2.4 Hecke Operators

In this section we define Hecke operators on level 1 modular forms and derive their basic properties. Later in this book, we will not give proofs of the analogous properties for Hecke operators on high-level modular forms, since the proofs are clearest in the level 1 case, and the general case is similar (the proofs are available in other books, e.g. [Lan95]).

For any positive integer n, let

$$S_n = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in M_2(\mathbb{Z}) \;:\; a \geq 1,\; ad = n,\; \text{and } 0 \leq b < d \right\}.$$

Note that the set S_n is in bijection with the set of sublattices of $\mathbb{Z}^2$ of index n, where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ corresponds to $L = \mathbb{Z} \cdot (a, b) + \mathbb{Z} \cdot (0, d)$, as one can see, e.g., by using Hermite normal form (the analogue of reduced row echelon form over $\mathbb{Z}$).

Recall from (1.3.1) that if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Q})$, then

$$f|[\gamma]_k = \det(\gamma)^{k-1}(cz + d)^{-k} f(\gamma(z)).$$

Definition 2.4.1 (Hecke Operator $T_{n,k}$)**.** The nth Hecke operator $T_{n,k}$ of weight k is the operator on functions on $\mathfrak{h}$ defined by

$$T_{n,k}(f) = \sum_{\gamma \in S_n} f|[\gamma]_k.$$

Remark 2.4.2. It would make more sense to write $T_{n,k}$ on the right, e.g., $f|T_{n,k}$, since $T_{n,k}$ is defined using a right group action. However, if n, m are integers, then $T_{n,k}$ and $T_{m,k}$ commute (by Proposition 2.4.4 below), so it does not matter whether we consider the Hecke operators of given weight k as acting on the right or left.

Proposition 2.4.3. *If f is a weakly modular function of weight k, then so is $T_{n,k}(f)$; if f is also a modular function (i.e., is holomorphic on $\mathfrak{h}$), then so is $T_{n,k}(f)$.*

Proof. Suppose $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Since γ induces an automorphism of $\mathbb{Z}^2$, the set

$$S_n \cdot \gamma = \{\delta\gamma : \delta \in S_n\}$$

is also in bijection with the sublattices of $\mathbb{Z}^2$ of index n. For each element $\delta\gamma \in S_n \cdot \gamma$, there is $\sigma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\sigma\delta\gamma \in S_n$ (the element σ transforms $\delta\gamma$ to Hermite normal form), and the set of elements $\sigma\delta\gamma$ is thus equal to S_n. Thus

$$T_{n,k}(f) = \sum_{\sigma\delta\gamma \in S_n} f|[\sigma\delta\gamma]_k = \sum_{\delta \in S_n} f|[\delta\gamma]_k = T_{n,k}(f)|[\gamma]_k,$$

so $T_{n,k}(f)$ is weakly modular.

Since f is holomorphic on $\mathfrak{h}$, each $f|[\delta]_k$ is holomorphic on $\mathfrak{h}$ for $\delta \in S_n$. A finite sum of holomorphic functions is holomorphic, so $T_{n,k}(f)$ is holomorphic. □

We will frequently drop k from the notation in $T_{n,k}$, since the weight k is implicit in the modular function to which we apply the Hecke operator. Henceforth we make the convention that if we write $T_n(f)$ and f is modular, then we mean $T_{n,k}(f)$, where k is the weight of f.

Proposition 2.4.4. *On weight k modular functions we have*

$$T_{mn} = T_m T_n \qquad \textit{if } (m,n) = 1, \tag{2.4.1}$$

and

$$T_{p^n} = T_{p^{n-1}} T_p \; - \; p^{k-1} T_{p^{n-2}}, \qquad \textit{if } p \textit{ is prime.} \tag{2.4.2}$$

