Congruence properties of modular forms

Gil Alon

November 15, 1998

Contents

In this treatise I discuss the subject of congruence properties of coefficients of modular forms from several points of view.
This subject has first emerged in the work of Ramanujan, who discovered the congruence (proved in chapter 2)

$$\sigma_{11}(n) \equiv \tau(n) \ (mod\ 691)$$

which holds for all positive integers n, where $\sigma_{11}(n) = \sum_{d|n} d^{11}$ and $\tau(n)$ is defined by

$$\sum_{n=1}^{\infty} \tau(n) q^n = q \prod_{n=1}^{\infty} (1-q^n)^{24}$$

It turns out that both $\sigma_{11}(n)$ and $\tau(n)$ are Fourier coefficients of complex functions called **modular forms**, the $\sigma_{11}(n)$ being coefficients of a form called an **Eisenstein series** and $\tau(n)$ being coefficients of a form called a **cusp form**.
In chapter 1 an introduction to the subject of modular forms is presented.
In chapter 2, congruences between cusp forms and Eisenstein series are discussed, following the presentation of a recent paper ([DG]).
In chapter 3, I discuss congruence properties of cusp forms, using the theory of l-adic representations. This presentation follows an article of H.P.F Swinnerton-Dyer ([SwD]).
It turns out that for certain cusp forms, one can attach for every prime number l a 2-dimensional l-adic representation of some galois group (that is, a homomorphism from this group to $GL_2(\mathbb{Z}_l)$), such that the p-th fourier coefficient of f is equal to the trace of the value of this representation on an element determined by p (the Frobenius automorphism attached to p). By studying the image of this representation one can obtain information on the l-adic properties of the coefficients.

In chapter 4, I use the ideas in [SwD] to discuss congruence properties of **pairs** of cusp forms. Here one is lead to the study of the direct product of the two attached representations, most of the time using group-theoretic considerations. It is in this chapter that most of my own work on the subject appears.

I am indebted to Prof. Ehud De-Shalit for having introduced to me many areas of number theory over the last 11 years, and for his intensive and dedicated support during this project.

Chapter 1

Modular forms

In this chapter we give the fundamental definitions and theorems concerning modular forms. Detailed proofs can be found in [Mi].

1.1 Action of $GL_2^+(\mathbb{R})$ on points and functions

1.1.1 The projective line

For a field F we define the projective line $\mathbb{P}^1(F)$ as the set of lines in F^2 passing through the origin, namely

$$\mathbb{P}^1(F) = (F^2 \setminus \{(0,0)\})/\sim$$

where $\sim$ is the equivalence relation

$$(x_1, x_2) \sim \lambda(x_1, x_2) \;\; \forall (x_1, x_2) \neq (0,0), \;\; \lambda \neq 0$$

We identify the points of $\mathbb{P}^1(F)$ with $F \cup \{\infty\}$ by

$$x \in F \longleftrightarrow [(x,1)] \in \mathbb{P}^1(F)$$

$$\infty \longleftrightarrow [(1,0)]$$

1.1.2 Action of $GL_2(F)$ on points

We let $GL_2(F)$ act on $\mathbb{P}^1(F)$ by

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)[(x,y)] = [(ax+by, cx+dy)]$$

One sees immediately that this action is well defined, and that it is indeed an action. Together with the identification of $\mathbb{P}^1(F)$ and $F \cup \{\infty\}$ we get an action of $GL_2(F)$ on $F \cup \{\infty\}$:

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z = \frac{az+b}{cz+d}$$

The last equality should be interpreted as follows: if $cz + d = 0$, the result is ∞. If $z = \infty$, the result is $\frac{a}{c}$, unless $c = 0$ in which case the result is again ∞.

1.1.3 Action of $GL_2^+(\mathbb{R})$ on the upper half-plane

Taking $F = \mathbb{C}$ we get an action of $GL_2(\mathbb{C})$ on $\mathbb{C} \cup \{\infty\}$, and in particular an action of $GL_2(\mathbb{R})$ on $C \cup \{\infty\}$. Since (taking $F = \mathbb{R}$) $GL_2(\mathbb{R})$ also acts on $\mathbb{R} \cup \infty$, it acts separately on $\mathbb{C} \setminus \mathbb{R}$.
Let $z \in \mathbb{C} \setminus \mathbb{R}$, $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL_2(\mathbb{R})$, then we have $\gamma z \in \mathbb{C} \setminus \mathbb{R}$, and

$$im(\gamma z) = im\frac{az+b}{cz+d} = im\frac{(az+b)(c\overline{z}+d)}{|cz+d|^2} = \frac{(ad-bc)im\ z}{|cz+d|^2} = \frac{det\ \gamma}{|cz+d|^2} im\ z$$

In particular if $det\ \gamma > 0$, then $im\ z$ and $im\ \gamma z$ have the same sign.
Define $\mathcal{H} = \{z \in \mathbb{C} : im\ z > 0\}$ (the upper half-plane). Then we have an action of $GL_2^+(\mathbb{R})$ on $\mathcal{H}$ and on $\mathcal{H} \cup \mathbb{R} \cup \{\infty\}$.

1.1.4 $k-$ action of $GL_2^+(\mathbb{R})$ on functions

Let $\mathcal{F}$ be the set of functions $\mathcal{H} \to \mathbb{C} \cup \{\infty\}$, and k an even integer. We define an action of $GL_2^+(\mathbb{R})$ on $\mathcal{F}$ as follows:

$$f|_{[\gamma]_k}(z) = f(\gamma z)\gamma'(z)^{\frac{k}{2}}$$

for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ we have $f|_{[\gamma]_k}(z) = f(\frac{az+b}{cz+d})(\frac{az+b}{cz+d})'^{\frac{k}{2}} = f(\frac{az+b}{cz+d})\frac{(det\ \gamma)^{\frac{k}{2}}}{(cz+d)^k}$.
By abuse of notation we shall often write $f(z)|_{[\gamma]_k}$ for $f|_{[\gamma]_k}(z)$.
This is an action since if $\alpha, \beta \in GL_2^+(\mathbb{R})$,

$$f(z)|_{[\alpha]_k}|_{[\beta]_k} = f(\alpha z)\alpha'(z)^{\frac{k}{2}}|[\beta]_k = f(\alpha\beta z)\alpha'(\beta z)^{\frac{k}{2}}\beta'(z)^{\frac{k}{2}} = f(\alpha\beta z)((\alpha\beta)'(z))^{\frac{k}{2}} = f(z)|[\alpha\beta]_k$$

We call it *the $k-$action of $GL_2^+(\mathbb{R})$ on $\mathcal{F}$.*

Definition 1.1.1. *let $\Gamma \subseteq GL_2^+(\mathbb{R})$ be a subgroup. An element of $\mathcal{F}$ which is stable under the $k-$ action of Γ is called an automorphic form of weight k for Γ.*

1.2 $SL_2(\mathbb{Z})$ and its congruence subgroups

Among the subgroups of $SL_2(\mathbb{Z})$, our interest will be in the group

$$SL_2(\mathbb{Z}) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : a, b, c, d \in \mathbb{Z},\ ad - bc = 1\}$$

and some of its subgroups. We will note $\Gamma(1) = SL_2(\mathbb{Z})$, and call it *the full modular group*. It is a group since if $ad - bc = 1$, $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$.
Finite-index subgroups of $\Gamma(1)$ will be called *modular groups.*
Let N be a positive integer, and $SL_2(\mathbb{Z}/N\mathbb{Z})$ the finite group of matrices of determinant 1 over $\mathbb{Z}/N\mathbb{Z}$. We have the homomorphism $\phi : SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/N\mathbb{Z})$ of reduction modulo N.
We denote by $\Gamma(N)$ the kernel of this homomorphism, namely

$$\Gamma(N) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) : a \equiv 1,\ b \equiv 0,\ c \equiv 0,\ d \equiv 1\ (modN)\}$$

We call $\Gamma(N)$ *the principal congruence modular group mod* N. It is indeed a modular group since by the first isomorphism theorem, $|\Gamma(1)/\Gamma(N)| = |Im\ \phi| < \infty$.

Definition 1.2.1. *A subgroup of* $\Gamma(1)$ *is called a congruence modular group if it contains* $\Gamma(N)$ *for some positive integer* N. *The minimal such* N *is called the level of* Γ.

Of course, if $\Gamma(1) \supseteq G \supseteq \Gamma(N)$ then $G = \phi^{-1}(\tilde{G})$ for some subgroup $\tilde{G} \subset SL_2(\mathbb{Z}/N\mathbb{Z})$.
The following are congruence subgroups:

$$\Gamma_0(N) = \{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) : c \equiv 0\ (mod\ N)\}$$

$$\Gamma_1(N) = \{\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)\} = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N) : a \equiv d \equiv 1\ (mod\ N)\}$$

1.2.1 $PSL_2(\mathbb{Z})$

The following claims are easy to verify

1. The center of $SL_2(\mathbb{Z})$ is ± 1
2. ± 1 Are the only matrices in $SL_2(\mathbb{Z})$ which act on $\mathcal{H}$ as the identity.

We define $PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z})/\{\pm 1\}$.
This group acts faithfully on $\mathcal{H}$. We can similarly define its k-action on functions. We have a projection $\pi : SL_2(\mathbb{Z}) \twoheadrightarrow PSL_2(\mathbb{Z})$. We shall also call subgroups of $PSL_2(\mathbb{Z})$ of the form $\pi\Gamma$, $\Gamma \subseteq SL_2(\mathbb{Z})$ congruence modular groups whenever Γ is a congruence modular group.

1.3 Modular forms

Let Γ be a congruence modular group of level N.

Definition 1.3.1. *Let f be a meromorphic function on $\mathcal{H}$. We call f a* ***meromorphic modular form*** *of weight k for Γ if:*

1. *f is an automorphic function of weight k for Γ*
2. *for all $\gamma \in SL_2(\mathbb{Z})$, one can expand*

$$f|_{[\gamma]_k} = \sum_{n=-\infty}^{\infty} a_n^{\gamma} {q_N}^n$$

where $q_N(z) = e^{\frac{2\pi i z}{N}}$ and $a_n^{\gamma} = 0$ for all sufficiently small n.

If f is also holomorphic and $a_n^{\gamma} = 0$ for $n < 0$ (for every γ) then we call f ***a modular form.***
If additionally $\forall \gamma, a_0^{\gamma} = 0$ we call f ***a cusp form.***

We note
$M_k(\Gamma)$ = The vector space over $\mathbb{C}$ of modular forms of weight k for Γ
$S_k(\Gamma)$ = The vector space over $\mathbb{C}$ of cusp forms of weight k for Γ
(The letter S in $S_k(\Gamma)$ stands for the German term *Spitzenform* for cusp-form).
Note that since f is an automorphic function for Γ, one has to verify the conditions for a_n^{γ} only for one representative of each coset in $SL_2(\mathbb{Z})/\Gamma$.
The restrictions on the coefficients $\{a_n^{\gamma}\}$ might seem unnatural. They will become more evident in the next section.
The following claim is very easy:

Claim 1.3.2. *Let $f_1 \in M_{k_1}(\Gamma)$, $f_2 \in M_{k_2}(\Gamma)$. Then $f_1 f_2 \in M_{k_1+k_2}(\Gamma)$. If additionally $f_1 \in S_{k_1}(\Gamma)$ then $f_1 f_2 \in S_{k_1+k_2}(\Gamma)$*

It turns out ([Mi], lemma 2.1.1.) that the infinite sum $\sum_{even\ k} M_k(\Gamma)$ is a direct sum, and hence it has a structure of a graded algebra over $\mathbb{C}$.

1.4 The Riemann surface of a congruence subgroup

Let $\Gamma \subseteq PSL_2(\mathbb{Z})$ be a congruence modular group. We will define a compact Riemann surface X_Γ which will be intimately connected with modular forms for Γ: modular forms for Γ will turn out to be merely q-differentials on X_Γ.
We will first give a structure of Riemann surface to $\Gamma\backslash\mathcal{H}$, and then add a finite number of points called *cusps* to make it compact.

1.4.1 $\Gamma\backslash\mathcal{H}$ as a Riemann surface

We give $\mathcal{H}$ the Euclidian topology, and $\Gamma\backslash\mathcal{H}$ the quotient topology. By [Mi] (theorem 1.5.2 and lemma 1.7.2), this is still a Hausdorff space.

Lemma 1.4.1. *For any point $z \in \mathcal{H}$ there exists a neighborhood U of z such that for every $\gamma \in \Gamma$,*

$$\gamma U \cap U \neq \emptyset \Leftrightarrow \gamma z = z$$

The lemma is obvious from the proof of [Mi], lemma 1.7.2.

Definition 1.4.2. *Let $z \in \mathcal{H}$. We say that z is an elliptic point with respect to Γ if there exists $\sigma \in \Gamma$ such that $\sigma z = z$. Otherwise, we call z an ordinary point.*

Denote by π the projection from $\mathcal{H}$ to $\Gamma\backslash\mathcal{H}$.
For $a = \pi(z_0) \in \Gamma\backslash\mathcal{H}$, we wish to give a neighborhood and a local parameter (V_a, t_a). Assume first that z_0 is an ordinary point (note that this does not depend on the representative z_0 of a), and let U be a neighborhood as in lemma 1.4.1. Put $V_a = \pi(U)$, then we have a bijection $t_a : V_a \to U$ given by

$$\forall z \in U,\ t_a(\pi(z)) = z$$

If z is an elliptic point, we can not give the same parameter since it will not be well-defined: It might happen that z and γz are in U for some $\gamma \in \Gamma_{z_0}$. (We denote by Γ_{z_0} the stabilizer of z_0 in Γ).

