6 Hecke operators

Recall the conjecture of Ramanujan that \( \tau(mn) = \tau(m)\tau(n) \) for \( m, n \) relatively prime. While this specific result was shown by Mordell, it was Hecke who developed a general theory to determine which modular forms have multiplicative Fourier coefficients. Note that the Eisenstein series \( E_k \) essentially have multiplicative Fourier coefficients by the following elementary exercise.

**Exercise 6.0.1.** If \( \gcd(m,n) = 1 \), show \( \sigma_k(m)\sigma_k(n) = \sigma_k(mn) \).

Namely, if we write

\[
E_k(z) = 1 - \frac{2}{B_k} \sum \sigma_{k-1}(n)q^n = \sum a_n q^n,
\]

the above exercise shows

\[
a_m a_n = \left(\frac{2}{B_k}\right)^2 \sigma_{k-1}(m)\sigma_{k-1}(n) = \left(\frac{2}{B_k}\right)^2 \sigma_{k-1}(mn) = -\frac{2}{B_k} a_{mn} = a_1 a_{mn}.
\]

Very roughly, the idea of Hecke is the following. Consider an operator

\[
U_p \left( \sum a_n q^n \right) = \sum a_{pn} q^n.
\]

If this operator preserves \( S_{12}(1) \), then

\[
U_p \Delta = \lambda_p \Delta
\]

for some \( \lambda_p \in \mathbb{C} \), since \( S_{12}(1) = \langle \Delta \rangle \). On the other hand,

\[
U_p \Delta = U_p(q - 24q^2 + 252q^3 - \cdots) = \tau(p)q - 24\tau(p)q^2 + \cdots
\]

so \( \lambda_p = \tau(p) \) and therefore we would have \( \tau(pn) = \tau(p)\tau(n) \) for all \( n \).

However, \( \tau \) is not totally multiplicative, i.e., \( \tau(p^2) \neq \tau(p)^2 \), so \( U_p \) does not preserve \( \Delta \). Instead, since Ramanujan predicted \( \tau(p^2) = \tau(p)^2 - p^{11} \), one would expect \( \tau(pn) = \tau(p)\tau(n) - p^{11}\tau(n/p) \) for \( p|n \). Thus one can guess the correct operator (for weight 12) should be

\[
T_p \left( \sum a_n q^n \right) = \sum a_{pn} q^n + \sum_{p|n} \left( a_{pn} + p^{11}a_{n/p} \right) q^n.
\]

Then one just needs to show \( T_p \) preserves \( S_{12}(1) \).

6.1 Hecke operators for \( \Gamma_0(N) \)

While we could define the Hecke operators directly by their action on Fourier expansions along the lines of (6.1), it will be helpful (and more motivated) to think of them as acting on lattices.

First let’s consider the case of \( \text{PSL}_2(\mathbb{Z}) \). Recall \( \text{PSL}_2(\mathbb{Z}) \backslash \mathfrak{H} \) parameterizes the space of lattices up to homothety (equivalence by \( \mathbb{C}^\times \)). Hence a weak modular form of weight 0 is simply a meromorphic function on equivalence classes of lattices.

What about weight \( k \)? Given \( f \in M_k(1) \), consider the lattice \( \Lambda = \langle 1, \tau \rangle \) with \( \tau \in \mathfrak{H} \). Put \( F(\Lambda) = f(\tau) \). Since \( \langle \omega_1, \omega_2 \rangle = \langle \omega'_1, \omega'_2 \rangle \) if and only if

\[
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix}
\]

83
one easily sees that $\langle 1, \tau \rangle = \langle 1, \tau' \rangle$ (with $\tau' \in \mathcal{S}$ also) if and only if $\tau' = \tau + d$ for some $d \in \mathbb{Z}$. Thus $f$ having period 1 implies $F(\Lambda)$ is well-defined on lattices of the form $\langle 1, \tau \rangle$.

Now consider an arbitrary lattice $\Lambda = \langle \omega_1, \omega_2 \rangle$. We can write $\Lambda = \omega_1 \langle 1, \omega_2/\omega_1 \rangle = \lambda \langle 1, \tau \rangle$ where $\lambda = \omega_1$ and $\tau = \omega_2/\omega_1$. Thus, to see how to extend $F$ to a function on all lattices it suffices to determine how $F$ should behave under multiplying a lattice $\Lambda = \langle 1, \tau \rangle$ by a scalar $\lambda$. It’s reasonable to ask that a scalar $\lambda$ should just transform $F$ by some factor, i.e.,

$$F(\lambda \Lambda) = c(\lambda)F(\Lambda).$$

The only condition we have is that this should be compatible with our definition of $F(\langle 1, \tau \rangle) = f(\tau)$. In other words, if $\lambda \langle 1, \tau' \rangle = \langle 1, \tau \rangle$, we need to ensure $F(\lambda \langle 1, \tau' \rangle) = F(\langle 1, \tau \rangle)$.

This boils down to the case where $\lambda = \tau$ and $\tau' = -\frac{1}{\tau}$. Here we require

$$c(\tau)f\left(-\frac{1}{\tau}\right) = F(\tau \langle 1, \frac{1}{\tau}\rangle) = F(\langle 1, \tau \rangle) = f(\tau).$$

But now the modularity of $f$ implies $c(\tau) = \frac{1}{\tau^k}$. In other words, the weight $k$ modular form $f$ can be viewed as a homogeneous function $F$ of degree $-k$ on the space of lattices of $\mathbb{C}$, i.e., a function $F$ such that

$$F(\lambda \Lambda) = \frac{1}{\lambda^k}F(\Lambda)$$

for $\lambda \in \mathbb{C}^\times$. Call the space of such $F$ by $\mathcal{L}(k)$.

Conversely, if $F$ is a function of lattices such that $F(\lambda \Lambda) = \lambda^{-k}F(\Lambda)$ with $k \geq 0$ even, then one can check $f(z) = F(\langle 1, z \rangle) \in M_k(1)$. Hence there is a bijection between $M_k(1)$ and $\mathcal{L}(k)$.

Then for PSL$_2(\mathbb{Z})$ we can define the $n$-th Hecke operator in terms of $F \in \mathcal{L}(k)$:

$$T_n(F(\Lambda)) = F\left(\sum_{\Lambda' \subseteq \Lambda, |\Lambda'|^2 = n} \Lambda'\right),$$

where the sum is over all $\Lambda' \subseteq \Lambda$ of index $n$. In other words $T_n$ averages $F$ over all sublattices of index $n$. It is clear that $T_n F$ is still a function of lattices, and it is homogeneous of degree $-k$. Clearly $T_1 F = F$.

Thus the correspondence between modular forms $f \in M_k(1)$ and $F \in \mathcal{L}(k)$ induces an action of the Hecke operators

$$T_n : M_k(1) \to M_k(1).$$

Let’s see how $T_n$ translates to an operator on $M_k(1)$.

