing the other complex square root negates these two points. Modulo this, (P_τ, Q_τ) is a canonical ordered basis of $E_{j(\tau)}[N]$ over $\mathbb{Z}/N\mathbb{Z}$. The x-coordinates of $\pm P_\tau$ and $\pm Q_\tau$ are given by $f_{1,0}(\tau)$ and $f_{0,1}(\tau)$ respectively, so the set of x-coordinates of the nonzero N-torsion points of $E_{j(\tau)}$ is $\{f_0^{\pm\overline{v}}(\tau)\}$.

Thus, up to change of sign, the function field $\mathbb{C}(X(N)) = \mathbb{C}(j, f_{1,0}, f_{0,1})$ describes an enhanced elliptic curve for $\Gamma(N)$, namely

$$(E_{j(\tau)}, \pm(P_\tau, Q_\tau)),$$

representing a point of the moduli space $S(N)$. Similarly, $j(\tau)$ and $f_1(\tau)$ describes the enhanced elliptic curve $(E_{j(\tau)}, \pm Q_\tau)$ of $\Gamma_1(N)$, representing a point in $S_1(N)$, and $j(\tau$ and $f_0(\tau)$ describes $(E_{j(\tau)}, \langle Q_\tau \rangle)$ representing a point in $S_0(N)$.

If we consider τ as a variable and view the invariant j as the j-invariant function in τ, then we can gather the family of elliptic curves of the form $E_{j(\tau)}$ in the **universal elliptic curve**

$$E_j : y^2 = 4x^3 - \frac{27j}{j-1728}x - \frac{27j}{j-1728}$$

which, since j surjects from $\mathbb{H}$ to $\mathbb{C}$ (see Proposition 1.7. in [24]), gives a complex elliptic curve for every complex number $j \neq 0, 1728$.

Proposition 5.2. *Let E_j be the universal elliptic curve described above. The field of meromorphic functions $\mathbb{C}(X(N))$ on $X(N)$ is the generated by the x-coordinates of the nonzero N-torsion points on E_j, i.e.,*

$$\mathbb{C}(X(N)) = \mathbb{C}(j, x(E_j[N])).$$

Moreover, the field extension $\mathbb{C}(j, x(E_j[N]))$ is Galois and the corresponding Galois group is isomorphic to $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm I\}$.

The quotient by $\{\pm I\}$ can be eliminated by adjoining the y-coordinates of the N-torsion points to $\mathbb{C}(j)$, giving the following.

Lemma 5.3. *The field extension $\mathbb{C}(j, E_j[N])/\mathbb{C}(j)$ is Galois and the corresponding Galois group is isomorphic to $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$.*

Proof. See Corollary 7.5.3. in [7]. □

Let $H = \mathrm{Gal}(\mathbb{C}(j, E_j[N])/\mathbb{C}(j))$. The ordered basis (P_τ, Q_τ) specifies a representation

$$\rho : H \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$
$$\rho(\sigma)\begin{pmatrix} P_\tau \\ Q_\tau \end{pmatrix} = \begin{pmatrix} \sigma(P_\tau) \\ \sigma(Q_\tau) \end{pmatrix}$$

which is injective with image $\rho(H) = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$.

Now we turn our attention to the study of modular curves as algebraic curves over the rational numbers, working with function fields over $\mathbb{Q}$ instead of $\mathbb{C}$. Since $\mathbb{Q}$ is a prime subfield of $\mathbb{C}$, much of the algebraic structure developed so far will carry over.

The equation defining the universal elliptic curve E_j has coefficients in $\mathbb{Q}(j)$. Viewing the curve as defined as defined over $\mathbb{Q}(j)$ means we are considering the points $(x, y) \in \overline{\mathbb{Q}(j)}^2$ that satisfy E_j. Thus, this includes the nonzero N-torison points $E_j[N]$ over $\mathbb{C}(j)$. The extension $\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j)$ is again Galois. Define

$$H_{\mathbb{Q}} = \mathrm{Gal}(\mathbb{Q}(\mu_N, j, E_j[N])/\mathbb{Q}(j))$$

where μ_N is the set of complex Nth roots of unity. We have a similar representation defined in terms of the ordered basis (P_τ, Q_τ),

$$\rho : H_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$

$$\rho(\sigma)\begin{pmatrix}P_\tau\\Q_\tau\end{pmatrix}=\begin{pmatrix}\sigma(P_\tau)\\\sigma(Q_\tau)\end{pmatrix},\quad \sigma\in H_{\mathbb{Q}}.$$

Lemma 5.4. *Let $\mu\in\mu_N$ and let $\sigma\in H_{\mathbb{Q}}$. Then*

$$\sigma(\mu)=\mu^{\det(\rho(\sigma))}.$$

Thus $\mu_N\subset\mathbb{Q}(j,E_j[N])$ so $H_{\mathbb{Q}}=\mathrm{Gal}(\mathbb{Q}(j,E_j[N])/\mathbb{Q}(j)$. The field extension $\mathbb{Q}(j,E_j[N])/\mathbb{Q}(j)$ is generated by $E_j[N]$ so ρ is injective, and by Lemma 5.4, ρ restricts to $H_{\mathbb{Q}(\mu_N)}\to\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$.

Theorem 5.5. *(Restriction Lemma). Let k and F be field extensions of f inside K such that the extension F over f is Galois. If $K=kF$ then K/k is Galois and there is a natural injection*

$$\mathrm{Gal}(K/k)\to\mathrm{Gal}(F/f)$$

which has image $\mathrm{Gal}(F/(k\cap F))$.

Proof. See Lemma 7.6.2. in [7]. □

The Restriction Lemma shows that $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ injects into $H_{\mathbb{Q}(\mu_N)}$, and since $\rho:H_{\mathbb{Q}(\mu_N)}\to\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is injective, ρ must be an isomorphism since the two groups are finite. The lemma further shows that $\mathbb{C}(j)\cap\mathbb{Q}(j,E_j[N])=\mathbb{Q}(\mu_N,j)$ and $\mathbb{Q}(j,E_j[N])\cap\overline{\mathbb{Q}}=\mathbb{Q}(\mu_N)$.

If we consider the inclusions $\mathbb{Q}(j)\subset\mathbb{Q}(\mu_N,j)\subset\mathbb{Q}(j,E_j[N])$, we see that $|H_{\mathbb{Q}}|=|H_{\mathbb{Q}(\mu_n)}||(\mathbb{Z}/N\mathbb{Z})^\times|=|\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})||(\mathbb{Z}/N\mathbb{Z})^\times|=|\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})|$, so the representation $\rho:H_{\mathbb{Q}}\to\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ surjects, and therefore is an isomorphism.

Let $\mathbf{K}$ be an intermediate field extension of $\mathbb{Q}(j,E_j[N])/\mathbb{Q}(j)$. We want to know whether or not $\mathbb{K}$ corresponds to an algebraic curve over $\mathbb{Q}$. To this end, let $K=\mathrm{Gal}(\mathbb{Q}(j,E_j[N])/K)\subset H_{\mathbb{Q}}$. Since $\det(\rho)$ specifies how $H_{\mathbb{Q}}$ permutes μ_N, we have

$$\mathbf{K}\cap\overline{\mathbb{Q}}=\mathbb{Q}\iff\mathbf{K}\cap\mathbb{Q}(\mu_N)=\mathbb{Q}\iff\det(\rho):K\to(\mathbb{Z}/N\mathbb{Z})^\times\text{ is surjective.}$$

That is $\mathbf{K}$ is the function field of an algebraic curve over $\mathbb{Q}$ if and only if $\det(\rho:K\to(\mathbb{Z}/N\mathbb{Z})^\times)$ is surjective.