Proof. Let L be a lattice of index mn. The quotient $\mathbb{Z}^2/L$ is an abelian group of order mn, and $(m,n) = 1$, so $\mathbb{Z}^2/L$ decomposes uniquely as a direct sum of a subgroup of order m with a subgroup of order n. Thus there exists a unique lattice L' such that $L \subset L' \subset \mathbb{Z}^2$, and L' has index m in $\mathbb{Z}^2$. The lattice L' corresponds to an element of S_m, and the index n subgroup $L \subset L'$ corresponds to multiplying that element on the right by some uniquely determined element of S_n. We thus have

$$\mathrm{SL}_2(\mathbb{Z}) \cdot S_m \cdot S_n = \mathrm{SL}_2(\mathbb{Z}) \cdot S_{mn},$$

i.e., the set products of elements in S_m with elements of S_n equal the elements of S_{mn}, up to $\mathrm{SL}_2(\mathbb{Z})$-equivalence. It then follows from the definitions that for any f, we have $T_{mn}(f) = T_n(T_m(f))$. Applying this formula with m and n swapped yields the equality $T_{mn} = T_m T_n$.

We will show that $T_{p^n} + p^{k-1} T_{p^{n-2}} = T_p T_{p^{n-1}}$. Suppose f is a weight k weakly modular function. Using that $f|[p]_k = (p^2)^{k-1} p^{-k} f = p^{k-2} f$, we have

$$\sum_{x \in S_{p^n}} f|[x]_k \; + \; p^{k-1} \sum_{x \in S_{p^{n-2}}} f|[x]_k = \sum_{x \in S_{p^n}} f|[x]_k \; + p \sum_{x \in pS_{p^{n-2}}} f|[x]_k.$$

Also

$$T_p T_{p^{n-1}}(f) = \sum_{y \in S_p} \sum_{x \in S_{p^{n-1}}} f|[x]_k|[y]_k = \sum_{x \in S_{p^{n-1}} \cdot S_p} f|[x]_k.$$

Thus it suffices to show that S_{p^n} disjoint union p copies of $pS_{p^{n-2}}$ is equal to $S_{p^{n-1}} \cdot S_p$, where we consider elements with multiplicities and up to left $\mathrm{SL}_2(\mathbb{Z})$-equivalence (i.e., the left action of $\mathrm{SL}_2(\mathbb{Z})$).

Suppose L is a sublattice of $\mathbb{Z}^2$ of index p^n, so L corresponds to an element of S_{p^n}. First suppose L is not contained in $p\mathbb{Z}^2$. Then the image of L in $\mathbb{Z}^2/p\mathbb{Z}^2 = (\mathbb{Z}/p\mathbb{Z})^2$ is of order p, so if $L' = p\mathbb{Z}^2 + L$, then $[\mathbb{Z}^2 : L'] = p$ and $[L : L'] = p^{n-1}$, and L' is the only lattice with this property. Second suppose that $L \subset p\mathbb{Z}^2$ if of index p^n, and that $x \in S_{p^n}$ corresponds to L. Then every one of the $p+1$ lattices $L' \subset \mathbb{Z}^2$ of index p contains L. Thus there are $p+1$ chains $L \subset L' \subset \mathbb{Z}^2$ with $[\mathbb{Z}^2 : L'] = p$.

The chains $L \subset L' \subset \mathbb{Z}^2$ with $[\mathbb{Z}^2 : L'] = p$ and $[\mathbb{Z}^2 : L] = p^{n-1}$ are in bijection with the elements of $S_{p^{n-1}} \cdot S_p$. On the other hand the union of S_{p^n}

with p copies of $pS_{p^{n-2}}$ corresponds to the lattices L of index p^n, but with those that contain $p\mathbb{Z}^2$ counted $p+1$ times. The structure of the set of chains $L \subset L' \subset \mathbb{Z}^2$ that we derived in the previous paragraph gives the result. □

Corollary 2.4.5. *The Hecke operator T_{p^n}, for prime p, is a polynomial in T_p. If n, m are any integers then $T_n T_m = T_m T_n$.*

Proof. The first statement is clear from (2.4.2), and this gives commutativity when m and n are both powers of p. Combining this with (2.4.1) gives the second statement in general. □

