Modular Forms

Gaurish Korpal[1]
gaurish.korpal@niser.ac.in

Winter Internship Project Report

[1] 4^{th} year Int. MSc. Student, National Institute of Science Education and Research, Jatni (Bhubaneswar, Odisha)

Certificate

Certified that the winter internship project report "Modular Forms" is the bona fide work of "Gaurish Korpal", 4^{th} year Int. MSc. student at National Institute of Science Education and Research, Jatni (Bhubaneswar, Odisha), carried out under my supervision during December 8, 2017 to December 30, 2017.

Place: Kozhikode
Date: December 30, 2017

Prof. M. Manickam
Supervisor
Director,
Kerala School of Mathematics,
Kunnamangalam, Kozhikode 673571

Abstract

The basics of classical one-variable theory of modular forms has been discussed. Notions of elliptic functions, fundamental domains and Eisenstein series has been introduced.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor *Prof. M. Manickam* for his motivation and immense knowledge. I am also thankful to *Sandeep E. M.*[1] for the enlightening discussions.

Last but not the least, I would like to thank

- Donald Knuth for TEX
- Michael Spivak for $\mathcal{AMS}$-TEX
- Sebastian Rahtz for TEX Live
- Leslie Lamport for LATEX
- American Mathematical Society for $\mathcal{AMS}$-LATEX
- Hàn Thế́ Thành for pdfTEX
  - Heiko Oberdiek for `hyperref` package
  - Steven B. Segletes for `stackengine` package
  - Axel Sommerfeldt for `subcaption` package
  - David Carlisle for `graphicx` package
  - Javier Bezos for `enumitem` package
  - Hideo Umeki for `geometry` package
  - Peter R. Wilson & Will Robertson for `epigraph` package
- Philipp Khl & Daniel Kirsch for Detexify (a tool for searching LATEX symbols)
- TeX.StackExchange community for helping me out with LATEX related problems

GeoGebra was used to create all figures for this document.

[1] PhD Student, KSoM-Kozhikode

Contents

Introduction

> *"There are five elementary arithmetical operations: addition, subtraction, multiplication, division, and... modular forms."*
>
> — Martin Eichler, *quoted by Andrew Wiles in NOVA Season 25 Episode 4 "The Proof" (First Aired: Oct 28, 1997)*

Modular forms are a particular class of *automorphic forms*. The word *automorphic* is made up of two Greek words, *auto* meaning self and *morphe* meaning shape. Hence the adjective automorphic is used in the contexts where "something" keeps the shape "similar to the original" under certain "changes of variables" [8, pp. 147]. An automorphic form is a well-behaved function from a topological group G to the complex numbers (or complex vector space) which is invariant under the action of a discrete subgroup $\Gamma \subset G$ of the topological group. When dealing with automorphic forms for the modular group, it is traditional to use the term *modular* instead of *automorphic*. Hence, modular forms are automorphic forms defined over the groups $SL_2(\mathbb{R})$ or $PSL_2(\mathbb{R})$ with the discrete subgroup being the modular group, or one of its congruence subgroups.

The story of modular forms began with elliptic functions, which are doubly periodic meromorphic complex functions. Elliptic functions were studied by Karl Weierstrass and date back to C. F. Gauss, and they led naturally to the study of elliptic curves, which are intimately related to modular forms. In the first chapter, we will see the basic theory of Weierestrass $\wp$-function and it's importance in the theory of elliptic functions and elliptic curves.

The full modular group is the special linear group $SL_2(\mathbb{Z})$ of 2×2 matrices with integer coefficients and unit determinant. The modular group is obtained by identifying the matrices A and $-A$ in the full modular group. Hence, the modular group is the projective special linear group $PSL_2(\mathbb{Z})$ of 2×2 matrices with integer coefficients and unit determinant. It acts on the upper-half of the complex plane by fractional linear transformations and is a subgroup of the group of isometries of the hyperbolic plane (Poincaré half-plane model). Given a topological space and a group acting on it, the images of a single point under the group action form an orbit of the action. The modular group acts on the upper half-plane, hence we can construct fundamental domains, i.e. a subset of the upper half-plane whose interior contains exactly one point from each of these orbits. Important subgroups of the modular group, called congruence subgroups, are given by imposing congruence relations on the associated matrices. In the second chapter we will discuss the properties (full) modular group, congruence subgroups and their fundamental domains.

The origins of the definition of modular forms date back to the first half of the nineteenth century, to the era of Carl Jacobi and Gotthold Eisenstein. In the third chapter we discuss the constructions of Eisenstein series and their importance in the theory of modular forms, hence justifying the belief that modular forms give the fifth operation of arithmetic. Then we conclude the chapter by illustrating the birth of modular forms from the elliptic theory via the modular invariant, called j-invariant.

The book by Neal Koblitz [1] has been used as the main reference for this report.

Chapter 1

Weierstrass's elliptic functions

> *"... it is true that a mathematician who is not somewhat of a poet, will never be a perfect mathematician."*
>
> —Karl Weierstrass, *in a letter to Sofia Kovalevskaya, August 27, 1883, as shared by Gösta Mittag-Leffler at the 2nd International Congress for Mathematicians in Paris.*

1.1 Introduction

Definition 1 (Lattice)**.** The set of all integral linear combinations of two given complex numbers ω_1 and ω_2, where ω_1 and ω_2 do not lie on the same line through the origin is called a *lattice on a complex plane* and is denoted by Λ.

$$\Lambda = \{m\omega_1 + n\omega_2 : m, n \in \mathbb{Z}\}$$

Remark 1. The choice of ω_1, ω_2 giving the lattice Λ is not unique. For example, $\omega_1' = \omega_1 + \omega_2$ and ω_2 give the same lattice. More generally, we can obtain new basis ω_1' and ω_2' for the lattice Λ by multiplying the column vector $\omega = \begin{pmatrix}\omega_1\\ \omega_2\end{pmatrix}$ by a 2×2 matrix A of determinant ± 1, i.e. $\omega' = A\omega$ where $\omega' = \begin{pmatrix}\omega_1'\\ \omega_2'\end{pmatrix}$. See, Definition 10 and [1, Problem I.5.1].

Remark 2. We shall always take ω_1, ω_2 in clockwise order; that is, we shall assume that ω_1/ω_2 has positive imaginary part.

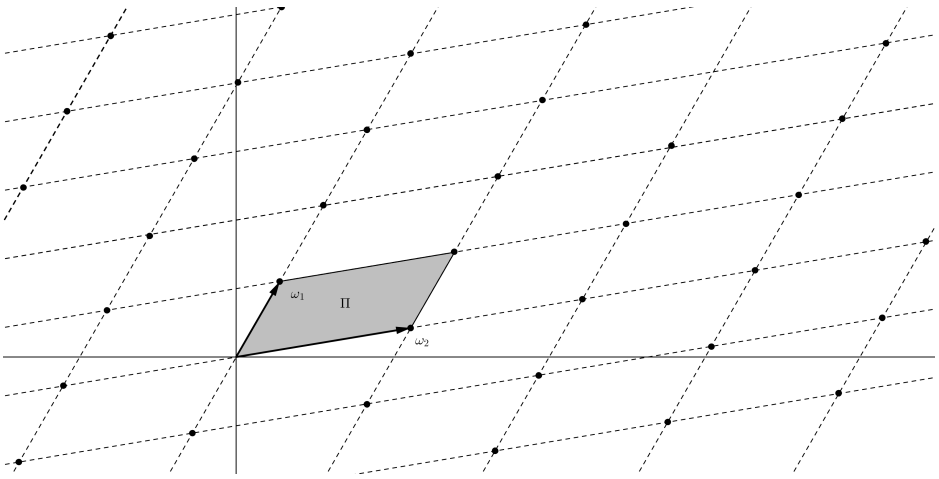


Figure 1.1: A lattice Λ and its fundamental parallelogram Π. Drawn in GeoGebra 4.0.34.0.

Definition 2 (Fundamental parallelogram)**.** The fundamental parallelogram for ω_1 and ω_2 is defined as

$$\Pi = \{a\omega_1 + b\omega_2 : 0 \le a \le 1, 0 \le b \le 1\}$$

Proposition 1. *Any number $z_0 \in \mathbb{C}$ can be written as the sum of an element in the lattice Λ and an element in Π, and in only one way unless the element of Π happens to lie on the boundary of Π.*

Proof. Since ω_1, ω_2 form a basis of $\mathbb{C}$ over $\mathbb{R}$, any number $z_0 \in \mathbb{C}$ can be written in the form $z_0 = x_0\omega_1 + y_0\omega_2$ for some $x_0, y_0 \in \mathbb{R}$. Then z_0 can be written as the sum of an element $\lambda = \lfloor x_0 \rfloor \omega_1 + \lfloor y_0 \rfloor \omega_2 \in \Lambda$ and an element $\varpi = (x_0 - \lfloor x_0 \rfloor)\omega_1 + (y_0 - \lfloor x_0 \rfloor)\omega_2 \in \Pi$, and in only one way unless x_0 or y_0 happens to be an integer, i.e., the element of Π happens to lie on the boundary of Π. □

Definition 3 (Elliptic function)**.** For a given lattice Λ, a meromorphic function on $\mathbb{C}$ is said to be an *elliptic function*[1] relative to Λ if $f(z + \lambda) = f(z)$ for all $\lambda \in \Lambda$. We denote the set of elliptic function relative to Λ by $\mathcal{E}_\Lambda$.

Remark 3. Instead of checking for each $\lambda \in \Lambda$, it suffices to check for $\lambda = \omega_1$ and $\lambda = \omega_2$. In other words, an elliptic function is periodic with two periods ω_1 and ω_2. Such a function is determined by its values on the fundamental parallelogram Π; and its values on opposite points of the boundary of Π are the same, i.e.,

$$f(a\omega_1 + \omega_2) = f(a\omega_1), \qquad f(\omega_1 + b\omega_2) = f(b\omega_2)$$

for $0 \le a, b \le 1$. Thus, we can think of an elliptic function $f(z)$ as a function on the set Π with opposite sides glued together. This set (more precisely, "complex manifold") is known as a "torus", denoted[2] by $\mathbb{C}/\Lambda$.

Proposition 2. *$\mathcal{E}_\Lambda$ is a subfield of the field of all meromorphic functions.*

Proof. One can easily check that the sum, difference, product and quotient of two elliptic functions relative to Λ is again an elliptic function relative to Λ. □

Remark 4. Also, the subfield $\mathcal{E}_\Lambda$ is closed under differentiation.

Proposition 3. *A function $f(z) \in \mathcal{E}_\Lambda$ which has no pole in the fundamental parallelogram Π must be a constant.*

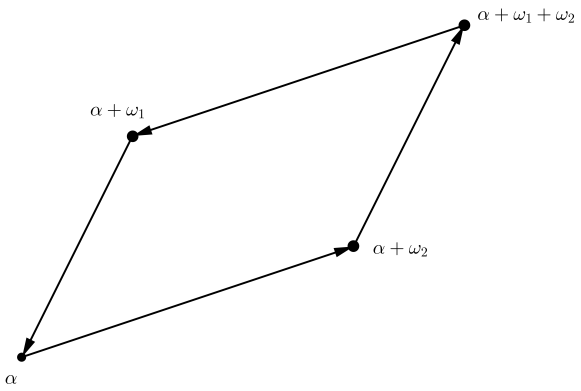
Proof. Since Π is compact, any such function must be bounded in Π, and hence due to periodicity on whole of $\mathbb{C}$. Since $f(z)$ is holomorphic in Π, and hence due to periodicity on whole of $\mathbb{C}$. By Liouville's theorem [5, Theorem IV.3.4], an entire function which is bounded on all of $\mathbb{C}$ must be constant. □

Proposition 4. *Suppose $f(z) \in \mathcal{E}_\Lambda$ has no poles on the boundary of $\alpha + \Pi = \{\alpha + z : z \in \Pi\}$. Then the sum of the residues of $f(z)$ in $\alpha + \Pi$ is zero.*

Proof. Consider the boundary C of $\alpha + \Pi$.

[1] For details about the history of this term, see [3, §2.1, 2.2 and 2.3].

[2] Compare it with the notation in Definition 18.



By the residue theorem [5, Theorem V.2.2], the sum of residues is equal to $\frac{1}{2\pi i}\int_C f(z)dz$. But the integral over opposite sides cancel, since the values of $f(z)$ at corresponding points are same (Remark 3), while dz has opposite signs, because the path of integration is in opposite directions on opposite sides. Thus the integral is zero, and so the sum of residues in zero. □

Remark 5. Since a meromorphic function can only have finitely many poles in a bounded region, it is always possible to choose an α such that the boundary of $\alpha + \Pi$ misses the poles on $f(z)$.

Corollary 1. *A non-constant elliptic function must have at least two poles (or a multiple pole) in* $\alpha + \Pi$.

Proof. If $f(z) \in \mathcal{E}_\Lambda$ is non-constant and has single simple pole, then the sum of residues would not be zero. □

Proposition 5. *Suppose* $f(z) \in \mathcal{E}_\Lambda$ *has no zeros or poles on the boundary of* $\alpha + \Pi$. *Let* $\{m_j\}$ *be the orders of various zeros in* $\alpha + \Pi$, *and let* $\{n_j\}$ *be the orders of various poles in* $\alpha + \Pi$. *Then* $\sum m_i = \sum n_j$.

Proof. By the argument principle [5, Theorem V.3.4], the difference $\sum m_i - \sum n_j$ is equal to $\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}dz$, where C is the boundary of $\alpha + \Pi$. Now using Remark 4 and Proposition 2, we conclude that $\frac{f'(z)}{f(z)} \in \mathcal{E}_\Lambda$. Hence again as in Proposition 4, we have this contour integral to be equal to zero. Hence completing the proof. □

1.2 Weierstrass $\wp$-function

Definition 4. The Weierstrass $\wp$-function is denoted as $\wp(z;\Lambda)$ or simply $\wp(z)$ is the lattice is fixed throughout the discussion. It's defined as

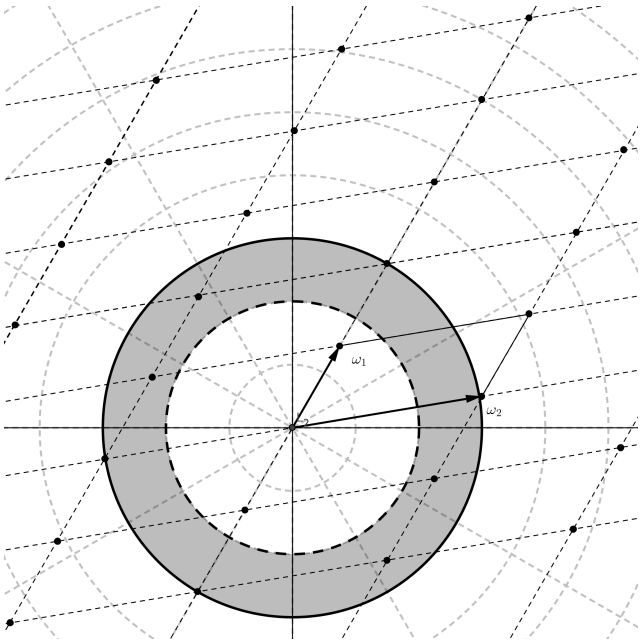
$$\wp(z) = \wp(z;\Lambda) = \frac{1}{z^2} + \sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right)$$

Lemma 1. *The series* $\displaystyle\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} \frac{1}{|\lambda|^r}$ *converges for* $r > 2$.

Proof. We can spilt the sum into sums over λ satisfying $n-1 < |\lambda| \leq n$, as $n = 1, 2, 3, \ldots$

$$\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} \frac{1}{|\lambda|^r} = \sum_{n=1}^{\infty} \sum_{n-1<|\lambda|\leq n} \frac{1}{|\lambda|^r} \tag{1.1}$$

Visually, inner sum is summing over annulus of inner radius $n-1$ and outer radius n, for example:



Now, since the number of integral points N enclosed in a very large region is equal to the area of the region plus some error term depending on the perimeter of the region [6, §8.1], here we have

$$N = (\pi n^2 - \pi(n-1)^2) + C(2\pi n + 2\pi(n-1)) = \pi(2n-1)(1+C2\pi)$$

for some constant C. Hence N is of the order n, i.e. $N = O(n)$. Thus we have

$$\sum_{\substack{\lambda\in\Lambda \\ n-1<|\lambda|\leq n}} \frac{1}{|\lambda|^r} \leq C' \frac{n}{n^r} = C' \frac{1}{n^{r-1}}$$

by choosing an appropriate constant C'. Now, using this inequality in (1.1) we get

$$\sum_{\substack{\lambda\in\Lambda \\ \lambda\neq 0}} \frac{1}{|\lambda|^r} \leq C' \sum_{n=1}^{\infty} \frac{1}{n^{r-1}}$$

which converges[3] for $r-1>1$. Hence completing the proof. □

Remark 6. For another proof using the fact from functional analysis that a vector space V is finite dimensional if and only if all norms are equivalent, see [3, Lemma 2.5.1].

Theorem 1. *$\wp(z) \in \mathcal{E}_\Lambda$ and its only pole is a double pole at each lattice point.*

Proof. We will divide the proof into four parts

Claim 1. *$\wp(z)$ is holomorphic on $\mathbb{C}-\Lambda$*

For $z \in \mathbb{C}-\Lambda$ we have

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{\lambda\in\Lambda \\ \lambda\neq 0}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right)$$

[3]This can be proved using Cauchy condensation test or integral test. That is,

$$\sum_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}} < \lim_{m\to\infty} \int_1^m \frac{dx}{x^{1+\varepsilon}} = \lim_{m\to\infty} \frac{1}{\varepsilon}\left(1 - \frac{1}{m^\varepsilon}\right) < \infty$$

.

First we write the summand over a common denominator:

$$\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} = \frac{2z\lambda - z^2}{(z-\lambda)^2\lambda^2} = O\left(\frac{1}{|\lambda|^3}\right), \quad \text{as } |\lambda| \to \infty$$

Therefore by Lemma 1, we conclude that the series $\wp(z)$ is absolutely convergent. Moreover, if $|z|$ is bounded above then the series $\wp(z)$ converges absolutely and uniformly, i.e. for z in any compact subset of $\mathbb{C} - \Lambda$, $\wp(z)$ is absolutely uniformly convergent. Hence by [5, Corollary VII.2.4], the function $\wp(z)$ is holomorphic on $\mathbb{C} - \Lambda$.

