ON THE DISCRIMINANT OF THE HECKE RING, $\mathbb{T}_k$, AND ITS INDEX IN THE RING OF INTEGERS OF $\mathbb{T}_k \otimes \mathbb{Q}$

1. Introduction

In the second section, we explore preliminary material associated to modular forms: we build a sufficient background by defining and examining crucial topics to the study of modular forms, like the modular group and its action on the upper half plane $\mathcal{H}$. We also explore the vector space of modular forms of weight k and its properties. In our exploration, we define define Eisenstein series, the discriminant function, as well as the modular invariant. Using the previously Eisenstien series, we find a basis for the space of modular forms of weight k and hence the dimension of such spaces (which turn out to depend only on the weight k). Following this, we gain knowledge about the space of all modular forms; in particular, that it's a graded ring.

In the third section, we discuss congruence subgroups and the weight k operators associated to them. We go on to define double cosets and the operators associated to them, allowing us to work with modular forms with respect to a congruence subgroup (compared to working with respect to the full modular group). This idea leads us directly into our fourth section on Hecke operators.

The fourth section is quite short and focuses directly on the Hecke operator, T_p, for a prime p. In particular, we describe its effect on a modular form through its effect of the form's q-expansion.

The fifth and final section of this paper works with modular forms mod ℓ for some prime ℓ. We define the space of modular forms mod ℓ, and discover that the concept of "weight" is no longer well defined mod ℓ, and we move to a new concept called filtration. We use the powerful concept of filtration to generate fruitful theory not only about the local components of the ring $R_k \otimes \overline{\mathbb{F}_\ell}$, but also the index of the ring of integers in it, where R_k denotes the Hecke ring mod ℓ.

We briefly summarize the notation used in this paper. When referring to groups and subgroups, the symbol $\trianglelefteq$ denotes a normal subgroup; that is, $H \trianglelefteq G$ means H is a normal subgroup of G. When we write $\mathbb{Z}/n\mathbb{Z}$ or $\mathbb{Z}_n$, we mean the same: the factor group of $\mathbb{Z}$ by the normal subgroup $n\mathbb{Z}$. $ker(f)$ denotes the kernel of the function f. I_2 denotes the 2-by-2 identity matrix. The set of units in $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n$ is represented by $(\mathbb{Z}_n)^*$. The third root of unity $e^{2\pi i/3}$ is denoted by ζ_3.

2. Preliminaries

Definition 2.1. *A function f: $D \subseteq \mathbb{C} \to \mathbb{C}$ is called holomorphic if f is analytic at every point p of D; equivalently, f is holomorphic if locally,*

$$f(x) = \sum_{n=0}^{\infty} a_n (z-p)^n, \quad a_i \in \mathbb{C}. \tag{2.1}$$

Date: May 6, 2020.

Example 2.2. *Some trivial examples of holomorphic functions on $\mathbb{C}$ are polynomials, $sin(z), cos(z)$, and e^z; one can verify these are holomorphic functions using the Cauchy-Riemann equations.*

Definition 2.3. *A function f: $D \subseteq \mathbb{C} \to \mathbb{C}$ is called meromorphic if it is holomorphic except on a discrete set of points, called poles; equivalently, a function is called meromorphic if locally, for every point $p \in D \setminus S$,*

$$f(z) = \sum_{n \geqslant k}^{\infty} a_n (z-p)^n, \quad a_i \in \mathbb{C}, \text{ for some } k \in \mathbb{Z}. \tag{2.2}$$

Remark 2.4. Clearly, if a function is holomorphic, it is meromorphic by taking $k = 0$. The converse is not always true; it is true if $f(z)$ has no poles.

Example 2.5. *Since all holomorphic functions are meromorphic, trivial examples of meromorphic functions are those in Example (2.2). An example of a meromorphic but not holomorpic function is $f(z) = 1/z$; we see this is meromorphic on $\mathbb{C}$, but not holomorphic on $\mathbb{C}$ since $f(z)$ is not analytic at 0.*

Definition 2.6. *Define $\mathcal{H} = \{z \in \mathbb{C} | Im(z) > 0\}$ as the upper half complex plane.*

Definition 2.7. *The modular group G is defined as the group of 2×2 matrices with entries in the integers and determinant 1, under matrix multiplication:*

$$G = \mathrm{PSL}_2(\mathbb{Z}) \cong \mathrm{SL}_2(\mathbb{Z}) \,/\, \{\pm 1\} = \{A \in M_{2x2}(\mathbb{Z}) \,|\, det(A) = 1\} \,/\, \{\pm 1\} \tag{2.3}$$

Remark 2.8. After defining the action of G on $\mathcal{H}$, it will become clear why we mod out by $\{\pm 1\}$ in the above definition: for any $A \in \mathrm{SL}_2(\mathbb{Z})$, the transformation Az is the same transformation as $-Az$. That is, the two matrices $-A, A \in \mathrm{SL}_2(\mathbb{Z})$ have the same effect on a fixed $z \in \mathcal{H}$.

Proposition 2.9. *The modular group G is generated by $S := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $T := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.*

Proposition 2.10. *The modular group G acts on $\tilde{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ by if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G$, then*

$$Az = \frac{az+b}{cz+d}, \quad \forall z \in \mathbb{C}. \tag{2.4}$$

Note that as $det(A) = 1$, c, d cannot simultaneously be zero, $Az \in \mathbb{C}, \forall z \in \mathbb{C}, \forall A \in G$.

Proof. Clearly, $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z = z, \forall z \in \tilde{\mathbb{C}}$. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$. Then the product $AB = \begin{bmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{bmatrix}$. Finally,

$$\begin{aligned} A(Bz) &= A\left(\frac{a'z+b'}{c'z+d'}\right) \\ &= \frac{a\left(\frac{a'z+b'}{c'z+d'}\right)+b}{c\left(\frac{a'z+b'}{c'z+d'}\right)+d} \end{aligned}$$

$$
\begin{aligned}
&= \frac{a(a'z+d')+b(c'z+d')}{c(a'z+b')+d(c'z+d')} \\
&= \frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')} \\
&= (AB)z.
\end{aligned}
$$

□

Proposition 2.11. [7, Theorem 1] *$D = \{z \in \mathcal{H} \mid Re(z) \leqslant 1/2 \text{ and } |z| \geqslant 1\} \subseteq \mathcal{H}$ is the fundamental domain for the action of G on $\mathcal{H}$.*

Proof. See the proof of Theorem 1 in [7] □

Definition 2.12. *Let $k \in \mathbb{Z}$ and $f : \mathcal{H} \to \mathbb{C}$ a meromorphic function. f is called weakly modular of weight k if f satisfies*

$$f(\begin{bmatrix} a & b \\ c & d \end{bmatrix} z) = (cz+d)^k f(z), \quad \forall \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}), \forall z \in \mathcal{H}. \tag{2.5}$$

An equivalent characterization of weakly modular functions is, given a meromorphic function f on $\mathcal{H}$, f is weakly modular of weight k if and only if the following two conditions are satisfied:

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

for all $z \in \mathcal{H}$.