5.2 Modular curves and modularity

In intermediate field extensions of $\mathbb{Q}(j,E_j[N])/\mathbb{Q}(j)$ that we are particularly interested in are the ones corresponding to X_0 and X_1. Let $\mathbf{K}_0=\mathbb{Q}(j,f_0)$ and $\mathbf{K}_1=\mathbb{Q}(j,f_1)$, with corresponding subgroups K_0 and K_1 of $H_{\mathbb{Q}}$. One can show that

$$\rho(K_0)=\left\{\begin{pmatrix}a&b\\0&d\end{pmatrix}\in\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})\right\}$$

and

$$\rho(K_1)=\left\{\pm\begin{pmatrix}a&b\\0&1\end{pmatrix}\in\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})\right\}.$$

In both cases $\det(\rho)$ surjects onto $(\mathbb{Z}/N\mathbb{Z})^\times$, so $\mathbf{K}_0$ and $\mathbf{K}_1$ are function fields of nonsingular projective algebraic curves over $\mathbb{Q}$. We denote these curves $X_0(N)_{\mathrm{alg}}$ and $X_1(N)_{\mathrm{alg}}$ respectively. To relate the two algebraic curves back to the complex modular curves $X_0(N)$ and $X_(N)$ we will need the following theorem.

Theorem 5.6. *Let k be a field and let C be a nonsingular projective algebraic curve over k defined by polynomials $\varphi_1, \dots, \varphi_m \in k[x_1, \dots, x_n]$. Let $k(C) = k(t)[u]/\langle p(u)\rangle$. Suppose $\mathbf{K}$ is a field containing k. Then viewing the polynomials $\varphi_1, \dots, \varphi_m$ as elements of $\mathbf{K}[x_1, \dots, x_n]$ defines a nonsingular algebraic curve C' over $\mathbf{K}$. Moreover, the function field of C' is given by $\mathbf{K}(C') = \mathbf{K}(t)[u]/\langle p(u)\rangle$.*

In the context of modular curves, let $k = \mathbb{Q}$ and $C = X_1(N)_{\mathrm{alg}}$. Let $p_1 \in \mathbb{Q}(j)[x]$ be the minimal polynomial of $f_1 = f_{0,1}$ over $\mathbb{Q}(j)$ so that we now have $\mathbb{Q}(C) = \mathbb{Q}(j, f_1) = \mathbb{Q}(j)[x]/\langle p_1\rangle$. Let $\mathbf{K} = \mathbb{C}$. Since $\mathbb{C}$ is algebraically closed, the theorem tells us that the points with coordinates in C satisfying the polynomials defining $X_1(N)_{\mathrm{alg}}$ over $\mathbb{Q}$ form a nonsingular algebraic curve C', denoted $X_1(N)_{\mathrm{alg},\mathbb{C}}$, over $\mathbb{C}$ which has function field $\mathbb{C}(j)[x]/\langle p_1\rangle$ and p_1 is the minimal polynomial of f_1 over $\mathbb{C}(j)$. Thus

$$\mathbb{C}(X_1(N)) = \mathbb{C}(j, f_1) = \mathbb{C}(j)[x]/\langle p_1\rangle.$$

By the Curves-Fields correspondence, $X_1(N) = X_1(N)_{\mathrm{alg},\mathbb{C}}$ up to isomorphism over $\mathbb{C}$.

Using f_0 in the place of f_1 in the above argument gives the analogous result for $X_0(N)$. We are now ready to state an algebraic version of the Modularity Theorem.

Theorem 5.7. *(Modularity Theorem, version X). Let $\mathcal{E}$ be an elliptic curve over Q. Then there exists a positive integer N and a surjective morphism over $\mathbb{Q}$ of curves over $\mathbb{Q}$,*

$$X_0(N)_{\mathrm{alg}} \to \mathcal{E}.$$

Definition 5.8. Let $\mathcal{E}$ be an elliptic curve over $\mathbb{Q}$. The smallest positive integer N that satisfies Theorem 5.7 is called the **analytic conductor** of $\mathcal{E}$.

Remark. The analytic conductor of an elliptic curve over $\mathbb{Q}$ is thus well defined on isomorphism classes of $\mathbb{Q}$ of elliptic curves over $\mathbb{Q}$. In fact, it is well defined on isogeny classes of elliptic curves over $\mathbb{Q}$.

Remark. The Jacobians of modular curves also have algebraic descriptions as nonsingular projective varieties defined over $\mathbb{Q}$. A version of the Modularity Theorem in terms of Jacobians is as follows.

Theorem 5.9. *(Modularity Theorem, version J). Let $\mathcal{E}$ be an elliptic curve over Q. Then there exists a positive integer N and a surjective morphism over $\mathbb{Q}$ of varieties over $\mathbb{Q}$,*

$$J_0(N)_{\mathrm{alg}} \to \mathcal{E}.$$

5.3 Galois representations and elliptic curves

As we will not need to work with Weierstrass equations, we drop the convention of using $\mathcal{E}$ to denote and elliptic curve, using E instead. Let E be an elliptic curve over $\mathbb{Q}$. Recall from section 1.4 that for any positive integer N have the N-torsion subgroup $E[N] \subset E$ as the kernel of the N-fold addition map $[N]$. Since the group law on E is defined by rational functions over $\mathbb{Q}$, the elements of $E[N]$ satisfy some set of polynomial equations with coefficients in $\mathbb{Q}$. Therefore their coordinates are algebraic numbers and so $E[N]$ lies in $E(\overline{Q})$. The same argument shows that $E[N]$ is stable under the action of $G_{\mathbb{Q}}$ on $\overline{Q}$.

In this way, we have a canonical action of $G_{\mathbb{Q}}$ on $E[N]$. Furthermore, for any $\sigma \in G_{\mathbb{Q}}$, the map $P \to \sigma(P)$ is a group automorphism. That is, $\sigma(P+Q) = \sigma(P) + \sigma(Q)$.

If we choose an ordered basis (P, Q) of $E[N]$ over $\mathbb{Z}/N\mathbb{Z}$, then since $E[p^n] \cong (\mathbb{Z}/p^n\mathbb{Z})^2$ (Theorem 3.9), the action of $G_{\mathbb{Q}}$ on $E[N]$ may be viewed as a representation

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}).$$

Let p be prime. The map $[N]$ restricted to p-power torsion subgroups gives maps

$$E[p] \longleftarrow E[p^2] \longleftarrow E[p^3] \longleftarrow \dots$$

Definition 5.10. Let E be an elliptic curve defined over $\mathbb{Q}$ and let p be a prime number. The **p-adic Tate module of E** is the (inverse) limit

$$\mathrm{Ta}_p(E) = \varprojlim_{n \in \mathbb{N}} E[p^n]$$

with transition morphisms given by $[N]|_{E[p^{n+1}]} : E[p^{n+1}] \to E[p^n]$.

Fix a basis (P_n, Q_n) of $E[p^n]$ for each $n \in \mathbb{Z}^+$ such that each basis is a lift of its predecessor, i.e.

$$[p]P_{n+1} = P_n \quad \text{and} \quad [p]Q_{n+1} = Q_n \quad \forall n \in \mathbb{Z}^+.$$

Any basis determines an isomorphism $E[p^n] \cong (\mathbb{Z}/p^n\mathbb{Z})^2$. Therefore, we have

$$\mathrm{Ta}_p(E) \cong \mathbb{Z}_p^2.$$

Similarly, we have isomorphisms

$$\mathrm{Aut}(E[p^n]) \cong \mathrm{GL}_2(\mathbb{Z}/p^n\mathbb{Z})$$

which combine to give

$$\mathrm{Aut}(\mathrm{Ta}_p(E)) \cong \mathrm{GL}_2(\mathbb{Z}_p).$$

For each n, the field $\mathbb{Q}(E[p^n])$ is Galois, so there are restriction maps

$$G_{\mathbb{Q}} \to \mathrm{Gal}(\mathbb{Q}[E[p^n]]), \quad \sigma \mapsto \sigma|_{\mathbb{Q}(E[p^n])}.$$

We also have inclusions

$$\mathrm{Gal}(\mathbb{Q}(E[p^n])/\mathbb{Q}) \to \mathrm{Aut}(E[p^n]).$$

These maps are compatible in the sense that for each n we have a commutative diagram

$$\begin{tikzcd}
& G_{\mathbb{Q}} \arrow[dl] \arrow[dr] & \\
\mathrm{Aut}(E[p^n]) & & \mathrm{Aut}(E[p^{n+1}]) \arrow[ll]
\end{tikzcd}$$

and so $\mathrm{Ta}_p(E)$ is a $G_{\mathbb{Q}}$-module. $G_{\mathbb{Q}}$ acts on $\mathrm{Ta}_p(E)$ to give a continuous homomorphism

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p).$$

Composing with the inclusion $\mathrm{GL}_2(\mathbb{Z}_p) \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_p)$, $\rho_{E,p}$ can be seen as an p-adic Galois representation of dimension 2.

