Computing dimensions of spaces of modular forms

Robin Visser

Supervisor: **Dr Dirk Basson**

HONOURS PROJECT, 2018

Stellenbosch University

Contents

Abstract

Let $\mathcal{M}_k(\Gamma)$ denote the space of modular forms of weight k for some finite index subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$. Noting that this forms a finite-dimensional $\mathbb{C}$ vector space, we are interested in computing the dimension of such spaces. Whilst this has been done in full generality using algebro-geometric theorems like Riemann-Roch, we take an alternate approach and use contour integrals to determine bounds on the dimension, as well as giving explicit results for some basic subgroups Γ.

Introduction

The study of modular forms span many different fields of mathematics, which include numerous applications in number theory, combinatorics and algebraic geometry, just to name a few [3]. Indeed, modular forms have been instrumental in resolving many classical problems, from calculating the number of ways to represent a positive integer n using r squares [2, p. 44] [6, p. 19], to being at the centre of Wiles's proof of Fermat's Last Theorem [5].

The modern definition of modular forms came about in the first half of the nineteenth century in the context of elliptic functions [3, p. 1]. Since then, many generalisations have been extensively studied over the past century, including Hilbert, Siegel and Drinfeld modular forms among numerous others [1, 7, 8].

For this project, we study the classical approach to modular forms and prove that the space of modular forms of weight k for some finite index subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ is finite-dimensional. This result is well-known for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and can be established by evaluating a contour integral around the fundamental domain. We generalise this idea by considering, as an example, a contour integral around the fundamental domain for $\Gamma = \Gamma(2)$.

We finally prove a general valence formula for any finite index subgroup Γ which we use to obtain explicit bounds on the dimension. Using these results, we calculate the dimension for some simple examples of congruence subgroups, as well as compare our results obtained with the general dimension formulae derived in the literature using Riemann surface theory [6, 24].

Definitions

We first recall some preliminary definitions:

Let $M_2(\mathbb{Z})$ denote the set of 2 by 2 matrices with integer entries. We define the **modular group** as the elements of $M_2(\mathbb{Z})$ with determinant 1.

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z}) : ad - bc = 1 \right\} \tag{1}$$

It can easily be proven [6, p. 2] that $\mathrm{SL}_2(\mathbb{Z})$ is generated by two matrices S and T where

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{2}$$

(see Appendix for proof by induction).

We denote the **upper half plane** on $\mathbb{C}$ as $\mathcal{H}$, defined as the set of complex numbers with strictly positive imaginary part.

$$\mathcal{H} = \{\tau \in \mathbb{C} \ : \ \mathrm{Im}(\tau) > 0\} \tag{3}$$

We can thus define a group action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$, known as **fractional linear transformations**, where

$$\gamma(\tau) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tau := \frac{a\tau + b}{c\tau + d} \qquad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}). \tag{4}$$

One can easily check that the above definition indeed defines a group action. Note that

$$\mathrm{Im}\left(\frac{a\tau + b}{c\tau + d}\right) = \frac{1}{2i}\left(\frac{a\tau + b}{c\tau + d} + \frac{a\bar{\tau} + b}{c\bar{\tau} + d}\right) = \frac{(ad - bc)\mathrm{Im}(\tau)}{|c\tau + d|^2}. \tag{5}$$

Thus, as all elements in $\mathrm{SL}_2(\mathbb{Z})$ have unit (and therefore positive) determinant, we obtain that if $\tau \in H$, then $\gamma(\tau) \in H$.

It is worth noting that the fractional linear transformation given by some matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ is equivalent to that given by $-\gamma$. Indeed, we have

$$\gamma(\tau) = \frac{a\tau + b}{c\tau + d} = \frac{-a\tau - b}{-c\tau - d} = (-\gamma)(\tau). \tag{6}$$

Thus, we do not have a direct correspondence between fractional linear transformations and matrices in $\mathrm{SL}_2(\mathbb{Z})$. This therefore suggests defining a group where we quotient out the sign of the matrix from $\mathrm{SL}_2(\mathbb{Z})$.

We thus define the **projective special linear group** of 2 by 2 matrices with integer coefficients and unit determinant as $\mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{I, -I\}$. This way, we have a natural 1-to-1 correspondence between elements of $\mathrm{PSL}_2(\mathbb{Z})$ and fractional linear transformations.

We can now finally define modular forms for $\mathrm{SL}_2(\mathbb{Z})$.

Definition 1: [6, p. 4] Let $k \in \mathbb{Z}$. A **modular form** of weight k for $\mathrm{SL}_2(\mathbb{Z})$ is a function $f : \mathcal{H} \to \mathbb{C}$ satisfying the following three conditions:

1. f is holomorphic on $\mathcal{H}$.
2. $f(\gamma(\tau)) = f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k f(\tau) \qquad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$
3. $f(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$.

Recall that a function is **holomorphic** if it is complex differentiable in a neighborhood of every point in its domain. We shall often refer to condition 2 above as the **modularity condition**.

We shall denote the set of all modular forms of weight k for $\mathrm{SL}_2(\mathbb{Z})$ as $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. One can easily note that if $f, g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, then by linearity of the modularity condition, we have $\alpha f + \beta g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for any $\alpha, \beta \in \mathbb{C}$. Thus, $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ forms a vector space over $\mathbb{C}$.

We can furthermore prove the following lemma regarding products of modular forms.

Lemma 2: [6, p. 8] Let f_1 be a modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$ and let f_2 be a modular form of weight k_2 for $\mathrm{SL}_2(\mathbb{Z})$. Then $f_1 f_2$ is a modular form of weight $k_1 + k_2$ for $\mathrm{SL}_2(\mathbb{Z})$.

Proof: We note that the product $f_1 f_2$ is clearly holomorphic on $\mathcal{H}$ and that $f_1 f_2(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$. Thus, all that remains is to prove the modularity condition. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Note that

$$f_1 f_2(\tau) = f_1(\tau) f_2(\tau) = (c\tau + d)^{k_1} f_1(\tau)(c\tau + d)^{k_2} f_2(\tau) = (c\tau + d)^{k_1 + k_2} f_1 f_2(\tau)$$

Thus, $f_1 f_2$ satisfies the modularity condition for weight $k_1 + k_2$ and therefore $f_1 f_2$ is a modular form of weight $k_1 + k_2$ for $\mathrm{SL}_2(\mathbb{Z})$. $\square$

When proving a given holomorphic function f is a modular form, we note that it is sufficient to check the modularity condition only for matrices S and T. Indeed we prove that the second condition is equivalent to the following simpler definition:

Lemma 3: [6, p. 14] Let f be a complex-valued function which is holomorphic on $\mathcal{H}$ and is bounded as $\mathrm{Im}(\tau) \to \infty$. Then f is a modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$ if and only if it satisfies the following equations

$$f(\tau + 1) = f(\tau) \qquad \text{and} \qquad f\left(\frac{-1}{\tau}\right) = \tau^k f(\tau). \tag{7}$$

Proof: We first note that substituting T and S into the modularity condition given in definition 1 yields the two equations respectively given in (7). Thus, if f is a modular form, then f satisfies (7).

Conversely, assume f satisfies both equations given in (7).

We proceed by showing in general that if the modularity condition is satisfied for some two matrices γ_1, $\gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$, then it is satisfied for $\gamma_1\gamma_2$ and for γ_1^{-1}.

Let $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Note that $\gamma_1^{-1} = \begin{pmatrix} d_1 & -b_1 \\ -c_1 & a_1 \end{pmatrix}$, as $\det(\gamma_1) = 1$. We thus have

$$\begin{aligned} f(\gamma_1\gamma_2\tau) &= (c_1(\gamma_2\tau) + d_1)^k f(\gamma_2\tau) \\ &= \left(c_1\left(\frac{a_2\tau + b_2}{c_2\tau + d_2}\right) + d_1\right)^k (c_2\tau + d_2)^k f(\tau) \\ &= (c_1(a_2\tau + b_2) + d_1(c_2\tau + d_2))^k f(\tau) \\ &= ((c_1a_2 + d_1c_2)\tau + c_1b_2 + d_1d_2)^k f(\tau) \end{aligned}$$

thus the modularity condition is satisfied for $\gamma_1\gamma_2$. Similarly checking γ_1^{-1}, we note

$$\begin{aligned} & & f(\gamma_1(\gamma_1^{-1}\tau)) &= (c_1(\gamma_1^{-1}\tau) + d_1)^k f(\gamma_1^{-1}\tau) \\ \implies & & f(\tau) &= \left(c_1\left(\frac{d_1\tau - b_1}{-c_1\tau + a_1}\right) + d_1\right)^k f(\gamma_1^{-1}\tau) \\ \implies & & (-c_1\tau + a_1)^k f(\tau) &= (c_1(d_1\tau - b_1) + d_1(-c_1\tau + a_1))^k f(\gamma_1^{-1}\tau) \\ \implies & & (-c_1\tau + a_1)^k f(\tau) &= (a_1d_1 - b_1c_1)^k f(\gamma_1^{-1}\tau) \end{aligned}$$

As the determinant is one, we obtain: $f(\gamma_1^{-1}\tau) = (-c_1\tau + a_1)^k f(\tau)$ which proves that the modularity condition is satisfied for γ_1^{-1}.

This therefore proves in general that if the modularity condition is satisfied for some finite set of matrices, then it is satisfied by the group generated by those matrices. Since $\langle S, T\rangle = \mathrm{SL}_2(\mathbb{Z})$, it is thus sufficient to check the modularity condition for just S and T. Thus f satisfies the modularity condition for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ $\square$

We conclude by noting some simple examples of modular forms.

Examples: For weight $k = 0$, we have that the modularity condition becomes $f(\frac{a\tau+b}{c\tau+d}) = f(\tau)$. Thus, any **constant function** is a modular form of weight 0 for $\mathrm{SL}_2(\mathbb{Z})$. In fact, we shall later prove that the only modular forms of weight 0 for $\mathrm{SL}_2(\mathbb{Z})$ are the constant functions.

Using $-I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ in the modularity condition yields

$$f(\tau) = (-1)^k f(\tau). \tag{8}$$

Thus, if k is odd, we have $f(\tau) = 0$ for all $\tau \in \mathcal{H}$, and hence there are no non-zero modular forms of odd weight for $\mathrm{SL}_2(\mathbb{Z})$.

For non-trivial examples of modular forms, we turn to Eisenstein series.

Eisenstein series

Let $k \geq 4$ be an even integer. We define [6, p. 4]

$$G_k(\tau) = \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \frac{1}{(m\tau+n)^k}. \tag{9}$$

To check that this defines a modular form, we verify each of the three conditions given in Definition 1.

First, note that $G_k(\tau)$ converges absolutely on $\mathcal{H}$ for $k > 2$. Hence the order of the summation does not matter, and thus the definition given in equation (9) is well-defined. We also note that the sum converges uniformly on compact subsets of $\mathcal{H}$, and thus G_k is holomorphic [6, p. 4].

For simplicity of notation, we define $Z' = \mathbb{Z}^2 \backslash \{(0,0)\}$. To prove the modularity condition, we use Lemma 3 and note

$$G_k(\tau+1) = \sum_{(m,n)\in Z'} \frac{1}{(m\tau+m+n)^k} = \sum_{(m,n)\in Z'} \frac{1}{(m'\tau+n')^k} = G_k(\tau)$$

where the second equality follows from absolute convergence, thus we may rearrange terms, noting that $(m,n) \mapsto (m, m+n)$ is a bijection from $\mathbb{Z}^2\backslash\{(0,0)\}$ to $\mathbb{Z}^2\backslash\{(0,0)\}$.

Furthermore, we have

$$\begin{aligned} G_k\left(\frac{-1}{\tau}\right) &= \sum_{(m,n)\in Z'} \frac{1}{(m(\frac{-1}{\tau})+n)^k} = \sum_{(m,n)\in Z'} \frac{\tau^k}{(-m+n\tau)^k} \\ &= \tau^k \sum_{(m,n)\in Z'} \frac{1}{(n\tau-m)^k} = \tau^k \sum_{(m',n')\in Z'} \frac{1}{(m'\tau+n')^k} \\ &= \tau^k G_k(\tau) \end{aligned}$$

where, as before, the second last equality follows from rearranging terms. Thus G_k satisfies the modularity condition for both S and T, and thus satisfies the condition for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Finally, to investigate the behaviour as $\text{Im}(\tau) \to \infty$, we can rewrite $G_k(\tau)$ by splitting the sum

$$\begin{aligned} G_k(\tau) &= \sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{1}{n^k} + \sum_{m \in \mathbb{Z} \setminus \{0\}} \sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \\ &= 2 \sum_{n=1}^{\infty} \frac{1}{n^k} + 2 \sum_{m=1}^{\infty} \sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \end{aligned} \tag{10}$$

where the first equality follows by extracting the terms with $m = 0$, and the second equality follows by noting that $n^{-k} = (-n)^{-k}$ and $(m\tau + n)^{-k} = (-m\tau - n)^{-k}$ as k is even.

For each $m \in \mathbb{Z}^+$, we define a function $b_m(\tau)$ denoting the inner sums in equation (10), where

$$b_m(\tau) = \sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \qquad \text{for all } \tau \in \mathcal{H}$$

We can prove that, for each $m \in \mathbb{Z}^+$, we have $b_m(\tau)$ converges to 0 as $\text{Im}(\tau) \to \infty$. As $G_k(\tau) = G_k(\tau + 1)$, we may assume that $|\text{Re}(\tau)| \leq 1$. Note that

$$\begin{aligned} \left| \frac{1}{(m\tau + n)^k} \right| = \frac{1}{|m\tau + n|^k} &\leq \frac{1}{\max(|\text{Re}(m\tau + n)|, |\text{Im}(m\tau + n)|)^k} \\ &= \frac{1}{\max(|m\text{Re}(\tau) + n|, |m\text{Im}(\tau)|)^k}. \end{aligned}$$

Note that, as $|\text{Re}(\tau)| \leq 1$, we have $m|\text{Re}(\tau)| \leq m$. Thus, when $|n| > |m| + |m\text{Im}(\tau)|$, we have $|m\text{Re}(\tau) + n| \geq |m\text{Im}(\tau)|$. Therefore, for $|n| > |m| + |m\text{Im}(\tau)|$, we have

$$\left| \frac{1}{(m\tau + n)^k} \right| \leq \frac{1}{\max(|m\text{Re}(\tau) + n|, |m\text{Im}(\tau)|)^k} = \frac{1}{|m\text{Re}(\tau) + n|^k}$$

and for $|n| \leq |m| + |m\text{Im}(\tau)|$, we have

$$\left| \frac{1}{(m\tau + n)^k} \right| \leq \frac{1}{\max(|m\text{Re}(\tau) + n|, |m\text{Im}(\tau)|)^k} \leq \frac{1}{|m\text{Im}(\tau)|^k}$$

Therefore, denoting $r = |m| + |m\text{Im}(\tau)|$, we obtain

$$\begin{aligned} |b_m(\tau)| = \left| \sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k} \right| &\leq \sum_{n \in \mathbb{Z}} \left| \frac{1}{(m\tau + n)^k} \right| \\ &\leq \sum_{|n| \leq r} \frac{1}{|m\text{Im}(\tau)|^k} + \sum_{|n| > r} \frac{1}{|m\text{Re}(\tau) + n|^k} \\ &= \frac{2(|m| + |m\text{Im}(\tau)|) + 1}{|m\text{Im}(\tau)|^k} + \sum_{|n| > r} \frac{1}{|m\text{Re}(\tau) + n|^k}. \end{aligned}$$

Letting $\text{Im}(\tau) \to \infty$, we have that the first term goes to 0, as $k \geq 4$. For the second term we note that $\sum_{|n| > m} |m\text{Re}(\tau) + n|^{-k}$ converges for $k \geq 4$, thus the tail of the infinite sum must converge to 0, noting that $r \to \infty$ as $\text{Im}(\tau) \to \infty$.

Therefore, we obtain

$$\lim_{\text{Im}(\tau) \to \infty} |b_m(\tau)| = 0$$

which thus implies

$$\lim_{\text{Im}(\tau) \to \infty} b_m(\tau) = 0$$

as absolute convergence implies convergence. Hence, each of the inner sums in equation (10) converges to 0. To prove that the entire second term converges to 0, we make use of the fact that $G_k(\tau)$ not only converges uniformly on compact sets, but in fact converges uniformly on infinite half-strips $\Omega_{A,B}$ [6, p. 8] defined as

$$\Omega_{A,B} = \{\tau \in \mathcal{H} \mid |\mathrm{Re}(\tau)| \leq A, \mathrm{Im}(\tau) \geq B\}$$

where A and B are two fixed positive real numbers. This uniform converges ensures that

$$\sum_{m=1}^{\infty} b_m(\tau) = \sum_{m=1}^{\infty} \sum_{n \in \mathbb{Z}} \frac{1}{(m\tau + n)^k}$$

converges to 0 as $\mathrm{Im}(\tau) \to \infty$.

Therefore, we obtain the following explicit limit value for $G_k(\tau)$ as $\mathrm{Im}(\tau) \to \infty$

$$\lim_{\mathrm{Im}(\tau) \to \infty} G_k(\tau) = 2 \sum_{n=1}^{\infty} \frac{1}{n^k} = 2\zeta(k) \tag{11}$$

where ζ denotes the Riemann-zeta function. In particular, we have that $G_k(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$, and thus satisfies condition 3 in Definition 1.

This therefore gives a non-trivial example of a modular form of weight $k \geq 4$ where k is even. Note that if k is odd, then G_k is simply the zero function as expected, noting the observation in equation (8).

Congruence subgroups

One can generalise the notion of a modular form for $\mathrm{SL}_2(\mathbb{Z})$ to subgroups Γ of $\mathrm{SL}_2(\mathbb{Z})$.

We first introduce the slash operator of weight k

$$(f|_k\gamma)(\tau) := (cz + d)^{-k} f(\gamma\tau) = \frac{1}{(c\tau + d)^k} f\left(\frac{a\tau + b}{c\tau + d}\right). \tag{12}$$

Note that in the special case of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, we have $(f|_k\gamma)(\tau) = f(\tau)$ for all $\tau \in \mathcal{H}$ and all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ by the modularity condition. We now define modular forms for arbitrary subgroups Γ of $\mathrm{SL}_2(\mathbb{Z})$.

Definition 4: [6, p. 17] Let $k \in \mathbb{Z}$ and let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. A modular form of weight k for Γ is a function $f : \mathcal{H} \to \mathbb{C}$ satisfying the following three conditions:

1. f is holomorphic on $\mathcal{H}$.
2. $f(\gamma(\tau)) = f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k f(\tau) \qquad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$
3. $(f|_k\gamma)(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty \qquad$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$

In the special case that $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, we easily observe that the above definition is identical to that given in Definition 1. We also observe that, if f is a modular form of weight k for some subgroup Γ, then it is also a modular form of weight k for any smaller subgroup Γ' where $\Gamma' \subseteq \Gamma$.

We shall denote the set of all modular forms of weight k for Γ as $\mathcal{M}_k(\Gamma)$. As with $\mathrm{SL}_2(\mathbb{Z})$, we have that $\mathcal{M}_k(\Gamma)$ forms a vector space over $\mathbb{C}$. One can also easily prove that if $f_1 \in \mathcal{M}_{k_1}(\Gamma)$

and $f_2 \in \mathcal{M}_{k_2}(\Gamma)$ then $f_1 f_2 \in \mathcal{M}_{k_1+k_2}(\Gamma)$.

Definition 5: [6, p. 13] Let N be a positive integer. We define the **principal congruence subgroup** of level N as

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\}. \tag{13}$$

Note that the matrix congruence is entry-wise. In other words, $a, c \equiv 1 \pmod N$ and $b, c \equiv 0 \pmod N$.

We shall prove that $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$ and obtain an explicit expression for the index.

Theorem 6: [6, p. 13] Let N be a positive integer. Then $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$ with the index given by

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right). \tag{14}$$

Proof: We define a natural homomorphism $\phi : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}\backslash N\mathbb{Z})$ given by the natural mapping

$$\phi \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} [a]_N & [b]_N \\ [c]_N & [d]_N \end{pmatrix}$$

where $[x]_N$ denotes the congruence class of x modulo N. It can easily be seen that ϕ is a homomorphism. Calculating the kernel, we obtain

$$\phi(\gamma) = I \iff \begin{pmatrix} [a]_N & [b]_N \\ [c]_N & [d]_N \end{pmatrix} = I \iff \gamma \in \Gamma(N).$$

Thus $\Gamma(N)$ is the kernel of ϕ, and thus $\Gamma(N)$ is a normal subgroup in $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, we can show that ϕ is surjective.

Firstly, if $N = 1$, then $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ consists of a single element whereby ϕ is a constant map and thus trivially surjective. We therefore now assume $N > 1$.

Indeed, let $\delta = \begin{pmatrix} [a]_N & [b]_N \\ [c]_N & [d]_N \end{pmatrix}$ be an arbitrary element in $\mathrm{SL}_2(\mathbb{Z}\backslash N\mathbb{Z})$. Thus, we have

$$ad - bc \equiv 1 \pmod N \qquad \Longrightarrow \qquad ad - bc + qN = 1$$

for some $q \in \mathbb{Z}$. Therefore $\gcd(a, b, N) = 1$ and in particular we have $\gcd(a, b)$ is coprime to N. Note that we may assume a, b, c, d are all positive integers, as $[x]_N = [x + N]_N$ for any integer x. We define $g = \gcd(a, b)$ and consider the following:

Let $\{p_{g_1}, p_{g_2}, \ldots, p_{g_k}\}$ be the set of all primes dividing g and let $\{p_{a_1}, p_{a_2}, \ldots, p_{a_l}\}$ be the set of all primes dividing a and not dividing g (note that either or both sets may be empty). Let t be an integer satisfying the following congruence conditions:

$$\begin{aligned} t &\equiv 0 \pmod{p_{a_i}} \quad \text{for all } i \in \{1, 2, \ldots, l\} \quad \text{and} \\ t &\equiv 1 \pmod{p_{g_i}} \quad \text{for all } i \in \{1, 2, \ldots, k\} \end{aligned}$$

Note that each of the prime moduli are pairwise distinct by definition. Thus, by the Chinese Remainder theorem, such a $t \in \mathbb{Z}$ must exist.

Now, define $b' = b + tN$. We now claim that $\gcd(a, b') = 1$. Indeed, assume for contradiction that there exists some prime q such that $q|a$ and $q|b'$. We consider two cases:

Case 1: q divides g. Thus, $q|\gcd(a, b)$ which implies $q|b$, hence $q|b' - b = tN$. However, by construction of t we have $t \equiv 1 \pmod{q}$. Thus $N \equiv tN \equiv 0 \pmod{q}$ and hence q divides N. Therefore $q|\gcd(a, b, N)$ which contradicts $\gcd(a, b, N) = 1$.

Case 2: q does not divide g. As $q|a$, we have by construction of t that $q|t$. Thus $q|b' - tN = b$. Therefore, as $q|a$ and $q|b$, this implies $q|\gcd(a, b)$, contradicting our assumption that $q \not| g$.

Therefore, we conclude that $\gcd(a, b') = 1$. Thus by Bézout's identity, there exist $x, y \in \mathbb{Z}$ such that $ax - b'y = 1$. Now, define $c \in \mathbb{Z}$ and $d \in \mathbb{Z}$ by:

$$\begin{aligned} c' &= c + y(1 - (ad - b'c)) \\ d' &= d + x(1 - (ad - b'c)) \end{aligned}$$

and define $\gamma = \begin{pmatrix} a & b' \\ c' & d' \end{pmatrix}$. We prove that $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Indeed

$$\begin{aligned} \det(\gamma) &= ad' - b'c' \\ &= a(d + c(1 - (ad - b'c))) - b'(c + y(1 - (ad - b'c))) \\ &= ad - b'c + (ax - b'y)(1 - (ad - b'c)) \\ &= ad - b'c + (1 - (ad - b'c)) = 1 \end{aligned}$$

and thus $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Furthermore, we note

$$\begin{aligned} c' &= c + y(1 - (ad - b'c)) \equiv c + yqN \equiv c \pmod{N} \\ d' &= d + x(1 - (ad - b'c)) \equiv d + xqN \equiv d \pmod{N}. \end{aligned}$$

Thus we finally obtain

$$\phi(\gamma) = \phi\begin{pmatrix} a & b' \\ c' & d' \end{pmatrix} = \phi\begin{pmatrix} a & b + tN \\ c' & d' \end{pmatrix} = \begin{pmatrix} [a]_N & [b + tN]_N \\ [c']_N & [d']_N \end{pmatrix} = \begin{pmatrix} [a]_N & [b]_N \\ [c]_N & [d]_N \end{pmatrix} = \delta.$$

Thus, this proves ϕ is surjective.

We therefore obtain by the first isomorphism theorem that ϕ induces the isomorphism:

$$\mathrm{SL}_2(\mathbb{Z})/\Gamma(N) \cong \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}).$$

Hence, noting that $|\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})| \leq |M_2(\mathbb{Z}/N\mathbb{Z})| = N^4$, we have that $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is a finite group and thus $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$.

Indeed, the index can specifically be given as the order of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ where:

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = |\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})| = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right) \tag{15}$$

where the order of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is given in the Appendix.

Definition 7: [6, p. 13] Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We say that Γ is a **congruence subgroup** if $\Gamma(N) \subseteq \Gamma$ for some positive integer N. In this case, Γ is a congruence subgroup of level N.

Note that since $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$, this implies that every congruence subgroup is a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Two important congruence subgroups of note are

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} (\text{mod } N) \right\} \tag{16}$$

where '*' simply means unspecified (i.e. the only condition for matrices in $\Gamma_0(N)$ are that N divides c)

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} (\text{mod } N) \right\}. \tag{17}$$

We conclude this section by simply quoting the index of $\Gamma_0(N)$ and $\Gamma_1(N)$ in $\mathrm{SL}_2(\mathbb{Z})$, where the relevant proofs can be obtained from [10, p. 22].

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] = N \prod_{p|N} \left(1 + \frac{1}{p}\right) \tag{18}$$

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(N)] = N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right) \tag{19}$$

q - expansions

Recall that any modular form for $\mathrm{SL}_2(\mathbb{Z})$ satisfies $f(\tau+1) = f(\tau)$. Thus f is always periodic with period 1. We can therefore describe f in terms of a new variable $q := \exp 2\pi i\tau$. We denote $\tilde{f}(q) = f(\tau)$.

We easily note this is well-defined. Indeed, let $q = \exp 2\pi i\tau$ and $q' = \exp 2\pi i\tau'$. Then

$$\begin{aligned} q = q' \quad &\implies \quad \exp 2\pi i\tau = \exp 2\pi i\tau' &&\implies \quad \tau = \tau' + n \qquad \text{for some } n \in \mathbb{Z} \\ &\implies \quad f(\tau) = f(\tau' + n) &&\implies \quad f(\tau) = f(\tau') \qquad \text{(by periodicity)} \\ &\implies \quad \tilde{f}(q) = \tilde{f}(q'). \end{aligned}$$

Also note, if $\tau = x + iy \in \mathcal{H}$, then

$$|q| = |\exp 2\pi i\tau| = |\exp 2\pi ix||\exp -2\pi y| = |\exp -2\pi y| \le 1 \tag{20}$$

where the last equality follows since $y > 0$. Thus, the change of variables sends the upper half plane $\mathcal{H}$, to the punctured open unit disc $D' = \{q \in \mathbb{C} \mid 0 < |q| < 1\}$.

We also note that as $\mathrm{Im}(\tau) \to \infty$, we have $q \to 0$. Hence the condition that $f(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$ implies that $\tilde{f}(q)$ is bounded as $q \to 0$. Therefore, by the Riemann Removable Singularity Theorem, we have that $\lim_{q\to 0} \tilde{f}(q)$ exists and furthermore, that the extension of $\tilde{f}$ to the point $q = 0$ defined by $\tilde{f}(0) := \lim_{q\to 0} \tilde{f}(q)$ is analytic on the unit disc $D = \{q \in \mathbb{C} \mid 0 \le |q| < 1\}$.

Therefore, f has a Fourier expansion, given as

$$f(\tau) = \tilde{f}(q) = \sum_{n=0}^{\infty} a_n(f) q^n \quad \text{where } q = e^{2\pi i\tau}. \tag{21}$$

As an example, we recall the Eisenstein series $G_k(\tau)$ of weight k for k even, $k \geq 4$. One can prove that the q-expansion of $G_k(\tau)$ can be given as [6, p. 5]

$$\tilde{G}_k(q) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n \tag{22}$$

where $\sigma_{k-1}(n)$ denotes the sum of the $(k-1)$-th powers of the positive divisors of n

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

We also note the **normalised Eisenstein series**, defined as $E_k(\tau) = G_k(\tau)/(2\zeta(k))$, where the constant coefficient in the q-expansion normalises to 1. Noting that ζ evaluated at the even integers gives [6, p. 10]

$$\zeta(k) = (-1)^{k/2+1}\frac{B_k(2\pi)^k}{2(k!)} \qquad \text{for } k \text{ even}, k \geq 2$$

we thus obtain

$$\begin{aligned}\tilde{E}_k(q) &= 1 + \frac{(2\pi i)^k}{\zeta(k)(k-1)!}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n \\ &= 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n\end{aligned}$$

Therefore, the coefficients in the q-expansion of the normalised Eisenstein series E_k are all rational with a common denominator. The first few coefficients for E_4 and E_6 are given as

$$\begin{aligned}\tilde{E}_4(q) &= 1 + 240q + 2160q^2 + 6720q^3 + \mathcal{O}(q^4) \\ \tilde{E}_6(q) &= 1 - 504q - 16632q^2 - 122976q^3 + \mathcal{O}(q^4)\end{aligned}$$

We are often interested in the set of modular forms which has a constant conefficient of 0 in their q-expansion. This leads to the following definition:

Definition 8: [6, p. 6] Let $k \in \mathbb{Z}$. A **cusp form** of weight k for $\mathrm{SL}_2(\mathbb{Z})$ is a modular form f of weight k for $\mathrm{SL}_2(\mathbb{Z})$ such that $a_0(f) = 0$ (i.e. the leading coefficient in the Fourier expansion of f is 0).