Claim 2. *$\wp(z)$ is meromorphic on $\mathbb{C}$ with a double pole at all lattice points and no other poles.*

Suppose $|z| < R$, and write

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{\lambda\in\Lambda \\ |\lambda|\leq 2R}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right) + \sum_{\substack{\lambda\in\Lambda \\ |\lambda|> 2R}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right)$$

The term in the second sum is $O(1/|\lambda|^3)$ uniformly for $|z| < R$, so by Lemma 1 this second sum defines a holomorphic function in $|z| < R$. Finally, note that the first sum exhibits double poles at the lattice points in the disc $|z| < R$.

Claim 3. *$\wp(z)$ is an even function*

We have $\wp(z) = \wp(-z)$ because

$$\wp(-z) = \frac{1}{z^2} + \sum_{\substack{\lambda\in\Lambda \\ \lambda\neq 0}} \left(\frac{1}{(-z-\lambda)^2} - \frac{1}{\lambda^2}\right) = \frac{1}{z^2} + \sum_{\substack{-\lambda\in\Lambda \\ \lambda\neq 0}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right) = \wp(z)$$

Since summing over $\lambda \in \Lambda$ is same as summing over $-\lambda \in \Lambda$.

Claim 4. *$\wp(z)$ is doubly periodic*

Differentiating the series $\wp(z)$ term by term we get

$$\wp'(z) = -2\sum_{\lambda\in\Lambda} \frac{1}{(z-\lambda)^3}$$

From this we observe that $\wp'(z)$ is doubly periodic, since replacing z by $z + \lambda_0$ for some fixed $\lambda_0 \in \Lambda$ just rearranges the terms in the sum. Thus $\wp'(z) \in \mathcal{E}_\Lambda$.

Therefore we have,

$$\wp'(z+\omega_i) - \wp'(z) = 0, \qquad \text{for } i = 1, 2$$

Hence we must have

$$\wp(z+\omega_i) - \wp(z) = C, \qquad \text{for } i = 1, 2$$

for some constant C. But substituting $z = -\frac{\omega_i}{2}$ and using the fact that $\wp(z)$ is an even function, we conclude that

$$C = \wp\left(\frac{\omega_i}{2}\right) - \wp\left(-\frac{\omega_i}{2}\right) = 0, \qquad \text{for } i = 1, 2$$

Hence proving our claim (Remark 3).

□

Remark 7. Since $\wp(z)$ has exactly one double pole in the the shifted fundamental parallelogram $\alpha + \Pi$, by Proposition 5 it has exactly two zeros there (or one double zero).

Proposition 6. *$\wp(z)$ takes every value $z_0 \in \mathbb{C} \cup \{\infty\}$ exactly twice on the torus $\mathbb{C}/\Lambda$ counting multiplicity; and the values assumed with multiplicity two are ∞, $e_1 = \wp(\omega_1/2), e_2 = \wp(\omega_2/2)$ and $e_3 = \wp((\omega_1+\omega_2)/2)$.*

Proof. We will divide the proof into several parts

Claim 1. *For any fixed $z_0 \in \mathbb{C}$, the elliptic function $\wp(z) - z_0$, has either two simple zeros or one double zero in Π.*

Since $\wp(z) - z_0$ has exactly one double pole in the the shifted fundamental parallelogram $\alpha + \Pi$, by Proposition 5 it has exactly two zeros there (or one double zero).

Claim 2. *The zeros of $\wp'(z)$ are precisely $\omega_1/2$, $\omega_2/2$ and $(\omega_1+\omega_2)/2$.*

$\wp'(z)$ is an odd function since

$$\wp'(-z) = -2\sum_{\lambda\in\Lambda} \frac{1}{(-z-\lambda)^3} = 2\sum_{-\lambda\in\Lambda} \frac{1}{(z-\lambda)^3} = -\wp'(z)$$

Since $\wp'(z) \in \mathcal{E}_\Lambda$ and an odd function, we have

$$\wp'\left(\frac{\omega_1}{2}\right) = \wp'\left(\frac{\omega_1}{2} - \omega_1\right) = \wp'\left(-\frac{\omega_1}{2}\right) = -\wp'\left(\frac{\omega_1}{2}\right)$$

Hence $\wp'\left(\frac{\omega_1}{2}\right) = 0$. Similarly, we can show $\wp'\left(\frac{\omega_2}{2}\right) = 0$ and $\wp'\left(\frac{\omega_1+\omega_2}{2}\right) = 0$. Since, $\wp'(z)$ has exactly one triple pole 0 in the the shifted fundamental parallelogram $-\frac{\omega_1+\omega_2}{2} + \Pi$, by Proposition 5 it has exactly three zeros there (or one triple zero). We have found the three zeros, hence proving our claim.

Claim 3. *The values assumed by $\wp(z)$ with multiplicity two are $\infty, e_1 = \wp(\omega_1/2), e_2 = \wp(\omega_2/2)$ and $e_3 = \wp((\omega_1+\omega_2)/2)$.*

0 is a double pole with residue zero and other three $\omega_1/2$, $\omega_2/2$ and $(\omega_1+\omega_2)/2$ are double roots of $\wp(z) - e_i$ for $i = 1, 2, 3$ since they are the only roots of $\wp'(z)$.

□

Remark 8. Thus, the Weierstrass $\wp$-function gives a two-to-one map from the torus $\mathbb{C}/\Lambda$ to the Riemann sphere $\mathbb{P}^1_{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ (Definition 6), except over the four *branch points* e_1, e_2, e_3, ∞, each of which has a single preimage in $\mathbb{C}/\Lambda$ (the quotient space).

1.3 Field of elliptic functions

Proposition 7. *The subfield $\mathcal{E}^+_\Lambda \subset \mathcal{E}_\Lambda$ of even elliptic functions for Λ is generated by $\wp$, i.e. $\mathcal{E}^+_\Lambda = \mathbb{C}(\wp)$.*

Proof. Let $f(z) \in \mathcal{E}^+_\Lambda$ and Π' be a fundamental parallelogram with two sides removed:

$$\Pi' = \{a\omega_1 + b\omega_2 : 0 \le a < 1, 0 \le b < 1\}$$

Then every point in $\mathbb{C}$ differs by a lattice element from exactly one point in Π'; that is Π' is a set of coset representatives for the additive group of complex numbers modulo the subgroup Λ. We will list zeros and poles in Π', omitting 0 from our list (even if it happens to be a zero or pole of $f(z)$). Following method will be used to list the zeros (and analogously poles) of $f(z)$:

Case 1. *$a \in \Pi'$, $a \neq 0$, is a zero of $f(z)$ which is not half of a lattice point, i.e. $a \neq \omega_1/2, \omega_2/2$ or $(\omega_1+\omega_2)/2$.*

We define the symmetric point a^* as follows:

$$a^* = \begin{cases} \omega_1 + \omega_2 - a & \text{if } a \text{ is in the interior of } \Pi' \\ \omega_1 - a & \text{if } a \text{ is on the side of } \omega_1 \\ \omega_2 - a & \text{if } a \text{ is on the side of } \omega_2 \end{cases}$$

We have $f(a^* - z) = f(-a - z)$ by double periodicity, and this is equal to $f(a+z)$ because $f(z)$ is an even function. Thus, if a is a zero of order m, i.e $f(a+z) = a_m z^m + a_{m+1} z^{m+1} + \ldots$, it follows that $f(a^* + z) = a_m(-z)^m + a_{m+1}(-z)^{m+1} + \ldots$, i.e. a^* is a zero of order m.

Hence in this case, when the zeros of $f(z)$ in Π' are not half-lattice points, we write all such a's repeated as per their multiplicity but choose only one from each pair of symmetrical zeros a and a^* for our list of zeros of $f(z)$.

Case 2. *$a \in \Pi'$, $a \neq 0$, is a zero of $f(z)$ which is half of a lattice point, i.e. $a = \omega_1/2, \omega_2/2$ or $(\omega_1+\omega_2)/2$.*

Let m be the order of a, i.e., $f(a+z) = a_m z^m + a_{m+1}^{m+1} + \ldots$. Now since $2a \in \Lambda$ we have

$$f(a+z) = f(a - 2a + z) = f(-a + z) = f(a - z)$$

since $f(z)$ is doubly periodic and an even function. Thus we have

$$a_m z^m + a_{m+1} z^{m+1} + \ldots = a_m(-z)^m + a_{m+1}(-z)^{m+1} + \ldots$$

Hence m is even (and odd higher coefficients are zero).

So, in this case we write all such a's repeated as per their multiplicity but choose only half as many times as the multiplicity of each of them.

Using above method we get the list $\{a_i\}$ of zeros of $f(z)$ in Π'. Similarly, we get the list $\{b_j\}$ of poles of $f(z)$ in Π' (i.e., only half of them appear).

Since all of the a_i and b_j are non-zero, the values $\wp(a_i)$ and $\wp(b_j)$ are finite. We define following elliptic function:

$$g(z) = \frac{\prod_i (\wp(z) - \wp(a_i))}{\prod_j (\wp(z) - \wp(b_j))}$$

Since $0 \in \Pi'$ is the only pole in the numerator or denominator of $g(z)$, it follows that the non-zero zeros of $g(z)$ must comes from the zeros of $\wp(z) - \wp(a_i)$, while the non-zero poles of $g(z)$ must come from the zeros of $\wp(z) - \wp(b_j)$. But by Proposition 6 we know that $\wp(z) - z_0$ (for a constant z_0) has a double zero at half-lattice points, and otherwise has pair of simple zeros at a and the symmetric point a^*. Hence, all these are the only zeros of $\wp(z) - z_0$ in Π'.

By our construction of $\{a_i\}$ and $\{b_j\}$, we see that $g(z)$ and $f(z)$ have the same order of zero or pole everywhere in Π', with the possible exception of the point 0. Let's choose α such that no lattice point and no zero or pole of $f(z)$ or $g(z)$ is on the boundary of $\alpha + \Pi$, then $\alpha + \Pi$ will contain precisely one lattice point λ_0. Let m_f denote the order of zero of $f(z)$ at λ_0 (m_f is negative if there is a pole), and m_g denote the analogous order of $g(z)$. Since $f(z)$ and $g(z)$ have same orders of zeros everywhere in $\alpha + \Pi$ with possible exception of λ_0, by Proposition 5 we get $m_f = m_g$. Hence, $g(z)$ has same zeros and poles as $f(z)$ (counting multiplicities).

It follows that $f(z) = cg(z)$ for some constant c. Since $g(z)$ is a rational function of $\wp(z)$, this completes the proof. □

Corollary 2. *For a fixed positive integer N, the even elliptic function $\wp(Nz)$ is a rational function in $\wp(z)$.*

Theorem 2. *$\mathcal{E}_\Lambda = \mathbb{C}(\wp, \wp')$, i.e. any elliptic function for Λ is a rational expression in $\wp(z;\Lambda)$ and $\wp'(z;\Lambda)$. More precisely, given $f(z) \in \mathcal{E}_\Lambda$, there exist two rational functions $g(X), h(X)$ such that $f(z) = g(\wp(z)) + \wp'(z)h(\wp(z))$.*

Proof. If $f(z)$ is an elliptic function for Λ, then so are the two even functions

$$\frac{f(z)+f(-z)}{2}, \quad \text{and} \quad \frac{f(z)-f(-z)}{2\wp'(z)}$$

Since $f(z)$ is equal to the first of these function plus $\wp'(z)$ times the second, the theorem follows from Proposition 7. □

1.4 Differential equation of Weierstrass $\wp$-function

Theorem 3. *$\wp(z)$ satisfies the differential equation*

$$(\wp'(z))^2 = 4(\wp(z)-e_1)(\wp(z)-e_2)(\wp(z)-e_3)$$

where $e_1 = \wp(\omega_1/2), e_2 = \wp(\omega_2/2)$ and $e_3 = \wp((\omega_1+\omega_2)/2)$.

Proof. As seen in the second claim of Proposition 6, $\wp'(z)$ has a triple pole at 0 and three simple zeros, hence as in the proof of Proposition 7 for the even function $(\wp'(z))^2$, there are three a_i's and no b_j's. Therefore we have

$$(\wp'(z))^2 = C(\wp(z)-e_1)(\wp(z)-e_2)(\wp(z)-e_3)$$

where C is some constant. To find C we compare the lowest power of z in the Laurent expansion at the origin. Since $\wp(z) - \frac{1}{z^2}$ is continuous at origin, we have $-\frac{2}{z^3}$ as the leading term of $\wp'(z)$. Hence the leading term of $(\wp'(z))^2$ is $\frac{4}{z^6}$. On the right-hand-side we have $-\frac{C}{z^6}$ as the leading term. We conclude that $C = 4$, thus completing the proof. □

Lemma 2. *We have the following series expansion*

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + x^4 + \dots$$

for $|x| < 1$.

Proof. For $|x| < 1$ we have the geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots$$

Now differentiating both sides we get

$$\frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + 4x^3 + \dots$$

□

Theorem 4. *$\wp(z)$ satisfies the differential equation*

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

where $g_2 = g_2(\Lambda) = 60\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} \frac{1}{\lambda^4}$ and $g_3 = g_3(\Lambda) = 140\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} \frac{1}{\lambda^6}$.

Proof. Firstly we will expand $\wp(z)$ and $(\wp'(z))^2$ near the origin. Since both are even functions, only even powers of z will appear.

Let $c = \min_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}}(|\lambda|)$ and $r < 1$. Assume that z is in the disc of radius rc about the origin. For each non-zero $\lambda \in \Lambda$, we expand the term corresponding to λ in the definition of $\wp(z)$. Using Lemma 2 with $x = \frac{z}{\lambda}$, 1 subtracted from both sides and both sides divided by λ^2, we get the series expansion for the summand of the $\wp(z)$. Hence we have

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} 2\frac{z}{\lambda^3} + 3\frac{z^2}{\lambda^4} + 4\frac{z^3}{\lambda^5} + \ldots + (k+2)\frac{z^{k+1}}{\lambda^{k+3}} + \ldots = \frac{1}{z^2} + \sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}} \frac{2z}{\lambda^3}\sum_{k=0}^{\infty}\frac{(k+2)}{2}\frac{z^k}{\lambda^k} \quad (1.2)$$

Now we write the sum of the absolute values of the terms in the inner sum in the form

$$\left|\sum_{k=0}^{\infty}\frac{(k+2)}{2}\frac{z^k}{\lambda^k}\right| \leq \sum_{k=0}^{\infty}\frac{(k+2)}{2}\frac{|z|^k}{|\lambda|^k} < 1 + \frac{3}{2}r + \frac{4}{2}r^2 + \frac{5}{2}r^3 + \ldots < \frac{1}{(1-r)^2}$$

since $|z| < r|\lambda|$. Thus we have

$$\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}}\left|\frac{2z}{\lambda^3}\sum_{k=0}^{\infty}\frac{(k+2)}{2}\frac{z^k}{\lambda^k}\right| < \frac{2z}{(1-r)^2}\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}}\frac{1}{|\lambda|^3} < \infty$$

using Lemma 1. Hence the double series in (1.2) is absolutely convergent for $|z| < rc$ and we can reverse the order of summation[4] to obtain:

$$\wp(z) = \frac{1}{z^2} + \sum_{k=0}^{\infty}\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}}\frac{(k+2)z^{k+1}}{\lambda^{3+k}} = \frac{1}{z^2} + \sum_{k=3}^{\infty}(k-1)z^{k-2}\sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}}\frac{1}{\lambda^k} \quad (1.3)$$

Now for $k > 2$ we define

$$G_k = G_k(\Lambda) = G_k(\omega_1, \omega_2) = \sum_{\substack{\lambda\in\Lambda\\ \lambda\neq 0}}\frac{1}{\lambda^k} = \sum_{\substack{m,n\in\mathbb{Z}\\ \text{not both zero}}}\frac{1}{(m\omega_1 + n\omega_2)^k}$$

where $G_k = 0$ for odd k since the term for λ cancels the term for $-\lambda$. Using this notation in (1.3), we get

$$\wp(z) = \frac{1}{z^2} + 3G_4z^2 + 5G_6z^4 + 7G_8z^6 + \ldots \quad (1.4)$$

Hence as expected, $\wp(z)$ has only even powers of z in its expansion.

We now use (1.4) to compute the first few terms in the expansions of $\wp'(z)$, $(\wp'(z))^2$, $(\wp(z))^2$ and $(\wp(z))^3$ as follows:

$$\wp'(z) = -\frac{2}{z^3} + 6G_4z + 20G_6z^3 + 42G_8z^5 + \ldots \quad (1.5)$$

$$(\wp'(z))^2 = \frac{4}{z^6} - 24G_4\frac{1}{z^2} - 80G_6 + (36G_4^2 - 168G_8)z^2 + \ldots \quad (1.6)$$

$$(\wp(z))^2 = \frac{1}{z^4} + 6G_4 + 10G_6z^2 + \ldots \quad (1.7)$$

$$(\wp(z))^3 = \frac{1}{z^6} + 9G_4\frac{1}{z^2} + 15G_6 + (21G_8 + 27G_4^2)z^2 + \ldots \quad (1.8)$$

Suppose that we can find a cubic polynomial $f(z) = ax^3 + bx^2 + cx + d$ such that the Laurent series expansion at 0 of the elliptic function $f(\wp(z))$ agrees with the Laurent expansion of

[4] For more details, see the note at the end of Chapter 7 of [4, pp. 197].

$(\wp'(z))^2$ through the negative powers of z. Then by Proposition 3, this difference is constant; and if we suitably choose d, we can make this constant zero. Hence we just need to find a, b, c, d such that

$$(\wp'(z))^2 = a(\wp(z))^3 + b(\wp(z))^2 + c\wp(z) + d$$

for the negative powers of the Laurent expansions of both sides. If we multiply (1.8) by a, (1.7) by b, (1.4) by c, and then add them all to the constant d, and finally equate the coefficients of z^{-6}, z^{-4} and z^{-2}, and the constant term to the corresponding coefficients in (1.6), we obtain:

$$\begin{cases} a = 4; \\ b = 0; \\ a(9G_4) + c = -24G_4; \\ a(15G_6) + b(2G_4) + d = -80G_6 \end{cases}$$

Thus, $c = -60G_4 = -60G_4(\Lambda) = -g_2(\Lambda) = -g_2$ and $d = -140G_6 = -140G_6(\Lambda) = -g_3(\Lambda) = -g_3$. We have thereby derived a second form of differential equation for $\wp(z)$. □

Corollary 3. *All G_k's can be expressed as polynomials in G_4 and G_6 with rational coefficients.*

Proof. By the above theorem we have

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6$$

Now using (1.6), (1.8) and (1.4) in the above equation and comparing the coefficients of z^2 on both sides we get $G_8 = \frac{3}{7}G_4^2$. Hence, $G_8 \in \mathbb{Q}[G_4, G_6]$. Next we observe that to find the expression for G_k we just need to compare the coefficients for z^{k-6} on both sides and these coefficients are the combinations of G_ℓ's where ℓ is an even number not greater than k. Hence by induction on k the result follows. □

Proposition 8. *The map $\vartheta : \mathbb{C}/\Lambda \to \mathbb{P}^2_{\mathbb{C}}$ defined as*

$$z \longmapsto \begin{cases} (\wp(z), \wp'(z), 1) & \text{for } z \neq 0 \\ (0, 1, 0) & \text{for } z = 0 \end{cases}$$

is an analytic one-to-one correspondence between the torus $\mathbb{C}/\Lambda$ and the elliptic curve $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$ in the complex projective plane[5] $\mathbb{P}^2_{\mathbb{C}}$.