Theorem 5.11. *Let p be a prime number and let E be an elliptic curve over $\mathbb{Q}$ with analytic conductor N. Then the following statements hold.*

i. The Galois representation $\rho_{E,p}$ is unramified at every prime $p \nmid pN$.

ii. For any prime $p \nmid pN$, let $\mathcal{P} \subset \overline{\mathbb{Z}}$ be a maximal ideal lying over (p). Then the characteristic equation of $\rho_{E,p}(\mathrm{Frob}_{\mathfrak{P}})$ is $x^2 - a_p(E)x + p = 0$.

iii. $\rho_{E,p}$ is irreducible.

Proof. See Theorem 9.4.1. in [7]. □

5.4 Galois representations and modularity

In this section we will associate Galois representations to modular curves and define the notion of modularity for a Galois representation. We will be brief here and only state key results, directing the reader to Sections 9.5 and 9.6 of [7] for details.

Let us first make a generalisation of the p-adic numbers to ideals lying over (p).

Definition 5.12. Let K be a number field over $\mathbb{Q}$ with ring of integers $\mathcal{O}_K$, and let p be a prime. We can write the factorisation of $p\mathcal{O}_K$ into maximal ideals as

$$p\mathcal{O}_K = \prod_{\lambda | p} \lambda^{e_\lambda}$$

where the notation $\lambda | p$ means that λ lies over (p). Then we can define the **λ-adic integers** as the limit

$$\mathcal{O}_{K,\lambda} = \varprojlim_n \{\mathcal{O}_K/\lambda^n\}.$$

We also define the **λ-adic numbers** as the field of fractions K_λ of $\mathcal{O}_{K,\lambda}$.

Define the **residue degree** $f_\lambda = [\mathcal{O}_K/\lambda : \mathbf{F}_p]$.

We can view $\mathbb{Z}_p$ as a subring of $\mathcal{O}_{K,\lambda}$, and $\mathbb{Q}_p$ as a subfield of K_λ with equality when $e_\lambda f_\lambda = 1$. We will use without proof that there is a ring homomorphism

$$K \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong \prod_{\lambda | p} K_\lambda. \tag{5.1}$$

Recall that $X_1(N)$ is a projective nonsingular algebraic curve over $\mathbb{Q}$. Thus, it is defined by polynomial equations in $\mathbb{Q}[x_1, \dots, x_n]$ which we can view as polynomials in $\mathbb{C}[x_1, \dots, x_n]$. This defined a nonsingular algebraic curve over $\mathbb{C}$ which we will denote with $\mathbb{C}$ in the subscript. This curve can also be viewed as a compact Riemann surface.

View $J_1(N)$ as the Jacobian of the curve $X_1(N)_{\mathbb{C}} = \mathcal{S}_2(\Gamma_1(N)^{\vee}/\mathrm{H}_1(X_1(N)_{\mathbb{C}}, Z)$. Letting g be the genus of $X_1(N)_{\mathbb{C}}$, the Jacobian $J_1(N)$ is isomorphic to the g-dimension complex torus obtained by integration modulo homology as outlined in Appendix A.

By the methods outlined in Section 5.2, we can also identify $\mathrm{Pic}^0(X_1(N))$ with a subgroup of the complex Picard group $\mathrm{Pic}^0(X_1(N)_{\mathbb{C}})$ which is isomorphic to the Jacobian by Abel's Theorem. Thus, there is an inclusion of p^n torsion

$$i_n : \mathrm{Pic}^0(X_1(N))[p^n] \to \mathrm{Pic}^0(X_1(N)_{\mathbb{C}})[p^n] \cong (\mathbb{Z}/p^n\mathbb{Z})^{2g}.$$

By Igusa's Theorem, $X_1(N)$ has good reduction at primes p that don't divide N. Therefore we have restriction maps

$$\pi_n : \mathrm{Pic}^0(X_1(N))[p^n] \to \mathrm{Pic}^0(\tilde{X}_1(N))[p^n]$$

.

Definition 5.13. The **p-adic Tate module of** $X_1(N)$ is the limit

$$\mathrm{Ta}_p(\mathrm{Pic}^0(X_1(N))) = \varprojlim_n \{\mathrm{Pic}^0(X_1(N))[p^n]\}.$$

Similarly to the case of elliptic curves, by choosing a compatible basis gives

$$\mathrm{Ta}_p(\mathrm{Pic}^0(X_1(N))) \cong \mathbb{Z}_p^{2g}.$$

Any automorphism $\sigma \in G_{\mathbb{Q}}$ defines an automorphism on degree-0 divisors on $X_1(N)$ by

$$\sum n_P P \mapsto \sum n_P \sigma(P).$$

For any $f \in \overline{Q}(X_1(N))$, we have equality of divisors $(\sigma(f)) = \sigma((f))$, so σ descends to a map on $\mathrm{Pic}^0(X_1(N))$. Since the extension $\mathbb{Q}(\mathrm{Pic}^0(X_1(N))[p^n]$ is Galois over $\mathbb{Q}$ for all positive integers n, the action

$$\mathrm{Pic}^0(X_1(N)) \times G_{\mathbb{Q}}\& \to \mathrm{Pic}^0(X_1(N))$$

$$(\sum n_P P, \sigma) \mapsto \sum n_P \sigma(P)$$

restricts to $\mathrm{Pic}^0(X_1(N))[p^n]$. For each n, the diagram

$$\begin{tikzcd}
 & G_{\mathbb{Q}} \arrow[dl] \arrow[dr] & \\
\mathrm{Aut}(\mathrm{Pic}^0(X_1(N))[p^n]) & & \mathrm{Aut}(\mathrm{Pic}^0(X_1(N))[p^{n+1}]) \arrow[ll]
\end{tikzcd}$$

commutes, and so the action of $G_{\mathbb{Q}}$ gives a continuous homomorphism

$$\rho_{X_1(N),p} : G_{\mathbb{Q}} \to \mathrm{GL}_{2g}(\mathbb{Z}_p).$$

Again we compose with the inclusion $\mathrm{GL}_2(\mathbb{Z}_p) \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_p)$ to obtain a $2g$-dimensional Galois representation.

Now let $f \in \mathcal{S}_2(N, \chi)$ be a normalised eigenform and recall the isomorphism $\mathbb{T}_{\mathbb{Z}}/I_f \xrightarrow{\sim} \mathbb{Z}[\{a_n(f)\}]$. Under this isomorphism, the Fourier coefficient $a_p(f)$ acts on the abelian variety A_f as $T_p + I_f$, and $\chi(p)$ acts on A_f as $\langle p \rangle + I_f$. The ring $\mathbb{Z}[\{a_n(f)\}]$ generates the number field K_f of f. The abelian variety also has an p-adic Tate module.

