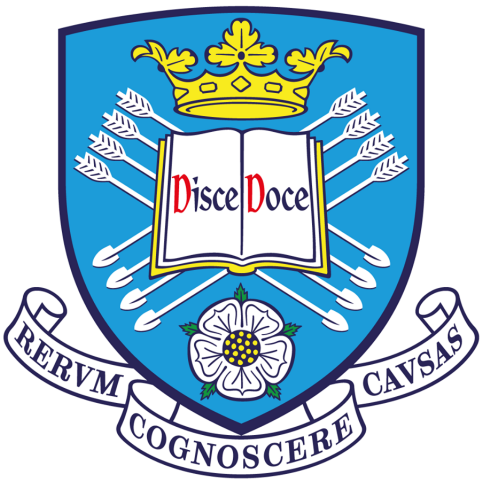


Congruences of Local Origin for Higher Levels

David Spencer

Submitted for the degree of Doctor of Philosophy
(Mathematics)

School of Mathematics and Statistics

June 2018

Supervisor: Prof. Neil Dummigan

University of Sheffield

Abstract

In this thesis we extend the work of Dummigan and Fretwell on congruences of "local origin". Such a congruence is one whose modulus is a divisor of a missing Euler factor of an L-function. The main congruences we will investigate are between the Hecke eigenvalues of a level N Eisenstein series of weight k and the Hecke eigenvalues of a level Np cusp form of weight k.

We first prove the existence of a congruence for weights $k \geq 2$. The proof will be an adaptation of the one used by Dummigan and Fretwell. We then show how the result can be further extended to the case of weight 1. The same method of proof cannot be used here and so we utilise the theory of Galois representations and make use of class field theory in order to prove the existence of a congruence in this case.

Inspired by an analogy with the weight 1 case, we prove the existence of a congruence between the Hecke eigenvalues of a weight k, level N cusp form and the Hecke eigenvalues of a paramodular Siegel newform of a particular level and weight. We will show how when $k = 2$ we end up with a scalar valued Siegel modular form and when $k > 2$ we end up with a vector valued Siegel modular form.

We will also consider the link with the Bloch-Kato conjecture in each case.

Contents

List of Tables

Introduction

There have been many interesting examples of congruences between modular forms studied over the years. Probably the first of such congruences that spring to mind is found in the work of Ramanujan. Ramanujan was among the first to begin studying modular forms in detail. He became quite interested in the discriminant function $\Delta(z)$ and spent a great deal of time studying its Fourier coefficients. These coefficients are denoted by $\tau(n)$. He stated many different results involving these coefficients, among which were certain congruences. The most famous in the literature is the following:

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

Here $\sigma_{11}(n) = \sum_{d|n} d^{11}$ is a power divisor sum. We therefore have values of a well known number theoretic function congruent to the coefficients of a rather mysterious function.

Naturally one might want to know why the modulus is 691 and also why $\sigma_{11}(n)$ appears. If, rather than simply studying Δ, we move to the larger space of all weight 12 modular forms we can explain both of these points. We see that the weight 12 Eisenstein series has the answer in its Fourier coefficients. We see both $\sigma_{11}(n)$ and $B_{12} = -\frac{691}{2730}$ appearing.

Of course there is actually something a little deeper going on here. The appearance of the Bernoulli number B_{12} is actually via the Riemann zeta function $\zeta(n)$. We see that 691 is a prime dividing the "rational part" of $\zeta(12)$. That is $691|\frac{\zeta(12)}{\pi^{12}} \in \mathbb{Q}$.

After the work of Ramanujan, many others became interested in congruences such as the Ramanujan 691 congruence. There are many generalisations of this congruence. Since a prime dividing the numerator of B_{12} gave a congruence between a weight 12 cusp form and a weight 12 Eisenstein series, one might naturally wonder whether, in general, a prime divisor of the numerator of B_k could be the modulus of a congruence between an Eisenstein series and a cusp form both of weight k and level 1. This is indeed the case.

We can also obtain congruences of "local origin". These are congruences between the Hecke eigenvalues of higher level cusp forms and the Hecke eigenvalues of level 1 Eisenstein series. What do we mean by "local origin"? Earlier we saw that a prime divisor of the numerator of the Bernoulli number B_k became the modulus of a congruence.

We can consider this as a divisor of the global Riemann zeta value. Recall that the Riemann zeta function has an Euler product and can be written as

$$\zeta(k) = \prod_p \frac{1}{1-p^{-k}}.$$

If we fix a particular prime p and omit the factor $(1-p^{-k})^{-1}$ and denote this new product by $\zeta_{\{p\}}(k)$ then

$$(1-p^{-k})^{-1}\zeta_{\{p\}}(k) = \zeta(k),$$

that is

$$p^k\zeta_{\{p\}}(k) = (p^k-1)\zeta(k).$$

We already know that primes dividing $\zeta(k)$ give level 1 congruences and it turns out that we should expect primes dividing p^k-1 to give level p congruences. This is exactly the result of Dummigan and Fretwell [DF].

Theorem 0.0.1. *Let p be prime and let $k \geq 4$ be an even integer. Suppose that $\ell > 3$ is a prime such that $ord_\ell((p^k-1)(B_k/2k)) > 0$, where B_k is the k-th Bernoulli number. Then there exists a normalised eigenform (for all T_q for primes $q \neq p$) $f = \sum_{n=1}^{\infty} a_n q^n \in S_k(\Gamma_0(p))$, and some prime ideal $\lambda|\ell$ in the field of definition $\mathbb{Q}_f = \mathbb{Q}(\{a_n\})$ such that*

$$a_q \equiv 1 + q^{k-1} \pmod{\lambda}$$

for all primes $q \neq p$

If ℓ divides $B_k/2k$ we may take $f \in S_k(\mathrm{SL}_2(\mathbb{Z}))$ and the congruence just becomes a generalisation of Ramanujan's congruence. The interesting case is when $\mathrm{ord}_\ell(B_k/2k) = 0$ and $\ell|(p^k-1)$; this is when, in general, we get a level p congruence. Such an ℓ occurs in the value at $s=k$ of the partial zeta function precisely because of the missing Euler factor. Following work of Harder [Har], using methods in Eisenstein cohomology, these congruences are known as congruences of local origin. We also note that the value $1+q^{k-1}$ appears since this is the Hecke eigenvalue when T_q is applied to E_k and so we have a congruence between a level p cusp form and a level 1 Eisenstein series.

One thing to note here is that we were starting at level 1 and then raising the level to level p. What about starting at level N? This is precisely the main work involved in this thesis. We first of all note that there is an analogue of Ramanujan's congruence at level N. That is, a congruence between the Hecke eigenvalues of a level N Eisenstein series and the Hecke eigenvalues of a level N cusp form. This is work of Dummigan in [D]. In this thesis we will be interested in the "local origin" analogue at higher levels. That is we wish to obtain a congruence between the Hecke eigenvalues of a level N Eisenstein series and the Hecke eigenvalues of a level Np cusp form. One difference now is that the Euler factors and moduli of congruences come from Dirichlet L-functions rather than the Riemann zeta function. In proving the congruence, we will adapt the proof used in [DF]. Thus we produce a linear combination of level Np Eisenstein series

that is a cusp form modulo ℓ and we will make use of the Deligne-Serre Lemma, among other results, to obtain a genuine characteristic zero cusp form of level Np congruent to a level N Eisenstein series. The main difficulty is obtaining a formula for the constant term at each cusp. In [DF] this is quite straightforward as there are only two cusps to consider. In the higher level case however there are many cusps. We note that Billerey and Menares prove a similar but weaker result to our main theorem in [BM]. In this paper they prove a formula for the constant term of the Eisenstein series at any cusp but one character is taken to be trivial. In this thesis we generalise this argument to the case where both characters can be non-trivial. They later proved the same constant term formula in [BM2] using a different method to the one in this thesis.

Once we have dealt with the case of weight $k \geq 2$ we move onto the more exotic case of weight 1. A lot of standard results about modular forms fail to hold in the case of weight 1. For example, there are no longer formulas for the dimensions of spaces of modular forms of a particular level. Also the usual Riemann-Roch theory fails and therefore we cannot apply the Deligne-Serre lemma in the same way that we will for weights $k \geq 2$. This forces a different approach. The one we will take is via Galois representations and class field theory. The results given in this section will be very similar to those of higher weight, however there will be certain restrictions.

One of the main results that we will rely on in this thesis is the case of Serre's modularity conjecture proven by Khare and Wintenberger. This result tells us how, given an odd, irreducible two-dimensional complex linear representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, we have an associated weight 1 cusp form. We note that, although we use this approach, the f that we will construct is an explicit theta series. Such series are known to be modular forms and so we actually didn't need the result of Khare and Wintenberger. Our main task will be to try to cook up a particular two-dimensional representation whose associated cusp form satisfies a congruence with a weight 1 Eisenstein series. In the case of weight 1 this basically boils down to the Hecke eigenvalue at q of the cusp form f being congruent to a sum of character values $\psi(q)$ and $\varphi(q)$ modulo some λ. It turns out that ψ and φ must be related to the character associated to an imaginary quadratic field K. The results are separated into three theorems. Firstly we consider the case where one character is trivial since this turns out to be more straightforward whilst also laying out most of the groundwork for the more general result. We then look at the general case where both characters could be non-trivial. This case is similar but we make an adjustment in the method. This adjustment will have the effect of increasing the level of the cusp form by the conductor of a particular character. We will then go on to prove a theorem stating that these are the only such local origin congruences in the case of weight 1. This will involve carefully considering the remaining cases and showing that any possible congruence actually leads to a contradiction.

In order to produce the two-dimensional representation required to prove our theorems, we will induce a one-dimensional representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$. We note that one way we could get the desired congruence (in the case where one character is trivial) is by inducing the trivial representation. This would give a representation with trace $1+\eta(q)$

when evaluated at a Frobenius element at q. Note that this is exactly what we want on the right hand side of the congruence. This however would give a reducible two-dimensional representation and we therefore would not be able to utilise the result of Khare-Wintenberger. We will instead end up inducing a non-trivial ray class character that is congruent to the trivial character modulo λ. Note that this induction is actually level raising for GL(1).

We then finish off the thesis by generalising this argument. By instead inducing a two-dimensional representation we will obtain a Siegel paramodular newform. This will then, via a chain of results, satisfy a congruence with a weight k cusp form of level N. We will first consider the scalar valued case. The method will involve beginning with a weight 2 classical cusp form of level N with associated Galois representation, ρ say, of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. We then consider the restriction, that is we view it as a representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$ where K is a real quadratic field. This is analogous to considering the trivial representation in the weight 1 case. Likewise, because this is obtained by restriction, inducing would give a reducible representation. We therefore replace this with a congruent representation that is trivial modulo λ. This representation will then be induced to give a Siegel paramodular newform. If we instead consider this process purely in terms of what is happening to the modular forms, we first begin with a classical cusp form and take a base change to a Hilbert modular form. We then use a level raising result proven by Taylor which is a generalisation of Ribet's result for classical forms. Finally we take a theta lift to get the Siegel paramodular newform. In the case where we begin with a weight k $(k > 2)$ classical cusp form, we will actually end up producing a vector valued Siegel modular form that satisfies a congruence. The method is largely the same. When we raise the level of the Hilbert modular form however, we can no longer use the same theta lifting result. At this stage we have to generalise the argument.

Notice that these results involve a congruence between two cusp forms rather than a cusp form and an Eisenstein series. Although this is quite different to the previous results, and involves a lot of new material, we will in fact see quite a few similarities between the two cases. For example we will see how the correct L-function to be considering here is a symmetric square L-function. We will also see how there is an analogue of the Euler factor in this case.

In each case that we consider, we will show that our results agree with the Bloch-Kato conjecture. In particular we will look in detail at what the Bloch-Kato conjecture says for the weight $k \geq 2$ case and the Siegel case. In the weight 1 case, although much of the theory about modular forms is more difficult, the Bloch-Kato conjecture is actually more straightforward. The Bloch-Kato conjecture is a far-reaching generalisation of results such as the analytic class number formula and the Birch and Swinnerton-Dyer conjecture. The Bloch-Kato conjecture gives us information about values at integer points of L-functions associated to motives. In this thesis we won't consider the full generality of the conjecture, in particular, we will not cover material on motives. In the case of our main theorem, it will involve Dirichlet L-functions and in the Siegel case,

as mentioned, it will involve symmetric square L-functions. The conjecture relates divisibility of certain L-values to divisibility of the order of certain Bloch-Kato Selmer groups. The existence of congruences such as the ones in this thesis allow us to construct non-trivial elements in these Bloch-Kato Selmer groups (or Tate-Shafarevich groups). In general this is quite interesting since not much is known about these groups. We note that the conjecture is known in the case of Dirichlet L-functions; it was proven by Huber and Kings.

Chapter 1 contains an overview of the theory of classical (elliptic) modular forms. Here we will see definitions of all the necessary objects including: Eisenstein series, cusp forms, Hecke operators, congruence subgroups, etc.

Chapter 2 contains a brief overview of (global) class field theory. We begin with a brief recap of the basics from algebraic number theory that are required. We then move on to discuss inertia groups, decomposition subgroups and Frobenius elements. We then state the main theorems of global class field theory: Artin Reciprocity and the Existence Theorem. We finish off with a discussion of ray class fields and the Hilbert class field as this will be essential in the work on weight 1.

Chapter 3 contains the work on local origin congruences for weights $k \geq 2$. As previously mentioned the main bulk of work in this section is proving a formula for the constant term of an Eisenstein series at any cusp. The method used is an adaptation of the method used to prove a simpler case in [BM]. Although Billerey and Menares later generalised the result in [BM2], they use a different method of proof to the one we use. They also go a little further and show that the formula holds in the case of weight 2. A discussion of this work is included. We then move on to prove the main theorem for weights $k \geq 2$. A discussion of Katz modular forms is required for this. The section is finished off with a comparison with the Bloch-Kato formula for a partial L-value.

Chapter 4 contains the weight 1 case. We begin by stating a simplified result whose proof will lay the groundwork for a generalisation. This simplified result is also most analogous to the higher weight case. We also state the generalised version of this result along with a statement that there are no other such congruences for weight 1. We then move onto background material on Galois representations and the modularity results necessary to link these representations with modular forms. We next discuss results on dihedral representations as these are the ones that lead to a congruence. Once we have all the necessary results we discuss comparisons with the higher weight case while proving various different cases of the simplified theorem. We then give a sketch proof of the generalised theorem pointing out the minor differences in the proof. We finish the chapter by discussing the remaining classes of projective representations. We use Schur covers to instead consider linear representations and then show that these could not be reducible modulo p for any p. Hence there are no other possibilities for a congruence.

Chapter 5 contains the generalisation of the method used in the weight 1 case. We begin with background material on Hilbert modular forms and Siegel modular forms. We then discuss base change from a classical weight 2 cusp form to a Hilbert modular form of parallel weight 2 over a real quadratic field K. This section involves a result

about the level of such a base change along with a conjecture about a certain case. We then discuss level raising congruences in both the classical case, as studied by Ribet, and the Hilbert case, as studied by Taylor. Next we discuss a theta lifting result of Johnson-Leung and Roberts that can be used to lift a Hilbert modular form to a Siegel paramodular newform. We then show that each of these results can be combined to give the existence of a congruence between a classical cusp form of weight 2 and a Siegel paramodular newform of weight 2. We also give an example using this method. We compare this with the Bloch-Kato conjecture in a similar way as in chapter 3. After this we consider the symmetric square L-function and show that in fact you can consider an Euler factor of this L-function in the same way as in chapter 3. That is, a prime dividing the Euler factor should give the modulus of a congruence. We also give an alternate argument for constructing a non-zero element in a Bloch-Kato Selmer group. This argument is similar to ones used in the theory of Galois deformations. Finally, to close out the chapter, we generalise the method to obtain a congruence between a weight k classical cusp form and a vector valued Siegel modular form. This involves introducing background material on L-parameters and L-packets along with the notion of a limit of discrete series representation. We then make use of the Satake isomorphism and Satake parameters in order to determine the Hecke eigenvalues of our Siegel modular form.

Acknowledgements

Over the years I have had the support and encouragement of many people, some of whom without I would not have had the confidence to pursue a PhD.

I would like to begin by thanking my supervisor Neil Dummigan. I am grateful for the many discussions we have had where his guidance has helped steer me towards many a solution where there seemingly were none. I would particularly like to thank him for his encouragement and patience while working through some of the more technical aspects of this thesis.

Secondly I would like to thank Dan Fretwell, without whom I would not have become a postgraduate student! His help, advice and encouragement throughout my undergraduate degree was the catalyst for me pursuing a PhD. I am also thankful for the many helpful discussions we have had throughout my time as a postgraduate student.

I would also like to thank the Number Theory group at Sheffield. When I began my PhD there were only a few students in the group but I am pleased to say that these numbers have more than doubled during my time as a postgraduate student. This helped to create an environment where there was always an opportunity to discuss number theory. I am grateful for the many seminar series run by both staff and students that have taken place over the last four years. They have helped to deepen my knowledge and understanding of the subject.

Last, but by no means least, I would like to thank my parents for their constant support and encouragement throughout my life. Their belief in me has never faltered.

Chapter 1

Classical Modular Forms

In this section we will see an overview of the results we will need from the theory of classical (elliptic) modular forms. We will see various definitions, examples and important results along the way. Many standard references will be used with [DiSh] being a primary source.

Although modular forms are objects belonging to the world of analysis, they still have significant uses in number theory. They are functions, defined on the upper half plane $\mathcal{H}$, which are "almost" invariant under an action by the matrix group $\mathrm{SL}_2(\mathbb{Z})$ and satisfy certain holomorphy conditions. By studying modular forms we can often find identities and congruences of number theoretical significance. They also played a part in the proof of Fermat's Last Theorem as there is a connection between elliptic curves and modular forms (The Modularity Theorem). We will briefly mention some of the links between modular forms and Elliptic curves but this is not a primary focus. There are many different types of modular form, such as Hilbert modular forms, Maass forms, Siegel modular forms and automorphic forms, but here we will only be concerned with classical modular forms. For more details on Hilbert modular forms see Section 5.1 and for details on Siegel modular forms see Section 5.2.

§ 1.1 First Definitions and Examples

Recall $\mathcal{H} = \{a + ib \in \mathbb{C} \,|\, b > 0\}$, the upper half plane. We may define an action by $\mathrm{SL}_2(\mathbb{R})$ on $\mathcal{H}$. Take $\tau \in \mathcal{H}$ and $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$. The matrix γ then acts on τ by a fractional linear transformation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(\tau) = \frac{a\tau + b}{c\tau + d}, \quad \tau \in \mathcal{H}.$$

Lemma 1.1.1. *The group* $\mathrm{SL}_2(\mathbb{R}) = \{A \in \mathrm{M}_2(\mathbb{R}) \,|\, \det(A) = 1\}$ *acts transitively on the upper half plane* $\mathcal{H}$.

Proof. We consider the orbit of the element $i \in \mathcal{H}$. We will show that i has full orbit. Let $\tau = a + bi$. Let $\gamma = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{b} & 0 \\ 0 & \sqrt{b}^{-1} \end{pmatrix} = \begin{pmatrix} \sqrt{b} & a\sqrt{b}^{-1} \\ 0 & \sqrt{b}^{-1} \end{pmatrix}$. Then $\gamma(i) = \frac{\sqrt{b}i + a\sqrt{b}^{-1}}{\sqrt{b}^{-1}} = a + bi = \tau$. □

Suppose we try to consider functions invariant under this action. So we are after functions $f : \mathcal{H} \to \mathbb{C}$ such that $f(\gamma z) = f(z)$ for all $\gamma \in \mathrm{SL}_2(\mathbb{R})$. The transitivity of this action tells us that the only such functions are the constant functions. In order to obtain results of number theoretical importance it makes sense to instead restrict to an action by the so called modular group, a subgroup of $\mathrm{SL}_2(\mathbb{R})$. The *modular group* is the group of 2×2 matrices with integer entries and determinant 1,

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z}, ad - bc = 1 \right\}.$$

This group is generated by the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We will see shortly that these matrices give some important properties that a modular form must satisfy. Now that we have an action on the upper half plane $\mathcal{H}$, it makes sense to ask what it means to be modular.

Definition 1.1.2. Let k be an integer. A meromorphic function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is *weakly modular of weight* k if

$$f(\gamma(\tau)) = (c\tau + d)^k f(\tau) \quad \text{for} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \quad \text{and} \quad \tau \in \mathcal{H}. \tag{1.1}$$

Notice that we are no longer asking for invariance but are only interested in invariance up to a factor of $(c\tau + d)^k$. Note that weak modularity of weight 0 is exactly $\mathrm{SL}_2(\mathbb{Z})$ invariance. We will not be interested in the theory of weakly modular functions (those invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$) but we make a brief remark about the history of these functions. The first example of a weakly modular function arose in the theory of elliptic curves. Recall that an elliptic curve over a field K is a non-singular projective curve E/K of genus 1 over K together with a point $\mathcal{O} \in E(K)$.

Suppose $K \subset \mathbb{C}$. Then $E(\mathbb{C})$ is a Lie group and has the structure of a torus, i.e., there is an isomorphism of Lie groups $E(\mathbb{C}) \cong \mathbb{C}/\Lambda$ for some lattice Λ. The details of this can be found in chapter 6 of Silverman [Sil]. It is well known that the isomorphism classes of elliptic curves are determined by the j-invariant of the curve. Suppose we have a curve with $\Lambda = \mathbb{Z} \oplus z\mathbb{Z} = \langle 1, z \rangle$. If we define $j(z)$ to be the j-invariant of the curve then it turns out that $j(z)$ is actually a weakly modular function. This shows that there are non-trivial examples of weakly modular functions, however this is actually the only interesting one.

Theorem 1.1.3. *The $\mathbb{C}$-algebra of weakly modular functions is isomorphic to $\mathbb{C}(j)$.*

A proof of this result is given on page 73 of [DiSh].

We therefore look for functions which are weakly modular of weight $k > 0$. We may consider the action of $-I$ where I is the identity matrix. Letting $\gamma = -I$ in (1.1) we see that $f(\tau) = (-1)^k f(\tau)$, showing that there are no weakly modular functions of odd weight except the zero function. By considering the action of the generators of the modular group we obtain two properties that any weakly modular function for $\mathrm{SL}_2(\mathbb{Z})$ must satisfy. We have

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}(\tau) = \tau + 1 \quad \text{and} \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}(\tau) = -1/\tau,$$

so in particular

$$f(\tau + 1) = f(\tau) \quad \text{and} \quad f(-1/\tau) = \tau^k f(\tau).$$

The first of these conditions tells us that f is $\mathbb{Z}$-periodic. If in addition we know that f is holomorphic on $\mathcal{H}$ and also at ∞, it follows that f must have a Fourier expansion. This expansion, often referred to as the q-expansion, is given by

$$f(\tau) = \sum_{n=0}^{\infty} a_n(f) q^n, \quad q = e^{2\pi i \tau}.$$

Before giving the full definition of a modular form we introduce some standard notation that will become useful later.

Definition 1.1.4. For $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ define the *factor of automorphy* $j(\gamma, \tau) \in \mathbb{C}$ for $\tau \in \mathcal{H}$ to be $j(\gamma, \tau) = c\tau + d$. For $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and any integer k define the *weight-k slash operator* $[\gamma]_k$ on functions $f : \mathcal{H} \to \mathbb{C}$ by

$$(f[\gamma]_k) = j(\gamma, \tau)^{-k} f(\gamma(\tau)), \quad \tau \in \mathcal{H}.$$

Remark 1.1.5. The condition of being weakly modular can now be written as $f[\gamma]_k = f$.

We are now ready to give the full definition of a modular form.

Definition 1.1.6. Let k be an integer. A function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is a *modular form of weight k* if

(1) f is holomorphic on $\mathcal{H}$,

(2) f is weakly modular of weight k,

(3) f is holomorphic at ∞.

The set of modular forms of weight k is denoted by $M_k(\mathrm{SL}_2(\mathbb{Z}))$. Note that being holomorphic at ∞ means that $\lim_{\mathrm{Im}(\tau)\to\infty} f(\tau)$ exists and is finite.

The first thing we might ask about this definition is whether such a thing even exists. Note that the j-invariant is not a modular form of weight 0 as it is not holomorphic at ∞; it has a simple pole there (it has q-expansion $j(z) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + 864299970q^3 + 20245856356q^4 + \dots$). In fact the condition of being a modular form of weight 0 is so nice that there are no interesting examples.

Corollary 1.1.7. *The $\mathbb{C}$-algebra of modular forms of weight 0 is isomorphic to $\mathbb{C}$, i.e., the only modular forms of weight 0 are the constant functions.*

We may also exclude a number of possible weights at this point.

Lemma 1.1.8. *If $k < 0$ or k is odd then we have $M_k(\mathrm{SL}_2(\mathbb{Z})) = \{0\}$.*

Proof. It is clear that if $k < 0$ then f would not be a holomorphic function. We also know that since $-I \in \mathrm{SL}_2(\mathbb{Z})$ the function f would have to satisfy $f(\tau) = (-1)^k f(\tau)$. Hence k cannot be odd. □

Fortunately there are many examples of modular forms; those functions which are only invariant up to a factor of $(c\tau + d)^k$ and satisfy the necessary holomorphy conditions. The simplest of these is the Eisenstein series.

Definition 1.1.9. Let $k > 2$ be an even integer and define the *Eisenstein series of weight k* to be

$$G_k(\tau) = \sum_{\substack{c,d \in \mathbb{Z} \\ (c,d) \neq (0,0)}} \frac{1}{(c\tau + d)^k}, \quad \tau \in \mathcal{H}.$$

Notice that this is a 2-dimensional analogue of the Riemann zeta function $\zeta(k) = \sum_{n=1}^{\infty} 1/n^k$. It is fairly easy to show that this series satisfies the necessary conditions to be a modular form and that its q-expansion is given by

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

where the coefficient $\sigma_{k-1}(n)$ is the power divisor function

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

The details of this can be found on pages 4-5 of [DiSh].

Notice that the coefficients of this q-expansion contain many things of number theoretic interest. This occurrence is not isolated, there are many examples of modular forms with interesting information contained in the Fourier coefficients.

This q-expansion motivates the following *normalised* Eisenstein series

$$E_k(\tau) = G_k(\tau)/2\zeta(k).$$

We may do this since $\zeta(k) \neq 0$ for $k > 2$. This normalised Eisenstein series has the following simpler looking q-expansion:

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

where B_k is the k-th Bernoulli number. Here we have used the fact that for even k,

$$\zeta(k) = \frac{(-1)^{\frac{k}{2}+1}B_k(2\pi)^k}{2(k!)}.$$

Notice how we have been avoiding the case of $k = 2$. The reason for this is that although we can define the same series when $k = 2$, it doesn't transform in the correct way to be a modular form. In particular, the series is not absolutely convergent. We can however organise the terms in a specific way in order to get conditional convergence. If we consider the sum

$$G_2(\tau) = \sum_{c\in\mathbb{Z}}\sum_{d\in\mathbb{Z}'_c}\frac{1}{(c\tau+d)^2}$$

where $\mathbb{Z}'_c = \mathbb{Z}/\{0\}$ if $c = 0$ and $\mathbb{Z}'_c = \mathbb{Z}$ otherwise. This series does converge conditionally and the terms are organised in such a way that we still obtain the same q-expansion, i.e.,

$$G_2(\tau) = 2\zeta(2) - 8\pi^2\sum_{n=1}^{\infty}\sigma(n)q^n,\ q = e^{2\pi i\tau},\ \sigma(n) = \sum_{\substack{d|n\\ d>0}} d.$$

The problem now however, is that the conditional convergence prevents this series from being weakly modular. It can be shown (through a non-trivial calculation) that

$$(G_2[\gamma]_2)(\tau) = G_2(\tau) - \frac{2\pi ic}{c\tau+d} \text{ for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

This can be corrected, but it again comes with a drawback. The function $G_2(\tau) - \pi/\mathrm{Im}(\tau)$ is weight-2 invariant under $\mathrm{SL}_2(\mathbb{Z})$ but it is not holomorphic. It is clear that the case of weight 2 is not an easy one to deal with and there are many obstacles to overcome. We will see later that in the case of level N, we can use methods similar to these in order to obtain a weight 2 Eisenstein series. This will be covered in Section 1.4.3.

These Eisenstein series allow us to obtain many different relations between power divisor sums that would be very difficult to deduce using another method. Before we see an example of this we need to know more about the structure of the space $M_k(\mathrm{SL}_2(\mathbb{Z}))$.

This space forms a vector space over $\mathbb{C}$, so addition and scalar multiplication is well-defined. Further we can actually multiply two modular forms and get another, this means that the set of all modular forms, denoted $M(\mathrm{SL}_2(\mathbb{Z})) = \oplus_{k=0}^{\infty} M_k(\mathrm{SL}_2(\mathbb{Z}))$, is a structure known as a *graded algebra*. Essentially if we have $f \in M_k(\mathrm{SL}_2(\mathbb{Z}))$ and $g \in M_l(\mathrm{SL}_2(\mathbb{Z}))$ then $fg \in M_{k+l}(\mathrm{SL}_2(\mathbb{Z}))$. This is important because it allows us to create new modular forms from ones we already know. In fact it turns out that the only modular forms we really need to know are E_4 and E_6.

Theorem 1.1.10. *There exists an isomorphism of $\mathbb{C}$-algebras:*

$$M(\mathrm{SL}_2(\mathbb{Z})) \cong \mathbb{C}[E_4, E_6].$$

In particular:

$$M_k(\mathrm{SL}_2(\mathbb{Z})) = \bigoplus_{4a+6b=k} \mathbb{C} E_4^a E_6^b.$$

There is also an important subspace within $M_k(\mathrm{SL}_2(\mathbb{Z}))$.

Definition 1.1.11. A *cusp form* of weight k is a modular form of weight k whose Fourier expansion has leading coefficient $a_0 = 0$, i.e.,

$$f(\tau) = \sum_{n=1}^{\infty} a_n q^n, \quad q = e^{2\pi i \tau}.$$

The set of cusp forms is denoted $S_k(\mathrm{SL}_2(\mathbb{Z}))$ and is a subspace of $M_k(\mathrm{SL}_2(\mathbb{Z}))$.

By Theorem 1.1.10 we see that each of the spaces $M_k(\mathrm{SL}_2(\mathbb{Z}))$ has a basis. In fact this basis is finite (non-obvious) and we have formulas for the dimension when $k \geq 2$ is an even integer:

$$\dim(M_k(\mathrm{SL}_2(\mathbb{Z}))) = \begin{cases} \lfloor \frac{k}{12} \rfloor + 1 & \text{if } k \not\equiv 2 \pmod{12} \\ \lfloor \frac{k}{12} \rfloor & \text{if } k \equiv 2 \pmod{12} \end{cases}$$

Also $\dim(S_k(\mathrm{SL}_2(\mathbb{Z})) = \dim(M_k(\mathrm{SL}_2(\mathbb{Z})) - 1$.

Let us consider the space $M_8(\mathrm{SL}_2(\mathbb{Z}))$. By the dimension formulas we see that this space is 1-dimensional. If we consider E_4^2 and E_8, these are both modular forms of weight 8. This means they must be linearly dependent but since they both have $a_0 = 1$ they must actually be equal. So

$$\left(1 + 240 \sum_{n=1}^{\infty} \sigma_3(n) q^n\right)^2 = 1 + 480 \sum_{n=1}^{\infty} \sigma_7(n) q^n.$$

Equating coefficients of q^n on both sides gives the identity

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

Here we see that we have something entirely number theoretical arising from an object belonging to the world of complex analysis. There are many more examples of identities such as this, arising in a similar manner. We can also obtain various interesting congruences including the famous Ramanujan 691 congruence:

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

where $\tau(n)$ is the Ramanujan tau function, whose values are the coefficients of the weight 12 cusp form

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

This function is known as the discriminant function and has important connections with elliptic curves. Ramanujan conjectured that the τ function should satisfy certain properties, including:

$$\begin{array}{cc} \tau(mn) = \tau(m)\tau(n) & \text{if } m,n \text{ are coprime} \\ \tau(p^r) = \tau(p)\tau(p^{r-1}) - p^{11}\tau(p^{r-2}) & \text{for any prime } p \text{ and integer } r > 2. \end{array}$$

This conjecture was correct and was originally proved by Mordell. This result can also be proved using the theory of Hecke operators. We will look at the Hecke operators later in Section 1.5.

§ 1.2 Congruence Subgroups

We can extend our study of modular forms to those with $\mathrm{SL}_2(\mathbb{Z})$ replaced by a congruence subgroup.

Definition 1.2.1. Let N be a positive integer. The *principal congruence subgroup of level N* is

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

We can extend this definition to a general congruence subgroup.

Definition 1.2.2. A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a *congruence subgroup* if $\Gamma(N) \subset \Gamma$ for some $N \in \mathbb{Z}^+$, in which case Γ is a congruence subgroup of *level N*.

If we consider the map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ it is clear that the kernel of this map is exactly $\Gamma(N)$ since this is exactly the definition of the matrices in $\Gamma(N)$. Hence it is clear that $\Gamma(N)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and that it has finite index. We have $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = |\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})|$ by the first isomorphism theorem. Since we know $|\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})|$ we have

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = N^3\prod_{p|N}\left(1-\frac{1}{p^2}\right),$$

where the product is over all prime divisors of N. From this it follows that every congruence subgroup has finite index in $\mathrm{SL}_2(\mathbb{Z})$.

Although there are many different congruence subgroups that we could work with, there are two standard congruence subgroups that are of particular interest in the theory of modular forms. These congruence subgroups are as follows:

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{N} \right\}$$

(where "$*$" means "unspecified") and

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

These congruence subgroups satisfy the following inclusions:

$$\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z}).$$

Now consider the map

$$\Gamma_1(N) \to \mathbb{Z}/N\mathbb{Z}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto b \pmod{N}.$$

It is clear that this map is a surjection since the upper right entry of a matrix in $\Gamma_1(N)$ is arbitrary. It is also clear, by considering the above inclusions, that it has kernel $\Gamma(N)$. Hence $\Gamma(N)$ is normal in $\Gamma_1(N)$ and we have $[\Gamma_1(N) : \Gamma(N)] = N$. It follows that

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(N)] = N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right),$$

since

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(N)] = \frac{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)]}{[\Gamma_1(N) : \Gamma(N)]}.$$

Now consider the map

$$\Gamma_0(N) \to (\mathbb{Z}/N\mathbb{Z})^\times, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \pmod{N}.$$

We see that it is a surjection by a similar argument to the previous case. Also it is clear that the kernel is $\Gamma_1(N)$ as we need the lower right entry to be 1 (mod N). Hence $\Gamma_1(N)$ is normal in $\Gamma_0(N)$ and we have $[\Gamma_0(N) : \Gamma_1(N)] = \varphi(N)$, where φ is the Euler totient function. This is a little less straightforward; for details see page 14 of [DiSh].

Using the same argument as before we have

$$\begin{aligned}
[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] &= \frac{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(N)]}{[\Gamma_0(N) : \Gamma_1(N)]} \\
&= \frac{N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right)}{\varphi(N)} \\
&= \frac{N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right)}{N \prod_{p|N} \left(1 - \frac{1}{p}\right)} \\
&= N \prod_{p|N} \frac{\left(\frac{p^2-1}{p^2}\right)}{\left(\frac{p-1}{p}\right)} \\
&= N \prod_{p|N} \frac{p^2 - 1}{p^2 - p} \\
&= N \prod_{p|N} \frac{p+1}{p} \\
&= N \prod_{p|N} \left(1 + \frac{1}{p}\right).
\end{aligned}$$

Note that $\Gamma(1) = \Gamma_0(1) = \Gamma_1(1) = \mathrm{SL}_2(\mathbb{Z})$ and so level 1 modular forms are exactly those we have considered.

Although $\mathrm{SL}_2(\mathbb{Z})$ contains the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, a particular congruence subgroup may not contain this matrix. Therefore a modular form for a congruence subgroup Γ is no longer necessarily $\mathbb{Z}$-periodic. Instead, any congruence subgroup Γ contains a translation matrix of the form

$$\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} : \tau \mapsto \tau + h$$

for some minimal $h \in \mathbb{Z}^+$. This follows because Γ contains $\Gamma(N)$ for some N, but h may properly divide N. Every function $f : \mathcal{H} \to \mathbb{C}$ that is weakly modular with respect to Γ is therefore $h\mathbb{Z}$ periodic. If f is holomorphic on $\mathcal{H}$ and at ∞ it therefore has a q-expansion given by

$$f(\tau) = \sum_{n=0}^{\infty} a_n q_h^n,\ q_h = e^{2\pi i \tau / h}.$$

Since congruence subgroups are smaller than $\mathrm{SL}_2(\mathbb{Z})$ it is easier for a function $f : \mathcal{H} \to \mathbb{C}$ to satisfy the transformation property and therefore you would expect more modular forms to be in the vector space $M_k(\Gamma)$. Note that $M_k(\Gamma)$ is standard notation to be

introduced in definition 1.2.4. The first thing we might ask is whether these spaces are still finite. Fortunately this is still the case and we do in general still have formulas for the dimensions of these spaces. This is covered in detail in Sections 3.5 and 3.6 of [DiSh]. In order to keep these vector spaces finite-dimensional, we might expect to have to impose more strict conditions for a function $f : \mathcal{H} \to \mathbb{C}$ to be a modular form. In fact, modular forms for congruence subgroups need to be holomorphic not only on $\mathcal{H}$ but at all cusps. In other words we no longer only have holomorphy at ∞ as a condition; there are other points that the function must be holomorphic at. These points are known as the cusps. The idea now is to adjoin not only ∞ to $\mathcal{H}$ but also $\mathbb{Q}$. We therefore define the extended upper half plane to be $\mathcal{H}^* = \mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$. We then consider the Γ-equivalence classes of points in $\mathbb{Q} \cup \{\infty\} \subset \mathcal{H}^*$. Such an equivalence class is called a cusp. Note that when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is only a single equivalence class (represented by ∞) and so this is well-defined. We will briefly discuss cusps in the next section when we talk about modular curves. For a more thorough description of cusps, see chapter 2 of [DiSh]. If we write any cusp $s \in \mathbb{Q} \cup \{\infty\}$ as $s = \alpha(\infty)$ for some $\alpha \in \Gamma$, holomorphy at s is naturally defined in terms of holomorphy at ∞ via the $[\alpha]_k$ operator. Note that s here is simply a representative for the equivalence class of s under the action of Γ.

Before moving on to define a modular form for the congruence subgroup Γ we will briefly discuss the action of a congruence subgroup on $\mathcal{H}^*$. We already know how $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathcal{H}$ (by fractional linear transformations) and similarly any congruence subgroup Γ acts in the same way. The question we might ask is how $\mathrm{SL}_2(\mathbb{Z})$ or any congruence subgroup Γ acts on $\mathbb{Q} \cup \{\infty\}$. We can consider $\mathbb{Q} \cup \{\infty\}$ as $\mathbb{P}^1(\mathbb{Q})$, the projective line with rational coordinates. We call $\mathbb{P}^1(\mathbb{Q})$ the set of *(rational) cusps*. It turns out that this view allows us to extend our definition and we may use the usual formula. In other words, a congruence subgroup Γ acts on $\mathcal{H}^*$ in the same way as it acts on $\mathcal{H}$. As we mentioned, there is only a single cusp in the case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. We have the following result of which the proof may illuminate the situation.

Lemma 1.2.3. $\mathrm{SL}_2(\mathbb{Z})$ *acts transitively on the set of cusps. In particular,* ∞ *is a representative of the single* Γ*-equivalence class of cusps.*

Proof. Write any given rational number as a/c with a, c coprime. We may use Euclid's algorithm to complete a and c to a matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Then

$$\gamma(\infty) = \frac{a}{c}.$$

This follows by considering the projective coordinates of ∞. We have

$$\gamma(\infty) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}.$$

□

We now define *weakly modular of weight k with respect to Γ* to mean meromorphic and weight-k invariant under Γ, that is, a meromorphic function f on $\mathcal{H}$ is weakly modular of weight k if

$$f[\gamma]_k = f \quad \text{for all } \gamma \in \Gamma.$$

The definition of a modular form with respect to a congruence subgroup is now very similar to that of a modular form for $\mathrm{SL}_2(\mathbb{Z})$ but we have the extra cusps to consider.