The space of all cusp forms of weight k is denoted $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$. Since $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ is closed under $\mathbb{C}$-linear combinations, we have that $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ is a subspace of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

One can prove that the lowest weight cusp form for $\mathrm{SL}_2(\mathbb{Z})$ occurs at weight $k = 12$. Indeed, we note that E_4^3 and E_6^2 are linearly independent in $\mathcal{M}_{12}(\mathrm{SL}_2(\mathbb{Z}))$, where both have leading coefficients of 1. We therefore define the cusp form $\Delta \in \mathcal{S}_{12}(\mathrm{SL}_2(\mathbb{Z}))$ where

$$\Delta = \frac{E_4^3 - E_6^2}{1728} \tag{23}$$

(we include the factor 1728 to normalise Δ, ensuring that the linear coefficient is 1). Doing a q-expansion of Δ yields

$$\tilde{\Delta}(q) = q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 + \mathcal{O}(q^6).$$

Finally, we remark that Δ has an infinite product representation given by [10, p. 99]

$$\tilde{\Delta}(q) = q \prod_{n=1}^{\infty} (1 - q^n)^{24}. \tag{24}$$

We shall study product representations of other cusp forms in more detail when computing the dimension of $\mathcal{M}_k(\Gamma)$ for various other subgroups Γ.

Fundamental domains

When working with a modular form f for Γ, we note that the function value at some point $z \in \mathbb{C}$ determines the function value at the set of all points $\Gamma z = \{\gamma z \mid \gamma \in \Gamma\}$ by the modularity condition. We are thus interested in a minimal set of points on the complex plane that fully determines f on the upper half plane $\mathcal{H}$.

Definition 9: Let f be a modular form of weight k for some finite index subgroup Γ. We define a **fundamental domain** of f as a subset $\mathcal{F} \subseteq \mathbb{C}$ satisfying the following two conditions:

1. $\Gamma\mathcal{F} = \{\gamma z \mid \gamma \in \Gamma \text{ and } z \in \mathcal{F}\} = \mathcal{H}$
2. $\text{int}(\mathcal{F}) \cap \text{int}(\gamma\mathcal{F}) = \emptyset \qquad \text{for all } \gamma \in \Gamma - \{I, -I\}$

Some authors [17, p. 14] also include the condition that $\mathcal{F}$ is the closure of its interior, and that $\mathcal{F}$ is connected.

We can restate the first condition as simply that every orbit induced under the group action of Γ on $\mathcal{H}$ has at least one representative element in $\mathcal{F}$. The second condition states that no two elements in the interior of $\mathcal{F}$ should belong to the same orbit (i.e. the only possible exceptions are the boundaries).

Clearly, there are infinitely many choices for a fundamental domain $\mathcal{F}$. Most authors [10, p.25], however, work with the following fundamental domain for $\text{SL}_2(\mathbb{Z})$, shown in Figure 1.

$$\mathcal{F} = \{\tau \in \mathcal{H} \mid |\text{Re}(\tau)| \leq \frac{1}{2} \text{ and } |\tau| \geq 1\} \tag{25}$$

To give a graphical picture of the first condition for the fundamental domain $\mathcal{F}$, we plot $\gamma(\mathcal{F})$ for various matrices $\gamma \in \text{SL}_2(\mathbb{Z})$, as shown in Figure 2. Note that our fundamental domain simply corresponds to $\mathcal{F} = I(\mathcal{F})$, however any other translate of $\mathcal{F}$ would suffice as a fundamental domain.

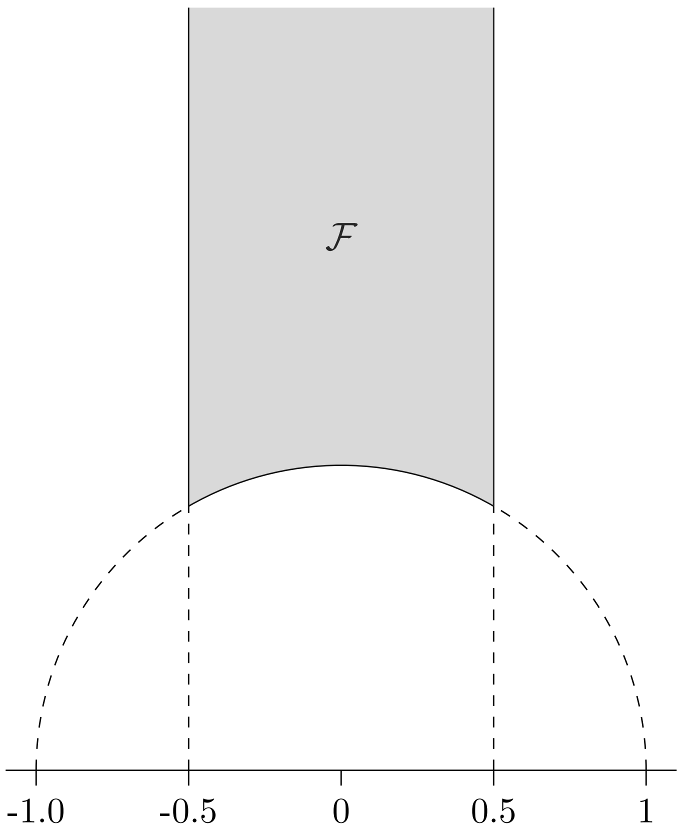

Figure 1: Fundamental Domain $\mathcal{F}$ for $\mathrm{SL}_2(\mathbb{Z})$

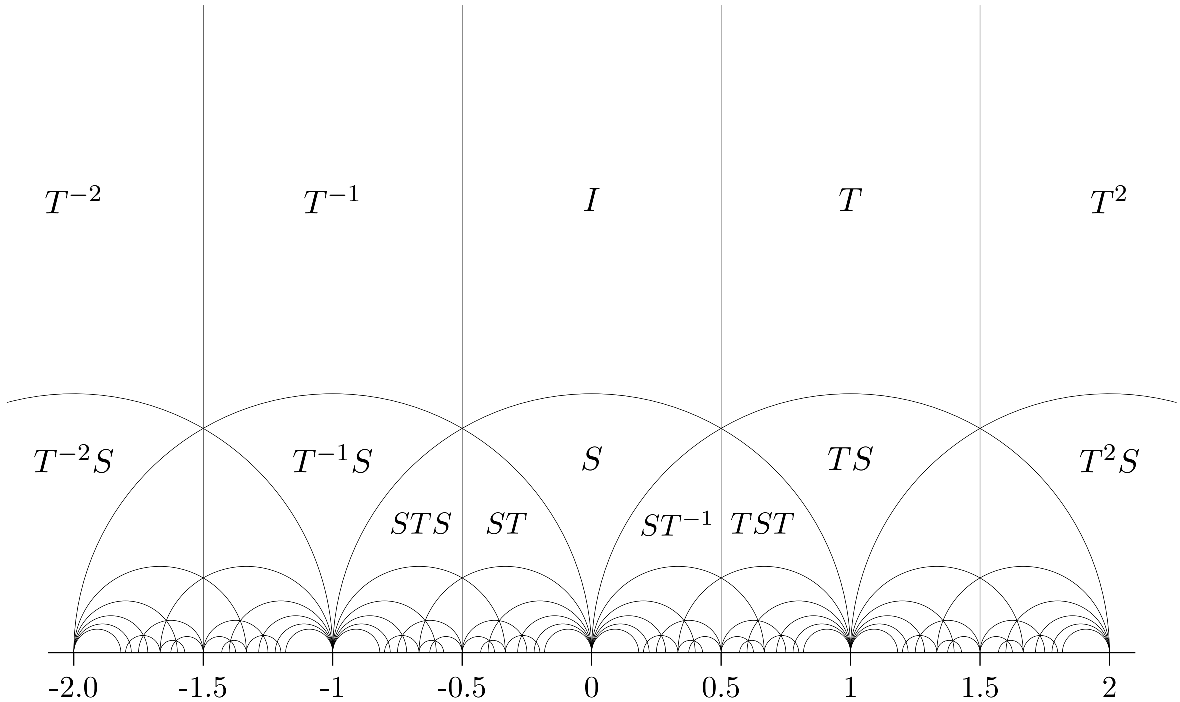

Figure 2: Partition of $\mathcal{H}$ into translates $\gamma(\mathcal{F})$ as γ ranges over $\mathrm{SL}_2(\mathbb{Z})$

Elliptic points

We remarked earlier the motivation for defining $\mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{I, -I\}$, noting that both matrices I and $-I$ act trivially at every point $\tau \in \mathcal{H}$. That is $I\tau = (-I)\tau = \tau$ for all $\tau \in \mathcal{H}$. We are furthermore interested in calculating the points $\tau \in \mathcal{H}$ for which there exist matrices $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ other than I and $-I$ such that $\gamma(\tau) = \tau$.

Definition 10: Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We define the **stabiliser subgroup** of Γ with respect to τ (also denoted the isotropy subgroup of τ) as

$$\Gamma_\tau = \{\gamma \in \Gamma \;:\; \gamma(\tau) = \tau\}.$$

We note that Γ_τ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})_\tau$. Our aim is to calculate $|\mathrm{SL}_2(\mathbb{Z})_\tau|$ for all $\tau \in \mathcal{H}$. One can easily show that, for any $\delta \in \mathrm{SL}_2(\mathbb{Z})$, we have $\mathrm{SL}_2(\mathbb{Z})_{\delta\tau} = \delta \mathrm{SL}_2(\mathbb{Z})_\tau \delta^{-1}$ [20, p. 225], and thus $|\mathrm{SL}_2(\mathbb{Z})_{\delta\tau}| = |\mathrm{SL}_2(\mathbb{Z})_\tau|$. Therefore, we may restrict our attention without loss of generality to points in the fundamental domain $\mathcal{F}$.

Let $\tau \in \mathcal{F}$. Clearly $\{I, -I\} \subseteq \mathrm{SL}_2(\mathbb{Z})_\tau$. Let us assume there exists some $\gamma_0 \in \Gamma$ such that $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})_\tau$ and $\gamma_0 \notin \{I, -I\}$. Thus

$$\gamma_0(\tau) = \tau \quad \Longrightarrow \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tau = \tau \quad \Longrightarrow \quad \frac{a\tau + b}{c\tau + d} = \tau$$
$$\Longrightarrow \quad c\tau^2 + (d-a)\tau - b = 0.$$

Assume for contradiction that $c = 0$. Then either $a = d$ or $\tau = b/(d-a)$. As $\tau \in \mathcal{H}$ (and thus $\tau \notin \mathbb{Q}$), we have the former case $a = d$, and therefore $b = 0$. Since $\det(\gamma_0) = 1$, this implies $ad = 1$ and hence $a = d = 1$ or $a = d = -1$. We therefore obtain

$$\gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \quad \text{or} \quad \gamma_0 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$$

which contradicts $\gamma_0 \notin \{I, -I\}$. Thus $c \neq 0$. Hence, we obtain

$$\tau = \frac{a - d \pm \sqrt{(a-d)^2 + 4bc}}{2c}.$$

As $\det(\gamma_0) = ad - bc = 1$, this yields

$$\tau = \frac{a - d \pm \sqrt{(a-d)^2 + 4(ad-1)}}{2c} = \frac{a - d \pm \sqrt{(a+d)^2 - 4}}{2c}.$$

Since $\tau \in \mathcal{H}$, we have

$$(a+d)^2 - 4 < 0 \quad \Longrightarrow \quad |a+d| < 2 \quad \Longrightarrow \quad a + d \in \{-1, 0, +1\}.$$

We therefore consider the different cases for $a + d$.
Case 1: $a + d = 0$. We therefore obtain

$$\tau = \frac{a - d \pm \sqrt{(a+d)^2 - 4}}{2c} = \frac{a}{c} \pm \frac{i}{c}.$$

As $\tau \in \mathcal{F}$, we require $\mathrm{Im}(\tau) \geq \frac{\sqrt{3}}{2}$, which thus implies $|c| = 1$. Furthermore, we have $|\mathrm{Re}(\tau)| \leq 1/2$, hence $a = 0$, which gives us $\tau = i$.

Case 2: $a + d = \pm 1$. We thus obtain

$$\tau = \frac{a - d \pm \sqrt{(a+d)^2 - 4}}{2c} = \frac{2a \pm 1}{2c} \pm \frac{\sqrt{3}}{2c} i.$$

As before, the condition $\mathrm{Im}(\tau) \geq \frac{\sqrt{3}}{2}$ implies that $|c| = 1$. Also, as $|\mathrm{Re}(\tau)| \leq 1/2$, this implies $|2a \pm 1| \leq 1$. However, as $2a \pm 1$ is odd, we thus have the equality $|2a \pm 1| = 1$ which therefore gives us

$$\tau = \pm \frac{1}{2} + \frac{\sqrt{3}}{2} i$$

We define $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2} i$ and note that $S\omega = \frac{1}{2} + \frac{\sqrt{3}}{2} i$. We have thus proven that if $\tau \in \mathcal{H}$ has non-trivial stabiliser subgroup $\mathrm{SL}_2(\mathbb{Z})_\tau$, then τ must either be in the orbit of i or ω.

We now explicitly calculate $\mathrm{SL}_2(\mathbb{Z})_i$. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})_i$ such that $\gamma \notin \{I, -I\}$. By the calculation above, we have $a = d = 0$. As $\det(\gamma) = 1$, this impies $bc = -1$ and thus $\gamma \in \{S, -S\}$. Noting that both S and $-S$ stabilise i, we thus obtain

$$\mathrm{SL}_2(\mathbb{Z})_i = \{I, -I, S, -S\}.$$

In a similar manner, we calculate $\mathrm{SL}_2(\mathbb{Z})_\omega$. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})_\omega$ such that $\gamma \notin \{I, -I\}$. We again consider two cases.

Case 1: $a + d = 1$. We therefore have

$$\omega = \frac{2a - 1}{2c} \pm \frac{\sqrt{3}}{2} i.$$

Thefore, if $c = 1$, we have $a = 0$ and $d = 1$. As before, since $\det(\gamma) = 1$, this yields $b = -1$. If $c = -1$, we have $a = 1$ and $d = 0$, and again since $\det(\gamma) = 1$, this yields $b = 1$.

Case 2: $a + d = -1$. We therefore have

$$\omega = \frac{2a + 1}{2c} \pm \frac{\sqrt{3}}{2} i.$$

If $c = 1$, we obtain $a = -1$, $d = 0$, and $b = -1$. Otherwise, if $c = -1$, we obtain $a = 0$, $d = -1$ and $b = 1$.

Therefore, considering all cases. we obtain

$$\gamma \in \left\{ \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \right\}.$$

One can easily verify all the above elements stabilise ω. Therefore, we finally obtain

$$\mathrm{SL}_2(\mathbb{Z})_\omega = \left\{ I, -I, \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, -\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, -\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \right\}$$

We thus have the following result

$$|\mathrm{SL}_2(\mathbb{Z})_\tau| = \begin{cases} 4 & \text{if } \gamma(\tau) = i \text{ for some } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\ 6 & \text{if } \gamma(\tau) = \omega \text{ for some } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\ 2 & \text{otherwise} \end{cases} \tag{26}$$

which can also stated as

$$|\mathrm{PSL}_2(\mathbb{Z})_\tau| = \begin{cases} 2 & \text{if } \gamma(\tau) = i \text{ for some } \gamma \in \mathrm{PSL}_2(\mathbb{Z}) \\ 3 & \text{if } \gamma(\tau) = \omega \text{ for some } \gamma \in \mathrm{PSL}_2(\mathbb{Z}) \\ 1 & \text{otherwise} \end{cases} \tag{27}$$

We conclude that $\mathrm{SL}_2(\mathbb{Z})_\tau$ and thus Γ_τ is always finite for any subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$.

Definition 11: [6, p. 48] Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. A point $\tau \in \mathcal{H}$ is an **elliptic point** if the stabiliser subgroup Γ_τ is non-trivial. That is, there exists $\gamma \in \Gamma_\tau$ such that $\gamma \notin \{I, -I\}$.

Let $\bar{\Gamma}_\tau = \{\pm I\}\Gamma_\tau/\{\pm I\}$. We furthermore denote $\tau \in \mathcal{H}$ as an **elliptic point of order 2** if $|\bar{\Gamma}_\tau| = 2$ and denote $\tau \in \mathcal{H}$ as an elliptic point of order 3 if $|\bar{\Gamma}_\tau| = 3$. From (27), we have that all elliptic points either have order 2 or 3 and that all elliptic points of order 2 and 3 lie in the orbit of i and ω respectively.

We end this section with a simple remark that if Γ is finite index in $\mathrm{SL}_2(\mathbb{Z})$, then $\Gamma\backslash\mathcal{H}$ contains finitely many elliptic points [20, p. 225]. Indeed, let R be a finite set of coset representatives for $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$. If $\Gamma\tau$ is an elliptic point, then we must have

$$\Gamma\tau \in \{\Gamma\gamma_r(i),\, \Gamma\gamma_r(\omega) \;:\; \gamma_r \in R\}.$$

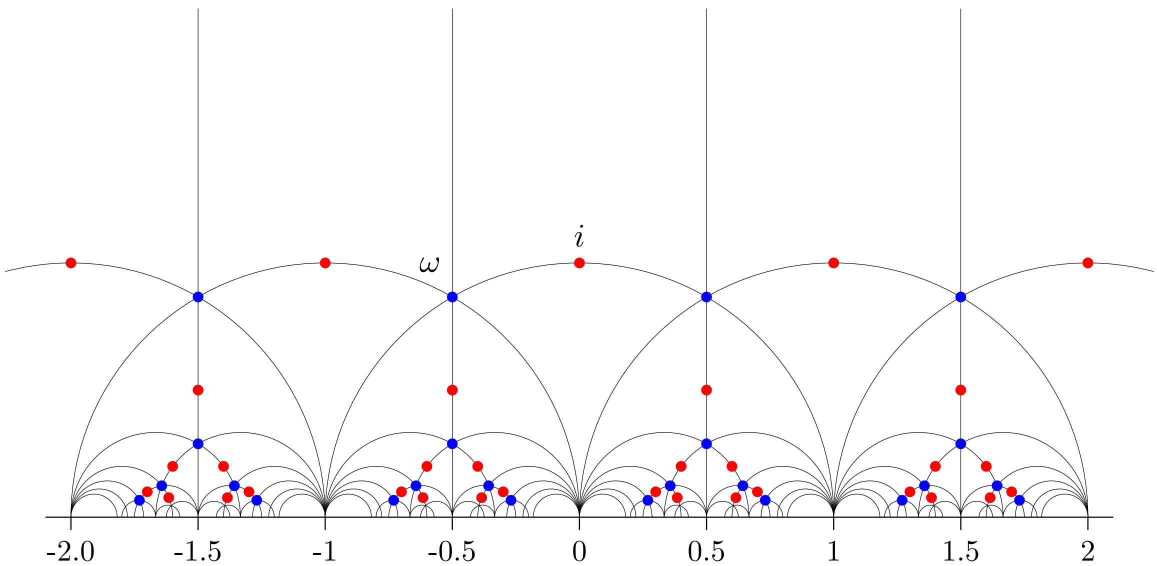

Figure 3: Some of the elliptic points denoted on $\mathcal{H}$. The points in the orbit of i are drawn in red, and the points in the orbit of ω are drawn in blue.

Modular curves and cusps

For any congruence subgroup Γ in $\mathrm{SL}_2(\mathbb{Z})$, we define the **modular curve** $Y(\Gamma)$ as the set of orbits induced by the action of Γ on $\mathcal{H}$ [6, p. 45]

$$Y(\Gamma) = \Gamma\backslash\mathcal{H} = \{\Gamma\tau \mid \tau \in \mathcal{H}\}. \tag{28}$$

One can prove that $Y(\Gamma)$ forms a Riemann surface which is Haussdorf [6, p. 47]. However, $Y(\Gamma)$ on its own is not compact. To compactify this surface, we need to extend the action of Γ on $\mathcal{H}$

to include the projective line over the the rationals $\mathbb{Q}$.

We note the projective line over $\mathbb{Q}$ is the set

$$\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}.$$

We noted previously that the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$ gives fractional linear transformations on $\mathcal{H}$. We can extend this action to $\mathbb{P}^1(\mathbb{Q})$ as well, where we define

$$\gamma\left(\frac{p}{q}\right) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{p}{q} := \frac{ap+bq}{cp+dq} \qquad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

and we define

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \infty = \frac{a}{c} \qquad \text{and} \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{-d}{c} = \infty.$$

We therefore have that $\mathrm{SL}_2(\mathbb{Z})$ also acts on $\mathbb{P}^1(\mathbb{Q})$. Indeed, for an arbitrary subgroup Γ, we let Γ act on $\mathcal{H} \cup \mathbb{P}^1(\mathbb{Q}) = \mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$. Taking the extended quotient induced by this action give us the compactified Riemann surface $X(\Gamma)$ [6, p. 58], defined as

$$X(\Gamma) = \Gamma\backslash\mathcal{H} \cup \mathbb{P}^1(\mathbb{Q}) = Y(\Gamma) \cup \Gamma\backslash\mathbb{P}^1(\mathbb{Q}) = Y(\Gamma) \cup \Gamma\backslash(\mathbb{Q} \cup \{\infty\}). \tag{29}$$

We therefore note that $X(\Gamma)$ consists of adjoining the points $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ to the modular curve $Y(\Gamma)$, thus compactifying the Riemann surface. We define these points $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ as the **cusps** of $X(\Gamma)$, which we shall also denote as $\mathrm{Cusps}(\Gamma)$.

The cusps are hence simply the set of orbits in $\mathbb{P}^1(\mathbb{Q})$ induced by the action of Γ. We first determine the cusps for $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 12: $X(\mathrm{SL}_2(\mathbb{Z}))$ contains a single cusp.

Proof: We simply prove that $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$. Indeed, let $c \in \mathbb{Q}$ be given, and denote $c = \frac{p}{q}$ where p and q are coprime. Thus, by Bézout's identity, there exists $r, s \in \mathbb{Z}$ such that $pr - qs = 1$. Now, define $\gamma := \begin{pmatrix} p & s \\ q & r \end{pmatrix}$. By definition of r and s we have that $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Note

$$\gamma(\infty) = \begin{pmatrix} p & s \\ q & r \end{pmatrix} \infty = \frac{p}{q} = c. \tag{30}$$

Thus, c is in the orbit of ∞, for all $c \in \mathbb{Q}$. Hence, the action of $\mathrm{SL}_2(\mathbb{Z})$ induces a single orbit on $\mathbb{P}^1(\mathbb{Q})$, and thus $\mathrm{SL}_2(\mathbb{Z})$ acts transitively. □

We can furthermore prove that for any finite-index subgroup Γ, $X(\Gamma)$ consists of finitely many cusps.

Theorem 13: [6, p. 58] Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Then $X(\Gamma)$ contains finitely many cusps.

Proof: [2, p. 34] We first note that the stabiliser of $\mathrm{SL}_2(\mathbb{Z})$ with respect to ∞ is

$$\mathrm{SL}_2(\mathbb{Z})_\infty = \left\{ \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{Z} \right\}. \tag{31}$$

Indeed, let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})_\infty$. Then

$$\gamma\infty = \infty \quad \implies \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\infty = \infty \quad \implies \quad \frac{a}{c} = \infty \quad \implies \quad c = 0.$$

Thus, as $\det(\gamma) = 1$, we have $ad = 1$ and hence $a = d = 1$ or $a = d = -1$. We therefore have $\gamma = \pm\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$. Conversely, it can easily be seen that any such γ of that form stabilises ∞, thus proving (31).

Thus, by the Orbit-Stabilizer theorem, we obtain a bijection between $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{Z})_\infty$ and $\mathbb{P}^1(\mathbb{Q})$ given by the map $\gamma\mathrm{SL}_2(\mathbb{Z})_\infty \mapsto \gamma(\infty)$.

We now consider the function $\phi : \Gamma\backslash\mathrm{SL}_2(\mathbb{Z}) \to \Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ defined by

$$\Gamma\gamma \mapsto \Gamma\gamma(\infty). \tag{32}$$

Note that ϕ is well-defined by associativity of the group action induced by $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, we have that ϕ is surjective, as $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$. Thus the order of $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ is no greater than the order of $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$, and as Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$ this implies $X(\Gamma)$ contains finitely many cusps. □

We shall proceed to define the **width** at each of the cusps of Γ. Let $s \in \mathbb{P}^1(\mathbb{Q})$ and let $\Gamma s \in \mathrm{Cusps}(\Gamma)$ be a cusp of Γ. We consider the stabiliser of s in Γ, defined as

$$\Gamma_s = \{\gamma \in \Gamma \mid \gamma(s) = s\}. \tag{33}$$

Note that by Proposition 12, there exists $\gamma_s \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma_s(\infty) = s$. We now prove that $\Gamma_s = \Gamma \cap \gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1}$. Indeed, for any $\gamma \in \Gamma$, we have

$$\begin{aligned} \gamma \in \Gamma_s \iff \gamma(s) = s \iff \gamma\gamma_s(\infty) = \gamma_s(\infty) &\iff \gamma_s^{-1}\gamma\gamma_s(\infty) = \infty \\ &\iff \gamma_s^{-1}\gamma\gamma_s \in \mathrm{SL}_2(\mathbb{Z})_\infty \\ &\iff \gamma \in \gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1}. \end{aligned}$$

Thus, we have $\Gamma_s = \Gamma \cap \gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1}$. This implies that Γ_s is a subgroup of $\gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1}$. We can furthermore define the natural map

$$\pi : \Gamma_s\backslash(\gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1}) \to \Gamma\backslash\mathrm{SL}_2(\mathbb{Z}) \tag{34}$$

where, for some $\delta \in \mathrm{SL}_2(\mathbb{Z})_\infty$, we define $\pi(\Gamma_s\gamma_s\delta\gamma_s^{-1}) = \Gamma\gamma_s\delta\gamma_s^{-1}$. Note that as $\Gamma_s \subseteq \Gamma$, we have that π is well-defined. Furthermore, we can prove π injective. Indeed, let $\delta_1, \delta_2 \in \mathrm{SL}_2(\mathbb{Z})_\infty$. Then

$$\begin{aligned} \pi(\Gamma_s\gamma_s\delta_1\gamma_s^{-1}) = \pi(\Gamma_s\gamma_s\delta_2\gamma_s^{-1}) &\implies \Gamma\gamma_s\delta_1\gamma_s^{-1} = \Gamma\gamma_s\delta_2\gamma_s^{-1} \\ &\implies \gamma_s\delta_1\delta_2^{-1}\gamma_s^{-1} \in \Gamma. \end{aligned}$$

Furthermore, as $\delta_1\delta_2^{-1} \in \mathrm{SL}_2(\mathbb{Z})_\infty$, we note

$$\gamma_s\delta_1\delta_2^{-1}\gamma_s^{-1} \in \gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1}. \tag{35}$$

Thus, we have

$$\gamma_s\delta_1\delta_2^{-1}\gamma_s^{-1} \in \Gamma \cap \gamma_s\mathrm{SL}_2(\mathbb{Z})_\infty\gamma_s^{-1} = \Gamma_s \implies \Gamma\gamma_s\delta_1\gamma_s^{-1} = \Gamma\gamma_s\delta_2\gamma_s^{-1}$$

thus proving π injective. As Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$, we therefore obtain that Γ_s has finite index in $\gamma_s \mathrm{SL}_2(\mathbb{Z})_\infty \gamma_s^{-1}$. We now conjugate by γ_s and define the subgroup

$$H_s = \gamma_s^{-1} \Gamma \gamma_s \cap \mathrm{SL}_2(\mathbb{Z})_\infty. \tag{36}$$

Thus, H_s is a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})_\infty$. Noting that

$$\mathrm{SL}_2(\mathbb{Z})_\infty = \{\pm I\} \times \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{Z} \right\} \cong \mathbb{Z}_2 \times \mathbb{Z} \tag{37}$$

we obtain three possible cases for H_s (as shown in the Appendix by calculation of the subgroups of $\mathbb{Z}_2 \times \mathbb{Z}$).

1. H_s is an infinite cyclic group generated by $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ where $h \geq 1$.
2. H_s is an infinite cyclic group generated by $\begin{pmatrix} -1 & h \\ 0 & -1 \end{pmatrix}$ where $h \geq 1$.
3. H_s is an infinite group isomorphic to $\mathbb{Z}_2 \times h\mathbb{Z}$, generated by $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ where $h \geq 1$.

We furthermore note that h is the index of $\{\pm I\}H_s$ in $\mathrm{SL}_2(\mathbb{Z})_\infty$. With this, we can now define the width at a cusp.

Definition 14: [2, p. 35] Let Γ be a finite-index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and let $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$. We define the **width** of the cusp $\mathfrak{c}$ as the index of $\{\pm I\}H_\mathfrak{c}$ in $\mathrm{SL}_2(\mathbb{Z})$ (i.e. the positive integer h as given above). We denote the width of $\mathfrak{c}$ (with respect to Γ) as $h_\Gamma(\mathfrak{c})$

We furthermore define the cusp $\mathfrak{c}$ as **irregular** if $H_\mathfrak{c}$ is an infinite cyclic group generated by $\begin{pmatrix} -1 & h_\Gamma(\mathfrak{c}) \\ 0 & -1 \end{pmatrix}$ (i.e. case 2 shown above), otherwise we define the cusp $\mathfrak{c}$ as **regular**.

We also remark that, if Γ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$, then we have $\Gamma = \gamma^{-1}\Gamma\gamma$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. This therefore implies that $H_\mathfrak{c}$ is the same for all cusps $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$, and thus the widths of all cusps are identical as well as all cusps either being all regular or all irregular.

To illustrate this, we consider the example of $\Gamma_1(4)$.

