Mod p Modular Forms

Andrew Mendelsohn

2018/19

CID: 01540461
Supervisor: Dr David Helm

Abstract

Contents

1 Preliminaries

1.1 Algebraic Number Theory

This subsection relies on the results and proofs of [Cox13], [Lan94], and [Sut16].

Let K be a number field and $\mathcal{O}_K$ denote its ring of integers. A Dedekind domain is an integrally closed, noetherian domain in which every (non-zero) prime ideal is maximal. In addition, in a Dedekind domain every fractional ideal has a unique factorization into prime ideals. Recall that $\mathcal{O}_K$ is a Dedekind domain.

Definition 1. Let $A \subset B$ be an inclusion of rings and $\mathfrak{p}$ a prime ideal in A. If $\mathfrak{q}$ is a prime ideal of B, we say $\mathfrak{q}$ lies above $\mathfrak{p}$ if $\mathfrak{q} \cap A = \mathfrak{p}$.

Let L be a finite extension of K and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_K$. Then $\mathfrak{p}\mathcal{O}_L$ is an ideal of $\mathcal{O}_L$, which is a Dedekind domain, so $\mathfrak{p}\mathcal{O}_L$ has a unique factorization into prime ideals of $\mathcal{O}_L$:

$$\mathfrak{p}\mathcal{O}_L = \prod_i \mathfrak{q}_i^{e_i} = \mathfrak{q}_1^{e_1} ... \mathfrak{q}_g^{e_g}, \tag{1}$$

for $\mathfrak{q}_i$ prime in $\mathcal{O}_L$ all lying above $\mathfrak{p}$.
We call e_i the *ramification index* of $\mathfrak{q}_i$ over $\mathfrak{p}$. We call $f_i = \left[\mathcal{O}_L/\mathfrak{q}_i : \mathcal{O}_K/\mathfrak{p}\right]$ the *inertial degree* of $\mathfrak{p}$ in $\mathfrak{q}_i$. Note that since both $\mathcal{O}_L/\mathfrak{q}_i$ and $\mathcal{O}_K/\mathfrak{p}$ are finite fields, $\mathcal{O}_L/\mathfrak{q}_i$ is a separable and normal extension of $\mathcal{O}_K/\mathfrak{p}$.

For the following results, the picture is as follows: we take a number field K and its ring of integers $\mathcal{O}_K$, together with L, a finite, Galois extension of K, and its ring of integers $\mathcal{O}_L$ (which is the integral closure of $\mathcal{O}_K$ in L) to obtain:

$$\begin{array}{ccc} K & \subset & L \\ \cup & & \cup \\ \mathcal{O}_K & \subset & \mathcal{O}_L \end{array}$$

Lemma 1. *Let L be a finite, Galois extension of K and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_k$. If $\mathfrak{q}_1$ and $\mathfrak{q}_2$ are prime ideals of $\mathcal{O}_L$ lying above $\mathfrak{p}$, there exists an element $\sigma \in Gal(L/K)$ such that $\sigma(\mathfrak{q}_1) = \mathfrak{q}_2$.*

Proof. Suppose that $\mathfrak{q}_1 \neq \sigma(\mathfrak{q}_2)$ for all $\sigma \in Gal(L/K)$. In a Dedekind domain, non-zero prime ideals are maximal, and in general distinct maximal ideals are coprime. Thus $\mathfrak{q}_1$ and $\mathfrak{q}_2$ are coprime, and we can apply the Chinese Remainder Theorem to find $x \in \mathcal{O}_L$ such that

$$x \equiv 0 \bmod \mathfrak{q}_1 \text{ and } x \equiv 1 \bmod \sigma(\mathfrak{q}_2)$$

for all $\sigma \in Gal(L/K)$. Denote $Gal(L/K)$ by G. The norm $N_{L/K}(x) = \prod_{\sigma \in G} \sigma(x)$ maps L to K and $\mathcal{O}_L$ to $\mathcal{O}_L$, so $N_{L/K}(x) \in K \cap \mathcal{O}_L = \mathcal{O}_K$. Moreover, since $x \equiv 0 \bmod \mathfrak{q}_1$, $N_{L/K}(x) \in \mathfrak{q}_1 \cap \mathcal{O}_K = \mathfrak{p}$. However, since $x \notin \sigma(\mathfrak{q}_2)$ for all $\sigma \in G$, we must have $\sigma^{-1}(x) \notin \mathfrak{q}_2$ for all $\sigma \in G$, which is equivalent to $\sigma(x) \notin \mathfrak{q}_2$ for all $\sigma \in G$. This is a contradiction, as $N_{L/K}(x) = \prod_{\sigma \in G} \sigma(x) \in \mathfrak{p} \subset \mathfrak{q}_2$, as $\mathfrak{q}_2$ lies above $\mathfrak{p}$. □

Note: this is equivalent to saying G acts transitively on the set of primes lying above $\mathfrak{p}$.

Corollary 1. *If L and K are as above, and $\mathfrak{q}_1, ..., \mathfrak{q}_g$ are the prime ideals of $\mathcal{O}_L$ lying above $\mathfrak{p} \subset \mathcal{O}_K$, then*

(i) *The $\mathfrak{q}_i$ have the same ramification index, e, for all i.*

(ii) *The inertial degrees f_i of $\mathfrak{p}$ in $\mathfrak{q}_i$ are equal, for all i.*

Proof.

(i) Let $\sigma \in G = Gal(L/K)$. Then σ fixes K, thereby fixing $\mathfrak{p}$. In addition, σ fixes $\mathcal{O}_L$: if $\alpha \in \mathcal{O}_L$, there exists a monic polynomial f with coefficients in $\mathcal{O}_K$ such that $f(\alpha) = 0$. Then $0 = \sigma(f(\alpha)) = f(\sigma(\alpha))$, since σ fixes elements of K; so $\sigma(\alpha)$ lies in L and is integral over $\mathcal{O}_K$, so must be an element of $\mathcal{O}_L$. This means that $\sigma(\mathcal{O}_L) \subseteq \mathcal{O}_L$. The same argument holds for σ^{-1}, i.e. $\sigma^{-1}(\mathcal{O}_L) \subseteq \mathcal{O}_L$, which implies that $\mathcal{O}_L \subseteq \sigma(\mathcal{O}_L)$. Combining both inclusions, we must have that σ fixes $\mathcal{O}_L$ setwise.
We have found that σ fixes both $\mathfrak{p}$ and $\mathcal{O}_L$, so it follows that σ fixes $\mathfrak{p}\mathcal{O}_L$. Define $\nu_i : \mathfrak{p}\mathcal{O}_L \mapsto e_i$ to be the function that gives the ramification index of $\mathfrak{q}_i$. Then

$$\begin{aligned} e_j &= \nu_j\big(\mathfrak{p}\mathcal{O}_L\big) \\ &= \nu_j\big(\sigma\big(\mathfrak{p}\mathcal{O}_L\big)\big) \\ &= \nu_j\big(\sigma\big(\prod_i \mathfrak{q}_i^{e_i}\big)\big) \\ &= \nu_j\big(\prod_i \sigma(\mathfrak{q}_i)^{e_i}\big) \\ &= \nu_j\big(\prod_i \mathfrak{q}_{\pi(i)}^{e_i}\big) \\ &= \nu_j\big(\prod_i \mathfrak{q}_i^{e_{\pi^{-1}(i)}}\big) \\ &= e_{\pi^{-1}(j)}, \end{aligned}$$

where π is the induced permutation on the set $\{1, .., g\}$ of indices of the $\mathfrak{q}_i$. Since G acts transitively on primes lying above $\mathfrak{p}$, the induced permutation is also transitive, and so the e_i are equal for all i.

(ii) As shown in (i), σ fixes $\mathcal{O}_L$, so we obtain an isomorphism $\mathcal{O}_L/\mathfrak{q}_i \xrightarrow{\sim} \mathcal{O}_L/\sigma(\mathfrak{q}_i)$. Thus

$$f_i = \big[\mathcal{O}_L/\mathfrak{q}_i : \mathcal{O}_K/\mathfrak{p}\big] = \big[\mathcal{O}_L/\sigma(\mathfrak{q}_i) : \mathcal{O}_K/\mathfrak{p}\big] = f_{\pi(i)},$$

and again by the transitivity of G on primes lying above $\mathfrak{p}$, we are done.

□

Definition 2. In the above situation, we say an ideal $\mathfrak{p} \subset \mathcal{O}_K$ *ramifies* if $e > 1$, and is *unramified* if $e = 1$. Alternatively, we say that L/K is unramified at $\mathfrak{p}$. If $e = 1$ and $f = 1$, we say $\mathfrak{p}$ *splits completely.*

Definition 3. Let L be a finite, Galois extension of K and $\mathfrak{q}$ a prime ideal of L. Define the *Decomposition Group* of $\mathfrak{q}$ to be the stabilizer of $\mathfrak{q}$ in G,

$$D_{\mathfrak{q}} = \{\sigma \in G : \sigma(\mathfrak{q}) = \mathfrak{q}\}, \tag{2}$$

and the *Inertia Group* of $\mathfrak{q}$ to be

$$I_{\mathfrak{q}} = \{\sigma \in G : \sigma(\alpha) \equiv \alpha \bmod \mathfrak{q}, \text{ for all } \alpha \in \mathcal{O}_L\} \tag{3}$$

Let $\sigma \in D_{\mathfrak{q}}$. Since any element of $D_{\mathfrak{q}}$ fixes $\mathfrak{q}$, and any element of G fixes $\mathcal{O}_L$, σ induces an automorphism on $\mathcal{O}_L/\mathfrak{q}$. Denote this automorphism by $\bar{\sigma}$. Since $\sigma \in G$, σ fixes $\mathcal{O}_k$ and so $\bar{\sigma}$ fixes $\mathcal{O}_K/\mathfrak{p}$. So $\bar{\sigma} \in Gal(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p}) = \bar{G}$. We then obtain a homomorphism $D_{\mathfrak{q}} \to \bar{G}$, given by mapping σ to $\bar{\sigma}$. Finally note that an element in the kernel of this map maps to the identity on $\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p}$, so is the identity on $\mathcal{O}_L/\mathfrak{q}$, and thus is an element of $I_{\mathfrak{q}}$. We have obtained an isomorphism $D_{\mathfrak{q}}/I_{\mathfrak{q}} \xrightarrow{\sim} \bar{G}$.

Proposition 1.

