MODULAR FORMS

SARAH FUJIMORI AND TAE KYU KIM

Abstract. Modular forms are complex analytic functions on the upper half plane that satisfy a condition with respect to the modular group. They are important for their connection to elliptic curves. In this paper, we define the modular group and modular forms, and introduce Eisenstein series and cusp forms. We then define congruence subgroups and explore the relationship between modular forms and elliptic curves over the complex numbers through their construction from the Weierstrass $\wp$-function to their classification using the j-invariant.

1. Introduction

Modular forms were first studied in connection with elliptic curves and elliptic integrals in the early 19th century, and they were later studied in the context of automorphic forms. Later, the famous Modularity Theorem, which related elliptic curves and modular forms, along with demands for development from other number theory encouraged further research in modular forms. A weaker variant of the Modularity Theorem proven by Andrew Wiles would be used to prove Fermat's Last Theorem in 1995, and the complete Modularity Theorem would be proven in 2001 by numerous mathematicians.

Modular forms can be thought of as having a multiplicative scaling behavior much like homogeneous functions: there is a constant k such that for all $\vec{x}$ in our domain and scalars α in a field, we have $f(\alpha\vec{x}) = \alpha^k f(\vec{x})$. While scalar multiplication will not be the transformation for modular forms, the defining identities of modular forms of weight k will look very similar. The transformations that modular forms satisfy are the "modular group" i.e. $\mathrm{SL}_2(\mathbb{Z})$ (or more generally, a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ with finite index). So the functions that transform under the modular group with a kind of scaling behavior are called modular forms.

Section 2 and 3 will define and provide important examples of modular forms, including the Eisenstein series and cusp forms. In section 4, we will introduce congruence subgroups. Section 5 discusses the relationship between complex elliptic curves and modular forms through their construction from the Weierstrass $\wp$-function to their classification using the j-invariant.

2. Modular forms

Before we define modular forms, we first introduce the modular group:

Definition 2.1. The *modular group* is the group of 2-by-2 matrices with integer entries and determinant 1,

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

The group operation is matrix multiplication. Each element in the modular group is a linear fractional transformation on the upper half plane $\mathcal{H} = \{\tau \in \mathbb{C} : \mathrm{Im}(\tau) > 0\}$. That is,

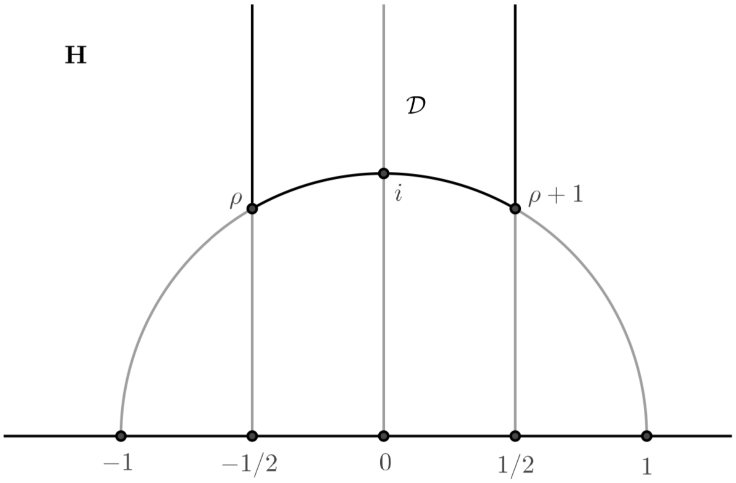


Figure 1. A Fundamental Domain

for an element $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have the map

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} (\tau) = \frac{a\tau + b}{c\tau + d}, \quad \tau \in \mathcal{H}\,.$$

We say that the modular group *acts* on the half plane $\mathcal{H}$. These transformations have natural group properties: having an identity transformation, having inverse transformations, and being associative. Any element in the modular group represents a transformation of the half plane, and the product of two matrices in the modular group represents the transformation that is obtained by consecutively applying the transformations represented by the two matrices.

Definition 2.2. For a group G acting on a space X, consider the images of a single point under the group action; these form **orbits** of the group action. A **fundamental domain** is a subset of X which contains exactly one point from this orbit.