Proof. The image of any non-zero point z of $\mathbb{C}/\Lambda$ is a point in the xy-plane (with complex coordinates) whose x- and y-coordinates satisfy the relationship $y^2 = f(x)$, where $f(x) = 4x^3 - g_2x - g_3$ (Theorem 4). Here $f(x) \in \mathbb{C}[x]$ is a cubic polynomial with distinct roots e_1, e_2 and e_3 (Theorem 3). Thus, every point $z \in \mathbb{C}/\Lambda$ maps to a point on the elliptic curve $y^2 = f(x)$ in $\mathbb{P}^2_{\mathbb{C}}$.

Now as in Remark 8, every x-value except the roots of $f(x)$ (and infinity) has precisely two z's such that $\wp(z) = x$. The y-coordinates $y = \wp'(z)$ coming from two z's are the two square roots of $f(x) = f(\wp(z))$. If, however, x happens to be a root of $f(x)$ then there is only one z value such that $\wp(z) = x$, and the corresponding y-coordinate is $y = \wp'(z) = 0$, so that again we are getting the solutions to $y^2 = f(x)$ for our given x. Hence, this map ϑ is a one-to-one correspondence between the torus and the elliptic curve (including the point at infinity).

Moreover, the map from $\mathbb{C}/\Lambda$ to the elliptic curve in $\mathbb{P}^2_{\mathbb{C}}$ is analytic, since near any point of $z \in \mathbb{C}/\Lambda$, $\vartheta(z)$ is given by a triplet of analytic functions. Near non-lattice points of $\mathbb{C}$ the maps is given by $z \mapsto (\wp(z), \wp'(z), 1)$ and near lattice points the map is given by $z \mapsto (\wp(z)/\wp'(z), 1, 1/\wp'(z))$ which is a triplet of analytic function near Λ. □

[5]For a short discussion on elliptic curves and projective planes, see [19, §1.8].

Chapter 2

Full modular group and its congruence subgroups

"The modular group takes its name form the fact that the points of the quotient space $SL_2(\mathbb{Z})\backslash\mathbb{H}$ are moduli (=parameters) for the isomorphism classes of elliptic curves over $\mathbb{C}$."

— Don Zagier, *The 1-2-3 of Modular forms*

2.1 Introduction

Definition 5 (Special linear group of degree 2 over a ring)**.** Let R be a commutative ring then the *special linear group of degree 2 over* R, $SL_2(R)$, is the set of 2×2 matrices with entries from R such that their determinant is equal to 1, along with multiplication of matrices as the group operation. That is,

$$SL_2(R) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad - bc = 1; a, b, c, d \in R \right\}$$

Definition 6 (Riemann sphere)**.** The set of extended complex numbers, denoted by $\mathbb{C}_\infty$, consist of the complex numbers together with ∞, i.e., $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$. Geometrically, the set of extended complex numbers is referred to as the *Riemann sphere* (or extended complex plane), equivalently the complex projective line $\mathbb{P}^1_{\mathbb{C}}$.

Definition 7 (Fractional linear transformation)**.** *Fractional linear transformation of the Riemann sphere* is the (left) group action of $SL_2(\mathbb{R})$ on the set $\mathbb{C}_\infty$. Given an element $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$ we define the group action $\varphi : SL_2(\mathbb{R}) \times \mathbb{C}_\infty \longrightarrow \mathbb{C}_\infty$ as:

$$(g, z) \longmapsto \begin{cases} \frac{az+b}{cz+d} & \text{for } z \in \mathbb{C} \\ \frac{a}{c} & \text{for } z = \infty \end{cases}$$

where $\left(g, \frac{-d}{c}\right) \mapsto \infty$ and if $c = 0$ then $(g, \infty) \mapsto \infty$. For the ease of writing, we will denote $\varphi(g, z)$ by $g \cdot z$.

Remark 9. One can easily check that the (left) group action axiom $g_1 g_2 \cdot z = g_1 \cdot (g_2 \cdot z)$, i.e.

$$\varphi(g_1 g_2, z) = \varphi(g_1, \varphi(g_2, z)),\ \forall g_1, g_2 \in SL_2(\mathbb{R}), z \in \mathbb{C}_\infty$$

is satisfied.

Definition 8 (Projective special linear group of degree 2 over $\mathbb{R}$)**.** Let $I \in SL_2(\mathbb{R})$ be the identity matrix, then the quotient group $SL_2(\mathbb{R})/\{I, -I\}$ acts faithfully on $\mathbb{C}$, i.e. each element other than the identity acts nontrivially. This quotient group is called *projective special linear group of degree 2 over* $\mathbb{R}$ and is denoted as $PSL_2(\mathbb{R})$.

Remark 10. One can easily check that $\pm I \in SL_2(\mathbb{R})$ are the only matrices which act trivially on $\mathbb{C}_\infty$, since $g \cdot z = z$ (for all z) implies that g is a scalar matrix of determinant 1.

2.2 Upper half-plane

Definition 9 (Upper half-plane)**.** The *upper half-plane* $\mathbb{H}$ is the set of complex numbers with positive imaginary part:

$$\mathbb{H} = \{z \in \mathbb{C} \ : \mathrm{Im}(z) > 0\}$$

Remark 11. The group $SL_2(\mathbb{R})$ acts on the set $\mathbb{H}$ by the fractional linear transformations since any $g \in SL_2(\mathbb{R})$ preserves $\mathbb{H}$. That is, $\mathrm{Im}(z) > 0$ implies $\mathrm{Im}(g \cdot z) > 0$ because

$$\mathrm{Im}(g \cdot z) = \frac{\mathrm{Im}(z)}{|cz + d|^2} \text{ for } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$$

Definition 10 (Full modular group)**.** The subgroup $SL_2(\mathbb{Z})$ of $SL_2(\mathbb{R})$ consisting of matrices with integer entries is called the *full modular group* and is denoted by Γ.

Remark 12. The group $SL_2(\mathbb{Z})$ is also called the *homogeneous modular group* [3, pp. 281].

Definition 11 (Modular group)**.** Let $I \in \Gamma$ be the identity matrix, then the quotient group $\Gamma/\{I, -I\}$ acts faithfully on $\mathbb{H}$, i.e. each element other than the identity acts nontrivially. This quotient group is called the *modular group* and is denoted as $\overline{\Gamma}$.

Remark 13. In general, if G is a subgroup of $SL_2(\mathbb{R})$, then $\overline{G}$ denotes $G/\{I, -I\}$ if $-I \in G$, otherwise if $-I \notin G$ then $\overline{G} = G$.

Definition 12 (Principal congruence subgroup of level N)**.** Let N be an integer greater than or equal to 1, we associate to N a reduction homomorphism of the full modular group:

$$\begin{aligned} \pi_N : & SL_2(\mathbb{Z}) \longrightarrow SL_2(\mathbb{Z}/N\mathbb{Z}) \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \begin{pmatrix} \overline{a} & \overline{b} \\ \overline{c} & \overline{d} \end{pmatrix} \end{aligned}$$

The kernel of π_N is called the *principal congruence subgroup of level* N, it is written $\Gamma(N)$. Thus

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

Remark 14. $\Gamma(N)$ is a normal subgroup of Γ. Moreover, one can prove that π_N is an onto map [15, Theorem 3.2], and hence $\Gamma/\Gamma(N) \cong SL_2(\mathbb{Z}/N\mathbb{Z})$.

Definition 13 (Congruence subgroup of level N)**.** A subgroup of Γ (or of $\overline{\Gamma}$) is called a *congruence subgroup of level* N is it contains $\Gamma(N)$ (or $\overline{\Gamma}(N)$, if we were considering matrices modulo $\pm I$).

Remark 15. $\overline{\Gamma}(2) = \Gamma(2)/\{I, -I\}$, whereas for $N > 2$ we have $\overline{\Gamma}(N) = \Gamma(N)$ because $-1 \not\equiv 1 \pmod{N}$, and so $-I \notin \Gamma(N)$.

Remark 16. A congruence subgroup of level N also has level N' for any multiple N' of N, because $\Gamma(N) \supset \Gamma(N')$.

Definition 14 (Congruence subgroup)**.** A subgroup Γ' of Γ is called a *congruence subgroup* if there exists a natural number $N \geq 1$ such that $\Gamma' \supset \Gamma(N)$.

Remark 17. Not all subgroups of Γ are congruence subgroups, see [15, §4] for examples of non-congruence subgroups. But in this report we shall never deal with non-congruence subgroups.

Definition 15 (Hecke congruence subgroup of level N)**.** The *Hecke congruence subgroup of level* N is given by

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \ : \ c \equiv 0 \pmod{N} \right\}$$

Remark 18. Since Hecke subgroup contains $\{I, -I\}$, we have $\overline{\Gamma_0}(N) = \Gamma_0(N)/\{I, -I\}$.

Remark 19. Another important congruence subgroup is

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N) : a \equiv 1 \pmod{N} \right\}$$

Then since $ad - bc = 1$ and $N|c$, it follows that $ad \equiv 1 \pmod{N}$, and hence $d \equiv 1 \pmod{N}$.

Definition 16 (Γ'-equivalent)**.** If Γ' is a subgroup of Γ, we say two points $z_1, z_2 \in \mathbb{H}$ are *Γ'-equivalent* if there exists $\gamma \in \Gamma'$ such that $z_2 = \gamma \cdot z_1$.

Remark 20. Whenever a group acts on a set, it divides set into equivalence classes, where two points of the set are said to be in the same equivalence class if there is an element of the group which takes one to the other.

2.3 Fundamental domain of the full modular group

Definition 17 (Fundamental domain)**.** A closed region[1] F in $\mathbb{H}$ is a *fundamental domain for the subgroup* Γ' *of* Γ is every $z \in \mathbb{H}$ is Γ'-equivalent to a point in F, but no two distinct points z_1, z_2 in the interior[2] of F are Γ'-equivalent.

Remark 21. In a fundamental domain, two boundary[3] points are permitted to be Γ'-equivalent.

Remark 22. Hence, each point in the space $\mathbb{H}$ has a point of its orbit[4] by $SL_2(\mathbb{Z})$ in the fundamental domain F and the only points in the fundamental domain that lie in the same orbit are on the boundary [15, §2].

Theorem 5. *The region $\mathcal{F}$ defined as*

$$\mathcal{F} = \left\{ z \in \mathbb{H} : -\frac{1}{2} \leq \mathrm{Re}(z) \leq \frac{1}{2} \text{ and } |z| \geq 1 \right\}$$

is a fundamental domain for Γ.

Proof. Consider the following two fractional linear transformations in Γ

$$s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

hence $s \cdot z = \frac{-1}{z}$ and $t \cdot z = z + 1$.

Note the $\mathcal{F}$ consists of the points outside the unit circle and inside the unit strip symmetrical around the imaginary axis.

[1] A connected set is called region. Usually, F will also be simple connected.

[2] Let S be a subset of a topological space X. The interior of a set S is the set of all interior points of S, where $x \in X$ is an interior point of S if x is contained in an open subset of S.

[3] Let S be a subset of a topological space X. Then the boundary of S is what is left after removing the interior of S from the closure of S.

[4] Let G act on X, then for each $x \in X$, its orbit is $\mathrm{Orb}_x = \{g \cdot x : g \in G\} \subset X$. The orbit of a point is a geometric concept: it is the set of places where the point can be moved by the group action [14, §3].

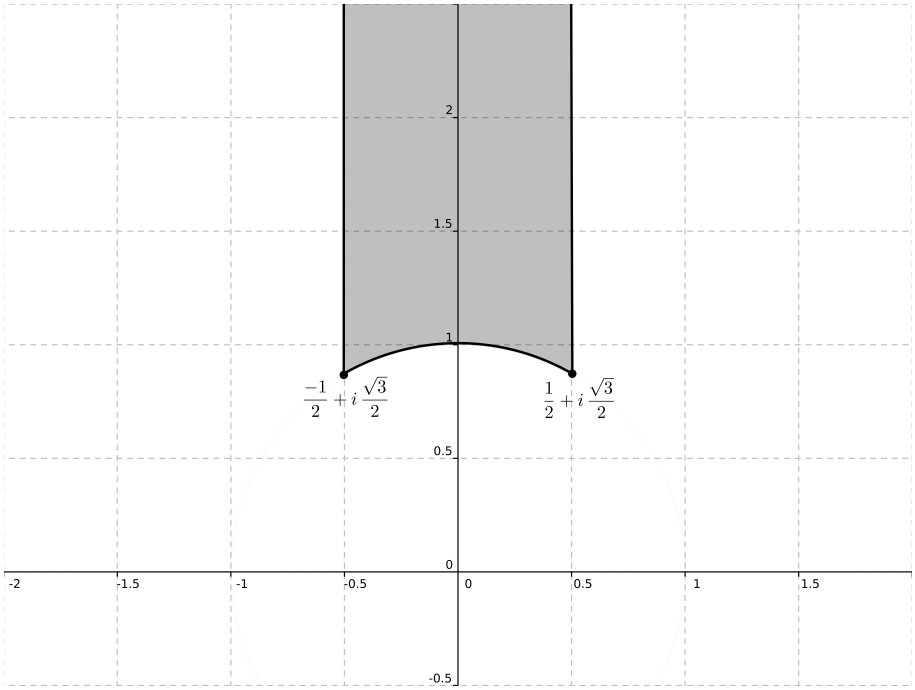


Figure 2.1: A fundamental domain $\mathcal{F}$ of Γ. Drawn using a: $((-0.5 \leq x) \wedge (x \leq 0.5) \wedge (x^2 + y^2 \geq 1) \wedge (y \geq 0))$ in GeoGebra 4.0.34.0.

Claim 1. *For a given $z \in \mathbb{H}$ there exists $\gamma \in \widetilde{\Gamma} = \langle s, t \rangle$ such that $\gamma \cdot z \in \mathcal{F}$*

Let $z \in \mathbb{H}$ be fixed. Then $\widetilde{\Gamma} = \langle s, t \rangle$ is a subgroup of Γ generated by s and t. If $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \widetilde{\Gamma}$, then since $ad - bc = 1$ we have

$$\operatorname{Im}(\gamma \cdot z) = \frac{\operatorname{Im}(z)}{|cz + d|^2} \tag{2.1}$$

Since $c, d \in \mathbb{Z}$ such that $ad - bc = 1$, we have $|cz + d| > 0$. Thus there exists a $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \widetilde{\Gamma}$ such that $\operatorname{Im}(\gamma \cdot z)$ is maximal. Without loss of generality, we may assume that $\gamma \cdot z$ is in the strip $-\frac{1}{2} \operatorname{Re}(\gamma \cdot z) \leq \frac{1}{2}$, since otherwise we can replace γ by $t^\beta \gamma \in \widetilde{\Gamma}$ for some suitable β. But then if $\gamma \cdot z$ won't be in $\mathcal{F}$, i.e. $|\gamma \cdot z| < 1$ then we would have $s\gamma \in \widetilde{\Gamma}$ such that as in (2.1)

$$\operatorname{Im}(s\gamma \cdot z) = \operatorname{Im}(s \cdot (\gamma \cdot z)) = \frac{\operatorname{Im}(\gamma \cdot z)}{|\gamma \cdot z + 0|^2} > \operatorname{Im}(\gamma \cdot z)$$

contradicting our choice of $\gamma \in \widetilde{\Gamma}$ so the $\operatorname{Im}(\gamma \cdot z)$ is maximal. Thus there exists $\gamma \in \widetilde{\Gamma}$ such that $\gamma \cdot z \in \mathcal{F}$.

Claim 2. *No two distinct points in the interior of $\mathcal{F}$ are Γ-equivalent.*

Let $z_1, z_2 \in \mathcal{F}$ (not necessarily distinct) be Γ-equivalent, i.e. there exist $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ (can be identity) such that $z_2 = \gamma \cdot z_1$. Without loss of generality, we may assume that $\operatorname{Im}(z_2) \geq \operatorname{Im}(z_1)$, otherwise we multiply by γ^{-1} on both sided of the equivalence relation. Then as in (2.1)

$$\operatorname{Im}(z_2) = \operatorname{Im}(\gamma \cdot z_1) = \frac{\operatorname{Im}(z_1)}{|cz_1 + d|^2} \geq \operatorname{Im}(z_1)$$

hence we must have $|cz_1 + d| \leq 1$. Since $|z_1| \geq 1$ and $c, d \in \mathbb{Z} \subset \mathbb{R}$ we must have c and d to be $0, 1$ or -1. So we are left with following cases:

(i) $c = 0, d = \pm 1$ *and* $|z_1| \geq 1$

Either γ or $-\gamma$ is a translation $t^\beta = \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}$. But such a γ will take a point in $\mathcal{F}$ to another point in $\mathcal{F}$ if and only if $\beta = 0$ (i.e. γ is $\pm I$) or $\beta = \pm 1$ and $\mathrm{Re}(z_1) = \pm\frac{1}{2}$. Hence either $z_1 = z_2$ or z_1 and z_2 lie on the vertical lines (boundary of $\mathcal{F}$).

(ii) $c = \pm 1, d = 0$ *and* $|z_1| = 1$

Then we have $\gamma = \begin{pmatrix} \pm\beta & \mp 1 \\ \pm 1 & 0 \end{pmatrix} = \pm \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \pm t^\beta s$ for some $\beta \in \mathbb{Z}$. Since z_1 is on boundary, s will flip z_1 symmetrically about imaginary axis. But such a γ will take a point in $\mathcal{F}$ to another point in $\mathcal{F}$ if and only if $\beta = 0$ and z_1, z_2 are on the unit circle (located symmetrically with respect to imaginary axis) or $\beta = \pm 1$ and $\mathrm{Re}(z_1) = \pm\frac{1}{2}$. Hence z_1 and z_2 are located symmetrically on the unit circle (boundary of $\mathcal{F}$).