First observe that the sublattices $\Lambda'$ of $\Lambda = \langle 1, \tau \rangle$ of index $n$ are precisely $\Lambda' = \langle 1, \tau' \rangle$, where

$$\begin{pmatrix} \tau' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \tau \\ 1 \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_n(\mathbb{Z}),$$

where $\mathcal{M}_n(\mathbb{Z})$ denotes the $2 \times 2$ integer matrices of determinant $n$. For such a $\Lambda'$ we have

$$F(\Lambda') = F\left((c\tau + d)(\frac{a\tau + b}{c\tau + d}, 1)\right) = (c\tau + d)^{-k} f(\frac{a\tau + b}{c\tau + d}) = f|_{\mu,k}(\tau).$$

84
Proof.

Proposition 6.1.2. Now we show it actually is an operator on so the above sum over coset representatives is well defined. It is also easy to see that

\[ f|_{\mu,k}(\tau) = (c\tau + d)^{-k} f \left( \frac{a\tau + b}{c\tau + d} \right) \quad \text{for} \quad \mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_n(\mathbb{Z}). \]

(Note in general for \( \mu = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2^+(\mathbb{Q}) \) (the superscript + means positive determinant) one typically defines the slash operator with a factor of \( \det(\mu)^{k/2} \), so our notation departs from the standard here, but seems the most reasonable for our present purpose.)

Given \( \mu_1, \mu_2 \in \mathcal{M}_n(\mathbb{Z}) \), one can check \( \mu_1(1,\tau) = \mu_2(1,\tau) \) if and only if \( \mu_2 = \gamma\mu_1 \) for some \( \gamma \in \text{SL}_2(\mathbb{Z}) \). Hence on the level of modular forms, we have

\[ T_n f = \sum_{\mu \in \text{SL}_2(\mathbb{Z}) \setminus \mathcal{M}_n(\mathbb{Z})} f|_{\mu,k}. \]

For our arithmetic purposes, it is better to introduce a normalization factor of \( n^{1-k} \) in the Hecke operator, which we will do below. The above discussion was just motivation for how to define Hecke operators for \( \text{PSL}_2(\mathbb{Z}) \).

While we will not go through the details (cf. [Kob93]), a similar argument can be made for the modular groups \( \Gamma_0(N) \) (as well as for \( \Gamma_1(N) \) and \( \Gamma(N) \)). The idea is that \( \Gamma_0(N) \setminus \mathcal{Q} \) parameterizes pairs \( (\Lambda, C) \) where \( \Lambda \) is a lattice in \( \mathbb{C} \) and \( C \) is a cyclic subgroup of \( \Lambda \) of order \( N \). Then one can identify modular forms \( f \in M_k(N) \) with homogeneous functions of degree \(-k\) on pairs \( (\Lambda, C) \) and define Hecke operators \( T_n \) similarly, though some care must be taken when \( \gcd(n,N) > 1 \).

Let

\[ \mathcal{M}_{n,N}(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_n(\mathbb{Z}) | c \equiv 0 \text{ mod } N \right\} / \{ \pm I \}. \]

Definition 6.1.1. Suppose \( \gcd(n,N) = 1 \). We define the \( n \)-th Hecke operator \( T_n \) on \( M_k(N) \) by

\[ T_n f = n^{k-1} \sum_{\mu \in \Gamma_0(N) \setminus \mathcal{M}_{n,N}(\mathbb{Z})} f|_{\mu,k}. \]

We will see the normalization factor \( n^{k-1} \) will make the action on Fourier coefficients nice. Observe that for \( f \in M_k(N), \mu \in \mathcal{M}_{n,N}(\mathbb{Z}) \) and \( \gamma \in \Gamma_0(N) \) we have

\[ f|_{\gamma\mu,k} = (f|_{\gamma,k})|_{\mu,k} = f|_{\mu,k}, \]

so the above sum over coset representatives is well defined. It is also easy to see that \( T_n \) is linear. Now we show it actually is an operator on \( M_k(N) \), as well as \( S_k(N) \).

Unless otherwise specified, we assume \( \gcd(n,N) = 1 \) in what follows.

Proposition 6.1.2. We have \( T_n : M_k(N) \to M_k(N) \) and \( T_n : S_k(N) \to S_k(N) \).

Proof. Let \( \gamma \in \Gamma_0(N) \). Then

\[ (T_n f)|_{\gamma,k} = n^{k-1} \left( \sum_{\mu \in \Gamma_0(N) \setminus \mathcal{M}_{n,N}(\mathbb{Z})} f|_{\mu,k} \right)|_{\gamma,k} = n^{k-1} \sum_{\mu \in \Gamma_0(N) \setminus \mathcal{M}_{n,N}(\mathbb{Z})} f|_{\mu\gamma,k} = n^{k-1} \sum_{\mu \in \Gamma_0(N) \setminus \mathcal{M}_{n,N}(\mathbb{Z})} f|_{\mu,k} = T_n f \]

85
since right multiplication by $\gamma \in \Gamma_0(N)$ preserves $\mathcal{M}_{n,N}(\mathbb{Z})$, and therefore simply permutes the cosets $\Gamma_0(N) \backslash \mathcal{M}_{n,N}(\mathbb{Z})$.

Note that each $f|_{\mu,k}$ is holomorphic on $\mathfrak{H}$, therefore $T_n f$ is. To see $T_n f$ is holomorphic at each cusp, let $\tau \in \text{PSL}_2(\mathbb{Z})$ and consider $f|_{\mu,k}|_{\tau} = f|_{\mu',k}$ where $\mu' = \mu \tau = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_n(\mathbb{Z})$ (but not necessarily $\mathcal{M}_{n,N}(\mathbb{Z})$). As $\text{Im}(z) \to \infty$, $f|_{\mu',k}(z) \to \lim_{\text{Im}(z) \to \infty} \frac{1}{(cz + d)^2} f \left( \frac{a}{c} \right)$ which has a finite limit because $f$ is holomorphic at the cusp $\frac{a}{c}$. Thus each $f|_{\mu'}$ is holomorphic at $i\infty$, and $T_n f$ is holomorphic at the cusps.

The same argument we used for holomorphy at the cusps shows that if $f$ is vanishes at the cusps, so does $T_n f$. \hfill \square

**Lemma 6.1.3.** A set of coset representatives for $\Gamma_0(N) \backslash \mathcal{M}_{n,N}(\mathbb{Z})$ is

$$\left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, d > 0, \text{ad} = n, \ 0 \leq b < n \right\}.$$

**Proof.** Take $\begin{pmatrix} a & b \\ cN & d \end{pmatrix} \in \mathcal{M}_{n,N}(\mathbb{Z})$ representing some coset in $\Gamma_0(N) \backslash \mathcal{M}_{n,N}(\mathbb{Z})$. Since $\gcd(n, N) = 1$, we know $\gcd(a, cN) = 1$. Then there exist $x, y \in \mathbb{Z}$ such that $\begin{pmatrix} x & y \\ cN & -a \end{pmatrix} \in \Gamma_0(N)$. Now observe

$$\begin{pmatrix} x & y \\ cN & -a \end{pmatrix} \begin{pmatrix} a & b \\ cN & d \end{pmatrix} = \begin{pmatrix} ax + cyN & bx + dy \\ bcN - ad & 0 \end{pmatrix},$$

which means we may assume $c = 0$ for our coset representative.