Let

$$\mathcal{D} = \left\{ z \in \mathcal{H} : |z| \geq 1, \frac{-1}{2} \geq \mathrm{Re}(z) \leq \frac{1}{2} \right\}.$$

This region is pictured in Figure 1. We claim the following:

Theorem 2.3. *D is a fundamental domain for the modular group acting on the upper half plane.*

We first prove the following proposition, which will be helpful for the proof of Theorem 2.3:

Proposition 2.4. *Let $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an element of $\mathrm{SL}_2(\mathbb{Z})$. Then, for some $\tau \in \mathcal{H}$, we have*

$$\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau + d|^2}.$$

Proof. By the definition of the group action of $\mathrm{SL}_2(\mathbb{Z})$, we have

$$\mathrm{Im}(\gamma z) = \mathrm{Im}\left(\frac{a\tau + b}{c\tau + d}\right)$$

Multiplying numerator and denominator by $\overline{c\tau + d}$ to rationalize this quantity, this is equal to

$$\mathrm{Im}\left(\frac{(a\tau + b)(\overline{c\tau + d})}{(c\tau + d)(\overline{c\tau + d})}\right) = \frac{\mathrm{Im}(a\tau + b)(\overline{c\tau + d})}{|c\tau + d|^2}$$

Let $\tau = x + yi$. Since $ad - bc = 1$, the imaginary part of $(a\tau + b)(\overline{c\tau + d})$ is

$$\mathrm{Im}(ax + ayi + b)(cx + d - cyi) = -acxy + acxy + ady - bcy = (ad - bc)y = y,$$

so our expression simplifies to

$$\frac{\mathrm{Im}(\tau)}{|c\tau + d|^2}$$

as desired. □

We now prove Theorem 2.3 with a proof from [BD16].

Proof of Theorem 2.3. Let z be a point in $\mathcal{H}$. We show that we can choose $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$|\gamma z| \geq 1, \frac{-1}{2} \geq \mathrm{Re}(\gamma z) \leq \frac{1}{2}.$$

Define the two matrices S and T as

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Note that the transformations that S and T correspond to are

$$S(\tau) = \tau + 1, T(\tau) = -\frac{1}{\tau}$$

By Proposition 2.4, for some $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, we have $\mathrm{Im}(\gamma z) = \frac{\mathrm{Im}(z)}{|cz+d|^2}$. Since c and d are integers, the quantity $|cz + d|^2$ has a minimum, which then maximizes the value of $\mathrm{Im}(\gamma z)$.

Let γ correspond to this minimum, and multiply γ by powers of S such that γz satisfies $\frac{-1}{2} \geq \mathrm{Re}(\gamma z) \leq \frac{1}{2}$. Note that this will not change the value of $\mathrm{Im}\, z$, since the transformation corresponding to S only changes $\mathrm{Re}(z)$.

We now show that $\mathrm{Im}\, \gamma z \geq 1$. Assume for contradiction that $\mathrm{Im}\, \gamma z < 1$; then, by Proposition 2.4, $\mathrm{Im}\, T(\gamma z) = \frac{\mathrm{Im}(\gamma z)}{|\gamma z|^2} > \mathrm{Im}(\gamma z)$. Recall that γ was constructed to maximize $\mathrm{Im}(\gamma z)$, and we just showed that $T(\gamma z)$ has larger imaginary part. This is a contradiction, and we conclude that $\mathrm{Im}(\gamma z) \geq 1$.

We have now shown that $\mathrm{Im}(\gamma z) \geq 1$ and $\frac{-1}{2} \geq \mathrm{Re}(\gamma z) \leq \frac{1}{2}$, and it remains to show that no two points in the interior of D are $\mathrm{SL}_2(\mathbb{Z})$-equivalent. Suppose that two points z_1 and z_2 satisfy $z_2 = \gamma z_1$, where $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$; then, we show that they must lie on the boundary of D.

Without loss of generality, assume that $\mathrm{Im}(z_2) \geq \mathrm{Im}(z_1)$. Then, by Proposition 2.4, we know that $|cz_1 + d|^2 = (c\,\mathrm{Re}(z_1) + d)^2 + (c\,\mathrm{Im}(z_1))^2 \leq 1$. Since z_1 lies in D, and since c and d

are integers, we have the following four cases (Note that it is very helpful to look at Figure 1 to analyze these cases):

(1) $c = 0$, $d = \pm 1$. This means that either γ or $-\gamma$ is a translation, so either z_1 and z_2 lie on the boundary lines $\mathrm{Re}(z) = \pm\frac{1}{2}$, or $\pm\gamma$ is the identity map.
(2) $c = \pm 1, d = 0$. This means that either $\gamma = \pm T$ and z_1 and z_2 lie on the portion of the unit circle that is part of the boundary of D, or $\gamma = \pm S^{\pm 1} T$, and z_1 and z_2 are the two roots of unity.
(3) $c = d = \pm 1$; then, $z_1 = \rho$.
(4) $c = -d = \pm 1$; then, $z_1 = \rho + 1$.