Proof of equivalent definitions. Assuming $f(\begin{bmatrix} a & b \\ c & d \end{bmatrix} z) = (cz+d)^k f(z)$, $\forall \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, then clearly (1) and (2) are satisfied since $S, T \in \mathrm{SL}_2(\mathbb{Z})$.

Conversely, suppose (1) and (2) are satisfied. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. It suffices to show that if $f(Az) = (cz+d)^k f(z)$ and $f(Bz) = (c'z+d')^k f(z)$, then $f\big((AB)z\big) = \big((ca'+dc')z + (cb'+dd')\big)^k f(z)$, which implies if f is weakly modular with respect to A and B, then it is weakly modular with respect to the group generated by them. Under assumption, then, if (1) and (2) are satisfied, f is weakly modular for the group generated by them and by Proposition 2.9, it is weakly modular with respect to $\mathrm{SL}_2(\mathbb{Z})$.

So, assume f is weakly modular with respect to A and B as above. Then,

$$
\begin{aligned}
f\big((AB)z\big) &= f\big(A(Bz)\big) \\
&= (c(Bz)+d)^k f(Bz) \\
&= (c\left(\frac{a'z+b'}{c'z+d'}\right)+d)^k (c'z+d')^k f(z) \\
&= \big((ca'+dc')z + (cb'+dd')\big)^k f(z)
\end{aligned}
$$

which is precisely what we wanted to show. (This direction's proof was based on the proof of (a) in Lemma 1.2.2 in [2]). □

Example 2.13. *Examples of weakly modular functions include constant functions and Eisenstein series, which are defined in Definition* 2.18.

Corollary 2.14. *The only weakly modular function of odd weight is the zero function.*

Proof. Let $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ and $f(z)$ a weakly modular function of odd weight k. Using Definition 2.12 and that $Az = \dfrac{-z+0}{0-1} = z$,

$$\begin{aligned} f(z) &= f(Az) \\ &= f(\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} z) \\ &= (-1)^k f(z). \end{aligned}$$

Clearly, if k is odd, then $f(z) = 0$. □

Let f be a weakly modular function. By (2) in the equivalent characterization of weakly modular functions in Definition 2.12, $f(z+1) = f(z), \forall z \in \mathcal{H}$. Because of this, f is equal to some function $g(q)$ where $q = e^{2\pi iz}$ and if f is holomorphic, then $g(q)$ is holomorphic on the unit disk minus the origin. Using the equality $|q| = e^{2\pi Im(z)}$, we see that $q \to 0$ if and only if $Im(z) \to \infty$ (the previous paragraph is due to [2, pg. 3]).

Thus, when f extends meromorphically (holomorphically) function at the origin, we say it is meromorphic (holomorphic) at infinity. By "extends meromorphically (holomorphically) at the origin," we mean if there exists some meromorphic (holomorphic) funtion h on the unit disk such that $h(z) = g(q)$ on the unit disk minus the origin.

Definition 2.15. *Let $k \in \mathbb{Z}$ and f a weakly modular function. f is called modular if f is holomorphic on $\mathcal{H}$ and at infinity, where we consider infinity to lie far in the imaginary direction.*

With this, one can characterize a modular form of weight k as a series

$$f(z) = \sum_{n=0}^{\infty} a_n (z-p)^n, \quad a_i \in \mathbb{C} \tag{2.6}$$

for all $p \in \mathcal{H}$, and supposing the second condition from the equivalent characterization of weakly modular in Definition 2.12 is satisfied, one can write $f(z)$ as a function of $q = e^{2\pi iz}$. Thus, a modular form of weight k is given by

$$f(z) = \sum_{n=0}^{\infty} a_n q^n = \sum_{n=0}^{\infty} a_n e^{2\pi niz}, \tag{2.7}$$

which converges absolutely for $|q| < 1$.

Definition 2.16. [2, Definition 1.1.3] *A modular form is called a cusp form if $a_0 = 0$ in its q-expansion; equivalently, a modular form is a cusp form if* $\lim_{\mathrm{Im}(z)\to\infty} f(z) = 0$.

It's well known that the space of modular forms of weight k and the space of cusp forms of weight k over the full modular group (commonly denoted $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ respectively) are vector spaces over $\mathbb{C}$, and that $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ is a subspace of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. One could also characterize the space of cusp forms of weight k as the kernel of the map $\phi : \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \to \mathbb{C}$ by $\phi : \sum_{n=0}^{\infty} a_n q^n \mapsto a_0$.

Remark 2.17. As in [2, pg. 4], one typically denotes the space of modular forms

$$\mathcal{M}(\mathrm{SL}_2(\mathbb{Z})) = \bigoplus_{k\in\mathbb{Z}} \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})),$$

which is a graded ring (the product of two modular forms of weight k and weight k' modular forms is a form of weight $k + k'$).

In addition, the space of cusp forms

$$\mathcal{S}(\mathrm{SL}_2(\mathbb{Z})) = \bigoplus_{k=0}^{\infty} \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$$

forms a graded ideal in $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$ ([2, pg. 6]).

Definition 2.18. *Let $k > 2$. The function*

$$G_k^*(z) = \sum_{(c,d)\in\mathbb{Z}^2\setminus(0,0)} \frac{1}{(cz+d)^k} \tag{2.8}$$

is called the Eisenstein series of weight k, where $G_k^*(\infty) = 2\zeta(k)$, where ζ denotes the Riemann zeta function given by $\zeta(k) = \sum_{d=1}^{\infty} 1/d^k$ (see [7, Proposition 4]*).*

Fact 2.19. [2, pg. 5] The Eisenstein series of weight k for all $k \in \mathbb{Z}_{\geqslant 3}$ is a modular form of weight k, and if one writes it in its q-expansion,

$$G_k^*(z) = \sum_{n=0}^{\infty} a_n q^n,$$

with $q = e^{2\pi i z}$, then $G_k^*(0) = 2\zeta(k)$ where ζ is the Riemann zeta function.

The Eisenstein series, G_k^*, is commonly normalized in two different ways: the first normalizes the constant term and the second normalizes the coefficient of q in the q-expansion for G_k^*. The former will is denoted E_k and the latter is denoted G_k.

The normalized Eisenstein series of weight k, G_k, can be expressed in the following way:

$$G_k(z) = \frac{1}{2}\zeta(1-k) + \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n \tag{2.9}$$

where $\sigma_{k-1}(n) = \sum_{\substack{m|n \\ m>0}} m^{k-1}$.

Moreover, the other normalized Eisenstein series of weight k, E_k, can be expressed in the following way:

$$E_k(z) = 1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n, \tag{2.10}$$

where $\sigma_{k-1}(n) = \sum_{\substack{m|n \\ m>0}} m^{k-1}$ and B_k represents the k^{th} Bernoulli number. This expression for E_k uses the identity $\frac{-2k}{B_k} = \frac{2}{\zeta(1-k)}$.