Definition 5.14. The **p-adic Tate module of** A_f is the limit

$$\mathrm{Ta}_p(A_f) = \varprojlim_n \{A_f[p^n]\} \cong \mathbb{Z}_p^{2d}.$$

The action of $\mathbb{Z}[\{a_n(f)\}]$ on A_f extends to an action on $\mathrm{Ta}_p(A_f)$.

Lemma 5.15. *The map* $\mathrm{Pic}^0(X_1(N))[p^n] \to A_f[p^n]$ *is surjective, and its kernel is stable under* $G_{\mathbb{Q}}$.

Proof. See Lemma 9.5.2. in [7]. □

This shows that $G_{\mathbb{Q}}$ also acts on $A_f[p^n]$, and therefore it also acts on $\mathrm{Ta}_p(A_f)$. Choosing compatible coordinates gives a Galois representation

$$\rho_{A_f,p} : G_{\mathbb{Q}} \to \mathrm{GL}_{2d}(\mathbb{Q}_p).$$

Let $V_p(A_f)$ be the tensor product $\mathrm{Ta}_p(A_f) \otimes \mathbb{Q}$. This is a module over $K_f \otimes_{\mathbb{Q}} \mathbb{Q}_p$.

Lemma 5.16. $V_p(A_f)$ *is a free module of rank* 2 *over* $K_f \otimes_{\mathbb{Q}} \mathbb{Q}_p$.

Proof. See Lemma 9.5.3. in [7]. □

$G_{\mathbb{Q}}$ acts $K_f \otimes_{\mathbb{Q}} \mathbb{Q}_p$-linearly on $V_p(A_f)$, and by the above lemma, $V_p(A_f) \cong (K_f \otimes_{\mathbb{Q}} \mathbb{Q}_p)^2$. Choosing a basis for $V_p(A_f)$ gives a homomorphism $G_{\mathbb{Q}} \to \mathrm{GL}_2(K_f \otimes_{\mathbb{Q}} \mathbb{Q}_p)$. Moreover, by the isomorphism 5.1 we can consider the decomposition

$$K_f \otimes_{\mathbb{Q}} \mathbb{Q}_p \cong \prod_{\lambda | p} K_{f,\lambda}.$$

For each λ lying over (p) we can compose the homomorphism with projection to obtain Galois representations

$$\rho_{f,\lambda} : G_{\mathbb{Q}} \to \mathrm{GL}_2(K_{f,\lambda}).$$

Definition 5.17. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Q}_p)$ be an irreducible Galois representation such that $\det(\rho) = \chi_p$. Then ρ is **modular** is there exists a newform $f \in \mathcal{S}_2(\Gamma_0(M_f))$ such that $K_{f,\lambda} = \mathbb{Q}_p$ for some maximal ideal λ of $\mathcal{O}_{K_f}$ lying over p such that $\rho_{f,\lambda} \sim \rho$.

In particular, consider representations of the form $\rho_{E,p}$ for an elliptic curve E. The Weil pairing shows that the action of an element of the absolute Galois group on the root of unity μ_{p^n} is given by the determinant. We also know that the action raises μ_{p^n} to the nth entry of $\chi_p(\sigma)$, so $\det(\rho_{E,p}) = \chi_p$. Thus, the representations arising from elliptic curves are obvious candidates to consider for modularity. In fact, there is the following version of the Modularity Theorem.

Theorem 5.18. *(Modularity Theorem, version R). Let E be an elliptic curve over $\mathbb{Q}$. Then $\rho_{E,p}$ is modular for some p.*

This version of the Modularity Theorem was proved for semistable curves in 1995 by [31] and [27] to prove Fermat's Last Theorem, and proved in generality by [3].

This in fact implies a stronger version of the Modularity Theorem.

Theorem 5.19. *(Modularity Theorem, strong version R). Let E be an elliptic curve over $\mathbb{Q}$ and let N be its conductor. Then there exists a newform $f \in \mathcal{S}_2(\Gamma_0(N))$ with number field $K_f = \mathbb{Q}$ and such that $\rho_{f,p} \sim \rho_{E,p}$.*

6 Modularity modulo prime powers

Modulo forms modulo a prime p are a central object of study in modern Arithmetic Geometry, perhaps most notable for its role in the proof of the Modularity Theorem. One of the natural ways to extend this theory is to study congruences of modular forms modulo a prime power p^m, as well as their associated Galois representations. Many additional difficulties arise in working over $\mathbb{Z}/p^{\mathbb{Z}}$ which for $p \geq 2$ is neither reduced not factorial.

6.1 Strong, weak, and dc-weak modularity

Throughout this chapter, let p be a fixed prime number and let N be a positive integer not divisible by p.

We begin by considering Galois representations modulo p^n. We must first formulate an appropriate notion of modularity modulo prime powers. To this end, we will introduce three notions of modularity due to Chen, Kiming, and Wiese [4]: "strong" modularity, "weak" modularity, and "dc-weak" modularity.

Let $f \in \mathcal{S}_k(Np^r, \chi)$ be a normalised cuspidal eigenform with Fourier expansion

$$f(\tau) = \sum_{n=1}^{\infty} a_n(f) q^n, \quad q = e^{2\pi i \tau}.$$

Due to Shimura, Deligne, and Serre [5] [6] [23], there exists a continuous Galois representation

$$\rho_{f,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Q}}_p)$$

attached to f which is unramified outside Np. Moreover,

$$\operatorname{tr} \rho(\mathrm{Frob}_p) = a_p(f) \quad \text{and} \quad \det \rho(\mathrm{Frob}_p) = p^{k-1}\chi(p)$$

for all primes $p \nmid Np$.

For a finite extension over K of $\mathbb{Q}_p$ and Λ an $\mathcal{O}_K$-lattice in K^2, by continuity and compactness $\rho_{f,p}$ descends to a representation

$$\rho_{f,\Lambda,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathcal{O}_K).$$

Let $\mathfrak{p}_K$ denote the maximal ideal of the ring of $\mathcal{O}_K$. Let π_K denote a uniformizer of $\mathcal{O}_K$ and let v_K be the valuation satisfying $v_K(\pi_K) = 1$. We would like to study the reduction of the representation $\rho_{f,\Lambda,p}$ modulo $\mathfrak{p}_K^n$. However, due to ramification, this is problematic as the exponent n is not invariant under base extension. That is, if we want two compare two eigenforms, one with coefficients in $\mathcal{O}_K/\mathfrak{p}_K^n$ and another with coefficients in $\mathcal{O}_K/\mathfrak{p}_K^{m'}$ then we need to work in a ring that contains both. For such a ring to exist, $\mathfrak{p}_K^n \cap \mathbb{Z}_p$ and $\mathfrak{p}_K^{n'} \cap \mathbb{Z}_p$ must both yield the same power of p. In order to deal with this, Wiese and Taixés i Ventosa [26] introduce the ring $\overline{\mathbb{Z}/p^n\mathbb{Z}}$ to be defined below.

Definition 6.1. Let n be a positive integer and let $a, b \in \overline{\mathbb{Z}_p}$. We say that a **is congruent to** b **modulo** p^n if $v_p(a-b) > n-1$. We write $a \equiv b \pmod{p^n}$.

Definition 6.2. Let $L/K/\mathbb{Q}_p$ be finite field extension and let n be a positive integer. Let

$$\gamma_{L/K}(n) = (n-1)e_{L/K} + 1$$

where $e_{L/K}$ is the ramification index of the extension L/K.