Thus, we have shown that every point in $\mathcal{H}$ is equivalent under the action of $\mathrm{SL}_2(\mathbb{Z})$ to a point in D, and that no two points in the interior of D are equivalent, so we are done. □

An important consequence of Theorem 2.3 is the following:

Theorem 2.5. *The modular group* $\mathrm{SL}_2(\mathbb{Z})$ *is generated by the two matrices*

$$S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Proof. Let Γ' be the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by S and T, let g be an element of $\mathrm{SL}_2(\mathbb{Z})$, and let $z \in \mathcal{H}$. Then, by Theorem 2.3, there exists some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma(gz) \in D$. But then, since no two points in the interior are equivalent, $\gamma g = \pm I$, where I is the identity matrix. Thus, since γ can be written as a product of S and T, g can as well, and this concludes the proof. □

Before we define weakly modular functions, we will first give some background on complex analysis:

Definition 2.6. Let $f(z)$ be a complex-valued function of a complex variable z. Then, f is **holomorphic** at a point $z \in \mathbb{C}$ if it is differentiable in a neighborhood of z.

Example. Polynomial functions are always holomorphic on $\mathbb{C}$.

Example. Some other examples of holomorphic functions are the trigonometric functions and the complex exponential function.

The term "holomorphic" is often used interchangeably with "analytic", which means that a function has a convergent power series for a neighborhood of a point. These notions turn out to be equivalent; we refer the reader to [Sha03] for a more in-depth treatment of power series.

We can also define functions that are holomorphic on all of a subset of $\mathbb{C}$, except for a few isolated points. These functions are called meromorphic functions.

Definition 2.7. A function f is **meromorphic** if it can be written as the ratio of two holomorphic functions.

Example. All rational functions are meromorphic on the complex plane

We now define weakly modular functions:

Definition 2.8. Let k be an integer. A meromorphic function $f : \mathcal{H} \to \mathbb{C}$ is **weakly modular of weight** k if for all $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and all $\tau \in \mathcal{H}$,

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

Since $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, which correspond to the automorphisms $\tau \mapsto \tau + 1$ and $\tau \mapsto -\frac{1}{\tau}$, this property is equivalent to the two functional equations

$$f(\tau + 1) = f(\tau), f\left(-\frac{1}{\tau}\right) = \tau^k f(\tau).$$

It is also important to note that modular forms are sometimes defined as having weight $2k$ instead of weight k. This is because of the following:

Proposition 2.9. *If k is odd, then the only function that can satisfy this condition is the zero function.*

Proof. Suppose k is odd. Consider the matrix $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, which corresponds to the transformation $\tau \mapsto \tau$. Then, f must satisfy $f(\tau) = (-1)^k f(\tau)$, so since $(-1)^k = -1$, $f(\tau)$ must be identically 0. □

One of the conditions that a modular form satisfies is being holomorphic at ∞, which we define below:

Definition 2.10. A function $f : \mathcal{H} \to \mathbb{C}$ is **holomorphic at** ∞ if $f(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$.

We can now define modular forms of weight k:

Definition 2.11. A function $f : \mathcal{H} \to \mathbb{C}$ is a **modular form of weight** k if

(1) f is holomorphic on $\mathcal{H}$,
(2) f is holomorphic at ∞,
(3) f is weakly modular of weight k.

The set of modular forms of weight k forms a ring, and it is denoted $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. Furthermore, we can add and multiply modular forms:

Proposition 2.12. *We have the following properties:*

(1) *If $f(\tau)$ and $g(\tau)$ are both modular forms of weight k, then the sum $f + g$ is also a modular form of weight k.*
(2) *If $f(\tau)$ and $g(\tau)$ are modular forms of weight l and m, respectively, then the product $(fg)(\tau)$ is a modular form of weight $l + m$.*

Proof. (1) Suppose $f(\tau)$ and $g(\tau)$ are both modular forms of weight k. Then, clearly the sum $f + g$ is holomorphic on $\mathcal{H}$ and at ∞. To show weak modularity, note that for any element $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of $\mathrm{SL}_2(\mathbb{Z})$,

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

so we conclude that $f + g$ is a modular form of weight k.

(2) Suppose $f(\tau)$ and $g(\tau)$ are modular forms of weight l and m. Then, clearly the product fg is holomorphic on $\mathcal{H}$ and at ∞. Additionally, for some element $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ of $\mathrm{SL}_2(\mathbb{Z})$,

$$fg(\gamma(\tau)) = \Big((c\tau + d)^l f(\tau)\Big)\left((c\tau + d)^m g(\tau)\right) = (c\tau + d)^{l+m} fg(\tau)$$

so we conclude that fg is a modular form of weight $l+m$.