Further, given $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathcal{M}_{n,N}(\mathbb{Z})$, we may replace it by the coset representative

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} a & b + dx \\ 0 & d \end{pmatrix},$$

where we choose $x$ so that $0 \leq b + dx < d$, i.e., we may just assume $0 \leq b < d$. Since $n = \text{ad} > 0$, $a$ and $d$ have the same sign, we may multiply by $-I$ if necessary to assume $a, d > 0$. This shows any coset has a representative of the desired form.

Now consider $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ and $\begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}$ such that $\text{ad} = \text{a'd'} = n, \ 0 \leq b < d$ and $0 \leq b' < d'$. Assume they represent the same coset, i.e.,

$$\begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} r & s \\ tN & u \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} ra & rb + sd \\ tNa & tNb + ud \end{pmatrix},$$

for some $\begin{pmatrix} r & s \\ tN & u \end{pmatrix} \in \Gamma_0(N)$. First we see this means $t = 0$, which means $r = u = 1$ (up to $\pm 1$). This means $a' = a, d' = d$, and $b' = b + sd$. However $0 \leq b' < d' = d$ implies we need $s = 1$, i.e., $\begin{pmatrix} r & s \\ tN & u \end{pmatrix} = I$, and therefore any two distinct matrices of the prescribed form represent distinct cosets. \hfill \square
Theorem 6.1.4. Let \( f(z) = \sum a_nq^n \in M_k(N) \) and suppose \( \gcd(m, N) = 1 \). Then

\[
(T_m f)(z) = \sum b_n q^n
\]

where

\[
b_n = \sum_{d \mid \gcd(m, n)} d^{k-1} a_{mn/d^2}.
\]

In particular, if \( m = p \) is prime, then

\[
b_n = \begin{cases} a_{pn} & \text{if } p \nmid n \\ a_{pn} + p^{k-1} a_{n/p} & \text{if } p \mid n \end{cases}
\]

so

\[
(T_p f)(z) = \sum_{n \not\equiv 0 \mod p} a_{pn} q^n + \sum_{n \equiv 0 \mod p} \left( a_{pn} + p^{k-1} a_{n/p} \right) q^n.
\]

Proof. By the previous lemma, we have

\[
(T_m f)(z) = m^{k-1} \sum_{a, b, d} f\left( \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \right)^k
\]

where the sum runs over \( a, b, d \in \mathbb{Z}_{\geq 0} \) such that \( ad = m \) and \( 0 \leq b < d \). Note

\[
f\left( \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \right)^k (z) = \frac{1}{d^k} f\left( \frac{az + b}{d} \right) = \frac{1}{d^k} \sum_{n=0}^{\infty} a_n e^{2\pi i \frac{az+b}{d} n} e^{2\pi i \frac{b}{d} n} = \frac{1}{d^k} \sum_{n=0}^{\infty} \zeta_d^{bn} a_n q^{\frac{a}{d} n}.
\]

For fixed \( a, d \) we will be considering the sum

\[
\sum_{0 \leq b < d} \frac{1}{d^k} f\left( \frac{az + b}{d} \right) = \frac{1}{d^k} \sum_{n=0}^{\infty} \left( \sum_{0 \leq b < d} \zeta_d^{bn} \right) a_n q^{\frac{a}{d} n}.
\]

Note that the inner sum \( \sum_{0 \leq b < d} \zeta_d^{bn} \) will just be a sum over all \( d \)-th roots of unity, and therefore vanish, unless \( d \mid n \), in which case it is just \( d \). Thus

\[
\sum_{0 \leq b < d} \frac{1}{d^k} f\left( \frac{az + b}{d} \right) = \frac{1}{d^{k-1}} \sum_{n=0}^{\infty} a_{dn} q^{an}.
\]

Hence

\[
(T_m f)(z) = m^{k-1} \sum_{ad = m} \sum_{0 \leq b < d} \frac{1}{d^k} f\left( \frac{az + b}{d} \right) = m^{k-1} \sum_{ad = m} \left( \frac{1}{d^{k-1}} \sum_{n=0}^{\infty} a_{dn} q^{an} \right)
\]

\[= \sum_{n=0}^{\infty} \sum_{a \mid m} a^{k-1} a_{mn/a} q^{an} = \sum_{n=0}^{\infty} \sum_{d \mid m} d^{k-1} a_{mn/d^2} q^n.
\]
Corollary 6.1.5. The Ramanujan tau function satisfies \( \tau(mn) = \tau(m)\tau(n) \) whenever \( \gcd(m, n) = 1 \) and \( \tau(p^r) = \tau(p)\tau(p^{r-1}) - p^{11}\tau(p^{r-2}) \).

Proof. Consider

\[
(T_p \Delta)(z) = \sum_{p|n} \tau(pn)q^n + \sum_{p|n} (\tau(pn) + p^{11}\tau(n/p)) q^n.
\]

On the other hand, \( T_p \) acts on \( S_{12}(1) = \mathbb{C}\Delta \) so

\[
(T_p \Delta)(z) = \lambda \Delta(z)
\]

for some \( \lambda \in \mathbb{C}^\times \). The 1st Fourier coefficients of \( \Delta \) and \( T_p \Delta \) are just 1 and \( \tau(p) \), so \( \lambda = \tau(p) \). Comparing the \( n \)-th Fourier coefficients, we see

\[
\begin{cases}
\tau(pn) = \lambda \tau(n) = \tau(p)\tau(n) & p \nmid n \\
\tau(pn) + p^{11}\tau(n/p) = \lambda \tau(n) = \tau(p)\tau(n) & p \nmid n.
\end{cases}
\]

The former equation proves the multiplicativity when \( m = p \) and the latter equation proves the recursion relation for \( \tau(p^r) \).

To obtain the general multiplicativity law \( \tau(mn) = \tau(m, n) \) for \( \gcd(m, n) = 1 \), we simply use the same argument as above with \( T_m \) instead of \( T_p \).

The above argument applies in a more general setting.

Exercise 6.1.6. Suppose \( f(z) = \sum a_nq^n \in S_k(N) \) (resp. \( M_k(N) \)) and \( \dim S_k(N) = 1 \) (resp. \( \dim M_k(N) = 1 \)). Show

(i) If \( \gcd(m, N) = \gcd(n, N) = \gcd(m, n) = 1 \), then \( a_1a_{mn} = a_ma_n \). In particular, if \( f(z) \neq 0 \) we must have \( a_1 \neq 0 \), and if we normalize \( f \) such that \( a_1 = 1 \), the Fourier coefficients are multiplicative.

(ii) If \( a_1 = 1 \) and \( p \nmid N \), then \( a_{pn} = a_pa_{p^{n-1}} - p^{k-1}a_{p^{n-2}} \) for \( n \geq 2 \).

What this means is we can use Hecke operators to compute values of Fourier coefficients from just knowing what the \( T_p \)'s are.