Proposition 2.4.6. *Suppose $f = \sum_{n\in\mathbb{Z}} a_n q^n$ is a modular function of weight k. Then*

$$T_n(f) = \sum_{m\in\mathbb{Z}} \left(\sum_{1\leq d \,|\, \gcd(n,m)} d^{k-1} a_{mn/d^2} \right) q^m.$$

In particular, if $n = p$ is prime, then

$$T_p(f) = \sum_{m\in\mathbb{Z}} \left(a_{mp} + p^{k-1} a_{m/p} \right) q^m,$$

where $a_{m/p} = 0$ if $m/p \notin \mathbb{Z}$.

The proposition is not that difficult to prove (or at least the proof is easy to follow), and is proved in [Ser73, §VII.5.3] by writing out $T_n(f)$ explicitly and using that $\sum_{0\leq b<d} e^{2\pi i b m/d}$ is d if $d \mid m$ and 0 otherwise. A corollary of Proposition 2.4.6 is that T_n preserves M_k and S_k.

Corollary 2.4.7. *The Hecke operators preserve M_k and S_k.*

Remark 2.4.8. Alternatively, for M_k this is Proposition 2.4.3, and for S_k we see from the definitions that if $f(i\infty) = 0$ then $T_n f$ also vanishes at $i\infty$.

Example 2.4.9. Recall that

$$E_4 = \frac{1}{240} + q + 9q^2 + 28q^3 + 73q^4 + 126q^5 + 252q^6 + 344q^7 + \cdots.$$

Using the formula of Proposition 2.4.6, we see that

$$T_2(E_4) = (1/240 + 2^3 \cdot (1/240)) + 9q + (73 + 2^3 \cdot 1)q^2 + \cdots.$$

Since M_k has dimension 1, and we have proved that T_2 preserves M_k, we know that T_2 acts as a scalar. Thus we know just from the constant coefficient of $T_2(E_4)$ that

$$T_2(E_4) = 9E_4.$$

More generally, for p prime we see by inspection of the constant coefficient of $T_p(E_4)$ that

$$T_p(E_4) = (1 + p^3)E_4.$$

In fact for any k one has that

$$T_n(E_k) = \sigma_{k-1}(n)E_k,$$

for any integer $n \geq 1$ and even weight $k \geq 4$.

Example 2.4.10. By Corollary 2.4.7, the Hecke operators T_n also preserve the subspace S_k of M_k. Since S_{12} has dimension 1 (spanned by Δ), we see that Δ is an eigenvector for every T_n. Since the coefficient of q in the q-expansion of Δ is 1, the eigenvalue of T_n on Δ is the nth coefficient of Δ. Since $T_{nm} = T_n T_m$ for $(n, m) = 1$ we have proved the non-obvious fact that the function $\tau(n)$ that gives the nth coefficient of Δ is a multiplicative function.

Remark 2.4.11. The Hecke operators respect $M_k = S_k \oplus \mathbb{C}E_k$, i.e., for all k the series E_k are eigenvectors for all T_n, and because (in this book) we normalize E_k so that the coefficient of q is 1, the eigenvalue of T_n on E_k is the coefficient $\sigma_{k-1}(n)$ of q^n in the q-expansion of E_k.

2.5 Computing Hecke Operators

In this section we describe an algorithm for computing matrices of Hecke operators on M_k.

Algorithm 2.5.1 (Hecke Operator)**.** *This algorithm computes a matrix for the Hecke operator T_n on the Victor Miller basis for M_k.*

