

Modular Forms

Harry Armitage

March 8, 2021

Contents

0	Introduction	2
1	Modular forms on $SL_2(\mathbb{Z})$	4
2	Hecke Operators and L-functions	14
3	Modular Forms on Congruence Subgroups of $SL_2(\mathbb{Z})$	29
4	Modular Curves and Riemann Surfaces	38

0 Introduction

Notation. We will write $\mathbb{H} := \{\tau \in \mathbb{C} : \Im(\tau) > 0\}$ for the complex upper half plane. This is acted on by two groups:

$$GL_2(\mathbb{R})^+ = \{g \in GL_2(\mathbb{R}) : \det(g) > 0\} \geq SL_2(\mathbb{Z}) = \{g \in GL_2(\mathbb{Z}) : \det(g) = 1\}$$

Lemma 0.1. $GL_2(\mathbb{R})^+$ acts on \mathbb{H} by Möbius transformations. This action is transitive.

Proof. Let $\tau \in \mathbb{H}$, $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R})^+$. We then write $g\tau = \frac{a\tau+b}{c\tau+d}$. This is an action on \mathbb{C} by theory about Möbius transformations. To see that $g\tau \in \mathbb{H}$, we check:

$$\Im(g\tau) = \frac{1}{2}(g\tau - \overline{g\tau}) = \det(g) \frac{\Im(\tau)}{|c\tau + d|^2}$$

Now for transitivity, let $\tau = x + iy \in \mathbb{H}$. Then $\tau = \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} i$. □

Definition 0.2. Let $k \in \mathbb{Z}$, and $f : \mathbb{H} \rightarrow \mathbb{C} \cup \{\infty\}$, and let $g \in GL_2(\mathbb{R})^+$. Then we define $f|_k[g] : \mathbb{H} \rightarrow \mathbb{C} \cup \{\infty\}$ by the formula

$$f|_k[g](\tau) = f(g\tau) \det(g)^{k-1} j(g, \tau)^{-k}$$

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

Lemma 0.3. This defines a right actions of $GL_2(\mathbb{R})^+$ on the set of functions $f : \mathbb{H} \rightarrow \mathbb{C} \cup \{\infty\}$.

Proof. Suppose $g, h \in GL_2(\mathbb{R})^+$. We need to show that $f|_k[gh] = (f|_k[g])|_k[h]$.

$$\begin{aligned} RHS(\tau) &= f|_k[g](h\tau) \det(h)^{k-1} j(h, \tau)^{-k} \\ &= f(gh\tau) \det(g)^{k-1} j(g, h\tau)^{-k} j(h, \tau)^{-k} \det(h)^{k-1} \\ LHS(\tau) &= f(gh\tau) \det(gh)^{k-1} j(gh, \tau) \end{aligned}$$

So we need to check that $j(g, h\tau)j(h, \tau) = j(gh, \tau)$.

Note that if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $g \begin{pmatrix} \tau \\ 1 \end{pmatrix} = \begin{pmatrix} a\tau + b \\ c\tau + d \end{pmatrix} = j(g, \tau) \begin{pmatrix} g\tau \\ 1 \end{pmatrix}$.

So $gh \begin{pmatrix} \tau \\ 1 \end{pmatrix} = j(gh, \tau) \begin{pmatrix} gh\tau \\ 1 \end{pmatrix} = gj(h, \tau) \begin{pmatrix} h\tau \\ 1 \end{pmatrix} = j(h, \tau)j(g, h\tau) \begin{pmatrix} gh\tau \\ 1 \end{pmatrix}$. □

Definition 0.4. Let $k \in \mathbb{Z}$, and let $\Gamma \leq SL_2(\mathbb{Z})$ be a finite index subgroup. Then a meromorphic function $f : \mathbb{H} \rightarrow \mathbb{C} \cup \{\infty\}$ is called a weakly modular function of weight k and level Γ if it satisfies $\forall \gamma \in \Gamma, f|_k[\gamma] = f$.

Motivating Examples

1. Modular forms were first studied in the context of elliptic functions. Suppose that E is an elliptic curve over \mathbb{C} , and let ω be a non-vanishing holomorphic differential on E . Then there's a unique holomorphic isomorphism of Riemann surfaces

$$\mathbb{C}/\Lambda \xrightarrow[\psi]{} E(\mathbb{C})$$

such that $\psi^*(\omega) = dz$. Here $\Lambda \subset \mathbb{C}$ is a lattice.

E can be defined by the equation $y^2 = x^3 - 60G_4(\Lambda)x - 140G_6(\Lambda)$ where $G_k(\Lambda) = \sum_{\omega \in \Lambda \setminus \{0\}} \omega^{-k}$.

This is absolutely convergent provided $k \geq 4$.

If $\tau \in \mathbb{H}$, then we can write $\Lambda_\tau = \mathbb{Z}\tau \oplus \mathbb{Z}$. This is a lattice, and the functions $G_k(\tau) = G_k(\Lambda_\tau)$ are examples of modular forms.

2. If $f : \mathbb{H} \rightarrow \mathbb{C}$ is a modular form, then f has a Fourier expansion $f(\tau) = \sum_{n \geq 0} a_n e^{2\pi i n \tau / h}$ for some natural number h , and complex numbers a_n . These Fourier coefficients often carry useful arithmetic information.

For example, consider $\theta(\tau) = \sum_{n \in \mathbb{Z}} e^{\pi i n^2 \tau}$. If $k \geq 2$ is an even integer, then θ^{2k} is a modular form of weight k . Its Fourier expansion is $\theta^{2k}(\tau) = \sum_{n \geq 1} r_{2k}(n) e^{\pi i n \tau}$ where $r_{2k}(n)$ is the number of ways of writing $n = x_1^2 + \dots + x_{2k}^2$, where $x_i \in \mathbb{Z}$.

By relating θ^{2k} to other modular forms with known Fourier series, we can then get information about the numbers $r_{2k}(n)$. For example, $r_4(n) = 8 \sum_{d|n, 4 \nmid d} d$.

3. Recall the Riemann zeta function $\zeta(s) = \sum_{n \geq 1} n^{-s}$. This function has some important properties:
 - a) It has a meromorphic continuation to all of \mathbb{C} .
 - b) It has a functional equation relating $\zeta(s)$ and $\zeta(1-s)$.
 - c) It has a representation as an Euler product $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$.

Any series $L(s) = \sum_{n \geq 1} a_n n^{-s}$ with $a_n \in \mathbb{C}$ which has properties analogous to these is called an L -function.

For example, if $N \in \mathbb{N}$ and $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$ is a character, we can define the Dirichlet L -function $L(\chi, s) = \sum_{(n, N)=1} \chi(n \bmod N) n^{-s}$. These functions can be used to prove Dirichlet's theorem on primes in arithmetic progression.

Modular forms can be used to construct L -functions with these properties. To find the right modular forms, we need to introduce Hecke operators.

4. The Langlands programme predicts relations between objects occurring in number theory and modular forms. This includes as a special case the Shimura-Taniyama-Weil conjecture, otherwise known as the modularity theorem. This asserts a bijection between elliptic curves over \mathbb{Q} up to isogeny and certain modular forms, given by $(L\text{-function of elliptic curve}) = (L\text{-function of modular form})$.

1 Modular forms on $SL_2(\mathbb{Z})$

Recall the definition, for $f : \mathbb{H} \rightarrow \mathbb{C}$, $k \in \mathbb{Z}$, $g \in GL_2(\mathbb{R})^+$, we have

$$f|_k[g](\tau) = \det(g)^{k-1} f(g\tau) j(g, \tau)^{-k}$$

We said f is *weakly modular of weight k and level $SL_2(\mathbb{Z})$* if f is meromorphic on \mathbb{H} and, for all $\gamma \in SL_2(\mathbb{Z})$, $f|_k[\gamma] = f$.

Note that $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL_2(\mathbb{Z})$ satisfies $f|_k[T](\tau) = f(\tau+1)$. So if f is a weakly modular function, then we can define a new function

$$\tilde{f} : \{q \in \mathbb{C} : 0 < |q| < 1\} \rightarrow \mathbb{C}; e^{2\pi i \tau} \mapsto f(\tau)$$

This function \tilde{f} is meromorphic, since f is.

Definition 1.1. We say that the weakly modular function f is:

- meromorphic at ∞ if \tilde{f} is meromorphic at 0.
- holomorphic at ∞ if \tilde{f} is holomorphic at 0.
- vanishes at ∞ if \tilde{f} is holomorphic and vanishes at 0.

If f is meromorphic at ∞ then \tilde{f} has a Laurent expansion $\tilde{f}(q) = \sum_{n \in \mathbb{Z}} a_n q^n$ valid in some region $\{0 < |q| < \varepsilon\}$, where $a_n \in \mathbb{C}$ and $a_n = 0$ if $n < 0$ and $|n|$ is sufficiently large.

We get a formula $f(\tau) = \sum_{n \in \mathbb{Z}} a_n q^n$ where $q = e^{2\pi i \tau}$. This is valid in some region $\{\tau \in \mathbb{H} : \Im \tau > R\}$, and is called the q -expansion of f . Then f is holomorphic at ∞ if and only if $a_n = 0$ when $n < 0$, and $f(\infty) = a_0$.

Definition 1.2. Let f be a weakly modular function of weight k and level $SL_2(\mathbb{Z})$. We say that f is

- a **modular function** if f is meromorphic at ∞ .
- a **modular form** if f is holomorphic in \mathbb{H} and holomorphic at ∞ .
- a **cuspidal modular form** if f is a modular form vanishing at ∞ .

all with weight k and level $SL_2(\mathbb{Z})$.

We write $M_k(SL_2(\mathbb{Z}))$ for the \mathbb{C} -vector space of modular forms of weight k and level $SL_2(\mathbb{Z})$. We write $S_k(SL_2(\mathbb{Z}))$ for the subspace of cuspidal modular forms.

Examples. If $\tau \in \mathbb{H}$, then $\Lambda_\tau = \mathbb{Z}\tau \oplus \mathbb{Z}$. if $k \in \mathbb{Z}$, then we can define $G_k(\tau) = \sum_{\omega \in \Lambda_\tau \setminus \{0\}} \omega^{-k}$.

If $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$, then $\Lambda_{\gamma\tau} = \mathbb{Z} \begin{pmatrix} a\tau+b \\ c\tau+d \end{pmatrix} \oplus \mathbb{Z} = j(\gamma, \tau)^{-1} \mathbb{Z}(a\tau+b) \oplus \mathbb{Z}(c\tau+d) = j(\gamma, \tau)^{-1} \Lambda_\tau$.

Finally, we find $G_k|_k[\gamma](\tau) = G_k(\gamma\tau) j(\gamma, \tau)^{-k} = \sum_{\omega \in \Lambda_{\gamma\tau} \setminus \{0\}} (\omega j(\gamma, \tau))^{-k} = \sum_{\omega \in \Lambda_\tau \setminus \{0\}} \omega^{-k} = G_k(\tau)$.

Proposition 1.3. Suppose $k \geq 4$ and k is even. Then $G_k(\tau)$ converges absolutely and uniformly on compact subsets of \mathbb{H} . Moreover, $G_k(\tau)$ is holomorphic at ∞ and $G_k(\infty) = 2\zeta(k)$. In particular, $G_k \in M_k(SL_2(\mathbb{Z}))$.

Remark. We have $-I = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in SL_2(\mathbb{Z})$ and $f|_k[-I] = f \cdot (-1)^k$, so if k were odd then $f \equiv 0$, and hence $M_k(SL_2(\mathbb{Z})) = 0$ when k is odd.

Proof. Fix $A \geq 1$. Define $\Omega_A = \{\tau \in \mathbb{H} : |\Re(\tau)| \leq A, \Im(\tau) \geq \frac{1}{A}\}$. We'll show uniform convergence of G_k in Ω_A . Note that if $\tau \in \Omega_A$, then for any $x \in \mathbb{R}$, $|\tau + x| \geq \frac{1}{A}$, and $|\tau + x| \geq \frac{1}{2}|x|$ if $|x| \geq 2A$. Hence $|\tau + x| \geq \sup(1/A, 1/2A^2|x|) \geq \frac{1}{2A^2} \sup(1, |x|)$ for any $x \in \mathbb{R}$.

If $\tau \in \Omega_A$, then:

$$\begin{aligned} \sum_{(m,n) \in \mathbb{Z}^2 \setminus \{0\}} |m\tau + n|^{-k} &= \sum_{(m,n)} |m|^{-k} |\tau + n/m|^{-k} \\ &\leq \sum_{(m,n)} \frac{|m|^{-k}}{(2A)^{-k}} \sum (1, |n/m|)^{-k} \\ &= \sum_{(m,n)} (2A)^k \sup(|m|^{-k}, |n|^{-k}) \\ &= \sum_{r \in \mathbb{N}} (2A)^k r^k 8r = (2A)^k 8\zeta(k-1) \end{aligned}$$

This shows absolute and uniform convergence.

To show that G_k is holomorphic at ∞ and $G_k(\infty) = 2\zeta(k)$, it's enough to show that

$$\lim_{\tau \in \Omega_1, \Im \tau \rightarrow \infty} G_k(\tau) = 2\zeta(k)$$

This limit equals $\sum_{(m,n)} \lim_{\Im \tau \rightarrow \infty} (m\tau + n)^{-k} = \sum_{n \in \mathbb{Z} \setminus \{0\}} n^{-k} = 2\zeta(k)$, as all terms with $m \neq 0$ vanish. \square

G_k is an example of an *Eisenstein series*.

Definition 1.4. We define the **normalised Eisenstein series** $E_k(\tau) = \frac{1}{2\zeta(k)} G_k(\tau) = 1 + \sum_{n \geq 1} a_n q^n$. We'll see that the a_n are rational numbers of bounded denominators.

Remark. If $f \in M_k(SL_2(\mathbb{Z}))$ and $g \in M_\ell(SL_2(\mathbb{Z}))$, then $fg \in M_{k+\ell}(SL_2(\mathbb{Z}))$. So $E_4^3, E_6^2 \in M_{12}(SL_2(\mathbb{Z}))$, and $E_4^3(\infty) = E_6^2(\infty)$, so $\Delta = \frac{E_4^3 - E_6^2}{1728} \in S_{12}(SL_2(\mathbb{Z}))$. We'll see shortly that $\Delta = \sum_{n \geq 1} b_n q^n$ where $b_1 = 1, b_n \in \mathbb{Z}$ for all $n \geq 1$.

We now study a fundamental domain for the action of $SL_2(\mathbb{Z})$ on \mathbb{H} . We will write $\Gamma(1) = SL_2(\mathbb{Z})$, and $\overline{\Gamma(1)} = SL_2(\mathbb{Z})/\langle -I \rangle$. This will make sense later.

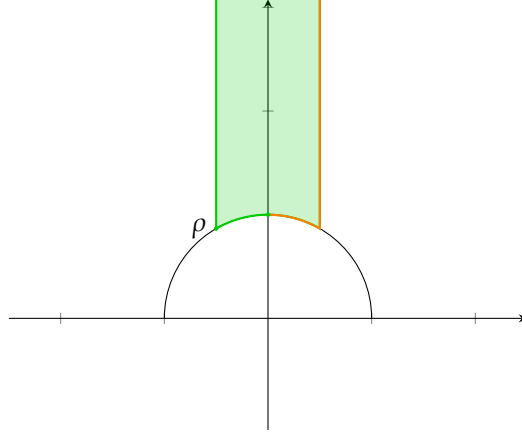
We write

$$\mathcal{F} = \{\tau \in \mathbb{H} : -\frac{1}{2} \leq \Re \tau \leq \frac{1}{2}, |\tau| \geq 1\}$$

and

$$\mathcal{F}' = \{\tau \in \mathcal{F} : \Re \tau < 1/2, |\tau| = 1 \implies \Re \tau \leq 0\}$$

In the following diagram, \mathcal{F} is all of the green + orange regions, whilst \mathcal{F}' is just the green area. We also define $\rho := \exp(2\pi i/3)$



We have elements $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \Gamma(1)$.

Proposition 1.5. \mathcal{F} is a fundamental domain for the action of $\overline{\Gamma(1)}$ on \mathbb{H} . More precisely, if $\tau \in \mathbb{H}$ there is $\gamma \in \overline{\Gamma(1)}$ such that $\gamma\tau \in \mathcal{F}$. If $\gamma\tau \in \mathcal{F}^\circ$, then γ is unique. Moreover, each $\tau \in \mathbb{H}$ is $\overline{\Gamma(1)}$ -conjugate to exactly one element of \mathcal{F}' .

Proof. We first prove that any $\tau \in \mathbb{H}$ is $\overline{\Gamma(1)}$ -conjugate to an element of \mathcal{F} . We proved earlier that if $\tau \in \mathbb{H}$ and $\gamma \in \Gamma(1)$, then $\Im \gamma(\tau) = \Im(\tau)/|c\tau + d|^2$.

If $\tau \in \mathbb{H}$, then $\Lambda_\tau = \mathbb{Z}_\tau \oplus \mathbb{Z}$ is a lattice. So as $(c, d) \in \mathbb{Z}^2 \setminus \{0\}$, the numbers $|c\tau + d|$ achieve a minimum. Consequently, the numbers $\Im \gamma(\tau)$ for $\gamma \in \Gamma(1)$ achieve a maximum. So wlog we may assume $\Im(\tau) \geq \Im(\gamma\tau)$ for all $\gamma \in \Gamma(1)$. Also wlog we may take $-\frac{1}{2} \leq \Re(\tau) \leq \frac{1}{2}$.

We then claim that these properties are sufficient for $\tau \in \mathcal{F}$. It is sufficient to show that $|\tau| \geq 1$. We have $\Im(S\tau) = \Im(\tau)/|\tau|^2 \leq \Im(\tau)$, and hence $|\tau|^2 \geq 1$, so we are done. \square

We slightly strengthen this with the following proposition:

Proposition 1.6.

1. For all $\tau \in \mathbb{H}$, there is a unique $\gamma \in \overline{\Gamma(1)}$ such that $\gamma\tau \in \mathcal{F}'$.
2. If $\tau \in \mathcal{F}'$, then $\text{Stab}_{\overline{\Gamma(1)}}(\tau) = \{I\}$, except $\text{Stab}_{\overline{\Gamma(1)}}(i) = \{I, S\}$ and $\text{Stab}_{\overline{\Gamma(1)}}(\rho) = \{1, ST, (ST)^2\}$.
3. $\overline{\Gamma(1)}$ is generated by S and T .

Proof. To prove the first two parts, it's enough to show that:

a) For all $\tau \in \mathbb{H}$, there is $\gamma \in \overline{\Gamma(1)}$ such that $\gamma\tau \in \mathcal{F}'$.

b) For all $\tau, \tau' \in \mathcal{F}'$ and $\gamma \in \overline{\Gamma(1)}$, $\gamma\tau' = \tau \implies \tau' = \tau$ and either $\begin{cases} \gamma = 1 \\ \tau = i, \gamma = S \\ \tau = \rho, \gamma = ST, (ST)^2 \end{cases}$.

a) was done above. For b), take $\tau, \tau' \in \mathcal{F}'$ such that $\tau' = \gamma\tau$. We have $\Im(\gamma\tau) = \Im(\tau)/|c\tau + d|^2$ where $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Without loss of generality, we have $\Im(\tau') = \Im(\gamma\tau) \geq \Im(\tau) \implies |c\tau + d| \leq 1$.

So $\Im(\tau) \geq \sqrt{3}/2 \implies |c\tau + d| \geq c\sqrt{3}/2$, and so $|c| \leq 1$, so we can assume $c = 1$ or 0 (if -1 , just multiply by $-I$, since we are in $\overline{\Gamma(1)} = \Gamma(1)/\langle -I \rangle$). We then split into cases:

1. $c = 0, \gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$. This forces $\gamma = I, \tau = \tau'$.
2. $c = 1, \gamma = \begin{pmatrix} a & b \\ 1 & d \end{pmatrix}$. Now $|\tau + d| \leq 1$. Then $\tau \in \mathcal{F}' \implies$ either $d = 0, |\tau| = 1$, or $d = 1, \tau = \rho$.