Defining the forms $f_1(z) = 60G_4(z)$ and $f_2(z) = 140G_6(z)$, one arrives at what is commonly known as the discriminant function, $\Delta^* : \mathcal{H} \to \mathbb{C}$, given by $\Delta^*(z) = (f_1(z))^3 - 27(f_2(z))^2$ ([2, pg. 6]) which is a modular form of weight 12 (as $(f_1(z))^3$ and $f_2(z))^2$ are forms of weight 12). It is easy to check that the first term in the q-expansion for Δ^* is zero, and so by Definition 2.16, we conclude that $\Delta^*(z)$ is a cusp form. We verify in the proof of Theorem 2.21 that Δ^* is not the zero function and is zero nowhere except at infinity.

It is often useful normalize the coefficient of q in the q-expansion of Δ^*, which we will denote as Δ and is described in the following way: $\Delta(z) = (1/1728)(E_4^3 - E_6^2)$. Since E_4 and E_6 have only rational coefficients, it follows that Δ does as well.

Defining Δ^* allows one to develop another common modular function, $j : \mathcal{H} \to \mathbb{C}$ given by $j(z) = 1728\frac{(f_1(z))^3}{\Delta^*(z)}$. The j function is known as the modular invariant since $j(Az) = j(z), \forall A \in \mathrm{SL}_2(\mathbb{Z})$ ([[2]]). Since the only zero of Δ^* is at infinity, one observes that j has a simple pole at infinity (which shows why it is not a modular form).

Lemma 2.20 ([7, Theorem 3])**.** *Some notation from the theorem in* [7] *is used. Let $p \in \mathcal{H}$, let f be a modular form, and let G denote the full modular group. Let $ord_p(f)$ be the integer s for which $f/(z-p)^s$ is nonzero. If f is a nonzero modular form of weight k, then the following formula is satisfied:*

$$ord_\infty(f) + \frac{1}{2}ord_i(f) + \frac{1}{3}ord_{\zeta_3}(f) + \sum_{p \in \mathcal{H}/G}^{*} ord_p(f) = k/12 \tag{2.11}$$

where $\zeta_3 = e^{2\pi i/3}$ is a third root of unity, and where $\sum^*$ means to take p not in the equivalence classes of neither i nor ζ_3.

Note that as f is a modular form, it has no poles. In particular, $ord_\infty(f), ord_p(f) \geqslant 0, \forall p \in \mathcal{H}/G$.

Theorem 2.21. [7, Theorem 4]
(1) If $k < 0$ or positive and odd, or if $k = 2$, then $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) = \{0\}$.
(2) Multiplication by Δ^* gives an isomorphism between $\mathcal{M}_{k-12}(\mathrm{SL}_2(\mathbb{Z}))$ and $S_k(\mathrm{SL}_2(\mathbb{Z}))$.

Proof. The following proof is due to [7], and we will be using some of their notation. Let f be a modular form of weight k, and again let G denote the full modular group. Since the left hand side of Formula 2.11 is nonnegative for modular forms, k must be nonnegative and hence the only modular forms of negative weight are the zero function.

If k is positive and odd, Corollary 2.14 showed that the only forms satisfying this are also the zero function.

If $k = 2$, then the right hand side of Formula 2.11 equals $1/6$. Multiplying each side by 6 gives:

$$6ord_\infty(f) + 3ord_i(f) + 2ord_{\zeta_3}(f) + 6\sum_{p\in\mathcal{H}/G}^{*} ord_p(f) = 1.$$

But $ord_\infty(f), ord_p(f) \in \mathbb{Z}_{\geqslant 0}, \forall p \in \mathcal{H}/G$, thus giving us a sum of nonnegative integers equal to 1, which is impossible. Thus, any modular form of weight 2 is the zero function. This proves (1).

For the sake of brevity, let $a = ord_\infty(f), b = ord_i(f)$, and $c = ord_{\zeta_3}(f)$. Recall that the discriminant function $\Delta^* = (60G_4(z))^3 - 27(140G_6(z))^2$. Since G_4 and G_6 are modular forms, they satisfy Formula 2.11. Moreover, letting $k = 4$ or 6 makes the right hand side of Formula 2.11 an element of $\mathbb{Q} \setminus \mathbb{Z}$ and so $\sum^{*}_{p\in\mathcal{H}/G} ord_p(f)$ must be zero. Applying the Formula to G_4 and multiplying through by 6, we have

$$6a + 3b + 2c = 2.$$

Clearly, the only solution is $(a, b, c) = (0, 0, 1)$.

Similarly, applying the formula to G_6 and multiplying through by 6, we have

$$6a + 3b + 2c = 3.$$

Clearly, the only solution is $(a, b, c) = (0, 1, 0)$. Together, this tells us G_4 has one zero at ζ_3 and G_6 has one zero at i, and so Δ^* cannot be the zero function because it is not zero at i. We've already seen that Δ^* is a cusp form of weight 12, and so applying Formula 2.11 to Δ^* gives

$$1 + b + c + \sum_{p\in\mathcal{H}/G}^{*} ord_p(f) = 1$$

implies that $b = c = \sum^{*}_{p\in\mathcal{H}/G} ord_p(f) = 0$ and proves that Δ^* is nonzero on $\mathcal{H}$ except at infinity (in fact, it proves it has a simple zero at infinity).

Let $h \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $g = h/\Delta^*$. Since h is a cusp form, it has a zero at infinity. Since Δ^* has a simple zero at infinity and nowhere else, g is holomorphic on $\mathcal{H}$ and at infinity. Clearly, g has weight $k - 12$. Thus, $g \in \mathcal{M}_{k-12}(\mathrm{SL}_2(\mathbb{Z}))$ and (2) is proven. □

Corollary 2.22. *If $k = 0, 4, 6, 8, 10$, then the space of modular forms of weight k is one dimensional with generators $1, G_4, G_6, G_4^2$, and G_4G_6 respectively.*

Proof. Recall that $S_k(\mathrm{SL}_2(\mathbb{Z})) = ker(\phi)$ where the linear map $\phi : \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \to \mathbb{C}$ is given by

$$\phi : \sum_{n=0}^{\infty} a_n q^n \mapsto a_0.$$

By the Rank-Nullity Theorem, $dim(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) = rank(\phi) + dim(\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))) \leqslant 1 + dim(\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})))$.

By Theorem 2.21, $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})) \cong \mathcal{M}_{k-12}(\mathrm{SL}_2(\mathbb{Z}))$. By Theorem 2.21 again, if $k = 0, 4, 6, 8$, or 10, $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z})) = \{0\}$, implying that $dim(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) \leqslant 1$. Then as $1, G_4, G_6, G_4^2$, and G_4G_6 are nonzero forms of weights $0, 4, 6, 8$, and 10

respectively shows that $dim(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) = 1$ and such nonzero forms generate (and in fact are a basis for) their respective space of modular forms. □

Theorem 2.23. [7, Corollary 2] *The space of modular forms of weight k has as a basis the polynomials $\{G_4^a G_6^b\}$ where $4a + 6b = k$ and $a, b \in \mathbb{Z}_{\geqslant 0}$.*

Proof. This proof is due to [7], and we will keep some of its notation. We first show that G_4, G_6 generate $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for all k. For (even) $k < 8$, G_4 and G_6 generate $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ by Corollary 2.22.