For the definition we need the following lemma ([Mi], Theorem 1.5.4, art. 1)

Lemma 1.4.3. *If z_0 is an elliptic point then Γ_{z_0} is a finite cyclic group*

Definition 1.4.4. *For $z_0 \in \mathcal{H}$, $e_z = |\Gamma_{z_0}|$. We call this number the order of z_0 with respect to Γ.*

Now there exists some $\rho \in SL_2(\mathbb{R})$ such that $\rho\mathcal{H} = K = \{z \in \mathbb{C} : |z| < 1\}$, and $\rho z_0 = 0$.
$(\Gamma_{z_0})^\rho$ is a finite cyclic group which acts faithfully on K and fixes 0, hence by Schwarz'es theorem it is the group of rotations in angels $\{\frac{2\pi k}{e_{z_0}} : \ 0 \leq k \leq e_{z_0} - 1\}$.
This gives an homeomorphism

$$(\Gamma_{z_0})^\rho \backslash K \cong K$$

by $x \mapsto x^{e_{z_0}}$.
We take U as before. Define $K_r = \{z \in \mathbb{C} : |z| < r\}$. Pick r small enough such that $\rho^{-1}K_r \subseteq U$ and define $U_r = \rho^{-1}K_r$. Then as before we have homeomorphisms

$$\Gamma\backslash\pi(U_r) \cong \Gamma_{z_0}\backslash U_r \cong (\Gamma_{z_0})^\rho\backslash K_r \cong K_{r^{e_{z_0}}}$$

and we can define $V_a = \pi(U_r)$, and

$$t_a(\pi(z)) = (\rho z)^{e_{z_0}} \quad \forall z \in U_r$$

1.4.2 Addition of cusps

$\Gamma\backslash\mathcal{H}$ is not compact. We compactify it by adding a finite number of points called cusps. First let us define an extension of the upper half-plane

$$\mathcal{H}^* = \mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$$

(note that all the added points are $SL_2(\mathbb{Z})$-equivalent).
We define a topology on $\mathcal{H}^*$ as follows: For $z \in \mathcal{H}$ we take the fundamental neighborhood system as in the topology of $\mathcal{H}$. For $z \in \mathbb{Q} \cup \{\infty\}$ we take $\sigma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\sigma z = \infty$ and take as the fundamental system of neighborhoods the sets $\{s \in \mathcal{H} : im(\sigma s) > l\} \cup \{z\}$ for all real $l > 0$.
We now consider the space $\Gamma\backslash\mathcal{H}^*$. Since Γ is a congruence modular group, it has a finite index in $SL_2(\mathbb{Z})$ and hence $\Gamma\backslash(\mathbb{Q} \cup \{\infty\})$ is finite. Hence our new space $\Gamma\backslash\mathcal{H}^*$ has only a finite number of points more than $\Gamma\backslash\mathcal{H}$.

Theorem 1.4.5. *$\Gamma\backslash\mathcal{H}^*$ is a compact hausdorff space*

This space is hausdorff by [Mi],lemma 1.7.7, and it is compact since for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ we know this by [Mi],theorem 4.1.2, and by [Mi], corollary 1.9.2 the property "$\Gamma\backslash\mathcal{H}^*$ is compact" carries over from Γ to subgroups of Γ of finite index.
We now extend the Riemann surface structure of $\Gamma\backslash\mathcal{H}$ to $\Gamma\backslash\mathcal{H}^*$.

Lemma 1.4.6. *Let $\Gamma \subseteq SL_2(\mathbb{Z})$ be a congruence modular group, $z \in \mathbb{Q} \cup \{\infty\}$ and $\sigma \in SL_2(\mathbb{Z})$ such that $\sigma z = \infty$. Then there exists $0 < h \in \mathbb{Z}$ such that*

$$\sigma\Gamma_z\sigma^{-1} \cdot \{\pm 1\} = \{\pm \left(\begin{smallmatrix} 1 & mh \\ 0 & 1 \end{smallmatrix}\right) | m \in \mathbb{Z}\}$$

moreover, if $l > h$ and $W = \{s \in \mathcal{H} : im(\sigma s) > l\} \cup \{z\}$, then

$$\gamma z = z \Leftrightarrow \gamma W \cap W \neq \emptyset$$

(The first part of the lemma can be found in [Mi], theorem 1.5.4, art. 2 and the second part follows from [Mi], corollary 1.7.5).

Let $a = \pi(z_0)$ be a cusp. we take $\sigma \in SL_2(\mathbb{Z})$ such that $\sigma z_0 = \infty$
We take h, l and W as in the lemma. Then

$$\Gamma\backslash\pi(W) \cong \Gamma_{z_0}\backslash W \cong \sigma\Gamma_{z_0}\sigma^{-1}\backslash\sigma W \cong \{\left(\begin{smallmatrix} 1 & mh \\ 0 & 1 \end{smallmatrix}\right) | m \in \mathbb{Z}\}\backslash(\{z : im(z) > l\} \cup \{\infty\})$$

The last set is homeomorphic to the disc K_r where $r = e^{\frac{-2\pi i l}{h}}$, by $\psi(z) = e^{\frac{2\pi i z}{h}}$ (we implicitly define $\psi(\infty) = 0$). We therefore define $V_a = \pi(W)$ and

$$t_a(\pi(z)) = \psi(\sigma z) = e^{\frac{2\pi i(\sigma z)}{h}}$$

Definition 1.4.7. *We denote the Riemann surface $\Gamma\backslash\mathcal{H}$ by X_Γ and call it* ***the Riemann surface of*** *Γ.*

1.4.3 The connection between modular forms and differentials

As promised, we give an interpretation of modular forms for Γ as holomorphic differentials on X_Γ. Let f be an automorphic modular form of weight $k = 2m$ ($m \in \mathbb{Z}$). We shall attach to it an m-differential ω_f on X_Γ.
We have defined in the previous section an atlas of local parameters $\{(U_a, t_a)\}_{a \in X_\Gamma}$ on X_Γ. Now we wish to add functions $\{\phi_a\}$ such that

$$\phi_a(x)(\frac{dt_a}{dt_b})^m(x) = \phi_b(x) \;\; \forall x \in U_a \cap U_b$$

Let $a = \pi(z_0) \in X_\Gamma$.
Recall that the neighborhood V_a of a is always of the form $\pi(U)$ where U is a neighborhood of z_0 in $\mathcal{H}^*$ satisfying

$$\gamma U \cap U \neq \emptyset \Leftrightarrow \gamma \in = \Gamma_{z_0}$$

note that we also have

$$\gamma \in \Gamma_{z_0} \Rightarrow \gamma U = U$$

(since if z_0 is ordinary, this is trivial, if it is elliptic then after a suitable conjugation Γ_{z_0} acts as a group of rotations on some disc K_r, and if a is a cusp then after conjugation Γ_{z_0} acts as a group of integer translations $(z \mapsto z+n)$ on $\{z : im\ z > l\} \cup \{\infty\}$).

We first assume that $z_0 \in \mathcal{H}$, that is, a is not a cusp. we define (following [Mi]) a function $g(z)$ on U by

$$g(z) = f(z)(\frac{d(t_a \circ \pi(z))}{dz})^{-m}$$

Then (since $\pi \circ \gamma = \pi$ and f is Γ-automorphic)

$$g(\gamma z) = f(\gamma z)(\frac{\frac{d(t_a \circ \pi \circ \gamma(z))}{dz}}{\frac{d(\gamma z)}{dz}})^{-m} = \gamma'(z)^{-m} f(z)(\frac{d(t_a \circ \pi(z))}{dz})^{-m}\gamma'(z)^m = g(z)$$

Therefore there exists a function ϕ_a on V_a satisfying $\phi_a \circ \pi(z) = g(z)$. By definition ϕ_a is meromorphic, since g is.
Assume now that a is a cusp. We first define $g(z)$ on $U \setminus \{z_0\}$ (this set is in $\mathcal{H}$) by the same equation $g(z) = f(z)(\frac{d(t_a \circ \pi(z))}{dz})^{-m}$ we get again that g is γ-invariant and hence defines a meromorphic function ϕ_a on $V_a \setminus \{a\}$.

Claim 1.4.8. *ϕ_a is meromorphic at a.*

Proof. Recall that $t_a \circ \pi(z) = e^{\frac{2\pi i(\sigma z)}{h}}$ where $\sigma \in SL_2(\mathbb{Z})$, $\sigma z_0 = \infty$, and

$$\sigma \Gamma_{z_0} \sigma^{-1} \cdot \{\pm 1\} = \{\pm \left(\begin{smallmatrix} 1 & mh \\ 0 & 1 \end{smallmatrix}\right) | m \in \mathbb{Z}\}$$

hence

$$\frac{d(t_a \circ \pi(z))}{dz} = \frac{2\pi i}{h} t_a \circ \pi(z) \frac{d(\sigma z)}{dz}$$

put $c = (2\pi i)^{-m}$, then

$$\phi_a \circ \pi(z) = cf(z)(\frac{d(\sigma z)}{dz})^{-m}(t_a \circ \pi(z))^{-m} = c(f|_{[\sigma^{-1}]_k})(\sigma z)(t_a \circ \pi(z))^{-m}$$

It remains to prove that $(f|_{[\sigma^{-1}]_k})(\sigma z)$ is a holomorphic function of $t_a \circ \pi(z)$ at z_0 Since f is a meromorphic modular form, we have a development

$$(f|_{[\sigma^{-1}]_k})(\sigma z) = \sum_{n=T}^{\infty} a_n e^{\frac{2\pi i \sigma z}{N}}$$

where N is the level of Γ. We have $\Gamma(N) \in \Gamma$, and since $\Gamma(N) \lhd \mathrm{SL}_2(\mathbb{Z})$, $\Gamma(N) \subseteq \sigma\Gamma\sigma^{-1}$. In particular we have $\left(\begin{smallmatrix}1 & N\\ 0 & 1\end{smallmatrix}\right) \in \sigma\Gamma\sigma^{-1}$. On the other hand, $\left(\begin{smallmatrix}1 & N\\ 0 & 1\end{smallmatrix}\right)\infty = \infty$, hence

$$\left(\begin{smallmatrix}1 & N\\ 0 & 1\end{smallmatrix}\right) \in (\sigma\Gamma\sigma^{-1})_\infty = \sigma\Gamma_{z_0}\sigma^{-1} \subseteq \{\pm\left(\begin{smallmatrix}1 & mh\\ 0 & 1\end{smallmatrix}\right) | m \in \mathbb{Z}\}$$

We deduce that $h \mid N$. On the other hand we have $\sigma^{-1}\left(\begin{smallmatrix}1 & h\\ 0 & 1\end{smallmatrix}\right)\sigma \in \Gamma$ (for convenience, we ignore the $\pm$ sign by looking at $\left(\begin{smallmatrix}1 & h\\ 0 & 1\end{smallmatrix}\right)$ as an element of $PSL_2(\mathbb{Z})$). Hence $f|_{[\sigma^{-1}\left(\begin{smallmatrix}1 & h\\ 0 & 1\end{smallmatrix}\right)\sigma]_k} = f \Rightarrow f|_{[\sigma^{-1}]}|_{\left(\begin{smallmatrix}1 & h\\ 0 & 1\end{smallmatrix}\right)} = f|_{[\sigma^{-1}]}$ If we put $\rho = f|_{[\sigma^{-1}]}$ we see that

$$\rho(z) = \rho|_{\left(\begin{smallmatrix}1 & h\\ 0 & 1\end{smallmatrix}\right)} = \rho(z+h)$$

Since

$$\rho(z) = \sum_{n=T}^{\infty} a_n e^{\frac{2\pi i n z}{N}}$$

we get

$$\sum_{n=T}^{\infty} a_n e^{\frac{2\pi i n}{N}} = \sum_{n=T}^{\infty} a_n e^{\frac{2\pi i n(z+h)}{N}} = \sum_{n=T}^{\infty} a_n e^{\frac{2\pi i n h}{N}} e^{\frac{2\pi i n z}{N}}$$

and hence, a_n must vanish whenever $N \nmid nh$. The result follows. □

Hence ϕ_a is a meromorphic function U.
We have thus defined an m-differential ω_f (the condition $\phi_a(x)(\frac{dt_a}{dt_b})^m(x) = \phi_b(x)$ can be checked easily). It turns out (see [Mi], p.49) that $(f \mapsto \omega_f)$ is an isomorphism of vector spaces between meromorphic modular forms of weight $k = 2m$ for Γ and meromorphic m-differentials on X_Γ. Furthermore, it satisfies (for f, g meromorphic modular forms of weights k_1, k_2) $\omega_{fg} = \omega_f\omega_g$.
By the above isomorphism we get that there always exist meromorphic modular forms of any even weight.

1.4.4 Calculation of dimensions

We now use X_Γ, the correspondence $f \mapsto \omega_f$ and the Riemann-Roch theorem to calculate the dimensions of $M_k(\Gamma)$ and $S_k(\gamma)$.