In the first case, $\gamma = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$, and so $\gamma\tau = a - \frac{1}{\tau}$. We have $\Re(\tau), \Re(\Gamma(\tau)) = a - \Re(\tau)$ both in $[-1/2, 0]$. The only possibilities are $\Re(\tau) = -\frac{1}{2}, a = -1, \tau = \rho, \gamma = (ST)^2$ and $\Re(\tau) = 0, a = 0, \tau = i, \gamma = S$.

In the second case, $d = 1, \tau = \rho, \gamma = \begin{pmatrix} a & b \\ 1 & 1 \end{pmatrix}$. Then $\gamma\rho = \frac{a\rho+b}{\rho+1}$. We have $\rho^2 + \rho + 1 = 0, \rho^2 = \rho - 1$, so $\gamma\rho = -\rho(a\rho + b) = -a\rho^{-1} - b\rho$.

We know that $|\rho + 1| = |\tau + d| = 1$, so $\Im(\gamma\rho) = \Im(\rho)/|\rho + 1| = \Im(\rho)$. So $\gamma\rho = \rho$, as ρ is the unique element of \mathcal{F}' of smallest imaginary part, and hence $\rho = -a\rho^{-1} - b\rho \implies a = 0, b = -1$, and so $\gamma = ST$.

For part 3., let's take $G = \langle S, T \rangle$. For all $\tau \in \mathbb{H}$, there is $\gamma \in G$ with $\gamma\tau \in \mathcal{F}$. Why? Without loss of generality, we can assume that, for all $\gamma \in G, \Im(\gamma\tau) \leq \Im(\tau)$, and moreover that $-\frac{1}{2} \leq \Re(\tau) \leq \frac{1}{2}$.

This implies that $\tau \in \mathcal{F}$, as $\Im(S\tau) = \Im(\tau)/|\tau|^2 \leq \Im(\tau) \implies |\tau| \geq 1$.

Choose $\tau \in \mathcal{F}^\circ$. Choose $\gamma \in \overline{\Gamma(1)}$. We'll show that $\gamma \in G$. Note that $\gamma\tau \in \mathbb{H}$, so there is $\delta \in G$ such that $\delta\gamma\tau \in \mathcal{F}$, so $\delta\gamma\tau \in \mathcal{F}^\circ$ and $\delta\gamma = I$, so $\delta = \gamma^{-1} \in G$. \square

If $P \in \overline{\Gamma(1)} \backslash \mathbb{H}$ (since $\overline{\Gamma(1)}$ acts on the left on \mathbb{H} , this is a left quotient. P is a $\overline{\Gamma(1)}$ -orbit, i.e. can be represented as $\overline{\Gamma(1)} \cdot \tau$ for some $\tau \in \mathbb{H}$), then we define $e_P = |\text{Stab}_{\overline{\Gamma(1)}}(\tau)|$.

We've just shown that $e_P = 1$ except for $\begin{cases} e_{\overline{\Gamma(1)} \cdot \rho} = e_\rho = 3 \\ e_{\overline{\Gamma(1)} \cdot i} = e_i = 2 \end{cases}$.

Suppose that f is a modular function of weight k and level $SL_2(\mathbb{Z})$. Then we define $v_P(f)$ to be the order of f at τ (where τ is a representative for P).

Note that this is independent of the specific choice of representative τ as, for any $\gamma \in \Gamma(1)$ we have $f(\gamma\tau)j(\gamma, \tau)^{-k} = f(\tau)$, and $j(\gamma, \tau)$ is holomorphic and non-vanishing.

We define $v_\infty(f) = \inf\{n \in \mathbb{Z} : a_n \neq 0\}$, where $f(\tau) = \sum_{n \in \mathbb{Z}} a_n q^n$ is the q -expansion of f . Equivalently, this is the order of \tilde{f} at $q = 0$.

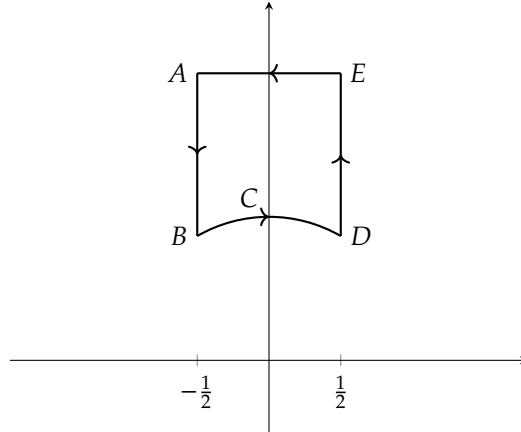
Theorem 1.7. Let f be a modular function of weight k and level $SL_2(\mathbb{Z})$. Assume that $f \neq 0$. Then:

$$v_\infty(f) + \sum_{P \in \overline{\Gamma(1)} \backslash \mathbb{H}} \frac{1}{e_P} v_P(f) = \frac{k}{12}$$

Proof. Let $U \subseteq \mathbb{C}$ be an open subset, and $\gamma \subseteq U$ a positively oriented simple closed contour, and $f : U \rightarrow \mathbb{C}$ a meromorphic function with no zeros or poles on γ . Then $\frac{1}{2\pi i} \oint_{\gamma} \frac{df}{f} = \sum_{\tau \in \text{Int}(\gamma)} v_{\tau}(f)$ - this is the argument principle.

Let's first prove the theorem assuming that f has no zeros or poles on the boundary of \mathcal{F} . Since f is meromorphic at infinity, there exists a $R > 0$ such that f has no zeros or poles on in $\{\tau \in \mathbb{H} : \Im(\tau) \geq R\}$.

We consider the contour $\gamma = ABCDEA$:



where $A = -\frac{1}{2} + iR$, $B = \rho$, $C = i$, $D = \rho + 1$, $E = \frac{1}{2} + iR$.

The argument principle gives $\frac{1}{2\pi i} \oint_{\gamma} \frac{df}{f} = \sum_{\tau \in \text{Int}(\mathcal{F})} v_{\tau}(f)$.

We can break up the integral into the different segments AB, BC, CD, DE, and EA, and make some observations:

- $f(\tau) = f(\tau + 1)$, so $\int_A^B \frac{df}{f} = \int_E^D \frac{df}{f} = -\int_D^E \frac{df}{f}$, so these paths cancel.
- The image of the path EA under the map $\tau \mapsto e^{2\pi i \tau}$ is a negatively oriented circle c going around $q = 0$, so $\frac{1}{2\pi i} \int_E^A \frac{df}{f} = \frac{1}{2\pi i} \oint_c \frac{d\tilde{f}}{\tilde{f}} = -v_0(\tilde{f}) = -v_{\infty}(f)$.
- The path from CD is the image of the path CB under S . So $\frac{1}{2\pi i} \int_D^C \frac{df}{f} = \frac{1}{2\pi i} \int_B^C \frac{d(f \circ S)}{(f \circ S)}$. We have $f(S\tau) = f(\tau)\tau^k$, so $\frac{d(f \circ S)}{f \circ S} = \frac{k d\tau}{\tau} + \frac{df}{f}$.

Hence this integral is $\frac{1}{2\pi i} \int_B^C \frac{k}{\tau} d\tau + \int_B^C \frac{df}{f}$, and so we have:

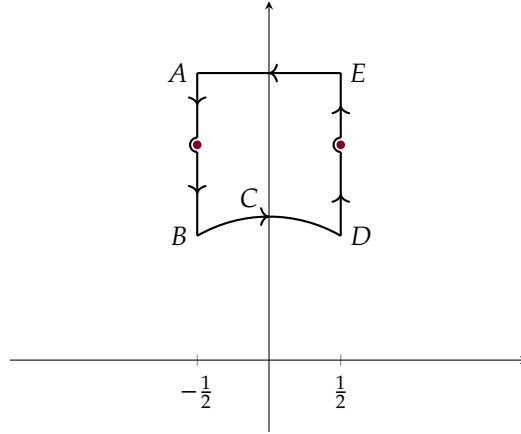
$$\frac{1}{2\pi i} \int_B^C \frac{df}{f} + \frac{1}{2\pi i} \int_C^D \frac{df}{f} = \frac{1}{2\pi i} \int_C^B \frac{k}{\tau} d\tau = \frac{k}{12}$$

Putting this all together, we have:

$$\frac{k}{12} - v_{\infty}(f) = \sum_{\tau \in \text{Int}(\mathcal{F})} v_{\tau}(f)$$

Since we're assuming all zeros and poles are in the interior and so have $e_p = 1$, adding in the ϵ 's for the result in the theorem doesn't change anything.

If there are zeros or poles on the boundary of \mathcal{F} , then we need a modified contour. First suppose that f has a zero or pole on the lines AB and DE , but nowhere else. Then we use the contour γ' :



where the small semicircles are chosen so that they avoid all zeros or poles of f , noting that the zeros and poles of a meromorphic function are isolated, and so that AB is mapped to ED by T , in order to still have $\int_B^A \frac{df}{f} + \int_D^E \frac{df}{f} = 0$. The rest of the proof goes through as before. We can make a similar modification if f has a zero/pole on BC .

The remaining case is when f has a zero or pole at ρ or i . In this case, we use the following observation: let $g : U \rightarrow \mathbb{C}$ be a meromorphic function defined in an open neighbourhood of $z = 0$.

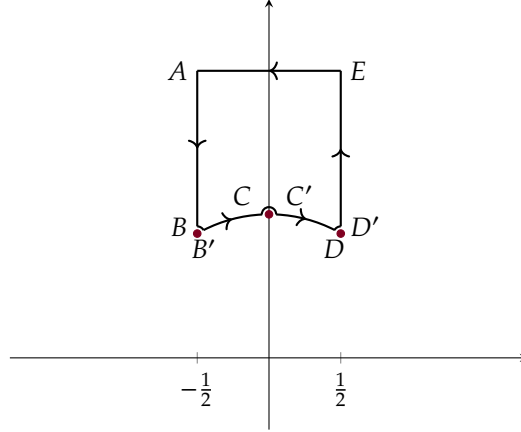
We consider the paths $\gamma_\epsilon : [0, 1] \rightarrow U$ given by $\gamma_\epsilon(t) = \epsilon e^{2\pi i(\theta_0 + t\theta)}$. Then:

$$\lim_{\epsilon \rightarrow 0} \frac{1}{2\pi i} \int_{\gamma_\epsilon} \frac{dg}{g} = \frac{\theta}{2\pi} v_0(g)$$

To show this, write $g(z) = z^n h(z)$ where $n = v_0(g)$ and $h(z)$ is holomorphic and non-vanishing at 0. Then

$$\frac{1}{2\pi i} \int_{\gamma_\epsilon} \frac{dg}{g} = \frac{1}{2\pi i} \int_{\gamma_\epsilon} \frac{ndz}{z} + \frac{1}{2\pi i} \int_{\gamma_\epsilon} \frac{dh}{h} \rightarrow \frac{\theta}{2\pi} + 0$$

Now suppose that f has zeros or poles at ρ or i , and at no other points on the boundary of \mathcal{F} . We consider a family of contours γ_ϵ given by replacing γ at B , C , and D by small arcs of radius ϵ .



Then the argument principle gives:

$$\frac{1}{2\pi i} \left[\int_A^B + \int_B^{B'} + \dots + \int_E^A \frac{df}{f} \right] = \sum_{\tau \in \text{Int}(\mathcal{F})} v_\tau(f)$$

It's still the case that $\frac{1}{2\pi i} \int_E^A \frac{df}{f} = -v_\infty(f)$, and that the paths AB and $D'E$ cancel. It's also still the case that $\frac{1}{2\pi i} \left[\int_{B'}^C + \int_{C'}^D \frac{df}{f} \right] = \frac{\alpha k}{2\pi}$, where α is the angle swept out by CB' , which tends to $k/12$ as $\varepsilon \rightarrow 0$.

We need to understand the remaining terms given by the paths BB' , CC' , DD' . Using our previous observation, we see that $\lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi i} \int_B^{B'} \frac{df}{f} = -\frac{1}{6}v_\rho(f)$. Similarly, we have $\lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi i} \int_C^{C'} \frac{df}{f} = -\frac{1}{2}v_i(f)$, $\lim_{\varepsilon \rightarrow 0} \frac{1}{2\pi i} \int_D^{D'} \frac{df}{f} = -\frac{1}{6}v_\rho(f)$.

We finally obtain an identity:

$$v_\infty(f) + \frac{1}{3}v_\rho(f) + \frac{1}{2}v_i(f) + \sum_{\tau \in \text{Int}(\mathcal{F})} v_\tau(f) = \frac{k}{12}$$

giving the result. □

Let's now apply this to some examples. Take $k = 4$, $f = E_4 \in M_4(SL_2(\mathbb{Z}))$. We get:

$$v_\infty(E_4) + \sum_{P \in \overline{\Gamma(1)} \setminus \mathbb{H}} \frac{1}{e_P} v_P(E_4) = \frac{1}{3}$$

and so $v_\rho(E_4) = 1$ and $v_P(E_4) \neq 0$ for $P \neq \overline{\Gamma(1)} \cdot \rho$. i.e. E_4 has a simple zero at ρ and no other zeros in \mathcal{F}' .

Now take $k = 6$, $f = E_6$. We get $LHS = \frac{1}{2}$, and so $v_i(E_6) = 1$, $v_P(E_6) = 0$ for all $P \neq \overline{\Gamma(1)} \cdot i$, i.e. E_6 has a simple zero at i and no other zeros in \mathcal{F}' .

We defined $\Delta = (E_4^3 - E_6^2)/1728 \in S_{12}(SL_2(\mathbb{Z}))$. Then $\Delta(i) = E_4(i)^3/1728 \neq 0$, and so Δ is actually a non-zero cuspidal modular form. We apply our formula to Δ , using that it is non-zero, and

get

$$v_\infty(\Delta) + \sum_{P \in \overline{\Gamma(1)}} \frac{1}{e_P} v_P(\Delta) = 1$$

We know Δ is cuspidal so $v_\infty(\Delta) \geq 1$, hence $v_\infty(\Delta) = 1$ and Δ is non-vanishing in \mathbb{H} .

Theorem 1.8. *Let $k \in \mathbb{Z}$ be an even integer. Then:*

1. *If $k < 0$ or $k = 2$, then $M_k(SL_2(\mathbb{Z})) = 0$. Moreover, $M_0(SL_2(\mathbb{Z})) = \mathbb{C}$ (identified with the constant functions).*
2. *If $4 \leq k \leq 10$ or $k = 14$, then $M_k(SL_2(\mathbb{Z})) = \mathbb{C} \cdot E_k$*
3. *If $k \geq 0$, then multiplication by Δ induces an isomorphism $M_k(SL_2(\mathbb{Z})) \xrightarrow{\sim} S_{k+12}(SL_2(\mathbb{Z}))$.*

Proof. We use the formula $v_\infty(f) + \sum_P \frac{1}{e_P} v_P(f) = \frac{k}{12}$, valid for any non-zero $f \in M_k(SL_2(\mathbb{Z}))$. If $k < 0$, $LHS \geq 0$, $RHS < 0$ and so there are no such f .

If $k = 2$, $RHS = 1/6$, $LHS = a + b/2 + c/3$ where $a, b, c \in \mathbb{Z}_{\geq 0}$.

Suppose $f \in M_0(SL_2(\mathbb{Z}))$ and f is not a scalar. Then there is $\lambda \in \mathbb{C}$ such that $f - \lambda$ is cuspidal and non-zero, so $v_\infty(f - \lambda) \geq 1$. But then $LHS > 0$, $RHS = 0$, and we have a contradiction. Hence $M_0(SL_2(\mathbb{Z})) = \mathbb{C}$.

Now suppose $f \in M_k(SL_2(\mathbb{Z}))$ and either $4 \leq k \leq 10$ or $k = 14$. Then there is $\lambda \in \mathbb{C}$ such that $f - \lambda E_k \in S_k(SL_2(\mathbb{Z}))$. If $f - \lambda E_k \neq 0$, we get $v_\infty(f - \lambda E_k) + \sum_P \frac{1}{e_P} v_P(f - \lambda E_k) = \frac{k}{12}$. If $k < 12$, then $RHS < 1$ and $LHS \geq 1$. If $k = 14$, then we will use part 3 and 1 to show $S_{14}(SL_2(\mathbb{Z})) = 0$, so $M_{14}(SL_2(\mathbb{Z})) = \mathbb{C}E_{14}$.

To prove the final part of the theorem, consider the described map $\times \Delta : M_k(SL_2(\mathbb{Z})) \rightarrow S_{k+12}(SL_2(\mathbb{Z}))$. It's injective as Δ is non-vanishing in \mathbb{H} , so $f\Delta = g\Delta \implies f = g$. It's surjective as Δ is non-vanishing and $v_\infty(\Delta) = 1$. This means that, if $f \in S_{k+12}(SL_2(\mathbb{Z}))$ then $v_\infty(f/\Delta) = v_\infty(f) - 1 \geq 0$, and so $f/\Delta \in M_k(SL_2(\mathbb{Z}))$.

□

Corollary 1.9. *For any $k \in \mathbb{Z}$, $k \geq 0$ even, we have*

$$\dim_{\mathbb{C}} M_k(SL_2(\mathbb{Z})) = \begin{cases} \lfloor \frac{k}{12} \rfloor + 1 & k \not\equiv 2 \pmod{12} \\ \lfloor \frac{k}{12} \rfloor & k \equiv 2 \pmod{12} \end{cases}$$

Proof. The theorem shows this is true for $0 \leq k \leq 14$. We have $M_k(SL_2(\mathbb{Z})) = \mathbb{C}E_k \oplus S_k(SL_2(\mathbb{Z}))$, just by subtracting a scalar multiple of E_k to get a cusp form, and so $\dim_{\mathbb{C}} M_{k+12}(SL_2(\mathbb{Z})) = 1 + \dim_{\mathbb{C}} M_k(SL_2(\mathbb{Z}))$, and the result follows by induction. □

Corollary 1.10. *Let $k \geq 0$ be even. Then $M_k(SL_2(\mathbb{Z}))$ is spanned as a \mathbb{C} -vector space by the elements $E_4^a E_6^b$ where $a, b \in \mathbb{Z}_{\geq 0}$ and $4a + 6b = k$.*

Proof. This holds when $k \leq 10$. We'll now show that if the corollary holds for k , then it holds for $k + 12$. This will give the general case by induction.

Choose $a, b \in \mathbb{Z}_{\geq 0}$ such that $4a + 6b = k + 12$. Then $E_4^a E_6^b \in M_{k+12}$ with leading term 1 in its q -expansion, so we have $M_{k+12}(SL_2(\mathbb{Z})) = S_{k+12}(SL_2(\mathbb{Z})) \oplus \mathbb{C}E_4^a E_6^b = \Delta M_k(SL_2(\mathbb{Z})) \oplus \mathbb{C}E_4^a E_6^b$.

Note that $\Delta = (E_4^3 - E_6^2)/1728$, so the result follows. \square

Definition 1.11. We define $j : \mathbb{H} \rightarrow \mathbb{C}$ by the formula $j(\tau) = E_4^3(\tau)/\Delta(\tau)$. This is a modular function of weight 0 and level $SL_2(\mathbb{Z})$.

If $\tau \in \mathbb{H}$, then $j(\tau)$ is the j -invariant of the elliptic curve $E_\tau = \mathbb{C}/\Lambda_\tau$.

Theorem 1.12.

1. j is holomorphic in \mathbb{H} and $v_\infty(j) = -1$.
2. j gives a bijection $\overline{\Gamma(1)} \backslash \mathbb{H} \xrightarrow{\sim} \mathbb{C}$.
3. Every modular function of weight 0 and level $SL_2(\mathbb{Z})$ is a rational function of j .