Exercise 6.1.7. Using the fact that \( \dim S_{16}(1) = 1 \), use the previous exercise to help compute the first 10 Fourier coefficients of \( E_4\Delta \).

Now you might ask, for what modular forms \( f \) are the Fourier coefficients multiplicative in the sense \( a_1a_{mn} = a_ma_n \) for \( \gcd(m, n) = 1 \)? By Exercise 6.0.1, we know this is true for the Eisenstein series \( E_k \), and now we have seen it is true for \( \Delta \), and by Exercise 6.1.6, any \( E_k\Delta \) where \( k = 4, 6, 8, 10, 14 \).

In general, if you have two power series (or Fourier expansions) \( f \) and \( g \) with multiplicative coefficients, the formal product \( fg \) will not have multiplicative Fourier coefficients, so there is no reason to expect all modular forms—or even products of modular forms with multiplicative Fourier coefficients—to have multiplicative coefficients. For instance, starting in weight 24 for the full modular group, we have a space of cusp forms of dimension greater than 1, namely \( S_{24}(1) = <\Delta^2, E_{12}\Delta> \). Since

\[
\Delta(z) = q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 - 6048q^6 + \cdots
\]

88
we compute
\[ \Delta^2(z) = q^2 - 48q^3 + 1080q^4 - 2944q^5 + 143820q^6 + \cdots. \]

Right away, we see two things: 1) the Fourier coefficients of \( q^2 \) and \( q^3 \) do not multiply to give the Fourier coefficient of \( q^6 \), and 2) \( a_1 = 0 \) so the multiplicative property \( a_1a_n = a_m a_n \) for \( \gcd(m, n) = 1 \) would mean \( a_n = 0 \) whenever \( n \) is divisible by two different primes, which is obviously not the case. (In fact something stronger is true, see Corollary 6.1.11 below.)

One can also check that
\[ E_{12} = 1 + \frac{65520}{691} (q + 2049q^2 + 177148q^3 + 4196353q^4 + 48828126q^5 + \cdots) \]
so
\[ E_{12} \Delta = q + \frac{65520}{691} \left( \frac{2039}{2730} q^2 + \frac{527191}{260} q^3 + \frac{525013709}{4095} q^4 + \cdots - \frac{666536766}{65} q^6 + \cdots \right) \]

While the first Fourier coefficient is 1 here, again one sees the products of the \( q^2 \) and \( q^3 \) coefficients does not give the \( q^6 \) coefficient (it’s not even the right sign!). So one cannot hope that the Fourier coefficients of \( E_k \Delta \) are multiplicative in general.

You might now ask, are there any elements of \( S_{24}(1) \) with multiplicative Fourier coefficients or does this property of having multiplicative Fourier coefficients only occur in small weight? We will see there always are cusp forms with multiplicative Fourier coefficients. First, let’s make an observation in the weight 12 case: while \( \frac{691}{65520} E_{12} \) and \( \Delta \) have multiplicative Fourier coefficients, one can’t expect this for any modular form in \( M_{12}(1) \). Namely if \( f \) and \( g \) have multiplicative Fourier coefficients, \( f + g \) will generally not. So while most forms in \( M_{12}(1) \) do not have multiplicative Fourier coefficients, \( M_{12}(1) \) is generated by two “nice” forms, \( \frac{691}{65520} E_{12} \) and \( \Delta \), which do.

Using the theory of Hecke operators, we will show that there exists such a “nice” basis for \( M_k(N) \) and \( S_k(N) \). The basic idea is to show that the Hecke operators \( T_n \) commute for \( \gcd(m, n) = 1 \) (and \( m, n \) prime to \( N \)). Then one defines an inner product \( \langle \cdot, \cdot \rangle \), called the Petersson inner product, on \( M_k(N) \) and shows each \( T_n \) is Hermitian with respect to \( \langle \cdot, \cdot \rangle \), i.e., \( \langle T_n f, g \rangle = \langle f, T_n g \rangle \). Then a well known theorem in linear algebra says that a commuting family of operators which are Hermitian with respect to \( \langle \cdot, \cdot \rangle \) can be simultaneously diagonalized by some orthonormal basis with respect \( f_1, \ldots, f_r \) to \( \langle \cdot, \cdot \rangle \). In other words, each \( f_i \) is an eigenform for \( T_n \), i.e., \( T_n f_i = \lambda f_i \) for some \( \lambda \in \mathbb{C} \).

**Definition 6.1.8.** Let \( f \in M_k(N) \) be nonzero. We say \( f \) is a (Hecke) eigenform if, for each \( n \) relatively prime to \( N \), there exists a (Hecke) eigenvalue \( \lambda_n \in \mathbb{C} \) such that \( T_n f = \lambda_n f \).

We say \( f \) is a normalized eigenform if the first nonzero Fourier coefficient is 1.

The arguments from Corollary 6.1.5 and Exercise 6.1.6 apply to show any Hecke eigenform has multiplicative Fourier coefficients. Precisely, work out

**Exercise 6.1.9.** Let \( f = \sum a_n q^n \in M_k(N) \) be a Hecke eigenform. Suppose \( \gcd(m, n) = \gcd(m, N) = \gcd(n, N) = 1 \). Then \( a_1 a_{m n} = a_m a_n \).

In other words, the basis \( f_1, \ldots, f_r \) which diagonalizes the \( T_n \)'s asserted above is a basis of Hecke eigenforms, and therefore a basis of \( M_k(N) \) with multiplicative Fourier coefficients in the sense of the previous exercise.

We remark that there is a technicality here we have ignored, namely that the Petersson inner product \( \langle f, g \rangle \) is only well defined when at least \( f \) or \( g \) is a cusp form. So the above argument will
only technically show that \(S_k(N)\) has a basis of eigenforms, but at least when \(N = 1\), we will see how this implies one can extend the basis of eigenforms of \(S_k(N)\) to a basis of eigenforms for \(M_k(N)\).

We remark that for an eigencusp form (a Hecke eigenform which is a cusp form) the first nonzero Fourier coefficient should be \(a_1\) in order for the Hecke operators to not force every Fourier coefficient to be zero.

**Lemma 6.1.10.** Let \(f = \sum a_n q^n \in M_k(1)\) be a Hecke eigenform and \(k > 0\). Then \(a_1 \neq 0\).

*Proof.* Write \(T_m f = \lambda_m f = \sum b_n q^n\). Then \(b_1 = a_m\). Consequently

\[ a_m = \lambda_m a_1. \]

In other words, if \(a_1 = 0\), then \(a_m = 0\) for all \(n \geq 0\). Consequently \(f(z) = a_0 \in M_0(1)\).

**Corollary 6.1.11.** Let \(f \in M_k(1)\). Then \(\Delta^2 f\) is not a Hecke eigenform.

In particular, this shows \(\Delta^2\) is not an eigenform.

Now let’s get to the first step in showing the existence of Hecke eigenforms, which is showing the Hecke operators are commutative.