(i) $|D_{\mathfrak{q}}| =$

(ii) $|I_{\mathfrak{q}}| = e$.

Proof.

(i)

(ii) $|\bar{G}| = \left[\mathcal{O}_L/\mathfrak{q} : \mathcal{O}_K/\mathfrak{p}\right] = f$. $\left[\mathrm{G} : D_{\mathfrak{q}}\right] = g$, where g is the number of primes lying above $\mathfrak{p}$.

□

Definition 4. The *Artin Symbol*

1.1.1 Frobenius Elements

Definition 5. Let K be a field of characteristic p. The *Frobenius map* on K is the map

$$F : K \to K,\ a \mapsto a^p \tag{4}$$

If $q = p^r$ for some r, consider the Frobenius map F on $\mathbb{F}_q$. Since $a^p \equiv a$ mod p by Fermat Euler, F fixes $\mathbb{F}_p$. Moreover,

$$\begin{aligned} F(x+y) &= (x+y)^p = x^p + y^p = F(x) + F(y) \text{ (by the binomial theorem)}, \\ F(xy) &= (xy)^p = x^p y^p = F(x)F(y), \\ F(\frac{x}{y}) &= (\frac{x}{y})^p = \frac{x^p}{y^p} = \frac{F(x)}{F(y)}, \text{ and} \\ F(0) &= 0,\ F(1) = 1. \end{aligned}$$

So F is an automorphism on $\mathbb{F}_q$ that fixes $\mathbb{F}_p$, so is an element of $Gal(\mathbb{F}_q/\mathbb{F}_p)$. Finally, F has order r, since if it had order $s < r$, we would have $x^{p^s} = x$, for all $x \in \mathbb{F}_q$. Obviously we also have $x^q = x$, for all $x \in \mathbb{F}_q$. But then $X^{p^s} - X$ would have $q > p^s$ roots. So F has order $r = \left[\mathbb{F}_q : \mathbb{F}_p\right] = |Gal(\mathbb{F}_q/\mathbb{F}_p)|$. We have shown the following:

Proposition 2. *Let $q = p^r$ and $F : \mathbb{F}_q \to \mathbb{F}_q, a \mapsto a^p$. Then F is an automorphism and generates $Gal(\mathbb{F}_q/\mathbb{F}_p)$.*

We can extend this notion to the case when the ground field has order a prime power. If $|K|$ has order p^r and L is a finite extension of K of order p^s, where $r|s$, then the automorphism on L given by

$$a \mapsto a^{p^r}$$

fixes the field K, and has order s/r. Since $Gal(\mathbb{F}_{p^s}/\mathbb{F}_{p^r})$ has order s/r and is cyclic (since it is a subgroup of $Gal(\mathbb{F}_{p^s}/\mathbb{F}_p)$), the Frobenius automorphism of L/K in this case again generates the Galois group.
We now begin to explore the nature of Frobenius automorphisms with regard to $Gal(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p})$: set $Nm(\mathfrak{p}) = |\mathcal{O}_K/\mathfrak{p}|$. Since $\mathcal{O}_K/\mathfrak{p}$ is finite, its order is a power of p, where p is the unique prime lying in $\mathbb{Z} \cap \mathfrak{p}$. Then there exists a unique automorphism on $\mathcal{O}_L/\mathfrak{q}$ given by

$$a \mapsto a^{Nm(\mathfrak{p})}$$

that fixes $\mathcal{O}_K/\mathfrak{p}$ and generates $Gal(\mathcal{O}_L/\mathfrak{q}/\mathcal{O}_K/\mathfrak{p})$. Using the notation we developed above, there exists $\bar{\sigma} \in \bar{G} : \langle \bar{\sigma} \rangle = \bar{G}$. Then, using the isomorphism $D_\mathfrak{q}/I_\mathfrak{q} \tilde{\to} \bar{G}$, there exists a coset of elements, $\sigma I_\mathfrak{q}$, that map to $\bar{\sigma}$, for some $\sigma \in D_\mathfrak{q}$. We can now reformulate the automorphism as

$$\sigma(a) \equiv a^{Nm(\mathfrak{p})} \bmod \mathfrak{q}, \text{ for } a \in \mathcal{O}_L.$$

A representative of this coset is called a *Frobenius element* of G, and denoted $Frob_\mathfrak{q}$. If $\mathfrak{p}$ is unramified, $e = 1$, and so the inertia group is trivial; then the coset of Frobenius elements consists of one unique element.

Proposition 3. *Let $\sigma \in G$ and $\mathfrak{p}$ be unramified. Then $Frob_{\sigma(\mathfrak{q})} = \sigma Frob_\mathfrak{q} \sigma^{-1}$.*

Proof.

$$\begin{aligned} Frob_\mathfrak{q}(a) &\equiv a^{Nm(\mathfrak{p})} \bmod \mathfrak{q} \Leftrightarrow \\ \sigma Frob_\mathfrak{q}(a) &\equiv \sigma(a)^{Nm(\mathfrak{p})} \bmod \sigma(\mathfrak{q}) \Leftrightarrow \\ \sigma Frob_\mathfrak{q}(\sigma^{-1}(a)) &\equiv \sigma(\sigma^{-1}(a))^{Nm(\mathfrak{p})} \bmod \sigma(\mathfrak{q}) \Leftrightarrow \\ \sigma Frob_\mathfrak{q}\sigma^{-1}(a) &\equiv a^{Nm(\mathfrak{p})} \bmod \sigma(\mathfrak{q}) \end{aligned}$$

By definition, we also have $Frob_{\sigma(\mathfrak{q})}(a) \equiv a^{Nm(\mathfrak{p})} \bmod \sigma(\mathfrak{q})$. Since $\mathfrak{p}$ is unramified, $D_\mathfrak{q} \cong \bar{G}$ and the Frobenius element is unique; so we must have $Frob_{\sigma(\mathfrak{q})} = \sigma Frob_\mathfrak{q} \sigma^{-1}$. □

Thus Frobenius elements are all conjugate to one another. If G is abelian, the conjugacy class has size 1 and consists of a single element. We will denote this element $Frob_\mathfrak{p}$.

1.1.2 Cebotarev's Density Theorem

We will need Cebotarev's Density Theorem in order to better understand the conjugacy class of Frobenius elements; in particular, their density in the Galois group follows from Cebotarev's result. To begin working towards a proof, we will let K be a number field, and $\mathcal{P}_K$ the set of finite primes of $\mathcal{O}_K$. Set $\mathcal{S}$ to be a subset of $\mathcal{P}_K$. Then we define the *Dirichlet Density* of $\mathcal{S}$ as

$$\delta(\mathcal{S}) = lim$$

1.2 Modular Forms

This section relies on Serre's *A Course in Arithmetic* and Robert Kurinczuk's lecture notes on modular forms [Kur17].

Let $SL_2(\mathbb{Z})$ act on the upper half of the complex plane, $\mathbb{H}$, as linear fractional transformations, i.e. for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ and $z \in \mathbb{H}$, $\gamma \cdot z = \frac{az+b}{cz+d}$.

Definition 6. A modular form of weight k and level 1 is a function

$$f : \mathbb{H} \to \mathbb{C}$$

that is holomorphic on $\mathbb{H} \cup \{\infty\}$ and satisfies the modular transformation law:

$$f(\gamma \cdot z) = (cz + d)^k f(z),$$

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

Note that by plugging in $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ one can deduce that there are no modular forms of odd weight, negative weight, or weight equal to 2.
We define the Eisenstein series of weight k to be

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

and the normalised Eisenstein series of weight k to be

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

where B_k denotes the kth Bernouilli number, defined by the following expression:

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!}$$

In particular, we have

$$E_4(z) = 1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n,$$

and

$$E_6(z) = 1 - 504 \sum_{n=1}^{\infty} \sigma_5(n) q^n$$

Any modular form of weight k can be written as a homogeneous polynomial in E_4 and E_6, where for any given term ${E_4}^a {E_6}^b$ we have $4a + 6b = k$. We will denote E_4 and E_6 by Q and R, respectively.
If E_2 does not quite satisfy the modular transformation rule, instead obeying

$$E_2(\frac{-1}{z}) = z^2 E_2(z) + \frac{12z}{2i\pi} \tag{5}$$

We will later see that its behaviour changes modulo p, and we will denote E_2 by P.

2 Modular Forms Reduced Mod p

This section closely follows the results of Serre [Jea72] and Swinnerton-Dyer [Pet72], with a little help from Lang's *Introduction to Modular Forms* [Lan01].

Definition 7. A modular form

$$f(z) = \sum_{n=0}^{\infty} a_n q^n$$

is p-integral if the p-adic valuation v_p is greater than or equal to zero when evaluated at the coefficients a_n for all n ≥ 0, i.e. $v_p(a_n) \geq 0$ for all n ≥ 0.

Note: This is equivalent to saying that the denominators of the coefficients are not divisible by p. Thus we can think of p-integral modular forms as being modular forms with coefficients in the ring of rational numbers with denominators coprime to p. Denote this ring by σ.

Definition 8. Let $f(z)$ be a p-integral modular form of weight k. Write

$$\tilde{f}(z) = \sum_{n=0}^{\infty} \tilde{a}_n q^n$$

for the reduction of $f(z)$ modulo p. Denote by M_k the σ-module of p-integral modular forms of weight k, and by $\tilde{M}_k$ the vector space over $\mathbb{F}_p$ of $\tilde{f}$, for f in M_k. The $\mathbb{F}_p$-algebra of modular forms mod p is denoted $\tilde{M}$, and is the direct sum of the $\tilde{M}_k$.

2.1 Derivation on the Space of Modular Forms

Definition 9. If

$$f(z) = \sum_{n=0}^{\infty} a_n q^n$$

is a p-integral modular form of weight k, let

$$\theta f = q\frac{df}{dq} = \sum_{n=0}^{\infty} n a_n q^n$$

Set $\partial f = 12\theta f - kPf$.

Theorem 1.

(i) *Let g and h be modular forms of weights m and n, respectively. Then ∂ satisfies $\partial(g \cdot h) = h \cdot \partial g + g \cdot \partial h$.*

(ii) *If f is a p-integral modular form of weight k, ∂f is a modular form of weight $k + 2$.*

Proof.