For (even) $k \geqslant 8$, the proof is by way of induction. We begin by noting that since 2 and 3 are relatively prime, some linear combination of them equals 1. In particular, for any positive even integer k, it equals some linear combination of 4 and 6 with coefficients in $\mathbb{Z}_{\geqslant 0}$. Since we are inducting on k for $k \geqslant 8$, the base case is when $k = 8$. By Corollary 2.22, $\mathcal{M}_8(\mathrm{SL}_2(\mathbb{Z}))$ is generated by G_4^2 and so the base case is true.

Now assume that $\mathcal{M}_r(\mathrm{SL}_2(\mathbb{Z}))$ is generated by G_4 and G_6 for all $8 \leqslant r < k - 1$. Choose some $a_1, b_1 \in \mathbb{Z}_{\geqslant 0}$ such that $4a_1 + 6b_1 = k$. The function $f := G_4^{a_1} G_6^{b_1}$ is a modular form of weight k. For any $g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, $\exists \alpha \in \mathbb{C}$ such that $g - \alpha f$ is a cusp form of weight k and hence is in $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$. By Theorem 2.21, $f - \alpha g = h\Delta^*$ for some $h \in \mathcal{S}_{k-12}(\mathrm{SL}_2(\mathbb{Z}))$. But $h\Delta^*$ is a cusp form of lower weight (hence a modular form of lower weight) and thus by our inductive hypothesis can be written as a sum of $G_4^s G_6^t$ with $s, t \in \mathbb{Z}_{\geqslant 0}$. Clearly, then, so can $g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

It remains to prove that the forms given by $G_4^a G_6^b$ with $4a + 6b = k$ are linearly independent. Suppose for the sake of contradiction that it was not the case. So, suppose

$$c_1 G_4^{a_1} G_6^{b_1} + c_2 G_4^{a_2} G_6^{b_2} + ... + c_m G_4^{a_m} G_6^{b_m} = 0, \quad c_i \in \mathbb{C}.$$

Assume $b_1 = max\{b_j : 1 \leqslant j \leqslant m\}$ (as if it were not, reorder the terms such that they are and rename the indices). Then divide throught by $G_4^{a_1} G_6^{b_1}$ to obtain

$$c_1 + c_2 G_4^{a_2 - a_1} G_6^{b_2 - b_1} + ... + c_m G_4^{a_m - a_1} G_6^{b_m - b_1} = 0, \quad c_i \in \mathbb{C}.$$

Since b_1 was the maximum of all the b_j, are and rename the indices). Then divide throught by $G_4^{a_1} G_6^{b_1}$ to obtain

$$c_1 + c_2 \frac{G_4^{a_2 - a_1}}{G_6^{b_1 - b_2}} + ... + c_m \frac{G_4^{a_m - a_1}}{G_6^{b_1 - b_m}} = 0, \quad c_i \in \mathbb{C}.$$

Using the fact that $4a_j - 4a_1 = 6b_1 - 6b_j$ for all $1 \leqslant j \leqslant m$, we have

$$c_1 + c_2 \Big(\frac{G_4^6}{G_6^4}\Big)^{(a_2 - a_1)/6} + ... + c_m \Big(\frac{G_4^6}{G_6^4}\Big)^{(a_m - a_1)/6} = 0, \quad c_i \in \mathbb{C},$$

implying that G_4^6/G_6^4 satisfies a nonzero algebraic equation. Satisfying such an equation implies G_4^6/G_6^4 is constant (as if it were not, it would have nontrivial dependence on a variable, but then it would not satisfy the equation). However, from Theorem 2.21, we saw the only zero of G_4 was at ζ_3 and the only zero of G_6 was at the complex number i, contradicting that G_4^6/G_6^4 is constant. □

Corollary 2.24. *The space of modular forms $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$ is generated by G_4 and G_6.*

Proof. By Theorem 2.23, for any $k \in \mathbb{Z}_{\geqslant 0}$, $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ has a basis $\{G_4^a G_6^b\}$ with $4a + 6b = k$. By construction, $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z})) = \bigoplus_{k=0}^{\infty} \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, so G_4 and G_6 certainly generate $\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))$. □

Corollary 2.25. [7, Corollary 1] *Let $k \geqslant 0$ even. The dimension of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ can be computed as follows:*

$$dim(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) = \begin{cases} \lfloor k/12 \rfloor & \text{if } k \equiv 2 \bmod 12 \\ \lfloor k/12 \rfloor + 1 & \text{if } k \not\equiv 2 \bmod 12 \end{cases}$$

Proof. The following proof is due to [2]. This is obviously true for even k with $0 \leqslant k < 12$. Note that when $k > 2$ and even, $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \neq \{0\}$, so the map ϕ in Corollary 2.22 is onto.

Theorem 2.21 proved that $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \cong \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$, and the proof of Corollary 2.22 yielded that $dim(\mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))) = dim(\mathcal{M}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))) - 1$. Thus, $dim(\mathcal{M}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))) = dim(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) + 1$. Replacing k by $k+12$ in the above formula yields the same. □

Although in Corollary 2.24 we only have a result about the basis for $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, Theorem 4 in Chapter 10 of [6] gives a more general result about the basis of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, produced below (omitting the proof).

Theorem 2.26. [6, Chapter 10, Theorem 4] *A basis β for $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ with coefficients in $\mathbb{Z}$ (which is also a basis for $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ with coefficients in $\mathbb{C}$) is:*

(1) If $k \equiv 0 \bmod 4$, then $\beta = \{E_4^a \Delta^b\}$ with $4a + 12b = k$.

(2) If $k \equiv 2 \bmod 4$, then $\beta = \{E_6 E_4^a \Delta^b\}$ with $4a + 12b = k - 6$.

3. Congruence Subgroups

Definition 3.1. [2, pg. 13] *Let $n \in \mathbb{Z}^+$. The principal congruence subgroup of level n is the subgroup $\Gamma(n) \subseteq \mathrm{SL}_2(\mathbb{Z})$ given by*

$$\Gamma(n) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{n} \right\} \tag{3.1}$$

Definition 3.2. [2, Definition 1.2.1] *A subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ is called a congruence subgroup if it contains $\Gamma(n)$ for some $n \in \mathbb{Z}$. In this case, we call Γ a congruence subgroup of level n.*

Two well-known congruence subgroups are:

$$\Gamma_0(n) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} \bar{a} & \bar{b} \\ 0 & \bar{d} \end{bmatrix} \pmod{n} \right\} \tag{3.2}$$

$$\Gamma_1(n) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \;\middle|\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & \bar{b} \\ 0 & 1 \end{bmatrix} \pmod{n} \right\}. \tag{3.3}$$

These congruence subgroups satisfy the relation

$$\Gamma(n) \subsetneq \Gamma_1(n) \subsetneq \Gamma_0(n).$$

Proposition 3.3. [2, pg. 13] *For any $n \in \mathbb{N}, \Gamma(n) \trianglelefteq \Gamma_1(n) \trianglelefteq \Gamma_0(n)$.*

Proof. We follow the proof in [2, pg. 13], and we will show successive normality separately: that is, we first show normality of $\Gamma(n)$ in $\Gamma_1(n)$, then normality of $\Gamma_1(n)$ in $\Gamma_0(n)$.