**Lemma 6.1.12.** Suppose \(\gcd(m, N) = \gcd(n, N) = 1\) and \(\gcd(m, n) = 1\). Then the Hecke operators on \(M_k(N)\) satisfy \(T_m T_n = T_{mn}\).

*Proof.* Take \(f(z) = \sum a_r q^r \in M_k(N)\). Write \(T_n f(z) = \sum b_r q^r\) and \(T_m(T_n f)(z) = \sum c_r q^r\). Then

\[ b_r = \sum_{d | \gcd(n, r)} d^{k-1} a_{nr/d^2} \]

so

\[ c_r = \sum_{e | \gcd(m, r)} e^{k-1} b_{mr/e^2} = \sum_{e | \gcd(m, r)} e^{k-1} \sum_{d | \gcd(n, mr/e^2)} d^{k-1} a_{mnr/d^2e^2}. \]

Since \(\gcd(m, n) = 1\), \(d\) runs over the divisors of \(\gcd(n, r/e)\), and therefore \(d' = de\) runs over the divisors of \(mn\) and

\[ c_r = \sum_{d' | \gcd(mn, r)} \left( d' \right)^{k-1} a_{mnr/(d')^2} \]

so \(T_m T_n = T_{mn}\).

**Lemma 6.1.13.** Suppose \(p \nmid N\). The Hecke operators on \(M_k(N)\) satisfy

\[ T_{p^r} T_{p^s} = \sum_{d | \gcd(p^r, p^s)} d^{k-1} T_{p^{r+s/d^2}}. \] (6.2)

*Proof.* The proof here is modeled on the one in [Apo90, Theorem 6.13], but there appear to me to be errors in the \(r = 1\) case of loc. cit. I believe I have corrected them.

We may as well assume \(r \leq s\).

First let’s show the \(r = 1\) case, which just says

\[ T_{p} T_{p^s} = T_{p^{s+1}} + p^{k-1} T_{p^{s-1}}. \]
For $f \in M_k(n)$,
\[
T_p^s f(z) = p^{s(k-1)} \sum_{0 \leq i \leq s} \sum_{0 \leq b < p^i} f \left( \frac{p^{s-i}b}{p^i} \right).
\]
In particular
\[
T_p g(z) = p^{(k-1)} g(pz) + p^{-1} \sum_{0 \leq b' < p} g \left( \frac{z + b'}{p} \right)
\]
so
\[
T_p T_p^s f(z) = p^{(s+1)(k-1)} \sum_{0 \leq i \leq s} \sum_{0 \leq b < p^i} f \left( \frac{p^{s+1-i}z + b}{p^i} \right)
\]
\[
+ p^{-1} p^{s(k-1)} \sum_{0 \leq b' < p} \sum_{0 \leq i \leq s} \sum_{0 \leq b < p^i} f \left( \frac{p^{s-i}(z + b') + pb}{p^{i+1}} \right).
\]
Note the $i = s$ term from (6.5) is
\[
p^{-1-s} \sum_{0 \leq b' < p} \sum_{0 \leq b < p^i} f \left( \frac{z + b' + pb}{p^{s+1}} \right) = p^{-1-s} \sum_{0 \leq b < p^{s+1}} f \left( \frac{z + b}{p^{s+1}} \right).
\]
Thus adding the $i = s$ term from (6.5) to (6.4) gives (6.3) with $s$ replaced by $s + 1$, i.e., they sum to $T_{p^{s+1}} f(z)$.

Now the remaining terms, i.e., the $i < s$ terms from (6.5), are
\[
p^{-1} p^{s(k-1)} \sum_{0 \leq b' < p} \sum_{0 \leq i \leq s-1} \sum_{0 \leq b < p^i} f \left( \frac{p^{s-1-i}z + b + p^{s-1-i}b'}{p^i} \right)
\]
If $i \leq \frac{s-1}{2}$, then $\frac{p^{s-1-i}b'}{p^i} \in \mathbb{Z}$, so by periodicity of $f$ there is no dependence on $b'$ and the contribution for such a fixed $i$ is
\[
p^{s(k-1)} p^{-i} \sum_{0 \leq b < p^i} f \left( \frac{p^{s-1-i}z + b + p^{s-1-i}b'}{p^i} \right).
\]
In fact, for any $i$, $b + p^{s-1-i}b'$ mod $p^i$ runs over the set of residue classes mod $p^i$ exactly $p$ times, so the contribution for any $i$ is given again by (6.6). This proves the $r = 1$ case.

Now we suppose (6.2) is true up to some fixed $r$. By the $r = 1$ case we have
\[
T_{p^{r+1}} T_{p^r} = (T_p T_{p^r}) T_{p^r} - p^{k-1} T_{p^{r-1}} T_{p^r}.
\]
By the inductive assumption that (6.2) is true for $r$, we also have
\[
T_p (T_{p^r} T_{p^s}) = \sum_{0 \leq i \leq r} p^{i(k-1)} T_p T_{p^{r+s-2i}}.
\]
Comparing these, and making another use of the $r = 1$ case, gives

$$T_{p^r+1} T_{p^s} = \sum_{0 \leq i \leq r} p^{i(k-1)} T_{p^{i+2}} - p^{k-1} T_{p^{i-1}} T_{p^s}$$

$$= \sum_{0 \leq i \leq r} \left( p^{i(k-1)} T_{p^{i+1}} + p^{(i+1)(k-1)} T_{p^{i-1}} \right) - p^{k-1} T_{p^{i-1}} T_{p^s}.$$

Expanding out the last term using the $r - 1$ case of (6.2) gives

$$T_{p^r+1} T_{p^s} = \sum_{0 \leq i \leq r+1} p^{i(k-1)} T_{p^{i+1}}.$$

The previous two lemmas combine to give the following.

**Corollary 6.1.14.** The Hecke operators $\{T_n\}_{\gcd(n,N)=1}$ on $M_k(N)$ commute.

The former lemma also shows knowing what the prime power Hecke operators do tells you what all Hecke operators do, and the latter lemma shows that the prime power Hecke operators are determined by the prime Hecke operators. In particular, we see

**Corollary 6.1.15.** If $f \in M_k(N)$ is an eigenform for each $T_p$ with $p \nmid N$ prime, i.e., $T_p f = \lambda_p f$ for some $\lambda_p \in \mathbb{C}$, then $f$ is a Hecke eigenform.

**Proof.** The $r = 1$ case of (6.2) says, for $j = s - 1 \geq 2$,

$$T_{p^j} = T_p T_{p^{j-1}} - p^{k-1} T_{p^{j-2}}.$$

If $T_p^j f = \lambda_p^j f$ for some $\lambda_p^j$ whenever $i < j$, then we see

$$T_{p^j} f = T_p (\lambda_p^{j-1} f) - p^{k-1} \lambda_p^{j-2} f = \lambda_p^j f$$

where

$$\lambda_p^j = \left( \lambda_p^{j-1} - p^{k-1} \lambda_p^{j-2} \right). \quad (6.7)$$

Hence by induction, we see if $T_p f = \lambda_p f$ then $T_{p^j} f = \lambda_p^j f$ for some $\lambda_p^j$—precisely, with $\lambda_p^j$ satisfying the recursion in (6.7).