Later, we'll give $\overline{\Gamma(1)} \backslash \mathbb{H} \sqcup \{\infty\}$ the structure of a compact Riemann surface. Part 3 of the above theorem says that j gives an isomorphism from this surface to the Riemann sphere.

Proof.

1. $j = E_4^3/\Delta$ is holomorphic in \mathbb{H} as Δ is non-vanishing in \mathbb{H} . We have $v_\infty(E_k^3) = 0$ and $v_\infty(\Delta) = 1$, so $v_\infty(E_4^3/\Delta) = 0 - 1 = -1$.
2. We need to show that for all $z \in \mathbb{C}$, there is a unique $\tau \in \mathbb{H}$ modulo $\overline{\Gamma(1)}$ such that $j(\tau) = z$, or equivalently, $E_4^3(\tau) - z\Delta(\tau) = 0$.

We have $v_\infty(E_4^3 - z\Delta) + \sum_P \frac{1}{e_P} v_P(E_4^3 - z\Delta) = 1$. The first term is 0, since the leading term of $E_4^3 - z\Delta$ is 1. Then we have $a + b/2 + c/3 = 1$ for $a, b, c \in \mathbb{Z}_{\geq 0}$. We see that there's exactly one P such that $v_P(E_4^3 - z\Delta) > 0$, and so have the desired statement.

3. Let f be a non-zero modular function of weight 0 and level $SL_2(\mathbb{Z})$. Then we can multiply f by terms of the form $j - j(\tau_0)$ to conclude that f is holomorphic in \mathbb{H} . Then there exists $n \geq 0$ such that $\Delta^n f$ is holomorphic at ∞ , and so $\Delta^n f \in M_{12n}(SL_2(\mathbb{Z}))$.

To prove the theorem, it's enough to show that, if $4a + 6b = 12$, then $E_4^a E_6^b / \Delta^n$ is a rational function of j .

Note that, if $4a + 6b = 12n$, then $2a + 3b = 6n$, so we can write $a = 3p, b = 2q$ for some $p, q \in \mathbb{Z}_{\geq 0}$, and then $p + q = n$. Hence $E_4^a E_6^b / \Delta^n = (E_4^3/\Delta)^p (E_6^2/\Delta)^q = j^p (E_6^2/\Delta)^q$. So it remains to show that E_6^2/Δ is a rational function of j .

By definition, $\Delta = (E_4^3 - E_6^2)/1728$, and so $E_6^2/\Delta = E_4^3/\Delta - 1728 = j - 1728$.

\square

Proposition 1.13. Let $k \geq 4$ be an even integer. Then the q -expansion of G_k is

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

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

Proof. We use the identity $\pi \cot(\pi\tau) = \frac{1}{\tau} + \sum_{n=1}^{\infty} \left(\frac{1}{\tau-n} + \frac{1}{\tau+n} \right)$ as holomorphic functions in \mathbb{H} . The RHS is absolutely convergent and uniformly convergent in compact subsets.

Note that $\pi \cot(\pi\tau) = \pi i \frac{e^{i\pi\tau} + e^{-i\pi\tau}}{e^{i\pi\tau} - e^{-i\pi\tau}} = \pi i \frac{q+1}{q-1} = -\pi i(1+q)(1+q+q^2+\dots)$.

So $-\pi i - 2\pi i \sum_{n \geq 1} q^n = \frac{1}{\tau} + \sum_{n=1}^{\infty} \left(\frac{1}{\tau-n} + \frac{1}{\tau+n} \right)$. Differentiating $k-1$ times gives:

$$-2\pi i \sum_{n \geq 1} (2\pi i)^{k-1} q^n = (-1)^{k-1} (k-1)! \left[\frac{1}{\tau^k} + \sum_{n=1}^{\infty} \left(\frac{1}{(\tau-n)^k} + \frac{1}{(\tau+n)^k} \right) \right] = (-1)^{k-1} (k-1)! \sum_{n \in \mathbb{Z}} (\tau+n)^{-k}$$

Hence, after rearranging, we have:

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

Applying this to $G_k(\tau)$, we have:

$$\begin{aligned} G_k(\tau) &= \sum_{(m,n) \in \mathbb{Z}^2 \setminus \{0\}} (m\tau + n)^{-k} = 2\zeta(k) + 2 \sum_{m \geq 1} \sum_{n \in \mathbb{Z}} (m\tau + n)^{-k} \\ &= 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k-1)!} \sum_{m,n \geq 1} n^{k-1} q^{nm} \\ &= 2\zeta(k) + 2 \frac{(2\pi i)^k}{(k-1)!} \sum_{n \geq 1} \sigma_{k-1}(n) q^n \end{aligned}$$

□

Corollary 1.14. $E_k(\tau)$ has q -expansion:

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

In particular, the coefficients are rationals, and integers when $k = 4$ or 6 , in which case:

$$E_4(\tau) = 1 + 240 \sum_{n \geq 1} \sigma_3(n) q^n; E_6(\tau) = 1 - 504 \sum_{n \geq 1} \sigma_5(n) q^n$$

Proof. Fact: when $k \in 2\mathbb{Z}_{>0}$, $\pi^k / \zeta(k)$ is a rational number. In particular, $\zeta(4) = \pi^4/90$, $\zeta(6) = \pi^6/945$. This gives the result. □

Proposition 1.15. The q -expansion of Δ is $q + \sum_{n \geq 1} a_n q^n$ where $a_n \in \mathbb{Z}$ for all $n \geq 2$. The q -expansion of j is $q^{-1} + \sum_{n \geq 0} b_n q^n$ where $b_n \in \mathbb{Z}$ for all $n \geq 0$.

Proof. Since $j = E_4^3 / \Delta$, it's enough to show that Δ has the claimed properties.

By definition, $\Delta = (E_4^3 - E_6^2) / 1728$. Write $E_4 = 1 + 240U$, $E_6 = 1 - 504V$, where U, V are the sum parts of **1.14**.

Then, expanding, $\Delta = \frac{5}{12}U + \frac{7}{12}V + P(U, V) = \frac{5}{12}(U - V) + (V + P(U, V))$ where $P \in \mathbb{Z}[U, V]$

So we need to show that $\sigma_3(n) \equiv \sigma_5(n) \pmod{12}$ for all $n \in \mathbb{N}$. It would be enough to show that $n^3 \equiv n^5 \pmod{3}$ and $\pmod{4}$ for all $n \in \mathbb{N}$, which is true. This shows that $b_n \in \mathbb{Z} \forall n \geq 2$. It remains to show that the leading term of Δ is q . It's $\frac{3 \cdot 240 + 2 \cdot 504}{1728} = 1$. □

Proposition 1.16. *Let $k \geq 0$ be an even integer. Then there is a basis f_1, \dots, f_n for the space $S_k(SL_2(\mathbb{Z}))$ such that:*

- a) *If $f_i = \sum_{n \geq 1} a_n(f_i)q^n$, then $a_n(f_i) \in \mathbb{Z}$ for all $n \in \mathbb{N}$, $i = 1, \dots, N$.*
- b) *We have $a_n(f_i) = \delta_{in}$ if $1 \leq i, n \leq N$.*

Proof. We can assume $S_k(SL_2(\mathbb{Z})) \neq 0$. We write $k = 12a + d$, where $4 \leq d \leq 14$. We know $\dim_{\mathbb{C}} S_k(SL_2(\mathbb{Z})) = N = \lfloor k/12 \rfloor$ if $k \not\equiv 2 \pmod{12}$, otherwise $\lfloor k/12 \rfloor - 1$. Hence $a = N$.

We then write $d = 4A + 6B$ for some $A, B \in \mathbb{Z}_{\geq 0}$. We then introduce the modular forms $g_i = \Delta^i E_4^A E_6^B E_6^{2(N-i)}$ for $i = 1, \dots, N$.

The g_i are modular forms of weight $12i + 4A + 6B + 12(N - i) = 12N + d = k$.

The leading term in the q -expansion of g_i is $q^i \cdot 1^A \cdot 1^B \cdot 1^{2(N-i)} = q^i$.

All the numbers $a_n(g_i)$ for $n \in \mathbb{N}$, $i = 1, \dots, N$ are integers.

So we have $g_1 = q + a_2(g_1)q^2 + \dots$, $g_2 = 0 + q^2 + a_3(g_2)q^3 + \dots$, and so on.

Hence we can perform row reduction over \mathbb{Z} to find modular forms f_1, \dots, f_n as in the statement of the proposition. \square

2 Hecke Operators and L-functions

Hecke operators are endomorphisms of $S_k(SL_2(\mathbb{Z}))$ and $M_k(SL_2(\mathbb{Z}))$. They're indexed by $n \in \mathbb{N}$. They exist for rather abstract reasons.

Definition 2.1. *Let G be a group, which may be infinite, and let $H \leq G$ be a subgroup. We say (G, H) is a **Hecke pair** if for all $g \in G$, the set $H \backslash HgH$ is finite, where $HgH = \{h_1gh_2 : h_1, h_2 \in H\}$, acted on the right by H .*

Lemma 2.2. *Let $H \leq G$ be groups. Then if $g \in G$, there is a bijection:*

$$H \cap g^{-1}Hg \backslash H \rightarrow H \backslash HgH$$

In particular (G, H) is a Hecke pair if and only if, for all $g \in G$, $[H : H \cap g^{-1}Hg] < \infty$.

Proof. We first write down the map $H \rightarrow H \backslash HgH$, given by $h \mapsto Hgh$. This map is surjective. We need to show that, if $h_1, h_2 \in H$, then $Hgh_1 = Hgh_2 \iff \exists x \in H \cap g^{-1}Hg$ such that $h_1 = xh_2$.

But $Hgh_1 = Hgh_2 \iff \exists x \in H$ such that $xgh_1 = gh_2$, i.e. $h_1 = g^{-1}x^{-1}gh_2 \iff \exists y \in H \cap g^{-1}Hg$ such that $h_1 = yh_2$.

To see this, note that if $h_1 = g^{-1}x^{-1}gh_2$, then $g^{-1}x^{-1}g \in H$, and $gg^{-1}x^{-1}gg^{-1} = x^{-1} \in H$. So if $y = g^{-1}x^{-1}g$, then $y \in H \cap g^{-1}Hg$, and $h_1 = yh_2$. On the other hand, if there exists $y \in H \cap g^{-1}Hg$ such that $h_1 = yh_2$, then we can write $y = g^{-1}x^{-1}g$, where $x \in H$, and then $h_1 = g^{-1}x^{-1}gh_2$. \square

Definition 2.3. *Let (G, H) be a Hecke pair. Then the Hecke algebra $\mathcal{H}(G, H)$ is defined to be the set of all functions $f : G \rightarrow \mathbb{C}$ satisfying the conditions:*

- 1. $\forall g \in G, h_1, h_2 \in H$, we have $f(h_1gh_2) = f(g)$.
- 2. f is nonzero on only finitely many elements H -double cosets.

If $f_1, f_2 \in \mathcal{H}(G, H)$, then we define their product:

$$(f_1 \cdot f_2)(g) = \sum_i f_1(g g_i^{-1}) f_2(g_i)$$

where g_i is any set of representatives for the decomposition $G = \sqcup_i H g_i$.

Remarks.

1. We can generalise this to the context where G is a locally compact topological group.
2. There is another natural definition for the product. The one we choose is well-adapted to right actions.

Lemma 2.4. *The multiplication in $\mathcal{H}(G, H)$ is well-defined and associative. Moreover, $[H]$ is the unit element.*

Notation. If $X \subset G$, then $[X] : G \rightarrow G$ is its characteristic function. So if $g \in G$, then $[H g H] \in \mathcal{H}(G, H)$.

Proof. By definition, $(f_1 \cdot f_2)(g) = \sum_i f_1(g g_i^{-1}) f_2(g_i)$. We first check that this sum is finite. Note that we can write $f_1(g) = \sum_{x_i} \lambda_i [H x_i H]$. Since (G, H) is a Hecke pair, we can write $[H x_i H] = \sum_j [H y_{ij}]$, hence $f_1 = \sum_{i,j} \lambda_i [H y_{ij}]$, where the sum is finite.

Similarly, we can write $f_2 = \sum_{k,\ell} \mu_k [H z_{k\ell}]$, where the sum is finite.

So:

$$(f_1 \cdot f_2)(g) = \sum_{i,j,k,\ell,r} \lambda_i \mu_k [H y_{ij}](g g_r^{-1}) [H z_{k\ell}](g_r) \quad (*)$$

We see this sum is finite since the non-zero terms have $g_r \in H z_{k\ell}$, i.e. $H z_{k\ell} = H g_r$.

We next check that $f_1 \cdot f_2$ is independent of the choice of g_i . Any other choice is given by $h_i g_i$ for some elements $h_i \in H$.

Then $(f_1 \cdot f_2)(g) = \sum_i f_1(g (h_i g_i)^{-1}) f_2(h_i g_i) = \sum_i f_1(g g_i^{-1} h_i^{-1}) f_2(h_i g_i) = \sum_i f_1(g g_i^{-1}) f_2(g_i)$.

We next check that $f_1 \cdot f_2$ is H -biinvariant. If $g \in G$, $h_1, h_2 \in H$, then:

$$(f_1 \cdot f_2)(h_1 g h_2) = \sum_i f_1(h_1 g h_2 g_i^{-1}) f_2(g_i) = \sum_i f_1(g (g_i h_2^{-1})^{-1}) f_2(g_i h_2^{-1})$$

Since $G = \sqcup_i H g_i$, $G = \sqcup_i H g_i h - 2^{-1}$, this is equal to $(f_1 \cdot f_2)(g)$

We next check that $f_1 \cdot f_2$ is supported on finitely many H -double cosets. It is enough to check that $f_1 \cdot f_2$ is supported on finitely many right H -cosets.

Using $(*)$, we see that if $(f_1 \cdot f_2)(g) \neq 0$, then there are i, j, k, ℓ, r such that $g_r \in H z_{k\ell}$ and $g g_r^{-1} \in H y_{ij}$, so there are i, j, k, ℓ such that $g \in H y_{ij} H z_{k\ell}$.

Since (G, H) is a Hecke pair, $H y_{ij} H z_{k\ell}$ is a union of finitely many H -right cosets. So we've shown that $f_1 \cdot f_2 \in \mathcal{H}(G, H)$.

We next show associativity. If $f_1, f_2, f_3 \in \mathcal{H}(G, H)$, then:

$$\begin{aligned} (f_1 \cdot (f_2 \cdot f_3))(g) &= \sum_i f_1(gg_i^{-1})(f_2 \cdot f_3)(g_i) = \sum_{i,j} f_1(gg_i^{-1})f_2(g_ig_j^{-1})f_3(g_j) \\ ((f_1 \cdot f_2) \cdot f_3)(g) &= \sum_j (f_1 \cdot f_2)(gg_j^{-1})f_3(g_j) = \sum_{i,j} f_1(gg_j^{-1}g_jg_i^{-1})f_2(g_ig_j^{-1})f_3(g_j) \end{aligned}$$

which are equal. Note that $G = \sqcup_i Hg_i \implies G = \sqcup_i Hg_ig_j^{-1}$.

It remains to show that $[H]$ is the unit in $\mathcal{H}(G, H)$. We need to check just that $[H] \cdot f = f$ for all $f \in \mathcal{H}(G, H)$.

$$([H] \cdot f)(g) = \sum_i [H](gg_i^{-1})f(g_i)$$

Now note that $[H](gg_i^{-1}) \neq 0 \iff gg_i^{-1} \in H \iff Hg = Hg_i$, so this is equal to $f(g)$. \square

Definition 2.5. Let V be a \mathbb{C} -vector space on which G acts on the right by linear maps. The subspace of H -invariants is $V^H := \{v \in V : \forall h \in H, vh = v\}$.

Proposition 2.6. Let V be a \mathbb{C} -vector space on which G acts on the right by \mathbb{C} -linear maps. Let (G, H) be a Hecke pair. Then V^H is a right $\mathcal{H}(G, H)$ -module under the action $v \in V, f \in \mathcal{H}(G, H)$ gives

$$v \cdot f = \sum_i f(g_i)(v \cdot g_i)$$

where $G = \sqcup_i Hg_i$.

Proof. We first check $v \cdot f$ is a well defined element of V . The sum is finite, as $f(g_i)$ is non-zero for only finitely many elements g_i . If we choose different coset representative h_ig_i with $h_i \in H$, then

$$\sum_i f(h_ig_i)(v \cdot h_ig_i) = \sum_i f(g_i)(v \cdot g_i)$$

as $v \in V^H$.

We next check that $v \cdot f$ is in V^H . If $h \in H$, then:

$$(v \cdot f) \cdot h = \sum_i f(g_i)v g_i h = \sum_i f(g_i h)v g_i h = v \cdot f$$

since $g_i h$ is also a transversal of G by H .

We next need to check that, if $v \in V^H$ and $f_1, f_2 \in \mathcal{H}(G, H)$, then $v \cdot (f_1 \cdot f_2) = (v \cdot f_1) \cdot f_2$. We compute:

$$\begin{aligned} LHS &= \sum_i (f_1 f_2)(g_i)(v \cdot g_i) \\ &= \sum_i \sum_j f_1(g_ig_j^{-1})f_2(g_j)(v \cdot g_i) \\ RHS &= \sum_j f_2(g_j)((v \cdot f_1) \cdot g_j) \\ &= \sum_j \sum_i f_2(g_j)f_1(g_ig_j^{-1})(v \cdot g_ig_j^{-1}g_j) \end{aligned}$$

using that, for fixed j , $g_i g_j^{-1}$ is also a transversal. \square

Now fix $k \in \mathbb{Z}$. Let $V_k = \{f : \mathbb{H} \rightarrow \mathbb{C} \text{ meromorphic}\}$, equipped with the weight k action of $GL_2(\mathbb{Q})^+ = GL_2(\mathbb{Q}) \cap GL_2(\mathbb{R})^+$.

Lemma 2.7. $(GL_2(\mathbb{Q})^+, \Gamma(1))$ is a Hecke pair.

Proof. If $N \in \mathbb{N}$, then we define $\Gamma(N) = \ker(SL_2(\mathbb{Z}) \rightarrow SL_2(\mathbb{Z}/N\mathbb{Z}))$. This is a finite index subgroup of $\Gamma(1)$. We need to check that, if $g \in GL_2(\mathbb{Q})^+$, then $[\Gamma(1) : \Gamma(1) \cap g^{-1}\Gamma(1)g] < \infty$. Fix $g \in GL_2(\mathbb{Q})^+$. We can find $N \in \mathbb{N}$ such that Ng and $Ng^{-1} \in M_2(\mathbb{Z})$. Then:

$$g\Gamma(N^2)g^{-1} \subset g(1 + N^2M_2(\mathbb{Z}))g^{-1} = 1 + (Ng)M_2(\mathbb{Z})(Ng^{-1}) \subset M_2(\mathbb{Z})$$

So $g\Gamma(N^2)g^{-1} \leq \Gamma(1)$, and hence $\Gamma(N^2) \leq g^{-1}\Gamma(1)g \cap \Gamma(1)$. \square

Hence $V_k^{\Gamma(1)}$, the space of weakly modular functions of weight k and level $\Gamma(1)$, is a $\mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma(1))$ -module.

Definition 2.8. For $n \in \mathbb{N}$, write T_n for the endomorphism of $V_k^{\Gamma(1)}$ induced by the element $[X_n] \in \mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma(1))$, where $X_n = \{g \in M_2(\mathbb{Z}) : \det(g) = n\}$.