Definition 1.2.4. Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ and let k be an integer. A function $f : \mathcal{H} \to \mathbb{C}$ is a *modular form of weight k with respect to Γ* if

(1) f is holomorphic,

(2) f is weight-k invariant under Γ,

(3) $f[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

If in addition,

(4) $a_0 = 0$ in the Fourier expansion of $f[\alpha]_k$ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$,

then f is a *cusp form* of weight k with respect to Γ. The modular forms of weight k with respect to Γ are denoted $M_k(\Gamma)$, the cusp forms $S_k(\Gamma)$.

We also note that whenever $\Gamma_1 \subset \Gamma_2$ we have $M_k(\Gamma_2) \subset M_k(\Gamma_1)$. In other words any modular form for the larger congruence subgroup is also a modular form for the smaller congruence subgroup. However there could be modular forms for the smaller congruence subgroup that are not modular forms for the larger congruence subgroup. This gives us the notion of oldforms and newforms. We will see more about these later when we discuss Hecke operators.

Since the easiest example of a modular form at level 1 was that of an Eisenstein series, we might wonder whether there is a generalisation to level N. By the above, it is clear that any modular form for $\mathrm{SL}_2(\mathbb{Z})$ is also a modular form for $\Gamma(N)$ (or any other congruence subgroup Γ). Hence the level 1 Eisenstein series "is" a level N Eisenstein series. As you may have guessed from the suggestive notation, this is regarded as an oldform. So the question is whether there is anything genuinely new at level N, i.e., are there any newforms? There is a generalisation, only the definition is somewhat more complicated. Before giving the details of this, we will take a brief detour and discuss modular curves.

§ 1.3 Modular Curves

Given a congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ we can form a quotient known as a *modular curve*. A modular curve is defined as the quotient space of orbits under Γ,

$$Y(\Gamma) = \Gamma\backslash\mathcal{H} = \{\Gamma\tau : \tau \in \mathcal{H}\}.$$

The modular curves for $\Gamma_0(N)$, $\Gamma_1(N)$, and $\Gamma(N)$ are denoted

$$Y_0(N) = \Gamma_0(N)\backslash\mathcal{H}, \quad Y_1(N) = \Gamma_1(N)\backslash\mathcal{H}, \quad Y(N) = \Gamma(N)\backslash\mathcal{H}.$$

These modular curves are in fact Riemann surfaces but they are not compact. The problem is the cusps that we mentioned in the previous section. We can however compactify the curves by adding the cusps. We then obtain compact Riemann surfaces. So we now have a quotient of the extended upper half plane $\mathcal{H}^*$. Recall that this was defined as $\mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$. This quotient is defined by $X(\Gamma) = \Gamma\backslash\mathcal{H}^*$. We have similar definitions as before for each of the congruence subgroups:

$$X_0(N) = \Gamma_0(N)\backslash\mathcal{H}^*, \quad X_1(N) = \Gamma_1(N)\backslash\mathcal{H}^*, \quad X(N) = \Gamma(N)\backslash\mathcal{H}^*.$$

In order to understand these modular curves it is necessary to understand the orbits $\{\Gamma\tau : \tau \in \mathcal{H}\}$. It is therefore necessary to introduce the notion of a *fundamental domain* for the congruence subgroup Γ. This is essentially a region of the upper half plane for which any point in $\mathcal{H}$ can be identified with one inside the region. In other words, there is some $\gamma \in \Gamma$ such that $\gamma\tau_1 = \tau_2$ with $\tau_1, \tau_2 \in \mathcal{H}$ and $\tau_2 \in \mathcal{D}$ where $\mathcal{D}$ represents a fundamental domain. Also the only points within the domain to be identified with each other lie on the boundary of the domain.

The simplest example of a fundamental domain comes in the case when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. In this case we have

$$\mathcal{D} = \{\tau \in \mathcal{H} : |\mathrm{Re}(\tau)| \leq 1/2, |\tau| \geq 1\}.$$

If we consider the modular curve $Y(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}$ and a map from $\mathcal{D}$ to this curve we have the following result.

Lemma 1.3.1. *The map $\pi : \mathcal{D} \to Y(1)$ surjects, where π is the natural projection $\pi(\tau) = \mathrm{SL}_2(\mathbb{Z})\tau$.*

Proof. See page 53 of [DiSh]. □

We know that this map cannot be injective since we have some identifications at the boundary of $\mathcal{D}$. However, as we previously mentioned, these are the only identifications.

Lemma 1.3.2. *Suppose τ_1 and τ_2 are distinct points in $\mathcal{D}$ and that $\tau_2 = \gamma\tau_1$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Then either*

(1) $\mathrm{Re}(\tau_1) = \pm 1/2$ *and* $\tau_2 = \tau_1 \mp 1$, *or*

(2) $|\tau_1| = 1$ *and* $\tau_2 = -1/\tau_1$.

Proof. See pages 53-54 of [DiSh]. □

If we map the fundamental domain $\mathcal{D}$ of $\mathrm{SL}_2(\mathbb{Z})$ via stereographic projection to the Riemann sphere this gives a triangle with a single vertex missing. This vertex is precisely the "cusp" at ∞. It becomes clear looking at this that adding in this single cusp will make the curve compact. In general, for a congruence subgroup Γ, there will be a number of cusps like this that need to be added to make the curve compact. In fact, this process will always be possible since the number of such cusps is always finite.

Lemma 1.3.3. *The modular curve $X(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}^*$ has one cusp. For any congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ the modular curve $X(\Gamma)$ has finitely many cusps.*

Proof. This proof was left as Exercise 2.4.1 in [DiSh]. We have proved the first part of the statement in Lemma 1.2.3. By the previous section we know that a congruence subgroup Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$. In other words $\mathrm{SL}_2(\mathbb{Z})$ is a finite union of cosets with $\cup_{j=1}^{d}\Gamma\gamma_j = \mathrm{SL}_2(\mathbb{Z})$. If there were infinitely many cusps, then there would be infinitely many elements of $\mathbb{Q}\cup\{\infty\}$ not in the same coset. This is a contradiction. So there are finitely many cusps. □

This fact will be very useful in our later work as it will be necessary to know the constant term of a particular Eisenstein series at each cusp. If there were infinitely many cusps then the method we use would not be possible. For those interested in the finer details of modular curves, see chapter 2 of [DiSh].

§ 1.4 Level N Eisenstein Series

Here we only give a detailed description of the Eisenstein series for $\Gamma_0(N)$ and $\Gamma_1(N)$. Although we give the definition of an Eisenstein series for $\Gamma(N)$ we will not look too deeply at the Fourier expansion (or the relevant proofs). The reason for this is that we simply need the Eisenstein series for $\Gamma(N)$ as a building block for the Eisenstein series we will be interested in. For those interested in seeing more about the Eisenstein series for $\Gamma(N)$, see Section 4.2 of [DiSh]. We can in fact view the space of Eisenstein series as a quotient of the full space of modular forms. For a congruence subgroup Γ and any integer k, we define the *weight k Eisenstein space* of Γ to be the quotient space of the modular forms by the cusp forms,

$$\mathcal{E}_k(\Gamma) = M_k(\Gamma)/S_k(\Gamma).$$

We will see later that this space can actually be redefined as a subspace of $M_k(\Gamma)$ complementary to $S_k(\Gamma)$. That is, the space of modular forms is made up precisely of Eisenstein series and cusp forms.

In order to define level N Eisenstein series we will need to consider modular forms with character. We therefore introduce the necessary character theory in order to do this.

1.4.1 Dirichlet Characters, Gauss Sums, and Eigenspaces

In order to ease notation, let G_N denote the multiplicative group $(\mathbb{Z}/N\mathbb{Z})^\times$ for any positive integer N.

Definition 1.4.1. A *Dirichlet character modulo* N is a homomorphism of multiplicative groups,

$$\chi : G_N \to \mathbb{C}.$$

Recall that $|G_N| = \varphi(N)$ where φ is the Euler totient function. Given any two Dirichlet characters χ and ψ modulo N, we can consider the product character defined by $(\chi\psi)(n) = \chi(n)\psi(n)$ for $n \in G_N$. This is again a Dirichlet character modulo N. We can also consider the set of Dirichlet characters modulo N; this forms a multiplicative group known as the *dual group* of G_N, denoted $\widehat{G}_N$, whose identity is the *trivial character modulo* N. This character simply maps every element to 1 and is usually denoted by $\mathbf{1}$ or $\mathbf{1}_N$ if N needs to be emphasized. The values taken by a Dirichlet character are roots of unity. This follows since G_N is a finite group. It therefore follows that the inverse of such a character is simply the complex conjugate, defined by $\overline{\chi}(n) = \overline{\chi(n)}$ for every $n \in G_N$.

It turns out that in fact there is a close link between these two groups.

Proposition 1.4.2. *Let $\widehat{G}_N$ be the dual group of G_N. Then $\widehat{G}_N$ is isomorphic to G_N. In particular, the number of Dirichlet characters modulo N is $\varphi(N)$.*

This isomorphism is noncanonical and involves arbitrary choices of which elements map to which characters. However we now know exactly how many Dirichlet characters there are modulo N. We also have certain orthogonality relations that are satisfied. We have

$$\sum_{n \in G_N} \chi(n) = \begin{cases} \varphi(N) & \text{if } \chi = \mathbf{1}, \\ 0 & \text{if } \chi \neq \mathbf{1}, \end{cases} \qquad \sum_{\chi \in \widehat{G}_N} \chi(n) = \begin{cases} \varphi(N) & \text{if } n = 1, \\ 0 & \text{if } n \neq 1. \end{cases}$$

Another important operation we can do with Dirichlet characters is lifting to a larger modulus. This will be something that we need to do in most of our main theorems. Consider any positive integer N and suppose $d|N$. Every Dirichlet character χ modulo d lifts to a Dirichlet character χ_N modulo N. This is simply defined by $\chi_N(n \pmod N) = \chi(n \pmod d)$. In other words for each $n \in G_N$ you look at its reduction $\overline{n}$ modulo d and then $\chi_N(n)$ is given the same value as is taken by $\chi(\overline{n})$.

One might wonder whether it is possible to go the other way; from a Dirichlet character modulo N to a Dirichlet character modulo d. This isn't always possible and so we assign a value to each character called the *conductor*. This is the smallest positive divisor such that a Dirichlet character modulo N factors through a character of smaller modulus.

We know that $(\mathbb{Z}/N\mathbb{Z})^\times$ is the invertible elements of $(\mathbb{Z}/N\mathbb{Z})$ and that a Dirichlet character can be defined as a function on this group. We might wonder if we can

extend the definition to include the non-invertible elements as well. That is, can we extend to a function $\chi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$? The answer is yes: we simply set $\chi(n) = 0$ for all non-invertible elements $n \in \mathbb{Z}/N\mathbb{Z}$. This can then be extended further to the whole of the integers. We can let $\chi : \mathbb{Z} \to \mathbb{C}$ by given by $\chi(n) = \chi(n \pmod N)$ for all $n \in \mathbb{Z}$. It is then clear that for any n with $\gcd(n, N) > 1$, we have $\chi(n \pmod N) = \chi(0) = 0$. Although this function is no longer a homomorphism, it is completely multiplicative, i.e., $\chi(nm) = \chi(n)\chi(m)$ for all n, m.

We may now modify the orthogonality relations slightly. If we sum over $n = 0$ to $N-1$ in the first relation and take $n \in \mathbb{Z}$ in the second we get

$$\sum_{n=0}^{N-1} \chi(n) = \begin{cases} \varphi(N) & \text{if } \chi = \mathbf{1}, \\ 0 & \text{if } \chi \neq \mathbf{1}, \end{cases} \quad \sum_{\chi \in \widehat{G}_N} \chi(n) = \begin{cases} \varphi(N) & \text{if } n \equiv 1 \pmod N, \\ 0 & \text{if } n \not\equiv 1 \pmod N. \end{cases}$$

We are now in a position to define a Gauss sum. The *Gauss sum* of a Dirichlet character χ modulo N is the complex number

$$g(\chi) = \sum_{n=0}^{N-1} \chi(n)\mu_N^n, \quad \mu_N = e^{2\pi i/N}.$$

If χ is primitive modulo N then for any integer m,

$$\sum_{n=0}^{N-1} \chi(n)\mu_N^{nm} = \overline{\chi}(m)g(\chi).$$

It therefore follows that the Gauss sum of a primitive character is non-zero. In fact the square of the absolute value works out to be N, the details of which are on page 118 of [DiSh]. Gauss sums will appear at several points throughout this thesis, particularly when we calculate the constant term of the Eisenstein series at other cusps.

We finish with a remark about the distribution of certain types of Dirichlet character.

Definition 1.4.3. Let χ be a Dirichlet character modulo N. If $\chi(-1) = 1$, we say that χ is an *even* Dirichlet character. If $\chi(-1) = -1$ we say that χ is *odd*.

Lemma 1.4.4. *Let N be a positive integer. If $N = 1$ or $N = 2$ then every Dirichlet character χ modulo N is even. If $N > 2$ then the number of Dirichlet characters modulo N is even, half of them being even, the other half being odd.*

Proof. See page 118 of [DiSh]. □

The reason we are interested in Dirichlet characters, and in particular modular forms with character, is because they decompose the space $M_k(\Gamma_1(N))$ into a direct sum

of subspaces. For a Dirichlet character χ modulo N we define the χ-*eigenspace* of $M_k(\Gamma_1(N))$,

$$M_k(N,\chi) = \{f \in M_k(\Gamma_1(N)) : f[\gamma]_k = \chi(d_\gamma)f \quad \text{for all} \quad \gamma \in \Gamma_0(N)\},$$

where d_γ denotes the lower right entry of γ. The eigenspace corresponding to the trivial character is $M_k(\Gamma_0(N))$. That is $M_k(N,\mathbf{1}) = M_k(\Gamma_0(N))$. Also the space $M_k(N,\chi)$ is just $\{0\}$ unless $\chi(-1) = (-1)^k$. We then have

$$M_k(\Gamma_1(N)) = \bigoplus_\chi M_k(N,\chi).$$

This result also holds for the cusp forms, and therefore also for the quotients (the space of Eisenstein series) as well,

$$\mathcal{E}_k(\Gamma_1(N)) = \bigoplus_\chi \mathcal{E}_k(N,\chi).$$

Consider the Riemann zeta function $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$. We could instead view this as essentially $\zeta(s) = \sum_{n=1}^\infty \frac{\mathbf{1}(n)}{n^s}$ where $\mathbf{1}(n)$ is the trivial Dirichlet character. We might wonder what happens if we had a different Dirichlet character. This in fact gives us what is known as the Dirichlet L-function. Given any Dirichlet character χ modulo N, there is an associated *Dirichlet L-function*,

$$L(s,\chi) = \sum_{n=1}^\infty \frac{\chi(n)}{n^s} = \prod_{p\in\mathcal{P}} (1-\chi(p)p^{-s})^{-1}, \ \mathrm{Re}(s) > 1,$$

where $\mathcal{P}$ is the set of primes. This L-function extends to a meromorphic function on the whole s-plane and the extension is entire unless the character is trivial. If the character is trivial then as we discussed above, the L-function is essentially the Riemann zeta function. The L-function satisfies a functional equation and its form depends on the value of $\chi(-1)$. If $\chi(-1) = 1$, the functional equation is

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right) N^s L(s,\chi) = \pi^{-(1-s)/2}\Gamma\left(\frac{1-s}{2}\right) g(\chi) L(1-s,\overline{\chi}),$$

and when $\chi(-1) = -1$ it is

$$\pi^{-(s+1)/2}\Gamma\left(\frac{s+1}{2}\right) N^s L(s,\chi) = -i\pi^{-(2-s)/2}\Gamma\left(\frac{2-s}{2}\right) g(\chi) L(1-s,\overline{\chi}),$$

where $\Gamma(s)$ is the gamma function from complex analysis, defined by

$$\Gamma(s) = \int_{t=0}^\infty e^{-t}t^s \frac{\mathrm{d}t}{t}, \ s \in \mathbb{C}, \ \mathrm{Re}(s) > 0.$$

1.4.2 Level N Eisenstein Series when $k \geq 3$

The definition of Eisenstein series for $\Gamma(N)$ isn't too far away from that of the Eisenstein series for level 1. We now simply look at a sum over certain congruence classes rather than over all integers. Let N be a positive integer and let $\overline{v} \in (\mathbb{Z}/N\mathbb{Z})^2$ be a row vector of order N. Here the overline denotes reduction modulo N. Let $\delta = \begin{bmatrix} a & b \\ c_v & d_v \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ with (c_v, d_v) a lift of $\overline{v}$ to $\mathbb{Z}^2$, and let $k \geq 3$ be an integer. Let ϵ_N be $1/2$ if $N \in \{1, 2\}$ and 1 if $N > 2$. Define the weight k Eisenstein series for $\Gamma(N)$ by

$$E_k^{\overline{v}}(\tau) = \epsilon_N \sum_{\substack{(c,d) \equiv v \pmod N \\ \gcd(c,d)=1}} (c\tau + d)^{-k}.$$

Note that if $N = 1$, then every pair of integers $(c, d) \equiv v \pmod 1$. Hence this definition reduces exactly to the level 1 (normalised) Eisenstein series when we set $N = 1$. We also have the non-normalised series

$$G_k^{\overline{v}}(\tau) = \sum_{(c,d) \equiv v \pmod N}{}' (c\tau + d)^{-k},$$

where $\sum'$ means we sum over non-zero pairs (c, d).

We are now in a position to define the Eisenstein series for $\Gamma_0(N)$ and $\Gamma_1(N)$. First of all we note that vectors modulo N of the form $v = \overline{(0, d)}$ satisfy

$$\overline{(0,d)\gamma} = \overline{(0, dd_\gamma)} \text{ for all } \gamma \in \Gamma_0(N),$$

where d_γ is the lower right entry of γ. If we take a sum over all $d \in (\mathbb{Z}/N\mathbb{Z})^\times$, this gives a sum of Eisenstein series

$$\sum_{d \in (\mathbb{Z}/N\mathbb{Z})^\times} G_k^{(0,d)}$$

lying in $M_k(\Gamma_0(N))$. We can also introduce a character modulo N into the sum to get something in $M_k(N, \chi)$, namely

$$\sum_{d \in (\mathbb{Z}/N\mathbb{Z})^\times} \overline{\chi}(d) G_k^{(0,d)}.$$

This kind of process can be generalised to get a basis for the space of Eisenstein series $\mathcal{E}_k(N, \chi)$.

Given any two primitive Dirichlet characters ψ modulo u and φ modulo v such that $uv = N$ and $\psi\varphi(-1) = (-1)^k$ (note that the characters are both raised to level uv here so that the product makes sense), we can consider a linear combination of the Eisenstein series for $\Gamma(N)$,

$$G_k^{\psi,\varphi}(\tau) = \sum_{c=0}^{u-1} \sum_{d=0}^{v-1} \sum_{e=0}^{u-1} \psi(c)\overline{\varphi}(d) G_k^{\overline{(cv, d+ev)}}(\tau).$$

If we take $\gamma = \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix} \in \Gamma_0(N)$, it is fairly easy to show that

$$(G_k^{\psi,\varphi}[\gamma]_k)(\tau) = (\psi\varphi)(d_\gamma)G_k^{\psi,\varphi}(\tau),$$

from which it follows that $G_k^{\psi,\varphi}(\tau) \in M_k(N, \psi\varphi)$. The details of this are given on page 127 of [DiSh]. The main thing that we will be interested in is what the Fourier expansion of this series looks like. To ease notation, let $C_k = \frac{(-2\pi i)^k}{(k-1)!}$.

Theorem 1.4.5. *The Eisenstein series $G_k^{\psi,\varphi}$ takes the form*

$$G_k^{\psi,\varphi}(\tau) = \frac{C_k g(\overline{\varphi})}{v^k} E_k^{\psi,\varphi}(\tau),$$

where $E_k^{\psi,\varphi}$ has Fourier expansion

$$E_k^{\psi,\varphi}(\tau) = \delta(\psi)L(1-k,\varphi) + 2\sum_{n=1}^{\infty} \sigma_{k-1}^{\psi,\varphi}(n)q^n,\ q = e^{2\pi i\tau}.$$

Here $\delta(\psi)$ is 1 if $\psi = \mathbf{1}_1$, and is 0 otherwise, and the generalised power sum in the Fourier coefficient is

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

Proof. See pages 127-129 of [DiSh]. □

Now that we know what the Eisenstein series look like at level N, we might be interested in which Eisenstein series we need in order to form a basis. Recall that at level one we simply needed E_4 and E_6. Given any positive integer N and any $k \geq 3$, let $A_{N,k}$ be the set of triples (ψ, φ, t) with ψ and φ primitive Dirichlet characters modulo u and v with $\psi\varphi(-1) = (-1)^k$, and t a positive integer such that $tuv|N$. Then $|A_{N,k}| = \dim(\mathcal{E}_k(\Gamma_1(N)))$. For any triple $(\psi, \varphi, t) \in A_{N,k}$ define

$$E_k^{\psi,\varphi,t}(\tau) = E_k^{\psi,\varphi}(t\tau).$$

Note that $E_k^{\psi,\varphi,t} \in M_k(\Gamma_1(tuv))$ and since $tuv|N$ we have $E_k^{\psi,\varphi,t} \in M_k(\Gamma_1(N))$. In fact these Eisenstein series are enough to form a basis.

Theorem 1.4.6. *Let N be a positive integer and let $k \geq 3$. The set*

$$\{E_k^{\psi,\varphi,t} : (\psi, \varphi, t) \in A_{N,k}\}$$

represents a basis of $\mathcal{E}_k(\Gamma_1(N))$. For any character χ modulo N, the set

$$\{E_k^{\psi,\varphi,t} : (\psi, \varphi, t) \in A_{N,k},\ \psi\varphi = \chi\}$$

represents a basis of $\mathcal{E}_k(N, \chi)$.

Note that as it stands the Eisenstein space is defined as a quotient and not as a subspace. Hence at the moment this set only represents a basis. When we redefine the Eisenstein space as a subspace, we will be able to take the span of these elements giving us a basis. If we take our character χ to be trivial this gives a basis for $\mathcal{E}_k(\Gamma_0(N))$.

1.4.3 Eisenstein Series of Weight 2

Recall that the level 1 Eisenstein series had to be modified in the case of weight 2. Since the level N Eisenstein series reduce to the level 1 Eisenstein series when we set $N = 1$, it is clear that we will also need to modify our level N Eisenstein series. In a similar way to that of the level 1 Eisenstein series it isn't too hard to obtain a series that is weakly modular. It is slightly harder to obtain a series that is holomorphic. Consider the following series:

$$g_2^{\overline{v}}(\tau) = G_2^{\overline{v}}(\tau) - \frac{\pi}{N^2 \mathrm{Im}(\tau)},\ \overline{v} \in (\mathbb{Z}/N\mathbb{Z})^2 \text{ of order } N.$$

This series is weight-2 invariant with respect to $\Gamma(N)$ but the term $-\frac{\pi}{N^2\mathrm{Im}(\tau)}$ obviously causes the series to be non-holomorphic. One way of getting rid of this problem would be to take a difference of two such series. Differences such as $g_2^{\overline{v_1}} - g_2^{\overline{v_2}}$ where v_1 and v_2 are cusps of $\Gamma(N)$ are modular forms since they are holomorphic and weakly modular and their Fourier coefficients are small enough. It follows that a basis for the Eisenstein space $\mathcal{E}_2(\Gamma(N))$ is given by the linear combinations of these series such that the coefficients sum to zero.

Theorem 1.4.7.

$$\mathcal{E}_2(\Gamma(N)) = \left\{ \sum_{\overline{v}} a_{\overline{v}} g_2^{\overline{v}} : \sum_{\overline{v}} a_{\overline{v}} = 0 \right\},$$

where the sums are taken over vectors $\overline{v}$ of order N in $(\mathbb{Z}/N\mathbb{Z})^2$.

For more details about this construction see Section 4.6 of [DiSh].

What about the Eisenstein series for $\Gamma_1(N)$ and its eigenspaces? Well it turns out that in this case not much is different to the higher weight Eisenstein series. Let ψ and φ be Dirichlet characters modulo u and v respectively with $uv = N$ and $\psi\varphi(-1) = (-1)^k$ and φ primitive. Consider the sums

$$G_2^{\psi,\varphi}(\tau) = \sum_{c=0}^{u-1} \sum_{d=0}^{v-1} \sum_{e=0}^{u-1} \psi(c)\overline{\varphi}(d) G_2^{\overline{(cv,d+ev)}}(\tau),$$

$$E_2^{\psi,\varphi}(\tau) = \delta(\psi) L(-1,\varphi) + 2\sum_{n=1}^{\infty} \sigma_1^{\psi,\varphi}(n) q^n,\ q = e^{2\pi i \tau}.$$

These are the same series as before with $k = 2$. If either of ψ or φ is non-trivial then the coefficients sum to 0 in $G_2^{\psi,\varphi}(\tau)$, and so we have

$$G_2^{\psi,\varphi} \in M_2(N, \psi\varphi), \quad G_2^{\psi,\varphi}(\tau) = \frac{C_2 g(\overline{\varphi})}{v^2} E_2^{\psi,\varphi}(\tau).$$

When both ψ and φ are trivial no sum $G_{2,\mathbf{1}_u,\mathbf{1}_v}(\tau)$ is a modular form. However this can be modified in a similar way to the level 1 weight 2 case. For any positive integer t,

$$G_{2,\mathbf{1}_1,\mathbf{1}_1}(\tau) - tG_{2,\mathbf{1}_1,\mathbf{1}_1}(t\tau) = \frac{1}{N^2} G_{2,t}(\tau)$$

where $G_{2,t} \in M_2(\Gamma_0(t))$ is the series given by $G_2(\tau) - tG_2(t\tau)$.

We can again ask which of these series are required for a basis. Let $A_{N,2}$ be the set of triples (ψ, φ, t) such that ψ and φ are primitive Dirichlet characters modulo u and v with $\psi\varphi(-1) = 1$, and t is an integer such that $1 < tuv|N$. Note that the triple $(\mathbf{1}_1, \mathbf{1}_1, 1)$ is excluded here. For any triple in $A_{N,2}$ define

$$E_2^{\psi,\varphi,t}(\tau) = \begin{cases} E_2^{\psi,\varphi}(t\tau) & \text{unless } \psi = \varphi = \mathbf{1}_1, \\ E_2^{\mathbf{1}_1,\mathbf{1}_1}(\tau) - tE_2^{\mathbf{1}_1,\mathbf{1}_1}(t\tau) & \text{if } \psi = \varphi = \mathbf{1}_1, \end{cases}$$

Theorem 1.4.8. *Let N be a positive integer. The set*

$$\{E_2^{\psi,\varphi,t} : (\psi, \varphi, t) \in A_{N,2}\}$$

represents a basis of $\mathcal{E}_2(\Gamma_1(N))$. For any character χ modulo N, the set

$$\{E_2^{\psi,\varphi,t} : (\psi, \varphi, t) \in A_{N,2},\ \psi\varphi = \chi\}$$

represents a basis of $\mathcal{E}_2(N, \chi)$.

1.4.4 Eisenstein Series of Weight 1 and Bernoulli Numbers

Now that we are working with Eisenstein series at level N, we can actually consider modular forms of weight 1. Although there were no modular forms of odd weight at level 1, we already know there are modular forms of odd weight at level N. The question is whether there are any modular forms of weight 1. In fact, there are. We just have to be clever about the functions we use. If we tried to use the same Eisenstein series as before with $k = 1$ we would immediately run into problems. At weight 2 we only managed to get conditional convergence. At weight 1 there is no convergence at all, no matter how we arrange the terms. We can however use a modified function such that we obtain a modular form of weight 1 satisfying the properties that we would like an Eisenstein series of weight 1 to satisfy. In particular, it lies in the correct vector space and has a similar q-expansion to that of an Eisenstein series. In order to do this we will first define a generalisation of the Bernoulli numbers. These will both be useful

in defining the weight 1 Eisenstein series and will also appear in our later work. If we consider the computation of power sums

$$\begin{aligned}
&1^0 + 2^0 + \cdots + n^0 = n,\\
&1^1 + 2^1 + \cdots + n^1 = \frac{1}{2}(n^2 + n),\\
&1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}(2n^3 + 3n^2 + n),\\
&\text{etc,}
\end{aligned}$$

then the Bernoulli numbers arise naturally. In order to study these, let n be a positive integer and let the k-th power sum up to $n-1$ be

$$S_k(n) = \sum_{m=0}^{n-1} m^k,\ k \in \mathbb{N}.$$

The power series with these sums as coefficients is their generating function

$$\mathbf{S}(n,t) = \sum_{k=0}^{\infty} S_k(n)\frac{t^k}{k!}.$$

This can also be written as

$$\mathbf{S}(n,t) = \frac{e^{nt}-1}{t}\frac{t}{e^t-1}.$$

The second term here is independent of n and actually contains the Bernoulli numbers. We have

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

where B_k is the k-th *Bernoulli number*.

The k-th *Bernoulli polynomial* is defined by

$$B_k(X) = \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \end{array} \right) B_j X^{k-j}.$$

These Bernoulli polynomials then have a generating function given by

$$\frac{te^{tX}}{e^t-1} = \sum_{k=0}^{\infty} B_k(X)\frac{t^k}{k!}.$$

We can use this generating function to define a generalised Bernoulli number, a special case of which involves Dirichlet characters.

Definition 1.4.9. Let χ be a Dirichlet character modulo u. The *generalised Bernoulli numbers attached to* χ are defined by

$$\sum_{c=0}^{u-1} \chi(c) \frac{te^{ct}}{e^{ut}-1} = \sum_{k=0}^{\infty} B_{k,\chi} \frac{t^k}{k!}.$$

Note that if $\chi = \mathbf{1}_1$, then $B_{k,\mathbf{1}_1} = B_k$. We also note that there is an explicit formula for the generalised Bernoulli numbers. We have

$$B_{k,\chi} = u^{k-1} \sum_{c=0}^{u-1} \chi(c) B_k(c/u).$$

Recall that there was a relation between the standard Bernoulli numbers and the Riemann zeta function. This can be generalised to obtain a relation between the generalised Bernoulli numbers and the L-function attached to the character χ. We have the following for all $k \geq 1$:

$$L(1-k, \chi) = -\frac{B_{k,\chi}}{k}.$$

This will become extremely important in our later work as it will appear as the constant term of an Eisenstein series whose Fourier coefficients satisfy a congruence.

We are now in a position to define the weight 1 Eisenstein series. Although we didn't go too deeply into the details of the weight 2 Eisenstein series we note that in fact the Weierstrass $\wp$-function defines a weight 2 Eisenstein series. Recall that this function arises in the theory of elliptic curves. In particular, every elliptic curve is a rational linear combination of $\wp$ and $\wp'$. These functions are associated to a lattice Λ. There is another function, also related to a lattice Λ which leads naturally to series of weight 1. The function in question is the *Weierstrass σ-function*

$$\sigma_\Lambda(z) = z \prod_{\omega \in \Lambda}{}' \left(1 - \frac{z}{\omega}\right) e^{z/\omega + \frac{1}{2}(z/\omega)^2}, \; z \in \mathbb{C}.$$

Taking the logarithmic derivative σ'/σ gives the *Weierstrass zeta function*, denoted by Z in order to avoid confusion with the Riemann zeta function,

$$Z_\Lambda(z) = \frac{1}{z} + \sum_{\omega \in \Lambda}{}' \left(\frac{1}{z-\omega} + \frac{1}{\omega} + \frac{z}{\omega^2}\right), \; z \in \mathbb{C}.$$

This function has simple poles with residue 1 at the lattice points. It isn't periodic with respect to Λ, instead, since $Z'_\Lambda = -\wp_\Lambda$ is periodic, if $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ then the quantities

$$\eta_1(\Lambda) = Z_\Lambda(z+\omega_1) - Z_\Lambda(z) \;\text{ and }\; \eta_2(\Lambda) = Z_\Lambda(z+\omega_2) - Z_\Lambda(z)$$

are lattice constants such that

$$Z_\Lambda(z + n_1\omega_1 + n_2\omega_2) = Z_\Lambda(z) + n_1\eta_1(\Lambda) + n_2\eta_2(\Lambda), \; n_1, n_2 \in \mathbb{Z}.$$

Under the normalising convention $\omega_1/\omega_2 \in \mathcal{H}$ the lattice constants satisfy the *Legendre relation* $\eta_2(\Lambda)\omega_1 - \eta_1(\Lambda)\omega_2 = 2\pi i$. The second lattice constant appears in the q-product expansion of σ specialised to $\Lambda = \Lambda_\tau$,

$$\sigma_{\Lambda_\tau}(z) = \frac{1}{2\pi i} e^{\frac{1}{2}\eta_2(\Lambda_\tau)z^2}(e^{\pi i z} - e^{-\pi i z}) \prod_{n=1}^{\infty} \frac{(1 - e^{2\pi i z}q^n)(1 - e^{-2\pi i z}q^n)}{(1-q^n)^2},$$

where $q = e^{2\pi i \tau}$. The logarithmic derivative is therefore

$$Z_{\Lambda_\tau}(z) = \eta_2(\Lambda_\tau)z - \pi i \frac{1 + e^{2\pi i z}}{1 - e^{2\pi i z}} - 2\pi i \sum_{n=1}^{\infty} \left(\frac{e^{2\pi i z}q^n}{1 - e^{2\pi i z}q^n} - \frac{e^{-2\pi i z}q^n}{1 - e^{-2\pi i z}q^n} \right).$$

For any vector $\overline{v} \in (\mathbb{Z}/N\mathbb{Z})^2$ of order N, the function of modular points

$$F_1^{\overline{v}}(\mathbb{C}/\Lambda, (\omega_1/N + \Lambda, \omega_2/N + \Lambda)) \\ = Z_\Lambda \left(\frac{c_v\omega_1 + d_v\omega_2}{N} \right) - \frac{c_v\eta_1(\Lambda) + d_v\eta_2(\Lambda)}{N}$$

is well-defined and degree-1 homogeneous with respect to $\Gamma(N)$. The corresponding function

$$g_1^{\overline{v}}(\tau) = \frac{1}{N} Z_{\Lambda_\tau} \left(\frac{c_v\tau + d_v}{N} \right) - \frac{c_v\eta_1(\Lambda_\tau) + d_v\eta_2(\Lambda_\tau)}{N^2}$$

is weakly modular of weight 1 with respect to $\Gamma(N)$. It can be shown (see page 139 of [DiSh]) that

$$g_1^{\overline{v}}(\tau) = G_1^{\overline{v}}(\tau) - \frac{C_1}{N} \left(\frac{c_v}{N} - \frac{1}{2} \right), \; 0 \le c_v < N.$$

Here the series $G_1^{\overline{v}}$ is analogous to $G_k^{\overline{v}}$ for $k \ge 3$. Since $g_1^{\overline{v}}$ is holomorphic and weakly modular with respect to $\Gamma(N)$ and its n-th Fourier coefficient grows as Cn, it is a weight 1 modular form with respect to $\Gamma(N)$.

We can extend this to $\Gamma_1(N)$ and $\Gamma_0(N)$ in a similar way as in the weight 2 case. Let ψ and φ be Dirichlet characters modulo u and v with $uv = N$ and φ primitive and $(\psi\varphi)(-1) = -1$. As before, consider the sums

$$G_1^{\psi,\varphi}(\tau) = \sum_{c=0}^{u-1} \sum_{d=0}^{v-1} \sum_{e=0}^{u-1} \psi(c)\overline{\varphi}(d) g_1^{\overline{(cv, d+ev)}}(\tau),$$

$$E_1^{\psi,\varphi}(\tau) = \delta(\varphi)L(0,\psi) + \delta(\psi)L(0,\varphi) + 2\sum_{n-1}^{\infty} \sigma_0^{\psi,\varphi}(n)q^n.$$

Notice the difference in the constant term of $E_1^{\psi,\varphi}$. The details of the calculation are given on page 140 of [DiSh]. As before we get

$$G_1^{\psi,\varphi} \in M_1(N, \psi\varphi), \quad G_1^{\psi,\varphi}(\tau) = \frac{C_1 g(\overline{\varphi})}{v} E_1^{\psi,\varphi}(\tau).$$

Using the same techniques as the previous basis theorems we may obtain a basis for the weight 1 Eisenstein subspace. There are however slight differences in this case. Let $A_{N,1}$ be the set of triples $(\{\psi, \varphi\}, t)$ such that ψ and φ, taken this time as an unordered pair, are primitive Dirichlet characters modulo u and v satisfying the parity condition $(\psi\varphi)(-1) = -1$, and t is a positive integer such that $tuv|N$. Suppose such a triple contained the same character ψ twice. If ψ was even then $\psi(-1)\psi(-1) = 1.1 = 1$. If ψ was odd then $\psi(-1)\psi(-1) = (-1).(-1) = 1$. Hence the parity condition shows that $A_{N,1}$ contains no triples $(\{\psi, \varphi\}, t)$ with the same character twice, so taking the characters in unordered pairs means that $A_{N,1}$ contains half as many elements as it would otherwise, and we have $|A_{N,1}| = \dim(\mathcal{E}_1(\Gamma_1(N))$. Since the Fourier coefficients of $E_1^{\psi,\varphi}$ are symmetric in ψ and φ, the series depends on the two characters only as an unordered pair, and it makes sense to define for each triple $(\{\psi, \varphi\}, t) \in A_{N,1}$

$$E_1^{\psi,\varphi,t}(\tau) = E_1^{\psi,\varphi}(t\tau).$$

Theorem 1.4.10. *Let N be a positive integer. The set*

$$\{E_1^{\psi,\varphi,t} : (\{\psi, \varphi\}, t) \in A_{N,1}\}$$

represents a basis of $\mathcal{E}_1(\Gamma_1(N))$. For any character χ modulo N, the set

$$\{E_1^{\psi,\varphi,t} : (\{\psi, \varphi\}, t) \in A_{N,1},\ \psi\varphi = \chi\}$$

represents a basis of $\mathcal{E}_1(N, \chi)$.

§ 1.5 Hecke Operators

The main aim of this section is to give a canonical basis for the space $S_k(\Gamma_1(N))$. One particularly useful way of learning more about the Fourier coefficients of modular forms is by making use of Hecke operators. The idea is to create some linear operators for which certain modular forms will be eigenforms and the Fourier coefficients will be precisely the Hecke eigenvalues. This rather clever idea came about while trying to prove various relations such as the Ramanujan congruence. In this section we will describe the theory of Hecke operators.

1.5.1 The Double Coset Operator

First, we fix congruence subgroups Γ_1 and Γ_2 of $\mathrm{SL}_2(\mathbb{Z})$. We can therefore view these congruence subgroups as subgroups of $\mathrm{GL}_2^+(\mathbb{Q})$, the group of 2-by-2 matrices with positive determinant and rational entries. Our aim is to construct certain maps between $M_k(\Gamma_1)$ and $M_k(\Gamma_2)$. That is we would like maps that turn modular forms for Γ_1 into modular forms for Γ_2. We can do this by making use of a particular type of coset. For any $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$, the set

$$\Gamma_1\alpha\Gamma_2 = \{\gamma_1\alpha\gamma_2 : \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}$$

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

We can act on such a double coset by left multiplication by Γ_1. This action partitions the double coset into orbits. These orbits have the form $\Gamma_1\beta$ with representative $\beta = \gamma_1\alpha\gamma_2$. Hence the orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$ is a disjoint union $\bigcup\Gamma_1\beta_j$ for some choice of representatives β_j. This is a finite union, the proof of which can be found on page 164 of [DiSh].

Now that we are working with elements of $\mathrm{GL}_2^+(\mathbb{Q})$ it makes sense to extend our definition of the weight-k slash operator. For any $\beta \in \mathrm{GL}_2^+(\mathbb{Q})$ and $k \in \mathbb{Z}$, the weight-k β operator on functions $f : \mathcal{H} \to \mathbb{C}$ is given by

$$(f[\beta]_k)(\tau) = (\det\beta)^{k-1} j(\beta,\tau)^{-k} f(\beta(\tau)),\ \tau \in \mathcal{H}.$$

We may now use this extended operator to define the double coset operators.

Definition 1.5.1. For congruence subgroups Γ_1 and Γ_2 of $\mathrm{SL}_2(\mathbb{Z})$ and $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$, the weight-$k$ $\Gamma_1\alpha\Gamma_2$ operator takes functions $f \in M_k(\Gamma_1)$ to

$$f[\Gamma_1\alpha\Gamma_2]_k = \sum_j f[\beta_j]_k$$

where $\{\beta_j\}$ are orbit representatives, i.e., $\Gamma_1\alpha\Gamma_2 = \bigcup_j \Gamma_1\beta_j$ is a disjoint union.

