Contents

Modular Forms

Jingming Cui

1 The Classical Modular Forms

1.1 Periodic functions

Take $X = \mathbb{R}$ and $\Gamma = \mathbb{Z}$, so $\mathbb{R}/\mathbb{Z}$ is the circle. A function $f\colon \mathbb{R} \to \mathbb{C}$ is periodic of period 1 if

$$f(x+n) = f(x) \quad \text{for all } n \in \mathbb{Z}.$$

Recall that $f(x)$ has the Fourier series representation

$$f(x) = \sum_{n\in\mathbb{Z}} a_n e(nx),$$

where

$$e(x) = \mathrm{e}^{2\pi \mathrm{i} x} = \cos 2\pi x + \mathrm{i} \sin 2\pi x$$

and coefficients are given by

$$a_n = \int_0^1 f(x) e(-nx) \mathrm{d}x.$$

Example 1.1. Consider $\{x\} \coloneqq x - [x]$, we have

$$\{x\} = \frac{1}{2} - \sum_{n=1}^{\infty} \frac{\sin 2\pi n x}{\pi n}$$

which is boundedly convergent (the partial sums are uniformly bounded), but not absolutely. Say precisely, the Fourier series converges to $\{x\}$ for $x \notin \mathbb{Z}$ and $1/2$ for $x \in \mathbb{Z}$. Clearly the Fourier partial sums are uniformly bounded (there exists a constant M such that $|S_N(x)| \leq M$ for all N and x). Hence the convergence is boundedly convergent. However, the series is not absolutely convergent. If we take $x = 1/4$, then $\sin(2\pi n/4) = \sin(\pi n/2)$ equals to 0 for even n and ± 1 for odd n. The series becomes

$$\sum_{n=1}^{\infty} \left| \frac{\sin(2\pi n x)}{\pi n} \right| = \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1},$$

which diverges.

Let $f\colon \mathbb{R} \to \mathbb{C}$ be any function of rapid decay at $\pm\infty$ so that the series

$$g(x) = \sum_{n\in\mathbb{Z}} f(x+n)$$

converges absolutely. Clearly the function g is periodic of period 1. Then $g(x)$ admits Fourier expansion

$$g(x) = \sum_{n\in\mathbb{Z}} a_g(n)e(nx),$$

then

$$\begin{aligned} a_g(n) &= \int_0^1 g(x)e(-nx)\mathrm{d}x \\ &= \int_0^1 \sum_{m\in\mathbb{Z}} f(x+m)e(-nx)\mathrm{d}x \\ &= \int_0^1 \sum_{m\in\mathbb{Z}} f(x+m)e(-n(x+m))\mathrm{d}(x+m) \\ &= \int_{\mathbb{R}} f(x)e(-nx)\mathrm{d}x \\ &= \hat{f}(n). \end{aligned}$$

If we set $x = 0$, we obtain the **Poisson summation formula**

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

The left-hand side sums the values of the function f at all integer points (the lattice $\mathbb{Z}$), so is called a geometric sum (or lattice sum). The right-hand side sums the values of the Fourier transform $\hat{f}$ at all integer frequencies, so is called a spectral sum.

The space $X = \mathbb{R}^k$ is acted on by $\Gamma = \mathbb{Z}^k$ as integral vector translations. The Poisson summation formula asserts that

$$\sum_{\mathbf{n}\in\mathbb{Z}^k} f(\mathbf{n}) = \sum_{\mathbf{n}\in\mathbb{Z}^k} \hat{f}(\mathbf{n})$$

where

$$\hat{f}(\mathbf{y}) = \int_{\mathbb{R}^k} f(\mathbf{x})e(-\mathbf{xy})\mathrm{d}\mathbf{x}$$

and $\mathbf{xy} = x_1y_1 + \cdots + x_ky_k$ is the scalar product.

1.2 Modular functions

Now consider

$$X = \mathbb{H} = \{z = x + \mathrm{i}y \mid x \in \mathbb{R}, y \in \mathbb{R}^+\}$$

and

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

Γ acts on $\mathbb{H}$ through the linear fractional transformations

$$\gamma z = \frac{az + b}{cz + d} \quad \text{if} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

If we write

$$z = \frac{z_1}{z_2} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

then γz can be regarded as matrix multiplication

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix},$$

which is precisely linear. This explains “linear fractional transformations".

For

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}),$$

we have

$$\frac{\mathrm{d}gz}{\mathrm{d}z} = \frac{1}{(cz + d)^2}.$$

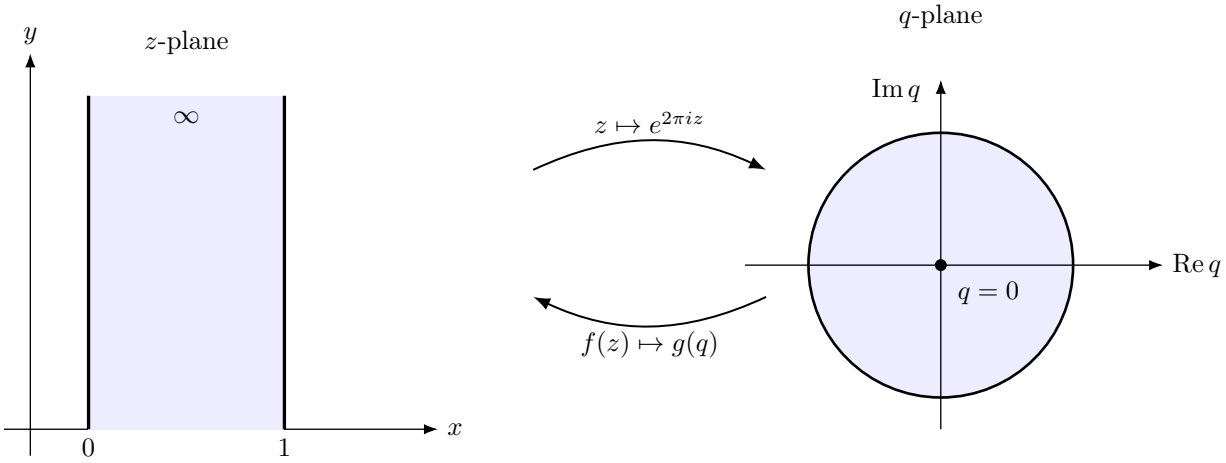
This implies that $\mathrm{d}z$ is a section of line bundle—specifically, the cotangent bundle. Now take a holomorphic 1-form $f(z)\mathrm{d}z$. We want it to be Γ-invariant. $f(\gamma z)\mathrm{d}(\gamma z) = f(z)\mathrm{d}z$ implies that $f(\gamma z) = (cz + d)^2 f(z)$. This is the transformation law for a modular form of weight 2. For even integer k, $f(z)(\mathrm{d}z)^{k/2}$ is Γ-invariant implies that $f(\gamma z) = (cz + d)^k f(z)$, which is the transformation law for a modular form of weight k. On the other hand, if k is odd, $f(z) = f((-I)z) = (-1)^k f(z)$ implies that $f \equiv 0$.

Definition 1.1. Let $f \colon \mathbb{H} \to \mathbb{C}$ be meromorphic. We say f is a **modular function of weight** $k \in 2\mathbb{Z}$ if

$$f(\gamma z) = (cz + d)^k f(z), \quad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma = \mathrm{SL}_2(\mathbb{Z}), \forall z \in \mathbb{H}.$$

Definition 1.2. Let f be a modular function of weight k. Then f is called **modular form** if f is holomorphic in $\mathbb{H} \cup \{\infty\}$.

Remark 1.1. f is holomorphic at ∞ iff $g(q)$ is holomorphic at 0. One may see the following picture. Note that f is periodic of period 1, it only suffices to consider $0 \leq x \leq 1$.



Definition 1.3. Let $k \geq 4$ even. We define the Eisenstein series:

$$G_k(z) = \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(mz+n)^k}, \quad z \in \mathbb{H}$$

Proposition 1.1. When $k > 2$, the series converges absolutely. When $k \geq 4$ even, G_k is a modular form of weight k.

Proof. We claim that for $\sigma > 2$,

$$f(\sigma) = \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{|mz+n|^\sigma}$$

converges absolutely.

Note that

$$\begin{aligned} f(\sigma) &= \sum_{\substack{m=0\\n\in\mathbb{Z}\setminus\{0\}}} \frac{1}{|n|^\sigma} + \sum_{\substack{m\neq 0\\n\in\mathbb{Z}}} \frac{1}{|mz+n|^\sigma} \\ &\ll 1 + 2\sum_{m\geq 1}\sum_{n\in\mathbb{Z}} \frac{1}{m^\sigma \left|z+\frac{n}{m}\right|^\sigma} \\ &= 1 + 2\sum_{m\geq 1}\frac{1}{m^\sigma}\sum_{n\in\mathbb{Z}} \frac{1}{\left|z+\frac{n}{m}\right|^{2\cdot\sigma/2}}. \end{aligned}$$

For $z = x + \mathrm{i}y$ where $y > 0$,

$$\left|x+\frac{n}{m}+\mathrm{i}y\right|^2 = \left(x+\frac{n}{m}\right)^2 + y^2.$$

Fix $m \geq 1$, $x + n/m$ becomes an arithmetic series with step $1/m$, and every internal $[\ell, \ell+1)$ contains exactly m terms (possibly one less at boundaries, but at most m). For every n with

$x + n/m \in [\ell, \ell + 1)$, we have

$$\left| x + \frac{n}{m} + \mathrm{i}y \right|^2 \geq \min\left\{ \ell^2 + y^2, (\ell+1)^2 + y^2 \right\}.$$

Hence

$$\sum_{n\in\mathbb{Z}} \frac{1}{\left|z + \frac{n}{m}\right|^{2\cdot\sigma/2}} \leq \sum_{\ell\in\mathbb{Z}} \frac{1}{\left(\min\left\{\ell^2 + y^2, (\ell+1)^2 + y^2\right\}\right)^{\sigma/2}} m.$$

Since the series

$$\sum_{\ell\in\mathbb{Z}} \frac{1}{\left(\min\left\{\ell^2 + y^2, (\ell+1)^2 + y^2\right\}\right)^{\sigma/2}}$$

converges, there exists a constant $C = C(y, \sigma)$ with

$$\sum_{n\in\mathbb{Z}} \frac{1}{\left|z + \frac{n}{m}\right|^{2\cdot\sigma/2}} \leq Cm,$$

and it follows that

$$1 + \sum_{m\geq 1} \frac{1}{m^\sigma} \sum_{n\in\mathbb{Z}} \frac{1}{\left|z + \frac{n}{m}\right|^{2\cdot\sigma/2}} \leq C \sum_{m\geq 1}^{\infty} \frac{m}{m^\sigma} = C \sum_{m=1}^{\infty} \frac{1}{m^{\sigma-1}}$$

converges.

Now let

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

then for $k \geq 4$,

$$\begin{aligned} G_k(\gamma z) &= \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{\left(m\frac{az+b}{cz+d} + n\right)^k} \\ &= (cz+d)^k \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(maz + mb + ncz + nd)^k} \end{aligned}$$

Now the map $(m, n) \mapsto (ma + nc, mb + nd)$ is a linear transformation on $\mathbb{Z}^2$ with determinant $ad - bc = 1$, hence it is a bijection from $\mathbb{Z}^2 \setminus \{(0, 0)\}$ onto itself. Since the series converges absolutely for $k > 2$, we can reorder the terms arbitrarily without changing the sum. Thus $G_k(\gamma z) = (cz + d)^k G_k(z)$.

One can use Fourier expansion to show that $G_k(z)$ is holomorphic in $\mathbb{H} \cup \{\infty\}$. A direct observation is that as $z \to \infty$, $G_k(z)$ looks like $2\zeta(k)$. □

For

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

we define $j(\gamma, z) = cz + d$. Let

$$\Gamma_\infty = \left\{ \pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z} \right\}.$$

The left action of Γ_∞ on Γ is given by

$$\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + cn & b + dn \\ c & d \end{pmatrix},$$

it follows that

$$\Gamma_\infty \backslash \Gamma = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z}, \gcd(c, d) = 1 \right\} \subset \mathrm{SL}_2(\mathbb{Z}).$$

Definition 1.4. We define the normalized Eisenstein series

$$E_k(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \frac{1}{j(\gamma, z)^k} = \frac{1}{2} \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1}} \frac{1}{(cz + d)^k}, \quad k \geq 4 \text{ even}.$$

Clearly we have

$$\begin{aligned} G_k(z) &= \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \frac{1}{(mz + n)^k} \\ &= \sum_{\ell=1}^{\infty} \sum_{\substack{m,n \in \mathbb{Z} \\ \gcd(m,n)=\ell}} \frac{1}{(mz + n)^k} \\ &= \sum_{\ell=1}^{\infty} \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1}} \frac{1}{(\ell cz + \ell d)^k} \\ &= \sum_{\ell=1}^{\infty} \frac{1}{\ell^k} \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1}} \frac{1}{(cz + d)^k} \\ &= 2\zeta(k) E_k(z). \end{aligned}$$

If f, g are modular functions of weight k, then clearly the product fg and g^2 are both modular functions of weight $2k$. This reflects the graded algebra structure.

Example 1.2. The j-invariant is defined by

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

which is a meromorphic modular function of weight 0.

1.3 The Fourier expansion of Eisenstein series

We have the well-known product representation for the sine function

$$\sin \pi z = \pi z \prod_{n>0} \left(1 - \frac{z}{n}\right)\left(1 + \frac{z}{n}\right).$$

Take the logarithmic derivative

$$\pi \frac{\cos \pi z}{\sin \pi z} = \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{z-n} + \frac{1}{z+n}\right).$$

Then this is equal to

$$\pi \frac{\frac{\mathrm{e}^{\mathrm{i}\pi z}+\mathrm{e}^{-\mathrm{i}\pi z}}{2}}{\frac{\mathrm{e}^{\mathrm{i}\pi z}-\mathrm{e}^{-\mathrm{i}\pi z}}{2\mathrm{i}}} = \pi\mathrm{i}\frac{e(z)+1}{e(z)-1} = -\pi\mathrm{i}(e(z)+1)\sum_{\ell\geq 0} e(\ell z) = -\pi\mathrm{i} - \sum_{\ell \geq 1} 2\pi\mathrm{i} e(\ell z).$$

We differentiate these expansions $k-1$ times, getting

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

for any $k > 2$ even. Hence we derive the Fourier expansion of $G_k(z)$ as follows

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

and by collecting terms with $dm = n$ we arrive at

$$\begin{aligned} G_k(z) &= 2\zeta(k) + 2\frac{(2\pi\mathrm{i})^k}{\Gamma(k)}\sum_{n=1}^{\infty}\left(\sum_{d|n} d^{k-1}\right) e(nz) \\ &= 2\zeta(k) + 2\frac{(2\pi\mathrm{i})^k}{\Gamma(k)}\sum_{n=1}^{\infty} \sigma_{k-1}(n) e(nz). \end{aligned}$$

Here $\sigma_s(n)$ is the arithmetic function given by

$$\sigma_s(n) = \sum_{d|n} d^s.$$

Since $G_k(z) = 2\zeta(k)E_k(z)$, it follows that

$$E_k(z) = 1 + \frac{(2\pi\mathrm{i})^k}{\Gamma(k)\zeta(k)}\sum_{n=1}^{\infty} \sigma_{k-1}(n) e(nz).$$

Modular forms can be regarded as higher-order generalizations of the Riemann Zeta Function. Indeed, the L-function of the normalized Eisenstein series E_k is

$$L(s, E_k) = \sum_{n\geq 1} \frac{\sigma_{k-1}(n)}{n^s} = \sum_{n\geq 1}\sum_{\ell | n} \frac{\ell^{k-1}}{n^s} = \sum_{\ell\geq 1}\sum_{m\geq 1} \frac{\ell^{k-1}}{m^s \ell^s} = \zeta(s)\zeta(s-k+1).$$

The normalized Eisenstein series $E_4(z)$ and $E_6(z)$ have the following Fourier expansions:

$$E_4(z) = 1 + 240e(z) + 2160e(2z) + \cdots$$
$$E_6(z) = 1 - 504e(z) - 16632e(2z) + \cdots$$

Calculating E_4^3 and E_6^2, the constant terms are both $1^3 = 1$ and $1^2 = 1$. When we subtract them, the constant term vanishes, which creates a cusp form:

$$\begin{aligned} E_4^3(z) - E_6^2(z) &= (1 + 720e(z) + \dots) - (1 - 1008e(z) + \cdots) \\ &= 1728e(z) + \cdots \end{aligned}$$

Then we can define a cusp form of weight 12,

$$\Delta(z) = \frac{1}{1728}(E_4^3 - E_6^2) = \frac{3375}{64\pi^{12}}G_4^3 - \frac{33075}{256\pi^{12}}G_6^2 = \sum_{n\geq 1}\tau(n)e(nz),$$

where $\tau(1) = 1$. The arithmetic function τ is called the Ramanujan function.

1.4 The modular group

Theorem 1.1. The modular group is generated by two matrices

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

Proof. Given arbitrary

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

we have

$$\begin{aligned} T^n\gamma &= \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+nc & b+nd \\ c & d \end{pmatrix}, \\ S\gamma &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -c & -d \\ a & b \end{pmatrix}. \end{aligned}$$

Moreover $S^2 = -I$.

If $c = 0$, $(a, d) = (1, 1)$ or $(-1, -1)$, it follows that

$$\gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = T^b$$

or

$$\gamma = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} = -I \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} = S^2 T^{-b}.$$

Now assume that $c \neq 0$. by replacing γ by

$$S^{\epsilon_1}\gamma = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$$

with $\epsilon_1 = \pm 1$, we obtain $|a_1| > |c_1|$. Since $ad - bc = 1$, $\gcd(a, c) = 1$. By the Euclidean algorithm, we can choose an integer n_1 such that left-multiplying by T^{n_1} yields

$$T^{n_1} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a_1 + c_1 n_1 & b_1 + n_1 d_1 \\ c_1 & d_1 \end{pmatrix},$$

where $|a_1 + c_1 n_1| < |c_1|$. By repeating this process,

$$T^{n_\ell} S^{\epsilon_\ell} \cdots T^{n_1} S^{\epsilon_1} \gamma = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix},$$

which is generated by S and T from above discussion. Indeed, At each step, c decreases strictly; since $|c| \geq 0$, after a finitely many steps we reach $c = 0$. □

Recall that we have defined $j(\gamma, z) = (cz + d)$. If we have

$$f(\gamma_1 z) = j(\gamma_1, z)^k f(z),$$
$$f(\gamma_2 z) = j(\gamma_2, z)^k f(z),$$

then

$$f(\gamma_1 \gamma_2 z) = j(\gamma_1 \gamma_2, z)^k f(z).$$

Indeed,

$$\begin{aligned} f(\gamma_1\gamma_2 z) &= j(\gamma_1, \gamma_2 z)^k f(\gamma_2 z) \\ &= j(\gamma_1, \gamma_2 z)^k j(\gamma_2, z)^k f(z) \\ &= \left((c_1(\gamma_2 z) + d_1)\,(c_2 z + d_2)\right)^k f(z) \\ &= \left(\left(c_1 \frac{a_2 z + b_2}{c_2 z + d_2} + d_1\right)(c_2 z + d_2)\right)^k f(z) \\ &= \left(c_1(a_2 z + b_2) + d_1(c_2 z + d_2)\right)^k f(z) \\ &= \left((c_1 a_2 + d_1 c_2) z + (c_1 b_2 + d_1 d_2)\right)^k f(z) \\ &= j(\gamma_1\gamma_2, z)^k f(z). \end{aligned}$$

That is, to verify that f is modular, it suffices to check the transformation property for the generators S and T.

Theorem 1.2. The set

$$D = \left\{ z = x + \mathrm{i}y : |x| < \frac{1}{2}, |z| > 1 \right\}$$

is a fundamental domain for the modular group $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, i.e., it has the following properties:

1. D is a domain in $\mathbb{H}$,
2. every orbit of Γ has a point in D or on the boundary ∂D,
3. distinct points in D are not in the same orbit of Γ.

Proof. Clearly

$$\mathrm{Im}(\gamma z) = \frac{\mathrm{Im}(z)}{|cz + d|^2}, \quad \forall \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Fix $z \in \mathbb{H}$, for arbitrary $N \in \mathbb{Z}_{>0}$, there exists finitely many $c, d \in \mathbb{Z}$ such that $|cz + d|^2 < N$. Then there exists $c, d \in \mathbb{Z}$ with $|cz + d|$ minimal, i.e., $\mathrm{Im}(\gamma z)$ maximal. That is, each orbit Γz has a point with largest height (the imaginary part). Such a point, say z, has the property that

$$|cz + d| \geq 1 \text{ for all } \gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \Gamma.$$

This maximal property is preserved by translations (which are represented by matrices with $c = 0$), so we can choose a maximal point of the orbit in the strip $|x| \leq \frac{1}{2}$. We shall show that the domain

$$D' = \left\{ z \in \mathbb{H} : |x| < \frac{1}{2}, |cz + d| > 1 \text{ for all } c, d \text{ with } c \neq 0 \text{ and } \gcd(c, d) = 1 \right\}$$

coincides with D. Indeed, $D' \subset D$ by choosing $c = 1, d = 0$. Conversely, if $z \in D$ and $c \neq 0$ then

$$|cz + d|^2 = c^2|z|^2 + 2cdx + d^2 > c^2 - |cd| + d^2 \geq 1,$$

so $D \subset D'$.

By applying T, every orbit can have a representative in $D \cup \partial D$. Distinct points in D are not in the same orbit of Γ. Indeed, if $z, \gamma z \in D = D'$, we have $|cz + d| > 1$, so $\mathrm{Im}(\gamma z) < \mathrm{Im}(z)$. Applying the same method to γ^{-1} yields $\mathrm{Im}(z) < \mathrm{Im}(\gamma z)$ since $\gamma^{-1}(\gamma z) = z$, which is a contradiction. Thus $D = D'$ is a fundamental domain. □

Observe that D has a parabolic vertex at ∞, an elliptic vertex at i of order $m(\mathrm{i}) = 2$, and two equivalent elliptic vertices at $\rho = \frac{-1+\mathrm{i}\sqrt{3}}{2}$ and $\rho' = \frac{1+\mathrm{i}\sqrt{3}}{2}$ of order $m(\rho) = m(\rho') = 3$.

Theorem 1.3 (Valence Formula)**.** Let $f \neq 0$ be a modular form of weight $k \geq 0$. Denote by $m_f(w)$ the order of f at w. If w is an elliptic fixed point (of the modular group) denote its order by $m(w)$, and put $m(w) = 1$ otherwise. We have

$$\sum_{w \pmod{\Gamma}} \frac{m_f(w)}{m(w)} = \frac{k}{12}.$$

Proof. Cut off the upper part of the standard fundamental polygon D along a horizontal segment of very large height. Make very small circular cuts at the elliptic vertices and at any singular point of f which lies on the boundary ∂D in such a way that the equivalent points have congruent cuts. Let P denote the resulting cut polygon.

Integrate f'/f along the boundary ∂P of the cut polygon, getting by Cauchy's theorem

$$\frac{1}{2\pi\mathrm{i}} \int_{\partial P} \frac{f'}{f}(z)\mathrm{d}z = \sum_{w \in P} m_f(w).$$

Note that $m(w) = 1$ for any $w \in P$. On the horizontal segment (we can assume that it does not contain singular points) we see from the Fourier expansions

$$f(z) = \sum_{n=m}^{\infty} a_n e(nz),$$
$$f'(z) = \sum_{n=m}^{\infty} 2\pi\mathrm{i}n a_n e(nz)$$

where $a_m \neq 0$ for $m = m_f(\infty)$ that

$$\frac{f'}{f}(z) = \sum_{\ell=0}^{\infty} b_\ell e(\ell z)$$

with $b_0 = 2\pi \mathrm{i} m_f(\infty)$. Integral over the horizontal segment, i.e., from $1/2$ to $-1/2$,

$$\begin{aligned}
\int_{\frac{1}{2}}^{-\frac{1}{2}} \frac{f'}{f}(z)\mathrm{d}z &= \int_{\frac{1}{2}}^{-\frac{1}{2}} \sum_{\ell=0}^{\infty} b_\ell e(\ell z)\mathrm{d}z \\
&= \sum_{\ell=0}^{\infty} \int_{\frac{1}{2}}^{-\frac{1}{2}} b_\ell e(\ell z)\mathrm{d}z \\
&= -2\pi \mathrm{i} m_f(\infty) + \sum_{\ell=1}^{\infty} \frac{b_\ell}{2\pi \mathrm{i}\ell} \mathrm{e}^{2\pi \mathrm{i}\ell z}\Big|_{\frac{1}{2}}^{-\frac{1}{2}} \\
&= -2\pi \mathrm{i} m_f(\infty).
\end{aligned}$$

The two integrals along equivalent vertical segments with congruent cuts at singular points cancel out by periodicity. At every elliptic vertex $w = \rho, \mathrm{i}, \rho'$ we have

$$\frac{f'}{f}(z) = \frac{m_f(w)}{z - w} + h(z)$$

where $h(z)$ is holomorphic in a small disc centered at w. Integral along the small arc,

$$\lim_{\epsilon \to 0} \frac{1}{2\pi \mathrm{i}} \int_{\gamma_\epsilon} \frac{f'}{f}(z)\mathrm{d}z = \frac{\theta_w}{2\pi} m_f(w).$$

For $w = \mathrm{i}$, the right-hand side is $\frac{1}{2} m_f(\mathrm{i})$; For $w = \rho, \rho'$, the right-hand side is $\frac{1}{6} m_f(\rho)$.

It remains to evaluate the integral along the arc of the circle $|z| = 1$ at the bottom of P with the elliptic points ρ, i, ρ' and other singular points being removed. The middle point i splits the arc into two equivalent arcs, say A and $A' = SA$. We have

$$\begin{aligned}
f(Sz) &= z^k f(z) \\
f'(Sz)z^{-2} &= kz^{k-1} f(z) + z^k f'(z) \\
\frac{f'}{f}(Sz)z^{-2} &= \frac{k}{z} + \frac{f'}{f}(z).
\end{aligned}$$

Hence

$$\frac{1}{2\pi i} \int_{A'} \frac{f'}{f}(z)\mathrm{d}z = -\frac{1}{2\pi i} \int_A \frac{f'}{f}(Sz)\mathrm{d}(Sz) = -\frac{1}{2\pi i} \int_A \left(\frac{k}{z} + \frac{f'}{f}(z) \right) \mathrm{d}z,$$

and

$$\frac{1}{2\pi i} \int_{A \cup A'} \frac{f'}{f}(z)\mathrm{d}z = -\frac{1}{2\pi i} \int_A \frac{k}{z}\mathrm{d}z = \frac{k}{12}$$

because the arc A has one-twelfth the length of the whole circle $|z| = 1$. Gathering together the above evaluations, one obtains

$$\sum_{w \in P} m_f(w) = \frac{k}{12} - \frac{1}{2} m_f(\mathrm{i}) - \frac{1}{3} m_f(\rho) - m_f(\infty),$$

which is precisely the Valence Formula. □

1.5 The linear space of modular forms

In this section, k must be even.

Let $\mathcal{M}_k$ denote the linear space of modular forms of weight $k > 0$. Clearly modular forms of different weights are linearly independent over $\mathbb{C}$, so the space of all modular forms is the direct sum of the $\mathcal{M}_k$,

$$\mathcal{M} = \bigoplus_{k \geq 0} \mathcal{M}_k.$$

The whole space $\mathcal{M}$ can also be considered as a graded algebra with respect to the inclusions

$$\mathcal{M}_k \mathcal{M}_\ell \subset \mathcal{M}_{k+\ell}.$$

First let us consider each space $\mathcal{M}_k$ for a few small, even values of k. Let a be the sum of the orders of the zeros at ∞ and all ordinary points in $\mathbb{H}$, i.e., $a = m_f(\infty) + \sum_{w \neq \mathrm{i}, \rho} m_f(w)$; b be the order of the zero at the elliptic point i, i.e., $b = m_f(\mathrm{i})$; c be the order of the zero at the elliptic point ρ, i.e., $c = m_f(\rho)$. The Valence Formula (Theorem 1.3) gives equation $a + b/2 + c/3 = k/12$.

Case $k = 0$. Note that $m_f(w)/m(w) \geq 0$ for all w, so $m_f(w) = 0$ for all w, and so f has no zeros in $\mathbb{H} \cup \{\infty\}$. We claim that $\mathcal{M}_0 = \mathbb{C}$. Clearly constant functions are modular forms of weight 0. For any $f \in \mathcal{M}_0$, we pick arbitrary $x_0 \in \mathbb{H} \cup \{\infty\}$. Then it suffices to show $f(z) = f(x_0)$ for all z. Let $g(z) = f(z) - f(x_0)$ and assume conversely $g \not\equiv 0$. Recall that constant function $z \mapsto f(x_0)$ is an element of $\mathcal{M}_0$. Hence $g \in \mathcal{M}_0$. But note that $g(z)$ has zero point at x_0, the Valence Formula (Theorem 1.3) implies that $\sum_{w \pmod \Gamma} \frac{m_g(w)}{m(w)} > 0$, a contradiction. Therefore $f(z) \equiv f(x_0)$, f is constant.

$$\mathcal{M}_0 = \mathbb{C}.$$

Case $k = 2$. Then $m_f(z) = 0$ for all z, so $f(z)$ is constant, and necessarily zero, and

$$\mathcal{M}_2 = 0.$$

Case $k = 4$. Clearly $G_4 \in \mathcal{M}_4$. Let $f \in \mathcal{M}_4$. $a + b/2 + c/3 = 1/3$ implies that $(a, b, c) = (0, 0, 1)$. It follows that $m_f(\rho) = 1$ and $m_f(z) = 0$ for all $z \not\equiv \rho \pmod \Gamma$. Hence f/G_4 has no zeros and no poles, and is a modular form of weight 0. It follows that $f/G_4 \in \mathbb{C}$, $f \in G_4\mathbb{C}$.

$$\mathcal{M}_4 = G_4(z)\mathbb{C}.$$

Case $k = 6$. Then $m_f(i) = 1$ and $m_f(z) = 0$ for $z \not\equiv i \pmod \Gamma$. The same argument as before shows that $f(z) = cG_6(z)$ and

$$\mathcal{M}_6 = G_6(z)\mathbb{C}.$$

Case $k = 8$. Then $m_f(\rho) = 2$ and $m_f(z) = 0$ for $z \not\equiv \rho \pmod{\Gamma}$. Hence $f(z) = cG_4(z)^2$ and

$$\mathcal{M}_8 = G_4(z)^2\mathbb{C}.$$

Case $k = 10$. Then $m_f(i) = 1$, $m_f(\rho) = 1$ and $m_f(z) = 0$ for other inequivalent points. Hence we argue as before to show that $f(z) = cG_4(z)G_6(z)$ and

$$\mathcal{M}_{10} = G_4(z)G_6(z)\mathbb{C}.$$

Case $k = 12$. Recall that $\Delta(z)$ does not vanish in $\mathbb{H}$ and it has a simple zero at $z = \infty$. Since $G_{12}(\infty) \neq 0$ there exists a constant c such that $f(z) - cG_{12}(z)$ vanishes at ∞, so

$$\frac{f(z) - cG_{12}(z)}{\Delta(z)} \in \mathcal{M}_0 = \mathbb{C}$$

and $f(z) = b\Delta(z) + cG_{12}(z)$. This proves that

$$\mathcal{M}_{12} = \Delta(z)\mathbb{C} \oplus G_{12}(z)\mathbb{C}.$$

Similarly by induction we prove that

$$\mathcal{M}_k = \Delta(z)\mathcal{M}_{k-12} \oplus G_k(z)\mathbb{C}, \qquad \text{if } k \geqslant 12.$$

From the above results we conclude

Theorem 1.4. The dimension of the space $\mathcal{M}_k$ is given by

$$\dim \mathcal{M}_k = \begin{cases} [k/12] & \text{if } k \equiv 2 \pmod{12} \\ [k/12] + 1 & \text{if } k \not\equiv 2 \pmod{12}. \end{cases}$$

Next we show that any $f \in \mathcal{M}_k$ is a polynomial in G_4 and G_6. This is already proved for $k \leq 10$. If $k \geq 12$ we subtract from f a monomial in G_4, G_6 of special type, namely we consider

$$f - cG_4^a G_6^b, \quad \text{where } 4a + 6b = k,$$

with a suitable constant c such that the resulting function vanishes at ∞. Then divide by Δ, getting a function in $\mathcal{M}_{k-12}$, and the proof is completed by induction.

We shall also prove that G_4, G_6 are algebraically independent. Suppose to the contrary that $P(G_4, G_6) = 0$ for some polynomial $P \neq 0$. Modular forms of different weights are linearly independent over $\mathbb{C}$, so we may assume P is homogeneous, it then follows that the function G_4^3/G_6^2 would verify a nontrivial algebraic equation with coefficients in $\mathbb{C}$. Thus G_4^3/G_6^2 would be constant, which is absurd because G_4 is zero at ρ but not G_6, and $G_4^3(\infty)/G_6^2(\infty) = 2\zeta(4)^3/\zeta(6)^2 = 49/20$.

Theorem 1.5. The Eisenstein series G_4, G_6 are algebraically independent and generate the space of all modular forms, that is

$$\mathcal{M} = \mathbb{C}[G_4, G_6].$$

A modular form $f \in \mathcal{M}_k$ is called a cusp form if it vanishes at ∞, i.e., its Fourier expansion has no constant term,

$$f(z) = \sum_{n \geq 1} a_n e(nz).$$

The linear space $\mathcal{S}_k$ of cusp forms together with $E_k(z)\mathbb{C}$ gives

$$\mathcal{M}_k = E_k(z)\mathbb{C} \oplus \mathcal{S}_k$$

for even $k \geq 4$.

2 Automorphic Forms in General

2.1 The hyperbolic plane

We denote by $\mathrm{GL}_2^+(\mathbb{R})$ the group of 2×2 real matrices of positive determinant. Observe that a linear fractional transformation g determines the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R})$$

up to a scalar because the matrices $\mathrm{diag}\{\alpha, \alpha\}$ with $\alpha \neq 0$ give the identity transformation. Dividing by a scalar, we can represent g by a matrix of determinant 1. Note that the action of

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

coincide with the action of

$$\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}.$$

The **projective special linear group**

$$\mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{I, -I\},$$

where

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

is identified with the linear fractional transformations.

Let $\mathbb{H}$ be the upper half-plane, and $G = \mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{H}$ by

$$G \times \mathbb{H} \to \mathbb{H}, \quad (g, z) \mapsto gz = \frac{az+b}{cz+d}, \quad \forall g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}), z \in \mathbb{H}.$$

Note that

$$\mathrm{Im}(gz) = \frac{y}{|cz+d|^2} > 0,$$

so the action is well-defined. In Euclidean geometry, translations and rotations do not change the distance between two points; we now aim to define a Riemannian metric on the upper half-plane $\mathbb{H}$, which is $\mathrm{SL}_2(\mathbb{R})$-invariant, that is, $\rho(gz, gw) = \rho(z, w)$.

Since

$$\frac{\mathrm{d}gz}{\mathrm{d}z} = \frac{a(cz+d) - (az+b)c}{(cz+d)^2} = \frac{1}{(cz+d)^2},$$

we obtain

$$\begin{cases} |\mathrm{d}gz| = \dfrac{1}{|j_g(z)|^2}|\mathrm{d}z|, \\ \mathrm{Im}(gz) = \dfrac{\mathrm{Im}(z)}{|j_g(z)|^2}, \end{cases}$$

it follows that

$$\frac{|\mathrm{d}gz|}{\mathrm{Im}(gz)} = \frac{|\mathrm{d}z|}{\mathrm{Im}(z)}.$$

Since $\frac{|\mathrm{d}z|}{\mathrm{Im}(z)}$ is invariant under the transformation, it serves as the natural metric for the hyperbolic plane. If we write $z = x + \mathrm{i}y$, we get $|\mathrm{d}z| = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$ and $\mathrm{Im}(z) = y$. Thus, the hyperbolic line element is:

$$\mathrm{d}s = \frac{|\mathrm{d}z|}{\mathrm{Im}(z)} = \frac{1}{y}\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}.$$

This metric is distinct from the Euclidean metric, since

$$\mathrm{d}s_{\mathbb{R}} = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}.$$

With the metric derived from this differential the upper half-plane becomes a Riemannian manifold. The distance function on $\mathbb{H}$ is given by

$$\rho(z, w) = \min_L \int_0^1 (x'(t)^2 + y'(t)^2)^{\frac{1}{2}} y(t)^{-1} \mathrm{d}t.$$

Given any two points $z, w \in \mathbb{H}$, we can always find a $g \in \mathrm{SL}_2(\mathbb{R})$ that maps both points onto the imaginary axis, such that $gz = \mathrm{i}y_1$ and $gw = \mathrm{i}y_2$ for some $y_1, y_2 \in \mathbb{R}^+$. Such matrix g can be obtained by solving a system of algebraic equations. Geometrically, the geodesics (shortest paths)

in the hyperbolic plane $\mathbb{H}$ are semicircles orthogonal to the real axis, or vertical half-lines (which can be viewed as semicircles of infinite radius centered at infinity). Since Möbius transformations map generalized circles to generalized circles and preserve angles, the transformation g maps the semicircular geodesic passing through z and w precisely onto the vertical geodesic (the imaginary axis) passing through $\mathrm{i}y_1$ and $\mathrm{i}y_2$. Because the hyperbolic metric $\mathrm{d}s = y^{-1}$, $\mathrm{d}z$ is invariant under g, the integral can be computed:

$$\rho(z,w) = \min_L \int_0^1 (x'(t)^2 + y'(t)^2)^{\frac{1}{2}} y(t)^{-1} \mathrm{d}t = \left| \int_{y_2}^{y_1} \frac{\mathrm{d}y}{y} \right|.$$

More explicitly, we have the relation

$$\cosh \rho(z,w) = 1 + 2u(z,w),$$

where

$$u(z,w) = \frac{|z-w|^2}{4 \operatorname{Im}(z) \operatorname{Im}(w)}.$$

Given an area $\Omega \subset \mathbb{H}$, we have

$$\mathrm{vol}(\Omega) = \int_\Omega \mathrm{d}\mu z = \int_\Omega \frac{1}{y^2} \mathrm{d}x\mathrm{d}y,$$

this gives hyperbolic measure.

Let $\mathcal{F}$ be the fundamental domain defined in Theorem 1.2,

$$\mathrm{vol}(\mathcal{F}) = \int_{\mathcal{F}} \frac{\mathrm{d}x\mathrm{d}y}{y^2} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x \int_{\sqrt{1-x^2}}^{\infty} \frac{\mathrm{d}y}{y^2} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{\sqrt{1-x^2}} \mathrm{d}x = \frac{\pi}{3}.$$

Indeed, if $\mathcal{T}$ be a hyperbolic triangle with interior angles α, β, γ, the hyperbolic area of $\mathcal{T}$ is given by (Gauss defect phenomenon)

$$\mathrm{vol}(\mathcal{T}) = \pi - \alpha - \beta - \gamma.$$

2.2 The classification of motions

The identity motion (every $z \in \mathbb{H}$ is a fixed point) forms a class by itself. Any other motion

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq \pm 1$$

has one or two fixed points in $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. There are three cases:

1. g has one fixed point on $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$,
2. g has two distinct fixed points on $\hat{\mathbb{R}}$,

3. g has one fixed point in $\mathbb{H}$ and the complex conjugate fixed point in the lower half-plane.

Accordingly g is called **parabolic**, **hyperbolic**, **elliptic**. By conjugating we can bring g to one of the following types:

1. $z \mapsto z + t$ (translation, fixed point ∞),

2. $z \mapsto pz$ (dilation, fixed points $0, \infty$),

3. $z \mapsto k(\theta)z$ (rotation, fixed point i),

where $t \in \mathbb{R}$, $p \in \mathbb{R}^+$, $\theta \in \mathbb{R}$ and

$$k(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

A parabolic motion moves points along horocycles (circles in $\mathbb{H}$ tangent to $\hat{\mathbb{R}}$). A hyperbolic motion moves points along hypercycles (the segments in $\mathbb{H}$ of circles in $\hat{\mathbb{C}}$ through the fixed points on $\hat{\mathbb{R}}$). The geodesic line through the fixed points of a hyperbolic motion is mapped to itself, not identically. Of the two fixed points one is repelling, the other is attracting. An elliptic motion moves points along circles centered at its fixed point in $\mathbb{H}$.

The number of fixed points of a rigid motion is invariant under conjugation; therefore the above classification applies naturally to the conjugacy classes. The same classification can be also described in terms of trace (which is an algebraic invariant of conjugation), namely if

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq \pm 1,$$

then

1. g is parabolic $\iff |a + d| = 2$,

2. g is hyperbolic $\iff |a + d| > 2$,

3. g is elliptic $\iff |a + d| < 2$.

Indeed, $gz = z$ implies that $az + b = z(cz + d)$, so $cz^2 + (d - a)z - b = 0$. If $c \neq 0$, it is well-known that

$$z_{1,2} = \frac{-(d-a) \pm \sqrt{\Delta}}{2c},$$

where $\Delta = (d - a)^2 - 4c(-b) = a^2 + d^2 - 2ad + 4bc = (a + d)^2 - 4$. The number of fixed points is exactly the number of roots of the quadratic equation. Furthermore, the absolute value of the trace is an invariant under conjugation, that is, $\mathrm{Tr}(hgh^{-1}) = \mathrm{Tr}(g)$.

2.3 Discrete groups — Fuchsian groups

We provide some basic definitions and properties without proof.

The group $M_2(\mathbb{R})$ of 2×2 real matrices is a vector space with a norm given by

$$||g|| = \sqrt{a^2 + b^2 + c^2 + d^2} \text{ if } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Notice that since $|| - g|| = ||g||$, this norm is well-defined on the projective quotient group $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{-I, I\}$.

Definition 2.1. A subgroup Γ of $\mathrm{PSL}_2(\mathbb{R})$ is **discrete** if for arbitrary $B > 0$,

$$\#\{\gamma \in \Gamma : ||\gamma|| < B\} < \infty.$$

Definition 2.2. A subgroup Γ of $\mathrm{PSL}_2(\mathbb{R})$ is said to act **discontinuously** on the hyperbolic plane $\mathbb{H}$ if every orbit Γz has no limit point on $\mathbb{H}$.

Theorem 2.1 (Poincaré)**.** A subgroup $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ is discrete iff it acts discontinuously on $\mathbb{H}$.

Example 2.1. $\mathrm{SL}_2(\mathbb{Z})$ is discrete.

Definition 2.3. A subgroup $\Gamma \subset \mathrm{PSL}_2(\mathbb{R})$ acting discontinuously on $\mathbb{H}$ is called a **Fuchsian group**.

Definition 2.4. A Fuchsian group Γ is said to be of **the first kind** if every point on the boundary $\partial\mathbb{H} = \hat{\mathbb{R}}$ is a limit (in the $\hat{\mathbb{C}}$-topology) of some orbit Γz.

In the $\hat{\mathbb{C}}$-topology, a subset $U \subset \hat{\mathbb{C}}$ is open if either $\infty \notin U$ and U is open in $\mathbb{C}$, or $\infty \in U$ and its complement $\hat{\mathbb{C}} \setminus U$ is a compact subset of $\mathbb{C}$.

Proposition 2.1. Every Fuchsian group of the first kind has a finite number of generators.

Definition 2.5. Let Γ be a Fuchsian group of the first kind. A set $\mathcal{F} \subset \mathbb{H}$ is called a fundamental domain for Γ if

1. $\mathcal{F}$ is a domain in $\mathbb{H}$,
2. distinct points in $\mathcal{F}$ are not equivalent,
3. any orbit of Γ contains a point in the closure $\overline{\mathcal{F}}$ with respect to the $\hat{\mathbb{C}}$-topology.

Theorem 2.2. Let Γ be a Fuchsian group of the first kind, then $\mathrm{vol}(\mathcal{F}) < \infty$.

Proposition 2.2. Suppose Γ is a Fuchsian group of the first kind and $w \in \mathbb{H}$ is not fixed by any $\gamma \in \Gamma$ other than the identity motion. Then the set

$$D = \{z \in \mathbb{H} : \rho(z, w) < \rho(z, \gamma w) \text{ for all } \gamma \in \Gamma, \gamma \neq 1\}$$

is a fundamental domain for Γ.

Proof. To begin with we prove D is connected open set. For any $\gamma \in \Gamma$, $\gamma \neq 1$, define

$$H_\gamma = \{z \in \mathbb{H} : \rho(z, w) < \rho(z, \gamma w)\}.$$

Then $D = \bigcap_{\gamma \neq 1} H_\gamma$. Note that the function $z \mapsto \rho(z, \gamma w) - \rho(z, w)$ is continuous, H_γ is the complete preimage of $(0, +\infty)$, so is open.

Since Γ acts on $\mathbb{H}$ discontinuously, Γw admits no limit point in $\mathbb{H}$. Then for arbitrary compact $K \subset \mathbb{H}$, $\#\{\gamma \in \Gamma : \gamma w \in K\} < \infty$. Set $z_0 \in D$ and $r_0 = \rho(z_0, w)$. Consider the hyperbolic disk $\overline{B}(z_0, r_0 + 1)$, there exists finite subset $F \subset \Gamma \setminus \{1\}$ with $\gamma w \in \overline{B}(z_0, 2r_0 + 2), \gamma \in F$ since $\overline{B}(z_0, 2r_0 + 2)$ is compact. We claim that for any $\gamma \notin F$, $\rho(z, w) < \rho(z, \gamma w)$ holds in some neighborhood of z_0. Indeed, if $\gamma \notin F$, we have $\rho(z_0, \gamma w) > 2r_0 + 2$. Then for $z \in B(z_0, 1)$,

$$\rho(z, \gamma w) \geq \rho(z_0, \gamma w) - \rho(z, z_0) > 2r_0 + 2 - 1 = 2r_0 + 1,$$

$$\rho(z, w) \leq \rho(z_0, w) + \rho(z, z_0) < r_0 + 1.$$

It follows that $\rho(z, w) < r_0 + 1 < 2r_0 + 1 < \rho(z, \gamma w)$, the above claim holds. Therefore,

$$D \cap B(z_0, 1) = \left(\bigcap_{\gamma \in F} H_\gamma \right) \cap B(z_0, 1)$$

is intersection of finitely many open sets, hence open. Since $z_0 \in D$ arbitrary, D is open.

D is nonempty. Indeed, for arbitrary $\gamma \neq 1$, $\rho(w, w) = 0 < \rho(w, \gamma w)$ since $\gamma w \neq w$ by assumption.

Since $\{z : \rho(z, w) = \rho(z, \gamma w\}$ is a geodesic, H_γ is convex for all γ, $D = \bigcap_{\gamma \neq 1} H_\gamma$, as an intersection of convex subsets, is also convex. It follows that D is connected.

Distinct points in D are not equivalent. Assume conversely, there exists $z_1, z_2 \in D$ and $1 \neq \gamma \in \Gamma$ with $z_2 = \gamma z_1$. Since $z_1 \in D$, for $\gamma^{-1} \neq 1$, we have

$$\rho(z_1, w) < \rho(z_1, \gamma^{-1} w).$$

On the other hand, since $z_2 = \gamma z_1$,

$$\rho(z_1, \gamma^{-1} w) = \rho(\gamma z_1, w) < \rho(\gamma z_1, \gamma w) = \rho(z_1, w),$$

a contradiction. Thus $\gamma = 1$.

Any orbit of Γ admits a representative in $\overline{D}$. it suffices to show for arbitrary $z \in \mathbb{H}$, there exists $\gamma_0 \in \Gamma$ such that $\gamma_0 z \in \overline{D}$. Define $f\colon \Gamma \to \mathbb{R}_{\geq 0}$ by $\gamma \mapsto \rho(\gamma z, w)$, we claim that there exists $\gamma_0 \in \Gamma$ with f minimal. Note that $f(\gamma) = \rho(\gamma z, w) = \rho(z, \gamma^{-1} w)$. Set $r = f(1) = \rho(z, w)$ and consider $\overline{B}(z, r)$, the set

$$S = \left\{\gamma \in \Gamma : \gamma^{-1} w \in \overline{B}(z, r)\right\} = \{\gamma \in \Gamma : f(\gamma) \leq r\}$$

is finite thanks to the discreteness of Γw. Furthermore for arbitrary $\gamma \notin S$, $f(\gamma) > r \geq f(\gamma_0)$. Therefore, there exists $\gamma_0 \in S$ such that $f(\gamma_0)$ minimal. For $1 \neq \eta \in \Gamma$, we have $f(\gamma_0) \leq f(\eta^{-1}\gamma_0)$, i.e., $\rho(\gamma_0 z, w) \leq \rho(\eta^{-1}\gamma_0 z, w)$. Then $\rho(\eta^{-1}\gamma_0 z, w) = \rho(\gamma_0 z, \eta w)$, and so $\rho(\gamma_0 z, w) \leq \rho(\gamma_0 z, \eta w)$, which is precisely $\gamma_0 z \in \overline{D}$. □

Let Γ be a Fuchsian group of the first kind, and let $\mathcal{F}$ be a fundamental domain for Γ. Suppose that the closed polygon $\overline{\mathcal{F}}$ is not compact. Then $\overline{\mathcal{F}}$ must have a vertex on $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, say $\mathfrak{a}$. Such $\mathfrak{a}$ is called a **cusp**. Equivalently, $\mathfrak{a}$ is fixed point of some parabolic element of Γ. A point $\mathfrak{a} \in \hat{\mathbb{R}}$ is a cusp for Γ iff its stabilizer $\Gamma_\mathfrak{a} = \{\gamma \in \Gamma : \gamma\mathfrak{a} = \mathfrak{a}\}$ is an infinite cyclic group consisting entirely of parabolic transformations (apart from the identity). Such a subgroup is generated by a single primitive parabolic element.

The stability group $\Gamma_\mathfrak{a}$ is cyclic infinite, say generated by $\gamma_\mathfrak{a}$. There exists $\sigma_\mathfrak{a} \in \mathrm{SL}_2(\mathbb{R})$ such that (change $\gamma_\mathfrak{a}$ to $-\gamma_\mathfrak{a}$ if necessary)

$$\sigma_\mathfrak{a}\infty = \mathfrak{a}, \quad \sigma_\mathfrak{a}^{-1}\gamma_\mathfrak{a}\sigma_\mathfrak{a} = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}.$$

We shall call $\sigma_\mathfrak{a}$ a scaling matrix of the cusp $\mathfrak{a}$.

We will make some explanations on the existence of $\sigma_\mathfrak{a}$. To begin with, we claim $\mathrm{SL}_2(\mathbb{R})$ acts transitively on $\hat{\mathbb{R}}$. It only suffices to show for arbitrary $x \in \hat{\mathbb{R}}$, there exists $g \in \mathrm{SL}_2(\mathbb{R})$ such that $g\infty = x$. If $x = \infty$, we may take identity. Now assume $x \in \mathbb{R}$, let

$$g = \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix},$$

then

$$g\infty = \lim_{z \in \infty} \frac{xz - 1}{z + 0} = x.$$

Now from the above claim, there exists $\tau \in \mathrm{SL}_2(\mathbb{R})$ with $\tau\infty = \mathfrak{a}$. Define $\gamma' = \tau^{-1}\gamma_\mathfrak{a}\tau$, we obtain

$$\gamma'\infty = \tau^{-1}\gamma_\mathfrak{a}\tau\infty = \tau^{-1}\gamma_\mathfrak{a}\mathfrak{a} = \tau^{-1}\mathfrak{a} = \infty.$$

This implies that

$$\gamma' = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}).$$

(This is because a matrix in $\mathrm{SL}_2(\mathbb{R})$ that fixes ∞ must be upper triangular.) Moreover, $\mathrm{Tr}(\gamma') = \mathrm{Tr}(\gamma_{\mathfrak{a}}) = \pm 2$ since $\gamma_{\mathfrak{a}}$ is a parabolic element. Combine $ad - 0b = ad = 1$, we obtain $(a, d) = (1, 1)$ or $(-1, -1)$. Change $\gamma_{\mathfrak{a}}$ to $\gamma_{\mathfrak{a}}^{-1}$ if necessary, we have

$$\gamma' = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$$

with $b > 0$. If we set

$$\delta = \begin{pmatrix} \sqrt{b} & 0 \\ 0 & 1/\sqrt{b} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}),$$

then

$$\delta^{-1}\gamma'\delta = \begin{pmatrix} 1/\sqrt{b} & 0 \\ 0 & \sqrt{b} \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{b} & 0 \\ 0 & 1/\sqrt{b} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Therefore, let $\sigma_{\mathfrak{a}} = \tau\delta$, one may easily verify it is precisely what we want.

2.4 Congruence groups

Let q be a positive integer. Then

$$\Gamma(q) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{q} \right\}$$

is called the **principal congruence group of level** q.

Theorem 2.3. It is a normal subgroup of the modular group $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$ of index

$$[\Gamma(1) : \Gamma(q)] = q^3 \prod_{p|q} \left(1 - \frac{1}{p^2}\right).$$

Proof. Consider the exact sequence

$$1 \longrightarrow \Gamma(q) \longrightarrow \mathrm{SL}_2(\mathbb{Z}) \longrightarrow \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}) \longrightarrow 1 \ .$$

Note that $\Gamma(q) \hookrightarrow \mathrm{SL}_2(\mathbb{Z})$ is inclusion, and the projection $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$ is defined by $M \mapsto M \bmod q$. Then $\mathrm{Ker}(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})) = \{M \in \mathrm{SL}_2(\mathbb{Z}) : M \equiv I \pmod{q}\}$, which is precisely $\Gamma(q)$.