Example: Let $\Gamma = \Gamma_1(4)$. Note that $-I \notin \Gamma$. We can computationally verify that Γ has index 12 in $\mathrm{SL}_2(\mathbb{Z})$ and thus has index 6 in $\mathrm{PSL}_2(\mathbb{Z})$. A set of coset representatives R for Γ in $\mathrm{SL}_2(\mathbb{Z})$ can be given as

$$R = \{I, -I, S, -S, ST, -ST, ST^{-1}, -ST^{-1}, ST^2, -ST^2, ST^2S, -ST^2S\} \tag{38}$$

where a set of coset representatives Q for Γ in $\mathrm{PSL}_2(\mathbb{Z})$ can be given as

$$Q = \{\pm I, \pm S, \pm ST, \pm ST^{-1}, \pm ST^2, \pm ST^2 S\}. \tag{39}$$

Noting the surjectivity of the map ϕ given in equation (32), we can thus find all cusps by calculating $\gamma(\infty)$ for all $\gamma \in R$. Note

$$\pm I(\infty) = \infty$$
$$\pm S(\infty) = \pm ST(\infty) = \pm ST^{-1}(\infty) = \pm ST^2(\infty) = 0$$

$$\pm ST^2S(\infty) = -\frac{1}{2}.$$

Note that as $T \in \Gamma$, we have that $1/2 = T(-1/2) = TST^2S(\infty)$. Thus $1/2$ and $-1/2$ correspond to the same cusp under the group action induced by Γ. It is also easy to see that if $\gamma'(\infty) = 0$, then $\gamma' \notin \Gamma$. Similarly, ∞ and $1/2$ are not in the same orbit as well as 0 and $1/2$ not being in the same orbit.

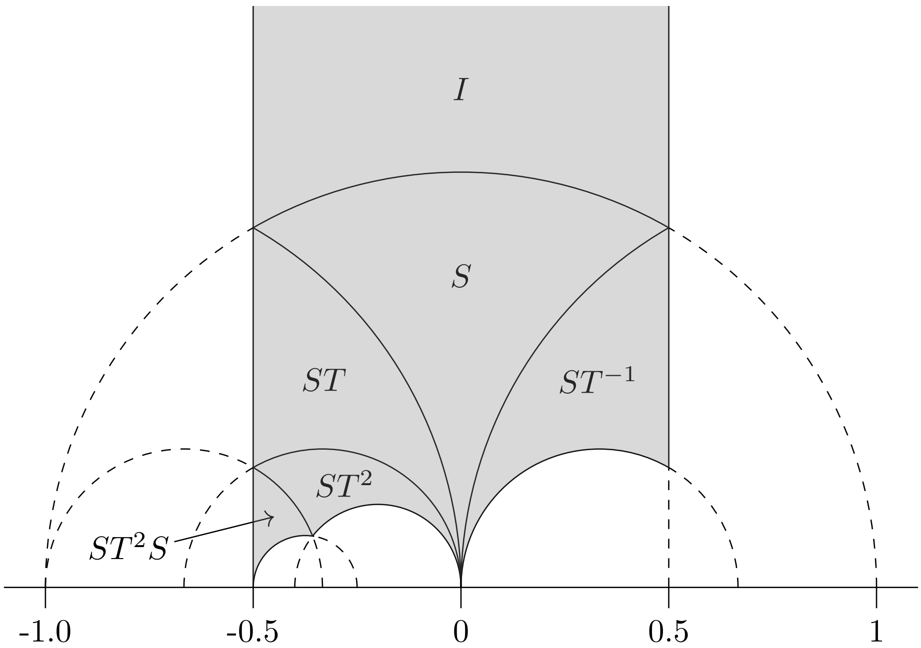

Figure 4: Fundamental Domain for $\Gamma_1(4)$. Note that there are six coset translates of I corresponding to the 6 cosets of $\Gamma_1(4)$ in $\mathrm{PSL}_2(\mathbb{Z})$. One can also visually notice the 3 cusps at ∞, 0 and $-1/2$ corresponding to the three sets of cosets $\{I\}, \{ST^{-1}, S, ST, ST^2\}$ and $\{ST^2S\}$ respectively.

We therefore have that $X(\Gamma_1(4))$ consists of exactly 3 cusps, namely

$$\mathrm{Cusps}(\Gamma_1(4)) = \{\Gamma(\infty), \Gamma(0), \Gamma(1/2)\}. \tag{40}$$

We now calculate the widths at each of the cusps and determine whether it is regular or irregular. For $c = \infty$, we have $I(\infty) = \infty$. Therefore

$$H_\infty = I^{-1}\Gamma I \cap \mathrm{SL}_2(\mathbb{Z})_\infty = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{Z} \right\} = \langle T \rangle. \tag{41}$$

Thus H_∞ is an infinite cyclic group generated by T, and therefore the cusp at ∞ is *regular* and has width $h_{\Gamma_0(4)}(\infty) = 1$. For $c = 0$. we have $S(\infty) = 0$. Thus

$$H_0 = S^{-1}\Gamma S \cap \mathrm{SL}_2(\mathbb{Z})_\infty. \tag{42}$$

To calculate H_0, we let $\gamma = \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})_\infty$. We note that

$$\begin{aligned} \gamma \in S^{-1}\Gamma S &\implies S\gamma S^{-1} \in \Gamma \\ &\implies \mp \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \Gamma \end{aligned}$$

$$\implies \pm\begin{pmatrix}1 & 0\\ -b & 1\end{pmatrix} \in \Gamma.$$

Note that the above condition holds if and only if b is a multiple of 4 with the positive sign chosen. Hence, we obtain that $H_0 = S^{-1}\Gamma S \cap \mathrm{SL}_2(\mathbb{Z})_\infty$ is an infinite cyclic group generated by $\begin{pmatrix}1 & 4\\ 0 & 1\end{pmatrix}$, therefore obtaining that 0 is a *regular* cusp with width $h_{\Gamma_0(4)}(0) = 4$.

We finally consider the cusp at 1/2. As before, we have that $ST^2S(\infty) = -1/2$. Thus

$$H_{1/2} = (ST^2S)^{-1}\Gamma(ST^2S) \cap \mathrm{SL}_2(\mathbb{Z}).$$

We let $\gamma = \pm\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})_\infty$. Note that

$$\begin{aligned}\gamma \in (ST^2S)^{-1}\Gamma(ST^2S) &\implies (ST^2S)\gamma(ST^2S)^{-1} \in \Gamma\\ &\implies \pm\begin{pmatrix}-1 & 0\\ 2 & -1\end{pmatrix}\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}\begin{pmatrix}-1 & 0\\ 2 & -1\end{pmatrix}^{-1} \in \Gamma\\ &\implies \pm\begin{pmatrix}-1 & 0\\ 2 & -1\end{pmatrix}\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}\begin{pmatrix}-1 & 0\\ -2 & -1\end{pmatrix} \in \Gamma\\ &\implies \pm\begin{pmatrix}1+2b & b\\ -4b & 1-2b\end{pmatrix} \in \Gamma\end{aligned}$$

From the definition of $\Gamma_1(4)$, b must therefore satisfy the following system of congruences

$$\pm(1+2b) \equiv 1 \pmod 4, \qquad \pm(-4b) \equiv 0 \pmod 4, \qquad \pm(1-2b) \equiv 1 \pmod 4.$$

Noting that the first and last condition are equivalent and that we always have $\mp 4b \equiv 0 \pmod 4$ for all $b \in \mathbb{Z}$, we thus have

$$\gamma \in (ST^2S)^{-1}\Gamma(ST^2S) \iff \pm(1+2b) \equiv 1 \pmod 4.$$

One can easily verify that this congruence is satisfied if and only if either b is even and the positive sign is chosen, or if b is odd and the negative sign is chosen. We therefore finally obtain

$$H_{1/2} = \left\{\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix} \mid b \in 2\mathbb{Z}\right\} \cup \left\{\begin{pmatrix}-1 & b\\ 0 & -1\end{pmatrix} \mid b \in 2\mathbb{Z}+1\right\} = \left\langle -\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\right\rangle.$$

Thus, $H_{1/2}$ is an infinite cyclic group generated by $-\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$. Therefore, the cusp at 1/2 is *irregular* with width $h_{\Gamma_0(4)}(\frac{1}{2}) = 1$.

It is worth noting that irregular cusps occur rather infrequently. Indeed, the cusp at 1/2 for $\Gamma_1(4)$ is the *only* irregular cusp that occurs out of all the congruence subgroups $\Gamma(N), \Gamma_0(N)$ and $\Gamma_1(N)$ [6, p. 103].

Eisenstein series revisited

We defined the Eisenstein series $G_k(\tau)$ and its normalisation $E_k(\tau)$, explicitly giving a non-trivial example of a modular form of even weight $k \geq 4$ for $\mathrm{SL}_2(\mathbb{Z})$. We now generalise this construction to other subgroups Γ of $\mathrm{SL}_2(\mathbb{Z})$.

Let $\gamma = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ We define the **factor of automorphy** of γ as a function $j(\gamma,\tau) \in \mathbb{C}$ where

$$j(\gamma,\tau) = (c\tau + d) \quad \text{for all } \tau \in \mathcal{H}.$$

We note that the factor of automorphy $j(\gamma, \tau)$ satisfies the following property.

Lemma 15: For any $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$, we have

$$j(\gamma_1, \gamma_2\tau) = j(\gamma_2, \tau)^{-1} j(\gamma_1\gamma_2, \tau)$$

Proof: Let $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. We thus have

$$\begin{aligned} j(\gamma_1, \gamma_2\tau) = c_1 \left(\frac{a_2\tau + b_2}{c_2\tau + d_2} \right) + d_2 &= (c_2\tau + d_2)^{-1}(c_1(a_2\tau + b_2) + d_1(c_2\tau + d_2)) \\ &= (c_2\tau + d_2)^{-1}((a_2c_1 + d_1c_2)\tau + (b_2c_1 + d_1d_2)) \\ &= j(\gamma_2, \tau)^{-1} j(\gamma_1\gamma_2, \tau) \end{aligned}$$

which proves the lemma. □

We also define P_∞ as the subgroup generated by T

$$P_\infty = \langle T \rangle = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} : b \in \mathbb{Z} \right\}.$$

We now note the following expression for $G_k(\tau)$. Letting $(m, n) = (gc, gd)$ where $g = \gcd(m, n)$ and c and d are coprime, we obtain

$$\begin{aligned} G_k(\tau) = \sum_{(m,n)\in Z'} \frac{1}{(m\tau + n)^k} &= \sum_{g=1}^{\infty} \sum_{\substack{(m,n)\in Z' \\ \gcd(m,n)=g}} \frac{1}{(m\tau + n)^k} \\ &= \sum_{g=1}^{\infty} \sum_{\substack{(c,d)\in Z' \\ \gcd(c,d)=1}} \frac{1}{g^k(c\tau + d)^k} \\ &= \sum_{g=1}^{\infty} \frac{1}{g^k} \sum_{\substack{(c,d)\in Z' \\ \gcd(c,d)=1}} \frac{1}{(m\tau + n)^k} \\ &= \zeta(k) \sum_{\substack{(c,d)\in Z' \\ \gcd(c,d)=1}} \frac{1}{(m\tau + n)^k} \end{aligned}$$

We can rewrite the final sum by noting that the function $\phi : P_\infty \backslash \mathrm{SL}_2(\mathbb{Z}) \to \{(c, d) \in \mathbb{Z}^2 : \gcd(c, d) = 1\}$ given by

$$\phi(P_\infty\gamma) = (c, d) \qquad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

induces a bijection between the two sets. Indeed, well-definedness can be verified by noting that

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + kc & b + kd \\ c & d \end{pmatrix} \qquad \text{for all } k \in \mathbb{Z}. \tag{43}$$

To prove injectivity, let $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Assuming $\phi(P_\infty\gamma_1) = \phi(P_\infty\gamma_2)$, this implies $c_1 = c_2$ and $d_1 = d_2$. Defining $c := c_1 = c_2$ and $d := d_1 = d_2$, we obtain

$$\gamma_1\gamma_2^{-1} = \begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} \begin{pmatrix} d & -b_2 \\ -c & a_2 \end{pmatrix} = \begin{pmatrix} 1 & a_2b_1 - a_1b_2 \\ 0 & 1 \end{pmatrix} \in P_\infty$$

which implies $P_\infty\gamma_1 = P_\infty\gamma_2$, thus proving injectivity. Finally, surjectivity is obtained from Bézout's identity. We therefore obtain:

$$G_k(\tau) = \zeta(k) \sum_{\substack{(c,d)\in Z' \\ \gcd(c,d)=1}} \frac{1}{(m\tau+n)^k} = \zeta(k) \sum_{\gamma\in P_\infty\backslash \mathrm{SL}_2(\mathbb{Z})} j(\gamma,\tau)^{-k}$$

where the final sum is well-defined since $j(h\gamma,\tau) = j(\gamma,\tau)$ for all $h \in P_\infty$, as shown in (43).

This therefore motivates defining an Eisenstein series for any congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$.

Definition 16: [17, p. 60] Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We define the **normalised Eisenstein series** of weight k attached to Γ and to the cusp ∞ as

$$E_{k,\Gamma,\infty}(\tau) = \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma,\tau)^{-k}. \tag{44}$$

We prove that $E_{k,\Gamma,\infty}$ is a modular form. Indeed, note that the natural map $(P_\infty\cap\Gamma)\backslash\Gamma \to P_\infty\backslash\mathrm{SL}_2(\mathbb{Z})$ is injective [17, p.60] and thus the sum converges by convergence of $E_k(\tau)$. Thus, $E_{k,\Gamma,\infty}$ is holomorphic on $\mathcal{H}$.

To prove the modularity condition, let $\gamma_0 \in \Gamma$. Thus [15, p. 53]

$$E_{k,\Gamma,\infty}(\gamma_0\tau) = \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma,\gamma_0\tau)^{-k} = j(\gamma_0,\tau)^k \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma\gamma_0,\tau)^{-k}$$

Now, since $\gamma_0 \in \Gamma$, we have that multiplication by γ_0 simply permutes the set of cosets. Thus, the set $\{\gamma\gamma_0 : \gamma \in (P_\infty\cap\Gamma)\backslash\Gamma\}$ forms a set of coset representatives for $(P_\infty\cap\Gamma)$ in Γ. Therefore

$$E_{k,\Gamma,\infty}(\gamma_0\tau) = j(\gamma_0,\tau)^k \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma,\tau)^{-k} = j(\gamma_0,\tau)^k E_{k,\Gamma,\infty}(\tau) \tag{45}$$

We finally note that $E_{k,\Gamma,\infty}$ is bounded as $\mathrm{Im}(\tau)\to\infty$, as the sum $E_{k,\Gamma,\infty}$ is simply a sub-series of E_k. Thus, we have $E_{k,\Gamma,\infty} \in \mathcal{M}_k(\Gamma)$.

We now investigate the behaviour of $E_{k,\Gamma,\infty}$ at each of the cusps of Γ. We first consider the behaviour as $\mathrm{Im}(\tau)\to\infty$. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. We note [17, p. 60]

$$\lim_{\mathrm{Im}(\tau)\to\infty} j(\gamma,\tau)^{-k} = \lim_{\mathrm{Im}(\tau)\to\infty} (c\tau+d)^{-k} = \begin{cases} d^{-k} & \text{if } c = 0 \\ 0 & \text{if } c \neq 0 \end{cases} \tag{46}$$

Now, as shown in (31), we have that $c = 0 \iff \gamma \in \{\pm I\}P_\infty$. Thus, we can evaluate

$$\begin{aligned}
\lim_{\mathrm{Im}(\tau)\to\infty} E_{k,\Gamma,\infty}(\tau) &= \lim_{\mathrm{Im}(\tau)\to\infty} \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma,\tau)^{-k} \\
&= \lim_{\mathrm{Im}(\tau)\to\infty} \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash(\{\pm I\}P_\infty\cap\Gamma)} j(\gamma,\tau)^{-k} \\
&= \begin{cases} 1 & \text{if } P_\infty\cap\Gamma = \{\pm I\}P_\infty\cap\Gamma \\ 2 & \text{if } P_\infty\cap\Gamma \subsetneq \{\pm I\}P_\infty\cap\Gamma \text{ and } k \text{ even} \\ 0 & \text{if } P_\infty\cap\Gamma \subsetneq \{\pm I\}P_\infty\cap\Gamma \text{ and } k \text{ odd} \end{cases}
\end{aligned}$$

We therefore obtain that, if k is even, then $E_{k,\Gamma,\infty}$ is nonzero at the cusp ∞. Let us now consider a cusp $\mathfrak{c}$ different from ∞, and let $\gamma_c \in \mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma\gamma_c(\infty) = \mathfrak{c}$. We note

$$\begin{aligned}\lim_{\mathrm{Im}(\tau)\to\infty} E_{k,\Gamma,\infty}|_k\gamma_c(\tau) &= \lim_{\mathrm{Im}(\tau)\to\infty} j(\gamma_c,\tau)^{-k} \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma,\gamma_c\tau)^{-k} \\ &= \lim_{\mathrm{Im}(\tau)\to\infty} \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma\gamma_c,\tau)^{-k}\end{aligned}$$

Now, as $\mathfrak{c}$ is a cusp different from ∞, we have, for any $\gamma \in \Gamma$

$$\gamma(\mathfrak{c}) \neq \infty \implies \gamma\gamma_s(\infty) \neq \infty \implies \gamma\gamma_s \notin \{\pm I\}P_\infty \implies \lim_{\mathrm{Im}(\tau)\to\infty} j(\gamma,\tau)^{-k} = 0$$

where the last equality follows by (46). Therefore, we have

$$\lim_{\mathrm{Im}(\tau)\to\infty} E_{k,\Gamma,\infty}|_k\gamma_c(\tau) = \lim_{\mathrm{Im}(\tau)\to\infty} \sum_{\gamma\in(P_\infty\cap\Gamma)\backslash\Gamma} j(\gamma\gamma_c,\tau)^{-k} = 0 \tag{47}$$

We thus have that, for k even, $E_{k,\Gamma,\infty}$ is non-zero at the cusp ∞ and zero at all other cusps $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$, $\mathfrak{c} \neq \infty$.

We can further generalise the construction given in Definition 16 to construct Eisenstein series that are non-vanishing at a single given cusp $\mathfrak{c}_0$ and vanishing at all other cusps. Indeed, this can be done by simply translating $E_{k,\Gamma,\infty}$ by an element γ_c where $\Gamma\gamma_c(\infty) = \mathfrak{c}_0$.

Definition 17: [17, p. 61] Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$, let $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$ and let $\gamma_c \in \mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma\gamma_c(\infty) = \mathfrak{c}$. We define the normalised Eisenstein series of weight k attached to Γ and to the cusp $\mathfrak{c}$ as

$$E_{k,\Gamma,\mathfrak{c}} := E_{k,\gamma_c^{-1}\Gamma\gamma_c,\infty}|_k\gamma_c^{-1} = \sum_{\gamma\in((\gamma_c P_\infty \gamma_c^{-1})\cap\Gamma)\backslash\Gamma} j(\gamma_c^{-1}\gamma,\tau)^{-k} \tag{48}$$

We therefore have, for any finite-index subgroup Γ, an explicit construction for an Eisenstein series attached to each cusp of Γ. Since we also have that $E_{k,\Gamma,\mathfrak{c}}$ iz non-zero at the cusp $\mathfrak{c}$ and zero at all other cusps, we note that the set of modular forms

$$\{E_{k,\Gamma,\mathfrak{c}} \ : \ \mathfrak{c} \in \mathrm{Cusps}(\Gamma)\}$$

is thus linearly independent in Γ for even weight k. This fact will be fundamental in establishing a lower bound on the dimension of $\mathcal{M}_k(\Gamma)$.

We conclude by expressing $E_{k,\Gamma,\mathfrak{c}}$ more concretely for the case $\Gamma = \Gamma_0(N)$. We thus define the Eisenstein series of weight k and level N as [15, p. 52]

$$E_{k,N}(\tau) := E_{k,\Gamma_0(N),\infty} = \sum_{\gamma\in P_\infty\backslash\Gamma_0(N)} j(\gamma,\tau)^{-k} = \frac{1}{2} \sum_{\substack{(c,d)\in\mathbb{Z}^2 \\ \gcd(c,d)=1 \\ c\equiv 0 \bmod N}} \frac{1}{(c\tau+d)^k} \tag{49}$$

For completeness, we can express the q-expansion of $E_{k,N}$ as follows [15, p. 61]

$$\tilde{E}_{k,N}(q) = 1 - \frac{2k\phi(N)}{N^k B_k} \prod_{p|N} \left(1 - \frac{1}{p^k}\right)^{-1} \sum_{n=1}^{\infty} \sigma_{k-1,N}(n) q^n$$

where

$$\sigma_{k,N}(n) = \sum_{d|N} \frac{\mu(N/\gcd(d,N))}{\phi(N/\gcd(d,N))} d^k.$$

Valence formula

We first recall some notions from complex analysis:

Definition 18: [4, p. 105] Let f be a complex-valued function which is meromorphic in some neighbourhood of a point $z_0 \in \mathbb{C}$. Then there exists a *unique* integer n such that the function $g(z)$ defined as

$$g(x) = \frac{f(z)}{(z - z_0)^n}$$

is holomorphic and non-zero in a neighbourhood around z_0. We define the **order** at the point z_0 as the integer n, denoted as $\mathrm{ord}_{z_0}(f) = n$.

Moreover, it can easily be seen that, if f_1 and f_2 and both complex-valued functions meromorphic in some neighbourhood of z_0, then

$$\mathrm{ord}_{z_0}(f_1 f_2) = \mathrm{ord}_{z_0}(f_1) + \mathrm{ord}_{z_0}(f_2).$$

Theorem 19: [4, p. 123] (Cauchy's Argument Principle) Let f be a meromorphic function on the complex plane. Let $D \subseteq \mathbb{C}$ be a non-zero and analytic region in the domain of f and suppose C is a positively oriented piece-wise smooth simple closed curve in D. Let E be the region enclosed by $\mathcal{C}$. Then

$$\frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f'(z)}{f(z)} dz = \sum_{z \in E} \mathrm{ord}_z(f).$$

Specifically, if z_0 is some point in the domain of f, we can define $\mathcal{C}_{n,z_0}$ as a positively oriented circle contour centred at z_0 with radius $\frac{1}{n}$. Then

$$\lim_{n \to \infty} \frac{1}{2\pi i} \int_{\mathcal{C}_{n,z_0}} \frac{f'(z)}{f(z)} dz = \mathrm{ord}_{z_0}(f). \tag{50}$$

One can also prove more generally [12, p. 116], if $\mathcal{C}_{n,z_0}$ is an arc of radius $\frac{1}{n}$ centred at z_0 with an angle of $\Delta\theta$, then

$$\lim_{n \to \infty} \frac{1}{2\pi i} \int_{\mathcal{C}_{n,z_0}} \frac{f'(z)}{f(z)} dz = \mathrm{ord}_{z_0}(f) \frac{\Delta\theta}{2\pi}. \tag{51}$$

We now note some further properties of orders for modular forms. Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We can furthermore define the order of some orbit Γz in $\Gamma \backslash \mathcal{H}$ as:

$$\mathrm{ord}_{\Gamma z}(f) = \mathrm{ord}_z(f) \tag{52}$$

One can easily verify that this is well-defined. Indeed, if $z_1, z_2 \in \mathcal{H}$ were in the same orbit, then there exists some $\gamma \in \Gamma$ where $\gamma(z_1) = z_2$. Thus

$$\mathrm{ord}_{z_1}(f) = \mathrm{ord}_{z_1}((cz+d)^k f) = \mathrm{ord}_{z_1}(f\gamma) = \mathrm{ord}_{z_2}(f) \tag{53}$$

noting that $(cz+d)^k$ is non-zero analytic on $\mathcal{H}$, and thus does not change the order. Therefore, the order of f at some point only depends on the orbit.

We finally define the order at a cusp. Let $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$ and let $\gamma_c \in \mathrm{SL}_2(\mathbb{Z})$ be such that $\gamma_c(\infty) = \mathfrak{c}$. Note that $f|_k \gamma_c(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$. Thus, similar to the q-expansion we did for f, we can do a q-expansion for $f|_k \gamma_c$.

Let $\tilde{h}_\Gamma(\mathfrak{c})$ be the least positive integer such that $\begin{pmatrix} 1 & \tilde{h}_\Gamma(\mathfrak{c}) \\ 0 & 1 \end{pmatrix} \in \gamma_c^{-1}\Gamma\gamma_c$. Noting that the width of a cusp is well-defined, we have the $\tilde{h}_\Gamma(\mathfrak{c})$ is well-defined and furthermore that

$$\tilde{h}_\Gamma(\mathfrak{c}) = \begin{cases} h_\Gamma(\mathfrak{c}) & \text{if } \mathfrak{c} \text{ is regular} \\ 2h_\Gamma(\mathfrak{c}) & \text{if } \mathfrak{c} \text{ is irregular} \end{cases}$$

by definition of $h_\Gamma(\mathfrak{c})$. Therefore

$$\begin{pmatrix} 1 & \tilde{h}_\Gamma(\mathfrak{c}) \\ 0 & 1 \end{pmatrix} \in \gamma_c^{-1}\Gamma\gamma_c \quad \Longrightarrow \quad \gamma_c \begin{pmatrix} 1 & \tilde{h}_\Gamma(\mathfrak{c}) \\ 0 & 1 \end{pmatrix} \in \Gamma\gamma_c \quad \Longrightarrow \quad \Gamma\gamma_c \begin{pmatrix} 1 & \tilde{h}_\Gamma(\mathfrak{c}) \\ 0 & 1 \end{pmatrix} = \Gamma\gamma_c.$$

We therefore obtain that

$$f|_k\gamma_c = f|_k\gamma_c \begin{pmatrix} 1 & \tilde{h}_\Gamma(\mathfrak{c}) \\ 0 & 1 \end{pmatrix}$$

and thus $f|_k\gamma_c(z) = f|_k\gamma_c(z+\tilde{h}_\Gamma(\mathfrak{c}))$. Thus, $f|_k\gamma_c$ is periodic with period $\tilde{h}_\Gamma(\mathfrak{c})$, which therefore implies we can do a q-expansion

$$q_\mathfrak{c} = \exp\left(2\pi i z/\tilde{h}_\Gamma(\mathfrak{c})\right)$$

where we define $\tilde{f}_c(q_c) = f|_k\gamma_c(z)$. As before, we note that this change of variables sends the upper half plane $\mathcal{H}$ to the punctured open unit disc D'. However, as $f|_k\gamma_c(z)$ is bounded as $\mathrm{Im}(z) \to \infty$, this implies $\tilde{f}_c(q_c)$ is bounded as $q \to 0$. Thus $\lim_{q\to 0} \tilde{f}_c(q_c)$ exists and we therefore have a Fourier expansion

$$f|_k\gamma_c = \tilde{f}_c(q_\mathfrak{c}) = \sum_{n=0}^{\infty} a_{\mathfrak{c},n}(f) q_\mathfrak{c}^n.$$

We can now define the **order** of f at the cusp $\mathfrak{c}$ as the least positive integer n such that $a_{\mathfrak{c},n}(f) \neq 0$. We denote this as $\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)$.

Definition 20: [6, p. 17] Let $k \in \mathbb{Z}$. A **cusp form** of weight k for Γ is a modular form f of weight k for Γ such that $a_{\mathfrak{c},0}(f) = 0$ for all $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$ (i.e. the leading coefficient in the Fourier expansion of f at each cusp is 0).

If $a_{\mathfrak{c},0}(f) = 0$ for some particular cusp $\mathfrak{c}$, then we often say that f vanishes at the cusp $\mathfrak{c}$. We denote the space of all cusp forms of weight k for Γ as $\mathcal{S}_k(\Gamma)$. Noting that $\mathcal{S}_k(\Gamma)$ is closed under $\mathbb{C}$-linear combinations, we obtain that $\mathcal{S}_k(\Gamma)$ is a subspace of $\mathcal{M}_k(\Gamma)$.

We now prove an important property about the orders at the cusps which shall be used later when proving the general valence formula.

Lemma 21: Let Γ and Γ' be two finite index subgroups of $\mathrm{SL}_2(\mathbb{Z})$ such that $\Gamma' \subseteq \Gamma$. Let f be a modular form of weight k for Γ. Let $\mathfrak{c}' \in \mathrm{Cusps}(\Gamma')$ and define $\mathfrak{c}$ as the cusp in Γ which is the image of $\mathfrak{c}'$ under the natural map $\pi : \mathrm{Cusps}(\Gamma') \to \mathrm{Cusps}(\Gamma)$ where $\Gamma' c \mapsto \Gamma c$ for all $c \in \mathbb{P}^1(\mathbb{Q})$. Then we have

$$\frac{\mathrm{ord}_{\Gamma',\mathfrak{c}'}(f)}{\tilde{h}_{\Gamma'}(\mathfrak{c}')} = \frac{\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)}{\tilde{h}_\Gamma(\mathfrak{c})}$$

Proof: We first note that, since $\Gamma' \subseteq \Gamma$, we have f is also a modular form of weight k for Γ'. This also implies that π is well-defined. Let $\gamma_{c'} \in \mathrm{SL}_2(\mathbb{Z})$ be such that $\Gamma'\gamma_{c'}(\infty) = \mathfrak{c}'$. By definition of $\mathfrak{c}$, we thus also have $\Gamma\gamma_{c'}(\infty) = \mathfrak{c}$. Thus

$$\begin{pmatrix} 1 & h_{\Gamma'}(\mathfrak{c}') \\ 0 & 1 \end{pmatrix} \in \{\pm I\}\gamma_c^{-1}\Gamma'\gamma_c \subseteq \{\pm I\}\gamma_c^{-1}\Gamma\gamma_c$$

As $\pm\gamma_c^{-1}\Gamma\gamma_c$ is generated by $\begin{pmatrix} 1 & h_\Gamma(\mathfrak{c}) \\ 0 & 1 \end{pmatrix}$, we therefore have that $h_\Gamma(\mathfrak{c})$ divides $h'_\Gamma(\mathfrak{c})$. Likewise, we also obtain that $\tilde{h}_\Gamma(\mathfrak{c})$ divides $\tilde{h}'_\Gamma(\mathfrak{c}')$.