We might wonder what happens when we choose different representatives $\{\beta_j\}$. In fact, this is not an issue at all. The double coset operators are well-defined and are independent of how the representatives $\{\beta_j\}$ are chosen. As mentioned briefly we want maps that transform modular forms for Γ_1 into modular forms for Γ_2. These maps are exactly the right ones.

Theorem 1.5.2. *For any $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$, the weight-$k$ double coset operator $[\Gamma_1\alpha\Gamma_2]_k$ defines a linear map $M_k(\Gamma_1) \to M_k(\Gamma_2)$. This map induces a linear map $S_k(\Gamma_1) \to S_k(\Gamma_2)$.*

Proof. A proof of this result is given on pages 165-166 of [DiSh]. □

There are a few interesting examples we can consider. Suppose that $\Gamma_1 \supset \Gamma_2$ and we take $\alpha = I$. Here the double coset operator is $f[\Gamma_1\alpha\Gamma_2]_k = f$. This gives us the natural inclusion of $M_k(\Gamma_1)$ into $M_k(\Gamma_2)$. Suppose instead that $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$. Then the double coset operator is simply $f[\Gamma_1\alpha\Gamma_2]_k = f[\alpha]_k$, giving the natural translation from $M_k(\Gamma_1)$ to $M_k(\Gamma_2)$.

1.5.2 The $\langle d\rangle$ and T_p Operators

Suppose we take $\Gamma_1 = \Gamma_2 = \Gamma$. Then Theorem 1.5.2 says we have an endomorphism of $M_k(\Gamma)$. This is the kind of map we will be most interested in. We will now introduce

two operators on the space $M_k(\Gamma_1(N))$. Take $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, let $\alpha \in \Gamma_0(N)$ and consider the double coset operator $[\Gamma_1(N)\alpha\Gamma_1(N)]_k$. Recall that $\Gamma_1(N)$ is normal in $\Gamma_0(N)$. This followed by considering the map

$$\Gamma_0(N)/\Gamma_1(N) \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^\times \text{ where } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \pmod N.$$

Since $\Gamma_1(N) \triangleleft \Gamma_0(N)$, we have $\alpha^{-1}\Gamma_1(N)\alpha = \Gamma_1(N)$ for all $\alpha \in \Gamma_0(N)$. By the above property, we have for each $f \in M_k(\Gamma_1(N))$,

$$f[\Gamma_1(N)\alpha\Gamma_1(N)]_k = f[\alpha]_k,\ \alpha \in \Gamma_0(N),$$

again in $M_k(\Gamma_1(N))$. It follows that the group $\Gamma_0(N)$ acts on $M_k(\Gamma_1(N))$, and since the subgroup $\Gamma_1(N)$ acts trivially, its really an action of the quotient $(\mathbb{Z}/N\mathbb{Z})^\times$. The action of $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, determined by $d \pmod N$ and denoted by $\langle d \rangle$, is

$$\langle d \rangle : M_k(\Gamma_1(N)) \to M_k(\Gamma_1(N))$$

given by

$$\langle d \rangle f = f[\alpha]_k \text{ for any } \alpha = \begin{pmatrix} a & b \\ c & \delta \end{pmatrix} \in \Gamma_0(N) \text{ with } \delta \equiv d \pmod N.$$

This operator is known as a *diamond operator* and is the first type of Hecke operator. These operators have an immediately useful property. For any character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}$, the space $M_k(N, \chi)$ is precisely the χ-eigenspace of the diamond operators,

$$M_k(N, \chi) = \{f \in M_k(\Gamma_1(N)) : \langle d \rangle f = \chi(d) f \text{ for all } d \in (\mathbb{Z}/N\mathbb{Z})^\times\}.$$

The second type of Hecke operator will again have $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, but this time we let $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$, with p prime. This operator is denoted by T_p. We have

$$T_p : M_k(\Gamma_1(N)) \to M_k(\Gamma_1(N)),\ p \text{ prime}$$

given by

$$T_p f = f[\Gamma_1(N) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma_1(N)]_k.$$

The double coset here is given by

$$\Gamma_1(N) \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \Gamma_1(N) = \left\{\gamma \in \mathrm{M}_2(\mathbb{Z}) : \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & p \end{pmatrix} \pmod N, \det\gamma = p\right\}.$$

The two types of Hecke operator here commute and the proof of this can be found on pages 169-170 in [DiSh]. Of the two operators here, the T_p operator is the one

we will be most interested in as this is the one that will appear in our main theorem. The following results will give us more information about how the T_p operator acts on modular forms. The required orbit representatives needed for the action are calculated on page 170 of [DiSh]. Here we just state the results.

Proposition 1.5.3. *Let $\in \mathbb{Z}^+$, let $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, and let $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ where p is prime. The operator $T_p = [\Gamma_1 \alpha \Gamma_2]$ on $M_k(\Gamma_1(N))$ is given by*

$$T_p f = \begin{cases} \sum_{j=0}^{p-1} f[\left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)]_k & \text{if } p|N, \\ \sum_{j=0}^{p-1} f[\left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)]_k + f[\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right)\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]_k & \text{if } p \nmid N, \text{ where } mp - nN = 1. \end{cases}$$

A similar result holds for the action on modular forms for $\Gamma_0(N)$ although the final orbit representative is replaced by $\beta_\infty = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$. This follows since $\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right) \in \Gamma_0(N)$ (We know that $mp - nN = 1$ so $\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and the bottom left entry is $0 \pmod N$). Therefore $f[\left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right)]_k = f$ as f is weakly modular with respect to $\Gamma_0(N)$. The next result gives us the action on Fourier coefficients.

Proposition 1.5.4. *Let $f \in M_k(\Gamma_1(N))$. Since $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_1(N), f$ has period 1 and hence has a Fourier expansion*

$$f(\tau) = \sum_{n=0}^{\infty} a_n(f) q^n,\ q = e^{2\pi i \tau}.$$

Then:

(1) Let $\mathbf{1}_N : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}$ be the trivial character modulo N. Then $T_p f$ has Fourier expansion

$$\begin{aligned}(T_p f)(\tau) &= \sum_{n=0}^{\infty} a_{np}(f) q^n + \mathbf{1}_N(p) p^{k-1} \sum_{n=0}^{\infty} a_n(\langle p \rangle f) q^{np} \\ &= \sum_{n=0}^{\infty} (a_{np}(f) + \mathbf{1}_N(p) p^{k-1} a_{n/p}(\langle p \rangle f)) q^n.\end{aligned}$$

That is,

$$a_n(T_p f) = a_{np}(f) + \mathbf{1}_N(p) p^{k-1} a_{n/p}(\langle p \rangle f) \text{ for } f \in M_k(\Gamma_1(N)).$$

(Here $a_{n/p} = 0$ when $n/p \notin \mathbb{N}$, $\mathbf{1}_N(p) = 1$ when $p \nmid N, \mathbf{1}_N(p) = 0$ when $p|N$.

(2) Let $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}$ be a character. If $f \in M_k(N, \chi)$ then also $T_p f \in M_k(N, \chi)$, and now its Fourier expansion is

$$\begin{aligned}(T_p f)(\tau) &= \sum_{n=0}^{\infty} a_{np}(f) q^n + \chi(p) p^{k-1} \sum_{n=0}^{\infty} a_n(f) q^{np} \\ &= \sum_{n=0}^{\infty} (a_{np}(f) + \chi(p) p^{k-1} a_{n/p}(f)) q^n.\end{aligned}$$

That is,

$$a_n(T_p f) = a_{np}(f) + \chi(p)p^{k-1}a_{n/p}(f) \text{ for } f \in M_k(N,\chi).$$

Proof. See page 172 of [DiSh]. □

Double coset operators take modular forms to modular forms and they also respect the subspace of cusp forms. That is, they take cusp forms to cusp forms. It follows that we may restrict T_p to $S_k(\Gamma_1(N))$. Recall that the space $S_{12}(\mathrm{SL}_2(\mathbb{Z}))$ was 1-dimensional with the discriminant function Δ lying in this space. This tells us that in fact Δ must be an eigenform for the T_p operator. In fact the Hecke eigenvalues are exactly the Fourier coefficients, i.e., the values of the Ramanujan τ function. This is enough to prove the conjectured relations that the τ function should satisfy.

It also turns out that the Eisenstein series are eigenvectors of the Hecke operators. By definition of $M_k(N,\chi)$ as an eigenspace, $\langle d\rangle E_k^{\psi,\varphi,t} = \chi(d)E_k^{\psi,\varphi,t}$. Also if we apply the Hecke operator T_p to this series we see the following

Theorem 1.5.5. *Let $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}$ be a Dirichlet character modulo N. Let ψ and φ be primitive Dirichlet characters modulo u and v, let t be a positive integer with $tuv|N$ and $(\psi\varphi)(-1) = (-1)^k$. Let p be prime and $k \geq 1$. Excluding the case $k = 2, \psi = \varphi = \mathbf{1}$,*

$$T_p E_k^{\psi,\varphi} = (\psi(p) + \varphi(p)p^{k-1})E_k^{\psi,\varphi} \quad \text{if } uv = N \text{ or if } p \nmid N.$$

Also,

$$T_p E_2^{\mathbf{1}_1,\mathbf{1}_1,t} = (1 + \mathbf{1}_N(p)p)E_2^{\mathbf{1}_1,\mathbf{1}_1,t} \text{ if } t \text{ is prime and } N \text{ is a power of } t \text{ or if } p \nmid N.$$

Proof. This proof is Exercise 5.2.5 in [DiSh]. We omit the proof of the case $k = 2$.

We start by showing that the generalised divisor sum $\sigma_{k-1}^{\psi,\varphi}(n)$ is multiplicative, i.e.,

$$\sigma_{k-1}^{\psi,\varphi}(nm) = \sigma_{k-1}^{\psi,\varphi}(n)\sigma_{k-1}^{\psi,\varphi}(m) \quad \text{when } \gcd(m,n) = 1.$$

If we let f be the function sending m to m^{k-1}, then we have

$$\sigma_{k-1}^{\psi,\varphi}(n) = (\varphi f * \psi).$$

Since ψ and φ are Dirichlet characters, it follows that ψ and φf are multiplicative. It then follows that the Dirichlet convolution is also multiplicative. Therefore $\sigma_{k-1}^{\psi,\varphi}$ is multiplicative.

We now let p be prime and let $n \geq 1$. Write $n = n'p^e$ with $p \nmid n'$ and $e \geq 0$. Then

$$\begin{aligned}
\sigma_{k-1}^{\psi,\varphi}(np) &= \sigma_{k-1}^{\psi,\varphi}(n')\sigma_{k-1}^{\psi,\varphi}(p^{e+1}) \\
&= \sigma_{k-1}^{\psi,\varphi}(n') \left(\sum_{m|p^{e+1}} \psi\left(\frac{p^{e+1}}{m}\right) \varphi(m) m^{k-1} \right) \\
&= \sigma_{k-1}^{\psi,\varphi}(n') \left(\psi(p^{e+1}) + \cdots + \psi(p)\varphi(p^e)(p^e)^{k-1} + \varphi(p^{e+1})(p^{e+1})^{k-1} \right) \\
&= \psi(p)\sigma_{k-1}^{\psi,\varphi}(n) + \varphi(p^{e+1})(p^{e+1})^{k-1}\sigma_{k-1}^{\psi,\varphi}(n').
\end{aligned}$$

Also

$$\begin{aligned}
\sigma_{k-1}^{\psi,\varphi}(n/p) &= \sigma_{k-1}^{\psi,\varphi}(n')\sigma_{k-1}^{\psi,\varphi}(p^{e}) \\
&= \sigma_{k-1}^{\psi,\varphi}(n') \left(\sum_{m|p^{e}} \psi\left(\frac{p^{e}}{m}\right) \varphi(m) m^{k-1} \right) \\
&= \sigma_{k-1}^{\psi,\varphi}(n') \left(\psi(p^{e}) + \cdots + \psi(p)\varphi(p^{e-1})(p^{e-1})^{k-1} + \varphi(p^{e})(p^{e})^{k-1} \right).
\end{aligned}$$

So

$$\begin{aligned}
\chi(p)p^{k-1}\sigma_{k-1}^{\psi,\varphi}(n/p) &= \psi(p)\varphi(p)p^{k-1}\sigma_{k-1}^{\psi,\varphi}(n/p) \\
&= \varphi(p)p^{k-1}\sigma_{k-1}^{\psi,\varphi}(n') \left(\psi(p^{e+1}) + \cdots + \psi(p)\varphi(p^e)(p^e)^{k-1} \right).
\end{aligned}$$

If we add and subtract $\varphi(p^{e+1})(p^{e+1})^{k-1}$ in the sum, this can be written as

$$\chi(p)p^{k-1}\sigma_{k-1}^{\psi,\varphi}(n/p) = \varphi(p)p^{k-1}\sigma_{k-1}^{\psi,\varphi}(n) - \varphi(p^{e+1})(p^{e+1})^{k-1}\sigma_{k-1}^{\psi,\varphi}(n').$$

Hence when $e \geq 0$ and $p \nmid N$ we have

$$\sigma_{k-1}^{\psi,\varphi}(np) + \chi(p)p^{k-1}\sigma_{k-1}^{\psi,\varphi}(n/p) = (\psi(p) + \varphi(p)p^{k-1})\sigma_{k-1}^{\psi,\varphi}(n).$$

But this precisely says that

$$a_n(T_p E_k^{\psi,\varphi}) = (\psi(p) + \varphi(p)p^{k-1})a_n(E_k^{\psi,\varphi}).$$

□

This result holds more generally for T_n where $n \in \mathbb{N}$, but we won't prove that here.

We now state the result saying that these Hecke operators commute.

Proposition 1.5.6. *Let d and e be elements of $(\mathbb{Z}/N\mathbb{Z})^\times$, and let p and q be prime. Then*

(1) $\langle d \rangle T_p = T_p \langle d \rangle$,

(2) $\langle d\rangle\langle e\rangle = \langle e\rangle\langle d\rangle = \langle de\rangle$,

(3) $T_pT_q = T_qT_p$.

Proof. See page 172 of [DiSh]. □

These results can all be extended to work with general n. In other words, we have operators $\langle n\rangle$ and T_n which still commute and still act on modular forms in similar ways. The following section will give the details.

1.5.3 The $\langle n\rangle$ and T_n Operators

Up to this point we have only considered $\langle d\rangle$ for $d \in (\mathbb{Z}/d\mathbb{Z})^\times$ and T_p for p prime. These definitions can be extended in a natural way.

For $n \in \mathbb{Z}^+$ with $(n, N) = 1$, $\langle n\rangle$ is determined by $n \pmod N$. If $(n, N) > 1$ then we simply define $\langle n\rangle = 0$, the zero operator on $M_k(\Gamma_1(N))$. For $n, m \in \mathbb{Z}^+$ we have $\langle nm\rangle = \langle n\rangle\langle m\rangle$. In other words the diamond operator $\langle n\rangle$ is totally multiplicative.

In order to define T_n we use the T_p's as building blocks. Set $T_1 = 1$. We already have T_p defined for primes p; for prime powers, define inductively

$$T_{p^r} = T_pT_{p^{r-1}} - p^{k-1}\langle p\rangle T_{p^{r-2}}, \text{ for } r \geq 2.$$

Note it can be shown that $T_{p^r}T_{q^s} = T_{q^s}T_{p^r}$ for distinct primes p and q. We then extend this definition multiplicatively to T_n for all n,

$$T_n = \prod T_{p_i^{e_i}} \text{ where } n = \prod p_i^{e_i}.$$

It follows by Proposition 1.5.6 that the T_n all commute and

$$T_{nm} = T_nT_m \text{ if } (n, m) = 1.$$

The formulas for the Fourier coefficients can be generalised in a fairly obvious way for these new operators. For a statement of the result, along with a proof, see page 179 of [DiSh].

1.5.4 The Petersson Inner Product and Adjoints of the Hecke Operators

Often we are interested in knowing about the space $S_k(\Gamma_1(N))$. In order to learn more about this space we can turn it into an inner product space. The inner product, known as the Petersson inner product, will be defined as an integral. This integral however only converges on the space of cusp forms, not on the larger space $M_k(\Gamma_1(N))$.

We define the *hyperbolic measure* on the upper half plane,

$$\mathrm{d}\mu(\tau) = \frac{\mathrm{d}x\mathrm{d}y}{y^2},\ \tau = x + iy \in \mathcal{H}.$$

It turns out that this measure is invariant under the automorphism group $\mathrm{GL}_2^+(\mathbb{R})$ of $\mathcal{H}$. This means that for all $\alpha \in \mathrm{GL}_2^+(\mathbb{R})$ and $\tau \in \mathcal{H}$ we have $\mathrm{d}\mu(\alpha(\tau)) = \mathrm{d}\mu(\tau)$. It therefore follows, since $\mathrm{SL}_2(\mathbb{Z})$ is a subgroup of $\mathrm{GL}_2^+(\mathbb{R})$, that $\mathrm{d}\mu$ is $\mathrm{SL}_2(\mathbb{Z})$-invariant. Since the set $\mathbb{Q} \cup \infty$ is countable and has measure zero, the measure $\mathrm{d}\mu$ suffices for integrating over the extended upper half plane $\mathcal{H}^* = \mathcal{H} \cup \mathbb{Q} \cup \infty$. Using what we know about the fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$, we can easily extend to a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}^*$. This is given by

$$\mathcal{D}^* = \{\tau \in \mathcal{H} : \mathrm{Re}(\tau) \leq 1/2, |\tau| \geq 1\} \cup \{\infty\}.$$

For any continuous bounded function $\varphi : \mathcal{H} \to \mathbb{C}$ and any $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, the integral $\int_{\mathcal{D}^*} \varphi(\alpha(\tau))\mathrm{d}\mu(\tau)$ converges. We therefore need to come up with a function with these properties. We omit the details here (they can be found in Section 5.4 of [DiSh]) but state the definition of the Petersson inner product.

Definition 1.5.7. Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. The *Petersson inner product*,

$$\langle , \rangle_\Gamma : S_k(\Gamma) \times S_k(\Gamma) \to \mathbb{C},$$

is given by

$$\langle f, g \rangle = \frac{1}{V_\Gamma} \int_{X_\Gamma} f(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k \mathrm{d}\mu(\tau).$$

Here $X(\Gamma)$ is the modular curve and V_Γ is the volume of the modular curve given by

$$V_\Gamma = \int_{X(\Gamma)} \mathrm{d}\mu(\tau).$$

It is clear that this inner product is linear in f, conjugate linear in g, Hermitian-symmetric, and positive definite. Even though this inner product is defined on $S_k(\Gamma)$ and doesn't converge on the larger space $M_k(\Gamma)$, the argument showing this actually only requires the product fg to vanish at each cusp. Therefore it suffices that only one of f and g be a cusp form. In particular we could consider one of f and g as an Eisenstein series and the other as a cusp form. This in fact always gives 0 and so in some sense we can consider the Eisenstein series and the cusp forms as orthogonal. This is exactly what we would like as we wish to define the Eisenstein space as the orthogonal complement of the cusp forms. We see this shortly in Section 1.5.6.

As with any inner product we might be interested in knowing about the adjoints of our Hecke operators using this inner product. Recall that if V is an inner product space and T is a linear operator on V, then the adjoint T^* is a linear operator on V defined by the condition

$$\langle Tv, w \rangle = \langle v, T^* w \rangle, \text{ for all } v, w \in V.$$

Also if an operator T commutes with its adjoint it is called normal. Finding the adjoints of the Hecke operators makes up the bulk of Section 5.5 of [DiSh]. Here we simply state the main result.

Theorem 1.5.8. *In the inner product space $S_k(\Gamma_1(N))$, the Hecke operators $\langle p\rangle$ and T_p for $p \nmid N$ have adjoints*

$$\langle p\rangle^* = \langle p\rangle^{-1} \text{ and } T_p^* = \langle p\rangle^{-1} T_p.$$

Thus the Hecke operators $\langle n\rangle$ and T_n for n relatively prime to N are normal.

Proof. See page 186 of [DiSh] □

We are now able to make use of the Spectral Theorem of linear algebra. We have a commuting family of normal operators on a finite-dimensional inner product space, therefore there is an orthogonal basis of simultaneous eigenvectors for the operators. Since such vectors are modular forms in this case, we call them *eigenforms*. The result is the following.

Theorem 1.5.9. *The space $S_k(\Gamma_1(N))$ has an orthogonal basis of simultaneous eigenforms for the Hecke operators $\{\langle n\rangle, T_n : (n, N) = 1\}$.*

1.5.5 Oldforms and Newforms

As mentioned previously at any given level N, we have the notion of oldforms and newforms. As can be guessed from the name, an *oldform* is a modular form that comes from a level $M|N$ with $M < N$. A *newform* is a modular form that is in $M_k(\Gamma_1(N))$ (Or any other space at level N) but does not come from a lower level. We now make this notion more precise.

The most trivial way we can view an oldform is by the observation that for $M|N$ we have $S_k(\Gamma_1(M)) \subset S_k(\Gamma_1(N))$, i.e., the inclusion of $S_k(\Gamma_1(M))$ into $S_k(\Gamma_1(N))$. This isn't the only way of embedding $S_k(\Gamma_1(M))$ into $S_k(\Gamma_1(N))$ however. We can compose with the *multiply-by-d* map, where d is any factor of N/M. For any such d, let

$$\alpha_d = \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix}$$

so that $(f[\alpha_d]_k)(\tau) = d^{k-1} f(d\tau)$ for $f : \mathcal{H} \to \mathbb{C}$. The linear map $[\alpha_d]_k$ is injective and takes $S_k(\Gamma_1(M))$ to $S_k(\Gamma_1(N))$. Naturally we need some way of distinguishing between the oldforms and genuine newforms. To do this we make use of a collection of maps.

Definition 1.5.10. For each divisor d of N, let i_d be the map

$$i_d : (S_k(\Gamma_1(Nd^{-1})))^2 \to S_k(\Gamma_1(N))$$

given by

$$(f, g) \mapsto f + g[\alpha_d]_k.$$

The subspace of *oldforms at level* N is

$$S_k(\Gamma_1(N))^{\text{old}} = \sum_{\substack{p|N \\ p \text{ prime}}} i_p((S_k(\Gamma_1(Np^{-1})))^2).$$

The subspace of *newforms at level* N is the orthogonal complement with respect to the Petersson inner product,

$$S_k(\Gamma_1(N))^{\text{new}} = (S_k(\Gamma_1(N))^{\text{old}})^{\perp}.$$

As one would hope, the Hecke operators respect this decomposition of the space $S_k(\Gamma_1(N))$.

Proposition 1.5.11. *The subspaces* $S_k(\Gamma_1(N))^{old}$ *and* $S_k(\Gamma_1(N))^{new}$ *are stable under the Hecke operators* T_n *and* $\langle n \rangle$ *for all* $n \in \mathbb{Z}^+$.

Proof. See pages 188-189 of [DiSh] □

Corollary 1.5.12. *The spaces* $S_k(\Gamma_1(N))^{old}$ *and* $S_k(\Gamma_1(N))^{new}$ *have orthogonal bases of eigenforms for the Hecke operators away from the level,* $\{T_n, \langle n \rangle : (n, N) = 1\}$.

1.5.6 Eigenforms and Eisenstein Series

Since eigenforms are a central part of the theory of modular forms it would be nice to know a little more about them. Corollary 1.5.12 told us that the spaces $S_k(\Gamma_1(N))^{\text{old}}$ and $S_k(\Gamma_1(N))^{\text{new}}$ have orthogonal bases of eigenforms for the Hecke operators. Let f be such an eigenform. It can be shown that if $f \in S_k(\Gamma_1(N))^{\text{new}}$ then in fact f is an eigenform for all T_n and $\langle n \rangle$. The details of this can be found in Section 5.8 of [DiSh].

Definition 1.5.13. A non-zero modular form $f \in M_k(\Gamma_1(N))$ that is an eigenform for the Hecke operators T_n and $\langle n \rangle$ for all $n \in \mathbb{Z}^+$ is a *Hecke eigenform* or simply an *eigenform*. The eigenform $f(\tau) = \sum_{n=0}^{\infty} a_n(f) q^n$ is *normalised* when $a_1(f) = 1$. a *newform* is a normalised eigenform in $S_k(\Gamma_1(N))^{\text{new}}$.

Theorem 1.5.14. *Let* $f \in S_k(\Gamma_1(N))^{new}$ *be a non-zero eigenform for the Hecke operators* T_n *and* $\langle n \rangle$ *for all* n *with* $(n, N) = 1$. *Then*

(1) f *is a Hecke eigenform, i.e., an eigenform for* T_n *and* $\langle n \rangle$ *for all* $n \in \mathbb{Z}^+$. *A suitable scalar multiple of* f *is a newform.*

(2) If $\tilde{f}$ *satisfies the same conditions as* f *and has the same* T_n*-eigenvalues, then* $\tilde{f} = cf$ *for some constant* c.

The set of newforms in the space $S_k(\Gamma_1(N))^{new}$ is an orthogonal basis of the space. Each such newform lies in an eigenspace $S_k(N,\chi)$ and satisfies $T_n f = a_n(f)f$ for all $n \in \mathbb{Z}^+$. That is, its Fourier coefficients are its T_n-eigenvalues.

Proof. See pages 195-196 of [DiSh] □

Now that we have considered the theory of Hecke operators and the Petersson inner product we are in a position to redefine the Eisenstein subspace. We redefine

$$\mathcal{E}_k(\Gamma(N)) = \{f \in \text{span}(\{E_k^{\overline{v}}(\tau) : v \in (\mathbb{Z}/N\mathbb{Z})^2\})|f \text{ is holomorphic}\}.$$

So now $\mathcal{E}_k(\Gamma(N))$ is a subspace of $M_k(\Gamma(N))$ linearly disjoint from the cusp forms $S_k(\Gamma(N))$, replacing the earlier definition as the quotient space $M_k(\Gamma(N))/S_k(\Gamma(N))$. That is

$$M_k(\Gamma(N)) = S_k(\Gamma(N)) \oplus \mathcal{E}_k(\Gamma(N)).$$

This decomposition is orthogonal, the details of this can be found on pages 206-207 of [DiSh]. For any congruence subgroup Γ at level N we define

$$\mathcal{E}_k(\Gamma) = \mathcal{E}_k(\Gamma(N)) \cap M_k(\Gamma),$$

and this applies in particular when $\Gamma = \Gamma_1(N)$. Redefine for any Dirichlet character χ modulo N

$$\mathcal{E}_k(N,\chi) = \mathcal{E}_k(\Gamma_1(N)) \cap M_k(N,\chi).$$

For any congruence subgroup Γ the Eisenstein space is linearly disjoint from the cusp forms and similarly for the eigenspaces, we thus have a direct sum,

$$M_k(\Gamma) = S_k(\Gamma) \oplus \mathcal{E}_k(\Gamma) \text{ and } M_k(N,\chi) = S_k(N,\chi) \oplus \mathcal{E}_k(N,\chi),$$

and the decompositions are orthogonal. The sets of Eisenstein series specified earlier as coset representatives for bases of the Eisenstein spaces as quotients are now actual bases of the Eisenstein spaces as complements.

Chapter 2

Class Field Theory

Class field theory is one of the most ground breaking areas of modern algebraic number theory. It gave the answers to some very important questions. Given a number field K, is there a way to classify all possible finite extensions of K purely in terms of the arithmetic in K? This question is very broad in scope and is not yet fully answered. A well known result concerning abelian extensions is the Kronecker-Weber Theorem. This theorem states that the abelian extensions of $\mathbb{Q}$ are all subfields within some cyclotomic extension of $\mathbb{Q}$, therefore they are expressible in terms of roots of unity. This therefore characterises all possible abelian extensions of $\mathbb{Q}$. We could then ask about all abelian extensions of any number field K. If we restrict to this case, then this question is answered by class field theory. Along the way many interesting questions will be answered. For example: Given a number field K and a finite extension L, in what way does the prime ideal $\mathfrak{p} \in K$ factorise in L? Further, which primes of K ramify in L? Given a number field K, what are all the abelian extensions of K? Suppose we want a particular prime to ramify in an abelian extension of K, what should this extension be? All of these questions can be answered by using (global) class field theory.

Ultimately class field theory is about the intricate link between the so called generalised ideal class groups and the Galois groups of abelian extensions of a number field, i.e., $\mathrm{Gal}(L/K)$. We will see, via a special map known as the Artin map, a tight correspondence known as the Artin reciprocity law. In fact this relation will be an isomorphism. This result is very broad and can be used to prove all previously known reciprocity laws such as Gauss' quadratic reciprocity law. The Artin reciprocity law tells us precisely how primes factorise in abelian extensions via mod $\mathfrak{m}$ behaviour for a particular modulus to be defined later. In other words, we know exactly how primes behave in an extension, purely by information from the base field. We will also see that we can in fact go the other way. We can choose our modulus in such a way that we are fixing the primes we would like to ramify; this then completely determines the abelian extension L/K!

Given this close relationship between the generalised ideal class groups and the Galois groups of abelian extensions, we might also wonder whether it is possible to determine the maximal abelian extension with given ramification. This is in fact possible and these fields are known as ray class fields. We could also consider the case of determining the maximal unramified abelian extension; the corresponding field is known as the Hilbert class field. Once we have described the main results of class field theory we will see more about these fields.

The following background material follows the structure of some notes by Fretwell [Fret]. For some motivation on why class field theory can be useful, see the introduction of these notes. There are many well known sources that can be used for class field theory such as [Mil], [Ch], [N] and [Cox]. Of course there are also many other useful sources.

§ 2.1 Recap of Basic Algebraic Number Theory

2.1.1 Number Fields

For a more thorough background any good text in algebraic number theory will suffice. For example [J] or [StTa] cover all of the necessary algebraic number theory. We begin with the definition of a number field.

Definition 2.1.1. A *number field* is a field $\mathbb{Q} \subset K \subset \mathbb{C}$ such that the extension $K/\mathbb{Q}$ has finite degree.

It is easy to show that any such extension must be algebraic. It follows that any element $\alpha \in K$ must satisfy a minimal polynomial over $\mathbb{Q}$. Clearing denominators, this polynomial has coefficients in $\mathbb{Z}$.

Definition 2.1.2. The *ring of integers* of a number field K is:

$$\mathcal{O}_K = \{\alpha \in K | f(\alpha) = 0 \text{ for some monic } f(x) \in \mathbb{Z}[x]\}.$$

The ring of integers is a Dedekind domain and is therefore Noetherian. This means that we always have a factorisation into irreducibles; although this need not be unique. We can however get unique factorisation if we consider factorisation of ideals into prime ideals.

Theorem 2.1.3. *Given a proper ideal $\mathfrak{a} \in \mathcal{O}_K$, there exist prime ideals $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_g$ of $\mathcal{O}_K$ and positive integers $e_1, e_2, \ldots, e_g$ such that*

$$\mathfrak{a} = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2}\ldots\mathfrak{p}_g^{e_g}.$$

Further, this factorisation is unique up to reordering.

Given a non-zero ideal $\mathfrak{a} \in \mathcal{O}_K$ we can consider the quotient ring $\mathcal{O}_K/\mathfrak{a}$. We get the following result:

Proposition 2.1.4. *If K is a number field and $\mathfrak{a}$ is a non-zero ideal of $\mathcal{O}_K$, then the quotient ring $\mathcal{O}_K/\mathfrak{a}$ is finite.*

Another benefit of $\mathcal{O}_K$ being a Dedekind domain is that every non-zero prime ideal is also maximal. This tells us that the quotient $\mathcal{O}_K/\mathfrak{p}$ is a field. Proposition 2.1.4 then tells us that this is a finite field. In fact it is a finite field of characteristic p and so must contain p^{f_i} elements for some f_i. Now consider one such $\mathfrak{p}_i$ from the factorisation of Theorem 2.1.3.

Definition 2.1.5. The finite field $\mathcal{O}_K/\mathfrak{p}_i$ is called the *residue field* of $\mathfrak{p}_i$. This is denoted by $\mathbb{F}_{\mathfrak{p}_i}$. The positive integer f_i is known as the *inertia degree*. The positive integer e_i is called the *ramification degree* of $\mathfrak{p}_i$. We say that $\mathfrak{a}$ ramifies in $\mathcal{O}_K$ (or simply in K) if some $e_i > 1$.

As we previously mentioned, there is a nice relation between each of these numbers.

Theorem 2.1.6. *Let K be a number field with ring of integers $\mathcal{O}_K$. Then for any prime ideal factorisation*

$$\mathfrak{a} = \mathfrak{p}_1^{e_1}\mathfrak{p}_2^{e_2}\dots\mathfrak{p}_g^{e_g}$$

we have $\sum_{i=1}^{g} e_i f_i = [K:\mathbb{Q}]$.

This is quite a useful result as it tells us a lot about how certain ideals can factorise. In general there are three types of behaviour that can occur. We say an ideal $\mathfrak{a}$ is *inert* if it is already a prime ideal of $\mathcal{O}_K$. We say that $\mathfrak{a}$ *splits* if $\mathfrak{a}$ is a product of two or more prime ideals. Further we say that $\mathfrak{a}$ splits completely if $e_i = f_i = 1$ for each i. Finally, as already mentioned, we say that $\mathfrak{a}$ *ramifies* if some $e_i > 1$. In general it is possible to see a combination of splitting and ramification together. However consider the case of K being a quadratic extension of $\mathbb{Q}$. Then there are only three possibilities for the factorisation type of an ideal $\mathfrak{a}$. We have either:

1 $\mathfrak{a} = \mathfrak{p}$ is inert. So $g = e = 1, f = 2$.

2 $\mathfrak{a} = \mathfrak{p}_1\mathfrak{p}_2$ splits. So $g = 2, e_1 = e_2 = f_1 = f_2 = 1$.

3 $\mathfrak{a} = \mathfrak{p}^2$ ramifies. So $g = 1, e = 2, f = 1$.

It turns out, as we will see later, that this result will give even fewer possibilities for possible factorisation types when working in a Galois extension.

For any ideal $\mathfrak{a}$ we define the *norm* to be $\mathrm{N}(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$. This norm is multiplicative and always gives integer values. Note that for a prime ideal $\mathfrak{p}$ we have $\mathrm{N}(\mathfrak{p}) = p^f$ where f is the inertia degree.

2.1.2 Relative Extensions of Number Fields

So far we have purely been considering the factorisation of ideals in K for some number field K. What about if we consider a (finite) extension L/K? We could still ask how ideals factorise further in L. In particular, given a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, what does $\mathfrak{p}\mathcal{O}_L$ look like? Given a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ we get an ideal $\mathfrak{p}\mathcal{O}_L$ of $\mathcal{O}_L$ with factorisation

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1^{e_1}\mathfrak{q}_2^{e_2}\dots\mathfrak{q}_g^{e_g}.$$

Each $\mathfrak{q}_i$ defines a prime ideal $\mathfrak{p} = \mathfrak{q}_i \cap \mathcal{O}_K$ of $\mathcal{O}_K$. Hence the finite field $\mathbb{F}_{\mathfrak{q}_i} = \mathcal{O}_L/\mathfrak{q}_i$ contains $\mathbb{F}_\mathfrak{p}$ as a subfield. It follows that $\mathbb{F}_{\mathfrak{q}_i}/\mathbb{F}_\mathfrak{p}$ is a finite extension of finite fields. This extension has degree p^{f_i} where p is the rational prime lying below $\mathfrak{q}_\mathfrak{i}$. In other words we have $\mathfrak{q}_i \cap \mathbb{Z} = \mathfrak{p} \cap \mathbb{Z} = p$.

As before we define the various numerical quantities in a similar fashion.

Definition 2.1.7. The positive integer f_i is called the *inertia degree* of $\mathfrak{q}_i$ in K. The positive integer e_i is called the *ramification index* of $\mathfrak{a}$ in L. We say that $\mathfrak{a}$ *ramifies* in $\mathcal{O}_L$ (or simply in L) if some $e_i > 1$.

Again, we end up with a similar relation as before.

Theorem 2.1.8. *Let L/K be a finite extension of number fields. For any prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ we have a factorisation*

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1^{e_1}\mathfrak{q}_2^{e_2}\dots\mathfrak{q}_g^{e_g}$$

satisfying $\sum_{i=1}^{g} e_i f_i = [L : K]$.

Similarly to the way we defined the norm earlier we can define a norm relative to the extension L/K.

Definition 2.1.9. Given a proper prime ideal $\mathfrak{q}$ of $\mathcal{O}_L$ we define the *relative norm* to be $N_{L/K}(\mathfrak{q}) = \mathfrak{p}^f$ where $\mathfrak{q} \cap \mathcal{O}_K = \mathfrak{p}$.

This definition can be extended multiplicatively to define a relative norm for any non-zero proper ideal of $\mathcal{O}_L$. This norm will turn out to be very important as it will appear in the statement of the Artin reciprocity law.

2.1.3 The Ideal Class Group

Given that certain number fields do not have unique factorisation, it was a natural question to wonder how far away from having unique factorisation a particular number field was. The reason a number field can fail to have unique factorisation is because certain ideals are not principal. This is because, in some sense, you are trying to consider multiples of elements that shouldn't be in whichever ring you are working

with. However if you were to work with an extension of that ring, the elements would genuinely exist there and you would end up with unique factorisation. If all ideals of your ring were principal, then you would obtain unique factorisation. This led to the idea of trying to consider when two ideals were different (in the sense of not being the same upto a multiple of a principal ideal). This can be achieved by considering a particular quotient group known as the ideal class group.

For a number field K, the ideals of $\mathcal{O}_K$ almost form a group under multiplication. The only issue is that we do not have inverses. The way to solve this is by introducing the notion of a fractional ideal.

Definition 2.1.10. A *fractional ideal* of a number field K is a non-zero $\mathcal{O}_K$-submodule of K.

For any $\alpha \in K^\times$ we write a fractional ideal in the form $\alpha^{-1}\mathfrak{a}$. We then have the following:

Theorem 2.1.11. *Every fractional ideal of K is invertible. Hence the set I_K of fractional ideals of K is an abelian group under multiplication. The principal fractional ideals form a subgroup denoted P_K.*

We note that the unique factorisation of an ideal $\mathfrak{a}$ of K into prime ideals of $\mathcal{O}_K$ can be extended to the fractional ideals. In other words, every fractional ideal has a unique factorisation into prime fractional ideals. We may now define the ideal class group.

Definition 2.1.12. The *ideal class group* C_K of a number field K is the abelian group I_K/P_K. The order of this group is called the *class number* of K denoted h_K or simply h.

A number field K has unique factorisation precisely when $h_K = 1$, in other words, all ideals are principal. We might wonder what the possibilities are for h_K; in particular, can it be infinite? Fortunately, it is always finite. This is due to the fact that the ideal class group is a finite abelian group (non-obvious). There are methods of determining the class number for a given number field K but we will not go into that here.

§ 2.2 The Main Theorems of Global Class Field Theory

Now that we have covered the necessary background on basic algebraic number theory we are ready to work our way towards the Artin Reciprocity Theorem and the Existence Theorem. In doing so we will restrict our attention to Galois extensions of number fields. This makes the results much nicer while not really restricting us too much as the problems we are often interested in involve Galois extensions.

2.2.1 The Action of the Galois Group and Frobenius Elements

Consider a Galois extension L/K of number fields. Given a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ we have the following factorisation:

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{q}_1^{e_1}\mathfrak{q}_2^{e_2}\dots\mathfrak{q}_g^{e_g}.$$

As we have already stated, we don't really need to know each $\mathfrak{q}_i$, we simply want to know the value of each e_i, f_i and g. This will tell us precisely the behaviour of the factorisation, i.e., does the prime split? Is it inert? Does it ramify? One way that we can do this is by considering the primes $\mathfrak{q}_i$ as a set. We can then study that set by using a group action. The particular group we will use is the Galois group $\mathrm{Gal}(L/K)$.

Lemma 2.2.1. *The Galois group* $\mathrm{Gal}(L/K)$ *acts on each of the sets* $X_\mathfrak{p} = \{\mathfrak{q}_1, \mathfrak{q}_2, \dots, \mathfrak{q}_g\}$ *of prime ideal divisors of* $\mathfrak{p}\mathcal{O}_L$. *Further, this action is transitive.*

Consider an element $\sigma \in \mathrm{Gal}(L/K)$. It is easy to see that the set $\sigma(\mathfrak{q}_i)$ is again a prime ideal of $\mathcal{O}_L$. Also we have that

$$\mathfrak{p}\mathcal{O}_L = \sigma(\mathfrak{p})\sigma(\mathcal{O}_L) = \sigma(\mathfrak{p}\mathcal{O}_L) = \sigma(\mathfrak{q}_1)^{e_1}\sigma(\mathfrak{q}_2)^{e_2}\dots\sigma(\mathfrak{q}_g)^{e_g}.$$

We then see, by unique factorisation, that $\sigma(\mathfrak{q}_i) = \mathfrak{q}_j \in X_\mathfrak{p}$ for some j. This shows that the operation on $X_\mathfrak{p}$ is well-defined. The other group action axioms can be checked easily. The proof of transitivity can be found in any good book on algebraic number theory.