The only question is the surjectivity of the residue class map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$. Note that an element in $\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$ does not necessarily have a preimage in $\mathrm{SL}_2(\mathbb{Z})$ under the natural lifting. For example, consider the matrix

$$\begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/6\mathbb{Z})$$

which satisfies $9 + 4 = 13 \equiv 1 \pmod 6$, but

$$\begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \notin \mathrm{SL}_2(\mathbb{Z}).$$

We claim that

$$\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}) = \langle T, T' \rangle,$$

where

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } T' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

For arbitrary

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}),$$

we have $ad - bc \equiv 1 \pmod q$, so $\gcd(\gcd(a, c), q) = 1$. From the Fundamental Theorem of Arithmetic, we may write $q = \prod_{p|q} p^{r_p}$. For every $p \mid q$, there exists x_p with $\gcd(ax_p + c, p) = 1$. Indeed, we can choose

$$x_p = \begin{cases} 1 & \text{if } p \mid c, \\ 0 & \text{if } p \nmid c. \end{cases}$$

Using the Chinese Remainder Theorem, we can find n with

$$n \equiv x_p \pmod{p^{r_p}}, \quad \forall p \mid q.$$

It follows that $c_1 = an + c$ satisfies $\gcd(c_1, q) = 1$. Thus there exists n with

$$T'^n \gamma = \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c_1 & d_1 \end{pmatrix},$$

where $\gcd(c_1, q) = 1$. Hence c_1 is invertible in $\mathbb{Z}/q\mathbb{Z}$, there exists k with $a + kc_1 \equiv 0 \pmod q$, it follows that

$$T^k T'^n \gamma = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c_1 & d_1 \end{pmatrix} \equiv \begin{pmatrix} 0 & b_2 \\ c_1 & d_1 \end{pmatrix} \pmod q,$$

where $-b_2c_1 \equiv 1 \pmod{q}$. By applying T',

$$T'^{d_1c_1}T^kT'^n\gamma = \begin{pmatrix} 1 & 0 \\ d_1c_1 & 1 \end{pmatrix} \begin{pmatrix} 0 & b_2 \\ c_1 & d_1 \end{pmatrix} \equiv \begin{pmatrix} 0 & b_2 \\ c_1 & 0 \end{pmatrix} \pmod{q}.$$

However, note that

$$\begin{pmatrix} 0 & -c_1^{-1} \\ c_1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -c_1^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c_1 & 1 \end{pmatrix},$$

we have $T'^{d_1c_1}T^kT'^n\gamma \in \langle T, T' \rangle$, so $\gamma \in \langle T, T' \rangle$. Thus $\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}) = \langle T, T' \rangle$. Since $T, T' \in \mathrm{SL}_2(\mathbb{Z})$, the natural projection is surjective.

By the short exact sequence we have

$$[\Gamma(1) : \Gamma(q)] = |\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})|.$$

Now we compute $|\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})|$. For every

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}),$$

we have $ad - bc \equiv 1 \pmod{q}$, which implies that $\gcd(\gcd(c,d), q) = 1$. First we count the number of c, d with $\gcd(\gcd(c,d), q) = 1$, and then count a, b. From $\gcd(\gcd(c,d), q) = 1$, there exists $x, y \in \mathbb{Z}$ with $\gcd(c,d)x + qy = 1$, and so there exists x_1, y_1 with $(x_1c + y_1d)x = 1 - qy$. Then

$$\begin{aligned}
\sum_{c=1}^{q}\sum_{d=1}^{q} \delta_{\gcd(\gcd(c,d),q)=1} &= \sum_{c=1}^{q}\sum_{d=1}^{q} \sum_{r \mid \gcd(\gcd(c,d),q)} \mu(r) \quad (\text{since } \mu * u = I) \\
&= \sum_{r \mid q} \mu(r) \sum_{\substack{c=1 \\ r \mid c}}^{q} \sum_{\substack{d=1 \\ r \mid d}}^{q} 1 \\
&= \sum_{r \mid q} \mu(r) \left(\frac{q}{r}\right)^2 \qquad (\text{since } \#\{c \in \{1, \cdots, q\} : r \mid c\} = \frac{q}{r}) \\
&= q^2 \sum_{r \mid q} \frac{\mu(r)}{r^2}.
\end{aligned}$$

Now we compute $\#(a,b)$ with $ad - bc \equiv 1 \pmod{q}$. Note that $cx_1x + dy_1x + qy = 1$ implies that $a_0d - b_0c \equiv 1 \pmod{q}$, where $a_0 = xy_1$ and $b_0 = -x_1x$. That is, (a_0, b_0) is a particular solution. Let (a,b) be arbitrary solution. Set $\alpha = a - a_0$ and $\beta = b - b_0$, then $\alpha d - \beta c \equiv 0 \pmod{q}$. Let $t_0 \equiv a_0\beta - b_0\alpha \pmod{q}$. Then we obtain

$$ct = a_0c\beta - b_0c\alpha = a_0c\beta - (a_0d - 1)\alpha = \alpha - a_0(c\beta - d\alpha) = \alpha - a_0(\alpha d - \beta c) \equiv \alpha \pmod{q}$$

and

$$dt = a_0 d\beta - b_0 d\alpha = (1 + b_0 c)\beta - b_0 d\alpha = \beta + b_0(c\beta - d\alpha) = \beta - b_0(\alpha d - \beta c) \equiv \beta \pmod{q}.$$

Therefore, the general solution has the form

$$a = a_0 + ct, \quad b = b_0 + dt, \quad t \in \mathbb{Z}/q\mathbb{Z}.$$

These solutions are different each other. Indeed, if $ct_1 \equiv ct_2, dt_1 \equiv dt_2 \pmod{q}$, then

$$t_1 - t_2 \equiv (a_0 d - b_0 c)(t_1 - t_2) = a_0 \cdot d(t_1 - t_2) - b_0 \cdot c(t_1 - t_2) \equiv 0 \pmod{q}.$$

It the follows that $\#(a, b) = q$. Thus

$$|\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})| = q \cdot q^2 \sum_{r|q} \frac{\mu(r)}{r^2} = q^3 \prod_{p|q} \left(1 + \frac{\mu(p)}{p^2}\right) = q^3 \prod_{p|q} \left(1 - \frac{1}{p^2}\right)$$

since $\mu(n)$ is multiplicative. □

Theorem 2.4. The number of inequivalent cusps of $\Gamma(q)$ is

$$h_q = \begin{cases} 3 & \text{if } q = 2, \\ \dfrac{1}{2} q^2 \displaystyle\prod_{p|q} \left(1 - \frac{1}{p^2}\right) & \text{if } q \geq 3. \end{cases}$$

Proof. We claim the set of cusps is precisely the $\Gamma(q)$-equivalence classes of the projective line

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

Take a parabolic element of $\Gamma(q)$:

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}), \quad a \equiv d \equiv 1 \pmod{q}, \quad b \equiv c \equiv 0 \pmod{q}, \quad |a + d| = 2.$$

Let $\mathfrak{a}$ be fixed point of γ. If $c = 0$, $\mathfrak{a} = \infty \in \mathbb{Q} \cup \{\infty\}$. If $c \neq 0$, $\gamma\mathfrak{a} = \mathfrak{a}$ implies that $\mathfrak{a} = -(d - a)/2c \in \mathbb{Q}$. Thus $\mathfrak{a} \in \mathbb{Q} \cup \{\infty\}$. On the other hand, given arbitrary $\mathfrak{b} \in \mathbb{Q} \cup \{\infty\}$, we construct an parabolic element in $\mathrm{SL}_2(\mathbb{Z})$ fixing $\mathfrak{b}$. If $\mathfrak{b} = \infty$, clearly it is fixed point of

$$\gamma_\infty = \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix}.$$

Now let $\mathfrak{b} = u/v \in \mathbb{Q}$ with $u, v \in \mathbb{Z}$ coprime, then

$$\gamma_\mathfrak{b} = \begin{pmatrix} 1 - quv & qu^2 \\ -qv^2 & 1 + quv \end{pmatrix}$$

is a parabolic element of $\Gamma(q)$ which fixes $\mathfrak{b}$. Therefore the set of cusps coincide with $\mathbb{Q} \cup \{\infty\}$.

Let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, $\mathfrak{a} = u_1/v_1$ and $\mathfrak{b} = u_2/v_2$. If $\gamma\mathfrak{a} = \mathfrak{b}$, there exists $\lambda \in \mathbb{Q}$ with

$$\gamma \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \lambda \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}.$$

Write $\lambda = r/s$ with $r, s \in \mathbb{Z}$ coprime. Let $d = \gcd(\gamma u_1, \gamma v_1)$. Since $\gcd(u_1, v_1) = 1$, there exists $x, y \in \mathbb{Z}$ with $u_1 y - v_1 x = 1$, then

$$\det \begin{pmatrix} \gamma u_1 & x \\ \gamma v_1 & y \end{pmatrix} = \det(\gamma) \det \begin{pmatrix} u_1 & x \\ v_1 & y \end{pmatrix} = 1 \cdot 1 = 1.$$

It follows that $d^2 \mid 1$, so $d = 1$. Since $s \mid \gamma u_1, \gamma v_1$, we obtain $s = 1$, so $\lambda = r \in \mathbb{Z}$. However, $\gcd(\lambda u_2, \lambda v_2) = |\lambda| = 1 = \gcd(\gamma u_1, \gamma v_1)$, we have $\lambda = \pm 1$.

Let $\mathfrak{a} = u_1/v_1$ and $\mathfrak{b} = u_2/v_2$, then $\mathfrak{a}$ and $\mathfrak{b}$ are equivalent (i.e., there exists $\gamma \in \Gamma(q)$ such that $\gamma\mathfrak{a} = \mathfrak{b}$) implies that

$$\gamma \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \pm \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}.$$

We claim further that $\mathfrak{a}$ and $\mathfrak{b}$ are equivalent iff

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \equiv \pm \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \pmod{q}.$$

If $\mathfrak{a}$ and $\mathfrak{b}$ are equivalent, then

$$\gamma \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \pm \begin{pmatrix} u_2 \\ v_2 \end{pmatrix},$$

take both sides modulo q we obtain

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \equiv \pm \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \pmod{q}.$$

For the converse, assume that

$$\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \equiv \epsilon \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \pmod{q}$$

holds for some ϵ, and let

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \epsilon \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}.$$

Since $\gcd(u_1, v_1) = 1$, there exists integers x, y with $u_1 y - v_1 x = 1$, and set

$$\sigma = \begin{pmatrix} u_1 & x \\ v_1 & y \end{pmatrix},$$

then $\det(\sigma) = 1$. Now assume

$$\tau = \begin{pmatrix} w_1 & x + qt_1 \\ w_2 & y + qt_2 \end{pmatrix}$$

for some $t_1, t_2 \in \mathbb{Z}$ satisfying $\tau \equiv \sigma \pmod{q}$, we show that such t_1, t_2 exists. Note that

$$\det(\tau) = w_1(y + qt_2) - w_2(x + qt_1) = (w_1 y - w_2 x) + q(w_1 t_2 - w_2 t_1).$$

From our assumption,

$$w_1 y - w_2 x = u_1 y - v_1 x + q(\cdots) = 1 + qk \quad k \in \mathbb{Z}.$$

We need $\det(\tau) = 1 + qk + q(w_1 t_2 - w_2 t_1) = 1$, so it suffices to show there exists $t_1, t_2 \in \mathbb{Z}$ such that $w_1 t_2 - w_2 t_1 = -k$. Clearly the assertion holds since $\gcd(w_1, w_2) = 1$. Therefore

$$\tau \sigma^{-1} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \epsilon \begin{pmatrix} u_2 \\ v_2 \end{pmatrix},$$

$\mathfrak{a}$ and $\mathfrak{b}$ are equivalent.

Thus the number of inequivalent cusps of $\Gamma(q)$ is precisely the number of orbits

$$S = \{(u, v) \in (\mathbb{Z}/q\mathbb{Z})^2 : \gcd(\gcd(u, v), q) = 1\}$$

under ± 1. When $q = 2$, there exists 3 cusps

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

When $q \geq 3$, from the proof of Theorem 2.3, we have $h_q = \frac{1}{2} q^2 \prod_{p|q} \left(1 - \frac{1}{p^2}\right)$. □

Any subgroup of the modular group which contains $\Gamma(q)$ is called a **congruence subgroup of level** q. Here are two examples;

$$\Gamma_0(q) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{q} \right\},$$

$$\Gamma_1(q) = \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{q} \right\}.$$

The first one is called the **Hecke congruence group**. It is not a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

Theorem 2.5. A set of representatives for $\Gamma_0(q)\backslash\Gamma_0(1)$ is given by

$$\begin{pmatrix} * & * \\ u & v \end{pmatrix} \in \Gamma_0(1) \quad \text{with } v \mid q \text{ and } u \pmod{q/v}.$$

Hence the index is

$$\nu_q = [\Gamma_0(1) : \Gamma_0(q)] = q \prod_{p \mid q} \left(1 + \frac{1}{p}\right).$$

Proof. First we show, for arbitrary

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(1),$$

in its coset $\Gamma_0(q)g$ there exists a unique representative

$$\begin{pmatrix} * & * \\ u & v \end{pmatrix} \in \Gamma_0(1)$$

such that $v \mid q$ and u is uniquely determined modulo q/v.

Let

$$h = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_0(q),$$

note that

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} * & * \\ \gamma a + \delta c & \gamma b + \delta d \end{pmatrix}.$$

Since $h \in \Gamma_0(q)$, $\gamma \equiv 0 \pmod{q}$. Meanwhile, $h \in \mathrm{SL}_2(\mathbb{Z})$ implies that $\alpha\delta - \beta\gamma = 1$, so $\alpha\delta \equiv 1 \pmod{q}$, and so $\gcd(\delta, q) = 1$. Consequently, $\gcd(\gamma b + \delta d, q) = \gcd(\delta d, q) = \gcd(d, q)$. (d, q) is invariant under left multiplication by $\Gamma_0(q)$. Set $v_0 = \gcd(d, q)$.

Now we want to find γ, δ with $\gamma b + \delta d = v_0$, $q \mid \gamma$ and $(\gamma, \delta) = 1$. Let $\gamma = q\gamma_1$, the equation $\gamma b + \delta d = v_0$ becomes $qb\gamma_1 + d\delta = v_0$. By dividing v_0, $(qb/v_0)\gamma_1 + (d/v_0)\delta = 1$. However, $g \in \Gamma_0(1)$ implies that $\gcd(b, d) = 1$, so $\gcd(qb, d) = \gcd(q, d) = v_0$, so $\gcd(qb/v_0, d/v_0) = 1$. Bézout's theorem yields one solution $(\gamma_1^{(0)}, \delta^{(0)}) \in \mathbb{Z}^2$. Hence all solutions are given by $(\gamma_1, \delta) = (\gamma_1^{(0)} + (d/v_0)t, \delta^{(0)} - (q/v_0)bt)$, $t \in \mathbb{Z}$. We now choose t so that $\gcd(\delta, q) = 1$. First, for every prime $\ell \mid (q/v_0)$, reducing $(q/v_0)b\gamma_1 + (d/v_0)\delta = 1$ modulo ℓ gives $(d/v_0)\delta \equiv 1 \pmod{\ell}$, hence $\ell \nmid \delta$. Therefore $\gcd(\delta, q/v_0) = 1$ for every solution. Next let $\ell \mid v_0$ with $\ell \nmid (q/v_0)$. Since $\ell \mid v_0 \mid d$ and $\gcd(b, d) = 1$, we have $\ell \nmid b$, and thus $\ell \nmid (q/v_0)b$. Hence modulo ℓ, there is exactly one residue class of t for which $\delta = \delta^{(0)} - (q/v_0)bt \equiv 0 \pmod{\ell}$, namely $t \equiv \delta^{(0)}((q/v_0)b)^{-1} \pmod{\ell}$. Choose t avoiding this residue class for every prime $\ell \mid v_0$ with $\ell \nmid (q/v_0)$. By the Chinese Remainder

Theorem, such a choice of t exists. For this choice we have $\gcd(\delta, v_0) = 1$. Combining the two parts, we obtain $\gcd(\delta, q) = 1$. Now set $\gamma = q\gamma_1$. Then $q \mid \gamma$, and any common divisor of γ and δ must divide both q and δ, so $\gcd(\gamma, \delta) = 1$. Therefore there exist $\alpha, \beta \in \mathbb{Z}$ such that $\alpha\delta - \beta\gamma = 1$, and hence

$$h = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_0(q).$$

Moreover, $\gamma b + \delta d = q\gamma_1 b + \delta d = v_0$.

u is uniquely determined modulo q/v. Fix $v = \gcd(d, q)$ and consider another pair (γ', δ'). We have $v = \gamma b + \delta d = \gamma' b + \delta' d$ and $ad - bc = 1$ implies that $\gcd(b, d) = 1$, we have $(\gamma', \delta') = (\gamma + dt, \delta - bt)$ for integer t. For the new matrix to still lie in $\Gamma_0(q)$, $\gamma' \equiv 0 \pmod{q}$, so $dt \equiv 0 \pmod{q}$. This implies that $(d/v)t \equiv 0 \pmod{q/v}$. But note that $\gcd(d/v, q/v) = 1$, $t \equiv 0 \pmod{q/v}$. Therefore $u' = \gamma' a + \delta' c = (\gamma + dt)a + (\delta - bt)c = u + t$ is unique modulo q/v.

Given (u, v) with $v \mid q$ and $\gcd(u, v) = 1$, there exists $x, y \in \mathbb{Z}$ such that $xv - yu = 1$. Therefore

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

We have established a one-to-one correspondence between a set of representatives for $\Gamma_0(q)\backslash\Gamma_0(1)$ and

$$\begin{pmatrix} * & * \\ u & v \end{pmatrix} \in \Gamma_0(1) \quad \text{with } v \mid q \text{ and } u \pmod{q/v}.$$

Consequently,

$$\nu_q = [\Gamma_0(1) : \Gamma_0(q)] = \sum_{v \mid q} \sum_{\substack{u \pmod{q/v} \\ \gcd(u,v)=1}} 1.$$

We claim that ν_q is multiplicative. That is, $(q_1, q_2) = 1$ implies that $\nu_{q_1 q_2} = \nu_{q_1}\nu_{q_2}$. From the proof of Theorem 2.3, natural homomorphism $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z})$ is surjective. Let

$$B(q) = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}) \right\}.$$

Then $\Gamma_0(q)$ is the complete preimage of $B(q)$ under the natural homomorphism. So $\nu_q = [\Gamma_0(1) : \Gamma_0(q)] = [\mathrm{SL}_2(\mathbb{Z}/q\mathbb{Z}) : B(q)]$. Thanks to the Chinese Remainder Theorem, we have $\mathrm{SL}_2(\mathbb{Z}/q_1q_2\mathbb{Z}) \cong \mathrm{SL}_2(\mathbb{Z}/q_1\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z}/q_2\mathbb{Z})$. Under this isomorphism, $B(q_1q_2) \cong B(q_1) \times B(q_2)$.

So

$$\begin{aligned}\nu_{q_1q_2} &= [\mathrm{SL}_2(\mathbb{Z}/q_1q_2\mathbb{Z}) : B(q_1q_2)] \\ &= [\mathrm{SL}_2(\mathbb{Z}/q_1\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z}/q_2\mathbb{Z}) : B(q_1) \times B(q_2)] \\ &= [\mathrm{SL}_2(\mathbb{Z}/q_1\mathbb{Z}) : B(q_1)] \cdot [\mathrm{SL}_2(\mathbb{Z}/q_2\mathbb{Z}) : B(q_2)] \\ &= \nu_{q_1}\nu_{q_2}.\end{aligned}$$

So let $q = p^\alpha$ and we compute ν_q. If $v = 1$, $p^\alpha/1 = p^\alpha$ and $\gcd(u, 1) = 1$ holds for all u, the number is p^α; if $v = p^k$ where $1 \leq k \leq \alpha - 1$, $p^\alpha/p^k = p^{\alpha-k}$ and $\gcd(u, p^k) = 1$ iff $(u, p) = 1$, the number is precisely $\varphi(p^{\alpha-k}) = p^{\alpha-k} - p^{\alpha-k-1}$; if $v = p^\alpha$, clearly the number is 1. Now sum over all $k = 0, 1, \cdots, \alpha$,

$$\nu_{p^\alpha} = p^\alpha + (p^{\alpha-1} - p^{\alpha-2}) + \cdots + (p^1 - p^0) + 1 = p^\alpha \left(1 + \frac{1}{p}\right).$$

It follows that for q general,

$$\nu_q = \prod_{p^\alpha || q} p^\alpha \left(1 + \frac{1}{p}\right) = q \prod_{p|q} \left(1 + \frac{1}{p}\right).$$

□

Decompose $\Gamma_0(q)$ into a disjoint union of cosets,

$$\Gamma_0(q)\backslash\Gamma = \left\{\delta : \Gamma = \bigsqcup_i \Gamma_0(q)\delta_i\right\},$$

then

$$\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H}) = [\Gamma_0(1) : \Gamma_0(q)]\,\mathrm{vol}(\Gamma\backslash\mathbb{H}) = \frac{\pi}{3} q \prod_{p|q} \left(1 + \frac{1}{p}\right).$$

Theorem 2.6. The cusps $u_1/v_1, u_2/v_2 \in \mathbb{Q} \cup \{\infty\}$, given in lowest terms, are equivalent under $\Gamma_0(q)$ iff there exists $v \mid q$ such that $\gcd(q, v_1) = \gcd(q, v_2) = v$ and $u_1v_1 \equiv u_2v_2 \pmod{\gcd(q, v^2)}$. Hence the number of inequivalent cusps is

$$h = \sum_{vw=q} \varphi(\gcd(v, w)).$$

Proof. Fixing arbitrary elements

$$\begin{pmatrix} u_1 & \overline{v_1} \\ v_1 & \overline{u_1} \end{pmatrix}, \begin{pmatrix} u_2 & \overline{v_2} \\ v_2 & \overline{u_2} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

(the existence of $\overline{u_1}, \overline{v_1}$ (resp. $\overline{u_2}, \overline{v_2}$) is from $\gcd(u_1, v_1) = \gcd(u_2, v_2) = 1$). We need to examine the statement

$$\exists \gamma \in \Gamma_0(q) : \begin{pmatrix} u_1 & \overline{v_1} \\ v_1 & \overline{u_1} \end{pmatrix} \infty = \gamma \begin{pmatrix} u_2 & \overline{v_2} \\ v_2 & \overline{u_2} \end{pmatrix} \infty.$$

This is clearly

$$\begin{aligned}
&\iff \exists \gamma \in \Gamma_0(q) : \exists n \in \mathbb{Z} : \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} u_1 & \overline{v_1} \\ v_1 & \overline{u_1} \end{pmatrix}^{-1} \gamma \begin{pmatrix} u_2 & \overline{v_2} \\ v_2 & \overline{u_2} \end{pmatrix} \\
&\iff \exists n \in \mathbb{Z} : \begin{pmatrix} u_1 & \overline{v_1} \\ v_1 & \overline{u_1} \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_2 & \overline{v_2} \\ v_2 & \overline{u_2} \end{pmatrix}^{-1} \in \Gamma_0(q) \\
&\iff n \in \mathbb{Z} : \begin{pmatrix} u_1 & \overline{v_1} + nu_1 \\ v_1 & \overline{u_1} + nv_1 \end{pmatrix} \begin{pmatrix} u_2 & -\overline{v_2} \\ -v_2 & \overline{u_2} \end{pmatrix} \in \Gamma_0(q) \\
&\iff \exists n \in \mathbb{Z} : \overline{u_2} v_1 - \overline{u_1} v_2 - n v_1 v_2 \equiv 0 \pmod{q} \\
&\iff \exists m, n \in \mathbb{Z} : \overline{u_2} v_1 - \overline{u_1} v_2 = qm + n v_1 v_2 \\
&\iff \overline{u_2} v_1 \equiv \overline{u_1} v_2 \pmod{\gcd(q, v_1 v_2)}.
\end{aligned}$$

Let us examine the last condition. Observe that by $\gcd(\overline{u_1}, v_1) = \gcd(\overline{u_2}, v_2) = 1$ the congruence forces $\gcd(q, v_1) \mid v_2$ and $\gcd(q, v_2) \mid v_1$, i.e., $\gcd(q, v_1) = \gcd(q, v_2)$. Denoting this common value by v and writing $q = vw$, the congruence becomes $\overline{u_2} v_1 / v \equiv \overline{u_1} v_2 / v \pmod{\gcd(w, v_1 v_2 / v)}$. Note that $\gcd(v_1/v, q/v) = \gcd(v_1, q)/v = 1$ and $\gcd(v_2/v, q/v) = \gcd(v_2, q)/v = 1$, i.e., both v_1/v and v_2/v are coprime to w. In particular, $v_1 v_2 / v = v(v_1/v)(v_2/v)$ shows that $\gcd(w, v_1 v_2/v) = \gcd(w, v)$ and the congruence further simplifies to $\overline{u_2} v_1 / v \equiv \overline{u_1} v_2 / v \pmod{\gcd(w, v)}$. Note also that $u_1 \overline{u_1}$ and $u_2 \overline{u_2}$ are $\equiv 1 \pmod{v}$, hence the congruence is equivalent to $u_1 v_1 / v \equiv u_2 v_2 / v \pmod{\gcd(w, v)}$. Multiplying this by v we obtain the congruence in the theorem.

For any decomposition $q = vw$ and any reduced residue class $u' \bmod \gcd(v, w)$ we pick some $u \in \mathbb{Z}$ such that $\gcd(u, v) = 1$ and $u \equiv u' \pmod{\gcd(v, w)}$. This exists by the Chinese Remainder Theorem, e.g., we can take $u \in \mathbb{Z}$ such that $u \equiv 1 \pmod{p}$ for any prime $p \mid v$ with $p \nmid w$ and also $u \equiv u' \pmod{\gcd(v, w)}$. By the above discussion, the resulting rational numbers u/v represent the $\Gamma_0(q)$-orbits of $\mathbb{Q} \cup \{\infty\}$, and their number equals $\sum_{q=vw} \varphi(\gcd(v, w))$. □

2.5 Double coset decomposition

Throughout Γ is a Fuchsian group of the first kind which contains parabolic motions. Select a complete set of inequivalent cusps for Γ, say $\mathfrak{a}, \mathfrak{b}, \cdots$, and the corresponding scaling matrices

$\sigma_{\mathfrak{a}}, \sigma_{\mathfrak{b}}, \cdots$ in $\mathrm{SL}_2(\mathbb{R})$. Recall that $\sigma_{\mathfrak{a}}^{-1}\mathfrak{a} = \sigma_{\mathfrak{b}}^{-1}\mathfrak{b} = \cdots = \infty$ and $\sigma_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} = \sigma_{\mathfrak{b}}^{-1}\Gamma_{\mathfrak{b}}\sigma_{\mathfrak{b}} = \cdots = B$, where

$$B = \left\{ \pm \begin{pmatrix} 1 & n \\ & 1 \end{pmatrix} : n \in \mathbb{Z} \right\}.$$

Theorem 2.7. For any pair of cusps $\mathfrak{a}, \mathfrak{b}$ for Γ the set $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ is partitioned into disjoint double cosets

$$\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} = \delta_{\mathfrak{a}\mathfrak{b}}B \cup \bigcup_{c>0} \bigcup_{d \pmod c} B \begin{pmatrix} * & * \\ c & d \end{pmatrix} B$$

where $\delta_{\mathfrak{a}\mathfrak{b}}B = B$ if $\mathfrak{a} = \mathfrak{b}$ or else it is the empty set, and the union is taken over $c > 0$ and $d \pmod c$ such that

$$\begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}.$$

Proof. Define

$$\Omega_\infty = \{\omega \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} : \omega\infty = \infty\}.$$

Assume that Ω_∞ is non-empty. For arbitrary $\omega \in \Omega_\infty$, i.e., there exists $\gamma \in \Gamma$ with $\omega = \sigma_{\mathfrak{a}}^{-1}\gamma\sigma_{\mathfrak{b}}$. Since $\omega\infty = \infty$, we have $(\sigma_{\mathfrak{a}}^{-1}\gamma\sigma_{\mathfrak{b}})\infty = \infty$, and so $\gamma(\sigma_{\mathfrak{b}}\infty) = \sigma_{\mathfrak{a}}\infty$. But $\sigma_{\mathfrak{a}}\infty = \mathfrak{a}$ and $\sigma_{\mathfrak{b}}\infty = \mathfrak{b}$, so $\gamma\mathfrak{b} = \mathfrak{a}$. This implies that the cusps $\mathfrak{a}, \mathfrak{b}$ are equivalent; hence $\mathfrak{a} = \mathfrak{b}$, so $\gamma\mathfrak{a} = \mathfrak{a}$ and then $\gamma \in \Gamma_{\mathfrak{a}}$. Clearly $\Omega_\infty = \{\omega \in \sigma_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} : \omega\infty = \infty\} = \sigma_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} = B$ if $\mathfrak{a} = \mathfrak{b}$. On the other hand if $\mathfrak{a} \neq \mathfrak{b}$, $\gamma\mathfrak{b} = \mathfrak{a}$ leads to a contradiction of our assumption $\mathfrak{a} \not\sim \mathfrak{b}$, hence $\Omega_\infty = \emptyset$.

Now consider elements in $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \setminus \Omega_\infty$. Let

$$\begin{pmatrix} a & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}},$$

since ω does not fix ∞, $c \neq 0$, without loss of generality we may assume $c > 0$. Take arbitrary

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in B,$$

the relation

$$\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & * \\ c & d \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a + cm & * \\ c & d + cn \end{pmatrix}.$$

shows that the double coset

$$B \begin{pmatrix} * & * \\ c & d \end{pmatrix} B$$

determines c uniquely and d modulo integral multiples of c.

Assume there exists another matrix

$$\omega' = \begin{pmatrix} a' & * \\ c & d' \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}},$$

where $d' \equiv d \pmod{c}$, we claim that $\omega' \in B\omega B$. Indeed there exists $n \in \mathbb{Z}$ with $d' = d + cn$, then

$$\omega' \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} * & * \\ c & d' \end{pmatrix} \in B\omega,$$

we obtain $\omega' \in B\omega \subset B\omega B$.

Every $\omega \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ either fix ∞ or do not fix ∞. If ω does not fix ∞, by multiplying if necessary, we have $c > 0$. So every ω belongs to one of these cosets. Clearly two different cosets do not meet. We finish the proof. □

Put

$$\mathcal{C}(\mathfrak{a}, \mathfrak{b}) = \left\{ c > 0 : \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\}.$$

Let $c(\mathfrak{a}, \mathfrak{b})$ denote the smallest element of $\mathcal{C}(\mathfrak{a}, \mathfrak{b})$, and put $c(\mathfrak{a}) = c(\mathfrak{a}, \mathfrak{a})$. Define $c_{\mathfrak{ab}} = \max\{c(\mathfrak{a}), c(\mathfrak{b})\}$.

Theorem 2.8. For any $X > 0$ we have

$$\sum_{0<c\le X} c^{-1} \# \left\{ d \pmod{c} : \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\} \le c_{\mathfrak{ab}}^{-1} X.$$

Hence for any $c \in \mathcal{C}(\mathfrak{a}, \mathfrak{b})$ we have

$$\# \left\{ d \pmod{c} : \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\} \le c_{\mathfrak{ab}}^{-1} c^2.$$

Proof. Without loss of generality we may assume $c(\mathfrak{a}) \ge c(\mathfrak{b})$. In $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$, choose arbitrary two matrices with $0 < c, c' \le X$:

$$\omega = \begin{pmatrix} * & * \\ c & d \end{pmatrix}, \quad \omega' = \begin{pmatrix} * & * \\ c' & d' \end{pmatrix}.$$

Moreover we may assume $0 \le d < c, 0 \le d' < c$. Let $\omega'' = \omega'\omega^{-1} \in (\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}})(\sigma_{\mathfrak{b}}^{-1}\Gamma\sigma_{\mathfrak{a}}) = \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{a}}$. Then

$$\omega'' = \begin{pmatrix} * & * \\ c' & d' \end{pmatrix} \begin{pmatrix} d & * \\ -c & * \end{pmatrix} = \begin{pmatrix} * & * \\ c'd - cd' & * \end{pmatrix}.$$

Set $c'' = c'd - cd'$. From the definition of $c(\mathfrak{a})$, $|c''| \ge c(\mathfrak{a})$.

If $c'' = 0$, $\omega'' \in B$, so

$$\omega'\omega^{-1} = \omega'' = \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}.$$

This implies that $c' = c$ and $d' = d + mc$, thus $m = 0$. We obtain $c = c'$ and $d = d'$.

If $c'' \neq 0$, $|c'd - cd'| \geq c(\mathfrak{a})$. It follows that

$$\left| \frac{d'}{c'} - \frac{d}{c} \right| \geq \frac{c(\mathfrak{a})}{cc'} \geq \frac{c(\mathfrak{a})}{cX}.$$

since $c' \leq X$. Let $\{r_1, r_2, \ldots, r_N\}$ be the set of fractions d/c corresponding to the elements in $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ with $0 < c \leq X$, such that $0 \leq r_1 < r_2 < \cdots < r_N < 1$. The length of the interval $[0, 1)$ is 1. By summing the gaps between successive points (and treating $r_{N+1} = r_1 + 1$), $\sum_{i=1}^{N}(r_{i+1} - r_i) = r_{N+1} - r_1 = 1$. Applying the bound from (2.37) to each gap $(r_{i+1} - r_i) \geq c(\mathfrak{a})/(c_i X)$:

$$1 \geq \sum_{i=1}^{N} \frac{c(\mathfrak{a})}{c_i X} = \frac{c(\mathfrak{a})}{X} \sum_{i=1}^{N} \frac{1}{c_i},$$

which gives

$$\sum_{i=1}^{N} \frac{1}{c_i} \leq \frac{X}{c(\mathfrak{a})}.$$

The sum $\sum 1/c_i$ is exactly the sum over c of c^{-1} times the number of $d \pmod c$. Thus

$$\sum_{i=1}^{N} \frac{1}{c_i} = \sum_{0<c\leq X} \frac{1}{c} \# \left\{ d \pmod c : \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\}.$$

Consequently

$$\sum_{0<c\leq X} c^{-1} \# \left\{ d \pmod c : \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\} \leq \frac{X}{c(\mathfrak{a})} = c_{\mathfrak{a}\mathfrak{b}}^{-1} X.$$

For a fixed c, taking $X = c$ and considering only the term for that c,

$$\frac{1}{c} \# \left\{ d \pmod c : \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\} \leq \sum_{0<t\leq c} t^{-1} \# \left\{ d \pmod t : \begin{pmatrix} * & * \\ t & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\} \leq c_{\mathfrak{a}\mathfrak{b}}^{-1} c.$$

Multiplying by c yields the assertion. □

Theorem 2.9. Let $\mathfrak{a}$ be a cusp for Γ, $z \in \mathbb{H}$ and $Y > 0$. We have

$$\# \left\{ \gamma \in \Gamma_{\mathfrak{a}} \backslash \Gamma : \operatorname{Im}(\sigma_{\mathfrak{a}}^{-1}\gamma z) > Y \right\} < 1 + \frac{10}{c(\mathfrak{a})Y}.$$

Proof. We only prove when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$.

Let $\Gamma' = \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{a}}$. In this case, the cusp is $\mathfrak{a}' = \infty$, the scaling matrix is $\sigma_{\mathfrak{a}'} = I$ and $\Gamma'_{\mathfrak{a}'} = B$. If we set $S' = \{\gamma' \in B\backslash\Gamma' : \mathrm{Im}(I^{-1}\gamma' z') > y\}$, we show that $\#S' = \#S$ where $S = \{\gamma \in \Gamma_{\mathfrak{a}}\backslash\Gamma : \mathrm{Im}(\sigma_{\mathfrak{a}}^{-1}\gamma z) > Y\}$ and $c(\mathfrak{a}) = c(\mathfrak{a}')$. For arbitrary $\gamma \in \Gamma$, let $f(\gamma) = \sigma_{\mathfrak{a}}^{-1}\gamma\sigma_{\mathfrak{a}} = \gamma' \in \Gamma'$. Since $\sigma_{\mathfrak{a}} \in \mathrm{SL}_2(\mathbb{R})$ admits an inverse, f is a group isomorphism. Two elements $\gamma_1, \gamma_2 \in \Gamma$ belongs to one left coset iff $\gamma_1\gamma_2^{-1} \in \Gamma_{\mathfrak{a}}$. It follows that $\sigma_{\mathfrak{a}}^{-1}(\gamma_1\gamma_2^{-1})\sigma_{\mathfrak{a}} \in \sigma_{\mathfrak{a}}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} = B$, so $(\sigma_{\mathfrak{a}}^{-1}\gamma_1\sigma_{\mathfrak{a}})(\sigma_{\mathfrak{a}}^{-1}\gamma_2\sigma_{\mathfrak{a}})^{-1} \in B$, i.e., $\gamma_1'(\gamma_2')^{-1} \in B$. That is, γ_1' and γ_2' belongs to the same left coset in Γ', we obtain a bijection between $\Gamma_{\mathfrak{a}}\backslash\Gamma$ and $B\backslash\Gamma'$. On the other hand, let $z' = \sigma_{\mathfrak{a}}^{-1}z$, $\gamma' z' = (\sigma_{\mathfrak{a}}^{-1}\gamma\sigma_{\mathfrak{a}})(\sigma_{\mathfrak{a}}^{-1}z) = \sigma^{-1}\gamma z$. Thus $\mathrm{Im}(\gamma' z') = \mathrm{Im}(\sigma_{\mathfrak{a}}^{-1}\gamma z)$, $\#S = \#S'$. Moreover, from the definition,

$$c(\mathfrak{a}) = \left\{ c > 0 : \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{a}} \right\},$$

and

$$c(\mathfrak{a}') = \left\{ c > 0 : \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in I^{-1}\Gamma' I = \Gamma' \right\},$$

we obtain $c(\mathfrak{a}) = c(\mathfrak{a}')$ since $\Gamma' = \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{a}}$.

Consequently, we may assume $\mathfrak{a} = \infty$, $\sigma_{\mathfrak{a}} = I$ and $\Gamma_{\mathfrak{a}} = B$. Then the strip $P = \{z \in \mathbb{H} : 0 \le \mathrm{Re}(z) < 1, \mathrm{Im}(z) > 0\}$ is a fundamental domain of $\Gamma_{\mathfrak{a}} = B$. Indeed, clearly P is a domain. Assume there exists $z_1, z_2 \in P$ with $z_2 = \gamma_n z_1$ for some non-trivial $\gamma_n \in B$, here

$$\gamma_n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in B.$$

Let $z_1 = x_1 + \mathrm{i}y_1$ and $z_2 = x_2 + \mathrm{i}y_2$, then $y_1 = y_2$ and $x_1 + n = x_2$, which contradicts to $0 < x_1, x_2 < 1$. The closure of P is $\overline{P} = \{z \in \mathbb{H} : 0 \le \mathrm{Re}(z) \le 1, \mathrm{Im}(z) > 0\}$. For arbitrary $z = x + \mathrm{i}y \in \mathbb{H}$, there exists unique integer $m = [x]$ with $m \le x < m + 1$. It follows that $\gamma_{-m} z \in \overline{P}$. That is, every orbit admits a representative in $\overline{P}$. Let $D = \{z \in \mathbb{H} : -1/2 < \mathrm{Re}(z) < 1/2, |z| > 1\}$ be the standard polygon of Γ, then $D \subset P$. For any $\omega \in \Gamma$, we have $\#\{B\gamma \in B\backslash\Gamma : \mathrm{Im}(\gamma\omega z) > Y\} = \#\{B\gamma \in B\backslash\Gamma : \mathrm{Im}(\gamma z) > Y\}$ since ω is invertible. When counting in $B\backslash\Gamma$, replacing z with another point on the Γ-orbit does not change the total count; therefore we may assume $z \in D$. Note that

$$\mathrm{Im}(\gamma z) = \frac{\mathrm{Im}(z)}{|cz + d|^2}, \quad \forall z \in \mathbb{H}, \forall \gamma \in \Gamma.$$

From the proof of Theorem 1.2 $\mathrm{Im}(\gamma z) \le \mathrm{Im}(z)$, $|cz + d| \ge 1$ holds for any $\gamma \in \Gamma$.

Assume $z = x + \mathrm{i}y$. $\mathrm{Im}(\gamma z) > Y$ implies that $Y < y/|cz+d|^2 \le y$. Since $|cz+d|^2 = (cx+d)^2 + (cy)^2$, we obtain

$$(cy)^2 \le (cx+d)^2 + (cy)^2 < \frac{y}{Y} \Longrightarrow c < (yY)^{-1/2}$$

and

$$(cx+d)^2 \le (cx+d)^2 + (cy)^2 < \frac{y}{Y} \Longrightarrow |cx+d| < \left(\frac{y}{Y}\right)^{1/2}.$$

Assume that $C \le c < 2C$, then $|cx+d| < (y/Y)^{1/2}$ implies that

$$\left|x + \frac{d}{c}\right| < \frac{1}{c}\left(\frac{y}{Y}\right)^{1/2} \le \frac{1}{C}\left(\frac{y}{Y}\right)^{1/2}.$$

Thanks to the proof of Theorem 2.8, when the denominators less or equal to $2C$ the distance between any two distinct fractions d_1/c_1 and d_2/c_2 is bounded below $\delta = c(\mathfrak{a})/(c_1c_2) > c(\mathfrak{a})/(2C)^2 = c(\mathfrak{a})/4C^2$. In an interval of length L, the number of points with distance at least δ is at most $1 + L/\delta$. We estimate that

$$1 + \frac{\frac{2}{C}(y/Y)^{1/2}}{\frac{c(\mathfrak{a})}{4C^2}} = 1 + \frac{8C}{c(\mathfrak{a})}\left(\frac{y}{Y}\right)^{1/2}.$$

Note that $y > Y$ implies that $(y/Y)^{1/2} > 1$ and $2C \ge c(\mathfrak{a})$, we obtain $2C/c(\mathfrak{a}) \ge 1$. Adding these bounds for $C = 2^{-1}(yY)^{-1/2}$ with $n \ge 1$, we get

$$\sum_{n\ge 1} \frac{10 \cdot 2^{-n}(yY)^{-1/2}}{c(\mathfrak{a})}\left(\frac{y}{Y}\right)^{1/2} = \frac{10}{c(\mathfrak{a})Y}\sum_{n\ge 1} 2^{-n} = \frac{10}{c(\mathfrak{a})Y}.$$

When $c = 0$, $\gamma \in B$, so $B\gamma = B$ is trivial coset. Finally adding 1 to account for $\Gamma_{\mathfrak{a}}$, we finish the proof. □

2.6 Multiplier systems

For a complex number $z \neq 0$ we choose its argument in $(-\pi, \pi]$. We denote the principal branch of the logarithm by $\log z$, so its real for positive z and

$$\log z = \log|z| + \mathrm{i}\arg z \quad \text{if } z \in \mathbb{C}^\times.$$

For any transformation matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}),$$

we define the **factor of automorphy** as $j(A, z) = cz + d$.

If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix},$$

then

$$Bz = \frac{a'z + b'}{c'z + d'}.$$

Hence

$$j(A, Bz) = c\frac{a'z + b'}{c'z + d'} + d = \frac{c(a'z + b') + d(c'z + d')}{c'z + d'}.$$

Multiplying by $j(B, z) = c'z + d'$, we get

$$j(A, Bz)\, j(B, z) = c(a'z + b') + d(c'z + d').$$

On the other hand,

$$AB = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix},$$

so

$$j(AB, z) = (ca' + dc')z + (cb' + dd').$$

Therefore

$$j(AB, z) = c(a'z + b') + d(c'z + d') = j(A, Bz)\, j(B, z).$$

This proves the chain rule

$$j(AB, z) = j(A, Bz)j(B, z).$$

Define

$$2\pi\omega(A, B) = -\arg j(AB, z) + \arg j(A, Bz) + \arg j(B, z).$$

The chain rule implies that

$$\arg j(AB, z) \equiv \arg j(A, Bz) + \arg j(B, z) \pmod{2\pi}$$

That is, $\omega(A, B) \in \mathbb{Z}$. On the other hand, for arbitrary

$$M = \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R}),$$

we have $j(M, z) = cz+d$ by definition. It follows that $\mathrm{Im}(j(M, z)) = cy$. If $c > 0$, $j(M, z) \in \mathbb{H}$ for all $z \in \mathbb{H}$, so $z \mapsto \arg(j(M, z))$ is continuous on $\mathbb{H}$; similarly for $c < 0$, $z \mapsto \arg(j(M, z))$ is continuous on $\mathbb{H}$; if $c = 0$, $j(M, z) = d \neq 0$ since $ad = 1$. Thus $z \mapsto -\arg j(AB, z) + \arg j(A, Bz) + \arg j(B, z)$ is continuous, which forces $-\arg j(AB, z) + \arg j(A, Bz) + \arg j(B, z)$ is a constant.

This definition is note depend on the choice of $z \in \mathbb{H}$. Moreover $\arg j(A, Bz), \arg j(B, z), \arg j(AB, z) \in (-\pi, \pi]$, we obtain $2\pi\omega(A, B) \in (-3\pi, 3\pi)$, but note that $\omega(A, B) \in \mathbb{Z}$, there must be $\omega(A, B) \in \{-1, 0, 1\}$.

For any real number k we define the **factor system of weight** k by setting

$$w(A, B) = e\left(k\omega(A, B)\right).$$

Note that $w(A, B)$ depends on $k \pmod 1$ and $w(A, B) = 1$ if k is an integer.

Let $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ be a discrete subgroup. A **multiplier system of weight** k **for** Γ is a function $\vartheta \colon \Gamma \to \mathbb{C}$ such that

$$|\vartheta(\gamma)| = 1$$

$$\vartheta(\gamma_1\gamma_2) = w(\gamma_1, \gamma_2)\vartheta(\gamma_1)\vartheta(\gamma_2).$$

Since $w(-I, -I) = e(k)$, the above properties imply that if $-I$ belongs to Γ, then $\vartheta(-1) = \pm e(-k/2)$. Throughout we shall require that

$$\vartheta(-1) = e(-k/2) \text{ if } -I \in \Gamma,$$

which is called the consistency condition (the other choice yields the zero automorphic form only).

2.7 Automorphic forms

Let $k \in \mathbb{Z}$. For a function $f : \mathbb{H} \to \mathbb{C}$, we define the **slash operator**, denoted as $f_{|A}$, by:

$$f_{|A}(z) = j(A, z)^{-k} f(Az) = (cz + d)^{-k} f\left(\frac{az + b}{cz + d}\right)$$

Clearly $f_{|AB} = (f_{|A})_{|B}$. For

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q),$$

we define $\chi(\gamma) \coloneqq \chi(d)$, where χ is Dirichlet character modulo q.

Definition 2.6 (automorphic form)**.** A holomorphic function $f : \mathbb{H} \to \mathbb{C}$ is called a **modular form of weight** k**, level** $\Gamma_0(q)$**, and nebentypus (or character)** χ if it is holomorphic at the cusps (we will introduce the definition of holomorphic at the cusps later) and satisfies

$$f_{|\gamma}(z) = \chi(\gamma) f(z), \quad \forall \gamma \in \Gamma_0(q).$$

The space of such modular forms is denoted by $\mathcal{M}_k(\Gamma_0(q), \chi)$.

Remark 2.1. In general we consider $k \in \mathbb{R}$. The slash operator is defined by

$$f_{|A}(z) = j(A,z)^{-k} f(Az),$$

where

$$j(A,z)^{-k} = \exp\left(-k \log(j(A,z))\right).$$

One may verify that

$$w(A,B) j(AB,z)^k = j(A,Bz)^k j(B,z)^k,$$

and

$$f_{|AB} = w(A,B)(f_{|A})_{|B}.$$

The general automorphic form is defined by replacing χ to general ϑ. But most of the time we only consider case $w = 1$, $\vartheta = \chi$ be Dirichlet character modulo q and $\Gamma = \Gamma_0(q)$.

Since $-I \in \Gamma_0(q)$, we can apply the slash operator definition

$$f_{|-I}(z) = j(-I,z)^{-k} f(z) = (-1)^{-k} f(z) = (-1)^k f(z).$$

On the other hand, applying the transformation law with the nebentypus gives

$$f_{|-I}(z) = \chi(-1) f(z).$$

Equating the two expressions yields

$$(-1)^k f(z) = \chi(-1) f(z).$$

For a non-zero modular form f, this implies that $(-1)^k = \chi(-1)$. Consequently, if the weight k is even (resp. odd), χ must be an even (resp. odd) character, i.e., $\chi(-1) = 1$ (resp. $\chi(-1) = -1$).

When a modular form $f \in \mathcal{M}_k(\Gamma_0(q), \chi)$ admits a non-trivial character χ, the strict period-1 translation symmetry is broken at the cusps. Let $\mathfrak{a}$ be a cusp of the group $\Gamma = \Gamma_0(q)$ and let $\sigma_{\mathfrak{a}} \in \mathrm{SL}_2(\mathbb{R})$ be the scaling matrix that maps ∞ to $\mathfrak{a}$. Let

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

represent the translation $z \mapsto z+1$. Therefore

$$\begin{aligned} f_{|\sigma_{\mathfrak{a}}}(z+1) &= f_{|\sigma_{\mathfrak{a}}}(\beta z) \\ &= j(\sigma_{\mathfrak{a}}, \beta z)^{-k} f(\sigma_{\mathfrak{a}} \beta z) \\ &= j(\sigma_{\mathfrak{a}}, \beta z)^{-k} f(\gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}} z) \qquad\qquad (\text{since } \sigma_{\mathfrak{a}}^{-1} \gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}} = \beta) \\ &= j(\sigma_{\mathfrak{a}}, \beta z)^{-k} j(\gamma_{\mathfrak{a}}, \sigma_{\mathfrak{a}} z)^{k} \chi(\gamma_{\mathfrak{a}}) f(\sigma_{\mathfrak{a}} z) \\ &= \chi(\gamma_{\mathfrak{a}}) f_{|\sigma_{\mathfrak{a}}}(z). \end{aligned}$$

Let $\chi(\gamma_{\mathfrak{a}}) = e(\kappa_{\mathfrak{a}})$, where $0 \leq \kappa_{\mathfrak{a}} < 1$. We define $g(z) = e(-\kappa_{\mathfrak{a}} z) f_{|\sigma_{\mathfrak{a}}}(z)$. Then we can easily verify that $g(z)$ is strictly periodic with period 1:

$$\begin{aligned} g(z+1) &= e(-\kappa_{\mathfrak{a}}(z+1)) f_{|\sigma_{\mathfrak{a}}}(z+1) \\ &= e(-\kappa_{\mathfrak{a}} z) e(-\kappa_{\mathfrak{a}}) \left(e(\kappa_{\mathfrak{a}}) f_{|\sigma_{\mathfrak{a}}}(z) \right) \\ &= e(-\kappa_{\mathfrak{a}} z) f_{|\sigma_{\mathfrak{a}}}(z) \\ &= g(z). \end{aligned}$$

We say that f is **holomorphic at** $\mathfrak{a}$ if $g(z)$ is holomorphic at ∞. In this case, $g(z)$ admits a standard Fourier series expansion in terms of $q = e(z)$:

$$g(z) = \sum_{n=0}^{\infty} a_n e(nz)$$

We obtain the generalized Fourier expansion for f at the cusp $\mathfrak{a}$:

$$f_{|\sigma_{\mathfrak{a}}}(z) = e(\kappa_{\mathfrak{a}} z) \sum_{n=0}^{\infty} a_n e(nz) = \sum_{n=0}^{\infty} a_n e((n+\kappa_{\mathfrak{a}}) z).$$

3 The Eisenstein and the Poincaré Series

3.1 Poincaré series

Set $\mathfrak{a} = \infty$, so $\sigma_{\mathfrak{a}} = I$ and $\Gamma_{\mathfrak{a}} = \Gamma_{\infty} = B$. Moreover set $\Gamma = \Gamma_0(q)$ be the Hecke congruence group. Given $p \colon \mathbb{H} \to \mathbb{C}$ be a holomorphic function which is periodic of period 1, we can define $\pi \colon \Gamma \times \mathbb{H} \to \mathbb{C}$ by

$$\pi(\gamma, z) = \overline{\chi(\gamma)} j_{\gamma}(z)^{-k} p(\gamma z).$$

It is Γ_{∞}-invariant. Indeed, let $\eta \in \Gamma_{\infty}$ and $\gamma \in \Gamma$. We have $j_{\eta\gamma}(z) = j_{\eta}(\gamma z) j_{\gamma}(z)$. For

$$\eta = \pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix},$$

we have $j_\eta(w) = 0 \cdot w \pm 1 = \pm 1$ for arbitrary $w \in \mathbb{C}$, it follows that $j_\eta(\gamma z)^{-k} = (\pm 1)^{-k} = (\pm 1)^k$. By definition $\chi(\eta\gamma) = \chi(\eta)\chi(\gamma)$. If

$$\eta = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix},$$

we have $\chi(\eta) = \chi(1) = 1$; if

$$\eta = \begin{pmatrix} -1 & -n \\ 0 & -1 \end{pmatrix} = -I \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix},$$

we have $\chi(\eta) = \chi(-I) = (-1)^k$ (this is standard condition of χ). Thus $\chi(\eta) = (\pm 1)^k$. Moreover, $\eta w = (\pm w \pm n)/\pm 1 = w + n$ for arbitrary w, so $p(\eta\gamma z) = p(\gamma z + n) = p(\gamma z)$ since p is of period 1. Consequently we have

$$\begin{aligned}
(\eta\gamma, z) &= \overline{\chi(\eta\gamma)} j_{\eta\gamma}(z)^{-k} p(\eta\gamma z) \\
&= \left((\pm 1)^k \overline{\chi(\gamma)}\right) \left((\pm 1)^k j_\gamma(z)^{-k}\right) p(\gamma z) \\
&= (\pm 1)^{2k} \overline{\chi(\gamma)} j_\gamma(z)^{-k} p(\gamma z) \\
&= 1 \cdot \pi(\gamma, z).
\end{aligned}$$

By averaging over $\Gamma_\infty \backslash \Gamma$, we define

$$\mathbf{P}(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \pi(\gamma, z), \quad k > 2.$$

It is called the **Poincaré series associated with the cusp ∞ at the generating function** p.

