10 HECKE OPERATORS

10.1. Introduction The motivation here is to find operators T which act on modular forms f of given weight 2k in such a way that the Fourier coefficients of f and Tf satisfy "useful" relationships. A crucial rôle is played by eigenforms f which satisfy

$$Tf = \lambda_T f.$$

One can set up the theory in rather greater generality. However we will follow Hecke who made the following definition.

Definition 10.1. For $k \in \mathbb{Z}$ and any n = 1, 2, ... the operator T_n is defined on the set of weakly modular functions of weight 2k by

$$(T_n f)(z) = n^{2k-1} \sum_{d|n} d^{-2k} \sum_{b=0}^{d-1} f\left(\frac{nz+bd}{d^2}\right).$$

10.2. The Fourier Expansion. For brevity we will use the notation $e(\alpha)$ for $e^{2\pi i\alpha}$ and q = e(z).

Theorem 10.1. Let f be a modular function of weight 2k, with Fourier expansion at ∞

$$f(z) = \sum_{m \in \mathbb{Z}} c(m)e(mz).$$
(1)

Then $T_n f$ has the Fourier expansion

$$(T_n f)(z) = \sum_{m \in \mathbb{Z}} \gamma_n(m) e(mz)$$

where

$$\gamma_n(m) = \sum_{d \mid (m,n)} d^{2k-1}c(mnd^{-2}).$$
1

10. HECKE OPERATORS

Proof. By definition the series (1) converges for q in a punctured disc centred at q = 0. In the interior of this annulus it converges absolutely. Thus

$$\begin{aligned} (T_n f)(z) &= \sum_{m \in \mathbb{Z}} \sum_{d|n} \left(\frac{n}{d}\right)^{2k-1} c(m) e\left(\frac{mnz}{d^2}\right) \frac{1}{d} \sum_{b=0}^{d-1} e(mb/d) \\ &= \sum_{m \in \mathbb{Z}} \sum_{d|(m,n)} \left(\frac{n}{d}\right)^{2k-1} c(m) e\left(\frac{mnz}{d^2}\right) \\ &= \sum_{d|n} \sum_{l \in \mathbb{Z}} \left(\frac{n}{d}\right)^{2k-1} c(ld) e(lnz/d) \\ &= \sum_{r|n} r^{2k-1} c(ln/r) e(lrz) \\ &= \sum_{m \in \mathbb{Z}} \sum_{r|(m,n)} r^{2k-1} c(mnr^{-2}) e(mz). \end{aligned}$$

In Definition 10.1, the general term can be rewritten as

$$f\left(\frac{(n/d)z+b}{d}\right).$$

Thus we are particularly interested in the behaviour of f under the action of $A = \begin{pmatrix} n/d & b \\ 0 & d \end{pmatrix}$.

Now det A = n. Thus we consider general $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with det A = n, a transformation of order n, and for our purposes we can associate Az with $z \mapsto \frac{az+b}{cz+d}$. Let $\Gamma(n)$ denote the set of all such transformations. Thus the modular group is $\Gamma(1)$. We say that $A, B \in \Gamma(n)$ are equivalent when there is a $C \in \Gamma$ such that A = CB. This is obviously an equivalence relation, and partitions $\Gamma(n)$ into equivalence classes. We could write $\Gamma(n)/\Gamma$ for the set of equivalence classes.

10.3. The structure of $\Gamma(n)$.

Theorem 10.2. Every equivalence class in $\Gamma(n)/\Gamma$ contains a representative $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ with d > 0

Proof. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a given member of the class. If c = 0, then we are done since $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ and $\begin{pmatrix} -a & -b \\ 0 & -d \end{pmatrix}$ represent the same transformation. If $c \neq 0$, then we choose r, s with (r, s) = 1 so that s/r = -a/c, and choose p, q so that ps - qr = 1. Hence

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} pa + qc & pb + qd \\ ra + sc & rb + sd \end{pmatrix},$$

$$ra + sc = 0$$
, det $\begin{pmatrix} p & q \\ r & s \end{pmatrix} = ps - qr = 1$.

Theorem 10.3. A complete set of non-equivalent elements of $\Gamma(n)$ is given by the $\sigma(n) = \sum_{d|n} d$ transformations $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ formed by taking ad = n, d > 0 and b ranging over a complete set of residues modulo d.