This group action will tell us a lot about the various factorisation types. Another result which helps make things easier involves the relation between the e_i's, f_i's and g. By the previous lemma, we see that in the case of a Galois extension, things become much simpler.

Corollary 2.2.2. *If* L/K *is a Galois extension then for any factorisation of a prime ideal* $\mathfrak{p} \subset \mathcal{O}_K$ *in* $\mathcal{O}_L$ *we have that* $e_1 = e_2 = \dots = e_g$ *(call the common value* e*) and* $f_1 = f_2 = \dots = f_g$ *(call the common value* f*). Hence* $efg = [L : K]$.

In particular our factorisation now looks much simpler; we have

$$\mathfrak{p}\mathcal{O}_L = (\mathfrak{q}_1\mathfrak{q}_2\dots\mathfrak{q}_g)^e.$$

One way we might think to study the set $X_\mathfrak{p}$ is via the stabilizer subgroups. From now on, we fix a $\mathfrak{p} \in \mathcal{O}_K$.

Definition 2.2.3. Let L/K be a Galois extension of number fields. Given a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ and a prime ideal $\mathfrak{q}$ of $\mathcal{O}_L$ such that $\mathfrak{q}|\mathfrak{p}\mathcal{O}_L$ we define the *decomposition group* to be:

$$D_\mathfrak{q} := \mathrm{Stab}(\mathfrak{q}) = \{\sigma \in \mathrm{Gal}(L/K) |\, \sigma(\mathfrak{q}) = \mathfrak{q}\}.$$

One of the first questions we might ask about this group is about the size of the group. Fortunately, since the group is defined as a stabilizer, we can make use of the Orbit-Stabilizer Theorem.

Lemma 2.2.4. *We have* $|D_\mathfrak{q}| = ef$ *for all* $\mathfrak{q}|\mathfrak{p}\mathcal{O}_L$.

Proof. Since the action is transitive, there is a single orbit. Namely $X_\mathfrak{p}$ containing g elements. The Orbit-Stabilizer Theorem then tells us that

$$|D_\mathfrak{q}| = \frac{|\mathrm{Gal}(L/K)|}{|\mathrm{orb}(\mathfrak{q})|} = \frac{[L:K]}{g} = \frac{efg}{g} = ef.$$

Here we have used the assumption that L/K is a Galois extension. □

Notice that this result gives us a way of finding the value of g. If we can determine any of the decomposition groups, we will be able to calculate its size. We can then simply calculate g by working out $\frac{[L:K]}{|D_\mathfrak{q}|}$. The question now is how to calculate e and f. This turns out to be a little harder; but not too much harder.

The idea is to try and cook up a map from $D_\mathfrak{q}$ into the Galois group of residue fields. This map will be created in such a way that the kernel will have size e.

First of all, notice that for any $\sigma \in \mathrm{Gal}(L/K)$, there is an induced isomorphism:

$$\tilde{\sigma} : \mathbb{F}_\mathfrak{q} \to \mathbb{F}_{\sigma(\mathfrak{q})}$$

$$x + \mathfrak{q} \mapsto \sigma(x) + \sigma(\mathfrak{q}).$$

Since $D_\mathfrak{q}$ is a subgroup of $\mathrm{Gal}(L/K)$ it makes sense to restrict this map to elements $\sigma \in D_\mathfrak{q}$. It is clear that if we do this, the induced isomorphism will be an automorphism of $\mathbb{F}_\mathfrak{q}$. Further, these automorphisms will fix elements of the subfield $\mathbb{F}_\mathfrak{p}$. Hence $\tilde{\sigma}$ is a well-defined element of $\mathrm{Gal}(\mathbb{F}_\mathfrak{q}/\mathbb{F}_\mathfrak{p})$.

We therefore obtain the following:

Theorem 2.2.5. *The map*

$$D_\mathfrak{q} \to \mathrm{Gal}(\mathbb{F}_\mathfrak{q}/\mathbb{F}_\mathfrak{p})$$

$$\sigma \mapsto \tilde{\sigma}$$

is an epimorphism of groups inducing an isomorphism:

$$D_\mathfrak{q}/I_\mathfrak{q} \cong \mathrm{Gal}(\mathbb{F}_\mathfrak{q}/\mathbb{F}_\mathfrak{p}),$$

where $I_\mathfrak{q} = \{\sigma \in D_\mathfrak{q} | \, \sigma(x) \equiv x \pmod{\mathfrak{q}}$ *for all* $x \in \mathcal{O}_L\}$.

Definition 2.2.6. The group $I_\mathfrak{q}$ above is called the *inertia group* of $\mathfrak{q}$.

We are now in a position to finally get our hands on the values of e and f.

Corollary 2.2.7. *We have that* $|I_\mathfrak{q}| = e$ *and* $|D_\mathfrak{q}/I_\mathfrak{q}| = f$ *for each* $\mathfrak{q}|\mathfrak{p}\mathcal{O}_L$.

Proof. The right hand side of the above isomorphism has order f by definition. We already know that $|D_\mathfrak{q}| = ef$, therefore the isomorphism gives $|I_\mathfrak{q}| = e$. □

We now have two subgroups of the Galois group $\mathrm{Gal}(L/K)$ whose sizes give us information about the factorisation types of ideals in the extension L/K. Although we know a lot from these subgroups, there is actually more information that we can extract.

A finite extension of finite fields has a cyclic Galois group whose canonical generator is the Frobenius automorphism $x \mapsto x^{|K|}$. In other words $\mathrm{Gal}(\mathbb{F}_\mathfrak{q}/\mathbb{F}_\mathfrak{p})$ is cyclic with generator being the Frobenius automorphism. However, we already know that this group is isomorphic to $D_\mathfrak{q}/I_\mathfrak{q}$. There must therefore be a unique element of $D_\mathfrak{q}/I_\mathfrak{q}$ that corresponds to the Frobenius automorphism. Recall that $\mathrm{N}(\mathfrak{p}) = |\mathcal{O}_K/\mathfrak{p}| = |\mathbb{F}_\mathfrak{p}|$. It therefore follows that this map is given by $x + \mathfrak{q} \mapsto x^{\mathrm{N}(\mathfrak{p})} + \mathfrak{q}$. Since the group $D_\mathfrak{q}/I_\mathfrak{q}$ is a quotient, this element will in general be a coset.

Theorem 2.2.8. *Let* $\mathfrak{p}$ *be a fixed prime ideal of* $\mathcal{O}_K$ *and let* $\mathfrak{q}$ *be a fixed prime ideal of* $\mathcal{O}_L$ *dividing* $\mathfrak{p}\mathcal{O}_L$. *Then there exists an element* $\sigma \in D_\mathfrak{q}$ *that satisfies* $\sigma(x) \equiv x^{\mathrm{N}(\mathfrak{p})} \pmod{\mathfrak{q}}$ *for all* $x \in \mathcal{O}_L$. *The set of such elements forms a coset of* $I_\mathfrak{q}$ *in* $D_\mathfrak{q}$. *If* $\mathfrak{p}$ *is unramified in* L *then* σ *is a unique element of* $D_\mathfrak{q}$.

Definition 2.2.9. The coset $\sigma(I_\mathfrak{q}) \in D_\mathfrak{q}/I_\mathfrak{q}$ in the above theorem is called the *Frobenius coset* of $\mathfrak{q}$ in L/K. For an unramified prime $\mathfrak{p}$, we call the unique element the *Frobenius element* of $\mathfrak{q}$. We denote this element by $\left(\frac{L/K}{\mathfrak{q}}\right)$ or simply $\mathrm{Frob}_\mathfrak{q}$ when the extension is understood.

Since this Frobenius element is related to the Frobenius automorphism which is a canonical generator for $\mathrm{Gal}(\mathbb{F}_\mathfrak{q}/\mathbb{F}_\mathfrak{p})$ we can easily determine the order of this element.

Lemma 2.2.10. *Let* $\mathfrak{p}$ *be unramified in* L. *The Frobenius element* $\left(\frac{L/K}{\mathfrak{q}}\right)$ *has order* f *in* $D_\mathfrak{q}$ *for all* $\mathfrak{q}$ *dividing* $\mathfrak{p}\mathcal{O}_L$.

Further, the Frobenius element is the identity automorphism if and only if $\mathfrak{p}$ *splits completely in* L.

Proof. Since we assume $\mathfrak{p}$ is unramified, it follows that $I_\mathfrak{q}$ is trivial and we have $D_\mathfrak{q} \cong \mathrm{Gal}(\mathbb{F}_\mathfrak{q}/\mathbb{F}_\mathfrak{p})$. Since the right hand side of this isomorphism is cyclic of order f, we conclude that $D_\mathfrak{q}$ is cyclic of order f. Since this is generated by the Frobenius element, we conclude that $\left(\frac{L/K}{\mathfrak{q}}\right)$ has order f.

The second claim follows since $\mathfrak{p}$ is unramified (i.e. we have $e = 1$) and an element of a group is the identity if and only if it has order 1.

□

We now know that each of the Frobenius elements for a given $\mathfrak{q}$ has the same order when $\mathfrak{p}$ is unramified. We might wonder whether each of the Frobenius elements are actually equal. In general, they won't be. We have the following result:

Theorem 2.2.11. *Let $\mathfrak{p}$ be unramified in L and let $\mathfrak{q}$ divide $\mathfrak{p}\mathcal{O}_L$. Then for all $\sigma \in \mathrm{Gal}(L/K)$ we have that*

$$\left(\frac{L/K}{\sigma(\mathfrak{q})}\right) = \sigma \left(\frac{L/K}{\mathfrak{q}}\right) \sigma^{-1}.$$

This tells us that in fact Frobenius elements have more structure to them than it first appears. They in fact form a conjugacy class in $\mathrm{Gal}(L/K)$. This result will be quite important in our later work on modular forms of weight 1. For now we state a result which explains why our extensions being abelian is so important.

Corollary 2.2.12. *If $\mathrm{Gal}(L/K)$ is abelian then the Frobenius elements are all equal for a given unramified $\mathfrak{p}$.*

Proof. We know that conjugacy classes of abelian groups consist of single elements. Hence Theorem 2.2.11 tells us that the Frobenius elements are equal, since they lie in the same conjugacy class. □

Definition 2.2.13. We call L/K an *abelian extension* if L/K is Galois and $\mathrm{Gal}(L/K)$ is abelian. In such an extension, we may denote the single Frobenius element attached to all $\mathfrak{q}|\mathfrak{p}\mathcal{O}_L$ by $\left(\frac{L/K}{\mathfrak{p}}\right)$, where $\mathfrak{p}$ is unramified.

We now notice that, in the case of an abelian extension, the Frobenius elements depend only on $\mathfrak{p}$, not on the $\mathfrak{q}$'s. In other words, there is only a dependence on an ideal coming from the base field. It turns out that we will be able describe the Frobenius elements in terms of congruence conditions.

2.2.2 The Artin Map for Abelian Extensions

Our main aim will be to construct a group homomorphism between a certain set of ideals (which will actually have a group structure) and the Galois group $\mathrm{Gal}(L/K)$. We know that the nicest type of situation is when we are considering unramified prime ideals $\mathfrak{p} \subseteq \mathcal{O}_K$ of an abelian extension L/K. We therefore restrict to this case. Consider the following map:

$$\{\text{unramified prime ideals } \mathfrak{p} \subseteq \mathcal{O}_K\} \to \mathrm{Gal}(L/K)$$

$$\mathfrak{p} \mapsto \left(\frac{L/K}{\mathfrak{p}}\right).$$

Considering this set alone will not get us very far since there is no group structure on the left hand side. We somehow need to come up with a set of ideals, not containing

any ramified primes, with a group structure. We obviously can't just take the set of fractional ideals I_K. We therefore need some way of modifying this set to cut out all of the ramified primes. This is where we need the notion of a modulus. The reason for this notation is because this will be the modulus of our congruence conditions.

Recall that a number field may be viewed as a subfield of the complex numbers. Each such number field has a number of embeddings into $\mathbb{C}$ (corresponding to the degree of the extension). If the embedding is genuinely contained in $\mathbb{C}$ it is called a *complex embedding* and if it is contained in $\mathbb{R}$ it is called a *real embedding*. We can extend the embeddings of $K \to \mathbb{C}$ to embeddings $L \to \mathbb{C}$. It is now possible that a real embedding of K can extend to give two conjugate embeddings of $\mathbb{C}$.

The importance of the real embeddings here is that they give us a notion of positivity. Consider the quadratic field $\mathbb{Q}(\sqrt{3})$. Here there are two real embeddings:

$$\sigma_1 : a + b\sqrt{3} \mapsto a + b\sqrt{3},$$

$$\sigma_2 : a + b\sqrt{3} \mapsto a - b\sqrt{3}.$$

How might we define positivity here? Consider $1 + \sqrt{3}$. We have $\sigma_1(1 + \sqrt{3}) > 0$, $\sigma_2(1 + \sqrt{3}) < 0$. Notice that this element is positive under one embedding but not the other. Consider however the element $\alpha = 19 + 2\sqrt{3}$. We now have $\sigma_1(\alpha) > 0$, $\sigma_2(\alpha) > 0$. Hence α is positive under all real embeddings. Such elements are called *totally positive*. This behaviour extends to other number fields. We now define a modulus for a general number field.

Definition 2.2.14. A *modulus* of a number field K is a formal product $\mathfrak{m} = \mathfrak{m}_0\mathfrak{m}_\infty$, where $\mathfrak{m}_0$ is an ideal of $\mathcal{O}_K$ and $\mathfrak{m}_\infty$ is a collection of real embeddings $K \to \mathbb{R}$.

We may now use a given modulus $\mathfrak{m}$ to create a nice subgroup of I_K. We note that two fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ are said to be coprime if they share no prime ideal factors.

Theorem 2.2.15. *Given a modulus $\mathfrak{m}$ of K, the set $I_K(\mathfrak{m}) = \{\mathfrak{a} \in I_K \,|\, \mathfrak{a}$ coprime to $\mathfrak{m}_0\}$ is a subgroup of I_K. It contains the set $P_{1,K}(\mathfrak{m}) = \{\langle\alpha\rangle \in P_K \,|\, \alpha \equiv 1 \pmod{\mathfrak{m}_0}$ and $\sigma(\alpha) > 0$ for all $\sigma \in \mathfrak{m}_\infty\}$ as a subgroup.*

Now consider an abelian extension L/K. This extension has a finite set of ramified primes (which may include infinite primes, i.e., real embeddings that extend to give conjugate complex embeddings). We may therefore put each of these ramified primes together to form a modulus $\mathfrak{m}$. We will then be able to define the Frobenius element on $I_K(\mathfrak{m})$. We have a map:

$$\Phi_\mathfrak{m} : I_K(\mathfrak{m}) \to \mathrm{Gal}(L/K).$$

Definition 2.2.16. Given an abelian extension of number fields L/K and a modulus $\mathfrak{m}$ of K divisible by all ramified primes of K in L the map $\Phi_\mathfrak{m}$ is called the *Artin map* of L/K with respect to $\mathfrak{m}$.

This map contains all of the information we would like to know on the splitting behaviour of ideals. It is clear that the map is a homomorphism, but it would be beneficial if we could form an isomorphism. This would immediately tell us that the Frobenius elements were completely determined by classes of ideals in K. In other words, we would know about the splitting of primes in the extension L purely from arithmetic information coming from K. This is exactly what the Artin Reciprocity Theorem tells us.

2.2.3 Artin Reciprocity

Since we wish to create an isomorphism, we are going to have to quotient $I_K(\mathfrak{m})$ by some subgroup. Recall that the ideal class group was defined as I_K/P_K. The first thing you might think of as a generalisation of this would be $I_K(\mathfrak{m})/P_K(\mathfrak{m})$. However $P_K(\mathfrak{m})$ would not satisfy the correct positivity criteria under each real embedding. This is where the subgroup $P_{1,K}(\mathfrak{m})$ comes in. We would expect this subgroup to lie in the kernel of a suitable Artin map. In fact, this won't in general be the full kernel, but the Artin Reciprocity Theorem tells us precisely what the kernel is.

Theorem 2.2.17 (Artin Reciprocity Law)**.** *Let L/K be an abelian extension of number fields. Suppose $\mathfrak{m}$ is a modulus of K divisible only by primes of K that ramify in L. Then:*

(1) The Artin map $\Phi_{\mathfrak{m}}$ is a surjective homomorphism.

(2) If the powers of the prime ideals in $\mathfrak{m}$ are big enough then we are able to guarantee that $\ker(\Phi_{\mathfrak{m}})$ *is a* congruence subgroup for $\mathfrak{m}$*, meaning that:*

$$P_{1,K}(\mathfrak{m}) \subseteq \ker(\Phi_{\mathfrak{m}}) \subseteq I_K(\mathfrak{m})$$

so that $I_K(\mathfrak{m})/\ker(\Phi_{\mathfrak{m}})$ is a generalised ideal class group for $\mathfrak{m}$ *(the definition of this is a quotient $I_K(\mathfrak{m})/H$ where H contains $P_{1,K}(\mathfrak{m})$).*

(3) Further, for such an $\mathfrak{m}$ we have that $\ker(\Phi_{\mathfrak{m}}) = P_{1,K}(\mathfrak{m})N_{L/K}(I_L(\mathfrak{m}))$*, giving an isomorphism:*

$$I_K(\mathfrak{m})/P_{1,K}(\mathfrak{m})N_{L/K}(I_L(\mathfrak{m})) \cong \mathrm{Gal}(L/K).$$

This theorem is very powerful and is one of the most important results of modern number theory. It tells us precisely how primes split in an extension L/K based purely on congruence conditions coming from the arithmetic of K. We can in fact give a rough converse to this result. We may first fix a modulus containing primes we would like to ramify in an extension L/K, this then completely determines the extension.

2.2.4 The Existence Theorem

We let K be a number field. Given a modulus $\mathfrak{m}$, can we find a number field L such that L/K is an abelian extension satisfying the behaviour contained in the Artin reciprocity law? As we have previously stated, the answer to this question is yes.

Theorem 2.2.18 (Existence Theorem)**.** *Let K be a number field and $\mathfrak{m}$ be any modulus of K. Then for each congruence subgroup H of $\mathfrak{m}$, there exists a unique number field L such that L/K is abelian, $\mathfrak{m}$ is divisible by the ramified primes of this extension and the Artin map induces an isomorphism:*

$$I_K(\mathfrak{m})/H \cong \mathrm{Gal}(L/K).$$

This theorem together with the Artin Reciprocity Theorem gives a correspondence between finite abelian extensions of K and generalised ideal class groups, with the choice of modulus corresponding to the choice of ramified primes. Now that we have this correspondence we might be interested in further questions. For example: Given a number field K, what is the maximal unramified abelian extension? What is the maximal abelian extension with given ramification? This leads to the theory of the Hilbert class field and ray class fields.

2.2.5 Ray Class Fields and the Hilbert Class Field

The Existence Theorem tells us that we are free to choose our modulus $\mathfrak{m}$ and also the congruence subgroup producing a number field with given properties. Fix a modulus $\mathfrak{m}$ and choose the congruence subgroup $H = P_{1,K}(\mathfrak{m})$. The Existence Theorem then tells us that there is a unique number field $K_{\mathfrak{m}}$ such that $K_{\mathfrak{m}}/K$ is an abelian extension whose ramified primes are exactly those appearing in $\mathfrak{m}$ and is such that $I_K(\mathfrak{m})/P_{1,K}(\mathfrak{m}) \cong \mathrm{Gal}(K_{\mathfrak{m}}/K)$. In fact, this field is unique (non-obvious).

Definition 2.2.19. The field $K_{\mathfrak{m}}$ is called the *ray class field of K* with respect to $\mathfrak{m}$. The group $C_{\mathfrak{m}} = I_K(\mathfrak{m})/P_{1,K}(\mathfrak{m})$ is called the *ray class group*.

Ray class groups (in particular, characters of this group) will become important in our later work on modular forms of weight 1.

It is clear that any other congruence subgroup H for $\mathfrak{m}$ will correspond to a field lying inside $K_{\mathfrak{m}}$ by Galois theory. Hence these fields really are the maximal abelian extensions with given ramification. We might wonder what the maximal abelian extension is without any ramified primes. This is the Hilbert class field. Recall that the primes appearing in the modulus $\mathfrak{m}$ are precisely those to ramify in the abelian extension. Therefore in order to guarantee that our abelian extension is unramified, we must take $\mathfrak{m}$ to be the trivial modulus $\mathfrak{1}$. The Existence Theorem then guarantees that we can find a field $K_{\mathfrak{1}}$ such that $K_{\mathfrak{1}}/K$ is the maximal unramified abelian extension of K.

Definition 2.2.20. The *Hilbert class field* of a number field K is the field K_1 defined above.

Using the Artin reciprocity law we have the following isomorphism:

$$I_K/P_K \cong \mathrm{Gal}(K_1/K).$$

The fact that the left hand side is now the classical ideal class group is due to the fact that our extra conditions on coprimality and positivity have effectively been dropped by using the trivial modulus. This gives some very nice results.

Lemma 2.2.21. *The Hilbert class field is a degree h_K extension of K. Also a prime ideal $\mathfrak{p} \in \mathcal{O}_K$ splits completely in the Hilbert class field if and only if $\mathfrak{p}$ is principal.*

Corollary 2.2.22. *The following are equivalent:*

(1) $K = K_1$.

(2) $h_K = 1$.

(3) $\mathcal{O}_K$ *is a principal ideal domain (or equivalently a unique factorisation domain).*

One may wonder how to actually construct ray class fields and Hilbert class fields. Unfortunately this is not always an easy task and not much is known beyond some of the more basic cases. For example, as previously mentioned, the Kronecker-Weber Theorem tells us about the maximal abelian extensions of $\mathbb{Q}$. In the case of an imaginary quadratic field $\mathbb{Q}(\sqrt{-d})$, we know how to construct these fields. We begin by constructing an elliptic curve with complex multiplication by the ring of integers of $\mathbb{Q}(\sqrt{-d})$. The j-invariant of this curve is adjoined to $\mathbb{Q}(\sqrt{-d})$ to get the Hilbert class field and we also adjoin certain torsion points related to Weber functions in order to construct the ray class fields. Details of this can be found in chapter 2 of [Sil2]. Beyond these two cases, not much is known.

Chapter 3

Congruences of Local Origin for Weights $k \geq 2$

As we have already mentioned, our aim is to make a generalisation of the congruence given in [DF]. This congruence itself is a generalisation of the famous Ramanujan 691 congruence. Recall that the Ramanujan congruence is one between the Hecke eigenvalues of a level 1 Eisenstein series and a level 1 cusp form with the modulus being a prime dividing $\zeta(12)/\pi^{12}$. The congruence in [DF] is then one between the Hecke eigenvalues of a level 1 Eisenstein series and a level p cusp form with modulus being a prime dividing an "Euler factor", i.e., a partial zeta value. There is then the level N generalisation of Ramanujan's congruence proven by Dummigan [D]. This is a congruence between the Hecke eigenvalues of a level N Eisenstein series and a level N cusp form. Here the modulus is now a prime dividing a certain Dirichlet L-value. We aim to prove the next logical step in these results. We prove the existence of a congruence between the Hecke eigenvalues of a level N Eisenstein series and a level Np cusp form, the modulus being a divisor of an Euler factor of a Dirichlet L-value.

Choose a weight $k \geq 2$, a level $N = uv > 1$ (with u and v coprime), and a Dirichlet character χ of conductor N. Let ψ and φ be primitive Dirichlet characters of conductors u, v respectively, with $\psi\varphi = \chi$, $uv = N$, and $\chi(-1) = (-1)^k$. Then we have $E_k^{\psi,\varphi}$ new at level N.

Theorem 3.0.1. *Let p be a prime with $p \nmid N$ and let $k \geq 2$ be an integer. Let $\lambda' \nmid 6N$ be a prime of the ring of integers of $\mathbb{Q}[\psi, \varphi]$ such that $ord_{\lambda'}\left((\varphi(p)p^k - \psi(p))(B_{k,\psi^{-1}\varphi}/k)\right) > 0$. Then there exists a normalised Hecke eigenform $f \in S_k(\Gamma_1(Np), \chi')$ (where χ' is χ raised to level Np) such that for all n with $n \nmid Np$,*

$$a_n(f) \equiv \sigma_{k-1}^{\psi,\varphi}(n) \pmod{\lambda},$$

where $\lambda|\lambda'$ is a prime of the ring of integers of the extension of $\mathbb{Q}(\psi, \varphi)$ generated by the a_n.

There are a few things to notice immediately about this result. One important feature is that we are including the case of weight 2. At level 1 this was an issue as it only worked in certain cases. A famous theorem of Mazur [M, Prop. 5.12(iii)] states that the congruence holds for some cuspidal eigenform $f \in S_2(\Gamma_0(p))$ if and only if ℓ divides the numerator of $(p-1)/12$. But $p^2-1 = (p-1)(p+1)$, so ℓ can divide p^2-1 (by dividing $(p+1)$) without there being a congruence. Note that the factor of $1/12$ comes from the value $\zeta(-1)$. However if we take $N=1$ and change the weight range to $k \geq 3$ we would have exactly Theorem 1.1 of [DF]. In particular the characters ψ and φ would be trivial giving $p^k - 1$ as the Euler factor. The generalised Bernoulli number $B_{k,\psi^{-1}\varphi}$ would also reduce to B_k.

We might wonder in what way each part of the theorem has been generalised. The first question we might ask would be regarding how the Euler factor has been obtained. Since we are now working with level N Eisenstein series, the moduli of our congruence will come from a prime dividing the value of a particular Dirichlet L-function. The particular L-function is $L(1-k, \psi^{-1}\varphi)$. The reason for this is that this L-value is the constant term of the level N Eisenstein series. Recall that Theorem 1.4.5 gave us the following Fourier expansion:

$$E_k^{\psi,\varphi}(\tau) = \delta(\psi)L(1-k,\varphi) + 2\sum_{n=1}^{\infty}\sigma_{k-1}^{\psi,\varphi}(n)q^n,\ q = e^{2\pi i\tau}.$$

This is in fact equivalent to

$$E_k^{\psi,\varphi}(\tau) = \delta(\psi)L(1-k,\psi^{-1}\varphi) + 2\sum_{n=1}^{\infty}\sigma_{k-1}^{\psi,\varphi}(n)q^n,\ q = e^{2\pi i\tau}.$$

This follows since $\delta(\psi) = 1$ if $\psi = \mathbf{1}_1$, in which case $L(1-k,\varphi) = L(1-k,\psi^{-1}\varphi)$, and $\delta(\psi) = 0$ otherwise, in which case the constant term vanishes. We now make use of the functional equation given in Section 1.4.1 to work with $L(k, \psi\varphi^{-1})$. In the same way as in [DF] we consider the missing Euler factor coming from this L-function. Recall from Section 1.4.1 that this L-function has the following Euler product expansion:

$$L(k,\psi\varphi^{-1}) = \prod_p \left(\frac{1}{1-(\psi\varphi^{-1})(p)p^{-k}}\right).$$

If we fix a particular prime p and pull out the factor $(1-(\psi\varphi^{-1})(p)p^{-k})^{-1}$ and denote this product by $L_{\{p\}}(k,\psi\varphi^{-1})$ then

$$(1-(\psi\varphi^{-1})(p)p^{-k})^{-1}L_{\{p\}}(k,\psi\varphi^{-1}) = L(k,\psi\varphi^{-1}),$$

that is

$$\varphi(p)p^k L_{\{p\}}(k,\psi\varphi^{-1}) = (\varphi(p)p^k - \psi(p))L(k,\psi\varphi^{-1}).$$

We can now use the functional equation to go back to the original L-function appearing as the constant term of the Eisenstein series. Ignoring various small factors and signs, the right hand side becomes

$$(\varphi(p)p^k - \psi(p))L(1-k,\psi^{-1}\varphi).$$

Recall also from Section 1.4.1 that $L(1-k,\psi^{-1}\varphi) = -\frac{B_{k,\psi^{-1}\varphi}}{k}$, where $B_{k,\psi^{-1}\varphi}$ is the generalised Bernoulli number defined by

$$\sum_{a=1}^{N}(\psi^{-1}\varphi)(a)\frac{te^{at}}{e^{Nt}-1} = \sum_{k=0}^{\infty} B_{k,\psi^{-1}\varphi}\frac{t^k}{k!}.$$

Just like before we might expect that dividing $L(1-k,\psi^{-1}\varphi)$ will give a level N congruence, whereas dividing $\varphi(p)p^k - \psi(p)$ should give a level Np congruence. This also addresses the appearance of the generalised Bernoulli numbers. The remaining generalisations in the theorem are the obvious choices for a generalisation.

The method we will use to prove Theorem 3.0.1 will be a generalisation of the proof of Theorem 1.1 of [DF]. We aim to produce a linear combination of Eisenstein series which vanishes modulo a prime divisor of $\varphi(p)p^k - \psi(p)$ at each cusp. This would then be lifted to a cusp form in characteristic 0 and replaced by a Hecke eigenform using the Deligne-Serre Lemma. In the level 1 case this proof was rather straightforward as there were only two cusps to consider and a particular Atkin-Lehner involution could be used to determine the constant term of the Eisenstein series at each cusp. At a general level N, there are more cusps to consider. It is therefore necessary to try and obtain a formula for the constant term of the level N Eisenstein series at any cusp. Once we have this information it will be possible to find the correct combination of Eisenstein series.

A similar result to Theorem 3.0.1 was proven by N. Billerey and R. Menares [BM]. Their work focused on proving that certain reducible Galois representations are modular. As a consequence this in fact proves a weaker version of Theorem 3.0.1 for $k \geq 3$. We state Theorem 2.1 from [BM].

Theorem 3.0.2. *Every odd representation which is the direct sum of two characters arises from a cuspidal eigenform.*

This result is essentially a translation of Theorem 3.0.1 into the language of Galois representations. However in proving this theorem, the level of the cusp form is not guaranteed to be Np, it could be at a higher level. Hence the result proven is weaker than Theorem 3.0.1.

We now look towards proving Theorem 3.0.1. The first step of which is proving a formula for the constant term of the level N Eisenstein series at any cusp.

§ 3.1 A Formula for the Constant Term of the Level N Eisenstein Series

In the process of proving Theorem 3.0.2, Billerey and Menares compute a formula for the constant term of a level N Eisenstein series with one of the two characters being trivial. The reason being that they could use a clever trick on the Galois representation side to consider the case where both characters are non-trivial. This therefore made

the computation of the constant term much simpler. The same method they use can however be applied in the general case. We therefore generalise the method to the case where both characters could be non-trivial. One thing to note however is that they are only working with the case $k \geq 3$. Recall that the weight 2 Eisenstein series only converges conditionally, in other words the order of summation matters. Since the calculation we will use involves working with the sums making up the Eisenstein series, and in particular, manipulating those sums, we will need to consider the case of weight 2 separately.

3.1.1 The Weight $k \geq 3$ Case

The constant term of the Eisenstein series at any given cusp is calculated by determining the action of the slash operator by an arbitrary matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Recall that there is only one equivalence class of cusps under the action of $\mathrm{SL}_2(\mathbb{Z})$. We will therefore be able to move from a given cusp to any other cusp via the slash operator with a matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. The proof of the formula however will require us to work with a matrix γ with top left entry coprime to N. This may sound like we are restricting the generality of the result; this is not the case however. First we state and prove the result; we then explain why this is not a restriction.

Lemma 3.1.1. *Let $N > 1$ and let $\gamma_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. Then we can choose $\gamma_2 = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_1(N)$ such that $\gamma_2\gamma_1$ has top left entry coprime to N.*

Proof. Suppose $N = p_1^{a_1}p_2^{a_2}\dots p_n^{a_n}$. Consider the matrix $\gamma_m = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^m \in \Gamma_1(N)$. Then $\gamma_m\gamma_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^m \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+mc & b+md \\ c & d \end{pmatrix}$. By Dirichlet's Theorem, as m varies, there are infinitely many primes in the arithmetic progression $a+mc$ since a and c are coprime. Since N only contains finitely many primes we may choose m such that $a+mc$ is a prime not appearing in N. Hence the top left entry of $\gamma_m\gamma_1$ is coprime to N. Hence we may take $\gamma_2 = \gamma_m$.

□

Since the Eisenstein series we will be using is a modular form for $\Gamma_1(N)$, it is invariant under the action of matrices $\gamma \in \Gamma_1(N)$. In particular, if we have $\gamma_1 \in \mathrm{SL}_2(\mathbb{Z})$ and $\gamma_2 \in \Gamma_1(N)$, then $f[\gamma_2\gamma_1]_k = (f[\gamma_2]_k)[\gamma_1]_k = f[\gamma_1]_k$ since $f[\gamma_2]_k = f$. Hence there is actually no restriction. We are now in a position to prove a formula for the constant term. The proof we are adapting is that of Proposition 1.2 of [BM]. In the following we use the notation $\overline{\varphi}$ instead of φ^{-1} in order to keep the notation cleaner. From now on we may use either interchangeably.

Proposition 3.1.2. *Let $N = uv$ with u and v coprime, and let $k \geq 2$ be an integer. Let ψ and φ be primitive Dirichlet characters of conductors u, v respectively. Let $M \geq 1$ be an integer coprime to N and $\gamma = \begin{pmatrix} a & \beta \\ b & \delta \end{pmatrix} \in SL_2(\mathbb{Z})$. Let α_M be the operator acting on complex-valued functions on the upper half plane $\mathcal{H}$ by $\alpha_M(f)(z) = f(Mz)$. Then the constant term of the q-expansion of $(\alpha_M E_k^{\psi,\varphi})[\gamma]_k$ is*

$$\begin{cases} 0 & \text{if } v \nmid b', \\ -\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})} \frac{\overline{\varphi}(M'a)\psi(\frac{-b'}{v})}{u^k M'^k} \frac{B_{k,\overline{\psi}\varphi}}{k} & \text{otherwise,} \end{cases}$$

where $b' = \frac{b}{gcd(b,M)}$ and $M' = \frac{M}{gcd(b,M)}$.

Proof. Recall from Section 1.4, we have

$$G_k^{\psi,\varphi} = \frac{C_k g(\overline{\varphi})}{v^k} E_k^{\psi,\varphi},$$

where

$$C_k = \frac{(-2\pi i)^k}{(k-1)!} \quad \text{and} \quad g(\overline{\varphi}) = \sum_{n=0}^{v-1} \overline{\varphi}(n) e^{\frac{2\pi i n}{v}}.$$

We also have

$$G_k^{\psi,\varphi}(\tau) = \sum_{c=0}^{u-1} \sum_{d=0}^{v-1} \sum_{e=0}^{u-1} \psi(c)\overline{\varphi}(d) G_k^{\overline{(cv,d+ev)}}(\tau),$$

where

$$G_k^{\overline{(cv,d+ev)}}(\tau) = \sum_{\substack{(f,g)\in\mathbb{Z}^2\setminus\{(0,0)\} \\ (f,g)\equiv(cv,d+ev) \pmod N}} \frac{1}{(f\tau+g)^k}.$$

Then

$$(\alpha_M(G_k^{\overline{(cv,d+ev)}}))[\gamma]_k(\tau) = \sum_{\substack{(f,g)\in\mathbb{Z}^2\setminus\{(0,0)\} \\ (f,g)\equiv(cv,d+ev) \pmod N}} \frac{1}{((fMa+gb)\tau + fM\beta + g\delta)^k}.$$

Hence the constant term of $(\alpha_M(G_k^{\overline{(cv,d+ev)}}))[\gamma]_k$ is given by

$$\Upsilon_{c,d,e} = \sum_{\substack{(f,g)\in\mathbb{Z}^2\setminus\{(0,0)\} \\ (f,g)\equiv(cv,d+ev) \pmod N \\ fMa+gb=0}} \frac{1}{(fM\beta+g\delta)^k}.$$

We first consider the case $a = 0$. If $a = 0$, then for $fMa + gb$ to be 0 we must have $g = 0$ as b must be non-zero (since $\gamma \in SL_2(\mathbb{Z})$). This means $d + ev \equiv 0 \pmod N$, or

in other words $\Upsilon_{c,d,e} = 0$ unless $d + ev \equiv 0 \pmod{N}$. But

$$\begin{aligned} d + ev \equiv 0 \pmod{N} &\Rightarrow d + ev \equiv 0 \pmod{v} \\ &\Rightarrow d \equiv 0 \pmod{v} \\ &\Rightarrow d = 0 \quad (\text{since} \quad 0 \leq d \leq v-1) \\ &\Rightarrow ev \equiv 0 \pmod{N} \\ &\Rightarrow e = 0 \quad (\text{since} \quad 0 \leq e \leq u-1). \end{aligned}$$

Hence $\Upsilon_{c,d,e} = 0$ unless $d = e = 0$. But now we must have $c = 0$ and $N = 1$ otherwise the contribution to the constant term of $G_k^{\psi,\varphi}$ is 0. Hence

$$\begin{aligned} \Upsilon_{0,0,0} = \sum_{\substack{f \in \mathbb{Z}\setminus\{0\} \\ f \equiv 0 \pmod{N}}} \frac{1}{(fM\beta)^k} &= \frac{1}{(M\beta)^k} \sum_{\substack{f \in \mathbb{Z}\setminus\{0\} \\ f \equiv 0 \pmod{1}}} \frac{1}{f^k} \\ &= \frac{1}{(M\beta)^k} \sum_{t \in \mathbb{Z}\setminus\{0\}} \frac{1}{t^k} \\ &= \frac{2}{M^k}\zeta(k), \end{aligned}$$

since $\beta = \pm 1$ (this follows since $a = 0$) and k is even (Here $N = 1$). Therefore the constant term Υ of $(\alpha_M G_k^{\psi,\varphi})[\gamma]_k$ is 0 if $N > 1$ and is $\frac{-C_k}{M^k}\frac{B_k}{k}$ when $N = 1$.

Now we consider the case $a \neq 0$. Given $g \equiv d + ev \pmod{N}, g \in \mathbb{Z}, d \neq 0$ the following conditions are equivalent:

$$\text{there exists } f \in \mathbb{Z}, f \equiv cv \pmod{N} \text{ such that } fMa + gb = 0; \tag{3.1}$$

$$M'a | g, v | b' \text{ and } u | cvM'a + gb', \tag{3.2}$$

where $M' = \frac{M}{\gcd(b,M)}$ and $b' = \frac{b}{\gcd(b,M)}$.

If the first condition holds, then $fMa + gb = 0$. But $fMa + gb = 0 \Leftrightarrow fM'a + gb' = 0$. Hence $M'a | gb'$ but $M'a$ and b' are coprime ($(M', b') = 1$ and $\gamma \in SL_2(\mathbb{Z})$), so $M'a | g$. Now suppose $f = cv + rN$ and $g = d + ev + sN$. Then we have $(cv + rN)M'a + (d + ev + sN)b' = 0$ and rearranging gives $(cvM'a + db' + evb') + (rM'a + sb')N = 0$. For this to be 0 we must have $cvM'a + db' + evb' \equiv 0 \pmod{N}$. In particular $db' \equiv 0 \pmod{v}$ and so either $v | b'$ or $d = 0$. But we assumed $d \neq 0$, so $v | b'$. Also we have $cvM'a + db' + evb' \equiv 0 \pmod{u}$, i.e., $u | cvM'a + gb'$.

On the other hand, if the second condition holds, put $f = \frac{-gb'}{M'a} \in \mathbb{Z}$. Then $f \in \mathbb{Z}$ satisfies $fM'a + gb' = 0$ and further $u | cvM'a + gb' \Rightarrow cv \equiv \frac{-gb'}{M'a} \pmod{u}$, i.e., $f \equiv cv \pmod{u}$. Here we have used the assumption that $(a, N) = 1$, which we can do due to Lemma 3.1.1. Also $f \equiv 0 \pmod{v}$, since $v | b'$, so $f \equiv cv \pmod{N}$.