Now, let $m = \tilde{h}_{\Gamma'}(\mathfrak{c}')/\tilde{h}_\Gamma(\mathfrak{c})$. Note that $f|_k\gamma_c$ is periodic with period $\tilde{h}_\Gamma(\mathfrak{c})$, thus doing a q-expansion of $f|_k\gamma_c$ for Γ, we note that

$$f|_k\gamma_c(q) = \sum_{n=0}^{\infty} a_{c,n}(f)q^n \quad \text{where } q = \exp\left(\frac{2\pi i z n}{\tilde{h}_\Gamma(\mathfrak{c})}\right)$$

Now, let $N = \operatorname{ord}_{\Gamma,c}(f)$. By definition of the order, We can therefore rewrite the q-expansion as

$$\begin{aligned}\tilde{f}|_k\gamma_c(q) = \sum_{n=0}^{\infty} a_{\mathfrak{c},n}(f)q = \sum_{n=N}^{\infty} a_{c,n}(f)q &= \sum_{n=N}^{\infty} a_{c,n}(f)\exp\left(\frac{2\pi i z n}{\tilde{h}_\Gamma(\mathfrak{c})}\right) \\ &= \sum_{n=N}^{\infty} a_{c,n}(f)\exp\left(\frac{2\pi i z n m}{\tilde{h}_{\Gamma'}(\mathfrak{c}')}\right) \\ &= \sum_{n=mN}^{\infty} a'_{c',n}(f)\exp\left(\frac{2\pi i z n}{\tilde{h}_{\Gamma'}(\mathfrak{c}')}\right)\end{aligned}$$

where the last equality follows by doing a q-expansion for Γ', where $a'_{c,n}(f) = 0$ if m does not divide n.

Noting that $a'_{c',mN}(f) = a_{c,n} \neq 0$, we therefore have that the order of f at the cusp c' for Γ' is

$$\operatorname{ord}_{\Gamma',c'}(f) = mN = \frac{h'_\Gamma}{h_\gamma}\operatorname{ord}_{\Gamma,c}(f)$$

which therefore implies

$$\frac{\operatorname{ord}_{\Gamma',c'}(f)}{\tilde{h}_{\Gamma'}(c')} = \frac{\operatorname{ord}_{\Gamma,c}(f)}{\tilde{h}_\Gamma(c)}.$$

This concludes the proof. □

Valence formula for $\mathbf{SL}_2(\mathbb{Z})$

By integrating around the fundamental domain $\mathcal{F}$ of $\mathrm{SL}_2(\mathbb{Z})$, we can derive a useful valence formula which will be used in a later section to derive bounds for the dimension of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

Theorem 22: [25, p. 30] Let f be a non-zero modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$. Let $\omega = e^{2\pi i/3}$ be the cubed root of unity in $\mathcal{H}$. Then

$$\operatorname{ord}_\infty(f) + \frac{1}{2}\operatorname{ord}_i(f) + \frac{1}{3}\operatorname{ord}_\omega(f) + \sum_{\substack{z\in \mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H} \\ z\neq\omega,i}} \operatorname{ord}_z(f) = \frac{k}{12} \tag{54}$$

where $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ denotes the set of all orbits induced by the group action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$.

Proof: [23, p. 85] Let f be a non-zero modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$. Note that, as f is holomorphic at infinity, there exists $R > 0$ such that for all $\mathrm{Im}(\tau) > R$ the value of $f(\tau)$ is nonzero.

For each $n \in \mathbb{N}$, we define a contour integral $\mathcal{C}_n$ around the fundamental domain $\mathcal{F}$ as shown in Figure 5. The imaginary part of the line segment E_nA_n is $R+n$ and the radius for each arc B_nB_n', C_nC_n' and D_nD_n' is $\frac{\epsilon}{n}$ where $\epsilon > 0$ is chosen sufficiently small so that no poles or zeros are contained within each circle centred at ω, i, and ω' with radius ϵ (where $w' = Sw = -\bar{\omega}$). As all poles/zeros are isolated, this can always be done.

In general, there may exist zeros or poles along with boundary of $\mathcal{F}$. For each zero/pole, we simply shift the contour around each zero/pole, such that it only occurs once in the interior of $\mathcal{C}_n$. For example, as shown in Figure 5, the points λ_1, λ_2 and λ_3 are all potential zeros/poles along the boundary of $\mathcal{F}$.

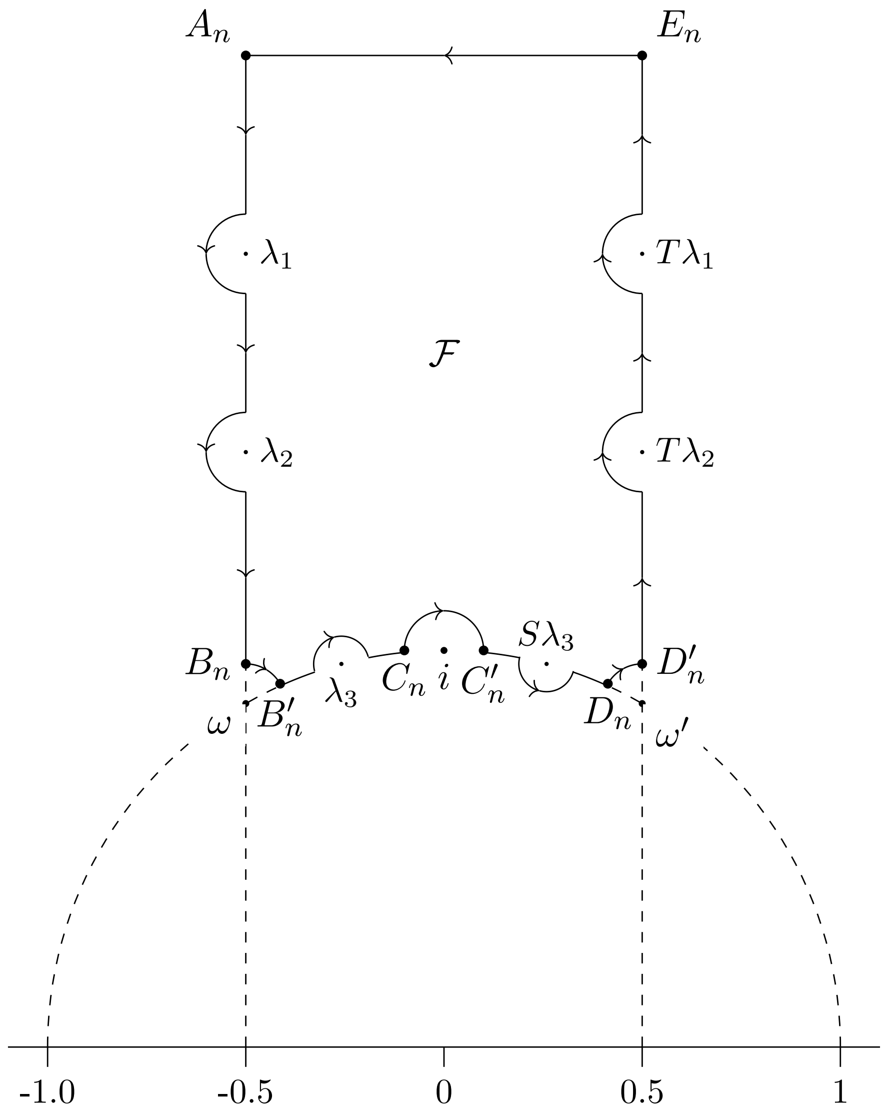

Figure 5: Contour integral of $\mathcal{C}_n$ around the fundamental domain $\mathcal{F}$ for $\mathrm{SL}_2(\mathbb{Z})$.

We now apply the argument principle to the region enclosed by the contour integral $\mathcal{C}_n$. Let $\mathcal{G}$ denote the region enclosed by $\mathcal{C}_n$. This yields

$$\frac{1}{2\pi i} \oint_{\mathcal{C}_n} \frac{f'(z)}{f(z)} dz = \sum_{z \in \mathrm{int}(\mathcal{G})} \mathrm{ord}_z(f).$$

We note that every orbit in $\mathrm{SL}_2(\mathbb{Z}) \backslash \mathcal{H}$ has exactly one representative in $\mathcal{F}$, with the exception of the boundary. However, as $\mathcal{C}_n$ is shifted around any zeros/poles along the boundary such that they occur exactly once in $\mathcal{G}$ (with the exception of i and ω), we obtain

$$\frac{1}{2\pi i}\oint_{\mathcal{C}_n}\frac{f'(z)}{f(z)}dz = \sum_{\substack{z\in\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}\\ z\neq\omega,i}} \mathrm{ord}_z(f). \tag{55}$$

We now calculate the integral over each piece-wise smooth segment of $\mathcal{C}_n$ separately. For A_nB_n and D'_nE_n, we note that $f(z)=f(z+1)$ and doing a change of variables $z\mapsto z+1$, we obtain

$$\int_{A_n}^{B_n}\frac{f'(z)}{f(z)}dz = \int_{A_n}^{B_n}\frac{f'(z+1)}{f(z+1)}dz = \int_{E_n}^{D'_n}\frac{f'(z)}{f(z)}dz = -\int_{D'_n}^{E_n}\frac{f'(z)}{f(z)}dz.$$

Thus, we have the cancellation

$$\int_{A_n}^{B_n}\frac{f'(z)}{f(z)}dz + \int_{D'_n}^{E_n}\frac{f'(z)}{f(z)}dz = 0 \tag{56}$$

To evaluate the integral along E_nA_n, we apply the change of variables $q=e^{2\pi iz}$ thus $dq = 2\pi ie^{2\pi iz}dz = 2\pi iqdz$. We also note

$$\frac{f'(z)}{f(z)} = \frac{\frac{d}{dq}\tilde{f}(q)\frac{dq}{dz}}{\tilde{f}(q)} = \frac{\tilde{f}'(q)}{\tilde{f}(q)}2\pi iq$$

Thus, denoting $\tilde{A}_n$ and $\tilde{B}_n$ as the points $\exp(2\pi iA_n)$ and $\exp(2\pi iE_n)$ respectively, we have

$$\int_{E_n}^{A_n}\frac{f'(z)}{f(z)}dz = \int_{\tilde{E}_n}^{\tilde{A}_n}\frac{\tilde{f}'(q)}{\tilde{f}(q)}2\pi iq\frac{dq}{2\pi iq} = \int_{\tilde{E}_n}^{\tilde{A}_n}\frac{\tilde{f}'(q)}{\tilde{f}(q)}dq$$

Hence, as $n\to\infty$, the integral approaches a small circle (with radius $e^{2\pi(R+n)}$) around $q=0$ in the **negative** orientation. Thus

$$\lim_{n\to\infty}\frac{1}{2\pi i}\int_{E_n}^{A_n}\frac{f'(z)}{f(z)}dz = -\mathrm{ord}_0(\tilde{f}) = -\mathrm{ord}_\infty(f). \tag{57}$$

To evaluate the three arcs $B_nB'_n$, $C_nC'_n$ and $D_nD'_n$, we apply equation (51). Note that the arc $B_nB'_n$ is centred at ω with negative orientation, and as $n\to\infty$, the local angle $\Delta\theta_B$ approaches $\frac{\pi}{6}$. Thus

$$\lim_{n\to\infty}\frac{1}{2\pi i}\int_{B_n}^{B'_n}\frac{f'(z)}{f(z)}dz = \frac{\mathrm{ord}_\omega(f)}{6}. \tag{58}$$

Similarly, noting that the arc $C_nC'_n$ approaches a half-circle as $n\to\infty$ and that the local angle $\Delta\theta_D$ approaches $\frac{\pi}{6}$, we also obtain (using equation (51))

$$\lim_{n\to\infty}\frac{1}{2\pi i}\int_{C_n}^{C'_n}\frac{f'(z)}{f(z)}dz = \frac{\mathrm{ord}_i(f)}{2} \qquad\text{and}\qquad \lim_{n\to\infty}\frac{1}{2\pi i}\int_{D_n}^{D'_n}\frac{f'(z)}{f(z)}dz = \frac{\mathrm{ord}_\omega(f)}{6}. \tag{59}$$

All that now remains is to the evaluate the integral from B'_n to C_n and from C'_n to D_n. For this, we note that $f(z)=z^{-k}f(-1/z)$. Using this transformation with the product rule for differentiation, we therefore obtain

$$\frac{f'(z)}{f(z)} = \frac{-kz^{-k-1}f(-1/z)+z^{-k}f'(-1/z)z^{-2}}{z^{-k}f(-1/z)} = -\frac{k}{z}+\frac{1}{z^2}\frac{f'(-1/z)}{f(-1/z)}. \tag{60}$$

Changing variables $u = -1/z$, we obtain $du = z^{-2}dz$, and thus

$$\begin{aligned}\int_{B'_n}^{C_n} \frac{f'(z)}{f(z)}dz &= \int_{B'_n}^{C_n} -\frac{k}{z} + \frac{1}{z^2}\frac{f'(-1/z)}{f(-1/z)}dz \\ &= -k\int_{B'_n}^{C_n} \frac{1}{z}dz + \int_{B'_n}^{C_n} \frac{1}{z^2}\frac{f'(-1/z)}{f(-1/z)}dz \\ &= -k\int_{B'_n}^{C_n} \frac{1}{z}dz + \int_{D_n}^{C'_n} \frac{f'(u)}{f(u)}du.\end{aligned}$$

To evaluate the first integral above, we note that $B'_n \to \omega$ and $C_n \to i$ as $n \to \infty$. Thus

$$\lim_{n\to\infty} \int_{B'_n}^{C_n} \frac{1}{z}dz = \int_{\omega}^{i} \frac{1}{z}dz$$

along an arc in the negatively oriented direction. Doing a change of variables $z = \exp i\theta$ where $dz = i\exp i\theta d\theta$ and noting that $\omega = \exp i(2\pi/3)$ and $i = \exp i(\pi/2)$, we have

$$\int_{\omega}^{i} \frac{1}{z}dz = \int_{2\pi/3}^{\pi/2} id\theta = i\left(\frac{\pi}{2} - \frac{2\pi}{3}\right) = -\frac{i\pi}{6}$$

We therefore obtain:

$$\begin{aligned}\lim_{n\to\infty} \int_{B'_n}^{C_n} \frac{f'(z)}{f(z)}dz + \lim_{n\to\infty} \int_{C'_n}^{D_n} \frac{f'(z)}{f(z)}dz &= -k\lim_{n\to\infty} \int_{B'_n}^{C_n} \frac{1}{z}dz \\ &= -k\left(-\frac{i\pi}{6}\right) = \frac{i\pi k}{6} \qquad (61)\end{aligned}$$

Putting all the separate contributions from (56) through to (59) and (61) together, we finally obtain

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{2\pi i}\oint_{\mathcal{C}_n} \frac{f'(z)}{f(z)}dz &= \lim_{n\to\infty} \frac{1}{2\pi i}\Bigg(\int_{A_n}^{B_n} \frac{f'(z)}{f(z)}dz + \int_{B_n}^{B'_n} \frac{f'(z)}{f(z)}dz + \int_{B'_n}^{C_n} \frac{f'(z)}{f(z)}dz + \int_{C}^{C'_n} \frac{f'(z)}{f(z)}dz \\ &\quad + \int_{C'_n}^{D_n} \frac{f'(z)}{f(z)}dz + \int_{D_n}^{D'_n} \frac{f'(z)}{f(z)}dz + \int_{D'_n}^{E_n} \frac{f'(z)}{f(z)}dz + \int_{E_n}^{A_n} \frac{f'(z)}{f(z)}dz\Bigg) \\ &= \frac{1}{2\pi i}\left(0 + \frac{i\pi k}{6}\right) - \mathrm{ord}_\infty(f) - \frac{\mathrm{ord}_\omega(f)}{6} - \frac{\mathrm{ord}_i(f)}{2} - \frac{\mathrm{ord}_\omega(f)}{6} \\ &= \frac{k}{12} - \frac{1}{2}\mathrm{ord}_\omega(f) - \frac{1}{3}\mathrm{ord}_i(f) - \mathrm{ord}_\infty(f) \qquad (62)\end{aligned}$$

Equating (62) with equation (55) obtained from the argument principle, yields the desired valence formula as given in equation (54). □

Valence formula for $\Gamma(2)$

In principle, one can do the same process we did for $\mathrm{SL}_2(\mathbb{Z})$ for any finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ to obtain a similar valence formula. To illustrate this in further generality, we compute a valence formula for $\Gamma(2)$.

To assist with the integration, we first prove the following lemma, which generalises the result obtained when integrating over the segments B'_nC_n and C'_nD_n in the $\mathrm{SL}_2(\mathbb{Z})$ case.

Lemma 23: Let f be a modular form of weight k for Γ. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two oriented curves in $\mathcal{H}$ such that $\gamma\mathcal{D}_1 = -\mathcal{D}_2$ for some $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$ (i.e. γ sends the curve $\mathcal{D}_1$ to the curve $\mathcal{D}_2$ with the opposite orientation) where $c \neq 0$. Then

$$\int_{\mathcal{D}_1} \frac{f'(z)}{f(z)} dz + \int_{\mathcal{D}_2} \frac{f'(z)}{f(z)} dz = -k \int_{\mathcal{D}_1} \frac{1}{z + d/c} dz. \tag{63}$$

Proof: As $\gamma \in \Gamma$, by the modularity condition we have that $f(\gamma z) = (cz+d)^k f(z)$. Thus, by the product rule for differentiation, we have

$$f'(z) = -ck(cz+d)^{-k-1} f(\gamma z) + (cz+d)^{-k} f'(\gamma z) \frac{d}{dz}(\gamma z).$$

Thus, we obtain

$$\begin{aligned} \frac{f'(z)}{f(z)} &= \frac{-ck(cz+d)^{-k-1} f(\gamma z) + (cz+d)^{-k} f'(\gamma z) \frac{d}{dz}(\gamma z)}{(cz+d)^{-k} f(\gamma z)} \\ &= \frac{-ck}{cz+d} + \frac{f'(\gamma z)}{f(\gamma z)} \frac{d}{dz}(\gamma z). \end{aligned} \tag{64}$$

Integrating both sides of (64) over $\mathcal{D}_1$ and doing a change of variables from $\gamma z \to z$ in the last integral yields:

$$\begin{aligned} & \int_{\mathcal{D}_1} \frac{f'(z)}{f(z)} dz = -k \int_{\mathcal{D}_1} \frac{c}{cz+d} dz + \int_{\mathcal{D}_1} \frac{f'(\gamma z)}{f(\gamma z)} \frac{d}{dz}(\gamma z) dz \\ \implies \quad & \int_{\mathcal{D}_1} \frac{f'(z)}{f(z)} dz + \int_{\mathcal{D}_2} \frac{f'(z)}{f(z)} dz = -k \int_{\mathcal{D}_1} \frac{1}{z + d/c} dz \end{aligned}$$

which proves the lemma. □

Theorem 24: Let f be a non-zero modular form of weight k for $\Gamma(2)$. Then

$$\mathrm{ord}_\infty(f) + \mathrm{ord}_0(f) + \mathrm{ord}_1(f) + \sum_{z \in \Gamma(2) \backslash \mathcal{H}} \mathrm{ord}_z(f) = \frac{k}{2} \tag{65}$$

Proof: We first establish a fundamental domain for $\Gamma(2)$. It can be verified that a set of coset representatives for $\Gamma(2)$ can be given as $R := \{I, T, S, TS, ST^{-1}, TST\}$. Indeed one can verify manually that $AB^{-1} \notin \Gamma$ for all $A, B \in R$, which thus verifies that each element in R yields a distinct coset. As $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(2)] = 6$, this confirms that R is a complete set of coset representatives. This hence provides a connected fundamental domain for $\Gamma(2)$, around which we can integrate.

We furthermore note that $\Gamma(2)$ has three cusps, where the natural representatives chosen (given our fundamental domain) are ∞, 0 and 1.

Let f be a non-zero modular form of weight k for $\Gamma(2)$. Since f is holomorphic at infinity, there exists $R > 0$ such that for all $\mathrm{Im}(\tau) > R$, the value of $f(\tau)$ is nonzero.

We define a contour integral $\mathcal{C}_n$ around the fundamental domain for each $n \in \mathbb{N}$, as shown in Figure 6 . As with the case for $\mathrm{SL}_2(\mathbb{Z})$, we can deal with any zeros or poles along the boundary by modifying the contour to circle around any zero/poles, ensuring that each orbit with non-zero order in $\Gamma(2)\backslash\mathcal{H}$ has exactly one representative in the region enclosed by $\mathcal{C}_n$, with the exception of $\Gamma(2)\omega$. We also let the imaginary part of A_nG_n be $R + n$ and let each arc be sufficiently

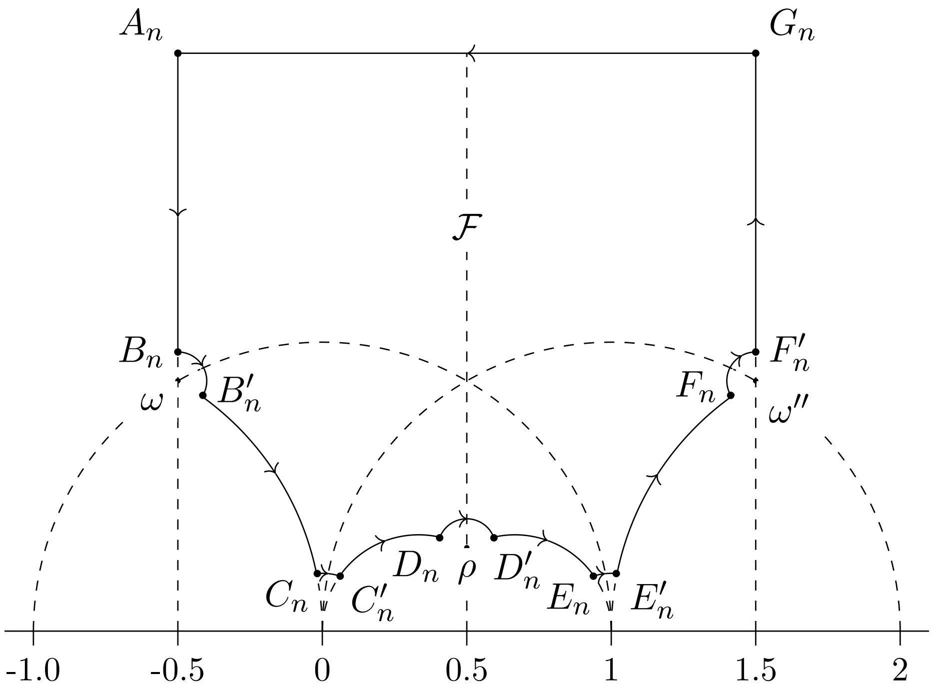

Figure 6: Contour integral of $\mathcal{C}_n$ for $\Gamma(2)$

small so as to isolate any other poles/zeros.

We shall assume without loss of generality that this modification is done accordingly for any zeros/poles along $\mathcal{C}_n$. (Note, this is not explicitly draw in on Figure 6.)

As with $\mathrm{SL}_2(\mathbb{Z})$, we first apply the argument principle to the region enclosed by $\mathcal{C}_n$. Using a similar argument to that for $\mathrm{SL}_2(\mathbb{Z})$, we obtain

$$\frac{1}{2\pi i}\oint_{\mathcal{C}_n}\frac{f'(z)}{f(z)}dz = \sum_{\substack{z\in\Gamma(2)\backslash\mathcal{H}\\ z\neq\omega}}\mathrm{ord}_z(f). \tag{66}$$

As before, we now simply integrate over each piece-wise smooth segment of $\mathcal{C}_n$ separately. Firstly, we consider the integral over A_nB_n and F'_nG_n. Noting that $T^2\in\Gamma(2)$, we have that $f(z)=f(z+2)$, and thus doing a change of variables $z\to z+2$, we obtain

$$\int_{A_n}^{B_n}\frac{f'(z)}{f(z)}dz = \int_{A_n}^{B_n}\frac{f'(z+2)}{f(z+2)}dz = \int_{G_n}^{F'_n}\frac{f'(z)}{f(z)}dz = -\int_{F'_n}^{G_n}\frac{f'(z)}{f(z)}dz. \tag{67}$$

Thus, we have the cancellation

$$\int_A^B\frac{f'(z)}{f(z)}dz + \int_{F'}^{G}\frac{f'(z)}{f(z)}dz = 0. \tag{68}$$

We now evaluate the integral along G_nA_n in a similar manner to the case for $\mathrm{SL}_2(\mathbb{Z})$. Here, we apply the change of variables $q=\exp\pi iz$, noting that f is periodic with period 2 since $T^2\in\Gamma(2)$. We therefore note

$$\frac{f'(z)}{f(z)} = \frac{\frac{d}{dq}\tilde{f}(q)\frac{dq}{dz}}{\tilde{f}(q)} = \frac{\tilde{f}'(q)}{\tilde{f}(q)}\pi iq.$$

Thus, denoting $\tilde{A}_n$ and $\tilde{G}_n$ as the points $\exp(\pi i A_n)$ and $\exp(\pi i G_n)$ respectively, we have

$$\int_{G_n}^{A_n} \frac{f'(z)}{f(z)} dz = \int_{\tilde{G}_n}^{\tilde{A}_n} \frac{\tilde{f}'(q)}{\tilde{f}(q)} \pi i q \frac{dq}{\pi i q} = \int_{\tilde{G}_n}^{\tilde{A}_n} \frac{\tilde{f}'(q)}{\tilde{f}(q)} dq. \tag{69}$$

Hence, as $n \to \infty$, the integral approaches a small circle (with radius $e^{2\pi(R+n)}$) around $q = 0$ in the **negative** orientation. Thus

$$\lim_{n\to\infty} \int_{G_n}^{A_n} \frac{f'(z)}{f(z)} dz = -\mathrm{ord}_0(\tilde{f}) = -\mathrm{ord}_\infty(f). \tag{70}$$

To evaluate the arcs around the other two cusps, we first do an initial change of variables to transform the arcs to horizontal segments in $\mathcal{H}$ with imaginary part tending towards infinity. We then apply a second change of variables, doing a q-substitution by using the definition of the order at the cusps. As shown in equation (60), we have

$$\frac{(f|_k S)'(z)}{(f|_k S)(z)} = \frac{-k}{z} + \frac{1}{z^2}\frac{f'(-1/z)}{f(-1/z)} = \frac{-k}{z} + \frac{1}{z^2}\frac{f'(Sz)}{f(Sz)}.$$

Thus, changing variables from z to Sz, we obtain

$$\begin{aligned}\int_{C_n}^{C_n'} \frac{f'(z)}{f(z)} dz = \int_{SC_n}^{SC_n'} \frac{f'(Sz)}{f(Sz)} d(Sz) &= \int_{SC_n}^{SC_n'} z^2 \left(\frac{(f|_k S)'(z)}{(f|_k S)(z)} + \frac{k}{z} \right) \frac{1}{z^2} dz \\ &= \int_{SC_n}^{SC_n'} \frac{(f|_k S)'(z)}{(f|_k S)(z)} dz + \int_{SC_n}^{SC_n'} \frac{k}{z} dz.\end{aligned}$$

To evaluate the second integral, we note that $|k/z|$ decreases to 0 as $\mathrm{Im}(z) \to \infty$. Thus, noting that SC_nSC_n' is a horizontal line segment in $\mathcal{H}$ with fixed length of 2, we obtain

$$\lim_{n\to\infty} \int_{SC_n}^{SC_n'} \frac{k}{z} dz = 0.$$

To evaluate the first integral, we do a q-expansion in a similar way to that done for the segment G_nA_n. Therefore, by definition of the order of the cusp at 0, noting that $S(\infty) = 0$, we obtain

$$\begin{aligned}\lim_{n\to\infty} \int_{C_n}^{C_n'} \frac{f'(z)}{f(z)} dz = \lim_{n\to\infty} \int_{SC_n}^{SC_n'} \frac{(f|_k S)'(z)}{(f|_k S)(z)} dz &= \lim_{n\to\infty} \int_{\tilde{SC}_n}^{\tilde{SC}_n'} \frac{(\tilde{f}|_k S)'(q)}{(\tilde{f}|_k S)(q)} dq \\ = -\mathrm{ord}_0(\tilde{f}|_k S) = -\mathrm{ord}_\infty(f|_k S) = -\mathrm{ord}_0(f).\end{aligned}$$

The arc E_nE_n' around the cusp 1 can be evaluated in a similar manner. We note that $TS(\infty) = 1$. Therefore, doing an initial substitution $z \mapsto TSz$, followed by a q-substitution yields

$$\begin{aligned}\lim_{n\to\infty} \int_{E_n}^{E_n'} \frac{f'(z)}{f(z)} dz = \lim_{n\to\infty} \int_{TSE_n}^{TSE_n'} \frac{(f|_k TS)'(z)}{(f|_k TS)(z)} dz &= \lim_{n\to\infty} \int_{\tilde{TSE}_n}^{\tilde{TSE}_n'} \frac{(\tilde{f}|_k TS)'(q)}{(\tilde{f}|_k TS)(q)} dq \\ = -\mathrm{ord}_0(\tilde{f}|_k TS) = -\mathrm{ord}_\infty(f|_k TS) = -\mathrm{ord}_1(f)\end{aligned}$$

We now evaluate the integral around the elliptic point $\rho = 1/2 + i/\sqrt{12} = (-ST^{-2}S)(\omega)$. As $-ST^{-2}S \in \Gamma(2)$, we have that ρ is in the orbit of ω. Proceeding similarly to the evaluation of the elliptic points in $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$, we note that the arc D_nD_n' tends towards a local angle of $\Delta\theta_D = 2\pi/3$. We thus obtain

$$\lim_{n\to\infty} \frac{1}{2\pi i} \int_{D_n}^{D_n'} \frac{f'(z)}{f(z)} dz = -\frac{2\mathrm{ord}_\omega(f)}{3}.$$

Evaluating the arcs around the other two elliptic points (both of which are in the orbit of ω and have a local angle of $2\pi/3$), yields

$$\lim_{n\to\infty}\frac{1}{2\pi i}\int_{C_n}^{C_n'}\frac{f'(z)}{f(z)}dz=-\frac{2\mathrm{ord}_\omega(f)}{3}\quad\text{and}\quad\lim_{n\to\infty}\frac{1}{2\pi i}\int_{E_n}^{E_n'}\frac{f'(z)}{f(z)}dz=-\frac{2\mathrm{ord}_\omega(f)}{3}.$$

To evaluate the integral along $B_n'C_n$ and $C_n'D_n$, we apply Lemma 23. We let $\gamma_1=\begin{pmatrix}1&0\\2&1\end{pmatrix}\in\Gamma(2)$ and note that $\gamma_1(B_n'C_n)=D_nC_n'$. Thus, we have

$$\int_{B_n'}^{C_n}\frac{f'(z)}{f(z)}dz+\int_{C_n'}^{D_n}\frac{f'(z)}{f(z)}dz=-k\int_{B_n'}^{C_n}\frac{1}{z+1/2}dz. \tag{71}$$

Similarly, we let $\gamma_2=\begin{pmatrix}1&-2\\2&-3\end{pmatrix}$, and note that $\gamma_2\in\Gamma(2)$ and $\gamma_2(E'F)=(ED')$. Thus, again by Lemma 23, we obtain

$$\int_{D_n'}^{E_n}\frac{f'(z)}{f(z)}dz+\int_{E_n'}^{F_n}\frac{f'(z)}{f(z)}dz=-k\int_{E_n'}^{F_n}\frac{1}{z-3/2}dz. \tag{72}$$

We now apply a change of variables in the first integral with $z\mapsto z-1/2$. Thus the contour of integration becomes a negatively orientated arc from $i\frac{\sqrt{3}}{2}$ to $1/2$. Similarly, we apply a change of variables in the second integral with $z\mapsto z+3/2$.