(iii) $c = d = \pm 1, \mathrm{Re}(z_1) = \frac{-1}{2}$ *and* $|z_1| = 1$

Then we have $\gamma = \begin{pmatrix} \pm\beta & \pm\beta \mp 1 \\ \pm 1 & \pm 1 \end{pmatrix} = \pm \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \pm t^\beta st$ for some $\beta \in \mathbb{Z}$. Since z_1 is on boundary, s will flip z_1 symmetrically about imaginary axis. But such a γ will take a point in $\mathcal{F}$ to another point in $\mathcal{F}$ if and only if $\beta = 0$ or $\beta = 1$. Hence either $z_1 = z_2 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ or $z_1 = z_2 - 1 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ (boundary of $\mathcal{F}$).

(iv) $c = -d = \pm 1, \mathrm{Re}(z_1) = \frac{1}{2}$ *and* $|z_1| = 1$

Then we have $\gamma = \begin{pmatrix} \pm\beta & \mp\beta \mp 1 \\ \pm 1 & \mp 1 \end{pmatrix} = \pm \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ hence $\gamma = \pm t^{\beta+1} sts$ for some $\beta \in \mathbb{Z}$. Since z_1 is on boundary, s will flip z_1 symmetrically about imaginary axis. But such a γ will take a point in $\mathcal{F}$ to another point in $\mathcal{F}$ if and only if $\beta = 0$ or $\beta = -1$. Hence either $z_1 = z_2 = \frac{1}{2} + i\frac{\sqrt{3}}{2}$ or $z_1 = z_2 - 1 = \frac{1}{2} + i\frac{\sqrt{3}}{2}$ (boundary of $\mathcal{F}$).

Combining the above four cases we complete the proof of our claim.

□

Remark 23. A key step in the proof of second claim of the above theorem was to be able to write γ in terms of s and t. The above factorization was done by computing st, $(st)^2$, ts and $(ts)^2$ before hand and comparing them with the matrix left after we bring out the factor t^β. For more details see, [15, Example 2.1].

Corollary 4. *Two distinct points z_1, z_2 on the boundary of $\mathcal{F}$ are Γ-equivalent only if $\mathrm{Re}(z_1) = \pm\frac{1}{2}$ and $z_2 = z_1 \pm 1$, or if z_1 is on the unit circle and $z_2 = -\frac{1}{z_1}$ (i.e. symmetrically located with respect to the imaginary axis).*

Proof. Follows from conclusions of the four cases discussed in second claim of the above theorem. □

Definition 18 (Quotient space of Γ on $\mathbb{H}$)**.** The set of Γ-equivalence classes[5] in $\mathbb{H}$ is called the *quotient space of* Γ *on* $\mathbb{H}$ and is denoted[6] by $\Gamma\backslash\mathbb{H}$.

[5]Given a set X and an equivalence relation $\sim$ on X, the *equivalence class* of an element $x \in X$ is the set $\{y \in X \mid y \sim x\}$ of elements which are equivalent to x. Further, the equivalence classes form a *partition* of X. This partition - the set of equivalence classes - is sometimes called the *quotient set* or the *quotient space* of X by $\sim$ and is denoted by $X/\sim$, as done in Remark 3.

[6]Here the set is $\mathbb{H}$ and the equivalence relation is the group action of Γ. The orbits of a group action on a set are called the *quotient space of the action on the set*, particularly when the orbits of the group action are the right

Remark 24. Visually, we identify the Γ-equivalent point on the boundary of $\mathcal{F}$, i.e. we fold $\mathcal{F}$ around the imaginary axis, gluing the point $\frac{1}{2}+iy$ to $-\frac{1}{2}+iy$ and the point $e^{2\pi i\theta}$ to $e^{2\pi i\left(\frac{1}{2}-\theta\right)}$ for $y \geq \frac{\sqrt{3}}{2}$ and $\frac{1}{6} \leq \theta \leq \frac{1}{3}$. The resulting set $\mathcal{F}$ with its sides glued is in one-to-one correspondence with the set of Γ-equivalence classes in $\mathbb{H}$, $\Gamma\backslash\mathbb{H}$.

Definition 19 (Isotropy subgroup)**.** Let G act on X. For each $x \in X$, its *isotropy subgroup* is $G_x = \{g \in G : g \cdot x = x\} \subset G$.

Remark 25. G_x is also called the stabilizer of x. The stabilizer of a point is an algebraic concept: it is the set of group elements that fix the point [14, §3].

Definition 20 (Negative-reciprocal map)**.** The fractional linear transformation s whose action on elements of $z \in \mathbb{H}$ is given by $s \cdot z = -\frac{1}{z}$, i.e. $s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and s is of order 4 in Γ.

Definition 21 (Translation by 1)**.** The fractional linear transformation t whose action on elements of $z \in \mathbb{H}$ is given by $t \cdot z = z + 1$, i.e. $t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, and t is of infinite order in Γ.

Corollary 5. *If $z \in \mathcal{F}$, then*

$$\Gamma_z = \begin{cases} \pm\{I, s\} & \text{if } z = i \\ \pm\{I, st, (st)^2\} & \text{if } z = -\frac{1}{2} + i\frac{\sqrt{3}}{2} = e^{\frac{2\pi i}{3}} \\ \pm\{I, ts, (ts)^2\} & \text{if } z = \frac{1}{2} + i\frac{\sqrt{3}}{2} = e^{\frac{\pi i}{3}} \\ \pm\{I\} & \text{otherwise} \end{cases}$$

Proof. Follows from factorization of matrices done in the four cases of the second claim of the above theorem. □

Remark 26. Note that, $s^2 = -I$, $(st)^3 = -I$ and $(ts)^3 = -I$. Thus,in $\overline{\Gamma}$ the elements s, st and ts generate cyclic subgroups of order $2, 3$ and 3 respectively. Moreover, these subgroups are isotropy subgroups of $i, e^{\frac{2\pi i}{3}}$ and $e^{\frac{\pi i}{3}}$ respectively.

Definition 22 (Elliptic points)**.** The points in $\mathbb{H}$ with non-trivial isotropy subgroups are called *elliptic points.*

Remark 27. From the above corollary we can conclude that $i, e^{\frac{2\pi i}{3}}$ and $e^{\frac{\pi i}{3}}$ are the examples of elliptic points.

Proposition 9. *The modular group $\overline{\Gamma}$ is generated by the two elements s and t.*

Proof. Let $\widetilde{\Gamma} = \langle s, t\rangle$ be the subgroup of Γ generated by s and t. Let z be any point in the interior of $\mathcal{F}$ and g be any element of Γ. Consider the point $g \cdot z \in \mathbb{H}$. From the claim 1 of Theorem 5 we know that there exists $\gamma \in \widetilde{\Gamma}$ such that $\gamma \cdot (g \cdot z) = \gamma g \cdot z \in \mathcal{F}$. But since z is in the interior of $\mathcal{F}$, by claim 2 of Theorem 5 and Corollary 5 it follows that $\gamma g = \pm I$, i.e. $g = \pm\gamma^{-1} \in \widetilde{\Gamma}$. This shows that any $g \in \Gamma$ is actually (upto a sign) in $\widetilde{\Gamma}$. Hence completing the proof. □

Remark 28. Thus, any element of $\overline{\Gamma}$ can be written in the form $s^{\alpha_1}t^{\beta_1}s^{\alpha_2}t^{\beta_2}\cdots s^{\alpha_\ell}t^{\beta_\ell}$, where all α_j, β_j are non-zero integers, except that we allow α_1 and/or β_1 to be zero. Since $s^2 = -I$, we may suppose that all of the α_j equal 1, except that $\alpha_1 = 0$ or 1.

cosets of a subgroup of a group, which arise from the action of the subgroup on the group by left translations, or respectively the left cosets as orbits under right translation, see [14, Example 2.10]. The notation $\Gamma\backslash\mathbb{H}$ rather than $\mathbb{H}/\Gamma$ is customarily used because the group Γ acts on the set $\mathbb{H}$ on the left. Compare it with the notation for torus in Remark 3.

2.4 Extended upper half-plane

Definition 23 (Extended upper half-plane)**.** We add to $\mathbb{H}$ a point at infinity and also all the rational numbers on the real axis, the set so formed is called the *extended upper half plane* and is denoted by $\overline{\mathbb{H}}$, i.e.

$$\overline{\mathbb{H}} = \mathbb{H} \cup \{\infty\} \cup \mathbb{Q}$$

Remark 29. The point at infinity should be visualized far up the positive imaginary axis; for this reason sometimes it's denoted by $i\infty$.

Definition 24 (Cusp)**.** The points $\{\infty\} \cup \mathbb{Q}$ are called *cusps*.

Remark 30. The full modular group Γ permutes the cusps transitively. That is, any fraction a/c in lowest terms can be completed to a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ by solving $ad - bc = 1$ for b and d; this matrix takes ∞ to a/c. Hence all rational numbers are in the same Γ-equivalence class as ∞.

If Γ' is a subgroup of Γ, then it also permutes the cusps but in general not transitively. That is, there is usually more than one Γ'-equivalence class among the cusps $\{\infty\} \cup \mathbb{Q}$.

Definition 25 (Cusp of Γ')**.** Let Γ' be a subgroup of Γ, then a *cusp of* Γ' is a Γ'-equivalence class of cusps.

Remark 31. We may choose any convenient representative of the equivalence class to denote the cusp. Thus, we say that "Γ has single cusp at ∞", where ∞ can be replaced by any rational number a/c in this statement.

Definition 26 (Topology of $\overline{\mathbb{H}}$)**.** The neighbourhood system [7] on $\overline{\mathbb{H}}$ is the collection $\{\mathcal{N}_z : z \in \overline{\mathbb{H}}\}$ where

1. for $z \in \mathbb{H}$ we define $\mathcal{N}_z = \{B(z,\delta) \cap \mathbb{H} : \delta > 0\}$ such that $B(z,\delta) = \{w \in \mathbb{C} : |w - z| < \delta\}$;
2. at ∞ we define $\mathcal{N}_\infty = \{N_\delta : \delta > 0\}$ such that $N_\delta = \{z \in \mathbb{H} : \mathrm{Im}(z) > \delta\} \cup \{\infty\}$; and
3. near a cusp $a/c \in \mathbb{Q} \subset \overline{\mathbb{H}}$ we define $\mathcal{N}_{a/c} = \{\gamma \cdot N_\delta : \delta > 0\}$ such that $N_\delta = \{z \in \mathbb{H} : \mathrm{Im}(z) > \delta\} \cup \{\infty\}$ and γ is obtained by completing a, c to a matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, i.e. γ transports N_δ to open disc which is tangent to the real axis at a/c.

Then $\mathcal{T} = \{U : z \in U \Rightarrow \exists V \in \mathcal{N}_z \text{ such that } V \subset U\}$ defines the topology[8] on $\overline{\mathbb{H}}$.

Remark 32. With the above topology, a sequence $\{z_j\}$ in $\overline{\mathbb{H}}$ approaches a/c means that the sequence $\{\gamma^{-1} \cdot z_j\}$ in $\overline{\mathbb{H}}$ approaches $i\infty$, i.e. the sequence $\{\mathrm{Im}(\gamma^{-1} \cdot z_j)\}$ in $\mathbb{H}$ approaches infinity in the usual sense. Hence, remember that the topology near the rational numbers a/c does not agree with the usual topology on the real line, i.e. a sequence of rational numbers which approaches a/c as real numbers will not approach a/c in the above topology.

Proposition 10. *Let $\mathbb{D}$ be a unit open disc centred at origin, i.e. $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$. Then the map $f : \mathbb{H} \cup \{\infty\} \to \mathbb{D}$ defined as*

$$f(z) = \begin{cases} e^{2\pi i z} & \text{if } z \in \mathbb{H} \\ 0 & \text{if } z = \infty \end{cases} \tag{2.2}$$

is continuous with respect to the topology defined above.

[7] For the definition and discussion see [5, Definition IX.4.16].
[8] For details see [5, Proposition IX.4.17].

Proof. Firstly, the map is well defined, since for $z \in \mathbb{H}$, i.e. $\operatorname{Im}(z) > 0$, we have $|f(z)| = |e^{2\pi i(\operatorname{Re}(z)+i\operatorname{Im}(z))}| = |e^{-2\pi \operatorname{Im}(z)}| < 1$. Secondly, N_δ is the inverse image of the open disk of radius $e^{-2\pi\delta}$ centred at origin, hence the inverse image of open set in $\mathbb{D}$ is an open set in $\mathbb{H} \cup \{\infty\}$. Since a function is said to be continuous if inverse image of every open set is open, f is continuous. □

Remark 33. In future discussions we will denote $e^{2\pi iz}$ by q, i.e. $f(z) = q$ for all $z \in \mathbb{H}$.

Definition 27 (Analytic structure on $\mathbb{H} \cup \{\infty\}$)**.** Given a function f on $\mathbb{H}$ of period 1 (i.e. $f(z) = q$ for $z \in \mathbb{H}$), it can be expressed as a power series[9] in the variable[10] q, i.e. it has a Fourier expansion of the form

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

then we say that $f(z)$

1. is *meromorphic at* ∞ if $a_n = 0$ for $n \ll 0$, i.e. have at most finitely many negative terms;
2. is *holomorphic at* ∞ if $a_n = 0$ for $n < 0$;
3. *vanishes at* ∞ if $a_n = 0$ for $n \leq 0$;

Remark 34. The above definition can be extended to function $f(z)$ of period N, i.e. $f(z) = q_N = e^{\frac{2\pi iz}{N}}$ for $z \in \mathbb{H}$, by replacing q by q_N in the Fourier expansion.

Definition 28 (Quotient space of Γ on $\overline{\mathbb{H}}$)**.** The set of Γ-equivalence classes in $\overline{\mathbb{H}}$ is called the *quotient space of* Γ *on* $\overline{\mathbb{H}}$ and is denoted by $\Gamma\backslash\overline{\mathbb{H}}$.

Remark 35. Let $\overline{\mathcal{F}}$ denote the fundamental domain $\mathcal{F}$ with Γ-equivalent boundary points identified (as in Remark 24) and with the cusp ∞ thrown in. Thus the points of $\overline{\mathcal{F}}$ are in one-to-one correspondence with Γ-equivalence classes in $\overline{\mathbb{H}}$, $\Gamma\backslash\overline{\mathbb{H}}$.

Definition 29 (Topology on $\overline{\mathcal{F}}$)**.** The topology on $\overline{\mathcal{F}}$ comes from the topology on $\overline{\mathbb{H}}$. That is, by a small disc around an interior point of $\mathcal{F}$ we mean a disc in the usual sense; by a small disc around ∞ we mean all points lying about the line $\operatorname{Im}(z) = \delta$, where δ is large; by a small disc around a boundary point $-\frac{1}{2} + iy$ we mean half-disc contained in $\mathcal{F}$ together with the half disc of same radius around $\frac{1}{2} + iy$ which is contained in $\mathcal{F}$; and so on.

Remark 36. Thus, $\overline{\mathcal{F}}$ has an analytic structure coming from the usual structure on $\mathbb{H}$, except at infinity where it comes from the usual structure at 0 after the change of variable (2.2).

2.5 Fundamental domain of the subgroups of full modular group

Theorem 6. *Let Γ' be a subgroup of Γ and the index of Γ' in Γ is n, i.e. $[\Gamma : \Gamma'] = n < \infty$, so that Γ can be written as a disjoint union of n cosets of Γ', i.e. $\Gamma = \bigcup_{i=1}^n \gamma_i\Gamma'$ with $\gamma_i \in \Gamma$. Then $\mathcal{F}' = \bigcup_{i=1}^n \gamma_i^{-1} \cdot \mathcal{F}$ is a fundamental domain for Γ'.*

Proof. We will divide the proof into two parts

[9] Power series with negative powers are called Laurent series. If $z = z_0$ is an isolated singularity of f and $f(z) = \sum_{-\infty}^{\infty} b_n(z - z_0)^n$ is its Laurent series expansion in $\{z : 0 < |z - z_0| < R\}$. Then $z = z_0$ is a pole of order m if and only if $b_{-m} \neq 0$ and $b_n = 0$ for $n \leq -(m + 1)$, or $z = z_0$ is a removable singularity if and only if $b_n = 0$ for $n \leq -1$. For details, see [5, Corollary V.1.18; Theorem V.1.2; Definition V.1.3].

[10] Suppose f is meromorphic in $\{z : 0 < |z - z_0| < R\}$, so that f has a power series expansion $f(z) = \sum_{n=-\infty}^{\infty} b_n(z - z_0)^n$ that converges in that annulus. Then the coefficients of the power series expansion of f are given by $b_n = \frac{1}{2\pi r^n}\int_0^{2\pi} f(z_0 + re^{i\theta})e^{-in\theta}d\theta$, i.e. $f(z) = \sum_{n=-\infty}^{\infty} \hat{f}(n)q^n$. For details, see [4, Theorem 3.7.1]. Note that in general, Fourier series don't converge to the original function.

Claim 1. *Every point $z \in \mathbb{H}$ is Γ'-equivalent to a point in $\mathcal{F}'$*

Let $z \in \mathbb{H}$. Since $\mathcal{F}$ is a fundamental domain for Γ, we can find $\gamma \in \Gamma$ such that $\gamma \cdot z \in \mathcal{F}$. Then for some i we have $\gamma = \gamma_i \gamma'$ with $\gamma' \in \Gamma'$, and hence $\gamma' \cdot z \in \gamma_i^{-1} \cdot \mathcal{F} \subset \mathcal{F}'$, as desired.

Claim 2. *No two distinct points in interior of $\mathcal{F}'$ are Γ'-equivalent*

On the contrary, suppose two distinct interior points of $\mathcal{F}'$ are Γ'-equivalent, i.e there exists $\gamma' \in \Gamma'$ (not identity), and some interior point $z \in \mathcal{F}'$ such that $\gamma' \cdot z \in \mathcal{F}'$ is also an interior point. The action of γ' is a bi-continuous map[11] on $\mathbb{H}$, so there exists an open set $U \subset \mathcal{F}'$ containing z such that $\gamma' \cdot z \in \gamma' \cdot U \subset \mathcal{F}'$.