If these equivalent conditions are satisfied, then we have

$$fM\beta + g\delta = \frac{1}{a}(fMa\beta + ga\delta) = \frac{1}{a}(ga\delta - gb\beta) = \frac{g}{a}.$$

Therefore the constant term Υ of $(\alpha_M G_k^{\psi,\varphi})[\gamma]_k$ is 0 when $v \nmid b'$ and is otherwise given by

$$\Upsilon = \sum_{d=0}^{v-1}\sum_{e=0}^{u-1}\overline{\varphi}(d) \sum_{\substack{g\equiv d+ev \pmod N \\ g\neq 0 \\ M'a|g}} \left(\frac{a}{g}\right)^k \sum_{\substack{c=0 \\ u|cvM'a+gb'}}^{u-1} \psi(c).$$

Let $g = M'at$. Then

$$\begin{aligned} M'a(cv+tb') \equiv 0 \pmod u &\Rightarrow cv+tb' \equiv 0 \pmod u \\ &\Rightarrow c \equiv \frac{-tb'}{v} \pmod u. \end{aligned}$$

Note that here we use the fact that u and v are coprime.

Hence

$$\begin{aligned} \Upsilon &= \sum_{d=0}^{v-1}\sum_{e=0}^{u-1}\overline{\varphi}(d) \sum_{\substack{t\equiv\frac{d+ev}{M'a} \pmod N \\ t\neq 0}} \left(\frac{1}{M't}\right)^k \sum_{\substack{c=0 \\ c\equiv\frac{-tb'}{v} \pmod u}}^{u-1} \psi(c) \\ &= \frac{\overline{\varphi}(M'a)}{M'^k}\sum_{d=0}^{v-1}\sum_{e=0}^{u-1} \sum_{\substack{t\equiv\frac{d+ev}{M'a} \pmod N \\ t\neq 0}} \frac{\overline{\varphi}(t)}{t^k} \sum_{\substack{c=0 \\ c\equiv\frac{-tb'}{v} \pmod u}}^{u-1} \psi(c) \\ &= \frac{\overline{\varphi}(M'a)\psi(\frac{-b'}{v})}{M'^k}\sum_{d=0}^{v-1}\sum_{e=0}^{u-1} \sum_{\substack{t\equiv\frac{d+ev}{M'a} \pmod N \\ t\neq 0}} \frac{\overline{\varphi}(t)\psi(t)}{t^k} \\ &= \frac{\overline{\varphi}(M'a)\psi(\frac{-b'}{v})}{M'^k} \sum_{t\in\mathbb{Z}\setminus\{0\}} \frac{\overline{\varphi}(t)\psi(t)}{t^k} \\ &= 2\frac{\overline{\varphi}(M'a)\psi(\frac{-b'}{v})}{M'^k} L(k,\psi\overline{\varphi}). \end{aligned}$$

We may now use the functional equation to get the desired form of the constant term. □

We must now consider the case of weight 2 separately.

3.1.2 The Weight 2 Case

The proof of Proposition 3.1.2 makes it clear that the conditional convergence of the weight 2 Eisenstein series will mean that this case needs careful consideration. In fact it requires quite a bit of analysis in order to prove, but luckily enough, the same formula still holds. Another paper by Billerey and Menares [BM2], submitted a couple of

months after the completion of the proof of Proposition 3.1.2, actually proves the same formula (See Proposition 4). They however prove the result in a similar but different way. They essentially run through mostly the same steps as in [BM, Proposition 1.2] but they combine this with the method used in [BD, Proposition 2.8]. In doing so they produced a unified and (slightly) simplified proof. In proving the formula however, they also consider the weight 2 case. We now describe the necessary results needed in order for Proposition 3.1.2 to hold in the case of weight 2. The following is due to Billerey and Menares and is not original material. We state the results without proof.

We first set some notation. For $\epsilon \geq 0$, we let

$$w^{2,\epsilon} = w^2|w|^{2\epsilon}, \quad w \in \mathbb{C}.$$

Let $y > 0$ be a positive real number. The notation $g_1 \ll_{y_0} g_2$ means that there exists a positive constant C, depending only on y_0, such that $|g_1(r)| \leq C|g_2(r)|$ for all r in the common domain of g_1, g_2.

Let

$$S_\epsilon(z) = \sum_{n\in\mathbb{Z}} \frac{1}{(z+n)^{2,\epsilon}}, \quad z \in \mathbb{C}\backslash\mathbb{R}.$$

We now state a series of lemmas which are required in order to prove that the constant term formula still holds in the case of weight 2.

Lemma 3.1.3. *Fix $y_0 > 0$. Then, we have that*

$$S_\epsilon(z) \ll_{y_0} \frac{1}{\Gamma(\epsilon)|y|^{1+2\epsilon}} + e^{-2\pi|y|}, \quad y = \mathrm{Im}(z), \quad |y| \geq y_0, \quad 0 < \epsilon \leq 1,$$

where for any real number $s > 0, \Gamma(s) = \int_0^\infty e^{-t}t^{s-1}\mathrm{d}t$.

Lemma 3.1.4. *For any $a_1, a_2, D \in \mathbb{Z}$ with $D \neq 0$, set*

$$\sigma_\epsilon(z; a_1, a_2, D) = \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ a_1+Dm\neq 0}} \frac{1}{(z(a_1+Dm)+a_2+Dn)^{2,\epsilon}}.$$

Then we have that

$$\lim_{\mathrm{Im}(z)\to\infty} \lim_{\epsilon\to 0^+} \sigma_\epsilon(z; a_1, a_2, D) = 0.$$

Lemma 3.1.5. *Proposition 3.1.2 holds in the case of weight 2.*

For proofs of each of these results see pages 10-12 of [BM2]. Now that we have a formula for the constant term of the level N Eisenstein series at any cusp, we are in a position to construct a linear combination of Eisenstein series vanishing modulo a prime divisor of the Euler factor at each cusp.

§ 3.2 Proving the Main Theorem

Our aim was to try and generalise the proof used in [DF]. Now that we have a formula for the constant term of the Eisenstein series at the cusps, we can try to find a particular linear combination of Eisenstein series that vanishes modulo λ' at each cusp. In [DF] the combination used was $E_k(z) - E_k(pz)$ where $E_k(z)$ is the level 1 Eisenstein series. Recall that $E_k(pz) = \alpha_p E_k(z)$. We would therefore hope that something similar would work in the general case of level N. It isn't exactly this but we only need to twist by a character. The following lemma uses the same notation as Theorem 3.0.1 and Proposition 3.1.2.

Lemma 3.2.1. *Suppose λ' is a prime dividing $\frac{B_{k,\overline{\psi}\varphi}}{k}\left(\varphi(p)p^k - \psi(p)\right)$. Here p is a prime with $(p, N) = 1$. Then the linear combination of Eisenstein series*

$$E = E_k^{\psi,\varphi} - \psi(p)\alpha_p E_k^{\psi,\varphi}$$

vanishes modulo λ' at each cusp.

Proof. We compute the constant term, say Υ, of the q-expansion of $E[\gamma]_k$ for $\gamma = \begin{bmatrix} a & \beta \\ b & \delta \end{bmatrix} \in SL_2(\mathbb{Z})$. By Proposition 3.1.2 we have,

$$\Upsilon = \begin{cases} -\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{v})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}\left(1 - \frac{\psi(p)\overline{\varphi}(p)}{p^k}\right) & \text{if } v|b' \text{ and } p \nmid b, \\ 0 & \text{otherwise.} \end{cases}$$

To see this, we consider various cases. Firstly, if $v \nmid b'$ then the constant term is 0. Now suppose $v|b'$ and $p|b$. For the part of the constant term of E contributed by $E_k^{\psi,\varphi}$ we have $M = 1$ and therefore $M' = 1$ and $b' = b$, so the constant term is $-\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{v})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}$. For the part contributed by $-\psi(p)\alpha_p E_k^{\psi,\varphi}$ we have $M = p$ and therefore $M' = 1$ and $b' = \frac{b}{p}$, so the constant term is $\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{pv})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}\psi(p) = \frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{v})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}\frac{\psi(p)}{\psi(p)}$. Hence the constant term of E in this case is 0. Now suppose $v|b'$ and $p \nmid b$. For the part of the constant term of E contributed by $E_k^{\psi,\varphi}$ we have $M = 1$ and therefore $M' = 1$ and $b' = b$, so the constant term is $-\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{v})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}$. For the part contributed by $-\psi(p)\alpha_p E_k^{\psi,\varphi}$ we have $M = p$ and therefore $M' = p$ and $b' = b$ (since $p \nmid b$), so the constant term is $\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{v})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}\frac{\psi(p)\overline{\varphi}(p)}{p^k}$. Hence the constant term of E is $-\frac{g(\psi\overline{\varphi})}{g(\overline{\varphi})}\frac{\overline{\varphi}(a)\psi(\frac{-b}{v})}{u^k}\frac{B_{k,\overline{\psi}\varphi}}{k}\left(1 - \frac{\psi(p)\overline{\varphi}(p)}{p^k}\right)$ in this case.

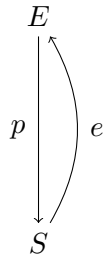
Therefore, under the assumption that λ' divides $\frac{B_{k,\overline{\psi}\varphi}}{k}\left(\varphi(p)p^k - \psi(p)\right)$, E vanishes modulo λ' at each cusp. □

We have now done most of the hard work in proving Theorem 3.0.1. The rest of the proof relies on some well known results which we now explain.

First of all, we will need to introduce the notion of Katz modular forms. Katz modular forms are a geometric formulation of modular forms whose properties agree with those of the classical modular forms. Rather than view a modular form as a holomorphic function on the upper half plane satisfying certain transformation properties, Katz instead considered them as a particular type of rule associated to elliptic curves and their invariant differentials. He also went a little further and considered them purely in terms of modular varieties. This is beyond the scope of what we need here. For precise details of Katz modular forms and some of the basic properties see [Ka, §1] or [E, §1]. Here we will just give a broad overview following the outline of [Ka] covering only the material required for the proof of Theorem 3.0.1.

The notion of an elliptic curve defined over a field can be extended to the notion of an elliptic curve defined over a commutative ring. In order to do this we first give the definition of an elliptic curve over a scheme. Although we won't need the full generality of the definition we include it for completeness.

Definition 3.2.2. By an elliptic curve over a scheme S, we mean a proper smooth morphism $p : E \to S$, whose geometric fibres are connected curves of genus one, together with a section $e : S \to E$.



We denote by $\underline{\omega}_{E/S}$ the invertible sheaf $p_*(\Omega^1_{E/S})$ on S, which is canonically dual (Serre duality) to the invertible sheaf $R^1p_*(\mathcal{O}_E)$ on S.

Most of this definition is outside the scope of what we need for the purposes of this thesis. However we note that if we take $S = \mathrm{Spec}(R)$, then this gives us the notion of an elliptic curve defined over a commutative ring R. Here $\mathrm{Spec}(R)$ is the spectrum of the ring R. The spectrum of R is the set of prime ideals of R. For more details of this construction, see [Ka].

The basic idea of Katz modular forms is that for an elliptic curve E over a commutative ring R we can consider a modular form f to be a particular kind of rule on the pair $(E/R, \omega)$ where ω is an "invariant differential". We use inverted commas here since this is slightly looser language than that used by Katz. Katz defined a modular form of weight k and level 1 to be a rule f which assigns to the pair $(E/R, \omega)$ an element $f(E/R, \omega) \in R$, such that the following conditions hold

(1) $f(E/R, \omega)$ depends only on the R-isomorphism class of the pair $(E/R, \omega)$,

(2) f is homogeneous of degree $-k$ in the second variable; for any $\lambda \in R^\times$, $f(E, \lambda\omega) = \lambda^{-k} f(E, \omega)$,

(3) The formation of $f(E/R, \omega)$ commutes with arbitrary extension of scalars $g : R \to R'$ (meaning $f(E_{R'}/R', \omega_{R'}) = g(f(E/R, \omega))$).

If we consider a ring R lying over a fixed ground ring R_0 in the previous definition, and only base changes by R_0-morphisms, this gives the notion of a modular form of weight k and level one defined over R_0, the R_0-module of which is denoted $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}), R_0)$. Notice how there is no condition on holomorphy of these forms, hence the unusual notation. We also note that Katz uses slightly different notation but we have opted to stick with notation that resembles that used in this thesis.

Since we have no condition on holomorphy yet it would be nice to try and incorporate this. In order to do this in this new setup we have to consider the q-expansion. This involves using the Tate curve. The Tate curve is a projective plane curve defined over the ring $\mathbb{Z}[[q]]$ of formal power series. Note that although the Tate curve is defined over $\mathbb{Z}[[q]]$, it is only an elliptic curve over $\mathbb{Z}((q))$. For more details see [Sil2, Chapter V, §3]. A modular form f of weight k defined over R_0 can then be evaluated on the pair $(\mathrm{Tate}(q), \omega_{\mathrm{can}})_{R_0}$ consisting of the Tate curve and its canonical differential, viewed as an elliptic curve with differential over $\mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$ (and not just over $R_0((q))$). The q-expansion of the modular form f is then given by the finite-tailed Laurent series

$$f((\mathrm{Tate}(q), \omega_{\mathrm{can}})_{R_0}) \in \mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0.$$

The modular form f is said to be holomorphic at ∞ if its q-expansion lies in the subring $\mathbb{Z}[[q]] \otimes_{\mathbb{Z}} R_0$; we now denote these modular forms with the familiar notation $M_k(\mathrm{SL}_2(\mathbb{Z}), R_0)$. The submodule of cusp forms $S_k(\mathrm{SL}_2(\mathbb{Z}), R_0)$ comprises of those modular forms whose q-expansion has constant term zero.

As we have previously mentioned, this notion of a Katz modular form in fact coincides with the notion of a classical modular form when we take $R = \mathbb{C}$. Although the language is more technical here, hopefully it is clear that the notions of q-expansion, holomorphicity and the transformation property all reduce to the same thing on both sides.

Similar definitions hold in the case of level N except now it involves a level N structure on the elliptic curve E. Although Katz uses a more general notion of level N structure, we will only consider the case which gives Katz modular forms for $\Gamma_1(N)$ since these are the forms we will consider in our proof. For details of the more general case, see [Ka, §1.2]. For our purposes, the level N structure will consist of an isomorphism $\varphi : \mathbb{Z}/N\mathbb{Z} \xrightarrow{\sim} C$ where C is a cyclic subgroup of $E[N]$. This is determined by $\varphi(1) = P$ where P is a point of order N. In this case we will obtain the R_0-module $\mathcal{M}_k(\Gamma_1(N), R_0)$. Recall these are the modular forms without holomorphicity. To consider the q-expansion in the level N case we have some extra conditions on the ring

R_0. In particular we require R_0 to contain $\frac{1}{N}$ and a primitive N-th root of unity ζ_N. We may then evaluate the modular form f on triples $(\mathrm{Tate}(q^N), \omega_{\mathrm{can}}, P)_{R_0}$ where P is a point of order N. The q-expansions of the modular form f are then the finitely many finite-tailed Laurent series

$$f((\mathrm{Tate}(q^N), \omega_{\mathrm{can}}, P)_{R_0}) \in \mathbb{Z}((q)) \otimes_{\mathbb{Z}} R_0$$

obtained by varying P over all points of order N. A modular form f defined over a ring R_0 containing $\frac{1}{N}$ and a primitive N-th root of unity ζ_N is said to be holomorphic at ∞ if all of the q-expansions lie in $\mathbb{Z}[[q]] \otimes_{\mathbb{Z}} R_0$. This gives the R_0-module $M_k(\Gamma_1(N), R_0)$. Again, the submodule of cusp forms is denoted $S_k(\Gamma_1(N), R_0)$. Note that each of these q-expansions corresponds to an expansion around a different cusp and a cusp form is one such that the constant term of each of these expansions is zero. In particular since we are considering all level N structures here (corresponding to $\Gamma_1(N)$) we have q-expansions at all cusps. In the case when $R = \mathbb{C}$ this will correspond to the q-expansions at each of the cusps of $X_1(N)$.

Our main proof will involve viewing classical modular forms as Katz modular forms. We wish to consider our Eisenstein series E as a Katz modular form whose reduction modulo λ' lies in the module $S_k(\Gamma_1(Np), \overline{\mathbb{F}}_\ell)$. In order to do this however, we will need to make use of the q-expansion principle. The principle was first introduced by Katz in 1972 and was later published in 1973. Corollaries 1.6.2 and 1.12.2 of [Ka] give the precise statements. The principle basically states that a modular form f is a Katz modular form over a ring R as long as its q-expansion at sufficiently many cusps has coefficients in the ring R. Clearly in our case this will mean that we have a Katz modular form with coefficients in $\mathbb{F}_{\lambda'}$, which we can then view as lying inside $\overline{\mathbb{F}}_\ell$; see the proof of Theorem 3.0.1 for more details.

Now that we have defined Katz modular forms we might wonder whether there is still a notion of Hecke operators acting on the spaces $M_k(\Gamma_1(Np), R)$. In the classical case we had double coset operators acting on holomorphic functions. Here our modular forms are no longer holomorphic functions. However, it turns out that we can still define Hecke operators acting on the spaces of Katz modular forms. These Hecke operators are defined in such a way that when we take $R = \mathbb{C}$, they agree with the Hecke operators in the classical case. For details see [Ka, §1.11].

We have now covered sufficient background material on Katz modular forms necessary to prove Theorem 3.0.1.

Proof of Theorem 3.0.1. Recall that we have the linear combination of Eisenstein series $E = E_k^{\psi,\varphi}(z) - \psi(p)\alpha_p E_k^{\psi,\varphi}(z) = E_k^{\psi,\varphi}(z) - \psi(p)E_k^{\psi,\varphi}(pz)$. We know that E is a classical modular form. We can view this as a Katz modular form in the module $M_k(\Gamma_1(Np), \mathbb{C})$.

Consider the $\mathbb{Z}[1/N]$-algebra R, with $R = \mathbb{Z}[\zeta_{Np}, \psi, \varphi]$ where ζ_{Np} is a primitive Np-th root of unity. Then, by an application of the q-expansion principle, we have $E \in M_k(\Gamma_1(Np), R)$. This follows since the coefficients of the q-expansion of E at each cusp lie in R.

By Lemma 3.2.1 we know that the holomorphic function E vanishes modulo λ' at each of the cusps (viewed as a classical modular form). It follows by the properties of Katz modular forms that this still holds when we view $E \in M_k(\Gamma_1(Np), R)$. That is, each of the q-expansions at different cusps has constant term divisible by λ'. It follows from the q-expansion principle that the reduction of E gives rise to an element $\overline{E} \in S_k(\Gamma_1(Np), \mathbb{F}_{\lambda'})$. Note that this can be viewed as a base change by a homomorphism from R to $\mathbb{F}_{\lambda'}$. The element $\overline{E}$ is a common eigenvector for each of the Hecke operators T_q for $q \nmid Np$. Note that we now have Hecke operators acting on a space of Katz modular forms. These Hecke operators behave in a way compatible with the classical Hecke operators acting on E viewed as a classical modular form. Since $\mathbb{F}_{\lambda'} \subset \overline{\mathbb{F}}_\ell$ (since $\lambda'|\ell$), we can view $\overline{E} \in S_k(\Gamma_1(Np), \overline{\mathbb{F}}_\ell)$.

We know by [E, Lemma 1.9] that the reduction map from $S_k(\Gamma_1(Np), \overline{\mathbb{Z}}_\ell)$ to $S_k(\Gamma_1(Np), \overline{\mathbb{F}}_\ell)$, where $\lambda'|\ell$, is surjective. Hence $\overline{E}$ is the reduction of some element $g \in S_k(\Gamma_1(Np), \mathcal{O}_{\lambda''})$, with $\mathcal{O}_{\lambda''}$ the ring of integers of some finite extension $K_{\lambda''}$ of $\mathbb{Q}_\ell$. That is $\overline{E}$ is the reduction of a characteristic 0 cusp form. This cusp form however may not be an eigenvector for each of the Hecke operators. Let $\mathbb{F}$ denote the residue field of $\mathcal{O}_{\lambda''}$. Each of the Hecke operators T_q for $q \nmid Np$ then commute and act on $S_k(\Gamma_1(Np), \mathbb{F})$ with $\overline{E}$ a common eigenvector, eigenvalue $\psi(q) + \varphi(q)q^{k-1}$ for T_q. By the Deligne-Serre lifting lemma [DeSe, Lemme 6.11], there exists a common eigenvector $f' \in S_k(\Gamma_1(Np), \mathcal{O}_\lambda)$ with $\mathcal{O}_\lambda$ the ring of integers in some finite extension K_λ of $K_{\lambda''}$, with eigenvalues congruent to $\psi(q) + \varphi(q)q^{k-1} \pmod{\lambda}$, where $\lambda|\lambda'$. As a consequence of Carayol's lemma (Proposition 1.10 of [E]) we see that f' arises from an $f \in S_k(\Gamma_1(Np), \chi')$ (rather than $f \in S_k(\Gamma_1(Np), \tilde{\chi})$ with $\tilde{\chi} \equiv \chi' \pmod{\lambda}$) with χ' as in the Theorem. This f is an eigenform for the Hecke operators T_q when $q \nmid Np$ with corresponding eigenvalues a_q satisfying

$$a_q \equiv \psi(q) + \varphi(q)q^{k-1} \pmod{\lambda},$$

for all $q \nmid Np$.

□

Remark 3.2.3. There are many ways that this work could be further generalised. For example we could try raising the level by a power of a prime, or a product of primes. Another way that we could generalise is by changing the modular forms that we are working with. Although it will not be covered in this thesis, it is worth noting that I am currently working with Daniel Fretwell and Catherine Hsu to try and generalise this work to the case of Hilbert modular forms. In the process of this work it has been brought to my attention that a constant term formula in the Hilbert case (see [O]) is proven without the condition that u and v be coprime. It should be possible to drop this condition also in my case, however I have not had the time to try and prove this.

§ 3.3 Comparison with the Bloch-Kato Formula for a Partial L-Value

The Bloch-Kato conjecture is a far reaching generalisation of the Birch and Swinnerton-Dyer conjecture for elliptic curves. Whereas the Birch and Swinnerton-Dyer conjecture relates arithmetic information associated to an elliptic curve and the behaviour of the Hasse-Weil L-function at $s = 1$, the Bloch-Kato conjecture goes further and gives us information about values at integer points of L-functions associated to motives. This thesis will not deal with the general conjecture; in particular we will not cover material on motives.

The aim of this section is to make use of the congruence we have found in Theorem 3.0.1 and construct a non-zero element in a Bloch-Kato Selmer group. This then gives the divisibility of the partial L-value by the modulus of the congruence. Although the case of the Bloch-Kato conjecture that we will be considering has already been proven, we can show that what we obtain agrees with the result. Naturally we would also like to go the other way. That is, given a non-zero element of a Bloch-Kato Selmer group, we would like divisibility of the partial L-value to give the existence of a congruence. In the case of the L-function being a Dirichlet L-function, this is exactly what we have proved in Theorem 3.0.1. Recall from the discussion after the statement of Theorem 3.0.1, that we interpreted $\mathrm{ord}_\lambda(\varphi(p)p^k - \psi(p)) > 0$ as there being a λ dividing the missing Euler factor at p of the Dirichlet L-function $L(1-k, \psi^{-1}\varphi)$. This is exactly the L-function that will appear in the Bloch-Kato conjecture. This condition implies the existence of a non-zero element in the λ part of the Bloch-Kato Selmer group. In more general cases however, divisibility of an L-function by λ need not imply the existence of a mod λ congruence.

Let k, p, ℓ, λ and $f = \sum_{n=1}^{\infty} a_n q^n \in S_k(\Gamma_1(Np), \chi)$ be as in Theorem 3.0.1 on page 48. Suppose that $p \neq \ell$ and let $L = \mathbb{Q}(\{a_n\})$. There exists a continuous representation attached to f given by:

$$\rho_f = \rho_{f,\lambda} : G_{\mathbb{Q}} \to \mathrm{GL}_2(L_\lambda),$$

unramified outside $Np\ell$, such that if $q \nmid Np\ell$ is a prime, and Frob_q is an arithmetic Frobenius element, then

$$\mathrm{Tr}(\rho_f(\mathrm{Frob}_q^{-1})) = a_q(f), \quad \det(\rho_f(\mathrm{Frob}_q^{-1})) = \chi(q)q^{k-1}.$$

We may conjugate so that ρ_f takes values in $\mathrm{GL}_2(\mathcal{O}_\lambda)$ and reduce modulo λ to get a continuous representation

$$\overline{\rho}_f = \overline{\rho}_{f,\lambda} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_\lambda).$$

This depends in general on a choice of invariant $\mathcal{O}_\lambda$-lattice. However the irreducible composition factors are well-defined.

In order to determine these composition factors we will need a couple of results. The first is the Cebotarev density theorem.

Theorem 3.3.1 (Cebotarev density theorem)**.** *The (lifts of) Frobenius elements are dense in the absolute Galois group* $G_{\mathbb{Q}}$*.*

What this result is really telling us is that we can completely determine a representation by considering the image of all the Frobenius elements. In fact, we don't actually need all of them; we may remove finitely many Frobenius elements and we can still determine the representation. This is useful since we only consider Frobenius elements Frob_q such that $q \nmid Np\ell$ here.

The second result that we require is the Brauer-Nesbitt theorem.

Theorem 3.3.2 (Brauer-Nesbitt)**.** *Let* ρ_1, ρ_2 *be semisimple* n*-dimensional Galois representations over a field* K *of characteristic* 0 *or* $\ell > n$*. Then* $\rho_1 \sim \rho_2$ *if and only if* $\mathrm{Tr}(\rho_1) = \mathrm{Tr}(\rho_2)$*.*

This theorem is essentially saying that we can determine when two Galois representations are equivalent purely by considering the character values. We note that the full theorem is a little stronger as the case $\ell \leq n$ is also considered but we do not require the full generality as we always have $\ell > n$.

Since we know the character values of $\overline{\rho}_f$ at all but finitely many Frobenius elements, we may combine the Cebotarev density theorem with the Brauer-Nesbitt theorem in order to fully determine the irreducible composition factors. Since $\mathrm{Tr}(\overline{\rho}_f(\mathrm{Frob}_q^{-1})) = \overline{a}_q = \psi(q)+\varphi(q)q^{k-1}$ in $\mathbb{F}_\lambda$ we see that the composition factors are the one-dimensional modules $\mathbb{F}_\lambda(\psi)$ and $\mathbb{F}_\lambda(1-k)(\varphi)$. Note that $\mathbb{F}_\lambda(\psi)$ is simply ψ viewed as a Galois representation (via Class Field Theory) on an $\mathbb{F}_\lambda$ vector space. Similarly $\mathbb{F}_\lambda(1-k)(\varphi)$ is the $(1-k)$-th Tate twist of φ viewed as a Galois representation on an $\mathbb{F}_\lambda$ vector space. The Tate twist here is simply multiplying φ by χ_ℓ^{1-k} where χ_ℓ is the ℓ-adic cyclotomic character.

Without loss of generality we may choose our invariant $\mathcal{O}_\lambda$-lattice such that

$$\overline{\rho}_f \sim \begin{bmatrix} \varphi\chi_\ell^{1-k} & * \\ 0 & \psi \end{bmatrix}.$$

Moreover an argument of Ribet [Rib3, Proposition 2.1] says that we may also choose our invariant $\mathcal{O}_\lambda$-lattice in such a way that $\overline{\rho}_f$ is realised on a space V such that

$$0 \longrightarrow \mathbb{F}_\lambda(1-k)(\varphi) \xrightarrow{\iota} V \xrightarrow{\pi} \mathbb{F}_\lambda(\psi) \longrightarrow 0$$

is a non-split extension of $\mathbb{F}_\lambda[G_{\mathbb{Q}}]$-modules. Choose a map $s : \mathbb{F}_\lambda(\psi) \to V$ that is $\mathbb{F}_\lambda$ linear and $x \in \mathbb{F}_\lambda(\psi)$. For $g \in G_{\mathbb{Q}}$ consider $g(s(g^{-1}(x))) - s(x)$. Note that

$$\pi(g(s(g^{-1}(x))) - s(x)) = g(\pi(s(g^{-1}(x)))) - \pi(s(x)) = g(g^{-1}(x)) - x = x - x = 0.$$

Hence $g(s(g^{-1}(x))) - s(x) \in \ker(\pi)$. Since the sequence is exact we see that $g(s(g^{-1}(x))) - s(x) \in \mathrm{im}(\iota)$. We therefore have a map $C : G_{\mathbb{Q}} \to \mathrm{Hom}(\mathbb{F}_\lambda(\psi), \mathbb{F}_\lambda(1-$

$k)(\varphi))$ defined by $C(g)(x) := \iota^{-1}(g(s(g^{-1}(x))) - s(x))$. Note that g acts on x by $g(x) = \psi(g)x$ and the action of g on $C(h)$ in $\mathrm{Hom}(\mathbb{F}_\lambda(\psi), \mathbb{F}_\lambda(1-k)(\varphi))$ is by g on the codomain and by g^{-1} on the domain. We also note that $C(g)(x) = \psi(g)^{-1}g(s(x)) - s(x) = g(s(g^{-1}(x))) - s(x)$. We omit the ι^{-1} to ease notation. We have

$$\begin{aligned} C(g)(x) + [gC(h)](x) &= g(s(g^{-1}(x))) - s(x) \\ &\quad + g\left[h(s(h^{-1}(g^{-1}(x)))) - s(g^{-1}(x))\right] \\ &= g(s(g^{-1}(x))) - s(x) \\ &\quad + gh(s(h^{-1}g^{-1}(x))) - g(s(g^{-1}(x))) \\ &= gh(s((gh)^{-1}(x))) - s(x) = C(gh)(x). \end{aligned}$$

We therefore see that C defines a cocycle as it satisfies the necessary condition.

We note that choosing a different map s would result in a different cocycle, but this would only differ by a coboundary. We therefore obtain a unique class $c := [C] \in H^1(G_\mathbb{Q}, \mathrm{Hom}(\mathbb{F}_\lambda(\psi), \mathbb{F}_\lambda(1-k)(\varphi))$ independent of the choice of x. This class is non-zero since the extension is non-split. The fact that c must be non-zero is proven by contradiction. If C were a coboundary (so the class c would be zero), we could adjust s in such a way that we have Galois equivariance, thus producing a splitting. Therefore since the extension is non-split, the class c must be non-zero.

We note that we may simplify $\mathrm{Hom}(\mathbb{F}_\lambda(\psi), \mathbb{F}_\lambda(1-k)(\varphi))$. We have

$$\begin{aligned} \mathrm{Hom}(\mathbb{F}_\lambda(\psi), \mathbb{F}_\lambda(1-k)(\varphi)) &\simeq (\mathbb{F}_\lambda(\psi))^* \otimes \mathbb{F}_\lambda(1-k)(\varphi) \\ &\simeq \mathbb{F}_\lambda(\psi^{-1}) \otimes \mathbb{F}_\lambda(1-k)(\varphi) \\ &\simeq \mathbb{F}_\lambda(1-k)(\psi^{-1}\varphi). \end{aligned}$$

We therefore have $c \in H^1(G_\mathbb{Q}, \mathbb{F}_\lambda(1-k)(\psi^{-1}\varphi))$. Let $V_\lambda = L_\lambda(1-k)(\psi^{-1}\varphi)$, let $M_\lambda = \mathcal{O}_\lambda(1-k)(\psi^{-1}\varphi)$, let $A_\lambda = V_\lambda/M_\lambda = (L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi)$ and let $A[\lambda]$ be the kernel of multiplication by λ in A_λ. Consider the inclusion $i : A[\lambda] \to A_\lambda$ and let $d := i_*(c) \in H^1(G_\mathbb{Q}, A_\lambda)$. Consider the following short exact sequence:

$$0 \longrightarrow A[\lambda] \xrightarrow{i} A_\lambda \xrightarrow{\text{“}\lambda\text{”}} A_\lambda \longrightarrow 0.$$

Here “λ” means multiplication by some uniformiser for λ. This short exact sequence gives rise to a long exact sequence in cohomology. We consider a piece of this sequence:

$$H^0(G_\mathbb{Q}, A_\lambda) \xrightarrow{\delta} H^1(G_\mathbb{Q}, A[\lambda]) \xrightarrow{i_*} H^1(G_\mathbb{Q}, A_\lambda).$$

We wish to know when i_* is injective. Since the sequence is exact, this is equivalent to knowing when the image of δ is trivial. This will be the case when $H^0(G_\mathbb{Q}, A_\lambda)$ is trivial. Since we can multiply by an appropriate power of λ, this is equivalent to $H^0(G_\mathbb{Q}, A[\lambda])$ being trivial. Consider $0 \neq x \in \mathbb{F}_\lambda(1-k)(\psi^{-1}\varphi)$. Then for $q \nmid Np\ell$,

$\mathrm{Frob}_q^{-1}(x) = (q^{k-1}\psi^{-1}\varphi(q))(x)$. Since $\ell \nmid N$ we may choose q such that

$$q \equiv \begin{cases} \zeta & (\mathrm{mod}\ \ell) \\ 1 & (\mathrm{mod}\ N), \end{cases}$$

where ζ is a primitive root modulo ℓ. Now if $(l-1) \nmid (k-1)$ then $\psi^{-1}\varphi(q) = 1$ since $q \equiv 1 \pmod{N}$ and $q^{k-1} \equiv \zeta^{k-1} \pmod{\ell}$. But then $\zeta^{k-1} \not\equiv 1 \pmod{\ell}$. It follows that $\mathrm{Frob}_q^{-1}(x) \neq x$ in $\mathbb{F}_\lambda$. For example, this condition is satisfied when $\ell > k$. Hence $H^0(G_{\mathbb{Q}}, A_\lambda)$ is trivial, so we see that i_* is injective and $d \neq 0$.

We would like d to belong to a Bloch-Kato Selmer group. We therefore now define these groups. Following [BlKa, §3], for $q \neq \ell$ let

$$H_f^1(G_{\mathbb{Q}_q}, V_\lambda) := \ker\left(H^1(G_{\mathbb{Q}_q}, V_\lambda) \to H^1(I_q, V_\lambda)\right).$$

Here $G_{\mathbb{Q}_q}$ has been identified with some decomposition subgroup at a prime above q, I_q is the inertia subgroup and the cohomology is for continuous cocycles and coboundaries. For $q = \ell$ let

$$H_f^1(G_{\mathbb{Q}_\ell}, V_\lambda) := \ker\left(H^1(G_{\mathbb{Q}_\ell}, V_\lambda) \to H^1(D_\ell, V_\lambda) \otimes_{\mathbb{Q}_\ell} B_{\mathrm{crys}}\right).$$

For a definition of Fontaine's ring B_{crys} see [BlKa, §1]. Let $H_f^1(G_{\mathbb{Q}}, V_\lambda)$ be the subspace of those elements of $H^1(G_{\mathbb{Q}}, V_\lambda)$ which, for all primes q, have local restriction lying in $H_f^1(G_{\mathbb{Q}_q}, V_\lambda)$. We have the natural exact sequence

$$0 \longrightarrow M_\lambda \longrightarrow V_\lambda \xrightarrow{\pi} A_\lambda \longrightarrow 0.$$

Let $H_f^1(G_{\mathbb{Q}_q}, A_\lambda) = \pi_* H_f^1(G_{\mathbb{Q}_q}, V_\lambda)$. Define the Selmer group $H_f^1(G_{\mathbb{Q}}, A_\lambda)$ to be the subgroup of elements of $H^1(G_{\mathbb{Q}}, A_\lambda)$ whose local restrictions lie in $H_f^1(G_{\mathbb{Q}_q}, A_\lambda)$ for all primes q. Since we have $\ell \nmid 6N$, in particular $\ell \neq 2$, we may omit $q = \infty$. More generally, given a finite set Σ of primes with $\ell \notin \Sigma$, we define $H_\Sigma^1(G_{\mathbb{Q}}, A_\lambda)$ to be the subgroup of elements of $H^1(G_{\mathbb{Q}}, A_\lambda)$ whose local restrictions lie in $H_f^1(G_{\mathbb{Q}_q}, A_\lambda)$ for all primes $q \notin \Sigma$.

Proposition 3.3.3. *Let* $\Sigma = \{q : q|N\}$. *Then* $d \in H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, A_\lambda)$.

Proof. For $q \nmid Np\ell$ (with $\lambda|\ell$) we have ρ_f unramified at q, that is $\rho_f|_{I_q}$ is trivial. It follows that the restriction of d to $H^1(I_q, A_\lambda)$ is 0 for such q. It follows from [Br, Lemma 7.4] that $d \in H_f^1(G_{\mathbb{Q}_q}, A_\lambda)$. If we assume that $\ell > k$ the representation ρ_f at ℓ is crystalline, we see that it satisfies the necessary local condition at ℓ; that is $d \in H_f^1(G_{\mathbb{Q}_\ell}, A_\lambda)$. This is a consequence of the second part of [DFG, Proposition 2.2]. Since the necessary local conditions are satisfied it follows that $d \in H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, A_\lambda)$. □

We now give a different way of producing a non-zero element of $H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, A_\lambda) = H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, (L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi))$. Let $L_{\Sigma\cup\{p\}}(k, \psi\varphi^{-1})$ be the partial Dirichlet L-function with Euler factors at primes $q \in \Sigma \cup \{p\}$ omitted. We now reformulate the λ-part of the Bloch-Kato conjecture, as in (59) of [DFG], similarly using the exact sequence as in their Lemma 2.1.

Conjecture 3.3.4 (Case of λ-part of Bloch-Kato)**.**

$$\begin{aligned}&\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})}{g(\psi\varphi^{-1})(2\pi i)^k}\right)\\&\quad=\mathrm{ord}_\lambda\left(\frac{\mathrm{Tam}^0_\lambda((L_\lambda/\mathcal{O}_\lambda)(k)\psi\varphi^{-1})\#H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}},(L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi))}{\#H^0(G_{\mathbb{Q}},(L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi))}\right).\end{aligned}$$

Note that here the period $g(\psi^{-1}\varphi)(2\pi i)^k$ follows from the calculation at the end of [DFG, §1.1.3]. Recall that $g(\psi^{-1}\varphi)$ is the Gauss sum of $\psi^{-1}\varphi$. We omit the definition of the Tamagawa factor $\mathrm{Tam}^0_\lambda((L_\lambda/\mathcal{O}_\lambda)(k)\psi\varphi^{-1})$, but note that (assuming $\ell > k+1$, recall that $\lambda|\ell$ here), its triviality is a direct consequence of [BlKa, Theorem 4.1(*iii*)]. In their notation we have $i = -k$ and $j = 1$. This case of the Bloch-Kato conjecture is actually known to be true. It was proven by Huber and Kings [HK, Theorem 5.4.1].

Recall that we have already shown that $H^0(G_{\mathbb{Q}},(L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi)) = H^0(G_{\mathbb{Q}},A_\lambda)$ is trivial. Since the Tamagawa factor is also trivial, we see that if $\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})}{g(\psi\varphi^{-1})(2\pi i)^k}\right) > 0$, then $\mathrm{ord}_\lambda\left(\#H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}},(L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi))\right) > 0$ and vice versa. Therefore if we can guarantee divisibility of the partial L-value by λ then we know that there must be a non-zero element in the Bloch-Kato Selmer group.

Proposition 3.3.5. $\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})}{g(\psi\varphi^{-1})(2\pi i)^k}\right) > 0\,.$

Proof. We will show that $\mathrm{ord}_{\lambda'}\left(\frac{L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})}{g(\psi\varphi^{-1})(2\pi i)^k}\right) > 0$. Since $\lambda|\lambda'$ this will imply that $\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})}{g(\psi\varphi^{-1})(2\pi i)^k}\right) > 0$. Concentrating on the $\{p\}$ part of $L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})$ here (recall $\ell \neq p$), we see that $\varphi(p)p^kL_{\{p\}}(k,\psi\varphi^{-1}) = (\varphi(p)p^k - \psi(p))L(k,\psi\varphi^{-1})$. Note that if $\lambda'|\varphi(p)p^k$ then we cannot have $\lambda'|\left(\varphi(p)p^k-\psi(p)\right)$. Also a generalisation of the Von Staudt-Clausen theorem, proven by Carlitz [Carl] tells us that the denominator of $L(1-k,\psi^{-1}\varphi)$ will not cancel any potential divisors of the Euler factor (see below for more details). Therefore if $\ell > k+1$ and $\lambda'|(\varphi(p)p^k-\psi(p))$, then $\mathrm{ord}_{\lambda'}\left(\frac{L_{\Sigma\cup\{p\}}(k,\psi\varphi^{-1})}{g(\psi\varphi^{-1})(2\pi i)^k}\right) > 0$. So the formula implies that $H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}},(L_\lambda/\mathcal{O}_\lambda)(1-k)(\psi^{-1}\varphi))$ contains a non-zero element. Of course here we are assuming that the factors in Σ do not cause any cancellation. □

We may use the functional equation of the Dirichlet L-function to consider $L(1-k,\psi^{-1}\varphi)$ instead of $L(k,\psi\varphi^{-1})$. We then have $L(1-k,\psi^{-1}\varphi) = -\frac{B_{k,\psi^{-1}\varphi}}{k}$. The generalisation of the Von Staudt-Clausen theorem proven by Carlitz then gives us conditions on this being an algebraic integer. Suppose we write $\frac{B_{k,\psi^{-1}\varphi}}{k} = \frac{\mathfrak{B}}{\mathfrak{D}}$ with $(\mathfrak{B},\mathfrak{D}) = 1$. If the conductor of $\psi^{-1}\varphi$, in this case N, contains at least two distinct primes then $\frac{B_{k,\psi^{-1}\varphi}}{k}$ is an algebraic integer as $\mathfrak{D}$ is the unit ideal. If instead $N = p^a$

for some a, then $\mathfrak{D}$ contains only prime ideal factors of p. In this case we have $\frac{B_{k,\psi^{-1}\varphi}}{k} \equiv \frac{1}{kN} C_k(\psi^{-1}\varphi) \pmod 1$ where

$$C_k(\psi^{-1}\varphi) = \sum_{r=1}^{N} \psi^{-1}\varphi(r) r^k.$$

Recall that since $\ell \nmid 6N$ there will be no cancellation in this case either.