We thus obtain

$$-k\int_{B'}^{C}\frac{1}{z+1/2}-k\int_{E'}^{F}\frac{1}{z-3/2}=-k\int_{-1/2}^{i\sqrt{3}/2}\frac{1}{z}\,dz+\int_{i\sqrt{3}/2}^{1/2}\frac{1}{z}\,dz. \tag{73}$$

Noting that the only pole of $\frac{1}{z}$ is at $z=0$, we transform the contour such that it forms a half circle in the negative orientation, as shown in Figure 7.

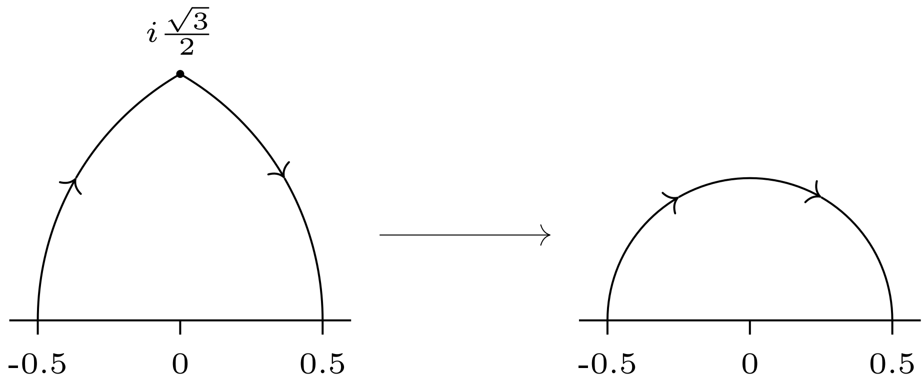

Figure 7: Transformation of the contour to a negatively orientated half circle in $\mathcal{H}$ centred at 0 with radius $r=1/2$.

We can now easily compute the integral by doing the standard change of variables $z=\frac{1}{2}\exp i\theta$, thus $dz=\frac{1}{2}i\exp i\theta\,d\theta$.

$$\int_{-1/2}^{i\sqrt{3}/2}\frac{1}{z}\,dz+\int_{i\sqrt{3}/2}^{1/2}\frac{1}{z}\,dz=\int_{\pi}^{0}i\,d\theta=-\pi i \tag{74}$$

Therefore, the total contribution of the integral over the segments $B_n'C_n, C_n'D_n, D_n', E_n$ and $E_n'F_n$ gives us

$$\int_{B_n'}^{C_n} \frac{f'(z)}{f(z)}dz + \int_{C_n'}^{D_n} \frac{f'(z)}{f(z)}dz + \int_{D_n'}^{E_n} \frac{f'(z)}{f(z)}dz + \int_{E_n'}^{F_n} \frac{f'(z)}{f(z)}dz = k\pi i$$

Finally, putting each of the separate contributions together gives us

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{2\pi i} \oint_{\mathcal{C}_n} \frac{f'(z)}{f(z)}dz &= \frac{1}{2\pi i}(0 + i\pi k) - \mathrm{ord}_\infty(f) - \mathrm{ord}_0(f) - \mathrm{ord}_1(f) - \mathrm{ord}_\omega(f) \\ &= \frac{k}{2} - \mathrm{ord}_\infty(f) - \mathrm{ord}_0(f) - \mathrm{ord}_1(f) - \mathrm{ord}_\omega(f) \end{aligned} \tag{75}$$

Equating (75) with equation (66) obtained from the argument principle yields

$$\begin{aligned} &\sum_{\substack{z\in\Gamma(2)\backslash\mathcal{H} \\ z\neq\omega}} \mathrm{ord}_z(f) = \frac{k}{2} - \mathrm{ord}_\infty(f) - \mathrm{ord}_0(f) - \mathrm{ord}_1(f) - \mathrm{ord}_\omega(f) \\ \implies \quad &\mathrm{ord}_\infty(f) + \mathrm{ord}_0(f) + \mathrm{ord}_1(f) + \sum_{z\in\Gamma(2)\backslash\mathcal{H}} \mathrm{ord}_z(f) = \frac{k}{2} \end{aligned}$$

where the final equality arises from absorbing $\mathrm{ord}_\omega(f)$ into the sum obtained from the argument principle. This hence proves the valence formula for $\Gamma(2)$. □

Valence formula for general Γ

In principle, a similar procedure to that done for $\Gamma(2)$ can be used to evaluate the contour integral around the fundamental domain for any finite index subgroup. Indeed, given a finite index subgroup Γ the main steps to follow are:

1. Determine a (connected) fundamental domain for Γ. This can be done either by manually searching for a set of coset representatives for $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$, or alternatively by using Farey symbols [13].
2. Integrate along a contour $\mathcal{C}$ around the fundamental domain.
3. Modify $\mathcal{C}$ accordingly to pass around zeros/poles on the boundary.
4. Identify which matrices $\gamma \in \Gamma$ send boundary segments of $\mathcal{F}$ to other boundary segments.
5. Use Lemma 23 to evaluate the integral along these boundary segments.
6. Use Equation (51) to evaluate the contours around the elliptic points.
7. Apply a q-substitution to evaluate the arcs around the cusps.
8. Equate the total contribution obtained from each piece-wise smooth segment with that obtained from using the argument principle on the contour integral.

Generalising the above procedure to arbitrary congruence subgroups, even for certain types of subgroups such as $\Gamma_0(p)$ where the fundamental domain is simple and well-known, is a highly non-trivial task. Instead of having to work with the fundamental domain directly, we can instead prove a generalised valence formula without any prior knowledge required of the fundamental domain.

We first prove the following group-theoretic result:

Lemma 25: [2, p. 35] Let G be a group acting transitively on a set X (i.e. there is only one orbit). Let H be a finite index subgroup of G. Note that H acts on X, and define G_x (respectively H_x) to be the stabilizer subgroup of G (respectively H) with respect to x.

Denoting R as a set of orbit representatives for the quotient $H\backslash X$, we get

$$\sum_{x\in R}[G_x : H_x] = [G : H]. \tag{76}$$

Note that if G and H are *finite* groups, then equation (76) is simply a consequence of the Orbit-Stabiliser theorem. For our purposes, however, we shall require the result for any groups G, H where H has finite index in G.

Proof: [2, p. 35] We first note that, by definition, H_x is a subgroup of G_x. We now define the map $\sigma : H_x\backslash G_x \to H\backslash G$ where $\sigma(H_x g) = Hg$ for all $g \in G_x$. Clearly σ is well-defined. Furthermore, let $g_1, g_2 \in G_x$ be such that $\sigma(H_x g_1) = \sigma(H_x g_2)$. We thus obtain

$$\begin{aligned}\sigma(H_x g_1) = \sigma(H_x g_2) &\implies Hg_1 = Hg_2\\ &\implies g_1 g_2^{-1} \in H \quad \text{and} \quad g_1 g_2^{-1} \in G_x\\ &\implies g_1 g_2^{-1} \in H_x\\ &\implies H_x g_1 = H_x g_2.\end{aligned}$$

Hence, σ is an injective map. As H has finite index in G, this implies H_x has finite index in G_x. We thus obtain $|H_x\backslash G_x| = |\mathrm{Im}(\sigma)| = |H\backslash HG_x|$

Now let $x_0 \in X$ be given, and define the map $\phi : H\backslash G \to H\backslash X$ where $\phi(Hg) = (Hgx_0)$. We first note that ϕ is surjective due to the group action of G on X being transitive.

Now, for each $x \in X$, we note that there exists $g_x \in G$ such that $g_x x_0 = x$ as G acts transitively on X. We also define two subsets F_{H_x} and F'_{H_x} of $H\backslash G$, where

$$F_{H_x} = \{Hg \in H\backslash G \mid Hgx = Hx\}$$

and

$$F'_{H_x} = \{Hg \in H\backslash G \mid Hgx_0 = Hx\}.$$

Note that, by definition, $F'_{H_x} = \phi^{-1}(\{Hx\})$. We now define a map $\tau : F_{H_x} \to F'_{H_x}$ where $\tau(Hg) = Hgg_x$. We note that this map is well-defined as

$$Hg \in F_{H_x} \implies Hgx = Hx \implies Hgg_x x_0 = Hx \implies Hgg_x \in F'_{H_x}$$

Furthermore, we note that τ is bijective. Indeed, if $g_1, g_2 \in G$ such that $\tau(Hg_1) = \tau(Hg_2)$, then

$$\begin{aligned}\tau(Hg_1) = \tau(Hg_2) &\implies Hg_1 g_x = Hg_2 g_x\\ &\implies g_1 g_x (g_2 g_x)^{-1} \in H\\ &\implies g_1 g_2^{-1} \in H\\ &\implies Hg_1 = Hg_2\end{aligned}$$

thus proving τ injective. To prove surjectivity, let $Hg \in F'_{H_x}$ and note that

$$\tau(Hgg_x^{-1}) = Hgg_x^{-1} g_x = Hg.$$

Thus, we have τ a bijective map and hence $|F'_{H_x}| = |F_{H_x}|$. We note that F_{H_x} is by definition the image of σ. Thus, we obtain

$$|F'_{H_x}| = |F_{H_x}| = |\mathrm{Im}(\sigma)| = [G_x : H_x] \tag{77}$$

for each $x \in X$. Finally, we conclude by summing over a set the coset represnetatives R for $H\backslash X$

$$[G : H] = |H\backslash G| = \sum_{x\in R} |\phi^{-1}(\{Hx\})| = \sum_{x\in R} |F'_{H_x}| = \sum_{x\in R} [G_x : H_x]$$

which thus proves the lemma. □

We are now ready to prove the main theorem, which generalises the valence formula obtained from $\mathrm{SL}_2(\mathbb{Z})$ and $\Gamma(2)$. Our proof shall mostly follow the one given by Bruin–Dahmen [2, p. 41].

Theorem 26: [2, p. 40] Let f be a non-zero modular form of integer weight k for some finite index subgroup Γ. For each cusp $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$, define

$$\epsilon_{\Gamma,\mathfrak{c}} = \begin{cases} 1 & \text{if } -I \notin \Gamma \text{ and } \mathfrak{c} \text{ is regular} \\ 2 & \text{if } -I \in \Gamma \text{ or } \mathfrak{c} \text{ is irregular} \end{cases}.$$

Then we have

$$\sum_{z\in\Gamma\backslash\mathcal{H}} \frac{\mathrm{ord}_z(f)}{|\Gamma_z|} + \sum_{\mathfrak{c}\in\mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)}{\epsilon_{\Gamma,\mathfrak{c}}} = \frac{k}{24}[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]. \tag{78}$$

We also have an analogous result for the projective special linear group $\mathrm{PSL}_2(\mathbb{Z})$ where we define

$$\bar{\epsilon}_{\Gamma,\mathfrak{c}} = \begin{cases} 1 & \text{if } \mathfrak{c} \text{ is regular} \\ 2 & \text{if } \mathfrak{c} \text{ is irregular} \end{cases}$$

and we have

$$\sum_{z\in\Gamma\backslash\mathcal{H}} \frac{\mathrm{ord}_z(f)}{|\bar{\Gamma}_z|} + \sum_{\mathfrak{c}\in\mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)}{\bar{\epsilon}_{\Gamma,\mathfrak{c}}} = \frac{k}{12}[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]. \tag{79}$$

Proof: [2, p. 41] We first prove that if γ_1 and γ_2 represent the same coset of Γ in $\mathrm{SL}_2(\mathbb{Z})$ then $(f|_k\gamma_1)(z) = (f|_k\gamma_2)(z)$.

Let $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Note that $\Gamma\gamma_1 = \Gamma\gamma_2$ implies $\gamma_1\gamma_2^{-1} \in \Gamma$. Thus, we have

$$f(\gamma_1\gamma_2^{-1}z) = ((c_1d_2 - c_2d_1)z + (a_2d_1 - b_2c_1))^k f(z). \tag{80}$$

Changing variables from z to $\gamma_2 z = \frac{a_2z+b_2}{c_2z+d_2}$ yields

$$\begin{aligned} f(\gamma_1\gamma_2^{-1}\gamma_2 z) &= \left((c_1d_2 - c_2d_1)\left(\frac{a_2z+b_2}{c_2z+d_2}\right) + (a_2d_1 - b_2c_1)\right)^k f(\gamma_2 z) \\ \implies \quad f(\gamma_1 z) &= \left(\frac{(c_1z+d_1)(a_2d_2 - b_2c_2)}{c_2z+d_2}\right)^k f(\gamma_2 z) \\ \implies \quad (c_1z+d_1)^{-k} f(\gamma_1 z) &= (c_2z+d_2)^{-k} f(\gamma_2 z) \\ \implies \quad (f|_k\gamma_1)(z) &= (f|_k\gamma_2)(z). \end{aligned}$$

Hence, the slash operator depends only on the coset γ represents and not on any particular representative.

We now prove that, for any $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$, we have

$$(f|_k\gamma)(\gamma_0 z) = (c_0 z + d_0)^k (f|_k\gamma\gamma_0)(z). \tag{81}$$

Note that

$$\begin{aligned}(f|_k\gamma)(\gamma_0 z) &= (c\gamma_0 z + d)^{-k} f(\gamma\gamma_0 z) \\ &= \left(c\frac{a_0 z + b_0}{c_0 z + d_0} + d\right)^{-k} f(\gamma\gamma_0 z) \\ &= (c_0 z + d_0)^k (c(a_0 z + b_0) + d(c_0 z + d_0))^{-k} f(\gamma\gamma_0 z) \\ &= (c_0 z + d_0)^k ((ca_0 + dc_0)z + (cb_0 + dd_0))^{-k} f(\gamma\gamma_0 z) \\ &= (c_0 z + d_0)^k (f|_k\gamma\gamma_0)(z)\end{aligned}$$

which thus proves (81).

Let R be a set of coset representatives for the quotient $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$. Note that since Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$, R is finite as $|R| = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$. We thus define the product F as

$$F(z) = \prod_{\gamma\in R} (f|_k\gamma)(z). \tag{82}$$

Note that $F(z)$ is invariant under a choice of coset representatives, as shown previously.

We prove that $F(z)$ is a modular form of weight $k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$ for $\mathrm{SL}_2(\mathbb{Z})$. Firstly, as R is finite, F is a finite product of holomorphic functions on $\mathcal{H}$. Thus F is holomorphic on $\mathcal{H}$.

To prove the modularity condition, let $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$ and define $Q = R\gamma_0 = \{\gamma'\gamma_0 | \gamma' \in R\}$. Note that Q is also a set of coset representatives for the quotient $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$. We thus note

$$\begin{aligned}F(\gamma_0 z) = \prod_{\gamma'\in R} (f|_k\gamma')(\gamma_0 z) &= \prod_{\gamma'\in R} (c_0 z + d_0)^k (f|_k(\gamma'\gamma_0)z) \\ &= (c_0 z + d_0)^{k[\mathrm{SL}_2(\mathbb{Z}):\Gamma]} \prod_{\gamma'\in Q} (f|_k(\gamma')z).\end{aligned}$$

Thus, F satisfies the modularity condition for any $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$ (as a modular form of weight $k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$).

To check that F is holomorphic at ∞, we simply note that $(f|_k\gamma)(z)$ is bounded as $\mathrm{Im}(z) \to \infty$. Thus the product over a set of finite coset representatives is bounded and hence $F(z)$ is bounded as $\mathrm{Im}(z) \to \infty$, therefore proving that $F(z)$ is a modular form of weight $k[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$ for $\mathrm{SL}_2(\mathbb{Z})$.

As $F(z)$ is a modular form for $\mathrm{SL}_2(\mathbb{Z})$, we can thus apply the valence formula obtained for $\mathrm{SL}_2(\mathbb{Z})$, by Theorem 22. This gives us

$$\mathrm{ord}_\infty(F) + \frac{1}{2}\mathrm{ord}_i(F) + \frac{1}{3}\mathrm{ord}_\omega(F) + \sum_{w\in W} \mathrm{ord}_w(F) = \frac{k}{12}[\mathrm{SL}_2(\mathbb{Z}) : \Gamma]. \tag{83}$$

We now rewrite equation (83), noting that the only elliptic points (points with non-trivial stabiliser) in $\mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ are i and ω. Thus, by (26), we have

$$\frac{1}{2}\text{ord}_\infty(F) + \sum_{z\in \text{SL}_2(\mathbb{Z})\backslash\mathcal{H}} \frac{\text{ord}_z(F)}{|\text{SL}_2(\mathbb{Z})_z|} = \frac{k}{24}[\text{SL}_2(\mathbb{Z}) : \Gamma] \tag{84}$$

We now calculate the order of F at a point $z \in \mathcal{H}$.

$$\begin{aligned}
\text{ord}_z(F) = \text{ord}_z\left(\prod_{\gamma\in R}(f|_k\gamma)(z)\right) &= \sum_{\gamma\in R}\text{ord}_z((f|_k\gamma)(z)) \\
&= \sum_{\gamma\in R}\text{ord}_z\left((cz+d)^{-k}f(\gamma z)\right) \\
&= \sum_{\gamma\in R}\text{ord}_{\gamma z}(f) \\
&= \sum_{w\in\Gamma\backslash\text{SL}_2(\mathbb{Z})z}[\text{SL}_2(\mathbb{Z})_w : \Gamma_w]\,\text{ord}_w(f) \qquad (85)
\end{aligned}$$

where the last equality follows from Lemma 25, where we have grouped together the terms γ where $\gamma z = w$ for each $w \in \Gamma\backslash\text{SL}_2(\mathbb{Z})z$.

We now note that $\text{SL}_2(\mathbb{Z})_w$ is finite and the cardinality $|\text{SL}_2(\mathbb{Z})_w|$ also only depends on the orbit of w. Hence, as $w = \gamma(z)$, we have $|\text{SL}_2(\mathbb{Z})_w| = |\text{SL}_2(\mathbb{Z})_z|$. We therefore obtain

$$[\text{SL}_2(\mathbb{Z})_w : \Gamma_w] = \frac{|\text{SL}_2(\mathbb{Z})_w|}{|\Gamma_w|} = \frac{|\text{SL}_2(\mathbb{Z})_z|}{|\Gamma_w|}.$$

Hence, dividing the identity obtaining in (85) by $|\text{SL}_2(\mathbb{Z})_z|$ yields

$$\frac{\text{ord}_z(F)}{|\text{SL}_2(\mathbb{Z})_z|} = \sum_{w\in\Gamma\backslash\text{SL}_2(\mathbb{Z})z}\frac{[\text{SL}_2(\mathbb{Z})_w : \Gamma_w]}{|\text{SL}_2(\mathbb{Z})_z|}\text{ord}_w(f) = \sum_{w\in\Gamma\backslash\text{SL}_2(\mathbb{Z})z}\frac{\text{ord}_w(f)}{|\Gamma_w|}.$$

This therefore allows us to evaluate the second term in equation (84), given as

$$\sum_{z\in\text{SL}_2(\mathbb{Z})\backslash\mathcal{H}}\frac{\text{ord}_z(F)}{|\text{SL}_2(\mathbb{Z})_z|} = \sum_{z\in\text{SL}_2(\mathbb{Z})\backslash\mathcal{H}}\sum_{w\in\Gamma\backslash\text{SL}_2(\mathbb{Z})z}\frac{\text{ord}_w(f)}{|\Gamma_w|} = \sum_{w\in\Gamma\backslash\mathcal{H}}\frac{\text{ord}_w(f)}{|\Gamma_w|}. \tag{86}$$

where the last equality follows by considering the natural bijective map ϕ from $\text{SL}_2(\mathbb{Z})\backslash\mathcal{H} \times \Gamma\backslash\text{SL}_2(\mathbb{Z})$ to $\Gamma\mathcal{H}$ given by $\phi(\text{SL}_2(\mathbb{Z})z, \Gamma\gamma) = \Gamma\gamma(z)$.

What remains is to evaluate the first term in equation (84). This is shown in Theorem 27 to be

$$\frac{1}{2}\text{ord}_\infty F = \sum_{c\in\text{Cusps}(\Gamma)}\frac{\text{ord}_{\Gamma,c}(f)}{\epsilon_{\Gamma,c}}. \tag{87}$$

We can therefore finally conclude that

$$\sum_{w\in\Gamma\backslash\mathcal{H}}\frac{\text{ord}_w(f)}{|\Gamma_w|} + \sum_{c\in\text{Cusps}(\Gamma)}\frac{\text{ord}_{\Gamma,c}(f)}{\epsilon_{\Gamma,c}} = \sum_{z\in\text{SL}_2(\mathbb{Z})\backslash\mathcal{H}}\frac{\text{ord}_z(F)}{|\text{SL}_2(\mathbb{Z})_z|} + \frac{1}{2}\text{ord}_\infty F = \frac{k}{24}[\text{SL}_2(\mathbb{Z}) : \Gamma] \tag{88}$$

where the final equality follows from equation (84). This therefore proves the general valence formula for arbitrary finite index subgroups Γ.

The projective analogue can be derived by simply considering two cases:

Case 1: $-I \in \Gamma$. We therefore have: $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma] = [\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]$. We also note that all cusps of $X(\Gamma)$ are regular, thus $\epsilon_{\Gamma,c} = 2$ and $\bar{\epsilon}_{\Gamma,c} = 1$, therefore $\epsilon_{\Gamma,c} = 2\bar{\epsilon}_{\Gamma,c}$ for all $c \in \mathrm{Cusps}(\Gamma)$. We also note that $-I$ trivially stabilises all points in $\mathcal{H}$, thus $-I \in \Gamma_z$ for all $z \in \Gamma\backslash\mathcal{H}$ and therefore $|\Gamma_z| = 2|\bar{\Gamma}_z|$.

Hence, multiplying equation (88) by 2, we obtain

$$\begin{aligned} & \sum_{w \in \Gamma\backslash\mathcal{H}} \frac{\mathrm{ord}_w(f)}{|\Gamma_w|/2} + \sum_{c \in \mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,c}(f)}{\epsilon_{\Gamma,c}/2} = \frac{k}{12}[\mathrm{SL}_2(\mathbb{Z}) : \Gamma] \\ \implies \quad & \sum_{w \in \Gamma\backslash\mathcal{H}} \frac{\mathrm{ord}_w(f)}{|\bar{\Gamma}_w|} + \sum_{c \in \mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,c}(f)}{\bar{\epsilon}_{\Gamma,c}/2} = \frac{k}{12}[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]. \end{aligned}$$

Case 2: $-I \notin \Gamma$. In this case, we note $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma] = 2[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]$. We also trivially obtain $\epsilon_{\Gamma,c} = \bar{\epsilon}_{\Gamma,c}$ by definition, and finally note that since $-I \notin \Gamma_z$, we have $|\Gamma_z| = |\bar{\Gamma}_z|$.

Therefore, simply rewriting equation (88), with the above observations, we obtain

$$\sum_{w \in \Gamma\backslash\mathcal{H}} \frac{\mathrm{ord}_w(f)}{|\bar{\Gamma}_w|} + \sum_{c \in \mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,c}(f)}{\bar{\epsilon}_{\Gamma,c}/2} = \frac{k}{12}[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]$$

which therefore proves the projective analogue of the theorem. □

All that remains now is to prove equation (87):

Theorem 27: [2, p. 42] Let Γ be a finite index subgroup in $\mathrm{SL}_2(\mathbb{Z})$ and let f be a modular form of weight k for Γ. Define F as given in equation (82). Then, we have

$$\frac{1}{2}\mathrm{ord}_\infty F = \sum_{\mathfrak{c} \in \mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)}{\epsilon_{\Gamma,\mathfrak{c}}}. \tag{89}$$

Proof: [2, p. 42] We shall prove the above theorem in several steps, given from (i) to (viii). First, we define the set $U \subseteq \mathrm{SL}_2(\mathbb{Z})$ where

$$U = \langle T \rangle = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{Z} \right\} \tag{90}$$

and consider the quotient

$$Z = \mathrm{SL}_2(\mathbb{Z})/U. \tag{91}$$

Step (i): We define the map $\phi : Z \to \mathbb{P}^1(\mathbb{Q})$ given by $\gamma U \mapsto \gamma(\infty)$. We first prove that ϕ is well-defined. Indeed let $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma_1 U = \gamma_2 U$. Thus

$$\begin{aligned} \gamma_1 U = \gamma_2 U \implies & \gamma_1^{-1}\gamma_2 \in U \\ \implies & \gamma_1^{-1}\gamma_2 \in \mathrm{Stab}_{\mathrm{SL}_2(\mathbb{Z})}(\infty) \quad \text{as } U \subseteq \mathrm{SL}_2(\mathbb{Z})_\infty \\ \implies & \gamma_1^{-1}\gamma_2(\infty) = \infty \\ \implies & \gamma_1(\infty) = \gamma_2(\infty). \end{aligned}$$

We also easily note that ϕ is also compatible with the group action induced by $\mathrm{SL}_2(\mathbb{Z})$. Indeed, for any $\gamma, \gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$ we have

$$\gamma_0\phi(\gamma U) = \gamma_0\gamma(\infty) = \phi(\gamma_0\gamma U).$$

Furthermore, we note ϕ is the *unique* map $Z \to \mathbb{P}^1(\mathbb{Q})$ that is compatible with the group action induced by $\mathrm{SL}_2(\mathbb{Z})$ and that sends U to ∞. Indeed, let ϕ' be such a map. We thus have that $\phi'(U) = \infty$ and

$$\phi'(\gamma U) = \gamma\gamma^{-1}\phi(\gamma U) = \gamma\phi(\gamma^{-1}(\gamma U) = \gamma\phi(U) = \gamma(\infty) = \phi(\gamma U). \tag{92}$$

Step (ii): We now proceed by defining the natural map $\phi_q : \Gamma\backslash Z \to \mathrm{Cusps}(\Gamma) = \Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ which we obtain by quotienting both sides of the map ϕ by Γ. The map ϕ_q is therefore given by $\phi_q(\Gamma\gamma U) = \Gamma\gamma(\infty)$. As before, we first prove ϕ_q is well-defined.

$$\begin{aligned}\Gamma\gamma_1 U = \Gamma\gamma_2 U &\implies \gamma_1 \in \Gamma\gamma_2 U \\ &\implies \gamma_1 = \delta\gamma_2\zeta \qquad \text{for some } \delta \in \Gamma, \zeta \in U \\ &\implies \gamma_1(\infty) = \delta\gamma_2\zeta(\infty) \\ &\implies \gamma_1(\infty) = \delta\gamma_2(\infty) \\ &\implies \Gamma\gamma_1(\infty) = \Gamma\gamma_2(\infty).\end{aligned}$$

We now prove that, for each $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$, the cardinality of $\phi_q^{-1}(\mathfrak{c})$ is $2/\epsilon_{\Gamma,\mathfrak{c}}$.

Let $\Gamma\gamma_1 U, \Gamma\gamma_2 U \in \Gamma\backslash Z$ such that $\phi_q(\Gamma\gamma_1 U) = \phi_q(\Gamma\gamma_2 U)$. Thus:

$$\begin{aligned}\phi_q(\Gamma\gamma_1 U) = \phi_q(\Gamma\gamma_2 U) &\implies \Gamma\gamma_1(\infty) = \Gamma\gamma_2(\infty) \\ &\implies \gamma_2^{-1}\delta\gamma_1 \in \mathrm{SL}_2(\mathbb{Z})_\infty \qquad \text{for some } \delta \in \Gamma \\ &\implies \gamma_2^{-1}\delta\gamma_1 \in \{\pm I\}U \\ &\implies \gamma_2^{-1}\delta\gamma_1 \in U \qquad \text{or} \quad \gamma_2^{-1}\delta\gamma_1 \in -U \\ &\implies \delta\gamma_1 \in \gamma_2 U \qquad \text{or} \quad \delta\gamma_1 \in -\gamma_2 U \\ &\implies \Gamma\gamma_1 U \subseteq \Gamma\gamma_2 U \quad \text{or} \quad \Gamma\gamma_1 U \subseteq -\Gamma\gamma_2 U\end{aligned}$$

Note that, if $-I \in \Gamma$, then we have $\Gamma = -\Gamma$ and thus $\Gamma\gamma_1 U = \Gamma\gamma_2 U$, therefore proving ϕ_q is injective. Otherwise, if $-I \notin \Gamma$, then we have two possibilities of either $\Gamma\gamma_1 U = \Gamma\gamma_2 U$ or $\Gamma\gamma_1 U = -\Gamma\gamma_2 U$, noting that $\Gamma\gamma_2 U \neq -\Gamma\gamma_2 U$, since $-I \notin U$.