Definition 1.4.9. *Let f be a meromorphic modular form of weight $k = 2m$ for Γ, $a = \pi(z_0) \in X_\Gamma$. We define $v_a(f)$ as follows:*
If a is not a cusp, we define

$$v_a(f) = \frac{1}{e_a} ord_{z-z_0}(f)$$

If a is a cusp, we take σ, h, t_a as in the definition of ω_f, develop

$$(f|_{[\sigma^{-1}]_k})(\sigma z) = \sum_{n=T}^{\infty} a_n e^{\frac{2\pi i \sigma z}{h}} = \sum_{n=T}^{\infty} a_n (t_a \circ \pi(z))^n$$

and define $v_a(f) = T$.
*We define the **divisor** of f as*

$$div(f) = \sum_{a \in X_\Gamma} v_a(f) a$$

Note that the sum in the definition of $div(f)$ contains only a finite number of summands. We regard $div(f)$ as an element of the free $\mathbb{Q}$-module generated by the points of X_Γ, i.e, $Div(X_\Gamma) \otimes \mathbb{Q}$.
The relation between $v_a(f)$ and $v_a(\omega_f)$ is summerized in the following theorem:

Theorem 1.4.10. *([Mi],theorem 2.3.3) we have*

$$div(f) = div(\omega_f) + \frac{k}{2} \sum_a (1 - \frac{1}{e_a}) a$$

$$deg(div(f)) = k(g-1) + \frac{k}{2} \sum_a (1 - \frac{1}{e_a})$$

where g is the genus of X_Γ, and we define $\frac{1}{e_a} = 0$ when a is a cusp.

For $d = \sum_a c_a a \in Div(X_\Gamma) \otimes \mathbb{Q}$ we define

$$[d] = \sum_a [c_a] a$$

Where $[x]$ denotes the integral part of a real number x.

Lemma 1.4.11. *Let f_0 be a meromorphic modular form of weight $k = 2m$ for Γ (we have seen that there always exists such a form). then we have isomorphisms of vector spaces*

1. $M_k(\Gamma) \cong L([div(f_0)])$
2. $S_k(\Gamma) \cong L([div(f_0)] - \sum_{v=1}^{t} b_v)$

Where $b_1, .., b_v$ are the cusps of Γ.

Proof. Denote by $\mathcal{A}_k$ the space of meromorphic modular forms of weight k for Γ and by $K(X_\Gamma)$ the field of meromorphic functions on X_Γ. then we have $K(X_\Gamma) \cong \mathcal{A}_0$, and $\omega_f \in K(X_\Gamma), div(f) = div(\omega_f)$ for $f \in \mathcal{A}_0$.
Now

$$M_k(\Gamma) = \{f | f \in \mathcal{A}_k, f = 0 \text{ or } div(f) \geq 0\}$$
$$= \{f f_0 | f \in \mathcal{A}_0,\ f = 0 \text{ or } div(f) + div(f_0) \geq 0\}$$
$$\cong \{f | f \in \mathcal{A}_0,\ f = 0 \text{ or } div(f) + div(f_0) \geq 0\}$$
$$\cong \{\phi \in K(X_\Gamma) | \phi = 0 \text{ of } div(\phi) + div(f_0) \geq 0\} = L([div(f_0)])$$

The second part is proved similarly (see [Mi], p.58). □

Corollary 1.4.12. *$M_k(\Gamma)$ is always finite-dimensional.*

We can now find the dimensions of $M_k(\Gamma)$ and $S_K(\Gamma)$. we use the following corollary of the Riemann-Roch theorem: If d is a divisor for which $deg(d) > 2g - 2$ then $dim\ L(d) = deg(d) - g + 1$.
It turns out ([Mi], p.59) that for $k \geq 4$ we always have

$$deg([div(f_0)]) > deg([div(f_0)] - \sum_{v=1}^{t} b_v) > 2g - 2$$

and therefore, summing up all the facts, we get-

Theorem 1.4.13. *For an even $k \geq 4$, let g be the genus of X_Γ, $e_1, .., e_r$ the orders of elliptic points on X_Γ, and t the number of cusps of Γ, then*

$$dim(S_k(\Gamma)) = (k-1)(g-1) + \sum_{i=1}^{r} [\frac{k}{2}(1 - \frac{1}{e_i})] + (\frac{k}{2} - 1)t$$

$$dim(M_k(\Gamma)) = dim(S_k(\Gamma)) + t$$

It can be shown by similar arguments that for $k = 2$, we have $dim(M_k(\Gamma)) = g$ and $dim(S_k(\Gamma)) = g + t - 1$.

1.5 Eisenstein Series

In this section we construct modular forms for $SL_2(\mathbb{Z})$ called Eisenstein series. They give us the first example of connection between modular forms and number theory, since their q-expansion coefficients are values of arithmetic functions of elementary number theory.
Eisenstein series are a particular case of a more general construction called **Poincare' series**.
Let $k \geq 4$ be an even integer. We define for $z \in \mathcal{H}$

$$G_k(z) = \sum_{(0,0)\neq(m,n)\in\mathbb{Z}^2} \frac{1}{(mz+n)^k}$$

Theorem 1.5.1. *([Ko],pp.109-111)*

1. *The series for G_k is absolutely convergent, uniformly on every compact subset of $\mathcal{H}$*

2. *G_k is a modular form for weight k for $SL_2(\mathbb{Z})$*

3. *The q-expansion of G_k is*

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

where $\zeta(k) = \sum_{n=1}^{\infty} n^{-k}$, B_k (the Bernoully numbers) are given by the Taylor expansion $\frac{x}{e^x-1} = \sum_{k=0}^{\infty} B_k\frac{x^k}{k!}$, and $\sigma_k(n) = \sum_{d|n} d^k$.

We divide G_k by $2\zeta(k)$ to get the so-called normalized Eisenstein series

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

The first examples are

$$E_4(z) = 1 + 240\sum\sigma_3(n)q^n$$

$$E_6(z) = 1 - 504\sum\sigma_5(n)q^n$$

$$E_8(z) = 1 + 480 \sum \sigma_7(n) q^n$$

$$E_{10}(z) = 1 - 264 \sum \sigma_9(n) q^n$$

$$E_{12}(z) = 1 + \frac{65520}{691} \sum \sigma_{11}(n) q^n$$

$$E_{14}(z) = 1 - 24 \sum \sigma_{13}(n) q^n$$

Using the Eisenstein series and claim 1.3.2 we can form many more modular forms. For example, let us construct a cusp form of weight 12 for $SL_2(\mathbb{Z})$:

$$\Delta(z) = \frac{{E_4}^3 - {E_6}^2}{1728} = q - 24q^2 + 252q^3 \pm ...$$

As another application of Eisenstein series, we use the fact that $dim\ M_8(SL_2(\mathbb{Z})) = 1$ (which we get by using theorem 1.4.13 , see [Mi],p.99) and get by comparing the constant summand in the q-expansion that ${E_4}^2 = E_8$, and hence by comparing the coefficients of q^n we get

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{m=1}^{n-1} \sigma_3(m)\sigma_3(n-m)$$

This identity is quite striking, because it involves only elementary functions, but no elementary proof of it is known.

1.6 Modular forms for $\Gamma_0(2)$

Using theorem 1.4.13 we get $dim\ M_2(\Gamma_0(2)) = 1, dim\ M_4(\Gamma_0(2))) = 2, dim\ M_6(\Gamma_0(2)) = 2$, and $dim\ M_8(\Gamma_0(2)) = 3$.
Hence, since $X_{Gamma_0(2)}$ has two cusps, 8 is the minimal weight k for which $dim\ S_k(\Gamma_0(2)) > 0$. We first wish to describe this space.
We define on $\mathcal{H}$

$$\eta(z) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n)$$

where $q = e^{2\pi iz}, q^{1/24} = e^{\pi iz/12}$.
Then ([Sh], pp.49-50) $S_8(\Gamma_0(2))$ is one-dimensional and is generated by

$$f_8 = (\eta(z)\eta(2z))^8.$$

By an identity of Euler (see [HW],theorem 353),

$$\eta(z) = q^{1/24} \sum_{n=-\infty}^{\infty} (-1)^n q^{\frac{n(3n+1)}{2}} = q^{1/24}(1 - q - q^2 + q^5 + q^7 - q^{12} - q^{15} \pm ...).$$

Whence

$$f_8(z) = q \left(\sum_{n=-\infty}^{\infty} (-1)^n q^{\frac{n(3n+1)}{2}} \right)^8 \left(\sum_{n=-\infty}^{\infty} (-1)^n q^{n(3n+1)} \right)^8 = q - 8q^2 + 28q^3 - 56q^4 \pm ...$$

We will now give a generator for $M_2(\Gamma_0(2))$:
We use the construction of Eisenstein series of weight $k = 2$ described in ([Mi],7.2) and get that the function

$$E_2(z) = 1 + 24 \sum_{n=1}^{\infty} (\sigma_1(n) - 2\sigma_1(\frac{n}{2}))q^n$$

where $\sigma_k(n) = \sum_{d|n} d^k$ and $\sigma_1(\frac{n}{2})$ is 0 when n is odd, is in $M_2(\Gamma_0(2))$.
We now use E_2 and the ordinary Eisenstein series $E_4, E_6, ...$ (which are in the corresponding $M_k(SL_2(\mathbb{Z}))$) to give a basis to each $M_k(\Gamma_0(2))$:

- For $M_4(\Gamma_0(2))$: ${E_2}^2, E_4$. We know that they are independent by inspection of their q-expansions at infinity.
- For $M_6(\Gamma_0(2))$: ${E_2}^3, E_2E_4$.
  We also have there the Eisenstein series E_6. By comparison of q-expansions we get the identity
  $$4{E_2}^3 - 3E_2E_4 = E_6$$
- For $M_8(\Gamma_0(2))$: ${E_4}^2, {E_2}^4, {E_2}^2E_4$.
  Here we can express f_8 by
  $$576f_8 = 5{E_2}^2E_4 - 4{E_2}^4 - {E_4}^2$$

Theorem 1.6.1. *For every even $k \geq 2$, Every form in $M_k(\Gamma_0(2))$ can be written uniquely as a polynomial in E_2 and E_4*

Proof. We argue by induction on k. We have already proved our claim for $k = 2, 4, 6, 8$. Suppose that $k \geq 10$, $f \in M_k(\Gamma_0(2))$ and that the theorem is true for smaller weights.
We look at the linear subspace $f_8 \cdot M_{k-8}(\Gamma_0(2))$. Its elements are (by the induction hypothesis) polynomials in E_2, E_4, and its co-dimension is 2, by theorem 1.4.13
It remains to find two more forms f_1, f_2 which are polynomials in E_2, E_4 and span the quotient subspace.
For any form f in $M_k(\Gamma_0(2))$ we define $f(\infty)$ to be the constant element in the q - expansion of f and $f(0)$ to be the constant element in the q - expansion of $f|_{\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]_k}$
Since $(f \mapsto (f(0), f(\infty)))$ is a linear transformation and $f_8 \cdot M_{k-8}(\Gamma_0(2))$ is in its kernel, a sufficient condition for f_1, f_2 to span the quotient space is that the matrix $\begin{pmatrix} f_1(\infty) & f_1(0) \\ f_2(\infty) & f_2(0) \end{pmatrix}$ is regular.

Such two forms are $E_4 {E_2}^{\frac{k-4}{2}}$ and ${E_4}^2 {E_2}^{\frac{k-8}{2}}$. Since $E_4(\infty) = E_4(0) = 1$ (E_4 is a modular form for $SL_2(\mathbb{Z})$), we have only to show that the above matrix is regular for $f_1 = E_4 {E_2}^2, f_2 = {E_4}^2$. But we have seen that these two, together with f_8 span $M_8(\Gamma_0(2))$, and if the above matrix were not regular, the dimension of the image of $M_8(\Gamma_0(2))$ by $(f \mapsto (f(0), f(\infty)))$ would be smaller than two and hence, since the kernel of this transformation is the one-dimensional $S_8(\Gamma_0(2))$, we would get $dim\ M_8(\Gamma_0(2)) < 3$, which is a contradiction.
Let us prove now that this representation is unique. If not, then we have

$$\sum_{i,j \geq 0} a_{i,j} E_2^i E_4^j = 0$$

where not all the $a_{i,j}$ are zero, and the number of nonzero summands in the sum is finite.
Since the sum $\sum_{even\ k} M_k(\Gamma)$ is direct for any congruence modular group Γ, we get that for any positive integer m,

$$\sum_{i,j:2i+4j=2m} a_{i,j} E_2^i E_4^j = 0$$

Upon dividing by E_2^m we get

$$\sum_{2j \leq m} a_{m-2j,j} (\frac{E_4}{{E_2}^2})^j = 0$$

If not all the above $a_{i,j}$ are 0 then for all $z \in \mathcal{H}$, $\frac{E_4(z)}{E_2(z)^2}$ is a root of the polynomial $\sum_{2j \leq m} a_{m-2j,j} x^j$. But as we have proved, $\frac{E_4}{E_2{}^2}$ is a non-constant meromorphic function on $\mathcal{H}$ and hence assumes an ifinite number of values when $z \in \mathcal{H}$, whereas a nonzero polynomial has only a finite number of roots. □

1.7 Fundamental domains

Let Γ be a congruence modular group. Let F be a closed subset of $\mathcal{H}$ and let U be the interior of F.