Define a map

$$\varphi : \Gamma_1(n) \to \mathbb{Z}_n$$

where for any $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_1(n)$,

$$\varphi(A) = b \,(\text{mod n}).$$

We have,

$$\begin{aligned} ker(\varphi) &= \{A \in \Gamma_1(n) \,|\, \varphi(A) \equiv 0 \ (\text{mod n})\} \\ &= \{A \in \Gamma_1(n) \,|\, b \equiv 0 \ (\text{mod n})\} \\ &= \Gamma(n). \end{aligned}$$

Since $ker(\varphi)$ is always a normal subgroup of $\Gamma_1(n)$, this shows that $\Gamma(n) \trianglelefteq \Gamma_1(n)$.

Define another map

$$\varphi^* : \Gamma_0(n) \to (\mathbb{Z}_n)^*$$

where for any $B = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \in \Gamma_0(n)$,

$$\varphi^*(B) = d_1 \ (\text{mod n}).$$

We have,

$$\begin{aligned} ker(\varphi^*) &= \{B \in \Gamma_0(n) \,|\, \varphi^*(A) \equiv 1 \ (\text{mod n})\} \\ &= \{B \in \Gamma_0(n) \,|\, d_1 \equiv 1 \ (\text{mod n})\} \end{aligned}$$

By definition of $\mathrm{SL}_2(\mathbb{Z})$, $det(B) = 1$, so $ad = 1$, implying that $ad \equiv 1$ (mod n). If we require that $d \equiv 1$ (mod n), then $a \equiv 1$ (mod n) as well. Then $ker(\varphi^*)$ is exactly $\Gamma_1(n)$. Again, as $ker(\varphi^*)$ is normal in $\Gamma_0(n)$, we have shown $\Gamma_1(n) \trianglelefteq \Gamma_0(n)$. □

Corollary 3.4. [2, pg. 14] *Let $\Gamma(n)$, $\Gamma_0(n)$, $\Gamma_1(n)$ be defined as in Equations (3.1), (3.2), (3.3) respectively. Then, $[\Gamma_1(n) : \Gamma(n)] = n$; $[\Gamma_0(n) : \Gamma_1(n)] = \phi(n)$, where $\phi(n)$ is the Euler totient Function.*

Proof. The maps defined in Proposition (3.3), φ and φ^*, are clearly onto maps. Using the First Isomorphism Theorem together with Proposition (3.3), we have that $\Gamma_1(n)/\Gamma(n) \cong \mathbb{Z}_n$ which implies that $|\Gamma_1(n)/\Gamma(n)| = [\Gamma_1(n) : \Gamma(n)] = |\mathbb{Z}_n| = n$. Similarly, we have that $|\Gamma_0(n)/\Gamma_1(n)| = [\Gamma_0(n) : \Gamma_1(n)] = |(\mathbb{Z}_n)^*| = \phi(n)$. □

Definition 3.5. [2, pg. 164] *Let Γ_1, Γ_2 be congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ (also subgroups of $\mathrm{GL}_2^+(\mathbb{Q})$). Then the set*

$$\Gamma_1 \alpha \Gamma_2 = \{\gamma_1 \alpha \gamma_2 | \gamma_1 \in \Gamma_1, \alpha \in \mathrm{GL}_2^+(\mathbb{Q}), \gamma_2 \in \Gamma_2\}$$

is a double coset in $\mathrm{GL}_2^+(\mathbb{Q})$.

The subgroup Γ_1 of $\mathrm{SL}_2(\mathbb{Z})$ acts on the double coset $\Gamma_1 \alpha \Gamma_2$ by left multiplication, allowing us to partition the set $\Gamma_1 \alpha \Gamma_2$ into disjoint orbits $\Gamma_1 \alpha \Gamma_2 = \bigsqcup_j \Gamma_1 \beta_j$, where $\beta_j = \gamma_1 a \gamma_2$ for some $\gamma_1 \in \Gamma_1, a \in \alpha$, and $\gamma_2 \in \Gamma_2$ (see [2, pg. 164]).

Definition 3.6. [2, pg. 14] *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. *The factor of automorphy* $j(A, z) \in \mathbb{C}$ *for* $z \in \mathcal{H}$ *is given by*

$$j(A, z) = cz + d.$$

Define the weight k *operator,* $f[A]_k$, *on functions* $f : \mathcal{H} \to \mathbb{C}$ *by*

$$(f[A]_k)(z) = j(A, z)^{-k} f(Az), \ A \in \mathrm{SL}_2(\mathbb{Z}). \tag{3.4}$$

The previously defined weight k operator $f[A]_k$ can be generalized to matrices $B \in \mathrm{GL}_2^+(\mathbb{Q})$ by $(f[B]_k)(z) = det(B)^{k-1} j(B, z)^{-k} f(Bz), \ A \in \mathrm{SL}_2(\mathbb{Z})$. This is in fact a generalization since this reduces to our definition when using matrices in $\mathrm{SL}_2(\mathbb{Z})$ (by definition, they always have determinant equal to 1).

Remark 3.7. In Definition 3.6, note that the functions on which $f[A]_k$ operates are not necessarily weakly modular; we can actually use the previously defined operator $f[A]_k$ on functions $f : \mathcal{H} \to \mathbb{C}$ to form an equivalent definition of weakly modular: a function is weakly modular of weight k if $f[A]_k \equiv f$, $\forall A \in \mathrm{SL}_2(\mathbb{Z})$. This "new" definition clearly coincides with Definition 2.12.

Definition 3.8. [2, pg. 165] *Let* Γ_1, Γ_2 *be congruence subgroups of* $\mathrm{SL}_2(\mathbb{Z})$, *and let* $A \in \mathrm{GL}_2^+(\mathbb{Q})$. *Define the weight* k *operator* $[\Gamma_1 A \Gamma_2]_k$ *on functions* $f : \mathcal{H} \to \mathbb{C}$ *by*

$$f[\Gamma_1 A \Gamma_2]_k = \sum_j f[\alpha_j]_k \tag{3.5}$$

where the α_j *are orbit representatives from the action of* Γ_1 *on* $\Gamma_1 A \Gamma_2$ *(see below Definition* 3.5*).*

Definition 3.9. [2, Definition 1.2.3] *Let* Γ *be a congruence subgroup and let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *We say a function* $f : \mathcal{H} \to \mathbb{C}$ *is a modular form of weight* k *with respect to* Γ *if the following three properties hold:*

1. f is holomorphic on $\mathcal{H}$.
2. $f(Az) = (cz + d)^k f(z), \forall A \in \Gamma$.
3. $f[A]_k$ is holomorphic at infinity $\forall A \in \mathrm{SL}_2(\mathbb{Z})$.