□

Thus, the sum

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

has an additional structure due to this property. This type of ring is called a **graded ring**.

Notice that due a consequence of being weakly modular of weight k, modular forms are 1-periodic, which means that they have a Fourier series expansion

$$\sum_{n=0}^{\infty} a_n q^n$$

where $q = e^{2\pi i\tau}$.

3. Eisenstein Series and Cusp Forms

Eisenstein series can be thought of as an analog of the Riemann zeta function in two dimensional lattices, and it turns out that they are modular forms. They are defined as follows:

Definition 3.1. Let k be a positive even integer. The **Eisenstein series of weight** k, $G_k(\tau) : \mathcal{H} \to \mathbb{C}$, is given by the series

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

Proposition 3.2. *$G_k(\tau)$ is a modular form of weight k.*

Proof. To see that the Eisenstein series is holomorphic on $\mathcal{H}$, take partial sums, and notice that those functions are holomorphic. Then, taking the limit as the number of terms approaches infinity yields a holomorphic function as well.

To prove weak modularity, we prove that $G_k(\tau)$ satisfies

$$G_k(\tau+1) = G_k(\tau), G_k\left(-\frac{1}{\tau}\right) = \tau^k G_k(\tau).$$

For the first property, we have

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

Note that as c and d run over the integers, c and $c+d$ do as well, so this is equal to $G_k(\tau)$, as desired.

For the second property,

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

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

Since d and $-c$ run over the integers as c and d do, this is equal to $\tau^k G_k(\tau)$ as desired. □

Proposition 3.3. *For an even integer $k > 2$, the Fourier series expansion of the Eisenstein series is*

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

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

Proof. We present the following proof from [DS05].

The following can be shown with standard complex analysis techniques, but we will omit it because we are not assuming knowledge of complex analysis. This is the cotangent sum:

$$\pi\cot\pi\tau = \frac{1}{\tau} + \sum_{n=1}^{\infty}\left[\frac{1}{\tau+n} + \frac{1}{\tau-n}\right] = \pi i - 2\pi i\sum_{m=0}^{\infty} q^m$$

where $q = e^{2\pi i\tau}$. Notice that differentiating the last two parts of the identity yields

$$-\frac{1}{\tau^2} + \sum_{n=1}^{\infty}\frac{-1}{(\tau+n)^2} + \frac{-1}{(\tau-n)^2} = (-2\pi i)(2\pi i)\sum_{m=0}^{\infty} mq^m$$

Differentiating again,

$$\frac{1}{\tau^3} + \sum_{n=1}^{\infty}\frac{2}{(\tau+n)^3)} + \frac{2}{(\tau+n)^3)} = -(2\pi i)^3\sum_{m=0}^{\infty} m^2q^m$$

We repeat this process to differentiate $k-1$ times, so that the exponent of the fractions on the left hand side is k. This results in the following identity:

$$\sum_{n\in\mathbb{Z}} -\frac{(k-1)!}{(\tau+n)^k} = -(2\pi i)^k\sum_{m=0}^{\infty} m^{k-1}q^m$$

Rearranging, we get that

$$\sum_{n\in\mathbb{Z}}\frac{1}{(\tau+n)^k} = \frac{(2\pi i)^k}{(k-1)!}\sum_{m=0}^{\infty} m^{k-1}q^m$$

Now note that for even integers $k > 2$, we can write $G_k(\tau)$ as

$$\sum_{(c,d)\in\mathbb{Z}\setminus\{(0,0)\}}\frac{1}{(c\tau+d)^k} = \sum_{d\neq 0}\frac{1}{d^k} + 2\sum_{c=1}^{\infty}\sum_{d\in\mathbb{Z}}\frac{1}{(c\tau+d)^k}$$

where the first sum is the case $c = 0$, and the second sum is multiplied by a factor of 2 because it covers the cases where c is positive or negative.

The quantity $\sum_{d\neq 0}\frac{1}{d^k}$ is equal to $2\zeta(k)$ because k is even, and for the double summation, we use the identity we derived above from the cotangent sum:

$$= 2\zeta(k) + 2\sum_{c=1}^{\infty}\left(\frac{(2\pi i)^k}{(k-1)!}\sum_{m=0}^{\infty} m^{k-1}q^{cm}\right)$$

We change the indexing of the double sums by summing $n = cm$ over integers and then summing over divisors $d = m$ of n:

$$= 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty}\sum_{d|n} d^{k-1}q^n$$

Let $\sigma_{k-1}(n)$ denote $\sum_{d|n} d^{k-1}$. Then, we have

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

as desired. □

We can divide $G_k(\tau)$ by $2\zeta(k)$ to make the constant term 1; this is called the normalized Eisenstein series and will make appearances in the Elliptic Curve section. Specifically, we have

$$E_k(\tau) = \frac{1}{2\zeta(k)}G_k(\tau) = 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n$$

where B_k are the Bernoulli numbers, a sequence of rational numbers that commonly appears in number theory.