By construction of $\mathcal{F}'$, there exists indices $1 \leq i, j \leq n$ such that $\gamma_i \cdot z \in \mathcal{F}$ and $\gamma_j \cdot (\gamma' \cdot z) = \gamma_j \gamma' \cdot z \in \mathcal{F}$. Since the action of γ_i is a bi-continuous map on $\mathbb{H}$, $\gamma_i \cdot U$ is an open set with non-empty intersection with $\mathcal{F}$. Let $V = \gamma_i \cdot U \cap \mathcal{F}^\circ$, where $\mathcal{F}^\circ$ denotes the interior of $\mathcal{F}$. Then the set V is open and non-empty. Take any $z_0 \in V$, then

$$\gamma' \gamma_i^{-1} \cdot z_0 \in \gamma' \gamma_i^{-1} \cdot V \subseteq \gamma' \cdot U \subset \mathcal{F}'$$

Let k be such that $\gamma' \gamma_i^{-1} \cdot z_0 \in \gamma_k^{-1} \mathcal{F}$, i.e. $z_0 \in \mathcal{F}^\circ$ is Γ-equivalent to $\gamma_k \gamma' \gamma_i^{-1} \cdot z_0 \in \mathcal{F}$. Since $\mathcal{F}$ is a fundamental domain for Γ, we conclude that $z_0 = \gamma_k \gamma' \gamma_i^{-1} \cdot z_0 \in \mathcal{F}^\circ$.

In general, the map $w \mapsto \gamma_k \gamma' \gamma_i^{-1} \cdot w$ is an automorphism of $\mathbb{H}$. Thus, there is an open set $W \subset \mathcal{F}^\circ$ containing z_0 such that $\gamma_k \gamma' \gamma_i^{-1} \cdot W \subset \mathcal{F}^\circ$. But then every element of W is Γ-equivalent to some element of $\mathcal{F}^\circ$. This can only happen if $\gamma_k \gamma' \gamma_i^{-1}$ acts trivially on W.

Since any two holomorphic functions agreeing on an open set must agree everywhere[12], we deduce that $\gamma_k \gamma' \gamma_i^{-1} = \pm I$ (by Corollary 5). Now

$$\gamma_i \cdot \Gamma' = \gamma_k \gamma' (\pm I) \cdot \Gamma' = \gamma_k \cdot \Gamma'$$

As the γ_i's are distinct coset representatives[13], we conclude that $\gamma_i = \gamma_k$ and so $\gamma' = \pm I$. In particular, $z = \gamma' \cdot z$. Hence z and $\gamma \cdot z$ are not distinct, contradicting our assumption and completing the proof.

□

Remark 37. There are many possible choices of γ_i in the coset decomposition of Γ by Γ' above.

Example 1 (Fundamental domain of $\Gamma(2)$)**.** By Remark 14 we know that $\Gamma/\Gamma(2) \cong SL_2(\mathbb{Z}/2\mathbb{Z})$. Since $SL_2(\mathbb{Z}/2\mathbb{Z})$ a group of order 6, we have $[\Gamma : \Gamma(2)] = 6$. Moreover, $SL_2(\mathbb{Z}/2\mathbb{Z})$ is a non-abelian group[14], thus $SL_2(\mathbb{Z}/2\mathbb{Z}) \cong S_3$. Hence, $\Gamma/\Gamma(2) \cong S_3$ and among the 6 cosets of $\Gamma(2)$, apart from identity, two are of order 3 and three are of order 2. So based on Remark 26 we choose

$$\gamma_1 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \qquad \gamma_2 = t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}; \qquad \gamma_3 = s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$$\gamma_4 = st = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}; \qquad \gamma_5 = ts = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}; \qquad \gamma_6 = t^{-1}st = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$$

where $\gamma_2^2, \gamma_3^2, \gamma_6^2 \in \Gamma(2)$ and $\gamma_4^3, \Gamma_5^3 \in \Gamma(2)$. Now by the above theorem $\mathcal{F}(2) = \bigcup_{i=1}^{6} \gamma_i^{-1} \cdot \mathcal{F}$ is a fundamental domain of $\Gamma(2)$.

Hence to get the complete picture of the $\mathcal{F}(2)$ we must glue together each of $\gamma_i \cdot \mathcal{F}$. Which we have as follows[15]

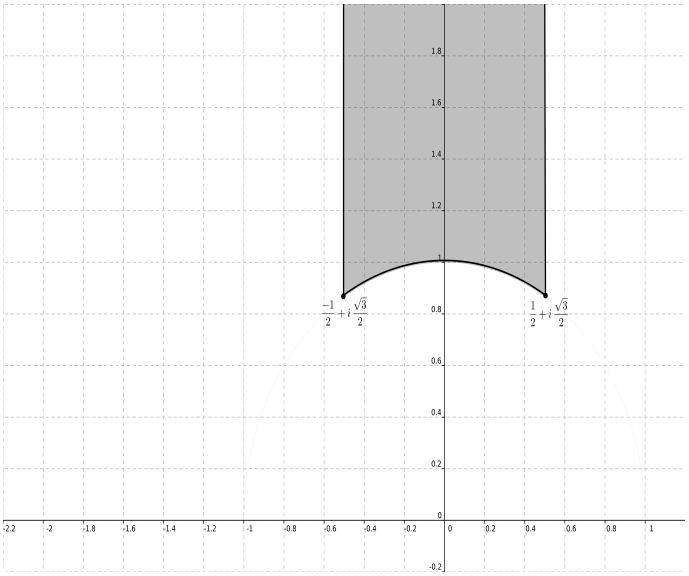
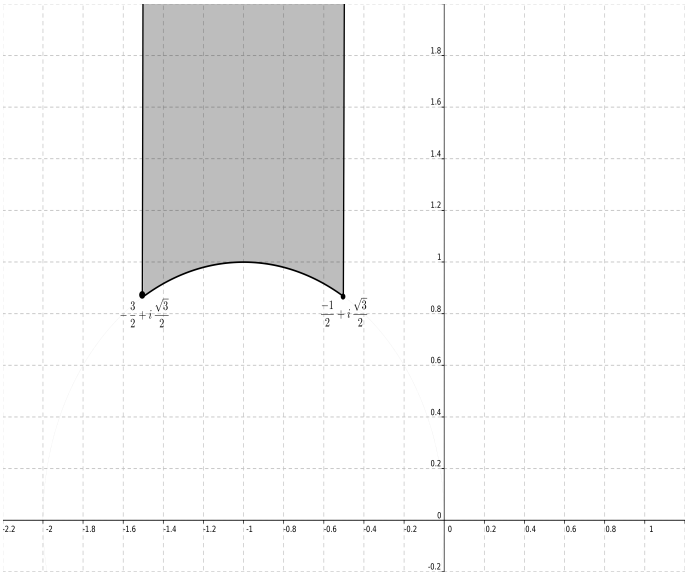
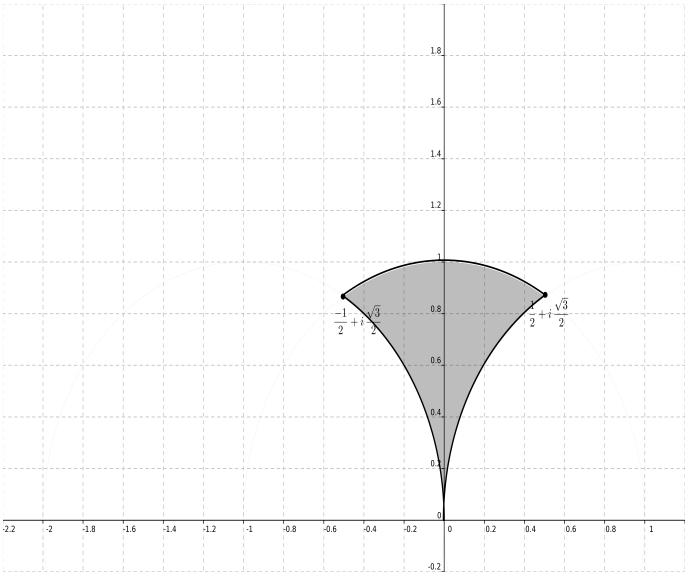
[11] Hence it's a homemorphism and in particular an open map, see [2, Theorem 3.2.6] and [7, Theorem 1.1.1].

[12] This is known as *identity theorem* in complex analysis, see [5, Theorem IV.3.7].

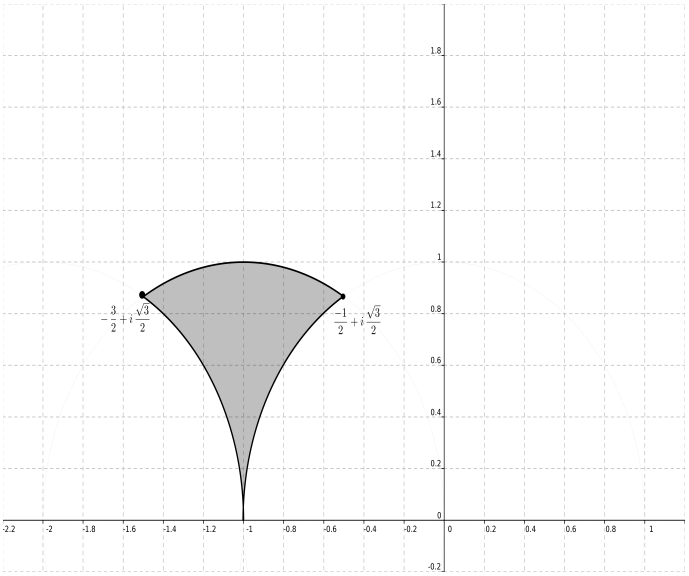
[13] Given a group G and a subgroup $H \subset G$, any two left cosets aH and bH either coincide or are disjoint. This fact is used to prove the Lagrange's theorem in group theory.

[14] Any non-abelian group of order six is isomorphic to the symmetric group of degree 3, see [16, Theorem 3.2].

[15] Recall that if $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $\gamma^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ since $ad - bc = 1$.

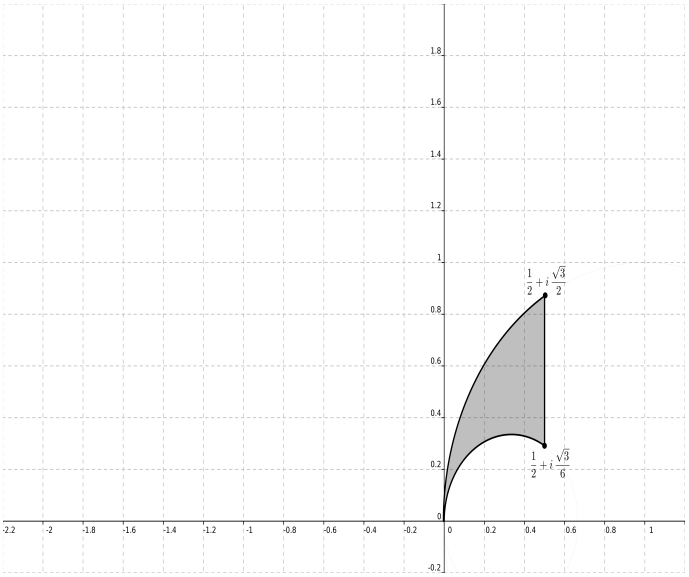
Part of fundamental domain	Drawing of the corresponding part
$\gamma_1^{-1} \cdot \mathcal{F} : z \mapsto z$ $\left\{z \in \mathbb{H} : -\frac{1}{2} \leq \mathrm{Re}(z) \leq \frac{1}{2} \text{ and } \lvert z \rvert \geq 1\right\}$	
$\gamma_2^{-1} \cdot \mathcal{F} : z \mapsto z - 1$ $\left\{z \in \mathbb{H} : -\frac{3}{2} \leq \mathrm{Re}(z) \leq -\frac{1}{2} \text{ and } \lvert z+1 \rvert \geq 1\right\}$	
$\gamma_3^{-1} \cdot \mathcal{F} : z \mapsto -\frac{1}{z}$ $\{z \in \mathbb{H} : \lvert z \rvert \leq 1, \lvert z-1 \rvert \geq 1 \text{ and } \lvert z+1 \rvert \geq 1\}$	

$\gamma_4^{-1} \cdot \mathcal{F} : z \mapsto -\frac{1}{z} - 1$



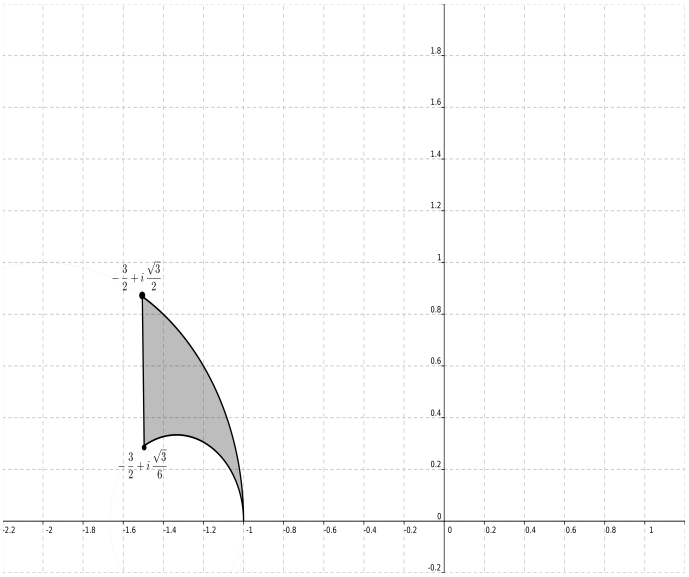
$\{z \in \mathbb{H} : |z+1| \leq 1, |z| \geq 1 \text{ and } |z+2| \geq 1\}$

$\gamma_5^{-1} \cdot \mathcal{F} : z \mapsto -\frac{1}{z-1}$



$\left\{z \in \mathbb{H} : \mathrm{Re}(z) \leq \frac{1}{2}, |z-1| \leq 1 \text{ and } |z-\frac{1}{3}| \geq \frac{1}{3}\right\}$

$\gamma_6^{-1} \cdot \mathcal{F} : z \mapsto -\frac{1}{z+1} - 1$



$\left\{z \in \mathbb{H} : \mathrm{Re}(z) \geq -\frac{3}{2}, |z+2| \leq 1 \text{ and } |z+\frac{4}{3}| \geq \frac{1}{3}\right\}$

To draw these regions we keep in mind following two properties of fractional linear transformations:

- it takes a circle or line to a circle or line and preserves symmetry about the real axis[16]
- it preserves angles between lines[17]

Since $\gamma_1^{-1} = I$, we have $\gamma_1^{-1} \cdot \mathcal{F} = \mathcal{F}$. Now, $\gamma_3^{-1} \cdot \mathcal{F}$ can be easily obtained from $\mathcal{F}$ since it's just a simple application of negative reciprocal map which has been discussed in detail in the proof of Theorem 5, i.e. gives a region where the boundaries are pieces of three different semicircles (see [8, pp. 153]). For $\gamma_5^{-1} \cdot \mathcal{F}$ note that $\gamma_5^{-1} \cdot \infty = 0$, $\gamma_5 \cdot e^{\frac{\pi i}{3}} = e^{\frac{\pi i}{3}}$ and $\gamma_5 \cdot e^{\frac{2\pi i}{3}} = \frac{1}{2} + i\frac{\sqrt{3}}{6}$ and join them as part of semicircles centred at rational points or straight lines perpendicular to real axis (see [7, Theorem 3.5.1]). Then, $\gamma_2^{-1} \cdot \mathcal{F}$ and $\gamma_4^{-1} \cdot \mathcal{F}$ are obtained by translating, one unit to the left, $\gamma_1^{-1} \cdot \mathcal{F}$ and $\gamma_3^{-1} \cdot \mathcal{F}$ respectively. Also, we obtain $\gamma_6^{-1} \cdot \mathcal{F}$ from $\gamma_5^{-1} \cdot \mathcal{F}$ by first reflecting about the imaginary axis and then translating by one unit to the left.

Hence, it follows that the boundary of any fundamental domain constructed in this way consists of vertical lines and arcs of circles centred at rational numbers on the real axis.

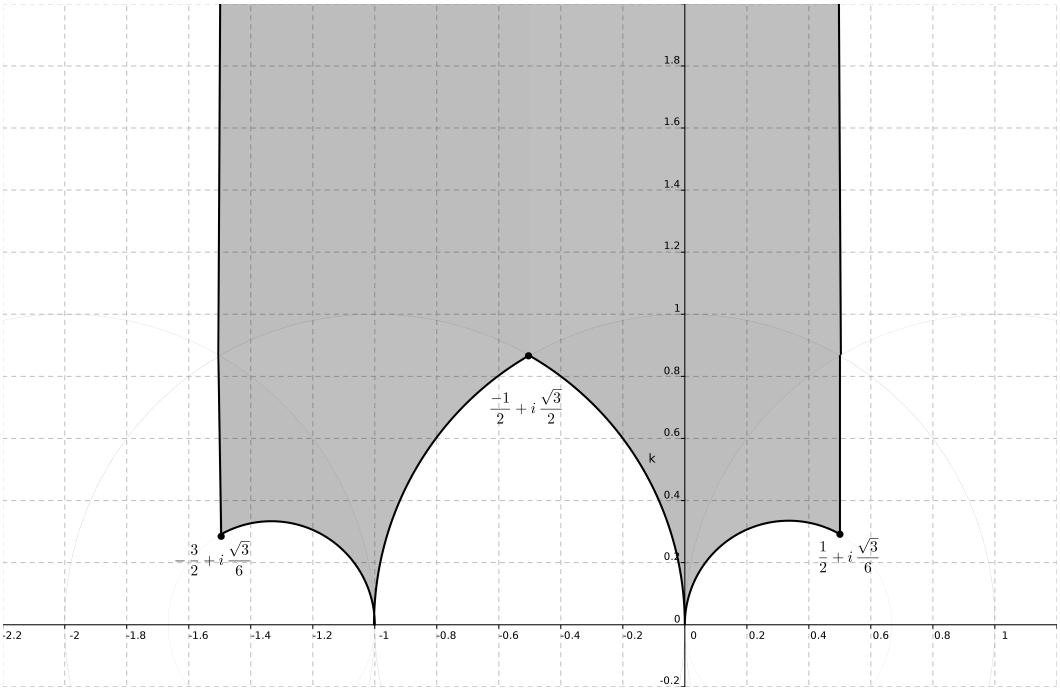


Figure 2.2: A fundamental domain $\mathcal{F}(2)$ of $\Gamma(2)$. Drawn using a: $((-0.5 \leq x) \wedge (x \leq 0.5) \wedge (x^2 + y^2 \geq 1) \wedge (y \geq 0))$; b: $((-1.5 \leq x) \wedge (x \leq -0.5) \wedge ((x+1)^2 + y^2 \geq 1) \wedge (y \geq 0))$; c: $((x^2 + y^2 \leq 1) \wedge (y \geq 0) \wedge ((x+1)^2 + y^2 \geq 1) \wedge ((x-1)^2 + y^2 \geq 1))$; d: $(((x+1)^2 + y^2 \leq 1) \wedge (y \geq 0) \wedge (x^2 + y^2 \geq 1) \wedge ((x+2)^2 + y^2 \geq 1))$; e: $(((x - 1/3)^2 + y^2 \geq 1/9) \wedge (y \geq 0) \wedge ((x-1)^2 + y^2 \leq 1) \wedge (x \leq 1/2))$; f: $(((x + 4/3)^2 + y^2 \geq 1/9) \wedge (y > 0) \wedge ((x+2)^2 + y^2 \leq 1) \wedge (x \geq (-3)/2))$ in GeoGebra 4.0.34.0.