(i) $\theta(g \cdot h) = q \cdot \frac{d}{dq}(g \cdot h) = q \cdot (g \cdot \frac{d}{dq}h + h \cdot \frac{d}{dq}g) = g \cdot \theta h + h \cdot \theta g$. Then

$$\begin{aligned}
\partial(g \cdot h) &= 12\theta(g \cdot h) - (m+n)P(g \cdot h) \\
&= 12g \cdot \theta(h) + 12h \cdot \theta(g) - mPgh - nPgh \\
&= g \cdot (12\theta(h) - nPh) + h \cdot (12\theta(g) - mPg) \\
&= g \cdot \partial h + h \cdot \partial g,
\end{aligned}$$

as required.

(ii) Recall P satisfies $P(\frac{-1}{z}) = z^2 P(z) + \frac{12z}{2i\pi}$. We will use this together with

$$f(\frac{-1}{z}) = z^k f(z) \tag{6}$$

to prove the result. Differentiating the left hand side of (6) with respect to z, we get $\frac{d}{dz}f(\frac{-1}{z}) = \frac{1}{z^2}f'(\frac{-1}{z})$. Differentiating the right hand side,

we obtain $\frac{1}{z^2}f'(\frac{-1}{z}) = kz^{k-1}f(z)+z^kf'(z)$, and so $f'(\frac{-1}{z}) = kz^{k+1}f(z)+z^{k+2}f'(z)$. Thus

$$\begin{aligned}
\partial f(\frac{-1}{z}) &= 12\theta f(\frac{-1}{z}) - kP(\frac{-1}{z})f(\frac{-1}{z}) \\
&= \frac{12}{2i\pi}\frac{d}{dz}f(\frac{-1}{z}) - kP(\frac{-1}{z})f(\frac{-1}{z}) \\
&= \frac{12}{2i\pi}(kz^{k+1}f(z) + z^{k+2}f'(z)) - k(z^2P(z) + \frac{12z}{2i\pi})z^kf(z) \\
&= \frac{12}{2i\pi}z^{k+2}f'(z)) - kz^{k+2}P(z)f(z) \\
&= z^{k+2}(12\theta f(z) - kP(z)f(z)) \\
&= z^{k+2}\partial f(z).
\end{aligned}$$

□

Corollary 2. *We have*

$$\partial Q = -4R, \tag{7}$$

$$\partial R = -6Q^2, \tag{8}$$

$$\partial \Delta = 0. \tag{9}$$

Before we prove this, we need a lemma:

Lemma 2. *We have*

$$\theta P = \frac{1}{12}(P^2 - Q) \tag{10}$$

$$\theta Q = \frac{1}{3}(PQ - R) \tag{11}$$

$$\theta R = \frac{1}{2}(PR - Q^2) \tag{12}$$

Proof. We show $\theta P - \frac{1}{12}P^2$ is a modular form of weight 4, and we then compare constant terms. Recall P satisfies (5). Differentiating both sides of this, we obtain

$$P'(\frac{-1}{z}) = 2z^3P(z) + z^4P'(z) + \frac{12z^2}{2i\pi}$$

Using this, we find

$$\begin{aligned}\theta P(\frac{-1}{z}) - \frac{1}{12}P^2(\frac{-1}{z}) &= \frac{1}{2i\pi}\frac{d}{dz}P(\frac{-1}{z}) - \frac{1}{12}(z^2P(z) + \frac{12z}{2i\pi})^2 \\ &= \frac{2z^3}{2i\pi}P(z) + \frac{z^4}{2i\pi}P'(z) - \frac{12z^2}{4\pi^2} - \frac{z^4}{12}P^2(z) \\ &\quad - \frac{2z^3}{2i\pi}P(z) + \frac{12z^2}{4\pi^2} \\ &= z^4\big(\theta P(z) - \frac{1}{12}P^2(z)\big),\end{aligned}$$

and so is modular of weight 4. The space of modular forms of weight 4 is one dimensional and spanned by Q, so $\theta P - \frac{1}{12}P^2$ is a scalar multiple of Q, with the scalar determined by the constant terms. θP is a cusp form and so has zero constant term, and the constant terms of P^2 and Q are both 1, and the result follows.
We know that ∂Q is a modular form of weight 6 and that $\partial Q = 12\theta Q - 4PQ$, so in order to prove (11) we show that the constant terms of $12\theta Q - 4PQ$ and $-4R$ are identical. The constant term of R is 1. θQ is a cusp form and so has zero constant term, and the constant term of PQ is 1. As the space of modular forms of weight 6 is spanned by R, we obtain the result.

Similarly, to show (12) note that $\partial R = 12\theta R - 6PR$ has weight 8. Observe that θR is a cusp form, PR has constant term 1, and that Q^2 has constant term 1. Thus $12\theta R - 6PR$ and Q^2 are both modular forms of weight 8 with identical constant terms, so the result follows. □

Proof of Corollary 2. (7) and (8) follow directly from (11) and (12), and the definition of ∂f.

$\theta\Delta$ is a modular form of weight 14 with constant term 0. As Δ is a cusp form, $P\Delta$ is also a cusp form (of weight 14) and hence also has zero constant term. Hence $\theta\Delta = P\Delta$. Finally, $\partial\Delta = 12\theta\Delta - 12P\Delta = 12(\theta\Delta - P\Delta) = 0$, as required. □

Grade the space of modular forms by weight. Recall there is an isomorphism of graded rings from the space of modular forms to the space of homogeneous polynomials $\mathbb{C}[X, Y]$, where Q maps to X and R maps to Y. The corresponding homogeneous polynomial to $f \in M_k$ will be denoted by $\Phi(X, Y) \in \sigma[X, Y]$. Then $\tilde{\Phi}(X, Y)$ is the polynomial in $\mathbb{F}_p[X, Y]$ obtained by reducing the coefficients of Φ modulo p, and the corresponding polynomial

to $\tilde{f} \in \tilde{M}_k$ is $\tilde{\Phi}(\tilde{X}, \tilde{Y}) \in \mathbb{F}_p[[q]]$. In the context of these polynomials, we will use X and Y interchangeably with Q and R, respectively.

Thus in order to determine the structure of $\tilde{M}$, we must determine the kernel of the map

$$\mathbb{F}_p[Q, R] \to \tilde{M}$$

.

We will denote this kernel by **a**.
The following result will prove highly useful; a proof can be found on pages 384-386 of [BS66]:

Lemma 3. *(Von Staudt) Let B_n denote the Bernouilli numbers, and $p > 3$. Then:*

(i) *If $p - 1|2\nu$, then $pB_{2\nu} \equiv -1$ mod p.*

(ii) *If $p - 1 \nmid 2\nu$, $\frac{B_{2\nu}}{2\nu}$ is p-integral, and $\frac{B_{2\nu}}{2\nu} \equiv \frac{B_{2\nu \bmod p-1}}{2\nu \bmod p-1}$ mod p.*

Until stated otherwise, the following results are all valid for p> 3.
Let A and B be the homogeneous polynomials such that

$$A(Q, R) = E_{p-1} \text{ and } B(Q, R) = E_{p+1}$$

Lemma 4. *A and B are polynomials in $\sigma[Q, R]$.*

Proof. Recall

$$E_{p-1}(z) = 1 - \frac{2p-2}{B_{p-1}} \sum_{n=1}^{\infty} \sigma_{p-2}(n)\ q^n,$$

where B_{p-1} is the p−1th Bernouilli number. With $2\nu =$ p -1, we can apply Lemma 3 (i) to deduce that p divides only the denominator of B_{p-1}; thus $\frac{2p-2}{B_{p-1}}$ is well defined modulo p, and $E_{p-1} \in \sigma[Q, R]$.

Using Lemma 3 (ii), we obtain $\frac{B_{p+1}}{p+1} \equiv \frac{B_2}{2}$ mod p. $B_2 = \frac{1}{6}$, so $\frac{B_{p+1}}{p+1} \equiv \frac{1}{12}$ mod p. As p $\neq$ 2 or 3, $\frac{2p+2}{B_{p+1}}$ is well defined modulo p, and so $E_{p+1} \in \sigma[Q, R]$. □

Lemma 5.

(i) *$\tilde{A}(\tilde{Q}, \tilde{R}) = 1$ and $\tilde{B}(\tilde{Q}, \tilde{R}) = \tilde{P}$.*

(ii) *$\partial\tilde{A} = \tilde{B}$ and $\partial\tilde{B} = -Q\tilde{A}$.*

(iii) *$\tilde{A}$ has no repeated factors and is coprime to $\tilde{B}$.*

Proof.

(i) Note that

$$\begin{aligned} E_{p-1}(z) &= 1 - \frac{2p-2}{B_{p-1}} \sum_{n=1}^{\infty} \sigma_{p-2}(n)\ q^n \\ &\equiv 1 \text{ mod p,} \end{aligned}$$

since p divides the denominator of B_{p-1} (by Lemma 3). Thus $\tilde{A}(\tilde{Q}, \tilde{R}) = \tilde{E}_{p-1} \equiv 1$ mod p. For the second result, note that Fermat-Euler implies that

$$\sigma_p(n) = \sum_{d|n} d^p \equiv \sum_{d|n} d \text{ mod p,}$$

i.e. $\sigma_p(n) \equiv \sigma_1(n)$ mod p. Recall also (from the proof of Lemma 3) that $\frac{B_{p+1}}{p+1} \equiv \frac{1}{12}$ mod p. Thus

$$\begin{aligned} E_{p+1} &= 1 - 2\ \frac{p+1}{B_{p+1}} \sum_{n=1}^{\infty} \sigma_p(n)\ q^n \\ &\equiv 1 - 24 \sum_{n=1}^{\infty} \sigma_1(n)\ q^n \text{ mod p} \\ &\equiv P \text{ mod p,} \end{aligned}$$

and so $\tilde{B}(\tilde{Q}, \tilde{R}) \equiv \tilde{E}_{p+1} \equiv \tilde{P}$.