If $p(z)$ is bounded, we have

$$|\mathbf{P}(z)| \ll \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} |j_\gamma(z)|^{-k} = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \frac{\mathrm{Im}(z)^{-k/2}}{\mathrm{Im}(\gamma z)^{-k/2}} = y^{-k/2} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \mathrm{Im}(\gamma z)^{k/2}.$$

Now we claim that $\mathrm{Im}(\gamma z)$ admits an upper bound. Fix $z = x + \mathrm{i}y \in \mathbb{H}$, we have $\mathrm{Im}(\gamma z) = y/|cz + d|^2$ for arbitrary

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma.$$

If $\gamma \in \Gamma_\infty$, $\gamma z = z + n$ implies that $\mathrm{Im}(\gamma z) = y$. If $\gamma \notin \Gamma_\infty$, we have $|c| \geq c(\infty) > 0$ from the definition of $c(\infty)$. Hence $|cz + d|^2 \geq (\mathrm{Im}(cz + d))^2 = (cy)^2$, so $\mathrm{Im}(\gamma z) = y/|cz + d|^2 \leq y/(c(\infty)^2 y^2) = 1/(c(\infty)^2 y)$. Consequently,

$$\mathrm{Im}(\gamma z) \leq \max\left\{y, \frac{1}{c(\infty)^2 y}\right\}$$

holds for all $\gamma \in \Gamma$. That is, there exists a constant $M(z) > 0$ such that $0 < \mathrm{Im}(\gamma z) \leq M(z)$ for all $\gamma \in \Gamma$. For every γ fixed, we have

$$\mathrm{Im}(\gamma z)^{k/2} = \frac{k}{2} \int_0^{\mathrm{Im}(\gamma z)} Y^{k/2-1} \mathrm{d}Y,$$

so

$$\sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \mathrm{Im}(\gamma z)^{k/2} = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \frac{k}{2} \int_0^{\mathrm{Im}(\gamma z)} Y^{k/2-1} \mathrm{d}Y = \frac{k}{2} \int_0^{M(z)} \#\{\gamma \in \Gamma_\infty \backslash \Gamma : \mathrm{Im}(\gamma z) > Y\} Y^{k/2-1} \mathrm{d}Y.$$

Thanks to Theorem 2.9,

$$\#\{\gamma \in \Gamma_\infty \backslash \Gamma : \mathrm{Im}(\gamma z) > Y\} \leq 1 + \frac{10}{c(\infty)Y},$$

then

$$\sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \mathrm{Im}(\gamma z)^{k/2} \leq \frac{k}{2} \int_0^{M(z)} \left(1 + \frac{10}{c(\infty)Y}\right) Y^{k/2-1} \mathrm{d}Y = \frac{k}{2} \int_0^{M(z)} Y^{k/2-1} \mathrm{d}Y + \frac{5k}{c(\infty)} \int_0^{M(z)} Y^{k/2-2} \mathrm{d}Y.$$

When $k > 2$ the two integers converge, so

$$\sum_{\gamma \in \Gamma_\infty \backslash \Gamma} |\pi(\gamma, z)| < \infty.$$

That is, $\mathbf{P}(z)$ converges absolutely.

Proposition 3.1. $\mathbf{P}(z)$ is an automorphic form. That is,

1. $\mathbf{P}_{|\gamma}(z) = \chi(\gamma)\mathbf{P}(z)$,
2. $\mathbf{P}(z)$ is holomorphic at every cusp.

Proof of (1). We have

$$\begin{aligned} \mathbf{P}_{|\tau}(z) &= j_\tau(z)^{-k} \mathbf{P}(\tau z) \\ &= j_\tau(z)^{-k} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \overline{\chi(\gamma)} j_{\gamma\tau^{-1}}(\tau z)^{-k} p(\gamma\tau z) \quad (\text{let } \gamma \mapsto \gamma\tau^{-1}). \end{aligned}$$

Note that $j_{\gamma\tau^{-1}}(\tau z) = j_\gamma(\tau^{-1}\tau z) j_{\tau^{-1}}(\tau z) = j_\gamma(z) j_{\tau^{-1}}(\tau z)$. On the other hand, $I = \tau^{-1}\tau$ implies that $1 = j_I(z) = j_{\tau^{-1}\tau}(z) = j_{\tau^{-1}}(\tau z) j_\tau(z)$, so $j_{\tau^{-1}}(\tau z) = j_\tau(z)^{-1}$. Then $j_{\gamma\tau^{-1}}(\tau z) = j_\gamma(z) j_\tau(z)^{-1}$, and hence $j_{\gamma\tau^{-1}}(\tau z)^{-k} = j_\gamma(z)^{-k} j_\tau(z)^k$. Moreover $\chi(\tau^{-1}) = \chi(\tau)^{-1} = \overline{\chi(\tau)}$ since $|\chi(\tau)| = 1$, we obtain

$$\begin{aligned} \mathbf{P}_{|\tau}(z) &= j_\tau(z)^{-k} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \overline{\chi(\gamma)} \chi(\tau) j_\gamma(z)^{-k} j_\tau(z)^k p(\gamma z) \\ &= \chi(\tau) \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \overline{\chi(\gamma)} j_\gamma(z)^{-k} p(\gamma z) \\ &= \chi(\tau) \mathbf{P}(z). \end{aligned}$$

□

3.2 Fourier expansion of Poincaré series

Recall we have double coset decomposition

$$\Gamma_0(q) = \Gamma_\infty \cup \bigcup_{\substack{c>0 \\ c\equiv 0 \pmod q}} \bigcup_{\substack{d \pmod c \\ \gcd(d,c)=1}} \Gamma_\infty \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Gamma_\infty.$$

It follows that

$$\begin{aligned}
\mathbf{P}(z) &= \overline{\chi(I)} j_I(z)^{-k} p(z) + \sum_{I \neq \gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_\infty} \sum_{\tau \in \Gamma_\infty} \overline{\chi(\gamma\tau)} j_{\gamma\tau}(z)^{-k} p(\gamma\tau z) \\
&= p(z) + \sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \sum_{\substack{d \pmod c \\ \gcd(d,c)=1}} \sum_{n\in\mathbb{Z}} \overline{\chi(d)} (cz+cn+d)^{-k} p\left(\frac{a}{c} - \frac{1}{c(cz+cn+d)}\right) \\
&= p(z) + \sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \sum_{\substack{d \pmod c \\ \gcd(d,c)=1}} \overline{\chi(d)} \sum_{n\in\mathbb{Z}} (cz+cn+d)^{-k} p\left(\frac{a}{c} - \frac{1}{c(cz+cn+d)}\right).
\end{aligned}$$

Say precisely, let

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ (c>0, q \mid c, \gcd(c,d)=1) \quad \text{and} \quad \tau = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in \Gamma_\infty,$$

we have

$$\gamma\tau z = \begin{pmatrix} a & an+b \\ c & cn+d \end{pmatrix} z = \frac{a(z+n)+b}{c(z+n)+d} = \frac{a}{c} - \frac{1}{c(cz+cn+d)}.$$

Moreover, $\chi(\gamma\tau) = \chi(cn+d) = \chi(d)$ since $q \mid c$.

Now we compute

$$\begin{aligned}
I_\gamma(z) &\coloneqq \sum_{n\in\mathbb{Z}} (cz+cn+d)^{-k} p\left(\frac{a}{c} - \frac{1}{c(cz+cn+d)}\right) \\
&= \sum_{n\in\mathbb{Z}} (cz+cn+d)^{-k} e\left(m\left(\frac{a}{c} - \frac{1}{c(cz+cn+d)}\right)\right) && \text{(Take } p(z)=e(mz)\text{)} \\
&= e\left(\frac{am}{c}\right) \sum_{n\in\mathbb{Z}} (cz+cn+d)^{-k} e\left(-\frac{m}{c^2\left(z+n+\frac{d}{c}\right)}\right) \\
&= e\left(\frac{am}{c}\right) \sum_{n\in\mathbb{Z}} \left(c\left(x+n+\frac{d}{c}\right) + ciy\right)^{-k} e\left(-\frac{m}{c^2\left(x+n+\frac{d}{c}+iy\right)}\right).
\end{aligned}$$

We can apply the Poisson summation formula to f, hence

$$\begin{aligned}
I_\gamma(z) &= e\left(\frac{am}{c}\right)\sum_{\widetilde{n}\in\mathbb{Z}}\int_{\mathbb{R}}\left(c\left(x+v+\frac{d}{c}\right)+c\mathrm{i}y\right)^{-k}e\left(-\frac{m}{c^2\left(x+v+\frac{d}{c}+\mathrm{i}y\right)}\right)e\left(-\widetilde{n}v\right)\mathrm{d}v\\
&= e\left(\frac{am}{c}\right)\sum_{\widetilde{n}\in\mathbb{Z}}\int_{\mathbb{R}}(c(v+\mathrm{i}y))^{-k}e\left(-\frac{m}{c^2(v+\mathrm{i}y)}-\widetilde{n}\left(v-x-\frac{d}{c}\right)\right)\mathrm{d}v \qquad \left(\text{Let } v\mapsto v-x-\frac{d}{c}\right)\\
&= e\left(\frac{am}{c}\right)\sum_{\widetilde{n}\in\mathbb{Z}}e\left(\widetilde{n}x+\frac{\widetilde{n}d}{c}\right)\int_{\mathbb{R}}(c(v+\mathrm{i}y))^{-k}e\left(-\frac{m}{c^2(v+\mathrm{i}y)}-\widetilde{n}v\right)\mathrm{d}v\\
&= e\left(\frac{am}{c}\right)\sum_{\widetilde{n}\in\mathbb{Z}}e\left(\widetilde{n}x+\frac{\widetilde{n}d}{c}\right)e\left(\widetilde{n}\mathrm{i}y\right)\int_{\mathbb{R}}(c(v+\mathrm{i}y))^{-k}e\left(-\frac{m}{c^2(v+\mathrm{i}y)}-\widetilde{n}(v+\mathrm{i}y)\right)\mathrm{d}v\\
&= \sum_{\widetilde{n}\in\mathbb{Z}}e\left(\tilde{n}(x+\mathrm{i}y)\right)e\left(\frac{am+\widetilde{n}d}{c}\right)\int_{\mathbb{R}}(c(v+\mathrm{i}y))^{-k}e\left(-\frac{m}{c^2(v+\mathrm{i}y)}-\widetilde{n}(v+\mathrm{i}y)\right)\mathrm{d}v\\
&= \sum_{\widetilde{n}\in\mathbb{Z}}e\left(\tilde{n}z\right)e\left(\frac{am+\widetilde{n}d}{c}\right)\int_{\mathbb{R}}(c(v+\mathrm{i}y))^{-k}e\left(-\frac{m}{c^2(v+\mathrm{i}y)}-\widetilde{n}(v+\mathrm{i}y)\right)\mathrm{d}v\\
&= \sum_{\widetilde{n}\in\mathbb{Z}}e\left(\tilde{n}z\right)e\left(\frac{am+\widetilde{n}d}{c}\right)\int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y}(cv)^{-k}e\left(-\frac{m}{c^2v}-\widetilde{n}v\right)\mathrm{d}v.
\end{aligned}$$

Denote

$$\mathcal{J}_c(m,n)=\int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y}(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\mathrm{d}v.$$

Since $k\in\mathbb{Z}$, the function

$$(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)$$

is holomorphic on the upper half-plane $\mathbb{H}$. Hence for any $y_1,y_2>0$, Cauchy's Theorem applied to the rectangle with vertices $\pm R+\mathrm{i}y_1$, $\pm R+\mathrm{i}y_2$ gives

$$\int_{-R+\mathrm{i}y_1}^{R+\mathrm{i}y_1}(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\mathrm{d}v-\int_{-R+\mathrm{i}y_2}^{R+\mathrm{i}y_2}(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\mathrm{d}v=-\int_{\text{vertical sides}}(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\mathrm{d}v.$$

Since

$$\left|(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\right|\ll|v|^{-k}$$

uniformly on the vertical sides, the right-hand side tends to 0 as $R\to\infty$. Therefore

$$\int_{-R+\mathrm{i}y_1}^{R+\mathrm{i}y_1}(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\mathrm{d}v=\int_{-R+\mathrm{i}y_2}^{R+\mathrm{i}y_2}(cv)^{-k}e\left(-\frac{m}{c^2v}-nv\right)\mathrm{d}v.$$

That is, $\mathcal{J}_c(m,n)$ does not depend on the choice of y.

We claim that if $n \le 0$, $\mathcal{J}_c(m,n) = 0$. Indeed, if $n = 0$, we have

$$\begin{aligned}
\left|\int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y} (cv)^{-k} e\left(-\frac{m}{c^2 v} - nv\right) \mathrm{d}v\right| &= \left|\lim_{y\to\infty} \int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y} (cv)^{-k} e\left(-\frac{m}{c^2 v} - nv\right) \mathrm{d}v\right| \\
&= \left|\lim_{y\to\infty} \int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y} (cv)^{-k} e\left(-\frac{m}{c^2 v}\right) \mathrm{d}v\right| \\
&\le \lim_{y\to\infty} \int_{-\infty}^{\infty} |c(x+\mathrm{i}y)|^{-k} \left|e\left(-\frac{m}{c^2(x+\mathrm{i}y)}\right)\right| \mathrm{d}x \quad (\text{Let } v = x+\mathrm{i}y) \\
&= \lim_{y\to\infty} \int_{-\infty}^{\infty} \frac{c^{-k}}{(x^2+y^2)^{k/2}} \exp\left(-\frac{2\pi m y}{c^2(x^2+y^2)}\right) \mathrm{d}x \\
&\le \lim_{y\to\infty} \int_{-\infty}^{\infty} c^{-k} \frac{1}{(x^2+y^2)^{k/2}} \mathrm{d}x \\
&= \lim_{y\to\infty} c^{-k} y^{1-k} \int_{-\infty}^{\infty} \frac{1}{(1+t^2)^{k/2}} \mathrm{d}t \qquad (\text{Let } x = yt) \\
&= 0
\end{aligned}$$

since $k > 2$. For case $n < 0$, similarly

$$\begin{aligned}
\left|\int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y} (cv)^{-k} e\left(-\frac{m}{c^2 v} - nv\right) \mathrm{d}v\right| &\le \lim_{y\to\infty} \int_{-\infty}^{\infty} |c(x+\mathrm{i}y)|^{-k} \left|e\left(-\frac{m}{c^2(x+\mathrm{i}y)}\right)\right| |e(-n(x+\mathrm{i}y))| \mathrm{d}x \\
&= \lim_{y\to\infty} e^{2\pi n y} c^{-k} \int_{-\infty}^{\infty} \frac{1}{(x^2+y^2)^{k/2}} \mathrm{d}x,
\end{aligned}$$

which clearly equals to 0. Hence $\mathcal{J}_c(m,n) = 0$ for all $n \le 0$.

For $n > 0$, if $m > 0$, we have

$$\begin{aligned}
\mathcal{J}_c(m,n) &= \int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y} (cv)^{-k} e\left(-\frac{m}{c^2 v} - nv\right) \mathrm{d}v \\
&= \int_{-\infty+\mathrm{i}y}^{+\infty+\mathrm{i}y} (cv)^{-k} \exp\left(-2\pi\mathrm{i}\frac{m}{c^2 v} - 2\pi\mathrm{i}nv\right) \mathrm{d}v \\
&= (-1)^k c^{-k} (2\pi\mathrm{i}n)^k \left(-\frac{1}{2\pi\mathrm{i}n}\right) \int_{\mathrm{Re}(w)=2\pi n y} w^{-k} \exp\left(w - \frac{4\pi^2 mn}{c^2 w}\right) \mathrm{d}w \quad (\text{Let } w = -2\pi\mathrm{i}nv) \\
&= (-1)^k c^{-k} (2\pi\mathrm{i}n)^{k-1} \int_{2\pi n y - \mathrm{i}\infty}^{2\pi n y + \mathrm{i}\infty} w^{-k} \exp\left(w - \frac{4\pi^2 mn}{c^2 w}\right) \mathrm{d}w \\
&= (-1)^k c^{-k} (2\pi\mathrm{i}n)^{k-1} \int_{2\pi n y - \mathrm{i}\infty}^{2\pi n y + \mathrm{i}\infty} w^{-k} \exp\left(w - \frac{x^2}{4w}\right) \mathrm{d}w \qquad \left(\text{Let } \frac{x^2}{4} = \frac{4\pi^2 mn}{c^2}\right) \\
&= (-1)^k c^{-k} (2\pi\mathrm{i}n)^{k-1} \cdot 2\pi\mathrm{i} \left(\frac{2}{x}\right)^{k-1} J_{k-1}(x) \\
&= \frac{2\pi}{\mathrm{i}^k c} \left(\frac{n}{m}\right)^{(k-1)/2} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right),
\end{aligned}$$

where $J_{k-1}(x)$ is the Bessel J-function. If $m = 0$, let $x \to 0$ in the above equation we obtain

$$\mathcal{J}_c(0,n) = \frac{(-2\pi\mathrm{i})^k}{\Gamma(k) c^k} n^{k-1} \quad (n > 0).$$

Consequently,

$$\begin{aligned}\mathbf{P}(z) &= e(mz) + \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \sum_{\substack{d \pmod c\\ \gcd(d,c)=1}} \overline{\chi(d)} \sum_{n>0} e(nz) e\left(\frac{am+nd}{c}\right) \mathcal{J}_c(m,n) \\ &= e(mz) + \sum_{n\geq 1} e(nz) \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \sum_{\substack{d \pmod c\\ \gcd(d,c)=1}} \overline{\chi(d)} e\left(\frac{m\overline{d}+nd}{c}\right) \mathcal{J}_c(m,n) \\ &= e(mz) + \sum_{n\geq 1} e(nz) \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} S_\chi(m,n;c)\mathcal{J}_c(m,n).\end{aligned}$$

Indeed, $\gamma \in \Gamma_0(q)$ implies that $ad - bc = 1$, so $ad \equiv 1 \pmod c$, and so $a \equiv \overline{d} \pmod c$, where $\overline{d}$ is the inverse of d modulo c. Furthermore,

$$\mathbf{P}(z) = e(mz) + 2\pi \mathrm{i}^{-k} \sum_{n\geq 1} \left(\frac{n}{m}\right)^{(k-1)/2} \left(\sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \frac{S_\chi(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right) e(nz)$$

for $m > 0$. We know $\mathbf{P}(z)$ is holomorphic at $\mathfrak{a} = \infty$.

Example 3.1. Let $p(z) = e(mz)$. If $m = 0$,

$$E(z) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \overline{\chi(\gamma)} j_\gamma(z)^{-k}$$

is called the Eisenstein series of weight k. We have

$$\begin{aligned}E(z) &= 1 + \sum_{n\geq 1} e(nz) \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} S_\chi(0,n;c) \left(\frac{2\pi}{\mathrm{i}c}\right)^k \frac{n^{k-1}}{\Gamma(k)} \\ &= 1 + \sum_{n\geq 1} \left(\frac{2\pi}{\mathrm{i}}\right)^k \frac{n^{k-1}}{\Gamma(k)} \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \frac{S_\chi(0,n;c)}{c^k} e(nz) \\ &= 1 + \sum_{n\geq 1} \eta(n) e(nz),\end{aligned}$$

here

$$\eta(n) = \left(\frac{2\pi}{\mathrm{i}}\right)^k \frac{n^{k-1}}{\Gamma(k)} \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \frac{S_\chi(0,n;c)}{c^k}.$$

If $m > 0$, we have

$$\begin{aligned}\mathbf{P}_m(z) &= \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \overline{\chi(\gamma)} j_\gamma(z)^{-k} e(m\gamma z) \\ &= e(mz) + \sum_{n\geq 1} e(nz) \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} S_\chi(m,n;c) \frac{2\pi}{\mathrm{i}^k c} \left(\frac{n}{m}\right)^{(k-1)/2} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right) \\ &= \sum_{n\geq 1} \left(\frac{n}{m}\right)^{(k-1)/2} e(nz) \left(\delta_{mn} + 2\pi \mathrm{i}^{-k} \sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \frac{S_\chi(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right).\end{aligned}$$

Note that if $m \geq 1$, the Fourier expansion of $\mathbf{P}_m(z)$ has no constant term. Hence:

Proposition 3.2. For $m \geq 1$, the Poincaré series is a cusp form.

3.3 The Hilbert space of cusp forms

Let $f, g \in \mathcal{M}_k(\Gamma_0(q), \chi)$ be automorphic forms, then

$$f_{|\gamma}(z) = j_\gamma(z)^{-k} f(\gamma z) = \chi(\gamma) f(z),$$
$$g_{|\gamma}(z) = j_\gamma(z)^{-k} g(\gamma z) = \chi(\gamma) g(z).$$

We show that $\mathrm{Im}(z)^k f(z)\overline{g(z)}$ is Γ-invariant. Indeed, given arbitrary

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q),$$

we have

$$\begin{aligned}
\mathrm{Im}(\gamma z)^k f(\gamma z)\overline{g(\gamma z)} &= \left(\frac{\mathrm{Im}(z)}{|cz+d|^2}\right)^k \chi(d)(cz+d)^k f(z)\overline{\chi(d)(cz+d)^k g(z)} \\
&= \left(\frac{\mathrm{Im}(z)}{|cz+d|^2}\right)^k \chi(d)(cz+d)^k f(z)\overline{\chi(d)}(\overline{cz+d})^k \overline{g(z)} \\
&= \mathrm{Im}(z)^k \frac{\chi(d)\overline{\chi(d)}(cz+d)^k(\overline{cz+d})^k}{|cz+d|^{2k}} f(z)\overline{g(z)} \\
&= \mathrm{Im}(z)^k f(z)\overline{g(z)}.
\end{aligned}$$

Define

$$\langle f, g\rangle = \int_{\Gamma\backslash\mathbb{H}} \mathrm{Im}(z)^k f(z)\overline{g(z)}\mathrm{d}\mu(z), \quad \mathrm{d}\mu(z) = \frac{\mathrm{d}x\mathrm{d}y}{y^2}.$$

Since $\mathrm{d}\mu(z)$ is Γ-invariant, the integral is well-defined. If f is a cusp form, it has exponential decay at cusps. Moreover, after removing the neighborhoods of the cusps, the fundamental domain reduces to a compact set $K \subset \mathbb{H}$, and the integral is finite on K. Hence

$$\|f\|^2 = \langle f, f\rangle = \int_{\Gamma\backslash\mathbb{H}} \mathrm{Im}(z)^k |f(z)|^2 \mathrm{d}\mu(z) < \infty.$$

Now,

$$\langle f, g\rangle = \int_{\gamma\backslash\mathbb{H}} \mathrm{Im}(z)^k f(z)\overline{g(z)}\mathrm{d}\mu(z) = \overline{\int_{\gamma\backslash\mathbb{H}} \mathrm{Im}(z)^k g(z)\overline{f(z)}\mathrm{d}\mu(z)} = \overline{\langle g, f\rangle},$$

so $\langle\ ,\ \rangle$ satisfies conjugate symmetry, and the linearity is clear. It follows that the Petersson inner product is well-defined. Therefore, $\mathcal{S}_k(\Gamma_0(q), \chi)$ equipped with $\langle\ ,\ \rangle$ is a finite dimension Hilbert space since $\dim_{\mathbb{C}} \mathcal{S}_k(\Gamma_0(q), \chi) < \infty$.

Let's consider the projection of $f \in \mathcal{M}_k(\Gamma_0(q), \chi)$ on $\mathbf{P}_m$.

$$\begin{aligned}
\langle f, \mathbf{P}_m \rangle &= \int_{\Gamma\backslash\mathbb{H}} \mathrm{Im}(z)^k f(z) \overline{\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \overline{\chi(\gamma)} j_\gamma(z)^{-k} e(m\gamma z)} \mathrm{d}\mu(z) \\
&= \int_{\Gamma\backslash\mathbb{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(z)^k f(z) \chi(\gamma) \overline{j_\gamma(z)}^{-k} \overline{e(m\gamma z)} \mathrm{d}\mu(z) \\
&= \int_{\Gamma\backslash\mathbb{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^k \left|j_\gamma(z)\right|^{2k} \overline{j_\gamma(z)}^{-k} \chi(\gamma) f(z) \overline{e(m\gamma z)} \mathrm{d}\mu(z) \\
&= \int_{\Gamma\backslash\mathbb{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^k j_\gamma(z)^k \chi(\gamma) f(z) \overline{e(m\gamma z)} \mathrm{d}\mu(z) \\
&= \int_{\Gamma\backslash\mathbb{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^k f(\gamma z) \overline{e(m\gamma z)} \mathrm{d}\mu(z) \\
&= \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\Gamma\backslash\mathbb{H}} \mathrm{Im}(\gamma z)^k f(\gamma z) \overline{e(m\gamma z)} \mathrm{d}\mu(z) \\
&= \int_{\Gamma_\infty\backslash\mathbb{H}} \mathrm{Im}(z)^k f(z) \overline{e(mz)} \mathrm{d}\mu(z) && \text{(Unfold)} \\
&= \int_0^\infty \int_0^1 y^k f(z) e\left(-m\overline{z}\right) \frac{\mathrm{d}x\mathrm{d}y}{y^2} \\
&= \int_0^\infty y^k \int_0^1 \sum_{n\geq 0} \hat{f}(n) e(nz) e\left(-m\overline{z}\right) \frac{\mathrm{d}x\mathrm{d}y}{y^2} \\
&= \int_0^\infty \sum_{n\geq 0} \hat{f}(n) \mathrm{e}^{-2\pi(n+m)y} \left(\int_0^1 \mathrm{e}^{2\pi\mathrm{i}(n-m)x} \mathrm{d}x \right) \frac{\mathrm{d}y}{y^2} \\
&= \int_0^\infty y^k \hat{f}(m) \mathrm{e}^{-4\pi m y} \frac{\mathrm{d}y}{y^2} \\
&= \hat{f}(m) \int_0^\infty y^{k-2} \mathrm{e}^{-4\pi m y} \mathrm{d}y \\
&= \hat{f}(m) \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}. && \left(\text{Recall } \Gamma(s) = \int_0^\infty \mathrm{e}^{-y} y^s \frac{\mathrm{d}y}{y}\right)
\end{aligned}$$

Then we obtain

Theorem 3.1.

$$\langle f, \mathbf{P}_m \rangle = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \hat{f}(m).$$

Corollary 3.1. The space of cusp forms $\mathcal{S}_k(\Gamma_0(q), \chi)$ is spanned by $\{\mathbf{P}_m\}_{m\geq 1}$.

Proof. Let $M = \mathrm{span}\{\mathbf{P}_m(z)\}_{m\geq 1} \subset \mathcal{S}_k(\Gamma_0(q), \chi)$. If $f \perp M$, we have $\langle f, \mathbf{P}_m \rangle = 0$, so $\hat{f}(m) = 0$ for all $m \geq 1$. Thus $f = \sum_{m\geq 1} \hat{f}(m) e(mz) = 0$. That is, $M^\perp = \{0\}$. Since $\mathcal{S}_k(\Gamma_0(q), \chi) = M \oplus M^\perp$, we obtain $M = \mathcal{S}_k(\Gamma_0(q), \chi)$. □

Let $\mathcal{F}$ be an orthonormal basis of $\mathcal{S}_k(\Gamma_0(q),\chi)$. Since for $m \geq 1$ we have $\mathbf{P}_m \in \mathcal{S}_k(\Gamma_0(q),\chi)$, we have

$$\begin{aligned}\mathbf{P}_m &= \sum_{f\in\mathcal{F}} \langle \mathbf{P}_m, f\rangle\, f \\ &= \sum_{f\in\mathcal{F}} \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}\overline{\hat{f}(m)} f(z) \\ &= \sum_{n\geq 1}\left(\sum_{f\in\mathcal{F}} \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}\overline{\hat{f}(m)}\hat{f}(n)\right) e(nz).\end{aligned}$$

On the other hand, recall that

$$\mathbf{P}_m(z) = \sum_{n\geq 1}\left(\frac{n}{m}\right)^{(k-1)/2} e(nz)\left(\delta_{mn} + 2\pi \mathrm{i}^{-k}\sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \frac{S_\chi(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right).$$

Hence

Theorem 3.2 (Petersson Trace Formula)**.**

$$\left(\frac{n}{m}\right)^{(k-1)/2}\left(\delta_{mn} + 2\pi \mathrm{i}^{-k}\sum_{\substack{c>0\\ c\equiv 0 \pmod q}} \frac{S_\chi(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right) = \sum_{f\in\mathcal{F}} \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}\overline{\hat{f}(m)}\hat{f}(n).$$

Example 3.2. Consider $\Gamma = \Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$, and χ be trivial character. Recall that

$$\begin{aligned}\Delta(z) &= (2\pi)^{12} e(z)\prod_{n\geq 1}(1-e(nz))^{24} \\ &= (2\pi)^{12}\sum_{n\geq 1}\tau(n)e(nz),\end{aligned}$$

where $\tau(n)$ is called Ramanujan τ-function. Since $\mathbf{P}_m(z) \in \mathcal{S}_{12}$, so we can assume $\mathbf{P}_m(z) = c(m)\Delta(z)$. Recall that

$$\langle f, \mathbf{P}_m\rangle = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}}\hat{f}(m),$$

then

$$c(m)\|\Delta\|^2 = \langle \Delta, c(m)\Delta\rangle = \langle \Delta, \mathbf{P}_m\rangle = \frac{\Gamma(11)}{(4\pi m)^{11}}(2\pi)^{12}\tau(m),$$

so

$$c(m) = \frac{2\pi\Gamma(11)\tau(m)}{(2\pi)^{11}}\frac{1}{\|\Delta\|^2}.$$

Then

$$\mathbf{P}_m(z) = (2\pi)\Gamma(11)\frac{\tau(m)}{(2m)^{11}}\frac{\Delta(z)}{\|\Delta\|^2}.$$

For

$$f(z) = \frac{\Delta(z)}{\|\Delta\|} = \sum_{n\geq 1} \frac{(2\pi)^{12}\tau(n)}{\|\Delta\|} e(nz).$$

Applying Petersson Trace Formula (Theorem 3.2),

$$\frac{\Gamma(11)}{(4\pi\sqrt{mn})^{11}} \frac{(2\pi)^{12}\tau(m)}{\|\Delta\|} \frac{(2\pi)^{12}\tau(n)}{\|\Delta\|} = \delta_{mn} + 2\pi \sum_{c>0} \frac{S(m,n;c)}{c} J_{11}\left(\frac{4\pi\sqrt{mn}}{c}\right).$$

Therefore

$$\tau(m)\tau(n) = \frac{\|\Delta\|^2}{4\pi^{13}\Gamma(11)} (mn)^{11/2} \left(\delta_{mn} + 2\pi \sum_{c>0} \frac{S(m,n;c)}{c} J_{11}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right).$$

Let

$$\nu = \frac{\|\Delta\|^2}{4\pi^{13}\Gamma(11)},$$

we have

$$\tau(m)\tau(n) = \nu (mn)^{11/2} \left(\delta_{mn} + 2\pi \sum_{c>0} \frac{S(m,n;c)}{c} J_{11}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right).$$

In particular for $n > 1$ we have

$$\tau(n) = 2\pi\nu n^{11/2} \sum_{c>0} c^{-1} S(1,n;c) J_{11}\left(\frac{4\pi\sqrt{n}}{c}\right).$$

Theorem 3.3 (Ramanujan–Petersson bound for $\tau(n)$, proved by Deligne)**.**

$$|\tau(n)| \leq n^{11/2} d(n).$$

4 Kloosterman Sums

4.1 The classical Kloosterman sums

The **classical Kloosterman sum** is defined as

$$S(m,n;c) = \sum_{\substack{d \pmod c \\ \gcd(d,c)=1}} e\left(\frac{m\overline{d} + nd}{c}\right) = \sum_{d \pmod c}^{*} e\left(\frac{m\overline{d} + nd}{c}\right),$$

where $d\overline{d} \equiv 1 \pmod c$. Note that

$$S(m,n;c) = \sum_{d \pmod c}^{*} e\left(\frac{m\overline{d} + nd}{c}\right) = \sum_{d \pmod c}^{*} e\left(\frac{md + n\overline{d}}{c}\right) = S(n,m;c).$$

Moreover

$$\begin{aligned}
\overline{S(m,n;c)} &= \sum_{d \pmod c}^{*} e\left(-\frac{m\overline{d}+nd}{c}\right) \\
&= \sum_{d \pmod c}^{*} e\left(\frac{m(-\overline{d})+n(-d)}{c}\right) \\
&= \sum_{d \pmod c}^{*} e\left(\frac{m\overline{(-d)}+n(-d)}{c}\right) \quad (\text{since } \gcd(-d,c)=1) \\
&= \sum_{d \pmod c}^{*} e\left(\frac{m\overline{d}+nd}{c}\right) \qquad (-d \mapsto d) \\
&= S(m,n;c).
\end{aligned}$$

That is, $S(m,n;c) \in \mathbb{R}$. If $\gcd(a,c)=1$ we also have

$$\begin{aligned}
S(m,an;c) &= \sum_{d \pmod c}^{*} e\left(\frac{m\overline{d}+and}{c}\right) \\
&= \sum_{d \pmod c}^{*} e\left(\frac{ma\overline{d}+an\overline{a}d}{c}\right) \quad (d \mapsto ad, \text{ which is bijective since } a \text{ is invertible}) \\
&= \sum_{d \pmod c}^{*} e\left(\frac{am\overline{d}+nd}{c}\right) \\
&= S(am,n;c).
\end{aligned}$$

We have the **Selberg identity**

$$S(m,n;c) = \sum_{d \mid \gcd(c,m,n)} dS(mnd^{-2},1,cd^{-1}).$$

Lemma 4.1 (orthogonality of additive characters)**.** Let q be a positive integer. For each $n \in \mathbb{Z}$, we have

$$\frac{1}{q}\sum_{a \pmod q} e\left(\frac{na}{q}\right) = \begin{cases} 1 & \text{if } q \mid n, \\ 0 & \text{if } q \nmid n. \end{cases}$$

Proof. If $q \mid n$, i.e., $n = qk$, then for every a we have

$$e\left(\frac{na}{q}\right) = e(ka) = 1.$$

It follows that

$$\frac{1}{q}\sum_{a \pmod q} e\left(\frac{na}{q}\right) = \frac{1}{q}\sum_{a \pmod q} 1 = \frac{1}{q}\cdot q = 1.$$

If $q \nmid n$, let $\zeta = e(n/q)$. Since $q \nmid n$, we have $\zeta \neq 1$ but $\zeta^q = 1$. So

$$\frac{1}{q} \sum_{a \pmod q} e\left(\frac{na}{q}\right) = \frac{1}{q} \sum_{a=0}^{q-1} \zeta^a = \frac{1}{q} \cdot \frac{1-\zeta^q}{1-\zeta} = 0.$$

□

Proof of the Selberg identity. From Lemma 4.1 it follows that

$$\begin{aligned}
S(m,n;c) &= S(n,m;c) \\
&= \sideset{}{^*}\sum_{x \pmod c} e\left(\frac{mx+n\overline{x}}{c}\right) \\
&= \sum_{x,y \pmod c} e\left(\frac{xm+ny}{c}\right) \cdot \mathbf{1}_{xy\equiv 1 \pmod c} \\
&= \sum_{x,y \pmod c} e\left(\frac{xm+ny}{c}\right) \frac{1}{c} \sum_{a \pmod c} e\left(\frac{a(xy-1)}{c}\right) \\
&= \frac{1}{c} \sum_{a \pmod c} \sum_{x,y \pmod c} e\left(\frac{x(m+ay)+ny-a}{c}\right) \\
&= \frac{1}{c} \sum_{d \mid c} \sum_{\substack{a \pmod{c/d} \\ \gcd(a,c/d)=1}} \sum_{x,y \pmod c} e\left(\frac{x(m+ady)+ny-ad}{c}\right).
\end{aligned}$$

We evaluate the x-sum by orthogonality, getting

$$S(m,n;c) = \sum_{d \mid c} \sum_{\substack{a \pmod{c/d} \\ \gcd(a,c/d)=1}} \sum_{\substack{y \pmod c \\ ady \equiv -m \pmod c}} e\left(\frac{ny-ad}{c}\right),$$

since

$$\sum_{x \pmod c} e\left(\frac{x(m+ady)}{c}\right) = \begin{cases} c & \text{if } m+ady \equiv 0 \pmod c, \\ 0 & \text{else,} \end{cases}$$

from above Lemma.

Since $ady \equiv -m \pmod c$, necessarily $d \mid m$, so we can set $d \mid \gcd(m,c)$. Then $ady \equiv -m \pmod c$ implies that $ady \equiv -d(m/d) \pmod{d \cdot (c/d)}$, so $ay \equiv -(m/d) \pmod{c/d}$. Since $\gcd(a, c/d) = 1$, we have $y \equiv -\overline{a}(m/d) \pmod{c/d}$. All the solutions of $y \equiv -\overline{a}(m/d) \pmod{c/d}$ have the form $y = -\overline{a}(m/d) + t(c/d)$ where $t \pmod d$. Thus

$$\begin{aligned}
\sum_{\substack{y \pmod c \\ ady \equiv -m \pmod c}} e\left(\frac{ny-ad}{c}\right) &= \sum_{t \pmod d} e\left(\frac{n\left(-\overline{a}(m/d)+t(c/d)\right)-da}{d(c/d)}\right) \\
&= e\left(\frac{-n\overline{a}(m/d)-da}{c}\right) \sum_{t \pmod d} e\left(\frac{nt}{d}\right).
\end{aligned}$$

Note that

$$\sum_{t \pmod d} e\left(\frac{nt}{d}\right) = \begin{cases} d & \text{if } d \mid n, \\ 0 & \text{if } d \nmid n, \end{cases}$$

we can also set $d \mid n$. That is, $d \mid \gcd(m, n, c)$.

We thus have

$$\begin{aligned} S(m,n;c) &= \sum_{d \mid \gcd(m,n,c)} \sum_{\substack{a \pmod{c/d} \\ \gcd(a,c/d)=1}} \sum_{\substack{y \pmod c \\ y \equiv -\overline{a}m/d \pmod{c/d}}} e\left(\frac{ny - ad}{c}\right) \\ &= \sum_{d \mid \gcd(m,n,c)} d \sum_{\substack{a \pmod{c/d} \\ \gcd(a,c/d)=1}} e\left(-\frac{(m/d)(n/d)\overline{a} + a}{c/d}\right) \\ &= \sum_{d \mid \gcd(m,n,c)} d \sum_{\substack{a \pmod{c/d} \\ \gcd(a,c/d)=1}} e\left(\frac{(m/d)(n/d)\overline{a} + a}{c/d}\right) \qquad (a \mapsto -a) \\ &= \sum_{d \mid \gcd(m,n,c)} dS(mnd^{-2}, 1; cd^{-1}). \end{aligned}$$

This establishes the Selberg identity for Kloosterman sums. □

In particular, if $\gcd(c, m, n) = 1$, then

$$S(m,n;c) = S(mn, 1, c).$$

If $c = qr$ and $\gcd(q, r) = 1$, then by the Chinese Remainder Theorem, we have twisted multiplicativity

$$S(m,n,qr) = S\left(\overline{q}m, \overline{q}n; r\right) S\left(\overline{r}m, \overline{r}n; q\right).$$

Indeed, since $c = qr$ and $\gcd(q, r) = 1$, we have $\mathbb{Z}/(qr)\mathbb{Z} \cong \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/r\mathbb{Z}$, and $(\mathbb{Z}/(qr)\mathbb{Z})^\times \cong (\mathbb{Z}/q\mathbb{Z})^\times \times (\mathbb{Z}/r\mathbb{Z})^\times$. Consequently, each invertible residue class d modulo qr uniquely corresponds to a pair of (d_q, d_r), where $d_q \in (\mathbb{Z}/q\mathbb{Z})^\times$, $d_r \in (\mathbb{Z}/r\mathbb{Z})^\times$, and $d \equiv d_q \pmod q$, $d \equiv d_r \pmod r$. Moreover $\overline{d} \equiv \overline{d_q} \pmod q$, $\overline{d} \equiv \overline{d_r} \pmod r$. Note that for arbitrary integral x, we have $x \equiv x(r\overline{r}) + x(q\overline{q}) \pmod{qr}$ for $q\overline{q} \equiv 1 \pmod r$ and $r\overline{r} \equiv 1 \pmod q$, then $x/(qr) \equiv (\overline{q}x)/r + (\overline{r}x)/q$

(mod 1). It follows that

$$\begin{aligned}
S(m,n,qr) &= \sideset{}{^*}\sum_{d \pmod{qr}} e\left(\frac{m\overline{d}+nd}{qr}\right) \\
&= \sideset{}{^*}\sum_{d \pmod{qr}} e\left(\frac{\overline{q}(m\overline{d}+nd)}{r}\right) e\left(\frac{\overline{r}(m\overline{d}+nd)}{q}\right) \\
&= \sideset{}{^*}\sum_{d \pmod{qr}} e\left(\frac{\overline{q}m\overline{d_r}+\overline{q}nd_r}{r}\right) e\left(\frac{\overline{r}m\overline{d_q}+\overline{r}nd_q}{q}\right) \\
&= \sideset{}{^*}\sum_{d_q \pmod{q}} \sideset{}{^*}\sum_{d_r \pmod{r}} e\left(\frac{\overline{q}m\overline{d_r}+\overline{q}nd_r}{r}\right) e\left(\frac{\overline{r}m\overline{d_q}+\overline{r}nd_q}{q}\right) \\
&= \left(\sideset{}{^*}\sum_{d_q \pmod{q}} e\left(\frac{\overline{r}m\overline{d_q}+\overline{r}nd_q}{r}\right)\right)\left(\sideset{}{^*}\sum_{d_r \pmod{r}} e\left(\frac{\overline{q}m\overline{d_r}+\overline{q}nd_r}{q}\right)\right) \\
&= S\left(\overline{q}m,\overline{q}n;r\right) S\left(\overline{r}m,\overline{r}n;q\right).
\end{aligned}$$

This allows one to reduce the problem of evaluating $S(m,n;c)$ for any c to that for prime power moduli.

Lemma 4.2. If $c = p^{2\alpha}$ with $\alpha \geq 1$ and $\gcd(c,2n) = 1$, then

$$S(n,n;c) = c^{1/2}\left(e\left(\frac{2n}{c}\right) + e\left(\frac{-2n}{c}\right)\right).$$

Proof. From definition

$$S(n,n;p^{2\alpha}) = \sideset{}{^*}\sum_{d \pmod{p^{2\alpha}}} e\left(\frac{nd+n\overline{d}}{p^{2\alpha}}\right).$$

Given $a \pmod{p^{2\alpha}}$ and $b \pmod{p^{\alpha}}$ and consider $a(1+bp^{\alpha})$. Since $1+bp^{\alpha}$ admits an inverse modulo $p^{2\alpha}$, the equation $a(1+bp^{\alpha}) \equiv d \pmod{p^{2\alpha}}$ has a unique solution $a \equiv d(1+bp^{\alpha})^{-1} \pmod{p^{2\alpha}}$ for fixed $d \pmod{p^{2\alpha}}$ and $b \pmod{p^{\alpha}}$. Moreover $\gcd(d,p^{2\alpha}) = \gcd(d,p) = 1$ implies that $\gcd(a,p) = 1$. Consequently, there exists p^{α} many solutions $a \pmod{p^{2\alpha}}$ and $b \pmod{p^{\alpha}}$ with $d \equiv a(1+bp^{\alpha}) \pmod{p^{2\alpha}}$ for every $d \pmod{p^{2\alpha}}$. Hence

$$\begin{aligned}
S(n,n;p^{2\alpha}) &= p^{-\alpha} \sum_{b \pmod{p^{\alpha}}} \sum_{\substack{a \pmod{p^{2\alpha}} \\ \gcd(a,p)=1}} e\left(\frac{na(1+bp^{\alpha})+n\overline{a}(1-bp^{\alpha})}{p^{2\alpha}}\right) \\
&= p^{-\alpha} \sideset{}{^*}\sum_{a \pmod{p^{2\alpha}}} e\left(\frac{n\left(a+\overline{a}\right)}{p^{2\alpha}}\right) \sum_{b \pmod{p^{\alpha}}} e\left(\frac{n\left(a-\overline{a}\right)b}{p^{\alpha}}\right) \\
&= \sideset{}{^*}\sum_{\substack{a \pmod{p^{2\alpha}} \\ a \equiv \overline{a} \pmod{p^{\alpha}}}} e\left(\frac{n\left(a+\overline{a}\right)}{p^{2\alpha}}\right). \qquad \text{(Orthogonality of additive characters)}
\end{aligned}$$

Note that $a \equiv \overline{a} \pmod{p^{\alpha}}$ iff $a^2 \equiv 1 \pmod{p^{\alpha}}$ iff $(a-1)(a+1) \equiv 0 \pmod{p^{\alpha}}$. That is, $p^{\alpha} \mid (a-1)(a+1)$. However $p \nmid \gcd(a-1,a+1)$, otherwise $p = 2$, contrary to $\gcd(p^{2\alpha},2n) = 1$. So

$p^\alpha \mid a-1$ or $p^\alpha \mid a+1$. If $p^\alpha \mid a-1$, then $a = 1+tp^\alpha \pmod{p^\alpha}$, it follows that $\overline{a} = 1-tp^\alpha \pmod{p^{2\alpha}}$; if $p^\alpha \mid a+1$, then $a = -1+tp^\alpha \pmod{p^\alpha}$, it follows that $\overline{a} = -1-tp^\alpha \pmod{p^{2\alpha}}$. Hence

$$\begin{aligned} S(n,n;p^{2\alpha}) &= \sum_{t \pmod{p^\alpha}} e\left(\frac{n(1+tp^\alpha+1-tp^\alpha)}{p^{2\alpha}}\right) + \sum_{t \pmod{p^\alpha}} e\left(\frac{n(-1+tp^\alpha-1-tp^\alpha)}{p^{2\alpha}}\right) \\ &= p^\alpha\left(e\left(\frac{2n}{p^{2\alpha}}\right)+e\left(\frac{-2n}{p^{2\alpha}}\right)\right) \\ &= \sqrt{c}\left(e\left(\frac{2n}{c}\right)+e\left(\frac{-2n}{c}\right)\right). \end{aligned}$$

□

Lemma 4.3. If $c = p^{2\alpha+1}$ with $\alpha \geq 1$ and $\gcd(c, 2n) = 1$, then

$$S(n,n;c) = 2\left(\frac{n}{c}\right)c^{1/2}\operatorname{Re}\left(\varepsilon_c e\left(\frac{2n}{c}\right)\right),$$

where

$$\left(\frac{n}{c}\right)$$

is the Legendre–Jacobi symbol and $\varepsilon_c = 1, i$ according to whether $c \equiv 1$ or $-1 \pmod 4$.

Proof. From definition we have

$$S(n,n;c) = \sideset{}{^*}\sum_{d \pmod{p^{2\alpha+1}}} e\left(\frac{n\overline{d}+nd}{p^{2\alpha+1}}\right)$$

For arbitrary $b \pmod{p^\alpha}$, $1+bp^{\alpha+1}$ is invertible modulo $p^{2\alpha+1}$. Thus for every $d \pmod{p^{2\alpha+1}}$ there exists a unique $a \pmod{p^{2\alpha+1}}$ with $d \equiv a(1+bp^{\alpha+1}) \pmod{p^{2\alpha+1}}$. Moreover $\gcd(a,p) = 1$ since $\gcd(d, p^{2\alpha+1}) = \gcd(d,p) = 1$. Thus the exists p^α many $a \pmod{p^{2\alpha+1}}$ and $b \pmod{p^\alpha}$ such that $d \equiv a(1+bp^{\alpha+1}) \pmod{p^{2\alpha+1}}$ for every d. Then

$$\begin{aligned} S(n,n;c) &= p^{-\alpha}\sum_{b \pmod{p^\alpha}}\ \sideset{}{^*}\sum_{a \pmod{p^{2\alpha+1}}} e\left(\frac{na(1+bp^{\alpha+1})+n\overline{a}(1-bp^{\alpha+1})}{p^{2\alpha+1}}\right) \\ &= p^{-\alpha}\sideset{}{^*}\sum_{a \pmod{p^{2\alpha+1}}} e\left(\frac{n\,(a+\overline{a})}{p^{2\alpha+1}}\right)\sum_{b \pmod{p^\alpha}} e\left(\frac{nb\,(a-\overline{a})}{p^\alpha}\right) \\ &= \sideset{}{^*}\sum_{\substack{a \pmod{p^{2\alpha+1}} \\ a \equiv \overline{a} \pmod{p^\alpha}}} e\left(\frac{n\,(a+\overline{a})}{p^{2\alpha+1}}\right). \qquad \text{(Orthogonality of additive characters)} \end{aligned}$$

Now $a \equiv \overline{a} \pmod{p^\alpha}$ iff $p^\alpha \mid (a-1)(a+1)$, either $p^\alpha \mid a-1$ or $p^\alpha \mid a+1$. If $a = 1+tp^\alpha \pmod{p^{2\alpha+1}}$, so $\overline{a} = 1-tp^\alpha+t^2p^{2\alpha} \pmod{p^{2\alpha+1}}$; if $a = -1+tp^\alpha \pmod{p^{2\alpha+1}}$, so $\overline{a} = -1-$

$tp^\alpha - t^2p^{2\alpha} \pmod{p^{2\alpha+1}}$. Hence

$$\begin{aligned}
S(n,n;p^{2\alpha+1}) &= \sum_{t \pmod{p^{\alpha+1}}} e\left(\frac{n(2+t^2p^{2\alpha})}{p^{2\alpha+1}}\right) + \sum_{t \pmod{p^{\alpha+1}}} e\left(\frac{n(-2-t^2p^{2\alpha})}{p^{2\alpha+1}}\right) \\
&= e\left(\frac{2n}{p^{2\alpha+1}}\right) \sum_{t \pmod{p^{\alpha+1}}} e\left(\frac{nt^2}{p}\right) + e\left(\frac{-2n}{p^{2\alpha+1}}\right) \sum_{t \pmod{p^{\alpha+1}}} e\left(\frac{-nt^2}{p}\right) \\
&= p^\alpha e\left(\frac{2n}{p^{2\alpha+1}}\right) \sum_{t \pmod{p}} e\left(\frac{nt^2}{p}\right) + p^\alpha e\left(\frac{-2n}{p^{2\alpha+1}}\right) \sum_{t \pmod{p}} e\left(\frac{-nt^2}{p}\right) \\
&= p^\alpha \left(\frac{n}{p}\right) \varepsilon_p p^{1/2} e\left(\frac{2n}{p^{2\alpha+1}}\right) + p^\alpha \left(\frac{-n}{p}\right) \varepsilon_p p^{1/2} e\left(\frac{-2n}{p^{2\alpha+1}}\right) \\
&= \left(\frac{n}{p}\right) c^{1/2} \left(\varepsilon_p e\left(\frac{2n}{c}\right) + \left(\frac{-1}{p}\right) \varepsilon_p e\left(\frac{-2n}{c}\right)\right) \\
&= \left(\frac{n}{p}\right) c^{1/2} \left(\varepsilon_p e\left(\frac{2n}{c}\right) + \overline{\varepsilon_p} e\left(\frac{-2n}{c}\right)\right) \\
&= 2\left(\frac{n}{c}\right) c^{1/2} \operatorname{Re}\left(\varepsilon_p e\left(\frac{2n}{c}\right)\right).
\end{aligned}$$

□

We put the Lemma 4.2 and Lemma 4.3 in a unified form

Proposition 4.1. If $c = p^\beta$ with $\beta \geq 2$ and $\gcd(c, 2n) = 1$, then

$$S(n,n;c) = 2\left(\frac{n}{c}\right) c^{1/2} \operatorname{Re}\left(\varepsilon_c e\left(\frac{2n}{c}\right)\right).$$

More generally, if $c = p^\beta$ with $\beta \geq 2$ and $\gcd(c, 2mn) = 1$, then $S(m,n;c) = 0$ unless $m = \ell^2 n \pmod{c}$. For $m = \ell^2 n \pmod{c}$, we have

$$S(m,n;c) = S(\ell^2 n, n; c) = S(\ell n, \ell n; c) = 2\left(\frac{\ell n}{c}\right) c^{1/2} \operatorname{Re}\left(\varepsilon_c e\left(\frac{2\ell n}{c}\right)\right)$$

by $\gcd(\ell, c) = 1$ and Proposition 4.1.

In any case we conclude

Corollary 4.1. If $c = p^\beta$ with $\beta \geq 2$ and $\gcd(c, 2mn) = 1$, then

$$|S(m,n;c)| \leq 2c^{1/2}.$$

Proposition 4.2. Let $c = 2^\beta$ with $\beta \geq 2$, and assume $\gcd(c, mn) = 1$. If $\beta = 2\alpha$ with $\alpha \geq 1$, then

$$S(m,n;2^{2\alpha}) = 2^\alpha \sum_{\substack{s \pmod{2^\alpha} \\ s \text{ odd} \\ ms^2 \equiv n \pmod{2^\alpha}}} e\left(\frac{ms + n\overline{s}}{2^{2\alpha}}\right).$$

If $\beta = 2\alpha + 1$ with $\alpha \geq 1$, then

$$S(m,n;2^{2\alpha+1}) = 2^{\alpha+1} \sum_{\substack{s \pmod{2^\alpha} \\ s \text{ odd} \\ ms^2 \equiv n+2^\alpha \pmod{2^{\alpha+1}}}} e\left(\frac{ms+n\overline{s}}{2^{2\alpha+1}}\right).$$

Consequently,

$$|S(m,n;2^\beta)| \leq 4 \cdot 2^{\beta/2}.$$

In 1948 A. Weil proved the Riemann hypothesis for curves over finite fields, from which he deduced that

$$|S(1,n;p)| \leq 2p^{1/2}.$$

Theorem 4.1. For any $c \geq 1, m, n$ we have

$$|S(m,n;c)| \leq \gcd(m,n,c)^{1/2} c^{1/2} \tau(c)$$

where $\tau(c)$ denotes the divisor function.