Hence, the boundary of $\mathcal{F}(2)$ is made up from the vertical lines $\mathrm{Re}(z) = -\frac{3}{2}$ and $\mathrm{Re}(z) = \frac{1}{2}$, the circles of radius 1 centred at 0 and -1, and the circles of radius $\frac{1}{3}$ centred at $\frac{1}{3}$ and $-\frac{4}{3}$.

Moreover, geometrically we see that $\Gamma(2)$ has three "cusps": ∞, 0 and -1. We can easily verify this using Remark 30, since for

- $\gamma_1^{-1} \cdot \mathcal{F}$ and $\gamma_2^{-1} \cdot \mathcal{F}$, $a/c = \infty$;
- $\gamma_3^{-1} \cdot \mathcal{F}$ and $\gamma_5^{-1} \cdot \mathcal{F}$, $a/c = 0$; and
- $\gamma_4^{-1} \cdot \mathcal{F}$ and $\gamma_6^{-1} \cdot \mathcal{F}$, $a/c = -1$

Remark 38. Examples of another fundamental domain of $\Gamma(2)$ can be found in [7, Example F, pp. 141]. More examples of fundamental domain drawings can be found here: $\Gamma_0(2)$ at [17, Example 3.4.11] and $\Gamma_0(3)$ at [8, pp. 184].

[16] In fact, the geodesics in the hyperbolic plane, i.e. Poincaré upper half plane, are semi-circles and straight lines orthogonal to real axis, see [7, Theorem 1.2.4].

[17] Any fractional linear transformation is conformal mapping, see [7, Theorem 1.3.2]

Chapter 3

Modular forms for the full modular group

"It's a joy to meet divisor functions as the coefficients of Fourier series, thus power series in q, representing modular forms. Thus identities such as $G_4^2 = \frac{7}{3}G_8$, which I originally mentioned for their amusement value, now become quite profound relationships connecting the divisor functions $\sigma_3(n)$ and $\sigma_7(n)$."

— Alf van der Poorten, *Notes on Fermat's Last Theorem*

3.1 Introduction

Definition 30 (Modular function of weight k for Γ)**.** Let $f(z)$ be a meromorphic function on the upper half-plane $\mathbb{H}$, and let k be an integer. If $f(z)$ satisfies the relation

$$f(\gamma \cdot z) = (cz + d)^k f(z) \qquad \text{for all} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

and is meromorphic at infinity (Definition 27) then $f(z)$ is called a *modular function of weight k for $\Gamma = SL_2(\mathbb{Z})$*.

Definition 31 (q-expansion)**.** Let $f(z)$ be a modular function then the Fourier series expansion

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

which has at most finitely many non-zero a_n with $n < 0$ is called its *q-expansion*.

Definition 32 (Modular form of weight k for Γ)**.** Let $f(z)$ be a holomorphic function on the upper half-plane $\mathbb{H}$, and let k be an integer. If $f(z)$ satisfies the relation

$$f(\gamma \cdot z) = (cz + d)^k f(z) \qquad \text{for all} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

and is holomorphic at infinity (Definition 27) then $f(z)$ is called a *modular form of weight k for $\Gamma = SL_2(\mathbb{Z})$*. The set of such functions is denoted by $M_k(\Gamma)$.

Definition 33 (Cusp-form of weight k for Γ)**.** Let $f(z)$ be a holomorphic function on the upper half-plane $\mathbb{H}$, and let k be an integer. If $f(z)$ satisfies the relation

$$f(\gamma \cdot z) = (cz + d)^k f(z) \qquad \text{for all} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

and vanishes at infinity (Definition 27) then $f(z)$ is called a *modular function of weight k for $\Gamma = SL_2(\mathbb{Z})$*. The set of such functions is denoted by $S_k(\Gamma)$.

Remark 39. Cusp-forms are sometimes also called *parabolic forms.*

Proposition 11. *If k is odd, there is no non-zero modular form of weight k for Γ.*

Proof. Let $f(z)$ be a modular form and $\gamma = -I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ then by the definition of modular form

$$f(\gamma \cdot z) = (-1)^k f(z) = -f(z)$$

but since $\gamma \cdot z = z$, we have $f(z) = -f(z)$ for all $z \in \overline{\mathbb{H}}$, i.e. $f(z) \equiv 0$. □

Proposition 12. *Let $f(z)$ be a meromorphic function on $\mathbb{H}$, and k be an even integer. Then $f(z)$ satisfies the relation*

$$f(\gamma \cdot z) = (cz + d)^k f(z) \qquad \text{for all} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \tag{3.1}$$

if and only if $f(z)$ satisfies the relations $f(z+1) = f(z)$ and $f(-1/z) = z^k f(z)$.

Proof. ($\Rightarrow$) Since the relation (3.1) is true for all $\gamma \in \Gamma$, in particular it's true for $\gamma = t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\gamma = s = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Which gives us the relations $f(z+1) = f(z)$ and $f(-1/z) = z^k f(z)$.

($\Leftarrow$) For $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we have $ad - bc = 1$ and hence

$$\frac{d}{dz}(\gamma \cdot z) = \frac{d}{dz}\left(\frac{az+b}{cz+d}\right) = \frac{1}{(cz+d)^2}$$

Hence we can rewrite (3.1) as

$$\left(\frac{d}{dz}(\gamma \cdot z)\right)^{k/2} f(\gamma \cdot z) = f(z)$$

that is, $f(z)(dz)^{k/2}$ is invariant when z is replaced by $\gamma \cdot z$. From this we see that if (3.1) holds for γ_1 and γ_2 then it holds for $\gamma_1\gamma_2$. Since all of $\overline{\Gamma}$ is generated by s and t (Proposition 9), this means that $f(z+1) = f(z)$ and $f(-1/z) = z^k f(z)$ imply (3.1). □

Remark 40. If $k = 0$ then the condition (3.1) says $f(z)$ is invariant under Γ, i.e it may be considered as a function on $\Gamma\backslash\mathbb{H}$ (Definition 18). If $k = 2$, then the differential form $f(z)dx$ on $\mathbb{H}$ is invariant under Γ, i.e. $f(\gamma \cdot z)d(\gamma \cdot z) = f(z)dz$, since $\frac{d}{dz}(\gamma \cdot z) = \frac{1}{(cz+d)^2}$.

3.2 Eisenstein series

Definition 34 (Eisenstein series)**.** Let k be an even integer greater than 2. For $z \in \mathbb{H}$ we define the Eisenstein series $G_k(z)$ as

$$G_k(z) = \sum_{m,n}{}' \frac{1}{(mz+n)^k}$$

where the summation is over pairs of integers m, n not both zero.

Remark 41. If we let Λ_z denote the lattice in $\mathbb{C}$ spanned by 1 and z, then we saw this in the proof of Theorem 4 (where $\mathrm{Im}(z) > 0$ by Remark 2). That is, the function $G_k(z)$ is what we denoted $G_k(\Lambda_z)$. The new point of view in this chapter is to think of $G_k(z) = G_k(\Lambda_z)$ as functions of z, not merely as coefficients in the Laurent expansion of the Weierstrass $\wp$-function. (The letter z was used in a different way in Theorem 4).

Lemma 3. *The two series*

$$\sum_{m,n}{}' \frac{1}{(|n|+|m|)^r} \qquad \text{and} \qquad \sum_{mz+n\in\Lambda_z^*} \frac{1}{|mz+n|^r}$$

where $\Lambda_z = \{mz+n : n,m \in \mathbb{Z}\}$ *and* $\Lambda_z^* = L - \{0\}$, *converge if* $r > 2$.

Proof. The question whether a double series converges absolutely is independent of the order of summation[1]. Here we shall first sum in m and then in n.

For the first series, the usual integral comparison can be applied[2]. For each $n \neq 0$ we have the partial sum

$$\begin{aligned}
\sum_{|m|\le N} \frac{1}{(|n|+|m|)^r} &= \frac{1}{|n|^r} + 2\sum_{1\le m\le N} \frac{1}{(|m|+|n|)^r} \\
&= \frac{1}{|n|^r} + 2\sum_{|n|+1\le k\le N} \frac{1}{k^r} \\
&\le \frac{1}{|n|^r} + 2\int_{|n|}^{\infty} \frac{dx}{x^r} \\
&\le \frac{1}{|n|^r} + \frac{C}{|n|^{r-1}}
\end{aligned}$$

Now taking $N \to \infty$ and $r > 2$ we get

$$\begin{aligned}
\sum_{m,n}{}' \frac{1}{(|n|+|m|)^r} &= \sum_{|m|\neq 0} \frac{1}{|m|^r} + \sum_{|n|\neq 0}\sum_{m\in\mathbb{Z}} \frac{1}{(|n|+|m|)^r} \\
&\le \sum_{|m|\neq 0} \frac{1}{|m|^r} + \sum_{|n|\neq 0} \left(\frac{1}{|n|^r} + \frac{C}{|n|^{r-1}}\right) \\
&< \infty
\end{aligned}$$

To prove that the second series also converges, it suffices to show that there is a constant C' for a given z, such that

$$|n| + |m| \le C'|n+mz| \qquad \text{for all} \quad n,m \in \mathbb{Z}$$

Let $z = x + iy$ with $y > 0$ then for $|n| \le 2|m||x|$

$$\begin{aligned}
|n| + |m| &\le |n+mx| + |m| \\
&\le |n+mx| + |my| \\
&\le 2\sqrt{(n+mx)^2 + (my)^2} \\
&= 2|n+mz|
\end{aligned}$$

where we used the fact that for any two positive numbers a and b, $a \le \sqrt{a^2+b^2}$ and $b \le \sqrt{a^2+b^2}$. □

Remark 42. For more details, see [4, Lemma 9.1.5].

Theorem 7. *Eisenstein series is a modular form of weight* k, *i.e.* $G_k \in M_k(\Gamma)$.

Proof. We will prove it in three steps

[1] For more details, see [4, pp. 197].

[2] Which says that $\sum_{n=1}^{\infty} \frac{1}{n^t} < 1 + \sum_{n=2}^{\infty} \int_{n-1}^{n} \frac{dx}{x^t} = 1 + \int_1^{\infty} \frac{dx}{x^t}$, it was also used in [18, §1.3.1].

Claim 1. *$G_k(z)$ is holomorphic on the upper half-plane*

By the second series of Lemma 3, we know that for $k > 2$ the series $G_k(z)$ converges absolutely and uniformly in every half-plane $\mathrm{Im}(z) \geq \delta > 0$, i.e. in any compact subset of $\mathbb{H}$, whenever $k \geq 4$. Hence by [5, Corollary VII.2.4], the function $G_k(z)$ is holomorphic on $\mathbb{H}$.

Claim 2. *$G_k(z)$ is holomorphic at infinity*

Observe that $G_k(z)$ approaches a finite limit as $z \to i\infty$:

$$\lim_{z \to i\infty} \sum_{m,n}{}' \frac{1}{(mz+n)^k} = \sum_{n \neq 0} \frac{1}{n^k} = 2\zeta(k)$$

where $\zeta(k)$ is the Riemann zeta-function, see [18, §1.3.4]. Hence the Fourier expansion of $G_k(z)$ (as discussed in Definition 27) has no negative terms.

Claim 3. *$G_k(z)$ satisfies the transformation relations:*

$$G_k(z+1) = G_k(z) \qquad \text{and} \qquad z^{-k} G_k(-1/z) = G_k(z)$$

$G_k(z)$ is periodic of period 1 because $n + m(z+1) = (n+m) + mz$, and we can rearrange the sum by replacing $n+m$ by n. Also, we have

$$\left(n + \left(\frac{-1}{z}\right)\right)^k = z^{-k}(nz - m)^k$$

and again we can rearrange the sum, this time replacing $(-m, n)$ by (n, m). The claim follows.

Combining the above three claims with Proposition 12, we complete the proof. □

Theorem 8. *Let k be an even integer greater than 2, and let $z \in \mathbb{H}$. Then the modular form $G_k(z)$ has q-expansion*

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

where $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$ and the Bernoulli numbers B_k (for even integers k) are defined by setting

$$\frac{x}{e^x - 1} = -\frac{x}{2} + \sum_{k \geq 0} B_k \frac{x^k}{k!}$$

since $B_{2r+1} = 0$ for ≥ 1 and $B_1 = \frac{-1}{2}$.

Proof. We will first prove the following two claims:

Claim 1. $\zeta(k) = -\dfrac{(2\pi i)^k}{2} \dfrac{B_k}{k!}$ *for any even integer $k \geq 2$*

From the product formula for the sine function [4, §5.3.2] we know that[3]

$$\pi \cot(\pi z) = \sum_{n=-\infty}^{\infty} \frac{1}{z+n} = \lim_{N \to \infty} \sum_{|n| \leq N} \frac{1}{z+n} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2}$$

[3] Otherwise we can do logarithmic differentiation of the product formula for sine obtained by Hadamard factorization theorem, see the Problem 2(i) in J. Murphy's "Homework 6 - Solution sketches." for Spring 2015 - Math 185 - Complex Analysis course at University of California, Berkeley. https://math.berkeley.edu/~murphy/185-Solutions6.pdf. *(accessed on 20 December 2017).*

Hence we have

$$z\cot(z) = 1 + 2\sum_{n=1}^{\infty} \frac{z^2}{z^2 - \pi^2 n^2}$$

Now observing that we have geometric series

$$\sum_{\ell=1}^{\infty} \left(\frac{z}{n\pi}\right)^{2\ell} = \frac{z^2}{\pi^2 n^2 - z^2}$$

we get

$$z\cot(z) = 1 - 2\sum_{n=1}^{\infty}\sum_{\ell=1}^{\infty} \left(\frac{z}{n\pi}\right)^{2\ell} = 1 - 2\sum_{k\geq 2} \zeta(k)\frac{z^k}{\pi^k} \tag{3.2}$$

where k takes even integer values.

On the other hand, in the definition of Bernoulli numbers, if we put $x = 2iz$, we obtain:

$$\frac{2iz}{e^{2iz} - 1} = -iz + \sum_{k\geq 0} B_k \frac{(2iz)^k}{k!}$$

Now observing that

$$\cot(z) = \frac{\cos(z)}{\sin(z)} = i\frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}} = i + \frac{2i}{e^{2iz} - 1}$$

and $B_0 = 1$ we get [9]:

$$z\cot(z) = 1 + 2\sum_{k\geq 2} \frac{(2\pi i)^k}{2} \frac{B_k}{k!} \frac{z^k}{\pi^k} \tag{3.3}$$

Comparing (3.2) and (3.3), we complete the proof of our claim.

Claim 2. $\displaystyle\sum_{n=-\infty}^{\infty} \frac{1}{(mz+n)^k} = \frac{(2\pi i)^k}{(k-1)!} \sum_{d=1}^{\infty} d^{k-1} q^{dm}$

As noted in previous case, we have

$$\pi\cot(\pi z) = \sum_{n=-\infty}^{\infty} \frac{1}{z+n}$$

Also we can re-write left hand side as:

$$\pi\cot(\pi z) = \pi\frac{\cos(\pi z)}{\sin(\pi z)} = \pi i - \frac{2\pi i}{1-q} = \pi i - 2\pi i \sum_{d=0}^{\infty} q^d$$

where $q = e^{2\pi i z}$. Hence we have

$$\pi i - 2\pi i \sum_{d=0}^{\infty} q^d = \sum_{n=-\infty}^{\infty} \frac{1}{z+n} \tag{3.4}$$

Now differentiating it with respect to z we get

$$-2\pi i \sum_{d=0}^{\infty} (2\pi i d) q^d = \sum_{n=-\infty}^{\infty} \frac{-1}{(z+n)^2} \qquad \Rightarrow (2\pi i)^2 \sum_{d=1}^{\infty} d q^d = \sum_{n=-\infty}^{\infty} \frac{1}{(z+n)^2}$$

Hence by successive differentiation of (3.4) $k-1$ times (where k is even) we get

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

Then replacing z by mz we get

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

hence completing the proof of our claim.

Now we will use the above two claims to complete the proof.

$$\begin{aligned}
G_k(z) = \sum_{m,n}{}' \frac{1}{(mz+n)^k} &= \sum_{n\neq 0}\frac{1}{n^k} + \sum_{m\neq 0}\sum_{n\in\mathbb{Z}}\frac{1}{(mz+n)^k} \\
&= 2\zeta(k) + 2\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty}\frac{1}{(mz+n)^k} \\
&= 2\zeta(k) + 2\sum_{m=1}^{\infty}\frac{(2\pi i)^k}{(k-1)!}\sum_{d=1}^{\infty} d^{k-1}q^{dm} \\
&= 2\zeta(k) - \frac{4k\zeta(k)}{B_k}\sum_{m=1}^{\infty}\sum_{d=1}^{\infty} d^{k-1}q^{dm}
\end{aligned}$$

If we put $dm = n$ then we get

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

□

Definition 35 (Normalized Eisenstein series)**.** The series obtained by dividing $G_k(z)$ by the constant $2\zeta(k)$ is called *normalized Eisenstein series*, $E_k(z)$, i.e.

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

3.3 Discriminant modular form

Definition 36 (Discriminant of a polynomial)**.** The discriminant of a polynomial $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n \in \mathbb{C}[x]$ such that $\alpha_1, \alpha_2, \ldots, \alpha_n$ are its roots, i.e.

$$f(x) = a_0(x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_n)$$

is given by $\text{disc}(f) = a_0^{n-1}\prod_{i<j}(\alpha_i-\alpha_j)^2$.

Remark 43. The discriminant of a polynomial is non-zero if and only if the roots are distinct.

Lemma 4. *The discriminant of* $f(x) = 4x^3 - ax - b \in \mathbb{C}[x]$ *is* $a^3 - 27b^2$.