If $\mathfrak{c} = \Gamma\gamma_1(\infty) = \Gamma\gamma_2(\infty)$ is a **regular** cusp, then by definition, we have that either case could occur and thus $\phi_q^{-1}(\mathfrak{c}) = 2$.

Conversely, if $c = \Gamma\gamma_1(\infty) = \Gamma\gamma_2(\infty)$ is an **irregular** cusp, then we note that

$$\gamma_2^{-1}\delta\gamma_1 \in \mathrm{SL}_2(\mathbb{Z})_\infty \implies \gamma_1^{-1}\delta\gamma_1 \in \gamma_1^{-1}\gamma_2\mathrm{SL}_2(\mathbb{Z})_\infty$$

However, as $\gamma_1^{-1}\delta\gamma_1 \in H_\mathfrak{c} = \left\langle \begin{pmatrix} -1 & h_\Gamma(\mathfrak{c}) \\ 0 & -1 \end{pmatrix} \right\rangle$, then we thus obtain that only one of $\gamma_1^{-1}\delta\gamma_1 \in \gamma_1^{-1}\gamma_2 U$ or $\gamma_1^{-1}\delta\gamma_1 \in -\gamma_1^{-1}\gamma_2 U$ may be obtained. Thus, we have $\phi_q^{-1}(\mathfrak{c}) = 1$.

We therefore obtain

$$\phi_q^{-1}(\mathfrak{c}) = \begin{cases} 1 & \text{if } -I \in \Gamma \text{ or } \mathfrak{c} \text{ is irregular} \\ 2 & \text{if } -I \notin \Gamma \text{ and } \mathfrak{c} \text{ is regular} \end{cases}$$

which thus proves that $\phi_q^{-1}(\mathfrak{c}) = 2/\epsilon_{\Gamma,\mathfrak{c}}$.

Step (iii): We now define the natural map $\sigma : \Gamma\backslash\mathrm{SL}_2(\mathbb{Z}) \to \Gamma\backslash Z$ defined by $\Gamma\gamma \to \Gamma\gamma U$ where $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Note that σ is clearly well-defined. To calculate the cardinality of the fibre of σ, we let $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$ be given, and calculate $\sigma^{-1}(\{\Gamma\gamma_0 U\})$. Indeed, let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\sigma(\Gamma\gamma) = \Gamma\gamma_0 U$. Thus:

$$\begin{aligned}\sigma(\Gamma\gamma) = \Gamma\gamma_0 U &\implies \Gamma\gamma U = \Gamma\gamma_0 U \\ &\implies \gamma_0^{-1}\delta\gamma \in U \\ &\implies \gamma_0^{-1}\delta\gamma \in U\end{aligned}$$

Let $h = h_\Gamma(\Gamma\gamma_0(\infty))$ be the width of the cusp $\Gamma\gamma_0(\infty)$ and define $H = \left\langle \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \right\rangle \subseteq U$. Note that h is the index of H in U. Let $\{\gamma_r \in R \,:\, r \in R\}$ be a set of coset representatives for H in U. We thus have

$$\begin{aligned}\gamma_0^{-1}\delta\gamma \in U &\implies \gamma_0^{-1}\delta\gamma \in \bigcup_{r\in R} H\gamma_r \\ &\implies \gamma_0^{-1}\delta\gamma \in H\gamma_r \quad \text{for some } r \in R.\end{aligned}$$

We now consider two cases: If $\Gamma\gamma_0(\infty)$ is **regular**, then we have $H \subseteq H_c \subseteq \gamma_0^{-1}\Gamma\gamma_0$. Thus

$$\gamma_0^{-1}\delta\gamma \in H\gamma_r \implies \gamma_0^{-1}\delta\gamma \in \gamma_0^{-1}\Gamma\gamma_0\gamma_r \implies \Gamma\gamma = \Gamma\gamma_0\gamma_r.$$

As there are h possible choices for γ_r, this proves $|\sigma^{-1}(\{\Gamma\gamma_0 U\})| \leq h$ To prove $|\sigma^{-1}(\{\Gamma\gamma_0 U\})| = h$, we let $\gamma_{r,1}, \gamma_{r,2} \in R$ such that $\Gamma\gamma_0\gamma_{r,1} = \Gamma\gamma_0\gamma_{r,2}$. Thus

$$\begin{aligned}\Gamma\gamma_0\gamma_{r,1} = \Gamma\gamma_0\gamma_{r,2} &\implies \gamma_0\gamma_{r,1}(\gamma_0\gamma_{r,2})^{-1} \in \Gamma \\ &\implies \gamma_0\gamma_{r,1}\gamma_{r,2}^{-1}\gamma_0^{-1} \in \gamma \\ &\implies \gamma_{r,1}\gamma_{r,2}^{-1} \in \gamma_0^{-1}\Gamma\gamma_0 \\ &\implies \gamma_{r,1}\gamma_{r,2}^{-1} \in H \quad (\text{as } \gamma_{r,1}\gamma_{r,2} \in \mathrm{SL}_2(\mathbb{Z})_\infty) \\ &\implies H\gamma_{r,1} = H\gamma_{r,2}\end{aligned}$$

We therefore have that $H\gamma_{r,1} \neq H\gamma_{r,2}$ implies $\Gamma\gamma_0\gamma_{r,1} \neq \Gamma\gamma_0\gamma_{r,2}$ Thus, as $H\backslash U$ consists of h distinct cosets, we have that $|\sigma^{-1}(\{\Gamma\gamma_0 U\})| = h$.

If If $\Gamma\gamma_0(\infty)$ is **irregular**, then we have $H \subseteq \{\pm I\}H_c \subseteq \gamma_0^{-1}\{\pm I\}\Gamma\gamma_0$. Thus:

$$\gamma_0^{-1}\delta\gamma \in H\gamma_r \implies \gamma_0^{-1}\delta\gamma \in \gamma_0^{-1}\{\pm I\}\Gamma\gamma_0\gamma_r \implies \Gamma\gamma = \Gamma\gamma_0\gamma_r \quad \text{or} \quad \Gamma\gamma = -\Gamma\gamma_0\gamma_r$$

Similarly as before, we have that $|\sigma^{-1}(\{\Gamma\gamma_0 U\})| \leq 2h$. To prove equality, as before, we note that the cosets $\Gamma\gamma_0\gamma_r$, $r \in R$ are all distinct. Furthermore, we prove that the sets $\{\Gamma\gamma_0\gamma_r \,:\, r \in R\}$ and $\{-\Gamma\gamma_0\gamma_r \,:\, r \in R\}$ are disjoint. Indeed, we have

$$\begin{aligned}\Gamma\gamma_0\gamma_{r,1} = -\Gamma\gamma_0\gamma_{r,2} &\implies \gamma_{r,1}\gamma_{r,2}^{-1} \in -\gamma_0^{-1}\Gamma\gamma_0 \\ &\implies \gamma_{r,1}\gamma_{r,2}^{-1} \in -H \quad (\text{as } \gamma_{r,1}\gamma_{r,2} \in \mathrm{SL}_2(\mathbb{Z})_\infty)\end{aligned}$$

Note that if $\gamma_{r,1} = \gamma_{r,2}$, then this implies $-I \in H_c$ which contradicts $\Gamma\gamma_0(\infty)$ being an irregular cusp. Otherwise, assuming $\gamma_{r,1} \neq \gamma_{r,2}$ we have $\gamma_{r,1}\gamma_{r,2}^{-1} \in H_c \cap U \subseteq H$, which contradicts $\gamma_{r,1}$ and $\gamma_{r,2}$ being distinct coset representatives of H in U.

Thus we obtain $|\sigma^{-1}(\{\Gamma\gamma_0 U\})| = 2h$ when $\Gamma\gamma_0(\infty)$ is irregular. We therefore have

$$|\sigma^{-1}(\{\Gamma\gamma_0 U\})| = \begin{cases} h & \text{if } \Gamma\gamma_0(\infty) \text{ regular} \\ 2h & \text{if } \Gamma\gamma_0(\infty) \text{ irregular} \end{cases}$$

which thus proves $|\sigma^{-1}(\{\Gamma\gamma_0 U\})| = \tilde{h}_\Gamma(\Gamma\gamma_0(\infty))$ by definition.

At this stage, a summary of the maps defined so far can be given by the diagram below:

$$
\begin{array}{ccccc}
 & & Z & \xrightarrow{\ \phi\ } & \mathbb{P}^1(\mathbb{Q}) \\
 & & \downarrow & & \downarrow \\
\Gamma\backslash \mathrm{SL}_2(\mathbb{Z}) & \xrightarrow{\ \sigma\ } & \Gamma\backslash Z & \xrightarrow{\ \phi_q\ } & \Gamma\backslash \mathbb{P}^1(\mathbb{Q}) = \mathrm{Cusps}(\Gamma)
\end{array}
$$

In summary, we have the unique map $\phi : Z \to \mathbb{P}^1(\mathbb{Q})$ which is compatible with the group action of $\mathrm{SL}_2(\mathbb{Z})$ and sends U to ∞. Quotienting both sides of this map with Γ give us ϕ_q which sends $\Gamma\gamma U$ to $\Gamma\gamma(\infty)$, where the fibre of ϕ_q over $\mathfrak{c}$, for any cusp $\mathfrak{c}$ has cardinality $2/\epsilon_{\Gamma,\mathfrak{c}}$. We then constructed the natural map $\sigma : \Gamma\ \mathrm{SL}_2(\mathbb{Z}) \to \Gamma\ Z$ where the fibre over any $x \in \Gamma\backslash Z$ has cardinality $\tilde{h}_\Gamma(\phi_q(x))$.

Step (iv): We now prove that, for any finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$, there exists a subgroup Γ' of $\mathrm{SL}_2(\mathbb{Z})$ such that Γ' is *normal* in $\mathrm{SL}_2(\mathbb{Z})$ and $\Gamma' \subseteq \Gamma$ where Γ' has finite index in Γ.

Note that in the case where Γ is a congruence subgroup, there always exists a positive integer N such that $\Gamma(N) \subseteq \Gamma$, where $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$, as shown in Theorem 6. In the more general case, we define Γ' as follows

$$\Gamma' = \bigcap_{\gamma \in \mathrm{SL}_2(\mathbb{Z})} \gamma^{-1}\Gamma\gamma. \tag{93}$$

Clearly, as $\gamma^{-1}\Gamma\gamma$ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})$ for each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, then Γ' is a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, it can easily be shown that Γ' is normal. Indeed, let $\delta \in \mathrm{SL}_2(\mathbb{Z})$. We note

$$\delta^{-1}\Gamma'\delta = \delta^{-1}\left(\bigcap_{\gamma \in \mathrm{SL}_2(\mathbb{Z})} \gamma^{-1}\Gamma\gamma\right)\delta = \bigcap_{\gamma \in \mathrm{SL}_2(\mathbb{Z})} (\gamma\delta)^{-1}\Gamma(\gamma\delta) = \Gamma'$$

as right-multiplication by δ is a bijection from $\mathrm{SL}_2(\mathbb{Z})$ to itself. We also note that Γ' is clearly contained in Γ since

$$\Gamma' = \bigcap_{\gamma \in \mathrm{SL}_2(\mathbb{Z})} \gamma^{-1}\Gamma\gamma = \Gamma \cap \left(\bigcap_{\gamma \in \mathrm{SL}_2(\mathbb{Z})-\{I\}} \gamma^{-1}\Gamma\gamma\right) \subseteq \Gamma.$$

Finally, we prove that Γ' has finite index in Γ. We define a map $\psi : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Sym}(\Gamma\backslash \mathrm{SL}_2(\mathbb{Z}))$ where $\mathrm{Sym}(\Gamma\backslash \mathrm{SL}_2(\mathbb{Z}))$ denotes the set of permutations of $\Gamma\backslash \mathrm{SL}_2(\mathbb{Z})$ (i.e. bijective functions on $\Gamma\backslash \mathrm{SL}_2(\mathbb{Z})$). We define ψ as:

$$\psi(\gamma) = f : \Gamma\backslash \mathrm{SL}_2(\mathbb{Z}) \to \Gamma\backslash \mathrm{SL}_2(\mathbb{Z}) \quad \text{where} \quad f(\Gamma\gamma_0) = \Gamma\gamma_0\gamma \quad \text{for all } \gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$$

One can easily verify that ψ is a group homomorphism. We finally prove that the kernel of ψ is Γ'. Note that:

$$
\begin{aligned}
\delta \in \Gamma' &\iff \delta \in \bigcap_{\gamma \in \mathrm{SL}_2(\mathbb{Z})} \gamma^{-1}\Gamma\gamma \\
&\iff \delta \in \gamma^{-1}\Gamma\gamma \qquad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\
&\iff \gamma\delta \in \Gamma\gamma \qquad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\
&\iff \Gamma\gamma\delta = \Gamma\gamma \qquad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\
&\iff \psi(\delta)\Gamma\gamma = \Gamma\gamma \qquad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\
&\iff \delta \in \ker(\psi)
\end{aligned}
$$

Hence, the kernel of the group homomorphism ψ is Γ', and thus:

$$\mathrm{SL}_2(\mathbb{Z})/\Gamma' = \mathrm{SL}_2(\mathbb{Z})/\ker(\psi) \cong \mathrm{Im}\psi \leq \mathrm{Sym}(\Gamma\backslash\mathrm{SL}_2(\mathbb{Z}))$$

As Γ is finite index in $\mathrm{SL}_2(\mathbb{Z})$, this implies $\mathrm{Sym}(\Gamma\backslash\mathrm{SL}_2(\mathbb{Z}))$ is a *finite* group and thus Γ' has finite index in $\mathrm{SL}_2(\mathbb{Z})$.

Step (v): Let Γ' be any finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ such that Γ' is normal in $\mathrm{SL}_2(\mathbb{Z})$, $\Gamma' \subseteq \Gamma$ and Γ' has finite index in Γ. Define the natural map $\pi : \Gamma'\backslash\mathrm{SL}_2(\mathbb{Z}) \to \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$ given by $\pi(\Gamma'\gamma) = \Gamma\gamma$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Note that π clearly well-defined as $\Gamma' \subseteq \Gamma$.

Now, let $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$ be given. We calculate the cardinality of $\pi^{-1}(\{\Gamma\gamma_0\})$. Let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\pi(\Gamma'\gamma) = \Gamma\gamma_0$. Also, noting the Γ' is a subgroup of Γ, let R denote a set of coset representatives for Γ' in Γ. Note that $|R| = [\Gamma : \Gamma']$. Thus, we have

$$\begin{aligned}
\pi(\Gamma'\gamma) = \Gamma\gamma_0 &\implies \Gamma\gamma = \Gamma\gamma_0 \\
&\implies \gamma\gamma_0^{-1} \in \Gamma \\
&\implies \gamma\gamma_0^{-1} \in \bigcup_{\gamma_r \in R} \Gamma'\gamma_r \\
&\implies \gamma\gamma_0^{-1} \in \Gamma'\gamma_r \quad \text{ for some } \gamma_r \in R \\
&\implies \Gamma'\gamma = \Gamma'\gamma_r\gamma_0 \quad \text{ for some } \gamma_r \in R
\end{aligned}$$

which proves that $|\pi^{-1}(\{\Gamma\gamma_0\})| \leq |R| = [\Gamma : \Gamma']$. We also note, if $\gamma_{r,1}, \gamma_{r,2} \in R$ such that $\Gamma'\gamma_{r,1}\gamma_0 = \Gamma'\gamma_{r,2}\gamma_0$, then

$$\begin{aligned}
\Gamma'\gamma_{r,1}\gamma_0 = \Gamma'\gamma_{r,2}\gamma_0 &\implies \gamma_{r,1}\gamma_0(\gamma_{r,2}\gamma_0)^{-1} \in \Gamma' \\
&\implies \gamma_{r,1}\gamma_{r,2}^{-1} \in \Gamma' \\
&\implies \Gamma'\gamma_{r,1} = \Gamma'\gamma_{r,2}
\end{aligned}$$

thus proving that distinct cosets $\Gamma'\gamma_{r,1} \neq \Gamma'\gamma_{r,2}$ results in $\Gamma'\gamma_{r,1}\gamma_0 \neq \Gamma'\gamma_{r,2}\gamma_0$. We therefore obtain that $|\pi^{-1}(\{\Gamma\gamma_0\})| = [\Gamma : \Gamma']$.

For the following two steps, we shall derive two different expressions for the sum

$$\sum_{\Gamma'\gamma \in \Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})} \mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f)$$

Step (vi): We rewrite F as a product over the cosets in $\Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})$. Noting the result in (v), we obtain

$$F^{[\Gamma:\Gamma']} = \left(\prod_{\gamma \in \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})} f|_k\gamma \right)^{[\Gamma:\Gamma']} = \prod_{\gamma \in \Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})} f|_k\gamma$$

noting that the slash operator $f|_k\gamma$ depends only on the coset. Now, as Γ' is a normal subgroup in $\mathrm{SL}_2(\mathbb{Z})$, we have that all cusps $\mathfrak{c} \in \mathrm{Cusps}(\Gamma')$ have the same width $h_{\Gamma'}(\mathfrak{c})$ and that all cusps are either all regular or all irregular. We can thus define $\tilde{h}_{\Gamma'} = \tilde{h}_{\Gamma'}(\mathfrak{c})$ for all $\mathfrak{c} \in \mathrm{Cusps}(\Gamma')$. We can therefore do a q-expansion where $q_{\Gamma'} = \exp(2\pi i z/\tilde{h}_{\Gamma'})$. We therefore obtain

$$\prod_{\gamma \in \Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})} f|_k\gamma(z) = \prod_{\gamma \in \Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})} \left(\sum_{n=0}^{\infty} a_{\mathfrak{c},n}(f) q_{\Gamma'}^n \right) = \prod_{\gamma \in \Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})} \left(\sum_{n=\mathrm{ord}_{\Gamma',c}(f)}^{\infty} a_{\mathfrak{c},n}(f) q_{\Gamma'}^n \right)$$

where the last equality follows by definition of the order of f at the cusps. In order for us to do a q-expansion of F at infinity, since F is a modular form for $\mathrm{SL}_2(\mathbb{Z})$, we must expand over $q = \exp(2\pi i z) = q_{\Gamma'}^{\tilde{h}_{\Gamma'}}$. Therefore, noting the first non-zero term in the q-expansion, we obtain

$$\begin{aligned} F^{[\Gamma:\Gamma']} &= \prod_{\gamma \in \Gamma' \backslash \mathrm{SL}_2(\mathbb{Z})} \left(\sum_{n=\mathrm{ord}_{\Gamma',\mathfrak{c}}(f)}^{\infty} a_{\mathfrak{c},n}(f) q_{\Gamma'}^{n} \right) \\ &= \left(\prod_{\gamma \in \Gamma' \backslash \mathrm{SL}_2(\mathbb{Z})} a_{\mathfrak{c},n}(f) \right) q_{\Gamma'}^{\sum \mathrm{ord}_{\Gamma',\mathfrak{c}}(f)} + \mathcal{O}\left(q_{\Gamma'}^{1+\sum \mathrm{ord}_{\Gamma',n}(f)} \right) \\ &= \left(\prod_{\gamma \in \Gamma' \backslash \mathrm{SL}_2(\mathbb{Z})} a_{\mathfrak{c},n}(f) \right) q^{\left(\sum \mathrm{ord}_{\Gamma',\mathfrak{c}}(f)\right)/\tilde{h}_{\Gamma'}} + \mathcal{O}\left(q^{1+\left(\sum \mathrm{ord}_{\Gamma',\mathfrak{c}}(f)\right)/\tilde{h}_{\Gamma'}} \right) \end{aligned}$$

where the sum in the exponent of q ranges over all $\gamma \in \Gamma' \backslash \mathrm{SL}_2(\mathbb{Z})$, and noting that only the first-term in the q-expansion is explicitly calculated. Taking orders on both sides, we obtain

$$\mathrm{ord}_{\infty}\left(F^{[\Gamma:\Gamma']} \right) = \frac{1}{\tilde{h}_{\Gamma'}} \sum_{\gamma \in \Gamma' \backslash \mathrm{SL}_2(\mathbb{Z})} \mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f)$$

which yields

$$\sum_{\Gamma'\gamma \in \Gamma' \backslash \mathrm{SL}_2(\mathbb{Z})} \mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f) = [\Gamma : \Gamma']\, \tilde{h}_{\Gamma'}\, \mathrm{ord}_{\infty}(F). \tag{94}$$

Step (vii): Using the results obtained from Lemma 21, as well as step (iii) and (v), we note

$$\begin{aligned}
\sum_{\gamma\in\Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})}\mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f) &= [\Gamma:\Gamma']\sum_{\gamma\in\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})}\mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f) && \text{by (iii)}\\
&= [\Gamma:\Gamma']\sum_{x\in\Gamma\backslash Z}\tilde{h}_{\Gamma}(x)\mathrm{ord}_{\Gamma',\phi_q(x)}(f) && \text{by (v)}\\
&= [\Gamma:\Gamma']\sum_{x\in\Gamma\backslash Z}\tilde{h}_{\Gamma'}(x)\mathrm{ord}_{\Gamma,\phi_q(x)}(f) && \text{by Lemma 21}\\
&= [\Gamma:\Gamma']\,\tilde{h}_{\Gamma'}\sum_{x\in\Gamma\backslash Z}\mathrm{ord}_{\Gamma,\phi_q(x)}(f)
\end{aligned}$$

where the last equality follows as $\tilde{h}_{\Gamma'}(\mathfrak{c})$ is common for all cusps $\mathfrak{c}$ of Γ'. This therefore yields

$$\sum_{\Gamma'\gamma\in\Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})}\mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f) = [\Gamma:\Gamma']\,\tilde{h}_{\Gamma'}\sum_{x\in\Gamma\backslash Z}\mathrm{ord}_{\Gamma,\phi_q(x)}(f). \tag{95}$$

Step (viii): Finally, we equate equations (94) and (95) to give

$$\begin{aligned}
[\Gamma:\Gamma']\,\tilde{h}_{\Gamma'}\,\mathrm{ord}_\infty(F) &= \sum_{\Gamma'\gamma\in\Gamma'\backslash\mathrm{SL}_2(\mathbb{Z})}\mathrm{ord}_{\Gamma',\Gamma\gamma(\infty)}(f) = [\Gamma:\Gamma']\,\tilde{h}_{\Gamma'}\sum_{x\in\Gamma\backslash Z}\mathrm{ord}_{\Gamma,\phi_q(x)}(f)\\
\implies \qquad \mathrm{ord}_\infty(F) &= \sum_{x\in\Gamma\backslash Z}\mathrm{ord}_{\Gamma,\phi_q(x)}(f)
\end{aligned}$$

noting that $[\Gamma:\Gamma']\,\tilde{h}_{\Gamma'}$ is non-zero. We now use the result in step (ii), noting that the sum on the right hand side only depends on $\phi_q(x)$. Thus

$$\mathrm{ord}_\infty(F) = \sum_{x\in\Gamma\backslash Z}\mathrm{ord}_{\Gamma,\phi_q(x)}(f) = \sum_{\mathfrak{c}\in\mathrm{Cusps}(\Gamma)}\frac{2}{\epsilon_{\Gamma,\mathfrak{c}}}\mathrm{ord}_{\Gamma,\mathfrak{c}}(f).$$

Hence, we obtain

$$\frac{1}{2}\mathrm{ord}_\infty F = \sum_{\mathfrak{c}\in\mathrm{Cusps}(\Gamma)}\frac{\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)}{\epsilon_{\Gamma,\mathfrak{c}}} \tag{96}$$

which concludes the proof. □

Applications

Sturm's bound

The valence formula given in Theorem 26 for arbitrary finite-index subgroups Γ provides a powerful tool for bounding the dimension of $\mathcal{M}_k(\Gamma)$. To show this, we first note the following lemma.

Lemma 28: [10, p. 53] (Sturm's bound) Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and let $f \in \mathcal{M}_k(\Gamma)$. If

$$\frac{\mathrm{ord}_{\Gamma,\infty}(f)}{\bar{\epsilon}_{\Gamma,\infty}} > \frac{k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]}{12} \tag{97}$$

then $f = 0$.

Proof: Since f is holomorphic on $\mathcal{H}$ and at each cusp, we have that $\mathrm{ord}_z(f) \geq 0$ for all $z \in \mathcal{H}$ and $\mathrm{ord}_{\Gamma,\mathfrak{c}}(f) \geq 0$ for all cusps $\mathfrak{c} \in \mathrm{Cusps}(\Gamma)$. Thus, the left hand side of (79) is non-negative. Therefore, if f satisfies inequality (97), then this contradicts the valence formula, and thus $f = 0$.

We remark that Sturm's bound is often given as the slightly stronger result [15, p. 88]

$$\sum_{\mathfrak{c} \in \mathrm{Cusps}(\Gamma)} \frac{\mathrm{ord}_{\Gamma,\mathfrak{c}}(f)}{\epsilon_{\Gamma,\mathfrak{c}}} > \frac{k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]}{12}. \tag{98}$$

However, the proof of this is identical to what we present above.

Using Sturm's bound, one can show that any modular form f of weight k for a finite index subgroup Γ is uniquely identified by knowing sufficiently many of the first few terms in its q-expansion.

Indeed, let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a finite index subgroup, and let N be the minimal positive integer such that $T^N \in \Gamma$ (note that, by definition, $N = \tilde{h}_\Gamma(\infty)$). Now, for any modular form f for Γ, we can Let $q = \exp \frac{2\pi i z}{N}$. We thus write it's Fourier expansion as $f = \sum a_n q^n$.

Theorem 29: [15, p. 89] Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Let $f, g \in \mathcal{M}_k(\Gamma)$ and denote their respective q-expansions at ∞ as

$$\tilde{f}(q) = \sum_{n=0}^{\infty} a_n q^n \quad \text{and} \quad \tilde{g}(q) = \sum_{n=0}^{\infty} b_n q^n$$

where $q = \exp(2\pi i z / \tilde{h}_\Gamma(\infty))$. If $a_n = b_n$ for all $0 \leq n \leq \bar{\epsilon}_{\Gamma,\infty} k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]/12$, then $f = g$.

Proof: Let $r = \bar{\epsilon}_{\Gamma,\infty} k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]/12$ and define the modular form $h(z) = f(z) - g(z)$. Note that $h \in \mathcal{M}_k(\Gamma)$. The q-expansion of h is thus

$$\tilde{h}(q) = \sum_{n=0}^{\infty} (a_n - b_n) q^n = q^{r+1} \sum_{n=r+1}^{\infty} (a_n - b_n) q^{n-r-1}. \tag{99}$$

Therefore, h has a zero at infinity of order $> r$. Thus, we have

$$\begin{aligned} \mathrm{ord}_{\Gamma,\infty}(f) &> \frac{\bar{\epsilon}_{\Gamma,\infty} k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]}{12} \\ \implies \quad \frac{\mathrm{ord}_{\Gamma,\infty}(f)}{\bar{\epsilon}_{\Gamma,\infty}} &> \frac{k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]}{12}. \end{aligned}$$

Therefore, by Sturm's bound, we have h is identically zero and hence $f = g$. $\square$

Corollary 30: [15, p. 89] Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Then

$$\dim \mathcal{M}_k(\Gamma) \leq \left\lfloor \frac{k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]}{12} \right\rfloor + 1. \tag{100}$$

Proof: We first assume the cusp at infinity is regular, thus $\bar{\epsilon}_{\Gamma,\infty} = 1$. Let $r = \left\lfloor \frac{k[\mathrm{PSL}_2(\mathbb{Z}):\Gamma]}{12} \right\rfloor$. We define a linear map $\phi : \mathcal{M}_k(\Gamma) \to \mathbb{C}^{r+1}$ where for any $f \in \mathcal{M}_k(\Gamma)$, we map f to the first $r+1$ of its Fourier coefficients, given as

$$\tilde{f}(q) = \sum_{n=0}^{\infty} a_n q^n \quad \mapsto \quad (a_0, a_1, a_2, \ldots, a_r).$$

We clearly note that ϕ is linear. Furthermore, by Theorem 29, we have that ϕ is injective. Thus, we obtain

$$\dim \mathcal{M}_k(\Gamma) = \dim(\mathrm{Im}\phi) \leq \dim(\mathbb{C}^{r+1}) = r + 1 = \left\lfloor \frac{k[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]}{12} \right\rfloor + 1$$

which proves the corollary. If the cusp at infinity is irregular, then doing the q-expansion in the usual way by substituting $q = \exp(2\pi i z/\tilde{h}_\Gamma(\infty))$ introduces a factor of 2, since $\tilde{h}_\Gamma(\infty) = 2h_\Gamma(\infty)$ for irregular cusps.

We now note that substituting $\begin{pmatrix} -1 & h_\Gamma(\infty) \\ 0 & -1 \end{pmatrix}$ into the modularity condition yields

$$f(z - h_\Gamma(\infty)) = (-1)^k f(z).$$

Therefore, if k is even, f is periodic with period $h_\Gamma(\infty) = \frac{1}{2}\tilde{h}_\Gamma(\infty)$ in which case we can q-expand over $q = \exp(2\pi i z/h_\Gamma(\infty))$ and thus proceed as if ∞ is regular. If k is odd, we can then define a new function g on $\mathcal{H}$ given by [14, p. 36]

$$g(z) = f(z) \exp\left(\frac{-\pi i z}{h_\Gamma(\infty)}\right) \qquad \text{for all } z \in \mathcal{H}.$$

We therefore note

$$\begin{aligned} g(z - h_\Gamma(\infty)) &= f(z - h_\Gamma(\infty)) \exp\left(\frac{-\pi i(z - h_\Gamma(\infty))}{h_\Gamma(\infty)}\right) \\ &= e^{\pi i} f(z - h_\Gamma(\infty)) \exp\left(\frac{-\pi i z}{h_\Gamma(\infty)}\right) \\ &= f(z) \exp\left(-\pi i z/h_\Gamma(\infty)\right) \\ &= g(z). \end{aligned}$$

We therefore have that $g(z)$ is periodic with period $h_\Gamma(\infty) = \frac{1}{2}\tilde{h}_\Gamma(\infty)$. Doing the argument as before with $g(z)$ therefore proves the corollary. $\square$

We thus have that $\dim \mathcal{M}_k(\Gamma)$ is always finite-dimensional for any finite index subgroup Γ and weight $k \in \mathbb{Z}$. This fundamental property lies at the heart of why modular forms are so widely studied. Indeed, both Loeffler and Sarnak refer to this as the "The unreasonable effectiveness of modular forms in number theory" [14, p. 46] [21].