(ii) By (i), $\theta\tilde{A}(\tilde{Q}, \tilde{R}) = 0$, so we obtain

$$\begin{aligned} \partial\tilde{A}(\tilde{Q}, \tilde{R}) &= 12\theta\tilde{A}(\tilde{Q}, \tilde{R}) - (p-1)\tilde{P}\tilde{A}(\tilde{Q}, \tilde{R}) \\ &= \tilde{P}\tilde{A}(\tilde{Q}, \tilde{R}) \\ &= \tilde{P} \\ &= \tilde{B}(\tilde{Q}, \tilde{R}) \end{aligned}$$

This means the q-expansion of $\partial A - B$ has coefficients divisible by p. Of course, $\partial A - B$ is a modular form of weight p+1, and so $\partial A - B \in$

p$\sigma[Q,R]$ and $\partial\tilde{A} = \tilde{B}$.
For the second result, observe

$$\begin{aligned}\partial\tilde{B}(\tilde{Q},\tilde{R}) &= 12\theta\tilde{B}(\tilde{Q},\tilde{R}) - (p+1)\tilde{P}\tilde{B}(\tilde{Q},\tilde{R}) \\ &= 12\theta\tilde{B}(\tilde{Q},\tilde{R}) - \tilde{P}\tilde{B}(\tilde{Q},\tilde{R}) \\ &= 12\theta\tilde{P} - \tilde{P}^2 \\ &= -\tilde{Q},\end{aligned}$$

by (10). Similarly to before, this means that p divides the coefficients of the q-expansion of $\partial B + QA$, which is a modular form (of weight $p+3$) and so has coefficients in $p\sigma[Q,R]$. Thus $\partial\tilde{B} = -Q\tilde{A}$.

(iii)

□

Note: (i) means that P becomes a modular form of weight p +1 modulo p.

Theorem 2. *The ideal* $\boldsymbol{a}$ *is equal to the principal ideal generated by* $\tilde{A} - 1$.

Proof. Recall $\mathbf{a}$ is the kernel of

$$\mathbb{F}_p[Q,R] \to \tilde{M},$$

so can also be thought of as the kernel of

$$\mathbb{F}_p[Q,R] \to \mathbb{F}_p[[q]],$$

given by replacing Q and R with $\tilde{Q}$ and $\tilde{R}$ respectively, since any modular form has a power series expansion in q. As in Lemma 4, $\tilde{A}(\tilde{Q},\tilde{R}) = 1$, hence $\tilde{A}-1 \in \mathbf{a}$. $\mathbf{a}$ is prime since $\mathbb{F}_p[[q]]$ is an integral domain. Let $\mathfrak{m}$ be a maximal ideal containing $\mathbf{a}$. We now have the chain of ideals

$$0 \subseteq (\tilde{A}-1) \subseteq \mathbf{a} \subseteq \mathfrak{m} \tag{13}$$

If $(\tilde{A}-1)$ is a prime ideal, we have obtained a chain of prime ideals of length three. They cannot all be prime, as this contradicts the Krull dimension of $\mathbb{F}_p[X,Y]$, which is two. Furthermore, $\mathbf{a}$ is not maximal, since the image is $\mathbb{F}_p[\tilde{Q},\tilde{R}]$, which is not a field. To complete the proof, we prove that $(\tilde{A}-1)$ is prime (equivalent to $\tilde{A}-1$ being irreducible) which will imply $(\tilde{A}-1) = \mathbf{a}$.
□

The final result of this section will be the relation between mod p modular forms and their weights:

Corollary 3. *(Kummer's Congruence) Let f and g be p-integral modular forms of weights k and l, respectively. If $f \equiv g$ mod p and $f \not\equiv 0$ mod p, then $k \equiv l$ mod (p−1).*

Proof. Without loss of generality, let $k \leq l$. If $f \equiv g$ mod p, then $f(\frac{-1}{z}) \equiv g(\frac{-1}{z})$ mod p, i.e.

$$z^k\big(f(z) - z^{l-k}g(z)\big) \equiv 0 \text{ mod p},$$

so we must have $f(z) - z^{l-k}g(z) \equiv 0$ mod p. Since $f \equiv g$ mod p, we must have $z^{l-k} \equiv 1$ mod p, and by Fermat-Euler we obtain $k - l = (\text{p}-1)\text{m}$, for some scalar m, i.e. $k \equiv l$ mod (p−1). □

2.2 Filtration

Let $\tilde{f}$ be a graded element in $\tilde{M}$, i.e. a linear combination of elements of various $\tilde{M}_k$ where the k are all congruent modulo p−1 (c.f. Corollary 3). We can multiply the summands by appropriate powers of $\tilde{A}$ in order to get every summand in the same $\tilde{M}_k$, so that $\tilde{f}$ belongs to a single $\tilde{M}_k$.

Definition 10. Let f be a graded element of $\tilde{M}$. Define the filtration of $\tilde{f}$ to be the lowest k such that $\tilde{f} \in \tilde{M}_k$. We denote the filtration by $w(\tilde{f})$.

Note: We can equivalently say that the filtration of a p-integral modular form f is the lowest weight k such that there exists a modular form g for which we have $f \equiv g$ mod p.

Lemma 6.

(i) *If f is a p-integral modular form of weight k, with $f = \phi(Q, R)$ for $\phi \in \sigma[Q, R]$ and $f \not\equiv 0$ mod p, the $w(\tilde{f}) < k \iff \tilde{A} | \tilde{\phi}$.*

(ii) *If $\tilde{f}$ is graded in $\tilde{M}$, we have $w(\theta\tilde{f}) \leq w(\tilde{f}) + p + 1$, with equality if and only if $w(\tilde{f}) \not\equiv 0$ mod p.*

Proof.

(i) Clearly if $\tilde{A} \nmid \tilde{\phi}$, then $w(\tilde{f})$ cannot be less than k, since in order to obtain the isobaric polynomial of degree k, ϕ, we have multiplied various

summands by $\tilde{A}$ to get every summand into the same $\tilde{M}_k$. Thus, if there exists a summand not divisible by $\tilde{A}$, the filtration of $\tilde{f}$ cannot be less than the degree of that summand, which is at least k.
Conversely, let $\tilde{A}|\tilde{\phi}$, and suppose $w(\tilde{f}) = k$. Then we must have $\tilde{\phi} = \tilde{A}\tilde{\psi}$, for some isobaric polynomial ψ corresponding to some modular form g of weight less than k. This implies

$$\tilde{f} = \tilde{\phi}(\tilde{Q}, \tilde{R}) = \tilde{A}(\tilde{Q}, \tilde{R})\tilde{\psi}(\tilde{Q}, \tilde{R}) = \tilde{\psi}(\tilde{Q}, \tilde{R}) = \tilde{g},$$

which contradicts $w(\tilde{f}) = k$.

(ii) Let $w(f) = k$ and f be as in (i). We have $\partial\tilde{\phi}(\tilde{Q}, \tilde{R}) = 12\theta\tilde{f} - k\tilde{P}\tilde{f}$, which is equivalent to

$$12\theta\tilde{f} = \tilde{A}(\tilde{Q}, \tilde{R})\partial\tilde{\phi}(\tilde{Q}, \tilde{R}) + k\tilde{B}(\tilde{Q}, \tilde{R})\tilde{f},$$

using the facts that $\tilde{A}(\tilde{Q}, \tilde{R}) = 1$ and $\tilde{B}(\tilde{Q}, \tilde{R}) = \tilde{P}$. Hence $12\theta\tilde{f}$ is the image of $\tilde{A}\partial\tilde{\phi}+k\tilde{B}\tilde{\phi}$ in $\tilde{M}$. Observe that both summands have filtration less than or equal to $w(\tilde{f})+p+1$: $\tilde{A}$ has filtration p-1 and $\partial\tilde{\phi}$ filtration $w(\tilde{f}) + 2$, and $\tilde{B}$ filtration p$+1$ and f filtration $w(\tilde{f})$. Since $w(\tilde{f}) = k$, we have by (i) that $\tilde{A} \nmid \tilde{\phi}$. Furthermore, Lemma 5 implies that $\tilde{A} \nmid \tilde{B}$. Combining these two results, we find that $w(\theta\tilde{f}) = w(\tilde{f}) + p + 1$ if and only if $\tilde{A} \nmid (\tilde{A}\partial\tilde{\phi} + k\tilde{B}\tilde{\phi})$ if and only if $p \nmid k$, i.e. if and only if $w(\tilde{f}) \not\equiv 0$ mod p.

□

We now deal with the cases of p$=$ 2 and p$=$ 3.

Theorem 3. *If $p = 2$ or $p = 3$, we have*

(i) $\tilde{P} = \tilde{Q} = \tilde{R} = 1$.

(ii) $\tilde{M} = \mathbb{F}_p[\tilde{\Delta}]$.

(iii) $\partial\tilde{M} = 0$.

Proof.

(i) 24, 240, and 504 are all divisible by both 2 and 3, and the q-expansions of P, Q, and R all begin with 1, and the result follows.

(ii) Δ can be written

$$\Delta = q\prod_{n=1}^{\infty}(1-q^n)^{24},$$

and using the binomial theorem we can see that $\tilde{\Delta} = q$.

(iii) As in the proof of Corollary 2, $\partial\Delta = 0$. Part (ii) of this theorem implies the result.

□

3 Galois Representations Attached to Modular Forms

3.1 Introducing ρ_l

In this section, we will take the existence of a Galois representation attached to the coefficients of a modular form as given. We will first explore the possible images of the representation, and then go on to use our theory of mod p modular forms to examine which primes are 'exceptional' for a given modular form. Before we proceed, we need to develop some notation. We will let l be a prime number, and let K_l be the maximal algebraic extension of $\mathbb{Q}$ ramified only at l. Further, we will take K_l^{ab} to be the maximal subfield of K_l which is abelian over $\mathbb{Q}$. Frob(p) will denote the conjugacy class of Frobenius elements in $\mathrm{Gal}(K_l/\mathbb{Q})$.

Theorem 4.

(i) *There exists an isomorphism* $Gal\big(K_l^{ab}/\mathbb{Q}\big) \cong \mathbb{Z}_l^*$, *where* $\mathbb{Z}_l^*$ *is the group of l-adic units. This in turn induces a character*

$$\chi_l : Gal\big(K_l/\mathbb{Q}\big) \to Gal\big(K_l^{ab}/\mathbb{Q}\big) \xrightarrow{\sim} \mathbb{Z}_l^*,$$

such that

$$\chi\big(Frob_p\big) = p, \tag{14}$$

for all $p \neq l$.

(ii) *Let* $f = \sum_{n=1}^{\infty} a_n q^n$ *be a normalized cusp form of weight* k *with integer coefficients, and dirichlet series*

$$\sum_{n=1}^{\infty} a_n n^{-s} = \prod_{i=1}^{\infty} \big(1 - a_p p^{-s} + p^{k-1-2s}\big)^{-1}$$

Then there exists a continuous homomorphism

$$\rho_l : Gal\big(K_l/\mathbb{Q}\big) \to GL_2\big(\mathbb{Z}_l\big),$$

that depends on f *such that* $\rho_l\big(Frob_p\big)$ *has characteristic polynomial*

$$x^2 - a_p x + p^{k-1} \tag{15}$$

for each $p \neq l$.