Proof. Let α, β, γ be the roots of $f(x)$, then we have

$$f(x) = 4(x-\alpha)(x-\beta)(x-\gamma)$$

so that if we differentiate we have

$$f'(x) = 4(x-\alpha)(x-\beta) + 4(x-\alpha)(x-\gamma) + 4(x-\beta)(x-\gamma)$$

Then we have

$$\begin{cases} f'(\alpha) = 4(\alpha-\beta)(\alpha-\gamma) \\ f'(\beta) = 4(\beta-\alpha)(\beta-\gamma) \\ f'(\gamma) = 4(\gamma-\alpha)(\gamma-\beta) \end{cases}$$

Taking the product we observe that

$$\operatorname{disc}(f) = 4^2(\alpha-\beta)^2(\alpha-\gamma)^2(\beta-\gamma)^2 = -\frac{1}{4}f'(\alpha)f'(\beta)f'(\gamma) \tag{3.5}$$

Also, since $f(x) = 4x^3 - ax - b$, on differentiation we get

$$f'(x) = 12x^2 - a$$

and using this in (3.5) we get

$$\begin{aligned} \operatorname{disc}(f) &= -\frac{1}{4}(12\alpha^2 - a)(12\beta^2 - a)(12\gamma^2 - a) \\ &= -\frac{1}{4}\left(12^3\alpha^2\beta^2\gamma^2 - 12^2(\alpha^2\beta^2 + \alpha^2\beta^2 + \beta^2\gamma^2)a + 12(\alpha^2+\beta^2+\gamma^2)a^2 - a^3\right) \\ &= -432\alpha^2\beta^2\gamma^2 + 36(\alpha^2\beta^2 + \alpha^2\beta^2 + \beta^2\gamma^2)a - 3(\alpha^2+\beta^2+\gamma^2)a^2 + \frac{1}{4}a^3 \end{aligned} \tag{3.6}$$

Now we note that

$$\begin{aligned} \alpha^2+\beta^2+\gamma^2 &= (\alpha+\beta+\gamma)^2 - 2(\alpha\beta+\alpha\gamma+\beta\gamma) \\ &= 0 - 2\left(\frac{-a}{4}\right) = \frac{a}{2} \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} \alpha^2\beta^2 + \alpha^2\beta^2 + \beta^2\gamma^2 &= (\alpha\beta+\alpha\beta+\beta\gamma)^2 - 2(\alpha^2\beta\gamma + \beta^2\alpha\gamma + \gamma^2\alpha\beta) \\ &= (\alpha\beta+\alpha\beta+\beta\gamma)^2 - 2\alpha\beta\gamma(\alpha+\beta+\gamma) \\ &= \left(\frac{-a}{4}\right)^2 - 2\frac{b}{4}(0) = \frac{a^2}{16} \end{aligned} \tag{3.8}$$

Hence using $\alpha\beta\gamma = b/4$, along with (3.7) and (3.8) in (3.6), we obtain

$$\begin{aligned} \operatorname{disc}(f) &= \left(-432 \times \frac{b^2}{16}\right) + \left(36 \times \frac{a^2}{16} \times a\right) - \left(3 \times \frac{a}{2} \times a^2\right) + \frac{1}{4}a^3 \\ &= a^3 - 27b^2 \end{aligned}$$

□

Remark 44. The definition of discriminant used here [1, Problem 2 of §I.6] is a non-standard definition, in general discriminant is defined to be $a_n^{2n-2}\prod_{i<j}(\alpha_i - \alpha_j)^2$. If we would have used that definition we would have got $16(a^3 - 27b^2)$, but since $\mathbb{C}$ is a field of characteristic zero, for unequal roots (which we need for elliptic curves) all we need to check is the value of $a^3 - 27b^2$ is not zero. Hence the current definition serves the purpose.

Theorem 9. *The discriminant of the elliptic curve corresponding to Λ_z is given by*

$$\Delta(z) = \frac{(2\pi)^{12}}{1728}\left(E_4(z)^3 - E_6(z)^2\right)$$

Proof. As seen in Proposition 8, the elliptic curve corresponding to Λ_z is the cubic polynomial $4x^3 - g_2(\Lambda_z)x - g_3(\Lambda_z)$. By Lemma 4 we know that

$$\Delta(z) = g_2(z)^3 - 27g_3(z)^2 \tag{3.9}$$

Also from Remark 41 and Theorem 4 we know that

$$g_2(z) = g_2(\Lambda_z) = 60G_4(z) \qquad \text{and} \qquad g_3(z) = g_3(\Lambda_z) = 140G_6(z)$$

Using the first claim of Theorem 8, we know that $\zeta(4) = \pi^4/90$ and $\zeta(6) = \pi^6/945$. Hence we can express g_2 and g_3 in terms of the normalized Eisenstein series E_4 and E_6 as follows

$$g_2(z) = \frac{4\pi^4}{3}E_4(z) \qquad \text{and} \qquad g_3(z) = \frac{8\pi^6}{27}E_6(z)$$

Using this in (3.9) we get

$$\Delta(z) = \frac{2^6\pi^{12}}{27}\left(E_4(z)^3 - E_6(z)^2\right) = \frac{(2\pi)^{12}}{1728}\left(E_4(z)^3 - E_6(z)^2\right)$$

□

Remark 45. From the proof it's clear that $g_2(z)$ and $g_3(z)$ are modular forms for Γ of weight 4 and 6 respectively.

Corollary 6. *$\Delta(z)$ is a cusp-form of weight 12 for Γ.*

Proof. It's clearly a modular form of weight 12 for Γ. Moreover, because both $E_4(z)$ and $E_6(z)$ have constant term $a_0 = 1$ in their 1-expansions, we see that the constant term for $\Delta(z)$ is zero. Hence $\Delta(z)$ is a cusp-form. □

Definition 37. The cusp-form

$$\Delta(z) = \frac{(2\pi)^{12}}{1728}\left(E_4(z)^3 - E_6(z)^2\right)$$

is called the *discriminant modular form.*

Remark 46. The Fourier coefficients of $(2\pi)^{-12}\Delta(z)$ define the Ramanujan τ-function:

$$(2\pi)^{-12}\Delta(z) = \sum_{n=1}^{\infty}\tau(n)q^n = q - 24q^2 + 252q^3 + \dots$$

The values of $\tau(n)$ are integers, see [2, Exercise 4.4.5].

3.4 Space of modular and cusp forms

Proposition 13. *The sets of modular functions, modular forms and cusp-forms of some fixed weight are complex vector spaces.*

Proof. Firstly we need to verify the closure under vector addition and scalar multiplication. Let V be the sets of modular functions, modular forms and cusp-forms of some fixed weight k.

Let $f(z), g(z) \in V$ and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then

$$f(\gamma \cdot z) = (cz+d)^k f(z) \qquad \text{and} \qquad g(\gamma \cdot z) = (cz+d)^k g(z)$$

Now if $h(z) = f(z) + g(z)$ then

$$h(\gamma \cdot z) = (cz+d)^k (f(z) + g(z)) = (cz+d)^k h(z)$$

and since analytic properties are preserved under addition we have verified additive closure. Similarly, we can verify that the elements of V form an abelian group under addition.

Let $z_0 \in \mathbb{C}$ and $f(z) \in V$, then

$$z_0 f(\gamma \cdot z) = z_0 (cz+d)^k f(z) = (cz+d)^k z_0 f(z)$$

and hence $z_0 f(z) \in V$ since analytic properties are preserved under scalar multiplication. Similarly, one can check that scalar multiplication gives this group a $\mathbb{C}$-module structure. □

Corollary 7. *The set of modular functions is a field if and only if the weight is* 0.

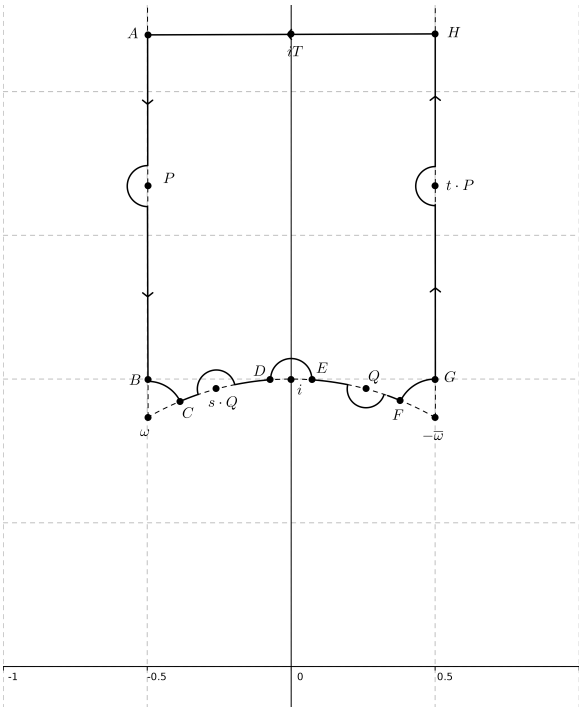
Proof. Note that the product of a modular function of weight k_1 and a modular functions of weight k_2 is a modular function of weight $k_1 + k_2$, and the quotient of a modular function of weight k_1 by a non-zero modular function of weight k_2 is a modular function of weight $k_1 - k_2$. In particular, the set of modular functions of weight zero is a field. □

Proposition 14 (Valence formula)**.** *Let $f(z)$ be a non-zero modular function of weight k for Γ. For $P \in \mathbb{H}$, let $v_P(f)$ denote the order of zero (or minus the order of pole) of $f(z)$ at the point P. Let $v_\infty(f)$ denote the index of the first non-vanishing term in the q-expansion of $f(z)$. Then*

$$v_\infty(f) + \frac{1}{2} v_i(f) + \frac{1}{3} v_\omega(f) + \sum_{\substack{P \in \Gamma \backslash \mathbb{H} \\ P \neq i, \omega}} v_P(f) = \frac{k}{12}$$

where $i = \sqrt{-1}$ and $\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

Proof. Let $\mathcal{C}$ be the following contour:



The top of $\mathcal{C}$ is a horizontal line from $H = \frac{1}{2} + iT$ to $A + -\frac{1}{2} + iT$, there T is take larger than the imaginary part of any of the zeros or poles of $f(z)$. Such a T exists, i.e $f(z)$ does not have poles or zeros with arbitrarily large imaginary part. This follows from the fact that the change of variables $q = e^{2\pi iz}$ makes $f(z)$ into a meromorphic function of q in a disc around $q = 0$. The rest of the contour follows around the boundary along circular arcs of small radius ε, where $\varepsilon \to 0$. This is done is such a way so as to include every Γ-equivalence class of zero or pole exactly once[4] inside $\mathcal{C}$, except that i and ω (and $s \cdot \omega = -\overline{\omega}$) are kept outside of $\mathcal{C}$ if they are zeros or poles. Here we have illustrated the case when zero and poles on the boundary of $\mathcal{F}$ consist of i, ω (and hence $s \cdot \omega$), one point P on the vertical boundary line (and hence also its Γ-equivalent point $t \cdot P$ on the opposite line), and one point Q on the unit circle part of the boundary (and hence also $s \cdot Q$).

According to the argument principle [5, Theorem V.3.4], we have:

$$\frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f'(z)}{f(z)} dz = \sum_{\substack{P \in \Gamma \backslash \mathbb{H} \\ P \neq i, \omega}} v_P(f) \tag{3.10}$$

On the other hand, we evaluate the integral (3.10) section by section, as follows:

Case 1. *$A \to B$ and $G \to H$*

The integral from A to B cancels the integral from G to H, because $f(z+1) = f(z)$ (Proposition 12) and the lines go in opposite directions. Hence we have:

$$\frac{1}{2\pi i} \int_{AB} \frac{f'(z)}{f(z)} dz + \frac{1}{2\pi i} \int_{GH} \frac{f'(z)}{f(z)} dz = 0 \tag{3.11}$$

Case 2. *$H \to A$*

We make the change of variables $q = e^{2\pi iz}$. Let $g(q) = f(z) = \sum a_n q^n$ be the q-expansion. Then using $f'(z) = g'(q)\frac{dq}{dz}$, we find that

$$\begin{aligned} \frac{1}{2\pi i} \int_{HA} \frac{f'(z)}{f(z)} dz &= \frac{1}{2\pi i} \int_{\frac{1}{2}+iT}^{-\frac{1}{2}+iT} \frac{f'(z)}{f(z)} dz \\ &= \frac{1}{2\pi i} \int_{e^{-2\pi T} e^{\pi i}}^{e^{-2\pi T} e^{-\pi i}} \frac{g'(q)}{g(q)} dq \\ &= \frac{1}{2\pi i} \int_{\widehat{\mathcal{C}}} \frac{g'(q)}{g(q)} dq \\ &= -v_\infty(f) \end{aligned} \tag{3.12}$$

where $\widehat{\mathcal{C}}$ is the circle of radius $e^{-2\pi T}$ centred at origin and traversed in a clockwise direction. It follows that this integral is minus the order of zero or pole of $g(q)$ at 0, and this is what we mean by $-v_\infty(f)$ (see Proposition 10).

Case 3. *$B \to C$, $D \to E$ and $F \to G$*

If $f(z)$ has Laurent expansion $c_m(z-a)^m + \ldots$ near a, with $c_m \neq 0$, then

$$\frac{f'(z)}{f(z)} = \frac{m}{z-a} + h(z)$$

[4] This is because $v_P(f)$ does not change if P is replaced by $\gamma \cdot P$ for $\Gamma \in \Gamma$. Hence it's enough to work with fundamental domain, $\mathcal{F}$.

where $h(z)$ is holomorphic at a. If $\widetilde{\mathcal{C}}$ is a circular arc of angle θ centred at a with small radius ϵ, then

$$\begin{aligned}\frac{1}{2\pi i}\int_{\widetilde{\mathcal{C}}}\frac{f'(z)}{f(z)}dz &= \frac{1}{2\pi i}\int_{\widetilde{\mathcal{C}}}\frac{m}{z-a}dz + \frac{1}{2\pi i}\int_{\widetilde{\mathcal{C}}}h(z)dz \\ &= \frac{1}{2\pi}\int_0^\theta \frac{m}{(a+\varepsilon e^{i\phi})-a}\varepsilon e^{i\phi}d\phi + \frac{1}{2\pi}\int_0^\theta h(a+\varepsilon e^{i\phi})\varepsilon e^{i\phi}d\phi \\ &= \frac{m\theta}{2\pi} + \frac{\varepsilon}{2\pi}\int_0^\theta h(a+\varepsilon e^{i\phi})e^{i\phi}d\phi\end{aligned}$$

where we traverse the curve in anticlockwise direction. Now as $\varepsilon \to 0$, the integral approaches $m\theta/2\pi$. We apply this to section $B \to C$, $D \to E$ and $F \to G$ to get (since $\omega = e^{\frac{2\pi i}{3}}$ and $i = e^{\frac{i\pi}{2}}$)

$$\frac{1}{2\pi i}\int_{BC}\frac{f'(z)}{f(z)}dz = -v_\omega(f)\times\frac{\pi}{3}\times\frac{1}{2\pi} = -\frac{v_\omega(f)}{6} \tag{3.13}$$

$$\frac{1}{2\pi i}\int_{DE}\frac{f'(z)}{f(z)}dz = -v_i(f)\times\pi\times\frac{1}{2\pi} = -\frac{v_i(f)}{2} \tag{3.14}$$

$$\frac{1}{2\pi i}\int_{FG}\frac{f'(z)}{f(z)}dz = -v_{-\overline{\omega}}(f)\times\frac{\pi}{3}\times\frac{1}{2\pi} = -v_\omega(f)\times\frac{1}{6} = -\frac{v_\omega(f)}{6} \tag{3.15}$$

where the negative sign is because all the arc are traversed in clockwise direction.

Case 4. *$C \to D$ and $E \to F$*

The transformation $s \in \Gamma$ takes the curve $C \to D$ to the curve $F \to E$. And by Proposition 12 we have

$$f(s\cdot z) = f\left(-\frac{1}{z}\right) = z^k f(z) \tag{3.16}$$

Differentiating this we obtain

$$f'(s\cdot z)\frac{d(s\cdot z)}{dz} = -\frac{1}{z^2}f'\left(-\frac{1}{z}\right) = z^k f'(z) + kz^{k-1}f(z) \tag{3.17}$$

We now divide (3.17) by (3.16) to get

$$\frac{f'(s\cdot z)}{f(s\cdot z)}d(s\cdot z) = \frac{f'(z)}{f(z)}dz + k\frac{dz}{z}$$

Hence using this we have

$$\begin{aligned}\frac{1}{2\pi i}\int_{CD}\frac{f'(z)}{f(z)}dz + \frac{1}{2\pi i}\int_{EF}\frac{f'(z)}{f(z)}dz &= \frac{1}{2\pi i}\int_{CD}\frac{f'(z)}{f(z)}dz - \frac{1}{2\pi i}\int_{FE}\frac{f'(z)}{f(z)}dz \\ &= \frac{1}{2\pi i}\int_{CD}\frac{f'(z)}{f(z)}dz - \frac{1}{2\pi i}\int_{s\cdot CD}\frac{f'(z)}{f(z)}dz \\ &= \frac{1}{2\pi i}\int_{CD}\frac{f'(z)}{f(z)}dz - \frac{1}{2\pi i}\int_{CD}\frac{f'(s\cdot z)}{f(s\cdot z)}d(s\cdot z) \\ &= -\frac{k}{2\pi i}\int_{CD}\frac{dz}{z} \\ &= -\frac{k}{2\pi i}\int_0^{\frac{\pi}{6}}\frac{-ie^{i\theta}d\theta}{e^{i\theta}} = \frac{k}{12}\end{aligned} \tag{3.18}$$

since $\varepsilon \to 0$ (so we integrate from ω to i along the section of unit circle centred at origin), with negative sign since curve is traversed in clockwise direction.

Now using (3.11), (3.12), (3.13), (3.14), (3.15) and (3.18) in (3.10), we complete the proof. □

Theorem 10. *Let k be any even integer, then*

(a) the only modular forms of weight 0 for Γ are constants, i.e. $M_0(\Gamma) = \mathbb{C}$.

(b) $M_k(\Gamma) = \{0\}$ if k is negative or $k = 2$.

(c) $M_k(\Gamma)$ is one-dimensional vector space, generated by E_k, if $k = 4, 6, 8, 10$ or 14; in other words, $M_k(\Gamma) = \mathbb{C}E_k$ for those values of k.

(d) $S_k(\Gamma) = \{0\}$ if $k < 12$ or $k = 14$; $S_{12}(\Gamma) = \mathbb{C}\Delta$; and for $k > 14$, $S_k(\Gamma) = \Delta M_{k-12}(\Gamma)$ (i.e. the cusp forms of weight k are obtained by multiplying modular forms of weight $k - 12$ by the function $\Delta(z)$)

(e) $M_k(\Gamma) = S_k(\Gamma) \oplus \mathbb{C}E_k$ for $k > 2$

Proof. Note that for a modular form all terms on the left in valence formula (Proposition 14) are non-negative.

(a) Let $f \in M_0(\Gamma)$, and c be any value taken by $f(z)$. Then $f(z) - c \in M_0(\Gamma)$ has a zero, i.e. one of the terms on the left in valence formula is strictly positive. Since the right side is 0, this can only happen if $f(z) - c$ is the zero function.