Lemma 2.9. The set $\Gamma(1) \setminus X_n$ is finite, and a set of representatives is:

$$\left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, d \in \mathbb{N}, ad = n, b \in \mathbb{Z}, 0 \leq b < d \right\}$$

Proof. If $\alpha \in X_n$, then $\mathbb{Z}^2 \alpha \leq \mathbb{Z}^2$ of index $\det(\alpha) = n$. There's a map:

$$\Gamma(1) \setminus X_n \rightarrow L_n := \{\Lambda \leq \mathbb{Z}^2 : [\mathbb{Z}^2 : \Lambda] = n\}$$

given by $\alpha \mapsto \mathbb{Z}^2 \alpha$. This map factors through the subgroup $\Gamma(1)$. We claim this map is bijective. It's surjective, since if $\Lambda \leq \mathbb{Z}^2$ has index n , then we can find $u, v \in \Lambda$ such that $\Lambda = \mathbb{Z}u \oplus \mathbb{Z}v$, and then $\Lambda = \mathbb{Z}^2 \alpha$, where $\alpha = \begin{pmatrix} u & v \\ v & v \end{pmatrix}$. After possibly swapping u, v , we can assume that $\det(\alpha) = n$.

It's injective since if $\alpha, \beta \in X_n$ and $\mathbb{Z}^2 \alpha = \mathbb{Z}^2 \beta$, then $\mathbb{Z}^2 = \mathbb{Z}^2 \beta \alpha^{-1}$. This forces $\beta \alpha^{-1} = \gamma \in SL_2(\mathbb{Z})$, and so $\beta \equiv \alpha$ modulo $\Gamma(1)$.

It remains to find a suitable α for each choice of $\Lambda \leq \mathbb{Z}^2$ of index n .

Take e_1, e_2 to be the standard basis of \mathbb{Z}^2 . Then $\mathbb{Z}e_2 \cap \Lambda \leq \mathbb{Z}e_2$ has finite index, say d . Then we have a short exact sequence:

$$0 \rightarrow \mathbb{Z}e_2/(\mathbb{Z}e_2 \cap \Lambda) \rightarrow \mathbb{Z}^2/\Lambda \rightarrow \mathbb{Z}^2/(\Lambda + \mathbb{Z}e_2) \rightarrow 0$$

Suppose $\mathbb{Z}^2/(\Lambda + \mathbb{Z}e_2) \cong \mathbb{Z}e_1/\mathbb{Z}e_1 \cap (\Lambda + \mathbb{Z}e_2)$ has order a . Then $ad = n$.

Note that $d \in \mathbb{N}$ is the least d such that $de_2 \in \Lambda$, and $a \in \mathbb{N}$ is the least a such that $ae_1 \in \Lambda + \mathbb{Z}e_2$. $a \in \mathbb{N}$ is the least a such that $ae_1 \in \Lambda + \mathbb{Z}e_2$. Equivalently, the least a such that there exists $b \in \mathbb{Z}$ with $ae_1 + be_2 \in \Lambda$. We can fix a choice of b by requiring $0 \leq b < d$. There's a unique choice of b with this property since, if $ae_1 + be_2 \in \Lambda$ and $ae_1 + ce_2 \in \Lambda$ with $c > b$, then $(c - b)e_2 \in \Lambda$ which would be a contradiction if $c - b < d$.

Λ is thus generated by (a, b) and $(0, d)$, and so $\Lambda = \mathbb{Z}^2 \alpha$ where $\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$. \square

Proposition 2.10. Let $n \in \mathbb{N}$ and f be a modular function of weight k and level $SL_2(\mathbb{Z})$ with q -expansion $f(\tau) = \sum_{m \in \mathbb{Z}} a_m q^m$. Then $T_n f$ is also a modular function with q -expansion $\sum_{m \in \mathbb{Z}} b_m q^m$, where $b_m = \sum_{a|(m,n)} a^{k-1} a_{mn/a^2}$.

Proof. By definition, $T_n f = f \cdot [X_n] = \sum_{g \in \Gamma(1) \setminus X_n} f|_k[g]$. More explicitly:

$$\begin{aligned} (T_n f)(\tau) &= \sum_{ad=n, 0 \leq b < d} f\left(\frac{a\tau + b}{d}\right) d^{-k} n^{k-1} \\ &= n^{k-1} \sum_{m \in \mathbb{Z}} \sum_{ad=n} a_m \sum_{0 \leq b < d} e^{2\pi i m \left(\frac{a\tau + b}{d}\right)} \\ &= n^{k-1} \sum_{m \in \mathbb{Z}} \sum_{ad=n} a_m e^{2\pi i \left(\frac{am\tau}{d}\right)} \sum_{0 \leq b < d} e^{2\pi i mb/d} \end{aligned}$$

Note that $\sum_{0 \leq b < d} e^{2\pi i mb/d} = \begin{cases} d & d|m \\ 0 & d \nmid m \end{cases}$. So we have:

$$\begin{aligned} (T_n f)(\tau) &= n^{k-1} \sum_{m \in \mathbb{Z}} \sum_{ad=n, d|m} d^{1-k} a_m e^{2\pi i am\tau/d} \\ &= \sum_{ad=n} a^{k-1} \sum_{m \in \mathbb{Z}} a_{dm} e^{2\pi i am\tau} \\ &= \sum_{m \in \mathbb{Z}} q^m \left(\sum_{a|(m,n)} a_{mn/a^2} a^{k-1} \right) \end{aligned}$$

It remains to show that $T_n f$ is a modular function. We need to check that $\sum_{a|(m,n)} a^{k-1} a_{nm/a^2} = 0$ when $m < 0$ and $|m| \gg 0$.

Call $a_{nm/a^2} = b_m$. Then suppose that there is $r \in \mathbb{N}$ such that $a_m = 0$ whenever $m < -r$. If $b_m \neq 0$, then there is $a|(n, m)$ such that $a_{nm/a^2} \neq 0$, so $nm/a^2 \geq -r$, so $m \geq -ra^2/n$. Since $a|n$, we must have $a \leq n$, and hence $m \geq -rn$. \square

Corollary 2.11. T_n preserves the spaces $M_k(SL_2(\mathbb{Z}))$ and $S_k(SL_2(\mathbb{Z}))$.

Proof. If f is a modular form then f is holomorphic in \mathbb{H} and $a_m = 0$ when $m < 0$. We know:

$$T_n(f) = \sum_{\substack{ad=n \\ 0 \leq b < d}} f|_k \left[\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \right]$$

And this is certainly holomorphic in \mathbb{H} .

We need to check $b_m = 0$ if $m < 0$. But b_m is a sum of terms $a^{k-1} a_{mn/a^2}$ which are 0 if $m < 0$.

If f is also cuspidal, then $a_0 = 0$, so $b_0 = \sum_{a|n} a^{k-1} a_0 = \sigma_{k-1}(n) a_0 = 0$. \square

Corollary 2.12. If f is a modular function, then $a_0(T_n f) = \sigma_{k-1}(n) a_0(f)$, and $a_1(T_n f) = a_n(f)$.

Lemma 2.13. Let (G, H) be a Hecke pair. Let $f_1 = \sum_i \lambda_i [Hx_i]$ and $f_2 = \sum_j \mu_j [Hy_j]$ be elements of $\mathcal{H}(G, H)$.

Then $f_1 \cdot f_2 = \sum_{i,j} \lambda_i \mu_j [Hx_i y_j]$.

Proof.

$$(f_1 \cdot f_2)(g) = \sum_k f_1(g g_k^{-1}) f_2(g_k) = \sum_{i,j,k} \lambda_i [Hx_i](g g_k^{-1}) \mu_j [Hy_j](g_k)$$

where $G = \bigsqcup_k H g_k$.

For each j , there is a unique $k(j)$ such that $g_{k(j)} \in Hy_j$. Then write $g_{k(j)} = h_j y_j$.

Then $(f_1 \cdot f_2)(g) = \sum_{i,j} \lambda_i \mu_j [Hx_i](g y_j^{-1} h_j^{-1}) = \sum_{i,j} \lambda_i \mu_j [Hx_i](g y_j^{-1}) = \sum_{i,j} \lambda_i \mu_j [Hx_i y_j](g)$. \square

Proposition 2.14. 1. If $n, m \in \mathbb{N}$ and $(n, m) = 1$, then $T_n \circ T_m = T_{nm}$.

2. Let $p, n \in \mathbb{N}$ with p prime. Then $T_{p^n} \circ T_p = T_{p^{n+1}} + p^{k-1} T_{p^{n-1}}$.

Proof. Let $X_+ = \{g \in M_2(\mathbb{Z}) : \det(g) > 0\} = \bigsqcup_{n \in \mathbb{N}} X_n \subseteq GL_2(\mathbb{Q})^+$.

Let $\mathcal{L}_+ = \{\Lambda \subseteq \mathbb{Z}^2 \text{ of finite index}\} = \bigsqcup_{n \in \mathbb{N}} \mathcal{L}_n$.

Last time we showed there is a bijection $SL_2(\mathbb{Z}) \backslash X_n \rightarrow \mathcal{L}_n$ given by $\alpha \mapsto \mathbb{Z}^2 \alpha$

Hence there's a bijection $SL_2(\mathbb{Z}) \backslash X_+ \rightarrow \mathcal{L}_+$.

If $f \in \mathcal{H}(GL_2(\mathbb{Q})^+, SL_2(\mathbb{Z}))$ is supported in X_+ , then we can identify f with a function $\varphi_f : \mathcal{L}_+ \rightarrow \mathbb{C}$ given by $\varphi_f(\mathbb{Z}^2 \alpha) = f(\alpha)$.

Now fix $n, m \in \mathbb{N}$. We can write $[X_n] = \sum_i [SL_2(\mathbb{Z}) x_i]$ and $[X_m] = \sum_j [SL_2(\mathbb{Z}) y_j]$.

Then $[X_n][X_m] = \sum_{i,j} [SL_2(\mathbb{Z}) x_i y_j]$.

For fixed j , we have $\mathbb{Z}^2 \geq \mathbb{Z}^2 y_j \geq \mathbb{Z}^2 x_i y_j$, and the $\mathbb{Z}^2 x_i y_j$ are exactly the subgroups of $\mathbb{Z}^2 y_j$ of index n .

For varying i and j , the $\mathbb{Z}^2 x_i y_j$ are exactly the subgroups of \mathbb{Z}^2 of index nm , each Λ appearing with multiplicity

$$\#\{\mathbb{Z}^2 \geq \Lambda' \geq \Lambda \mid [\mathbb{Z}^2 : \Lambda'] = m \text{ and } [\Lambda' : \Lambda] = n\}$$

When $(n, m) = 1$, each subgroup $\Lambda \leq \mathbb{Z}^2$ of index nm has a unique $\mathbb{Z}^2 \geq \Lambda' \geq \Lambda$ such that $[\mathbb{Z}^2 : \Lambda'] = m$ and $[\Lambda' : \Lambda] = n$. We find $[X_n] \cdot [X_m] = [X_{nm}]$, hence $T_n \circ T_m = T_{nm}$.

Now consider p a prime and $n \in \mathbb{N}$. We want to compute $[X_p] \cdot [X_{p^n}] = f$. Then f corresponds to $\varphi_f : \mathcal{L}_{p^{n+1}} \rightarrow \mathbb{C}$, given by $\varphi_f(\Lambda) = \#\{\mathbb{Z}^2 \geq \Lambda' \geq \Lambda\}$, where $[\mathbb{Z}^2 : \Lambda'] = p^n$, $[\Lambda' : \Lambda] = p$.

We compute $\varphi_f(\Lambda)$.

Case 1: \mathbb{Z}^2/Λ is cyclic. Then \mathbb{Z}^2/Λ has a unique subgroup of order p , so $\varphi_f(\Lambda) = 1$.

Case 2: $\mathbb{Z}^2/\Lambda \cong \mathbb{Z}/p^a \mathbb{Z} \oplus \mathbb{Z}/p^b \mathbb{Z}$ where $a, b \geq 1$ and $a + b = n + 1$. Equivalently, $\Lambda \leq p\mathbb{Z}^2$. In this case \mathbb{Z}^2/Λ has $p + 1$ subgroups of order p , so $\varphi_f(\Lambda) = p + 1$.

So $\varphi_f : \mathcal{L}_{p^{n+1}} \rightarrow \mathbb{C}$ is equal to $1 + p\delta_{\Lambda \leq p\mathbb{Z}^2}$.

hence $[X_p] \cdot [X_{p^n}] = [X_{p^{n+1}}] + p[pX_{p^{n+1}}]$, and so $T_{p^n} \circ T_p = T_{p^{n+1}} + p \cdot p^{2k-2} \cdot p^{-k} T_{p^{n-1}} = T_{p^{n+1}} + p^{k-1} T_{p^{n-1}}$. \square

Corollary 2.15. 1. For all primes p , T_{p^n} is a polynomial in T_p .

2. For all $n, m \in \mathbb{N}$, T_n and T_m commute.

Proof. 1. By induction on n , using the identity for $T_{p^{n+1}}$.

2. We know that if $n = \prod_i p_i^{a_i}$, then $T_n = \prod_i T_{p_i^{a_i}}$. We also know that T_{p^a}, T_{q^b} commute for any primes $p, q; a, b \in \mathbb{N}$.

□

Proposition 2.16. Let $k \geq 4$ be even. Then, for all $n \in \mathbb{N}$, $T_n E_k = \sigma_{k-1}(n) E_k$.

I.e., the E_k are eigenvectors for the linear operators T_n , with eigenvalues $\sigma_{k-1}(n)$.

This has the consequence that, if we want to diagonalise the actions of T_n on M_k , it's enough to diagonalise the actions of T_n on S_k , because of the decomposition $M_K(SL_2(\mathbb{Z})) = S_k(SL_2(\mathbb{Z})) \oplus \mathbb{C}E_k$.

Proof. To show E_k is an eigenvector of T_n , it's enough to show it's an eigenvector of the T_p where p ranges over the primes.

The eigenvalue is found to be $\sigma_{k-1}(n)$ by the identity $a_0(T_n E_k) = \sigma_{k-1}(n) a_0(E_k)$ (and $a_0(E_k) \neq 0$).

Let p be prime. We'll show G_k is an eigenvector of T_p .

$$(T_p G_k)(\tau) = \sum_{\substack{ad=p \\ 0 \leq b < d}} G_k \Big|_k \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} = \sum_{\substack{ad=p \\ 0 \leq b < d}} \sum_{\substack{\omega \in \Lambda_{(a\tau+b)/d} \\ \omega \neq 0}} \omega^{-k} p^{k-1} d^{-k}$$

where $\Lambda_\tau = \mathbb{Z}\tau \oplus \mathbb{Z}$.

Note $\Lambda_{(a\tau+b)/d} \frac{1}{d} (\mathbb{Z}(a\tau+b) \oplus d\mathbb{Z})$. These $\mathbb{Z}(a\tau+b) \oplus d\mathbb{Z}$ are precisely the subgroups of Λ_τ of index p , and hence:

$$(T_p G_k)(\tau) = \sum_{\substack{ad=p \\ 0 \leq b < d}} p^{k-1} \sum_{\substack{\omega \in \Lambda_{(a\tau+b)/d} \\ \omega \neq 0}} (d\omega)^{-k} = \sum_{\substack{\Lambda' \leq \Lambda_\tau \\ [\Lambda_\tau : \Lambda'] = p}} p^{k-1} \sum_{\omega \in \Lambda' \setminus \{0\}} \omega^{-k}$$

Think of some $\omega \in \Lambda_\tau \setminus \{0\}$. How many $\Lambda' \leq \Lambda_\tau$ of index p contain ω ?

Case 1 $\omega \notin p\Lambda_\tau$. Then $\Lambda' = \mathbb{Z}\omega + p\Lambda_\tau$ is the unique such subgroup.

Case 2 $\omega \in p\Lambda_\tau$. Then every such $\Lambda' \leq \Lambda_\tau$ contains ω .

Hence:

$$\begin{aligned} (T_p G_k)(\tau) &= p^{k-1} \left[\sum_{\omega \in \Lambda_\tau \setminus \{0\}} \omega^{-k} + \sum_{\omega \in p\Lambda_\tau \setminus \{0\}} p \omega^{-k} \right] \\ &= p^{k-1} \left[G_k(\tau) + \sum_{\omega \in \Lambda_\tau \setminus \{0\}} p(p\omega)^{-k} \right] \\ &= (1 + p^{k-1}) G_k(\tau) \\ &= \sigma_{k-1}(p) G_k(\tau) \end{aligned}$$

□

Our focus now is on $S_k(SL_2(\mathbb{Z}))$.

Lemma 2.17.

1. $\frac{dx dy}{y^2}$ where $\tau = x + iy$ is a $GL_2(\mathbb{R})^+$ -invariant area form.
2. If $f, g : \mathbb{H} \rightarrow \mathbb{C}$ are smooth and invariant under the weight k action of a subgroup $G \leq GL_2(\mathbb{R})^+$, then $\omega(f, g) = f(\tau)\overline{g(\tau)}y^k \frac{dy dx}{y^2}$ satisfies $\alpha^* \omega(f, g) = \det(\alpha)^{2-k} \omega(f, g)$ for all $\alpha \in G$.

Proof.

1. $\tau = x + iy, d\tau = dx + idy, d\bar{\tau} = dx - idy$, so $d\tau d\bar{\tau} = -2idxdy$

If $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R})^+$, then

$$g^*(d\tau) = d \frac{a\tau + b}{c\tau + d} = \frac{a(c\tau + d) - c(a\tau + b)}{(c\tau + d)^2} d\tau = \frac{\det(g)}{j(g\tau)^2} d\tau$$

$$\text{Hence } g^* \left(\frac{d\tau d\bar{\tau}}{y^2} \right) = \frac{\det(g)^2 d\tau d\bar{\tau}}{|j(g, \tau)|^4} = \left(\frac{\det(g)y}{|j(g\tau)|^2} \right)^{-2} = \frac{d\tau d\bar{\tau}}{y^2}.$$

2. Take $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$. Then $\alpha^* \omega(f, g) = f(\alpha\tau)\overline{g(\alpha\tau)} \Im(\alpha\tau)^k \frac{dx dy}{y^2}$.

We know $f(\alpha\tau)j(\alpha, \tau)^{-k} \det(\alpha)^{k-1} = f(\tau)$, and similarly for g , so

$$\alpha^* \omega(f, g) = \det(\alpha)^{2-k} \omega(f, g)$$

□

So if f, g are invariant under the weight k action of $\Gamma(1)$, we have a $\Gamma(1)$ -invariant area form $\omega(f, g)$.

Definition 2.18. If $\varphi : \mathbb{H} \rightarrow \mathbb{C}$ is a $\Gamma(1)$ -invariant function, then:

$$\int_{\Gamma(1) \backslash \mathbb{H}} \varphi \frac{dx dy}{y^2} = \int_{\mathcal{F}} \varphi \frac{dx dy}{y^2}$$

when this integral is absolutely convergent.

Definition 2.19. Let $f, g \in S_k(\Gamma(1))$. Then we define

$$\langle f, g \rangle = \int_{\Gamma(1) \backslash \mathbb{H}} \omega(f, g)$$

Lemma 2.20. For any $f, g \in S_k(\Gamma(1))$, then integral defining $\langle f, g \rangle$ is absolutely convergent. This defines an inner product on the \mathbb{C} -vector space $S_k(\Gamma(1))$.

Proof. It's enough to show that

$$1. \int_{\Gamma(1) \setminus \mathbb{H}} \frac{dx dy}{y^2} < \infty$$

2. If $f \in S_k(\Gamma(1))$, then $|f(\tau)|y^{k/2}$ is bounded in \mathcal{F} .

To show (1), we observe that $\mathcal{F} \subseteq \{\tau \in \mathbb{H} : -\frac{1}{2} \leq \Re \tau \leq \frac{1}{2}, \Im \tau \geq \frac{1}{2}\}$, so:

$$\int_{\Gamma(1) \setminus \mathbb{H}} \frac{dx dy}{y^2} \leq \int_{x=-\frac{1}{2}}^{\frac{1}{2}} \int_{y=\frac{1}{2}}^{\infty} \frac{dx dy}{y^2} < \infty$$

To show (2), we observe that $|f(\tau)|y^{k/2} \leq \sum_{n \geq 1} |a_n| e^{-2\pi n y} y^{k/2} \rightarrow 0$ as $y \rightarrow \infty$.