Proof. We will leave these results unproven. □

Note that (15) implies that the trace of $\rho_l\big(Frob_p\big)$ is a_p, and that the norm of $\rho_l\big(Frob_p\big)$ is p^{k-1}. More generally than this, we have

$$det \circ \rho_l = \chi_l^{k-1}$$

Since χ_l maps into $\mathbb{Z}_l^*$, the image of $det \circ \rho_l$ is $(k-1)$th powers in $\mathbb{Z}_l^*$. Denote by $\tilde{\rho}_l$ the map

$$\tilde{\rho}_l : Gal\big(K_l/\mathbb{Q}\big) \to GL_2\big(\mathbb{Z}_l\big) \to GL_2\big(\mathbb{F}_l\big), \tag{16}$$

induced by reducing ρ_l mod l. More generally, we will use a tilde to denote reduction mod l.

Lemma 7. *The set of matrices*

$$H_2 = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big) \,\middle|\, \begin{array}{l} a \equiv d \equiv 1 \ mod\ l \\ b \equiv c \equiv 0 \ mod\ l \end{array} \right\}$$

is generated by $I + lu$, *for* $u \in U \coloneqq \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \right\}$.

Proof. Label the three matrices in U as u_1, u_2, and u_3 respectively. Note that each $I + lu_i$ is an element of H_2: all three $I + lu_i$ reduce to the identity mod l, and $I + lu_1$ and $I + lu_2$ clearly have determinant 1. To see this for $I + lu_3$, observe that

$$\left| \begin{pmatrix} 1+l & -l \\ l & 1-l \end{pmatrix} \right| = (1+l)(1-l) + l^2 = 1 - l^2 + l^2 = 1.$$

The claim is that

$$H_2 = \langle I + lu_1, I + lu_2, I + lu_3 \rangle = \left\langle \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ l & 1 \end{pmatrix}, \begin{pmatrix} 1+l & -l \\ l & 1-l \end{pmatrix} \right\rangle$$

Since each $I + lu_i$ is in H_2, we have $\langle I + lu_1, I + lu_2, I + lu_3 \rangle \subseteq H_2$. To conclude the proof, we show that H_2 and $\langle I + lu_1, I + lu_2, I + lu_3 \rangle$ have the same cardinality.
We can think of H_2 as the kernel of $SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big) \to SL_2\big(\mathbb{Z}/l\mathbb{Z}\big)$, and so obtain

$$\Big[SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big) : SL_2\big(\mathbb{Z}/l\mathbb{Z}\big)\Big] = |H_2|$$

Using the formula $|\mathrm{SL}_2\left(\mathbb{Z}/l^e\mathbb{Z}\right)| = l^{3e}\left(1 - \frac{1}{l^2}\right)$, we find that $|SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big)| = l^6(1 - 1/l^2) = l^4(l^2 - 1)$ and $|SL_2\big(\mathbb{Z}/l\mathbb{Z}\big)| = l^3(1 - 1/l^2) = l(l^2 - 1)$, so that

$$|H_2| = \Big[SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big) : SL_2\big(\mathbb{Z}/l\mathbb{Z}\big)\Big] = l^4(l^2 - 1)/l(l^2 - 1) = l^3$$

It is clear that $I + lu_1$ and $I + lu_2$ both have order l; it remains to check that $I + lu_3$ has order l, and we will be done. To this end, observe that

$$\begin{pmatrix} 1+l & -l \\ l & 1-l \end{pmatrix}^n = \begin{pmatrix} 1+nl & -ln \\ ln & 1-nl \end{pmatrix}$$

so $I + lu_3$ has order l in $SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big)$. □

Theorem 5. *Let $l > 3$ and G be a subgroup $GL_2\big(\mathbb{Z}_l\big)$ closed in the l-adic topology. If $\tilde{G}$ contains $SL_2\big(\mathbb{F}_l\big)$, then G contains $SL_2\big(\mathbb{Z}_l\big)$.*

Proof. Let G_n denote the image of G in $GL_2\big(\mathbb{Z}/l^n\mathbb{Z}\big)$. We need to prove that $SL_2\big(\mathbb{Z}/l^n\mathbb{Z}\big) \subset G_n$ for all $n > 0$. We will rely on two inductive arguments: firstly, we will show that G_n contains the kernel of the map

$$\varphi : SL_2\big(\mathbb{Z}/l^n\mathbb{Z}\big) \to SL_2\big(\mathbb{Z}/l^{n-1}\mathbb{Z}\big),$$

for $n \geq 2$. After this, we will use this result and induction to prove the theorem.
Denote the kernel of φ by H_n and let $n = 2$, so

$$\begin{aligned} H_2 &= \ker\Big(SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big) \to SL_2\big(\mathbb{Z}/l\mathbb{Z}\big) \Big) \\ &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2\big(\mathbb{Z}/l^2\mathbb{Z}\big) \,\middle|\, \begin{matrix} a \equiv d \equiv 1 \bmod l \\ b \equiv c \equiv 0 \bmod l \end{matrix} \right\}, \end{aligned}$$

which by Lemma 7 is equal to $\langle I + lu_1, I + lu_2, I + lu_3 \rangle$, with u_i as in the lemma. We will show that G_2 contains the images of each $I + lu_i$. Since $I + u_i$ is in $SL_2\big(\mathbb{Z}\big)$ and $\tilde{G}$ contains $SL_2\big(\mathbb{F}_l\big)$, there exists a matrix $b \in G$ such that $b \equiv I + u_i \bmod l$. Alternatively, we can write $b = I + u_i + lv$, for some matrix v with entries in $\mathbb{Z}_l$. We can then see that

$$\begin{aligned} b^l &= (I + u_i + lv)^l = I + l(u_i + lv) + ... + (u_i + lv)^l \\ &\equiv I + lu_i \bmod l^2, \end{aligned}$$

since each term (except for $I + lu_i$) has either a factor of l^2 or u_i^2, and $u_i^2 = 0$ for each i. Then $H_2 \subset G_2$.
Now assume $H_{n-1} \subset G_{n-1}$. Take an element of H_n, say $I + l^{n-1}w$, where w has entries in $\mathbb{Z}_l$; then $I + l^{n-2}w \bmod l^{n-1}$ is in H_{n-1}. By induction we have $I + l^{n-2}w \bmod l^{n-1} \in G_{n-1}$. Similarly to before, there exists an element $c \in G$ such that $c = I + l^{n-2}w + l^{n-1}x$, where x has entries in $\mathbb{Z}_l$. Again we have

$$\begin{aligned} c^l &= (I + l^{n-2}w + l^{n-1}x)^l = I + l(l^{n-2}w + l^{n-1}x) + ... + (l^{n-2}w + l^{n-1}x)^l \\ &= I + l^{n-1}w + l^n x + ... + (l^{n-2}w + l^{n-1}x)^l \\ &\equiv I + l^{n-1}w \bmod l^n. \end{aligned}$$

So $H_n \subset G_n$, as required.
To finish the proof, we proceed by induction, noting that we have assumed in the statement of the theorem that G_1 contains $SL_2\big(\mathbb{F}_l\big)$, so the $l = 1$ case holds. Fix n and suppose that $SL_2\big(\mathbb{Z}/l^{n-1}\mathbb{Z}\big) \subset G_{n-1}$. Set K to be the kernel of $G_n \to G_{n-1}$. We obtain the following diagram:

$$\begin{array}{ccccc} H_n & \overset{\varphi_1}{\hookrightarrow} & SL_2\big(\mathbb{Z}/l^n\mathbb{Z}\big) & \overset{\varphi_2}{\twoheadrightarrow} & SL_2\big(\mathbb{Z}/l^{n-1}\mathbb{Z}\big) \\ \big\downarrow \phi_1 & & \downarrow f & & \big\downarrow \phi_2 \\ \mathrm{K} & \overset{\psi_1}{\hookrightarrow} & G_n & \overset{\psi_2}{\twoheadrightarrow} & G_{n-1}, \end{array}$$

which commutes. We need to show that the middle vertical map, f, is injective. Let a be an element in the kernel of f. Then $f(a) = I$. As the diagram commutes, we have

$$I = \psi_2(f(a)) = \phi_2(\varphi_2(a)),$$

and by the injectivity of ϕ_2, we must have $\varphi_2(a) = I$. So a is in the kernel of φ_2, which is precisely H_n. Using the commutativity of the diagram again, we have

$$\psi_1(\phi_1(a)) = f(\varphi_1(a)) = I,$$

and since both ψ_1 and ϕ_1 are injective, we find that $a = I$ and f is injective, proving the theorem. □

Note: this theorem implies that in order to determine

Definition 11. A prime number l is an *exceptional* prime for the cusp form f if the image of ρ_l does not contain $SL_2\big(\mathbb{Z}_l\big)$.

Note: In light of this definition, we can rewrite Theorem 5 as the following: If $l > 3$, l is exceptional for f if and only if the image of $\tilde{\rho}_l$ does not contain $SL_2\big(\mathbb{F}_l\big)$.

Recall that $\rho_l : Gal\big(K_l/\mathbb{Q}\big) \to GL_2\big(\mathbb{Z}_l\big)$, and that $det \circ \rho_l = \chi_l^{k-1}$. Thus if l is *not* exceptional for f, the image of ρ_l is the preimage of $(\mathbb{Z}_l^*)^{k-1}$ in $GL_2\big(\mathbb{Z}_l\big)$, under the determinant map. The theorem tells us that it suffices to look at the image of $\tilde{\rho}_l$ instead of ρ_l. We will hunt for exceptional primes; we start by finding some subgroups of $GL_2\big(\mathbb{F}_l\big)$ which do not contain $SL_2\big(\mathbb{F}_l\big)$.

3.2 The Images of $\tilde{\rho}_l$

Let V be a 2 dimensional vector space over $\mathbb{F}_l$. We begin by defining two important subgroups of $GL_2\big(\mathbb{F}_l\big)$:

Definition 12. A *Borel* subgroup is any subgroup of $GL_2\big(\mathbb{F}_l\big)$ conjugate to the subgroup of non-singular (i.e. determinant non-zero) upper triangular matrices.

Remark: there is a bijection between Borel subgroups and one-dimensional subspaces of V, where for each one-dimensional subspace there exists a unique Borel subgroup, made up of matrices with the subspace as an eigenspace.