Definition 3.4. A **cusp form of weight** k is a modular form with $a_0 = 0$ in its Fourier expansion.

Definition 3.5. Define the functions $g_2(\tau)$ and $g_3(\tau)$ as

$$g_2(\tau) = 60G_4(\tau), \quad g_3(\tau) = 140G_6(\tau).$$

Then, define the **modular discriminant** $\Delta : \mathcal{H} \to \mathbb{C}$ as $\Delta(\tau) = (g_2(\tau))^3 - 27(g_3(\tau))^2$.

Proposition 3.6. *Δ is a cusp form of weight 12.*

Proof. We can easily see that Δ is a modular form of weight 12: recall that a product of modular forms of weight k_1 and k_2 is a modular form of weight k_1k_2, so $g_2^3(\tau)$ and $g_3^2(\tau)$ are both modular forms of weight 12, and thus a linear combination of them is also a modular form of weight 12.

It remains to show that Δ is a cusp form. Proposition 3.3 tells us the q-expansion of $G_4(\tau)$ and $G_6(\tau)$, so all we need to do is compute the constant terms and show that they are equal to 0. The constant term of $g_2^3(\tau)$ is $(60(2\zeta(4)))^3$, and the constant term of $g_3^2(\tau)$ is $(140(2\zeta(6))^2$. We know $\zeta(4) = \frac{\pi^4}{90}$ and $\zeta(6) = \frac{\pi^6}{945}$, so it can be verified easily that the constant coefficient of Δ is 0. □

4. Congruence Subgroups

While our current definition of a modular form gives rise to many interesting modular forms, extending our definition so that a function only has to satisfy the weakly modular property for a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ rather than the entire group opens up the possibility for many more. As the name suggests, a congruence subgroup is a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ created by imposing congruence conditions on the entries of the matrix:

Definition 4.1. Let N be a positive integer. The **principal congruence subgroup of level** N is the group

$$\Gamma(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \underset{2}{\mathrm{SL}}(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod N \right\}.$$

Then, we can define a general congruence subgroup as follows:

Definition 4.2. A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a **congruence subgroup** if there exists an integer N such that $\Gamma(N) \subset \Gamma$.

We now define modular forms with respect to a congruence subgroup:

Definition 4.3. Let Γ be a congruence subgroup, and for $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$, let

$$f[\gamma]_k = (c\tau + d)^{-k} f(\gamma(\tau)).$$

Then $f : \mathcal{H} \to \mathbb{C}$ is a **modular form with respect to** Γ if

(1) f is holomorphic on $\mathcal{H}$,
(2) $f[\gamma]_k$ is holomorphic at ∞ for all $\gamma \in \Gamma$,
(3) $f[\gamma]_k = f(\tau)$ for all $\tau \in \mathbb{C}$, $\gamma \in \Gamma$.

There is a certain important type of congruence subgroup called $\Gamma_0(N)$ that we will also want to introduce:

Definition 4.4. The subgroup $\Gamma_0(N)$ is defined by

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

Remark 4.5. Modular forms for this subgroup can be used to prove that any integer can be written as a sum of four integer squares (which we will omit here).

5. Elliptic Curves

One of the main reasons we care about modular forms is for their relationship to elliptic curves over the complex numbers. In this section, we discuss and explore this connection.

While we have previously defined the modular forms over $\mathbb{Z}^2 - \{(0,0)\}$, we can consider them as a function over lattices L in $\mathbb{C}$ generated by complex numbers ω_1, ω_2 with $\omega_1/\omega_2 \in \mathcal{H}$. We can also denote lattices by their generators, $L = L(\omega_1, \omega_2)$. The Eisenstein series can be written in terms of lattices:

$$G_{2k}(L) = \sum_{\omega \in L - \{(0,0)\}} \frac{1}{\omega^{2k}}. \tag{1}$$

We shall abbreviate $G_{2k}(\omega_1, \omega_2) = G_{2k}(L(\omega_1, \omega_2))$. Equation 1 immediately tells us that for any $\lambda \in \mathcal{H}$, $G_{2k}(\lambda\omega_1, \lambda\omega_2) = \lambda^{-2k} G_{2k}(\omega_1, \omega_2)$. In particular, we can choose $\lambda = \omega_2^{-1}$ so that $G_{2k}(\omega_1/\omega_2, 1) = \omega_2^{2k} G_{2k}(\omega_1, \omega_2)$. We can rewrite this as

$$G_{2k}(\omega_1, \omega_2) = \omega_2^{-2k} G_{2k}\left(\frac{\omega_1}{\omega_2}, 1\right). \tag{2}$$