Chapter 4

The Weight 1 Case

Now that we have proved our main result we wish to extend this to the case of weight 1 modular forms. This is not straightforward however, as there are several differences in this case as we have already seen partially in Section 1.4.4. One such difference is that we no longer have formulas for the dimensions of the spaces $M_1(\Gamma_1(N), \chi)$ and $S_1(\Gamma_1(N), \chi)$. This is down to certain terms in the Riemann-Roch formula not necessarily canceling each other out anymore. There are however some conjectured formulas for the dimension of $S_1(\Gamma_1(q), \chi)$ by Trotabas; See conjecture 2.1 of [Tr]. In fact, this same issue means we will not be able to make use of the same method of proof as we used for Theorem 3.0.1. The Deligne-Serre Lemma [DeSe] is not applicable in the case of weight 1. This is because the reduction map from $S_1(\Gamma_1(Np), \bar{\mathbb{Z}}_\ell)$ to $S_1(\Gamma_1(Np), \bar{\mathbb{F}}_\ell)$ is no longer surjective. Evidence of this was found by Buzzard [Bu]. In this paper he finds a mod 199 weight 1 cusp form of level 82 which does not lift to a cusp form of characteristic zero. There are many more examples of these kinds of modular forms in Schaeffer's thesis [Sch]. He refers to these particular modular forms as ethereal forms.

Although there are differences we still have methods for trying to generalise our main statement. Recall the construction of the weight 1 Eisenstein series from Section 1.4.4. Although this construction is rather different, we still end up with something that is very similar to the Eisenstein series of higher weight. In particular it is a modular form inside the correct vector space, has a similar Fourier expansion to higher weight Eisenstein series, and also has the Hecke eigenvalues you would expect of a weight 1 Eisenstein series. One difference now is that the constant term has a slightly different form, namely $\delta(\varphi)L(1-k, \psi) + \delta(\psi)L(1-k, \varphi)$ instead of $\delta(\psi)L(1-k, \varphi)$, where k is the weight of the corresponding Eisenstein series. We can naively work in the same way as before and ask the question: If a prime ℓ divides $\varphi(p)p - \psi(p)$, does there exist a congruence modulo λ (with $\lambda|\ell$), of Hecke eigenvalues, between a level N Eisenstein series and a level Np cusp form?

We will be looking first at the simplified case where we take one character to be trivial.

First assume that φ is trivial, i.e., we are looking at the series $E_1^{\psi,1}$. Hence we are interested in the Euler factor associated to $L(k, \psi\varphi^{-1}) = L(k, \psi)$. In exactly the same way as before we obtain the factor $p - \psi(p)$ (note that $k = 1$ and φ is trivial). Similarly if we take ψ to be trivial we obtain the factor $\varphi(p)p - 1$. We might expect divisors of these factors to be moduli of congruences. Also note that the Hecke eigenvalues (at a prime q) of these Eisenstein series are now simply $\psi(q) + 1$ and $1 + \varphi(q)$ respectively. We may expect to have to place some restrictions on our characters given the differences of the weight 1 case. But how can we decide on these restrictions? It turns out that the construction of a particular weight 1 cusp form will give us the answer to this. A special case of the result we will prove is the following:

Theorem 4.0.1. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field with discriminant N and associated quadratic character $\eta : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. Let p be a prime such that $p \nmid N$. Let $\ell \nmid 6N$ be a prime such that $ord_\ell((p - \eta(p))h) > 0$ where h is the class number of K. If p splits in K then there is a normalised Hecke eigenform $f \in S_1(\Gamma_1(Np), \eta')$ (where η' is congruent to η modulo λ) such that for all q with q prime,*

$$a_q(f) \equiv 1 + \eta(q) \pmod{\lambda},$$

where $\lambda|\ell$ is a prime of $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$. If p is inert in K then there is a normalised Hecke eigenform $f \in S_1(\Gamma_1(Np^2), \eta')$ (where η' is congruent to η modulo λ) such that for all q with q prime,

$$a_q(f) \equiv 1 + \eta(q) \pmod{\lambda},$$

where $\lambda|\ell$ is a prime of $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$.

Notice the very similar structure of this result to Theorem 3.0.1. The Euler factor $\varphi(p)p^k - \psi(p)$ now simplifies to $p - \eta(p)$; this follows because we are considering $\varphi = \mathbf{1}_1$, $\psi = \eta$ and $k = 1$. The class number of K is now the simplification of $B_{k,\psi^{-1}\varphi}/k$, both of which come from $L(1 - k, \psi^{-1}\varphi)$. Note that this is because $L(1 - k, \psi^{-1}\varphi)$ is now $L(0, \eta)$ which gives the class number. The Hecke eigenvalues of the Eisenstein series are now much simpler due to the fact that one character is trivial and $k = 1$; $\psi(q) + q^{k-1}\varphi(q)$ is simply $\eta(q) + 1$. The main difference with this result is the second half where we consider an inert prime in K. In this case the level of the cusp form is Np^2 and so we have raised the level by p^2. Although this isn't something we considered in the higher weight case it is a special consequence of the particular construction that we use. This is the simplest case and its proof will serve as a template for the more general result. As long as we have primitive Dirichlet characters ψ and φ, of conductors u and v respectively, with $uv = N$ and $\psi \equiv \eta\varphi \pmod{\lambda}$, then we obtain a congruence.

Theorem 4.0.2. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field with discriminant N and associated quadratic character $\eta : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. Suppose ψ and φ are primitive Dirichlet characters of conductors u and v respectively with $uv = N$. Let p be a prime such that $p \nmid N$. Let $\ell \nmid 6N$ be a prime such that $ord_\ell((\varphi(p)p - \psi(p))h) > 0$ where h is the class number of K. Assume that $\psi \equiv \eta\varphi \pmod{\lambda}$. If p splits in K*

then there is a normalised Hecke eigenform $f \in S_1(\Gamma_1(Nvp), \epsilon)$ *(where* ϵ *is congruent to* $\psi\varphi$ *modulo* λ*) such that for all* q *with* q *prime,*

$$a_q(f) \equiv \psi(q) + \varphi(q) \pmod{\lambda},$$

where $\lambda|\ell$ *is a prime of* $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$*. If* p *is inert in* K *then there is a normalised Hecke eigenform* $f \in S_1(\Gamma_1(Nvp^2), \epsilon)$ *(where* ϵ *is congruent to* $\psi\varphi$ *modulo* λ*) such that for all* q *with* q *prime,*

$$a_q(f) \equiv \psi(q) + \varphi(q) \pmod{\lambda},$$

where $\lambda|\ell$ *is a prime of* $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$*.*

Notice how Theorem 4.0.1 is a special case of this result. The differences here are that η is now related to two Dirichlet characters ψ and φ via a congruence condition and the level of the cusp forms have now been raised by an additional factor of v. This choice is arbitrary however, we could have just as easily raised the level by u instead. The details of this will become clear when we work through the proof.

We will later show that these two results cover everything that is happening in the weight 1 case. Basically if we have any Dirichlet character that isn't associated to a quadratic field then there isn't a congruence.

Theorem 4.0.3. *Suppose* ψ *and* φ *are primitive Dirichlet characters of conductors* u *and* v *respectively with* $uv = N$*, so* $E_1^{\psi,\varphi} \in M_1(\Gamma_1(N), \psi\varphi)$*. Let* $f \in S_1(\Gamma_1(M), \epsilon)$ *have associated Galois representation* ρ_f*. There exists a congruence between the Hecke eigenvalues of* $E_1^{\psi,\varphi}$ *and* f *only if* $\tilde{\rho}_f(G_\mathbb{Q}) \cong D_n$ *for some* n*.*

In order to prove these results we will need some background on the theory of Galois representations. Once we have the necessary background we will combine this with some class field theory in order to prove Theorem 4.0.1. The proof of Theorem 4.0.2 will then be a fairly straightforward adaptation of the method used to prove Theorem 4.0.1. We will then need to consider Schur covers of certain groups in order to complete the proof of Theorem 4.0.3.

§ 4.1 Galois Representations and Weight One Modular Forms

As with the higher weight cases we aim to try and construct a weight 1 cusp form whose Hecke eigenvalues satisfy a congruence. We already know that we will not be able to apply the same method as before. We therefore need to come up with a new construction without using the Deligne-Serre Lemma. Our aim is to use the theory of Galois representations. We will make use of the work of Khare-Wintenberger [KhWi] proving that a certain type of Galois representation is modular; that is, it gives rise to a newform of weight 1. This will be made precise once we have some basic definitions regarding Galois representations.

Let $\bar{\mathbb{Q}}/\mathbb{Q}$ be an algebraic closure of $\mathbb{Q}$ and let $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ denote the absolute Galois group. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{C})$ be a two dimensional continuous complex linear representation of $G_{\mathbb{Q}}$. Such a representation is called a *Galois representation*. Note that we will assume our representations are continuous from now on. This essentially corresponds to having open kernel, and hence, finite image. It then follows that ρ factors through $\mathrm{Gal}(K/\mathbb{Q})$ for some finite Galois extension $K/\mathbb{Q}$. We let $\epsilon = \det(\rho)$, which is a one dimensional representation of $G_{\mathbb{Q}}$.

If we take any embedding of $\bar{\mathbb{Q}}$ in $\mathbb{C}$ then complex conjugation on $\mathbb{C}$ induces an automorphism of $\bar{\mathbb{Q}}$. Any such automorphism is known as a Frobenius element at infinity or simply a complex conjugation. Suppose χ is a one dimensional linear representation of $G_{\mathbb{Q}}$. If $\chi(c) = -1$, where c is a complex conjugation, we say that χ is odd. Note that all such c are conjugate, so the condition $\chi(c) = -1$ is independent of the choice of embedding of $\bar{\mathbb{Q}}$ into $\mathbb{C}$.

Let N be the Artin conductor, and $L(s, \rho)$ the Artin L-function, of the representation ρ. The Artin L-function is defined by an Euler product over prime ideals $\mathfrak{p}$:

$$L(s, \rho) = \prod_{\mathfrak{p}} \det[I - \mathrm{N}(\mathfrak{p})^{-s} \rho(\mathrm{Frob}_{\mathfrak{p}}^{-1})]^{-1}$$

For a definition of the Artin conductor and some basic properties of the L-function, see [Cog]. The conductor of the representation $\epsilon = \det(\rho)$ divides N and so it may be viewed as a Dirichlet character mod N

$$\epsilon : (\mathbb{Z}/N\mathbb{Z})^{\times} \to \mathbb{C}^{\times}.$$

Note that we may do this by using class field theory since $(\mathbb{Z}/N\mathbb{Z})^{\times}$ is the Galois group of a cyclotomic quotient of $G_{\mathbb{Q}}$ (namely $\mathrm{Gal}(\mathbb{Q}_{\zeta_N}/\mathbb{Q})$). The representation ϵ is odd if and only if this Dirichlet character satisfies $\epsilon(-1) = -1$.

Now that we have set up the basic notation we are in a position to discuss the work of Khare-Wintenberger. The main result of the paper [KhWi] was a proof of Serre's modularity conjecture. This result involves mod p representations. Although we will not be directly interested with these representations, we will consider them when proving Theorem 4.0.3. We therefore give a basic overview of the conjecture. Suppose $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F})$ is an absolutely irreducible, continuous, two-dimensional, odd, mod p representation, with $\mathbb{F}$ a finite field of characteristic p. Such a representation is said to be of *Serre type*. Given a normalised eigenform

$$f = q + a_2 q^2 + a_3 q^3 + \dots$$

of some level N, weight k, and character $\chi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{F}^{\times}$, a theorem of Deligne, Eichler, Serre and Shimura attaches to f a representation $\rho_f : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathcal{O})$, where $\mathcal{O}$ is the ring of integers in a finite extension of $\mathbb{Q}_\ell$. For primes $p \nmid N\ell$, we have

$$\mathrm{Tr}(\rho_f(\mathrm{Frob}_p^{-1})) = a_p$$

and

$$\det(\rho_f(\mathrm{Frob}_p^{-1})) = p^{k-1}\chi(p).$$

The reduction of this representation modulo the maximal ideal of $\mathcal{O}$ gives a mod ℓ representation $\overline{\rho}_f$ of $G_{\mathbb{Q}}$. Serre's conjecture states that for ρ of Serre type, there is an eigenform f of level $\mathrm{N}(\rho)$, weight $k(\rho)$ and character $\chi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{F}^{\times}$ such that $\overline{\rho}_f \simeq \rho$. In their paper, Khare and Wintenberger prove this conjecture and then go on to deduce some results about representations in characteristic 0. These are the representations we will be interested in while proving Theorem 4.0.1.

The result that we are interested in, namely modularity of irreducible, two-dimensional, complex, odd Galois representations, is intimately tied in with the problem of solving Artin's conjecture on L-functions. For any $\rho : G_{\mathbb{Q}} \to \mathbb{C}^{\times}$ such that ϵ is odd, define:

$$\Lambda(s, \rho) = N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, \rho).$$

The function Λ then extends to a meromorphic function on the whole complex plane, and satisfies the following functional equation:

$$\Lambda(s, \rho) = W(\rho).\Lambda(1 - s, \rho^*).$$

Here $W(\rho)$ is a certain complex number with absolute value 1 known as the Artin root number and ρ^* is the complex conjugate representation. The Artin conjecture (in the two-dimensional case) states that the function $\Lambda(s, \rho)$ actually has an analytic continuation to the whole complex plane.

Over the years many different people have proven different cases of Artin's conjecture. Many of the two-dimensional cases were known by 1981. The final (odd) two-dimensional case was proven in 2009 by Khare and Wintenberger. Langlands had made a general conjecture, which in our case states that ρ arises from a cuspidal automorphic representation π of $\mathrm{GL}_2(\mathbb{A}_{\mathbb{Q}})$. From this we get that $L(s, \rho) = L(s, \pi)$. It is known that $L(s, \pi)$ has an analytic continuation to the whole complex plane and therefore one may deduce Artin's conjecture from this. In fact Theorem 10.1(i) of [KhWi] implies Langlands' conjecture, and therefore Artin's conjecture. This then provides a converse to a theorem of Deligne-Serre which attaches a complex representation to a newform of weight one. That is they proved the modularity of ρ and deduced Artin's conjecture from this. The result of Khare-Wintenberger is Corollary 10.2(ii) of [KhWi] and the result of Deligne-Serre is Theorem 2 from Section 3 of [S]. We state both here.

Theorem 4.1.1 (Khare-Wintenberger)**.** *Let ρ be an irreducible two-dimensional complex linear representation of $G_{\mathbb{Q}}$ with conductor N and $\epsilon = det(\rho)$ odd. Suppose $L(s, \rho) = \sum_{n=1}^{\infty} a_n n^{-s}$, and let $f(z) = \sum_{n=1}^{\infty} a_n q^n$. Then $f \in S_1(\Gamma_1(N), \epsilon)$ is a normalised newform.*

Theorem 4.1.2 (Deligne-Serre)**.** *Let $f \in S_1(\Gamma_1(N), \epsilon)$ be a normalised newform. Then there exists an irreducible two-dimensional complex linear representation ρ of $G_{\mathbb{Q}}$ such that $L_f(s) = L(s, \rho)$. Further, the conductor of ρ is N, and $det(\rho) = \epsilon$.*

Note that the even two-dimensional case is still an open problem but we only deal with odd representations so this is not an issue. The case that Khare and Wintenberger proved was the case when the projective image of ρ is non-solvable, that is, the projective image is A_5. The cases of solvable projective image were proven by various people. The case when the projective image is S_4 was proven by Tunnell [Tu] in 1981. The case when the projective image is A_4 was proven by Langlands [L] in 1980. The cases when the projective image is dihedral or cyclic were proven by Artin and Hecke. We will see more about the projective images of complex representations later, in particular, why these are the only possible images.

Before these cases had been proven it was necessary to rely on a technical condition in order to ensure that two-dimensional, odd, irreducible complex Galois representations were modular. This condition can be found in [S] from 1977; we state the condition here.

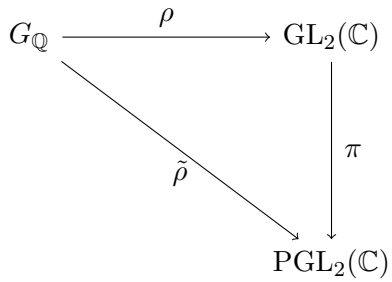
Condition 4.1.3. There exists a positive integer M such that, for all one dimensional linear representations χ of $G_{\mathbb{Q}}$ with conductor coprime to M, $\Lambda(s, \rho \otimes \chi)$ is a holomorphic function of s for $s \neq 0, 1$.

Notice that this condition is simply the Artin conjecture for $\rho \otimes \chi$. Since ρ is two-dimensional, $\det(\rho \otimes \chi) = \det(\rho)\chi^2$, so if ρ is odd, then so is $\rho \otimes \chi$. The representation ρ was known to satisfy Condition 4.1.3 if it was reducible or induced from a one-dimensional representation. Assuming this condition was enough to give the result proven by Khare-Wintenberger for this ρ. This was originally a result of Weil-Langlands. Note that for this representation ρ we have that $\mathrm{Tr}(\rho(\mathrm{Frob}_p^{-1})) = a_p$ and $\det(\rho(\mathrm{Frob}_p^{-1})) = \chi(p)$.

The theorems of Khare-Wintenberger and Deligne-Serre are of great importance since they give a concrete link between cusp forms of weight 1 and two-dimensional Galois representations. Since we know quite a bit about Galois representations this will make studying weight 1 cusp forms much easier. After looking at a little more theory we will be in a position to construct a weight 1 cusp form whose Hecke eigenvalues will satisfy a congruence as stated in Theorem 4.0.1. This cusp form will be constructed from a particular type of Galois representation making use of Theorem 4.1.1. It turns out that the existence of a congruence will depend on the projective image of the representation ρ. We therefore need to know some background material on projective representations.

4.1.1 Projective Representations

Firstly we note that a two-dimensional representation ρ gives rise to a projective linear representation, $\tilde{\rho}$, of $G_{\mathbb{Q}}$:



where $\mathrm{PGL}_2(\mathbb{C}) = \mathrm{GL}_2(\mathbb{C})/\mathbb{C}^\times$. The image of $\tilde{\rho}$ is a finite subgroup of $\mathrm{PGL}_2(\mathbb{C})$, and is therefore one of the following:

(1) C_n - cyclic of order n;

(2) D_n - dihedral of order $2n, n \geq 2$;

(3) the alternating groups A_4, A_5, or the symmetric group S_4.

This classification is well known. A purely algebraic proof can be found in Section 1.1 of Dolgachev's notes [Do]. Notice that these are exactly the cases that were checked for the Artin conjecture.

First of all we consider the case that $\mathrm{im}(\tilde{\rho})$ is cyclic. Suppose we have $a, b \in \mathrm{im}(\rho)$. Then $\pi(a), \pi(b) \in \mathrm{im}(\tilde{\rho})$. Since $\mathrm{im}(\tilde{\rho})$ is cyclic, we have $\pi(a) = g^r$ and $\pi(b) = g^s$ for some $g \in \mathrm{im}(\tilde{\rho})$ and $r, s \in \mathbb{Z}$. Hence under the preimage of π we see that $a = \lambda_a g^r$ and $b = \lambda_b g^s$ for some $\lambda_a, \lambda_b \in \mathbb{C}^\times$. It therefore follows that $ab = ba$ and $\mathrm{im}(\rho)$ is abelian. From this we deduce that ρ is reducible. We therefore wouldn't be able to use Theorem 4.1.1 to obtain a weight 1 cusp form. Hence we ignore this case. The case we will be interested in to begin with is case (2), the representations whose projective image is dihedral. It will turn out that this is the only case in which we can obtain a congruence as stated in Theorem 4.0.1 and Theorem 4.0.2. The remaining cases will all be considered when we prove Theorem 4.0.3 since none of these cases lead to a congruence. We note that, as implied in the argument above, a projective representation $\tilde{\rho}$ can be lifted to a linear representation ρ. We defer the details of this until Section 4.2.3 where we will need to consider the liftings in proving Theorem 4.0.3. For the proof of Theorem 4.0.1, and then the more general proof for Theorem 4.0.2, it is simply enough to know that such a lift exists.

4.1.2 Dihedral Representations

Now that we know the basics of Galois representations and their projective image, we can study the case that we are interested in. For now we simply assume results about liftings, for those wanting to know the details, consult Section 4.2.3. The following discussion, up to the statement of Proposition 4.1.4, is largely based on [S, §7]. Let

$\tilde{\rho}$ be a two-dimensional projective representation of $G_{\mathbb{Q}}$, and let ρ be some lifting of $\tilde{\rho}$. (This is possible by Theorem 4.2.11 since $\mathbb{Q}$ is a global field.) We say that $\tilde{\rho}$ (or ρ) is *dihedral* if $\tilde{\rho}(G_{\mathbb{Q}}) \subset \mathrm{PGL}_2(\mathbb{C})$ is isomorphic to the dihedral group D_n of order $2n$, for some $n \geq 2$. If ρ is dihedral, then it is an irreducible representation. We would therefore be able to use such a representation to construct a weight 1 cusp form using Theorem 4.1.1. In constructing the dihedral representation ρ, there will be a few technical details that need to be checked which we now discuss.

Let C_n be a cyclic subgroup of D_n of order n; if $n \geq 3$, C_n is uniquely determined. Suppose $\tilde{\rho}$ is a dihedral representation. Consider the composition

$$\omega : G_{\mathbb{Q}} \xrightarrow{\quad\tilde{\rho}\quad} D_n \longrightarrow D_n/C_n = \{\pm 1\} \ .$$

This is again a representation of $G_{\mathbb{Q}}$. It is a one-dimensional linear representation of order 2, which corresponds to some quadratic extension $K/\mathbb{Q}$. Let $G_K = \mathrm{Gal}(\bar{\mathbb{Q}}/K) \subset G_{\mathbb{Q}}$. Then $\tilde{\rho}(G_K) = C_n$, and $\rho|_{G_K}$ is reducible:

$$\rho|_{G_K} = \chi \oplus \chi^{'},$$

say, for some one-dimensional representations $\chi, \chi^{'}$ of G_K. If σ lies in the non-identity coset of $G_{\mathbb{Q}}/G_K$, then $\chi^{'} = \chi_\sigma$, where $\chi_\sigma(\gamma) = \chi(\sigma\gamma\sigma^{-1}), \gamma \in G_K$. Further, $\rho = \mathrm{Ind}_{K/\mathbb{Q}}(\chi)$, the representation of $G_{\mathbb{Q}}$ induced by χ.

Conversely, suppose we start with a quadratic number field $K/\mathbb{Q}$, corresponding to a character ω of $G_{\mathbb{Q}}$, and a one-dimensional linear representation χ of G_K. Let $\rho = \mathrm{Ind}_{K/\mathbb{Q}}(\chi)$, and let $\tilde{\rho}$ be the associated projective representation of $G_{\mathbb{Q}}$. If σ generates $\mathrm{Gal}(K/\mathbb{Q})$, let χ_σ be as above. Let $\mathfrak{f}$ be the conductor of χ, and d_K the discriminant of $K/\mathbb{Q}$.

Proposition 4.1.4. *With the above notations*

(1) The following are equivalent:

- *(1) ρ is irreducible;*
- *(2) ρ is dihedral;*
- *(3) $\chi \neq \chi_\sigma$.*

(2) The conductor of ρ is $|d_K|.N_{K/\mathbb{Q}}(\mathfrak{f})$.

(3) The representation $det(\rho) = \omega\chi_{\mathbb{Q}}$ of $G_{\mathbb{Q}}$ is odd if and only if either:

- *(1) K is imaginary*
  or
- *(2) K is real and χ has signature $+,-$ at infinity; that is, if $c, c^{'} \in G_K$ are Frobenius elements at the two real places of K, then $\chi(c) \neq \chi(c^{'})$. So $\{\chi(c), \chi(c')\} = \{1, -1\}$.*

(4) If $\tilde{\rho}(G_{\mathbb{Q}}) = D_n$, then n is the order of $\chi^{-1}.\chi_\sigma$.

For a proof see page 239 of [S]. Note that the determinant of ρ has the form $\omega\chi_{\mathbb{Q}}$ where ω is the quadratic character associated to the quadratic field K and $\chi_{\mathbb{Q}}$ is the restriction of χ to a character of $\mathbb{Q}$. A detailed description of $\chi_{\mathbb{Q}}$ is given in the proof by Serre [S, p. 239]. We give a brief description in the following section when we choose a particular χ, see the map given in (4.1).

It therefore follows that if we use a one-dimensional representation χ of G_K and induce to a representation of $G_{\mathbb{Q}}$, we will obtain a dihedral representation providing that $\chi \neq \chi_\sigma$. This representation will then be irreducible and, combining Proposition 4.1.4 and Theorem 4.1.1, we see that we are able to construct a weight 1 cusp form of level $|d_K|.N_{K/\mathbb{Q}}(\mathfrak{f})$ where $\mathfrak{f}$ is the conductor of χ. In detail, suppose $\rho = \mathrm{Ind}_{K/\mathbb{Q}}(\chi)$ is a dihedral representation of $G_{\mathbb{Q}}$. If $\epsilon = \det(\rho)$ is odd, and we put

$$L(s,\rho) = \sum_{n=1}^{\infty} a_n n^{-s}, \quad f(z) = \sum_{n=1}^{\infty} a_n q^n,$$

then, by Theorem 4.1.1, $f(z) \in S(\Gamma_1(N), \epsilon = \omega\chi_{\mathbb{Q}})$ where $N = |d_K|.N_{K/\mathbb{Q}}(\mathfrak{f})$. There are several examples given in [S]. We are now in a position to tie all of the theory together and prove Theorem 4.0.1.

§ 4.2 Proving the Weight 1 Theorems

4.2.1 Proof of Theorem 4.0.1

Before moving on to construct a particular cusp form we take a brief pause to explain the strategy. Although we are seeking an irreducible Galois representation in order to make use of the modularity we are also seeking the existence of a congruence modulo ℓ. That is, we require the representation to be reducible modulo ℓ. This will correspond to the requirement that $\mathrm{Tr}(\mathrm{Frob}_q^{-1}) = a_q$ where a_q is the q-th Hecke eigenvalue of the cusp form f, and $\det(\mathrm{Frob}_q^{-1}) = \eta(q)$ where η is the quadratic character given in Theorem 4.0.1. Over the course of this section we will see exactly how to guarantee that this occurs. We note that we will also be able to observe the congruence simply by knowing the Hecke eigenvalues of the cusp form that we will construct.

Let us first consider an imaginary quadratic field, that is $K = \mathbb{Q}(\sqrt{-d})$ for some square-free $d > 0$. Associated to this field, we have the quadratic character $\eta : \mathrm{Gal}(K/\mathbb{Q}) \to \{\pm 1\}$, or equivalently $\eta : (\mathbb{Z}/|d_K|)^\times \to \{\pm 1\}$. Now take a Hecke (ray class) character $\chi : I_K(\mathfrak{m})/P_{1,K}(\mathfrak{m}) \to \mathbb{C}^\times$, where $\mathfrak{m}$ is a particular modulus. Later we will be interested in the case where $\mathfrak{m}$ consists of a single prime ideal but for now we will work with a general modulus. Note that, via class field theory, we may view χ as a character of

G_K. We then have

$$L(s,\chi) = \sum_{\substack{\mathfrak{a}\in\mathcal{O}_K \\ \mathfrak{a}\text{ integral} \\ (\mathfrak{a},\mathfrak{m})=1}} \frac{\chi(\mathfrak{a})}{\mathrm{N}(\mathfrak{a})^s} = \prod_{\substack{\mathfrak{p}\in\mathcal{O}_K \\ \mathfrak{p}\text{ prime} \\ (\mathfrak{p},\mathfrak{m})=1}} \frac{1}{1-\chi(\mathfrak{p})\mathrm{N}(\mathfrak{p})^{-s}}.$$

We would like $\rho = \mathrm{Ind}_{K/\mathbb{Q}}(\chi)$ to be an irreducible representation. We therefore, for now, assume that $\chi \neq \chi_\sigma$. We will split the proof of Theorem 4.0.1 into three separate cases, each using a different ray class character, and we will check this condition in each case. We then use a general result about Artin L-functions to tell us that $L(s,\rho) = L(s,\mathrm{Ind}_{K/\mathbb{Q}}\chi) = L(s,\chi)$. It then follows (from Proposition 4.1.4 and Theorem 4.1.1) that

$$f = \sum_{\substack{\mathfrak{a}\in\mathcal{O}_K \\ \mathfrak{a}\text{ integral}}} \chi(\mathfrak{a}) q^{\mathrm{N}(\mathfrak{a})}$$

is a modular form of weight 1, level $|d_K|.\mathrm{N}(\mathfrak{m})$ and character $\eta\chi_{\mathbb{Q}}$. Here we note that we have a map

$$I_{\mathbb{Q}}(m)/P_{1,\mathbb{Q}}(m) \xrightarrow{i} I_K(\mathfrak{m})/P_{1,K}(\mathfrak{m}) \tag{4.1}$$

where $m = \mathbb{Z}\cap\mathfrak{m}$. Here χ is a character of the ray class group on the right and $\chi_{\mathbb{Q}}$ is the restriction defined as a character of the group on the left. We will take a particular character χ such that $\chi_{\mathbb{Q}}$ becomes trivial modulo ℓ. That is, the determinant of ρ, and therefore the character of f becomes η. This suggests that, as alluded to earlier, we should consider the Eisenstein series $E_1^{1,\eta}$ of level $|d_K|$. Unlike the higher weight case, we may now observe the congruence at all primes q, including those that divide N and p. Since η is the quadratic character associated to K, the Hecke eigenvalue of the Eisenstein series $E_1^{1,\eta}$ (or $E_1^{\eta,1}$) at a prime q is given by

$$1+\eta(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ 1 & \text{if } q \text{ ramifies in } \mathcal{O}_K, \\ 2 & \text{if } q \text{ splits in } \mathcal{O}_K. \end{cases}$$

We note that the Hecke eigenvalues for the cusp form f can simply be read off from the Fourier expansion. The Hecke eigenvalue at a prime q for the cusp form f is given by

$$a_f(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ \chi(\mathfrak{q}) & \text{if } q = \mathfrak{q}^2 \text{ ramifies in } \mathcal{O}_K, \\ \chi(\mathfrak{q}_1)+\chi(\mathfrak{q}_2) & \text{if } q = \mathfrak{q}_1\mathfrak{q}_2 \text{ splits in } \mathcal{O}_K. \end{cases}$$

It is clear that at a prime q which is inert in $\mathcal{O}_K$ both the Hecke eigenvalue of the Eisenstein series and the cusp form f is 0. Therefore there will automatically be a congruence at such a prime. Can we ensure that there is a congruence also at split

or ramified primes? Note that for a ramified prime q we are really considering the extension of the Dirichlet character $\eta : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ to $\eta : (\mathbb{Z}/N\mathbb{Z}) \to \mathbb{C}^\times$, where $\eta(q) = 0$ for all ramified primes. We therefore want to ensure that $\chi(\mathfrak{q}) \equiv 1 \pmod{\lambda}$ or $\chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) \equiv 2 \pmod{\lambda}$ for some prime λ. Here λ will be a prime of a large enough coefficient field (i.e. containing the values of χ). Note that the Hecke character takes its values in $\mathbb{C}^\times$, in other words, the d-th roots of unity for some d which will be a divisor of the order of the ray class group. It is known that for ℓ coprime to d, these values will be inequivalent modulo λ. It is clear that if we can guarantee that the values taken by χ are all congruent to 1 modulo λ then the congruence will be satisfied. If we have $\ell = d$ then this will be possible since we have $x^\ell - 1 \equiv (x-1)^\ell \pmod{\ell}$. That is, all ℓ-th roots of unity will be congruent to 1 modulo λ. Hence one necessary condition is that we must have ℓ dividing the order of the ray class group $C_\mathfrak{m} = I_K(\mathfrak{m})/P_{1,K}(\mathfrak{m})$.

There is a known formula for the size of a ray class group. This formula holds for any number field K, although it is not always easy to calculate. In our case (K being imaginary quadratic) it is simple.

Recall the definition of a modulus $\mathfrak{m} = \mathfrak{m}_0\mathfrak{m}_\infty$. We can write the finite part $\mathfrak{m}_0$ as

$$\mathfrak{m}_0 = \prod_\mathfrak{p} \mathfrak{p}^{m(\mathfrak{p})}$$

where $m(\mathfrak{p})$ is 1 if $\mathfrak{p}$ divides $\mathfrak{m}$ and 0 otherwise.

Theorem 4.2.1. *For every modulus $\mathfrak{m}$ of K there is an exact sequence*

$$0 \to U/U_{1,K}(\mathfrak{m}) \to P_K(\mathfrak{m})/P_{1,K}(\mathfrak{m}) \to C_\mathfrak{m} \to C \to 0$$

and canonical isomorphisms

$$P_K(\mathfrak{m})/P_{1,K}(\mathfrak{m}) \cong \prod_{\substack{\mathfrak{p}\ real \\ \mathfrak{p}|\mathfrak{m}}} \{\pm\} \times \prod_{\substack{\mathfrak{p}\ finite \\ \mathfrak{p}|\mathfrak{m}}} (\mathcal{O}_K/\mathfrak{p}^{\mathfrak{m}(\mathfrak{p})})^\times \cong \prod_{\substack{\mathfrak{p}\ real \\ \mathfrak{p}|\mathfrak{m}}} \{\pm\} \times (\mathcal{O}_K/\mathfrak{m}_0)^\times,$$

where

$$U = \mathcal{O}_K^\times, \text{ the group of units in } K,$$

$$U_{1,K}(\mathfrak{m}) = U \cap P_{1,K}(\mathfrak{m}).$$

Therefore $C_\mathfrak{m}$ is a finite group of order

$$h_\mathfrak{m} = \frac{h_K . 2^{r_0} . \mathrm{N}(\mathfrak{m}_0) . \prod_{\mathfrak{p}|\mathfrak{m}_0} \left(1 - \frac{1}{\mathrm{N}(\mathfrak{p})}\right)}{[U : U_{1,K}(\mathfrak{m})]}$$

where r_0 is the number of real places of $\mathfrak{m}$ and h_K is the class number of K.

A proof of this formula is given on page 147 of [Mil]. The idea of the proof is to calculate the order of the groups in the exact sequence and then use this to calculate the order of $C_{\mathfrak{m}}$. Note that when K is imaginary quadratic there will be no real places in $\mathfrak{m}$. Also the unit group U is either $\{\pm 1\}, \{\pm 1, \pm i\}$ or $\{\pm 1, \pm\zeta, \pm\zeta^2\}$ where ζ is a cube root of unity. Note that the last two cases only occur when $K = \mathbb{Q}(i)$ or $K = \mathbb{Q}(\sqrt{-3})$ respectively. The factor $[U : U_{1,K}(\mathfrak{m})]$ is therefore either $1, 2, 3, 4$ or 6. In each case the only primes involved are 2 or 3 and these are exactly the primes we avoid. This factor therefore does not cancel out any potential moduli.

We also note that our Eisenstein series $E_1^{1,\eta}$ has level $|d_K|$ and the cusp form will have level $|d_K|.\mathrm{N}(\mathfrak{m})$. The level of the cusp form will depend on our choice of modulus. If we take our modulus to be $\mathfrak{m} = \mathfrak{p}$ for a prime $\mathfrak{p}$ lying above a rational prime p which splits in K, the cusp form will have level $|d_K|p$. If we took the modulus to be $(m) = \mathfrak{p}$ with $\mathfrak{p}$ lying above an inert prime p, then the level of the corresponding cusp form would be $|d_K|p^2$. These statements follow because the conductor of ρ is given by Proposition 4.1.4 and then Theorem 4.1.1 gives the level of the corresponding cusp forms.

Let us now return to the Eisenstein series and consider the Euler factors which arise. This should allow us to find some conditions for the modulus ℓ. First of all considering $E_1^{1,\eta}$, the corresponding Euler factor is $\eta(p)p - 1$. Since η is the quadratic character we have associated to K

$$\eta(p)p - 1 = \begin{cases} -(p+1) & \text{if } p \text{ is inert in } \mathcal{O}_K, \\ p - 1 & \text{if } p \text{ splits in } \mathcal{O}_K. \end{cases}$$

If we now consider the series $E_1^{\eta,1}$ we have $p - \eta(p)$ as our Euler factor. We then have

$$p - \eta(p) = \begin{cases} p+1 & \text{if } p \text{ is inert in } \mathcal{O}_K, \\ p - 1 & \text{if } p \text{ splits in } \mathcal{O}_K. \end{cases}$$

Note that whether we consider $E_1^{1,\eta}$ or $E_1^{\eta,1}$, the primes dividing the Euler factor will be the same. This suggests that we should expect to find a congruence with a cusp form of higher level with modulus $\lambda|\ell$ if ℓ divides the Euler factor, in particular if it divides either $p+1$ or $p-1$. Also, as was mentioned in the discussion following Theorem 4.0.1, we now expect a prime dividing $L(0, \eta)$ to give a congruence between the Hecke eigenvalues of an Eisenstein series and a cusp form of the same level. As we will see, this is where the class number appears. We now split the proof of Theorem 4.0.1 into three separate propositions each dealing with a congruence at a different level.

We will first deal with the case that $\ell|h_K$.

Proposition 4.2.2. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field with discriminant N and associated quadratic character $\eta : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. Let $\ell \nmid 6N$ be a prime such that $\ell|h_K$. Then there exists a normalised Hecke eigenform $f \in S_1(\Gamma_1(N), \eta')$ (where η' is congruent to η modulo λ) such that for all q with q prime,*

$$a_q(f) \equiv 1 + \eta(q) \pmod{\lambda},$$

where $\lambda|\ell$ is a prime of $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$.

Proof. Since $\ell|h_K$, the order of $C_K \simeq I_K/P_K$, we may choose $\chi : C_K \to \mathbb{C}^\times$ of exact order ℓ. Via the isomorphism $C_K \simeq \mathrm{Gal}(H/K)$ where H is the Hilbert class field, we may also view χ as a character of $\mathrm{Gal}(H/K)$. Then for a prime ideal $\mathfrak{q}$ we have $\chi([\mathfrak{q}]) = \chi(\mathrm{Frob}_\mathfrak{q})$ since $[\mathfrak{q}] \mapsto \mathrm{Frob}_\mathfrak{q}$ via the Artin isomorphism. We write $\chi(\mathfrak{q})$ for $\chi([\mathfrak{q}])$. We would like to show that $\chi \neq \chi_\sigma$. We have

$$\chi_\sigma(\mathfrak{q}) = \chi_\sigma(\mathrm{Frob}_\mathfrak{q}) \overset{\text{by def of } \chi_\sigma}{=} \chi(\sigma \mathrm{Frob}_\mathfrak{q} \sigma^{-1}) \overset{\text{by Theorem 2.2.11}}{=} \chi(\mathrm{Frob}_{\sigma\mathfrak{q}}).$$

Now $\mathfrak{q}\sigma(\mathfrak{q}) \in P_K$ so $\chi(\mathfrak{q}\sigma(\mathfrak{q})) = 1$. So $\chi(\mathfrak{q})\chi_\sigma(\mathfrak{q}) = \chi(\mathfrak{q})\chi(\sigma(\mathfrak{q})) = \chi(\mathfrak{q}\sigma(\mathfrak{q})) = 1$. It follows that $\chi_\sigma = \chi^{-1}$. Since ℓ is odd and χ has order ℓ, $\chi \neq \chi^{-1}$. Hence $\chi \neq \chi_\sigma$, as required. This argument is taken from Serre [S, p.241]. Hence by Proposition 4.1.4, $\mathrm{Ind}_{K/\mathbb{Q}}\chi$ is irreducible. By Proposition 4.1.4 and Theorem 4.1.1, there is an associated cusp form $f \in S_1(\Gamma_1(N), \eta')$. Here $\eta' = \eta\chi_\mathbb{Q}$. Since χ takes values in ℓ-th roots of unity that are all congruent to 1 (mod λ), we see that the restriction $\chi_\mathbb{Q}$ is also trivial modulo λ. Hence η' reduces to η modulo λ. Recall that the Hecke eigenvalues of the Eisenstein series $E_1^{1,\eta}$ at a prime q, are given by

$$1 + \eta(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ 1 & \text{if } q \text{ ramifies in } \mathcal{O}_K, \\ 2 & \text{if } q \text{ splits in } \mathcal{O}_K. \end{cases}$$

Similarly the Hecke eigenvalue at a prime q for the cusp form f is given by

$$a_f(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ \chi(\mathfrak{q}) & \text{if } q = \mathfrak{q}^2 \text{ ramifies in } \mathcal{O}_K, \\ \chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) & \text{if } q = \mathfrak{q}_1\mathfrak{q}_2 \text{ splits in } \mathcal{O}_K. \end{cases}$$

Since the values of χ are ℓ-th roots of unity, which reduce to 1 (mod λ), we have that $a_q(f) \equiv 1 + \eta(q)$ (mod λ) for all primes q.