This function satisfies the following properties:

i. For $n = 1$, $\gamma_{L/K}(1) = 1$.

ii. If L/K is unramified then $\gamma_{L/K}(n) = n - 1 + 1 = n$.

iii. If we have finite extensions $M/L/K/\mathbb{Q}_p$, then $\gamma_{M/K}(n) = \gamma_{M/L}(\gamma_{L/K}(n))$.

iv. If we have finite extensions $L/K/\mathbb{Q}_p$ then $\gamma_{L/K}(n)$ is the smallest integer such that the embedding $\mathcal{O}_K \hookrightarrow \mathcal{O}_L$ induces an injective map $\mathcal{O}_K/(\pi_K^n) \hookrightarrow \mathcal{O}_L/(\pi_L^{\gamma_{L/K}(n)})$.

v. If $a, b \in K/\mathbb{Q}_p$ then

$$v_K(a-b) \geq \gamma_{K/\mathbb{Q}_p}(n) \iff v_p(a-b) > n-1 \iff a \equiv b \pmod{p^n}.$$

By property iv. we have ring extensions

$$\mathbb{Z}/p^n\mathbb{Z} \hookrightarrow \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_p}(n)}) \hookrightarrow \mathcal{O}_L/(\pi_L^{\gamma_{L/\mathbb{Q}_p}(n)}),$$

and by property v. we can view congruence modulo p^n as equality in the residue ring $\mathcal{O}_K/((\pi_K^{\gamma_{K/\mathbb{Q}_p}(n)}$. We define

$$\overline{\mathbb{Z}/p^n\mathbb{Z}} = \varinjlim_{K} \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_p}(n)})$$

where K runs through all subextensions $\mathbb{Q}_p \subset K \subset \overline{\mathbb{Q}_p}$ of finite degree over $\mathbb{Q}_p$. We always consider $\overline{\mathbb{Z}/p^n\mathbb{Z}}$ with the discrete topology.

The projections

$$\mathcal{O}_K \to \mathcal{O}_K/(\pi_K^{\gamma_{K/\mathbb{Q}_p}(n)})$$

induce surjective ring homomorphisms

$$\pi_n : \overline{\mathbb{Z}_p} \to \overline{\mathbb{Z}/p^n\mathbb{Z}}.$$

Thus we can also view congruence modulo p^n as equality in $\overline{\mathbb{Z}/p^n\mathbb{Z}}$.

Definition 6.3. A representation $G \to \mathrm{GL}_n(F)$ is **absolutely irreducible** if it is irreducible over the algebraic closure of F.

We define the reductions

$$\rho_{f,\Lambda,p,m} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Z}/p^m\mathbb{Z}})$$

for any positive integer m. By the Tchebotarov Density Theorem, a continuous Galois representation $\rho_{p,m} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{Z}/p^m\mathbb{Z}})$ is determined uniquely up to isomorphism by $\operatorname{tr} \rho_{p,m}(\mathrm{Frob}_p$ for all but finitely many primes p if the residual representation is absolutely irreducible. Thus, if there is only one choice of $\mathcal{O}_K$-lattice Λ such that $\rho_{f,\Lambda,p,1}$ is absolutely irreducible, then $\rho_{f,\Lambda,p,m}$ is determined uniquely up to isomorphism. In such a case, we denote the representation by $\rho_{f,p,m}$ and we say that $\rho_{f,p,m}$ is the **mod p^m Galois representation attached to f**.

Let $\mathbb{T}_k(\Gamma)$ be the **full Hecke algebra** generated as a ring by all the Hecke operators acting faithfully on $\mathcal{S}_k(\Gamma)$.

Definition 6.4. Let K be a number field. An **order** of K is a subring $\mathcal{O} \subset \mathcal{O}_K$ which is also a $\mathbb{Z}$-module of rank $[K : \mathbb{Q}]$.

The following definitions are due to Chen, Kiming, and Wiese [4] but the form they are stated in is from [28].

Definition 6.5. Let R be a ring.

A **weak** Hecke eigenform of weight k with respect to Γ with coefficients in R is a ring homomorphism $f : \mathbb{T}_k(\Gamma) \to R$.

f is called **strong** if there exists and order $\mathcal{O}$ in a number field and a ring homomorphism $\pi : \mathcal{O} \to R$ such that f factors as $\mathbb{T}_k(\Gamma) \to \mathcal{O} \xrightarrow{\pi} R$.

Embedding the order $\mathcal{O}$ into $\mathbb{C}$ gives an eigenform $\mathbb{T}_k(\Gamma) \to \mathbb{C}$ in the classical sense defined in Chapter 1. Thus, the strong Hecke eigenforms are obtained by applying π to the coefficients of an eigenform.

Definition 6.6. Let R be a ring. Let $S_{\leq b}(\Gamma) = \bigoplus_{k=1}^{b} \mathcal{S}_k(\Gamma)$ and let $\mathbb{T}_{\leq b}(\Gamma)$ be the full Hecke algebra acting faithfully on $S_{\leq b}(\Gamma)$. A ring homomorphism $f : \mathbb{T}_{\leq b}(\Gamma) \to R$ is called a **dc-weak eigenform with respect to Γ of weights $\leq b$**.

Remark. If R is a finite field or if $R = \overline{F}_p$, then the Deligne-Serre lifting lemma implies that the three notions coincide [28].

A Hecke eigenform with coefficients in $\overline{\mathbb{Z}/p^m\mathbb{Z}}$ is called a **modulo p^m Hecke eigenform**.

6.2 Level raising and level lowering

In this section we review generalisations of the level raising and level lowering theorems from modulo p to modulo p^n.

The level raising theorem due to Ribet [20] is:

Theorem 6.7. *Let ρ be a modular Galois representation of level N. Let $p \nmid pN$ be a prime satisfying at least one of*

$$\operatorname{tr} \rho(\mathrm{Frob}_p) \equiv pm(p+1) \pmod{p}.$$

Then ρ is a newform of level pN.

Let $f \in \mathcal{S}_2(\Gamma_0(N))$. The simplest generalisation of level raising modulo p^n is the problem of identifying primes $p \nmid N$ such that there exists a newform $g \in \mathcal{S}_2(\Gamma_0(Np))$ such that f and g are congruent modulo p^n. By congruence, we mean that we have congruence of Fourier coefficients $a_m(f) \equiv a_m(g) \pmod{p^n}$ for all m.

Theorem 3.1 of [28] gives the following generalisation.

Theorem 6.8. *Let R be a local topological ring with maximal ideal $\mathfrak{m}$. Let*

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(R)$$

be a modular Galois representation, associated with a weak eigenform

$$f : \mathbb{T}_2(N) \to R.$$

Suppose that the residual representation

$$\overline{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(R/\mathfrak{m})$$

is absolutely irreducible. Moreover, if $\mathrm{char}(R/\mathfrak{m}) = 2$ then further assume that $\overline{\rho}$ is unramified at 2 with scalar Frobenius. Let $p \nmid N$ be prime be such that ρ is unramified at p and

$$\mathrm{tr}\ \rho(\mathrm{Frob}_p) = \pm(p+1).$$

Then the image of θ is a finite ring, $R/\mathfrak{m}$ is a finite field, and ρ is associated with a weak eigenform

$$\theta' : \mathbb{T}(Np) \to R$$

which is a newform at p.

For level lowering, we again first look at lowering mod p. The main theorem, also due to Ribet, implies that Fermat's Last Theorem follows from the Taniyama-Shimura-Weil Conjecture [18].

Theorem 6.9. *(Ribet). Let ρ be an irreducible two-dimensional representation of $G_{\mathbb{Q}}$ over a finite field of characteristic $c > 2$. Suppose that ρ is modular of level N where N is square-free, and that there is a prime $q|N$, $q \neq c$ at which ρ is not finite. The ρ is modular of level N/p.*

The following generalisation is due to Dummigan [8].