The estimate $|S(m,n;p)| \ll p^{1-\delta}$ is historically motivated by Kloosterman's study of representation numbers of positive definite diagonal quaternary quadratic forms; in fact, Kloosterman introduced these sums in his 1926 work on the circle method for $ax^2 + by^2 + cz^2 + dt^2$.

4.2 Power-moments of Kloosterman sums

Define

$$V_k(p) = \sideset{}{^*}\sum_{a \pmod p} S^k(a,1;p).$$

Then

$$\begin{aligned}
(p-1)V_k(p) &= \sideset{}{^*}\sum_{b \pmod p} \sideset{}{^*}\sum_{a \pmod p} S^k(a,1;p) \\
&= \sideset{}{^*}\sum_{b \pmod p} \sideset{}{^*}\sum_{a \pmod p} S^k(1,a;p) \\
&= \sideset{}{^*}\sum_{b \pmod p} \sideset{}{^*}\sum_{a \pmod p} S^k(1,ab;p) \\
&= \sideset{}{^*}\sum_{b \pmod p} \sideset{}{^*}\sum_{a \pmod p} S^k(a,b;p) \\
&= \sum_{b \pmod p} \sum_{a \pmod p} S^k(a,b;p) - S^k(0,0;p) - \sideset{}{^*}\sum_{a \pmod p} S^k(a,0;p) - \sideset{}{^*}\sum_{b \pmod p} S^k(a,b;p).
\end{aligned}$$

Note that

$$S(0,0;p) = \sideset{}{^*}\sum_{a \pmod p} 1 = p-1$$

and

$$S(a,0;p) = \sideset{}{^*}\sum_{d \pmod p} e\left(\frac{a\overline{d}}{p}\right) = \sideset{}{^*}\sum_{d \pmod p} e\left(\frac{ad}{p}\right) = \sum_{d \pmod p} e\left(\frac{ad}{p}\right) - 1 = 0 - 1 = -1,$$

it follows that

$$(p-1)V_k(p) = \sum_{a \pmod p} \sum_{b \pmod p} S^k(a,b;p) - (p-1)^k - 2(-1)^k(p-1)$$

Since

$$S(a,b;p) = \sideset{}{^*}\sum_{d \pmod p} e\left(\frac{a\overline{d}+bd}{p}\right),$$

we have

$$S^k(a,b;p) = \sideset{}{^*}\sum_{x_1 \pmod p} \cdots \sideset{}{^*}\sum_{x_k \pmod p} e\left(\frac{a(x_1+\cdots+x_n)+b\left(\overline{x_1}+\cdots+\overline{x_k}\right)}{p}\right).$$

It follows that

$$\begin{aligned}
&\sum_{a \pmod p} \sum_{b \pmod p} S^k(a,b;p) \\
&= \sum_{a \pmod p} \sum_{b \pmod p} \sum_{x_1 \pmod p}^{*} \cdots \sum_{x_k \pmod p}^{*} e\left(\frac{a(x_1+\cdots+x_n)+b\left(\overline{x_1}+\cdots+\overline{x_k}\right)}{p}\right) \\
&= \sideset{}{^*}\sum_{x_1 \pmod p} \cdots \sideset{}{^*}\sum_{x_k \pmod p} \left(\sum_{a \pmod p} e\left(\frac{a\left(\overline{x_1}+\cdots+\overline{x_k}\right)}{p}\right)\right)\left(\sum_{b \pmod p} e\left(\frac{b(x_1+\cdots+x_k)}{p}\right)\right) \\
&= p^2\nu_k(p),
\end{aligned}$$

where

$$\nu_k(p) = \#\left\{x_1,\cdots,x_k \in (\mathbb{Z}/p\mathbb{Z})^\times : x_1+\cdots+x_k = 0, \overline{x_1}+\cdots+\overline{x_k} = 0\right\},$$

form the orthogonality of additive characters.

If $k=1$, we have $x_1 \equiv 0 \pmod p$ and $\overline{x_1} \equiv 0 \pmod p$. Such x_1 does not exist, so $\nu_1(p) = 0$.

Thus $(p-1)V_1(p) = p-1$, it follows that $V_1(p) = 1$. Indeed, one may compute directly

$$\begin{aligned} V_1(p) &= \sideset{}{^*}\sum_{a \pmod p} S(a,1;p) \\ &= \sideset{}{^*}\sum_{a \pmod p} \sideset{}{^*}\sum_{d \pmod p} e\left(\frac{a\overline{d}+d}{p}\right) \\ &= \sideset{}{^*}\sum_{d \pmod p} e\left(\frac{d}{p}\right) \sideset{}{^*}\sum_{a \pmod p} e\left(\frac{a\overline{d}}{p}\right) \\ &= -\sideset{}{^*}\sum_{d \pmod p} e\left(\frac{d}{p}\right) \\ &= 1. \end{aligned}$$

If $k=2$, we have $x_1 + x_2 \equiv 0 \pmod p$ and $\overline{x_1} + \overline{x_2} \equiv 0 \pmod p$. That is, $x_2 \equiv -x_1 \pmod p$, so $\nu_2(p) = p-1$. Hence $V_2(p) = p^2 - p - 1$.

If $k=3$, we have $x_1 + x_2 + x_3 \equiv 0 \pmod p$ and $\overline{x_1} + \overline{x_2} + \overline{x_3} \equiv 0 \pmod p$. Then from the first equation (by multiplying $x_1x_2x_3$) $x_3 \equiv -x_1 - x_2 \pmod p$, from the second equation $x_1x_2 + x_3(x_1+x_2) \equiv 0 \pmod p$. It follows that $x_1x_2 - (x_1+x_2)^2 \equiv 0 \pmod p$, then $x_1^2 + x_2^2 + x_1x_2 \equiv 0 \pmod p$. Consequently,

$$\begin{cases} x_1 + x_2 + x_3 \equiv 0 \pmod p \\ \overline{x_1} + \overline{x_2} + \overline{x_3} \equiv 0 \pmod p \end{cases} \iff \begin{cases} x_3 \equiv -(x_1+x_2) \pmod p \\ x_1^2 + x_2^2 + x_1x_2 \equiv 0 \pmod p \end{cases}.$$

Note that $x_1^2 + x_2^2 + x_1x_2 \equiv 0 \pmod p$ iff $x_1^2\left(1 + (x_2\overline{x_1}) + (x_2\overline{x_1})^2\right) \equiv 0 \pmod p$ iff $(x_2\overline{x_1})^2 + (x_2\overline{x_1}) + 1 \equiv 0 \pmod p$. Therefore

$$\begin{cases} x_1 + x_2 + x_3 \equiv 0 \pmod p \\ \overline{x_1} + \overline{x_2} + \overline{x_3} \equiv 0 \pmod p \end{cases} \iff \begin{cases} x_3 \equiv -(x_1+x_2) \pmod p \\ y \equiv x_2\overline{x_1} \pmod p \\ y^2 + y + 1 \equiv 0 \pmod p \end{cases}.$$

The number of solutions of $y \in (\mathbb{Z}/p\mathbb{Z})^\times$ satisfying $y^2 + y + 1 \equiv 0 \pmod p$ is

$$1 + \left(\frac{\Delta}{p}\right),$$

where $\Delta = -3$. Hence

$$\nu_3(p) = \#x_1 \cdot \#y = (p-1)\left(1 + \left(\frac{-3}{p}\right)\right)$$

implies that

$$V_3(p) = \left(\frac{-3}{p}\right)p^2 + 2p + 1.$$

For $k = 4$, we obtain

$$\begin{cases} x_1 + x_2 + x_3 + x_4 \equiv 0 \pmod p \\ \overline{x_1} + \overline{x_2} + \overline{x_3} + \overline{x_4} \equiv 0 \pmod p \end{cases} \iff \begin{cases} x_1 + x_2 \equiv u \pmod p \\ \overline{x_1} + \overline{x_2} \equiv v \pmod p \end{cases},$$

where $u \equiv -x_3 - x_4 \pmod p$ and $v \equiv -\overline{x_3} - \overline{x_4} \pmod p$, with $u, v \in \mathbb{Z}/p\mathbb{Z}$. Let $\eta(u, v)$ denote the number of solutions $(x_1, x_2) \in \left((\mathbb{Z}/p\mathbb{Z})^\times\right)^2$ to the system on the right-hand side, for given $u, v \in \mathbb{Z}/p\mathbb{Z}$.

Note that if (x_1, x_2) is a solution counted by $\eta(u, v)$, then (x_3, x_4) is a solution counted by $\eta(-u, -v)$. Conversely, if (x_1, x_2) is a solution for $\eta(u, v)$ and (x_3, x_4) is a solution for $\eta(-u, -v)$, then (x_1, x_2, x_3, x_4) satisfies $x_1 + x_2 + x_3 + x_4 \equiv 0 \pmod p$ and $\overline{x_1} + \overline{x_2} + \overline{x_3} + \overline{x_4} \equiv 0 \pmod p$. Hence

$$\begin{aligned} \nu_4(p) &= \sum_{v \pmod p} \sum_{u \pmod p} \eta(u, v)\, \eta(-u, -v) \\ &= \sum_{v \pmod p} \sum_{u \pmod p} \eta^2(u, v), \end{aligned}$$

since the map $(x, y) \mapsto (-x, -y)$ establishes a bijection between the solution sets of $\eta(u, v)$ and $\eta(-u, -v)$.

If $u = 0$, $x_1 + x_2 \equiv 0 \pmod p$ implies that $\overline{x_1} + \overline{x_2} \equiv 0 \pmod p$, i.e., $v \equiv 0 \pmod p$. For the converse given arbitrary $x_1 \in \mathbb{Z}/p\mathbb{Z}$, one may choose $x_2 = -x_1$ and then obtain a solution. Therefore $\eta(0, 0) = p - 1$ and $\eta(0, v) = 0$ if $v \neq 0$. Similarly $\eta(u, 0) = 0$ if $u \neq 0$. If $u, v \neq 0$, we have $\overline{x_1} + \overline{x_2} \equiv v \pmod p$, then $(x_1 + x_2)\overline{x_1 x_2} \equiv v \pmod p$. So $x_1 x_2 \equiv u\overline{v} \pmod p$. Therefore, $\eta(u, v)$ is the number of roots of equation $x^2 - ux + u\overline{v} \equiv 0 \pmod p$, which is precisely

$$1 + \left(\frac{1 - 4\overline{uv}}{p}\right), \quad (u, v \neq 0).$$

Consequently,

$$\begin{aligned} \nu_4(p) &= \eta^2(0, 0) + \sum_{u,v \pmod p}^{*} \eta^2(u, v) \\ &= (p-1)^2 + \sum_{u,v \pmod p}^{*} \left(1 + \left(\frac{1 - 4\overline{uv}}{p}\right)\right)^2 \\ &= (p-1)^2 + (p-1) \sum_{t \pmod p}^{*} \left(1 + \left(\frac{1 - 4\bar{t}}{p}\right)\right)^2 \end{aligned}$$

since for arbitrary fixed $t \in (\mathbb{Z}/p\mathbb{Z})^\times$ there exists $p - 1$ ordered pairs (u, v) with $uv = t$, $u, v \pmod p$. Note that $t \mapsto 1 - 4\bar{t}$ admits inverse $t \mapsto 4\overline{(1 - t)}$, it is a bijection from $(\mathbb{Z}/p\mathbb{Z})^\times$ to

$(\mathbb{Z}/p\mathbb{Z}) \setminus \{1\}$. Thus

$$\begin{aligned}
\nu_4(p) &= (p-1)^2 + (p-1) \sum_{\substack{t \pmod p \\ t \neq 1}} \left(1 + \left(\frac{t}{p}\right)\right)^2 \\
&= (p-1)^2 + (p-1) \left(\sum_{\substack{t \pmod p \\ t \neq 1}} 1 + 2 \sum_{\substack{t \pmod p \\ t \neq 1}} \left(\frac{t}{p}\right) + \sum_{\substack{t \pmod p \\ t \neq 1}} \left(\frac{t}{p}\right)^2 \right) \\
&= (p-1)^2 + (p-1)\left((p-1) + (-2) + (p-2)\right) \\
&= 3(p-1)(p-2).
\end{aligned}$$

Then

$$V_4(p) = 2p^3 - 3p^2 - 3p - 1.$$

For arbitrary $a \neq 0$ we have $|S(a,1;p)|^4 \leq V_4(p) < 16p^3$, so $|S(a,1;p)| < 2p^{3/4}$. If $p \nmid 2mn$, we then obtain $|S(m,n;p)| = |S(mn,1;p)| < 2p^{3/4}$.

Note that

$$\frac{V_4(p)}{V_2(p)} = \frac{2p^3 - 3p^2 - 3p - 1}{p^2 - p - 1} = 2p - 1 - \frac{2p+2}{p^2 - p - 1} > 2p - 2$$

holds for $p \geq 5$. Assume that $M = \max_{a \neq 0} |S(a,1;p)|^2$, then for every a we have $|S(a,1;p)|^4 \leq M|S(a,1;p)|^2$, it follows that

$$V_4(p) = \sum |S(a,1;p)|^4 \leq M \sum |S(a,1;p)|^2 = MV_2(p).$$

Therefore, there exists a with

$$|S(a,1;p)|^2 \geq \frac{V_4(p)}{V_2(p)} > 2p - 2,$$

i.e., $|S(a,1;p)| > (2p-2)^{1/2}$. This shows that the constant 2 in Weil's bound cannot be replaced by any number $< \sqrt{2}$.

From Weil's bound,

$$\left| \frac{S(a,1;p)}{\sqrt{p}} \right| \leq 2,$$

we then have

$$p^{-1/2} S(a,1;p) = 2\cos\theta_p(a),$$

where $0 \leq \theta_p(a) \leq \pi$ is called the Kloosterman sum angle. Chebyshev polynomials are defined recursively by

$$U_0(x) = 1, \quad U_1(x) = x, \quad xU_k(x) = U_{k-1}(x) + U_{k+1}(x).$$

One can show that

$$U_k(x) = \sum_{0 \le \ell \le k/2} \frac{(-1)^\ell (k-\ell)!}{\ell!(k-2\ell)!} x^{k-2\ell}.$$

Another way of defining the polynomials $U_k(x)$ is by

$$U_k(2\cos\theta) = \frac{\sin(k+1)\theta}{\sin\theta}.$$

The generating series of $U_k(x)$ is

$$\sum_{k=0}^{\infty} U_k(x) y^k = (1 - xy + y^2)^{-1}.$$

Theorem 4.2 (N. Katz)**.** For $k \ge 1$ and $p > 2$ we have

$$\left| \sum_{a \pmod p}^{*} U_k(2\cos\theta_p(a)) \right| \le \frac{1}{2}(k+1)p^{1/2}.$$

The key feature of the Katz estimate lies in the fact that the Tchebyshev polynomials $U_k(x)$ form an orthonormal basis in $\mathcal{L}^2([0,\pi],\mu)$ with respect to the Sato–Tate measure

$$\mu = 2\pi^{-1}(\sin\theta)^2 \mathrm{d}\theta.$$

Thus Theorem 4.2 implies that the Kloosterman angles $\theta_p(a)$ are asymptotically equidistributed for the Sato–Tate measure as a ranges over the primitive classes $\pmod p$ and $p \to \infty$. In other words, for any continuous function f on $[0,\pi]$ we have

$$\lim_{p\to\infty} \frac{1}{p-1} \sum_{a \pmod p}^{*} f(\theta_p(a)) = \frac{2}{\pi} \int_0^\pi f(\theta)(\sin\theta)^2 \mathrm{d}\theta.$$

A much harder question is, how is the Kloosterman sum angle distributed with respect to the modulus?

Conjecture. For a fixed $a \ne 0$ the Kloosterman angles $\theta_p(a)$ are asymptotically equidistributed for the Sato–Tate measure as p ranges over primes, i.e.,

$$\lim_{x\to\infty} \frac{1}{\pi(x)} \sum_{p \le x} f(\theta_p(a)) = \frac{2}{\pi} \int_0^\pi f(\theta)(\sin\theta)^2 \mathrm{d}\theta.$$

4.3 The Salié sums

Define the **Salié sum** by

$$T(m,n;c) = \sum_{d \pmod c}^{*} \left(\frac{d}{c}\right) e\left(\frac{m\overline{d} + nd}{c}\right).$$

Let

$$g(n,c) = \sum_{t \pmod c} e\left(\frac{nt^2}{c}\right).$$

Lemma 4.4. If $\gcd(c, 2n) = 1$, then

$$g(n,c) = \left(\frac{n}{c}\right)\varepsilon_c c^{1/2},$$

where

$$\varepsilon_c = \begin{cases} 1 & \text{if } c \equiv 1 \pmod 4, \\ \mathrm{i} & \text{if } c \equiv -1 \pmod 4. \end{cases}$$

Proof. Write $c = qr^2$, where q square-free, then

$$\begin{aligned}
g(n,c) &= \sum_{t \pmod{qr^2}} e\left(\frac{nt^2}{qr^2}\right) \\
&= \sum_{x \pmod{qr}} \sum_{y \pmod r} e\left(\frac{n(x+qry)^2}{qr^2}\right) \quad (t = x + qry \text{ where } x \pmod{qr} \text{ and } y \pmod r) \\
&= \sum_{x \pmod{qr}} e\left(\frac{nx^2}{qr^2}\right) \sum_{y \pmod r} e\left(\frac{2nxy}{r}\right) \\
&= r \sum_{\substack{x \pmod{qr} \\ r \mid x}} e\left(\frac{nx^2}{qr^2}\right) \qquad \text{(Orthogonality of additive characters)} \\
&= r \sum_{x \pmod q} e\left(\frac{nx^2}{q}\right) \qquad (r \mapsto rx) \\
&= rg(n,q).
\end{aligned}$$

It suffices to consider $g(n,q)$ where q is square-free. In this case,

$$\begin{aligned}
g(n,q) &= \sum_{x \pmod q} e\left(\frac{nx^2}{q}\right) \\
&= \sum_{y \pmod q} e\left(\frac{ny}{q}\right) \cdot \#\{x \pmod q : x^2 \equiv y \pmod q\}.
\end{aligned}$$

Thanks to the Chinese Remainder Theorem, let $q = p_1 \cdots p_s$, then

$$\begin{aligned}
\#\{x \pmod q : x^2 \equiv y \pmod q\} &= \prod_{i=1}^{s} \#\{x \pmod{p_i} : x^2 \equiv y \pmod{p_i}\} \\
&= \prod_{p \mid q} \left(1 + \left(\frac{y}{p}\right)\right).
\end{aligned}$$

It follows that

$$\begin{aligned}
g(n,q) &= \sum_{y \pmod q} e\left(\frac{ny}{q}\right) \prod_{p\mid q}\left(1+\left(\frac{y}{p}\right)\right) \\
&= \sum_{y \pmod q} e\left(\frac{ny}{q}\right) \sum_{d\mid q}\left(\frac{y}{d}\right) \\
&= \sum_{d\mid q} \sum_{y \pmod q} e\left(\frac{ny}{q}\right)\left(\frac{y}{d}\right) \\
&= \sum_{d\mid q} \sum_{\ell \pmod{q/d}} \sum_{t \pmod d} e\left(\frac{nd\ell}{q}\right) e\left(\frac{nt}{q}\right)\left(\frac{t}{d}\right) \quad (y = d\ell + t) \\
&= \sum_{d\mid q} \sum_{t \pmod d} \left(\frac{t}{d}\right) e\left(\frac{nt}{q}\right) \sum_{\ell \pmod{q/d}} e\left(\frac{\ell n}{q/d}\right) \\
&= \sum_{\substack{d\mid q \\ (q/d)\mid n}} \sum_{t \pmod d} \left(\frac{t}{d}\right) e\left(\frac{nt}{q}\right) \qquad \text{(Orthogonality of additive characters)} \\
&= \sum_{t \pmod q} \left(\frac{t}{q}\right) e\left(\frac{nt}{q}\right). \qquad \text{(}\gcd(q,2n)=1 \text{ implies } \gcd(q,n)=1, \text{ so } q=d\text{)}
\end{aligned}$$

Recall that we have

$$\sum_{a \pmod q} \chi(a) e\left(\frac{na}{q}\right) = \overline{\chi(n)} \sum_{a \pmod q} \chi(a) e\left(\frac{a}{q}\right),$$

since $\gcd(q,n)=1$. Say precisely, the number na runs through a complete residue system $\bmod q$ with a. Also, $|\chi(n)|^2 = \chi(n)\overline{\chi(n)} = 1$, so

$$\chi(a) = \chi(n)\overline{\chi(n)}\chi(a) = \overline{\chi(n)}\chi(na).$$

Therefore

$$\sum_{a \pmod q} \chi(a) e\left(\frac{na}{q}\right) = \overline{\chi(n)} \sum_{a \pmod q} \chi(na) e\left(\frac{na}{q}\right) = \overline{\chi(n)} \sum_{a \pmod q} \chi(a) e\left(\frac{a}{q}\right).$$

Note that the Legendre–Jacobi symbol is a Dirichlet character, so

$$g(n,q) = \left(\frac{n}{q}\right) \sum_{t \pmod q} \left(\frac{t}{q}\right) e\left(\frac{t}{q}\right) = \left(\frac{n}{q}\right) g(1,q).$$

Consequently,

$$\begin{aligned} g(n,c) &= rg(n,q) \qquad (c=qr^2) \\ &= r\left(\frac{n}{q}\right) g(1,q) \\ &= \left(\frac{n}{qr^2}\right) rg(1,q) \\ &= \left(\frac{n}{c}\right) g(1,c) \\ &= \left(\frac{n}{c}\right) \sqrt{c}\varepsilon_c. \end{aligned}$$

□

Suppose $\gcd(c,2n)=1$. Let us consider

$$F(x) = \sum_{d \pmod c}^{*} \left(\frac{d}{c}\right) e\left(\frac{m\overline{d}+ndx^2}{c}\right), \quad \text{where } x \pmod c.$$

Note that $F(1)=T(m,n;c)$. By the Fourier inversion,

$$F(x) = \frac{1}{c} \sum_{y \pmod c} \hat{F}(y) e\left(\frac{xy}{c}\right),$$

where

$$\hat{F}(y) = \sum_{x \pmod c} F(x) e\left(-\frac{xy}{c}\right) = \sum_{d \pmod c}^{*} \left(\frac{d}{c}\right) e\left(\frac{m\overline{d}}{c}\right) \sum_{x \pmod c} e\left(\frac{ndx^2-xy}{c}\right).$$

Complete the square,

$$\begin{aligned} e\left(\frac{ndx^2-xy}{c}\right) &= e\left(\frac{nd\left(x^2-2\cdot\overline{2nd}xy+\left(\overline{2nd}y\right)^2\right)}{c}\right) e\left(\frac{-nd\overline{4}\left(\overline{nd}\right)^2y^2}{c}\right) \\ &= e\left(\frac{nd\left(x-\overline{2nd}y\right)^2}{c}\right) e\left(\frac{-\overline{4nd}y^2}{c}\right). \end{aligned}$$

It follows that

$$\begin{aligned} \hat{F}(y) &= \sum_{d \pmod c}^{*} \left(\frac{d}{c}\right) e\left(\frac{m\overline{d}}{c}\right) e\left(\frac{-\overline{4nd}y^2}{c}\right) \sum_{\tilde{x} \pmod c} e\left(\frac{nd\tilde{x}^2}{c}\right) \quad \left(\tilde{x}=x-\overline{2nd}y\right) \\ &= \sum_{d \pmod c}^{*} \left(\frac{d}{c}\right) e\left(\frac{m\overline{d}}{c}\right) e\left(\frac{-\overline{4nd}y^2}{c}\right) g(nd,c). \end{aligned}$$

Since

$$g(nd,c) = \left(\frac{nd}{c}\right) g(1,c) = \left(\frac{d}{c}\right)\left(\frac{n}{c}\right) g(1,c) = \left(\frac{d}{c}\right) g(n,c),$$

we have

$$\begin{aligned}\hat{F}(y) &= \sum_{d \pmod c}^{*} \left(\frac{d}{c}\right) e\left(\frac{m\overline{d}}{c}\right) e\left(\frac{-\overline{4nd}y^2}{c}\right) \left(\frac{d}{c}\right) g(n,c) \\ &= \left(\sum_{d \pmod c}^{*} e\left(\frac{(4nm-y^2)d}{c}\right)\right) g(n,c). \qquad \left(\overline{4nd} \mapsto d\right)\end{aligned}$$

The Ramanujan sum can be computed as follows:

$$\begin{aligned}\sum_{d \pmod c}^{*} e\left(\frac{nd}{c}\right) &= \sum_{d \pmod c} e\left(\frac{nd}{c}\right) \sum_{\ell \mid \gcd(d,c)} \mu(\ell) \\ &= \sum_{\ell \mid c} \mu(\ell) \sum_{\substack{d \pmod c \\ \ell \mid d}} e\left(\frac{nd}{c}\right) \\ &= \sum_{\substack{\ell \mid c \\ (c/\ell) \mid n}} \mu(\ell) \frac{c}{\ell} \\ &= \sum_{r \mid \gcd(n,c)} r\mu\left(\frac{c}{r}\right). \qquad \left(r = \frac{c}{\ell}\right)\end{aligned}$$

so

$$\hat{F}(y) = g(n,c) \sum_{d \mid \gcd(4mn-y^2,c)} d\mu\left(\frac{c}{d}\right).$$

Then

$$\begin{aligned}F(x) &= \frac{1}{c} \sum_{y \pmod c} \sum_{d \mid \gcd(4nm-y^2,c)} d\mu\left(\frac{c}{d}\right) g(n,c) e\left(\frac{xy}{c}\right) \\ &= \frac{g(n,c)}{c} \sum_{d \mid c} d\mu\left(\frac{c}{d}\right) \sum_{\substack{y \pmod c \\ d \mid (4nm-y^2)}} e\left(\frac{xy}{c}\right).\end{aligned}$$

Let $y = \ell d + y'$ with $y' \pmod d$ and $\ell \pmod{c/d}$, $d \mid 4nm - y^2$ iff $y'^2 \equiv 4nm \pmod d$. So

$$\begin{aligned}\sum_{\substack{y \pmod c \\ d \mid (4nm-y^2)}} e\left(\frac{xy}{c}\right) &= \sum_{\substack{y' \pmod d \\ y'^2 \equiv 4nm \pmod d}} \sum_{\ell \pmod{c/d}} e\left(\frac{x(\ell d + y')}{c}\right) \\ &= \sum_{\substack{y \pmod d \\ y^2 \equiv 4nm \pmod d}} e\left(\frac{xy}{c}\right) \sum_{\ell \pmod{c/d}} e\left(\frac{x\ell}{c/d}\right) \\ &= \frac{c}{d} \sum_{\substack{y \pmod d \\ y^2 \equiv 4mn \pmod d \\ (c/d) \mid x}} e\left(\frac{xy}{c}\right). \qquad \text{(Orthogonality of additive characters)}\end{aligned}$$

If $\gcd(x,c)=1$, then $(c/d) \mid x$ implies that $d=c$, it follows that

$$\begin{aligned} F(x) &= \frac{g(n,c)}{c} c \sum_{\substack{y \pmod c \\ y^2 \equiv 4mn \pmod c}} e\left(\frac{xy}{c}\right) \\ &= \left(\frac{n}{c}\right) g(1,c) \sum_{\substack{y \pmod c \\ y^2 \equiv mn \pmod c}} e\left(\frac{2xy}{c}\right). \quad (y \mapsto 2y) \end{aligned}$$

Set $x=1$, we obtain

Lemma 4.5. If $\gcd(c,2n)=1$, we have

$$T(m,n;c) = \left(\frac{n}{c}\right) \varepsilon_c \sqrt{c} \sum_{\substack{y \pmod c \\ y^2 \equiv mn \pmod c}} e\left(\frac{2y}{c}\right).$$

Corollary 4.2. If $\gcd(c,2mn)=1$ we have

$$|T(m,n;c)| \le c^{1/2} \tau(c).$$

Proof. Let $c = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$, then

$$\begin{aligned} |T(m,n;c)| &= \sqrt{c} \left| \sum_{\substack{y \pmod c \\ y^2 \equiv mn \pmod c}} e\left(\frac{2y}{c}\right) \right| \\ &\le \sqrt{c} \sum_{\substack{y \pmod c \\ y^2 \equiv mn \pmod c}} 1 \\ &= \sqrt{c} \prod_{i=1}^{s} \# \left\{ y \pmod{p_i^{\alpha_i}} : y^2 \equiv mn \pmod{p_i^{\alpha_i}} \right\} \end{aligned}$$

thanks to the Chinese Remainder Theorem. Since $\gcd(c,2mn)=1$, $p_i \mid c$ implies that $p_i \nmid mn$. In $\mathbb{F}_p$, a quadratic equation has at most two roots, so $\#\{y \pmod p : y^2 \equiv mn \pmod p\} \le 2$. If $p \mid y$, then $p \mid y^2$ implies that $p \mid mn$, a contradiction, hence $p \nmid y$. For the polynomial $f(y) = y^2 - mn$, we have $f'(y) = 2y \not\equiv 0 \pmod p$. By Hensel's Lemma, $\# \left\{ y \pmod{p_i^{\alpha_i}} : y^2 \equiv mn \pmod{p_i^{\alpha_i}} \right\} = \#\{y \pmod p : y^2 \equiv mn \pmod p\} \le 2$. Consequently,

$$|T(m,n;c)| \le \sqrt{c} \prod_{i=1}^{s} 2 \le \sqrt{c} \prod_{i=1}^{s} (\alpha_i + 1) = \sqrt{c} \tau(c).$$

□

Corollary 4.3. If $\gcd(c,2n)=1$ we have

$$T(n,n;c) = \left(\frac{n}{c}\right) g(1,c) \sum_{\substack{ab=c \\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{a}}{b} - \frac{\overline{b}}{a}\right)\right),$$

here $b\overline{b} \equiv 1 \pmod a$ and $a\overline{a} \equiv 1 \pmod b$.

Proof. From Lemma 4.5,

$$T(n,n;c) = \left(\frac{n}{c}\right) g(1,c) \sum_{\substack{y \pmod c \\ y^2 \equiv n^2 \pmod c}} e\left(\frac{2y}{c}\right).$$

Since we have

$$\sum_{\substack{y \pmod c \\ y^2 \equiv n^2 \pmod c}} e\left(\frac{2y}{c}\right) = \sum_{\substack{y \pmod c \\ y^2 \equiv 1 \pmod c}} e\left(\frac{2ny}{c}\right)$$

by $y \mapsto ny$ since $\gcd(c,n) = 1$. Write $c = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$, then by the Chinese Remainder Theorem, solutions to $y^2 \equiv 1 \pmod c$ is precisely solutions to

$$\begin{cases} y^2 \equiv 1 \pmod{p_1^{\alpha_1}}, \\ \cdots \\ y^2 \equiv 1 \pmod{p_s^{\alpha_s}}. \end{cases}$$

The equation $y^2 \equiv 1 \pmod{p_i^{\alpha_i}}$ is equivalent to $(y-1)(y+1) \equiv 0 \pmod{p_i^{\alpha_i}}$. From $\gcd(c,2) = 1$ and $\gcd(y-1, y+1) \mid 2$, we obtain $p_i \nmid \gcd(y-1, y+1)$, which implies that $\min\{v_{p_i}(y-1), v_{p_i}(y+1)\} = 0$, where v_{p_i} is p_i-adic valuation. Moreover $p_i^{\alpha_i} \mid (y-1)(y+1)$ iff $v_{p_i}(y-1) + v_{p_i}(y+1) \geq \alpha_i$. Thus $p_i^{\alpha_i} \mid (y-1)$ or $p_i^{\alpha_i} \mid (y+1)$. Hence the solutions of $y^2 \equiv 1 \pmod{p_i^{\alpha_i}}$ are $y = 1 \bmod p_i^{\alpha_i}$ or $y = -1 \bmod p_i^{\alpha_i}$. Write $b = \prod_{i \in I} p_i^{\alpha_i}$ where $i \in I$ such that $y = 1 \bmod p_i^{\alpha_i}$ and $a = \prod_{j \in J} p_j^{\alpha_j}$ where $j \in J$ is such that $y = -1 \bmod p_j^{\alpha_j}$. By the Chinese Remainder Theorem, for every decomposition $c = ab$ with $\gcd(a,b) = 1$, there exists unique residue class modulo c satsifying $y \equiv -1 \pmod a$ and $y \equiv 1 \pmod b$, and such y can be written as $y = a\overline{a}(1) + b\overline{b}(-1) \bmod c$. Therefore,

$$\sum_{\substack{y \pmod c \\ y^2 \equiv 1 \pmod c}} e\left(\frac{2ny}{c}\right) = \sum_{\substack{ab=c \\ \gcd(a,b)=1}} e\left(\frac{2n}{ab}\left(a\overline{a} - b\overline{b}\right)\right) = \sum_{\substack{ab=c \\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{a}}{b} - \frac{\overline{b}}{a}\right)\right).$$

□

5 Bounds for the Fourier Coefficients of Cusp Forms

5.1 General estimates

Let f be a cusp form for Γ with respect to ϑ. (In particular, we may consider $\Gamma = \Gamma_0(q)$ and $\vartheta = \chi$, where χ is Dirichlet character modulo q.) Then at ∞, f admits Fourier expansion

$$f(z) = \sum_{n \geq 1} a(n) e(nz) = \sum_{n \geq 1} \tilde{a}(n) n^{(k-1)/2} e(nz).$$

Claim 5.1. $F(z) = (\mathrm{Im}(z))^{k/2}|f(z)|$ is bounded on $\mathbb{H}$.

Proof. Indeed, F is Γ-invariant.

$$\begin{aligned} F(\gamma z) &= (\mathrm{Im}(\gamma z)^{k/2}|f(\gamma z)| \\ &= \left(\frac{\mathrm{Im}(z)}{|j(\gamma,z)|^2}\right)^{k/2} |\vartheta(\gamma)||j(\gamma,z)|^k|f(z)| \\ &= (\mathrm{Im}(z))^{k/2}|f(z)| \\ &= F(z). \end{aligned}$$

Therefore it only suffices to show F is bounded in a fundamental domain $\mathcal{F}$. Write $\mathcal{F} = K \bigcup_{\mathfrak{a}} U_{\mathfrak{a}}$, where K compact and $U_{\mathfrak{a}}$ is a neighborhood of cusp $\mathfrak{a}$. Clearly $F(z) = (\mathrm{Im}(z))^{k/2}|f(z)|$ is continuous, so it is bounded in K. Given a cusp $\mathfrak{a}$ and let $\sigma_{\mathfrak{a}}$ be the scaling matrix. Then

$$f_{\mathfrak{a}}(z) = j(\sigma_{\mathfrak{a}}, z)^{-k} f(\sigma_{\mathfrak{a}} z) = \sum_{n\geq 1} a_{\mathfrak{a}}(n) e(nz)$$

since f is a cusp form. It follows that $f_{\mathfrak{a}}(z) = e^{2\pi \mathrm{i} z} h_{\mathfrak{a}}\left(e^{2\pi \mathrm{i} z}\right)$ for some $h_{\mathfrak{a}}$ which is holomorphic at 0. Thus $|f_{\mathfrak{a}}| = \left|e^{2\pi \mathrm{i} z}\right| \left|h_{\mathfrak{a}}\left(e^{2\pi \mathrm{i} z}\right)\right| \ll e^{-2\pi y}$, where $z = x + \mathrm{i}y$. Hence

$$\begin{aligned} F(\sigma_{\mathfrak{a}} z) &= \left(\mathrm{Im}(\sigma_{\mathfrak{a}} z)\right)^{k/2} |f(\sigma_{\mathfrak{a}} z)| \\ &= \left(\frac{\mathrm{Im}(z)}{|j(\sigma_{\mathfrak{a}}, z)|^2}\right)^{k/2} |j(\sigma_{\mathfrak{a}}, z)|^k |f_{\mathfrak{a}}(z)| \\ &= (\mathrm{Im}(z))^{k/2} |f_{\mathfrak{a}}(z)| \\ &\ll y^{k/2} e^{-2\pi y} \qquad\qquad (z = x + \mathrm{i}y) \\ &\ll 1. \end{aligned}$$

Thus F is bounded at the neighborhood of cusp $\mathfrak{a}$. We finish the proof. □

Claim 5.2. Let f be a modular form with multiplier ϑ. If $F(z)$ is bounded, then f is a cusp form.

Proof. Let $\mathfrak{a}$ be arbitrary cusp, and $\sigma_{\mathfrak{a}}$ be the scaling matrix. Define $f_{\mathfrak{a}}(z) = j(\sigma_{\mathfrak{a}}, z)^{-k} f(\sigma_{\mathfrak{a}} z)$. Since

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1} \Gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}},$$

there exists $\gamma_{\mathfrak{a}} \in \Gamma_{\mathfrak{a}}$ such that $\sigma_{\mathfrak{a}} T = \gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}}$, then

$$\begin{aligned}
f_{\mathfrak{a}}(z+1) &= f_{\mathfrak{a}}(Tz) \\
&= j(\sigma_{\mathfrak{a}}, Tz)^{-k} f(\sigma_{\mathfrak{a}} Tz) \\
&= j(\sigma_{\mathfrak{a}}, Tz)^{-k} f(\gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}} z) \\
&= j(\sigma_{\mathfrak{a}}, Tz)^{-k} \vartheta(\gamma_{\mathfrak{a}}) j(\gamma_{\mathfrak{a}}, \sigma_{\mathfrak{a}} z)^k f(\sigma_{\mathfrak{a}} z) \\
&= (j(\sigma_{\mathfrak{a}}, Tz) j(T, z))^{-k} \vartheta(\gamma_{\mathfrak{a}}) (j(\sigma_{\mathfrak{a}}, z)^{-1} j(\gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}}, z))^k f(\sigma_{\mathfrak{a}} Tz) \\
&= j(\sigma_{\mathfrak{a}} T, z)^{-k} \vartheta(\gamma_{\mathfrak{a}}) (j(\sigma_{\mathfrak{a}}, z)^{-1} j(\sigma_{\mathfrak{a}} T, z))^k f(\sigma_{\mathfrak{a}} Tz) \\
&= \vartheta(\gamma_{\mathfrak{a}}) j(\sigma_{\mathfrak{a}}, z)^{-k} f(\sigma_{\mathfrak{a}} Tz) \\
&= \vartheta(\gamma_{\mathfrak{a}}) f_{\mathfrak{a}}(z).
\end{aligned}$$

Since $|\vartheta(\gamma_{\mathfrak{a}})| = 1$, we can write $\vartheta(\gamma_{\mathfrak{a}}) = e(\kappa_{\mathfrak{a}})$ with $0 \leq \kappa_{\mathfrak{a}} < 1$. It follows that $e(-\kappa_{\mathfrak{a}} z) f_{\mathfrak{a}}(z)$ has period 1. Thus $e(-\kappa_{\mathfrak{a}} z) f_{\mathfrak{a}}(z)$ admits Fourier expansion at the cusp ∞

$$e(-\kappa_{\mathfrak{a}} z) f_{\mathfrak{a}}(z) = \sum_{n \geq 0} a_{\mathfrak{a}}(n) e(nz),$$

then

$$f_{\mathfrak{a}}(z) = \sum_{n \geq 0} a_{\mathfrak{a}}(n) e((n + \kappa_{\mathfrak{a}}) z).$$

Therefore F is bounded implies that there exists $C > 0$ such that $y^{k/2} |f_{\mathfrak{a}}(z)| \leq C$ for all $z = x + \mathrm{i}y \in \mathbb{H}$. Then $|f_{\mathfrak{a}}(z)| \leq C y^{-k/2}$, so $f_{\mathfrak{a}}(z) \to 0$ as $y \to \infty$. If $\kappa_{\mathfrak{a}} \neq 0$, then

$$f_{\mathfrak{a}}(z) = \sum_{n \geq 0} a_{\mathfrak{a}}(n) e((n + \kappa_{\mathfrak{a}}) z)$$

has no constant term; if $\kappa_{\mathfrak{a}} = 0$,

$$f_{\mathfrak{a}} = \sum_{n \geq 0} a_{\mathfrak{a}}(n) e(nz),$$

here

$$\begin{aligned}
|a_{\mathfrak{a}}(0)| &= \left| \int_0^1 f_{\mathfrak{a}}(x + \mathrm{i}y) \mathrm{d}x \right| \\
&\leq \int_0^1 |f_{\mathfrak{a}}(x + \mathrm{i}y)| \mathrm{d}x \\
&\leq \int_0^1 C y^{-k/2} \mathrm{d}x \\
&= C y^{-k/2} \to 0
\end{aligned}$$

as $y \to \infty$. Thus $|a_{\mathfrak{a}}(0)| = 0$. Consequently f is a cusp form. □

Now,

$$\begin{aligned}\int_0^1 |f(z)|^2 \mathrm{d}x &= \int_0^1 \left(\sum_{n\geq 1} a(n)e(nz)\right)\left(\sum_{m\geq 1} \overline{a(m)e(mz)}\right)\mathrm{d}x \\ &= \sum_{n\geq 1} a(n) \sum_{m\geq 1} \overline{a(m)} \left(\int_0^1 \mathrm{e}^{2\pi \mathrm{i}(n-m)x}\mathrm{d}x\right) \mathrm{e}^{-2\pi ny}\mathrm{e}^{-2\pi my} \\ &= \sum_{n\geq 1} |a(n)|^2 \mathrm{e}^{-4\pi ny} \\ &= \sum_{n\leq N} |a(n)|^2 \mathrm{e}^{-4\pi ny} + \sum_{n>N} |a(n)|^2 \mathrm{e}^{-4\pi ny} \\ &\ll y^{-k}\end{aligned}$$

from Claim 5.1. Therefore

$$\sum_{n\leq N} |a(n)|^2 \mathrm{e}^{-4\pi Ny} \leq \sum_{n\leq N} |a(n)|^2 \mathrm{e}^{-4\pi ny} \ll y^{-k}, \quad \forall N \geq 1.$$

Choose $y = 1/N$, we obtain

Theorem 5.1. For any $N \geq 1$ we have

$$\sum_{n\leq N} |a(n)|^2 \ll N^k.$$

where the implied constant depends on f.

In addition,

$$\sum_{n\leq N} |\tilde{a}(n)|^2 \ll N.$$

Indeed, from $|\tilde{a}(n)|^2 = |a(n)|^2/n^{k-1}$, we have

$$\begin{aligned}\sum_{n\leq N} |\tilde{a}(n)|^2 &= \sum_{n\leq N} \frac{|a(n)|^2}{n^{k-1}} \\ &= \frac{1}{N^{k-1}} \sum_{n\leq N} |a(n)|^2 + (k-1)\int_1^N \frac{1}{t^k}\left(\sum_{n\leq t} |a(n)|^2\right)\mathrm{d}t \\ &\ll \frac{N^k}{N^{k-1}} + \int_1^N \frac{t^k}{t^k}\mathrm{d}t \\ &\ll N\end{aligned}$$

by Abel summation. That is, on average $\tilde{a}(n)$ is of size 1. Later we will find that

$$\sum_{n\leq N} |\tilde{a}(n)|^2 = c_f N + O\left(N^{3/5-\delta}\right)$$

due to Bingrong Huang. The error term was later improved further by him. See [Hua21], [Hua24].

By Cauchy–Schwarz,

$$\sum_{n\le N}|a(n)|\le\left(\sum_{n\le N}|a(n)|^2\right)^{1/2}\left(\sum_{n\le N}1\right)^{1/2}\ll N^{k/2}N^{1/2}\ll N^{(k+1)/2}.$$

We have the following estimate:

Corollary 5.1. For any $N\ge 1$ we have

$$\sum_{n\le N}|a(n)|\ll N^{(k+1)/2}.$$

Theorem 5.2. For any real α and $N\ge 1$ we have

$$\sum_{n\le N}a(n)e(\alpha n)\ll N^{k/2}\log 2N.$$

where the implied constant depends only on f (not on α).

This is equivalent to

$$\sum_{n\le N}\tilde{a}(n)e(\alpha n)\ll N^{1/2}\log 2N.$$

where the implied constant depends only on f (not on α).

Proof. Note that

$$f(z)=\sum_{n\ge 1}a(n)e(nz),$$

where

$$a(n)=\int_0^1 f(z)e(-nz)\mathrm{d}x.$$

Then

$$\begin{aligned}\sum_{n\le N}a(n)e(\alpha n)&=\int_0^1 f(z)\sum_{n\le N}e(-nz)e(\alpha n)\mathrm{d}x\\&=\int_{-\alpha}^{1-\alpha}f(z+\alpha)\sum_{n\le N}e(-nz)\mathrm{d}x\quad(x\mapsto x+\alpha)\\&=\int_{-\alpha}^{1-\alpha}f(z+\alpha)\frac{e(-Nz)-1}{1-e(z)}\mathrm{d}x.\end{aligned}$$

Since

$$\left|\frac{e(-Nz)-1}{1-e(z)}\right|=\left|\frac{\mathrm{e}^{2\pi Ny}\mathrm{e}^{-2\pi\mathrm{i}Nx}-1}{1-e(z)}\right|\le\frac{\left|\mathrm{e}^{2\pi ny}\mathrm{e}^{-2\pi\mathrm{i}Nx}\right|+1}{|1-e(z)|}\ll\frac{\mathrm{e}^{2\pi Ny}}{|1-e(z)|},$$

we have

$$\sum_{n\le N} a(n)e(\alpha n) \ll \int_{-\alpha}^{1-\alpha} |f(z+\alpha)| \frac{\mathrm{e}^{2\pi Ny}}{|1-e(z)|}\mathrm{d}x$$
$$\ll y^{-k/2}\mathrm{e}^{2\pi Ny}\int_0^1 \frac{1}{|1-e(z)|}\mathrm{d}x$$

since $x \mapsto |1-e(x+\mathrm{i}y)|^{-1}$ is of period 1.

Now

$$|1-e(x+\mathrm{i}y)|^2 = \left(1-\mathrm{e}^{-2\pi y}\cos(2\pi x)\right)^2 + \left(\mathrm{e}^{-2\pi y}\sin(2\pi x)\right)^2$$
$$= \left(1-\mathrm{e}^{-2\pi y}\right)^2 + 4\mathrm{e}^{-2\pi y}\sin^2(\pi x).$$

Note that when $0 < y \le 1$, we have $1-\mathrm{e}^{-2\pi y} \asymp y$ and $|\sin(\pi x)| \asymp \|x\|$, where $\|x\|$ is the distance between x and the nearest integer. Then $|1-e(x+\mathrm{i}y)|^2 \asymp y^2 + \|x\|^2$. It follows that

$$\int_0^1 \frac{\mathrm{d}x}{|1-e(x+\mathrm{i}y)} \ll \int_0^1 \frac{\mathrm{d}x}{\sqrt{y^2+\|x\|^2}}$$
$$\le \int_0^1 \frac{\mathrm{d}x}{y+\|x\|}$$
$$= 2\int_0^{1/2} \frac{\mathrm{d}x}{y+x}$$
$$\ll \log(2+y^{-1}).$$

Therefore we have

$$\sum_{n\le N} a(n)e(\alpha n) \ll y^{-k/2}\mathrm{e}^{2\pi Ny}\log(2+y^{-1}).$$

If y is too small, then $y^{-k/2}$ is too large; if y is too large, then $e^{2\pi Ny}$ is too large. It is natural to take $Ny \asymp 1$, i.e., $y = 1/N$, then

$$\sum_{n\le N} a(n)e(\alpha n) \ll N^{k/2}\log 2N.$$

Now we prove the equivalence of the two assertions. First from

$$\sum_{n\le N} a(n)e(\alpha n) \ll N^{k/2}\log 2N$$

and $\tilde{a}(n) = a(n)n^{-(k-1)/2}$, we have

$$\begin{aligned}
\sum_{n\le N}\tilde{a}(n)e(\alpha n) &= \sum_{n\le N}a(n)n^{-(k-1)/2}e(\alpha n)\\
&= \frac{1}{N^{(k-1)/2}}\sum_{n\le N}a(n)e(\alpha n) - \int_1^N\left(\sum_{n\le t}a(n)e(\alpha n)\right)\mathrm{d}t^{-(k-1)/2}\\
&\ll N^{k/2}\log 2N\cdot N^{-(k-1)/2} + \int_1^N t^{k/2}\log 2t\cdot t^{-(k+1)/2}\mathrm{d}t\\
&= N^{1/2}\log 2N + \int_1^N t^{-1/2}\log 2t\mathrm{d}t\\
&\ll N^{1/2}\log 2N + N^{1/2}\log 2N\\
&\ll N^{1/2}\log 2N.
\end{aligned}$$

On the other hand, if the assertion

$$\sum_{n\le N}\tilde{a}(n)e(\alpha n) \ll N^{1/2}\log 2N$$

holds, we have

$$\begin{aligned}
\sum_{n\le N}a(n)e(\alpha n) &= \sum_{n\le N}\tilde{a}(n)n^{(k-1)/2}e(\alpha n)\\
&= N^{(k-1)/2}\sum_{n\le N}\tilde{a}(n)e(\alpha n) - \int_1^N\left(\sum_{n\le t}\tilde{a}(n)e(\alpha n)\right)\mathrm{d}t^{(k-1)/2}\\
&\ll N^{1/2}\log 2N\cdot N^{(k-1)/2} + \int_1^N t^{k/2-1}\log 2t\mathrm{d}t\\
&\ll N^{k/2}\log 2N.
\end{aligned}$$

□

Remark 5.1. It's not true

$$\sum_{n\le N}a(n)e(\alpha n) = O\left(N^{k/2-\eta}\right)\quad(\eta>0).$$

Otherwise, we will deduce that

$$\sum_{n\le N}\tilde{a}(n)e(\alpha n) = O\left(N^{1/2-\eta}\right),$$

which implies that

$$\sum_{n\le N}|\tilde{a}(n)|^2 = \int_0^1\left|\sum_{n\le N}\tilde{a}(n)e(\alpha n)\right|^2\mathrm{d}\alpha = O(N^{1-2\eta}),$$

contradiction to

$$\sum_{n\le N}|\tilde{a}(n)|^2 = c_f N + O\left(x^{3/5-\delta}\right).$$

Corollary 5.2. For any $q \ge 1$ and $a \pmod q$ we have

$$\sum_{\substack{n\le N\\ n\equiv a \pmod q}} a(n) \ll N^{k/2}\log 2N$$

where the implied constant depends only on f.

Equivalently,

$$\sum_{\substack{n\le N\\ n\equiv a \pmod q}} \tilde{a}(n) \ll N^{1/2}\log 2N$$

where the implied constant depends only on f.

Proof.

$$\begin{aligned}\sum_{n\le N} a(n)\delta_{n\equiv a \pmod q} &= \sum_{n\le N} a(n)\frac{1}{q}\sum_{b \pmod q} e\left(\frac{b(n-a)}{q}\right)\\ &= \frac{1}{q}\sum_{b \pmod q} e\left(\frac{-ab}{q}\right)\sum_{n\le N} a(n)e\left(\frac{bn}{q}\right)\\ &\ll \frac{1}{q}\sum_{b \pmod q} N^{k/2}\log 2N\\ &\ll N^{k/2}\log 2N.\end{aligned}$$

□

Theorem 5.3. For any complex numbers $(\alpha_m), (\beta_n)$, we have

$$\sum_{m<n\le N}\alpha_m\beta_n a(n-m) \ll N^{k/2}\|\alpha\|\|\beta\|$$

where $\|\alpha\|, \|\beta\|$ denote the ℓ_2-norms of $\alpha = (\alpha_m)$ and $\beta = (\beta_n)$ respectively, i.e.,

$$\|\alpha\| = \sqrt{\sum_i |\alpha_i|^2}, \quad \|\beta\| = \sqrt{\sum_i |\beta_i|^2}$$

respectively.

Proof. Note that

$$f(z) = \sum_{n\ge 1} a(n)e(nz),$$

and

$$a(n-m)=\int_0^1 f(z)e(-(n-m)z)\mathrm{d}x.$$

If $n-m\le 0$, let $a(n-m)=0$, then

$$\begin{aligned}
\sum_{m\le N}\sum_{n\le N}\alpha_m\beta_n a(n-m)&=\int_0^1 f(z)\sum_{m\le N}\alpha_m e(mz)\sum_{n\le N}\beta_n e(-nz)\mathrm{d}x\\
&\ll\int_0^1|f(z)|\left|\sum_{m\le N}\alpha_m e(mz)\right|\left|\sum_{n\le N}\beta_n e(-nz)\right|\mathrm{d}x\\
&\ll y^{-k/2}\int_0^1\left|\sum_{m\le N}\alpha_m e(mz)\right|\left|\sum_{n\le N}\beta_n e(-nz)\right|\mathrm{d}x && \text{(Claim 5.1)}\\
&\ll y^{-k/2}\left(\int_0^1\left|\sum_{m\le N}\alpha_m e(mz)\right|^2\mathrm{d}x\right)^{1/2}\left(\int_0^1\left|\sum_{n\le N}\beta_n e(-nz)\right|^2\mathrm{d}x\right)^{1/2} && \text{(Cauchy–Schwarz)}\\
&\ll y^{-k/2}\left(\sum_{m\le N}|\alpha_m|^2\mathrm{e}^{-4\pi my}\right)^{1/2}\left(\sum_{n\le N}|\beta_n|^2\mathrm{e}^{4\pi ny}\right)^{1/2}\\
&\ll N^{k/2}\left(\sum_{m\le N}|\alpha_m|^2\mathrm{e}^{-4\pi m/N}\right)^{1/2}\left(\sum_{n\le N}|\beta_n|^2\mathrm{e}^{4\pi n/N}\right)^{1/2} && \left(y=\frac{1}{N}\right)\\
&\ll N^{k/2}\left(\sum_{m\le N}|\alpha_m|^2\right)^{1/2}\left(\sum_{n\le N}|\beta_n|^2\right)^{1/2}.
\end{aligned}$$

□

The assertion

$$\sum_{n\le N}\tilde{a}(n)e(\alpha n)\ll N^{1/2}\log 2N$$

is called the **Wilton's bound**. Note that

$$|\tilde{a}(n)|^2\ll\sum_{n\le N}|\tilde{a}(n)|^2\ll N$$

implies that $|\tilde{a}(n)|\ll n^{1/2}$. The pointwise bound obtained by Wilton's method is

$$|\tilde{a}(n)|\ll n^{1/3}\log 2n,$$

which improve the trivial estimate $|\tilde{a}(n)|\ll n^{1/2}$.