As in [2, pg. 17], we denote the space of modular forms of weight k with respect to Γ as $\mathcal{M}_k(\Gamma)$. Moreover, $\mathcal{S}_k(\Gamma)$ are the cusp forms of weight k with respect to Γ. The space of modular forms with respect to Γ is the set

$$\mathcal{M}(\Gamma) = \bigoplus_{k=0}^{\infty} \mathcal{M}_k(\Gamma)$$

which forms a graded ring. The set of cusp forms with respect to Γ

$$\mathcal{S}(\Gamma) = \bigoplus_{k=0}^{\infty} \mathcal{S}_k(\Gamma)$$

is a graded ideal in $\mathcal{M}(\Gamma)$.

4. Hecke Operators

Definition 4.1. *Consider the double coset given by $\Gamma_1(n)\alpha\Gamma_1(n)$, where $\alpha = \begin{bmatrix} 1 & 0 \\ 0 & p \end{bmatrix}$ and p is a prime.*

The Hecke operator, $T_p : \mathcal{M}(\Gamma_1(n)) \to \mathcal{M}(\Gamma_1(n))$, is given by

$$\begin{aligned} T_p(f)(z) &= f[\Gamma_1(n)\alpha\Gamma_1(n)]_k(z) \\ &= \sum_j f[\beta_j]_k \end{aligned}$$

where β_j are distinct orbit representatives.

Lemma 4.2. [2, Proposition 5.2.1]

$$T_p(f) = \begin{cases} \sum_{j=0}^{p-1} f[\begin{bmatrix} 1 & j \\ 0 & p \end{bmatrix}]_k, & \text{if } p|n \\ \sum_{j=0}^{p-1} f[\begin{bmatrix} 1 & j \\ 0 & p \end{bmatrix}]_k + f[\begin{bmatrix} a & b \\ n & p \end{bmatrix}\begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix}]_k, & \text{if } p \nmid n, \text{ where } ap - nb = 1. \end{cases}$$

Proposition 4.3. *Let $f(z) = \sum_{j=0}^{\infty} a_m q^m \in \mathcal{M}(\Gamma_1(n))$ and let p be a prime not dividing n. The effect of the Hecke operator T_p on $f(z)$ can be characterized as*

$$T_p : \sum_{m=0}^{\infty} a_m q^m \mapsto \sum_{m=0}^{\infty} a_{mp} q^m + p^{k-1} \sum_{m=0}^{\infty} a_m q^{mp}$$

Proof. □

5. Modular Forms mod ℓ

Definition 5.1. *Let ℓ be a prime and let v_ℓ be the ℓ-adic valuation of $\mathbb{Q}$. That is, if for any $a \in \mathbb{Q}$ one writes $a = \ell^t(b/c)$ where $t \in \mathbb{Z}$ and ℓ divides neither b nor c. Then,*

$$v_\ell(a) = t.$$

If an element a of $\mathbb{Q}$ has nonnegative ℓ-adic valuation, a is said to be ℓ-integral.

If $f(z) \in \mathbb{Q}[[z]]$ in which all its coefficients are ℓ-integral, then one can consider its reduction mod ℓ. That is, if we write all of the coefficients of $f(z)$ in the most reduced form, since they are all ℓ-integral, the denominator cannot be divisible by ℓ. Thus, when we reduce such a form mod ℓ, no coefficients are left undefined; we never divide by 0.
Let

$$f(z) = \sum_{n=0}^{\infty} a_n q^n \in \mathbb{Q}[[x]]$$

with a_i ℓ-integral for all i. Then it's reduction in $\mathbb{F}_\ell[[z]]$ is given by

$$\overline{f(z)} = \sum_{n=0}^{\infty} \overline{a_i} q^n,$$

where $\overline{a_i} \equiv a_i \pmod{\ell}$.
Let Γ be a congruence subgroup of level n. Fix a prime ℓ not diving n. Let S be the set of all $f \in \mathcal{M}_k(\Gamma)$ whose q-coefficients at infinity are rational and ℓ-integral.

Definition 5.2. *The space of modular forms mod ℓ of weight k and level n, $\widetilde{\mathcal{M}_k(\Gamma)}$, is the set*

$$\left\{\overline{g(z)} \in \mathbb{F}_\ell[[z]] \;\middle|\; g(z) \in S \subseteq \mathcal{M}_k(\Gamma)\right\}.$$

The set $\widetilde{\mathcal{M}_k(\Gamma)}$ is a vector space over $\overline{\mathbb{F}_\ell}$. Denote $\widetilde{\mathcal{M}(\Gamma)}$ as the sum of the $\widetilde{\mathcal{M}_k(\Gamma)}$. Swinnerton-Dyer and Serre (see [9] and [8]) showed that for all k,

$$\widetilde{\mathcal{M}_k(\Gamma)} \subseteq \widetilde{\mathcal{M}_{k+\ell-1}(\Gamma)}$$

and so the sum is not a direct sum. Since we have the above relationship between modular forms of weight k and weight $k+\ell-1$ mod ℓ, the previously fruitful notion of "weight" is no longer well-defined mod ℓ. We then divert to the *filtration* of a modular form mod ℓ.

Definition 5.3. *Let Γ be a congruence subgroup of level n and $f(z) \in \widetilde{\mathcal{M}_k(\Gamma)}$. Then the value*

$$w(f) = inf\{j \mid f(z) \in \mathcal{M}_j(\Gamma)\}$$

is called the filtration of f.

Proposition 5.4. *Let $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ (i.e. work in level 1) and let $\ell = 2$ or 3. Then, $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} = \mathbb{F}_\ell[\overline{\Delta}]$.*

Proof. Since we are working in $\mathbb{Q}[[z]]$, we refer to Theorem 2.26 which gives us a basis for $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ with coefficients in $\mathbb{Z}$. Since all the coefficients are rational and ℓ-integral (since they are all integers), it is also a basis for $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ with rational and ℓ-integral coefficients.

Using Fact **??** and normalizing, we have

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

where $\sigma_{k-1}(n)$ is as defined in Fact **??**. Then,

$$E_4 = 1 + 240\sum_{n=0}^{\infty}\sigma_3(n)q^n \equiv 1 \pmod{2 \text{ or } 3}.$$

$$E_6 = 1 - 504\sum_{n=0}^{\infty}\sigma_5(n)q^n \equiv 1 \pmod{2 \text{ or } 3}.$$

By Theorem 2.26, for any $f(z) \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, either

$$(1)\ f(z) = \sum {}^{d_{a,b}}\!/_{(2\pi)^{12}} E_4^a \Delta^b$$

$$(2)\ f(z) = \sum {}^{d_{a,b}}\!/_{(2\pi)^{12}} E_6 E_4^a \Delta^b$$

for some $d_{a,b} \in \mathbb{Z}$. Reducing both cases mod 2 or 3, we are left with

$$\overline{f(z)} = \sum {}^{\overline{d_{a,b}}}\!/_{(2\pi)^{12}} (\overline{\Delta})^b$$

This implies $f(z) \in \widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))}$ is a sum of powers of $\overline{\Delta}$ with coefficients in $\mathbb{F}_\ell$. Equivalently, this means $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} \subseteq \mathbb{F}_\ell[\overline{\Delta}/(2\pi)^{12}]$. The reserve containment is obvious after noting that Δ has integer (hence ℓ-integral) coefficients. □

From this point, unless otherwise stated, we will work in level 1 (i.e. $\Gamma = \mathrm{SL}_2(\mathbb{Z})$) and assume that $\ell \geqslant 3$ is prime, since the by the previous proposition, the cases of $\ell = 2$ or 3 are trivial.