Definition 13. A *Cartan* subgroup is a maximal semi-simple commutative subgroup. There are two kinds:
1. A *split* Cartan subgroup is a subgroup conjugate to the group of non-singular diagonal matrices. Thus there is a bijection between split Cartan subgroups and unordered pairs of distinct one-dimensional subspaces of V, where a pair of subspaces corresponds to the set of matrices with those both subspaces as eigenspaces.
2. A *non-split* Cartan subgroup is defined as follows: take a quadratic extension of $\mathbb{F}_l$ and extend V by it. Call this vector space $\mathrm{V}^{(2)}$. Set w' to be a one-dimensional subspace of $\mathrm{V}^{(2)}$ not induced by a one-dimensional subspace of V, and w'' to be the conjugate of w' over $\mathbb{F}_l$. Then the non-split Cartan subgroup corresponding to w' **or** w'' is the set of elements which have both w' **and** w'' as eigenspaces.

We make three remarks on the definitions:
1. A split Cartan subgroup is isomorphic to $\mathbb{Z}/(l-1)\mathbb{Z} \times \mathbb{Z}/(l-1)\mathbb{Z}$. This is because its elements are non-zero, diagonal, and two-dimensional, so it is characterised by two independent non-zero elements in $\mathbb{F}_l$.
2. An element of a non-split Cartan subgroup is determined by its eigenvalue with respect only to w' (since w'' is its conjugate). We then find (?) that a non-split Cartan subgroup is isomorphic to $\mathbb{F}_{l^2}^{\times}$.
3. As conjugation by an element of the normalizer of a Cartan subgroup fixes the Cartan subgroup, it can have only two effects on subgroup's eigenspaces: it either fixes the eigenspaces, or swaps them. If it fixes the eigenspaces, it must already be in the Cartan subgroup. From this it follows (?) that a Cartan subgroup has index two in its normaliser.

Proposition 4. *$SL_2(\mathbb{F}_l)$ is generated by the matrices $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix}$.*

Proof. $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix}$ are clearly elements of $SL_2(\mathbb{F}_l)$. Multiplying a matrix by $\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix}$ on the right or the left is equivalent to performing elementary row and column operations on the matrix. It is known that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ generate $SL_2(\mathbb{Z})$. So we show elementary row and column operations transform $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ into $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and we will be done.

But this is clear, since subtracting the left column of $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ from its right column yields the result. □

Lemma 8. *Let G be a subgroup of $GL_2(\mathbb{F}_l)$ and H be the image of G in $PGL_2(\mathbb{F}_l)$. If l divides $|G|$, then G is contained in a Borel subgroup of $GL_2(\mathbb{F}_l)$, or $SL_2(\mathbb{F}_l) \subset G$. If $l \nmid |G|$, one of the following holds:*

(i) *H is cyclic and G is contained in a Cartan subgroup,*

(ii) *H is dihedral, G is contained in the normaliser of a Cartan subgroup, but not in the Cartan subgroup, and l is odd, or*

(iii) *$H \cong A_4, S_4,$ or A_5.*

Proof. Let l divide $|G|$. Pick an element of order l in G, say a. If a had two linearly independent eigenvectors, it is diagonalizable, and a diagonal matrix over $\mathbb{F}_l$ does not have order l. So there is a unique one-dimensional subspace of V, say W, which is an eigenspace for a. If every element of G has W as an eigenspace, G is contained within the same Borel subgroup as a. Suppose there exists an element of G which does not fix W, i.e. maps W to some other one-dimensional subspace W'. Call this element b. Observe that bab^{-1} is an element of G of order l. Moreover, since b^{-1} maps W' to W and a fixes W, W' is an eigenspace for bab^{-1}, and so is the unique eigenspace for bab^{-1} (for order reasons). If W and W' are generated by vectors u and v, respectively, fix u and v as a basis for V. We can then write

$$a = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}, \; bab^{-1} = \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix},$$

for some c and d in $\mathbb{F}_l \setminus \{0\}$. By Proposition 4, these matrices generate $SL_2(\mathbb{F}_l)$, and so G contains $SL_2(\mathbb{F}_l)$.
Now suppose that $l \nmid |G|$, so that $l \nmid |H|$. This implies that every element of H is semi-simple, and so every (non-identity) element has two eigenvectors over the algebraic closure of $\mathbb{F}_l(?)$. If two elements of H share an eigenvector, they have both eigenvectors in common: otherwise, suppose a and b are elements of H with only one eigenvector in common. We can find a basis in which a and b have the form

$$a = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}, \; b = \begin{pmatrix} \alpha & \beta \\ 0 & \delta \end{pmatrix},$$

where $c, d, \alpha, \beta,$ and δ are non-zero. Then the commutator

$$[a, b] = a^{-1}b^{-1}ab = \begin{pmatrix} 1 & \alpha^{-1}\beta(1 - c^{-1}d) \\ 0 & 1 \end{pmatrix}$$

is non-identity since $c \neq d$, and so has order l, contradicting our assumption on the order of H. H fixes its set of eigenvectors(?), so let $\xi_1, ..., \xi_\nu$ be representatives for the orbits of H on its set of eigenvectors, and for each ξ_i set μ_i to be the size of the stabilizer of ξ_i in H. Let $|H| = h$. By the Orbit-Stabilizer Lemma, we have $|O(\xi_i)| = |H|/|Stab_H(\xi_i)| = h/\mu_i$. Each non-identity element of H has two eigenvectors, so we obtain $2(|H|-1) = 2|H|-2$ pairs of a non-trivial element of H with one of its eigenvectors. We can also calculate this as follows: pick an orbit representative, ξ_i. Pair this representative up with the non-trivial elements fixing it - there are $\mu_i - 1$ of these. Then, since the orbits partition the group, picking a representative of each orbit and summing the (?) gives us the following formula:

$$2h - 2 = h(\mu_1 - 1)/\mu_1 + ... + h(\mu_\nu - 1)/\mu_\nu,$$

which can be rewritten as

$$2(1 - h^{-1}) = (1 - \mu_i^{-1}) + ... + (1 - \mu_\nu^{-1}).$$

Suppose $\nu = 4$. Then since $\mu_i > 1, \sum_{n=1}^{4} \mu_i^{-1} \leq 2$. But $h^{-1} \neq 0$, so we must have $\nu \leq 3$. If we only have two orbits of eigenvectors, $\nu = 2$ and so we must have $h = \mu_1 = \mu_2$. Let $\nu = 3$. If h is odd, then μ_i is also odd, since $\mu_i | h$. Then $\sum_{n=1}^{3} \mu_i^{-1} \leq 1$, which implies $h^{-1} \leq 0$, so h must be even. Now let h be even. Then clearly the solutions have the form $\mu_1 = \mu_2 = 2, \mu_3 = \frac{h}{2}$. We also get the following three cases:

(i) $h = 12,\ \mu_1 = 2,\ \mu_2 = \mu_3 = 3,$

(ii) $h = 24,\ \mu_1 = 2,\ \mu_2 = 3,\ \mu_3 = 4,$

(iii) $h = 60,\ \mu_1 = 2,\ \mu_2 = 3,\ \mu_3 = 5.$

We now associate each of the cases to a subgroup of $PGL_2(\mathbb{F}_l)$.
When $\nu = 2$, there are only two orbits of eigenvectors; since every element of H has two eigenvectors, in this case every element has the same eigenvectors, which means that the matrices form a cyclic group. By the correspondence between Cartan subgroups and eigenvectors, every element of G must lie in

the same Cartan subgroup of $GL_2(\mathbb{F}_l)$.
Let $\nu = 3$ and assume we are in the general case of H even. If $H > 4$, the Orbit-Stabiliser Lemma implies that the orbit of ξ_3 has size 2. Thus H has a cyclic subgroup, say H_0, of index 2, so H_0 is normal. The preimage of H_0 in $GL_2(\mathbb{F}_l)$ must be contained in a Cartan subgroup, since its elements are semi-simple and commute. The remaining elements of G must interchange the two eigenvectors of the orbit of ξ_3, since otherwise they would map to H_0. One can then see that conjugation of H_0 by one of these elements fixes H_0; so G is contained in the normaliser of a Cartan subgroup, but not in the Cartan subgroup itself. [pretty sure this works for h=4; sd says need a 'similar argument']
In case (i) above, H permutes the $|O(\xi_3)| = 12/3 = 4$ elements of the orbit of ξ_3. This action is faithful, since each non-identity element has only two eigenvectors, so cannot fix all elements of $O(\xi_3)$. This means H injects into S_4 and has order 12, so must be isomorphic to A_4.
In case (ii), no element of order three can be in the stabiliser of a representative of the orbits of ξ_1 or ξ_3, since their orders are powers of two. So all elements of order three must have their eigenvectors in the orbit of ξ_2. $\mu_2 = 3$, so two non-identity elements fix each element of $O(\xi_2)$, and since two elements sharing an eigenvector implies they share both eigenvectors, we can pair up each element of $O(\xi_2)$ with the other eigenvector fixed by the elements that fix that eigenvector, resulting in four pairs. If a non-trivial element of H, say α, fixed all four pairs, it would have order two because(?). Then since the $O(\xi_2)$ is the collection of eigenvectors of elements of H of order 3, no element of $O(\xi_2)$ is an eigenvector of α, so it must interchange the elements of each pair. Then it would commute with the elements in the stabiliser of ξ_2, which have order 3. Call one of these stabilising elements a.Then $a\alpha \in H$ would have order 6 - but H has no elements of order 6, so no such α exists. We have found that H acting as permutations on the four pairs has trivial kernel, and so $H \cong S_4$.
In case (iii), each stabiliser has prime order, so each element of H has prime order. Since two eigenvectors associated to elements of the same order are equivalent under H(?), we must have that any two cyclic subgroups of the same order are conjugate. Thus any proper normal subgroup contains all the elements of a given order. By applying the Sylow theorems, we deduce that there are 24 elements of order 5, 20 elements of order 3, and 15 elements of order 2. Since the order of a subgroup divides the order of a group, counting the possible orders of normal subgroups yields only the trivial subgroup.