Proof. By Theorem 10.2, given $B \in \Gamma(n)$ there is an $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ equivalent to B, and ad = n, d > 0. We show that

$$A_1 = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix} \text{ is equivalent to } A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$
(2)

if and only if

$$a_1 = a, d_1 = d, b_1 \equiv b \pmod{d}.$$
 (3)

Suppose (2). Then $A_1 = \begin{pmatrix} u & v \\ w & x \end{pmatrix} A = \begin{pmatrix} ua & ub + vd \\ wa & wb + xd \end{pmatrix}$ with ux - vw = 1. Since $ad = n \neq 0, w = 0$. Then $ux = 1, u = x = \pm 1$. As $xd > 0, u = x = 1, a_1 = a, b_1 = b + vd$. Suppose (s). Let $C = \begin{pmatrix} 1 & \frac{b_1 - b}{d} \\ 0 & 1 \end{pmatrix}$. Then, by (2), $CA = A_1$.

Definition 10.2. Let \mathcal{A}_n denote the set of transformations $A \in \Gamma(n)$, $Az = \frac{az+b}{d}$ with $ad = n, d > 0, 0 \le b < d$.

The set \mathcal{A}_n is a set of representatives for $\Gamma(n)/\Gamma$. Moreover if

$$A' = \begin{pmatrix} a & b + \lambda d \\ 0 & d \end{pmatrix}, \quad C = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix},$$

then A' = CA and for any weakly modular function f(A'z) = f(CAz) = f(Az). Thus we could take any complete set of residues modulo d for the b in our application.

By Definition 10.1,

$$(T_n f)(z) = \frac{1}{n} \sum_{A \in \mathcal{A}_n} a^{2k} f(Az).$$
(4)

Theorem 10.4. Suppose that $A_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \in \Gamma(n)$ and $V_1 \in \Gamma$. Then there are $A_2 = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix} \in \mathcal{A}_n$ and $V_2 \in \Gamma$ such that $A_1V_1 = V_2A_2$. Moreover, if $c_1 = 0$, $V_j = \begin{pmatrix} \alpha_j & \beta_j \\ \gamma_j & \delta_j \end{pmatrix}$, then $a_1(\gamma_2A_2z + \delta_2) = a_2(\gamma_1z + \delta_1)$.

Proof. Obviously $A_1V_1 \in \Gamma(n)$. Choose $A_2 \in \mathcal{A}_n$ equivalent to A_1V_1 in accordance with Theorem 10.2. Then there is a $V_2 \in \Gamma$ such that $V_2A_2 = A_1V_1$ as required. Suppose $c_1 = 0$. Then

$$A_1V_1 = \begin{pmatrix} a_1\alpha_1 + b_1\gamma_1 & a_1\beta_1 + b_1\delta_1 \\ d_1\gamma_1 & d_1\delta_1 \end{pmatrix},$$

$$V_2A_2 = \begin{pmatrix} \alpha_2a_2 & \alpha_2b_2 + \beta_2d_2 \\ \gamma_2a_2 & \gamma_2b_2 + \delta_2d_2 \end{pmatrix}.$$

Hence

$$d_1\gamma_1 = \gamma_2 a_2, \ d_1\delta_1 = \gamma_2 b_2 + \delta_2 d_2$$
$$\frac{n}{a_1}\gamma_1 = \gamma_2 \frac{n}{d_2}, \quad \delta_2 = \frac{d_1}{d_2}\delta_1 - \frac{b_2}{d_2}\gamma_2$$
$$\gamma_2 = \frac{d_2}{a_1}\gamma_1, \quad \delta_2 = \frac{a_2}{a_1}\delta_1 - \frac{b_2}{a_1}\gamma_1$$

Thus $a_1(\gamma_2 A_2 z + \delta_2) = a_1 \gamma_2 \frac{a_2 z + b_2}{d_2} + (a_2 \delta_1 - b_2 \gamma_1) = \gamma_1(a_2 z + b_2) - b_2 \gamma_1 + a_2 \delta_1.$

10.4. Properties of T_n . We are now in a position to prove the first desirable property of T_n , namely that if f is weakly modular of weight of 2k, then so is $T_n f$.

Theorem 10.5. Suppose that f is weakly modular of weight 2k and $V = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma$. Then

$$(T_n f)(Vz) = (\gamma z + \delta)^{2k} (T_n f)(z).$$

Proof. By (4),

$$(T_n f)(Vz) = \frac{1}{n} \sum_{A \in \mathcal{A}_n} a^{2k} f(AVz).$$

Given A and V, by Theorem 10.4 there are A_1 , V_1 such that $A_1 \in \mathcal{A}_n$, $AV = V_1A_1$, $V_1 \in \Gamma$ and $a(\gamma_1A_1z + \delta_1) = a_1(\gamma z + \delta)$. Thus

$$a^{2k}f(AVz) = a^{2k}f(V_1A_1z) = (a(\gamma_1A_1z + \delta_1))^{2k}f(A_1z) = a_1^{2k}(\gamma_1z + \delta_1)^{2k}f(A_1z).$$

Moreover, if $A'V = V'_1A'_1$, then A is equivalent to A' if and only if A_1 is equivalent to A'_1 . Thus A_1 runs over \mathcal{A}_n as A does.