□

As we have previously mentioned, the appearance of h_K is related to $L(0, \eta)$, that is, we can explain this by looking at the Eisenstein side. In the higher weight cases, dividing the constant term of the Eisenstein series, i.e., dividing the L-value, gave a congruence between a level N Eisenstein series and a level N cusp form. It turns out that the same is true here. The constant term now is given by $L(0, \eta)$ or $L(0, \eta^{-1})$ depending on whether we are considering $E_1^{1,\eta}$ or $E_1^{\eta,1}$. But since η is a quadratic character, $L(0, \eta) = L(0, \eta^{-1})$. Since we are working with a quadratic field we have

$$L(0, \eta) = \frac{\zeta_K(0)}{\zeta(0)} = -2\zeta_K(0),$$

where ζ_K is the Dedekind zeta function. We can calculate this value using the following formula:

$$\lim_{s\to 0} s^{-r}\zeta_K(s) = -\frac{h_K.R(K)}{w(K)},$$

where $R(K)$ is the regulator of K, $w(K)$ is the number of roots of unity contained in K and r is the rank of the unit group. Since K is an imaginary quadratic field $R(K) = 1$ and $r = 0$. Hence

$$\lim_{s\to 0} s^{-0}\zeta_K(s) = \lim_{s\to 0} \zeta_K(s) = -\frac{h_K}{w(K)}.$$

We then have

$$L(0,\eta) = \begin{cases} \frac{h_K}{2} & \text{if } K = \mathbb{Q}(i), \\ \frac{h_K}{3} & \text{if } K = \mathbb{Q}(\sqrt{-3}), \\ h_K & \text{otherwise.} \end{cases}$$

This explains the appearance of the class number h_K on the Eisenstein side. Again since $\ell > 3$, the denominators in the first two cases will not cancel out any potential moduli.

We now consider the case $\ell|(p-1)$ for p a prime that splits in K.

Proposition 4.2.3. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field with discriminant N and associated quadratic character $\eta : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. Let p be a prime that splits in K such that $p \nmid N$. Let $\ell \nmid 6N$ be a prime such that $\ell|(p-1)$. Then there is a normalised Hecke eigenform $f \in S_1(\Gamma_1(Np), \eta')$ (where η' is congruent to η modulo λ) such that for all q with q prime,*

$$a_q(f) \equiv 1 + \eta(q) \pmod{\lambda},$$

where $\lambda|\ell$ is a prime of $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$.

Proof. We follow the same strategy as for the proof of Proposition 4.2.2. Our modulus is now $\mathfrak{m} = \mathfrak{p}$ where $\mathfrak{p} \in \mathcal{O}_K$ is a prime lying above p. We see from the formula in Theorem 4.2.1 that

$$\begin{aligned} h_\mathfrak{p} &= \frac{h_K.2^{r_0}.\mathrm{N}(\mathfrak{p}).\left(1 - \frac{1}{\mathrm{N}(\mathfrak{p})}\right)}{[U : U_{1.K}(\mathfrak{p})]} \\ &= \frac{h_K.(\mathrm{N}(\mathfrak{p}) - 1)}{[U : U_{1.K}(\mathfrak{p})]} \\ &= \frac{h_K.(p-1)}{[U : U_{1.K}(\mathfrak{p})]}. \end{aligned}$$

Since $\ell|(p-1)$, we see that ℓ divides the order of $I_K(\mathfrak{p})/P_{1,K}(\mathfrak{p})$. We therefore may choose $\chi : I_K(\mathfrak{p})/P_{1,K}(\mathfrak{p}) \to \mathbb{C}^\times$ of exact order ℓ. By Proposition 4.2.2 we assume the order does not divide the class number, i.e., it divides $p-1$ but not h_K. Recall

that $[U : U_{1,K}(\mathfrak{p})] = 1, 2, 3, 4$ or 6 and so does not cancel out any potential moduli (since $\ell \nmid 6N$). We would like to show that $\chi \neq \chi_\sigma$. The character χ factors through a quotient of the ray class group, namely $P_K(\mathfrak{p})/P_{1,K}(\mathfrak{p}) \simeq (\mathcal{O}_K/\mathfrak{p})^\times \simeq \mathbb{F}_p^\times$ since p splits as say $\mathfrak{p}\overline{\mathfrak{p}}$. We now consider the action of σ on this quotient. Recall that we have $\mathrm{Gal}(K/\mathbb{Q}) = \{1, \sigma\}$. It is clear that σ maps $\mathfrak{p}$ to $\overline{\mathfrak{p}}$, therefore it maps $(\mathcal{O}_K/\mathfrak{p})^\times$ to $(\mathcal{O}_K/\overline{\mathfrak{p}})^\times$. It follows that χ and χ_σ are characters that factor through different quotients, and have different conductors (one has conductor $\mathfrak{p}$, the other $\overline{\mathfrak{p}}$). Hence it follows that $\chi \neq \chi_\sigma$. Therefore $\mathrm{Ind}_{K/\mathbb{Q}}(\chi)$ is irreducible. By Proposition 4.1.4 and Theorem 4.1.1, there is an associated cusp form $f \in S_1(\Gamma_1(Np), \eta')$. Here $\eta' = \eta\chi_\mathbb{Q}$. Since χ takes values in ℓ-th roots of unity that are all congruent to 1 (mod λ), we see that the restriction $\chi_\mathbb{Q}$ is also trivial modulo λ. Hence η' reduces to η modulo λ. (Note that $\mathfrak{m} = \mathfrak{p}$ here and $\mathrm{N}(\mathfrak{p}) = p$). Recall that the Hecke eigenvalues of the Eisenstein series $E_1^{1,\eta}$ at a prime q, are given by

$$1 + \eta(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ 1 & \text{if } q \text{ ramifies in } \mathcal{O}_K, \\ 2 & \text{if } q \text{ splits in } \mathcal{O}_K. \end{cases}$$

Similarly the Hecke eigenvalue at a prime q for the cusp form f is given by

$$a_f(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ \chi(\mathfrak{q}) & \text{if } q = \mathfrak{q}^2 \text{ ramifies in } \mathcal{O}_K, \\ \chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) & \text{if } q = \mathfrak{q}_1\mathfrak{q}_2 \text{ splits in } \mathcal{O}_K. \end{cases}$$

Since the values of χ are ℓ-th roots of unity, which reduce to 1 (mod λ), we have that $a_q(f) \equiv 1 + \eta(q) \pmod{\lambda}$ for all primes q. □

We now consider the case $\ell|(p+1)$ for p a prime that is inert in K.

Proposition 4.2.4. *Let $K = \mathbb{Q}(\sqrt{-d})$ be an imaginary quadratic field with discriminant N and associated quadratic character $\eta : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. Let p be a prime that is inert in K such that $p \nmid N$. Let $\ell \nmid 6N$ be a prime such that $\ell|(p+1)$. Then there is a normalised Hecke eigenform $f \in S_1(\Gamma_1(Np^2), \eta')$ (where η' is congruent to η modulo λ) such that for all q with q prime,*

$$a_q(f) \equiv 1 + \eta(q) \pmod{\lambda},$$

where $\lambda|\ell$ is a prime of $\mathbb{Q}(\{a_n(f)\}) = \mathbb{Q}(\zeta_\ell)$.

Proof. We follow the same strategy as for the proof of Proposition 4.2.2 and Proposition 4.2.3. Our modulus is now $\mathfrak{m} = \mathfrak{p}$ where $\mathfrak{p} \in \mathcal{O}_K$ is a prime lying above p. We

see from the formula in Theorem 4.2.1 that

$$\begin{aligned} h_{\mathfrak{p}} &= \frac{h_K.2^{r_0}.\mathrm{N}(\mathfrak{p}).\left(1-\frac{1}{\mathrm{N}(\mathfrak{p})}\right)}{[U:U_{1,K}(\mathfrak{p})]} \\ &= \frac{h_K.(\mathrm{N}(\mathfrak{p})-1)}{[U:U_{1,K}(\mathfrak{p})]} \\ &= \frac{h_K.(p^2-1)}{[U:U_{1,K}(\mathfrak{p})]} \\ &= \frac{h_K.(p-1)(p+1)}{[U:U_{1,K}(\mathfrak{p})]}. \end{aligned}$$

Since $\ell|(p+1)$ with $\ell > 3$, we see that ℓ divides the order of $I_K(\mathfrak{p})/P_{1,K}(\mathfrak{p})$. We therefore may choose $\chi : I_K(\mathfrak{p})/P_{1,K}(\mathfrak{p}) \to \mathbb{C}^\times$ of exact order ℓ. By Proposition 4.2.2 we may assume the order does not divide the class number, i.e., it divides $p+1$ but not h_K. Recall that $[U:U_{1,K}(\mathfrak{p})] = 1, 2, 3, 4$ or 6 and so does not cancel out any potential moduli (since $\ell \nmid 6N$). We would like to show that $\chi \neq \chi_\sigma$. The character χ again factors through the same quotient as in the proof of Proposition 4.2.3 but the modulus has changed giving us a different finite field. We now factor through $P_K(\mathfrak{p})/P_{1,K}(\mathfrak{p}) \simeq (\mathcal{O}_K/\mathfrak{p})^\times \simeq \mathbb{F}_{p^2}^\times$. Again we consider the action of σ on this quotient. Since $\mathfrak{p}$ is now an inert prime, this quotient is fixed under the action of σ, namely it acts as a non-trivial automorphism on the quotient. The only non-trivial automorphism of $\mathbb{F}_{p^2}^\times$ is the p-th power Frobenius map. Hence if we consider some ideal $\mathfrak{q} \in P_K(\mathfrak{p})/P_{1,K}(\mathfrak{p})$ we have $\chi(\sigma(\mathfrak{q})) = \chi(\mathfrak{q}^p)$. Hence $\chi(\mathfrak{q}) = \chi(\sigma(\mathfrak{q})) \Rightarrow \chi(\mathfrak{q}) = \chi(\mathfrak{q}^p)$. That is $\chi(\mathfrak{q})^{p-1} = 1$. Hence $\chi = \chi_\sigma$ if and only if ℓ is a prime dividing $(p-1)$. Since $\ell|(p+1)$ we see that $\chi \neq \chi_\sigma$. Hence $\mathrm{Ind}_{K/\mathbb{Q}}\chi$ is irreducible. By Proposition 4.1.4 and Theorem 4.1.1, there is an associated cusp form $f \in S_1(\Gamma_1(Np^2), \eta')$. Here $\eta' = \eta\chi_{\mathbb{Q}}$. Since χ takes values in ℓ-th roots of unity that are all congruent to $1 \pmod{\lambda}$, we see that the restriction $\chi_{\mathbb{Q}}$ is also trivial modulo λ. Hence η' reduces to η modulo λ. (Note that $\mathfrak{m} = \mathfrak{p}$ here and $\mathrm{N}(\mathfrak{p}) = p^2$). Recall that the Hecke eigenvalues of the Eisenstein series $E_1^{1,\eta}$ at a prime q, are given by

$$1 + \eta(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ 1 & \text{if } q \text{ ramifies in } \mathcal{O}_K, \\ 2 & \text{if } q \text{ splits in } \mathcal{O}_K. \end{cases}$$

Similarly the Hecke eigenvalue at a prime q for the cusp form f is given by

$$a_f(q) = \begin{cases} 0 & \text{if } q \text{ is inert in } \mathcal{O}_K, \\ \chi(\mathfrak{q}) & \text{if } q = \mathfrak{q}^2 \text{ ramifies in } \mathcal{O}_K, \\ \chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) & \text{if } q = \mathfrak{q}_1\mathfrak{q}_2 \text{ splits in } \mathcal{O}_K. \end{cases}$$

Since the values of χ are ℓ-th roots of unity, which reduce to $1 \pmod{\lambda}$, we have that $a_q(f) \equiv 1 + \eta(q) \pmod{\lambda}$ for all primes q. $\square$

Remark 4.2.5. Note that in the case where p is an inert prime, the Euler factor is $p+1$ but the order of the ray class group has $(p^2-1)=(p+1)(p-1)$ in the numerator. We would therefore expect that dividing $p+1$ should give a congruence when considering the Eisenstein side, and, as we have shown, this is indeed true. We could have asked what would happen if $\ell|(p-1)$ instead. In this case ℓ would still have divided the order of the ray class group but not the Euler factor. Hence we shouldn't have expected a congruence. This is verified by the fact that $\chi=\chi_\sigma$ in the case where $\ell|(p-1)$. That is $\mathrm{Ind}_{K/\mathbb{Q}}\chi$ is not irreducible in this case.

We have now proved Theorem 4.0.1. Our next task will be to prove the generalisation of this result where we no longer restrict ourselves to having a trivial character.

4.2.2 The Generalisation of Theorem 4.0.1

We now no longer want to assume that one of our characters is trivial and the other quadratic. Suppose we have an Eisenstein series $E_1^{\psi,\varphi}\in M_1(\Gamma_1(N),\psi\varphi)$. Here suppose ψ is primitive with conductor u, φ is primitive with conductor v and $\psi\varphi$ has conductor $N=uv$. We know by Proposition 4.1.4, that a dihedral representation comes from the induction of a character χ of $G_K=\mathrm{Gal}(\bar{\mathbb{Q}}/K)$ where K is a quadratic field. Associated to this dihedral representation is a cusp form f. We will first show that in order for a congruence to exist we must have $\psi\equiv\eta\varphi \pmod{\lambda}$ where η is the character associated to the quadratic field K.

Proposition 4.2.6. *Let ψ and φ be as above. Let K be an imaginary quadratic field of discriminant N with associated quadratic character $\eta:(\mathbb{Z}/N\mathbb{Z})^\times\to\mathbb{C}^\times$. Let χ be a ray class character for K and f_χ the associated cusp form. Then*

$$a_q(f_\chi)\equiv\psi(q)+\varphi(q)\pmod{\lambda},$$

where $\lambda|\ell$ is a prime of $\mathbb{Q}(\{a_n(f_\chi)\})$ if and only if $\psi\equiv\eta\varphi \pmod{\lambda}$.

Proof. Recall that for an inert prime q in K, the Hecke eigenvalue of f_χ is 0. We therefore would need $\psi(q)+\varphi(q)\equiv 0 \pmod{\lambda}$ at such a prime. We immediately see that we must have $\psi(q)\equiv-\varphi(q) \pmod{\lambda}$. Hence we must have $\psi=\alpha\varphi$ for some α such that $\alpha(q)\equiv-1 \pmod{\lambda}$ for all inert primes q. We may define $\alpha:(\mathbb{Z}/N\mathbb{Z})^\times\to\mathbb{C}^\times$ by $\alpha=\frac{\psi}{\varphi}$ (where ψ and φ are viewed as functions on $(\mathbb{Z}/N\mathbb{Z})^\times$). Hence α is a Dirichlet character modulo N.

Suppose we consider the reduction $\overline{\alpha}$ viewed as a character $\overline{\alpha}:G_{\mathbb{Q}}\to\overline{\mathbb{F}}_\ell^\times$. For any $\tau\in G_{\mathbb{Q}}\backslash G_K$ we have $\overline{\alpha}(\tau)=-1$. For $\sigma\in G_K$ we have $\sigma=(\sigma\tau^{-1})\tau$. Hence $\overline{\alpha}(\sigma)=\overline{\alpha}(\sigma\tau^{-1})\overline{\alpha}(\tau)=(-1)(-1)=1$. Hence we see that $\overline{\alpha}$ is a character which takes the values ±1. Also since the character factors through the quotient $G_{\mathbb{Q}}/G_K=\mathrm{Gal}(K/\mathbb{Q})$ which is cyclic of order 2, we see that $\overline{\alpha}$ is in fact quadratic. In particular it is exactly the quadratic character η associated to K. Hence we deduce that $\psi\equiv\eta\varphi \pmod{\lambda}$. □

Here we have proven that $\psi \equiv \eta\varphi \pmod{\lambda}$ is a necessary condition. We will now try and construct a cusp form satisfying a congruence in a similar manner to the method used to prove Theorem 4.0.1 using this condition. That is, we will prove that the condition is sufficient. Rather than give separate results with proofs as we did when proving Theorem 4.0.1, we simply give a sketch proof detailing the differences. We first investigate the new Euler factors. Associated to the Eisenstein series $E_1^{\psi,\varphi}$ is the Euler factor $\varphi(p)p - \psi(p)$. We may consider the Euler factor modulo λ.

Suppose first that p is an inert prime in $\mathcal{O}_K$. Then $\eta(p) = -1$ and we see that $\psi(p) \equiv -\varphi(p) \pmod{\lambda}$. Hence

$$\begin{aligned}\varphi(p)p - \psi(p) &\equiv \varphi(p)p + \varphi(p) \pmod{\lambda} \\ &\equiv \varphi(p)(p+1) \pmod{\lambda}.\end{aligned}$$

Suppose instead we consider a prime p that splits in $\mathcal{O}_K$. We now have $\eta(p) = 1$ hence $\psi(p) \equiv \varphi(p) \pmod{\lambda}$. Hence

$$\begin{aligned}\varphi(p)p - \psi(p) &\equiv \varphi(p)p - \varphi(p) \pmod{\lambda} \\ &\equiv \varphi(p)(p-1).\end{aligned}$$

Note that in both cases we see $(p+1)$ and $(p-1)$ appearing as before. We therefore see that ℓ divides $(p+1)$ in the inert case and $(p-1)$ in the split case since $\lambda|\ell$. We could also consider the Eisenstein series $E_1^{\varphi,\psi}$. However since ψ and φ are related via a congruence condition, we will end up with the same factors. We would therefore expect that dividing $p-1$ in the split case or dividing $p+1$ in the inert case should give a congruence just as in Theorem 4.0.1. Whether there is a congruence will depend on our choice of ray class character. It will turn out that we only need a slight modification.

If we again consider a potential congruence, now using what we know about the characters ψ and φ, we will be able to determine which ray class character we will need to use. In the proof of Proposition 4.2.6, we considered a potential congruence at inert primes in order to determine conditions for the Dirichlet characters ψ and φ. We now want to consider the split primes. For a ray class character χ (i.e. a character of G_K) we want

$$\psi(q) + \varphi(q) \equiv \chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) \pmod{\lambda} \quad \text{if } q = \mathfrak{q}_1\mathfrak{q}_2 \text{ splits in } \mathcal{O}_K.$$

We already know that $\psi \equiv \eta\varphi \pmod{\lambda}$. Hence at a split prime q we have $\psi(q) \equiv \eta(q)\varphi(q) \pmod{\lambda} \equiv \varphi(q) \pmod{\lambda}$. We therefore need

$$2\varphi(q) \equiv \chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) \pmod{\lambda} \quad \text{if } q = \mathfrak{q}_1\mathfrak{q}_2 \text{ splits in } \mathcal{O}_K.$$

If we take $\chi = \chi_1\chi_2$, where χ_1 is the ray class character we use in Theorem 4.0.1, which depends on which particular case we are in, we then need to determine χ_2. Note that in all cases $\chi_1 \equiv 1 \pmod{\lambda}$. We note that before we simply had 2 on the LHS of the congruence so it was enough to guarantee that the ℓ-th roots of unity were congruent

to 1 (mod λ) in order for the congruence to be satisfied. However, since φ is arbitrary, we now have more possibilities for the LHS. If we take $\chi_2 = \varphi \circ N_{K/\mathbb{Q}}$, where $N_{K/\mathbb{Q}}$ is the norm map, then we see that the congruence will be satisfied. This follows since

$$\begin{aligned}\chi(\mathfrak{q}_1) + \chi(\mathfrak{q}_2) &= \chi_1(\mathfrak{q}_1)\chi_2(\mathfrak{q}_1) + \chi_1(\mathfrak{q}_2)\chi_2(\mathfrak{q}_2) \\ &\equiv (\varphi \circ N_{K/\mathbb{Q}})(\mathfrak{q}_1) + (\varphi \circ N_{K/\mathbb{Q}})(\mathfrak{q}_2) \pmod{\lambda} \\ &\equiv 2\varphi(q) \pmod{\lambda}.\end{aligned}$$

Here the second line follows since $\chi_1(\mathfrak{q}_1) \equiv \chi_1(\mathfrak{q}_2) \equiv 1 \pmod{\lambda}$ and the third follows since $N_{K/\mathbb{Q}}(\mathfrak{q}_1) = N_{K/\mathbb{Q}}(\mathfrak{q}_2) = q$. The conductor of χ_2 will depend on the conductor of φ. Assume the conductor of φ is given by $v = p_1 \dots p_n$. Since $v|N$, these are ramified primes in K. In particular v factors as $\mathfrak{p}_1^2 \dots \mathfrak{p}_n^2$ in $\mathcal{O}_K$. The conductor of χ_2 will then be $\mathfrak{v} = \mathfrak{p}_1 \dots \mathfrak{p}_n$. Taking the norm of any factor we see that $N_{K/\mathbb{Q}}(\mathfrak{p}_i) = p_i$. This explains why we are adding a factor of v to the level of the cusp form in Theorem 4.0.2 since $N_{K/\mathbb{Q}}(\mathfrak{v}) = p_1 \dots p_n$.

We can also consider the congruence at the ramified primes. Since $uv = N$, the ramified primes are exactly those appearing in the prime factorisation of u and v. We again use the extended definition of a Dirichlet character. That is we have $\psi : (\mathbb{Z}/u\mathbb{Z}) \to \mathbb{C}^\times$ and $\varphi : (\mathbb{Z}/v\mathbb{Z}) \to \mathbb{C}^\times$. For any non-invertible element $q \in (\mathbb{Z}/u\mathbb{Z})$ we have $\psi(q) = 0$ and for any non-invertible element $q \in (\mathbb{Z}/v\mathbb{Z})$ we have $\varphi(q) = 0$. We want to satisfy

$$\psi(q) + \varphi(q) \equiv \chi(\mathfrak{q}) \pmod{\lambda} \quad \text{if } q = \mathfrak{q}^2 \text{ ramifies in } \mathcal{O}_K.$$

Suppose first that $q|u$. We have $\psi(q) = 0$ but $\varphi(q)$ depends on whether $q|v$ We therefore need to satisfy

$$\varphi(q) \equiv \chi(\mathfrak{q}) \pmod{\lambda}.$$

Since $\chi_1(\mathfrak{q}) \equiv 1 \pmod{\lambda}$ and $N_{K/\mathbb{Q}}(\mathfrak{q}) = q$, so $\chi_2(\mathfrak{q}) = \varphi(q)$, we see that the congruence is satisfied for such a ramified prime regardless of whether $q|v$.

Now suppose $q|v$. We have $\varphi(q) = 0$ and $\psi(q)$ depends on whether $q|u$. By the argument above the RHS of the congruence will be congruent to $\varphi(q) \pmod{\lambda}$. Hence the RHS is congruent to $0 \pmod{\lambda}$. We therefore need

$$\psi(q) \equiv 0 \pmod{\lambda}.$$

However we know that $\psi(q) \equiv \eta(q)\varphi(q) \pmod{\lambda} \equiv 0 \pmod{\lambda}$. It follows that the congruence holds for these primes as well.

In conclusion we see that the only modification we have made is to the ray class character χ. We have simply multiplied by another ray class character $\chi_2 = \varphi \circ N_{K/\mathbb{Q}}$ with conductor $\mathfrak{v} = \mathfrak{p}_1 \dots \mathfrak{p}_n$. This has the effect of raising the level of the cusp form f in each case by $v = p_1 \dots p_n$ since $N_{K/\mathbb{Q}}(\mathfrak{v}) = v$. Note that this choice is arbitrary, we could just have easily had $\chi_2 = \psi \circ N_{K/\mathbb{Q}}$ due to the fact that $\varphi \equiv \eta\psi \pmod{\lambda}$ as well since η is quadratic. This would instead have the effect of raising the level by u. We note that in Theorem 4.0.1, χ_2 was trivial since we had a trivial character (which

we could choose to be φ arbitrarily and so $v = 1$ in this case). This choice of χ_2 for Theorem 4.0.1 was enough because we had already chosen a ray class character that would guarantee the congruence held mod λ. We also note that in each case we still obtain the same conditions on when $\chi = \chi_\sigma$. This is because we can simply consider what is happening to χ_1. This follows since we have $\chi_2 = \chi_{2_\sigma}$ (This can be seen by considering the norms of ideals in $\mathcal{O}_K$). If $\chi_1 \neq \chi_{1\sigma}$, then we see that $\chi \neq \chi_\sigma$. We also note that modulo λ the ray class character χ reduces to $\chi_2 = \varphi \circ \mathrm{N}_{K/\mathbb{Q}}$. It follows that in this case $\chi_{\mathbb{Q}} \equiv \varphi^2 \pmod{\lambda}$. Hence the character of the induced representation, and therefore the character of the cusp form, will be $\epsilon = \eta\chi_{\mathbb{Q}} \equiv \eta\varphi^2 \pmod{\lambda} \equiv \psi\varphi \pmod{\lambda}$. Other than these changes we could state results similar to Propositions 4.2.2, 4.2.3 and 4.2.4 along with proofs following largely the same steps. We have therefore given a (sketch) proof of Theorem 4.0.2.

Although we will not cover the details of the Bloch-Kato conjecture in the case of weight 1, we note that the exact same argument as in Section 3.3 still applies in this case. Although the standard results of modular forms break down in the case of weight 1, so in some sense the theory is harder, the Bloch-Kato conjecture is actually more straightforward. Here the conjecture is basically a consequence of Dirichlet's analytic class number formula.

In order to prove Theorem 4.0.3, we will need to see some preliminaries on lifting projective representations.

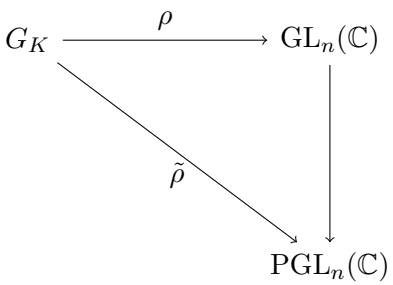
4.2.3 Liftings of Projective Representations

Let K be a global or local field. We assume throughout that our non-Archimedean local fields have finite residue field. Let $\bar{K}/K$ be a separable closure of K, and let $G_K = \mathrm{Gal}(\bar{K}/K)$. Let $\tilde{\rho}$ be a projective representation of G_K:

$$\tilde{\rho} : G_K \to \mathrm{PGL}_n(\mathbb{C}) = \mathrm{GL}_n(\mathbb{C})/\mathbb{C}^\times.$$

We will assume throughout that all representations of G_K are continuous.

Definition 4.2.7. A *lifting* of $\tilde{\rho}$ is a (continuous) linear representation $\rho : G_K \to \mathrm{GL}_n(\mathbb{C})$ such that the diagram



commutes.

Note that although we have given a general definition here, we will only be interested in the case $K = \mathbb{Q}$ and $n = 2$.

If ρ is a lifting of $\tilde{\rho}$, then so is a twist by any one dimensional linear representation χ of G_K, i.e., $\chi \otimes \rho$ is a lifting of $\tilde{\rho}$; further any lifting of $\tilde{\rho}$ is of this form, for some χ.

We would like to know when a particular projective representation can be lifted. It turns out that this is related to cohomology. We give the basic definitions that we will need. In the following G is a group with A a G-module.

Definition 4.2.8. (1) The group of *i-cochains* of G with coefficients in A is the set of functions from $G^i \to A$:

$$C^i(G, A) = \{f : G^i \to A\}.$$

(2) The *i-th differential* $d^i = d^i_A : C^i(G, A) \to C^{i+1}(G, A)$ is the map

$$\begin{aligned} d^i(f)(g_0, g_1, \ldots, g_i) &= g_0.f(g_1, \ldots, g_i) \\ &+ \sum_{j=1}^{i} (-1)^j f(g_0, \ldots, g_{j-2}, g_{j-1}g_j, g_{j+1}, \ldots, g_i) + (-1)^{i+1} f(g_0, \ldots, g_{i-1}). \end{aligned}$$

Definition 4.2.9. (1) We set $Z^i(G, A) = \ker d^i$, the group of *i-cocylces* of G with coefficients in A

(2) We set $B^0(G, A) = 0$ and $B^i(G, A) = \mathrm{im} d^{i-1}$ for $i \geq 1$. We refer to $B^i(G, A)$ as the group of *i-coboundaries* of G with coefficients in A.

Definition 4.2.10. We define the *i-th cohomology group* of G with coefficients in A to be

$$H^i(G, A) = Z^i(G, A)/B^i(G, A).$$

These are the definitions for group cohomology. There are some technical differences when working with profinite groups (as we will be doing) but the basic ideas still work. Refer to [Sha] for the technical aspects of Galois cohomology. We may consider $\mathbb{C}^\times$ as a G_K module, on which G_K acts trivially. Let $H^2(G_K, \mathbb{C}^\times)$ denote the 2-cohomology group of the profinite group G_K with coefficients in $\mathbb{C}^\times$. Suppose for each $g \in G_K$ we fix a lifted element $P(g)$ where we have lifted from $\mathrm{PGL}_2(\mathbb{C})$ to $\mathrm{GL}_2(\mathbb{C})$. The lifts then satisfy

$$P(gh) = \alpha(g, h)P(g)P(h)$$

for some $\alpha(g, h) \in \mathbb{C}^\times$. In fact the map $\alpha : G_K \times G_K \to \mathbb{C}^\times$ is a 2-cocycle and satisfies the relation

$$\alpha(g, hk)\alpha(h, k) = \alpha(g, h)\alpha(gh, k)$$

for all $g, h, k \in G_K$. The cocycle α depends on the choice of lift P; a different choice $Q(g) = \delta(g)P(g)$ will result in a different cocycle

$$\beta(g, h) = \delta(gh)\delta^{-1}(g)\delta^{-1}(h)\alpha(g, h).$$

Hence P defines a unique class in $H^2(G_K, \mathbb{C}^\times)$. If $H^2(G_K, \mathbb{C}^\times)$ is non-trivial, this may prevent the lifting process as it leads to an extension problem. Thankfully this is not an issue in our case.

Theorem 4.2.11 (Tate)**.** *Let K be a global or local field. Then $H^2(G_K, \mathbb{C}^\times) = 1$.*

Corollary 4.2.12. *Every projective representation of G_K has a lifting.*

For a proof of Theorem 4.2.11 see Section 6.5 of [S].

4.2.4 Proof of Theorem 4.0.3

In this section we will prove Theorem 4.0.3, that is, we will show that there are no other congruences in the weight 1 case. Note that we have already covered all possible congruences in the dihedral case in proving Theorem 4.0.2. We now consider the remaining cases where our projective Galois representation $\tilde{\rho}$ has image isomorphic to one of A_4, S_4 or A_5. Since representation theory works much the same way in characteristic p for p a prime unless p divides the order of the group, we expect there to be only a handful of possible moduli for congruences. Namely 2 and 3 in the cases of A_4 and S_4 and $2, 3$ and 5 for A_5. Note that earlier, we avoided 2 and 3 as they were considered small primes. Here we won't consider 2, but we will be interested in potential congruences modulo 3 or 5.

Our first task is to consider the 2-dimensional, odd, irreducible (in characteristic 0) projective representations of A_4, S_4 and A_5. It turns out that it is quite hard to come up with these representations and in fact it is easier to consider certain linear representations that are in one-to-one correspondence with these projective representations. This leads us to the notion of a Schur cover.

In the early 20th century, Issai Schur began studying projective representations, developing the theory for finite groups. He wrote two papers, one in 1904, one in 1907, laying the foundations of the theory. In his 1911 paper [Schur] he applied the work he had been doing to the case of the symmetric and alternating groups. The basic idea is that each of the symmetric and alternating groups have a corresponding covering group of which linear representations give you the projective representations you want to know about.

4.2.5 Schur multipliers and Schur Covers

In this section we cover the basic material needed in order to study the projective representations of A_4, S_4 and A_5. The following background material is from [HoHu]. In order to make the definition of the Schur multiplier more intuitive we first restate our definition of a projective representation.

Definition 4.2.13. Let V be a finite dimensional complex vector space. A (complex) projective representation of a group G on V is a function P from G into $\mathrm{GL}(V)$, the group of automorphisms of V, such that

(1) $P(1_G)$ is the identity linear transformation of V; and

(2) given elements x and y in G, there is a non-zero complex number $\alpha(x,y)$ such that
$$P(x)P(y) = \alpha(x,y)P(xy).$$

Suppose the dimension of V is d. If we choose a basis for V, we obtain a projective matrix representation $P : G \to \mathrm{GL}_d(\mathbb{C})$.

Each linear transformation $P(g)$ is invertible. It therefore follows from (1) that, for all g in G,
$$\alpha(g,1) = 1 = \alpha(1,g). \tag{4.2}$$

Using associativity of composition and of group multiplication to evaluate $P(x)P(y)P(z)$ gives that
$$\alpha(x,yz)\alpha(y,z) = \alpha(x,y)\alpha(xy,z) \tag{4.3}$$

for all x, y and z in G. Any map $\alpha : G \times G \to \mathbb{C}^\times := \mathbb{C}\backslash\{0\}$ satisfying conditions (4.2) and (4.3) is said to be a 2-cocycle. Notice that this is the same relation that came up when we were considering cohomology.

Definition 4.2.14. A linear representation of a group G on a finite dimensional vector space V is a homomorphism $R : G \to \mathrm{GL}(V)$. Thus a linear representation is a projective representation with trivial 2-cocycle.

It is clear that this notion is another way of viewing the familiar notion that two elements u, v in a projective space $P(V)$ are equivalent under the relation $u \sim v$ if and only if $u = \lambda v$ for some $\lambda \in \mathbb{C}^\times$. An element f in $\mathrm{GL}(V)$ induces an action f^* on $P(V)$ by
$$f^*[v] = [f(v)].$$

Thus f^* may be regarded as an element of $\mathrm{PGL}(V) = \mathrm{GL}(V)/\mathbb{C}^\times I$ where I is the identity matrix. In matrix terms, suppose V has dimension d, let $Z(d)$ be the set of non-zero multiples of the identity matrix, and define $\mathrm{PGL}_d(\mathbb{C})$ to be the quotient group $\mathrm{GL}_d(\mathbb{C})/Z(d)$. The isomorphic groups $\mathrm{PGL}(V)$ and $\mathrm{PGL}_d(\mathbb{C})$ are known as projective linear groups. Given a projective representation P of degree d for the group G, the map $P' : G \to \mathrm{PGL}_d(\mathbb{C})$ which takes an element g of G to the coset $P(g)Z(d)$ is a homomorphism. Conversely, any such homomorphism P' gives a projective representation in the original sense, by choosing a representative $P(g)$ for the coset $P'(g)$ (with the convention that P(1) = I). The homomorphism P' determines a 2-cocycle

α by this process, but the association to P' of P and α is not unique. If we make a different choice $Q(g)$ for $P(g)$, then

$$Q(g) = \delta(g)P(g)$$

for all g in G, where $\delta(1) = 1$ and $\delta(g)$ is in $\mathbb{C}^\times$. If β is the 2-cocycle associated with Q, it is related to α by the rule:

$$\beta(x, y) = \delta(x)\delta(y)(\delta(xy))^{-1}\alpha(x, y)$$

for all x and y in G. Two 2-cocycles related in this way are said to be cohomologous. Denote the cohomology class of a 2-cocycle α by $[\alpha]$. The set of such classes forms an abelian group, the Schur multiplier, denoted $H^2(G, \mathbb{C}^\times)$ or $M(G)$, under the operation

$$[\alpha][\beta] = [\alpha\beta],$$

where

$$(\alpha\beta)(x, y) = \alpha(x, y)\beta(x, y)$$

for all x and y in G.

Hopefully it is clear that the Schur multiplier is the second cohomology group of G with coefficients in $\mathbb{C}^\times$ where we consider $\mathbb{C}^\times$ as a G-module with trivial G-action. It would be nice if we knew a little more about the structure of the Schur multiplier. Luckily the following theorem addresses this very issue.

Theorem 4.2.15. *For any finite group G, the Schur multiplier has finite exponent dividing the order of G. Furthermore, if a cohomology class has order e, then there is a representative of that class which takes only e-th roots of unity as its values. Thus $M(G)$ is a finite group.*

A proof of this result is given on pages 3 and 4 of [HoHu]. We can now produce a central extension of $M(G)$ by G containing $M(G)$ as its commutator subgroup. In other words we have a central extension

$$1 \to A \to C \to G \to 1,$$

where C is known as a representation group or Schur covering group of G.

Definition 4.2.16. A Schur covering group for G is any group C satisfying the following conditions:

(1) C has central subgroup A contained in the commutator subgroup of C.

(2) $C/A \cong G$.

(3) $A \cong M(G)$.

Such a group always exists and can be constructed in a particular way as is proven in [HoHu, Theorem 1.2, p.5]. One thing to note from this construction is that we choose a set of symbols $r(g)$ in bijective correspondence with the elements of G. Now that we have these covering groups we need to find a way to relate these to projective representations. The next theorem will address this issue. Note that this result relies on the choice of symbols $r(g)$.

Theorem 4.2.17. *Let C be a Schur covering group of G. Given a projective representation P of G, there is a function $\delta : G \to \mathbb{C}^\times$ and a linear representation R of C such that, for all g in G,*

$$P(g) = \delta(g)R(r(g)).$$

Proof. See pages 7 and 8 of [HoHu]. □

We also have a converse to this result.

Theorem 4.2.18. *Let C be a Schur covering group for G and let $\lambda : A \to \mathbb{C}^\times$ be any homomorphism with A as above. Suppose that R is a linear representation of C such that $R(a) = \lambda(a)I$ for each a in A. Define P by $P(g) = R(r(g))$ for all g in G. Then P is a projective representation whose associated cocycle is α, where*

$$\alpha(x, y) = \lambda\left(\Phi(x, y)\right)$$

for all x and y in G.

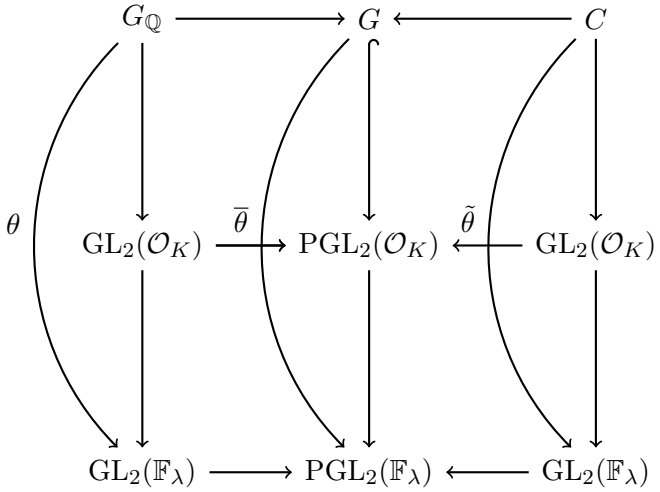
Proof. The proof of the previous result can simply be reversed. □

The definition of Φ here is not too important (it appears in Theorem 1.2 of [HoHu]). The most important thing is that we now have a bijective correspondence between projective representations of a group G and linear representations of its Schur covering group. This will allow us to more easily study the cases we are interested in.