In general, we can find another basis by setting $(\omega_1', \omega_2') = (a\omega_1 + b\omega_2, c\omega_1 + d\omega_2)$ for integers a, b, c, d with $ad - bc = \pm 1$. Choosing a basis so that $\operatorname{Im}(\omega_1'/\omega_2') > 0$ guarantees $ad - bc = 1$. The Eisenstein series is independent of the basis. Thus, we also have

$$G_{2k}(\omega_1, \omega_2) = G_{2k}(a\omega_1 + b\omega_2, c\omega_1 + d\omega_2) = (c\omega_1 + d\omega_2)^{-2k} G_{2k}\left(\frac{a\omega_1 + b\omega_2}{c\omega_1 + d\omega_2}, 1\right). \tag{3}$$

Combining (2) and (3) gives us

$$\begin{aligned}(c\omega_1 + d\omega_2)^{-2k} G_{2k}\left(\frac{a\omega_1 + b\omega_2}{c\omega_1 + d\omega_2}, 1\right) &= \omega_2^{-2k} G_{2k}\left(\frac{\omega_1}{\omega_2}, 1\right)\\ G_{2k}\left(\frac{a\omega_1/\omega_2 + b}{c\omega_1/\omega_2 + d}, 1\right) &= (c\omega_1/\omega_2 + d)^{2k} G_{2k}\left(\frac{\omega_1}{\omega_2}, 1\right).\end{aligned}$$

If we set $\tau = \omega_1/\omega_2$, by the abuse of notation $G_{2k}(\tau, 1) = G_{2k}(\tau)$, we have

$$G_{2k}\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^{2k} G_{2k}(\tau).$$

Thus, we have proved: (1) every lattice with two generators in $\mathbb{C}$ can be written in the form $\lambda L(1, \tau)$; (2) the weak modularity condition holds for the lattice definition of modular forms. The first fact allowed us to convert the lattice definition into a form with τ and $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, which shows that the lattice definition can be reduced to our original definition for modular forms.

Definition 5.1 (Weierstrass elliptic function)**.** For any lattice L in the complex plane, define the *Weierstrass elliptic function* to be

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in L - \{(0,0)\}} \left(\frac{1}{(z+\omega)^2} - \frac{1}{\omega^2}\right).$$

Remark 5.2. In general, an *elliptic function* is a meromorphic function that is doubly-periodic on the complex plane.

Remark 5.3. The $\wp$-function is not a modular form since it has poles on L.

Where did the equation for Definition 5.1 come from? If we were to naively construct a doubly-periodic function, we might want to imitate the form of Eisenstein series and define

$$f(z) = \sum_{\omega \in L} \frac{1}{(z+\omega)^3}$$

as this converges absolutely. However, it turns out that if we are careful to make sure that the resulting series converges, we can integrate $-2f(z)$ term by term to obtain $\wp(z)$, which is of order 2. Integrating $\frac{-2}{(z+\omega)^3}$ with respect to z gives $\frac{1}{(z+\omega)^2} + C(\omega)$ where $C(\omega)$ is a constant that only depends on ω. We set $C(\omega) = \frac{1}{\omega^2}$ so that the series converges. While we do not prove it here, there are no elliptic functions with order 1, so in a sense, $\wp(z)$ is the most fundamental type of elliptic function. In fact, we have the following result, which we will state without proof:

Proposition 5.4. *All elliptic functions are expressible as a rational function in $\wp(z)$ and $\wp'(z)$.*

We wish to show that $\wp(z)$ is doubly-periodic. By definition, we have $\wp'(z) = -2f(z)$, which is doubly-periodic. As $\mathbb{C} - L$ is connected, we can integrate to get that for any $\omega \in L$, $\wp(z+\omega) = \wp(z) + D(\omega)$ where $D(\omega)$ is a constant that depends only on ω. This is close to being doubly periodic. The final observation we need is to see that $\wp(z)$ is even because $\wp(z) = \wp(-z)$. Then we can choose our basis ω_1, ω_2 so that $\omega_i/2 \notin L$ for $i = 1, 2$. Now, we have

$$\wp(\omega_i/2) = \wp(-\omega_i/2) + D(\omega_i) = \wp(\omega_i/2) + D(\omega_i),$$

so $D(\omega_i) = 0$ for $i = 1, 2$. Then $\wp(z + \omega_i) = \wp(z)$ for $i = 1, 2$ implies that $\wp(z+\omega) = \wp(z)$ for any $\omega \in L$. Thus, $\wp(z)$ is doubly-periodic as desired.