5.2 Estimates by Kloosterman sums

Recall the Poincaré series is defined by

$$\mathbf{P}_m(z) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \overline{\chi(\gamma)} j_\gamma(z)^{-k} e(m\gamma z),$$

it admits Fourier expansion

$$\mathbf{P}_m(z) = \sum_{n\geq 1} p(m,n)e(nz)$$

since the Poincaré series is a cusp form. In Corollary 3.1, we proved that $\mathcal{S}_k(\Gamma_0(q),\chi)$ is spanned by $\{\mathbf{P}_m\}_{m\geq 1}$. We wish to improve $|\tilde{a}(n)| \ll n^{1/2}$. From the Petersson trace formula (Theorem 3.2),

$$p(m,n) = \left(\frac{n}{m}\right)^{(k-1)/2} \left(\delta_{mn} + 2\pi \mathrm{i}^{-k} \sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \frac{S_\chi(m,n;c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right).$$

Meanwhile,

$$p(m,n) = \sum_{f\in\mathcal{F}} \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \overline{\hat{f}(m)}\hat{f}(n),$$

where $\mathcal{F}$ is an orthonormal basis of $\mathcal{S}_k(\Gamma_0(q),\chi)$. Note that

$$\frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \left|\hat{f}(n)\right|^2 \leq \sum_{f\in\mathcal{F}} \frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \left|\hat{f}(n)\right|^2 = p(n,n),$$

then

$$\left|\hat{f}(n)\right|^2 \leq \frac{(4\pi n)^{k-1}}{\Gamma(k-1)} p(n,n).$$

Now

$$p(n,n) = 1 + 2\pi \mathrm{i}^{-k} \sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \frac{S_\chi(n,n;c)}{c} J_{k-1}\left(\frac{4\pi n}{c}\right).$$

Note that

$$J_{k-1}(x) \ll \min\{x^{k-1}, x^{-1/2}\}.$$

Consequently, for arbitrary $1/2 \leq \delta < 1$, we have

$$J_{k-1}(x) \ll x^{2\delta-1} \quad (x>0).$$

Suppose

$$\sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \frac{|S_\chi(n,n;c)|}{c^{2\delta}} \ll n^\varepsilon,$$

it follows that

$$p(n,n) \ll 1 + \sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \frac{|S_\chi(n,n;c)|}{c} J_{k-1}\left(\frac{4\pi n}{c}\right) \ll n^{2\delta-1+\varepsilon}.$$

Thus

$$\left|\hat{f}(n)\right|^2 \ll n^{k-1}p(n,n) \ll n^{k+2\delta-2+\varepsilon}.$$

That is,

Proposition 5.1. Let $k > 2$, and suppose that for some $1/2 \le \delta < 1$ one has

$$\sum_{\substack{c>0 \\ c\equiv 0 \pmod q}} \frac{|S_\chi(n,n;c)|}{c^{2\delta}} \ll n^{\varepsilon}.$$

Then for every $f \in \mathcal{S}_k(\Gamma_0(q),\chi)$,

$$\left|\hat{f}(n)\right| \ll n^{k/2-1+\delta+\varepsilon}.$$

Equivalently, from $\tilde{a}(n) = \hat{f}(n)n^{-(k-1)/2}$,

$$|\tilde{a}(n)| \ll n^{\delta-1/2+\varepsilon}.$$

Indeed, this Proposition holds for general Γ and ϑ.

5.3 Fourier coefficients of cusp forms of half integer weight

Let f be a cusp form in $\mathcal{S}_k(\Gamma_0(N),\vartheta)$ with

$$\vartheta(\tau) = \left(\overline{\varepsilon_d}\left(\frac{c}{d}\right)\right)^{2k}, \quad \tau = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4), \quad \varepsilon_d = \begin{cases} 1 & \text{if } d \equiv 1 \pmod 4, \\ \mathrm{i} & \text{if } d \equiv 3 \pmod 4. \end{cases}$$

Then f admits Fourier expansion

$$f(z) = \sum_{n\ge 1} a(n)e(nz).$$

By using the Weil bound of Salié sums (Corollary 4.2),

$$|T(n,n;c)| \le c^{1/2}\tau(c).$$

Combining with Proposition 5.1 (take $\delta = 3/4 + \eta$ where η is sufficiently small), we obtain

$$a(n) \ll n^{k/2-1/4+\varepsilon}.$$

Remark 5.2. If $n = m^2$, $a(n) \ll n^{k/2-1/4}$ is best possible. For example, let χ modulo 4 be a Dirichlet character, $\chi(-1) = -1$. Let

$$\theta(z, 4) = \sum_{m=-\infty}^{\infty} \chi(m) m e(m^2 z), \quad z \in \mathbb{H}.$$

θ is a cusp form for $\Gamma_0(8)$ of weight $k = 3/2$. If n is not square, $a(n) = 0$; if n is a square, say $n = m^2$, then $a(n) = a(m^2) = 2\chi(m)m$, so $\left|a(m^2)\right| = m = \sqrt{n}$, and so $a(n) = n^{k/2-1/4}$.

Theorem 5.4. Let $k \geq 5/2$, $4 \mid N$, $f \in \mathcal{S}_k(\Gamma_0(N), \vartheta)$. Then for n square-free,

$$a(n) \ll n^{k/2-1/4-1/222} \tau(n) \log n.$$

where the implied constant depends on f.

Conjecture (Ramanujan Conjecture)**.**

$$a(n) \ll n^{(k-1)/2+\varepsilon}.$$

Proof of Theorem. Let p be any prime. Note that $\Gamma_0(N) \supset \Gamma_0(pN)$, we have $\mathcal{S}_k(\Gamma_0(N), \vartheta) \subset \mathcal{S}_k(\Gamma_0(pN), \vartheta)$. Let $\|\cdot\|$ (resp. $\|\cdot\|_{pN}$) be the norm in $\mathcal{S}_k(\Gamma_0(N), \vartheta)$ (resp. $\mathcal{S}_k(\Gamma_0(pN), \vartheta)$). Given

$$f(z) = \sum_{n \geq 1} a(n) e(nz) \in \mathcal{S}_k(\Gamma_0(N), \vartheta).$$

To normalize f to have norm 1, let $f_{pN} = f/\|f\|_{pN}$. Take an orthonormal basis $\mathcal{F}$ of $\mathcal{S}_k(\Gamma_0(pN), \vartheta)$ containing f_{pN}. From the Petersson Trace Formula (Theorem 3.2),

$$\frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \sum_{g_{pN} \in \mathcal{F}} |\widehat{g_{pN}}(n)|^2 = 1 + 2\pi \mathrm{i}^{-k} \sum_{\substack{c>0 \\ c \equiv 0 \pmod{pN}}} c^{-1} S(n, n; c) J_{k-1}\left(\frac{4\pi n}{c}\right).$$

Note that

$$\frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \frac{|a(n)|^2}{\|f\|_{pN}^2} = \frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \left|\widehat{f_{pN}}(n)\right|^2 \leq \frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \sum_{g_{pN} \in \mathcal{F}} |\widehat{g_{pN}}(n)|^2$$

since the Fourier coefficient $\widehat{f_{pN}}(n)$ of f_{pN} is of the form $a(n)/\|f\|_{pN}$. On the other hand, thanks to Theorem 2.5,

$$[\Gamma_0(N) : \Gamma_0(pN)] = \frac{[\Gamma_0(1) : \Gamma_0(pN)]}{[\Gamma_0(1) : \Gamma_0(N)]} = \frac{(pN) \prod_{\ell \mid pN} \left(1 + \frac{1}{\ell}\right)}{N \prod_{\ell \mid N} \left(1 + \frac{1}{\ell}\right)} = p\left(1 + \frac{1}{p}\right) = p + 1.$$

(Later we will choose P such that $N(\log n)^2 < P \leq n^{1/6}$, this implies that $p \nmid N$, so $[\Gamma_0(N) : \Gamma_0(pN)] = p + 1$ holds.) Then $\|f\|_{pN}^2 = [\Gamma_0(N) : \Gamma_0(pN)]\|f\|^2 = (p+1)\|f\|^2$. Consequently,

$$\frac{\Gamma(k-1)}{(4\pi n)^{k-1}} \frac{|a(n)|^2}{(p+1)\|f\|^2} \leq 1 + 2\pi \mathrm{i}^{-k} \sum_{\substack{c>0 \\ c \equiv 0 \pmod{pN}}} c^{-1} S(n, n; c) J_{k-1}\left(\frac{4\pi n}{c}\right).$$

Choose P such that $N(\log n)^2 < P \leq n^{1/6}$. It follows that

$$\sum_{\substack{p\sim P\\ p\nmid n}} (\log p)\frac{\Gamma(k-1)}{(4\pi n)^{k-1}}\frac{|a(n)|^2}{(p+1)\|f\|^2} \leq \sum_{\substack{p\sim P\\ p\nmid n}} \log p + 2\pi \mathrm{i}^{-k} \sum_{\substack{c>0\\ c\equiv 0 \pmod N}} \frac{S(n,n;c)}{c}\left(\sum_{\substack{p\sim P\\ p\nmid n\\ p\mid c}} \log p\right) J_{k-1}\left(\frac{4\pi n}{c}\right),$$

where $\sum_{p\sim P}$ denotes the sum over primes $P < p < 2P$. Let

$$\omega(c) = \sum_{\substack{p\sim P\\ p\nmid n\\ p\mid c}} \log p,$$

and

$$S = \sum_{\substack{c>0\\ c\equiv 0 \pmod N}} \frac{S(n,n;c)}{c}\omega(c) J_{k-1}\left(\frac{4\pi n}{c}\right).$$

From Mertens Theorem,

$$\sum_{p\leq P} \frac{\log p}{p} \sim \log P,$$

we have

$$\sum_{p\sim P} \frac{\log p}{p} \gg 1$$

since

$$\sum_{\substack{p\sim P\\ p\mid n}} \frac{\log p}{p} \leq \frac{\log P}{P}\cdot\frac{\log n}{\log P} = \frac{\log n}{P} \ll \frac{1}{\log n},$$

so

$$\sum_{\substack{p\sim P\\ p\nmid n}} (\log p)\frac{\Gamma(k-1)}{(4\pi n)^{k-1}}\frac{|a(n)|^2}{(p+1)\|f\|^2} \gg \frac{|a(n)|^2}{n^{k-1}} \sum_{\substack{p\sim P\\ p\nmid n}} \frac{\log p}{p+1} \gg \frac{|a(n)|^2}{n^{k-1}}.$$

On the other hand, Chebyshev Theorem gives

$$\sum_{p\leq x} \log p \asymp x,$$

which implies that

$$\sum_{\substack{p\sim P\\ p\nmid n}} \log p \leq \sum_{p\sim P} \log p \ll P.$$

Therefore

$$|a(n)| \ll n^{(k-1)/2}\left(P^{1/2} + |S|^{1/2}\right).$$

Now we claim: the key range in the c-sum in S is $n^{1-\delta} = C < c < D = n^{1+\delta}$. Write $S = S^\flat + S^\sharp + S^*$, where $S^\flat$ (resp. $S^\sharp, S^*$) denote the sum over $c \le C$ (resp. $c \ge D, C < c < D$). Recall that we have Weil bound of Kloosterman sum (Theorem 4.1)

$$|S(n,n;c)| \ll \gcd(n,c)^{1/2} c^{1/2} \tau(c),$$

and

$$|J_{k-1}(x)| \ll \min\{x^{k-1}, x^{-1/2}\} \ll \min\{x^{3/2}, x^{-1/2}\}$$

where $k \ge 5/2$. Moreover, if we write $c = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$, then

$$\omega(c) \le \sum_{p|c} \log p = \sum_{i=1}^{s} \log p_i \le \log c.$$

Thus,

$$\begin{aligned}
S^\flat &\ll \sum_{c \le C} \frac{\gcd(n,c)^{1/2} c^{1/2} \tau(c)}{c} \log c \left(\frac{n}{c}\right)^{-1/2} \\
&\ll n^{-1/2} \sum_{c \le C} \gcd(n,c)^{1/2} \tau(c) \log c \\
&\ll n^{-1/2} \sum_{c \le C} \left(\sum_{\substack{d|n \\ d|c}} d^{1/2} \right) \tau(c) \log c \\
&= n^{-1/2} \sum_{d|n} d^{1/2} \sum_{\substack{c \le C \\ d|c}} \tau(c) \log c \\
&= n^{-1/2} \sum_{d|n} d^{1/2} \sum_{m \le C/d} \tau(dm) \log(dm) \qquad (\text{Write } c = dm) \\
&\ll n^{-1/2} C \tau(n) (\log n)^2 \sum_{d|n} d^{-1/2}
\end{aligned}$$

since $\tau(dm) \le \tau(d)\tau(m) \le \tau(n)\tau(m)$, $\log(dm) \ll \log n$ and $\sum_{m \le x} \tau(m) \ll x \log x$. Then

$$\begin{aligned}
S^\flat &\ll n^{-1/2} C \tau(n) (\log n)^2 \prod_{p|n} (1 + p^{-1/2}) \\
&\ll n^{-1/2} C \tau(n) (\log n)^2 \tau(n) \qquad \left(1 + p^{-1/2} \le 2, \text{ and } \tau(n) = \prod_{p|n} 2\right) \\
&= n^{1/2-\delta} \left(\tau(n) \log n\right)^2 \qquad \left(C = n^{1-\delta}\right)
\end{aligned}$$

Similarly,

$$\begin{aligned}
S^\sharp &\ll n^{3/2}\sum_{c\geq D}\gcd(n,c)^{1/2}\tau(c)\log c\cdot c^{-2}\\
&\leq n^{3/2}\sum_{d|n}d^{1/2}\sum_{\substack{c\geq D\\ d|c}}\tau(c)\log c\cdot c^{-2}\\
&= n^{3/2}\sum_{d|n}d^{-3/2}\sum_{m\geq D/d}\tau(dm)\log(dm)\cdot(dm)^{-2}\quad(\text{Write } c=dm)\\
&\leq n^{3/2}\tau(n)\sum_{d|n}d^{-3/2}\sum_{m\geq D/d}\tau(m)\log(dm)m^{-2}.
\end{aligned}$$

Using the estimate

$$\sum_{m\leq x}\tau(m)\ll x\log x,$$

for $j=0,1,2,\cdots$ we have

$$\sum_{2^jD/d\leq m<2^{j+1}D/d}\tau(m)\leq\sum_{m<2^{j+1}D/d}\tau(m)\ll\left(2^{j+1}D/d\right)\log\left(2^{j+1}D/d\right).$$

It follows that

$$\begin{aligned}
\sum_{2^jD/d\leq m<2^{j+1}D/d}\tau(m)\log(dm)m^{-2} &\ll\left(2^jD/d\right)^{-2}\log\left(2^{j+1}D\right)\left(2^{j+1}D/d\right)\log\left(2^{j+1}D/d\right)\\
&= 2^{-j}\frac{d}{D}\log\left(2^{j+1}D\right)\log\left(2^{j+1}D/d\right).
\end{aligned}$$

Consequently

$$\begin{aligned}
\sum_{m\geq D/d}\tau(m)\log(dm)m^{-2} &\ll\frac{d}{D}\sum_{j\geq0}2^{-j}\left(\log(2^{j+1}D)\right)^2\\
&\ll\frac{d}{D}\left((\log D)^2\sum_{j\geq0}2^{-j}+\sum_{j\geq0}2^{-j}(j+1)^2\right)\\
&\ll\frac{d}{D}(\log D)^2,
\end{aligned}$$

which implies that

$$S^\sharp\ll n^{3/2}D^{-1}\tau(n)(\log n)^2\sum_{d|n}d^{-1/2}\ll D^{-1}n^{3/2}(\tau(n)\log n)^2.$$

It only suffices to estimate

$$S^*=\sum_{\substack{C<c<D\\ c\equiv0\pmod N}}\omega(c)c^{-1}S(n,n;c)J_{k-1}\left(\frac{4\pi n}{c}\right).$$

Write $c = qr$, where q denotes the largest factor of c coprime with nN, so $4 \mid N \mid r$ and $r \mid (nN)^\infty$ (that is, every prime factor of r divides nN). Here q is good factor and r is bad factor. It follows that

$$S^* = \sum_{\substack{r \\ r \mid (nN)^\infty \\ N \mid r}} r^{-1} \sum_{\substack{q \\ C<qr<D \\ \gcd(q,nN)=1}} \omega(c) q^{-1} S(n,n;qr) J_{k-1}\left(\frac{4\pi n}{qr}\right).$$

From the definition

$$\omega(c) = \sum_{\substack{p \sim P \\ p \nmid n \\ p \mid c}} \log p,$$

but the prime factor of r are all from nN, and $p \nmid N$ by $p > P > N$, we then obtain $\omega(c) = \omega(q)$. Let

$$T_r = \sum_{\substack{q \\ C<qr<D \\ \gcd(q,nN)=1}} \omega(q) q^{-1} S(n,n;qr) J_{k-1}\left(\frac{4\pi n}{qr}\right),$$

then $S^* = \sum_r r^{-1} T_r$.

We claim: it only suffices to treat r at $r < R = n^{2\delta}$. Indeed, assume $r > R$. Since $r \mid (nN)^\infty$ and $\gcd(q, nN) = 1$, immediately we have $\gcd(q, r) = 1$, $\gcd(n, qr) = \gcd(n, r)$ and $\tau(qr) = \tau(q)\tau(r)$. Recall that the Weil bound of Kloosterman sum (Theorem 4.1) gives

$$S(n,n;qr) \ll (qr)^{1/2} \tau(qr) \gcd(n,qr)^{1/2}.$$

Then

$$|S(n,n;qr)| \ll \gcd(n,r)^{1/2} q^{1/2} r^{1/2} \tau(q)\tau(r).$$

It follows that

$$\begin{aligned}
|T_r| &\ll \tau(r) \gcd(n,r)^{1/2} r^{1/2} \sum_{\substack{q \\ C<qr<D \\ \gcd(q,nN)=1}} q^{-1/2} \tau(q) \log q \left| J_{k-1}\left(\frac{4\pi n}{qr}\right) \right| \\
&\ll \tau(r) \gcd(n,r)^{1/2} r^{1/2} \sum_{\substack{q \\ C<qr<D \\ \gcd(q,nN)=1}} q^{-1/2} \tau(q) \log q \min\left\{ \left(\frac{n}{qr}\right)^{-1/2}, \left(\frac{n}{qr}\right)^{3/2} \right\} \\
&\ll \tau(r) \gcd(n,r)^{1/2} r^{1/2} \sum_{\substack{q \\ C<qr<D \\ \gcd(q,nN)=1}} \tau(q) \log q \min\left\{ \left(\frac{r}{n}\right)^{1/2}, \frac{n^{3/2}}{q^2 r^{3/2}} \right\}.
\end{aligned}$$

For the case $r \le n$, note that

$$\left(\frac{r}{n}\right)^{1/2} = \frac{n^{3/2}}{q^2 r^{3/2}}$$

implies that $q = n/r$. It follows that

$$|T_r| \ll \tau(r)\gcd(n,r)^{1/2}r^{1/2}\left(\left(\frac{r}{n}\right)^{1/2}\sum_{\substack{q\le n/r\\ C<qr<D\\ \gcd(q,nN)=1}}\tau(q)\log q + \frac{n^{3/2}}{r^{3/2}}\sum_{\substack{q\ge n/r\\ C<qr<D\\ \gcd(q,nN)=1}}\frac{\tau(q)\log q}{q^2}\right)$$

$$\ll \tau(r)\gcd(n,r)^{1/2}r^{1/2}\left(\left(\frac{r}{n}\right)^{1/2}\sum_{q\le n/r}\tau(q)\log q + \frac{n^{3/2}}{r^{3/2}}\sum_{q\ge n/r}\frac{\tau(q)\log q}{q^2}\right)$$

From the fact

$$\sum_{m\le x}\tau(m)\log m \ll x(\log x)^2$$

and

$$\sum_{m\ge x}\frac{\tau(m)\log m}{m^2} \ll x^{-1}(\log x)^2,$$

we have

$$|T_r| \ll \tau(r)\gcd(n,r)^{1/2}n^{1/2}(\log n)^2.$$

For the case $r > n$, $q > n/r$ always holds. So for all $q \ge 1$ we have

$$\min\left\{\left(\frac{r}{n}\right)^{1/2}, \frac{n^{3/2}}{q^2r^{3/2}}\right\} = \frac{n^{3/2}}{q^2r^{3/2}}.$$

So

$$|T_r| \ll \tau(r)\gcd(n,r)^{1/2}r^{1/2}\frac{n^{3/2}}{r^{3/2}}\sum_{q\ge 1}\frac{\tau(q)\log q}{q^2} \ll \tau(r)\gcd(n,r)^{1/2}\frac{n^{3/2}}{r} \le \tau(r)\gcd(n,r)^{1/2}n^{1/2}.$$

Consequently, whether $r \le n$ or $r > n$, we always have

$$|T_r| \ll \tau(r)\gcd(n,r)^{1/2}n^{1/2}(\log n)^2.$$

Hence,

$$\sum_{r>R} r^{-1}|T_r| \ll n^{1/2}(\log n)^2\sum_{\substack{R<r<D\\ r|(nN)^\infty}}\tau(r)\gcd(n,r)^{1/2}r^{-1},$$

since $C < qr < D$ and $q \ge 1$ implies $r < D$. Note that

$$\sum_{\substack{R<r<D\\ r|(nN)^\infty}}\tau(r)\gcd(n,r)^{1/2}r^{-1} \le R^{-1/2}\sum_{r|(nN)^\infty}\tau(r)\gcd(n,r)^{1/2}r^{-1/2}.$$

Since $r = \prod_{p|nN} p^{\nu_p}$ and $r \mapsto \tau(r)\gcd(n,r)^{1/2}r^{-1/2}$ is multiplicative,

$$\sum_{r|(nN)^\infty}\tau(r)\gcd(n,r)^{1/2}r^{-1/2} = \prod_{p|nN}\sum_{\nu\ge 0}\tau(p^\nu)\gcd(n,p^\nu)^{1/2}p^{-\nu/2}.$$

If $p \mid N$, since N is fixed,

$$\prod_{p|N} \sum_{\nu \geq 0} \tau(p^\nu) \gcd(n, p^\nu)^{1/2} p^{-\nu/2} \ll_N 1.$$

If $p \mid n$ and $p \nmid N$, since n is square-free, $\gcd(n, p^\nu) = p$ where $\nu \geq 1$. It follows that

$$\sum_{\nu \geq 0} \tau(p^\nu) \gcd(n, p^\nu)^{1/2} p^{-\nu/2} = 1 + \sum_{\nu \geq 1} (\nu + 1) p^{1/2} p^{-\nu/2} \leq 4$$

as $p \geq 19$. Thus

$$\prod_{\substack{p|n \\ p \nmid N}} \sum_{\nu \geq 0} \tau(p^\nu) \gcd(n, p^\nu)^{1/2} p^{-\nu/2} \ll_N 4^{\omega(n)} = \tau(n)^2.$$

Consequently

$$\sum_{r > R} r^{-1} |T_r| \ll R^{-1/2} n^{1/2} (\log n)^2 \tau(n)^2.$$

Hence it remains to treat r with $r < R$. We suppress the r-dependence in the notation (formally, one may think of r as fixed, or $r = 1$ after renormalizing the notation).

Note that $r \mid (nN)^\infty$ and $\gcd(q, nN) = 1$ implies that $\gcd(q, r) = 1$, the twisted multiplicativity of Kloosterman sum gives

$$\begin{aligned}
T_r &= \sum_{\substack{q \\ C < qr < D \\ \gcd(q, nN) = 1}} \omega(q) q^{-1} S\left(n\overline{q}, n\overline{q}; r\right) T\left(n\overline{r}, n\overline{r}; q\right) J_{k-1}\left(\frac{4\pi n}{qr}\right) \\
&= \sum_{s \pmod r}^{*} S\left(n\overline{s}, n\overline{s}; r\right) \sum_{\substack{q \equiv s \pmod r \\ C < qr < D \\ \gcd(q, nN) = 1}} \omega(q) q^{-1} T\left(n\overline{r}, n\overline{r}; q\right) J_{k-1}\left(\frac{4\pi n}{qr}\right) \\
&\ll \sum_{s \pmod r}^{*} \left|S\left(n\overline{s}, n\overline{s}; r\right)\right| \left| \sum_{\substack{q \equiv s \pmod r \\ C < qr < D \\ \gcd(q, nN) = 1}} \omega(q) q^{-1} T\left(n\overline{r}, n\overline{r}; q\right) J_{k-1}\left(\frac{4\pi n}{qr}\right) \right| \\
&\ll \sum_{s \pmod r}^{*} r \left| \sum_{\substack{q \equiv s \pmod r \\ C < qr < D \\ \gcd(q, nN) = 1}} \omega(q) q^{-1} T\left(n\overline{r}, n\overline{r}; q\right) J_{k-1}\left(\frac{4\pi n}{qr}\right) \right|.
\end{aligned}$$

From Corollary 4.3,

$$T\left(n\overline{r}, n\overline{r}; q\right) = \varepsilon_q \left(\frac{nr}{q}\right) q^{1/2} \sum_{\substack{ab = q \\ \gcd(a, b) = 1}} e\left(2n\overline{r}\left(\frac{\overline{a}}{b} - \frac{\overline{b}}{a}\right)\right).$$

If $\gcd(a,b)=1$, there exists $\overline{a},\overline{b}$ such that $a\overline{a}\equiv 1 \pmod b$ and $b\overline{b}\equiv 1 \pmod a$, then $a\overline{a}+b\overline{b}\equiv 1 \pmod a$ and $a\overline{a}+b\overline{b}\equiv 1 \pmod b$, and so $a\overline{a}+b\overline{b}\equiv 1 \pmod{ab}$. That is, $1/(ab)\equiv \overline{b}/a+\overline{a}/b \pmod 1$. Therefore

$$e\left(\frac{1}{ab}\right)=e\left(\frac{\overline{b}}{a}\right)e\left(\frac{\overline{a}}{b}\right).$$

Using $\gcd(a,b)=1$ and $\gcd(r,q)=1$, we have $\gcd(ar,b)=1$, similarly

$$e\left(\frac{\overline{ar}}{b}\right)=e\left(-\frac{\overline{b}}{ar}\right)e\left(\frac{1}{abr}\right).$$

So

$$\begin{aligned}\sum_{\substack{ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{ar}}{b}-\frac{\overline{br}}{a}\right)\right) &= \sum_{\substack{a<b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{ar}}{b}-\frac{\overline{br}}{a}\right)\right)+\sum_{\substack{a\geq b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{ar}}{b}-\frac{\overline{br}}{a}\right)\right)\\ &= \sum_{\substack{a<b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{-\overline{b}}{ar}-\frac{\overline{br}}{a}\right)+\frac{2n}{abr}\right)+\sum_{\substack{a\geq b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{ar}}{b}+\frac{\overline{a}}{br}\right)-\frac{2n}{abr}\right)\\ &= \sum_{\substack{a<b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{-\overline{b}}{ar}-\frac{\overline{br}}{a}\right)+\frac{2n}{abr}\right)+\sum_{\substack{a<b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\left(\frac{\overline{b}}{ar}+\frac{\overline{br}}{a}\right)-\frac{2n}{abr}\right)\\ &= 2\,\mathrm{Re}\sum_{\substack{a<b\\ ab=q\\ \gcd(a,b)=1}} e\left(2n\frac{\overline{b}}{ar}+2n\frac{\overline{br}}{a}-\frac{2n}{abr}\right).\end{aligned}$$

($q=1$ gives no contribution since $\omega(1)=0$.) Consequently,

$$\begin{aligned}&\sum_{\substack{q\equiv s \pmod r\\ C<qr<D\\ \gcd(q,nN)=1}} \omega(q)q^{-1}T\left(n\overline{r},n\overline{r};q\right)J_{k-1}\left(\frac{4\pi n}{qr}\right)\\ &= \sum_{\substack{a<b,C<abr<D\\ ab\equiv s \pmod r\\ \gcd(a,b)=1,\gcd(ab,nN)=1}} 2\varepsilon_{ab}\left(\frac{nr}{ab}\right)\mathrm{Re}\left(\omega(ab)\frac{1}{ab}(ab)^{1/2}e\left(\frac{2n\overline{b}}{ar}+\frac{2n\overline{br}}{a}-\frac{2n}{abr}\right)J_{k-1}\left(\frac{4\pi n}{abr}\right)\right)\\ &= \varepsilon_s\left(\frac{r}{s}\right)\left(\frac{2r}{n}\right)^{1/2}\mathrm{Re}\sum_{\substack{a<b,C<abr<D\\ ab\equiv s \pmod r\\ \gcd(a,b)=1,\gcd(ab,nN)=1}} \omega(ab)\left(\frac{n}{ab}\right)e\left(\frac{2n\overline{b}}{ar}+\frac{2n\overline{br}}{a}\right)\left(\frac{2n}{abr}\right)^{1/2}e\left(\frac{-2n}{abr}\right)J_{k-1}\left(\frac{4\pi n}{abr}\right)\\ &= \varepsilon_s\left(\frac{r}{s}\right)\left(\frac{2r}{n}\right)^{1/2}\mathrm{Re}\sum_{\substack{a<b,C<abr<D\\ ab\equiv s \pmod r\\ \gcd(a,b)=1,\gcd(ab,nN)=1}} \omega(ab)\left(\frac{n}{ab}\right)e\left(\frac{2n\overline{b}}{ar}+\frac{2n\overline{br}}{a}\right)j\left(\frac{2n}{abr}\right),\end{aligned}$$

where

$$j\left(\frac{2n}{abr}\right)=\left(\frac{2n}{abr}\right)^{1/2}e\left(\frac{-2n}{abr}\right)J_{k-1}\left(\frac{4\pi n}{abr}\right),$$

that is,

$$j(x) = x^{1/2}e(-x)J_{k-1}(2\pi x).$$

Indeed, for the third equality, from $\gcd(ab, nN) = 1$ and $4 \mid N$ we have $2 \nmid ab$. On the other hand $ab \equiv s \pmod r$ and $4 \mid r$, we obtain $ab \equiv s \pmod 4$, then $\varepsilon_{ab} = \varepsilon_s$ and

$$\left(\frac{nr}{ab}\right) = \left(\frac{n}{ab}\right)\left(\frac{r}{ab}\right) = \left(\frac{n}{ab}\right)\left(\frac{r}{s}\right).$$

Hence,

$$\begin{aligned}
&\sum_{\substack{q\equiv s \pmod r\\ C<qr<D\\ \gcd(q,nN)=1}} \omega(q)q^{-1}T\left(n\overline{r}, n\overline{r}; q\right) J_{k-1}\left(\frac{4\pi n}{qr}\right)\\
&\ll \left(\frac{r}{n}\right)^{1/2} \sum_{a<A=(D/r)^{1/2}} \left(\frac{n}{a}\right) \sum_{\substack{a<b<D/(ar)\\ ab\equiv s \pmod r\\ C<abr}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right) j\left(\frac{2n}{abr}\right)\\
&= \left(\frac{r}{n}\right)^{1/2} \sum_{a<A} \left(\frac{n}{a}\right) \sum_{\substack{\max\{a,C/(ar)\}<b\le D/(ar)\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right) j\left(\frac{2n}{abr}\right).
\end{aligned}$$

Summation by parts,

$$\begin{aligned}
&\left(\frac{r}{n}\right)^{1/2} \sum_{a<A} \left(\frac{n}{a}\right) \sum_{\substack{\max\{a,C/(ar)\}<b\le D/(ar)\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right) j\left(\frac{2n}{abr}\right)\\
&= \left(\frac{r}{n}\right)^{1/2} \sum_{a<A} \left(\frac{n}{a}\right) \int_{\max\{a,C/(ar)\}}^{D/(ar)} j\left(\frac{2n}{a\xi r}\right) \mathrm{d}\left(\sum_{\substack{\max\{a,C/(ar)\}<b\le \xi\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right)\right).
\end{aligned}$$

We have

$$\begin{aligned}
&\int_{\max\{a,C/(ar)\}}^{D/(ar)} j\left(\frac{2n}{a\xi r}\right) \mathrm{d}\left(\sum_{\substack{\max\{a,C/(ar)\}<b\le \xi\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right)\right)\\
&= \sum_{\substack{\max\{a,C/(ar)\}<b\le D/(ar)\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right) j\left(\frac{2n}{a\frac{D}{ar}r}\right)\\
&\quad - \int_{\max\{a,C/(ar)\}}^{D/(ar)} \sum_{\substack{\max\{a,C/(ar)\}<b\le \xi\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right) \frac{\partial}{\partial\xi} j\left(\frac{2n}{a\xi r}\right) \mathrm{d}\xi\\
&\ll \max_B \left|\sum_{\substack{\max\{a,C/(ar)\}<b\le B\\ ab\equiv s \pmod r}} \omega(ab)\left(\frac{n}{b}\right) e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right)\right| \left(\left|j\left(\frac{2n}{a\frac{D}{ar}r}\right)\right| + \int_{\max\{a,C/(ar)\}}^{D/(ar)} \left|\frac{\partial}{\partial\xi} j\left(\frac{2n}{a\xi r}\right)\right| \mathrm{d}\xi\right).
\end{aligned}$$

The chain rule gives

$$\left|\frac{\partial}{\partial\xi}j\left(\frac{2n}{a\xi r}\right)\right| = \left|j'\left(\frac{2n}{a\xi r}\right)\left(-\frac{2n}{ar\xi^2}\right)\right| \ll \left|j'\left(\frac{2n}{a\xi r}\right)\right|\frac{n}{ar\xi^2}.$$

Using the estimate $j'(x) \ll 1$,

$$\int_{\max\{a,C/(ar)\}}^{D/(ar)}\left|\frac{\partial}{\partial\xi}j\left(\frac{2n}{a\xi r}\right)\right|\mathrm{d}\xi \ll \frac{n}{ar}\int_{\max\{a,C/(ar)\}}^{D/(ar)}\xi^{-2}\mathrm{d}\xi \ll \frac{n}{ar\max\{a,C/(ar)\}} \le \frac{n}{C}.$$

Moreover $\max\{a,C/(ar)\} \ge C/(ar)$ and $|j(2n/D)| \ll 1$. Combine with $n/C = n^{\delta} \ge 1$ we get $|j(2n/D)| \ll n/C$. It follows that

$$\left|j\left(\frac{2n}{D}\right)\right| + \int_{\max\{a,C/(ar)\}}^{D/(ar)}\left|\frac{\partial}{\partial\xi}j\left(\frac{2n}{a\xi r}\right)\right|\mathrm{d}\xi \ll \frac{n}{C}.$$

Therefore,

$$\sum_{\substack{q\equiv s \pmod r\\ C<qr<D\\ \gcd(q,nN)=1}}\omega(q)q^{-1}T\left(n\overline{r},n\overline{r};q\right)J_{k-1}\left(\frac{4\pi n}{qr}\right)$$

$$\ll C^{-1}(nr)^{1/2}\sum_{\substack{a<A\\ \gcd(a,nr)=1}}\max_{B}\left|\sum_{\substack{\max\{a,C/(ar)\}<b<B\\ ab\equiv s \pmod r\\ \gcd(a,b)=1,\gcd(b,nN)=1}}\omega(ab)\left(\frac{n}{b}\right)e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right)\right|.$$

By $\gcd(a,b)=1$, $\omega(ab)=\omega(a)+\omega(b)$, define

$$V_a = \sum_{\substack{\max\{a,C/(ar)\}<b<B\\ ab\equiv s \pmod r\\ \gcd(a,b)=1,\gcd(b,nN)=1}}\omega(b)\left(\frac{n}{b}\right)e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right)$$

and

$$V_a' = \sum_{\substack{\max\{a,C/(ar)\}<b<B\\ ab\equiv s \pmod r\\ \gcd(a,b)=1,\gcd(b,nN)=1}}\left(\frac{n}{b}\right)e\left(2n(1+r\overline{r})\frac{\overline{b}}{ar}\right).$$

Consequently,

$$\sum_{\substack{q\equiv s \pmod r\\ C<qr<D\\ \gcd(q,nN)=1}}\omega(q)q^{-1}T\left(n\overline{r},n\overline{r};q\right)J_{k-1}\left(\frac{4\pi n}{qr}\right) \ll C^{-1}(nr)^{1/2}\sum_{\substack{a<A\\ \gcd(a,nr)=1}}\left(|V_a|+\omega(a)|V_a'|\right).$$

First we estimate V_a. Write $b=p\ell$, then

$$V_a = \sum_{\substack{p\sim P\\ p\nmid anN}}(\log p)\left(\frac{n}{p}\right)\sum_{\substack{\max\{a,C/(ar)\}/p<\ell<B/p\\ ap\ell\equiv s \pmod r\\ \gcd(a,\ell)=1,\gcd(\ell,nN)=1}}\left(\frac{n}{\ell}\right)e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right).$$

If $a \le P$. For fixed p, note that $\gcd(a,b)=1$, $p \mid b$, $\gcd(ab,nN)=1$ and $r \mid (nN)^\infty$ implies that $\gcd(a,r)=1$ and $\gcd(p,r)=1$, then ap is invertible modulo r. So $ap\ell \equiv s \pmod r$ iff $\ell = s\overline{ap} \pmod r$. Since s is invertible modulo r, we have $\gcd(\ell_0,r)=1$. Since $N \mid r$, if $\ell \equiv \ell_0 \pmod r$, immediately we have $\gcd(\ell,N)=1$. Next we deal with $\gcd(a,\ell)=1$. Fix $\ell \equiv \ell_0 \pmod r$. Since $\gcd(a,r)=1$, there exist a residue classes modulo ar satisfying $\lambda \equiv \ell_0 \pmod r$. Set $\Lambda_{a,p} = \{\lambda \pmod{ar} : \lambda \equiv \ell_0 \pmod r, \gcd(\lambda,a)=1\}$, clearly we have $\#\Lambda_{a,p} \le a$. On the other hand if ℓ belongs to some fixed residue class $\ell \equiv \lambda \pmod{ar}$, then $e\left(2n(1+r\overline{r})\overline{p\ell}/(ar)\right)$ is fixed. Consequently

$$\sum_{\substack{\max\{a,C/(ar)\}/p<\ell<B/p \\ ap\ell\equiv s \pmod r \\ \gcd(a,\ell)=1,\gcd(\ell,nN)=1}} \left(\frac{n}{\ell}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) = \sum_{\lambda\in\Lambda_{a,p}} e\left(2n(1+r\overline{r})\frac{\overline{p\lambda}}{ar}\right) \sum_{\substack{\max\{a,C/(ar)\}/p<\ell<B/p \\ \ell\equiv\lambda \pmod{ar}}} \left(\frac{n}{\ell}\right).$$

Now we only need to estimate

$$\sum_{\substack{\max\{a,C/(ar)\}/p<\ell<B/p \\ \ell\equiv\lambda \pmod{ar}}} \left(\frac{n}{\ell}\right).$$

Write $\ell = \lambda + arm$. From $\gcd(a,n)=1$ and $r \mid (nN)^\infty$, all of the common prime factors come from r. Here we only consider $r < R = n^{2\delta}$ where $\delta < 1$ and n square-free, then $r < n$ for n sufficiently large, and so r cannot be divided by all of the prime factors of n, it follows that $n/\gcd(n,r) > 1$. In the residue class $\ell \equiv \lambda \pmod{ar}$, the Pólya–Vinogradov inequality gives

$$\sum_{\substack{\max\{a,C/(ar)\}/p<\ell<B/p \\ \ell\equiv\lambda \pmod{ar}}} \left(\frac{n}{\ell}\right) \ll n^{1/2}\log n.$$

Hence

$$\sum_{\substack{\max\{a,C/(ar)\}/p<\ell<B/p \\ ap\ell\equiv s \pmod r \\ \gcd(a,\ell)=1,\gcd(\ell,nN)=1}} \left(\frac{n}{\ell}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) \ll an^{1/2}\log n.$$

Moreover

$$\sum_{\substack{p\sim P \\ p\nmid anN}} (\log p) \le \sum_{p\sim P} \log p \ll P$$

implies that

$$V_a \ll aPn^{1/2}\log n \quad (a \le P).$$

If $P < a < A$ we regard V_a as a bilinear form in ℓ and p. Let $L = B/P$, $P_1 = \max\{P, \max\{a, C/(ar)\}/\ell\}$

and $P_2 = \min\{2P, B/\ell\}$. Since $P < p < 2P$ and $\ell < B/p$, we have $\ell < L$. Then

$$
\begin{aligned}
|V_a|^2 &= \left| \sum_{\ell<L} \left(\frac{n}{\ell}\right) \sum_{\substack{P_1<p<P_2 \\ ap\ell\equiv s \pmod r}} (\log p) \left(\frac{n}{p}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) \right|^2 \\
&\le \left(\sum_{\ell<L} 1^2\right) \sum_{\ell<L} \left| \left(\frac{n}{\ell}\right) \sum_{\substack{P_1<p<P_2 \\ ap\ell\equiv s \pmod r}} (\log p) \left(\frac{n}{p}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) \right|^2 \\
&\le L \sum_{\ell<L} \left| \left(\frac{n}{\ell}\right) \sum_{\substack{P_1<p<P_2 \\ ap\ell\equiv s \pmod r}} (\log p) \left(\frac{n}{p}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) \right|^2 \\
&\le L \sum_{\ell<L} \left| \sum_{\substack{P_1<p<P_2 \\ ap\ell\equiv s \pmod r}} (\log p) \left(\frac{n}{p}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) \right|^2 \\
&= L \sum_{\ell<L} \sum_{\substack{P_1<p<P_2 \\ ap\ell\equiv s \pmod r}} \sum_{\substack{P_1<p'<P_2 \\ ap'\ell\equiv s \pmod r}} (\log p)(\log p') \left(\frac{n}{p}\right)\left(\frac{n}{p'}\right) e\left(2n(1+r\overline{r})\frac{\overline{p\ell}}{ar}\right) e\left(-2n(1+r\overline{r})\frac{\overline{p'\ell}}{ar}\right).
\end{aligned}
$$

From $ap\ell \equiv s \pmod r$ and $ap'\ell \equiv s \pmod r$, we have $a\ell(p-p') \equiv 0 \pmod r$. Since $\gcd(ab, nN) = 1$, $b = p\ell$ and $r \mid (nN)^\infty$, then $\gcd(a, r) = \gcd(\ell, r) = 1$. Consequently $a\ell(p-p') \equiv 0 \pmod r$ implies that $p \equiv p' \pmod r$. That is, $ap\ell \equiv s \pmod r$ and $ap'\ell \equiv s \pmod r$ are equivalent under condition $p \equiv p' \pmod r$. Let

$$
I(p,p') = \left\{ \ell : \begin{array}{l} \dfrac{\max\{a, C/(ar)\}}{p} < \ell < \dfrac{B}{p}, \\ \dfrac{\max\{a, C/(ar)\}}{p'} < \ell < \dfrac{B}{p'}, \\ ap\ell \equiv s \pmod r, \\ \gcd(\ell, ar) = 1, \\ \gcd(a, p\ell) = \gcd(a, p'\ell) = 1. \end{array} \right\}.
$$

Therefore

$$
|V_a|^2 \ll L(\log P)^2 \sum_{\substack{P<p,p'<2P \\ p\equiv p' \pmod r}} \left| \sum_{\ell\in I(p,p')} e\left(2n(1+r\overline{r})\frac{(\overline{p}-\overline{p'})\,\overline{\ell}}{ar}\right) \right|.
$$

When $p = p'$, $\overline{p} - \overline{p'} = 0$, so

$$
|V_a|^2 \ll L(\log P)^2 \sum_{\substack{P<p,p'<2P \\ p=p'}} \left| \sum_{\ell\in I(p,p')} 1 \right| \ll L(\log P)^2 \sum_{\substack{P<p,p'<2P \\ p=p'}} L \ll L^2 P (\log P)^2.
$$

When $p \neq p'$, from $p \equiv p' \pmod{r}$ we have $r \mid p - p'$ Clearly $\overline{p} - \overline{p'} \equiv (p' - p)\overline{pp'} \pmod{ar}$, so

$$e\left(2n(1+r\overline{r})\frac{\left(\overline{p}-\overline{p'}\right)\overline{\ell}}{ar}\right) = e\left(2n(1+r\overline{r})\frac{(p'-p)\overline{pp'\ell}}{ar}\right) = e\left(2n(1+r\overline{r})\frac{(p'-p)}{r}\frac{\overline{pp'\ell}}{a}\right).$$

From $\gcd(a, nr) = 1$, we have $\gcd(a, n) = \gcd(a, r) = 1$. On the other hand $\gcd(a, b) = 1$ and $b = p\ell$ implies that $\gcd(p, a) = \gcd(p', a) = 1$. Therefore $\overline{pp'}$ is invertible modulo a. Note that $1 + r\overline{r} \equiv 2 \pmod{a}$ (since $r\overline{r} \equiv 1 \pmod{a}$) and $4 \mid N$, $\gcd(a, nN) = 1$ implies that $2 \nmid a$, it follows that $1 + r\overline{r}$ is invertible modulo a. Hence

$$\gcd\left(2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}, a\right) = \gcd\left(\frac{p'-p}{r}, a\right) = \gcd(p'-p, a).$$

Now we estimate

$$\sum_{\ell \in I(p,p')} e\left(2n(1+r\overline{r})\frac{(p'-p)}{r}\frac{\overline{pp'\ell}}{a}\right).$$

From $ap\ell = s \pmod{r}$ and $\gcd(ap, r) = 1$, all ℓ satisfying this congruence lie in a fixed residue class modulo r. Let ℓ vary in the interval given by $I(p, p')$. The length of this interval $\leq L$, since $p, p' > P$, and the upper bounds are at most $B/P = L$. Divide this interval into several complete intervals of length ar, and at most two endpoint intervals of length less than ar. First consider a complete interval. Since $\gcd(a, r) = 1$, in such a complete interval, the integers ℓ satisfying $\ell \equiv s\overline{ap} \pmod{r}$, each residue class modulo a appearing exactly once. Therefore, when we add the condition $\gcd(\ell, a) = 1$ in such a complete interval, $\ell \bmod a$ runs exactly once through all invertible residue classes modulo a. Hence the sum over the complete interval equals

$$\begin{aligned}
\left|\sideset{}{^*}\sum_{x \pmod{a}} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\overline{x}}{a}\right)\right| &= \left|\sideset{}{^*}\sum_{x \pmod{a}} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}x}{a}\right)\right| \\
&= \left|\mu\left(\frac{a}{\gcd\left(a, 2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\right)}\right)\frac{\varphi(a)}{\varphi\left(\frac{a}{\gcd\left(a, 2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\right)}\right)}\right| \\
&\leq \frac{\varphi(a)}{\varphi\left(\frac{a}{\gcd\left(a, 2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\right)}\right)}.
\end{aligned}$$

Write

$$a = \prod_q q^{\alpha_q}, \quad \frac{a}{\gcd\left(a, 2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\right)} = \prod_q q^{\beta_q},$$

it follows that

$$\frac{\varphi(a)}{\varphi\left(\frac{a}{\gcd\left(a, 2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\right)}\right)} = \prod_q \frac{\varphi(q^{\alpha_q})}{\varphi(q^{\beta_q})} \leq \prod_q q^{\alpha_q - \beta_q} = \gcd\left(a, 2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\right) = \gcd(p'-p, a).$$

Moreover

$$\left|\sum_{\ell\in I(p,p')} 1\right| \ll L$$

implies that

$$\frac{1}{a}\left(\sum_{x \pmod a}^{*} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\overline{x}}{a}\right)\right)\left(\sum_{\ell\in I(p,p')} 1\right) \ll \gcd(p-p',a)a^{-1}L \le \gcd(p-p',a)a^{-1}L\log L.$$

Next we estimate the endpoint intervals. Take arbitrary endpoint interval with length $< ar$ and only consider ℓ with $\ell \equiv s\overline{ap} \pmod r$. Since length of the interval $< ar$, there exists at most a many such ℓ. By the orthogonality of additive characters,

$$\sum_{\substack{\ell \text{ in the endpoint interval}\\ \ell\equiv s\overline{ap} \pmod r\\ \gcd(\ell,a)=1}} e\left(2n(1+r\overline{r})\frac{(p'-p)}{r}\frac{\overline{pp'\ell}}{a}\right)$$

$$= \sum_{\substack{\ell \text{ in the endpoint interval}\\ \ell\equiv s\overline{ap} \pmod r\\ \gcd(\ell,a)=1}} \left(\sum_{x \pmod a}^{*} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\overline{x}}{a}\right)\cdot\frac{1}{a}\sum_{m \pmod a} e\left(\frac{m(x-\ell)}{a}\right)\right)$$

$$= \frac{1}{a}\sum_{m \pmod a}\left(\sum_{x \pmod a}^{*} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\overline{x}+mx}{a}\right)\right)\left(\sum_{\substack{\ell \text{ in the endpoint interval}\\ \ell\equiv s\overline{ap} \pmod r}} e\left(-\frac{m\ell}{a}\right)\right).$$

For $m = 0$, the Ramanujan sum gives

$$\sum_{\substack{\ell \text{ in the endpoint interval}\\ \ell\equiv s\overline{ap} \pmod r\\ \gcd(\ell,a)=1}} e\left(2n(1+r\overline{r})\frac{(p'-p)}{r}\frac{\overline{pp'\ell}}{a}\right) \ll \gcd(p'-p,a)^{1/2}a^{1/2}\tau(a)\log a.$$

For $m \neq 0$, the Weil bound of Kloosterman sum (Theorem 4.1) gives

$$\left|\sum_{x \pmod a}^{*} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\overline{x}+mx}{a}\right)\right| \ll \gcd(p-p',a)^{1/2}a^{1/2}\tau(a).$$

On the other hand, when $-a/2 < m \le a/2$, we have

$$\left|\sum_{\ell\in I(p,p')} e\left(-\frac{m\ell}{a}\right)\right| \ll \sum_{\ell\in I(p,p')} 1 \ll L.$$

Also, all of the ℓ in the interval satisfies $\ell \equiv s\overline{ap} \pmod r$ forms a finite arithmetic series,

$$\left|\sum_{\ell\in I(p,p')} e\left(-\frac{m\ell}{a}\right)\right| \ll \frac{1}{|1-e\left(-mr/a\right)|} = \frac{a}{2|\sin(\pi mr/a)|} \ll \frac{a}{|m|}$$

implies that

$$\left|\sum_{\ell\in I(p,p')} e\left(-\frac{m\ell}{a}\right)\right| \ll \min\left\{L, \frac{a}{|m|}\right\}.$$

So

$$\sum_{\substack{m \pmod a\\ m\neq 0}} \left|\sum_{\ell\in I(p,p')} e\left(-\frac{m\ell}{a}\right)\right| \ll \sum_{1\le|m|\le a/2} \min\left\{L, \frac{a}{|m|}\right\} \ll a\log a.$$

Note that there are at most two endpoint intervals,

$$\frac{1}{a}\sum_{m \pmod a}\left(\sideset{}{^*}\sum_{x \pmod a} e\left(\frac{2n(1+r\overline{r})\frac{p'-p}{r}\overline{pp'}\overline{x}+mx}{a}\right)\right)\left(\sum_{\substack{\ell \text{ in the endpoint interval}\\ \ell\equiv s\overline{ap} \pmod r}} e\left(-\frac{m\ell}{a}\right)\right)$$

$$\ll \gcd(p-p',a)^{1/2}a^{1/2}\tau(a)\log a.$$

Hence

$$\sum_{\ell\in I(p,p')} e\left(2n(1+r\overline{r})\frac{\left(\overline{p}-\overline{p'}\right)\overline{\ell}}{ar}\right) \ll \gcd(p-p',a)a^{-1}L\log L + \gcd(p-p',a)^{1/2}a^{1/2}\tau(a)\log a.$$

Consequently here

$$|V_a|^2 \ll L(\log P)^2 \sum_{\substack{P<p,p'<2P\\ p\neq p'}} \left(\gcd(p-p',a)a^{-1}L\log L + \gcd(p-p',a)^{1/2}a^{1/2}\tau(a)\log a\right).$$

Now

$$\begin{aligned}\sum_{\substack{P<p,p'<2P\\ p\neq p'}} \gcd(p-p',a) &\le \sum_{\substack{P<p,p'<2P\\ p\neq p'}} \sum_{\substack{d|a\\ d|(p-p')}} d\\ &\le \sum_{d|a} d\#\{P<p,p'<2P : p\neq p', p\equiv p' \pmod d\}.\end{aligned}$$

Fix p prime where $P<p<2P$ and we count p' with $P<p'<2P$ and $p'\equiv p \pmod d$. From $p'\equiv p \pmod d$, p' belongs to one fixed residue class modulo d, In any interval of length P, the number of integers in a fixed residue class modulo d is at most $P/d+1$. Therefore, for each fixed p, the number of p' satisfying the condition is at most $\ll P/d+1$. On the other hand $\#\{P<p<2P\}\le P$. Hence

$$\#\{P<p,p'<2P : p\neq p', p\equiv p' \pmod d\} \le \#\{P<p,p'<2P : p\equiv p' \pmod d\} \ll \frac{P^2}{d}+P.$$

It follows that

$$\sum_{\substack{P<p,p'<2P\\ p\neq p'}} \gcd(p-p',a) \ll \sum_{\substack{d|a\\ d<P}} d\left(\frac{P^2}{d}+P\right) \ll P^2\tau(a).$$

Similarly

$$\sum_{\substack{P<p,p'<2P\\ p\neq p'}} \gcd(p-p',a)^{1/2} \leq \sum_{\substack{d\mid a\\ d<P}} d^{1/2}\left(\frac{P^2}{d}+P\right) \ll P^2\tau(a).$$

Thus

$$|V_a|^2 \ll a^{-1}L^2P^2\tau(a)(\log P)^2\log L + a^{1/2}LP^2\tau(a)^2(\log P)^2\log a \ll \left(L^2P + a^{-1}L^2P^2 + a^{1/2}LP^2\right)(\tau(a)\log n)^2,$$

then

$$\begin{aligned} V_a &\ll \left(LP^{1/2} + a^{-1/2}LP + a^{1/4}L^{1/2}P\right)\tau(a)\log n \\ &\ll \left(LP^{1/2} + a^{1/4}L^{1/2}P\right)\tau(a)\log n \qquad \left(P<a \text{ implies that } a^{-1/2}LP \leq LP^{1/2}\right) \\ &= \left(BP^{-1/2} + a^{1/4}B^{1/2}P^{1/2}\right)\tau(a)\log n. \qquad \left(L=\frac{B}{P}\right) \end{aligned}$$

The same estimate holds true for V_a' by similar arguments.