1. [Compute dimension] Compute $d = \dim(M_k) - 1$ using Corollary 2.2.6.
2. [Compute basis] Using the algorithm implicit in Lemma 2.3.1, compute the reduced row echelon basis $f_0, \ldots, f_d$ for M_k modulo q^{dn+1}.
3. [Compute Hecke operator] Using Proposition 2.4.6, compute for each i the image $T_n(f_i) \pmod{q^{d+1}}$.
4. [Write in terms of basis] The elements $T_n(f_i) \pmod{q^{d+1}}$ uniquely determine linear combinations of $f_0, f_1, \ldots, f_d \pmod{q^d}$. These linear combinations are easy to find once we compute $T_n(f_i) \pmod{q^{d+1}}$, since our basis of f_i is in reduced row echelon form. The linear combinations are just the coefficients of the power series $T_n(f_i)$ up to and including q^d.
5. [Write down matrix] The matrix of T_n acting from the right relative to the basis $f_0, \ldots, f_d$ is the matrix whose rows are the linear combinations found in the previous step, i.e., whose rows are the coefficients of $T_n(f_i)$.

Proof. First note that we need only compute a modular form f modulo q^{dn+1} in order to compute $T_n(f)$ modulo q^{d+1}. This follows from Proposition 2.4.6, since in the formula the dth coefficient of $T_n(f)$ involves only a_{dn}, and smaller-indexed coefficients of f. Uniqueness in Step 4 follows from Lemma 2.3.1 above. □

Example 2.5.2. We compute in detail the Hecke operator T_2 on M_{12} using the above algorithm.

1. [Compute dimension] We have $d = 2 - 1 = 1$.

2. [Compute basis] We compute up to (but not including) the coefficient of $q^{dn+1} = q^{1\cdot 2+1} = q^3$. As given explicitly in the proof of Lemma 2.3.1, we have

$$F_4 = 1 + 240q + 2160q^2 + \cdots \quad \text{and} \quad F_6 = 1 - 504q - 16632q^2 + \cdots .$$

   Thus M_{12} has basis

$$F_4^3 = 1+720q+179280q^2+\cdots \quad \text{and} \quad \Delta = (F_4^3 - F_6^2)/1728 = q - 24q^2 + \cdots$$

   Subtracting 720Δ from F_4^3 yields the echelon basis, which is

$$f_0 = 1 + 196560q^2 + \cdots \quad \text{and} \quad f_1 = q - 24q^2 + \cdots .$$

   SAGE can do the arithmetic involved in the above calculation as follows:

```
sage: R = QQ[['q']]      # power series ring
sage: q = R.0            # generator of the power series ring
sage: F4 = 1 + 240*q + 2160*q^2 + O(q^3)
sage: F6 = 1 - 504*q - 16632*q^2 + O(q^3)
sage: F4^3
1 + 720*q + 179280*q^2 + O(q^3)
sage: Delta = (F4^3 - F6^2)/1728; Delta
q - 24*q^2 + O(q^3)
sage: F4^3 - 720*Delta
1 + 196560*q^2 + O(q^3)
```

3. [Compute Hecke operator] In each case letting a_n denote thoe nth coefficient of f_0 or f_1, respectively, we have

$$\begin{aligned} T_2(f_0) &= T_2(1 + 196560q^2 + \cdots) \\ &= (a_0 + 2^{11}a_0)q^0 + (a_2 + 2^{11}a_{1/2})q^1 + \cdots \\ &= 2049 + 196560q + \cdots \end{aligned}$$

   and

$$\begin{aligned} T_2(f_1) &= T_2(q - 24q^2 + \cdots) \\ &= (a_0 + 2^{11}a_0)q^0 + (a_2 + 2^{11}a_{1/2})q^1 + \cdots \\ &= 0 - 24q + \cdots \end{aligned}$$

4. [Write in terms of basis] We read off at once that

$$T_2(f_0) = 2049f_0 + 196560f_1 \quad \text{and} \quad T_2(f_1) = 0f_0 + (-24)f_1$$

5. [Write down matrix] Thus the matrix of T_2, acting from the right on the basis f_0, f_1, is
$$T_2 = \begin{pmatrix} 2049 & 196560 \\ 0 & -24 \end{pmatrix}.$$

As a consistency check note that the characteristic polynomial of the computed T_2 is $(x-2049)(x+24)$, and that $2049 = 1 + 2^{11}$ is the sum of the 11th powers of the divisors of 2.