Theorem 6.10. *(Dummigan). Let p be a prime and let $p + 2 > k \geq 2$. Let $q \nmid N$ be a prime such that p is not congruent to 1 modulo q. Let $f \in \mathcal{S}_k(\Gamma_1(Np))$ be an eigenform with let λ be a prime of the coefficient field of f above q. Further assume that the residual Galois representation of f modulo λ is irreducible. The if for some positive integer m the Galois representation of f modulo λ^m is unramified at p, then there exists a weak eigenform g of weight k with respect to $\Gamma_1(N)$ such that $a_m(f) = a_m(g)$ for all m coprime to p.*

A Jacobians and Abel's Theorem

This section assumes a foundational understanding of Riemann surface theory although we will review some of the main ideas that we will need. For an introductory text on the subject, see [1] or [9].

Let X be a compact Riemann surface of genus $g \geq 1$ and let $x_0 \in X$. Consider the map that sends a point $x \in X$ to the function of holomorphic differentials $\omega \in \Omega^1_{\text{hom}}(X)$ given by

$$\omega \mapsto \int_{x_0}^{x} \omega.$$

Of course, different paths from x_0 to x might give different values for the integral, but any two such paths differ by a loop obtained by travelling forwards along one path and then backwards along the other. Thus, for this map to be well-defined we will need to quotient away integration over loops in X. We formalise this idea below.

Definition A.1. Let X be a compact Riemann surface. A (Weil) **divisor** D on X is a finite formal sum with integer coefficients of the form

$$D = \sum_{x \in X} n_x x, \quad n_x \in \mathbb{Z}$$

for which $n_x = 0$ for all but finitely many elements $x \in X$. The **set of divisors on** X is denoted $\text{Div}(X)$.

$\text{Div}(X)$ forms a group under addition given by

$$\sum_{x \in X} n_x x + \sum_{x \in X} n'_x x = \sum_{x \in X} (n_x + n'_x) x.$$

This group is clearly abelian and generated by the elements of X with no additional relations, so $\text{Div}(X)$ is in fact the free abelian group on the points of X.

Definition A.2. The **degree** of a divisor is

$$\deg(D) = \sum_{x \in X} n_x.$$

The subgroup of degree 0 divisors is denoted

$$\text{Div}^0(X) = \{D \in \text{Div}(X) : \deg(D) = 0\}.$$

The map sending a divisor to its degree $\deg : \text{Div}(X) \to \mathbb{Z}$ is a group homomorphism since $\deg\left(\sum_{x \in X} n_x x\right) + \deg\left(\sum_{x \in X} n'_x x\right) = \deg\left(\sum_{x \in X} (n_x + n'_x) x\right) = (n_x + n'_x)$.

Let $\mathbf{M}(X)$ denote the field of meromorphic functions on X and let $\mathbf{M}(X)^\times$ denote the multiplicative group of nonzero meromorphic functions. For every meromorphic function $f \in \mathbf{M}(X)$ and $x \in X$ define

$$\text{ord}_x(f) = \begin{cases} k, & \text{if } f \text{ has a zero of order } k \text{ at } x \\ -k, & \text{if } f \text{ has a pole of order } k \text{ at } x \\ \infty, & \text{if } f \text{ is identically zero in a neighbourhood of } x \\ 0, & \text{otherwise.} \end{cases}$$

To each nonzero meromorphic function $f \in \mathbf{M}(X)$ we associate a divisor

$$\operatorname{div}(f) = \sum_{x \in X} \operatorname{ord}_x(f)x.$$

The map sending nonzero meromorphic functions to their corresponding divisors

$$\operatorname{div} : \mathbf{M}(X)^\times \to \operatorname{Div}(X)$$

is a group homomorphism. This can easily be seen by considering the sums of their Laurent series about each point.

Definition A.3. Let X be a compact Riemann surface and let $D_1, D_2 \in \operatorname{Div}(X)$. The D_1 and D_2 are **linearly equivalent** if there exists a meromorphic function f on X such that $\operatorname{div}(f) = D_1 - D_2$.

The divisors of meromorphic functions are called **principal**.

It is easy to see that linear equivalence defines an equivalence relation. Every divisor is linearly equivalent to itself by letting f be a nonzero constant map. If $\operatorname{div}(f) = D_1 - D_2$ then $\operatorname{div}(-f) = D_2 - D_1$, so we have symmetry. Finally if $\operatorname{div}(f) = D_1 - D_2$ and $\operatorname{div}(f') = D_2 - D_3$ then $\operatorname{div}(ff') = D_1 - D_3$ so we also have transitivity.

Proposition A.4. *Let X be a compact Riemann surface and let f be a meromorphic function on X. The* $\deg(f) = 0$.

Proof. Let $\omega = df/f$. Then ω is holomorphic everywhere except at the zeros and poles of f. If $x \in X$ is a zero of order m then $\operatorname{Res}_x(\omega) = m$. On the other hand if $x \in X$ is a pole of order m then $\operatorname{Res}_x(\omega) = -m$. By the Residue theorem, the residues sum to 0 so $\deg(f) = 0$. □

It follows that the set of principal divisors is a subgroup of the group of degree 0 divisors. The subgroup of principal divisors is the set of all divisors linearly equivalent to 0, denoted

$$\operatorname{Div}^P(X) = \{D \in \operatorname{Div}^0(X) : D = \operatorname{div}(f) \text{ for some } f \in \mathbf{M}(X)\}.$$

Definition A.5. The (full) **Picard group of** X is the quotient group

$$\operatorname{Pic}(X) = \operatorname{Div}(X)/\operatorname{Div}^P(X).$$

The **degree-**0 **Picard group of** X is the quotient group

$$\operatorname{Pic}^0(X) = \operatorname{Div}^0(X)/\operatorname{Div}^P(X).$$

Since $\operatorname{Div}(X)/\operatorname{Div}^0(X) \cong \mathbb{Z}$, there is a short exact sequence

$$0 \longrightarrow \operatorname{Pic}^0(X) \longrightarrow \operatorname{Pic}(X) \longrightarrow \mathbb{Z} \longrightarrow 0.$$

The degree-0 Picard group captures the extent to which degree-0 divisors fail to be divisors of meromorphic functions on X. If $g \geq 1$ and we fix a basepoint $x_0 \in X$ then there is an embedding $X \to \operatorname{Pic}^0(X)$ defined by

$$x \mapsto [x - x_0]$$

where $[x - x_0]$ denotes the linear equivalence class $x - x_0 + \mathrm{Div}^P(X)$ and the subtraction is of divisors in $\mathrm{Div}^0(X)$.

Now we want to construct maps from the degree-0 Picard group into the set of linear functions of holomorphic differentials in X modulo integration on loops as we described at the beginning of this section. To do this, we will need to introduce the first homology group $\mathrm{H}_1(X, \mathbb{Z})$ of a compact Riemann surface X. Since we will only need the first homology group, we will only introduce the 1-chains, 1-cycles, and 1-boundaries.

A 1-chain on a Riemann surface X is a formal finite linear combination with integer coefficients

$$c = \sum_{i=1}^{n} n_i c_i, \quad n_i \in \mathbb{Z}$$

where the $c_i : [0,1] \to X$ are paths. Integration of a closed differential form ω over a 1-chain is defined by

$$\int_c \omega = \sum_{i=1}^{n} n_i \int_{c_j} \omega.$$

Let $C_1(X)$ denote the set of all 1-chains on X. $C_1(X)$ forms an abelian group with addition as the group operation.

We define a boundary operator

$$\partial : C_1(X) \to \mathrm{Div}(X)$$

as follows. Suppose $c : [0,1] \to X$ is a path. If $c(0) = c(1)$, then set $\partial(\gamma) = 0$. Otherwise let $\partial(c)$ be the divisor $c(1) - c(0)$. For an arbitrary 1-chain $c = \sum_{i=1}^{n} n_i c_i$, set $\partial(c) = \sum_{i=1}^{n} n_i \partial(c_i)$. It is immediate from this definition that $\mathrm{im}(\partial) \subset \mathrm{Div}^0(X)$.