Thus H is simple, and must be isomorphic to A_5, which is the only simple group of order 60. $\square$

Corollary 4. *Let ρ_l be a continuous homomorphism $Gal(K_l/\mathbb{Q}) \to GL_2(\mathbb{Z}_l)$ such that $det \circ \rho_l = \chi_l^{k-1}$ for some even k. Let $G \subset GL_2(\mathbb{F}_l)$ be the image of $\tilde{\rho}_l$ and let H be the image of G in $PGL_2(\mathbb{F}_l)$. If G does not contain $SL_2(\mathbb{F}_l)$, then either*

(i) *G is contained in a Borel subgroup of $GL_2(\mathbb{F}_l)$, or*

(ii) *G is contained in the normalizer of a Cartan subgroup, but not in the Cartan subgroup itself, or*

(iii) *$H \cong S_4$.*

Proof. A subgroup of a split Cartan subgroup with eigenspaces W_1 and W_2, say, is a subgroup of the Borel subgroups corresponding to either W_1 or W_2. If we deal with the case of a non-split Cartan subgroup, and show that $H \ncong A_4$ or A_5, then by the above lemma we will be done. Let C be a non-split Cartan subgroup, so $C \cong \mathbb{Z}/(l^2-1)\mathbb{Z}$. Since C is commutative, the homomorphism $Gal\big(K_l/\mathbb{Q}\big) \to C$ must factor through $Gal\big(K_l^{ab}/\mathbb{Q}\big)$. Recall from Theorem 4 that $Gal\big(K_l^{ab}/\mathbb{Q}\big) \cong \mathbb{Z}_l^*$, which has order prime to l. Combining this with $C \cong \mathbb{Z}/(l^2-1)\mathbb{Z}$, we find that the order of $\mathbb{Z}_l^*$ must divide $l-1$, and so its image must lie in the set of scalar matrices, and hence in a Borel subgroup. Thus the non-split case is covered by (i).
Now let $l > 2$. We obtain the following commutative diagram

$$\begin{array}{ccccc} Gal\big(K_l/\mathbb{Q}\big) & \to & G & \stackrel{det}{\to} & \mathbb{F}_l^* \\ & & \downarrow & & \downarrow \\ & & H & \to & \mathbb{F}_l^*/\big(\mathbb{F}_l^*\big)^2 \end{array}$$

We assumed that the image of G is $(k-1)$th powers in $\mathbb{F}_l^*$ and that k is even. Then H surjects onto $\mathbb{F}_l^*/\big(\mathbb{F}_l^*\big)^2$, since the image of H maps to elements of $\mathbb{F}_l^*$ that are not even powers. Since there are $\frac{l-1}{2}$ quadratic residues mod l, we have $\mathbb{F}_l^*/\big(\mathbb{F}_l^*\big)^2 \cong \mathbb{Z}/2\mathbb{Z}$, which implies H has a subgroup of order 2 - but neither A_4 nor A_5 have such a subgroup. $\square$

Corollary 5. *Let $f = \sum_{n=1}^{\infty} a_n q^n$ be a cusp form of weight k with $a_1 = 1$ and integer coefficients, with a Dirichlet series that has an Euler product, as in Theorem 4. Also, let ρ_l be as in the theorem. If the image of $\tilde{\rho}_l$ does*

not contain $SL_2\big(\mathbb{F}_l\big)$ *(so* l *is an exceptional prime for* f*), then we have the following congruences for the coefficients of* f*:*

(i) $\exists m \in \mathbb{Z}$ *such that*

$$a_n = n^m \sigma_{k-1-2m}(n) \ mod\ l,$$

for all n *prime to* l*,*

(ii) $a_n \equiv 0$ *mod* l*, whenever* n *is a quadratic non-residue mod* l*, and*

(iii) $p^{1-k}a_p^2 \equiv$ *0, 1, 2, or 4 mod* l *for all primes* $p \neq l$*.*

Each case follows from the corresponding section of the preceding corollary.

Proof.

(i) Without loss of generality, assume that the Borel subgroup involved consists of the non-singular upper triangular matrices, so that for a general element σ of $Gal\big(K_l/\mathbb{Q}\big)$ we have

$$\tilde{\rho}_l(\sigma) = \begin{pmatrix} \alpha(\sigma) & \beta(\sigma) \\ 0 & \delta(\sigma) \end{pmatrix}$$

α is a continuous homomorphism $Gal\big(K_l/\mathbb{Q}\big) \to \mathbb{F}_l^*$, so we must have $\alpha = \tilde{\chi}_l^m$ for some $m \in \mathbb{Z}$. Furthermore, note that

$$\alpha\delta = det \circ \tilde{\rho}_l(\sigma) = \tilde{\chi}_l^{k-1}$$

by Theorem 4, which implies that $\delta = \tilde{\chi}_l^{k-1-m}$. Plugging in $Frob_p$ for σ and using (14) and (15), we see that

$$\begin{aligned} \tilde{a}_p &= tr(\tilde{\rho}_l(Frob_p)) = \alpha(Frob_p) + \delta(Frob_p) \\ &= \tilde{\chi}_l^m(Frob_p) + \tilde{\chi}_l^{k-1-m}(Frob_p) \\ &= p^m + p^{k-1-m}, \end{aligned} \tag{17}$$

i.e. $a_p \equiv p^m + p^{k-1-m} \mod l$ for $p \neq l$. The result follows from the relation between the Dirichlet series and the Euler product.

(ii) Since every element of $GL_2(\mathbb{F}_2)$ is contained in a Borel or Cartan subgroup, assume $l > 2$. Denote by C a Cartan subgroup and by N its normaliser. We have a homomorphism

$$Gal\big(K_l/\mathbb{Q}\big) \to N \to N/C,$$

and by our third remark on Cartan subgroups, we have $N/C \cong \mathbb{Z}/2\mathbb{Z}$. Since G is contained in the normalizer of a Cartan subgroup, but not in the Cartan subgroup itself, the above homomorphism is surjective. $Gal\big(K_l/\mathbb{Q}\big)$ must factor through $Gal\big(K_l^{ab}/\mathbb{Q}\big)$ since $\mathbb{Z}/2\mathbb{Z}$ is commutative. Recall that $Gal\big(K_l^{ab}/\mathbb{Q}\big) \cong \mathbb{Z}_l^*$, and that the only continuous homomorphism from $\mathbb{Z}_l^*$ onto $\mathbb{Z}/2\mathbb{Z}$ has kernel consisting of the squares. Thus we obtain that $\tilde{\rho}_l(Frob_p) \in C \Leftrightarrow$ p is a quadratic residue mod l. If $\alpha \in N \setminus C$, so p is a quadratic non-residue, it swaps the associated eigenspaces of C, so can be put in the form $\begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$, which has zero trace. The result follows by (15).

(iii) In this case $H \cong S_4$, so every element has order 1, 2, 3, or 4. We can then write the eigenvalues of an element of H as μ, μ^{-1}, where μ is a 2nd, 4th, 6th, or 8th root of unity respectively(?why not as 1st, 2nd, 3rd, 4th roi? doesnt his way give a sign ambiguity? i suppose it doesnt matter much). Then an element of G has eigenvalues $\lambda\mu, \lambda\mu^{-1}$. We now check the various cases:

(a) Let $\mu^2 = 1$. Then the characteristic polynomial, $f(x)$, has repeated root $-\lambda$, so $f(x) = x^2 + 2\lambda x + \lambda^2$. From this and (15) we obtain the congruences $p^{k-1} \equiv \lambda^2 \bmod l$ and $a_p \equiv -2\lambda \bmod l$. Thus $a_p^2 \equiv 4\lambda^2 \bmod l$, and so $a_p^2 p^{1-k} \equiv 4 \bmod l$.

(b) Let $\mu^4 = 1$. If $\mu^2 = 1$, we are back at case (a), so let $\mu^2 = -1$. Then $f(x)$ has roots $\lambda\mu, -\lambda\mu$ and we have $f(x) = x^2 + \lambda^2$. Similarly as in (a), we get $p^{k-1} \equiv \lambda^2 \bmod l$ and $a_p \equiv 0 \bmod l$. Thus $a_p^2 p^{1-k} \equiv 0 \bmod l$.

(c) Let $\mu^6 = 1$; either $\mu^3 = 1$ or $\mu^3 = -1$. If $\mu^3 = 1$, we end up with roots $\lambda\mu, \lambda\mu^2$ and $f(x) = x^2 + \lambda x + \lambda^2$. The congruences $p^{k-1} \equiv \lambda^2 \bmod l$ and $a_p \equiv -\lambda \bmod l$ result. Thus $a_p^2 \equiv \lambda^2 \bmod l$, and so $a_p^2 p^{1-k} \equiv 1 \bmod l$.
If $\mu^3 = -1$, $f(x)$ has roots $\lambda\mu, -\lambda\mu^2$, and we have $f(x) = x^2 - \lambda x +$

λ^2. This time, our congruences are $p^{k-1} \equiv \lambda^2 \bmod l$ and $a_p \equiv \lambda \bmod l$. Thus $a_p^2 \equiv \lambda^2 \bmod l$, and again we find $a_p^2 p^{1-k} \equiv 1 \bmod l$.

(d) Let $\mu^8 = 1$. If $\mu^4 = 1$ we are back in case (b), so let $\mu^4 = -1$. Then $\mu^2 = i$ or $-i$. We obtain $f(x) = x^2 - \lambda x(\mu + \mu^7) + \lambda^2$. The congruences are $p^{k-1} \equiv \lambda^2 \bmod l$ and $a_p \equiv \lambda(\mu + \mu^7) \bmod l$. Thus $a_p^2 \equiv \lambda^2(\mu^2 + \mu^6 + 2) \equiv \lambda^2(\mu^2 + \mu^{-2} + 2) \equiv 2\lambda^2 \bmod l$, and so we find $a_p^2 p^{1-k} \equiv 2 \bmod l$.

□

3.3 The Exceptional Primes

Corollary 5 shows us that for exceptional primes there exist congruences on the coefficients of modular forms; but are there similar congruences for non-exceptional primes?