Now let $n$ be relatively prime to $N$, and write $n = p_1^{e_1} \ldots p_r^{e_r}$. By Lemma 6.1.12, we have

$$T_n f = \left( \prod_{i=1}^r T_{p_i^{e_i}} \right) f = \prod_{i=1}^r \lambda_{p_i^{e_i}} f.$$
6.2 Petersson inner product

We know that $M_k(\Gamma)$ is a finite dimensional complex vector space, and therefore can be made into a Hilbert space, i.e., an inner product space. The standard way to make a function space into a Hilbert space is with the $L^2$ inner product. Namely, if $f, g \in L^2(X)$ for some space $X$, then the inner product is given by

$$\langle f, g \rangle = \int_X f(x)g(x)dx.$$ 

While modular forms are not $L^2$ on $\mathfrak{H}$, they are essentially $L^2$ on relevant Riemann surface $X = \Gamma \backslash \mathfrak{H}$. We say essentially here, because they are of course not actually functions on $\Gamma \backslash \mathfrak{H}$ due to the automorphism transformation factor except in the uninteresting (constant) case of $k = 0$.

Let’s see what happens when we naively try to make $f \in M_k(\Gamma)$ into a function on $X = \Gamma \backslash \mathfrak{H}$. For simplicity, consider $\Gamma = \text{PSL}_2(\mathbb{Z}) = \langle S, T \rangle$. We already have $f(Tz) = f(z+1) = f(z)$ so $f$ descends to a function on the infinite cylinder $\langle T \rangle \backslash \mathfrak{H}$. We would like to modify $f$ to a function $F = F_f$ which still satisfies $F(z+1) = F(z)$ but also satisfies $F(Sz) = F(-1/z) = F(z)$.

Since the invariance under $T$ is a transformation rule in $x = \text{Re}(z)$ which we don’t want to mess up, we might just try to impose invariance under $S$ by modifying $y = \text{Im}(z)$. Since $f(S \cdot iy) = f(-1/iy) = f(i/y) = (iy)^k f(iy)$, it makes sense to consider the function (which will be called a Maass form)

$$F(z) := y^{k/2} f(z) = \text{Im}(z)^{k/2} f(z). \quad (6.8)$$

Then we still have $F(z+1) = F(z)$ for all $z$ and now

$$F(-1/iy) = F(i/y) = y^{-k/2} f(i/y) = \pm i^k y^{k/2} f(i/y) = (-1)^{k/2} F(f(iy),$$

so at least $F(-1/iy) = F(iy)$ when $k \equiv 0 \text{ mod } 4$. (One could let $F(z) = (iy)^{k/2} f(z)$ so that we actually have $F(-1/iy) = F(iy)$, but one typically considers $F$ as defined in (??), since the sign $(-1)^{k/2}$ will not matter in the end.) It doesn’t quite satisfy $F(-1/z) = F(z)$ for all $z$ but it will be close, and we can say how close.

In general, for any $f \in M_k(\Gamma)$, we can associate to $f$ the Maass form $F$ defined by (6.8). Recalling that

$$|j(\gamma, z)|^2 = \frac{\text{Im}(z)}{\text{Im}(\gamma z)} \quad (6.9)$$

(cf. (3.1)), we see

$$F(\gamma z) = \text{Im}(\gamma z)^{k/2} f(\gamma z) = \text{Im}(\gamma z)^{k/2} j(\gamma, z)^k f(z) = j_0(\gamma, z)^k \text{Im}(z)^{k/2} f(z) = j_0(\gamma, z)^k F(z)$$

where

$$j_0(\gamma, z) = \frac{j(\gamma, z)}{|j(\gamma, z)|}.$$ 

In other words, the Maass form $F$ (not holomorphic) is not quite invariant under $\Gamma$, but it is up to a factor $j_0(\gamma, z)$ of absolute value 1. This is good enough for our purposes.* Namely, if also

*We remark that one can make $f$ invariant under $\Gamma$ at the expense of working on a larger space than $\mathfrak{H}$. Namely, one has the surjective map $\text{PSL}_2(\mathbb{R}) \rightarrow \mathfrak{H}$ given by $\gamma \mapsto \gamma \cdot i$. Since $\text{SO}(2)$ stabilizes $i$, one can view $\text{PSL}_2(\mathbb{R})/\text{SO}(2) \cong \mathfrak{H}$. Thus one can lift $f$ to a function on $\text{PSL}_2(\mathbb{R})$ by $f(\gamma) = f(\gamma \cdot i)$ and define the automorphic form $\phi_f(\gamma) = j(\gamma, i)^{k/2} f(\gamma)$. This makes $\phi_f$ invariant under $\Gamma$ so it is a function on $\Gamma \backslash \text{PSL}_2(\mathbb{R})$. Hence, viewed as functions on $\text{PSL}_2(\mathbb{R})$, the passage from modular forms to automorphic forms trades right $\text{SO}(2)$ invariance for left $\Gamma$ invariance.
\(g \in M_k(\Gamma)\) and \(G(z) = \text{Im}(z)^{k/2} g(z)\) is the associated Maass form, it means the product

\[
F(\gamma z)G(\gamma z) = j_0(\gamma, z)^k j_0(\gamma, z)^{-k} F(z)G(z) = F(z)G(z)
\]
is invariant under \(\Gamma\).

Consequently, we can define the inner product

\[
\langle f, g \rangle = \int_X F(x)G(x)\,d\omega = \int_{\Gamma\backslash \mathcal{H}} y^k F(z)G(z)\,d\omega,
\]
where \(d\omega\) is an area measure on the Riemann surface \(X = \Gamma\backslash \mathcal{H}\). Now in practice, you want to express this as an integral on a fundamental domain for \(\Gamma\) inside \(\mathcal{H}\), so we need to know what the hyperbolic measure should be.

**Lemma 6.2.1.** The measure \(\frac{dx\,dy}{y^2} = \frac{dz\,	ext{Im}(z)^2}{|cz + d|^2}\) is an invariant measure on the hyperbolic plane \(\mathcal{H}\), i.e., for \(\gamma \in \text{PSL}_2(\mathbb{R})\) and we have

\[
\int_{\mathcal{H}} f(\gamma z) \frac{dx\,dy}{y^2} = \int_{\mathcal{H}} f(z) \frac{dx\,dy}{y^2},
\]
for any integrable (w.r.t. to this measure) function \(f\) on \(\mathcal{H}\).