(b) If $k < 0$ or $k = 2$, there is no way that the sum of non-negative terms on the left of valence formula could equal $k/12$.

(c) When $k = 4, 6, 8, 10$, or 14 we note that there is only one possible way of choosing the $v_P(f)$ so that the valence formula holds:

- for $k = 4$, we must have $v_\omega(f) = 1$, all other $v_P(f) = 0$;
- for $k = 6$, we must have $v_i(f) = 1$, all other $v_P(f) = 0$;
- for $k = 8$, we must have $v_\omega(f) = 2$, all other $v_P(f) = 0$;
- for $k = 10$, we must have $v_\omega(f) = v_i(f) = 1$, all other $v_P(f) = 0$;
- for $k = 14$, we must have $v_\omega(f) = 2$, $v_i(f) = 1$, all other $v_P(f) = 0$.

Let $f_1(z)$, $f_2(z)$ be non-zero elements of $M_k(\Gamma)$. Since there is only one possible way of choosing the $v_P(f)$, $f_1(z)$ and $f_2(z)$ have the same zeros. Hence the weight zero modular function $f_1(z)/f_2(z)$ is actually a modular form. By part (a), f_1 and f_2 are proportional. Choosing $f_2(z) = E_k(z)$ gives the part (c).

(d) For $f \in S_k(\Gamma)$ we have $v_\infty(f) > 0$, and all other terms on the left of the valence formula are non-negative. As seen in part (c), this is not possible for $k < 12$ or $k = 14$, hence $S_k(\Gamma) = \{0\}$ for these values of k.

Let $f \in S_{12}(\Gamma)$, the valence formula implies that the only zero of f is at infinity. Let $f_1(z)$, $f_2(z)$ be non-zero elements of $S_{12}(\Gamma)$. Since there is only one possible way of choosing the $v_P(f)$, $f_1(z)$ and $f_2(z)$ have the same zeros. Hence the weight zero modular function $f_1(z)/f_2(z)$ is actually a modular form. By part (a), f_1 and f_2 are proportional. And we know that Δ is cusp-form of weight 12, hence $f_2 = \Delta$ implies that $S_{12}(\Gamma) = \mathbb{C}\Delta$.

The valence formula implies that the only zero of Δ is at infinity. Hence for any $k > 14$ and any $f \in S_k(\Gamma)$, the modular function f/Δ is actually a modular form, i.e. $f/\Delta \in M_{k-12}(\Gamma)$. This completes the proof of the assertion in part (d).

(e) Since E_k does not vanish at infinity, given $f \in M_k(\Gamma)$ we can always subtract a suitable multiple of E_k so that the resulting $f - cE_k \in M_k(\Gamma)$ vanishes at infinity, i.e. $f - cE_k \in S_k(\Gamma)$.

□

Corollary 8. *We have following two identities*[5]

(a) $\sigma_7(n) = \sigma_3(n) + 120\sum_{m=1}^{n-1}\sigma_3(m)\sigma_3(n-m)$

(b) $11\sigma_9(n) = 21\sigma_5(n) - 10\sigma_3(n) + 5040\sum_{m=1}^{n-1}\sigma_3(n)\sigma_5(n-m)$

Proof. These identities follows by comparing the coefficients in the q-series expansions of both sides of the equations $E_4^2 = E_8$ and $E_4E_6 = E_{10}$, respectively ([2, Exercise 4.3.9]). □

Corollary 9 (Dimension formula)**.** *For $k \geq 0$,*

$$\dim_{\mathbb{C}}(M_k(\Gamma)) = \begin{cases} \lfloor \frac{k}{12} \rfloor & \text{if } k \equiv 2 \pmod{12} \\ \lfloor \frac{k}{12} \rfloor + 1 & \text{if } k \not\equiv 2 \pmod{12} \end{cases}$$

Proof. The formula is true for $k \leq 10$ by part (b) and part (c) of Theorem 10. By part (d) and (e) of Theorem 10, for $k \geq 12$ we have

$$\begin{aligned} \dim_{\mathbb{C}}(M_k(\Gamma)) &= 1 + \dim_{\mathbb{C}}(S_k(\Gamma)) \\ &= 1 + \dim_{\mathbb{C}}(M_{k-12}(\Gamma)) \end{aligned}$$

This proves the result by induction (since we have already verified for all the even remainders of 12). □

Corollary 10. *Any $f \in M_k(\Gamma)$ (k is even integer) can be written in the form*

$$f(z) = \sum_{4i+6j=k} c_{i,j} E_4(z)^i E_6(z)^j$$

Proof. We use induction on k. For $k = 4, 6, 8, 10, 14$ we note that $E_4, E_6, E_4^2, E_4E_6, E_4^2E_6$, respectively is an element of $M_k(\Gamma)$, and so, by part (c) of Theorem 10, E_4 and E_6 must span $M_k(\Gamma)$. Now suppose that $k = 12$ or $k > 14$. Since k is even and $\gcd(4, 6) = 2$, it is possible[6] to find i and j such that $4i + 6j = k$, in which case $E_4^iE_6^j \in M_k(\Gamma)$. Given $f \in M_k(\Gamma)$, by the same argument as in the proof of part (e) of Theorem 10, we can find $c \in \mathbb{C}$ such that $f - cE_4^iE_6^j \in S_k(\Gamma)$. By part (d) of Theorem 10, we can write f in the form

$$f = cE_4^iE_6^j + \Delta f_1 = cE_4^iE_6^j + \frac{(2\pi)^{12}}{1728}(E_4^3 - E_6^2)f_1$$

where $f_1 \in M_{k-12}(\Gamma)$. Now by the induction assumption (with k replaced by $k-12$), we obtain the desired polynomial[7] for f. □

3.5 j-invariant

Definition 38 (j-invariant)**.** It is a modular function of weight zero defined as

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

[5] This justifies the belief that modular forms give the fifth operation of arithmetic [3, 5.5].
[6] This is linear Diophantine equation, see [19, §2.1.1].
[7] We already proved a special case for this in first chapter, see Corollary 3.

Proposition 15. *The function j gives a bijection between $\Gamma\backslash\overline{\mathbb{H}}$ and the Riemann sphere $\mathbb{P}^1_{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.*

Proof. Recall that $\Gamma\backslash\overline{\mathbb{H}}$ is equivalent to the fundamental domain $\mathcal{F}$ with Γ-equivalent boundary points identified, $\Gamma\backslash\mathbb{H}$, and the point at infinity included (Remark 35). And we have

$$j(z) = \frac{1728 g_2(z)^3}{\Delta(z)}$$

In the proof of part (d) of Theorem 10 we saw that $\Delta(z)$ has a simple zero at infinity and no other zero. Since $g_2(z)$ does not vanish at infinity (Remark 45), this means that $j(z)$ has a simple pole at infinity and is holomorphic on $\mathbb{H}$. Thus, j takes ∞ to ∞.

For any $c \in \mathbb{C}$ the modular form $1728 g_2^3 - c\Delta \in M_{12}(\Gamma)$ must vanish at exactly one point $P \in \Gamma\backslash\mathbb{H}$, because when $k = 12$ exactly one of the terms on the left of valence formula is strictly positive. Dividing by Δ, we see that this means that $j(z) - c = 0$ for exactly one value of $z \in \Gamma\backslash\mathbb{H}$. Thus, j on $\Gamma\backslash\mathbb{H}$ is a bijection with $\mathbb{C}$. □

Proposition 16. *The modular functions of weight 0 for Γ are precisely the rational functions of j.*

Proof. Since $j(z)$ is a modular function of weight 0, by Corollary 7 we know that a rational function of $j(z)$ is also a modular function of weight 0.

Conversely, suppose that $f(z)$ is a modular function of weight zero for Γ. If z_j are the poles of $f(z)$ in $\Gamma\backslash\mathbb{H}$, counted with multiplicity, then

$$g(z) = f(z) \prod_j (j(z) - j(z_j))$$

is a modular function of weight 0 with no poles in $\mathbb{H}$. Hence it's sufficient to prove that $g(z)$ is a rational function of j. So, without loss of generality, we may assume that $f(z)$ has no poles in $\mathbb{H}$. Now we can multiply by a suitable power of $\Delta(z)$ to cancel the pole of $f(z)$ at ∞. Thus, for some k we will have $\Delta(z)^k f(z) \in M_{12}(\Gamma)$. By Corollary 10, we can write $f(z)$ as a linear combination of modular functions of the form $E_4^i E_6^j / \Delta^k$ (where $4i + 6j = 12k$), so it suffices to show that such a modular modular function is a rational expression in j. Since $4i + 6j$ is divisible by 12, we must have $i = 3i_0$ and $j = 2j_0$ for some i_0 and j_0, i.e. $i_0 + j_0 = k$. Since

$$j(z) = 1728 \frac{E_4(z)^3}{E_4(z)^3 - E_6(z)^2} \qquad \text{and} \qquad \Delta(z) = \frac{(2\pi)^{12}}{1728} \left(E_4(z)^3 - E_6(z)^2\right)$$

we have

$$\frac{E_4(z)^3}{\Delta(z)} = \frac{1}{(2\pi)^{12}} j(z) \qquad \text{and} \qquad \frac{E_6(z)^2}{\Delta(z)} = \frac{1}{(2\pi)^{12}} j(z) - \frac{1728}{(2\pi)^{12}}$$

Hence, $E_4^{3i_0} E_6^{2j_0} / \Delta^k$ is a product of such factors, i.e. a rational expression in j. This completes the proof. □

Remark 47. We saw in section 2.3 the fundamental domain of $SL_2(\mathbb{Z})$. Recall that in first chapter also we had a fundamental domain, a parallelogram $\Pi \subset \mathbb{C}$ for the lattice Λ (Definition 2). In that case, the group was Λ, the action of $g \in \Lambda$ on a point $z \in \mathbb{C}$ was simply $g \cdot z = g + z$. Every $z \in \mathbb{C}$ is Λ-equivalent to a point in Π, and no two points in the interior of Π are Λ-equivalent. In that situation we found it useful to glue the boundary of Π by identifying Λ-equivalent points. We obtained a torus (Remark 3), and we found that the map $z \mapsto (\wp(z), \wp'(z))$ gives an analytic isomorphism from the torus $\mathbb{C}/\Lambda$ to the elliptic curve $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$ in $\mathbb{P}^2_{\mathbb{C}}$ (Proposition 8). We discussed an analogous situation in Remark 24. In our present situation, j-invariant functions give an analytic isomorphism between $\overline{\mathcal{F}} = \Gamma\backslash\overline{\mathbb{H}}$ (Definition 28) and the projective line $\mathbb{P}^1_{\mathbb{C}}$.

Conclusion

> *"Modular forms are functions on the complex plane that are inordinately symmetric. They satisfy so many internal symmetries that their mere existence seem like accidents. But they do exist."*
>
> — Barry Mazur, *stated in NOVA Season 25 Episode 4 "The Proof" (First Aired: Oct 28, 1997)*

Modular forms occur naturally in connection with problems arising in many other areas of mathematics. Following are the few examples:

- Modularity theorem states that all rational elliptic curves arise from modular forms. Yutaka Taniyama stated a preliminary (slightly incorrect) version of the conjecture in 1950's, and a precise conjecture was formulated by Goro Shimura. In 1967, André Weil published a paper which provided a strong theoretical evidence for the conjecture. The theorem was proved for a large class of elliptic curves by Andrew Wiles in 1995, with a key result obtained by the joint work with Richard Taylor, completing the proof of Fermat's Last Theorem. The modularity theorem was proved completely by Fred Diamond[8], Brian Conrad[9], Richard Taylor and Christophe Breuil[10] in 2001 [10].

- In the past several years our understanding of the arithmetic of the number of partitions of a natural number n, $p(n)$, has increased dramatically. All of the advances have arisen from a single source: the fact that values of the partition function are intimately related to the arithmetic of modular forms. This connection has allowed the application of deep methods of Pierre Deligne, Jean-Pierre Serre, and Goro Shimura to the study of $p(n)$. In fact, there are much deeper connections between partitions and *modular* objects [11].

- There is a deeper relationship between the ζ-function and modular forms than a specialization appearing as a factor of the Eisenstein series. The Riemann ζ-function is an example of an L-series, an object which may be associated to both modular forms and elliptic curves, and which also provides connection between them. The zeta function is the Mellin transform of the Jacobi theta function, a weight 1/2 modular form. This fact, observed and exploited by Bernhard Riemann, is at the root of all later developments relating modular/automorphic forms and L-functions. Specifically regarding the zeroes of the zeta-function, G .H. Hardy used it in his proof thathere are infinitely many zeroes on the critical line[11]. The idea of using families of modular forms to study Riemann Hypothesis is an important one, employed by Henryk Iwaniec among others [12].

[8] Diamond, F. "On deformation rings and Hecke rings." *Annals of Mathematics* (Second Series) 144, no. 1 (1996), 137–166. doi:10.2307/2118586

[9] Conrad, B., Diamond, F. and Taylor, R. "Modularity of certain potentially Barsotti-Tate Galois representations." *Journal of the American Mathematical Society* 12, no. 2 (1999), 521–567. doi:10.1090/S0894-0347-99-00287-8

[10] Breuil, C., Conrad, B., Diamond, F. and Taylor, R. "On the modularity of elliptic curves over **Q**: wild 3-adic exercises." *Journal of the American Mathematical Society* 14, no. 4 (2001), 843–939. doi:10.1090/S0894-0347-01-00370-8

[11] Emerton, M. (https://mathoverflow.net/users/2874/emerton), comment on "Modular forms and the Riemann Hypothesis" URL (version: 2013-10-03): https://mathoverflow.net/q/14083.

- In 1950s and 1970s, Martin Eichler & Goro Shimura (for weight 2), Pierre Deligne (for weight > 2) and Pierre Deligne & Jean-Pierre Serre (for weight 1) proved that any new-form[12] is connected to a Galois representation. A converse of these theorems, conjectured by Jean-Pierre Serre in 1987, was proved in 2009 by Chandrashekhar Khare & Jean-Pierre Wintenberger[13] [8, §17.1].

- On March 14, 2016, Maryna Viazovska posted to the arXiv[14] a solution of the sphere packing problem in eight dimensions. The magic ingredient in Viazovskas proof is a certain special function. Viazovska constructs this function explicitly in terms of modular forms by using an unexpected integral transform, which establishes a new connection between modular forms and discrete geometry [13].

For more details regarding applications and generalizations we can refer to the book by J. H. Bruinier, G. van der Geer, G. Harder, and D. Zagier[15]

[12] These are a type of cusp-form for $\Gamma_1(N)$, see [8, §16.1].

[13] Khare, C. and Wintenberger, J-P. "Serre's modularity conjecture (I)." *Inventiones Mathematicae* 178, no. 3 (2009), 485–504. doi:10.1007/s00222-009-0205-7 and Khare, C. and Wintenberger, J-P. "Serre's modularity conjecture (II)", *Inventiones Mathematicae* 178, no. 3 (2009), 505–586. doi:10.1007/s00222-009-0206-6

[14] Viazovska, M. "The sphere packing problem in dimension 8." (last revised 4 Apr 2017) https://arxiv.org/abs/1603.04246

[15] *The 1-2-3 of modular forms*, Universitext, Springer-Verlag, Berlin, Heidelberg, 2008.

Bibliography

[1] Koblitz, N. *Introduction to Elliptic Curves and Modular Forms* (Graduate Texts in Mathematics 97). New York: Springer-Verlag, 1993.

[2] Murty, M. R., Dewar, M. and Graves, H. *Problems in the Theory of Modular Forms* (IMSc Lecture Notes in Mathematics 1). New Delhi: Hindustan Book Agency, 2015.

[3] Hellegouarch, Y. *Invitation to the Mathematics of Fermat-Wiles.* Massachusetts: Academic Press, 2002.

[4] Stein, E. M. and Shakarchi, R. *Complex Analysis* (Princeton Lectures in Analysis II). Princeton and Oxford: Princeton University Press, 2003.

[5] Conway, J. B. *Functions of One Complex Variable* (Graduate Texts in Mathematics 11). New York: Springer-Verlag, 1973.

[6] Gel'fond, A. O. and Linnik, Yu. V. *Elementary Methods in the Analytic Theory of Numbers.* Oxford: Pergamon Press, 1966.

[7] Katok, S. *Fuchsian groups.* Chicago and London: The University of Chicago Press, 1992.

[8] Ash, A. and Gross, R. *Summing It Up: From One Plus One to Modern Number Theory.* Princeton and Oxford: Princeton University Press, 2016.

[9] Sury, B. "Bernoulli Numbers and the Riemann Zeta Function." *Resonance* 8, no. 7 (2003), 54–62. http://www.ias.ac.in/article/fulltext/reso/008/07/0054-0062

[10] Darmon, H. "A Proof of the Full Shimura-Taniyama-Weil Conjecture is Announced." *Notices of the American Mathematical Society* 46, no. 11 (1999), 1397–1401. http://www.ams.org/notices/199911/comm-darmon.pdf

[11] Ahlgren, S. and Ono, K. "Adding and counting: The arithmetic of partitions." *Notices of the American Mathematical Society* 48, no. 9 (2001), 978–984. http://www.ams.org/notices/200109/fea-ahlgren.pdf

[12] Conrey, J. B. "The Riemann Hypothesis." *Notices of the American Mathematical Society* 50, no. 3 (2003), 341–353. http://www.ams.org/notices/200303/fea-conrey-web.pdf

[13] Cohn, H. "A conceptual breakthrough in sphere packing." *Notices of the American Mathematical Society* 64, no. 2 (2017), 102–115. http://dx.doi.org/10.1090/noti1474

[14] Conrad, K. "Group actions." unpublished essay available at http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/gpaction.pdf. *(accessed on 11 December 2017)*

[15] Conrad, K. "$SL_2(\mathbb{Z})$." unpublished essay available at http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/SL(2,Z).pdf. *(accessed on 11 December 2017)*

[16] Conrad, K. "Groups of order 4 and 6." unpublished essay available at http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/groupsorder4and6.pdf. *(accessed on 15 December 2017)*

[17] Martin, K. "Modular Forms." lecture notes available at http://www2.math.ou.edu/~kmartin/mfs/mfs.pdf. *(accessed on 16 December 2017)*

[18] Korpal, G. "Prime Numbers." *Summer Internship Project Report*, guided by Prof. K. Srinivas (05 June 2017 – 15 July 2017)

[19] Korpal, G. "Diophantine Equations." *Summer Internship Project Report*, guided by Prof. S. A. Katre (18 May 2015 – 16 June 2015)

Prepared in LaTeX 2_ε by Gaurish Korpal