Note - here we're using that f is cuspidal. Then we have:

$$\omega(f, g) = f(\tau) y^{k/2} \overline{g(\tau)} y^{k/2} \frac{dx dy}{y^2} \implies \langle f, g \rangle \text{ is absolutely convergent}$$

If $\langle f, f \rangle = 0$, then $\int_{\Gamma(1) \setminus \mathbb{H}} |f(\tau)|^2 y^k \frac{dx dy}{y^2} = 0 \implies f \equiv 0$, and so $\langle \cdot, \cdot \rangle$ is an inner product. \square

Definition 2.21. Suppose that $\Gamma \leq \Gamma(1)$ is a subgroup of finite index. Let $\varphi : \mathbb{H} \rightarrow \mathbb{C}$ be a smooth function, invariant under translation by Γ . Then we define

$$\int_{\Gamma \setminus \mathbb{H}} \varphi \frac{dx dy}{y^2} = \int_{\Gamma(1) \setminus \mathbb{H}} \sum_{\gamma \in \Gamma \setminus \Gamma(1)} (\varphi \circ \gamma) \frac{dx dy}{y^2}$$

Definition 2.22. Suppose $\Gamma \leq \Gamma(1)$ is a subgroup of $\Gamma(1)$ of finite index and let $f, g : \mathbb{H} \rightarrow \mathbb{C}$ be invariant under the weight k action of Γ . Then we define

$$\langle f, g \rangle = [\Gamma(1) : \Gamma]^{-1} \int_{\Gamma \setminus \mathbb{H}} \omega(f, g) = [\Gamma(1) : \Gamma] \int_{\Gamma \setminus \mathbb{H}} f \bar{g} y^k \frac{dx dy}{y^2}$$

when this integral is absolutely convergent.

Remarks.

1. If $\Gamma' \leq \Gamma \leq \Gamma(1)$ is another finite index subgroup, then $\langle f, g \rangle$ defined using Γ is the same as $\langle f, g \rangle$ defined using Γ' .
2. If $f : \mathbb{H} \rightarrow \mathbb{C}$ is invariant under the weight k action of $\Gamma(1)$, and $\alpha \in GL_2(\mathbb{Q})^+$, then $f|_k(\alpha)$ is invariant under the weight k action of the group $\alpha^{-1}\Gamma(1)\alpha \cap \Gamma(1)$, which has finite index in $\Gamma(1)$.

Proposition 2.23. Let $f, g : \mathbb{H} \rightarrow \mathbb{C}$ be smooth functions which are invariant under the weight k action of $\Gamma(1)$. Then we have:

$$\langle f|_k[\alpha], g|_k[\alpha] \rangle = \det(\alpha)^{k-2} \langle f, g \rangle$$

whenever both sides of this equality are defined.

Proof.

$$\begin{aligned} \langle f|_k[\alpha], g|_k[\alpha] \rangle &= \int_{\alpha^{-1}\Gamma(1)\alpha \cap \Gamma(1) \setminus \mathbb{H}} f(\alpha\tau) j(\alpha, \tau)^{-k} \det(\alpha)^{k-1} \overline{g(\alpha\tau) j(\alpha, \tau)^{-k} \det(\alpha)^{k-1}} y^k \frac{dx dy}{y^2} \\ &= \int_{\alpha^{-1}\Gamma(1)\alpha \cap \Gamma(1) \setminus \mathbb{H}} f(\alpha\tau) g(\alpha\tau) |j(\alpha, \tau)|^{-2k} \det(\alpha)^{2k-2} y^k \frac{dx dy}{y^2} \end{aligned}$$

We have a bijection $\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H} \xrightarrow{\alpha^{-1}} \Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha \backslash \mathbb{H}$, as if $\gamma \in \alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1)$ and $\tau \in \mathbb{H}$, then $\alpha^{-1}(\gamma \cdot \tau) = (\alpha^{-1}\gamma\alpha) \cdot (\alpha^{-1}\tau)$.

So we may make a change of variables in the integral to get:

$$\begin{aligned} &= \int_{\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H}} f(\tau) \overline{g(\tau)} |j(\alpha, \alpha^{-1}\tau)|^{-2k} \det(\alpha)^{2k-2} \Im(\alpha^{-1}\tau)^k \frac{dx dy}{y^2} \\ &= \int_{\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H}} f(\tau) \overline{g(\tau)} |j(\alpha, \alpha^{-1}\tau)|^{-2k} \det(\alpha)^{2k-2} \frac{\det(\alpha)^{-k} \Im(\tau)}{|j(\alpha^{-1}, \tau)|^{2k}} \frac{dx dy}{y^2} \end{aligned}$$

Note if $\alpha, \beta \in GL_2(\mathbb{R})^+$, then $j(\alpha\beta, \tau) = j(\beta, \tau)j(\alpha, \beta\tau)$.

$$\begin{aligned} &= \int_{\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H}} f(\tau) \overline{g(\tau)} \det(\alpha)^{k-2} \Im(\tau) \frac{dx dy}{y^2} \\ &= \frac{\langle f, g \rangle \det(\alpha)^{k-2}}{[\Gamma(1) : \alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1)]^{-1}} \end{aligned}$$

The only thing left to do is to show that $[\Gamma(1) : \Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha] = [\Gamma(1) : \alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1)]$.

This is true since there is a bijective map $\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H} \xrightarrow{\alpha^{-1}} \Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha \backslash \mathbb{H}$ which pulls $\frac{dx dy}{y^2}$ to $\frac{dx dy}{y^2}$.

Hence

$$\begin{aligned} \int_{\Gamma(1) \backslash \mathbb{H}} \frac{dx dy}{y^2} &= [\Gamma(1) : \Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha] \int_{\Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha \backslash \mathbb{H}} \frac{dx dy}{y^2} \\ &= [\Gamma(1) : \Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha]^{-1} \int_{\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H}} \frac{dx dy}{y^2} \\ &= \frac{[\Gamma(1) : \Gamma(1) \cap \alpha\Gamma(1)\alpha^{-1}]}{[\Gamma(1) : \Gamma(1) \cap \alpha^{-1}\Gamma(1)\alpha]} \int_{\alpha\Gamma(1)\alpha^{-1} \cap \Gamma(1) \backslash \mathbb{H}} \frac{dx dy}{y^2} \end{aligned}$$

And since the integral is nonzero, we are done. \square

Proposition 2.24. *For all $n \in \mathbb{N}$, the Hecke operator T_n on $S_k(\Gamma(1))$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle$.*

Proof. We need to show that, for all $f, g \in S_k(\Gamma(1))$, $\langle T_n f, g \rangle = \langle f, T_n g \rangle$. It's enough to do this for prime $n = p$.

We claim we can find matrices $x_i \in X_p$ such that $X_p = \bigsqcup_i \Gamma(1)x_i = \bigsqcup_i \Gamma(1)px_i^{-1}$.

Indeed, note that the adjugate map $\text{Adj} : X_p \rightarrow X_p$ is a bijection, as it preserves determinant. First choose representatives a_i, b_i so that $X_p = \bigsqcup_i \Gamma(1)a_i = \bigsqcup_i b_i\Gamma(1)$.

Since every matrix in X_p has Smith normal form $\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$, so X_p consists of a single $\Gamma(1)$ -double coset.

So we can choose $u_i, v_i \in \Gamma(1)$ such that $a_i a_i = b_i v_i = x_i$. Then $X_p = \bigsqcup_i \Gamma(1)x_i = \bigsqcup_i x_i \Gamma(1)$. Also $X_p = \text{Adj}(X_p) = \bigsqcup_i \gamma(1)px_i^{-1}$.

We then compute

$$\begin{aligned}
\langle T_p f, g \rangle &= \left\langle \sum_i f|_k[x_i], g \right\rangle \\
&= \sum_i \langle f|_k[x_i], g \rangle \\
&= \sum_i \langle f|_k[x_i]|_k[x_i^{-1}], g|_k[x_i^{-1}] \rangle \\
&= \sum_i \langle f, g|_k[x_i^{-1}] \rangle p^{k-2} \\
&= \langle f, \sum_i g|_k[p x_i^{-1}] \rangle \\
&= \langle f, T_p g \rangle
\end{aligned}$$

□

Corollary 2.25. *The operators T_n on $S_k(\Gamma(1))$ can be simultaneously diagonalised. Their eigenvalues are real numbers.*

Proof. T_n self-adjoint implies that T_n are diagonalisable with real eigenvalues, and since they commute, they can be simultaneously diagonalised. □

Definition 2.26. *We say that $f \in M_k(\Gamma(1))$ is a normalised eigenform if f is an eigenvector for all Hecke operators T_n and $a_1(f) = 1$.*

Lemma 2.27. *If $f \in M_k(SL_2(\mathbb{Z}))$, then if f is an eigenvector for all Hecke operators T_n and $k > 0$, then there's a unique scalar multiple of f which is a normalised eigenform.*

Proof. We have $a_n(f) = a_1(T_n(f))$. If $a_1(f) = 0$ and T_n has eigenvalue α_n on f , then $a_n(f) = a_1(T_n(f)) = \alpha_n(a_1(f)) = 0$, so $a_n(f) = 0$ for all $n \in \mathbb{N}$. Therefore $a_1(f) \neq 0$, and $f/a_1(f)$ is a normalised eigenform. □

We've seen $G_k(\tau) = 2\zeta(k) + \frac{2(2\pi i)^k}{(k-1)!} \sum_{n \geq 1} \sigma_{k-1}(n) q^n$ is an eigenvector for all Hecke generators T_n . The corresponding normalised eigenform is

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

Proposition 2.28. *The eigenvalues of T_n on $M_k(SL_2(\mathbb{Z}))$ are algebraic integers lying in a number field independent of n .*

Proof. We proved earlier that there's a basis f_1, \dots, f_N for $S_k(SL_2(\mathbb{Z}))$ such that

1. $\forall n \in \mathbb{N}, a_n(f_i) \in \mathbb{Z}$
2. $a_j(f_i) = \delta_{ij}$

Consequently, for any $f \in S_k(SL_2(\mathbb{Z}))$, we have $f = \sum_{i=1}^N a_i(f) f_i$.

Now consider the matrix of T_n with respect to this basis. We have $a_m(T_n f) = \sum_{a|(m,n)} a^{k-1} a_{mn/a^2}(f)$. So $T_n(f)$ has integer q -expansion coefficients if f does. We see that the matrix of T_n with respect to the basis f_1, \dots, f_N has integer entries. Hence the eigenvalues of T_n are algebraic integers.

If f is a normalised eigenform, then $f = \sum_{i=1}^N a_i(f) f_i$, so the eigenvalues of T_n on f lie in the field generated by $a_1(f), \dots, a_N(f)$. \square

To any weight k , we can associate the sequences $(a_1(f_i), a_2(f_i), \dots)$ of Hecke eigenvalues. Since the operators T_n are polynomials in the T_p for p prime, it's equivalent to give the sequences $(a_p(f_i))_p$.

First examples arise when $S_k(SL_2(\mathbb{Z}))$ is 1-dimensional - this happens when $k = 12, 16, 18, 20, 22, 26$. When $k = 12$, Δ is a normalised eigenform. Ramanujan proved that:

$$\Delta = q \prod_{n \geq 1} (1 - q^n)^{24} = \sum_{n \geq 1} \tau(n) q^n$$

He conjectured that $\tau(nm) = \tau(n)\tau(m)$ whenever n, m are coprime, and if p is prime, then $\tau(p)\tau(p^n) = \tau(p^{n+1}) + p^{11}\tau(p^{n-1})$. These relations follow from $T_{nm} = T_n T_m$ and $T_p T_{p^n} = T_{p^{n+1}} + p^{k-1} T_{p^{n-1}}$.

In the case $k = 24$, $\dim_{\mathbb{C}} S_k(SL_2(\mathbb{Z})) > 1$. We know we can find a basis g_1, g_2 for $S_{24}(SL_2(\mathbb{Z}))$ such that $g_1 = q + O(q^3)$, $g_2 = q^2 + O(q^3)$. Let's compute the matrix of T_2 with respect to this basis. We know $T_2(\sum_{n \geq 1} a_n q^n) = \sum_{n \geq 1} \left(\sum_{a|(2,n)} a^{k-1} a_{2n/a^2} \right) q^n$. In particular, $a_1(T_2(f)) = a_2(f)$, and $a_2(T_2(f)) = a_4(f) + 2^{23} a_1(f)$. So $T_2(f) = a_2(f) g_1 + (a_4(f) + 2^{23} a_1(f)) g_2$.

After a computation, we see that matrix of T_2 with respect to our chosen basis is $\begin{pmatrix} 0 & 1 \\ 20468736 & 1080 \end{pmatrix}$, which has eigenvalues $12(45 \pm \sqrt{144169})$. All of the eigenvalues of T_n on $S_{24}(SL_2(\mathbb{Z}))$ will lie in $\mathbb{Q}(\sqrt{144169})$.

Lemma 2.29. *Let p be a prime number. Then we have an identity of formal power series:*

$$1 + \sum_{n \geq 1} \tau(p^n) X^n = (1 - \tau(p)X + p^{11}X^2)^{-1}$$

Proof.

$$\left(\sum_{n \geq 0} \tau(p^n) X^n \right) (1 - \tau(p)X + p^{11}X^2) = 1 + \sum_{n \geq 1} \left(\tau(p^{n+2}) - \tau(p)\tau(p^{n+1}) + p^{11}\tau(p^n) \right) X^{n+2}$$

But we know $\tau(p)\tau(p^{n+1}) = \tau(p^{n+2}) + p^{11}\tau(p^n)$. \square

We can write this generating function as $\frac{1}{(1-\alpha_p X)(1-\beta_p X)}$ where α_p, β_p are the roots of the polynomial $1 - \tau(p)X + p^{11}X^2$, so $\alpha_p, \beta_p = \frac{-\tau(p) \pm \sqrt{\tau(p)^2 - 4p^{11}}}{2p^{11}}$.

There are two possibilities: $\tau(p)^2 - 4p^{11} \leq 0$, in which case α_p, β_p are complex conjugates of absolute value $p^{11/2}$, or $\tau(p)^2 > 4p^{11}$, and α_p, β_p are real numbers of distinct absolute values. Ramanujan conjectured that the first case always holds.

Conjecture 2.30 (Ramanujan-Peterson). *Let f be a cuspidal normalised eigenform of weight k and level $SL_2(\mathbb{Z})$. Then for every prime p , we have $|a_p(f)| \leq 2p^{(k-1)/2}$.*

This conjecture is behind many applications of modular forms to number theory. Later in the course we'll prove the formula

$$r_{24}(p) = \frac{16}{691}\sigma_{11}(p) + \frac{33152}{691}\tau(p)$$

where p is an odd prime and $r_{24}(n)$ is the number of ways of writing n as a sum of 24 integer squares.

The Ramanujan-Peterson gives the estimate $r_{24}(p) = \frac{16}{691}p^{11} + O(p^{11/2})$.

Another reason for the importance of Hecke operators is the possibility of constructing L -functions. First recall the Riemann zeta function $\zeta(s) = \sum_{n \geq 1} n^{-s}$. It has the following properties:

1. (Euler Product) $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$.
2. (Analytic Continuation) $\zeta(s)$ has a meromorphic continuation to all of \mathbb{C} with a simple pole at $s = 1$ and no other poles.
3. (Functional Equation) Let $\xi(s) = \pi^{-s/2} \Gamma(s/2) \zeta(s)$. Then $\xi(s) = \xi(1 - s)$.
4. (Special Values) The values of $\zeta(s)$ at integers should be related to arithmetic in some way.

Recall: $\Gamma(s) = \int_{y=0}^{\infty} e^{-y} y^s \frac{dy}{y}$. This integral converges absolutely and uniformly in compact subsets of the right half plane $\Re(s) > 0$. Integration by parts gives $s\Gamma(s) = \Gamma(s + 1)$, and so $\Gamma(s)$ can be continued meromorphically to \mathbb{C} .

A Dirichlet series $\sum_{n \geq 1} a_n n^{-s}$ with properties similar to 1,2,3 is called an **L -function**. Further examples are given by the Dirichlet L -function $L(\chi, s) = \sum_{n \geq 1, (n, N)=1} \chi(n \bmod N) n^{-s}$ where $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \rightarrow \mathbb{C}$ is a character.

Definition 2.31. Let $f \in M_k(SL_2(\mathbb{Z}))$ be a non-zero modular form and let its q -expansion be $f(\tau) = \sum_{n \geq 0} a_n q^n$. The associated L -function is $L(f, s) = \sum_{n \geq 1} a_n n^{-s}$.

For example, when f is the normalised eigenform associated to G_k , we have $L(f, s) = \sum_{n \geq 1} \sigma_{k-1}(n) n^{-s} = \sum_{n, m \geq 1} m^{k-1} (mn)^{-s} = \zeta(s) \zeta(s + 1 - k)$.

In this case, there is a functional equation relating $L(f, s)$ and $L(f, k - s)$.

Lemma 2.32. Let $f \in S_k(SL_2(\mathbb{Z}))$ be non-zero. Then $L(f, s)$ converges absolutely on the region $\Re(s) > k/2 + 1$.

Proof. On the example sheet, we show $|a_n(f)| = O(n^{k/2})$, from which we have absolute convergence. □

Theorem 2.33. Let $f \in S_k(SL_2(\mathbb{Z}))$ be a non-zero cuspidal modular form. Then:

1. $L(f, s)$ has an analytic continuation to \mathbb{C} .
2. Let $\Lambda(f, s) = (2\pi)^{-s} \Gamma(s) L(f, s)$. Then $\Lambda(f, s)$ admits an analytic continuation to \mathbb{C} and satisfies $\Lambda(f, s) = i^k \Lambda(f, k - s)$.

Proof. We define $F(s) = \int_0^\infty f(iy)y^s \frac{dy}{y}$, the Mellin transform of $f(iy)$. The function $F(s)$ converges absolutely in \mathbb{C} and defines a holomorphic function there. We know that $|f(iy)| = O(e^{-2\pi y})$ as $y \rightarrow \infty$. This shows that $\int_1^\infty f(iy)y^s dy/y$ has the desired properties.

We know $f(-1/\tau)\tau^{-k} = f(\tau)$ and hence $f(iy) = f(i/y)(iy)^{-k}$. This shows that $|f(i/y)| = |f(iy)(iy)^k| = O(e^{-\pi y})$ as $y \rightarrow \infty$, and so $\int_0^1 f(iy)y^s dy/y$ has the required properties.

Formally, we can compute:

$$\begin{aligned} F(s) &= \int_0^\infty a_n e^{-2\pi n y} y^s dy/y \\ &= \sum_{n \geq 1} a_n \int_0^\infty e^{-2\pi n y} y^s dy/y \\ &= \sum_{n \geq 1} a_n \int_0^\infty e^{-y} \left(\frac{y}{2\pi n}\right)^s dy/y \\ &= \sum_{n \geq 1} a_n n^{-s} \Gamma(s) (2\pi)^{-s} \\ &= (2\pi)^{-s} \Gamma(s) L(f, s) \\ &= \Lambda(f, s) \end{aligned}$$

The absolute convergence of $L(f, s)$ in the region $\Re(s) > k/2 + 1$ implies that this computation is valid in the same region.