Definition 1.7.1. *We call F a fundamental domain for Γ if:*

1. $F = \overline{U}$
2. $\mathcal{H} = \bigcup_{\gamma \in \Gamma} \gamma F$
3. *For $\gamma \in \Gamma \backslash Z(\Gamma)$, $\gamma U \cap U = \emptyset$*

Theorem 1.7.2. *([Ko],proposition III.1.1) the region*

$$F = \{z \in \mathcal{H} : -\frac{1}{2} \leq Re(z) \leq \frac{1}{2} \text{ and } |z| \geq 1\}$$

is a fundamental domain for $SL_2(\mathbb{Z})$

Corollary 1.7.3. *Any congruence modular group has a fundamental domain*

Proof. Let Γ be a congruence modular group, then since $[SL_2(\mathbb{Z}) : \Gamma] < \infty$, We can take representatives $\alpha_1, ..., \alpha_n \in SL_2(\mathbb{Z})$ so that $SL_2(\mathbb{Z}) = \bigcup_{i=1}^{n} \Gamma \alpha_i$. We set $F_2 = \cup \alpha_i F$, then

$$H = \bigcup_{\gamma \in \Gamma} \bigcup_{1 \leq i \leq n} \gamma \alpha_i F = \bigcup_{\gamma \in \Gamma} \gamma F_2$$

For $U_2 = int(F_2)$ we have

$$U_2 = int(\cup \alpha_i \overline{U}) = int(\cup \alpha_i U) = \cup \alpha_i U$$

The last equality is true since the $\alpha_i U$ are open and do not intersect each other. Hence

$$F_2 = \cup \alpha_i F = \cup \alpha_i \overline{U} = \cup \overline{\alpha_i U} = \overline{\cup \alpha_i U}$$

The last equality is true by the same consideration.
For $\gamma \in \Gamma \backslash Z(\Gamma)$ we have

$$\gamma U_2 \cap U_2 = \bigcup_{1 \leq i,j \leq n} \gamma \alpha_i U \cap \alpha_j U$$

and $\gamma \alpha_i U \cap \alpha_j U = \emptyset$ since if not, we would have ${\alpha_j}^{-1} \gamma \alpha_i U \cap U = \emptyset \Rightarrow {\alpha_j}^{-1} \gamma \alpha_i \in Z(\Gamma) \Rightarrow \gamma \alpha_i {\alpha_j}^{-1} \in Z(\Gamma) \Rightarrow \alpha_i {\alpha_j}^{-1} \in \Gamma$, Contradicting the definition of the α_i. □

1.8 The Petersson inner product

In this section we define a scalar product between elements of $M_k(\Gamma)$ and elements of $S_k(\Gamma)$ which will be invariant under the k- action of Γ.
Let Γ be a congruence modular group. let F be a fundamental domain for Γ, and $f, g \in M_k(\Gamma)$ with at least one of f, g in $S_k(\Gamma)$. We define

$$< f, g >= \frac{1}{[SL_2(\mathbb{Z}) : \pm\Gamma]} \int_F f(z)\overline{g(z)} y^k \frac{dxdy}{y^2}$$

(where $z = x + iy$).

Theorem 1.8.1. *1. The above integral is absolutely convergent and its value does not depend on the choice of F.*

2. *If Γ' is another congruence modular group such that $f, g \in M_k(\Gamma')$ and one of f, g is in $S_k(\Gamma')$ then the definition of $< f, g >$ is independent of whether we consider f, g in $M_k(\Gamma)$ or in $M_k(\Gamma')$.*

3. *$< , >$ satisfies the following conditions*

   - *$< , >$ is linear (over $\mathbb{C}$) in the first variable and anti-linear in the second ($< f, cg >= \overline{c} < f, g >$)*
   - *$< g, f >= \overline{< f, g >}$*
   - *$< f, f > > 0$ for $0 \neq f \in S_k(\Gamma)$*

4. *for $\alpha \in GL_2^+(\mathbb{Q})$, putting $\Gamma' = \alpha^{-1} \Gamma \alpha \cap \Gamma$ we have:*

   (a) Γ' is a congruence modular group

   (b) $[\alpha]_k$ takes $M_k(\Gamma)$ to $M_k(\Gamma')$, $S_k(\Gamma)$ to $S_k(\Gamma')$.

(c) $< f, g >=< f|_{[\alpha]_k}, g|_{[\alpha]_k} >$

For the proof see [Ko],chapter III, propositions 17,45,46,47.
The Petersson scalar product inspires an abstract definition of Eisenstein series for congruence modular groups rather than $SL_2(\mathbb{Z})$.

Lemma 1.8.2. *Let* $f \in S_k(\mathrm{SL}_2(\mathbb{Z}))$ *for an even* $k \geq 4$. *Then* $< G_k, f >= 0$

The lemma results from [Mi], theorem 2.6.10 and equality 4.1.5.

Definition 1.8.3. *For an arbitrary congruence modular group* Γ *we call any form in* $M_k(\Gamma)$ *which is perpendicular to all the elements of* $S_k(\Gamma)$ *an Eisenstein series.*

We denote by $N_k(\Gamma)$ the vector space of Eisenstein series of weight k for Γ. Then we have $dim\ N_k(\Gamma) = dim\ M_k(\Gamma) - dim\ S_k(\Gamma)$. For $k \geq 4$, by theorem 1.4.13, this equals the number of cusps of Γ.

1.9 Hecke operators

In this section we let $\Gamma = SL_2(\mathbb{Z})$.
We shall define an infinite sequence of linear operators $T(n)$ on $M_k(\Gamma)$. Our definition will be specific for Γ, although there is a more general construction which is valid for any congruence modular group.
let $f \in M_k(\Gamma)$.

Definition 1.9.1. $(T(n)f)(z) = \sum_{\substack{a\geq 1,\ ad=n \\ 0\leq b<d}} d^{-k} f(\frac{az+b}{d}) = \sum_{\substack{a\geq 1,\ ad=n \\ 0\leq b<d}} f|_{\left[\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right]_k}(z)$

Claim 1.9.2. $T(n)f \in M_k(\Gamma)$. *If additionally* f *is a cusp form then so is* $T(n)f$.

Obviously $T(n)$ is a linear operator.

Proposition 1.9.3. *1. All the* $T(n)$ *commute.*

2. $T(m)T(n) = T(mn)$ *whenever* $(m, n) = 1$
   $T(p)T(p^n) = T(p^{n+1}) + p^{k-1}T(p^{n-1})$ *for prime* p *and* $n \geq 1$

The proofs can be found in [Se], section VII.5.3.
We shall now discuss cusp forms which are eigenvectors for **all** the $T(n)$'s simultaneously.

Definition 1.9.4. *We say that $f \in M_k(\Gamma)$ is a Hecke eigenform if it is an eigenvector for all the $T(n)$'s.*

This phenomenon obviously happens when $dim S_k(\Gamma) = 1$, but it can be also shown that for all k, $S_k(\Gamma)$ has a basis of Hecke eigenforms.

Theorem 1.9.5. *([Se], section III.5.4) Let $f \in M_k(\Gamma)$ be a Hecke eigenform, $f(z) = \sum_{n=0}^{\infty} c(n)q^n$ and assume that $c(1) = 1$. then $\forall n \geq 1, T(n)f = c(n) \cdot f$.*

Together with proposition 1.9.3 we get-

Corollary 1.9.6. *Under the assumptions of the above theorem, we have*
$c(m)c(n) = c(mn)$ *whenever* $(m, n) = 1$
$c(p)c(p^n) = c(p^{n+1}) + p^{k-1}c(p^{n-1})$ *for prime* p *and* $n \geq 1$

Chapter 2

Congruences between Eisenstein series and cusp forms

Let B_k be the k^{th} Bernoulli number and $\frac{N_k}{D_k}$ the reduced fraction of $\frac{-B_k}{2k}$. In this chapter we show (following [DG]) that certain congruences between the Eisenstein series E_k and cusp forms hold mod. N_k.

2.1 The case $dim\ S_k(SL_2(\mathbb{Z})) = 1$

Definition 2.1.1. *Let $a, b \in \mathbb{Q}, 1 < N \in \mathbb{Z}$. We say that $a \equiv b(modN)$ if $a - b \in N\mathbb{Z}_{(N)}$, i.e, in the reduced fraction $a - b = \frac{n}{d}$, we have $(N, d) = 1$ and $N \mid n$*

For $k = 16, 18, 20, 22, 26$ we have $dim\ S_k(SL_2(\mathbb{Z})) = 1$.
We then have the following

Theorem 2.1.2. *For the weights k listed above, let f_k denote the unique cusp form of weight k normalized so that its first Fourier coefficient is 1. Then*

$$f_k \equiv \frac{-B_k}{2k} E_k\ modN_k$$

Proof. We have

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

We can find natural numbers t, s such that $4t + 6s = k$. Define $f = E_k - {E_4}^t {E_6}^s$, then f is a cusp form of weight k.

Since E_4, E_6 have integer coefficients we have $f \equiv E_k \ (mod\ 1)$ (by this we mean that $f - E_k$ has integer coefficients). Since $dim\ S_k(SL_2(\mathbb{Z})) = 1$, f is a multiple of f_k. by comparing the coefficient of q we get that $f = (-\frac{2k}{B_k} + a)f_k$ for some integer a. hence $f = (\frac{D_k}{N_k} + a)f_k$. Together we get

$$E_k \equiv (\frac{D_k}{N_k} + a)f_k(mod\ 1)$$

$$\Rightarrow \frac{N_k}{D_k}E_k \equiv (1 + a\frac{N_k}{D_k})f_k \equiv f_k(modN_k)$$

□

Corollary 2.1.3. *For $k = 12$ we have the cusp form*

$$\Delta(z) = \sum_{n=0}^{\infty} \tau(n)q^n$$

and (since $\tau(1) = 1$ and $N_{12} = 691$) we get the congruence

$$\tau(n) \equiv \sigma_{11}(n) \ \ mod\ 691$$

first observed by Ramanujan.

2.2 Congruences for $\Gamma_0(2)$

We wish to find a congruence which is similar to Ramanujan's congruence

$$\tau(n) \equiv \sigma_{11}(n) \ \ mod\ 691$$

but pertains to some $M_k(\Gamma_0(2))$, by a method similar to the one used in [DG].

Lemma 2.2.1. *$E_k(z)$ and $E_k(2z)$ form a basis to the space of Eisenstein series in $M_k(\Gamma_0(2))$*

Proof. $E_k(2z)$ is in $M_k(\Gamma_0(2))$ by [Ko],chapter 3,proposition 17.
$E_k(z)$ and $E_k(2z)$ are surely independent because the q-expansion coefficients of $E_k(2z)$ at infinity vanish for odd powers of q. □

We take for example $k = 12$. Let f be an Eisenstein series in $M_{12}(\Gamma_0(2)))$ with rational coefficients such that the denominators of its q-expansion coefficients are not divisible by 619 (otherwise we can't take f modulo 619).
Then f is a linear combination over $\mathbb{Q}$ of $E_{12}(z)$ and $E_{12}(2z)$.

$$f(z) = aE_{12}(z) + bE_{12}(2z)$$

moreover, the denominators of a and b are either not divisible by 619.
We may write

$$E_{12}(z) \equiv \Delta(z) \ (mod\ 619)$$

Meaning that the corresponding coefficients are congruent, and also

$$E_{12}(2z) \equiv \Delta(2z) \ (mod\ 619)$$

(note that also $\Delta(2z) \in S_{12}(\Gamma_0(2)))$. Therefore:

$$f(z) \equiv a\Delta(z) + b\Delta(2z) \ (mod\ 619)$$

Thus we have found a congruence between each such Eisenstein series and a cusp form in $S_{12}(\Gamma_0(2)))$. However, those congruences are not 'new' for us.

Chapter 3

l-adic Representations

In the previous chapter we obtained congruences between coefficients of Eisenstein series and those of cusp forms. In this chapter we find congruences among coefficients of cusp forms. To this end we state some results of the theory of l-adic representations which (in a sense) will enable us to find all such congruences.

3.1 l-adic representations for Hecke eigenvevtors in $S_k(SL_2(\mathbb{Z})) \cap \mathbb{Z}[[q]]$

Let l be a prime number. Denote by $\mathbb{K}_l$ the maximal algebraic extension of $\mathbb{Q}$ ramified only at l, and by $\mathbb{K}_l^{ab}$ the maximal subfield of $\mathbb{K}_l$ abelian over $\mathbb{Q}$.
By class-field theory, there is a canonical isomorphism

$$Gal(\mathbb{K}_l^{ab}/\mathbb{Q}) \cong \mathbb{Z}_l^*$$

This induces a canonical isomorphism

$$\chi_l : Gal(\mathbb{K}_l/\mathbb{Q}) \to Gal(\mathbb{K}_l^{ab}/\mathbb{Q}) \to \mathbb{Z}_l^*$$

For any prime $p \neq l$ denote by $Frob(p)$ any element of the conjugacy class of Frobenius elements of p in $Gal(\mathbb{K}_l/\mathbb{Q})$, then we have

$$\chi_l(Frob(p)) = p$$

Theorem 3.1.1. *Let $f(z) = \sum_{n=1}^{\infty} a_n q^n$ be a modular form in $S_k(SL_2(\mathbb{Z})) \cap \mathbb{Z}[[q]]$ such that $a_1 = 1$ and f is an eigenvector of all the Hecke operators T_n,*

Then there is a continuous homomorphism

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

such that for every $p \neq l$ we have:

$$trace(\rho_l(Frob(p))) = a_p$$

and

$$det(\rho_l(Frob(p))) = p^{k-1}$$

For the proof, see [De].
The above two equalities can be taken mod. l, and then the theorem relates to the induced map

$$\tilde{\rho}_l : Gal(\mathbb{K}_l/\mathbb{Q}) \to GL_2(\mathbb{Z}_l) \to GL_2(\mathbb{F}_l)$$

which satisfies the above equalities mod. l.
Since Frobenius elements are dense in $Gal(\mathbb{K}_l/\mathbb{Q})$ it is clear that the relation between $a_p \ mod\ l$ and $p^{k-1} \ mod\ l$ is determined by the image of ρ_l in $GL_2(\mathbb{F}_l)$.
The programme (of Serre and Swinnerton-Dyer) is as follows: First we obtain information on the subgroups of the finite group $GL_2(\mathbb{F}_l)$, and find which of them can be a continuous image of $Gal(\mathbb{K}_l/\mathbb{Q})$ satisfying $det(\tilde{\rho}_l(Frob(p))) = p^{k-1} \ mod\ l$. Then for each such subgroup we find the congruences it induces on a_p and p mod. l. And last, we develop a theory of *modular forms mod l* which enables us to eliminate most of the congruences, and have only a finite list of possible ones.