Recall that we have proved:

$$T_r \ll \sideset{}{^*}\sum_{s \pmod r} rC^{-1}(nr)^{1/2} \sum_{\substack{a<A\\ \gcd(a,nr)=1}} (|V_a| + \omega(a)|V_a'|),$$

From the above discussion,

$$\begin{aligned} \sum_{\substack{a<A\\ \gcd(a,nr)=1}} (|V_a| + \omega(a)|V_a'|) &= \sum_{\substack{a\leq P\\ \gcd(a,nr)=1}} (|V_a| + \omega(a)|V_a'|) + \sum_{\substack{P<a<A\\ \gcd(a,nr)=1}} (|V_a| + \omega(a)|V_a'|) \\ &\ll \sum_{a\leq P} aPn^{1/2}(\log n)^2 + \sum_{a\leq P} \log n \cdot an^{1/2}\log n + \left(DP^{-1/2} + D^{7/8}P^{1/2}\right)(\log n)^2 \\ &\ll n^{1/2}P^3(\log n)^2 + \left(DP^{-1/2} + D^{7/8}P^{1/2}\right)(\log n)^2. \end{aligned}$$

Hence

$$\begin{aligned} |T_r| &\ll \sideset{}{^*}\sum_{s \pmod r} rC^{-1}(nr)^{1/2}\left(n^{1/2}P^3 + DP^{-1/2} + D^{7/8}P^{1/2}\right)(\log n)^2 \\ &\ll r^2C^{-1}(nr)^{1/2}\left(n^{1/2}P^3 + DP^{-1/2} + D^{7/8}P^{1/2}\right)(\log n)^2. \end{aligned}$$

Now, choose $C = n^{110/111}, D = n^{112/111}$ and $P = n^{14/111}$. It follows that $n^{1/2}P^3 = n^{195/222} < n^{210/222} = DP^{-1/2}$. Thus

$$|T_r| \ll r^{5/2}C^{-1}\left(DP^{-1/2} + D^{7/8}P^{1/2}\right)n^{1/2}(\log n)^2,$$

and so

$$\begin{aligned}\sum_{\substack{r<R\\ r|(nN)^\infty}} r^{-1}|T_r| &\ll C^{-1}\left(DP^{-1/2}+D^{7/8}P^{1/2}\right)n^{1/2}(\log n)^2\sum_{\substack{r<R\\ r|(nN)^\infty}} r^{3/2}\\ &\le C^{-1}\left(DP^{-1/2}+D^{7/8}P^{1/2}\right)n^{1/2}(\log n)^2R^2\sum_{\substack{r<R\\ r|(nN)^\infty}} r^{-1/2}\\ &\ll R^2C^{-1}\left(DP^{-1/2}+D^{7/8}P^{1/2}\right)n^{1/2}(\tau(n)\log n)^2.\end{aligned}$$

Indeed we claim that

$$\sum_{r|(nN)^\infty} r^{-1/2} \ll_N \tau(n)^2.$$

Write $r=\prod_{p|nN}p^{\nu_p}$, then

$$\sum_{r|(nN)^\infty} r^{-1/2}=\sum_{\nu_p\ge 0}\prod_{p|nN}p^{-\nu_p/2}=\prod_{p|nN}\left(\sum_{\nu\ge 0}p^{-\nu/2}\right)=\prod_{p|nN}\frac{1}{1-p^{-1/2}}.$$

On the one hand

$$\prod_{p|N}\frac{1}{1-p^{-1/2}}\ll_N 1$$

since the left-hand side depends only on N. On the other hand,

$$\prod_{\substack{p|n\\ p\nmid N}}\frac{1}{1-p^{-1/2}}\le\prod_{\substack{p|n\\ p\nmid N}}\frac{1}{1-2^{-1/2}}<\prod_{\substack{p|n\\ p\nmid N}}4\le 4^{\omega(n)}.$$

Since n is square-free, $4^{\omega(n)}=\tau(n)^2$. Consequently

$$\sum_{r|(nN)^\infty} r^{-1/2} \ll_N \tau(n)^2.$$

Combining with

$$\begin{aligned}S^\flat &\ll Cn^{-1/2}(\tau(n)\log n)^2,\\ S^\sharp &\ll D^{-1}n^{3/2}(\tau(n)\log n)^2,\\ \sum_{r>R} r^{-1}|T_r| &\ll R^{-1/2}n^{1/2}(\tau(n)\log n)^2,\end{aligned}$$

we have

$$S\ll\left(Cn^{-1/2}+D^{-1}n^{3/2}+R^{-1/2}n^{1/2}+R^2C^{-1}\left(DP^{-1/2}+D^{7/8}P^{1/2}\right)n^{1/2}\right)(\tau(n)\log n)^2.$$

Furthermore choose $R=n^{2/111}$, immediately we have

$$S\ll n^{1/2-1/111}(\tau(n)\log n)^2.$$

From

$$|a(n)| \ll n^{(k-1)/2}\left(P^{1/2}+|S|^{1/2}\right),$$

we obtain

$$|a(n)| \ll n^{k/2-1/4-1/222}\tau(n)\log n.$$

□

5.4 Spectral analysis of the diagonal symbol

Note that either k is even or else the space $\mathcal{S}_k(\Gamma_0(q))$ is zero. For notational convenience we put $\ell = k-1$. Recall that the Petersson Trace Formula (Theorem 3.2),

$$\Delta_\ell(m,n) = \sum_{f\in\mathcal{F}} \overline{\psi_f(m)}\psi_f(n) = \delta_{mn} + 2\pi \mathrm{i}^{-k} \sum_{c\equiv 0 \pmod q} c^{-1} S(m,n;c) J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)$$

where $\mathcal{F}$ is an orthonormal basis in $\mathcal{S}_k(\Gamma_0(q))$ and $\psi_f(n)$ are the normalized Fourier coefficients given by $\psi_f(n) = \left(\Gamma(k-1)/(4\pi n)^{k-1}\right)^{1/2}\hat{f}(n)$.

We count $\Delta_\ell(m,n)$ with weight $g(\ell)$. It is of interest to average separately over $k \equiv 0 \pmod 4$ or $k \equiv 2 \pmod 4$. This implies that $\ell \equiv 3 \pmod 4$ and $\ell \equiv 1 \pmod 4$, so it only need to set $a \in \{1,3\}$ and consider

$$\sum_{\ell\equiv a \pmod 4} g(\ell)\Delta_\ell(m,n).$$

Note that

$$\sum_{\ell\equiv a \pmod 4} g(\ell)\Delta_\ell(m,n) = \delta_{mn} \sum_{\ell\equiv n \pmod 4} g(\ell) + 2\pi \sum_{c\equiv 0 \pmod q} \frac{S(m,n;c)}{c} \mathrm{i}^{a+1} \sum_{\ell\equiv a \pmod 4} g(\ell) J_\ell(m,n).$$

since $\mathrm{i}^{-\ell-1} = \mathrm{i}^{-a-1} = \mathrm{i}^{a+1}$. We are led to studying the Neumann series

$$G(x) = \sum_{\ell\equiv a \pmod 4} g(\ell) J_\ell(x).$$

Applying

$$J_\ell(x) = \int_{-1/2}^{1/2} e(\ell t)\mathrm{e}^{-\mathrm{i}x\sin 2\pi t}\mathrm{d}t.$$

Then

$$\begin{aligned}
\sum_{\ell\in\mathbb{Z}} g(\ell)e(\alpha\ell)J_\ell(x) &= \int_{-1/2}^{1/2} \mathrm{e}^{-\mathrm{i}x\sin 2\pi t} \sum_{\ell\in\mathbb{Z}} g(\ell)e((t+\alpha)\ell)\mathrm{d}t \\
&= \sum_{u\in\mathbb{Z}} \int_{-1/2}^{1/2} \hat{g}(t+\alpha+u)\mathrm{e}^{-\mathrm{i}x\sin 2\pi t}\mathrm{d}t \quad \text{(Poisson summation formula)} \\
&= \int_{\mathbb{R}} \hat{g}(t)\mathrm{e}^{-\mathrm{i}x\sin 2\pi(t-\alpha)}\mathrm{d}t. \qquad (t \mapsto t+\alpha+u)
\end{aligned}$$

From orthogonality of additive characters,

$$\begin{aligned} G(x) &= \sum_{\ell\in\mathbb{Z}} g(\ell) J_\ell(x) \frac{1}{4}\sum_{d=1}^{4} e\left(\frac{(\ell-a)d}{4}\right) \\ &= \frac{1}{4}\sum_{d=1}^{4} e\left(-\frac{ad}{4}\right) \sum_{\ell\in\mathbb{Z}} g(\ell) e\left(\frac{d\ell}{4}\right) J_\ell(x). \end{aligned}$$

Let $\alpha = d/4$,

$$G(x) = \int_{\mathbb{R}} \hat{g}(t) \frac{1}{4}\sum_{d=1}^{4} e\left(-\frac{ad}{4}\right) \mathrm{e}^{\mathrm{i}x\sin 2\pi(d/4-t)} \mathrm{d}t.$$

That is

$$4G(x) = \int_{\mathbb{R}} \hat{g}(t) c(t) \mathrm{d}t,$$

where

$$\begin{aligned} c(t) &= \frac{1}{4}\sum_{d=1}^{4} e\left(-\frac{ad}{4}\right) \mathrm{e}^{\mathrm{i}x\sin 2\pi(d/4-t)} \mathrm{d}t \\ &= \mathrm{i}^{-a} 2\mathrm{i}\sin(x\cos(2\pi t)) - 2\mathrm{i}\sin(x\sin(2\pi t)). \end{aligned}$$

For t small, using Taylor

$$\begin{aligned} \sin 2\pi t &= 2\pi t + O(|t|^3), \\ \cos 2\pi t &= 1 - 2\pi^2 t^2 + O(|t|^4). \end{aligned}$$

Then

$$\begin{aligned} c(t) &= -2\mathrm{i}\sin\left(2\pi t x + O(x|t|^3)\right) + 2\mathrm{i}^{1-a}\sin\left(x - 2\pi^2 t^2 x + O(x|t|^4)\right) \\ &= -2\mathrm{i}\sin 2\pi t x + O(x|t|^3) + 2\mathrm{i}^{1-a}\sin(x - 2\pi^2 t^2 x) + O(x|t|^4) \\ &= -2\mathrm{i}\sin 2\pi t x + O(x|t|^3) + 2\mathrm{i}^{1-a}\sin(x - 2\pi^2 t^2 x). \end{aligned}$$

We then obtain

$$4G(x) = 2\mathrm{i}^{1-a}\int_{\mathbb{R}} \hat{g}\sin(x - 2\pi^2 x t^2)\mathrm{d}t - 2\mathrm{i}\int_{\mathbb{R}} \hat{g}(t)\sin(2\pi x t)\mathrm{d}t + O\left(x\int_{\mathbb{R}} |\hat{g}|\,|t|^3 \mathrm{d}t\right).$$

First, $-2\mathrm{i}\sin(2\pi x t) = e(-x) - e(xt)$ implies that

$$\begin{aligned} 2\mathrm{i}\int_{\mathbb{R}} \hat{g}(t)\sin(2\pi x t)\mathrm{d}t &= \int_{\mathbb{R}} \hat{g}(t) e(-xt)\mathrm{d}t - \int_{\mathbb{R}} \hat{g}(t) e(xt)\mathrm{d}t \\ &= g(x) - g(-x) \\ &= g(x). \end{aligned}$$

Next, note that $\sin(x-2\pi^2 xt^2) = \mathrm{Im}\left(\mathrm{e}^{\mathrm{i}x}\mathrm{e}^{-2\pi^2\mathrm{i}xt^2}\right)$. It follows that

$$\begin{aligned}\int_{\mathbb{R}} \hat{g}(t)\mathrm{e}^{-2\pi^2\mathrm{i}xt^2}\mathrm{d}t &= \int_{\mathbb{R}} g(\xi)\left(\int_{\mathbb{R}} \mathrm{e}^{2\pi\mathrm{i}\xi t}\mathrm{e}^{-2\pi^2\mathrm{i}xt^2}\mathrm{d}t\right)\mathrm{d}\xi \\ &\int_{\mathbb{R}} g(\xi)\left(\mathrm{e}^{\mathrm{i}\xi/(2x)}\int_{\mathbb{R}} \mathrm{e}^{-2\pi^2\mathrm{i}x(t-\xi/(2\pi x))^2}\right)\mathrm{d}\xi \\ &= \frac{1}{\sqrt{2\pi x}}\int_{\mathbb{R}} g(\xi)\mathrm{e}^{\mathrm{i}\left(\xi^2/(2x)-\pi/4\right)}\mathrm{d}\xi.\end{aligned}$$

By multiplying $\mathrm{e}^{\mathrm{i}x}$ and take imaginary part,

$$\begin{aligned}2\int_{\mathbb{R}} \hat{g}(t)\sin(x-2\pi^2xt^2)\mathrm{d}t &= \frac{2}{\sqrt{2\pi x}}\int_{\mathbb{R}} g(\xi)\sin\left(x+\frac{\xi^2}{2x}-\frac{\pi}{4}\right)\mathrm{d}\xi \\ &= \frac{2}{\sqrt{2\pi x}}\int_0^\infty g(\xi)\sin\left(x+\frac{\xi^2}{2x}-\frac{\pi}{4}\right)\mathrm{d}\xi \\ &= \int_0^\infty g\left(\sqrt{2xy}\right)\sin\left(x+y-\frac{\pi}{4}\right)\frac{\mathrm{d}y}{\sqrt{\pi y}}. \quad \left(y=\frac{\xi^2}{2x}\right)\end{aligned}$$

Conclusion,

Lemma 5.1. Suppose g is a smooth, real function compactly supported on $\mathbb{R}^+$ and $a = \pm 1$. Then

$$4G(x) = g(x) + \mathrm{i}^{1-a}\int_0^\infty g\left(\sqrt{2xy}\right)\sin\left(x+y-\frac{\pi}{4}\right)\frac{\mathrm{d}y}{\sqrt{\pi y}} + O\left(x\int_{-\infty}^{\infty}|\hat{g}(t)|\,|t|^3\mathrm{d}t\right).$$

The treatment above of the Bessel-type Neumann series arising from the weighted Petersson kernel, in particular the use of Poisson summation to obtain the corresponding oscillatory integral, is closely related to the method in [ILS00]. Similar analyses of the Petersson kernel in Rankin–Selberg problems and related weight-aspect averages can be found in [LLY06]. Related spectral averaging techniques also appear in [BK17a], [BK17b].

6 Hecke Operators

6.1 Introduction

Hecke operators are averaging operators over a suitable finite collection of double cosets with respect to a group.

Let $k \in \mathbb{Z}$ and $A \in \mathrm{GL}_2^+(\mathbb{R})$, the group of 2×2 real matrices of positive determinant. Recall that the slash operator is defined by

$$f_{|A}(z) = \det(A)^{k/2}j_A(z)^{-k}f(Az), \quad A = \begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix}, \quad j_A(z) = \gamma z + \delta.$$

For $0 \neq a \in \mathbb{R}$, we have

$$f_{|aA} = \left(\frac{a}{|a|}\right)^k f_{|A}.$$

Indeed, note that $\det(aA) = a^2 \det(A)$ and $j_{aA}(z) = a(\gamma z + \delta)$ and $(aA)(z) = Az$. Thus

$$\begin{aligned} f_{|aA}(z) &= \det(aA)^{k/2} j_{aA}(z)^{-k} f((aA)z) \\ &= \left(\frac{a}{|a|}\right)^k f_{|A}. \end{aligned}$$

The slash operation is associative: $f_{|AB} = (f_{|A})_{|B}$. Let $g = f_{|A}$, then

$$\begin{aligned} (g_{|B})(z) &= \det(B)^{k/2} j_B(z)^{-k} g(Bz) \\ &= \det(B)^{k/2} j_B(z)^{-k} (f_{|A})(Bz) \\ &= \det(B)^{k/2} j_B(z)^{-k} \det(A)^{k/2} j_A(Bz)^{-k} f(ABz) \\ &= \det(AB)^{k/2} (j_A(Bz) j_B(z))^{-k} f(ABz) \\ &= \det(AB)^{k/2} (j_{AB}(z))^{-k} f(ABz) \\ &= f_{|AB}(z). \end{aligned}$$

6.2 Hecke operators T_n

Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and $n \in \mathbb{Z}^+$. Define

$$G_n = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z}, ad - bc = n \right\}.$$

Immediately we have $G_1 = \Gamma$ and $G_n = \Gamma G_n = G_n \Gamma$.

Lemma 6.1. The collection

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

forms a complete set of right coset representatives of G_n modulo Γ, i.e., we have the disjoint partition

$$G_n = \bigcup_{\rho \in \Delta_n} \Gamma \rho.$$

Proof. Clearly $G_n \supseteq \bigcup_{\rho \in \Delta_n} \Gamma \rho$, need to prove $G_n \subseteq \bigcup_{\rho \in \Delta_n} \Gamma \rho$. That is to prove, for arbitrary

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G_n$$

we have $A \in \Gamma\Delta_n$. Let $\gamma = c/\gcd(a,c)$ and $\delta = -a/\gcd(a,c)$. Then $\gamma a + \delta c = 0$ and $\gcd(\gamma,\delta) = 1$, and so there exists $\tau \in \Gamma$ such that

$$\tau = \begin{pmatrix} * & * \\ \gamma & \delta \end{pmatrix}.$$

It follows that

$$\tau A = \begin{pmatrix} * & * \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}.$$

Here $a'd' = \det(\tau A) = n$ and we can also assume $d' > 0$ by replacing τ by $-\tau$ if necessary. By choosing u with $0 \le b' - ud' < d'$,

$$\begin{pmatrix} 1 & -u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} a' & b' - ud' \\ 0 & d' \end{pmatrix}$$

implies that

$$\begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a' & b' - ud' \\ 0 & d' \end{pmatrix}.$$

Therefore

$$A = \tau^{-1} \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a' & b' - ud' \\ 0 & d' \end{pmatrix} \in \Gamma\Delta_n.$$

Now we prove the partition is disjoint. For $\rho, \rho' \in \Delta_n$ and $\rho \neq \rho'$, this is to prove $\Gamma\rho' \cap \Gamma\rho = \emptyset$. So it only suffices to show $\tau'\rho' \neq \tau\rho$ for all $\tau, \tau' \in \Gamma$, i.e., there exists no $\tau_0 \in \Gamma$ such that $\rho' = \tau_0\rho$. Assume conversely, let

$$\tau_0 = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma, \quad \rho = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \quad \rho' = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}.$$

If $\tau_0\rho = \rho'$, then

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b + \beta d \\ \gamma a & \gamma b + \delta d \end{pmatrix} = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}.$$

$\gamma a = 0$ implies that $\gamma = 0$ since we have $ad = n$. From $\alpha a = a'$ and $\gamma b + \delta d = \delta d = d'$ we obtain $n = a'd' = \alpha\delta ad = \alpha\delta n$, so $\alpha = \delta = \pm 1$. By $a, d, a', d' > 0$, $\alpha = \delta = 1$, then $a = a'$ and $d = d'$. Now $\alpha b + \beta d = b'$ implies that $b + \beta d = b'$. Since $b + \beta d = b'$ and $b = b'$, we have $\beta d = 0$, hence $\beta = 0$, which implies that

$$\tau_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$\rho = \rho'$, a contradiction. □

Note that Δ_n is a finite set with the number of elements

$$\sigma(n) = \sum_{d|n} d.$$

From Lemma 6.1, for any $\rho \in \Delta_n, \tau \in \Gamma$ there exist unique $\tau' \in \Gamma, \rho' \in \Delta_n$ such that $\rho\tau = \tau'\rho'$. But the converse is false. For example, if we consider

$$\tau' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \rho' = \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix},$$

set

$$\rho_1 = \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix}, \quad \tau_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\rho_2 = \begin{pmatrix} 1 & 1 \\ 0 & n \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix},$$

where $n > 1$, then $\rho_2\tau_2 = \rho_1\tau_1 = \tau'\rho'$.

Let $\chi\colon \mathbb{Z} \to \mathbb{C}$ be any function. This yields a function on $\mathrm{GL}_2(\mathbb{Z})$ by setting

$$\chi(\rho) = \overline{\chi(a)} \text{ if } \rho = \begin{pmatrix} a & * \\ * & * \end{pmatrix}.$$

Define the operator T_n on functions $f\colon \mathbb{H} \to \mathbb{C}$ by

$$T_n f = n^{k/2-1} \sum_{\rho \in \Delta_n} \overline{\chi(\rho)} f_{|\rho}.$$

Let

$$\rho = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then

$$\begin{aligned}(T_n f)(z) &= n^{k/2-1} \sum_{\rho \in \Delta_n} \overline{\chi(\rho)} f_{|\rho} \\ &= n^{k/2-1} \sum_{ad=n} \sum_{0 \le b < d} \chi(a) n^{k/2} d^{-k} f\left(\frac{az+b}{d}\right) \\ &= \frac{1}{n} \sum_{ad=n} \chi(a) a^k \sum_{0 \le b < d} f\left(\frac{az+b}{d}\right),\end{aligned}$$

which implies that

$$T_n = \frac{1}{n} \sum_{ad=n} \chi(a) a^k \sum_{0 \le b < d} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}.$$

6.3 The Hecke operators on periodic functions

From now on we assume that $\chi(1) = 1$ and $\chi(-1) = (-1)^k$. Let

$$\Gamma_\infty = \left\{ \pm \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} : u \in \mathbb{Z} \right\},$$

then $f_{|\tau} = \chi(\tau) f$ for any $\tau \in \Gamma_\infty$. That is to say, f is periodic of period 1, i.e., $f(z+1) = f(z)$.

Theorem 6.1. The operator T_n maps periodic functions to periodic functions, i.e.,

$$T_n \colon \mathcal{M}_k(\Gamma_\infty, \chi) \to \mathcal{M}_k(\Gamma_\infty, \chi).$$

Proof. It only suffices to show $(T_n f)(z+1) = (T_n f)(z)$. Indeed,

$$\begin{aligned}
(T_n f)(z+1) &= \frac{1}{n} \sum_{ad=n} \sum_{0 \le b < d} \chi(a) a^k f\left(\frac{a(z+1)+b}{d} \right) \\
&= \frac{1}{n} \sum_{ad=n} \sum_{0 \le b < d} \chi(a) a^k f\left(\frac{az+(b+a)}{d} \right) \\
&= \frac{1}{n} \sum_{ad=n} \sum_{0 \le b' < d} \chi(a) a^k f\left(\frac{az+b'}{d} \right) \\
&= (T_n f)(z).
\end{aligned}$$

□

Proposition 6.1. Suppose f is given by the series

$$f(z) = \sum_{m=0}^{\infty} a(m) e(mz)$$

which converges absolutely in $\mathbb{H}$. Then $T_n f$ is given by the series

$$(T_n f)(z) = \sum_{m=0}^{\infty} a_n(m) e(mz)$$

whose coefficients are

$$a_n(m) = \sum_{d \mid \gcd(m,n)} \chi(d) d^{k-1} a(mnd^{-2}).$$

Proof. We have

$$\begin{aligned}
(T_n f)(z) &= \frac{1}{n} \sum_{ad=n} \chi(a) a^k \sum_{0 \le b < d} f\left(\frac{az+b}{d} \right) \\
&= \frac{1}{n} \sum_{ad=n} \chi(a) a^k \sum_{0 \le b < d} \sum_{m=0}^{\infty} a(m) e\left(m \frac{az+b}{d} \right).
\end{aligned}$$

From

$$\sum_{b \pmod d} e\left(m\frac{b}{d}\right) = \begin{cases} d & \text{if } d \mid m, \\ 0 & \text{if } d \nmid m, \end{cases}$$

by letting $m = d\ell$ we have

$$\begin{aligned}(T_n f)(z) &= \frac{1}{n}\sum_{\ell=0}^{\infty} a(d\ell) \sum_{ad=n} \chi(a) a^k e(a\ell z) d \\ &= \sum_{\ell=0}^{\infty} a(d\ell) \sum_{ad=n} \chi(a) a^{k-1} e(a\ell z) \\ &= \sum_{m=0}^{\infty} e(mz) \left(\sum_{\substack{ad=n \\ a\ell=m}} \chi(a) a^{k-1} a(d\ell) \right) \qquad (a\ell = m) \\ &= \sum_{m=0}^{\infty} e(mz) \left(\sum_{d \mid \gcd(m,n)} \chi(d) d^{k-1} a\left(\frac{mn}{d^2}\right) \right).\end{aligned}$$

□

Corollary 6.1. For $n \geq 1$ we have

$$a_n(0) = \sigma_{k-1}(n, \chi) a(0)$$

where

$$\sigma_s(n, \chi) = \sum_{d \mid n} \chi(d) d^s.$$

Corollary 6.2. For $m, n \geq 1$ we have

$$a_n(m) = a_m(n).$$

From now on we assume that χ is completely multiplicative, i.e. $\chi(a_1 a_2) = \chi(a_1)\chi(a_2)$ if $a_1, a_2 \in \mathbb{Z}$.

Theorem 6.2. For any $m, n \geq 1$ we have

$$T_m T_n = \sum_{d \mid \gcd(m,n)} \chi(d) d^{k-1} T_{mnd^{-2}}.$$

Proof. We have

$$\begin{aligned} mn T_m T_n &= \sum_{a_1 d_1 = m} \sum_{b_1 \pmod{d_1}} \chi(a_1) a_1^k \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix} \sum_{a_2 d_2 = n} \sum_{b_2 \pmod{d_2}} \chi(a_2) a_2^k \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix} \\ &= \sum_{\substack{a_1 d_1 = m \\ a_2 d_2 = n}} \chi(a_1 a_2)(a_1 a_2)^k \sum_{\substack{b_1 \pmod{d_1} \\ b_2 \pmod{d_2}}} \begin{pmatrix} a_1 a_2 & a_1 b_2 + b_1 d_2 \\ 0 & d_1 d_2 \end{pmatrix}.\end{aligned}$$

Let $a_1b_2 + b_1d_2 = b$, $\delta = \gcd(a_1, d_2)$. Then $a_1 = a_1'\delta$ and $d_2 = d_2'\delta$ with $\gcd(a_1', d_2') = 1$. Also let $b' = a_1'b_2 + b_1d_2'$, then $b = \delta b'$. Moreover set $a_1'a_2 = a$ and $d_1d_2' = d$, then $\chi(a_1a_2)(a_1a_2)^k = \chi(\delta)\chi(a)\delta^k a^k$ and

$$\begin{pmatrix} a_1a_2 & a_1b_2 + b_1d_2 \\ 0 & d_1d_2 \end{pmatrix} = \begin{pmatrix} \delta a & \delta b' \\ 0 & \delta d \end{pmatrix} = \begin{pmatrix} a & b' \\ 0 & d \end{pmatrix}.$$

Since $\delta = \gcd(a_1, d_2)$ and $a_1 \mid m$, $d_2 \mid n$, it follows that $\delta \mid \gcd(m, n)$. Moreover $ad = mn/\delta^2$. We claim,

$$\sum_{a_1d_1=m} \sum_{a_2d_2=n} = \sum_{\delta \mid \gcd(m,n)} \sum_{ad=mn/\delta^2}.$$

From above discussion, given (a_1, a_2, d_1, d_2) we obtain (δ, a, d) satisfying $\delta \mid \gcd(m, n)$ and $ad = mn/\delta^2$. For the converse, fix $\delta \mid \gcd(m, n)$ and $ad = mn/\delta^2$, let $d_1 = \gcd(m/\delta, d), a_1' = m/(\delta d_1), d_2' = d/d_1$ and $a_2 = n/(\delta d_2')$, then $\gcd(a_1', d_2') = 1$ and $a_1d_1 = m$, $a_2d_2 = n$. Next, we claim,

$$\sum_{b_1 \pmod{d_1}} \sum_{b_2 \pmod{d_2}} = \sum_{b' \pmod{d}} \sum_{\delta \text{ times}}.$$

Indeed, $b_2 \pmod{d_2}$ implies that $b_2 \pmod{d_2'}$ δ times. From $\gcd(a_1', d_2') = 1$, we have $a_1'b_2 \pmod{d_2'}$ δ times. When $b_1 \pmod{d_1}$ and $x \pmod{d_2'}$, $x + b_1d_2' \pmod{d_1d_2'}$ since every integer $0, \cdots, d_1d_2' - 1$ can be uniquely written as $x + b_1d_2'$ where $0 \le x < d_2'$ and $0 \le b_1 < d_1$. So $b' \pmod{d_1d_2'}$ δ times. Renaming b' as b, we obtain,

$$\begin{aligned} mnT_mT_n &= \sum_{\delta \mid \gcd(m,n)} \sum_{ad=mn/\delta^2} \chi(\delta)\chi(a)\delta^k a^k \cdot \delta \sum_{b \pmod{d}} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \\ &= \sum_{\delta \mid \gcd(m,n)} \chi(\delta)\delta^{k+1} \sum_{ad=mn/\delta^2} \chi(a)a^k \sum_{b \pmod{d}} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}. \end{aligned}$$

Note that

$$mn\delta^{-2}T_{mn\delta^{-2}} = \sum_{ad=mn\delta^{-2}} \chi(a)a^k \sum_{b \pmod{d}} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix},$$

we conclude

$$T_mT_n = \sum_{\delta \mid \gcd(m,n)} \chi(\delta)\delta^{k-1}T_{mn\delta^{-2}}.$$

□

Corollary 6.3. The Hecke operators commute:

$$T_mT_n = T_nT_m.$$

Now

$$T_{m/d}T_{n/d} = \sum_{r|\gcd(m/d,n/d)} \chi(r)r^{k-1}T_{mn/(dr)^2}.$$

It follows that

$$\begin{aligned}\sum_{d|\gcd(m,n)} \mu(d)\chi(d)d^{k-1}T_{m/d}T_{n/d} &= \sum_{d|\gcd(m,n)} \mu(d)\chi(d)d^{k-1} \sum_{r|\gcd(m/d,n/d)} \chi(r)r^{k-1}T_{mn/(dr)^2} \\ &= \sum_{d|\gcd(m,n)} \sum_{r|\gcd(m/d,n/d)} \mu(d)\chi(dr)(dr)^{k-1}T_{mn/(dr)^2}.\end{aligned}$$

From $\mu * u = I$, for fixed dr,

$$\sum_{d|dr} \mu(d) = \begin{cases} 1 & \text{if } dr = 1, \\ 0 & \text{if } dr > 1. \end{cases}$$

It follows that

$$\sum_{d|\gcd(m,n)} \sum_{r|\gcd(m/d,n/d)} \mu(d)\chi(dr)(dr)^{k-1}T_{mn/(dr)^2} = T_{mn}.$$

That is,

$$T_{mn} = \sum_{d|\gcd(m,n)} \mu(d)\chi(d)d^{k-1}T_{m/d}T_{n/d}.$$

Hence in particular $\gcd(m,n) = 1$, $T_{mn} = T_mT_n$. Now take $m = p^\nu$ and $n = p$, from $\gcd(p^\nu, p) = p$ we obtain

$$T_{p^\nu p} = \mu(1)\chi(1)1^{k-1}T_{p^\nu}T_p + \mu(p)\chi(p)p^{k-1}T_{p^\nu/p}T_{p/p},$$

that is,

$$T_{p^{\nu+1}} = T_{p^\nu}T_p - \chi(p)p^{k-1}T_{p^{\nu-1}}.$$

Since Hecke operators commute,

$$T_{p^{\nu+1}} = T_pT_{p^\nu} - \chi(p)p^{k-1}T_{p^{\nu-1}}.$$

Immediately,

$$\left(1 - T_pX + \chi(p)p^{k-1}X^2\right)\left(\sum_{\nu=0}^{\infty} T_{p^\nu}X^\nu\right) = \sum_{\nu=0}^{\infty} T_{p^\nu}X^\nu - \sum_{\nu=0}^{\infty} T_pT_{p^\nu}X^{\nu+1} + \sum_{\nu=0}^{\infty} \chi(p)p^{k-1}T_{p^\nu}X^{\nu+2} = 1,$$

so

$$\sum_{\nu=0}^{\infty} T_{p^\nu}X^\nu = \left(1 - T_pX + \chi(p)p^{k-1}X^2\right)^{-1}.$$

In particular take $X = p^{-s}$,

$$\sum_{\nu=0}^{\infty} T_{p^\nu}p^{-\nu s} = \left(1 - T_pp^{-s} + \chi(p)p^{k-1-2s}\right)^{-1}.$$

When $\chi = 1$,

$$T_{p^{\nu+1}} = T_p T_{p^\nu} - p^{k-1} T_{p^{\nu-1}}.$$

We prove

$$p^{\nu(1-k)/2} T_{p^\nu} = U_\nu \left(p^{(1-k)/2} T_p \right)$$

by induction. For $\nu = 0, 1$, the assertion holds clearly. Now assume the assertion holds for ν and $\nu - 1$, from

$$U_{\nu+1}(x) = x U_\nu(x) - U_{\nu-1}(x),$$

set $x = p^{(1-k)/2} T_p$ we obtain

$$\begin{aligned} U_{\nu+1}\left(p^{(1-k)/2} T_p \right) &= p^{(1-k)/2} T_p \cdot U_\nu \left(p^{(1-k)/2} T_p \right) - U_{\nu-1}\left(p^{(1-k)/2} T_p \right) \\ &= p^{(1-k)/2} T_p \cdot p^{\nu(1-k)/2} T_{p^\nu} - p^{(\nu-1)(1-k)/2} T_{p^{\nu-1}} \\ &= p^{(\nu+1)(1-k)/2} T_{p^{\nu+1}}. \end{aligned}$$

We're done.

Let $\mathcal{T}$ denote the algebra over $\mathbb{C}$ generated by all T_n (called the **Hecke algebra**). By the above properties we conclude that $\mathcal{T}$ is a commutative algebra generated by the operators T_p for primes p.

6.4 The Hecke operators for the modular group

Next we consider the Hecke operators T_n on the linear space $\mathcal{M}_k(\Gamma)$ of modular forms of weight k, where $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. Then $f_{|-I} = f$ implies that $(-1)^k f = f$, we assume that k is an even integer. Since $f \in \mathcal{M}_k(\Gamma)$, for $\tau \in \Gamma$ we have $f_{|\tau} = f$, it follows that $f_{|\tau\rho} = (f_{|\tau})_{|\rho} = f_{|\rho}$. Consequently,

$$T_n f = n^{k/2-1} \sum_{\rho \in \Gamma \backslash G_n} f_{|\rho}.$$

Theorem 6.3. The Hecke operator T_n maps a modular form to a modular form and a cusp form to a cusp form:

$$\begin{aligned} T_n &: \mathcal{M}_k(\Gamma) \longrightarrow \mathcal{M}_k(\Gamma), \\ T_n &: \mathcal{S}_k(\Gamma) \longrightarrow \mathcal{S}_k(\Gamma). \end{aligned}$$

Proof. By the correspondence $\rho\tau = \tau'\rho'$, we have

$$\begin{aligned}
(T_n f)_{|\tau} &= \left(n^{k/2-1} \sum_{\rho \in \Delta_n} f_{|\rho} \right)_{|\tau} \\
&= n^{k/2-1} \sum_{\rho \in \Delta_n} (f_{|\rho})_{|\tau} \\
&= n^{k/2-1} \sum_{\rho \in \Delta_n} f_{|\rho\tau} \\
&= n^{k/2-1} \sum_{\rho' \in \Delta_n} f_{|\tau'\rho'} \\
&= n^{k/2-1} \sum_{\rho' \in \Delta_n} (f_{|\tau'})_{|\rho'} \\
&= n^{k/2-1} \sum_{\rho' \in \Delta_n} f_{|\rho'} \\
&= T_n f.
\end{aligned}$$

So $T_n f$ satisfies the modular transformation law. Clearly $f_{|\rho}$ is holomorphic for every ρ, so $T_n f$ holomorphic. Thanks to Proposition 6.1, $T_n f$ is holomorphic at every cusp.

If f is a cusp form, the Fourier coefficients of $T_n f$ is given by $a_n(0) = \sigma_{k-1}(n)a(0) = 0$, so $T_n f$ is a cusp form. □

Recall that $\mathcal{S}_k(\Gamma)$ is spanned by $\{P_m\}_{m \geq 1}$ when $k > 2$, where

$$P_m(z) = \sum_{\tau \in \Gamma_\infty \backslash \Gamma} j_\tau(z)^{-k} e(m\tau z).$$

Therefore $T_n P_m$ is a linear combination of various Poincaré series (not unique).

Theorem 6.4. For $k > 2$, $m \geq 0$ and $n \geq 1$ we have

$$T_n P_m = \sum_{d | \gcd(m,n)} (n/d)^{k-1} P_{mnd^{-2}}.$$

Proof. We have

$$P_m(z) = \sum_{\tau \in \Gamma_\infty \backslash \Gamma} j_\tau(z)^{-k} e(m\tau z) = \sum_{\tau \in \Gamma_\infty \backslash \Gamma} \left(e(m)_{|\tau} \right)(z),$$

it follows that

$$T_n P_m(z) = n^{k/2-1} \sum_{\rho \in \Delta_n} \left(\sum_{\tau \in \Gamma_\infty \backslash \Gamma} e(m)_{|\tau} \right)_{|\rho} = n^{k/2-1} \sum_{\rho \in \Delta_n} \sum_{\tau \in \Gamma_\infty \backslash \Gamma} e(m)_{|\tau\rho}.$$

Let $\mathcal{H}$ be any set of right representatives of $\Gamma_\infty\backslash\Gamma$, $\mathcal{G}$ be any set of right representatives of $\Gamma\backslash G_n$ (note that one can choose $\mathcal{G}=\Delta_n$). Then $\mathcal{H}\Delta_n$ is a set of right coset representative of $\Gamma_\infty\backslash G_n$,

$$G_n=\bigcup_{\rho\in\Delta_n}\Gamma\rho=\bigcup_{\rho\in\Delta_n}\bigcup_{\tau\in\mathcal{H}}\Gamma_\infty\tau\rho.$$

Moreover $\Delta_n\mathcal{H}$ is a set of right coset representative of $\Gamma_\infty\backslash G_n$ (This claim will be proved later). Hence

$$\begin{aligned}T_nP_m(z)&=n^{k/2-1}\sum_{\rho\in\Delta_n}\sum_{\tau\in\mathcal{H}}j_{\rho\tau}(z)^{-k}e(m\rho\tau z)\\&=n^{k/2-1}\sum_{ad=n}\sum_{b\pmod d}\sum_{\tau\in\mathcal{H}}\left(j_\rho(\tau z)j_\tau(z)\right)^{-k}e(m\rho\tau z).\end{aligned}$$

From

$$j_\rho(\tau z)=d,\quad\text{where }\rho=\begin{pmatrix}a&b\\0&d\end{pmatrix},$$

we have

$$\begin{aligned}T_nP_m(z)&=n^{k-1}\sum_{ad=n}\sum_{\tau\in\mathcal{H}}d^{-k}j_\tau(z)^{-k}e\left(m\frac{a\tau z}{d}\right)\sum_{b\pmod d}e\left(m\frac{b}{d}\right)\\&=n^{k-1}\sum_{ad=n}\sum_{d\mid m\pmod d}d^{1-k}\sum_{\tau\in\mathcal{H}}j_\tau(z)^{-k}e\left(\frac{am}{d}\tau z\right)\\&=\sum_{d\mid\gcd(m,n)}\left(\frac{n}{d}\right)^{k-1}\sum_{\tau\in\mathcal{H}}j_\tau(z)^{-k}e\left(\frac{nm}{d^2}\tau z\right)\\&=\sum_{d\mid\gcd(m,n)}\left(\frac{n}{d}\right)^{k-1}P_{mnd^{-2}}(z).\end{aligned}$$

Now we prove: $\Delta_n\mathcal{H}$ is a set of right coset representative of $\Gamma_\infty\backslash G_n$. Let

$$\rho_1=\begin{pmatrix}a_1&b_1\\0&d_1\end{pmatrix},\quad\rho_2=\begin{pmatrix}a_2&b_2\\0&d_2\end{pmatrix}\in\Delta_n$$

and $\tau_1,\tau_2\in\mathcal{H}$. Assume $\rho_1\tau_1\in\Gamma_\infty\rho_2\tau_2$, it follows that for some $u\in\mathbb{Z}$,

$$\rho_1\tau_1=\begin{pmatrix}1&u\\0&1\end{pmatrix}\begin{pmatrix}a_2&b_2\\0&d_2\end{pmatrix}\tau_2=\begin{pmatrix}a_2&b_2+ud_2\\0&d_2\end{pmatrix}\tau_2.$$

Therefore

$$\frac{1}{n}\begin{pmatrix}d_2&-b_2-ud_2\\0&a_2\end{pmatrix}\begin{pmatrix}a_1&b_1\\0&d_1\end{pmatrix}=\tau_2\tau_1^{-1}\in\mathrm{SL}_2(\mathbb{Z}).$$

That is,

$$\text{LHS} = \begin{pmatrix} \frac{1}{a_2} & \frac{-b_2-ud_2}{n} \\ 0 & \frac{1}{d_2} \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix} = \begin{pmatrix} \frac{a_1}{a_2} & \frac{b_1}{a_2} - \frac{d_1(b_2+ud_2)}{n} \\ 0 & \frac{d_1}{d_2} \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

So $a_2 \mid a_1$, $d_2 \mid d_1$. By $a_1 d_1 = a_2 d_2 = n$ and $a_1 a_2 > 0$, we obtain $a_1 = a_2 = a$ and $d_1 = d_2 = d$. Now

$$\begin{pmatrix} \frac{a_1}{a_2} & \frac{b_1-b_2-ud}{a} \\ 0 & \frac{d_1}{d_2} \end{pmatrix} \in \Gamma_\infty,$$

so $\tau_2 \tau_1^{-1} \in \Gamma_\infty$, and so $\tau_1 \sim \tau_2$, we have $\tau_1 = \tau_2$ since $\tau_1, \tau_2 \in \mathcal{H}$. Consequently,

$$\begin{pmatrix} a & b_1 \\ 0 & d \end{pmatrix} = \begin{pmatrix} a & b_2 + ud \\ 0 & d \end{pmatrix}.$$

By $0 \le b_1 < d$, we have $0 \le b_2 + ud_2 < d$, $b_1 = b_2$ and $u = 0$. So for different pair $(\rho_1, \tau_1) \ne (\rho_2, \tau_2)$, $\rho_1 \tau_1 \notin \Gamma_\infty \rho_2 \tau_2$. Now we prove for arbitrary $\tau \in \mathcal{H}$ and $\rho \in \Delta_n$, there exists $\tau' \in \mathcal{H}$ and $\rho' \in \Delta_n$ such that $\tau\rho \in \Gamma_\infty \rho' \tau'$. We have $\tau\rho = \rho_0 \tau_0$, where

$$\rho_0 = \begin{pmatrix} a_0 & b_0 \\ 0 & d_0 \end{pmatrix},$$

where $a_0 d_0 = n$ (not necessarily $0 \le b_0 < d_0$) and $\tau_0 \in \mathrm{SL}_2(\mathbb{Z})$. Let $\tau_0 = \gamma_\infty \tau'$, where $\gamma_\infty \in \Gamma_\infty$ and $\tau' \in \mathcal{H}$ since $\tau_0 \in \Gamma = \mathrm{SL}_2(\mathbb{Z})$. Assume that

$$\tau\rho = (\rho_0 \gamma_\infty)\tau' = \begin{pmatrix} a_0 & b' \\ 0 & d_0 \end{pmatrix} \tau'$$

for some b', there exists v makes $0 \le b'' \coloneqq b' - v d_0 < d_0$. Therefore

$$\tau\rho = \begin{pmatrix} 1 & v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_0 & b'' \\ 0 & d_0 \end{pmatrix} \tau' \in \Gamma_\infty \rho' \tau'.$$

We're done. □

For $m = 0$ we get the Eisenstein series

$$E_k(z) = \sum_{\tau \in \Gamma_\infty \backslash \Gamma} j_\tau(z)^{-k} = \frac{1}{2} \sum_{\gcd(c,d)=1} (cz+d)^{-k}.$$

Then

$$T_n E_k = \sum_{d \mid n} \left(\frac{n}{d}\right)^{k-1} E_k = \sigma_{k-1}(n) E_k.$$

E_k is an eigenfunction of all the Hecke operators T_n with eigenvalue $\sigma_{k-1}(n)$.

Corollary 6.4. For $m, n \geq 1$ we have

$$m^{k-1} T_n P_m = n^{k-1} T_m P_n.$$

Proof. We have

$$\begin{aligned} m^{k-1} T_n P_m &= m^{k-1} \sum_{d \mid \gcd(m,n)} (n/d)^{k-1} P_{mnd^{-2}} \\ &= n^{k-1} \sum_{d \mid \gcd(m,n)} (m/d)^{k-1} P_{mnd^{-2}} \\ &= n^{k-1} T_m P_n. \end{aligned}$$

□

Recall that Theorem 3.1

$$\langle f, P_m \rangle = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} \hat{f}(m).$$

Then

$$\begin{aligned} \langle T_n f, P_m \rangle &= \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} a_n(m), \\ \langle T_m f, P_n \rangle &= \frac{\Gamma(k-1)}{(4\pi n)^{k-1}} a_m(n). \end{aligned}$$

Proposition 6.2. For $m, n \geq 1$ and $f \in \mathcal{M}_k(\Gamma)$ we have

$$m^{k-1} \langle T_n f, P_m \rangle = n^{k-1} \langle T_m f, P_n \rangle .$$

Let $\mathcal{S}$ be a finite dimensional Hilbert space over $\mathbb{C}$. Any linear operator $T \colon \mathcal{S} \to \mathcal{S}$ has the adjoint operator $T^* \colon \mathcal{S} \to \mathcal{S}$ defined by

$$\langle Tf, g \rangle = \langle f, T^* g \rangle .$$

T is said to be **normal** if it commutes with T^*. If $T = T^*$ then T is said to be **self-adjoint**.

Theorem 6.5. The Hecke operators T_n acting on the space of cusp forms for the modular group are self-adjoint, i.e.

$$\langle T_n f, g \rangle = \langle f, T_n g \rangle \text{ for all } f, g \in \mathcal{S}_k(\Gamma).$$

Proof. $\mathcal{S}_k(\Gamma)$ is spanned by the Poincaré series, so it is enough to prove $\langle T_n P_m, P_\ell \rangle = \langle P_m, T_n P_\ell \rangle$. Thanks to Proposition 6.2 and Corollary 6.4,

$$\langle T_n P_m, P_\ell \rangle = \frac{n^{k-1}}{\ell^{k-1}} \langle T_\ell P_m, P_n \rangle = \frac{n^{k-1}}{\ell^{k-1}} \frac{\ell^{k-1}}{m^{k-1}} \langle T_m P_\ell, P_n \rangle = \frac{n^{k-1}}{\ell^{k-1}} \frac{\ell^{k-1}}{m^{k-1}} \frac{m^{k-1}}{n^{k-1}} \langle T_n P_\ell, P_m \rangle = \langle T_n P_\ell, P_m \rangle .$$

On the other hand, note that the Fourier coefficients of Poincaré series, and so the inner products, are real, we have $\langle T_n P_\ell, P_m \rangle = \langle P_m, T_n P_\ell \rangle$. Immediately we conclude that $\langle T_n P_m, P_\ell \rangle = \langle P_m, T_n P_\ell \rangle$. □

Proposition 6.3. Let $\mathcal{S}$ be a finite dimensional Hilbert space over $\mathbb{C}$ and let $\mathcal{T}$ be a commuting family of normal operators $T \colon \mathcal{S} \to \mathcal{S}$. Then there exists an orthonormal basis $\mathcal{F}$ of $\mathcal{S}$ which consists of common eigenfunctions of all the operators in $\mathcal{T}$.

Theorem 6.6 (Hecke)**.** the space $\mathcal{S}_k(\Gamma)$ of cusp forms for the modular group there exists an orthonormal basis $\mathcal{F}$ which consists of eigenfunctions of all the Hecke operators T_n.

Proof. Applying Proposition 6.3 by $\mathcal{S} = \mathcal{S}_k(\Gamma)$ and $\mathcal{T} = \{T_n\}$. Recall the commutativity of Hecke operators has been proved in Corollary 6.3. □

Let $f \in \mathcal{S}_k(\Gamma)$ be a Hecke eigenform,

$$T_n f = \lambda(n) f \text{ for } n = 1, 2, 3, \cdots$$

Suppose f has Fourier expansion

$$f(z) = \sum_{m=1}^{\infty} a(m) e(mz).$$

Recall Proposition 6.1,

$$a_n(m) = \sum_{d | \gcd(n,m)} d^{k-1} a\left(\frac{mn}{d^2}\right).$$

The m-th Fourier coefficient of $T_n f$ is

$$\sum_{d | \gcd(n,m)} d^{k-1} a\left(\frac{mn}{d^2}\right),$$

and the m-th Fourier coefficient of $\lambda(n) f$ is

$$\lambda(n) a(m),$$

we obtain

$$\lambda(n) a(m) = \sum_{d | \gcd(n,m)} d^{k-1} a\left(\frac{mn}{d^2}\right).$$

For $m = 1$ gives

$$a(n) = \lambda(n) a(1).$$

Hence $a(1) \neq 0$, as otherwise f would vanish identically. Therefore the Fourier coefficients of a Hecke cusp form are proportional to the eigenvalues of the Hecke operators.

As an example, consider the space $\mathcal{S}_{12}(\Gamma)$, which is one-dimensionally spanned by $\Delta(z)$; therefore $\Delta(z)$ is automatically a simultaneous eigenfunction of all the Hecke operators, namely

$$T_n\Delta(z) = \tau(n)\Delta(z),$$

where $\tau(n)$ is the Ramanujan function.

6.5 The Hecke operators with a character

Let

$$\Gamma_0(q) = \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma : \gamma \equiv 0 \pmod{q} \right\}$$

be the Hecke congruence subgroup, and χ be a Dirichlet character modulo q. It follows that

$$T_n f = n^{k/2-1} \sum_{\rho \in \Delta_n} \overline{\chi(\rho)} f_{|\rho} = n^{k/2-1} \sum_{\rho \in \Delta_n^q} \overline{\chi(\rho)} f_{|\rho},$$

where

$$\Delta_n^q = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \Delta_n : \gcd(a,q) = 1 \right\}.$$

Recall given

$$\rho = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \Delta_n^q, \quad \tau = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_0(q),$$

there exists

$$\rho' = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} \in \Delta_n, \quad \tau' = \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

such that $\rho\tau = \tau'\rho'$. It follows that

$$\begin{pmatrix} \alpha a + \gamma b & * \\ \gamma d & \delta d \end{pmatrix} = \begin{pmatrix} \alpha' a' & * \\ \gamma' a' & \gamma' b' + \delta' d' \end{pmatrix}.$$

We have $a' = \gcd(\alpha a + \gamma b, \gamma d)$ and $\gamma' = \gamma d / \gcd(\alpha a + \gamma b, \gamma d)$. Since $\gamma \equiv 0 \pmod{q}$, we have $q \mid \gamma d$. On the other hand $\alpha a + \gamma b \equiv \alpha a \pmod{q}$. From $\alpha\delta \equiv 1 \pmod{q}$, $\gcd(\alpha, q) = 1$, combining with $\gcd(a,q) = 1$ get $\gcd(a\alpha, q) = 1$. Hence $\gcd(\alpha a + \gamma b, q) = 1$. It follows that $\gcd(\gcd(\alpha a + \gamma b, \gamma d), q) = 1$. We have $q \mid \gamma'$. Now $\alpha'\delta' - \beta'\gamma' = 1$ implies $\alpha'\delta' \equiv 1 \pmod{q}$, then $\gcd(\alpha', q) = 1$. Since $\alpha' a' \equiv \alpha a \pmod{q}$, we have $\gcd(a', q) = 1$. That is, $\rho' \in \Delta_n^q$ and $\tau' \in \Gamma_0(q)$. Moreover, $\alpha a \equiv \alpha' a' \pmod{q}$ implies that $\chi(\rho)\chi(\tau) = \chi(\tau')\chi(\rho')$ if $\rho\tau = \tau'\rho'$ for $(\rho, \tau) \in \Delta_n^q \times \Gamma_0(q)$.

As before we require that the character $\chi \pmod{q}$ satisfies the consistency condition $\chi(-1) = (-1)^k$; otherwise $\mathcal{M}_k(\Gamma_0(q), \chi) = 0$.