We will now introduce three different operators on the space of modular forms mod ℓ. The first operator is Atkin's U_ℓ operator (as in [3, pg. 255]), and it takes $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))}$ to itself. We denote it by U. For each of the following operators, we describe their effect of a modular form $f(z)$ by describing its effect on the q-expansion of $f(z)$. So, let $f(z) = \sum_{n=0}^{\infty} a_n q^n \in \widetilde{\mathcal{M}(\mathrm{SL}_2(\mathbb{Z}))}$.

$$(i)\ U : \sum_{n=0}^{\infty} a_n q^n \mapsto \sum_{n=0}^{\infty} a_{nl} q^n$$

$$(ii)\ V : \sum_{n=0}^{\infty} a_n q^n \mapsto \sum_{n=0}^{\infty} a_n q^{nl}$$

$$(iii)\ \theta : \sum_{n=0}^{\infty} a_n q^n \mapsto \sum_{n=0}^{\infty} n a_n q^n$$

Proposition 5.5. [3, Fact 2.2] *The following describe some of the relationships between the previously introduced operators.*

$$(i)\ f|\,VU = f, \quad \forall f \in \widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))},$$
$$(ii)\ ker(\theta) = Im(V),$$
$$(iii)\ Im(\theta) = ker(U).$$

Proof. □

Fact 5.6. [4, Fact 1.7] If $f \in \widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))}$, then $w(f|V) = \ell w(f)$.

Fact 5.7. [4, Fact 1.4] The operator θ maps $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))}$ to $\widetilde{\mathcal{M}_{k+\ell+1}(\mathrm{SL}_2(\mathbb{Z}))}$. In particular, $w(f|\theta) \leqslant w(f) + \ell + 1$.

Lemma 5.8. [4, Lemma 1.9] *Let $f \in \widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))}$. Then $w(f|U) \leqslant {}^{(w(f)\,-\,1)}/_{\ell} + \ell$.*

Proof. □

Definition 5.9. [4, Definition 2.1] *The set $\{\lambda_p\}$ is called a system of eigenvalues if there is some nonzero eigenform f such that $f|T_p = \lambda_p f$, for all primes p.*

Definition 5.10. [4, Section 3] *Let $\mathbb{T}_k$ be the subring of $End_{\mathbb{C}}(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})))$ generated by the Hecke Operators. The subring $\mathbb{T}_k$ is called the Hecke ring.*

Definition 5.11. [4, Section 3] *Let R_k be the subring of $End_{\mathbb{F}_\ell}(\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))})$ generated by the Hecke Operators. The subring R_k is called the Hecke ring mod ℓ.*

Remark 5.12. The above two definitions can be generalized to any level n congruence subgroup, but for the purposes of this paper and results presented, we are limiting ourselves to level 1.

The ring $R_k \cong \mathbb{T}_k/\ell\mathbb{T}_k$ is an Artin ring, hence the ring $R_k \otimes \overline{\mathbb{F}_\ell}$ is also an Artin ring. As such, $R_k \otimes \overline{\mathbb{F}_\ell}$ has a finite number of maximal ideals and can be decomposed into a finite direct product of local Artin rings, which we will call local components. Moreover, each of the maximal ideals in $R_k \otimes \overline{F_\ell}$ is generated by the operators $T_p - \lambda_p$ where $\{\lambda_p\}$ is the system of eigenvalues. In other words, each maximal ideal in $R_k \otimes \overline{\mathbb{F}_\ell}$ is associated uniquely to a system of eigenvalues (see [4, Section 3]).

Definition 5.13. [4, Definition 3.1] *Let A_k^j be a local component of $R_k \otimes \overline{\mathbb{F}_\ell}$ and m_j its unique maximal ideal. Let $S_k^j \subseteq \widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} \otimes \overline{\mathbb{F}_\ell}$ be the generalized eigenspace for the system of eigenvalues $\{\lambda_p\}$ associated to the maximal ideal m_j; that is, S_k^j is composed of those forms in $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} \otimes \overline{\mathbb{F}_\ell}$ that are annihilated by some power of m_j. Since m_j is generated by $T_p - \lambda_p$, one could equivalently define S_k^j to be the set of the forms in $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} \otimes \overline{\mathbb{F}_\ell}$ which are annihilated by some power of the operators $T_p - \lambda_p$ for all p.*

As in [3], the images of the operators in a local component will again be denoted by T_p or U. We say a local component A_k^j of $R_k \otimes \overline{\mathbb{F}_\ell}$ is U-nilpotent if the image of operator U is nilpotent in it.

Lemma 5.14. [4, Lemma 3.2] *Let $S_k^j \subseteq \widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} \otimes \overline{\mathbb{F}_l}$ be the generalized eigenspace associated to A_k^j. Suppose S_k^j has a form f such that $w(f) > \ell + 1$. Then A_k^j is U-nilpotent.*

Proof. Suppose that S_k^j has a form of filtration greater than $\ell + 1$, yet A_k^j is not U-nilpotent.

Recall that in an Artin ring, all prime ideals are maximal. In particular, the nilradical and the Jacobson radical are equal. Since A_k^j is a local Artin ring, the nilradical is simply the unique maximal ideal. Thus, any element of A_k^j is either nilpotent or a unit. By assumption, U is not a nilpotent element meaning that U is a unit.

Since U is a unit, it is bijective on $\widetilde{\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))} \otimes \overline{\mathbb{F}_l}$, and in particular, on the generalized eigenspace S_k^j.

Let g be an element of maximal filtration in S_k^j. As U is bijective, $\exists g^* \in S_k^j$ such that $g^*|U = g$. Thus, $w(g^*) \geqslant w(g^*|U) = w(g) > \ell + 1$. As $w(g^*) > \ell + 1$ and the previous proposition said applying U decreases the filtration, we have that $w(g^*) > w(g^*|U) = w(g)$, a contradiction. □

In particular, the previous Lemma implies that if a local component A_k^j of $R_k \otimes \overline{\mathbb{F}_\ell}$ is not U-nilpotent, then it only has forms of filtration at most $\ell + 1$.