Definition 3.1.2. *We say that l is exceptional for f if the image of $\tilde{\rho}_l$ does not contain $SL_2(\mathbb{F}_l)$*

This definition is slightly different from the one used in [SwD], but the two definitions coincide for $l > 3$.

3.2 The possible images of $\tilde{\rho}_l$

First we classify the subgroups of $GL_2(\mathbb{F}_l)$.

Definition 3.2.1. *1. A Borel subgroup of $GL_2(\mathbb{F}_l)$ is any subgroup conjugate to the group of non-singular upper triangular matrices*

2. *A split Cartan subgroup is any subgroup of $GL_2(\mathbb{F}_l)$ conjugate to the group of non-singular diagonal matrices*

3. *A non-split Cartan subgroup is any subgroup of $GL_2(\mathbb{F}_l)$ conjugate in $GL_2(\mathbb{F}_{l^2})$ (but not in $GL_2(\mathbb{F}_l)$) to a cyclic group of order l^2-1 of diagonal elements.*

Lemma 3.2.2. *([SwD], p.12). Let G be a subgroup of $GL_2(\mathbb{F}_l)$. If the order of G is divisible by l, then either G is contained in a Borel subgroup of $GL_2(\mathbb{F}_l)$ or G contains $SL_2(\mathbb{F}_l)$.*
If the order of G is prime to l, let H be the image of G in $PGL_2(\mathbb{F}_l)$, then

1. *H is cyclic and G is contained in a Cartan subgroup, or*

2. *H is dihedral and G is contained in the normalizer of a Cartan subgroup but not in the Cartan subgroup itself (and the index of the Cartan subgroup in this normalizer is 2), or*

3. *H is isomorphic to A_4, S_4 or A_5*

In case 2, l must be > 2; in case 3, l must be > 3, > 3 or > 5 respectively.

Corollary 3.2.3. *([SwD], p.15). Suppose that l is exceptional for f.*
For the image G of $\tilde{\rho}_l$ in $GL_2(\mathbb{F}_l)$, only the following cases among the cases from the preceding lemma are possible:

1. *G is contained in a Borel subgroup*

2. *G is contained in the normalizer of a Cartan subgroup, but not in the Cartan subgroup itself*

3. *$H \cong S_4$*

For demonstration of the proof of the corollary, let us prove for example that the case where G is contained in a Cartan subgroup can be neglected.
If G is contained in a split Cartan subgroup then it is contained in a Borel subgroup. If it is contained in a non-split Cartan subgroup, say $G \subseteq C$. Since C is commutative, $\tilde{\rho}_l : Gal(\mathbb{K}_l/\mathbb{Q}) \to C$ factors through $Gal(\mathbb{K}_l^{ab}/\mathbb{Q}) \cong \mathbb{Z}_l^*$ (see lemma 4.3.6). Denote the map from $\mathbb{Z}_l^*$ to C by τ, then $ker\ \tau \supseteq (\mathbb{Z}_l^*)^{l^2-1}$. But $(\mathbb{Z}_l^*)^{l^2-1} = 1 + l\mathbb{Z}_l$ ($\subseteq$ is clear, and if $x \in 1 + l\mathbb{Z}_l$ then there is a solution in $\mathbb{Z}_l^*$ to the equation $y^{l^2-1} = x$, by Hensel's lemma). Hence

$$Im\ \tilde{\rho}_l = Im\ \tau \cong \mathbb{Z}_l^*/Ker\ \tau \subseteq (\mathbb{Z}_l^*)/(1 + \mathbb{Z}_l) \cong \mathbb{F}_l^*$$

Hence $Im\ \tilde{\rho}_l$ is cyclic of order dividing $l-1$, and hence it is contained in a split Cartan subgroup, and we are back in the previous case.

3.3 Classification of the possible congruences

Theorem 3.3.1. *Suppose that l is exceptional for f. then the three cases listed in corollary 3.2.3 imply respectively the following congruences:*

1. *There exists an integer m such that*

$$a_p \equiv p^m + p^{k-1-m} \ mod\ l$$

   for all the primes $p \neq l$

2. *$a_p \equiv 0$ mod l whenever p is a quadratic non-residue mod l*

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

Proof. The full proof is given is [SwD],p.17. We cite here the proof of case 1 in order to demonstrate the connection between $Im\ \tilde{\rho}_l$ and the congruences.
In case 1, G is contained in a Borel subgroup of $GL_2(\mathbb{F}_l)$. We may assume without loss of generality that the elements of G are upper triangular matrices. We write

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

Now α is a continuous homomorphism and must therefore (see lemma 4.3.6) be equal to $\tilde{\chi}_l^m$ for some integer m. Moreover $\alpha\delta = \tilde{\chi}_l^{k-1}$ by theorem 3.1.1, so that $\delta = \tilde{\chi}_l^{k-1-m}$. Taking $\sigma = Frob(p)$ we obtain

$$a_p \equiv trace(\tilde{\rho}_l(Frob(p))) = \alpha(\sigma) + \delta(\sigma) \equiv p^m + p^{k-1-m}\ (mod\ l)$$

□

3.4 Modular forms mod l

We denote by $\mathbb{Z}_{(l)}$ the local ring of $\mathbb{Q}$ at l, that is

$$\mathbb{Z}_{(l)} = \{\frac{m}{n} : m, n \in \mathbb{Z},\ l \nmid n\}$$

We can reduce elements of $\mathbb{Z}_{(l)}$ modulo l (we denote this by $\sim$).
We set $M_k = M_k(SL_2(\mathbb{Z})) \cap \mathbb{Z}_{(l)}[[q]]$, and

$$\tilde{M}_k = \{\sum_{n=0}^{\infty} \tilde{a}_n q^n : \sum_{n=0}^{\infty} a_n q^n \in M_k\}$$

We define $\tilde{M}$ = the sum of all the $\tilde{M}_k$ for even k. This sum (as we shall see) is not direct.
We write

$$P = 1 - 24\sum_{n=1}^{\infty}\sigma_1(n)q^n$$

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

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

P is not a modular form.
We define operators

$$\theta f = q\frac{df}{dq}, \quad \partial f = 12\theta f - kPf$$

where f is a modular form of weight k.

Lemma 3.4.1. *([SwD],p.19-20)*

1. *Let f be a modular form of weight k. then ∂f is a modular form of weight $k+2$.*
2. *∂ is a derivation on the graded algebra of modular forms such that $\partial Q = -4R$ and $\partial R = -6Q^2$.*

For a modular form $f = \sum a_n q^n$, let $\mathbb{A}$ be the additive group generated by the a_n, then f has a unique expression as an isobaric element of $\mathbb{A}[Q,\Delta] \oplus R\mathbb{A}[Q,\Delta]$ (An isobaric polynomial is a polynomial whose monomials are all of the same weight, where the weight of Q is 4, of Δ is 12 etc.). This can be easily seen because $\mathrm{SL}_2(\mathbb{Z})$ has only one cusp (hence $f - \mathbb{A}R$ includes a cusp form) and $S_k(\mathrm{SL}_2(\mathbb{Z})) = \Delta M_{k-12}(\mathrm{SL}_2(\mathbb{Z}))$ for $k \geq 16$.
Since $2^6 \cdot 3^3 \cdot \Delta = Q^3 - R^2$, the above f can be uniquely expressed as an isobaric element of $\mathbb{A}[\frac{1}{2},\frac{1}{3}][Q,R]$.

We need the following lemma on the l-adic nature of Bernoulli numbers

Lemma 3.4.2. *(von Staudt-Kummer) let k be an even positive integer.*

1. *If $(l-1) \mid k$ then $lb_k \equiv -1 \pmod{l}$*

2. *If* $(l-1) \nmid k$ *then* $\frac{B_k}{2k}$ *is* l*-integral and its residue class modulo* l *only depends on* $k \bmod (l-1)$

A proof can be found in [BS].

This lemma implies that for even $k = l \pm 1$ and $l > 3$, the coefficients of

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

are in $\mathbb{Z}_{(l)}$.
We denote by A and B the two isobaric polynomials in $\mathbb{Z}_{(l)}[Q, R]$ such that

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

Theorem 3.4.3. *([SwD],p.23).Suppose that* $l > 3$. *then*

1. $\tilde{A}(\tilde{Q}, \tilde{R}) = 1$ *and* $\tilde{B}(\tilde{Q}, \tilde{R}) = \tilde{P}$;
2. *The homomorphism* $\mathbb{F}_l[Q, R] \to \tilde{M}$ *given by* $C(Q, R) \mapsto C(\tilde{Q}, \tilde{R})$ *induces an isomorphism* $\tilde{M} \cong \mathbb{F}_l[Q, R]/(\tilde{A} - 1)$.

Corollary 3.4.4. 1. $\tilde{P} \in \tilde{M}_{l+1}$

2. θ *is a derivation on* $\tilde{M}$
3. *For* $0 \leq i \leq l-2$ *we have a filtration*

$$\tilde{M}_i \subseteq M_{i+(l-1)} \subseteq M_{i+2(l-1)} \subseteq \ldots$$

and if we define $\tilde{N}_i = \sum_{k=0}^{\infty} \tilde{M}_{i+k(l-1)}$ *then*

$$\tilde{M} = \sum_{i=0}^{l-2} \tilde{N}_i$$

is a direct sum.

Proof. (1) is clear since $\tilde{P} = \tilde{B}(\tilde{Q}, \tilde{R})$ and $\tilde{B}$ is isobaric of weight $l+1$.
(2) results from (1) and the definition of ∂.

The first part of (3) is because $\tilde{A}(\tilde{Q},\tilde{R}) = 1$ and $\tilde{A}$ is isobaric of weight $l-1$
Finally, suppose that

$$\sum_{i=0}^{l-2} \tilde{f}_i = 0,\ f_i \in \tilde{N}_i$$

Choose isobaric polynomials P_i such that $P_i(\tilde{Q},\tilde{R}) = \tilde{f}_i$.

Then by the last theorem there exists a polynomial C such that $\sum_{i=0}^{l-2} P_i = C(\tilde{A}-1)$.

We write $C = \sum_{i=0}^{l-2} C_i$ where C_i includes only monomials whose weight equals i modulo $l-1$. Then by comparing weights we get $P_i = (\tilde{A}-1)C_i$. Hence $\tilde{f}_i = (\tilde{A}(\tilde{Q},\tilde{R}) - 1)C_i(\tilde{Q},\tilde{R}) = 0$. □

Definition 3.4.5. *An element of one the $\tilde{N}_i$ is called a graded element.*

For a graded element $\tilde{f} \in \tilde{N}_i$ we define $\omega(\tilde{f}) = min\{k : \tilde{f} \in \tilde{M}_k\}$. It is clear by the last corollary that $\omega(\tilde{f}) \equiv i\ (mod\ l-1)$.

Lemma 3.4.6. *([SwD],p.25) let $\tilde{f}$ be a graded element, then $\omega(\theta\tilde{f}) \leq \omega(\tilde{f}) + l + 1$, with equality only if $l \nmid \omega(\tilde{f})$*

Corollary 3.4.7. *If $l \nmid \omega(\tilde{f})$ and t is a positive integer such that $(\omega(\tilde{f}))\ mod\ l) + t < l$, then*

$$\omega(\theta^t \tilde{f}) = \omega(\tilde{f}) + t(l+1)$$

Proof. Clear by consequent applications of the lemma. □

3.5 Extensive elimination of congruences

Using the theory of modular forms mod. l, we can now eliminate most of the possible congruences which have appeared in theorem 3.3.1, and in particular to prove that the number of exceptional primes for a modular form (satisfying the conditions of theorem 3.1.1) is finite.

Lemma 3.5.1. *case (1) of theorem 3.3.1 can happen only if either $2m < l \leq k+1$ or $m = 0$ and l divides the numerator of $\frac{B_k}{2k}$.*
case (2) can happen only if $l < 2k$.

Proof. We will show here the first part. The full proof can be found in [SwD],lemma 8. For the proof we define

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

This definition is not the same as in chapter 1 - The normalization is different. By lemma 3.4.2, G_k is in M_k whenever $(l-1) \nmid k$.