The group of 1-cycles is

$$Z_1(X) = \ker(\partial)$$

and the 1-boundaries is the subgroup

$$B_1(X) = \{c \in Z_1(X) : \int_c \omega = 0 \text{ for all closed differential forms } \omega\}.$$

The first homology group of X is

$$H_1(X, \mathbb{Z}) = Z_1(X)/B_1(X).$$

Definition A.6. Two cycles are said to be **homologous** if they differ by a boundary.

Proposition A.7. *Let Y be a smooth k-manifold and let ω be a closed k-form on Z. Suppose $f, g : Y \to Z$ are homotopic maps. Then $\int_Y f^*\omega = \int_Y g^*\omega$.*

Thus if two loops $c_1, c_2 : \mathbb{S}^1 \to X$ are homotopic, then $c_1 - c_2 \in B_1(X)$. That is, c_1 and c_2 are homologous.

To work with the homology group, we will want to fix a basis. View X as a sphere with g handles. Let $a_1, \dots, a_g$ be longitudinal loops and $b_1, \dots, b_g$ be latitudinal loops around each handle. Then $\mathrm{H}_1(X, \mathbb{Z})$ is generated by the a_i and b_j and we have $\mathrm{H}_1(X, \mathbb{Z}) \cong \mathbb{Z}^{2g}$.

Via isomorphism we can view the first homology group of X as

$$\mathrm{H}_1(X,\mathbb{Z}) \cong \mathbb{Z}\int_{a_1} \oplus \cdots \oplus \mathbb{Z}\int_{a_g} \oplus \mathbb{Z}\int_{b_1} \oplus \cdots \oplus \mathbb{Z}\int_{b_g}$$

and consider $\mathrm{H}_1(X,\mathbb{Z})$ as a subset of the dual space $\Omega^1_{\mathrm{hol}}(X)^\vee = \mathrm{Hom}_{\mathbb{C}}(\Omega^1_{\mathrm{hol}}(X),\mathbb{C})$, the set of $\mathbb{C}$-linear maps from the set of holomorphic differentials on X to $\mathbb{C}$. We will use without proof that

$$\Omega^1_{\mathrm{hol}}(X)^\vee \cong \mathbb{R}\int_{a_1} \oplus \cdots \oplus \mathbb{R}\int_{a_g} \oplus \mathbb{R}\int_{b_1} \oplus \cdots \oplus \mathbb{R}\int_{b_g}.$$

Since $\dim_{\mathbb{C}}(\Omega^1_{\mathrm{hol}}(X) = g$, it follows from thh Riemann bilinear relations that the first homology group $\mathrm{H}_1(X,\mathbb{Z})$ is a non-degenerate lattice in $\Omega^1_{\mathrm{hol}}(X)^\vee \cong \mathbb{R}^{2g} \cong \mathbb{C}^g$.

Definition A.8. The **Jacobian** of X is the quotient group

$$\mathrm{Jac}(X) = \Omega^1_{\mathrm{hol}}(X)^\vee / \mathrm{H}_1(X,\mathbb{Z}).$$

Remark. The Jacobian of X is the connected component of the identity in $\mathrm{Pic}(X)$ and is therefore an abelian variety.

Definition A.9. Fix any basepoint $x_0 \in X$. Then the **Abel-Jacobi map** is

$$\begin{aligned} u : X &\to \mathrm{Jac}(X) \\ x &\mapsto \int_{x_0}^{x} \end{aligned}$$

where by abuse of notation $\int_{x_0}^{x}$ denotes its equivalence class in $\mathrm{Jac}(X)$.

The Abel-Jacobi map can be extended linearly to $\mathrm{Div}(X)$ setting

$$u\left(\sum_{x\in X} n_x x\right) = \sum_{x\in X} n_x \int_{x_0}^{x}.$$

If we restrict to $\mathrm{Div}^0(X)$ then u is independent of choice of basepoint x_0 by the following proposition.

Proposition A.10. *Let u be the Abel-Jacobi map with respect to $x_0 \in X$. Let u' be the Abel-Jacobi map with respect to $x_0' \in X$. Let $D \in \mathrm{Div}^0(X)$. Then $u(D) = u'(D)$.*

Proof. If $D \in \mathrm{Div}^0(X)$ then we can write D as the sum $D = \sum_{i=1}^n y_i - \sum_{i=1}^n z_i$ where the y_i and z_i are points in X. Therefore by linearity it suffices to show that $u(y-z) = u'(y-z)$ for any $y, z \in X$. Evaluating the Abel-Jacobi maps we have

$$\begin{aligned} u(y-z) - u'(y-z) &= \left(\int_{x_0}^{y} - \int_{x_0}^{z}\right) - \left(\int_{x_0'}^{y} - \int_{x_0'}^{z}\right) \\ &= \int_{x_0}^{y} + \int_{y}^{x_0'} + \int_{x_0'}^{z} + \int_{z}^{x_0}. \end{aligned}$$

The sum of the integrals above can be reduced to a single integral around a closed path, i.e. a loop. Therefore $u(y-z) - u'(y-z) \in \mathrm{H}_1(X,\mathbb{Z})$ so it lies in the equivalence class of 0 in $\mathrm{Jac}(X)$. □

Corollary A.11. *Fix any basepoint $x_0 \in X$. The map $u : \mathrm{Div}^0(X) \to \mathrm{Jac}(X)$ given by*

$$\sum_{x \in X} n_x x \mapsto \sum_{x \in X} n_x \int_{x_0}^{x}$$

is well-defined.

Theorem A.12. *(Abel's Theorem).*

$$\ker(\mathrm{Div}^0(X) \xrightarrow{u} \mathrm{Jac}(X)) = \mathrm{Div}^P(X).$$

In particular, u induces an isomorphism

$$\mathrm{Pic}^0(X) \xrightarrow{\sim} \mathrm{Jac}(X)$$
$$\left[\sum_{x \in X} n_x x\right] \mapsto \sum_{x \in X} n_x \int_{x_0}^{x}.$$

Proof. See Theorem 21.7 of [9]. □

If $g \geq 1$ then the embedding of X in $\mathrm{Pic}(X)$ followed by the isomorphism of Abel's Theorem gives an embedding

$$X \to \mathrm{Jac}(X), \quad x \mapsto \int_{x_0}^{x}.$$

An immediate consequence of Abel's Theorem is that the 1-chains with finite nonzero coefficients that sum to 0 make up the entirety of $\Omega^1_{\mathrm{hol}}(X)^\vee$.

Index of Notation

A_f	the abelian variety associated to f
$E[N], \mathcal{E}[N]$	the N-torsion subgroup of E or $\mathcal{E}$
E_j	a universal elliptic curve
$G_{\mathbb{Q}}$	the absolute Galois group
G_k	the Eisenstein series of weight k
$H_{\mathbb{Q}}$	$\text{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j))$
$K_{N,d}$	the kernel of $\pi_{N,d}$
T_n	the Hecke operator, n any positive integer
T_p	the Hecke operator, p prime
$[\Gamma_1 \alpha \Gamma_2]_k$	the weight-k double coset operator
$[\gamma]_k$	the weight-k operator
Δ	the discriminant function
$\Delta_{\min}(E)$	the global minimal discriminant of E
$\Gamma(N)$	the principal congruence subgroup of level N
Γ	a principal congruence subgroup
$\Gamma_0(N)$	the congruence subgroup $\Gamma_0(N)$
$\Gamma_1(N)$	the congruence subgroup $\Gamma_1(N)$
$\mathbb{Q}_\ell$	the p-adic numbers
$\mathbb{Z}_\ell$	the p-adic integers
χ	a Dirichlet character modulo N
χ_p	the p-adic cyclotomic character
$\langle\ ,\ \rangle_\Gamma$	the Petersson inner product with respect to Γ
$\langle d \rangle, \langle n \rangle$	the diamond Hecke operator
$\mathbb{P}^n(k)$	the n-dimensional projective space over k
$\text{Div}(C)$	the divisor group of C
$\text{Div}^0(C)$	the degree-0 divisor group of C