The following proposition will allow us to connect the $\wp$-function to complex elliptic curves:

Proposition 5.5. *The Weierstrass $\wp$-function satisfies the differential relation*

$$\wp'(z)^2 = 4\wp^3(z) - g_2(L)\wp(z) - g_3(L).$$

Proof. We will only sketch an outline as the proof requires more computation. We look at the Taylor expansion of the difference of the two sides, $F(z) = \wp'(z)^2 - 4\wp^3(z) + g_2(L)\wp(z) + g_3(L)$, and see that it has no pole at $z = 0$, implying that $F(z)$ has no poles on L. As $F(z)$ is a combination of elliptic functions, it must be an elliptic function with order 0. But this would imply that $F(z)$ is constant as there are no non-constant elliptic functions with order less than 2. We can check that $F(z) = 0$ by seeing that $F(0) = 0$. □

Remark 5.6. The Weierstrass $\wp$-function is the inverse of the elliptic integral

$$u = \int_y^\infty \frac{\mathrm{d}s}{\sqrt{4s^3 - g_2(L)s - g_3(L)}}$$

so that $y = \wp(u)$. This can be shown by differentiating both sides then using the differential relation (5).

Now we introduce elliptic curves.

Definition 5.7 (Complex Projective Plane)**.** The *complex projective plane* is the set of ordered triplets $(a, b, c) \in \mathbb{C}^3$ under the equivalence relation $(a, b, c) \sim (a', b', c')$ if there exists a nonzero complex number λ such that $(a', b', c') = (\lambda a, \lambda b, \lambda c)$.

Definition 5.8 (Elliptic Curve)**.** An *elliptic curve* over a field k is a non-singular, complete curve of genus 1 with a distinguished point. When the characteristic of the field is not 2 or 3, the curve can be seen as the locus of the equation

$$Y^2Z = X^3 + aXZ^2 + bZ^3$$

where a and b is in k and $4a^3 + 27b^2 \neq 0$. This equation is called the *Weierstrass equation.*

The condition on a and b will prevent double roots, guaranteeing that E is of genus 1 and thus topologically a torus. Intuitively, we have a point at infinity, $(0, 1, 0)$. All other points have nonzero Z-value, so we can just think about the equation as being in two variables, X/Z and Y/Z.

We have the following embedding (map) from the complex torus $\mathbb{C}/L$ to the complex projective plane:

$$z \mapsto \begin{cases} (\wp(z) : \wp'(z) : 1), & z \neq 0 \\ (0 : 1 : 0) & \end{cases},$$

which gives us the form for the elliptic curve

$$E(L) : Y^2Z = 4X^3 - g_2(L)XZ^2 - g_3(L)Z^3.$$

If we use a different lattice $L' = \lambda L$, then we have $g_2(L') = \lambda^{-4}g_2(L)$ and $g_3(L') = \lambda^{-6}g_3(L)$, which can be absorbed into Z by substituting $Z' = \lambda^{-2}Z$. So the elliptic curve that we get from a lattice is unique up to homothety of the lattice (two lattices are homothetic if one can be turned and rescaled into the other).

Definition 5.9. The j-invariant is a modular form of weight zero:

$$j(z) = \frac{1728g_2(z)^3}{\Delta(z)} = 1728\frac{E_4(z)^3}{E_4(z)^3 - E_6(z)^2}.$$

We claim that $A = g_2(L)$ and $B = g_3(L)$ for some appropriate lattice L. In the accompanying proof, we use the fact that any lattice in $\mathbb{C}$ with two generators, $L = \{m\omega_1 + n\omega_2\}$ where $z = \omega_1/\omega_2 \in \mathcal{H}$, can be expressed as $L = \lambda L_z = \{m\lambda z + n\lambda\}$ where $\lambda = \omega_2$. Essentially, L is a complex multiple of L_z.