We have an identity of the form

$$a_p \equiv p^m + p^{m'} \ (mod\ l)$$

We can assume without loss of generality that $0 \leq m < m' < l-1$ and $m + m' \equiv k - 1 \ (mod\ l-1)$ ($m \neq m'$ since their sum is odd).
We conclude that

$$a_n \equiv n^m \sigma_{m'-m}(n) \ (mod\ l) \ \ \forall l \mid n$$

This is proved using the recursion formulae of corollary 1.9.6: First for $n = p^u$ (for a prime p) by induction on u and then using $a_{nm} = a_n a_m$ for $(n, m) = 1$.
Hence

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

We need the extra θ on each side to annihilate the coefficients of q^n for $l \mid n$.
(This is illegitimate only when $m = 0, m' = l-2$ for then the constant term in $G_{m'-m+1}$ is not l-integral. But then we have $a_p \equiv 1 + p^{l-2}$ and we can write instead $pa_p \equiv p + 1$, obtain as usual $na_n \equiv \sigma_1(n)$ and deduce that $\theta \tilde{f} = \theta^{l-1} \tilde{G}_2 = \theta^{l-1} \tilde{G}_{l+1}$. We will not give the details of this case (see [SwD]) since it is similar to the general case.)
We now look at the filtration ω of both sides. Assume that $l > k + 1$. On one hand, since $\omega(\tilde{f}) = k$ we have

$$\omega(\theta \tilde{f}) = \omega(\tilde{f}) + l + 1 = k + l + 1$$

on the other hand, by corollary 3.4.7 we have

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

By comparing we get

$$0 = ml + l + 2 + m' - k - l - 1 = ml + 1 + m' - k$$

Since $l > k$, we must have $m = 0$, $m' = l - 1$. Hence

$$\theta \tilde{f} = \theta \tilde{G}_k$$

We deduce that $\tilde{f} = \tilde{G}_k$ since if $l \nmid k$ then by lemma 3.4.6 θ is injective on $\tilde{M}_k$.
Now by examining the constant term we get that l divides the numerator of the constant term of G_k, which is $\frac{-B_k}{2k}$. □

Corollary 3.5.2. *For the 6 hecke eigenforms* $\Delta, Q\Delta, R\Delta, Q^2\Delta, QR\Delta, Q^2R\Delta$ *(of weights* $12, 16, 18, 20, 22, 26$ *respectively):*

1. *All the exceptional primes of the first case of theorem 3.3.1 are* < 23 *or prime divisors of the numerator of the corresponding* $\frac{B_k}{2k}$
2. *The only exceptional primes of the second case are* 23 *for* Δ *and* 31 *for* $Q\Delta$
3. *With the possible exception of* $l = 59$ *for* $Q\Delta$*, there are no exceptional primes of the third case.*

Proof. We have seen in the previous lemma how the range for the exceptional primes of the first and the second case is reduced. It is further reduced to the above range by numerical checking.
For the third case, for a given function $f(z) = \sum_{n=0}^{\infty} a_n q^n$, if l is exceptional for f then we have for any prime p, $p = l$ or l divides one of a_p^2, $a_p^2 - p^{k-1}$, $a_p^2 - 2p^{k-1}$, $a_p^2 - 4p^{k-1}$ and this gives a finite list of primes. Varying the prime p we can sieve out non-exceptional primes, and calculation shows that we are only left with $p = 59$ for $Q\Delta$. □

Chapter 4

Congruence properties of pairs of forms

For each of the six Hecke eigenforms $\Delta, Q\Delta, R\Delta, Q^2\Delta, QR\Delta, Q^2R\Delta$ we know by the previous chapter (up to one uncertainty) which are the exceptional primes, and what are the congruence relations for them.
It is natural to ask, what are the relations for *two* such forms.
As before, this will depend on the image of some l-adic representation modulo l. Now we have two representations, one for each form, and we are interested in their joint image in $GL_2(\mathbb{F}_l) \times GL_2(\mathbb{F}_l)$ (see below).
A basic tool in this analysis will be the following elementary group-theoretic lemma:

Lemma 4.0.3. *(Goursat's lemma) Let G_1 and G_2 be groups, and let H be a subgroup of $G_1 \times G_2$.*
Then there exist subgroups $A_1 \leq G_1$, $A_2 \leq G_2$ normal subgroups $B_1 \lhd A_1, B_2 \lhd A_2$ and an isomorphism $\phi : A_1/B_1 \to A_2/B_2$ such that $H = \bigcup_{\alpha \in A_1/B_1} \alpha \times \phi(\alpha)$.

Proof. Denote by π_1 and π_2 the projections from $G_1 \times G_2$ on G_1 and G_2.
set $A_1 = \pi_1(H)$, $A_2 = \pi_2(H)$

$$B_1 = \{x \in A_1 : \ (x, 1) \in H\}$$

and

$$B_2 = \{y \in A_2 : \ (1, y) \in H\}$$

Then for $x \in B_1$, $s \in A_1$, we have $(s, t) \in H$ for some $t \in A_2$, and hence

$$(x^s, 1) = (x^s, 1^t) = (x, 1)^{(s,t)} \in H \Rightarrow x^s \in B_1$$

Therefore $B_1 \lhd A_1$, and similarly $B_2 \lhd A_2$.
Now, for every $x \in A_1$ there exists $y \in A_2$ such that $(x, y) \in H$.
We define $f(x) = yB_2$, this is a legal definition since if $y' \in A_2$ satisfies $(x, y') \in H$ then by division

$$(1, y^{-1}y') \in H \Rightarrow y^{-1}y' \in B_2$$

One sees immediately that $f : A_1 \to A_2/B_2$ is a surjective homomorphism. we also have $Ker\ f = \{x : (x, 1) \in H\} = B_1$. By the first isomorphism theorem, f induces an isomorphism $\phi : A_1/B_1 \to A_2/B_2$ and by the definitions we have

$$H = \bigcup_{\alpha \in A_1/B_1} \alpha \times \phi(\alpha)$$

□

4.1 Definitions

Let $f_1, .., f_n$ be n distinct forms of our sextet $\Delta, Q\Delta, R\Delta, Q^2\Delta, QR\Delta, Q^2R\Delta$ (we are going to study the case $n = 2$), and let $l > 3$ be a prime.

Define

$$R(l; f_1, .., f_n) = \{(a_p^1 \ mod\ l, \dots, a_p^n \ mod\ l, p \ mod\ l) : p \neq l \ is \ prime\} \subseteq \{0, 1, .., l-1\}^{n+1}$$

$$S(l; f_1, .., f_n) = \{(a_p^1 \ mod\ l, \dots, a_p^n \ mod\ l) : p \neq l \ is \ prime\} \subseteq \{0, 1, .., l-1\}^{n}$$

Where $f_i(z) = \sum_{n=0}^{\infty} a_n^i q^n$, and k_i is the weight of f_i.

Let

$$\rho_i : Gal(\mathbb{K}_l/\mathbb{Q}) \to GL_2(\mathbb{F}_l)$$

be the associated Serre-Deligne representations of f_i, already taken modulo l.

Define $\rho : Gal(\mathbb{K}_l/\mathbb{Q}) \to GL_2(\mathbb{F}_l) \times ... \times GL_2(\mathbb{F}_l)$ by

$$\rho(x) = (\rho_1(x), .., \rho_n(x))$$

Define also

$$\mathcal{A} = \{(\sigma_1, .., \sigma_n) \in GL_2(\mathbb{F}_l) \times ... \times GL_2(\mathbb{F}_l) : \exists x \in \mathbb{F}_l, det(\sigma_i) = x^{k_i - 1}, \forall 1 \leq i \leq n\}$$

We have, by the properties of the associated representation, $Im\ \rho \subseteq \mathcal{A}$.

Definition 4.1.1. *Assume that l is non-exceptional for each f_i.*
We say that l is ***exceptional*** *for $(f_1, .., f_n)$ if $Im\ \rho \neq \mathcal{A}$*
Equivalently, l is ***non-exceptional*** *for $(f_1, .., f_n)$ if $SL_2(\mathbb{F}_l) \times .. \times SL_2(\mathbb{F}_l) \subseteq Im\ \rho$.*

Lemma 4.1.2. *If l is non-exceptional for $(f_1, .., f_n)$ then*

$$R(l; f_1, .., f_n) = \{0, 1, .., l-1\}^n \times \{1, .., l-1\}$$

Proof. We look at the image of

$$\rho \times \tilde{\chi}_l : \ Gal(\mathbb{K}_l/\mathbb{Q}) \to GL_2(\mathbb{F}_l) \times ... \times GL_2(\mathbb{F}_l) \times \mathbb{F}_l^*$$

For $l > 3$, the commutator of $SL_2(\mathbb{F}_l)$ is $SL_2(\mathbb{F}_l)$, hence the image of the commutator of $Gal(\mathbb{K}_l/\mathbb{Q})$ in $GL_2(\mathbb{F}_l) \times ... \times GL_2(\mathbb{F}_l) \times \mathbb{F}_l^*$ contains $SL_2(\mathbb{F}_l) \times ... \times SL_2(\mathbb{F}_l) \times \{1\}$. Since $\tilde{\chi}_l$ is onto $\mathbb{F}_l^*$, we get that the image of $\rho \times \tilde{\chi}_l$ contains $\mathrm{SL}_2(\mathbb{F}_l) \times ... \times SL_2(\mathbb{F}_l) \times \mathbb{F}_l^*$. Now for $(a_1, ..., a_n, a_{n+1}) \in \{0, 1, .., l-1\}^n \times \{1, .., l-1\}$ define

$$\mu = (\left(\begin{smallmatrix} a_1 & -1 \\ 1 & 0 \end{smallmatrix}\right), ..., \left(\begin{smallmatrix} a_n & -1 \\ 1 & 0 \end{smallmatrix}\right), a_{n+1})$$

Since Frobenius elements are dense in $Gal(\mathbb{K}_l/\mathbb{Q})$ and there exists a prime $p \neq l$ such that $\rho \times \tilde{\chi}_l(Frob(p)) = \mu$. Hence $a_p^i = trace\ \rho_i(Frob(p)) = trace\ \left(\begin{smallmatrix} a_i & -1 \\ 1 & 0 \end{smallmatrix}\right) = a_i$ and $p \equiv \tilde{\chi}_l(Frob(p)) = a_{n+1}\ (mod\ l)$, as required. □

4.2 Division into cases

From now on we focus on the case $n = 2$.

First we need a lemma on matrices:

Lemma 4.2.1. *Let $\mathbb{F}$ be a finite field, char $\mathbb{F} \neq 2$. Then $SL_2(\mathbb{F})$ does not have a subgroup of index 2*

Proof. It suffices to show that $S = \{g^2 : g \in SL_2(\mathbb{F})\}$ generates $SL_2(\mathbb{F})$ as a group. For every $x \in \mathbb{F}$ we have (since *char* $\mathbb{F} \neq 2$) $\left(\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}1 & 0\\ x & 1\end{smallmatrix}\right) \in S$. Now

$$\left(\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ y & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & z\\ 0 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}xy+1 & x\\ y & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & z\\ 0 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}xy+1 & xyz+x+z\\ y & yz+1\end{smallmatrix}\right)$$

Since more than half of the elements of $SL_2(\mathbb{F})$ are of this form, we are done. □

Throughout this section, we assume that l **is non-exceptional for** f_1 .
By Goursat's lemma, since $Im\ \rho \subseteq GL_2(\mathbb{F}_l) \times GL_2(\mathbb{F}_l)$, We have

$$Im\ \rho = \bigcup_{\alpha \in A_1/B_1} \alpha \times \phi(\alpha)$$

where $A_i = Im\ \rho_i$, $B_i \lhd A_i$, and $\phi : A_1/B_1 \to A_2/B_2$ is an isomorphism.
By assumption, we have

$$A_1 = Im\ \rho_1 = \{\sigma \in GL_2(\mathbb{F}_l) | det(\sigma) \in (\mathbb{F}_l)^{k_1}\}$$

In particular, $SL_2(\mathbb{F}_l) \lhd Im\ \rho_1 = A_1$.
We define $PB_1 = (\pm B_1)/\{\pm 1\}$.
By simplicity of $PSL_2(\mathbb{F}_l)$ for $l > 3$ ([Di]), and since $PB_1 \cap PSL_2(\mathbb{F}_l) \lhd PSL_2(\mathbb{F}_l)$, we may distinguish two cases:

- **Case 1:** $\pm B_1 \supseteq SL_2(\mathbb{F}_l)$. In this case, by lemma 4.2.1, we have $B_1 \supseteq SL_2(\mathbb{F}_l)$ and hence $SL_2(\mathbb{F}_l) \times \{\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\} \subseteq Im\ \rho$.

- **Case 2:** $B_1 \cap SL_2(\mathbb{F}_l) \subseteq \{\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right)\}$.

Definition 4.2.2. *In case 1 we say that f_1 and f_2 are* ***independent*** *modulo l*
In case 2 we call them ***dependent.***

We note that although the notion of dependence (or independence) of two objects in mathematics is usually symmetric, here there is an asymmetry because we require for the definition that l is non-exceptional *for* f_1. However, it will turn out from the following lemma that if l is also non-exceptional for f_2 then this definition is symmetric (i.e, f_1, f_2 (in)dependent $\Leftrightarrow$ f_2, f_1 (in)dependent).

