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 will rely on the results of Cox's *Primes of the Form* $x^2 + ny^2$ and Lang's *Algebraic Number Theory*.

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 $\mathcal{O}_L$:

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

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 $\mathfrak{q}_i$.
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 $N_{L/K}(x) = \prod\limits_{\sigma\in G} \sigma(x) \in \mathfrak{p} \subset \mathfrak{q}_2$, 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.

$\square$

Definition 2. In the above situation, we say an ideal $\mathfrak{p} \subset \mathcal{O}_K$ *ramifies* if $e > 1$, and is *unramified* if $e = 1$.

Definition 3. Let L be a finite, Galois extension of K and $\mathfrak{q}$ a prime ideal of L. We 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 $D_{\mathfrak{q}} \tilde{\rightarrow} \bar{G}$, given by mapping σ to $\bar{\sigma}$. Finally note that

Lemma 2. *There exists a*

1.2 Cebotarev's Density Theorem

1.3 Modular Forms

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

Definition 4. A modular form of weight k and level 1 is a function

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

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

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

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

We define the Eisenstein series of weight k to be

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

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.
E_2 does not quite satisfy the modular transformation rule, instead obeying

$$E_2(\frac{-1}{z}) = z^2E_2(z) + \frac{12z}{2i\pi} \tag{4}$$

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 and Swinnerton-Dyer, with a little help from Lang's *Introduction to Modular Forms*.

Definition 5. A modular form

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

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

$$f(\frac{-1}{z}) = z^k f(z) \tag{5}$$

to prove the result. Differentiating the left hand side of (5) 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*

$$\partial Q = -4R, \tag{6}$$

$$\partial R = -6Q^2, \tag{7}$$

$$\partial \Delta = 0. \tag{8}$$

Before we prove this, we need a lemma:

Lemma 3. *We have*

$$\Theta P = \frac{1}{12}(P^2 - Q) \tag{9}$$

$$\Theta Q = \frac{1}{3}(PQ - R) \tag{10}$$

$$\Theta R = \frac{1}{2}(PR - Q^2) \tag{11}$$

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 (4). Differentiating both sides of this, we obtain

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

Using this, we find

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

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 (10) 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 (11) 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. (6) and (7) follow directly from (10) and (11), 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 $\mathbf{a}$.

The following result will prove highly useful; a proof can be found on pages 384-386 of *Number Theory*, by Borevich and Shafarevich:

Lemma 4. *(Von Staudt) Let B_n denote the Bernouilli numbers, and $p > 3$. Then:*

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

(ii) *If $p-1 \not| B_{2\nu}$, $\frac{B_{2\nu}}{2\nu}$ is p-integral, and $\frac{B_{2\nu}}{2\nu} \equiv \frac{B_{2\nu \text{ mod } p-1}}{2\nu \text{ mod } p-1}$ 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 5. *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 4 (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 4 (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 6.

(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 2). 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 (9). 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

$$0 \subseteq (\tilde{A} - 1) \subseteq \mathbf{a} \subseteq \mathfrak{m} \tag{12}$$

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 $k - l =$ (p−1)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 8. 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 7.

(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 6 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 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*

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

such that

$$\chi\big(Frob(p)\big) = p,$$

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*

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

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*

$$x^2 - a_p x + p^{k-1} \tag{13}$$

for each $p \neq l$.

Proof. We will leave these results unproven. □

Note that (13) 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

$$\tilde{\rho}_l : Gal\big(K_l/\mathbb{Q}\big) \to GL_2\big(\mathbb{Z}_l\big) \to GL_2\big(\mathbb{F}_l\big), \tag{14}$$

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

Lemma 8. *The set of matrices*

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

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

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 8 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. Suppose that $SL_2\big(\mathbb{Z}/l^{n-1}\mathbb{Z}\big) \subset G_{n-1}$. We know that H_{n-1} is contained in G_{n-1}. $\square$

Note: this theorem implies that in order to determine

Definition 9. A prime number l is an *exceptional* prime for the cusp form f if the image of ρ_l does not contain $SL_2(\mathbb{Z}_l)$.

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 $SL_2(\mathbb{F}_l)$.

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

$$e^x \tag{15}$$

This is eq.15