To get the functional equation, we compute:

$$\begin{aligned} \Lambda(f, s) &= \int_0^1 f(iy)y^s dy/y + \int_1^\infty f(iy)y^s dy/y \\ &= \int_0^1 f(i/y)(iy)^{-k} y^s dy/y + \int_1^\infty f(iy)y^s dy/y \\ &= \int_1^\infty f(iy)i^k y^{k-s} dy/y + \int_1^\infty f(iy)y^s dy/y \end{aligned}$$

and hence $\Lambda(f, s) = i^k \Lambda(f, k-s)$. □

Theorem 2.34. *Let $f \in S_k(SL_2(\mathbb{Z}))$ be a normalised eigenform. Then $L(f, s)$ has the Euler product*

$$L(f, s) = \prod_p (1 - a_p(f)p^{-s} + p^{k-1-2s})^{-1}$$

Proof. We know $a_{nm}(f) = a_n(f)a_m(f)$ when n, m coprime. Hence $\sum_{n \geq 1} a_n(f)n^{-s} = \prod_p \sum_{k \geq 0} a_{p^k}(f)p^{-ks}$.

We need to show that $\sum_{k \geq 1} a_{p^k}(f)p^{-ks} = (1 - a_p(f) + p^{k-1-2s})^{-1}$. We saw last time that this follows from the relation for $T_p T_{p^k}$. □

Now, given $f \in S_k(SL_2(\mathbb{Z}))$ a normalised eigenform, we factor, for each prime p :

$$(1 - a_p X + p^{k-1} X^2) = (1 - \alpha_p X)(1 - \beta_p X)$$

Definition 2.35. For each $m \geq 1$, we define the associated symmetric power L -function

$$L(\text{Sym}^m, f, s) = \prod_p \prod_{i=0}^m (1 - \alpha_p^i \beta_p^{m-i} p^{-s})^{-1}$$

Lemma 2.36. $L(\text{Sym}^m, f, s)$ converges in some right half-plane.

Proof. We saw in the previous lecture that $|a_p(f)| = O(p^{k/2})$.

This implies that $|\alpha_p|, |\beta_p| = O(p^{k/2})$, and hence $|\alpha_p^i \beta_p^{m-i}| = O(p^{mk/2})$. It follows that $\prod_p (1 - \alpha_p^i \beta_p^{m-i} p^{-s})^{-1}$ is absolutely convergent in some right half-plane.

We thus deduce that $L(\text{Sym}^m, f, s)$ has the same property. \square

These L -functions were first considered in the 1960s by Langlands and Serre. The idea was to associate to each prime p the conjugacy class of the matrix $t_p = \begin{pmatrix} \alpha_p & 0 \\ 0 & \beta_p \end{pmatrix}$.

Then associate to any algebraic representation $R : GL_2(\mathbb{C}) \rightarrow GL_N(\mathbb{C})$ the L -function $L(R, f, s) = \prod_p \det(1 - R(t_p)p^{-s})^{-1}$. The irreducible representatives of GL_2 are the symmetric powers $\text{Sym}^m : GL_2(\mathbb{C}) \rightarrow GL_{m+1}(\mathbb{C})$ of the identity representations (up to a twist).

Conjecture 2.37. For all $m \geq 1$, $L(\text{Sym}^m, f, s)$ has an analytic continuation and satisfies a functional equation $\Lambda(\text{Sym}^m, f, s) = \Lambda(\text{Sym}^m, f, m(k-1) + 1 - s)$, where Λ is an explicit completed L -function.

Proposition 2.38 (Langlands). If, for all $m \geq 1$, $L(\text{Sym}^m, f, s)$ has an analytic continuation to \mathbb{C} , then the Ramanujan-Petersen conjecture holds for f .

The idea behind this proposition made another appearance in the proof of the Weil conjectures by Deligne.

Conjecture 2.39 (Sato-Tate). Suppose the Ramanujan-Petersen conjecture holds, so $|a_p(f)| \leq 2p^{(k-1)/2}$, and so $a_p/(2p^{(k-1)/2}) \in [-1, 1]$. Then these numbers are equidistributed throughout $[-1, 1]$ as $p \rightarrow \infty$ with respect to the Sato-Tate measure $\frac{2}{\pi} \sqrt{1-t^2} dt$.

More precisely, for all continuous functions $g : [-1, 1] \rightarrow \mathbb{R}$,

$$\lim_{X \rightarrow \infty} \frac{1}{\pi(X)} \sum_{p < X} g(a_p/(2p^{(k-1)/2})) \rightarrow \frac{2}{\pi} \int_{-1}^1 g(t) \sqrt{1-t^2} dt$$

where $\pi(x)$ is the prime counting function.

This has the consequence that the Ramanujan-Petersen conjecture is sharp.

Proposition 2.40 (Serre). Suppose for all $m \geq 1$, $L(\text{Sym}^m, f, s)$ has an analytic continuation to \mathbb{C} which is non-vanishing on the line $\Re(s) = 1 + m \frac{k-1}{2}$. Then the Sato-Tate conjecture holds for f .

Recall the formula for an odd prime p : $r_{24}(p) = \frac{16}{691}(1 + p^{11}) + \frac{33152}{691}\tau(p)$.

The Sato-Tate conjecture implies that the normalised error terms

$$(r_{24}(p) - \frac{16}{691}(1 + p^{11})) \times \frac{691}{33152 \times 2 \times p^{11/2}} \in [-1, 1]$$

are equidistributed.

The Sato-Tate conjecture was proved in 2010 by Barnet-Lamb, Geraghty, Harris, and Taylor. The complete analytic continuation of $L(\text{Sym}^m, f, s)$ was established in 2019 by Newton and Thorne.

The next step is to study modular forms $M_k(\Gamma), S_k(\Gamma)$ where Γ is any congruence subgroup of $SL_2(\mathbb{Z})$.

3 Modular Forms on Congruence Subgroups of $SL_2(\mathbb{Z})$

Definition 3.1. For all $N \in \mathbb{N}$, we define $\Gamma(N) = \ker(SL_2(\mathbb{Z}) \rightarrow SL_2(\mathbb{Z}/N\mathbb{Z}))$.

A **congruence subgroup** $\Gamma \leq SL_2(\mathbb{Z})$ is any subgroup containing $\Gamma(N)$ for some $N \in \mathbb{N}$.

Important examples:

- $\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : c \equiv 0 \pmod{N} \right\}$
- $\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : c \equiv 0, a \equiv 1 \pmod{N} \right\}$

The homomorphism $SL_2(\mathbb{Z}) \rightarrow SL_2(\mathbb{Z}/N\mathbb{Z})$ is surjective. This implies that $\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/N\mathbb{Z})^\times$, given by $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \pmod{N}$.

We've also seen $\Gamma = \langle S, \Gamma(2) \rangle$.

Definition 3.2. A **weakly modular function of weight k and level Γ** where Γ is a congruence subgroup is a meromorphic function on \mathbb{H} such that, for all $\gamma \in \Gamma$, $f|_k[\gamma] = f$.

Recall $SL_2(\mathbb{Z}) \cap \mathbb{H}$ had the fundamental domain \mathcal{F} .

Definition 3.3. A **cusp** of a congruence subgroup Γ is a Γ -orbit on the set $\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\} \subseteq \widehat{\mathbb{C}}$.

Note $GL_2(\mathbb{Q})^+$ acts on $\mathbb{P}^1(\mathbb{Q})$ by Möbius transformations.

Lemma 3.4. $SL_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$. Equivalently, $SL_2(\mathbb{Z})$ has a unique cusp.

Proof. It's enough to show that, if $\frac{a}{c} \in \mathbb{Q}$ with $a, b \in \mathbb{Z}, (a, b) = 1$, then there is $\gamma \in SL_2(\mathbb{Z})$ with $\gamma \cdot 0 = \frac{a}{c}$. If $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, then $\gamma \cdot 0 = B/D$.

Since $(a, b) = 1$, we can find $c, d \in \mathbb{Z}$ such that $ad - bc = 1$. So we can take $\gamma = \begin{pmatrix} -b & a \\ -d & c \end{pmatrix}$. □

Corollary 3.5. If $\Gamma \leq SL_2(\mathbb{Z})$ is a congruence subgroup, then Γ has only finitely many cusps.

Proof. Let $\Gamma_\infty = \text{Stab}_{\Gamma(1)}(\infty) = \left\{ \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} : b \in \mathbb{Z} \right\}$. There's a bijection $\Gamma(1)/\Gamma_\infty \rightarrow \mathbb{P}^1(\mathbb{Q})$, given by $\gamma\Gamma_\infty \mapsto \gamma \cdot \infty$.

The set of cusps for Γ can be identified with the double quotient $\Gamma \backslash \Gamma(1)/\Gamma_\infty$. □

For example, we compute the cusps of $\Gamma = \langle S, \Gamma(2) \rangle$. We want to compute $\Gamma \backslash \Gamma(1) / \Gamma_\infty = \langle S \rangle \backslash SL_2(\mathbb{Z}/2\mathbb{Z}) / \Gamma_\infty$. Now $SL_2(\mathbb{Z}/2\mathbb{Z}) / \Gamma_\infty = \mathbb{F}_2^2 \setminus \{0\}$.

So $\Gamma \backslash \Gamma(1) / \Gamma_\infty = \{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \}, \{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \}$. Representatives are given by 1 and $\begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$.

Γ has two cusps, corresponding to $\Gamma \cdot \infty$ and $\Gamma \cdot -1$.

We'll see later that if Γ is a congruence subgroup, then $\Gamma \backslash \mathbb{H} \sqcup \Gamma \backslash \mathbb{P}^1(\mathbb{Q})$ can be given the structure of a compact Riemann surface.

Definition 3.6. If Γ is a congruence subgroup, the width of the cusp ∞ is:

$$\min\{h \in \mathbb{N} : \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \Gamma\} = [\Gamma_\infty : \Gamma_\infty \cap \{\pm 1\}\Gamma]$$

Let h be the width of the cusp ∞ , define $q_h = e^{2\pi i \tau / h}$.

If f is a weakly modular function of weight k and level Γ , then $f(\tau + h) = f(\tau)$ for all $\tau \in \mathbb{H}$, so we can define a meromorphic function $\tilde{f} : \{q_h \in \mathbb{C} : 0 < |q_h| < 1\}$ by the formula $\tilde{f}(e^{2\pi i \tau / h}) = f(\tau)$.

Definition 3.7. We say f is meromorphic at ∞ if \tilde{f} is meromorphic at 0, and holomorphic (resp. vanishes) at infinity if \tilde{f} is holomorphic (resp. vanishes) at $q_h = 0$.

If f is meromorphic at ∞ , then \tilde{f} admits a q_h -expansion $\tilde{f}(q_h) = \sum_{n \in \mathbb{Z}} a_n q_h^n$, and hence $f(\tau) = \sum_{n \in \mathbb{Z}} a_n q_h^n$.

Definition 3.8. Let f be a weakly modular function of weight k and level Γ . We say f is meromorphic at the cusp $\Gamma \cdot z$, $z \in \mathbb{P}^1(\mathbb{Q})$, if choosing $\alpha \in SL_2(\mathbb{Z})$ such that $\alpha \cdot \infty = z$, $f|_k[\alpha]$ is meromorphic at α , when viewed as a weakly modular function of level $\alpha^{-1}\Gamma\alpha$.

Similarly, we define what it means for f to be holomorphic or vanish at the cusp $\Gamma \cdot z$. We define the width of the cusp $\Gamma \cdot z$ to be the width of the cusp ∞ for the group $\alpha^{-1}\Gamma\alpha$.

Lemma 3.9.

1. The width of a cusp is well defined.
2. The meromorphy/holomorphy/vanishing of a weakly modular function at a cusp is well-defined.

Proof.

1. We need to first show that if $\alpha\infty = \beta\infty = z$, then the widths given by these choices are equal.

We must have $\beta = \alpha S$ for some $S \in \Gamma_\infty$

Then $[\Gamma_\infty : \{\pm 1\}\beta^{-1}\Gamma\beta \cap \Gamma_\infty] = [\Gamma_\infty : \{\pm 1\}S^{-1}\alpha^{-1}\Gamma\alpha S \cap \Gamma_\infty] = [\Gamma_\infty : \{\pm 1\}\alpha^{-1}\Gamma\alpha \cap \Gamma_\infty]$.

Next we show that, if $\gamma \in \Gamma$ and $\alpha \in SL_2(\mathbb{Z})$, then $[\Gamma_\infty : \{\pm 1\}\alpha^{-1}\gamma^{-1}\Gamma\gamma\alpha \cap \Gamma_\infty] = [\Gamma_\infty : \{\pm 1\}\alpha^{-1}\Gamma\alpha \cap \Gamma_\infty]$, but this is true since $\gamma^{-1}\Gamma\gamma = \Gamma$.

2. Suppose $\alpha, \beta \in SL_2(\mathbb{Z})$ satisfy $\alpha\infty = z = \beta\infty$. Then $\beta = \alpha\delta$ for some $\delta \in \Gamma_\infty$, say $\delta = \pm \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$. Then $f|_k[\beta](\tau) = f|_k[\alpha\delta](\tau) = f|_k[\alpha](\tau + a)(\pm 1)^k$.

If $f|_k[\alpha](\tau) = \sum_{n \in \mathbb{Z}} a_n q_h^n$, then $f|_k[\beta](\tau) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i \tau/h} e^{2\pi i a/h} (-1)^k$.

This shows that the meromorphy at ∞ of $f|_k[\alpha]$ is independent of choice of α .

Now suppose $\gamma \in \Gamma$ and $\gamma z = z'$. If $\alpha\infty = z$, then $\gamma\alpha\infty = z'$, and hence $f|_k[\gamma\alpha] = f|_k[\alpha]$, and again meromorphy at ∞ is independent of the choice of z .

□

Remark. The q -expansion of a weakly modular function f is not well defined independently of choices. By “the” q -expansion of a weakly modular function which is meromorphic at ∞ , we mean the q -expansion at ∞ .

Definition 3.10. A **modular function of weight k and level Γ** , for Γ a congruence subgroup and $k \in \mathbb{Z}$, is a weakly modular function that is meromorphic at every cusp of Γ .

f is a **modular form** if f is holomorphic in \mathbb{H} and holomorphic at every cusp of Γ , and a **cuspidal modular form** if f is a modular form vanishing at every cusp.

We write $M_k(\Gamma)$ (resp. $S_k(\Gamma)$) for the vector spaces of modular (resp. cuspidal modular) forms.

Remark. An equivalent definition is that f is a modular form if f is holomorphic in \mathbb{H} and for all $\alpha \in \Gamma(1)$, $f|_k[\alpha]$ is holomorphic at ∞ .

Lemma 3.11.

1. If $f \in M_k(\Gamma)$ and $g \in M_\ell(\Gamma)$, then $fg \in M_{k+\ell}(\Gamma)$.
2. If $\Gamma' \leq \Gamma$ is another congruence subgroup, then $M_k(\Gamma) \leq M_k(\Gamma')$.
3. If Γ, Γ' are congruence subgroups and $\alpha \in GL_2(\mathbb{Q})^+$, and $f \in M_k(\Gamma)$ and if $\Gamma' \leq \alpha^{-1}\Gamma\alpha$, then $f|_k[\alpha] \in M_k(\Gamma')$.

Moreover, these hold if M_k is replaced by S_k everywhere.

Proof.

1. Clear - same as in the case $\Gamma = \Gamma(1)$.
2. Is a special case of 3.
3. $f|_k[\alpha]$ is holomorphic in \mathbb{H} and is invariant under the weight k action of $\alpha^{-1}\Gamma\alpha$. We need to check that $f|_k[\alpha]$ is holomorphic at the cusps of Γ' . Equivalently, we need to check that, for all $\beta \in \Gamma(1)$, $f|_k[\alpha\beta]$ is holomorphic at ∞ .

We can write $\alpha\beta\infty = \gamma\infty$ for some $\gamma \in \Gamma(1)$, and then $\alpha\beta = \gamma\delta$ for some $\delta \in \text{Stab}_{GL_2(\mathbb{Q})^+}(\infty)$, say $\delta = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ with $a, b, d \in \mathbb{Q}$, and $ad > 0$.

Then $f|_k[\alpha\beta](\tau) = f|_k[\gamma\delta](\tau) = f|_k[\gamma](\frac{a\tau+b}{d})(ad)^{k-1}d^{-k}$. Since $f|_k[\gamma]$ is holomorphic at ∞ , we know $\lim_{\Im \tau \rightarrow \infty} f|_k[\gamma](\tau)$ exists. It follows that $\lim_{\Im \tau \rightarrow \infty} f|_k[\gamma](\frac{a\tau+b}{d})$ also exists, and so $f|_k[\alpha\beta]$ is holomorphic at ∞ .

□

Examples. Recall $\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : c \equiv 0 \pmod{N} \right\}$.

If $f \in M_k(\Gamma_0(M))$ and $L \in \mathbb{N}$, then $f(L\tau) \in M_k(\Gamma_0(ML))$.

Why? $L\tau = \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix} \tau$. By the lemma, we need to check that $\Gamma_0(ML) \leq \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}^{-1} \Gamma_0(M) \begin{pmatrix} L & 0 \\ 0 & 1 \end{pmatrix}$.

So if $k \in 2\mathbb{N}$, $k \geq 4$, p prime, then $G_k(\tau), G_k(p\tau) \in M_k(\Gamma_0(p))$.

Example. If $\tau \in \mathbb{H}$, we define

$$\theta(\tau) = \sum_{n \in \mathbb{Z}} e^{\pi i n^2 \tau}$$

This defines absolutely and uniformly in compact subsets of \mathbb{H} , so defines a holomorphic function. We have $\theta(\tau + 2) = \theta(\tau)$.

Proposition 3.12 (Poisson Summation Formula). *Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a continuous function such that $\sum_{n \in \mathbb{Z}} |f(n+x)|$ converges uniformly in \mathbb{R} and $\sum_{n \in \mathbb{Z}} |\widehat{f}(n)|$ converges, where*

$$\widehat{f}(n) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i n t} dt$$

Then we have:

$$\sum_{n \in \mathbb{Z}} f(n) = \sum_{n \in \mathbb{Z}} \widehat{f}(n)$$

Proof. Let $\phi(x) = \sum_{n \in \mathbb{Z}} f(n+x)$. Then ϕ is continuous. Its Fourier expansion $\phi(x) = \sum_{n \in \mathbb{Z}} \widehat{\phi}(n) e^{2\pi i n x}$, where $\widehat{\phi}(n) = \int_0^1 \phi(t) e^{2\pi i n t} dt = \sum_{n \in \mathbb{Z}} \int_0^1 f(n+t) e^{-2\pi i n t} dt = \int_{-\infty}^{\infty} f(t) e^{-2\pi i n t} dt = \widehat{f}(n)$.

If $\psi : \mathbb{R} \rightarrow \mathbb{C}$ is continuous and periodic, so $\psi(x) = \psi(x+1)$, and its series of Fourier coefficients is absolutely convergent, then the Fourier series converges uniformly.

We then evaluate at $x = 0$: $\phi(0) = \sum_{n \in \mathbb{Z}} f(n) = \sum_{n \in \mathbb{Z}} \widehat{f}(n)$. □

We can use this to prove more properties of the θ function. We then apply this to $f_y(x) = e^{-\pi x^2 y}$, as then $\theta(iy) = \sum_{n \in \mathbb{Z}} f_y(n)$.

We compute:

$$\begin{aligned} \widehat{f_y}(s) &= \int_{-\infty}^{\infty} e^{-\pi x^2 y} e^{-2\pi i x s} dx \\ &= \int_{-\infty}^{\infty} e^{-\pi(x\sqrt{y} + is/\sqrt{y})^2} e^{-\pi s^2/y} dx \\ &= \int_{-\infty}^{\infty} e^{-\pi(x + is/\sqrt{y})^2} e^{-\pi s^2/y} dx / \sqrt{y} \\ &= \frac{e^{-\pi s^2/y}}{\sqrt{y}} \int_{-\infty}^{\infty} e^{-\pi x^2} dx \\ &= e^{-\pi s^2/y} / \sqrt{y} \end{aligned}$$

So $\theta(iy) = \sum_{n \in \mathbb{Z}} e^{-\pi n^2/y} \sqrt{y}^{-1} = \theta(1/y) \sqrt{y}^{-1}$. So $\theta(\tau) = \theta(-1/\tau) \sqrt{\tau/i}^{-1}$, where $\sqrt{\cdot}$ is the branch of the square root defined on $\Re(z) > 0$, which takes positive real values on $\mathbb{R}_{>0}$.