We now apply the above theorems to certain spaces.

Dimension of $\mathcal{M}_k(\mathbf{SL}_2(\mathbb{Z}))$

We first consider the simplest case, applying the above results to $\Gamma = \mathrm{SL}_2(\mathbb{Z})$.

Theorem 31: [23, p. 88] Let k be an integer and consider $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, the space of modular forms of weight k for $\mathrm{SL}_2(\mathbb{Z})$. The dimension of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ can be given as

$$\dim\, \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \begin{cases} 0 & \text{if } k < 0 \text{ or } k \text{ is odd} \\ \left\lfloor \frac{k}{12} \right\rfloor & \text{if } k \geq 0,\ k \equiv 2 \pmod{12} \\ \left\lfloor \frac{k}{12} \right\rfloor + 1 & \text{if } k \geq 0,\ k \not\equiv 2 \pmod{12} \end{cases} \tag{101}$$

Proof: [23, p. 88] We first consider $k < 0$. Note that Corollary 30 directly implies that, if $k < 0$, then $\dim\, \mathcal{M}_k(\Gamma) = 0$ for any finite-index subgroup Γ, thus the claim clearly holds for $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, since $-I \in \mathrm{SL}_2(\mathbb{Z})$, we have $\dim\, \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = 0$ for all k odd.

We now assume $k \geq 0$ and k even. We first consider the small weight cases:

Case 1: $k = 0$. We have $\mathbb{C} \subseteq \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, thus $\dim\, \mathcal{M}_0(\mathrm{SL}_2(\mathbb{Z})) \geq 1$. Furthermore, by the valence formula, we obtain $\dim\, \mathcal{M}_0(\mathrm{SL}_2(\mathbb{Z})) \leq 1$. Thus, we have $\dim(M_0(\mathrm{SL}_2(\mathbb{Z})) = 1$.

Case 2: $k = 2$. By the valence formula for $\mathrm{SL}_2(\mathbb{Z})$, we have

$$\mathrm{ord}_{i\infty}(f) + \frac{1}{2}\mathrm{ord}_i(f) + \frac{1}{3}\mathrm{ord}_\omega(f) + \sum_{\substack{z \in \mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H} \\ z \neq \omega, i}} \mathrm{ord}_z(f) = \frac{2}{12}. \tag{102}$$

Note that, as f is holomorphic on $\mathcal{H}$ and at the cusp ∞, we have that the orders at each point is a non-negative integer. Therefore, the lowest positive value the left hand side of (102) can attain is $1/3$. Hence, there are no solutions to equation (102), and thus the only modular form f of weight 2 for $\mathrm{SL}_2(\mathbb{Z})$ is $f = 0$. Thus $\dim\, \mathcal{M}_2(\mathrm{SL}_2(\mathbb{Z})) = 0$.

Case 3: $k \in \{4, 6, 8, 10\}$. As $k < 12$, by the valence formula, we obtain $\dim\, \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \leq 1$. However, since $k \geq 4$ and k even, we have the existence of a non-trivial modular form G_k. Thus $\mathbb{C}G_k \subseteq \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. Therefore $\dim\, \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = 1$.

In order to evaluate the dimension for higher weight k, we construct an isomorphism between $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $\mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$ using the cusp form $\Delta \in \mathcal{S}_{12}(\mathrm{SL}_2(\mathbb{Z}))$ introduced in (23).

Lemma 32: Let $k \geq 0$ be an even integer and define the function $\phi : \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \to \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$ defined as

$$\phi(f) = \Delta f \quad \text{for all } f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})).$$

Then ϕ defines an isomorphism between $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $\mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$.

Proof: We first note that, by Lemma 2, $\Delta f \in \mathcal{M}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$ for all $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ as Δ is a weight 12 modular form for $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, as $\Delta \in \mathcal{S}_{12}(\mathrm{SL}_2(\mathbb{Z}))$, we have $\mathrm{ord}_\infty(\Delta) \geq 1$ and thus

$$\mathrm{ord}_\infty(\Delta f) = \mathrm{ord}_\infty(\Delta) + \mathrm{ord}_\infty(f) \geq 1$$

as f is holomorphic at ∞. Thus $\Delta f \in \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$, which therefore proves that ϕ is well-defined.

We now note that ϕ is clearly linear. Indeed, for any $\alpha, \beta \in \mathbb{C}$ and $f, g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, we have

$$\phi(\alpha f + \beta g) = \Delta(\alpha f + \beta g) = \alpha \Delta f + \beta \Delta f = \alpha \phi(f) + \beta \phi(g).$$

We also note that, as Δ is a cusp form, we have $\mathrm{ord}_\infty(\Delta) \geq 1$. Thus, by the valence formula

$$\frac{1}{2}\mathrm{ord}_i(\Delta) + \frac{1}{3}\mathrm{ord}_\omega(\Delta) + \sum_{\substack{z \in \mathrm{SL}_2(\mathbb{Z}) \backslash \mathcal{H} \\ z \neq \omega, i}} \mathrm{ord}_z(\Delta) = 1 - \mathrm{ord}_\infty(\Delta) \leq 0.$$

As the left hand is non-negative, this implies $\mathrm{ord}_\infty(\Delta) = 1$ and thus

$$\frac{1}{2}\mathrm{ord}_i(\Delta) + \frac{1}{3}\mathrm{ord}_\omega(\Delta) + \sum_{\substack{z \in \mathrm{SL}_2(\mathbb{Z}) \backslash \mathcal{H} \\ z \neq \omega, i}} \mathrm{ord}_z(\Delta) = 1 - \mathrm{ord}_\infty(\Delta) = 0.$$

We therefore obtain that Δ has a *simple* zero at infinity and that $\mathrm{ord}_z(\Delta) = 0$ for all $z \in \mathcal{H}$, thus Δ is non-zero everywhere on $\mathcal{H}$.

We now prove that ϕ is injective. Indeed, let $f_1, f_2 \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ such that $\phi(f_1) = \phi(f_2)$. Then

$$\begin{aligned} \phi(f_1) = \phi(f_2) &\implies \Delta f_1 = \Delta f_2 \implies \Delta(f_1 - f_2) = 0 \\ &\implies f_1 - f_2 = 0 \implies f_1 = f_2 \end{aligned}$$

where the second last equality follows since $\Delta(z) \neq 0$ for all $z \in \mathcal{H}$.

To prove ϕ surjective, let $g \in \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$ and define $f = g/\Delta$. Since g and Δ are both holomorphic on $\mathcal{H}$ and $\Delta(z) \neq 0$ for all $z \in \mathcal{H}$, we have that f is holomorphic on $\mathcal{H}$. Furthermore, by a similar argument given in Lemma 2, we have that f satisfies the modularity condition of weight $k + 12 - 12 = k$ for $\mathrm{SL}_2(\mathbb{Z})$. Finally, since $g \in \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$, we have $\mathrm{ord}_\infty(g) \geq 1$, and thus

$$\mathrm{ord}_\infty(f) = \mathrm{ord}_\infty(g/\Delta) = \mathrm{ord}_\infty(g) - \mathrm{ord}_\infty(\Delta) \geq 0 \tag{103}$$

which proves f is holomorphic at ∞. Thus $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and therefore $\phi(f) = g$, proving surjectivity. □

Therefore, by this isomorphism, we have

$$\dim \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \dim \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z})) \qquad \text{for all } k \geq 0.$$

This proves the first of two steps which wil enable us to calculate recursively the dimension of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for all k.

Now, let $k \geq 4$ with k even. We consider the linear map $\sigma : \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \to \mathbb{C}$ which evaluates the leading coefficient in the q-expansion of f. That is $\sigma(f) = a_0(f) = f(\infty)$. We clearly note that f is linear. Furthermore, computing the kernel of σ, we obtain

$$f \in \ker(\sigma) \iff a_0(f) = 0 \iff f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})).$$

Thus we have $\ker(\sigma) = \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$. We finally note that σ is surjective, as

$$\sigma(\alpha E_k) = a_0(\alpha E_k) = \alpha a_0(E_k) = \alpha \quad \text{for all } \alpha \in \mathbb{C}.$$

We therefore obtain

$$\dim \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \dim \ker \sigma + \dim \mathrm{Im}\, \sigma = \dim \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})) + 1. \tag{104}$$

Thus

$$\dim \mathcal{M}_{k+12}(\mathrm{SL}_2(\mathbb{Z})) = \dim \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z})) + 1 = \dim \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) + 1$$

for all $k \geq 0$, k even. This therefore gives a simple recursive expression, relating the dimension of $\mathcal{M}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$ with that of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. Having computed the dimension for all $k < 12$, we thus conclude by induction that the dimension of $\mathcal{M}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$ is as given in (101). □

Dimension of $\mathcal{M}_k(\Gamma_0(4))$

We shall furthermore use the valence formula to obtain the exact dimension of $\mathcal{M}_k(\Gamma_0(4))$. We first recall the Dedekind eta function, which plays an important role in the theory of modular forms.

Definition 33: [10, p. 98] We define the **Dedekind eta function** $\eta(\tau)$ for $\tau \in \mathcal{H}$ as

$$\eta(\tau) = e^{\frac{\pi\tau i}{12}} \prod_{n=1}^{\infty} \left(1 - e^{2n\pi i\tau}\right) = q^{\frac{1}{24}} \prod_{n=1}^{\infty} (1 - q^n) \tag{105}$$

where the second equality follows by the standard $q = \exp(2\pi i\tau)$ substitution. One can prove that $\eta(\tau)$ is a modular form of weight 1/2, satisfying the following transformation property [10, p. 100]

$$\eta\left(-\frac{1}{z}\right) = \sqrt{\frac{z}{i}}\,\eta(z) \qquad \text{for all } z \in \mathcal{H}$$

We remark that, from the infinite product representation of Δ given in (24), we have $\Delta(q) = \eta^{24}(q)$.

For the purposes of calculating the dimension of $\mathcal{M}_k(\Gamma_0(4))$, we shall instead work with $F_\eta(\tau) = \eta^{12}(2\tau)$. Note that

$$F_\eta(\tau) = \left(e^{\frac{2\pi\tau i}{12}} \prod_{n=1}^{\infty} \left(1 - e^{4n\pi i\tau}\right)\right)^{12} = e^{2\pi i\tau} \prod_{n=1}^{\infty} \left(1 - e^{4n\pi i\tau}\right)^{12} = q \prod_{n=1}^{\infty} \left(1 - q^{2n}\right)^{12}.$$

We can prove that $F_\eta(\tau)$ is a modular form of weight 6 for $\Gamma_0(4)$.

Let $\{x_n\}_1^\infty$ be a sequence of positive real numbers. Recall that an infinite product over $(1 - x_n)$ for $n \geq 1$ converges if and only if the sum over x_n for $n \geq 1$ converges [11, p. 219]. That is

$$\prod_{n=1}^{\infty} (1 + x_n) \text{ converges} \qquad \Longleftrightarrow \qquad \sum_{n=1}^{\infty} x_n \text{ converges.}$$

We therefore note

$$\sum_{n=1}^{\infty} 12q^{2n} = 12 \sum_{n=1}^{\infty} (q^2)^n = \frac{q^2}{1 - q^2}$$

noting that the infinite geometric series converges as $0 \leq |q| < 1$. We furthermore note that this convergence is uniform over any compact disc $D_r = \{|q| \leq r < 1\}$ Thus, the infinite product as defined in $F_\eta(\tau)$ converges for all $\tau \in \mathcal{H}$ and is holomorphic on $\mathcal{H}$.

To prove that $F_\eta(\tau)$ satisfies the modularity condition of weight 6 for $\Gamma_0(4)$, it suffices to prove it for a set of generators of $\Gamma_0(4)$. As proven in the Appendix, we have $\langle -I, T, U \rangle = \Gamma_0(4)$. As $F_\eta(\tau)$ has a well-defined q-expansion by definition, we have that $F_\eta(\tau + 1) = F_\eta(\tau)$. Furthermore, since $k = 6$ is even, we have that $F_\eta(\tau)$ trivially satisfies the modularity condition for $-I$.

It therefore only remains to prove the modularity condition of weight 6 for $U = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}$. We therefore prove

$$F_\eta\left(\frac{\tau}{4\tau+1}\right) = (4\tau+1)^6 F_\eta(\tau). \tag{106}$$

We note that

$$\begin{aligned}
F_\eta\left(\frac{\tau}{4\tau+1}\right) = \eta^{12}\left(\frac{2\tau}{4\tau+1}\right) &= \eta^{12}\left(-\left(-\frac{1}{2\tau}-2\right)^{-1}\right) \\
&= \left(\sqrt{i\left(\frac{4\tau+1}{2\tau}\right)}\,\eta\left(-\frac{1}{2\tau}-2\right)\right)^{12} \\
&= -\left(\frac{4\tau+1}{2\tau}\right)^6 \eta^{12}\left(-\frac{1}{2\tau}\right) \\
&= -\left(\frac{4\tau+1}{2\tau}\right)^6 \left(\sqrt{\frac{2\tau}{i}}\right)^{12} \eta^{12}(2\tau) \\
&= (4\tau+1)^6 \eta^{12}(2\tau)
\end{aligned}$$

which proves the transformation property in (106). Finally, we check the behaviour of $F_\eta(\tau)$ at the three cusps of $\Gamma_0(4)$. By definition of F_η, we have the q-exansion at ∞ as

$$\tilde{F}_\eta(q) = q\prod_{n=1}^{\infty}\left(1-q^{2n}\right)^{12} = q - 12q^3 + 54q^5 + \mathcal{O}(q^6)$$

Thus, F_η has a zero of order 1 at the cusp ∞. To check the order at the cusp 0, we q-expand $F_\eta|_k S$, noting that $S(\infty) = 0$. We thus obtain

$$\begin{aligned}
F_\eta|_k S(\tau) = \tau^{-6}\eta^{12}\left(-\frac{2}{\tau}\right) &= \tau^{-6}\eta^{12}\left(-\frac{1}{\tau/2}\right) \\
&= \tau^{-6}\left(\sqrt{\frac{\tau/2}{i}}\right)^{12} \eta^{12}\left(\frac{\tau}{2}\right) \\
&= -\frac{1}{2^6}\eta^{12}\left(\frac{\tau}{2}\right) \\
&= -\frac{1}{2^6}q^{1/4} + \frac{3}{2^4}q^{3/4} - \frac{27}{2^5}q^{5/4} + \mathcal{O}(q^{6/4})
\end{aligned}$$

We note that the cusp at 0 has width 4, thus q-expanding around $q_0 = q^{1/4}$ yields

$$\tilde{F}_\eta|_k S(q_0) = -\frac{1}{2^6}q_0 + \frac{3}{2^4}q_0^3 - \frac{27}{2^5}q_0^5 + \mathcal{O}(q_0^6).$$

Thus, F_η has a zero of order 1 at the cusp 0. Finally, we q-expand at the cusp 1/2, noting that $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}(\infty) = 1/2$.

$$\begin{aligned}
F_\eta|_k \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}(\tau) &= (2\tau+1)^{-6}\eta^{12}\left(\frac{2\tau}{2\tau+1}\right) \\
&= (2\tau+1)^{-6}\eta^{12}\left(1-\frac{1}{2\tau+1}\right) \\
&= -(2\tau+1)^{-6}\eta^{12}\left(-\frac{1}{2\tau+1}\right)
\end{aligned}$$

$$= \eta^{12}(2\tau + 1)$$
$$= -F_\eta(\tau)$$

Thus, noting that the cusp at 1/2 has width 1, we therefore get the q-expansion

$$\tilde{F}_\eta|_k \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} (q) = -q + 12q^3 - 54q^5 + \mathcal{O}(q^6).$$

Therefore, F_η has a zero of order 1 at the cusp 1/2. Hence, as F_η vanishes at all the cusps, we conclude that $F_\eta \in \mathcal{S}_6(\Gamma_0(4))$.

Using this cusp form, we can now evaluate the dimension of $\mathcal{M}_k(\Gamma_0(4))$.

Theorem 34: [15, p. 90] Let k be an integer. The dimension of $\mathcal{M}_k(\Gamma_0(4))$ is

$$\dim \mathcal{M}_k(\Gamma_0(4)) = \begin{cases} 0 & \text{if } k < 0 \text{ or } k \text{ is odd} \\ \frac{k}{2} + 1 & \text{otherwise} \end{cases}. \tag{107}$$

Proof: We proceed in a similar way to the proof for $\mathrm{SL}_2(\mathbb{Z})$. As before, by the valence formula, we have that $\dim \mathcal{M}_k(\Gamma_0(4)) = 0$ for $k < 0$, and since $-I \in \Gamma_0(4)$, we have $\dim \mathcal{M}_k(\Gamma_0(4)) = 0$ for k odd. Again, noting that $\dim \mathcal{M}_0(\Gamma_0(4)) \leq 1$ by the valence formula and that $\mathbb{C} \subseteq \mathcal{M}_0(\Gamma_0(4))$, we obtain $\dim \mathcal{M}_0(\Gamma_0(4)) = 1$.

We now calculate the dimension of $\mathcal{M}_k(\Gamma_0(4))$ for weights $k = 2$ and $k = 4$. First, we note that $[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma_0(4)] = 6$. Thus, by Corollary 30, we obtain

$$\dim \mathcal{M}_k(\Gamma_0(4) \leq \frac{k}{2} + 1.$$

Therefore, $\dim \mathcal{M}_2(\Gamma_0(4) \leq 2$ and $\dim \mathcal{M}_4(\Gamma_0(4)) \leq 3$. We now compute a set of linearly independent modular forms in $\mathcal{M}_2(\Gamma_0(4))$. Indeed, we have both $E_{2,2} \in \mathcal{M}_2(\Gamma_0(4))$ and $E_{2,4} \in \mathcal{M}_2(\Gamma_0(4))$. We note their respective q-expansions are

$$\tilde{E}_{2,2}(q) = 1 + 24q + \mathcal{O}(q^2)$$
$$\tilde{E}_{2,4}(q) = 1 + 8q + \mathcal{O}(q^2)$$

Thus, $E_{2,2}$ and $E_{2,4}$ are linearly independent in $\mathcal{M}_2(\Gamma_0(4))$. Therefore, we have $\dim \mathcal{M}_2(\Gamma_0(4)) = 2$. For weight $k = 4$, we obtain three linearly independent modular forms $E_4, E_{4,2}$ and $E_{4,4}$. Their q-expansions are

$$\tilde{E}_4(q) = 1 + 240q + 2160q^2 + \mathcal{O}(q^3)$$
$$\tilde{E}_{4,2}(q) = 1 - 16q + 112q^2 + \mathcal{O}(q^3)$$
$$\tilde{E}_{4,4}(q) = 1 - 16q^2 + 112q^4 + \mathcal{O}(q^5)$$

Calculating the determinant of the 3x3 matrix formed by the first three terms in the q-expansion yields

$$\det \begin{pmatrix} 1 & 240 & 2160 \\ 1 & -16 & 112 \\ 1 & 0 & -16 \end{pmatrix} = 65536 \neq 0.$$

Thus, as the determinant is non-zero, we obtain three linearly independent modular forms of weight 4 for $\Gamma_0(4)$. Therefore $\dim \mathcal{M}_4(\Gamma_0(4)) = 3$.

As with $\mathrm{SL}_2(\mathbb{Z})$, we now construct an isomorphism between $\mathcal{M}_k(\Gamma_0(4))$ and $\mathcal{S}_{k+6}(\Gamma_0(4))$ given by $f \mapsto F_\eta f$. Well-definedness and linearity can easily be checked in a similar manner to the case for $\mathrm{SL}_2(\mathbb{Z})$ Now, since $F_\eta \in \mathcal{S}_6(\Gamma_0(4))$, we have

$$\mathrm{ord}_\infty(F_\eta) \geq 1, \quad \mathrm{ord}_0(F_\eta) \geq 1, \quad \text{and} \quad \mathrm{ord}_{1/2}(F_\eta) \geq 1$$

However, by the valence formula for $\Gamma_0(4)$, we have

$$\mathrm{ord}_\infty(F_\eta) + \mathrm{ord}_0(F_\eta) + \mathrm{ord}_{1/2}(F_\eta) \leq 3$$

Thus, F_η has a simple zero at each of the cusps $\infty, 0$ and $1/2$ and is non-zero everywhere on $\mathcal{H}$. As with the proof for $\mathrm{SL}_2(\mathbb{Z})$, these two properties allow us to claim that $f \mapsto F_\eta f$ is an injective and surjective map from $\mathcal{M}_k(\Gamma_0(4))$ to $\mathcal{S}_{k+6}(\Gamma_0(4))$.

We therefore obtain

$$\dim \mathcal{M}_k(\Gamma_0(4)) = \dim \mathcal{S}_{k+6}(\Gamma_0(4)) \qquad \text{for all } k \geq 0.$$

To relate the space of modular forms with the space of cusps forms for a given weight k, we now consider the linear map $\sigma : \mathcal{M}_k(\Gamma_0(4)) \to \mathbb{C}^3$ which maps f to its limit values at the cusps. That is

$$\sigma(f) = (f(\infty), f(0), f(1/2)) = (a_{\infty,0}(f), a_{\infty,1}(f), a_{\infty,1/2}(f)).$$

By definition, we have that the kernel of σ is $\mathcal{S}_k(\Gamma_0(4))$. We also have that σ is surjective, noting that the generalised Eisenstein series $E_k, E_{k,2}$ and $E_{k,4}$ always form a linearly independent set of modular forms for all $k \geq 4$, k even. We thus obtain

$$\dim \mathcal{M}_k(\Gamma_0(4)) = \dim \ker(\sigma) + \dim \mathrm{Im}(\sigma) = \dim \mathcal{S}_k(\Gamma_0(4)) + 3.$$

Therefore

$$\dim \mathcal{M}_{k+6}(\Gamma_0(4)) = \dim \mathcal{S}_{k+6}(\Gamma_0(4)) + 3 = \dim \mathcal{M}_k(\Gamma_0(4)) + 3$$

for all $k \geq 0$, k even. Therefore, as with $\mathrm{SL}_2(\mathbb{Z})$, this gives a recursive expression for the dimension of $\mathcal{M}_{k+6}(\Gamma_0(4))$ given the dimension of $\mathcal{M}_k(\Gamma_0(4))$. Noting that the dimension of $\mathcal{M}_2(\Gamma_0(4))$ and $\mathcal{M}_4(\Gamma_0(4))$ is 2 and 3 respectively, we therefore conclude by induction that the dimension for $\mathcal{M}_k(\Gamma_0(4))$ is as given in (107).

Further applications

We are not only limited to calculating the dimension for $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $\mathcal{M}_k(\Gamma_0(4)$. In fact, an explicit formula for the dimension of $\mathcal{M}_k(\Gamma)$ for several other subgroups Γ can be obtained using similar techniques to that shown previously.

To demonstrate we note the dimension for $\mathcal{M}_k(\Gamma_0(6))$ can be given as

$$\dim \mathcal{M}_k(\Gamma_0(6)) = \begin{cases} 0 & \text{if } k < 0 \text{ or } k \text{ odd} \\ k+1 & \text{otherwise} \end{cases}. \tag{108}$$

We shall simply summarise the steps applied for $\mathrm{SL}_2(\mathbb{Z})$ and $\Gamma_0(4)$ to $\Gamma_0(6)$.

As before, we first prove the dimension for small weights k. Indeed, we have $\dim \mathcal{M}_0(\Gamma_0(6)) = 1$ and $\dim \mathcal{M}_2(\Gamma_0(6)) = 3$. A lower bound of 3 can be attained by noting the three linear independent modular forms of weight 2 obtained from the generalised Eisenstein series construction

$E_{2,2}, E_{2,3}$, and $E_{2,6}$.

Conversely, we obtain an upper bound of 3 on the dimension of $\mathcal{M}_2(\Gamma_0(6)$ by Corollary 30, noting that we have the index $[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma_0(6)] = 12$.

To obtain the dimension for higher weights k by induction, we need to use a cusp form of appropriately small weight. Indeed, one can prove that [16, p. 4852]

$$F_6 := (\eta(\tau)\eta(2\tau)\eta(3\tau)\eta(6\tau))^2 \in \mathcal{S}_4(\Gamma_0(6)).$$

As before, we can prove that the linear map from $\mathcal{M}_k(\Gamma_0(6))$ to $\mathcal{S}_{k+4}(\Gamma_0(6))$ given by $f \mapsto F_6 f$ is an isomorphism. We note that $\Gamma_0(6)$ has four cusps (where a set of suitable representatives can be given as $\{0, 1/2, 1/3, \infty\}$). Thus, by the valence formula, we obtain that F_6 must have a *simple* zero at each cusp and must be non-zero everywhere on $\mathcal{H}$,

Finally, we note that

$$\dim \mathcal{M}_k(\Gamma_0(6)) = \dim \mathcal{S}_k(\Gamma_0(6)) + 4$$

by considering the linear map from $\mathcal{M}_k(\Gamma_0(6))$ to $\mathbb{C}^4$ where f maps to the 4 limit values at each of the four cusps. Noting that this map is surjective as $E_k, E_{k,2}, E_{k,3}$ and $E_{k,6}$ are all linearly independent, we obtain

$$\dim \mathcal{M}_{k+4}(\Gamma_0(6)) = \dim \mathcal{S}_{k+4}(\Gamma_0(6)) + 4 = \mathcal{M}_k(\Gamma_0(6)) + 4.$$

Therefore, by induction, we obtain an explicit dimension formulae for $\mathcal{M}_k(\Gamma_0(6))$.

Generalising this process to arbitrary subgroups Γ, we note that one of the non-trivial steps is calculating a non-zero cusp form for Γ of sufficiently low weight. To aid us in this regard, we can refer to Martin's list of multiplicative η-quotients [16, p.4852] which construct cusp forms of varying weight and level from the Dedekind eta function $\eta(\tau)$.

We can thus compute the dimension for some further examples of subgroups. We provide three more explicit examples:

For $\mathcal{M}_k(\Gamma_0(9))$, we have

$$\dim \mathcal{M}_k(\Gamma_0(9)) = \begin{cases} 0 & \text{if } k < 0 \text{ or } k \text{ odd} \\ k+1 & \text{otherwise} \end{cases}$$

using the cusp form $\eta^8(3\tau) \in \mathcal{S}_4(\Gamma_0(9))$. For $\mathcal{M}_k(\Gamma_1(2))$, we have

$$\dim \mathcal{M}_k(\Gamma_1(2)) = \begin{cases} 0 & \text{if } k < 0 \text{ or } k \text{ odd} \\ \lfloor \frac{k}{4} \rfloor + 1 & \text{otherwise} \end{cases}$$

using the cusp form $(\eta(\tau)\eta(2\tau))^8 \in \mathcal{S}_8(\Gamma_1(2))$. For $\mathcal{M}_k(\Gamma_1(5)$, we have

$$\dim \mathcal{M}_k(\Gamma_1(5)) = \begin{cases} 0 & \text{if } k < 0 \text{ or } k \text{ odd} \\ k+1 & \text{otherwise} \end{cases}$$

using the cusp form $(\eta(\tau)\eta(5\tau))^4 \in \mathcal{S}_4(\Gamma_1(5))$. Indeed, Martin [16] provides a total of 43 cusp forms all arising from eta quotients of the form $\eta(t_1\tau)^{r_1}\eta(t_2\tau)^{r_2}\ldots\eta(t_s\tau)^{r_s}$, all of which can be used to calculate the dimension for different congruence subgroups of varying level.

To summarise this process for a general congruence subgroup Γ, we do the following:

1. Using η quotients or otherwise, compute a cusp form Ψ of sufficiently low weight l for Γ.
2. Use the valence formula (or alternatively product representations) to prove $\Psi(\tau) \neq 0$ for all $\tau \in \mathcal{H}$.
3. Using the cusp form $\Psi \in \mathcal{S}_l(\Gamma)$, show that there is an isomorphism between $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_{k+l}(\Gamma)$ given by the mapping $f \mapsto \Psi f$.
4. Compute a set of ϵ_∞ linearly independent modular forms using generalised Eisenstein series, where ϵ_∞ denotes the number of cusps of Γ.
5. Use the valence formula and Eisenstein series to compute the exact dimension of $\mathcal{M}_k(\Gamma)$ for weights $k < l$.
6. Apply
$$\dim \mathcal{M}_{k+l}(\Gamma) = \dim \mathcal{S}_{k+l}(\Gamma) + \epsilon_\infty = \dim \mathcal{M}_k(\Gamma) + \epsilon_\infty$$
recursively to obtain a general formula for $\mathcal{M}_k(\Gamma)$ of arbitrary integer weight k.

Asymptotic analysis

We now compare our results obtained with the exact dimension formulae.

For any finite index subgroup Γ, one can consider the compact Riemann surface $X(\Gamma)$ in a topological way as a torus with g holes (or equivalently as a sphere with g handles). We denote the value g as the **genus** of $X(\Gamma)$.

We note the following useful result on computing the genus.