Theorem 10.6. If f is a weakly modular function, then so is $T_n f$. If $f \in M_k$, then so is $T_n f$. If $f \in M_k^0$, then so is $T_n f$.

Proof. The first part is immediate from the previous theorem. The second part then follows from Theorem 10.1 since if c(m) has its support on the non-negative integers, so does $\gamma_n(m)$. Finally, in Theorem 10.1, if the support of the c(m) is \mathbb{N} , then so is that of $\gamma_n(m)$.

We can now begin to explore the structure of the T_n . These operators turn out to be "multiplicative" in the usual sense.

Theorem 10.7. Suppose that $m, n \in \mathbb{N}$ and (m, n) = 1. Then $T_m T_n = T_{mn}$ and T_m and T_n commute.

Proof. By (4),

$$(T_m(T_nf))(z) = \frac{1}{mn} \sum_{A' \in \mathcal{A}_m} a'^{2k} \sum_{A \in \mathcal{A}_n} a^{2k} f(A'Az).$$

We have

$$A'' = A'A = \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} a'a & a'b + b'd \\ 0 & d'd \end{pmatrix}.$$

The number d'd runs over a set of (positive) divisors of mn as d' and d do of m and n respectively. Since a'|m and d|n we have (a',d) = 1. Every residue class modulo d'd can be written in the form u + vd with $0 \le u < d$, $o \le v < d'$. Choose b so that $a'b \equiv u$ (mod d). Then choose b' so that $b' \equiv v + (u - a'b)/d \pmod{d'}$. Then $a'b + b'd \equiv u + vd \pmod{d'd}$. Since there are only d'd choices for b and b' it follows that a'b + b'd ranges over a complete set of resides modulo d'd as b' and b do modulo d' and d respectively. Finally a'a = mn/(d'd). Thus

$$(T_m(T_n f))(z) = \frac{1}{mn} \sum_{A'' \in \mathcal{A}_{mn}} (a'')^{2k} f(A'' z) = (T_{mn} f)(z).$$

That T_m and T_n commute is immediate.

To avoid suffices becoming too cumbersome we use the notation

$$T(n) = T_n.$$

The structure of the Hecke operators is quite surprising as the following theorem shows.

Theorem 10.8. Let p be a prime and let $r \in \mathbb{N}$ and suppose that f is weakly modular of weight 2k. Then

$$T(p)(T(p^{r})f) = T(p^{r+1})f + p^{2k-1}T(p^{r-1})f.$$

Proof. By Definition 10.1,

$$(T(p^{r})f)z = \sum_{s=0}^{r} p^{2k(r-s)-r} \sum_{b=0}^{p^{s}-1} f\left(\frac{p^{r-s}z+b}{p^{s}}\right).$$

Moreover

$$(T(p)g)z = p^{2k-1}g(pz) + p^{-1}\sum_{b=0}^{p-1}g\left(\frac{z+b}{p}\right).$$

Thus

We recall Theorem 10.1. When f is a modular function, so that

$$f(z) = \sum_{m \in \mathbb{Z}} c(m) e(mz)$$

in a punctured disc (for q) centred at 0 we have

$$(T_n f)(z) = \sum_{m \in \mathbb{Z}} \gamma_n(m) e(mz)$$

with

$$\gamma_n(m) = \sum_{d \mid (m,n)} d^{2k-1} c(mnd^{-2}).$$

We can carry out several evalations. Thus

$$\gamma_n(0) = c(0)\sigma_{2k-1}(n), \tag{5}$$

$$\gamma_n(1) = c(n),\tag{6}$$

$$\gamma_p(m) = \begin{cases} c(mp) & \text{when } p \nmid m, \\ c(mp) + p^{2k-1}c(m/p) & \text{when } p|m. \end{cases}$$
(7)

10.5. Eigenfunctions. Let f be a modular form of weight 2k with k > 0 and not identically 0. We investigate the possibility that

$$T(n)f = \lambda(n)f$$
 for all $n \in \mathbb{N}$ where $\lambda(n) \in \mathbb{C}$. (8)

Definition 10.3. A modular form f of weight 2k with k > 0 and not identically 0 which satisfies (8) is called an eigenfunction or eigenform of T(n). The complex numbers $\lambda(n)$ are the eigenvalues of f.