It then follows that $\theta^8(\tau+2) = \theta^8(\tau)$, and $\theta^8(-1/\tau)\tau^{-4} = \theta^8(\tau)$, or equivalently that $\theta^8|_4 \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \theta^8$, and $\theta^8|_4 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \theta^8$.

Proposition 3.13. $\theta^8 \in M_4(\Gamma)$, where $\Gamma \leq \Gamma(1)$ is the subgroup of matrices γ such that $\gamma \equiv 1 \pmod{2}$ or $\gamma \equiv S \pmod{2}$.

Proof. We know θ^8 is holomorphic in \mathbb{H} . We know from ex. sheet 1 that Γ is generated by $\pm \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

This shows that θ^8 is invariant under the weight 4 action of Γ . It remains to check that θ^8 is holomorphic at the cusps $\Gamma \cdot \infty$ and $\Gamma \cdot 1$.

Since $\theta^8 = \left(1 + \sum_{n \geq 1} q_2^{n^2}\right)^8$, we see that θ^8 is holomorphic at ∞ .

To show holomorphy at $\Gamma \cdot 1$, we need to show $\theta^8|_4 \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ is holomorphic at ∞ . We'll show that $\theta^8|_4 \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ vanishes at ∞ .

We start by looking at $\theta(\tau+1) = \sum_{n \in \mathbb{Z}} (-1)^n q_2^{n^2}$. We thus have $\theta(\tau) + \theta(\tau+1) = 2\theta(4\tau)$.

So $\theta(-1/\tau) + \theta((\tau-1)/\tau) = 2\theta(-4/\tau)$, and hence:

$$\theta\left(\frac{\tau-1}{\tau}\right)\sqrt{\tau/i}^{-1} = \theta(-4/\tau)\sqrt{\tau/4i}^{-1} - \theta(-1/\tau)\sqrt{\tau/i}^{-1} = \theta(\tau/4) - \theta(\tau)$$

It then follows that $\theta^8|_4 \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = (\theta(\tau/4) - \theta(\tau))^8$. Hence θ^8 is holomorphic at vanishes at the cusp $\Gamma \cdot 1$. \square

One application of θ functions is the expression:

$$\theta^k(\tau) = \left(\sum_{n \in \mathbb{Z}} q_2^{n^2}\right)^k = \sum_{n \in \mathbb{Z}_{\geq 0}} r_k(n) q_2^n$$

where $r_k(n)$ is the number of ways of writing n as a sum of k squares.

For $k = 8$, we know already that $\theta^8 \in M_4(\Gamma)$. What other elements are there? We know $M_4(\Gamma(1)) \leq M_4(\Gamma)$, so $E_4 \in M_4(\Gamma)$. We also have $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \Gamma \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}^{-1} \leq \Gamma(1)$ via a direct computation.

This shows that $E_4|_4 \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} = \frac{1}{2}E_4(\frac{\tau+1}{2}) \in M_4(\Gamma)$.

We will prove later on that $\dim_{\mathbb{C}} M_4(\Gamma) = 2$, so $E_4(\tau), E_4(\frac{\tau+1}{2})$ span $M_4(\Gamma)$.

So $\theta^8(\tau) = \lambda E_4(\tau) + \mu E_4(\frac{\tau+1}{2})$ for $\lambda, \mu \in \mathbb{C}$. We can either compute λ, μ via q expansions, or using vanishing at cusps.

We know $\theta^8(\infty) = 1$ and $\theta^8|_4 \left[\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \right](\infty) = 0$.

We also have $E_4(\infty) = 1$ and $E_4|_4 \left[\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \right](\infty) = E_4(\infty) = 1$, whilst $E_4(\frac{\tau+1}{2}) = 1 + \sum_{n \geq 1} \sigma_3(n)(-1)^n q_2^n$, so $E_4(\frac{\tau+1}{2})|_{\tau=\infty} = 1$. We just need to compute $2E_4|_4 \left[\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \right](\infty) = 2 \cdot 2^3 E_4(2\tau) = 16E_4(2\tau)$, so at infinity this is 16.

So $\mu = -\lambda/16$, $\lambda + \mu = 1$, and hence $\lambda = 16/15$, $\mu = -1/15$. We conclude that $\theta^8 = \frac{16}{15}E_4(\tau) - \frac{1}{15}E_4(\frac{\tau+1}{2})$, and so $r_8(n) = \frac{16}{15} \cdot 240 \cdot \sigma_3(n/2) - \frac{1}{15}(-1)^n 240 \sigma_3(n) = 2^8 \sigma_3(n/2) - (-1)^n 2^4 \sigma_3(n)$. So if p is an odd prime, then e.g. $r_8(p) = 16\sigma_3(p)$.

Lemma 3.14. *If $\Gamma \leq \Gamma(1)$ is a congruence subgroup, then $(GL_2(\mathbb{Q})^+, \Gamma)$ is a Hecke pair.*

Proof. We need to show that, for all $g \in GL_2(\mathbb{Q})^+$, $\Gamma \backslash \Gamma g \Gamma$ is finite. Equivalently, $g^{-1}\Gamma g \cap \Gamma \backslash \Gamma$ is finite.

Choose $N \in \mathbb{N}$ such that $\Gamma(N) \leq \Gamma$ and $Ng, Ng^{-1} \in M_2(\mathbb{Z})$.

We have $g\Gamma(N^3)g^{-1} \subseteq 1 + gN^3M_2(\mathbb{Z})g^{-1} \subseteq 1 + (Ng)NM_2(\mathbb{Z})(Ng^{-1}) \subseteq 1 + NM_2(\mathbb{Z})$. Hence $g\Gamma(N^3)g^{-1} \leq \Gamma(N)$, so $\Gamma(N^3) \leq g^{-1}\Gamma g \cap \Gamma$. \square

Definition 3.15. *Let $k \in \mathbb{Z}$, and let $\mathcal{M}_k = \bigcup_{\Gamma \leq \Gamma(1)} M_k(\Gamma)$, and similarly \mathcal{S}_k .*

Proposition 3.16. *$GL_2(\mathbb{Q})^+$ acts on \mathcal{M}_k and \mathcal{S}_k by the weight k action.*

Proof. We need to show that, if $f \in M_k(\Gamma)$ and $g \in GL_2(\mathbb{Q})^+$, then there exists $\Gamma' \leq \Gamma(1)$ such that $f|_k[g] \in M_k(\Gamma')$. $f|_k[g]$ is invariant under the weight k action of $g^{-1}\Gamma g$.

We can always find a congruence subgroup $\Gamma' \leq g^{-1}\Gamma g$, and we've already proven in this case that $f|_k[g] \in M_k(\Gamma')$. \square

Corollary 3.17. *$\mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma)$ acts on $M_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$.*

We've seen that $\mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma(1))$ is commutative, and $\mathcal{S}_k(\Gamma(1))$ decomposes with multiplicity one under the action of this Hecke algebra. However, in general, $\mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma)$ is not commutative.

Recall $\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) : c \equiv 0 \pmod{N} \right\}$ and $\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) : a \equiv d \equiv 1 \pmod{N}, c \equiv 0 \pmod{N} \right\}$.

For any $N \in \mathbb{N}$, $SL_2(\mathbb{Z}) \rightarrow SL_2(\mathbb{Z}/N\mathbb{Z})$ is surjective, and in particular, $\Gamma_1(N) \triangleleft \Gamma_0(N)$, with $\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/N\mathbb{Z})^\times$, mapping to $d \pmod{N}$.

If $d \in \mathbb{N}$ and $(d, N) = 1$, then we write $\langle d \rangle$ for the endomorphism of $M_k(\Gamma_1(N))$ induced by $\left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_0(N) : D \equiv d \pmod{N} \right\} \in \mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma_1(N))$.

If p is prime, then we write T_p for the endomorphism of $M_k(\Gamma)$ induced by $[\Gamma_1(N) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma_1(N)] \in \mathcal{H}(GL_2(\mathbb{Q})^+, \Gamma_1(N))$. If $N = 1$, then this is the same as T_p .

In general, if $p \nmid N$, then T_p behaves in a similar way to the $N = 1$ case, but if $p|N$, then T_p is of a different nature.

Let's first discuss the operators $\langle d \rangle$. We can think about the action of these operators in another way: since $\Gamma_1(N) \triangleleft \Gamma_0(N)$, then $\Gamma_0(N)/\Gamma_1(N)$ acts on $\mathcal{M}_k^{\Gamma_1(N)} = M_k(\Gamma(N))$ via the weight k action.

Lemma 3.18. *If $f \in M_k(\Gamma_1(N))$ and $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_0(N)$, and $D \equiv d \pmod{N}$, then $f|_k[D] = \langle d \rangle(f)$.*

Proof. To compute the RHS, we decompose $\left\{ \begin{pmatrix} X & Y \\ Z & W \end{pmatrix} \in \Gamma_0(N) : w \equiv d \pmod{N} \right\}$ as a disjoint union $\bigsqcup_i \Gamma_1(N)x_i$.

Then $\langle d \rangle(f) = \sum_i f|_k[x_i]$. In fact this double coset is $\Gamma_1(N) \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, so $\langle d \rangle(f) = f|_k[D] = \text{LHS}$. \square

Lemma 3.19. *If A is a finite abelian group acting by linear maps on a \mathbb{C} -vector space V , then $V = \bigoplus_{\chi \in \text{Hom}(A, \mathbb{C}^\times)} V_\chi$, where $V_\chi = \{v \in V : \forall a \in A, a \cdot v = \chi(a)v\}$.*

Corollary 3.20.

$$M_k(\Gamma_1(N)) = \bigoplus_{\chi} M_k(\Gamma_1(N), \chi)$$

and similarly for $S_k(\Gamma_1(N))$, where the sums run over the set of homomorphisms $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$, and $M_k(\Gamma_1(N), \chi) = \{f \in M_k(\Gamma_1(N)) : \forall d \in \mathbb{Z}, (d, N) = 1 \implies \langle d \rangle f = \chi(d)f\}$.

Proposition 3.21.

1. If $p|N$ then $\Gamma_1(N) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma_1(N) = \bigsqcup_{j=0}^{p-1} \Gamma_1(N) \begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}$.
2. If $p \nmid N$, then $\Gamma_1(N) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma_1(N) = \bigsqcup_{j=0}^{p-1} \Gamma_1(N) \begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix} \sqcup \Gamma_1(N) \begin{pmatrix} p^a & r \\ pN & p^j \end{pmatrix}$.

Proof. For any $g \in GL_2(\mathbb{Q})^+$, there is a bijection $g^{-1}\Gamma_1(N)g \cap \Gamma_1(N) \backslash \Gamma_1(N) \rightarrow \Gamma_1(N) \backslash \Gamma_1(N)g\Gamma_1(N)$, given by $\gamma \mapsto \Gamma_1(N)g\gamma$. I got fed up with latexing matrices here - the proof is a big computation (lecture 17, 41:40). \square

So, if $f \in M_k(\Gamma_1(N))$, then if $p|N$ we have:

$$T_p f = \sum_{j=0}^{p-1} f|_k \left[\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \gamma_j \right]$$

where $\gamma_0, \dots, \gamma_{p-1}$ in $\Gamma_1(N)$ project to representatives for the quotient

$$\{SL_2(\mathbb{Z}/N\mathbb{Z}) : c = 0, b = 0\} \setminus \{SL_2(\mathbb{Z}/N\mathbb{Z}) : c = 0\}$$

We can take $\gamma_j = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}$.

If $p \nmid N$, then we have

$$T_p f = \sum_{j=0}^p f|_k \left[\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \gamma_j \right]$$

where $\gamma_0, \dots, \gamma_p \in \Gamma_1(N)$ project to representatives for the quotient

$$\{SL_2(\mathbb{Z}/p\mathbb{Z}) : b = 0\} \setminus SL_2(\mathbb{Z}/p\mathbb{Z})$$

We can take $\gamma_j = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}$ for $0 \leq j \leq p-1$, and $\gamma_p = \begin{pmatrix} p^a & r \\ N & s \end{pmatrix}$ where $a \in \mathbb{N}$ such that $p^a \equiv 1 \pmod{N}$, and $p^a s - Nr = 1$.

By definition, $\langle d \rangle(f) = f|_k[S]$ for any $S = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_0(N)$ such that $D \equiv d \pmod{N}$.

Lemma 3.22. $\langle d \rangle$ and T_p commute.

Corollary 3.23. The operators T_p preserve the subspaces $M_k(\Gamma_1(N), \chi)$ and $S_k(\Gamma_1(N), \chi)$ for characters $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$.

Proof of lemma. We just treat the case $p \nmid N$. Choose $\delta \in \Gamma_0(N)$ with entries A, B, C, D , congruent to the identity matrix mod p^2 . Then:

$$\langle d \rangle(T_p(f)) = \sum_{j=0}^p f|_k \left[\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \gamma_j \delta \right]$$

$$T_p(\langle d \rangle(f)) = \sum_{j=0}^p f|_k \left[\delta \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \gamma_j \right]$$

We can rewrite $\langle d \rangle(T_p(f)) = \sum_{j=0}^p f|_k \left[\delta \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}^{-1} \delta^{-1} \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \gamma_j \delta \right]$.

This will equal $T_p(\langle d \rangle(f))$ if the matrices $\varepsilon_j = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}^{-1} \delta^{-1} \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \gamma_j \delta$ lie in $\Gamma_1(N)$ and project to representatives for the quotient

$$\{SL_2(\mathbb{Z})/p\mathbb{Z} : b = 0\} \setminus SL_2(\mathbb{Z}/p\mathbb{Z})$$

Now $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}^{-1} \delta^{-1} \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} = \begin{pmatrix} D & -pB \\ -C/p & A \end{pmatrix} \in \Gamma_0(N)$ and congruent to $I \pmod{p}$. We see that $\varepsilon_j \in \Gamma_1(N)$ and congruent to $\gamma_j \pmod{p}$. \square

Proposition 3.24. Let $f \in M_k(\Gamma_1(N))$ and write $f(\tau) = \sum_{n \geq 0} a_n(f) q^n$. Then:

$$T_p(f) = \sum_{n \geq 0} [a_{np}(f) + p^{k-1} \delta_{p \nmid N} a_{n/p}(\langle p \rangle f)] q^n$$

where $\delta_{p \nmid N}$ is 1 if $p \nmid N$ and 0 otherwise, and $a_{n/p} = 0$ if $p \nmid n$.

Proof.

$$\begin{aligned}
T_p(f) &= \sum_{j=0}^{p-1} f|_k \left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix} \right] + \delta_{p \nmid N} f|_k \left[\begin{pmatrix} p^a & r \\ pN & ps \end{pmatrix} \right] \\
&= p^{k-1} p^{-k} \sum_{j=0}^{p-1} f\left(\frac{\tau+j}{p}\right) + \delta_{p \nmid N} f|_k \left[\begin{pmatrix} p^a & r \\ N & ps \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right] \\
&= \frac{1}{p} \sum_{n \geq 0} \sum_{j=0}^{p-1} a_n(f) e^{2\pi i n \tau / p} e^{2\pi i n j / p} + \delta_{p \nmid N} (\langle p \rangle(f))|_k \left[\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \right] \\
&= \sum_{n \geq 0} a_{np}(f) q^n + \delta_{p \nmid N} \sum_{n \geq 0} a_n(\langle p \rangle(f)) q^{np}
\end{aligned}$$

□

Corollary 3.25. *If p, q are primes, then T_p, T_q commute.*

Proof. WLOG $p \neq q$.

$$\begin{aligned}
a_n(T_p(T_q(f))) &= a_{np}(T_q(f)) + p^{k-1} \delta_{p \nmid a_{n/p}} (\langle p \rangle T_q(f)) \\
&= a_{npq}(f) + \delta_{q \nmid N} a_{np/q}(\langle q \rangle(f)) q^{k-1} + p^{k-1} \delta_{p \nmid N} (a_{nq/p}(\langle p \rangle(f)) + a_n/pq(\langle q \rangle \langle p \rangle f) \delta_{q \nmid N} q^{k-1}) \\
&= a_{npq}(f) + q^{k-1} \delta_{q \nmid N} a_{np/q}(\langle q \rangle(f)) + p^{k-1} \delta_{p \nmid N} a_{nq/p}(\langle p \rangle(f)) + p^{k-1} q^{k-1} \delta_{p \nmid N} \delta_{q \nmid N} a_{n/pq}(\langle p \rangle \langle q \rangle(f))
\end{aligned}$$

which is symmetric in p, q .

□

Proposition 3.26. *Define an inner product $\langle \cdot, \cdot \rangle$ on $S_k(\Gamma_1(N))$ by*

$$\langle f, g \rangle = [\Gamma(1) : \Gamma_1(N)]^{-1} \int_{\Gamma_1(N) \backslash \mathbb{H}} f(\tau) \overline{g(\tau)} y^k dx dy / y^2$$

If $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, its adjoint is $\langle d \rangle^ = \langle d^{-1} \rangle$. If $p \nmid N$, the adjoint of T_p is $T_p^* = \langle p^{-1} \rangle T_p$.*

Corollary 3.27. *The operators $\langle d \rangle$ and T_p for $p \nmid N$ are simultaneously diagonalisable on $S_k(\Gamma_1(N))$.*

Remark. If $p|N$, then T_p is not diagonalisable on $S_k(\Gamma_1(N))$ in general. The simultaneous eigenspaces of $\langle d \rangle, T_p$ for $p \nmid N$ not 1-dimensional - both of these defects are explained by the presence of “oldforms”. Suppose $N > 1$ and $f \in S_k(\Gamma(1))$ is a normalised eigenform. Then $f(\tau), f(N\tau) = g(\tau) \in S_k(\Gamma_1(N))$ are linearly independent. If $p \nmid N$ is a prime, then f, g are eigenvectors of T_p with the same eigenvalue, $a_p(f)$.

$T_p f$ is the same considered either in $S_k(\Gamma(1))$ or in $S_k(\Gamma_1(N))$, so $T_p f = a_p(f)f$. We have $g(\tau) \sum_{n \geq 1} a_n(f) q^{Nn} = \sum_{n \geq 1} a_{n/N}(f) q^n$, so $a_n(T_p g) = a_{np}(g) + p^{k-1} a_{n/p}(g) = a_{np/N}(f) + p^{k-1} a_{n/pN}(f) = a_{n/N}(T_p f) = a_p(f) a_{n/N}(f) = a_p(f) a_n(g)$.

Definition 3.28. *Let $N \in \mathbb{N}$. The old subspace $S_k^{old}(\Gamma_1(N)) \leq S_k(\Gamma_1(N))$ is the span of all modular forms of type $f(M\tau)$ where $M|N, M > 1$, and $f \in S_k(\Gamma_1(N/M))$.*

The new subspace $S_k^{new}(\Gamma_1(N))$ is the orthogonal complement with respect to the inner product of the old subspace.

Theorem 3.29. All of the operators $\langle d \rangle, T_p$ preserve the direct sum decomposition $S_k(\Gamma_1(N)) = S_k^{old}(\Gamma_1(N)) \oplus S_k^{new}(\Gamma_1(N))$.

The simultaneous eigenspaces of the operators $\langle d \rangle, T_p (p \nmid N)$ on $S_k^{new}(\Gamma_1(N))$ are 1-dimensional.