Proposition 5.10. *For any $A, B \in \mathbb{C}$ such that $A^3 \neq 27b^2$ there exists $L = \lambda L_z$ such that*

$$g_2(L) = A, \quad g_3(L) = B.$$

Proof. By definition of g_2 and g_3, we immediately get that $g_2(\lambda L_z) = \lambda^{-4}g_2(L_z)$ and similarly $g_3(\lambda L_z) = \lambda^{-6}g_3(L_z)$. We can re-express our desired result in terms of the normalized Eisenstein series: there exist λ and z such that

$$E_4(z) = \frac{3}{4\pi^4}g_2(z) = \frac{3}{4\pi^4}\lambda^4 g_2(L) = \frac{3}{4\pi^4}\lambda^4 A$$

and similarly

$$E_6(z) = \frac{27}{8\pi^6}g_3(z) = \frac{27}{8\pi^6}\lambda^6 g_3(L) = \frac{27}{8\pi^6}\lambda^6 B.$$

Then the condition $A^3 \neq 27B^2$ becomes $a^3 \neq b^2$, which implies $\frac{b^2}{a^3} \neq 1$. By the definition of the j-invariant, we have $\frac{E_6^2}{E_4^3} = 1 - \frac{1728}{j}$. A property of $j(z)$ is that it takes on all finite values on $\mathcal{H}$. Then $\frac{E_6(z)^2}{E_4(z)^3}$ can be any value we want except 1; in particular, for any a, b with $\frac{b^2}{a^3} \neq 1$ we can have $\frac{E_6(z)^2}{E_4(z)^3} = \frac{b^2}{a^3}$. Then choose λ so that $E_4(z) = \lambda^4 a$. This implies

$$E_6(z)^2 = \frac{b^2E_4(z)^3}{a^3} = \lambda^{12}b^2,$$

or

$$E_6(z) = \pm\lambda^6 b.$$

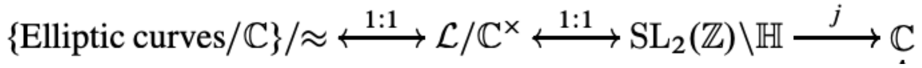
If we have a positive sign, then our λ and z satisfy our desired conditions. If we have a negative sign, then we can replace λ by $i\lambda$; this changes the sign for $E_6(z)$ but not for $E_4(z)$. Thus we are done. □

The above proposition directly proves the following proposition.

Proposition 5.11. *Every elliptic curve E over $\mathbb{C}$ is isomorphic to $E(L)$ for some lattice L.*

We also have (without proof):

Proposition 5.12. *Lattices L_1, L_2 are homothetic if and only if $j(L_1) = j(L_2)$.*



$$\begin{tikzcd}\{\text{Elliptic curves}/\mathbb{C}\}/\approx \arrow[r, leftrightarrow, "1:1"] & \mathcal{L}/\mathbb{C}^{\times} \arrow[r, leftrightarrow, "1:1"] & \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H} \arrow[r, "j"] & \mathbb{C}\end{tikzcd}$$

Figure 2. Classifying Complex Elliptic Curves using the j-invariant and the modular group

Thus, we can classify elliptic curves up to isomorphism by their j-invariants, which can be found by looking at the g_2 and g_3-invariants of the curves. We can further realize that two lattices being homothetic means that there is an invertible transformation on the complex plane to itself. Each lattice can be read off as $\tau = \omega_1/\omega_2 \in \mathcal{H}$. Thus, we can see that the lattices under homothety is isomorphic to $\mathcal{H}$ under equivalence by action of $\mathrm{SL}_2(\mathbb{Z})$ (more precise, to the orbit of $\mathcal{H}$ by the group action $\mathrm{SL}_2(\mathbb{Z})$). Figure 2. from [Mil06] shows this chain of bijections.

The sets are:

(1) $\{\text{Elliptic curves}/\,\mathbb{C}\}/\approx=$ the set of elliptic curves under some equivalence relation we will not describe here,
(2) $L/\,\mathbb{C}^{\times} =$ the set of lattices over the complex plane,
(3) $\mathrm{SL}_2(\mathbb{Z}) \setminus \mathcal{H} =$ the orbits of $\mathcal{H}$ under the group action $\mathrm{SL}_2(\mathbb{Z})$,
(4) $\mathbb{C} =$ the complex plane.

The bijections are:

(1) Every elliptic curve can be associated with a lattice; every equivalence of elliptic curves is associated with exactly one lattice up to homothety,
(2) Every lattice is associated with a $\tau = \omega_1/\omega_2$, and the homothety on lattices mirrors the action by $\mathrm{SL}_2(\mathbb{Z})$ on the half plane,
(3) The j-invariant maps the half-plane surjectively onto $\mathbb{C}$ (which we haven't proved).

References

[BD16] Peter Bruin and Sander Dahmen. Modular forms, 2016.

[DS05] Fred Diamond and Jerry Michael Shurman. *A first course in modular forms*, volume 228. Springer, 2005.

[Mil06] J.S. Milne. *Elliptic Curves*. BookSurge Publishers, 2006.

[Sha03] BV Shabat. Introduction to complex analysis-excerpts, 2003.