Let f be such a modular form and c(n) be its Fourier coefficient. By (6), $c(n) = \lambda(n)c(1)$ for $n \in \mathbb{N}$. If c(1) = 0, then f would be identically c(0) which is impossible for k > 0. Hence to be an eigenfunction $c(1) \neq 0$. In view of this we can normalise f by dividing through by the constant c(1). Then c(1) = 1. In that case $\lambda(n) = c(n)$ for every $n \in \mathbb{N}$.

Theorem 10.9. Suppose that f is a normalised eigenform of weight 2k. Then the coefficients c(n) of its Fourier series are multiplicative for $n \ge 1$ and

$$c(p)c(p^{r}) = c(p^{r+1}) + p^{2k-1}c(p^{r-1}) \quad (r \in \mathbb{N}).$$

Proof. Suppose that $m, n \in \mathbb{N}$ and (m, n) = 1. By Theorem 10.7,

$$\begin{split} c(mn)f(z) &= (T(mn)f)(z) = (T(m)(T(n)f))(z) \\ &= (T_m(c(n)f))(z) = c(n)(T_mf)(z) = c(n)c(m)f(z). \end{split}$$

Moreover, by Theorem 10.8,

$$\begin{aligned} c(p)c(p^{r})f(z) &= c(p^{r})(T(p)f)(z) = (T(p)(c(p^{r})f))(z) = (T(p)(T(p^{r})f))(z) \\ &= (T(p^{r+1}) + p^{2k-1}T(p^{r-1}))f)(z) = (c(p^{r+1}) + p^{2k-1}c(p^{r-1}))f(z). \end{aligned}$$

We would like to consider the Dirichlet series whose coefficients are the coefficients of the Fourier expansion of an eigenform. First of all we need to establish sufficient information to ensure that the Dirichlet series converges.

Theorem 10.10. Suppose that f is a cusp form of weight 2k. Then its Fourier coefficients a_n satisfy

 $a_n \ll n^k$

where the implicit constant may depend on f.

Proof. An immediate consequence of f being analytic on $\widetilde{\mathbb{H}}$ is that the series

$$\sum_{n=1}^{\infty} a_n q^n$$

has radius of convergence at least 1. Thus $f(z)/q \ll 1$ uniformly for $|q| \leq \frac{1}{2}$, i.e. $f(z) \ll e^{-2\pi y}$ for $y \geq \frac{\log 2}{2\pi}$. Let $\phi(z) = |f(z)|y^k$. Then for $A \in \Gamma$ we have $\phi(z) = |cz + d|^{-2k}|F(Az)|(\Im z)^k = |cz + d|^{-2k}|f(Az)||cz + d|^{2k}(\Im Az)^k = f(Az)(\Im Az)^k$. Thus ϕ is invariant under Γ . Moreover $\frac{\log 2}{2\pi} < \frac{1}{2}$. Hence $\phi(z) \ll y^k e^{-2\pi y} \ll 1$ uniformly on \mathbb{D} and hence uniformly on \mathbb{H} . Thus $f(z) \ll y^{-k}$ uniformly on \mathbb{H} .

By Cauchy's integral formula,

$$a_n = \frac{1}{2\pi i} \int_{\mathcal{C}_y} \widetilde{f}(q) q^{-n-1} dq$$

where C_y is the circle parameterised by $q = e^{2\pi i x - 2\pi y}$ $0 \le x \le 1$. Then

$$a_n = \int_0^1 f(x+iy)e(-nx-iny)dx \ll y^{-k}e^{2\pi ny}$$

The choice y = 1/n establishes the desired conclusion.

Theorem 10.11. Suppose that f is a non-cusp modular form f of weight 2k > 0 and not identically 0. Then its Fourier coefficient a_n satisfies

$$a_n \asymp n^{2k-1} \quad (n > n_0(k, f)).$$

Again the implicit constants may depend on f.

The symbol \asymp is used to mean that the ratio of the two sides of the expression lies between two constants.