Lemma 4.2.3. *1. If f_1 and f_2 are independent modulo l then*

(a) $R(f_1, f_2, l) = S(f_1, l) \times R(f_2, l) = \{0, 1, .., l-1\} \times R(f_2, l)$

(b) *If additionally l is non-exceptional for f_2 too, then l is non-exceptional for (f_1, f_2).*

2. *If f_1 and f_2 are dependent modulo l then*

   (a) *PB_1 is cyclic of order dividing $|\mathbb{F}_l^{*_1 - 1}|$*

   (b) *B_1 is composed of scalar matrices only.*

   (c) *l is non-exceptional for f_2, too.*

Proof. (1a) We have $SL_2(\mathbb{F}_l) \times \{\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\} \subseteq Im\ \rho$.
We look at the image of

$$\rho \times \tilde{\chi}_l : \ Gal(\mathbb{K}_l/\mathbb{Q}) \to GL_2(\mathbb{F}_l) \times GL_2(\mathbb{F}_l) \times \mathbb{F}_l^*$$

For $l > 3$, the commutator of $SL_2(\mathbb{F}_l)$ is $SL_2(\mathbb{F}_l)$, hence the image of the commutator of $Gal(\mathbb{K}_l/\mathbb{Q})$ in $GL_2(\mathbb{F}_l) \times GL_2(\mathbb{F}_l) \times \mathbb{F}_l^*$ contains $SL_2(\mathbb{F}_l) \times \{\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\} \times \{1\}$.
If $(x, y) \in R(f_2, l)$, and $z \in \mathbb{F}_l^*$ then there exists p such that $a_p^2 = x$ and $p = y$. Denote $\sigma = Frob(p)$. We have

$$(\rho_1(\sigma), \rho_2(\sigma), y) \in Im\ (\rho \times \tilde{\chi}_l)$$

we can find $\tau \in \mathrm{GL}_2(\mathbb{F}_l)$ such that

$$tr(\tau) = z,\ det(\tau) = det(\rho_1(\sigma))$$

We also have

$$(\tau, \rho_2(\sigma), y) \in Im\ (\rho \times \tilde{\chi}_l)$$

Since Frobenius elements are dense in $Gal(\mathbb{K}_l/\mathbb{Q})$, we can find a prime p_2 for which

$$(\rho \times \tilde{\chi}_l)(Frob(p_2)) = (\tau, \rho_2(\sigma), y)$$

Then modulo l, for $p = p_2$ we have

$$(a_p^1, a_p^2, p) = (tr(\tau), \mathrm{tr}(\rho_2(\sigma)), y) = (z, x, y)$$

this proves (1a).

If l is non-exceptional for f_2 then we also have $SL_2(\mathbb{F}_l) \subseteq B_2$ (i.e f_2 and f_1 are independent), since $A_1/B_1 \cong A_2/B_2$ and the other possibility for B_2 is too small, by part (2a) of this lemma.
Therefore we also have

$\{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)\} \times SL_2(\mathbb{F}_l) \subseteq Im\ \rho \Rightarrow SL_2(\mathbb{F}_l) \times SL_2(\mathbb{F}_l) \subseteq Im\ \rho.$

For (2a), *det* is an embedding $PB_1 \hookrightarrow \mathbb{F}_l^{*k_1-1}$.

Let us show that $SL_2(\mathbb{F}_l) \lhd A_2 = Im\ \rho_2$. We have $A_1/B_1 \cong A_2/B_2$, $A_1 \supseteq SL_2(\mathbb{F}_l)$ and $B_1 \subseteq \mathbb{F}_l^* I$. Hence, since $|SL_2(\mathbb{F}_l)| = l^3 - l$ we get that $l(l+1) \mid (|A_1/B_1|)$. If $SL_2(\mathbb{F}_l)$ is not contained in $Im\ \rho_2$ then by corollary 3.2.3, $Im\ \rho_2$ is either contained in a Borel subgroup, in the normalizer of a Cartan subgroup or is isomorphic to S_4. If it is contained in a Borel subgroup, then its size divides $l(l-1)^2$ (number of regular upper triangular matrices over $\mathbb{F}_l$), hence

$$l(l+1) \mid (|A_1/B_1|) \mid (|A_2/B_2|) \mid (|A_2|) \mid l(l-1)^2$$

This can not be for $p > 3$ since it implies $l+1 \mid (l-1)^2$ and hence $l+1 \mid 4l$.
If $Im\ \rho_2$ is contained in the normalizer of a cartan subgroup then its size divides $2(l^2-1)$ and similarly we get a contradiction. The remaining case is obvious. This proves (2c).

It remains to prove (2b). We pick a generator σ of PB_1. Since $B_1 \lhd A_1$, for each $x \in A_1$ there exists an integer n for which $\sigma^x = \sigma^n$.
We define $f(x) = n$, and note that f is a homomorphism from A_1 to the multiplicative group of $\mathbb{Z}/r\mathbb{Z}$, where $r = |PB_1|$. now

$$[A_1 : (Ker\ f \cap SL_2(\mathbb{F}_l))] \leq [A_1 : Ker\ f][A_1 : SL_2(\mathbb{F}_l)] \leq |(\mathbb{Z}/r\mathbb{Z})^*| l < rl \leq l^2$$

$$\Rightarrow |Ker\ f \cap SL_2(\mathbb{F}_l)| > \frac{|A_1|}{l^2} \geq \frac{|SL_2(\mathbb{F}_l)|}{l^2} = (l^3-l)/l^2 > 2$$

and hence (by simplicity of $PSL_2(\mathbb{F}_l)$- see [Di], and lemma 4.2.1), $SL_2(\mathbb{F}_l) \subseteq Ker\ f$. Hence σ commutes with all the elements of $PSL_2(\mathbb{F}_l)$ and is therefore scalar.

□

Corollary 4.2.4. *If l is non-exceptional for f_1 but is exceptional for f_2 then f_1 and f_2 are independent modulo l.*

4.3 Calculation of $R(l; f_1, f_2)$

We have 3 possibilities:

- l is exceptional for exactly one of f_1, f_2: In this case they are independent modulo l, and we can calculate $R(l; f_1, f_2)$ by lemma 4.2.3 - if (without loss of generality) l is non-exceptional for f_1 then $R(l; f_1, f_2) = \{0, 1, .., l-1\} \times R(f_2, l)$

- l is exceptional for both: This happens in what is called in [SwD] the first case. We can immediately obtain the set $R(l; f_1, f_2)$ since $a_p^1 \ mod \ l$ and $a_p^2 \ mod \ l$ are determined by $p \ mod \ l$.

- l is non-exceptional for both: We shall prove the following theorem

Theorem 4.3.1. *Suppose that l is non-exceptional for f_1 and for f_2. then*

1. *If $l > k_1 + k_2 + 1$ then l is non-exceptional for (f_1, f_2) (and hence $R(l; f_1, f_2) = \{0, 1, .., l-1\}^2 \times \{1, .., l-1\}$)*

2. *If $R(l; f_1, f_2) = \{0, 1, .., l-1\}^2 \times \{1, .., l-1\}$ then l is non-exceptional for (f_1, f_2)*

3. *If l is exceptional for (f_1, f_2) then either*

$$a_p^1 \equiv p^{\frac{k_1-k_2}{2}} a_p^2 \ (mod \ l)$$

or

$$a_p^1 \equiv \left(\frac{l}{p}\right) p^{\frac{k_1-k_2}{2}} a_p^2 \ (mod \ l)$$

We shall prove the theorem in the following pages.

Lemma 4.3.2. *If $l > 2$ and $R(l; f_1, f_2) = \{0, 1, .., l-1\}^2 \times \{1, .., l-1\}$ then l is non-exceptional for (f_1, f_2)*

Proof. We have $R(l; f_i) = \{0, 1, .., l-1\} \times \{1, .., l-1\}$ and hence by [SwD] l is non-exceptional for f_1 and for f_2.
By lemma 4.2.3 it suffices to show that f_1 and f_2 are independent mod l.
As before we have the decomposition

$$Im \ \rho = \bigcup_{\alpha \in A_1/B_1} \alpha \times \phi(\alpha)$$

By hypothesis, We can pick an element of $Im \ \rho_2$ with any trace in $\mathbb{F}_l$, and determinant in $\mathbb{F}_l^{*k_2-1}$. We will now pick such an element whose characteristic polynomial factors over $\mathbb{F}_l$ into two distinct linear factors.
To this end let us prove first that there exist $x \in \mathbb{F}_l^*$ such that $x^2 - 4 \in \mathbb{F}_l^{*2}$. This is trivial since

$$x^2 - 4 = y^2 \Leftrightarrow (x-y)(x+y) = 4$$

and we can take for example $x = 5/2, y = 3/2$
Take $\sigma_0 = \rho_2(\tilde{\sigma_0}) \in Im \ \rho_2$ for which $det\sigma_0 = 1$, $trace\sigma_0 = x$. Now every $\sigma = \rho_2(\tilde{\sigma}) \in$

$Im\ \rho_2$ with the same characteristic polynomial is conjugate in $GL_2(\mathbb{F}_l)$ to this σ_0, say $\sigma = \sigma_0^\alpha$. Hence

$$G = \{\beta : \sigma_0^\beta = \sigma\}$$

is a coset of

$$G_0 = \{\beta : \sigma_0^\beta = \sigma_0\}$$

We will now show that this coset intersects $Im\ \rho_2$:
We have

$$G \cap \text{Im } \rho_2 = \alpha G_0 \cap Im\ \rho_2 = \alpha(G_0 \cap \alpha^{-1}\text{Im } \rho_2)$$

Now our claim follows since G_0 includes all the scalar matrices, $Im\ \rho_2$ contains $\mathrm{SL}_2(\mathbb{F}_l)$ and also includes an element whose determinant is a quadratic non-residue mod. l (since $k_2 - 1$ is odd and $det(Im\ \rho_2) = (\mathbb{F}_l^*)^{k_2-1}$).

Hence we can assume $\alpha \in \text{Im } \rho_2$, say $\alpha = \rho_2(\tilde{\alpha})$.
We get:

$$\rho_2(\tilde{\sigma}) = \sigma = \sigma_0^\alpha = \rho_2(\tilde{\sigma_0})^{\rho_2(\tilde{\alpha})} = \rho_2(\tilde{\sigma_0}^{\tilde{\alpha}})$$
$$\Rightarrow \tilde{\sigma}(\tilde{\sigma_0}^{\tilde{\alpha}})^{-1} \in ker\ \rho_2$$

Suppose now that f_1 and f_2 are dependent mod l. We know that $\rho_1(ker\ \rho_2) = B_1$, and by lemma 2.3 we get that $\rho_1(\tilde{\sigma}(\tilde{\sigma_0}^{\tilde{\alpha}})^{-1})$ is a nonzero scalar matrix.

$$\exists b \in \mathbb{F}_l^* : \rho_1(\tilde{\sigma}) = b \cdot \rho_1(\tilde{\sigma_0}^{\tilde{\alpha}})$$
$$\Rightarrow trace(\rho_1(\tilde{\sigma})) = b \cdot trace(\rho_1(\tilde{\sigma_0}))$$

Surely there are at most $l - 1$ possibilities for $trace(\rho 1(\tilde{\sigma}))$, hence

$$|\{s : (s, x, 1) \in R(l; f_1, f_2)\}| < l$$

Contradicting our assumption. □

Theorem 4.3.3. *Suppose that l is non-exceptional for each of f_1, f_2 but is exceptional for (f_1, f_2). Then in the decomposition*

$$Im\ \rho = \bigcup_{\alpha \in A_1/B_1} \alpha \times \phi(\alpha)$$

there exists $\sigma \in \mathrm{GL}_2(\mathbb{F}_l)$ and a homomorphism $\psi : (\mathbb{F}_l^*)^{k_1} \to (\mathbb{F}_l^*)I/B_2$ such that

$$\phi(xB_1) = x^\sigma \cdot \psi(det\ x)$$

Proof. For each $i = 1, 2$ we have two possibilities: either $-1 \notin B_1$ and then we have a natural embedding as a normal subgroup $SL_2(\mathbb{F}_l) \hookrightarrow A_i/B_i$, or $-1 \in B_1$, and then $PSL_2(\mathbb{F}_l) \hookrightarrow A_i/B_i$.
By lemma 4.2.1 and a counting argument, SL_2 and PSL_2 can not be *both* embedded in the same A_i/B_i (if they were both embedded, then since the product of their sizes is bigger than the size of $GL_2(\mathbb{F}_l)$, the intersection of their images would be non-trivial, hence it would be the image of the simple group PSL_2, contradicting lemma 4.2.1) whence (since $A_1/B_1 \sim A_2/B_2$) we have $-1 \in B_1 \Leftrightarrow -1 \in B_2$.
Let G_i be the image of SL_2 (resp. PSL_2) in A_i/B_i. By lemma 4.2.1 and simplicity of PSL_2, we have $\phi(G_1) = G_2$.
Suppose first that $-1 \notin B_1$. Since every automorphism of $SL_2(\mathbb{F}_l)$ is induced by an inner automorphism of $GL_2(\mathbb{F}_l)$ (see [Di]), there exists $\sigma \in GL_2(\mathbb{F}_l)$ so that

$$\forall a \in SL_2(\mathbb{F}_l),\ \phi(aB_1) = a^\sigma B_2.$$

If $-1 \in B_1$, then again (Since every automorphism of $PSL_2(\mathbb{F}_l)$ is induced by an inner automorphism of $PGL_2(\mathbb{F}_l)$, and since $-1 \in B_2$), we get a σ with exactly the same property.
Now look at a general $x \in A_1$.
For each $a \in SL_2$, on one hand

$$\phi(a^x B_1) = \phi(aB_1)^{\phi(xB_1)} = (a^\sigma)^{\phi(xB_1)} B_2 = a^{\phi(xB_1)\sigma} B_2,$$

and on the other hand since $a^x \in SL_2$,

$$\phi(a^x B_1) = (a^x)^\sigma B_2 = a^{\sigma x} B_2$$

By comparing, we get

$$a^{x^\sigma \phi(xB_1)^{-1}} \in B_2$$

Since $det\ a = 1$ we get that

$$a^{x^\sigma \phi(xB_1)^{-1}} = \pm 1$$

thus, the image of $x^\sigma \phi(xB_1)^{-1}$ in PGL_2 commutes with all the elements of PSL_2 and hence is scalar.
So there exists a function $\psi : A_1/B_1 \to (\mathbb{F}_l^*)I/B_2$ such that

$$\phi(x) = x^\sigma \psi(x).$$

Moreover, ψ is an homomorphism and $\psi|_{G_1}$ is constantly 1, and from here we deduce the theorem. □