**Proof.** Write \(\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\), and let

\[
w = \gamma z = \frac{(ad + bc)x + bd + acy^2 + iy}{c^2x^2 + c^2y^2 + d^2} = u + iv.
\]

Note

\[
\frac{dw}{dz} = \frac{1}{|cz + d|^2} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y},
\]

Then the Jacobian determinant

\[
\begin{vmatrix}
\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\
\frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}
\end{vmatrix}^{-1} = \begin{vmatrix}
|cz + d|^{-2} & 0 \\
0 & |cz + d|^{-2}
\end{vmatrix}^{-1} = |j(\gamma, z)|^4.
\]

Since \(|j(\gamma, z)|^2 = \frac{\text{Im}(z)}{\text{Im}(\gamma z)} = \frac{y}{\text{Im}(\gamma z)}\), we have

\[
\int_{\mathcal{H}} f(\gamma z) \frac{dx\,dy}{y^2} = \int_{\mathcal{H}} f(w)|j(\gamma, z)|^4 \frac{du\,dv}{y^2} = \int_{\mathcal{H}} f(w) \frac{du\,dv}{v^2}.
\]

\[\square\]

**Definition 6.2.2.** Let \(f, g \in M_k(\Gamma)\). The Petersson inner product of \(f\) with \(g\) is defined to be

\[
\langle f, g \rangle := \frac{1}{[\text{PSL}_2(\mathbb{Z}) : \Gamma]} \int_{\Gamma\backslash \mathcal{H}} y^k f(z)g(z)\frac{dx\,dy}{y^2}
\]
(6.10)

whenever the integral converges.
The above discussion shows that the integrand is invariant under $\Gamma$, so the defining integral makes sense, provided it converges. Consequently, we can also write this integral as

$$\langle f, g \rangle := \frac{1}{|PSL_2(\mathbb{Z}) : \Gamma|} \int_{\mathcal{F}} y^k f(z) \overline{g(z)} \frac{dx dy}{y^2}$$

(6.11)

where $\mathcal{F}$ is any fundamental domain for $\Gamma$.

The normalization factor $|PSL_2(\mathbb{Z}) : \Gamma|^{-1}$ roughly cancels out the volume of $\Gamma \backslash \mathfrak{H}$, defined to be

$$\text{vol}(\Gamma \backslash \mathfrak{H}) = \int_{\Gamma \backslash \mathfrak{H}} \frac{dx dy}{y^2}.

\textbf{Lemma 6.2.3.} Let $\Gamma$ be a congruence subgroup in $PSL_2(\mathbb{Z})$. Then $\text{vol}(\Gamma \backslash \mathfrak{H})$ is finite.

\textit{Proof.} Let $\mathcal{F}'$ be a fundamental domain for $\Gamma$. By Lemma 3.4.9,

$$\mathcal{F}' = \bigcup_{i=1}^{[PSL_2(\mathbb{Z}) : \Gamma]} \alpha_i \mathcal{F},$$

where $\mathcal{F}$ is the standard fundamental domain for $PSL_2(\mathbb{Z})$ and $\alpha_i \in PSL_2(\mathbb{Z})$. Since the area measure is invariant under the action of $PSL_2(\mathbb{R})$,

$$\text{vol}(\mathcal{F}') = \sum \text{vol}(\alpha_i \mathcal{F}) = [PSL_2(\mathbb{Z}) : \Gamma] \text{vol}(\mathcal{F}),$$

so it suffices to show $\mathcal{F}$ has finite volume. Note

$$\text{vol}(\mathcal{F}) = \int_{-1/2}^{1/2} \int_{\sqrt{1-x^2}}^{\infty} \frac{dx dy}{y^2} \leq \int_{-1/2}^{1/2} \int_{1/2}^{\infty} dx dy = 2 \int_{-1/2}^{1/2} dx = 2.$$

\hfill $\Box$

The proof is of course valid not just for congruence subgroups, but any finite index subgroup of $PSL_2(\mathbb{Z})$, and consequently any subgroup of $PSL_2(\mathbb{R})$ commensurable with $PSL_2(\mathbb{Z})$.

\textbf{Exercise 6.2.4.} Compute $\text{vol}(X_0(1)) = \text{vol}(PSL_2(\mathbb{Z}) \backslash \mathfrak{H})$.

While not that interesting, the above lemma implies the Petersson inner product converges for $f, g \in M_0(\Gamma) = \mathbb{C}$. The issue in general is that, say for the standard fundamental domain of $PSL_2(\mathbb{Z})$, the $y^k$ can grow too fast for the Petersson inner product to converge unless $f(z)\overline{g(z)} \to 0$ as $y \to \infty$. Similarly, if we tend to a cusp in $\mathbb{Q}$, $f(z)\overline{g(z)}$ may grow too fast (cf. Section 4.3) for the inner product to converge. This suggests (when $k > 0$) we need $f$ or $g$ to be a cusp form for $(f,g)$ to converge.

\textbf{Proposition 6.2.5.} Let $f, g \in M_k(\Gamma)$. The Petersson inner product $\langle f, g \rangle$ converges if either $f \in S_k(\Gamma)$ or $g \in S_k(\Gamma)$.

\textit{Proof.} We assume $f \in S_k(\Gamma)$. The case $g \in S_k(\Gamma)$ is similar.

Let $\mathcal{F}$ be a fundamental domain for $\Gamma \backslash \mathfrak{H}$ and $\overline{\mathcal{F}}$ be its closure in $\overline{\mathfrak{H}}$. Write $\overline{\mathcal{F}} = K \cup \bigcup U_i$, where this is a finite union of subsets with $K$ compact in $\mathfrak{H}$ and each $U_i \subset \overline{\mathfrak{H}}$ containing exactly one cusp, say $z_i$. Since $y^{k-2}f(z)\overline{g(z)}$ is bounded on $K$ it suffices to show this is bounded on each $U_i$.

As in the proof of Lemma 4.5.8, we know $f(z) \to 0$ exponentially fast as $z \to z_i$, whereas $y^{k-2}g(z)$ has at most a finite order pole at $z_i$. Thus $y^{k-2}f(z)\overline{g(z)}$ is bounded on each $U_i$, and because $\text{vol}(\overline{\mathcal{F}}) = \text{vol}(\mathcal{F}) < \infty$, the Petersson inner product converges. \hfill $\Box$

95
What this means is that the Petersson inner product defines an inner product (it is clearly sesquilinear, i.e., linear in the first vector and anti, or conjugate, linear in the second) on the space of cusp forms $S_k(N)$. This inner product almost extends to an inner product on $M_k(N)$, but $\langle f, g \rangle$ need not converge when both $f$ and $g$ are not cusp forms.

While we have failed to make $M_k(N)$ a Hilbert space, we have at least made $S_K(N)$ one. Most importantly, the inner product commutes with the Hecke operators.

**Theorem 6.2.6.** The Hecke operators $\{T_n\}$ on $S_k(N)$ are hermitian with respect to the Petersson inner product, i.e.,

$$\langle T_n f, g \rangle = \langle f, T_n g \rangle$$

for $f, g \in S_k(N)$.

**Proof.** Fix a fundamental domain $F$ for $\Gamma_0(N)$. For simplicity of notation, we extend $\langle f, g \rangle$ by the (6.11) to any functions on $\mathfrak{H}$ for which the integral converges.