Proof. When k = 1 there are no modular forms of weight 2k. When $k \ge 2$ and dim $M_k = 1$ every modular form can be written as $\lambda G_k(z)$ where $\lambda \in \mathbb{C} \setminus \{0\}$, and the non-constant terms in the Fourier expansion of G_k are of the form $C_k \sigma_{2k-1}(n)$. Moreover for $k \ge 2$, $n^{2k-1} \le \sigma_{2k-1}(n) \le n^{2k-1} \zeta(2k-1)$ and the conclusion follows at once. When dim $M_k > 1$, $k \ge 6$ and every non-cusp modular form f can be written as $f(z) = \lambda G_k(z) + \mu g(z)$ where $\lambda \neq 0$ and g is a cusp form. The conclusion then follows from the observations above regarding the coefficients of G_k and Theorem 10.10.

Before proceeding further we display some eigenforms. First consider the modular form

$$E_k(z) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) e^{2\pi i n z}.$$

It is convenient to normalise this to give

$$E_k^* = c(0) + \sum_{n=1}^{\infty} \sigma_{2k-1}(n) e^{2\pi i n z}$$

where

$$c(0) = -\frac{B_{2k}}{4k}$$

It can be verified that

$$\sum_{j=0}^{\min(r,s)} p^{j(2k-1)} \sigma_{2k-1}(p^{r+s-2j}) = \sigma_{2k-1}(p^r) \sigma_{2k-1}(p^s)$$

and hence that when m > 0,

$$\sum_{d \mid (m,n)} d^{2k-1} \sigma_{2k-1}(mnd^{-2}) = \sigma_{2k-1}(m) \sigma_{2k-1}(n).$$

Thus, by (5),

$$(T_n E_k^*)(z) = \sigma_{2k-1}(n) E_k^*$$

Hence E_k^* is a normalised eigenform, and so λG_k is an eigenform for any non-zero complex number λ . By (5) any normalised non-cusp form of weight 2k has to satisfy $\lambda(n) = \sigma_{2k-1}(n)$ and then $c(n) = \sigma_{2k-1}(n)$. Thus there are no other non-cusp eigenforms.

By Theorem 10.6, if f is a cusp form of weight 2k, then so is $T_n f$. Suppose that dim $M_k^0 = 1$. 1. Then $T_n f = \lambda(n) f$ for some complex number $\lambda(n)$. Thus when dim $M_k^0 = 1$ every cusp form of weight 2k is an eigenform. By the remark after Definition 10.3 if f is normalised, then $\lambda(n) = c(n)$, the Fourier coefficient of f. Moreover, by Theorem 10.9 the coefficients are multiplicative and satisfy the recurrence relation of that theorem. The normalised cusp form $\Delta(z)/(2\pi)^{12}$ belongs to M_6^0 and dim $M_6^0 = 1$ and so is an eigenform. Hence Ramanujan's function $\tau(n)$ is multiplicative and satisfies

$$\tau(p)\tau(p^r) = \tau(p^{r+1}) + p^{2k-1}\tau(p^{r-1}).$$

By Theorem 10.10 it also satisfies

$$\tau(n) \ll n^6.$$

Theorem 10.12. Suppose that f is a normalised eigenform of weight 2k, and

$$f(z) = c(0) + \sum_{n=1}^{\infty} c(n)q^n$$

is its Fourier expansion at ∞ . Then the Dirichlet series

$$\Phi_f(s) = \sum_{n=1}^{\infty} c(n) n^{-s}$$

converges absolutely and locally uniformly for $\sigma > \sigma_k$, where $\sigma_k = k + 1$ if f is a cusp form and $\sigma_k = 2k$ otherwise. Moreover, when $\sigma > \sigma_k$,

$$\Phi_f(s) = \prod_p \left(1 - c(p)p^{-s} + p^{2k-1-2s} \right)^{-1}.$$

Proof. The convergence is immediate from the previous two theorems. By Theorem 10.9 the coefficients are multiplicative. Hence Φ_f has an Euler product valid for $\sigma > \sigma_k$,

$$\Phi_f(s) = \prod \left(1 + \sum_{m=1}^{\infty} c(p^m) p^{-ms} \right).$$

Moreover

$$(1 - c(p)p^{-s} + p^{2k-1-2s}) \left(1 + \sum_{m=1}^{\infty} c(p^m)p^{-ms} \right)$$

= $1 - c(p)p^{-s} + p^{2k-1-2s} + \sum_{m=1}^{\infty} \left(c(p^m)p^{-ms} - c(p)c(p^m)p^{-ms-s} + c(p^m)p^{2k-1-ms-2s} \right)$
= $1 - c(p)p^{-s} + p^{2k-1-2s} + \sum_{r=1}^{\infty} \left(c(p^{r+1})p^{-rs-s} - c(p)c(p^r)p^{-rs-s} + c(p^{r-1})p^{2k-1-rs-s} \right)$

and by the last part of Theorem 10.9 this is 1.