Theorem 6.7. The Hecke operator $T_n = T_n^\chi$ maps automorphic forms to automorphic forms and cusp forms to cusp forms:

$$\begin{aligned} T_n \colon \mathcal{M}_k(\Gamma_0(q), \chi) &\longrightarrow \mathcal{M}_k(\Gamma_0(q), \chi), \\ T_n \colon \mathcal{S}_k(\Gamma_0(q), \chi) &\longrightarrow \mathcal{S}_k(\Gamma_0(q), \chi). \end{aligned}$$

Proof. By the correspondence $\rho\tau = \tau'\rho'$, we have

$$\begin{aligned} (T_n f)_{|\tau} &= \left(n^{k/2-1} \sum_{\rho \in \Delta_n^q} \overline{\chi(\rho)} f_{|\rho} \right)_{|\tau} \\ &= n^{k/2-1} \sum_{\rho \in \Delta_n^q} \overline{\chi(\rho)} f_{|\rho\tau} \\ &= n^{k/2-1} \sum_{\rho' \in \Delta_n^q} \overline{\chi(\rho)} f_{|\tau'\rho'} \\ &= n^{k/2-1} \sum_{\rho' \in \Delta_n^q} \overline{\chi(\rho)} \chi(\tau') f_{|\rho'} \\ &= n^{k/2-1} \sum_{\rho' \in \Delta_n^q} \chi(\tau) \overline{\chi(\rho')} f_{|\rho'} \\ &= \chi(\tau)(T_n f). \end{aligned}$$

So $T_n f$ satisfies the modular transformation law. Clearly $f_{|\rho}$ is holomorphic for every ρ, so $T_n f$ holomorphic. Thanks to Proposition 6.1, $T_n f$ is holomorphic at every cusp.

Now assume f admits Fourier expansion

$$f(z) = \sum_{m \geq 0} a(m) e(mz).$$

We have

$$\begin{aligned} T_n f(z) &= n^{k/2-1} \sum_{\rho \in \Delta_n^q} \overline{\chi(\rho)} \left(\sum_{m \geq 0} a(m) e(mz) \right)_{|\rho} \\ &= n^{k/2-1} \sum_{\substack{ad=n \\ \gcd(a,q)=1}} \sum_{b \pmod d} \chi(a) n^{k/2} d^{-k} \sum_{m \geq 0} a(m) e\left(m \frac{az+b}{d} \right) \\ &= \sum_{\substack{ad=n \\ \gcd(a,q)=1}} \chi(a) n^{k-1} d^{-k} \sum_{m \geq 0} a(m) \left(\sum_{b \pmod d} e\left(\frac{mb}{d} \right) \right) e\left(\frac{ma}{d} z \right) \\ &= \sum_{\substack{ad=n \\ \gcd(a,q)=1, d \mid m}} \chi(a) n^{k-1} d^{1-k} \sum_{m \geq 0} a(m) e\left(\frac{ma}{d} z \right). \end{aligned}$$

Let $m = d\ell$, then

$$T_n f(z) = \sum_{\substack{ad=n \\ \gcd(a,q)=1}} \chi(a) n^{k-1} d^{1-k} \sum_{\ell \geq 0} a(d\ell) e(a\ell z).$$

Therefore

$$a_n(0) = \sum_{\substack{ad=n \\ \gcd(a,q)=1}} \chi(a) n^{k-1} d^{1-k} a(0) = a(0) \sum_{\substack{ad=n \\ \gcd(a,q)=1}} \chi(a) a^{k-1}.$$

Consequently $a(0) = 0$ implies that $a_n(0) = 0$. T_n sends a cusp form to a cusp form. □

By definition,

$$T_p f = p^{k/2-1} \sum_{\rho \in \Delta_p^q} \overline{\chi(\rho)} f_{|\rho}.$$

Given

$$\rho = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \Delta_p^q,$$

we have $ad = p$ and $a > 0$, which implies that $(a, d) = (1, p)$ or $(p, 1)$. If

$$\rho = \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix},$$

then $\overline{\chi(\rho)} = 1$, and

$$f_{|\rho}(z) = p^{-k/2} f\left(\frac{z+b}{p}\right).$$

If

$$\rho = \begin{pmatrix} p & b \\ 0 & 1 \end{pmatrix},$$

we have

$$f_{|\rho}(z) = p^{k/2} f(pz).$$

Consequently,

$$T_p = \frac{1}{p} \sum_{b \pmod p} \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix} + \chi(p) p^{k-1} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}.$$

If $p \mid q$, this expression reduces to

$$T_p = \frac{1}{p} \sum_{b \pmod p} \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix}.$$

Moreover from Theorem 6.2, we have

$$T_{p^\nu} = (T_p)^\nu \quad \text{if } p \mid q.$$

Theorem 6.8. Suppose $k > 2$ and $\chi(-1) = (-1)^k$. For all $m \geq 0$ and $n \geq 1$ we have

$$T_n P_m = \sum_{d|\gcd(m,n)} \chi\left(\frac{n}{d}\right)\left(\frac{n}{d}\right)^{k-1} P_{mnd^{-2}}.$$

For $m = 0$ we see that the Eisenstein series

$$E_k(z) = \sum_{\tau \in \Gamma_\infty \backslash \Gamma_0(q)} \overline{\chi(\tau)} j_\tau(z)^{-k}$$

is an eigenfunction of all T_n with eigenvalue

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

If $n \mid q^\infty$, then $\chi(n/d) = 0$ unless $n = d$. So $T_n P_m$ vanishes unless $n \mid m$, in which case $T_n P_m = P_{m/n}$. If $\gcd(n, q) = 1$, we can write

$$T_n P_m = \chi(n) n^{k-1} \sum_{d|\gcd(m,n)} \overline{\chi(d)} d^{1-k} P_{mnd^{-2}}.$$

Hence

Corollary 6.5. If $m, n \geq 1$ and $\gcd(mn, q) = 1$, then

$$\chi(m) m^{k-1} T_n P_m = \chi(n) n^{k-1} T_m P_n.$$

Our next aim is to find the adjoint of T_n. We use the concept of absolute inner product. Recall that

$$(f, g)(z) = y^k f(z) \overline{g(z)} \mathrm{d}\mu z.$$

Let ϑ be a multiplier system of weight k for a group Γ (here Γ is arbitrary). If $f, g \in \mathcal{S}_k(\Gamma, \vartheta)$, then $(f, g)(z)$ is a Γ-invariant measure, and the inner product is just the volume of a fundamental domain for this measure,

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

Given

$$\sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}),$$

we have

$$\begin{aligned}(f_{|\sigma}, g_{|\sigma})(z) &= \Big((\det \sigma)^{k/2}(cz + d)^{-k} f(\sigma z), (\det \sigma)^{k/2}(cz + d)^{-k} g(\sigma z)\Big) \\ &= y^k (\det \sigma)^k |cz + d|^{-2k} f(\sigma z) \overline{g(\sigma z)} \mathrm{d}\mu z.\end{aligned}$$

From

$$\mathrm{Im}(\sigma z) = \frac{\det\sigma \cdot y}{|cz+d|^2},$$

and $\mathrm{d}\mu(\sigma z) = \mathrm{d}\mu z$, we obtain

$$(f_{|\sigma}, g_{|\sigma})(z) = \mathrm{Im}(\sigma z)^k f(\sigma z)\overline{g(\sigma z)}\mathrm{d}\mu(\sigma z) = (f,g)(\sigma z).$$

Therefore, if $D \subset \mathbb{H}$ is measurable, then

$$\int_D (f_{|\sigma}, g_{|\sigma})(z) = \int_D (f,g)(\sigma z) = \int_{\sigma D} (f,g)(z)$$

by replacing $\sigma z \mapsto z$. For any subgroup $\Gamma' \subset \Gamma$ of finite index, we have

$$\langle f, g\rangle = \frac{1}{[\Gamma:\Gamma']}\int_{\Gamma'\backslash\mathbb{H}} (f,g)(z).$$

For computational convenience we wish to have Γ' sufficiently small so that the multiplier $\vartheta\colon \Gamma \to \mathbb{C}$ becomes trivial on Γ', i.e., $f, g \in \mathcal{S}_k(\Gamma')$. This is certainly possible if Γ is a congruence group and ϑ is a character (take $\mathrm{Ker}\,\vartheta = \{\gamma \in \Gamma : \vartheta(\gamma) = 1\}$). Furthermore, for any given $\sigma \in \mathrm{GL}_2^+(\mathbb{Q})$ one can choose Γ' small enough, but of finite index, such that the σ-slash operator maps $\mathcal{S}_k(\Gamma, \vartheta)$ into $\mathcal{S}_k(\Gamma')$, i.e.,

$$f_{|\sigma} \in \mathcal{S}_k(\Gamma') \text{ if } f \in \mathcal{S}_k(\Gamma, \vartheta).$$

Say precisely, for $\sigma \in \mathrm{GL}_2^+(\mathbb{Q})$ and $\gamma \in \Gamma'$ we have

$$\left(f_{|\sigma}\right)_{|\gamma} = f_{|\sigma\gamma} = \left(f_{|\sigma\gamma\sigma^{-1}}\right)_{|\sigma}.$$

Hence if we have $\sigma\gamma\sigma^{-1} \in \Gamma$ and $\vartheta(\sigma\gamma\sigma^{-1}) = 1$, then $f_{|\sigma\gamma\sigma^{-1}} = \vartheta(\sigma\gamma\sigma^{-1})f = f$. That is, $\left(f_{|\sigma}\right)_{|\gamma} = f_{|\sigma}$, $f_{|\sigma}$ is invariant under the action of Γ'-slash operators. Thus we can choose $\Gamma' \subset \sigma^{-1}\left(\Gamma \cap \mathrm{Ker}\,\vartheta\right)\sigma$.

For example, recall that

$$\Gamma(q) = \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod q\right\}$$

is the principal congruence subgroup. If $\sigma \in G_n$ and $\Gamma = \Gamma_0(q)$, then $\Gamma' = \Gamma(nq)$ is the group to take. Indeed the Dirichlet character is trivial on $\Gamma(q)$. Take arbitrary $\gamma \in \Gamma(nq)$, write $\gamma = I + nqA$ where $A \in M_2(\mathbb{Z})$. Since $\sigma \in G_n$, we have $\sigma^{-1} = 1/n\,\mathrm{adj}(\sigma)$, then

$$\sigma\gamma\sigma^{-1} = \sigma(I+nqA)\sigma^{-1} = I + q(\sigma A\,\mathrm{adj}(\sigma)).$$

That is, $\sigma\gamma\sigma^{-1} \in \Gamma(q) \subset \Gamma_0(q)$. Consequently

$$\left(f_{|\sigma}\right)_{|\gamma} = f_{|\sigma\gamma} = \left(f_{|\sigma\gamma\sigma^{-1}}\right)_{|\sigma} = \vartheta(\sigma\gamma\sigma^{-1}) f_{|\sigma} = f_{|\sigma},$$

which implies that $f_{|\sigma} \in \mathcal{S}_k(\Gamma(nq))$.

Proposition 6.4. Let Γ be a congruence group and $\sigma \in \mathrm{GL}_2^+(\mathbb{Q})$. Then the σ-slash operator is an isometry of $\mathcal{S}_k(\Gamma, \vartheta)$ with respect to the absolute inner product, i.e.,

$$\left\langle f_{|\sigma}, g_{|\sigma} \right\rangle = \langle f, g \rangle \quad \text{for all } f, g \in \mathcal{S}_k(\Gamma, \vartheta)$$

Proof. Let Γ' be a sufficiently small subgroup of Γ of finite index such that $f_{|\sigma} \in \mathcal{S}_k(\Gamma')$ if $f \in \mathcal{S}_k(\Gamma, \vartheta)$, and let $D' \subset \mathbb{H}$ denote the fundamental domain of Γ'. Then

$$\begin{aligned}
\left\langle f_{|\sigma}, g_{|\sigma} \right\rangle &= \int_{\Gamma\backslash\mathbb{H}} (f_{|\sigma}, g_{|\sigma})(z) \\
&= [\Gamma : \Gamma']^{-1} \int_{\Gamma'\backslash\mathbb{H}} (f_{|\sigma}, g_{|\sigma})(z) \\
&= [\Gamma : \Gamma']^{-1} \int_{D'} (f, g)(\sigma z).
\end{aligned}$$

Now we claim: $\sigma D'$ is the fundamental domian for the conjugate group $\Gamma'' = \sigma\Gamma'\sigma^{-1}$. Note that $\Gamma z = \left(\sigma\Gamma\sigma^{-1}\right)(\sigma z)$. Given $w \in \mathbb{H}$ and let $z = \sigma^{-1}w$, since D' is the fundamental domain of Γ', there exists $\gamma \in \Gamma'$ such that $\gamma z \in D'$. It follows that $\left(\sigma\gamma\sigma^{-1}\right)(\sigma w) = \sigma(\gamma z) \in \sigma D'$. So every orbit of Γ'' admits a representative in $\sigma D'$. If $w_1, w_2 \in \sigma D'$ belongs to the same Γ''-orbit, there exists $\delta \in \Gamma'$ with $w_2 = \sigma\delta\sigma^{-1}w_1$. Write $w_1 = \sigma z_1$ and $w_2 = \sigma z_2$, we obtain $z_2 = \delta z_1$, which implies that $z_1 = z_2$, so $w_1 = w_2$.

Choose Γ' small enough that $\Gamma'' \subset \Gamma$, from $\mathrm{vol}(\sigma D') = \mathrm{vol}(D')$

$$[\Gamma : \Gamma''] = \frac{\mathrm{vol}(\Gamma''\backslash\mathbb{H})}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} = \frac{\mathrm{vol}(\Gamma'\backslash\mathbb{H})}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} = [\Gamma : \Gamma'].$$

Hence

$$\begin{aligned}
\left\langle f_{|\sigma}, g_{|\sigma} \right\rangle &= [\Gamma : \Gamma']^{-1} \int_{\sigma D'} (f, g)(z) \\
&= [\Gamma : \Gamma'']^{-1} \int_{\sigma D'} (f, g)(z) \\
&= \langle f, g \rangle .
\end{aligned}$$

□

Replace g by $g_{|\sigma^{-1}}$ we obtain

$$\left\langle f_{|\sigma}, g \right\rangle = \left\langle f, g_{|\sigma^{-1}} \right\rangle .$$

Let $\sigma\sigma' = \det(\sigma)I$, $g_{|\sigma^{-1}} = g_{|(\det\sigma)\sigma^{-1}} = g_{|\sigma'}$, which yields

$$\left\langle f_{|\sigma}, g\right\rangle = \left\langle f, g_{|\sigma'}\right\rangle.$$

Theorem 6.9. If $\gcd(n,q) = 1$ and $f, g \in \mathcal{S}_k(\Gamma_0(q), \chi)$, then

$$\langle T_n f, g\rangle = \chi(n)\langle f, T_n g\rangle.$$

In other words, the adjoint of T_n is $T_n^* = \overline{\chi(n)}T_n$; thus T_n is normal while $\overline{\chi(n)}^{1/2}T_n$ is self-adjoint.

Proof. Given

$$\rho = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \Delta_n^q,$$

we have $ad = n$, then $a \mid n$. By assumption $\gcd(n,q) = 1$, so $\gcd(a,q) = 1$. Therefore $\Delta_n^q = \Delta_n$. Consequently

$$\begin{aligned}
\langle T_n f, g\rangle &= \left\langle n^{k/2-1}\sum_{\rho\in\Delta_n^q}\overline{\chi(\rho)}f_{|\rho}, g\right\rangle \\
&= n^{k/2-1}\sum_{\rho\in\Delta_n}\left\langle\overline{\chi(\rho)}f_{|\rho}, g\right\rangle \\
&= n^{k/2-1}\sum_{\rho\in\Delta_n}\left\langle\chi(n)\chi(\rho')f, g_{|\rho'}\right\rangle \\
&= n^{k/2-1}\sum_{\rho\in\Delta_n}\chi(n)\left\langle f, \overline{\chi(\rho')}g_{|\rho'}\right\rangle
\end{aligned}$$

where $\rho\rho' = nI$. For any $\tau, \tau' \in \Gamma_0(q)$, set $\rho'' = \tau'\rho'\tau^{-1}$, then

$$\chi(\tau')g_{|\rho'} = (g_{|\tau'})_{|\rho'} = g_{|\tau'\rho'} = g_{|\rho''\tau} = (g_{|\rho''})_{|\tau} = \chi(\tau)g_{|\rho''},$$

it follows that $g_{|\rho'} = \chi(\tau)\overline{\chi(\tau')}g_{|\rho''}$. On the other hand $\chi(\rho'')\chi(\tau) = \chi(\tau')\chi(\rho')$ implies that $\overline{\chi(\rho')}\chi(\tau)\overline{\chi(\tau')} = \chi(\rho'')$, so $\overline{\chi(\rho')}g_{|\rho'} = \overline{\chi(\rho'')}g_{|\rho''}$. Therefore,

$$\left\langle f, \overline{\chi(\rho')}g_{|\rho'}\right\rangle = \left\langle f, \overline{\chi(\rho'')}g_{|\rho''}\right\rangle.$$

If $\rho'' = \rho$,

$$\langle T_n f, g\rangle = n^{k/2-1}\sum_{\rho\in\Delta_n}\chi(n)\left\langle f, \overline{\chi(\rho)}g_{|\rho}\right\rangle = \chi(n)\langle f, T_n g\rangle.$$

We need to construct $\tau, \tau' \in \Gamma_0(q)$ such that $\rho = \rho'' = \tau'\rho'\tau^{-1}$. Let

$$\tau = \begin{pmatrix} \alpha & \beta \\ q & \delta \end{pmatrix}, \quad \tau' = \begin{pmatrix} \alpha' & \beta' \\ q & \delta' \end{pmatrix}.$$

Then $\rho'' = \tau'\rho'\tau^{-1}$ iff

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ q & \delta \end{pmatrix} = \begin{pmatrix} \alpha' & \beta' \\ q & \delta' \end{pmatrix} \begin{pmatrix} d & -b \\ 0 & a \end{pmatrix},$$

that is,

$$\begin{pmatrix} a\alpha + bq & a\beta + b\delta \\ dq & d\delta \end{pmatrix} = \begin{pmatrix} \alpha' d & -b\alpha' + a\beta' \\ dq & -bq + a\delta' \end{pmatrix}.$$

From $\det \tau = \det \tau' = 1$, only need to determine $\alpha, \delta, \alpha', \delta'$. From above equation, $a\alpha + bq = \alpha' d$ implies that $\alpha' = (a\alpha + bq)/d$, $d\delta = -bq + a\delta'$ implies that $\delta' = (d\delta + bq)/a$. Moreover we need $\alpha\delta \equiv 1 \pmod{q}$ and $\alpha'\delta' = n^{-1}(ad\alpha\delta + bq(a\alpha + d\delta + bq)) = \alpha\delta + bq(a\alpha + d\delta + bq)/n \equiv 1 \pmod{q}$. Note that if $\alpha\delta \equiv 1 \pmod{q}$ and $(a\alpha + d\delta + bq)/n \in \mathbb{Z}$ holds, then immediately we obtain $\alpha'\delta' \equiv 1 \pmod{q}$. Consequently, we want

$$\begin{cases} a\alpha + d\delta + bq \equiv 0 \pmod{n}, \\ a\alpha + bq \equiv 0 \pmod{d}, \\ d\delta + bq \equiv 0 \pmod{a}, \\ \alpha\delta \equiv 1 \pmod{q}. \end{cases}$$

Moreover $a, d \mid n$, so $a\alpha + d\delta + bq \equiv 0 \pmod{n}$ implies that $a\alpha + bq \equiv 0 \pmod{d}$ and $d\delta + bq \equiv 0 \pmod{a}$. So these conditions hold iff

$$\begin{cases} a\alpha + d\delta + bq \equiv 0 \pmod{n}, \\ \alpha\delta \equiv 1 \pmod{q}. \end{cases}$$

If n is square-free, $ad = n$ implies that $\gcd(a, d) = 1$. So there exists $\alpha, \delta' \in \mathbb{Z}$ such that $a\alpha + d\delta' = -bq$. Note that all the solutions has the form $\alpha = \alpha_0 + td$, since $\gcd(d, q) = 1$, we can choose $t \equiv d^{-1}(1 - \alpha_0) \pmod{q}$ such that $\alpha \equiv 1 \pmod{q}$, i.e., α is invertible modulo q. Since $\gcd(n, q) = 1$, from the Chinese Remainder Theorem, one can find δ satisfying

$$\begin{cases} \delta \equiv \delta' \pmod{n}, \\ \delta \equiv \alpha^{-1} \pmod{q}. \end{cases}$$

Therefore,

$$\begin{cases} a\alpha + d\delta + bq \equiv 0 \pmod{n}, \\ \alpha\delta \equiv 1 \pmod{q}. \end{cases}$$

If n is not square-free, the algebra of Hecke operators is commutative and generated by T_p. We use induction to prove $T_{p^r}^* = \overline{\chi(p^r)}T_{p^r}$. Assume the assertion holds for $r-2$ and $r-1$. Since

$$T_{p^r} = T_p T_{p^{r-1}} - \chi(p)p^{k-1}T_{p^{r-2}},$$

we have

$$\begin{aligned}
(T_{p^r})^* &= (T_p T_{p^{r-1}})^* - \left(\chi(p)p^{k-1}T_{p^{r-2}}\right)^* \\
&= T_{p^{r-1}}^* T_p^* - \overline{\chi(p)}p^{k-1}T_{p^{r-2}}^* \\
&= \overline{\chi(p^r)}T_{p^{r-1}}T_p - \overline{\chi(p^{r-1})}p^{k-1}T_{p^{r-2}} \\
&= \overline{\chi(p^r)}\left(T_{p^{r-1}}T_p - \chi(p)p^{k-1}T_{p^{r-2}}\right) \\
&= \overline{\chi(p^r)}T_{p^r}.
\end{aligned}$$

Now let $n = p_1^{r_1}\cdots p_n^{r_n}$, then $T_n = T_{p_1^{r_1}}\cdots T_{p_n^{r_n}}$, so $T_n^* = T_{p_1^{r_1}}^*\cdots T_{p_n^{r_n}}^*$, we finish the proof. □

Theorem 6.10. In $\mathcal{S}_k(\Gamma_0(q),\chi)$, there exists an orthonormal basis $\mathcal{F}$ which consists of eigenfunctions of all the Hecke operators T_n with $\gcd(n,q)=1$.

Proof. Clearly from Corllary 6.3, Proposition 6.3 and Theorem 6.9. □

Set $\gcd(n,q)=1$. If $T_n f = \lambda_f(n)f$, then

$$\lambda_f(n)\langle f,f\rangle = \langle \lambda_f(n)f, f\rangle = \langle T_n f, f\rangle = \chi(n)\langle f, T_n f\rangle = \chi(n)\langle f, \lambda_f(n)f\rangle = \chi(n)\overline{\lambda_f(n)}\langle f,f\rangle.$$

That is,

$$\lambda_f(n) = \chi(n)\overline{\lambda_f(n)}.$$

6.6 An overview of newforms

Let $f \in \mathcal{S}_k(\Gamma_0(q),\chi)$ be a Hecke eigencuspform, i.e., a cuspform which is a common eigenfunction of all T_n with $\gcd(n,q)=1$. Say

$$T_n f = \lambda(n)f \text{ if } \gcd(n,q)=1.$$

Suppose f has the Fourier expansion

$$f(z) = \sum_{m=1}^{\infty} a(m)e(mz).$$

Recall that the m-th Fourier coefficient of $T_n f$ satisfies

$$a_n(m) = \sum_{d|\gcd(m,n)} \chi(d) d^{k-1} a\left(\frac{mn}{d^2}\right), \quad \gcd(n,q) = 1.$$

Hence,

$$\lambda(n)a(m) = \sum_{d|\gcd(m,n)} \chi(d) d^{k-1} a\left(\frac{mn}{d^2}\right), \quad \gcd(n,q) = 1.$$

Let $m = 1$,

$$\lambda(n)a(1) = a(n), \quad \gcd(n,q) = 1.$$

However, we cannot deduce that $a(1) \neq 0$, because the condition $\gcd(n,q) = 1$ does not allow us to control all the coefficients $a(m)$. In fact there exist Hecke eigencuspforms $f \neq 0$ with $a(1) = 0$. One can consider

$$\Delta(z) = \sum_{n=1}^{\infty} \tau(n) e(nz),$$

it follows that

$$\Delta(2z) = \sum_{n=1}^{\infty} \tau(n) e(2nz) = \sum_{m=1}^{\infty} a(m) e(mz),$$

where

$$a(m) = \begin{cases} \tau(m/2) & \text{if } 2 \mid m, \\ 0 & \text{if } 2 \nmid m. \end{cases}$$

In particular $a(1) = 0$.

Let q^* denote the conductor of χ. Take any positive integers q', d such that $q^* \mid q'$ and $q'd \mid q$. Let $\chi' \pmod{q'}$ be the character induced by χ. Then

$$f(z) \in \mathcal{S}_k(\Gamma_0(q'), \chi') \quad \Longrightarrow \quad f(dz) \in \mathcal{S}_k(\Gamma_0(q), \chi).$$

More precisely, let

$$\sigma_d = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix},$$

then $g(z) \coloneqq f(dz) = f(\sigma_d z)$. Note that

$$\begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} \alpha & \beta d \\ \gamma/d & \delta \end{pmatrix} \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix},$$

and let

$$\tau = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma_0(q), \quad \tau' = \begin{pmatrix} \alpha & \beta d \\ \gamma/d & \delta \end{pmatrix} \in \Gamma_0(q').$$

It follows that

$$d^{k/2}g_{|\tau} = (f_{|\sigma_d})_{|\tau} = f_{\tau'\sigma_d} = (f_{|\tau'})_{|\sigma_d} = \chi'(\tau')f_{|\sigma_d} = \chi(\tau)f_{|\sigma_d} = \chi(\tau)d^{k/2}g.$$

Since $z \mapsto dz$ is holomorphic, $g(z)$ is holomorphic on $\mathbb{H}$. Moreover, if

$$f(z) = \sum_{n=1}^{\infty} a(m)e(mz),$$

then

$$f(dz) = \sum_{m=1}^{\infty} a(m)e(mdz) = \sum_{m\equiv 0 \pmod d} a\left(\frac{m}{d}\right)e(mz),$$

which implies that $f(dz)$ has the first Fourier coefficient zero if $d \neq 1$.

We denote by $\mathcal{S}_k^\flat(\Gamma_0(q), \chi)$, the linear subspace of $\mathcal{S}_k(\Gamma_0(q), \chi)$ spanned by all forms of type $f(dz)$ where $d \mid q$ and $f \in \mathcal{S}_k(\Gamma_0(q'), \chi')$ for some $q' < q$ such that $q'd \mid q$. Here $\chi' \pmod{q'}$ is the character which is induced by $\chi \pmod q$. Then we define $\mathcal{S}_k^\sharp(\Gamma_0(q), \chi)$ the linear subspace of $\mathcal{S}_k(\Gamma_0(q), \chi)$ orthogonal to $\mathcal{S}_k^\flat(\Gamma_0(q), \chi)$ with respect to the inner product. Thus we have the orthogonal decomposition

$$\mathcal{S}_k(\Gamma_0(q), \chi) = \mathcal{S}_k^\flat(\Gamma_0(q), \chi) \oplus \mathcal{S}_k^\sharp(\Gamma_0(q), \chi).$$

Proposition 6.5. The subspace $\mathcal{S}_k^\flat, \mathcal{S}_k^\sharp$ are stable under the Hecke operators with $\gcd(n, q) = 1$.

Proof. For $\mathcal{S}_k^\flat$, let $g(z) = f(dz) = d_0^{-k/2} f_{|\sigma_{d_0}}(z)$, where $f \in \mathcal{S}_k(\Gamma_0(q'), \chi')$, $q'd_0 \mid q$ and

$$\sigma_{d_0} = \begin{pmatrix} d_0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\begin{aligned} T_n\left(f_{|\sigma_{d_0}}\right) &= n^{k/2-1} \sum_{\rho\in\Delta_n} \overline{\chi(\rho)} \left(f_{|\sigma_{d_0}}\right)_{|\rho} \\ &= n^{k/2-1} \sum_{ad=n} \sum_{b \pmod d} \chi(a) f_{\left|\sigma_{d_0}\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)\right.} \\ &= n^{k/2-1} \sum_{ad=n} \sum_{b \pmod d} \chi(a) f_{\left|\left(\begin{smallmatrix} a & d_0 b \\ 0 & d \end{smallmatrix}\right)\sigma_{d_0}\right.}. \end{aligned}$$

Since $d_0 \mid q$, $d \mid n$ and $\gcd(n, q) = 1$, we have $\gcd(d_0, d) = 1$, so $d_0 b \pmod d$ as $b \pmod d$. Hence $T_n\left(f_{|\sigma_{d_0}}\right) = (T_n f)_{|\sigma_{d_0}}$. From $f \in \mathcal{S}_k(\Gamma_0(q'), \chi')$, we have $T_n f \in \mathcal{S}_k(\Gamma_0(q'), \chi')$. Thus

$$T_n g = T_n\left(d_0^{-k/2} f_{|\sigma_{d_0}}\right) = d_0^{-k/2} T_n\left(f_{|\sigma_{d_0}}\right) = d_0^{-k/2} (T_n f)_{|\sigma_{d_0}} = (T_n f)(d_0 z).$$

For arbitrary $g \in \mathcal{S}_k^\flat$, $g = \sum c_i g_i$ where $g_i(z) = f_i(d_i z)$, which implies that $T_n g \in \mathcal{S}_k^\flat$.

For $h \in \mathcal{S}_k^\sharp$ and $g \in \mathcal{S}_k^\flat$, then

$$\langle T_n h, g\rangle = \chi(n)\,\langle h, T_n g\rangle = \chi(n)0 = 0.$$

Therefore $T_n h \in \mathcal{S}_k^\sharp$ since $T_n h \in \mathcal{S}_k(\Gamma_0(q'), \chi')$. □

The Hecke eigencuspforms in the space $\mathcal{S}_k^\sharp(\Gamma_0(q), \chi)$ are called **newforms**.

The multiplicity-one principle holds true in the space $\mathcal{S}_k^\sharp(\Gamma_0(q), \chi)$. That is, for each sequence $\{\lambda(n)\}_{\gcd(n,q)=1}$ there is at most one cuspform $f \neq 0$ (up to a scalar) which is an eigenfunction of T_n with prescribed numbers $\lambda(n)$ for $\gcd(n, q) = 1$ as eigenvalues. Recall that $T_n T_m = T_m T_n$, use this for $\gcd(n, q) = 1$ and for all m,

$$T_n(T_m f) = T_m(T_n f) = T_m(\lambda(n) f) = \lambda(n)(T_m f).$$

That is, $T_m f$ and f have the same eigenvalue $\lambda(n)$ for all T_n with $\gcd(n, q) = 1$. By the multiplicity-one principle, $T_m f = \lambda(m) f$ for some $\lambda(m)$. This equation holds for all m.

Proposition 6.6. Suppose f is a newform, then $T_n f = \lambda(n) f$ holds for all n with $\gcd(n, q) = 1$. The first coefficient in the Fourier expansion does not vanish, so one can normalize f by setting $a(1) = 1$. In this case

$$a(n) = \lambda(n) \text{ for all } n.$$

Proof. Immediately from the fact

$$\lambda(n)a(1) = a(n), \quad \gcd(n, q) = 1.$$

□

6.7 Hecke eigencuspforms for a primitive character

Now consider case $\chi \pmod q$ is a primitive character. Thus in this case $\mathcal{S}_k^\flat$ is zero and $\mathcal{S}_k^\sharp$ is the whole space $\mathcal{S}_k(\Gamma_0(q), \chi)$.

For any d, the operator A_d defined by

$$A_d = \frac{1}{d} \sum_{b \pmod d} \begin{pmatrix} 1 & b/d \\ 0 & 1 \end{pmatrix}$$

acts on the Fourier series

$$f(z) = \sum_m a(m) e(mz)$$

by selecting the terms with $m \equiv 0 \pmod{d}$, i.e., we have

$$\begin{aligned}(A_d f)(z) &= \frac{1}{d} \sum_{b \pmod{d}} f\left(z + \frac{b}{d}\right) \\ &= \frac{1}{d} \sum_{b \pmod{d}} \sum_{m} a(m) e(mz) e\left(m\frac{b}{d}\right) \\ &= \sum_{m \equiv 0 \pmod{d}} a(m) e(mz).\end{aligned}$$

Hence the sifting operator S_q defined by

$$S_q = \sum_{d \mid q} \mu(d) A_d$$

acts on the Fourier series by selecting the terms with $\gcd(m, q) = 1$, i.e., we have

$$\begin{aligned}(S_q f)(z) &= \sum_{d \mid q} \mu(d) \sum_{m \equiv 0 \pmod{d}} a(m) e(mz) \\ &= \sum_{m=1}^{\infty} a(m) e(mz) \sum_{d \mid \gcd(m,q)} \mu(d) \\ &= \sum_{\gcd(m,q)=1} a(m) e(mz).\end{aligned}$$

Remark. One can show that $S_q \colon \mathcal{S}_k(\Gamma_0(q), \chi) \to \mathcal{S}_k(\Gamma_0(qq_1), \chi)$, where q_1 is the product of prime factors of q.

Proof. For $d \mid q$ squarefree, we have

$$T_d = \frac{1}{d} \sum_{b \pmod{d}} \begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix}.$$

Therefore,

$$\begin{aligned}(T_d f)(z) &= \frac{1}{d} \sum_{b \pmod{d}} \sum_{m} a(m) e\left(m\frac{z+b}{d}\right) \\ &= \sum_{m \equiv 0 \pmod{d}} a(m) e\left(\frac{m}{d} z\right).\end{aligned}$$

Let $(V_d f)(z) = f(dz)$, then $A_d = V_d T_d$. Indeed,

$$((V_d T_d) f)(z) = (V_d(T_d f))(z) = (T_d f)(dz) = \sum_{m \equiv 0 \pmod{d}} a(m) e(mz) = (A_d f)(z).$$

We claim $V_d\colon \mathcal{S}_k(\Gamma_0(q),\chi) \to \mathcal{S}_k(\Gamma_0(qd),\chi)$. Let $g(z) = f(dz)$, for arbitrary

$$\gamma = \begin{pmatrix} a & b \\ c & h \end{pmatrix} \in \Gamma_0(dq).$$

Then $dq \mid c$, which implies that $q \mid c/d$. Moreover

$$d\gamma z = d\frac{az+b}{cz+h} = \frac{adz+bd}{cz+h}.$$

On the other hand,

$$\gamma' = \begin{pmatrix} a & bd \\ c/d & h \end{pmatrix} \in \Gamma_0(q)$$

satisfies

$$\gamma'(dz) = \frac{adz+bd}{cz+h} = d\gamma z.$$

Since $f(\gamma' dz) = \chi(h)((c/d)(dz)+h)^k f(dz)$, we have $g(\gamma z) = f(d\gamma z) = \chi(h)(cz+h)^k g(z)$. Moreover, $f(dz)$ is clearly holomorphic. If

$$f(z) = \sum_{m\geq 1} a(m)e(mz),$$

then

$$g(z) = f(dz) = \sum_{m\geq 1} a(m)e(mdz)$$

also has no nonzero constant term. That is, $g \in \mathcal{S}_k(\Gamma_0(qd),\chi)$. Now $T_d\colon \mathcal{S}_k(\Gamma_0(q),\chi) \to \mathcal{S}_k(\Gamma_0(q),\chi)$ and $V_d\colon \mathcal{S}_k(\Gamma_0(q),\chi) \to \mathcal{S}_k(\Gamma_0(qd),\chi)$, we have $A_d\colon \mathcal{S}_k(\Gamma_0(q),\chi) \to \mathcal{S}_k(\Gamma_0(qd),\chi) \subset \mathcal{S}_k(\Gamma_0(qq_1),\chi)$, we finish the proof. □

Lemma 6.2. If $f \in \mathcal{S}_k(\Gamma_0(q),\chi)$ has the Fourier coefficients $a(n) = 0$ for all n with $\gcd(n,q) = 1$, then $f = 0$. In other words, the sifting operator S_q has zero kernel.

Proof. Given

$$\tau = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q),$$

from the relation $f_{|\tau} = \chi(\tau)f$ written at the point $(z-d)/c$, we get

$$z^{-k} f\left(\frac{a}{c} - \frac{1}{cz}\right) = \chi(d) f\left(-\frac{d}{c} + \frac{z}{c}\right).$$

Expanding both sides into Fourier series, we obtain a kind of Poisson's summation formula,

$$z^{-k} \sum_{m\geq 1} a(m) e\left(\frac{am}{c}\right) e\left(-\frac{m}{cz}\right) = \chi(d) \sum_{n\geq 1} a(n) e\left(-\frac{dn}{c}\right) e\left(\frac{nz}{c}\right)$$

for any $0 < c \equiv 0 \pmod{q}$ and $ad \equiv 1 \pmod{c}$. Hence, summing over $a \pmod{c}$ with $(a,c) = 1$ for $c = q$, we get

$$z^{-k} \sum_{m \geq 1} a(m) \left(\sideset{}{^*}\sum_{a \pmod{q}} e\left(\frac{am}{q}\right) \right) e\left(-\frac{m}{qz}\right) = \sum_{n \geq 1} a(n) \left(\sideset{}{^*}\sum_{d \pmod{q}} \chi(d) e\left(-\frac{dn}{q}\right) \right) e\left(\frac{nz}{q}\right).$$

Here

$$S(m, 0; q) = \sideset{}{^*}\sum_{a \pmod{q}} e\left(\frac{am}{q}\right)$$

be the Ramanujan sum, and

$$S_\chi(-n, 0; q) = \sideset{}{^*}\sum_{d \pmod{q}} \chi(d) e\left(-\frac{dn}{q}\right)$$

be the Gauss sum. In Appendix A, we proved

$$\chi(a)\tau\left(\overline{\chi}\right) = \sum_{b \pmod{m}} \overline{\chi(b)} e\left(\frac{ab}{m}\right).$$

Put χ be $\overline{\chi}$ and $a = -n$, we have

$$S_\chi(-n, 0; q) = \overline{\chi(-n)}\tau(\chi).$$

Since χ is primitive character,

$$S_\chi(-n, 0; q) = \begin{cases} \overline{\chi(-n)}\tau(\chi) & \text{if } \gcd(n, q) = 1, \\ 0 & \text{if } \gcd(n, q) > 1. \end{cases}$$

But $a(n) = 0$ when $\gcd(n, q) = 1$, so every term on the right-hand side vanishes, we conclude that every term on the left-hand side also vanishes (the uniqueness of Fourier series expansion), i.e., $a(m)S(m, 0; q) = 0$ for every m. If q is squarefree, by Lemma A.1,

$$S(m, 0; q) = \mu(q/\gcd(m, q))\varphi(\gcd(m, q)) \neq 0.$$

Therefore $a(m) = 0$ for every m, proving the Lemma in this case.

If q is not squarefree then the assertion of the Lemma remains true, but our proof needs substantial modifications. To this end we generalize the above relation between the Ramanujan and Gauss sums. First we multiply the Poisson formula by a function $\psi(a)$, and only then we sum over the primitive classes. We obtain

$$z^{-k} \sum_{m=1}^{\infty} a(m)S(m) e\left(-\frac{m}{qz}\right) = \sum_{n=1}^{\infty} a(n)S_\chi(n) e\left(\frac{nz}{q}\right)$$

where

$$S(m) = \sum_{a \pmod q}^{*} e\left(\frac{am}{q}\right)\psi(a)$$
$$S_\chi(n) = \sum_{d \pmod q} \chi(d) e\left(-\frac{dn}{q}\right)\psi(a)$$

and $ad \equiv 1 \pmod q$. For a given m we want $S(m) \neq 0$, whereas $S_\chi(n) = 0$ for all n with $\gcd(n,q) \neq 1$. We can permit $\psi(a)$ to depend on m. To construct such ψ we factor the modulus $q = q_1 q_2$ so that q_1 has primes in m and q_2 is coprime with m. It follows that $\gcd(q_1, q_2) = 1$ and $\gcd(m, q_2) = 1$. We put $q_0 = \prod_{p|\gcd(m,q)} p$. Take $\overline{q_1}$ and $\overline{q_2}$ satisfies $q_1\overline{q_1} \equiv 1 \pmod{q_2}$ and $q_2\overline{q_2} \equiv 1 \pmod{q_1}$. By the Chinese Remainder Theorem, we use only the residue classes $a \pmod q$ of type

$$a = bq_1\overline{q_1} + \left(1 + c\frac{q_1}{q_0}\right) q_2\overline{q_2}$$

where b runs over the classes modulo q_2 with $\gcd(b, q_2) = 1$ and c runs over the classes modulo q_0 with $\gcd(1 + cq_1/q_0, q_1) = 1$. For $\psi(a)$ we choose a character of conductor q_2 restricted to the above classes. From

$$a = bq_1\overline{q_1} + \left(1 + c\frac{q_1}{q_0}\right) q_2\overline{q_2},$$

we then have

$$e\left(\frac{am}{q}\right) = e\left(\frac{bm\overline{q_1}}{q_2}\right) e\left(\frac{m\overline{q_2}}{q_1}\left(1 + c\frac{q_1}{q_0}\right)\right),$$

which implies that

$$S(m) = \sum_{b \pmod{q_2}} \psi(b) e\left(\frac{bm\overline{q_1}}{q_2}\right) \sum_{\substack{c \pmod{q_0} \\ \gcd(1+cq_1/q_0, q_1)=1}} e\left(\frac{m\overline{q_2}}{q_1}\left(1 + c\frac{q_1}{q_0}\right)\right).$$

In Appendix A we proved:

$$\chi(a)\tau(\overline{\chi}) = \sum_{b \pmod m} \overline{\chi(b)} e\left(\frac{ab}{m}\right),$$

take $m = q_2$, $a = m\overline{q_1}$ and $\chi = \overline{\psi}$, we obtain

$$\sum_{b \pmod{q_2}} \psi(b) e\left(\frac{bm\overline{q_1}}{q_2}\right) = \overline{\psi(m\overline{q_1})}\tau(\psi) = \tau(\psi)\psi(q_1)\overline{\psi(m)}.$$

Now compute

$$\sum_{\substack{c \pmod{q_0} \\ \gcd(1+cq_1/q_0, q_1)=1}} e\left(\frac{m\overline{q_2}}{q_1}\left(1 + c\frac{q_1}{q_0}\right)\right) = \sum_{\substack{c \pmod{q_0} \\ \gcd(1+cq_1/q_0, q_1)=1}} e\left(\frac{m\overline{q_2}}{q_1}\right) e\left(\frac{mc\overline{q_2}}{q_0}\right).$$

Since $q_0 = \prod_{p|\gcd(m,q)} p$, we have $q_0 \mid m$, so $mc\overline{q_2}/q_0 \in \mathbb{Z}$. Consequently,

$$\sum_{\substack{c \pmod{q_0} \\ \gcd(1+cq_1/q_0, q_1)=1}} e\left(\frac{m\overline{q_2}}{q_1}\left(1+c\frac{q_1}{q_0}\right)\right) = e\left(\frac{m\overline{q_2}}{q_1}\right) \cdot \#\left\{c \pmod{q_0} \mid \gcd\left(1+c\frac{q_1}{q_0}, q_1\right) = 1\right\}.$$

Write $q_1 = \prod_{p|q_1} p^{\alpha_p}$, since $q_0 = \prod_{p|q_1} p$, we have $q_1/q_0 = \prod_{p|q_1} p^{\alpha_p - 1}$. $\gcd(1 + cq_1/q_0, q_1) = 1$ iff for every $p \mid q_1$ we have $p \nmid (1+cq_1/q_0)$. Since q_0 is squarefree, given $c \pmod{q_0}$ is equivalent to give all the $c \pmod{p}$ where $p \mid q_1$. If $p^2 \mid q_1$, $\alpha_p \geq 2$, so $q_1/q_0 \equiv 0 \pmod{p}$, and so $1+cq_1/q_0 \equiv 1 \pmod{p}$. Thus $p \nmid (1 + cq_1/q_0)$. That is, all the residue classes modulo p satisfy the condition. If $p \parallel q_1$, i.e., $\alpha_p = 1$, we have $p \nmid q_1/q_0$. $1 + cq_1/q_0 \equiv 0 \pmod{p}$ iff $c \equiv -(q_1/q_0)^{-1} \pmod{p}$, so there are $p - 1$ admissible classes. By the Chinese Remainder Theorem,

$$\#\left\{c \pmod{q_0} \mid \gcd\left(1+c\frac{q_1}{q_0}, q_1\right) = 1\right\} = \prod_{\substack{p|q_1 \\ p^2 \nmid q_1}} p \cdot \prod_{p \parallel q_1} (p-1) = q_0 \prod_{p \parallel q_1} \left(1 - \frac{1}{p}\right).$$

Hence

$$S(m) = \tau(\psi)\psi(q_1)\overline{\psi(m)} e\left(\frac{m\overline{q_2}}{q_1}\right) q_0 \prod_{p \parallel q_1} \left(1 - \frac{1}{p}\right) \neq 0$$

as desired. In order to compute $S_\chi(n)$ we factor $\chi \pmod{q}$ into the primitive characters $\chi_1 \pmod{q_1}$ and $\chi_2 \pmod{q_2}$. Since

$$a = bq_1\overline{q_1} + \left(1 + c\frac{q_1}{q_0}\right) q_2\overline{q_2},$$

and $ad \equiv 1 \pmod{q}$, d must satisfy $d \equiv \overline{b} \pmod{q_2}$ and $d \equiv \overline{(1 + cq_1/q_0)} \pmod{q_1}$. By the Chinese Remainder Theorem,

$$d = \overline{b}q_1\overline{q_1} + \overline{\left(1 + c\frac{q_1}{q_0}\right)} q_2\overline{q_2}.$$

We obtain

$$\begin{aligned} S_\chi(n) &= \sum_{b \pmod{q_2}} \psi(b)\overline{\chi(b)} e\left(-\frac{n\overline{b}\overline{q_1}}{q_2}\right) \sum_{\substack{c \pmod{q_0} \\ \gcd(1+cq_1/q_0, q_1)=1}} \overline{\chi_1\left(1+c\frac{q_1}{q_0}\right)} e\left(-n\frac{\overline{q_2}}{q_1}\overline{\left(1+c\frac{q_1}{q_0}\right)}\right) \\ &= \sum_{b \pmod{q_2}} \overline{\psi(b)}\chi_2(b) e\left(-bn\frac{\overline{q_1}}{q_2}\right) \sum_{\substack{c \pmod{q_0} \\ \gcd(1+cq_1/q_0, q_1)=1}} \overline{\chi_1\left(1+c\frac{q_1}{q_0}\right)} e\left(-n\frac{\overline{q_2}}{q_1}\overline{\left(1+c\frac{q_1}{q_0}\right)}\right) \end{aligned}$$

by $b \mapsto \overline{b}$. Now $\overline{\psi}\chi_2$ is primitive character modulo q_2. From Appendix A,

$$\chi(a)\tau(\overline{\chi}) = \sum_{b \pmod{m}} \overline{\chi(b)} e\left(\frac{ab}{m}\right)$$

vanishes when χ is primitive and $\gcd(a,m)>1$. So when $\gcd(n,q_2)\neq 1$,

$$\sum_{b \pmod{q_2}} \overline{\psi(b)}\chi_2(b)e\left(-bn\frac{\overline{q_1}}{q_2}\right)=0.$$

Now consider the c-sum

$$\sum_{\substack{c \pmod{q_0}\\ \gcd(1+cq_1/q_0,q_1)=1}} \overline{\chi_1\left(1+c\frac{q_1}{q_0}\right)}e\left(-n\frac{\overline{q_2}}{q_1}\overline{\left(1+c\frac{q_1}{q_0}\right)}\right).$$

Let $x=1+cq_1/q_0$, then clearly $x\equiv 1 \pmod{q_1/q_0}$. For the converse, every $x\equiv 1 \pmod{q_1/q_0}$ can be uniquely written as $x=1+cq_1/q_0$ where $c \pmod{q_0}$. The c-sum becomes

$$\sum_{\substack{x \pmod{q_1}, x\equiv 1 \pmod{q_1/q_0}\\ \gcd(x,q_1)=1}} \overline{\chi_1(x)}e\left(-n\frac{\overline{q_2}}{q_1}\overline{x}\right).$$

Let $d^*=\gcd(n,q_1)$, $n=d^*n^*$ and $q_1=d^*q^*$. Since $q_0=\prod_{p|q_1}p$, $\gcd(q_0,d^*)=\prod_{p|\gcd(q_1,n)}p$. Then $c \pmod{q_0}$ implies that

$$c=c_1\frac{q_0}{\gcd(q_0,d^*)}+c_2,\quad c_1 \pmod{\gcd(q_0,d^*)} \text{ and } c_2 \pmod{\frac{q_0}{\gcd(q_0,d^*)}}.$$

Then

$$x=c_1\frac{q_0}{\gcd(q_0,d^*)}\cdot\frac{q_1}{q_0}+c_2\frac{q_1}{q_0}+1.$$

When c_2 is fixed, x satisfies

$$x\equiv 1+c_2\frac{q_1}{q_0} \pmod{\frac{q_1}{\gcd(q_0,d^*)}}.$$

It follows that

$$\overline{x}-\overline{1+c_2\frac{q_1}{q_0}}\equiv 0 \pmod{\frac{q_1}{\gcd(q_0,d^*)}}.$$

Thus

$$e\left(-n\frac{\overline{q_2}}{q_1}\overline{x}\right)=e\left(-n\frac{\overline{q_2}}{q_1}\left(\overline{x}-\overline{1+c_2\frac{q_1}{q_0}}+\overline{1+c_2\frac{q_1}{q_0}}\right)\right)=e\left(-n\frac{\overline{q_2}}{q_1}\overline{1+c_2\frac{q_1}{q_0}}\right)$$

from $\gcd(q_0,d^*)\mid d^*=\gcd(n,q_1)\mid n$. Consequently the c-sum becomes

$$\sum_{c_2 \pmod{\frac{q_0}{\gcd(q_0,d^*)}}} e\left(-n\frac{\overline{q_2}}{q_1}\overline{1+c_2\frac{q_1}{q_0}}\right)\sum_{\substack{x \pmod{q_1}\\ x\equiv c_2\frac{q_1}{q_0}+1 \pmod{\frac{q_1}{\gcd(q_0,d^*)}}}}\overline{\chi_1(x)},$$

the x-sum vanishes as $\gcd(n,q_1)\neq 1$, from χ_1 primitive character. The c-sum also vanishes as $\gcd(n,q_1)\neq 1$.

Therefore, $S_\chi(n)=0$ as $\gcd(n,q)=1$. Then $a(m)=0$ for all $m\geq 1$, $f=0$.

We still have a problem if χ_2 is the only primitive character to modulus q_2. This is the case for $q_2 = 3$ and $q_2 = 4$. If $q_2 = 3$ we take ψ the principal character, and the above argument still works because

$$\sum_{b \pmod 3} \psi(b) e\left(-bm\frac{\overline{q_1}}{3}\right) = e\left(\frac{m\overline{q_1}}{3}\right) + e\left(\frac{2m\overline{q_1}}{3}\right) = e\left(\frac{1}{3}\right) + e\left(\frac{2}{3}\right) = -1 \neq 0.$$

If $q_2 = 4$, the unique primitive character modulo 4 is

$$\chi_2(b) = \chi_4(b) = \left(\frac{-1}{b}\right).$$

From $\mathcal{S}_k(\Gamma_0(q), \chi) \subset \mathcal{S}_k(\Gamma_0(2q), \chi)$, and $\gcd(n, 2q) = 1$ implies that $\gcd(n, q) = 1$, we can apply the above arguments for the modulus $2q$ in place of q. We take

$$\psi(b) = \chi_8(b) = \left(\frac{2}{b}\right).$$

Since χ_8 and $\chi_8\chi_4$ are both primitive, the previous result is valid. This completes the proof of Lemma for any primitive character $\chi \pmod q$. □

Lemma 6.2 implies the multiplicity-one principle. We get

Lemma 6.3. A Hecke eigencuspform $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$ can be normalized to have its first Fourier coefficient $a(1) = 1$, and the normalized form is determined uniquely by the eigenvalues of T_n with $\gcd(n, q) = 1$.

Theorem 6.11. If $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$ is an eigenfunction of every T_n with $\gcd(n, q) = 1$, then it is an eigenfunction of the whole algebra $\mathcal{T} = \{T_m : m \in \mathbb{N}\}$, i.e.,

$$T_m f = \lambda(m) f \quad \text{for all } m \in \mathbb{N}.$$

Let W be the **Fricke involution**, defined by

$$Wf = f_{|\omega} \quad \text{where } \omega = \begin{pmatrix} 0 & -1 \\ q & 0 \end{pmatrix},$$

i.e.,

$$(Wf)(z) = q^{-k/2} z^{-k} f(-1/(qz)).$$

We have

$$W^2 = (-1)^k.$$

Moreover,

$$W \colon \mathcal{S}_k(\Gamma_0(q), \chi) \longrightarrow \mathcal{S}_k\left(\Gamma_0(q), \overline{\chi}\right).$$

The Fricke involution is an isometry, i.e., we have

$$\langle Wf, Wg\rangle = \langle f, g\rangle\,.$$

Theorem 6.12. If $\gcd(n,q)=1$ we have

$$WT_n^{\chi} = \chi(n)T_n^{\overline{\chi}}W.$$

Let K be the **complex conjugation operator** defined by

$$(Kf)(z) = \overline{f}\,(-\overline{z})\,.$$

Note that

$$f(z)=\sum a(n)e(nz) \quad \Longrightarrow \quad (Kf)(z)=\sum \overline{a(n)}e(nz).$$

K is NOT linear over $\mathbb{C}$. Indeed

$$K(\lambda f) = \overline{\lambda}Kf \quad \text{if } \lambda\in\mathbb{C}.$$

We have

$$K^2=1, \quad WK=(-1)^kKW$$

$$K\colon \mathcal{S}_k(\Gamma_0(q),\chi)\longrightarrow \mathcal{S}_k\left(\Gamma_0(q),\overline{\chi}\right)$$

$$\langle Kf, Kg\rangle = \langle f, g\rangle$$

$$KT_n^{\chi} = T_n^{\overline{\chi}}K \text{ for all } n.$$

Set

$$\overline{W} = KW,$$

then

$$\overline{W}^2=1, \quad \overline{W}(\eta f)=\overline{\eta}\overline{W}f \text{ if } \eta\in\mathbb{C}$$

$$\overline{W}\colon \mathcal{S}_k(\Gamma_0(q),\chi)\longrightarrow \mathcal{S}_k(\Gamma_0(q),\chi)$$

$$T_n\overline{W} = \chi(n)\overline{W}T_n \text{ if } \gcd(n,q)=1.$$

This implies that

$$T_n\left(\overline{W}f\right) = \lambda(n)\left(\overline{W}f\right), \quad \gcd(n,q)=1.$$

By the multiplicity-one principle, one deduce

$$\overline{W}f = \eta f.$$

Theorem 6.13. A Hecke eigencuspform $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$ is also an eigenfunction of the involution operator $\overline{W}$, i.e.,

$$\overline{W} f = \eta f.$$

We have $|\eta| = 1$. Indeed, $f = \overline{W}^2 f = \overline{W}(\eta f) = \overline{\eta}\overline{W} f = \overline{\eta} \cdot \eta f = |\eta|^2 f$.