We can assume $\sigma = 1$, possibly after replacing ρ_2 by $\sigma\rho_2\sigma^{-1}$

Corollary 4.3.4. *Under the above assumptions (th. 4.3.3),*

$$Im\ \rho \subseteq \{(A, xA) : A \in GL_2(\mathbb{F}_l),\ x \in \mathbb{F}_l{}^*\}$$

Corollary 4.3.5. *Under the above assumptions (th. 4.3.3) only the two following congruences are possible (for all the primes $p \neq l$):*

1. $a_p^1 = p^{\frac{k_1-k_2}{2}} a_p^2\ mod\ l$
2. $a_p^1 = \left(\frac{l}{p}\right) p^{\frac{k_1-k_2}{2}} a_p^2\ mod\ l$

moreover, ρ is determined accordingly,

$$\rho_1(\sigma) = \tilde{\chi}_l^{\frac{k_1-k_2}{2}} \rho_2(\sigma)$$

or

$$\rho_1(\sigma) = \tilde{\chi}_l^{\frac{k_1-k_2+l-1}{2}} \rho_2(\sigma)$$

where $\tilde{\chi}_l$ is the canonical character

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

taken modulo l.

Proof. Define $\omega : Gal(K_l/\mathbb{Q}) \to \mathbb{F}_l{}^*$ by $\rho_1(\sigma) = \omega(\sigma)\rho_2(\sigma)$. Then ω is a homomorphism: $\omega(\sigma_1\sigma_2)\rho_2(\sigma_1\sigma_2) = \rho_1(\sigma_1\sigma_2) = \rho_1(\sigma_1)\rho_1(\sigma_2) = \omega(\sigma_1)\rho_2(\sigma_1)\omega(\sigma_2)\rho_2(\sigma_2) = \omega(\sigma_1)\omega(\sigma_2)\rho_2(\sigma_1\sigma_2)$.
We now need the following standard lemma:

Lemma 4.3.6. *Let C be a commutative group, and $\phi : Gal(K_l/\mathbb{Q}) \to C$ a homomorphism. Then ϕ factors through $Gal(K_l^{ab}/\mathbb{Q})$*
(i.e., if π denotes the restriction map from $Gal(K_l/\mathbb{Q})$ to $Gal(K_l^{ab}/\mathbb{Q})$, there exists a homomorphism $\psi : Gal(K_l^{ab}/\mathbb{Q}) \to C$ such that $\phi = \psi \circ \pi$. moreover, if ϕ is continuous then so is ψ).

Proof. Denote $A = ker\ \phi$, $L = K_l{}^A$. Since

$$Gal(L/\mathbb{Q}) = Gal(K_l/\mathbb{Q})/A \cong Im\ \phi \subseteq C$$

$L/\mathbb{Q}$ is an abelian extension, and by definition, $L \subseteq K_l^{ab}$, hence by Galois theory $ker\ \phi = A \supseteq Gal(K_l/K_l^{ab})$, therefore ϕ factors through $Gal(K_l/\mathbb{Q})/Gal(K_l/K_l^{ab}) \cong Gal(K_l^{ab}/\mathbb{Q})$. □

Remark 4.3.7. *This lemma immediately implies that $Gal(K_l^{ab}/\mathbb{Q})$ is the abelianisation of $Gal(K_l/\mathbb{Q})$.*

We get back to the proof of corollary 4.3.4.
Using the lemma, we can view ω as a continuous homomorphism from $Gal(K_l^{ab}/\mathbb{Q}) \cong \mathbb{Z}_l^*$ to $\mathbb{F}_l^*$. Hence there exists some integer m such that $\omega = \tilde{\chi}_l^m$. substitution gives $\rho_1 = \tilde{\chi}_l^m \rho_2$.
For a prime $p \neq l$ and $\sigma = Frob(p)$ we have:

- $\tilde{\chi}_l(\sigma) = p \pmod{l}$
- $\det\ \rho_i(\sigma) = p^{k_i}$
- $\det\ \rho_1(\sigma) = \tilde{\chi}_l(\sigma)^{2m} \det\ \rho_2(\sigma)$

From here we deduce that $p^{k_1} = p^{2m} \cdot p^{k_2}$ for all the primes $p \neq l$. Hence modulo l, $m = \frac{k_1 - k_2}{2}$ or $m = \frac{k_1 - k_2 + l - 1}{2}$.
Since $a_p^i = trace\ \rho_i(Frob(p))$, the results follow. □

Lemma 4.3.8. *If a congruence of the form*

$$a_p^1 \equiv p^r a_p^2\ (mod\ l)$$

holds for any prime $p \neq l$ and $2r \equiv (k_1 - k_2)\ (mod\ l-1)$, then we have

$$a_n^1 \equiv n^r a_n^2\ (mod\ l)$$

for all $p \nmid n$.

Proof. By corollary 1.9.6 we have $a_u^i a_v^i = a_{uv}^i$ whenever $(u, v) = 1$
$a_p^i a_{p^s}^i = a_{p^{s+1}}^i + p^{k_i - 1} a_{p^{s-1}}^i$ for prime p and $s \geq 1$ we first prove the assertion for $n = p^s$ by induction on s. For $s = 1$ we have it by assumption. assume the assertion true for $1, 2, ..s$ then

$$a_{p^{s+1}}^1 = a_p^1 a_{p^s}^1 - p^{k_1 - 1} a_{p^{s-1}}^1 =$$
$$\equiv p^r a_p^2 p^{rs} a_{p^s}^2 - p^{k_1 - 1} p^{r(s-1)} a_{p^{s-1}}^2 =$$
$$p^{r(s+1)} a_p^2 a_{p^s}^2 - p^{k_1 + r(s-1) - k_2} p^{k_2 - 1} a_{p^{s-1}}^2\ (mod\ l)$$

we have $r(s+1) \equiv (k_1 - k_2 + r(s-1))\ (mod\ l-1)$ and hence

$$a_{p^{s+1}}^1 \equiv p^{r(s+1)}(a_p^2 a_{p^s}^2 - p^{k_2 - 1} a_{p^{s-1}}^2)) = p^{r(s+1)} b_{p^{s+1}}^2$$

as desired.
Now the general conclusion follows easily from the fact that $a_u^i a_v^i = a_{uv}^i$ whenever $(u, v) = 1$ □

Corollary 4.3.9. *If $l > k_1 + k_2 + 1$ and l is non-exceptional for f_1 and for f_2 then it is non-exceptional for (f_1, f_2).*

Proof. By the previous corollary we have

$$a_p^1 \equiv p^{\frac{k_1 - k_2 + \epsilon(l-1)}{2}} a_p^2 \ (mod\ l)$$

where $\epsilon = 0$ or 1. By 4.3.8 this gives

$$a_n^1 \equiv n^{\frac{k_1 - k_2 + \epsilon(l-1)}{2}} a_n^2 \ (mod\ l),\ \forall l \nmid n$$

Therefore,

$$\theta \tilde{f}_1 = \theta^{\frac{k_1 - k_2 + \epsilon(l-1)}{2} + 1} \tilde{f}_2$$

Since

$$\omega(\tilde{f}_2) + \frac{k_1 - k_2 + \epsilon(l-1)}{2} + 1 \leq \frac{k_1 + k_2 + \epsilon(l-1)}{2} + 1 \leq$$
$$\leq \frac{l - 2 + \epsilon(l-1)}{2} + 1 < l$$

we get by corollary 3.4.7 that

$$\omega(\theta^{\frac{k_1 - k_2 + \epsilon(l-1)}{2} + 1} f_2) = k_2 + (l+1)(\frac{k_1 - k_2 + \epsilon(l-1)}{2} + 1)$$

on the other hand, by the same lemma, $\omega(\theta f_1) = k_1 + (l+1)$. therefore

$$k_1 - k_2 = (l+1)(\frac{k_1 - k_2 + \epsilon(l-1)}{2})$$

$$(k_1 - k_2)(l-1) + \epsilon(l+1)(l-1) = 0$$

$$k_2 - k_1 = \epsilon(l+1)$$

which is impossible since $l > k_1 + k_2$. □

4.4 Examples

For $l = 11$, by the table in [SwD], the only two forms for which l is non- exceptional are Δ and $QR\Delta$. By [SwD], p.24, we have

$$QR = 1\ (mod\ 11)$$

(by this we mean that the corresponding coefficients are congruent mod. 11) and hence

$$\Delta = QR\Delta \ (mod\ 11)$$

the weights of those forms are 12 and 22, and so (by Fermat's small theorem) this is an example of the second case in corollary 4.3.4:

$$a_p^1 = p^{\frac{22-12+11-1}{2}} a_p^2 \ (mod\ 11).$$

The only other examples are:

- $l = 13$:$a_p^1 = p \cdot a_p^2$ where $f_1 = Q\Delta, f_2 = Q^2R\Delta$
- $l = 13$:$a_p^1 = p^2 \cdot a_p^2$ where $f_1 = Q\Delta, f_2 = \Delta$
- $l = 13$:$a_p^1 = p \cdot a_p^2$ where $f_1 = Q^2R\Delta, f_2 = \Delta$
- $l = 17$:$a_p^1 = p^2 \cdot a_p^2$ where $f_1 = Q^2\Delta, f_2 = Q\Delta$
- $l = 19$:$a_p^1 = p^2 \cdot a_p^2$ where $f_1 = QR\Delta, f_2 = R\Delta$
- $l = 23$:$a_p^1 = p^2 \cdot a_p^2$ where $f_1 = Q^2R\Delta, f_2 = QR\Delta$

We know that there are no other examples since a computer check for all the pairs of forms and all $l \leq k_1 + k_2 + 1$ yealded that $R(l; f_1, f_2) = \{0, 1, .., l-1\}^2 \times \{1, .., l-1\}$. Let us now prove the above identities. We use the following easy lemma-

Lemma 4.4.1. *([SwD], lemma 6) suppose that $\tilde{f}_1$ and $\tilde{f}_2$ are both in $\tilde{M}_k$. then they are equal if and only if their first $[\frac{k}{12}] + 1$ coefficients coincide.*

Now suppose that we want to prove an identity of the form $a_p^1 \equiv p^r \cdot a_p^2 \ (mod\ l)\ \forall p \neq l$ where $r = \frac{k_1-k_2}{2}$ or $\frac{k_1-k_2+l-1}{2}$ (this is the case in all the above identities). Assume (for convenience) that $k_2 > k_1$. In view of lemma 4.3.8 we should try to prove that $a_n^1 = n^r \cdot a_n^2 \ (mod\ l)$ for all $l \nmid n$. That is, $\theta \tilde{f}_1 = \theta^{r+1} \tilde{f}_2$ by lemmas 3.4.1 and 3.4.6 we have:

$$\omega(\theta^{r+1} \tilde{f}_2) \leq k_2 + (r+1)(l+1)$$
$$\omega(\theta \tilde{f}_1) \leq k_1 + (l+1)$$
$$\omega(\theta \tilde{f}_1) \equiv \omega(\theta^{r+1} \tilde{f}_2) \ (mod\ l-1)$$

Hence it suffices to check that the identity holds for

$$n \leq \frac{(k_2 + (r+1)(l+1))}{12}$$

This has been verified for the above identities using a computer.

Bibliography

[BS] Z.I.Borevic and I.R.Safarevic: Number theory. (English translation, New York, 1966)

[De] P.Deligne: Formes modulaires et representations l-adiques, Seminaire Bourbaki, februar 1969, n.355

[Di] J.Dieudonne': La geometrie des groupes classiques. Berlin-Heidelberg-Gottingen: Springer 1955

[DG] B.Datskovsky and P.Guerzhoy: On Ramanujan Congruences for Modular Forms of Integral and Half-Integral Weights, Proc. A.M.S. 124 (1996), pp. 2283-2291

[HW] G.H.Hardy and E.M.Wright: An introduction to the theory of numbers

[Ko] Neal Koblitz: Introduction to Elliptic curves and Modular Forms, Springer-Verlag New York Berlin Heidelberg Tokyo

[Mi] Toshitsune Miyake: Modular Forms, Springer-Verlag Berlin Heidelberg 1989

[Se] J.P.Serre: Cours d'Arithmetique, Presses universitaires de France

[Sh] Shimura, G. (1971) Introduction to the arithmetic theory of automorphic functions, Iwanami Shoten and Princeton Univ. Press

[SwD] H.P.F Swinnerton-Dyer: On l-adic representations and congruences for coefficients of modular forms

[Ri] K.A.Ribet: On l-adic Representations Attached to Modular Forms