Theorem 35: [6, p. 68] Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Let d be the index $[\mathrm{PSL}_2(\mathbb{Z}) : \bar{\Gamma}]$, let ϵ_2 and ϵ_3 denote the number of elliptic points of $X(\Gamma)$ of order 2 and order 3 respectively, and let ϵ_∞ denote the number of cusps of $X(\Gamma)$. Then

$$g = 1 + \frac{d}{12} - \frac{\epsilon_2}{4} - \frac{\epsilon_3}{3} - \frac{\epsilon_\infty}{2}. \tag{109}$$

Proof sketch: [22, p. 39] Proving the above theorem formally is outside the scope of this project. However, to give an overview, we consider a triangulation of the Riemann surface $X(\Gamma)$ using coset translates of the fundamental domain $\mathcal{F}$ for $\mathrm{SL}_2(\mathbb{Z})$.

Indeed, we define the natural projection map $\pi : \Gamma\backslash\mathcal{H} \to \Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$ which sends all points from the fundamental domain of Γ to the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$. We can triangulate $X(\mathrm{SL}_2(\mathbb{Z}))$ using two triangular faces at the three vertices i, ω and ∞. Thus, the number of faces, edges and vertices for $\mathrm{SL}_2(\mathbb{Z})$ yields

$$F_{\mathrm{SL}_2(\mathbb{Z})} = 2, \qquad E_{\mathrm{SL}_2(\mathbb{Z})} = 3 \qquad V_{\mathrm{SL}_2(\mathbb{Z})} = 3$$

We now pull back this triangulation for $\mathrm{SL}_2(\mathbb{Z})$ using π^{-1} to obtain a triangulation for Γ. Denoting the faces, edges and vertices of this triangulation by F_Γ, E_Γ and V_Γ respectively, we obtain

$$F_\Gamma = 2d,$$
$$E_\Gamma = 3d,$$

$$V_\Gamma = \epsilon_\infty + d - \frac{1}{2}(d-\epsilon_2) + d - \frac{2}{3}(d-\epsilon_3)$$

where the number of vertices is calculated by carefully counting the number of elliptic points $z \in \Gamma\backslash\mathcal{H}$ of order 2 and order 3 such that $f(z) = i$ and $f(z) = \omega$ respectively.

Using the Euler characteristic χ for a Riemann surface, we have

$$V_\Gamma - E_\Gamma + F_\Gamma = \chi = 2 - 2g \tag{110}$$

which therefore yields

$$\begin{aligned} & 2g = 2 - 2d + 3d - \epsilon_\infty - d + \frac{1}{2}(d-\epsilon_2) - d + \frac{2}{3}(d-\epsilon_3) \\ \implies \quad & 2g = 2 + \frac{d}{6} - \frac{\epsilon_2}{2} - \frac{2\epsilon_3}{3} - \epsilon_\infty \\ \implies \quad & g = 1 + \frac{d}{12} - \frac{\epsilon_2}{4} - \frac{\epsilon_3}{3} - \frac{\epsilon_\infty}{2} \end{aligned}$$

which therefore proves the theorem. □

For completeness, we state without proof the dimension formula for arbitrary congruence subgroups Γ.

Theorem 36: [6, p. 88,91] Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and let $k \in \mathbb{Z}$. Let g denote the genus of $X(\Gamma)$, let ϵ_2 and ϵ_3 denote the number of elliptic points of $X(\Gamma)$ of order 2 and order 3 respectively, and let ϵ_∞ denote the number of cusps of $X(\Gamma)$. Then, the dimension of $\mathcal{M}_k(\Gamma)$ can be given as

$$\dim \mathcal{M}_k(\Gamma) = \begin{cases} (k-1)(g-1) + \left\lfloor \frac{k}{4} \right\rfloor \epsilon_2 + \left\lfloor \frac{k}{3} \right\rfloor \epsilon_3 + \frac{k}{2}\epsilon_\infty & \text{if } k \geq 2, k \text{ even} \\ 1 & \text{if } k = 0 \\ (k-1)(g-1) + \left\lfloor \frac{k}{3} \right\rfloor \epsilon_3 + \frac{k}{2}\epsilon_\infty^{\text{reg}} + \frac{k-1}{2}\epsilon_\infty^{\text{irr}} & \text{if } k \geq 3,\ k \text{ odd},\ -I \notin \Gamma \\ 0 & \text{if } k \text{ odd},\ -I \in \Gamma \\ 0 & \text{if } k < 0 \end{cases}$$

Proving the above theorem in its full generality involves a large amount of Riemann surface theory as well as applying the Riemann-Roch theorem. We shall not present any details of the proof here, but simply note that these arguments can all be found in Chapter 3 of Diamond, Shurman [6, p. 65-108], Section 2.6 of Shimura [24, p. 45-50] and Section 2.5 of Miyake [19, p. 57-61] among others.

We note that the dimension is asymptotically linear in k. Indeed, we have:

$$\dim \mathcal{M}_k(\Gamma) \sim k\left(g - 1 + \frac{\epsilon_2}{4} + \frac{\epsilon_3}{3} + \frac{\epsilon_\infty}{2}\right).$$

However, noting the expression for the genus given in equation (109), this yields

$$\dim \mathcal{M}_k(\Gamma) \sim \frac{k[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma]}{12}.$$

Therefore, not only is $k[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma]/12 + 1$ an upper bound for the dimension, but indeed it provides the exact asymptotic behaviour of the dimension as $k \to \infty$. To further investigate the exact discrepancy between the upper bound obtained in the valence formula and the explicit dimension, we calculate

$$\begin{aligned}\frac{k[\mathrm{PSL}_2(\mathbb{Z}):\Gamma]}{12}+1-\dim\mathcal{M}_k(\Gamma)&=\frac{kd}{12}+1-(k-1)(g-1)-\left\lfloor\frac{k}{4}\right\rfloor\epsilon_2-\left\lfloor\frac{k}{3}\right\rfloor\epsilon_3-\frac{k}{2}\epsilon_\infty\\&=g+k\left(\frac{\epsilon_2}{4}+\frac{\epsilon_3}{3}+\frac{\epsilon_\infty}{2}\right)-\left\lfloor\frac{k}{4}\right\rfloor\epsilon_2-\left\lfloor\frac{k}{3}\right\rfloor\epsilon_3-\frac{k}{2}\epsilon_\infty\\&=g+\epsilon_2\left(\frac{k}{4}-\left\lfloor\frac{k}{4}\right\rfloor\right)+\epsilon_3\left(\frac{k}{3}-\left\lfloor\frac{k}{3}\right\rfloor\right)\\&\leq g+\frac{3}{4}\epsilon_2+\frac{2}{3}\epsilon_3\end{aligned}$$

We therefore note that the main discrepancy between the upper bound and the dimension arises from the genus g. This is not too surprising as our methods for calculating a bound on the dimension were purely elementary and did not use any algebro-geometric techniques. This hence demonstrates the limitations of doing the contour method integral as it does not see the correction term involving the genus g.

For the principal congruence subgroups $\Gamma(N)$, we have the following result [6, p. 107]

$$\epsilon_2(\Gamma(N))=\epsilon_2(\Gamma(N))=0\quad\text{for all } N>1$$

which implies the discrepenacy between the upper bound and the dimension is *exactly* the genus g. Thus, the dimension formula simplifies to

$$\dim\mathcal{M}_k(\Gamma(N))=\frac{k[\mathrm{PSL}_2(\mathbb{Z}):\bar{\Gamma}(N)]}{12}+1-g.$$

For $\Gamma_1(N)$, one can also calculate [6, p. 107]

$$\epsilon_2(\Gamma_1(N))=\begin{cases}1 & \text{if } N=2\\0 & \text{if } N>2\end{cases}$$

and

$$\epsilon_3(\Gamma_1(N))=\begin{cases}1 & \text{if } N=3\\0 & \text{if } N=2 \text{ or } N>3\end{cases}.$$

We therefore also obtain the following dimension formula for $\Gamma_1(N)$ when $N>3$

$$\dim\mathcal{M}_k(\Gamma_1(N))=\frac{k[\mathrm{PSL}_2(\mathbb{Z}):\bar{\Gamma}_1(N)]}{12}+1-g.$$

For $\Gamma_0(N)$, one obtains a more involved expression involving the elliptic points of order 2 and 3 in the dimension (see Figure 3.3 and Figure 3.4 at [6, p. 107,108] for the precise dimension formulae). Noting that we do have exact results for the dimension at our disposal, we would nevertheless be interested in congruence subgroups which have dimension exactly the upper bound obtained in valence formula.

Proposition 37: Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. If $\Gamma=\Gamma_0(N)$ where $N\in\{2,3,4,6,8,9,12,16,18\}$ or $\Gamma=\Gamma_1(N)$ where $N\in\{2,3,4,5,6,7,8,9,10,12\}$ or $\Gamma=\Gamma(N)$ where $N\in\{3,4,5\}$, then

$$\dim\mathcal{M}_k(\Gamma)=\left\lfloor\frac{k[\mathrm{PSL}_2(\mathbb{Z}):\bar{\Gamma}(N)]}{12}\right\rfloor+1\qquad\text{for all } k\geq 4, k \text{ even}.$$

One can verify the above claim either by manually calculating the values $d,g,\epsilon_2,\epsilon_3,\epsilon_\infty$ for each congruence subgroup Γ or by computational verification. We note that any congruence subgroup Γ which has equality with the upper bound must have genus $g=0$. The converse, however, is not true, as the discrepancy between the upper bound and the dimension may also involve ϵ_2 and ϵ_3.

Conclusion

In conclusion, we have established in Theorem 26 a general valence formula for any finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$. This generalises the well-known valence formula for $\mathrm{SL}_2(\mathbb{Z})$ [23, p. 85]. Furthermore, our techniques are fully applicable to any finite index subgroup Γ, whereas several other authors [2, p.40] [17, p.32] restrict only to congruence subgroups.

Using the steps mentioned on page 54, one can obtain tight bounds on the dimension for many different finite index subgroups, where the valence formula provides an upper bound whilst constructions from Eisenstein series provide a lower bound. In several cases (for small level N), these bounds can give an explicit dimension formula, as shown for $\mathrm{SL}_2(\mathbb{Z}), \Gamma_0(4), \Gamma_0(6)$ and others. However, even in other cases where explicit formulae are not attainable, we can also establish tight enough bounds which are often good enough for many applications of modular forms.

Comparing our results to those already established using Riemann surface theory [6, p. 88], we note that the upper bound obtained in the valence formula provides an asymptotic estimate for the dimension, noting that further discrepancies arise primarily from the genus.

Acknowledgements

I would like to thank my supervisor, Dr Dirk Basson, for his immense support and helpful suggestions. I would also like to thank the Stellenbosch Mathematics Department for providing this valuable opportunity to learn new mathematics outside the scope of material presented in lectures.

References

[1] Andrianov, A. (2009) *Introduction to Siegel Modular Forms and Dirichlet Series*, Springer-Verlag, New York.

[2] Bruin, P., Dahmen, S. (2016) *Modular Forms*, [online] [Accessed 27 September 2018] Available at: `http://www.few.vu.nl/ sdn249/modularforms16/Notes.pdf`

[3] Bruinier, J.H., van der Geer, G., Harder, G., Zagier, D. (2008) *The 1-2-3 of modular forms* Springer-Verlag, Berlin, Heidelberg.

[4] Conway, J.B. (1986) *Functions of One Complex Variable I*, Springer-Verlag, New York.

[5] Cornell, G., Silverman, J.H., Stevens, G. (1997) *Modular Forms and Fermats Last Theorem*, Springer-Verlag, New York.

[6] Diamond, F., Shurman, J. (2005) *A First Course in Modular Forms*, Springer-Verlag, New York.

[7] Freitag, E. (1990) *Hilbert Modular Forms*, Springer-Verlag, Berlin, Heidelberg.

[8] Gekeler, E. (1988) *On the coefficients of Drinfeld modular forms*, Invent. Math., Vol. 93, No. 3, pp. 667-700.

[9] Jermann, J. (2006) *The valence formula and the space of modular forms*, [online] [Accessed 5 October 2018] Available at: `https://www2.math.ethz.ch/education/bachelor/seminars/ws0607/modular-forms/valence_formula_and_the_space_of_modular_forms.pdf`

[10] Kilford, L.J.P. (2008) *Modular Forms: A Classical and Computational Introduction*, 2nd ed, Imperial College Press.

[11] Knopp, K. (1990) *Theory and Application of Infinite Series*, Dover Publications.

[12] Koblitz, N. (1993) *Introduction to Elliptic Curves and Modular Forms*, 2nd ed. Springer-Verlag New York.

[13] Kulkarni, R.S. (1991) *An arithmetic-geometric method in the study of the subgroups of the modular group.* Amer. J. Math., Vol. 113, No. 6, pp. 1053-1133.

[14] Loeffler, D. (2012) *Modular Forms*, [online] [Accessed 23 September 2018] Available at: `https://www.mathematik.tu-darmstadt.de/media/algebra/homepages/voelz_1/Modular_Forms_Loeffler_2011.pdf`, University of Warwick.

[15] Martin, K. (2011) *Modular Forms Course Notes* [online] [Accessed 27 September 2018] Available at: `http://www2.math.ou.edu/ kmartin/mfs/mfs.pdf`

[16] Martin, Y. (1996) *Multiplicative η-quotients* Trans. Amer. Math. Soc., Vol. 348, No. 12, pp.4825-4856.

[17] Masdeu, M. (2015) *Modular Forms (MA4H9)* [online] [Accessed 27 September 2018] Available at: `http://homepages.warwick.ac.uk/ ~masmat/files/teaching/modforms.pdf`, University of Warwick.

[18] Milne, J.S. (2017) *Modular Functions and Modular Forms* (v1.31) [online] [Accessed 26 September 2018] Available at: `https://www.jmilne.org/math/CourseNotes/MF.pdf`

[19] Miyake, T. (1989) *Modular Forms*, Springer-Verlag Berlin Heidelberg.

[20] Murty, M.R., Dewar, M., Graves, H. (2016) *Problems in the Theory of Modular Forms*, Springer Singapore.

[21] Sarnak, P. (2011) *Randomness in Number Theory*, [online] [Accessed 24 October 2018] Available at: `https://publications.ias.edu/sites/default/files/Mahler%20Lecture%202%20-Randomness%20in%20number%20theory.pdf`, Mahler Lectures

[22] Schultz, D. (2015) *Notes on Modular Forms* [online] [Accessed 13 September 2018] Available at: `https://faculty.math.illinois.edu/ ~schult25/ModFormNotes.pdf`

[23] Serre, J.P. (1973) *A Course in Arithmetic*, Springer-Verlag, New York.

[24] Shimura, G. (1971) *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press, Princeton, N.J.

[25] Silverman, J.H. (1994) *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag, New York.

Appendix

Generators of $\mathrm{SL}_2(\mathbb{Z})$

Let S and T be two matrices in $\mathrm{SL}_2(\mathbb{Z})$ defined by:

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

We shall prove that S and T generate the full modular group $\mathrm{SL}_2(\mathbb{Z})$.

We define $\langle S, T \rangle$ as the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ generated by the matrices S and T. Clearly, as $S, T \in \mathrm{SL}_2(\mathbb{Z})$, we have $\langle S, T \rangle \subseteq \mathrm{SL}_2(\mathbb{Z})$. Now, for each $l \in \mathbb{Z}$, $l \geq 0$, we define

$$A_l = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ |c| = l \right\}.$$

We prove by strong induction that $A_l \subseteq \langle S, T \rangle$ for all $l \geq 0$.

Base case: We first consider $l = 0$. and let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A_0$. Thus, by definition, we have $c = 0$, and thus $ad = 1$, since $\det(\gamma) = 1$. We therefore obtain that either $a = d = 1$ or $a = d = -1$. In the former case, we easily observe that

$$\gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = T^b.$$

In the latter case, noting that $S^2 = -I$, we obtain

$$\gamma = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} = S^2 T^{-b}.$$

Thus, for both cases, we have $\gamma \in \langle S, T \rangle$. Hence $A_0 \subseteq \langle S, T \rangle$.

Induction hypothesis: We now assume for induction that $A_l \subseteq \langle S, T \rangle$ for all $0 \leq l \leq k$ for some $k \geq 0$. Now, let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A_{k+1}$. As $k + 1 > 0$, we have that $|c| > 0$.

Thus, by the division algorithm, we can obtain two integers $q, r \in \mathbb{Z}$ such that $d = qc + r$ where $|r| < c$. Note that

$$\begin{aligned} \gamma T^{-q} S &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} a & -qa + b \\ c & -qc + d \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -qa + b & -a \\ -qc + d & -c \end{pmatrix} = \begin{pmatrix} -qa + b & -a \\ r & -c \end{pmatrix}. \end{aligned}$$

Note that, as $|r| < c$, we have $\gamma T^{-q} S \in \langle S, T \rangle$ by the induction hypothesis. Thus, $\gamma \in \langle S, T \rangle$. We therefore obtain $A_{k+1} \subseteq \langle S, T \rangle$.

Therefore, by induction, we conclude that $A_l \subseteq \langle S, T \rangle$ for all $l \geq 0$. Finally, note that for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, we have $\gamma \in A_{|c|} \subseteq \langle S, T \rangle$, and thus $\langle S, T \rangle = \mathrm{SL}_2(\mathbb{Z})$, which proves that S and T generate the modular group $\mathrm{SL}_2(\mathbb{Z})$. $\square$

Generators of $\Gamma_0(4)$

To calculate a set of generators for $\Gamma_0(4)$, we proceed in a similar manner to the proof for $\mathrm{SL}_2(\mathbb{Z})$. We let $-I, T$ and U be three matrices in $\mathrm{SL}_2(\mathbb{Z})$ defined by

$$-I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}.$$

We shall prove that $-I, T$ and U generate the subgroup $\Gamma_0(4)$. We note that $-I, T, U \in \Gamma_0(4)$, thus $\langle -I, T, U \rangle \subseteq \Gamma_0(4)$.

As with $\mathrm{SL}_2(\mathbb{Z})$, for each $l \in \mathbb{Z}$, $l \geq 0$, we define A_l as

$$A_l = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ |c| = l \right\}$$

We prove by strong induction that $A_l \subseteq \langle S, T \rangle$ for all $l \geq 0$.

Base case: Let $\gamma \in A_0$. By the same argument as with $\mathrm{SL}_2(\mathbb{Z})$, we have either

$$\gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = T^b \quad \text{or} \quad \gamma = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} = -I \cdot T^{-b}$$

Thus $A_0 \subseteq \langle -I, T, U \rangle$.

Induction hypothesis: We now assume for induction that $A_l \subseteq \langle -I, T, U \rangle$ for all $0 \leq l \leq k$.

We now prove that there exists an integer q such that $|d - qc| < |c/2|$. Indeed, apply the division algorithm to obtain two integers q' and r such that $d = q'c + r$ where $0 \leq r < |c|$. If $r < |c/2|$, then we simply let $q = q'$, which gives the desired inequality.

Otherwise, assuming $r \geq |c/2|$, let $q = q' + \mathrm{sgn}(c)$ (where $\mathrm{sgn}(c) = 1$ if $c > 0$ and $\mathrm{sgn}(c) = -1$ if $c < 0$). We therefore have

$$|d - qc| = |d - (q' + \mathrm{sgn}(c))c| = |d - (d - r) - |c|| = |r - |c|| \leq |c/2|$$

as $r \geq |c/2|$. We finally note that equality cannot occur. Indeed, we have $\gcd(c, d) = 1$. As $4|c$, this implies d is odd. Thus $|d - qc|$ is odd, whereas $|c/2|$ is even, thus implying a strict inequality.

We therefore have the existence of an integer q such that $|d - qc| < |c/2|$. We note that

$$\begin{aligned} \gamma T^{-q} U^{\pm 1} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -q \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \pm 4 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a & -qa + b \\ c & -qc + d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \pm 4 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a \pm 4(b - aq) & b - aq \\ c \pm 4(d - cq) & d - cq \end{pmatrix} \end{aligned}$$

Note that $|4(d - cq)| < |2c|$. Therefore, depending on the sign of c, we have that either $|c + 4(d - cq)| < |c|$ or $|c - 4(d - cq)| < |c|$. We can therefore apply the induction hypothesis to obtain $\gamma T^{-q} U^{\pm 1} \in \langle -I, T, U \rangle$ which implies $\gamma \in \langle -I, T, U \rangle$. Therefore, $A_{k+1} \subseteq \langle -I, T, U \rangle$.

Thus, by induction, we have that $A_l \subseteq \langle -I, T, U \rangle$ for all $l \leq 0$. As with $\mathrm{SL}_2(\mathbb{Z})$, we therefore easily conclude that $-I$, T and U generate the congruence subgroup $\Gamma_0(4)$. $\square$

Order of $\mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$

Let N be a positive integer. We prove that the order of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is given by [6, p. 13]

$$|\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})| = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right). \tag{111}$$

Note that, for $N = 1$, the claim is trivial. We now first prove the above claim for prime powers $N = p^e$ where $e \geq 1$. We proceed by induction on e.

Base case: Let $e = 1$. Let $M_2(\mathbb{Z}/p\mathbb{Z})$ denote the set of 2 by 2 matrices of entries in the finite field $\mathbb{Z}/p\mathbb{Z}$.

We define the general linear group over $M_2(\mathbb{Z}/p\mathbb{Z})$ as the set of invertible 2 by 2 matrices over $\mathbb{Z}/p\mathbb{Z}$. As some matrix $A \in M_2(\mathbb{Z}/p\mathbb{Z})$ is invertible if and only if $\det(A) \neq 0$, we therefore have

$$\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{Z}/p\mathbb{Z}) : ad - bc \neq 0 \right\}. \tag{112}$$

Counting the elements in $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$, we consider some $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$. Note that γ is invertible if and only if the set of columns of γ are linearly independent.

For the first column $\begin{pmatrix} a \\ c \end{pmatrix}$, there are no restrictions a priori other than it not being the zero vector. This therefore gives $p^2 - 1$ possible choices for $\begin{pmatrix} a \\ c \end{pmatrix}$. For the second column $\begin{pmatrix} b \\ d \end{pmatrix}$, it must only not be a linear multiple of the first column. As there are p different multiples of the first column, this therefore yields $p^2 - p$ possible choices for $\begin{pmatrix} b \\ d \end{pmatrix}$. Therefore, the total number of elements in $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ can be given as

$$|\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})| = (p^2 - 1)(p^2 - p).$$

We now consider a map $\phi : \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z}) \to (\mathbb{Z}/p\mathbb{Z})^*$ given by the determinant map

$$\phi \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

As the determinant of all matrices in $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ is non-zero, ϕ is well-defined. We also note that ϕ is a homomorphism as the determinant is multiplicative, and that ϕ is surjective by Bézout's identity.

Note that the kernel of ϕ is $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ by definition. Therefore, we have

$$|\mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})| = \frac{|\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})|}{|(\mathbb{Z}/p\mathbb{Z})^*|} = \frac{(p^2 - 1)(p^2 - p)}{(p - 1)} = p^3 - p$$

Induction hypothesis: Let us now assume that we have $|\mathrm{SL}_2(\mathbb{Z}/p^e\mathbb{Z})| = p^{3e} - p^{3e-2}$ for some $e \geq 1$. We consider the reduction map $\pi : \mathrm{SL}_2(\mathbb{Z}/p^{e+1}\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/p^e\mathbb{Z})$ defined by the natural projection of reducing each entry modulo p^e. Since p^e divides p^{e+1}, this is well-defined. We also note that π is surjective as the natural map $\mathrm{SL}_2\mathbb{Z} \to \mathrm{SL}_2(\mathbb{Z}/p^e\mathbb{Z})$ surjects.

We now compute the kernel of ϕ. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/p^{e+1}\mathbb{Z})$. Thus

$$\begin{aligned} \gamma \in \ker \pi \quad &\iff \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{p^e} \\ &\iff \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} kp^e + 1 & lp^e \\ mp^e & np^e + 1 \end{pmatrix} \end{aligned}$$

for some $k, l, m, n \in \mathbb{Z}$. We now count the number of possibilities for k, l, m, n which yield distinct $\gamma \in \mathrm{SL}_2(\mathbb{Z}/p^{e+1}\mathbb{Z})$. As a, b, c, d are all considered modulo p^{e+1}, this implies a priori only p possibilities each for k, l, m, n. We also note, since $\gamma \in \mathrm{SL}_2(\mathbb{Z}/p^{e+1}\mathbb{Z})$, we have

$$\begin{aligned} ad - bc \equiv 1 \pmod{p^{e+1}} \quad &\iff \quad (kp^e + 1)(np^e + 1) - (lp^e)(mp^e) &\equiv 1 \pmod{p^{e+1}} \\ &\iff \quad (kn - lm)p^{2e} + (k+n)p^e + 1 &\equiv 1 \pmod{p^{e+1}} \\ &\iff \quad k + n \equiv 0 \pmod{p} \end{aligned}$$

Thus, for γ to have determinant 1 modulo p^{e+1}, we require $p|(k+n)$. Cleraly, there are exactly p such pairs (as any choice for k yields a unique choice for n (modulo p^{e+1})). Thus, we have p possibilities for (k, n), p possibilities for l, and p possibilities for m, all independent from each other. We thus have

$$|\ker \pi| = p^3. \tag{113}$$

Therefore, by the first isomorphism theorem, we obtain

$$|\mathrm{SL}_2(\mathbb{Z}/p^{e+1}\mathbb{Z})| = |\ker \pi||\mathrm{SL}_2(\mathbb{Z}/p^e\mathbb{Z})| = p^3(p^{3e} - p^{3e-2}) = p^{3(e+1)} - p^{3(e+1)-2}$$

which therefore completes the induction.

To obtain a general expression for an arbitrary positive integer N, we first note that, if R and S are two arbitrary rings, then we have an isomophism between $\mathrm{SL}_2(R \times S)$ and $\mathrm{SL}_2(R) \times \mathrm{SL}_2(S)$ given by

$$\begin{pmatrix} (a_1, a_2) & (b_1, b_2) \\ (c_1, c_2) & (d_1, d_2) \end{pmatrix} \mapsto \left(\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \right)$$

Furthermore, given a prime factorisation for $N = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k}$, then by the Chinese Remainder Theorem, we have [6, p. 401]

$$\mathbb{Z}/N\mathbb{Z} \cong \prod_{i=1}^{k} \mathbb{Z}/p_i^{e_i}\mathbb{Z}.$$

Thus, we have

$$\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) \cong \mathrm{SL}_2\left(\prod_{i=1}^{k} \mathbb{Z}/p_i^{e_i}\mathbb{Z} \right) \cong \prod_{i=1}^{k} \mathrm{SL}_2(\mathbb{Z}/p_i^{e_i}\mathbb{Z}).$$

Taking the orders from both sides, we finally obtain

$$\begin{aligned} |\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})| = \prod_{i=1}^{k} |\mathrm{SL}_2(\mathbb{Z}/p_i^{e_i}\mathbb{Z})| = \prod_{i=1}^{k} (p_i^{3e_i} - p_i^{3e_i - 2}) &= \left(\prod_{i=1}^{k} p_i^{e_i} \right)^3 \prod_{i=1}^{k} \left(1 - \frac{1}{p_i^2} \right) \\ &= N^3 \prod_{p|N} \left(1 - \frac{1}{p^2} \right) \end{aligned}$$

which therefore completes the proof. □

Subgroups of $\mathbb{Z}_2 \times \mathbb{Z}$

We calculate all the possible subgroups of $G = \mathbb{Z}_2 \times \mathbb{Z}$. Let H be a subgroup of G. Note that $H = \{(0, 0\}$ and $H = \{(0, 0), (1, 0)\}$ are two trivial subgroups of G. Otherwise, we may assume H contains some element $(a, b) \in G$ where $b \neq 0$. Thus we also have $(-a, -b) \in H$. We may therefore assume that H contains some element (a, b) where $b > 0$.

We define $h \geq 1$ as the smallest positive integer such that either $(0, h)$ or $(1, h)$ is in H. We now consider various cases:

Case 1: $(0, h) \in H$ and $(1, h) \notin H$. We thus have $\langle (0, h) \rangle \subseteq H$. Now, assume that $(x, qh + r) \in H$, where $q, h, r \in \mathbb{Z}$ where $0 < r < h$. Therefore $(x, qh + r) - q(0, h) \in H$, and thus $(x, r) \in H$. However, this contradicts the definition of h, as $0 < r < h$.

Thus, we have that if $(x, y) \in H$, then h must divide y. Now assume $(1, qh) \in H$. Thus $(1, qh) - (q - 1)(0, h) \in H$ and hence $(1, h) \in H$, which contradicts $(1, h) \notin H$.

Therefore $H = \langle (0, h) \rangle$.

Case 2: $(0, h) \notin H$ and $(1, h) \in H$. We thus have $\langle (1, h) \rangle \subseteq H$. As before, we assume that $(x, qh + r) \in H$, where $q, h, r \in \mathbb{Z}$ and $0 < r < h$. Therefore $(x, qh + r) - q(1, h) \in H$ which implies $(x - q, r) \in H$, therefore contradicting the definition of h.

Thus, we have $(x, y) \in H \implies h|y$. Now assume $(x, qh) \in H$. Thus $(x, qh) - (q - 1)(1, h) \in H$ which implies $(x - q + 1, h) \in H$. Since $(0, h) \notin H$, this implies that if q is odd, then x must be odd. Similarly, if q is even, then x must be even. Thus $H \subseteq \langle (1, h) \rangle$.

We therefore obtain $H = \langle (1, h) \rangle$.

Case 3: $(0, h) \in H$ and $(1, h) \in H$. Once again, assume that $(x, qh + r) \in H$, where $q, h, r \in \mathbb{Z}$ and $0 < r < h$. Therefore $(x, qh + r) - q(0, h) \in H$ and thus $(x, r) \in H$, contradicting the definition of h.

Therefore $(x, y) \in H \implies h|y$, which gives us $\langle (1, h), (0, h) \rangle \subseteq H$ and thus $H = \langle (1, 0), (0, h) \rangle$.

We therefore conclude that the only finite index subgroups of $\mathbb{Z}_2 \times \mathbb{Z}$ are $\langle (0, h) \rangle$, $\langle (1, h) \rangle$, and $\langle (1, 0), (0, h) \rangle$ where $h \geq 1$. $\square$