First we claim that for any $\tau = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{PSL}_2(\mathbb{R})$, we have

$$\langle f|_{\tau,k}, g \rangle = \langle f, g|_{\tau^{-1},k} \rangle,$$

where we extend $|_{\tau,k}$ to $\tau \in \text{GL}_2(\mathbb{R})^+$ = $\{ \alpha \in \text{GL}_2(\mathbb{R}) : \det(\alpha) > 0 \}$ by (4.10). (The condition of positive determinant ensures $\tau$ maps $\mathfrak{H}$ to $\mathfrak{H}$.)

This follows since

$$g^k f(z)|_{\tau,k} g(z) = \frac{\text{Im}(z)^k}{(cz+d)^k} f(\tau z) g(z) = \text{Im}(w)^k (cz+d)^k f(w) g(\tau^{-1} w),$$

where we put $w = \tau z$ and used (6.9). Then observing $(cz+d)^k = j(\tau, z)^k = j(\tau^{-1}, w)^{-k}$ shows this equals

$$\text{Im}(w)^k f(w) g(w)|_{\tau^{-1},k}$$

which implies our claim.

Recall

$$T_n f = n^{k-1} \sum_{\mu \in \text{SL}_2(\mathbb{Z}) \setminus \mathcal{M}_n(\mathbb{Z})} f|_{\mu,k} = n^{k-1} \sum_{i=1}^r f|_{\mu_i,k}$$

Where $\mu_1, \ldots, \mu_r$ is a set of representatives for $\text{SL}_2(\mathbb{Z}) \setminus \mathcal{M}_n(\mathbb{Z})$. Let $\nu = \begin{pmatrix} \sqrt{n} & 0 \\ 0 & \sqrt{n} \end{pmatrix}$ and write $\mu_i = \nu \tau_i$, where $\tau_i \in \text{PSL}_2(\mathbb{R})$.

Then

$$\langle T_n f, g \rangle = n^{k-1} \sum \langle f|_{\mu_i,k}, g \rangle = n^{k-1} \sum \langle (f|_{\nu,k})|_{\tau_i,k}, g \rangle = n^{k-1} \sum \langle (f|_{\nu,k}), g|_{\tau_i^{-1},k} \rangle.$$

Note $f(z)|_{\nu,k} = n^{-k/2} f(z)$. Hence

$$\langle T_n f, g \rangle = n^{k-1} \sum n^{-k/2} \langle f, g|_{\tau_i^{-1},k} \rangle = n^{k-1} \sum \langle f, g|_{\tau_i^{-1},\nu,k} \rangle = \langle f, T_n g \rangle.$$

Here the last equality follows because $\tau_i^{-1}\nu$ also runs over a set of representatives for $\text{SL}_2(\mathbb{Z}) \setminus \mathcal{M}_n(\mathbb{Z})$, whence

$$T_n g = n^{k-1} \sum g|_{\tau_i^{-1},\nu,k}.$$
Corollary 6.2.7. The space $S_k(N)$ has a basis consisting of Hecke eigenforms.

Proof. Since $\{T_n\}$ is a family of commuting operators which are hermitian with respect to $\langle \cdot, \cdot \rangle$, a well known theorem in linear algebra tells us that an orthonormal basis with respect to $\langle \cdot, \cdot \rangle$ is a basis of eigenvectors for all $\{T_n\}$.

By Exercise 6.1.9, we know the normalized ($a_1 = 1$) Hecke eigenforms in $S_k(N)$ have Fourier coefficients $a_n$ which are multiplicative for $n$ relatively prime to the level $N$. Hence the cusp forms are generated by forms whose Fourier coefficients are some kind of arithmetic sequences.

There are some more questions we can ask at this point to push our theory further.

- Does the whole space of modular forms $M_k(N)$ actually have a basis of Hecke eigenforms?
- Can we define the Hecke operators $T_n$ on $M_k(N)$ when $\gcd(n, N) > 1$?
- Can we actually find a basis for $S_k(N)$ (or $M_k(N)$) whose Fourier coefficients $a_n$ are multiplicative for all $n$?

Here we will just briefly discuss these questions.

First, observe that Theorem 6.2.6 is actually valid for $f, g \in M_k(N)$ when $\langle T_n f, g \rangle$ converges. One can use this to deduce that $M_k(N)$ does have a basis of eigenforms, and in fact we already know it at least one case.

Example 6.2.8. The space $M_k(1)$ has a basis of Hecke eigenforms. To see this, recall $M_k(1) = \mathbb{C}E_k \oplus S_k(1)$. By Exercise 6.0.1, we know $E_k$ is an eigenform, hence the basis of eigenforms for $S_k(1)$ can be extended to a basis of eigenforms for $M_k(1)$.

In fact, the Eisenstein series (for $M_k(N)$, or more generally for $M_k(\Gamma)$) are orthogonal to the space of cusp forms in the sense that $\langle E, f \rangle = 0$ whenever $E$ is an Eisenstein series and $f$ is a cusp form. Often one uses this relation to define Eisenstein series: the space of Eisenstein series for $M_k(\Gamma)$ will be the elements $E \in M_k(\Gamma)$ such that $\langle E, f \rangle = 0$ for all $f \in S_k(\Gamma)$.

Next, one can define Hecke operators $T_p$ on $M_k(N)$ when $\gcd(n, N) > 1$, however a definition in terms of $M_{n,N}(\mathbb{Z})$ is somewhat trickier as the analogue Lemma 6.1.3 is not as nice. For simplicity, let’s just talk about $T_p$ where $p | N$, since all the $T_{n'}$’s can be constructed out of just the $T_p$’s. One can define

$$T_p f = \sum_{0 \leq j < p} f\left|\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right|_k,$$

and if $f(z) = \sum a_n q^n$, one obtains

$$T_p f(z) = \sum a_{pn} q^n.$$

The definition of Hecke operators can be made more uniform (including for arbitrary congruence subgroups $\Gamma$) by working in terms of double cosets (see, e.g., [DS05] or [Kob93]).

While the Hecke operators $T_p$ for $p | N$ can be defined without much trouble, one does not in general have a basis of eigenforms for $S_k(N)$ for all $T_n$. Roughly the issue is the following. Say $N = pq$ with $p, q$ distinct primes and $f \in S_k(q)$ is a Hecke eigenform. In particular, $f$ is an eigenform for $T_p$ (on $S_k(q)$) Then automatically $f \in S_k(N)$. However $T_p$ on $S_k(N)$ acts differently, so $f$ is no longer an eigenform for $T_p$. 

97
More generally, if \( d | N \) and \( f(z) \in S_k(N/d) \), then \( f(dz) \in S_k(N) \) and we the subspace generated by all such elements the space \( S^\text{old}_k(N) \) of \emph{old forms}. Define the space of \emph{new forms} \( S^\text{new}_k(N) \) to be its orthogonal complement in \( S_k(N) \) (w.r.t. the Petersson inner product). Then it is true that \( S^\text{new}_k(N) \) has a basis of eigenforms for \( \text{all } T_p \), as opposed to just for \( p \nmid N \).

Finally we remark that the theory of Hecke operators for an arbitrary congruence subgroup \( \Gamma \) is similar.

See [Ste07] for how to computationally find a basis of Hecke eigenforms for a given space \( M_k(\Gamma) \).