Example 2.5.3. The Hecke operator T_2 on M_{36} with respect to the echelon basis is:

$$\begin{pmatrix} 34359738369 & 0 & 6218175600 & 9026867482214400 \\ 0 & 0 & 34416831456 & 5681332472832 \\ 0 & 1 & 194184 & -197264484 \\ 0 & 0 & -72 & -54528 \end{pmatrix}$$

It has characteristic polynomial

$$(x - 34359738369) \cdot (x^3 - 139656x^2 - 59208339456x - 1467625047588864),$$

where the cubic factor is irreducible.

Using the SAGE modular forms functions [[**TODO: warning – the ones used below do not exist yet!!!**]] we compute the above as follows:

```
sage: M = ModularForms(1,36)
sage: M.basis()
...  victor miller basis ...
sage: t = M.T(2).matrix(); t
... above matrix on vm basis ...
age: f = t.charpoly(); f.factor()
... factored form ...
```

The following is a famous and simple to state open problem about Hecke operators on modular forms of level 1. It generalizes our above observation that the characteristic polynomial of T_2 on M_k, for $k = 12, 36$, factors as a product of a linear factor and an irreducible factor.

Conjecture 2.5.4 (Maeda)**.** *The characteristic polynomial of T_2 on S_k is irreducible for any k.*

Kevin Buzzard observed that in many specific cases the Galois group of the characteristic polynomial of T_2 is the full symmetric group (see [Buz96]). See also [FJ02] for more evidence for Maeda's conjecture and connections to other problems of interest. [[**Todo: Isn't there something from a recent Berkeley grad student?**]]

2.5.1 Complexity of Computing Fourier Coefficients

Just how difficult is it to compute prime-indexed coefficients of the q-expansion

$$\begin{aligned}\Delta &= \sum_{n=1}^{\infty} \tau(n) q^n \\ &= q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 - 6048q^6 - 16744q^7 \\ &\quad + 84480q^8 - 113643q^9 - 115920q^{10} + 534612q^{11} - \\ &\quad 370944q^{12} - 577738q^{13} + 401856q^{14} + 1217160q^{15} + \\ &\quad 987136q^{16} - 6905934q^{17} + 2727432q^{18} + 10661420q^{19} + \cdots\end{aligned}$$

of the Δ-function?

Theorem 2.5.5 (Edixhoven et al.)**.** *Let p be a prime. There is an algorithm to compute $\tau(p)$, for prime p, that is polynomial-time in* $\log(p)$. *More generally, if $f = \sum a_n q^n$ is a Hecke eigenform in some space $M_k(\Gamma_1(N))$, where $k \geq 2$, then there is an algorithm to compute a_p in time polynomial in* $\log(p)$.

Bas Edixhoven, Jean-Marc Couveignes and Robin de Jong have proved that $\tau(p)$ can be computed in polynomial time; their approach involves sophisticated techniques from arithmetic geometry (e.g., étale cohomology, motives, Arakelov theory). *This is work in progress and has not been written up in detail yet.* The ideas they use are inspired by the ones introduced by Schoof, Elkies and Atkin for quickly counting points on elliptic curves over finite fields (see [Sch95]).

Edixhoven describes the strategy as follows:

1. We compute the mod ℓ Galois representation ρ associated to Δ. In particular, we produce a polynomial f such that $\mathbb{Q}[x]/(f)$ is the fixed field of $\ker(\rho)$. This is then used to obtain $\tau(p) \pmod{\ell}$ and do a Schoof-like algorithm for computing $\tau(p)$.

2. We compute the field of definition of suitable points of order ℓ on the modular Jacobian $J_1(\ell)$ to do part 1. (This modular Jacobian is the Jacobian of a model of $\Gamma_1(\ell)\backslash\mathfrak{h}^*$ over $\mathbb{Q}$.)

3. The method is to approximate the polynomial f in some sense (e.g., over the complex numbers, or modulo many small primes r), and use an estimate from Arakelov theory to determine a precision that will suffice.