$S_k^{new}(\Gamma_1(N))$ has a unique basis of normalised eigenforms, i.e. forms $f \in S_k^{new}(\Gamma_1(N))$ which are eigenvectors for all operators $\langle d \rangle, T_p$ which satisfy $a_1(f) = 1$.

Definition 3.30. A **newform** of weight k and level $\Gamma_1(N)$ is a normalised eigenform in $S_k^{new}(\Gamma_1(N))$.

Example. $S_2(\Gamma_1(1)) = 0$. If p is a prime, then $S_2(\Gamma_1(p)) = S_2^{new}(\Gamma_1(p))$. The first prime p for which this space is nonzero is $p = 11$, in which case it has dimension 1.

Recall $\eta(\tau) = q_{24} \prod_{n \geq 1} (1 - q^n)$. Then $\eta(\tau)^2 \eta(11\tau)^2$ is the unique newform in $S_2(\Gamma_1(11))$.

Theorem 3.31. Let $f \in S_k(\Gamma_1(N))$ be a newform, say $f(\tau) = \sum_{n \geq 1} a_n q^n$, and define $L(f, s) = \sum_{n \geq 1} a_n n^{-s}$. Then:

1. $L(f, s)$ is absolutely convergent in $\Re(s) > k/2 + 1$.
2. There is a character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \rightarrow \mathbb{C}^\times$ such that $f \in S_k(\Gamma_1(N), \chi)$.
3. There's an Euler product $L(f, s) = \prod_p (1 - a_p p^{-s} + \chi(p) p^{k-1-2s})^{-1}$ - here we make the convention that $\chi(p) = 0$ if $p \mid N$.
4. $L(f, s)$ has an analytic continuation to \mathbb{C} and if we define

$$\Lambda(f, s) = N^{s/2} (2\pi)^{-s} \Gamma(s) L(f, s)$$

then there's a sign $\varepsilon_f \in \{\pm 1\}$ such that $\Lambda(f, s) = \varepsilon_f \Lambda(f, k - s)$.

The Shimura-Taniyama-Weil conjecture (now the Modularity Theorem) asserts that there is a bijection between elliptic curves E/\mathbb{Q} up to isogeny and newforms $f \in S_2(\Gamma_1(N), \mathbb{1})$ such that, for all $n \in \mathbb{N}$, $a_n(f) \in \mathbb{Z}$, given by $E \mapsto f$ if, for all but finitely many primes p , $a_p(f) = p + 1 - |\tilde{E}(\mathbb{F}_p)|$.

4 Modular Curves and Riemann Surfaces

The goal of this section is to endow $\Gamma \backslash \mathbb{H}$ and $\Gamma \backslash \mathbb{H} \sqcup \Gamma \backslash \mathbb{P}^1(\mathbb{Q})$ with the structure of Riemann surfaces. These quotients are called modular curves.

Definition 4.1. A **Riemann surface** is a compact Hausdorff topological space X together with an atlas $\mathcal{A} = \{(U_i, V_i, \phi_i : U_i \rightarrow V_i) : i \in I\}$ where

1. $U_i \subset X$ is open and $V_i \subset \mathbb{C}$ is open.
2. For all $i, j \in I$, the map $\phi_i(U_i \cap U_j) \rightarrow \phi_j(U_i \cap U_j)$ is holomorphic.
3. $\bigcup_{i \in I} U_i = X$.

If X, Y are Riemann surfaces, then a morphism $f : X \rightarrow Y$ is a continuous map such that, for all charts $\phi_i : U_i \rightarrow V_i$ of X and $\psi_j : U'_j \rightarrow V_j$ of Y , the induced map $V_i \cap f^{-1}(U'_j) \rightarrow V'_j$ is holomorphic.

Examples. $X = \mathbb{C}, \mathcal{A} = \{(\mathbb{C}, \mathbb{C}, \text{id} : \mathbb{C} \rightarrow \mathbb{C})\}$. If Y is another Riemann surface, then a morphism $f : Y \rightarrow \mathbb{C}$ is a continuous function such that, for every chart $\phi_i : U_i \rightarrow V_i$ of Y , the map $f \circ \phi_i^{-1} : V_i \rightarrow \mathbb{C}$ is holomorphic.

$X = \widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. We have charts $(\mathbb{C}, \mathbb{C}, \text{id} : \mathbb{C} \rightarrow \mathbb{C})$, and $(\mathbb{C}^\times \cup \{\infty\}, \mathbb{C}, z \mapsto 1/z)$. Then $\phi_2 \circ \phi_1^{-1} : \mathbb{C}^\times \rightarrow \mathbb{C}^\times$ is $z \mapsto 1/z$.

The first goal: we can make $Y(\Gamma) = \Gamma \backslash \mathbb{H}$ for $\Gamma \leq \Gamma(1)$ a congruence subgroup into a connected Hausdorff topological space. We give \mathbb{H} its usual topology and $Y(\Gamma)$ the quotient topology. So, if $\pi : \mathbb{H} \rightarrow Y(\Gamma)$ is the quotient map then a subset $U \subset Y(\Gamma)$ is open if and only if $\pi^{-1}(U)$ is open.

Lemma 4.2. *Suppose $A, B \subset \mathbb{H}$ are compact subsets. Then $\{\gamma \in \Gamma(1) : \gamma A \cap B \neq \emptyset\}$ is finite.*

Proof. $\Omega_X = \{\tau \in \mathbb{H} : |\Re \tau| \leq X, \Im \tau \geq 1/X\}$. We showed that there is $C_X > 0$ such that, for all $\tau \in \Omega_X$ and $t \in \mathbb{R}$, $|\tau + t| \geq C_X \sup(1, |t|)$.

If $c, d \in \mathbb{Z}$, then $|c\tau + d| \geq C_X \sup(|c|, |d|)$.

So, if $\tau \in \Omega_X$, then $\Im(\gamma\tau) \leq \frac{\Im(\tau)}{C_X^2 \sup(|c|, |d|)^2}$, and hence $\sup(|c|, |d|)^2 \leq \frac{\Im(\tau)}{C_X^2 \Im(\gamma\tau)}$.

Now suppose that $A, B \subseteq \Omega_X$. If $\tau \in A$, $\gamma\tau \in B$, then $\sup(|c|, |d|)$ is bounded above by a constant only depending on A and B .

If $\gamma \in \Gamma(1)$ and $\gamma A \cap B \neq \emptyset$, then there are only finitely many possibilities for (c, d) . Now observe that γ, δ have the same bottom row, then $\delta = \varepsilon\gamma$ for some $\varepsilon \in \Gamma_\infty = \{SL_2(\mathbb{Z}) : a = d = 1, c = 0\} = \text{Stab}_{SL_2(\mathbb{Z})}(\infty)$.

Since A, B are compact, there can only be finitely many ε such that $\varepsilon\gamma A \cap B \neq \emptyset$. \square

Proposition 4.3. *$Y(\Gamma)$ is Hausdorff.*

Proof. Let $x, y \in \mathbb{H}$ be such that x, y are not Γ -conjugate. It will be enough to find open sets $W_x \ni x$ and $W_y \ni y$ of \mathbb{H} such that, if $\gamma \in \Gamma(1)$ then $\gamma W_x \cap W_y = \emptyset$.

We begin by choosing open subsets $U_x \ni x$ and $U_y \ni y$, and compact subsets $K_x \supseteq U_x$ and $K_y \supseteq U_y$ of \mathbb{H} .

The previous lemma tells us that the set $N = \{\gamma \in \Gamma : K_x \cap \gamma K_y \neq \emptyset\}$ is finite. For all $\gamma \in N$, we can choose open sets $A_\gamma \ni x$, $B_\gamma \ni \gamma y$ such that $A_\gamma \cap B_\gamma = \emptyset$, as \mathbb{H} is Hausdorff.

We now take $W_x = U_x \cap \bigcap_{\gamma \in N} A_\gamma$, an open set containing x , and $W_y = U_y \cap \bigcap_{\gamma \in N} \gamma^{-1} B_\gamma$, an open set containing y .

Then for all $\gamma \in \Gamma$, $W_x \cap \gamma W_y = \emptyset$. Indeed, $W_x \subset U_x \subset K_x$ and $W_y \subset U_y \subset K_y$, so if $W_x \cap \gamma W_y \neq \emptyset$, then $\gamma \in N$. But if $\gamma \in N$, then $W_x \subset A_\gamma$ and $\gamma W_y \subset B_\gamma$, and $A_\gamma \cap B_\gamma = \emptyset$, so $W_x \cap \gamma W_y = \emptyset$. \square

The next goal is to define the topology of $X(\Gamma) = \Gamma \backslash \mathbb{H} \sqcup \Gamma \backslash \mathbb{P}^1(\mathbb{Q})$.

Definition 4.4. We set $\mathbb{H}^* = \mathbb{H} \sqcup \mathbb{P}^1(\mathbb{Q})$. We let $GL_2(\mathbb{Q})^+$ act on \mathbb{H}^* by the usual action on \mathbb{H} and $\mathbb{P}^1(\mathbb{Q})$.

We define a topology on \mathbb{H}^* by declaring \mathbb{H} to be an open subspace with its usual topology, and by giving $z = \alpha\infty$ the basis of open nbhds $\alpha V_R \cup \{z\} = \alpha(V_R \cup \{\infty\})$, where $V_R = \{\tau \in \mathbb{H} : \Im(\tau) > R\}$. Note that $\alpha(V_r \cup \{\infty\})$ only depends on z and R , and not on the choice of α .

We give $X(\Gamma)$ the quotient topology

If $\gamma \in \Gamma(1)$, then $\gamma : \mathbb{H}^* \rightarrow \mathbb{H}^*$ is a homeomorphism. We define $X(\Gamma) = \Gamma \backslash \mathbb{H}^*$, endowed with the quotient topology. We write $\pi : \mathbb{H}^* \rightarrow X(\Gamma)$ for the quotient map.

Why is the topology defined this way? If $R > 2$, then for all $\gamma \in \Gamma(1)$, $\gamma V_R \cap V_R \neq \emptyset \implies \gamma \in \Gamma_\infty = \text{Stab}_{\Gamma(1)}(\infty)$. Indeed, if $\tau \in \gamma V_R \cap V_R$, say $\tau = \gamma \tau'$ for some $\tau' \in V_R$, then there is $\delta, \varepsilon \in \Gamma_\infty$ such that $\delta \tau, \varepsilon \tau'$ lie in \mathcal{F} .

We can further require $\delta \tau, \varepsilon \tau$ have real part in $[-\frac{1}{2}, \frac{1}{2}]$. It follows that $\delta \tau = \varepsilon \tau'$, since they are $\Gamma(1)$ -conjugate. We thus have $\tau = \delta^{-1} \varepsilon \gamma \tau$, so $\delta^{-1} \varepsilon \gamma = \pm 1$, so $\gamma \in \Gamma_\infty$.

We have a commutative diagram:

$$\begin{array}{ccc} \mathbb{H}^* & \longrightarrow & X(\Gamma) \\ \uparrow & & \uparrow \\ V(R) & \longrightarrow & V_R/\Gamma \cap \Gamma_\infty \end{array}$$

The map $V_R/\Gamma \cap \Gamma_\infty \rightarrow X(\Gamma)$ is injective, continuous, and open, so a homeomorphism onto its image.

We have a continuous bijection $V_R/\Gamma \cap \Gamma_\infty \rightarrow D_R = \{q_h \in \mathbb{C} : |q_h| < e^{-2\pi R/h}\}$ given by the formula $\infty \mapsto 0, \Gamma \cap \Gamma_\infty \cdot \tau \mapsto e^{2\pi i \tau/h}$. This map is open by the open mapping theorem from complex analysis. Hence $\pi(V_R)$ is an open neighbourhood of $\pi(\infty)$ in $X(\Gamma)$, which is homeomorphic to the disc $D_R \subseteq \mathbb{C}$. Similarly, if $\alpha \in \Gamma(1)$ and $\Gamma \cdot \alpha \infty$ is a cusp of Γ of width h , then $\pi(\alpha V_R)$ for any $R > 0$ is an open neighbourhood of $\pi(\alpha \infty)$ which is homeomorphic to the disc D_R by the map which sends $\alpha \infty \mapsto 0$, and $\Gamma \cdot \tau$ for $\tau \in \alpha V_R \mapsto e^{2\pi i \alpha^{-1} \tau/h}$.

Theorem 4.5. $X(\Gamma)$ is connected, Hausdorff, and compact.

Proof. \mathbb{H}^* is connected - if $\mathbb{H}^* = U \sqcup V$ for U, V open, then $\mathbb{H} = (U \cap \mathbb{H}) \cup (V \cap \mathbb{H})$, so without loss of generality $\mathbb{H} \subseteq U$, and $V \subset \mathbb{P}^1(\mathbb{Q})$. If $z \in \mathbb{P}^1(\mathbb{Q})$, then $\alpha V_R \subset U$ for some R , but $\alpha V_R \cap \mathbb{H} = \alpha V_R \cap U \neq \emptyset$, so $V = \emptyset$, and so \mathbb{H}^* is connected.

For $X(\Gamma)$ is Hausdorff, it's enough to take the pairs $\Gamma \cdot \tau, \Gamma \cdot \alpha \infty$, and $\Gamma \alpha \infty, \Gamma \beta \infty$, and construct disjoint open neighbourhoods in each case.

In the first case, we can find an open set $U_\tau \subset \mathbb{H}$ and a compact set $K_\tau \subset \mathbb{H}$ such that $\tau \in U_\tau \subset K_\tau$. We saw last time that we can find a constant $C_\tau > 0$ such that $\Im(\gamma \tau) \leq C_\tau$ for all $\gamma \in \Gamma(1)$ and $z \in K_\tau$.

If $R > C_\tau$, then $\gamma K_\tau \cap V_R = \emptyset$ for all $\gamma \in \Gamma(1)$, and hence $\gamma U_\tau \cap \alpha V_R = \emptyset$ for all $\gamma \in \Gamma$.

This means that $\pi(U_\tau)$ and $\pi(\alpha V_R)$ are disjoint.

If $\Gamma \alpha \infty, \Gamma \beta \infty$ are distinct cusps, then $\alpha V_R, \beta V_R$ are distinct when $R > 2$. If $\gamma \in \Gamma$ and $\gamma \alpha V_R \cap \beta V_R \neq \emptyset$, then there is $\delta \in \Gamma_\infty$ such that $\gamma \alpha = \beta \delta$, and so we have $\gamma \alpha \infty = \beta \infty$.

Hence $\pi(\alpha V_R), \pi(\beta V_R)$ are disjoint.

For compactness, we first show that $\mathcal{F}^* = \mathcal{F} \cup \{\infty\}$ is compact. It's enough to show that, for all $R > 0$, $\mathcal{F}^* - V_R$ is compact. But $\mathcal{F}^* - V_R = \{\tau \in \mathbb{H} : |\tau| \geq 1, \Re \tau \in [-\frac{1}{2}, \frac{1}{2}], \Im \tau \leq R\}$. This is closed and bounded in \mathbb{H} so is compact.

If $\gamma_1, \dots, \gamma_n$ is a transversal for $\Gamma \backslash \Gamma(1)$, then $\pi(\bigcup_{i=1}^n \gamma_i \mathcal{F}^*) = X(\Gamma)$. So $X(\Gamma)$ is a continuous image of a compact set, so compact. \square

We now show how to give $X(\Gamma)$ the structure of a Riemann surface, by giving explicit charts.

We let I be a set of representatives for the Γ -orbits in \mathbb{H}^* . For each $x \in I$, we construct a chart $\phi_x : U_x \rightarrow V_x$ where $U_x \subset X(\Gamma)$ is an open neighbourhood of $\pi(x)$, and $V_x \subset \mathbb{C}$ is an open subset.

If $x \in \mathbb{P}^1(\mathbb{Q})$, say $x = \alpha\infty$ for some $\alpha \in \Gamma(1)$, then we define $U_x = \pi(\alpha V_z)$, $V_x = \{q_h \in \mathbb{C} : |q_h| < e^{-2\pi R/h}\}$, and $\phi_x(\pi(\alpha\infty)) = 0$, $\phi_x(\pi(\alpha\tau)) = e^{2\pi i\tau/\mathbb{H}}$ if $\tau \in V_z$. We've already seen that ϕ_x is a homeomorphism.

If $x \in \mathbb{H}$, we first show that we can find an open set $W_x \subset \mathbb{H}$ with the following properties:

1. $x \in W_x$ and W_x is stable under $\text{Stab}_\Gamma(x)$.
2. If $y \in W_x$ and $y \neq x$, then $\text{Stab}_\Gamma(y) = \Gamma \cap \{\pm 1\}$.

To do this, choose $M_x \subset \mathbb{H}$ open and $K_x \subset \mathbb{H}$ compact, such that $x \in M_x \subset K_x$, and let $N = \{\gamma \in \Gamma : \gamma K_x \cap K_x \neq \emptyset\}$. Then N is finite, and contains $\text{Stab}_\Gamma(x)$. If $\gamma \in N - \text{Stab}_\Gamma(x)$, then $\gamma x \neq x$, so we can find $A_\gamma, B_\gamma \subset \mathbb{H}$ open such that $x \in A_\gamma$, $\gamma x \in B_\gamma$, and $A_\gamma \cap B_\gamma = \emptyset$.

We define $W_x = M_x \cap \bigcap_{\gamma \in N - \text{Stab}_\Gamma(x)} (A_\gamma \cap \gamma^{-1}B_\gamma)$. This is an open subset of \mathbb{H} containing x . If $\gamma \in \Gamma$ and $W_x \cap \gamma W_x \neq \emptyset$, then $\gamma \in N$. If $\gamma \notin \text{Stab}_\Gamma(x)$, then $W_x \cap \gamma W_x \subset A_\gamma \cap B_\gamma = \emptyset$.

Consider the Möbius transformation $\psi_x(z) = \frac{z-x}{z-\bar{x}}$. Then $\psi_x(x) = 0$, $\psi_x(\infty) = \infty$, $\psi_x \text{Stab}_\Gamma(x) \psi_x^{-1}$ stabilises both 0 and ∞ .

Any Möbius transformation which stabilises 0 and ∞ and which is of finite order must be given by $z \mapsto \zeta z$ for some root of unity ζ . We can assume, after shrinking W_x , that $\psi_x(W_x)$ is an open disc D_ε for some $0 < \varepsilon < 1$. Then W_x is stable under $\text{Stab}_\Gamma(x)$.

If $y \in W_x$ and $y \neq x$ and $\delta \in \Gamma$ and $\delta y = y$. Then $\delta W_x \cap W_x \neq \emptyset$, so $\delta \in \text{Stab}_\Gamma(x)$, so $\psi_x \delta \psi_x^{-1}$ stabilises $\psi_x(y) \neq 0$, and so $\psi_x \delta \psi_x^{-1}$ is trivial, giving $\delta \in \Gamma \cap \{\pm 1\}$.

We conclude that $\pi(W_x)$ is an open subset of $X(\Gamma)$ which contains $\pi(x)$ and which is homeomorphic to $W_x / \text{Stab}_\Gamma(x) \cong D_\varepsilon / \psi_x \text{Stab}_\Gamma(x) \psi_x^{-1}$.

The group $\psi_x \text{Stab}_\Gamma(x) \psi_x^{-1} / \Gamma \cap \{\pm 1\}$ is cyclic of some order n_x , generated by $z \mapsto e^{2\pi i/n_x} \cdot z$. It follows that $D_\varepsilon / \psi_x \text{Stab}_\Gamma(x) \psi_x^{-1}$ is homeomorphic to $D_{\varepsilon^{n_x}}$, via the map $z \mapsto z^{n_x}$.

We can now write down our chart around the point $\pi(x)$: we take $U_x = \pi(W_x)$, $V_x = D_{\varepsilon^{n_x}}$, $\phi_x : U_x \rightarrow V_x$ defined by $\phi_x(\pi(z)) = \psi_x(z)^{n_x}$ when $z \in W_x$.