Lemma 9. *Let $f = \sum_{n=1}^{\infty} a_n q^n$ be a cusp form of weight k with $a_1 = 1$ and integer coefficients, with a Dirichlet series that has an Euler product. If l is not exceptional for f, and N, N^* are non-empty open sets in $\mathbb{Z}_l, \mathbb{Z}_l^*$ respectively, then $\{p \text{ prime} \mid p \in N^* \text{ and } a_p \in N\}$ has positive density.*

Proof. Consider the map

$$(\rho_l, \chi_l) : Gal(K_l/\mathbb{Q}) \to GL_2(\mathbb{Z}_l) \times \mathbb{Z}_l^*$$

The projection onto $GL_2(\mathbb{Z}_l)$ contains $SL_2(\mathbb{Z}_l)$ since l is not exceptional. Then the projection onto $GL_2(\mathbb{Z}_l)$ of the commutator subgroup contains the commutator subgroup of $SL_2(\mathbb{Z}_l)$. Let $l > 3$. The commutator subgroup of $SL_2(\mathbb{F}_l)$ is normal in $SL_2(\mathbb{F}_l)$, which is simple, so the commutator subgroup is the whole of $SL_2(\mathbb{F}_l)$. Then by Theorem 5, the commutator subgroup of $SL_2(\mathbb{Z}_l)$ contains $SL_2(\mathbb{Z}_l)$, and so is the whole of $SL_2(\mathbb{Z}_l)$. If $l = 2$ or 3, let $\sigma \in Gal(K_l/\mathbb{Q})$ such that $\rho_l(\sigma) \in SL_2(\mathbb{Z}_l)$. We then have $\chi_l^{k-1}(\sigma) = det \circ \rho_l(\sigma) = 1$. Since k is even, $\chi_l(\sigma) = 1$ as $\mathbb{Z}_l$ contains no non-trivial roots of unity of odd order. Thus $SL_2(\mathbb{Z}_l) \times 1$ is contained in the image of (ρ_l, χ_l). □

From now on we assume f and ρ_l are as in Theorem 4, and that l is exceptional for f. We use the following lemma to restrict the number of primes for which each part of Corollary 5 can occur:

Lemma 10. *Congruences as in* (i) *can only occur if either* $2m < l < k$ or $m = 0$ *and* l *divides the numerator of* B_k*, and* (ii) *can only occur if* $l < 2k$*. In the case of* (iii)*, we cannot find the specific primes themselves, but we can find a finite set of primes guaranteed to contain all exceptional primes.*

Proof. Let $l > 3$. Recall (i) was the statement that $\exists m \in \mathbb{Z}$ such that

$$a_n = n^m \sigma_{k-1-2m}(n) \bmod l, \tag{18}$$

for all n prime to l. It suffices to consider $a_p \equiv p^m + p^{k-1-m} \bmod l$ for $p \neq l$. Here the exponents depend on their value mod $l-1$, so reducing them (and possibly switching them) we obtain

$$a_p \equiv p^m + p^{m'} \bmod l,$$

with $0 \leq m < m' < l-1$ and $m + m' \equiv k-1 \bmod l$. Since $k-1$ is odd, $m + m'$ is odd and hence $m \neq m'$. Applying this to (18) we find that $a_n \equiv n^m \sigma_{m'-m}(n) \bmod l$, if n is coprime to l. Thus we can write

$$\theta \tilde{f} = \theta^{m+1} \tilde{G}_{m'-m+1}, \tag{19}$$

noting that the terms involving powers of n, where $l|n$, disappear. This formula breaks down in the instance of $m = 0, m' = l-2$, since the constant term of G_{l-1} is not reducible by l. Note, however, that by Lemma (3) the constant term of G_{l-1} multiplied by l *is* reducible mod l. Using $a_p \equiv p^m + p^{k-1-m} \bmod l$, we obtain $pa_p \equiv 1 + p \bmod l$, and the congruence $na_n \equiv \sigma_1(n) \bmod l$ follows, where l and n are coprime. We can then write

$$\theta \tilde{f} = \theta^{l-1} \tilde{G}_2 = \theta^{l-1} \tilde{G}_{l+1}. \tag{20}$$

Recall $w(\theta \tilde{f}) \leq w(\tilde{f}) + l + 1$. If $3 < k < l-1$, we must have $w(\tilde{G}_k) = k$, so for $m' - m > 1$ we have

$$w(\theta^{m+1} \tilde{G}_{m'-m+1}) = m' - m + 1 + (m+1)(l+1)$$

Comparing this with the left, we find

$$m' - m + 1 + (m+1)(l+1) \leq k + l + 1, i.e.$$
$$m' + ml + 1 \leq k$$

if $1 < m' - m < l - 2$. If we have $k < l$, we must have $k - 1 \leq m + m'$. Combining this with our discussion of the filtration, we get $k-1 \leq m+m' \leq ml + m' \leq k - 1$, and so must have $m = 0$, $m' = k - 1$, and $w(\tilde{f}) = k$. Plugging these into (19) we obtain $\theta(\tilde{f} - \tilde{G}_k) = 0$. Then $\tilde{f} - \tilde{G}_k = 0$ or has filtration $k > 0$. If it has filtration k, we obtain (is the filtration of 0 0 or -inf?) $0 = w(\theta(\tilde{f} - \tilde{G}_k)) = w(\tilde{f} - \tilde{G}_k) + l + 1 = k + l + 1$, which is a contradiction as k and l are non-zero. So l divides the constant term of $f - G_k$, and since f is a cusp form, l must divide the numerator of B_k.
There are two exceptions left to consider: those of $m' - m = 1$, and $m = 0, m' = l - 2$. Since there are no modular forms of weight two, in the first of these cases we must have $w(\tilde{G}_{m'-m+1}) = w(\tilde{G}_2) = l + 1$. Similarly to before, taking the filtration of either side of (19) we find $(m + 1)(l + 1) \leq k$, if $m' - m = 1$. For the second case, from (20) we find that $l^2 - 1 \leq k$, if $m = 0, m' = l - 2$. In either instance, we must have $l < k$.
For the second case, that of $a_n \equiv 0$ for n a quadratic non-residue mod l, observe that $n^{\frac{l-1}{2}} \equiv 1 \bmod l$ if and only if n is a quadratic residue mod l. We can then express this case in terms of the theta operator as follows:

$$\theta \tilde{f} = \theta^{\frac{l+1}{2}} \tilde{f}$$

If $l > 2k$, we must have $w(\tilde{f}) = k(why?)$. Then the left hand side of the above equation has filtration $k+l+1$, while the right hand side has filtration $k + \frac{1}{2}(l + 1)^2$. These are not equal for any primes, so we have obtained a contradiction. Finally, $l \neq 2k$, since k is even and l is odd.
Case (iii) stipulates that $p^{1-k}a_p^2 \equiv 0, 1, 2,$ or 4 mod l for all primes $p \neq l$. Pick $p > 2$ such that $a_p \neq 0$. If l is such an exceptional prime, we must have $l = p$ or one that of the following is divisible by l: $a_p^2, a_p^2 - p^{k-1}, a_p^2 - 2p^{k-1}$, or $a_p^2 - 4p^{k-1}$. Since k is even, these are all non-zero and finite, so there is a finite number of divisors and hence one obtains a finite list which contains all exceptional primes. □

We are finally able to classify the exceptional primes of each type, for cusp forms satisfying the constraints of Theorem (4). There are six (known) such forms: $\Delta, Q\Delta, R\Delta, Q^2\Delta, QR\Delta,$ and $Q^2R\Delta$, which have weights 12, 16, 18, 20, 22, and 26 respectively.
For type (i) exceptional primes, the results can be expressed in a table, which lists the value of m for a given modular form and fixed choice of l:

Form	k	2	3	5	7	11	13	17	19	23	Other
Δ	12	0	0	1	1	-					691
$Q\Delta$	16	0	0	1	1	1	-				3617
$R\Delta$	18	0	0	2	1	1	1	-			43867
$Q^2\Delta$	20	0	0	1	2	1	1	-	-		283,617
QRΔ	22	0	0	2	1	-	1	1	-		131,593
$Q^2R\Delta$	26	0	0	2	2	1	-	1	1	-	657,931

The $k = 2$ column follows as the coefficients a_p (where p is prime) of Δ are even. When $l = 3$, we make use of the property of the tau function noted in [D H65], that if n and 3 are coprime, then $\tau(n) \equiv \sigma(n) \bmod 3$. Plugging in $p \neq 3$, we see that $a_p = \tau(p) \equiv 1 + p \bmod 3$, and so must have $m = 0$. Theorem (3) implies this for the rest of the column. For the remaining cases where $l < k$, one proceeds as follows: take, for example, the form Δ and $l = 5$, and consider $p = 2$. We have $a_2 = -24$, and so (17) implies

$$-24 \equiv 2^m + 2^{12-1-m} \bmod 5,$$

which simplifies to

$$1 \equiv 2^m + 2^{3-m} \bmod 5.$$

Following the constraints applied in Lemma (10) (i), plug in values of m satisfying $0 \leq m < m' \leq 3$, and after possibly adjusting the congruence to have the required form, one finds that 1 is the only value of m satisfying the congruence. This method works similarly for $l < k$ with small primes p, and also shows whether there are no such values m(here sd says to check the equation in thetas for values in the table - why? have we not just found the values of m?).

There are only two type (ii) primes: Δ has exceptional prime $l = 23$, and $Q\Delta$ has exceptional prime $l = 31$. The previous lemma indicates $l < 2k$, so we need to check $l < k$ such that l isn't type (i), and also $l = 2k - 3$ and $l = 2k - 1$.

$$e^x \tag{21}$$

This is eq.21

Bibliography

[BS66] Z. I. Borevich and I. R. Shafarevich. *Number Theory*. Academic Press Inc, 1966. ISBN: 9780121178505.

[Cox13] David Cox. *Primes of the Form $x^2 + ny^2$*. Pure and Applied Mathematics: a Wiley Series of Texts, Monographs, and Tracts. Wiley, 2013. ISBN: 9781118390184.

[D H65] D. H. Lehmer. "The Primality of Ramanujan's Tau-Function". In: *The American Mathematical Monthly* 72 (1965), 15–18.

[Jea72] Jean-Pierre Serre. "Congruences et Formes Modulaires". In: *Séminaire N. Bourbaki* (1971-1972), 319–338.

[Kur17] Robert Kurinczuk. *Modular Forms Lecture Notes*. 2017. URL: %7Bhttp://wwwf.imperial.ac.uk/~dhelm/M4P58/ModularForms2.pdf%7D.

[Lan94] Serge Lang. *Algebraic Number Theory*. Graduate Texts in Mathematics. Springer, 1994. ISBN: 9781461269229.

[Lan01] Serge Lang. *Introduction to Modular Forms*. A Series of Comprehensive Studies in Mathematics. Springer, 2001. ISBN: 3540078339.

[Pet72] Peter Swinnerton Dyer. "On l-adic Representations and Congruences for Coefficients of Modular Forms". In: *Modular Functions of One Variable III*. Vol. 350. Lecture Notes in Mathematics. 1972, pp. 1–55.

[Sut16] Andrew Sutherland. *MIT Number Theory I Lecture Notes*. 2016. URL: %7Bhttp://math.mit.edu/classes/18.785/2016fa/lectures.html%7D.