Definition 5.15. *Let R be a local ring with unique maximal ideal m. The* Zariski tangent dimension *is the dimension of m/m^2 over R/m.*

Fact 5.16. [4] The following fact can be found in the proof of Theorem 3.4 of [4]. Some generalized eigenspace of $R_k \otimes \overline{\mathbb{F}_\ell}$ must have an element f of filtration $w(f)$ satisfying

$$\ell+1<w(f)\leqslant\begin{cases}2\ell & \ell\geqslant 13 \text{ and } k\geqslant 2\ell^2\\ 3\ell-1 & \ell=7 \text{ or } 11 \text{ and } k\geqslant 3\ell^2-\ell\\ 3\ell+3 & \ell=5 \text{ and } k\geqslant 3\ell^2+3\ell\end{cases}$$

Lemma 5.17. [4, Lemma 4.2] *Let $S_k^j\subseteq\mathcal{M}_k\widetilde{(\mathrm{SL}_2(\mathbb{Z}))}\otimes\overline{\mathbb{F}_\ell}$ be the generalized eigenspace associated to the local component A_k^j in $R_k\otimes\overline{\mathbb{F}_\ell}$. Let m_j be the unique maximal ideal in A_k^j. If S_k^j has a form f of filtration s with ${}^k/_\ell\geqslant s>\ell+1$, then $dim_{A_k^j/m_j}({}^{m_j}/_{m_j^2})\geqslant 2$.*

Proof. □

Theorem 5.18. *If we are in any of the following cases*

1. $\ell\geqslant 13$ and $k\geqslant 2\ell^2$,
2. $\ell=7$ or 11 and $k\geqslant 3\ell^2-\ell$,
3. $\ell=5$ and $k\geqslant 3\ell^2+3\ell$

then $dim_{A_k^j/m_j}({}^{m_j}/_{m_j^2})\geqslant 2$ for at least one $j\in\{1,2,...,n\}$.

Proof. Assume we are in one of the cases listed in the Theorem. Then, by Fact 5.16, at least one of the generalized eigenspaces of $R_k\otimes\overline{\mathbb{F}_\ell}$ must have a form f of filtration satisfying $\ell+1<w(f)\leqslant {}^k/_\ell$. Then by the lemma, the associated local component(s) to said generalized eigenspace(s) must have Zariski tangent dimension at least 2. □

Proposition 5.19. *Any S_k^j contains a simultaneous eigenform f of filtration $w(f)$ such that $w(f)\leqslant\ell^2+\ell$.*

Proof. If S_k^j is not U-nilpotent, then we've already seen it has forms of filtration at most $\ell+1<\ell^2+\ell$.

**********(2) implies $w(f|\theta^2)\leqslant w(f|\theta)+\ell+1\leqslant(\ell+1)+\ell+1=2(\ell+1)$. In the same way, we see that $w(f|\theta^r)\leqslant r(\ell+1)\leqslant(\ell-1)(\ell+1)=\ell^2-\ell<\ell^2+\ell$.********** □

Theorem 5.20. [3, Theorem 4.5] *If $k>\ell^3+\ell^2$, then $dim_{A_k^j/m_j}({}^{m_j}/_{m_j^2})\geqslant 2$, $\forall j$.*

Proof. The theorem becomes trivial if we show when $k>\ell^3+\ell^2$, there's a form f with filtration satisfying $\ell+1<w(f)\leqslant\ell^2+\ell$. The last proposition says that there's always a form g of filtration at most $\ell^2+\ell$. Moreover, if $w(g)<\ell+1$, then $\ell+1<w(g|V)\leqslant\ell^2+\ell$. □

We will keep the same notation as in [3]. Let $\mathbb{T}_k$ denote the commutative subring of $End_{\mathbb{C}}(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})))$ generated by the Hecke operators. Moreover, let $\mathcal{O}_k$ be the ring of integers in $\mathbb{T}_k\otimes\mathbb{Q}$.

Theorem 5.21. [3, Theorem 5.1] *Fix a prime ℓ. If we are in any of the following cases:*

1. $\ell\geqslant 13$ and $k\geqslant 2\ell^2$,
2. $\ell=7$ or 11 and $k\geqslant 3\ell^2-\ell$,
3. $\ell=5$ and $k\geqslant 3\ell^2+3\ell$

then ℓ divides $[\mathcal{O}_k:\mathbb{T}_k]$.

Proof. 1. If A is a local Artin ring, then principal ideal ring (PIR) is equivalent to having Zariski tangent dimension less than or equal to 1.
2. $\mathcal{O}_k$ is isomorphic to a finite direct product of Dedekind domains.
3. $R_k \cong \mathbb{T}_k/\ell\mathbb{T}_k$.
The proof is by way of contradiction, so assume that ℓ doesn't divide the index $[\mathcal{O}_k : \mathbb{T}_k]$. Then, the map $\phi : \mathcal{O}_k/\mathbb{T}_k \to \mathcal{O}_k/\mathbb{T}_k$ given by multiplication by ℓ is an isomorphism, and so $\mathcal{O}_k = \ell\mathcal{O}_k + \mathbb{T}_k$. Note that $\mathbb{T}_k$ is a subring and $\ell\mathcal{O}_k$ is an ideal in $\mathcal{O}_k$. We have,

$$\mathcal{O}_k/\ell\mathcal{O}_k = (\ell\mathcal{O}_k + \mathbb{T}_k)/\ell\mathcal{O}_k \cong \mathbb{T}_k/(\ell\mathcal{O}_k \cap \mathbb{T}_k) = \mathbb{T}_k/\ell\mathbb{T}_k.$$

where the isomorphism is due to the Second Isomorphism Theorem for rings.

By fact 2, $\mathcal{O}_k$ is isomorphic to a finite direct product of Dedekind domains, so the factor ring $\mathcal{O}_k/\ell\mathcal{O}_k \cong \mathbb{T}_k/\ell\mathbb{T}_k \cong R_k$ is a PIR.

This means that the Artin ring R_k is also a PIR, and so are all of its local components, hence so are the local components of $R_k \otimes \overline{\mathbb{F}_\ell}$.

Then, by fact 1, each of the local components has Zariski tangent dimension less than or equal to 1, contradicting our previous theorem, which says under the given conditions of k, it has Zariski tangent dimension greater than 1. □

Theorem 5.22. [4, Theorem 3.5] *Let ℓ a prime relatively prime to the level (and if the level is not 1, $\ell \neq 2, 3$). Then $ord_\ell(d(\mathbb{T}_k))$ grows linearly with k.*

Theorem 5.23. [5, Theorem 5.4] *Let ℓ be a prime and n any positive integer. If k is sufficiently large, then ℓ^n divides $[\mathcal{O}_k : \mathbb{T}_k]$.*

References

[1] Rolf Busam and Eberhard Freitag, *Complex Analysis*. Springer-Verlag, Berlin-Heidelberg, 2005.
[2] Fred. Diamond and Jerry Shurman, *A First Course in Modular Forms*. Graduate Texts in Mathematics **228**. Springer-Verlag, New York, NY, 2005.
[3] Naomi Jochnowitz, *A Study of the Local Components of the Hecke Algebra mod ℓ*. Trans. Amer. Math. Soc., **270** (1982), pp. 253-267.
[4] Naomi Jochnowitz, *Congruences between systems of eigenvalues of modular forms*. Trans. Amer. Math. Soc. **270** (1982) 269-285.
[5] Naomi Jochnowitz, *The Index of the Hecke Ring, T_k in the Ring of Integers of $T_k \otimes \mathbb{Q}$*. Duke Math. J. **46** (1979) 253-267.
[6] Serge Lang, *Introduction to Modular Forms*. Springer-Verlag, Berlin-Heidelberg, 1987.
[7] Jean-Pierre Serre, *A Course in Arithmetic*. Graduate Texts in Mathematics **7**. Springer-Verlag, New York, NY, 1973.
[8]
[9]