Recall that Theorem 6.2 we proved

$$T_m T_n = \sum_{d \mid \gcd(m,n)} \chi(d) d^{k-1} T_{mnd^{-2}},$$

which implies that for all $m, n \geq 1$,

$$\lambda(m)\lambda(n) = \sum_{d \mid \gcd(m,n)} \chi(d) d^{k-1} \lambda(mnd^{-2}).$$

In particular, for any m and $d \mid q$ we obtain $\lambda(dm) = \lambda(d)\lambda(m)$. Hence

$$(A_d f)(z) = \sum_{n \equiv 0 \pmod{d}} \lambda(n) e(nz) = \sum_{n} \lambda(nd) e(dnz) = \lambda(d) f(dz);$$

in other words

$$A_d f = \lambda(d) d^{-k/2} f_{|\left(\begin{smallmatrix} d & \\ & 1 \end{smallmatrix}\right)} \quad \text{if } d \mid q.$$

Theorem 6.14. The eigenvalue of the involution $\overline{W}$ of a normalized Hecke eigencuspform $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$ is given by

$$\eta = \tau\left(\overline{\chi}\right) \lambda(q) q^{-k/2}.$$

Theorem 6.15. For $p \mid q$ the eigenvalues $\lambda(p)$ of T_p in the space $\mathcal{S}_k(\Gamma_0(q), \chi)$ satisfy

$$|\lambda(p)| = p^{(k-1)/2}.$$

7 Automorphic L-functions

7.1 The Hecke L-functions

Let $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$. Consider simultaneously the Fourier expansions at cusps ∞ and 0, say

$$f(z) = \sum_{1}^{\infty} a(n) e(nz)$$

$$g(z) = \sum_{1}^{\infty} b(n) e(nz)$$

where

$$g = f_{|\omega} \text{ with } \omega = \begin{pmatrix} & -1 \\ q & \end{pmatrix}.$$

Recall that $g \in \mathcal{S}_k(\Gamma_0(q), \overline{\chi})$. To these Fourier series we associate the Hecke L-functions

$$L_f(s) = \sum_1^\infty a(n) n^{-s}$$

$$L_g(s) = \sum_1^\infty b(n) n^{-s}.$$

In the half-plane we can show that

$$(2\pi)^{-s}\Gamma(s)L_f(s) = \int_0^\infty f(\mathrm{i}y) y^{s-1} \mathrm{d}y$$

$$(2\pi)^{-s}\Gamma(s)L_g(s) = \int_0^\infty g(\mathrm{i}y) y^{s-1} \mathrm{d}y.$$

Since f and g have exponential decay at 0 and ∞, we have

Theorem 7.1. The L-functions $L_f(s)$ and $L_g(s)$ associated to the cusp forms $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$ and $g = f_{|\omega} \in \mathcal{S}_k(\Gamma_0(q), \overline{\chi})$ have analytic continuation over $\mathbb{C}$, and they are entire functions. Actually the so-called complete L-functions

$$\Lambda_f(s) = \left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s) L_f(s)$$

$$\Lambda_g(s) = \left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s) L_g(s)$$

are entire functions, bounded on vertical strips.

Theorem 7.2. The complete L-functions satisfy the functional equation

$$\Lambda_f(s) = \mathrm{i}^k \Lambda_g(k-s).$$

If f is an eigenfunction of the operator $\overline{W} = KW$, say

$$\overline{W} f = \eta f,$$

then $\overline{b(n)} = \eta a(n)$, so the functional equation becomes

$$\left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s) L_f(s) = \mathrm{i}^k \eta \overline{\left(\frac{\sqrt{q}}{2\pi}\right)^{k-s}} \Gamma(k-s) \overline{L_f(k-\overline{s})}.$$

In particular if χ is real character and f has real Fourier coefficients and is an eigenfunction of W, say

$$Wf = \pm \mathrm{i}^k f,$$

then

$$\left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s)L_f(s) = \pm\left(\frac{\sqrt{q}}{2\pi}\right)^{k-s} \Gamma(k-s)L_f(k-s).$$

Theorem 7.3 (Hecke)**.** Suppose f and g are given by the Fourier series

$$f(z) = \sum_0^\infty a_n e(nz)$$

$$g(z) = \sum_0^\infty b_n e(nz)$$

with coefficients $a_n, b_n \ll n^\alpha$ for $n \geq 1$, where α is a positive constant. Put

$$L_f(s) = \sum_1^\infty a_n n^{-s}$$

$$L_g(s) = \sum_1^\infty b_n n^{-s}$$

and

$$\Lambda_f(s) = \left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s)L_f(s)$$

$$\Lambda_g(s) = \left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s)L_g(s)$$

where q is a given positive number (not necessarily an integer). Put

$$\omega = \begin{pmatrix} & -1 \\ q & \end{pmatrix}$$

and $\left(f_{|\omega}\right)(z) = \left(\sqrt{q}z\right)^{-k} f(-1/(qz))$, where k is a given positive integer. Then the following assertions are equivalent.

(A_1) The functions f, g are connected by

$$g = f_{|\omega}.$$

(A_2) Both $\Lambda_f(s)$ and $\Lambda_g(s)$ have meromorphic continuation over the whole $\mathbb{C}$,

$$\Lambda_f(s) + a_o s^{-1} + b_o \mathrm{i}^k (k-s)^{-1}$$

$$\Lambda_g(s) + b_o s^{-1} + a_o \mathrm{i}^{-k} (k-s)^{-1}$$

are entire and bounded on vertical strips, and they satisfy

$$\Lambda_f(s) = \mathrm{i}^k \Lambda_g(k-s).$$

7.2 Twisting automorphic forms and L-functions

Theorem 7.4. Let $f \in \mathcal{M}_k(\Gamma_0(q), \chi)$, where χ is a character of conductor q^* with $q^* \mid q$. Let ψ be a primitive character of conductor r. If f has the Fourier series

$$f(z) = \sum_0^\infty a(n)e(nz),$$

then the twisted series

$$f_\psi(z) = \sum_0^\infty \psi(n)a(n)e(nz)$$

belongs to $\mathcal{M}_k(\Gamma_0(N), \chi\psi^2)$, where $N = \operatorname{lcm}(q, q^*r, r^2)$. Moreover if f is a cusp form, so is f_ψ.

Theorem 7.5. Let $f \in \mathcal{S}_k(\Gamma_0(q), \chi)$, where χ is a character to modulus q. Let ψ be a primitive character of conductor r with $\gcd(q, r) = 1$. Then

$$f_{\psi|\omega_N} = w(\psi)g_{\overline{\psi}}$$

where

$$\omega_N = \begin{pmatrix} & -1 \\ N & \end{pmatrix}$$

for $N = qr^2$, $g = f_{|\omega}$ for

$$\omega = \omega_q = \begin{pmatrix} & -1 \\ q & \end{pmatrix}$$

and

$$w(\psi) = \chi(r)\psi(q)\tau(\psi)^2 r^{-1}.$$

Sketch of Proof. We have

$$\begin{pmatrix} 1 & u/r \\ & 1 \end{pmatrix} \begin{pmatrix} & -1 \\ N & \end{pmatrix} = r \cdot \begin{pmatrix} & -1 \\ q & \end{pmatrix} \begin{pmatrix} r & -v \\ -uq & * \end{pmatrix} \begin{pmatrix} 1 & u/r \\ & 1 \end{pmatrix}.$$

We choose $uvq \equiv -1 \pmod r$, this identity yields

$$f_{\left|\left(\begin{smallmatrix} 1 & u/r \\ & 1 \end{smallmatrix}\right)\omega_N\right.} = \chi(r)g_{\left|\left(\begin{smallmatrix} 1 & u/r \\ & 1 \end{smallmatrix}\right)\right.}.$$

Then

$$\sum_{u \pmod r}^* \mathrm{LHS}\,\overline{\psi(u)} = \sum_{v \pmod r}^* \mathrm{RHS}\,\psi(-vq).$$

□

Now let $L_f(s,\psi)$ denote the Hecke L-function associated with the twisted form f_ψ, i.e.,

$$L_f(s,\psi) = \sum_1^\infty \psi(n)a(n)n^{-s},$$

and put

$$\Lambda_f(s,\psi) = \left(\frac{\sqrt{N}}{2\pi}\right)^s \Gamma(s)L_f(s,\psi).$$

By Theorems 7.5 and 7.3 we deduce

Theorem 7.6. Suppose $f \in \mathcal{S}_k(\Gamma_0(q),\chi)$, ψ is a primitive character of conductor r with $\gcd(q,r) = 1$, $N = qr^2$ and $w(\psi)$ is given in Theorem 7.5. Then $\Lambda_f(s,\psi)$ is an entire function, bounded in vertical strips, and it satisfies the functional equation

$$\Lambda_f(s,\psi) = \mathrm{i}^k w(\psi)\Lambda_g\left(k-s,\overline{\psi}\right)$$

where $g = f_{|\omega}$ for

$$\omega = \omega_q = \begin{pmatrix} & -1 \\ q & \end{pmatrix}.$$

7.3 Converse theorems

Theorem 7.7 (Hecke)**.** Let k be a positive, even integer. Suppose f is given by Fourier series

$$f(z) = \sum_0^\infty a_n e(nz)$$

with the coefficients $a_n \ll n^\alpha$ for all $n \geq 1$. Then $f \in \mathcal{M}_k(\Gamma_0(1))$ if and only if the function

$$\Lambda_f(s) = (2\pi)^{-s}\Gamma(s)\sum_1^\infty a_n n^{-s}$$

can be analytically continued over the whole $\mathbb{C}$, $\Lambda_f(s) + a_o s^{-1} + \mathrm{i}^k a_o(k-s)^{-1}$ is entire and bounded in vertical strips, and satisfies the functional equation

$$\Lambda_f(s) = \mathrm{i}^k \Lambda_f(k-s).$$

Theorem 7.8 (Weil)**.** Let k be a positive, even integer, and χ a character to modulus $q \geq 1$. Suppose f, g are given by the Fourier series

$$f(z) = \sum_0^\infty a_n e(nz)$$

$$g(z) = \sum_0^\infty b_n e(nz)$$

with coefficients $a_n, b_n \ll n^\alpha$ for $n \geq 1$, where α is a positive constant. Suppose $\Lambda_f(s), \Lambda_g(s)$ defined by

$$\Lambda_f(s) = \left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s) L_f(s)$$

$$\Lambda_g(s) = \left(\frac{\sqrt{q}}{2\pi}\right)^s \Gamma(s) L_g(s)$$

satisfy (A_2) in Theorem 7.3. Furthermore let $\mathcal{R}$ be a set of prime numbers coprime with q which meets every primitive residue class, i.e., for any $c > 0$ and any a with $\gcd(a, c) = 1$ there exists $r \in \mathcal{R}$ such that $r \equiv a \pmod{c}$. Suppose for any primitive character ψ of conductor $r \in \mathcal{R}$ the functions $\Lambda_f(s, \psi), \Lambda_g(s, \psi)$ defined by

$$\Lambda_f(s, \psi) = \left(\frac{\sqrt{N}}{2\pi}\right)^s \Gamma(s) L_f(s, \psi).$$

$$\Lambda_g(s, \psi) = \left(\frac{\sqrt{N}}{2\pi}\right)^s \Gamma(s) L_g(s, \psi).$$

with $N = qr^2$ are entire, bounded in vertical strips and staisfy the functional equation

$$\Lambda_f(s, \psi) = \mathrm{i}^k w(\psi) \Lambda_g\left(k - s, \overline{\psi}\right)$$

with $w(\psi)$ given by

$$w(\psi) = \chi(r)\psi(q)\tau(\psi)^2 r^{-1}.$$

Then $f \in \mathcal{M}_k(\Gamma_0(q), \chi)$, $g \in \mathcal{M}_k\left(\Gamma_0(q), \overline{\chi}\right)$ and $g = f_{|\omega}$, where

$$\omega = \omega_q = \begin{pmatrix} & -1 \\ q & \end{pmatrix}.$$

Moreover f, g are cusp forms if $L_f(s)$ or $L_g(s)$ converges absolutely on some line $\mathrm{Re}(s) = \sigma$ with $0 < \sigma < k$.

Theorem 7.9 (Criterion for Cusp Forms II)**.** Suppose f is an automorphic form of weight $k > 0$ for a group Γ such that

$$y^k f(z) \longrightarrow 0 \quad \text{as} \quad y \longrightarrow 0$$

uniformly on $\mathbb{H}$. Then f is a cusp form.

8 Elliptic Curves

8.1 Introduction

Definition 8.1. Let a, b, c be the side length of a right triangle. A right triangle is **rational** if it has rational side length, i.e., $a, b, c \in \mathbb{Q}$. A right triangle is **primitive** if $a, b, c \in \mathbb{Z}$ are coprime.

Lemma 8.1. If the right triangle is primitive, then

$$\begin{cases} a = u^2 - v^2 \\ b = 2uv \\ c = u^2 + v^2 \end{cases}$$

for $u, v \in \mathbb{Z}$, $u > v > 0$.

Definition 8.2. $D \in \mathbb{Q}_{>0}$ is **congruent** if there is a rational triangle with area D.

Note that we only need to consider D square-free.

5 is the area of right triangle with side length $(9/6, 40/6, 41/6)$; 6 is the area of right triangle with side length $(3, 4, 5)$; 7 is the area of right triangle with side length $(24/5, 35/12, 337/60)$.

Oldest Problem I. Prove that there is algorithm for deciding in a finite number of steps whatever a given positive integer is congruent or not.

Oldest Problem II. Prove that every integer of the form $8n+5, 8n+6, 8n+7$ $(n = 0, 1, 2, \cdots)$ is congruent.

Lemma 8.2. $D \in \mathbb{Z}_{>0}$ is congruent iff $Dy^2 = x^3 - x$ for some $x, y \in \mathbb{Q}$ with $y \neq 0$.

Proof. By Lemma 8.1, Area $= uv(u^2 - v^2)$. So D is congruent iff $Dw^2 = uv(u^2 - v^2)$ for some $u, v, w \in \mathbb{Z}$ with $w \neq 0$. Set $x = u/v$ and $y = w/v^2$ to get desired formula. □

Theorem 8.1. There is no solution to $w^2 = uv(u^2 - v^2)$ in $\mathbb{Z}$ with $w \neq 0$. That is, 1 is mot congruent.

Proof. Use infinite decent. □

Let K be a field with $\mathrm{char}(K) \neq 2$, and let $\overline{K}$ be its algebraic closure.

Definition 8.3. (i) An **elliptic curve** E/K is the projective closure of a plane affine curve $y^2 = f(x)$, where $f(x) \in K[x]$ is monic cubic with different roots in $\overline{K}$.

(ii) For any field extension L/K, define $E(L) = \{(x, y) \in L^2 \mid y^2 = f(x)\} \cup \{\mathcal{O}\}$.

For example, the affine Fermat curve $x^n + y^n = 1$ has projective closure $X^n + Y^n = Z^n$. Ponits with $Z \neq 0$ correpsonds to affine ppints via $x = X/Z$, $y = Y/Z$. The points with $Z = 0$ are the points at infinity. In particular, when n is odd, $(-1, 1, 0)$ lies on the projective Fermat curve, and it may be chosen as the point $\mathcal{O}$.

The n-dimension projective space $\mathbb{P}^n$ is defined by $\mathbb{P}^n(K) = \left(K^{n+1} \setminus \{(0, 0, \cdots, 0)\}\right) / \sim$, where $(x_0, x_1, \cdots, x_n) \sim (\lambda x_0, \lambda x_1, \cdots, \lambda x_n)$ for $\lambda \in K^\times$. The equivalence class of $(x_0, x_1, \cdots, x_n)$ is written as $(x_0 : x_1 : \cdots : x_n)$. The projective plane $\mathbb{P}^2(K) = \{(x : y : 1) \mid (x, y) \in K^2\} \cup \{(x : y : 0) \mid (x, y) \in \mathbb{P}^1(K)\} = \mathbb{A}^2(K) \cup \mathbb{P}^1(K)$.

Example 8.1. $E\colon y^2 = x^3 - x$, then $E(\mathbb{Q}) = \{(0, 0), (1, 0), (-1, 0)\} \cup \{\mathcal{O}\}$ by Theorem 8.1.

Proposition 8.1. E is an elliptic curve over K, then $E(K(t)) = E(K)$.

Definition 8.4. A plane affine algebriac curve $C = \{f(x, y) = 0\} \subset \mathbb{A}^2$ (f irreducible) is **rational** if it has rational parameterzation. That is, there exists $\varphi(t), \psi(t) \in K(t)$ such that

(i) $f(\varphi(t), \psi(t)) = 0$;

(ii) The map $\mathbb{A}^1 \to \mathbb{A}^2$ given by $t \mapsto (\varphi(t), \psi(t))$ is injective on $\mathbb{A}^1 \setminus \{\text{finite set}\}$.

Example 8.2. Any non-singular plane conic is rational. For example, $x^2 + y^2 = 1$, one can choose

$$(x, y) = \left(\frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2}\right).$$

Example 8.3. Any singular plane cubic is rational. For example, $y^2 = x^3$, one can choose

$$(x, y) = (t^2, t^3).$$

For $y^2 = x^2(x + 1)$,

$$(x, y) = (t^2 - 1, t(t^2 - 1)).$$

Over an algebraically closed field K with $\mathrm{char}(K) \neq 2$, every elliptic curve with full rational 2-torsion may be written in Legendre form

$$C\colon Y^2Z = X(X - Z)(X - \lambda Z),$$

or, on the affine chart $Z \neq 0$,

$$C\colon y^2 = x(x - 1)(x - \lambda),$$

where $\lambda \in K \setminus \{0, 1\}$. The condition $\lambda \neq 0, 1$ ensures that the cubic polynomial $x(x - 1)(x - \lambda)$ has distinct roots.

An elliptic curve can also be written in Weierstrass form:

$$E\colon Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3.$$

on the affine chart $Z \neq 0$,

$$E\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

The genus $g(C) \in \mathbb{Z}_{\geq 0}$ is an invariant of a smooth projective curve C. If $K = \mathbb{C}$, $g(C)$ is the topological genus of the corresponding compact Riemann surface. A smooth plane curve $C \subseteq \mathbb{P}^2$ of degree d has genus $(d-1)(d-2)/2$. In particular, a smooth plane cubic has genus 1.

Proposition 8.2. Let C be a smooth projective curve over an algebraically closed field, C is rational iff $g(C) = 0$.

Let C be a smooth projective curve over K. Then C is an elliptic curve over K if and only if $g(C) = 1$ and C is equipped with a chosen K-rational point $\mathcal{O} \in C(K)$.

Let C be a smooth projective curve over a number field K.

(i) If $g(C) = 0$ and $C(K) \neq \varnothing$, then $C(K)$ is parametrized by $\mathbb{P}^1(K)$.

(ii) (Mordell–Weil Theorem) If $g(C) = 1$ and C has a chosen K-rational point, then C is an elliptic curve, and $C(K)$ is a finitely generated abelian group.

(iii) (Faltings Theorem) If $g(C) > 1$, then $C(K)$ is finite.

By the Mordell–Weil Theorem $E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$. It is unknown whether r is bounded as E ranges over all elliptic curves over $\mathbb{Q}$. The current record is $r \geq 29$ due to Elkies and Klagbrun in 2024; the previous record was $r \geq 28$ due to Elkies in 2006.

8.2 Weierstrass equations

Theorem 8.2. Let K be a field with $\text{char}(K) = 0$ or finite field. Every elliptic curve E is isomorphic over K to a curve in Weierstrass form via isomorphic mapping $\mathcal{O}_E \mapsto (0:1:0)$.

Proposition 8.3. Let E/K and E'/K be two elliptic curves with Weierstrass form, then $E \cong E'$ over K iff the Weierstrass equations are related by a substitution of the form

$$\begin{cases} x = u^2x' + r \\ y = u^3y' + u^2sx' + t \end{cases}$$

where $u, r, s, t \in K$ and $u \neq 0$.

Let E satisfy the Weierstrass equation

$$E\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{$*$}$$

Then E is smoooth iff $\Delta(a_1, \cdots, a_6) \neq 0$. If $\mathrm{char}(K) \neq 2$, we can complete the square; if $\mathrm{char}(K) \neq 3$, we can get rid of x^2 term. Hence if $\mathrm{char}(K) \neq 2, 3$, then

$$E\colon y^2 = x^3 + ax + b.$$

In this case, $\Delta = -16(4a^3 + 27b^2)$.

Definition 8.5. The j-invariant of E is

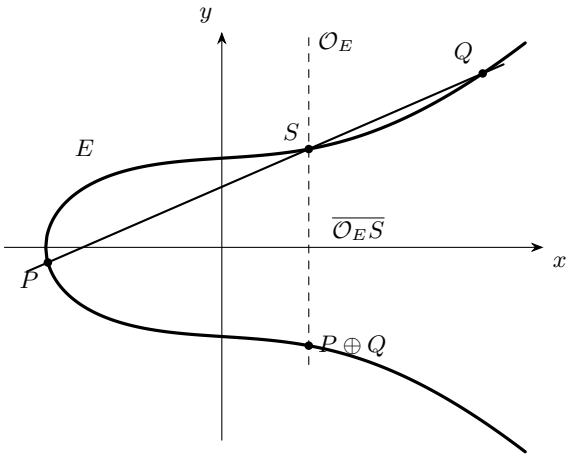
$$j(E) = 1728\frac{4a^3}{4a^3 + 27b^2}.$$

Corollary 8.1. Let E, E' be two elliptic curves over K. $E \cong E'$ implies that $j(E) = j(E')$. The converse holdds if $K = \overline{K}$.

8.3 Group law

Any line in $\mathbb{P}^2$ will meet the curve at exactly three points, counting multiplicity.

For two points P, Q on an elliptic curve E, take the line through P and Q. It meets E at a third point S. The line through S and $\mathcal{O}_E$ meets E at $P \oplus Q$. If $P = Q$, use the tangent at P.



Theorem 8.3. $(E, \oplus)$ is an abelian group.

Assume that $E\colon y^2 = x^3 + Ax + B$, and $P = (x_1, y_1)$, $Q = (x_2, y_2)$. If $P = \mathcal{O}_E$, then $P \oplus Q = Q$; if $Q = \mathcal{O}_E$, then $P \oplus Q = P$; if $x_1 = x_2$ and $y_1 = -y_2$, then $P \oplus Q = \mathcal{O}_E$. In other

cases, let

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} & \text{if } P \neq Q, \\ \dfrac{3x_1^2 + A}{2y_1} & \text{if } P = Q. \end{cases}$$

Then $P \oplus Q = (x_3, y_3)$, where $x_3 = \lambda^2 - x_1 - x_2$ and $y_3 = \lambda(x_1 - x_3) - y_1$.

Corollary 8.2. $E(K) = \big\{(x, y) \in K^2 \mid x, y \text{ satisfy } (*)\big\} \cup \{\mathcal{O}_E\}$ is an abelian group.

Indeed,

(i) $E(\mathbb{C}) \cong \mathbb{C}/\Lambda \cong \mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$, where Λ is a lattice.

(ii)

$$E(\mathbb{R}) = \begin{cases} \mathbb{Z}/2\mathbb{Z} \times \mathbb{R}/\mathbb{Z} & \text{if } \Delta > 0, \\ \mathbb{R}/\mathbb{Z} & \text{if } \Delta < 0. \end{cases}$$

(iii) If $K = \mathbb{F}_q$, then $|\#E(\mathbb{F}_q) - (q+1)| \leq 2\sqrt{q}$ (Hasse's Theorem).

(iv) If K is a local field, i.e., $[K : \mathbb{Q}_p] < \infty$, then $E(K)$ has a subgroup of finite index $\cong (\mathcal{O}_K, +)$.

8.4 Elliptic curves over $\mathbb{C}$

Definition 8.6. A **lattice** $\Lambda = \{\mathbb{Z}w_1 + \mathbb{Z}w_2\}$ is an additive subgroup of $\mathbb{C}$ generated by $w_1, w_2 \in \mathbb{C}$ that are linearly independent.

Example 8.4. $\{1, \tau\}$ is a lattice, where τ is a root of $x^2 + bx + c$ with $b, c \in \mathbb{Z}$ and $\Delta < 0$.

Definition 8.7. An **elliptic function** f is a meromorphic function with $f(z+w) = f(z)$, where $z \in \mathbb{C}$ and $w \in \Lambda$.

Definition 8.8. Weight k Eisenstein series

$$G_k(z) = \sum_{w \in \Lambda^*} \frac{1}{w^k}, \quad \Lambda = \Lambda \setminus \{0\}.$$

Definition 8.9. Weierstrass $\wp$-function

$$\wp(z) = \wp(z, \Lambda) = \frac{1}{z^2} + \sum_{w \in \Lambda^*} \left(\frac{1}{(z-w)^2} - \frac{1}{w^2} \right).$$

Remark 8.1. $\wp(z)$ has a pole of order 2 at $z \in \Lambda$. And we have

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

Laurant series expansion for $\wp(z)$ at $z = 0$:

$$\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty}(2n+1)G_{2n+2}(\Lambda)z^{2n}.$$

Theorem 8.4.

$$\wp'(z)^2 = 4\wp(z)^3 - g_2(\Lambda)\wp(z) - g_3(\Lambda), \quad \text{where } \begin{cases} g_2(\Lambda) = 60G_4(\Lambda), \\ g_3(\Lambda) = 140G_6(\Lambda). \end{cases}$$

Theorem 8.5 (Uniformization Theorem)**.** Every elliptic curve $E/\mathbb{C}$ arise from a lattice Λ.

Remark 8.2.

$$\Phi\colon \mathbb{C}/\Lambda \simeq E_\Lambda(\mathbb{C}), \quad z \longmapsto (\wp(z), \wp'(z)).$$

8.5 Elliptic curves over number field

Definition 8.10. An **isogeny** $\phi\colon E_1 \to E_2$ is a non-constant morphism with $\phi(\mathcal{O}_{E_1}) = \mathcal{O}_{E_2}$. We say that E_1 and E_2 are **isogenous** if there exists an isogeny from E_1 to E_2.

Remark 8.3. For $n \in \mathbb{Z}_{>0}$, define

$$[n]\colon E \longrightarrow E, \quad P \longmapsto P \oplus \cdots \oplus P.$$

If $n < 0$, define $[n] = [-1] \circ [-n]$.

Definition 8.11. The n-torsion subgroup of E is $E[n] = \operatorname{Ker}\left(E \xrightarrow{[n]} E\right)$.

Fact. We have $E[n] \cong (\mathbb{Z}/n\mathbb{Z})^2$, where $\operatorname{char}(K) \nmid n$. If $p = \operatorname{char}(K)$, then

$$E[p] \cong \begin{cases} 0 & \text{(supersigular case)}, \\ \mathbb{Z}/p\mathbb{Z} & \text{(ordinary case)}. \end{cases}$$

Proposition 8.4. Let E be an elliptic curve with $\operatorname{char}(K) \neq 2$. For $y^2 = f(x) = (x - e_1)(x - e_2)(x - e_3)$, where $e_1, e_2, e_3 \in K$ distinct. Then we have $E[2] = \{\mathcal{O}_E, (e_1, 0), (e_2, 0), (e_3, 0)\} \cong (\mathbb{Z}/2\mathbb{Z})^2$.

For example, $E\colon y^2 = x^3 - x$, then $E(\mathbb{Q}) = \{\mathcal{O}, (0,0), (1,0), (-1,0)\} \cong (\mathbb{Z}/2\mathbb{Z})^2$.

Definition 8.12. Let E be an elliptic curve over $\mathbb{Q}$, here

$$E/\mathbb{Q}\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Then define

$$\tilde{E}/\mathbb{F}_p\colon y^2 + \tilde{a_1}xy + \tilde{a_3}y = x^3 + \tilde{a_2}x^2 + \tilde{a_4}x + \tilde{a_6}.$$

Note $\tilde{E}/\mathbb{F}_p$ is an elliptic curve iff $\tilde{\Delta} \neq 0$ iff $\mathrm{ord}_p(\Delta) = 0$.

Definition 8.13. Let E be an elliptic curve over $\mathbb{Q}$. $E/\mathbb{Q}$ has **good reduction** at p if $\tilde{E}$ is nonsingular, **bad reduction** otherwise.

Theorem 8.6 (Week Mordell–Weil)**.** Let E be an elliptic curve over number field K. For $n \in \mathbb{Z}_{\geq 2}$, we have $|E(K)/nE(K)| < \infty$.

Theorem 8.7 (Mordell–Weil)**.** $E(K)$ is a finitely generated abelian group. Indeed we have $E(K) \cong E(K)_{\mathrm{tors}} \oplus \mathbb{Z}^r$.

8.6 The Birch and Swinnerton-Dyer Conjecture

Let K be a number field. Let $\varphi\colon E \to E'$ be an isogeny of elliptic curves over K, there exists exact sequence of $\mathrm{Gal}\left(\overline{K}/K\right)$-modules

$$0 \longrightarrow E[\varphi]\left(\overline{K}\right) \longrightarrow E\left(\overline{K}\right) \longrightarrow E'\left(\overline{K}\right) \longrightarrow 0\ .$$

Then we get long exact sequence

$$\cdots \longrightarrow E(K) \xrightarrow{\varphi} E'(K) \xrightarrow{\delta} H^1(K, E[\varphi]) \longrightarrow H^1(K, E) \xrightarrow{\varphi_*} H^1(K, E') \longrightarrow \cdots$$

Thus we get

$$\begin{array}{ccccccccc} 0 & \longrightarrow & E'(K)/\varphi E(K) & \xrightarrow{\delta} & H^1(K, E[\varphi]) & \longrightarrow & H^1(K, E)[\varphi_*] & \longrightarrow & 0 \\ & & \downarrow & & \downarrow{\scriptstyle \mathrm{res}_v} & \nearrow & \downarrow{\scriptstyle \mathrm{res}_v} & & \\ 0 & \longrightarrow & E'(K_v)/\varphi E(K_v) & \xrightarrow{\delta_v} & H^1(K_v, E[\varphi]) & \longrightarrow & H^1(K_v, E)[\varphi_*] & \longrightarrow & 0 \end{array}$$

The φ-Selmer group is defined by

$$\begin{aligned} S^{(\varphi)}(E'/K) &= \mathrm{Ker}\left(H^1(K, E[\varphi]) \longrightarrow \prod_{v \in M_K} H^1(K_v, E)\right) \\ &= \left\{\alpha \in H^1(K, E[\varphi]) \mid \mathrm{res}_v(\alpha) \in \mathrm{Im}(\delta_v), \forall v \in M_K\right\}. \end{aligned}$$

Tate–Shafarevich group:

$$Ш(E/K) = \mathrm{Ker}\left(H^1(K, E) \longrightarrow \prod_{v \in M_K} H^1(K_v, E)\right),$$

we obtain

$$0 \longrightarrow E'(K)/\varphi E(K) \longrightarrow S^{(\varphi)}(E/K) \longrightarrow Ш(E/K)[\varphi_*] \longrightarrow 0$$

Conjecture (Tate–Shafarevich)**.** $Ш(E/K)$ is finite.

If we take $\varphi = [n]$, then we have

$$0 \longrightarrow E(K)/nE(K) \longrightarrow S^{(n)}(E/K) \longrightarrow Ш(E/K)[n] \longrightarrow 0 \ .$$

Let E be an elliptic curve over $\mathbb{Q}$, define

$$L(E,s) = \prod_p L_p(E,s),$$

where

$$L_p(E,s) = \begin{cases} \left(1 - a_p p^{-s} + p^{1-2s}\right)^{-1} & \text{if } E \text{ has good reduction mod } p,\ a_p = p+1-\#\tilde{E}(\mathbb{F}_p), \\ (1-p^{-s})^{-1} & \text{split multiplicative reduction,} \\ (1+p^{-s})^{-1} & \text{non-split multiplicative reduction,} \\ 1 & \text{additive reduction.} \end{cases}$$

Theorem 8.8 (Wiles, Taylor–Wiles, Breuil–Conrad–Diamond–Taylor)**.** Every elliptic curve $E/\mathbb{Q}$ is modular. Equivalently, there exists a weight 2 newform f_E such that

$$L(E,s) = L(f_E,s).$$

In particular, $L(E,s)$ admits analytic continuation to all of $\mathbb{C}$ and satisfies a functional equation.

Conjecture (BSD)**.**

$$\mathrm{ord}_{s=1} L(E,s) = \mathrm{rank}\, E(\mathbb{Q}).$$

Conjecture (Strong BSD)**.**

$$\lim_{s\to 1} \frac{L(E,s)}{(s-1)^r} = \frac{\Omega_E \cdot \mathrm{Reg}(E(\mathbb{Q})) \cdot |Ш(E/\mathbb{Q})| \cdot \prod_p c_p}{|E(\mathbb{Q})_{\mathrm{tors}}|^2},$$

where $c_p = [E(\mathbb{Q}_p) : E_0(\mathbb{Q}_p)]$ is the Tamagawa number,

$$\mathrm{Reg}(E/\mathbb{Q}) = \det\left(\langle P_i, P_j\rangle\right)_{1\le i,j\le r}, \quad \text{where } \begin{cases} E(\mathbb{Q})/E(\mathbb{Q})_{\mathrm{tors}} = \langle P_1, \cdots, P_r\rangle, \\ \langle P, Q\rangle = \frac{1}{2}\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q), \end{cases}$$

and

$$\Omega_E = \int_{E(\mathbb{R})} \frac{\mathrm{d}x}{|2y + a_1 x + a_3|}.$$

A Ramanujan sum and Gauss sum

First we consider the Ramanujan sum

$$S(m,0;q) = \sum_{a \pmod q}^{*} e\left(\frac{am}{q}\right).$$

Lemma A.1. We have

$$S(m,0;q) = \mu\left(\frac{q}{\gcd(m,q)}\right)\frac{\varphi(q)}{\varphi\left(\frac{q}{\gcd(m,q)}\right)}.$$

Proof. See p.66,

$$S(m,0;q) = \sum_{s|\gcd(m,q)} s\mu\left(\frac{q}{s}\right).$$

Write $q = \prod_p p^{\alpha_p}$ and $\gcd(m,q) = \prod_p p^{\beta_p}$, where $0 \le \beta_p \le \alpha_p$. Then $s \mid \gcd(m,q)$ can be written as $s = \prod_p p^{\gamma_p}$ where $0 \le \gamma_p \le \beta_p$, so $q/s = \prod_p p^{\alpha_p - \gamma_p}$. Consequently,

$$\sum_{s|\gcd(m,q)} s\mu\left(\frac{q}{s}\right) = \prod_{p|q}\left(\sum_{\gamma=0}^{\beta_p} p^{\gamma}\mu\left(p^{\alpha_p-\gamma}\right)\right).$$

Note that $\sum_{\gamma=0}^{\beta_p} p^{\gamma}\mu\left(p^{\alpha_p-\gamma}\right)$ contributes only when $\gamma = \alpha_p$ and $\gamma = \alpha_p - 1$.

We now prove: for each p we have

$$\sum_{\gamma=0}^{\beta_p} p^{\gamma}\mu\left(p^{\alpha_p-\gamma}\right) = \mu\left(p^{\alpha_p-\beta_p}\right)\frac{\varphi\left(p^{\alpha_p}\right)}{\varphi\left(p^{\alpha_p-\beta_p}\right)}.$$

If $\beta_p \le \alpha_p - 2$, then $\alpha_p - \gamma \ge 2$, and so $\sum_{\gamma=0}^{\beta_p} p^{\gamma}\mu(p^{\alpha_p-\gamma}) = 0$; on the other hand $\alpha_p - \beta_p \ge 2$ implies that $\mu\left(p^{\alpha_p-\beta_p}\right) = 0$. If $\beta_p = \alpha_p - 1$, then $\sum_{\gamma=0}^{\beta_p} p^{\gamma}\mu(p^{\alpha_p-\gamma}) = -p^{\alpha_p-1}$; moreover $\mu(p) = -1$ and $\varphi\left(p^{\alpha_p}\right)/\varphi(p) = p^{\alpha_p-1}$. If $\beta_p = \alpha_p$, then $\sum_{\gamma=0}^{\beta_p} p^{\gamma}\mu(p^{\alpha_p-\gamma}) = p^{\alpha_p} - p^{\alpha_p-1}$; on the other hand $\mu(1) = 1$ and $\varphi\left(p^{\alpha_p}\right) = p^{\alpha_p} - p^{\alpha_p-1}$.

Consequently,

$$S(m,0;q) = \prod_{p|q}\mu\left(p^{\alpha_p-\beta_p}\right)\frac{\varphi\left(p^{\alpha_p}\right)}{\varphi\left(p^{\alpha_p-\beta_p}\right)} = \mu\left(\frac{q}{\gcd(m,q)}\right)\frac{\varphi(q)}{\varphi\left(\frac{q}{\gcd(m,q)}\right)}.$$

□

Lemma A.2. Let $\chi \pmod m$ be primitive Dirichlet character, $m > 1$, then for arbitrary a, b we have

$$\frac{1}{m}\sum_{c \pmod m}\chi(ac+b) = \begin{cases}\chi(b) & \text{if } m \mid a,\\ 0 & \text{if } m \nmid a.\end{cases}$$

Proof. First prove

$$\sum_{c \pmod{p^\nu}} \chi(ac+b) = \begin{cases} p^\nu \chi(b) & \text{if } p^\nu \mid a, \\ 0 & \text{if } p^\nu \nmid a. \end{cases}$$

If $p^\nu \mid a$, then $ac + b \equiv b \pmod{p^\nu}$ holds for all c. Therefore,

$$\sum_{c \pmod{p^\nu}} \chi(ac+b) = \sum_{c \pmod{p^\nu}} \chi(b) = p^\nu \chi(b).$$

Now assume $p^\nu \nmid a$. Let $a = p^r a_0$ where $0 \le r = v_p(a) < \nu$ and $\gcd(a_0, p) = 1$. Then a_0 is invertible modulo p^ν, which implies that $c \mapsto a_0 c$ gives

$$\sum_{c \pmod{p^\nu}} \chi(ac+b) = \sum_{c \pmod{p^\nu}} \chi(p^r c+b).$$

If $r = 0$, $\gcd(a, p^\nu) = 1$, so $c \mapsto ac + b$ is a permutation modulo p^ν, so

$$\sum_{c \pmod{p^\nu}} \chi(ac+b) = \sum_{c \pmod{p^\nu}} \chi(c) = 0.$$

If $1 \le r < \nu$, when c runs over the residue classes modulo p^ν, $p^r c$ runs over all residue classes modulo p^ν that are divided by p^r, each appearing exactly p^r times.

$$\sum_{c \pmod{p^\nu}} \chi(p^r c+b) = p^r \sum_{c \pmod{p^{\nu-r}}} \chi(p^r c+b)$$

If $p \mid b$, then for all c we have $p \mid p^r c + b$, so $\chi(p^r c + b) = 0$. Now assume $p \nmid b$. In this case b is invertible modulo p^ν. Therefore $p^r c + b = b\left(1 + p^r b^{-1} c\right)$, we have

$$\sum_{c \pmod{p^{\nu-r}}} \chi(p^r c+b) = \chi(b) \sum_{c \pmod{p^{\nu-r}}} \chi\left(1 + p^r b^{-1} c\right).$$

When c runs over the residue classes modulo $p^{\nu-r}$, $1 + p^r b^{-1} c$ precisely run bijectively over the subgroup $H_r = \{u \pmod{p^\nu} \mid u \equiv 1 \pmod{p^r}\}$. Consequently

$$\sum_{c \pmod{p^{\nu-r}}} \chi\left(1 + p^r b^{-1} c\right) = \sum_{u \in H_r} \chi(u).$$

Since χ is primitive modulo p^ν, its restriction to H_r is nontrivial. Indeed, if χ were trivial on H_r, then χ would factor through $(\mathbb{Z}/p^\nu\mathbb{Z})^\times / H_r \cong (\mathbb{Z}/p^r\mathbb{Z})^\times$, and hence would be induced from a character modulo p^r, contradicting primitivity. We obtain

$$\sum_{u \in H_r} \chi(u) = 0.$$

It follows that

$$\sum_{c \pmod{p^\nu}} \chi(ac+b) = 0.$$

From the Chinese Remainder Theorem,

$$\sum_{c \pmod{m}} \chi(ac+b) = \prod_{p \mid m} \sum_{c_p \pmod{p^{\nu_p}}} \chi_p(ac_p+b),$$

where each χ_p is primitive modulo p^{ν_p}. If $m \mid a$, then for every $p \mid m$ we have $p^{\nu_p} \mid a$, it follows that

$$\sum_{c \pmod{m}} \chi(ac+b) = \prod_{p \mid m} p^{\nu_p} \chi_p(b) = m\chi(b).$$

If $m \nmid a$, there exists some $p \mid m$ such that $p^{\nu_p} \nmid a$, then

$$\sum_{c_p \pmod{p^{\nu_p}}} \chi_p(ac_p+b) = 0.$$

Therefore

$$\sum_{c \pmod{m}} \chi(ac+b) = 0.$$

We finish the proof. □

Now let $\chi \pmod{m}$ be Dirichlet character and consider the Gauss sum

$$\tau(\chi) = \sum_{b \pmod{m}} \chi(b) e\left(\frac{b}{m}\right).$$

For $\gcd(a,m) = 1$, by multiplying $\overline{\chi(a)}$ and summing over $\chi \pmod{m}$,

$$\begin{aligned}
\sum_{\chi \pmod{m}} \overline{\chi(a)}\tau(\chi) &= \sum_{\chi \pmod{m}} \overline{\chi(a)} \sum_{b \pmod{m}} \chi(b) e\left(\frac{b}{m}\right) \\
&= \sum_{b \pmod{m}} e\left(\frac{b}{m}\right) \sum_{\chi \pmod{m}} \chi(b)\overline{\chi(a)} \\
&= \sum_{\substack{b \pmod{m} \\ b \equiv a \pmod{m}}} e\left(\frac{b}{m}\right) \varphi(m)
\end{aligned}$$

by the orthogonality. It follows that

$$\sum_{\chi \pmod{m}} \overline{\chi(a)}\tau(\chi) = \varphi(m) e\left(\frac{a}{m}\right),$$

which implies that

$$e\left(\frac{a}{m}\right) = \frac{1}{\varphi(m)} \sum_{\chi \pmod{m}} \overline{\chi(a)}\tau(\chi), \quad \gcd(a,m) = 1.$$

By definition, if $\gcd(a,m)=1$, then

$$\begin{aligned}\chi(a)\tau\left(\overline{\chi}\right) &= \sum_{b \pmod m} \chi(a)\overline{\chi(b)} e\left(\frac{b}{m}\right) \\ &= \sum_{b \pmod m} \overline{\chi(b)} e\left(\frac{ab}{m}\right). \quad (b \mapsto ab \text{ since } \gcd(a,m)=1)\end{aligned}$$

If $\tau\left(\overline{\chi}\right) \neq 0$, one can write

$$\chi(a) = \frac{1}{\tau\left(\overline{\chi}\right)} \sum_{b \pmod m} \overline{\chi(b)} e\left(\frac{ab}{m}\right), \quad \gcd(a,m)=1,$$

which is precisely the Fourier expansion of χ in terms of additive characters. The condition $\gcd(a,m)=1$ can be dropped if χ is primitive. Indeed, if $\gcd(a,m)>1$ we have $\chi(a)\tau\left(\overline{\chi}\right)=0$ since $\chi(a)=0$. On the other hand, note that

$$e\left(\frac{a(b+m/\gcd(a,m))}{m}\right) = e\left(\frac{ab}{m}\right) e\left(\frac{a}{\gcd(a,m)}\right) = e\left(\frac{ab}{m}\right)$$

since $a/\gcd(a,m) \in \mathbb{Z}$, $b \mapsto e\left(ab/m\right)$ is periodic modulo $m/\gcd(a,m)$. Therefore

$$\sum_{b \pmod m} \overline{\chi(b)} e\left(\frac{ab}{m}\right) = \sum_{c \pmod{\gcd(a,m)}} e\left(\frac{ac}{m}\right) \sum_{\substack{b \pmod m \\ b \equiv c \pmod{\gcd(a,m)}}} \overline{\chi(b)}.$$

Since χ is primitive, $\gcd(a,m)<m$ implies $m \nmid \gcd(a,m)$, we have

$$\sum_{c \pmod m} \overline{\chi(\gcd(a,m)c+r)} = 0$$

thanks to Lemma A.2. On the other hand, when c runs over the residue classes modulo m, $\gcd(a,m)c+r$ runs over the residue classes modulo m satisfying $b \equiv r \pmod{\gcd(a,m)}$ exactly $\gcd(a,m)$ times. So

$$\sum_{c \pmod m} \overline{\chi(\gcd(a,m)c+r)} = \gcd(a,m) \sum_{\substack{b \pmod m \\ b \equiv r \pmod{\gcd(a,m)}}} \overline{\chi(b)} = 0,$$

which implies that

$$\sum_{b \pmod m} \overline{\chi(b)} e\left(\frac{ab}{m}\right) = 0.$$

Lemma A.3. Suppose the character $\chi \pmod m$ is induced by the primitive character $\chi^* \pmod{m^*}$. Then

$$\tau(\chi) = \mu\left(\frac{m}{m^*}\right) \chi^*\left(\frac{m}{m^*}\right) \tau(\chi^*).$$

If $\chi \pmod m$ is primitive, then

$$|\tau(\chi)| = \sqrt{m}.$$

Proof. We have

$$\begin{aligned}
\tau(\chi) &= \sideset{}{^*}\sum_{a \pmod m} \chi^*(a) e\left(\frac{a}{m}\right) \\
&= \sum_{a \pmod m} \chi^*(a) e\left(\frac{a}{m}\right) \sum_{d \mid \gcd(a,m)} \mu(d) \qquad \left(1_{\gcd(a,m)=1} = \sum_{d \mid \gcd(a,m)} \mu(d)\right) \\
&= \sum_{d \mid m} \mu(d) \sum_{\substack{a \pmod m \\ d \mid a}} \chi^*(a) e\left(\frac{a}{m}\right) \\
&= \sum_{d \mid m} \mu(d) \sum_{b \pmod{m/d}} \chi^*(db) e\left(\frac{db}{m}\right) \qquad (\text{Let } a = db) \\
&= \sum_{d \mid m} \mu(d)\chi^*(d) \sum_{b \pmod{m/d}} \chi^*(b) e\left(\frac{b}{m/d}\right).
\end{aligned}$$

If $\chi^*(d) = 0$, the corresponding d-term vanishes. We now consider $\chi^*(d) \neq 0$, in this case $\gcd(d, m^*) = 1$. Since $m^* \mid m$, we have $m^* \mid m/d$, one can write $m/d = m^*\ell$. Therefore

$$\sum_{b \pmod{m/d}} \chi^*(b) e\left(\frac{b}{m/d}\right) = \sum_{b \pmod{m^*\ell}} \chi^*(b) e\left(\frac{b}{m^*\ell}\right).$$

Write $b = c + km^*$, where $c \pmod{m^*}$ and $k = 0, 1, \cdots, \ell - 1$. From $\chi^*(c + km^*) = \chi^*(c)$, we have

$$\sum_{b \pmod{m/d}} \chi^*(b) e\left(\frac{b}{m/d}\right) = \sum_{c \pmod{m^*}} \chi^*(c) e\left(\frac{c}{m^*\ell}\right) \sum_{k=0}^{\ell-1} e\left(\frac{k}{\ell}\right).$$

If $\ell > 1$, then

$$\sum_{k=0}^{\ell-1} e\left(\frac{k}{\ell}\right) = 0.$$

So

$$\sum_{b \pmod{m/d}} \chi^*(b) e\left(\frac{b}{m/d}\right) = \begin{cases} 0 & \text{if } \ell > 1, \\ \tau(\chi^*) & \text{if } \ell = 1. \end{cases}$$

When $\ell = 1$, $m/d = m^*$ implies that $d = m/m^*$. Consequently

$$\tau(\chi) = \mu\left(\frac{m}{m^*}\right) \chi^*\left(\frac{m}{m^*}\right) \tau(\chi^*).$$

Now assume $\chi \pmod m$ primitive. From the definition,

$$\begin{aligned}
|\tau(\chi)|^2 = \tau(\chi)\overline{\tau(\chi)} &= \sum_{a,b \pmod m} \chi(a)\overline{\chi(b)} e\left(\frac{a-b}{m}\right) \\
&= \sum_{b \pmod m}^{*} \sum_{a \pmod m} \chi(a)\overline{\chi(b)} e\left(\frac{a-b}{m}\right) \\
&= \sum_{a \pmod m} \chi(a) \sum_{b \pmod m}^{*} e\left(\frac{(a-1)b}{m}\right) \quad (a \mapsto ab) \\
&= \sum_{a \pmod m} \chi(a) \sum_{\substack{d \mid m \\ d \mid (a-1)}} d\mu\left(\frac{m}{d}\right) \qquad \text{(See the proof of Lemma A.1)} \\
&= \sum_{a \pmod m} \chi(a) \sum_{\substack{d \mid m \\ a \equiv 1 \pmod d}} d\mu\left(\frac{m}{d}\right) \\
&= \sum_{d \mid m} d\mu\left(\frac{m}{d}\right) \sum_{\substack{a \pmod m \\ a \equiv 1 \pmod d}} \chi(a) \\
&= m
\end{aligned}$$

since

$$\sum_{\substack{a \pmod m \\ a \equiv 1 \pmod d}} \chi(a) = 0$$

if $d < m$. Indeed, applying Lemma A.2 with coefficient $a = d$ and constant term $b = 1$, we get

$$\sum_{c \bmod m} \chi(dc+1) = 0.$$

Since, as c runs modulo m, $dc+1$ runs through all residue classes $a \bmod m$ satisfying $a \equiv 1 \pmod d$, each exactly d times, we obtain

$$\sum_{\substack{a \bmod m \\ a \equiv 1 \pmod d}} \chi(a) = 0.$$

That is, $|\tau(\chi)| = \sqrt{m}$. □

References

[Apo76] Tom M. Apostol (1976), *Introduction to Analytic Number Theory*, Undergraduate Texts in Mathematics, Springer, New York.

[BK17a] Jack Buttcane and Rizwanur Khan (2017), "A mean value of a triple product of L-functions," *Mathematische Zeitschrift* 285(1–2), 565–591. doi:10.1007/s00209-016-1721-y

[BK17b] Jack Buttcane and Rizwanur Khan (2017), "On the fourth moment of Hecke–Maass forms and the random wave conjecture," *Compositio Mathematica* 153(7), 1479–1511. doi:10.1112/S0010437X17007199

[Cha20] Sagun Chanillo (2020), *Notes on Kloosterman Sums*, lecture notes / technical report, Rutgers University. doi:10.13140/RG.2.2.31416.16644/2

[DS05] Fred Diamond and Jerry Shurman (2005), *A First Course in Modular Forms*, Graduate Texts in Mathematics 228, Springer, New York.

[Gol06] Dorian Goldfeld (2006), *Automorphic Forms and L-Functions for the Group* $\mathrm{GL}(n, R)$, Cambridge Studies in Advanced Mathematics 99, Cambridge University Press, Cambridge.

[Hua21] Bingrong Huang (2021), "On the Rankin–Selberg problem," *Mathematische Annalen* 381(3–4), 1217–1251. doi:10.1007/s00208-021-02186-7

[Hua24] Bingrong Huang (2024), "On the Rankin–Selberg problem, II," *The Quarterly Journal of Mathematics* 75(1), 1–10. doi:10.1093/qmath/haad037

[Iwa97] Henryk Iwaniec (1997), *Topics in Classical Automorphic Forms*, Graduate Studies in Mathematics 17, American Mathematical Society, Providence, RI.

[IK04] Henryk Iwaniec and Emmanuel Kowalski (2004), *Analytic Number Theory*, American Mathematical Society Colloquium Publications 53, American Mathematical Society, Providence, RI.

[ILS00] Henryk Iwaniec, Wenzhi Luo and Peter Sarnak (2000), "Low lying zeros of families of L-functions," *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* 91, 55–131. doi:10.1007/BF02698741

[Klo26] H. D. Kloosterman (1926), "On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$," *Acta Mathematica* 49(3–4), 407–464.

[LLY06] Yuk-Kam Lau, Jianya Liu and Yangbo Ye (2006), "A new bound $k^{2/3+\varepsilon}$ for Rankin–Selberg L-functions for Hecke congruence subgroups," *International Mathematics Research Papers* 2006, Article ID 35090, 78 pp. doi:10.1155/IMRP/2006/35090

[Ser73] Jean-Pierre Serre (1973), *A Course in Arithmetic*, Graduate Texts in Mathematics 7, Springer, New York.

[Sil09] Joseph H. Silverman (2009), *The Arithmetic of Elliptic Curves*, 2nd ed., Graduate Texts in Mathematics 106, Springer, Dordrecht.

[ST15] Joseph H. Silverman and John Tate (2015), *Rational Points on Elliptic Curves*, 2nd ed., Undergraduate Texts in Mathematics, Springer, Cham.

[Xi23] Ping Xi (2023), *An elementary proof of the Selberg identity for Kloosterman sums*, preprint. arXiv:2306.16929