$\mathrm{Div}^P(C)$	the principal divisors on C
$\overline{k}(C)$	the function field of C
$\overline{k}[C]$	the coordinate ring if C
$\overline{k}[C]_P$	the local ring of C at P
$\pi_{N,d}$	the natural projection $(\mathbb{Z}/N\mathbb{Z})^\times \to (\mathbb{Z}/d\mathbb{Z})^\times$
$a_n(E)$	the multiplicative extension of the modified solution count of an elliptic curve E
$a_n(f)$	the nth coefficient of f
c_4, c_6	Weierstrass coefficients
e_N	the Weil pairing
h^*	induced reverse map of Picard groups
h_*	induced forward map of Picard groups
$j(\gamma, \tau)$	the factor of automorphy
j	the modular invariant or the invariant of a Weierstrass equation

Sets of modular forms and cusp forms

$\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$	the set of modular forms of weight k
$\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$	the set of cusp forms of weight k
$\mathcal{M}_k(\Gamma)$	the set of modular forms of weight k with respect to Γ
$\mathcal{S}_k(\Gamma)$	the set of cusp forms of weight k with respect to Γ
$\mathcal{M}_k(N, \chi)$	the set of modular forms of weight k, character χ and level N
$\mathcal{S}_k(N, \chi)$	the set of cusp forms of weight k, character χ and level N

$\hat{\mathbb{C}}$	the extended complex plane
$\mathbb{H}$	the complex upper half-plane
$\mathbb{H}^*$	the extended upper half-plane
$\mathrm{SL}_2(\mathbb{Z})$	the special linear group of degree 2 over $\mathbb{Z}$
$\widehat{(\mathbb{Z}/N\mathbb{Z})}^\times$	the group of Dirichlet characters modulo N; the dual group of $(\mathbb{Z}/N\mathbb{Z})^\times$
$(\mathbb{Z}/N\mathbb{Z})^\times$	the multiplicative group of integers modulo N
$\mathbb{F}_q$	the field of q elements
$\overline{k}$	the algebraic closure of k

References

[1] Lars Ahlfors. *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*. McGraw-Hill, Inc., 1953.

[2] Michael Francis Atiyah and I. G. MacDonald. *Introduction to commutative algebra.* Addison-Wesley-Longman, 1969, pp. I–IX, 1–128. ISBN: 978-0-201-40751-8.

[3] Christophe Breuil et al. "On the modularity of ellptic curves over $\mathbb{Q}$: wild 3-adic exercises". In: *Journal of the American Mathematical Society* 14.4 (2001), pp. 843–939.

[4] Imin Chen, Ian Kiming, and Wiese Gabor. "On modular Galois representations modulo prime powers". In: *International Journal of Number Theory* 9.1 (2013), pp. 91–113.

[5] Pierre Deligne. *Formes modulaires et représentations ℓ-adiques*. fr. Séminaire Bourbaki. Lecture Notes in Math. Springer-Verlag, 1971.

[6] Pierre Deligne and Jean-Pierre Serre. "Formes modulaires de poids 1". fr. In: *Annales scientifiques de l'École Normale Supérieure* 4e série, 7.4 (1974), pp. 507–530. DOI: 10.24033/asens.1277. URL: http://www.numdam.org/articles/10.24033/asens.1277/.

[7] Fred Diamond and Jerry Shurman. *A First Course in Modular Forms*. Graduate Texts in Mathematics. Springer, 2005.

[8] Luis Dieulefait and Xavier Taixes i Ventosa. *Congruences between modular forms and lowering the level mod l^n*. 2008. DOI: 10.48550/ARXIV.0801.0104. URL: https://arxiv.org/abs/0801.0104.

[9] Otto Forster. *Lectures on Riemann Surfaces*. Springer New York, 1981.

[11] Robin Hartshorne. *Algebraic Geometry*. Graduate Texts in Mathematics. Springer Science+Business Media, Inc., 1977.

[12] Kenneth Ireland and Michael Rosen. *A Classical Introduction to Modern Number Theory*. Springer New York, 1990.

[13] Serge Lang. *Introduction to Modular Forms*. Springer-Verlag Berlin Heidelberg, 1987.

[14] J. Lehner and A.O.L. Atkin. "Hecke Operators on $\Gamma_0(m)$." In: *Mathematische Annalen* 185 (1970), pp. 134–160. URL: http://eudml.org/doc/161948.

[15] J.S. Milne. *Elliptic Curves*. BookSurge Publishers, 2006, pp. 238+viii. ISBN: 1-4196-5257-5.

[16] James S. Milne. *Algebraic Number Theory (v3.08)*. Available at www.jmilne.org/math/. 2020.

[17] K. Ranestad et al. *The 1-2-3 of Modular Forms: Lectures at a Summer School in Nordfjordeid, Norway*. Universitext. Springer Berlin Heidelberg, 2008.

[18] Kenneth A. Ribet. "From the Taniyama-Shimura conjecture to Fermat's last theorem". en. In: *Annales de la Faculté des sciences de Toulouse : Mathématiques* Ser. 5, 11.1 (1990), pp. 116–139. URL: http://www.numdam.org/item/AFST_1990_5_11_1_116_0/.

[19] Kenneth A. Ribet. "Galois representations and modular forms". In: (1995). DOI: 10.48550/ARXIV.MATH/9503219. URL: https://arxiv.org/abs/math/9503219.

[20] Kenneth A. Ribet. "Raising the Levels of Modular Representations". In: 1990.

[21] Pierre Samuel. *Algebraic Theory of Numbers*. Dover Publications, 2013.

[22] Jean-Pierre Serre. “A Course in Arithmetic”. In: *Graduate Texts in Mathematics* (1973), p. 26.

[23] Jean-Pierre Serre. “Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$”. In: *Duke Mathematical Journal* 54 (1987), pp. 179–230. English translation: Alex Ghitza. *On the two-dimensional modular representations of* $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. URL: `https://aghitza.org/publications/translation-serre-duke/`.

[24] Joseph H. Silverman. *The Arithmetic of Elliptic Curves.* Graduate Texts in Mathematics. Springer-Verlag New York Inc., 1986.

[25] P. Stevenhagen, H. W. Lenstra, and Jr. *Chebotarëv and his density theorem.* 1995.

[26] Xavier Taixés i Ventosa and Gabor Wiese. “Computing Congruences of Modular Forms and Galois Representations Modulo Prime Powers”. In: *Arithmetic, Geometry, Cryptography and Coding Theory 2009.* Ed. by David Kohel and Robert Rolland. Providence, RI: American Mathematical Society, 2010, pp. 145–166.

[27] Richard Taylor and Andrew Wiles. “Ring-Theoretic Properties of Certain Hecke Algebras”. In: *Annals of Mathematics* 141.3 (1995), pp. 553–572.

[28] Panagiotis Tsaknias and Gabor Wiese. “Topics on Modular Galois Representations Modulo Prime Powers”. In: *Algorithmic and Experimental Methods in Algebra, Geometry, and Number Theory.* Ed. by Gebhard Böckle and Gunter Malle. Springer International Publishing AG, 2017, pp. 741–763.

[29] André. Weil. *Foundations of Algebraic Geometry.* American Mathematical Society, 1946.

[30] Jared Weinstein. “Reciprocity laws and Galois representations: recent breakthroughs”. In: *Bulletin of the American Mathematical Society* 53 (2015), pp. 1–39.

[31] Andrew Wiles. “Modular Elliptic Curves and Fermat’s Last Theorem”. In: *Annals of Mathematics* 141.3 (1995), pp. 443–551.