4.2.6 Completing the Proof

We are now in a position to prove Theorem 4.0.3. We first show that considering the linear representation associated to the Schur cover is in fact equivalent to considering

the projective representation. Consider the following diagram:



If we had a congruence ($\ell > 2$) this would imply that ψ and φ would be the composition factors of the residual representation in the lower left of the above diagram. We would therefore have a line invariant under the GL_2 action. This invariance would still hold in PGL_2. It follows that this line must also be invariant in the lower right of the above diagram. In other words, the existence of a congruence implies that there exists a line $L \subset \mathbb{F}_\lambda^2$, which is a 1-dimensional subspace, satisfying

$$\theta(g)L = L \quad \text{for all } g \in G_{\mathbb{Q}} \Rightarrow \overline{\theta}(g)L = L \quad \text{for all } g \in G \Rightarrow \tilde{\theta}(g)L = L \quad \text{for all } g \in C.$$

It therefore follows that we may consider reducibility of the linear representation associated to the Schur cover in order to determine if a congruence exists.

We need to know what the Schur multipliers and Schur covers are for each of A_4, S_4 and A_5. It turns out that each of these groups has the same Schur multiplier, namely $\mathbb{Z}/2\mathbb{Z}$. We then have the following Schur coverings:

$$1 \to \mathbb{Z}/2\mathbb{Z} \to \mathrm{SL}(2,3) \to \mathrm{PSL}(2,3) \cong A_4 \to 1$$
$$1 \to \mathbb{Z}/2\mathbb{Z} \to \mathrm{GL}(2,3) \to \mathrm{PGL}(2,3) \cong S_4 \to 1$$
$$1 \to \mathbb{Z}/2\mathbb{Z} \to \mathrm{SL}(2,5) \to \mathrm{PSL}(2,5) \cong A_5 \to 1$$

Note that there is a second non-isomorphic Schur covering group for S_4 but we need only work with one of them so we ignore this extra group. We therefore now need to know about the irreducible 2-dimensional linear representations of $\mathrm{SL}(2,3), \mathrm{GL}(2,3)$ and $\mathrm{SL}(2,5)$. Character tables for each of these groups can be found online. We present these below.

The character table given in Table 4.1 is for $\mathrm{SL}(2,3)$ and can be found online at [2]. This group has size 24 and therefore we should only expect there to be potential congruences modulo 2 or 3. Since we are looking for potential congruences we simply need

to consider the traces (i.e. character values) modulo 2 and 3 of the two-dimensional representations compared to a sum of two one-dimensional representations. This is simply because if a congruence exists, then the irreducible two-dimensional representation will reduce to a sum of two characters (one-dimensional representations) and the traces would have to match modulo p. Suppose we call the one-dimensional representations χ_1, χ_2 and χ_3 respectively. We may then form a table comparing the traces of sums of characters against the traces of the two-dimensional representations. These are given in Table 4.2.

It is fairly clear that there can be no possible congruence modulo 3 here. Also note that we are avoiding congruences modulo 2. So we conclude that there can be no congruences when our projective image is A_4.

The character table given in Table 4.3 is for $\mathrm{GL}(2,3)$ and can be found online at [3]. We now use the same process as before. Let the one-dimensional representations be χ_1 and χ_2. The comparisons of traces are given in Table 4.4.

Again we should only be expecting potential congruences modulo 2 or 3 here. However it is clear, looking at the $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ column, that the two-dimensional representations cannot match the possible traces modulo 2 (even though we are avoiding this anyway). It is also fairly clear that the only match modulo 3 is between the second row $(\chi_1 + \chi_2)$ and the fourth row (the first two-dimensional representation). Therefore this two-dimensional representation reduces to $\begin{pmatrix} \chi_1 & * \\ 0 & \chi_2 \end{pmatrix}$ modulo 3. However if we compare the character values of this particular two-dimensional representation with the unique two-dimensional representation of S_4, we notice that these values match up. In particular, this representation must factor through S_4. Now S_4 has a normal subgroup of order 4, namely the Klein four-group. If we take the quotient of S_4 by this subgroup we get S_3 and this particular two-dimensional representation descends to a faithful representation of this group. Hence the two-dimensional representation of $\mathrm{GL}(2,3)$ that we are interested in must factor through S_3. We therefore cannot possibly have full projective image S_4. Hence we conclude that there are no possible congruences when the projective image is S_4.

It turns out that the only one-dimensional representation of $\mathrm{SL}(2,5)$ is the trivial representation. Therefore there is no possibility for the reduction modulo p of a two-dimensional representation leading to a congruence when we have projective image A_5. We therefore do not give the character table for $\mathrm{SL}(2,5)$, but note that it can be found here [4].

The only other possibility that we could have for a congruence is the so called "ethereal forms". These forms are mod p forms that cannot be lifted to characteristic 0 and are the subject of Schaeffer's thesis [Sch]. Since these forms are in characteristic p it would only make sense to try and reduce a particular two-dimensional Galois representation modulo p in order to try and find a congruence. However it turns out that this is not possible. Proposition 8.1.3 of [Sch] says that these forms are cusp forms. We then have

the following which is Theorem 2.4.1 of [Sch].

Theorem 4.2.19. *Let $k \geq 1$ and $f \in M_k(N, \chi; \overline{\mathbb{F}}_p)$ be a newform. There exists a continuous semisimple representation*

$$\rho_f : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{F}}_{\mathrm{p}})$$

unramified outside Np, such that for all $\ell \nmid Np$ and any Frobenius σ_ℓ above ℓ we have

$$\mathrm{Tr}\rho_f(\sigma_\ell) = a(f; T_\ell) \quad \textit{and} \quad \det\rho_f(\sigma_\ell) = \chi(\ell)\ell^{k-1}.$$

Furthermore, this representation is irreducible only if f is a cusp form.

We note that the case $k \geq 2$ follows from results of Eichler-Shimura and Deligne. Since all ethereal forms are cusp forms, this implies that the associated Galois representation is irreducible in characteristic p. Therefore there cannot be any congruences in this case either.

Representation/conjugacy class representative and size	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ (size 1)	$\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ (size 1)	$\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ (size 6)	$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ (size 4)	$\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$ (size 4)	$\begin{pmatrix}-1 & 1\\ 0 & -1\end{pmatrix}$ (size 4)	$\begin{pmatrix}-1 & -1\\ 0 & -1\end{pmatrix}$ (size 4)
trivial one-dimensional (χ_1)	1	1	1	1	1	1	1
non-trivial one-dimensional (χ_2)	1	1	1	ω	ω^2	ω^2	ω
non-trivial one-dimensional (χ_3)	1	1	1	ω^2	ω	ω	ω^2
two-dimensional	2	-2	0	-1	-1	1	1
two-dimensional	2	-2	0	$-\omega$	$-\omega^2$	ω^2	ω
two-dimensional	2	-2	0	$-\omega^2$	$-\omega$	ω	ω^2
three-dimensional	3	3	-1	0	0	0	0

Table 4.1: Character Table for $\mathrm{SL}(2,3)$

Here ω denotes a primitive cube root of unity.

Sums of characters and two-dimensional representations/conjugacy class representative and size	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ (size 1)	$\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$ (size 1)	$\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ (size 6)	$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ (size 4)	$\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$ (size 4)	$\begin{pmatrix}-1 & 1\\ 0 & -1\end{pmatrix}$ (size 4)	$\begin{pmatrix}-1 & -1\\ 0 & -1\end{pmatrix}$ (size 4)
$\chi_1 + \chi_1$	2	2	2	2	2	2	2
$\chi_1 + \chi_2$	2	2	2	$1+\omega$	$1+\omega^2$	$1+\omega^2$	$1+\omega$
$\chi_1 + \chi_3$	2	2	2	$1+\omega^2$	$1+\omega$	$1+\omega$	$1+\omega^2$
$\chi_2 + \chi_2$	2	2	2	2ω	$2\omega^2$	$2\omega^2$	2ω
$\chi_2 + \chi_3$	2	2	2	$\omega+\omega^2$	$\omega+\omega^2$	$\omega+\omega^2$	$\omega+\omega^2$
$\chi_3 + \chi_3$	2	2	2	$2\omega^2$	2ω	2ω	$2\omega^2$
two-dimensional	2	-2	0	-1	-1	1	1
two-dimensional	2	-2	0	$-\omega$	$-\omega^2$	ω^2	ω
two-dimensional	2	-2	0	$-\omega^2$	$-\omega$	ω	ω^2

Table 4.2: Comparison of traces of sums of characters of $\mathrm{SL}(2,3)$ and two-dimensional representations

Representation/conjugacy class representative and size	$\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$ (size 1)	$\begin{pmatrix}-1 & 0\\0 & -1\end{pmatrix}$ (size 1)	$\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$ (size 6)	$\begin{pmatrix}0 & 1\\1 & -1\end{pmatrix}$ (size 6)	$\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}$ (size 6)	$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$ (size 8)	$\begin{pmatrix}-1 & 1\\0 & -1\end{pmatrix}$ (size 8)	$\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$ (size 12)
trivial (χ_1)	1	1	1	1	1	1	1	1
non-trivial one-dimensional (χ_2)	1	1	1	-1	-1	1	1	-1
two-dimensional	2	2	2	0	0	-1	-1	0
two-dimensional	2	-2	0	$\sqrt{-2}$	$-\sqrt{-2}$	-1	1	0
two-dimensional	2	-2	0	$-\sqrt{-2}$	$\sqrt{-2}$	-1	1	0
three-dimensional	3	3	-1	-1	-1	0	0	1
three-dimensional	3	3	-1	1	1	0	0	-1
four-dimensional	4	-4	0	0	0	1	-1	0

Table 4.3: Character Table for $\mathrm{GL}(2,3)$

Reduction/conjugacy class representative and size	$\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$ (size 1)	$\begin{pmatrix}-1 & 0\\0 & -1\end{pmatrix}$ (size 1)	$\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$ (size 6)	$\begin{pmatrix}0 & 1\\1 & -1\end{pmatrix}$ (size 6)	$\begin{pmatrix}0 & 1\\1 & 1\end{pmatrix}$ (size 6)	$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$ (size 8)	$\begin{pmatrix}-1 & 1\\0 & -1\end{pmatrix}$ (size 8)	$\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$ (size 12)
$\chi_1 + \chi_1$	2	2	2	2	2	2	2	2
$\chi_1 + \chi_2$	2	2	2	0	0	2	2	0
$\chi_2 + \chi_2$	2	2	2	-2	-2	2	2	-2
two-dimensional	2	2	2	0	0	-1	-1	0
two-dimensional	2	-2	0	$\sqrt{-2}$	$-\sqrt{-2}$	-1	1	0
two-dimensional	2	-2	0	$-\sqrt{-2}$	$\sqrt{-2}$	-1	1	0

Table 4.4: Comparison of traces of sums of characters of $\mathrm{GL}(2,3)$ and two-dimensional representations

Chapter 5

Congruences Between Genus 1 and Genus 2 Cusp Forms

Now that we have proven a congruence in the case of weight 1, the next step would be to either consider similar congruences with prime powers as the moduli of a congruence, or to consider congruences involving different types of modular forms. In attempting to adapt the method used in the weight 1 case, it will turn out that we are considering the latter case.

One of the main ingredients of the weight 1 case was the one-dimensional representation of G_K (i.e. the Hecke character). This representation was induced to $G_{\mathbb{Q}}$ to give a two-dimensional representation whose associated modular form (a weight 1 cusp form) satisfied a congruence. We now aim to start with a two-dimensional representation and induce this to get a four-dimensional representation. The modular form attached to this representation will be a Siegel cusp form. Again it will turn out that this modular form will satisfy a congruence.

The first question we might be interested in asking is: Which two dimensional representation should we induce? The correct representation will be the one attached to a Hilbert modular form. The induction of the Galois representation will be equivalent to taking a theta lift of the Hilbert modular form. Just as in the weight 1 case where we needed $\chi \neq \chi_\sigma$, we will need the two-dimensional Galois representation associated to the Hilbert modular form to be non Galois-invariant. We will first cover the case which leads to a scalar valued Siegel cusp form before moving on to the more general case of vector valued Siegel modular forms.

We start by choosing a weight 2 cusp form $f \in S_2(\Gamma_0(N))$. We will then take a base change of this to get a Hilbert cusp form which, as we will show later, will usually have level norm N^2. The level norm will only be lower if N contains primes that ramify in the field K that we base change to. As a base change, this modular form is Galois-invariant. We therefore use a result of Taylor, which is a generalisation of a result of Ribet, to raise the level of the Hilbert modular form. If we raise the level by

a prime that splits in K, then this step will ensure that the Hilbert modular form is no longer Galois-invariant. We will then be able to use a result of Johnson-Leung and Roberts to take a theta lift of this Hilbert modular form to obtain a Siegel paramodular cusp form whose level and Hecke eigenvalues we will know. This cusp form will then satisfy a congruence, moreover it will be compatible with the congruence satisfied by the Hilbert modular forms involved in the level raising.

Before we move on to explaining this process we first need to cover some background material. First we will discuss the more standard material on Hilbert modular forms and Siegel modular forms. We then move on to discuss base change, level raising and theta lifts.

§ 5.1 Hilbert Modular Forms

There are many references for the theory of Hilbert modular forms such as Freitag [Frei], van der Geer [Geer], Goren [Goren] and The 1-2-3 of Modular Forms [1-2-3]. Here we follow the treatment given by Dembélé and Voight [DemVoi].

When working with classical (elliptic) modular forms we often consider the action of the modular group $\mathrm{SL}_2(\mathbb{Z})$ on the upper half plane. Of course, $\mathrm{SL}_2(\mathbb{Z})$ is a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$, which is itself a subgroup of $\mathrm{GL}_2(\mathbb{R})$. Here we can consider $\mathbb{Z}$ as the ring of integers of $\mathbb{Q}$. Let K be a totally real field with $[K:\mathbb{Q}] = n$ and let $\mathcal{O}_K$ be the ring of integers of K.

For simplicity we assume that the narrow class number of K is 1. Recall that the class group was defined as $C_K = I_K/P_K$. The narrow class group is defined to be $C_K^+ = I_K/P_K^+$ where P_K^+ is the totally positive principal fractional ideals. The narrow class number is then the order of this group.

When $n = 1$ we are in the case of classical modular forms. When $n \geq 2$, let $v_1, v_2, \ldots, v_n : K \to \mathbb{R}$ be the real places of K. Given an element $x \in K$, write $v_i(x) = x_i$ for the i-th embedding of x. Given a matrix $\gamma \in M_2(K)$ write $v_i(\gamma) = \gamma_i \in M_2(\mathbb{R})$.

Consider the group

$$\mathrm{GL}_2^+(K) = \{\gamma \in \mathrm{GL}_2(K) : \det\gamma_i > 0 \text{ for } i = 1, \ldots, n\}.$$

Since we have n different embeddings it is natural to consider an action on $\mathcal{H}^n$ (n copies of the upper half plane $\mathcal{H}$) by coordinatewise fractional linear transformations

$$z \mapsto \gamma z = (\gamma_i z_i)_i = \left(\frac{a_i z_i + b_i}{c_i z_i + d_i}\right)_{i=1,\ldots,n}.$$

As one might expect, we could consider the action of the Hilbert modular group $\mathrm{SL}_2(\mathcal{O}_K) \subset \mathrm{GL}_2^+(\mathcal{O}_K) \subset \mathrm{GL}_2^+(K)$. We also have the notion of congruence subgroups

in this new setting. Suppose $\mathfrak{N} \in \mathcal{O}_K$ is a non-zero ideal. Then we define

$$\Gamma_0(\mathfrak{N}) = \left\{ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathcal{O}_K) : c \in \mathfrak{N} \right\} \subset \mathrm{GL}_2^+(\mathcal{O}_K) \subset \mathrm{GL}_2^+(K).$$

We almost have the tools necessary to give the definition of a Hilbert modular form. First however, we define the automorphy factors.

Definition 5.1.1. For $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ and $z \in \mathcal{H}$, define

$$j(\gamma, z) = \det(\gamma)^{-1/2}(cz + d).$$

Definition 5.1.2. A Hilbert modular form of weight $(k_1, \ldots, k_n)$ and level $\mathfrak{N}$ is a holomorphic function $f : \mathcal{H}^n \to \mathbb{C}$ such that

$$f(\gamma z) = \left(\frac{a_1 z_1 + b_1}{c_1 z_1 + d_1}, \ldots, \frac{a_n z_n + b_n}{c_n z_n + d_n} \right) = \left(\prod_{i=1}^{n} j(\gamma_i, z_i)^{k_i} \right) f(z)$$

for all $\gamma \in \Gamma_0(\mathfrak{N})$. If $k_1 = k_2 = \cdots = k_n = k$, then f is said to be of parallel weight k.

We will be interested in the case where $K = \mathbb{Q}(\sqrt{d})$ is a real quadratic field and f is of parallel weight 2. This will be the case from now on, unless otherwise stated. In other words, we will consider functions $f : \mathcal{H}^2 \to \mathbb{C}$ such that

$$\begin{aligned} f(\gamma z) = \left(\frac{a_1 z_1 + b_1}{c_1 z_1 + d_1}, \frac{a_2 z_2 + b_2}{c_2 z_2 + d_2} \right) &= j(\gamma_1, z_1)^2 j(\gamma_2, z_2)^2 f(z) \\ &= \frac{(c_1 z_1 + d_1)^2 (c_2 z_2 + d_2)^2}{\det(\gamma_1)\det(\gamma_2)} f(z). \end{aligned}$$

In this case we denote the space of Hilbert modular forms of parallel weight 2 and level $\mathfrak{N}$ by $M_2(\mathfrak{N})$ and we denote the cusp forms by $S_2(\mathfrak{N})$. For modular forms of non-parallel weight, these spaces would be denoted $M_{k_1,k_2}(\mathfrak{N})$ and $S_{k_1,k_2}(\mathfrak{N})$. Again, these spaces are finite dimensional $\mathbb{C}$-vector spaces.

We note that the definition of a Hilbert modular form makes no reference to holomorphy at the cusps. This is because it automatically follows from Koecher's principle [Geer, §1] that we have holomorphy at the cusps. Just as in the classical case, there is an orthogonal decomposition of $M_2(\mathfrak{N}) = S_2(\mathfrak{N}) \oplus \mathcal{E}_2(\mathfrak{N})$ where $\mathcal{E}_2(\mathfrak{N})$ is spanned by the Eisenstein series of level $\mathfrak{N}$.

Hilbert modular forms, just like classical modular forms, have a Fourier expansion. This expansion however, is a little more complicated. For a fractional ideal $\mathfrak{a}$ of K let,

$$\mathfrak{a}_+ = \{x \in \mathfrak{a} : x_i > 0 \text{ for } i = 1, \ldots, n\}.$$

In other words $\mathfrak{a}_+$ is all the elements of $\mathfrak{a}$ that are positive under all real embeddings of K. If $\mathfrak{d}$ denotes the different ideal of K, then the inverse different $\mathfrak{d}^{-1} = \{x \in K :$

$\mathrm{Tr}(xy) \in \mathbb{Z}$ for all $y \in \mathcal{O}_K\}$ is a fractional ideal of K containing $\mathcal{O}_K$. The different ideal $\mathfrak{d}$ is then the inverse of this ideal. A Hilbert modular form $f \in M_2(\mathfrak{N})$ then admits the following Fourier expansion

$$f(z) = a_0 + \sum_{\mu \in (\mathfrak{d}^{-1})_+} a_\mu e^{2\pi i \mathrm{Tr}(\mu z)}.$$

Notice how the Fourier coefficients are no longer simply indexed by integers, they are now indexed by elements of an ideal. Suppose $f \in M_2(\mathfrak{N})$ and $\mathfrak{n}$ is a non-zero ideal of $\mathcal{O}_K$. Since we assumed that K has narrow class number 1, we may write $\mathfrak{n} = \nu\mathfrak{d}^{-1}$ for some $\nu \in \mathfrak{d}_+$. If we define $a_\mathfrak{n} = a_\nu$ then the transformation rule implies that $a_\mathfrak{n}$ does not depend on the choice of ν. We therefore call $a_\mathfrak{n}$ the Fourier coefficient of f at $\mathfrak{n}$.

Just as in the classical case there are Hecke operators acting on the spaces $M_2(\mathfrak{N})$ and $S_2(\mathfrak{N})$. Again these operators are pairwise commuting diagonalisable operators. They are now however indexed by ideals rather than integers. Given a prime ideal $\mathfrak{p} \nmid \mathfrak{N}$ and a totally positive generator p of $\mathfrak{p}$ we have

$$(T_\mathfrak{p} f)(z) = \mathrm{N}(\mathfrak{p}) f(pz) + \frac{1}{\mathrm{N}(\mathfrak{p})} \sum_{a \in \mathbb{F}_\mathfrak{p}} f\left(\frac{z+a}{p}\right),$$

where $\mathbb{F}_\mathfrak{p} = \mathcal{O}_K/\mathfrak{p}$ is the residue field of $\mathfrak{p}$. We could also write this in terms of the usual double coset decomposition. We have

$$(T_\mathfrak{p} f)(z) = \sum_{a \in \mathbb{P}^1(\mathbb{F}_\mathfrak{p})} (f|_k \pi_a)(z)$$

where $\pi_\infty = \begin{bmatrix} p & 0 \\ 0 & 1 \end{bmatrix}$ and $\pi_a = \begin{bmatrix} 1 & a \\ 0 & p \end{bmatrix}$ for $a \in \mathbb{F}_\mathfrak{p}$.

As in the classical case, if $f \in S_2(\mathfrak{N})$ is an eigenform, normalised so that $a_{(1)} = 1$, then $T_\mathfrak{n} f = a_\mathfrak{n} f$, and each $a_\mathfrak{n}$ is an algebraic integer.

We can also easily generalise the notion of an L-function attached to f. Associated to an eigenform $f \in S_2(\mathfrak{N})$ is the L-function

$$L(f,s) = \sum_\mathfrak{n} \frac{a_\mathfrak{n}}{\mathrm{N}(\mathfrak{n})^s}.$$

We also have an associated $\mathfrak{l}$-adic Galois representation

$$\rho_{f,\mathfrak{l}} : \mathrm{Gal}(\overline{K}/K) \to \mathrm{GL}_2(\overline{\mathcal{O}}_{K,\mathfrak{l}})$$

for primes $\mathfrak{l}$ of $\mathcal{O}_K$ such that, for any prime $\mathfrak{p} \nmid \mathfrak{N}\mathfrak{l}$, we have

$$\mathrm{Tr}(\rho_{f,\mathfrak{l}}(\mathrm{Frob}_\mathfrak{p})) = a_\mathfrak{p}(f) \quad \text{and } \det(\rho_{f,\mathfrak{l}}(\mathrm{Frob}_\mathfrak{p})) = \mathrm{N}(\mathfrak{p}).$$

Note that in this section we have assumed that the narrow class number is 1. This simply made it easy to consider the Fourier expansion of a Hilbert modular form and also to define the Hecke operators. We however note that the definitions can be generalised to the case where the narrow class number is not 1. For those wishing to see the details of this see [DemVoi, §7-9]. We also use this condition later when discussing the level raising result for Hilbert modular forms.

§ 5.2 Siegel Modular Forms

As with Hilbert modular forms, there are many references for Siegel modular forms. Among these are Andrianov [A], Kohnen [Ko], Buzzard [Bu2] and [1-2-3]. Here we broadly follow the outline given in Fretwell's thesis [Fret2], omitting any details irrelevant to this thesis.

Siegel modular forms are a different kind of generalisation of classical modular forms. Whereas Hilbert modular forms involved working over a totally real field and looking at an action on $\mathcal{H}^n$, we instead consider an action of a different group on a higher dimensional upper half plane. Consider the symplectic group of genus g

$$\mathrm{Sp}_{2g}(\mathbb{R}) = \{\gamma \in M_{2g}(\mathbb{R}) : \gamma J \gamma^T = J\},$$

where:

$$J = \begin{bmatrix} 0 & I_g \\ -I_g & 0 \end{bmatrix}.$$

This group can be thought of as a higher dimensional version of $\mathrm{SL}_n(\mathbb{R})$. Just as we have $\mathrm{SL_n}(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$, with $\mathrm{SL}_n(\mathbb{R})$ the kernel of the determinant map of $\mathrm{GL}_n(\mathbb{R})$, we have $\mathrm{Sp}_{2g}(\mathbb{R}) \subset \mathrm{GSp}_{2g}(\mathbb{R})$ where

$$\mathrm{GSp}_{2g}(\mathbb{R}) = \{\gamma \in M_{2g}(\mathbb{R}) : \gamma J \gamma^T = \mu(\gamma) J, \mu(\gamma) \in \mathbb{R}^\times\}.$$

The group $\mathrm{GSp}_{2g}(\mathbb{R})$ is known as the group of similitudes of $\mathrm{Sp}_{2g}(\mathbb{R})$. The similitude map $\mu : \mathrm{GSp}_{2g}(\mathbb{R}) \to \mathbb{R}^\times$ is analogous to the determinant map for $\mathrm{GL}_n(\mathbb{R})$ and in particular $\mathrm{Sp}_{2g}(\mathbb{R})$ is the kernel of this map. We note that $\mu(\gamma)$ is sometimes also known as the multiplier of γ.

Now that we have a generalisation of the special linear and general linear groups, we next generalise the upper half plane. The Siegel upper half space of genus g is given by

$$\mathcal{H}_g = \{Z \in M_g(\mathbb{C}) : Z^T = Z, \mathrm{Im}(Z) > 0\}.$$

Here $\mathrm{Im}(Z) > 0$ means that $\mathrm{Im}(Z)$ is positive definite. Note that when $g = 1$, this reduces to the usual upper half plane $\mathcal{H}$. In this case we already know that $\mathrm{Sp}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})$ acts transitively on the upper half plane $\mathcal{H}$ by fractional linear transformations. As one might expect, this result generalises.

Lemma 5.2.1. *The group* $\mathrm{Sp}_{2g}(\mathbb{R})$ *acts transitively on* $\mathcal{H}_g$ *by fractional linear transformations, i.e., if* $Z \in \mathcal{H}_g$ *and* $\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{R})$ *is written in* $g \times g$ *blocks then*

$$(\gamma, Z) \mapsto \gamma Z = \frac{AZ + B}{CZ + D}$$

defines a transitive group action.

This action is the obvious generalisation of the classical case to higher dimension. As usual we define the automorphy factor by $j(\gamma, Z) = (CZ + D)$.

Definition 5.2.2. A holomorphic function $F : \mathcal{H}_g \to \mathbb{C}$ is a classical Siegel modular form of genus g and weight k for $\mathrm{Sp}_{2g}(\mathbb{Z})$ if

(1) $F(\gamma Z) = \det(j(\gamma, Z))^k F(Z)$ for all $\gamma \in \mathrm{Sp}_{2g}(\mathbb{Z})$.

(2) If $g = 1$ then F is holomorphic at infinity.

As with Hilbert modular forms, we do not need to check holomorphicity at infinity when $g \geq 2$ because of Koecher's principle [Geer, §1]. We note that this definition is of scalar valued Siegel modular forms. We will see shortly a more general definition of Siegel modular form known as a vector valued Siegel modular form.

We denote the space of classical Siegel modular forms of weight k and genus g for $\mathrm{Sp}_{2g}(\mathbb{Z})$ by $M_k(\mathrm{Sp}_{2g}(\mathbb{Z}))$. As usual, these spaces are finite dimensional. Just as with classical modular forms, we require $k \geq 0$ in order for these spaces to be non-trivial. We also require that kg be even in order for the spaces to be non-trivial. Note that this fits with the classical case if we consider $g = 1$.

As mentioned earlier there is a more general notion of Siegel modular form known as a vector valued Siegel modular form. Let

$$\rho : \mathrm{GL}_g(\mathbb{C}) \to \mathrm{GL}(V)$$

be a finite dimensional irreducible complex representation.

Definition 5.2.3. A holomorphic function $F : \mathcal{H}_g \to V$ is a Siegel modular form of genus g and weight ρ for $\mathrm{Sp}_{2g}(\mathbb{Z})$ if

(1) $F(\gamma Z) = \rho(j(\gamma, Z))F(Z)$ for all $\gamma \in \mathrm{Sp}_{2g}(\mathbb{Z})$.

(2) If $g = 1$ then F is holomorphic at infinity.

Remark 5.2.4. If we take $\rho = \det^k$, then we obtain the scalar valued Siegel modular forms defined above.

Similarly to the classical Siegel modular forms we denote the $\mathbb{C}$-vector space of Siegel modular forms of genus g and weight ρ for $\mathrm{Sp}_{2g}(\mathbb{Z})$ by $M_\rho(\mathrm{Sp}_{2g}(\mathbb{Z}))$.

As would be expected, Siegel modular forms have a Fourier expansion. This expansion, just like that of Hilbert modular forms, is more complicated than in the classical case. First we fix a genus g. Let S_g be the set of $g \times g$ half integral matrices with integral diagonal elements. If $F \in M_\rho(\mathrm{Sp}_{2g}(\mathbb{Z}))$ then the Fourier expansion of F is given by

$$F(Z) = \sum_{T \in S_g} a(T) e^{2\pi i \mathrm{Tr}(TZ)}.$$

The first thing to notice about this expansion is that, like with Hilbert modular forms, the coefficients are no longer indexed by integers. They are now indexed by matrices $T \in S_g$. For T that are not positive semi definite, we have $a(T) = 0$. This is analogous to $a_n = 0$ for $n < 0$ in the classical case. We note that these matrices actually parametrise quadratic forms in g variables with integer coefficients.

Now that we have the Fourier expansion we can consider what it means for a Siegel modular form to be a cusp form.

Definition 5.2.5. A Siegel modular form $F \in M_\rho(\mathrm{Sp}_{2g}(\mathbb{Z}))$ is a cusp form if $a(T) = 0$ for all $T \in S_g$ such that T is positive semi-definite, but not definite.

As in the classical case we can consider congruence subgroups of $\mathrm{Sp}_{2g}(\mathbb{Z})$. For the purposes of this thesis we will not be interested in Siegel modular forms for the standard congruences subgroups. These are the obvious generalisations of the genus 1 congruence subgroups, e.g., $\Gamma_0^{(g)}(N)$ is the set of matrices in $\mathrm{Sp}_{2g}(\mathbb{Z})$ with bottom left $g \times g$ block congruent to 0 modulo N. For those wanting to know the details, see [Fret2].

5.2.1 Genus 2 Siegel Modular Forms

In this thesis we will only need to consider genus 2 Siegel modular forms, so from now on we restrict to this case.

Suppose we have an irreducible representation $\rho : \mathrm{GL}_2(\mathbb{C}) \to \mathrm{GL}(V)$. We note that it is enough to consider irreducible representations since the space $M_\rho(\mathrm{Sp}_4(\mathbb{Z})) = M_{\rho_1}(\mathrm{Sp}_4(\mathbb{Z})) \oplus M_{\rho_2}(\mathrm{Sp}_4(\mathbb{Z}))$ if ρ is reducible (This also holds for higher genus). It is known that the representation ρ is parametrised by its highest weight. It is known that for integers $j, k \geq 0$, the irreducible representation of highest weight $(j+k, k)$ has an explicit description as the representation $\mathrm{Symm}^j(\mathbb{C}^2) \otimes \det^k$, where $\mathrm{GL}_2(\mathbb{C})$ is acting via matrix multiplication on $\mathbb{C}^2$. If $\rho = \mathrm{Symm}^j(\mathbb{C}^2) \otimes \det^k$ we write $M_{j,k}(\mathrm{Sp}_4(\mathbb{Z}))$ and $S_{j,k}(\mathrm{Sp}_4(\mathbb{Z}))$ for the spaces of Siegel modular forms and Siegel cusp forms of genus 2.

We now consider the subgroup that we will be interested in working with; the paramodular group.

Definition 5.2.6. The paramodular group of level N is given by

$$K(N) = \begin{bmatrix} \mathbb{Z} & N\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ \mathbb{Z} & \mathbb{Z} & \mathbb{Z} & \frac{1}{N}\mathbb{Z} \\ \mathbb{Z} & N\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ N\mathbb{Z} & N\mathbb{Z} & N\mathbb{Z} & \mathbb{Z} \end{bmatrix} \cap \mathrm{Sp}_4(\mathbb{Q}).$$

Although this group looks very different to the groups we are used to working with, it is a very important one. This is the group that appears in the paramodularity conjecture; the genus 2 analogue of the modularity theorem. This result was conjectured by Brumer and Kramer [BrKr, Conjecture 1.1].

Although the theory is more complicated, there is a theory of newforms for the paramodular group. We omit the details but refer the reader to [RS1].

As would be expected there is a family of Hecke operators acting on the spaces of Siegel modular forms. Before defining these, we generalise the weight-k slash operator to genus 2. If $F : \mathcal{H}_2 \to V$ and $\alpha \in \mathrm{GSp}_4^+(\mathbb{Q})$, then the weight (j,k) slash operator is defined by

$$(F|_{j,k}\alpha)(Z) = \mu(\alpha)^{j+k-3}\rho(j(\alpha,Z))^{-1}F(\alpha Z).$$

For each prime p we may define Hecke operators T_p and T_{p^2}. The process is similar to the classical case. We decompose the double cosets

$$K(N)\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{bmatrix} K(N) = \coprod_i K(N)\mu_i,$$

$$K(N)\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p^2 & 0 \\ 0 & 0 & 0 & p \end{bmatrix} K(N) = \coprod_i K(N)\eta_i.$$

Each of these decompositions has finitely many representatives.

If $F \in S_{j,k}(K(N))$ we set

$$T_p(F) = \sum_i F|_{j,k}\mu_i,$$

$$T_{p^2}(F) = \sum_i F|_{j,k}\eta_i.$$

The operators T_p and T_{p^2} are then the Hecke operators at the prime p. There are more Hecke operators than this, but we only work with these operators. More specifically we will mostly be interested with the T_p operator. As with the classical case, we can find bases for the spaces $S_{j,k}(K(N))$ consisting of eigenforms for the Hecke operators. The

Hecke operators also preserve the newspace and we can again find a basis consisting of eigenforms for the Hecke operators.

We can also consider the Galois representation attached to such a modular form. This will be important as this will be the representation obtained by inducing the representation attached to the level raised Hilbert modular form. The trace of this representation will then correspond to the Hecke eigenvalues of the Siegel paramodular newform. These eigenvalues will then satisfy a congruence. Part 1 of the following result is a restatement of part of Theorem 1 of [W]. Part 2 is a restatement of Theorem 2 of [W2]. Note that this was originally stated in [W, Theorem IV] with an additional assumption that was later dropped.

Proposition 5.2.7. *Suppose F is a newform in $S_{j,k}(K(N))$. Let L be a number field containing all the Hecke eigenvalues $\lambda(p)$ for T_p and $\mu(p)$ for T_{p^2} for all primes $p \nmid N$.*

(1) For any prime λ' of $\mathcal{O}_L$, there exists a finite extension E of L (and $E_{\lambda'}$ of $L_{\lambda'}$), and a 4-dimensional semisimple Galois representation

$$\rho_{F,\lambda'} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_4(E_{\lambda'}),$$

unramified outside $\{N, \ell\}$ (where $\lambda'|\ell$), such that for each prime $p \notin \{N, \ell\}$,

$$\det(I - \rho_{F,\lambda'}(\mathrm{Frob}_p^{-1} p^{-s})) = L_p(s, F, \mathrm{spin})^{-1}$$

where $L_p(s, F, \mathrm{spin}) = P_p(p^{-s})^{-1}$, with

$$P_p(X) = 1 - \lambda(p)X + (\lambda(p)^2 - \mu(p) - p^{j+2k-4})X^2 - \lambda(p)p^{j+2k-3}X^3 + p^{2j+4k-6}X^4,$$

the Euler factor at p in the (shifted) spinor L-function of F.

(2) Further, the representation $\rho_{F,\lambda'}$ is symplectic. In particular the image of $\rho_{F,\lambda'}$ is contained in $GSp_4(E_{\lambda'})$.

For those wanting to know the details of the spinor L-function and other L-functions associated to Siegel modular forms, see [1-2-3, §20-21].

§ 5.3 Base Change

In this section the main result will tell us the level and Hecke eigenvalues of the base change Hilbert modular form given the level and Hecke eigenvalues of the classical modular form we start with. Before we state this result however, we first need to know that a base change actually exists. In fact we do always have a base change, and further, it is unique; see the Theorem in Gelbart's article on page 194 of [MFFLT]. Although this result is in the language of automorphic representations, all we need is the first part of the theorem which is easy to translate into the language of modular forms. It says that given a classical cusp form, there is a unique base change that is also cuspidal.

We now consider the base change from a classical modular form to a Hilbert modular form. As was mentioned at the beginning of this section, starting with a classical weight 2 modular form of level N and taking a base change will usually result in a Hilbert modular form of level norm N^2.

Theorem 5.3.1. *Let $f \in S_2(\Gamma_0(N))$ and let $K = \mathbb{Q}(\sqrt{d})$ with $d > 0$ and squarefree. Assume that $(N, d_K) = 1$ where d_K is the discriminant of K. Then the base change of f to K will be a Hilbert modular form f' defined over K with $f' \in S_2(\mathfrak{N})$ where $\mathfrak{N} = (N)$. Further the Hecke eigenvalues of f' at a prime $\mathfrak{p}$ are given by*

$$a_{\mathfrak{p}}(f') = \begin{cases} a_p(f) & \text{if } p = \mathfrak{p}\overline{\mathfrak{p}} \text{ splits in } \mathcal{O}_K, \\ a_p(f)^2 - 2p & \text{if } p = \mathfrak{p} \text{ is inert in } \mathcal{O}_K, \\ a_p(f) & \text{if } p = \mathfrak{p}^2 \text{ ramifies in } \mathcal{O}_K. \end{cases}$$

Before we prove this result, there are a few definitions and results we will need which we now state. Firstly, rather than considering the level of the Hilbert modular form, we can instead consider the conductor of the associated Galois representation since these will be equal. This is not a straightforward result. In fact it requires the combination of two other results. It follows from Théorème(A) of [Cara] and Theorem 1 of [Cas] which makes use of the local Langlands correspondence.

We therefore need to know a little about the conductor of a Galois representation. We first consider the case of a finite Galois extension L/K. The discriminant $\mathfrak{d}$ of this extension can be expressed by the following product decomposition

$$\mathfrak{d} = \prod f(\chi)^{\chi(1)},$$

where χ varies over the irreducible characters of the Galois group $G = \mathrm{Gal}(L/K)$. Each of the $f(\chi)$'s are ideals given by

$$\mathfrak{f}(\chi) = \prod_{\mathfrak{p} \nmid \infty} \mathfrak{p}^{f_{\mathfrak{p}}(\chi)}$$

with

$$f_{\mathfrak{p}}(\chi) = \sum_{i \geq 0} \frac{1}{[G_0 : G_i]} \mathrm{codim} V^{G_i},$$

where V is a representation with character χ and G_i is the i-th ramification group of $L_{\mathfrak{P}}/K_{\mathfrak{p}}$. The ideal $\mathfrak{f}(\rho)$ is the Artin conductor and each $f_{\mathfrak{p}}(\rho)$ is a local Artin conductor. We now define the ramification groups G_i. For $\sigma \in G$ we define,

$$i_G(\sigma) = v_L(\sigma x - x),$$

where x is an element such that $\mathcal{O}_L = \mathcal{O}_K[x]$, and v_L is the normalised valuation of L. The i-th ramification group is then defined as

$$G_i = \{\sigma \in G | \, i_G(\sigma) \geq i + 1\}.$$

We note that if L/K is unramified, then $i_G(\sigma) = 0$ for all $\sigma \in G, \sigma \neq 1$.

In our case we are working with two-dimensional representations, ρ and ρ' say, which are representations of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ and $\mathrm{Gal}(\bar{\mathbb{Q}}/K)$ respectively. We can therefore use the same process to calculate the conductor except we now need to take the appropriate inverse limit over all such finite extensions. We now prove Theorem 5.3.1.

Proof of Theorem 5.3.1. The Galois representation associated to f, ρ say, will be a representation of $G_{\mathbb{Q}}$. The Galois representation associated to f', ρ' say, will be the restriction of ρ to G_K. We therefore aim to find the conductor of ρ' given the conductor of ρ.

First we consider a prime p that splits in K, as say $p = \mathfrak{p}\bar{\mathfrak{p}}$. We aim to show that $f_{\mathfrak{p}}(\rho') = f_{\bar{\mathfrak{p}}}(\rho') = f_p(\rho)$. We note that since $K_{\mathfrak{p}}$ and $K_{\bar{\mathfrak{p}}}$ are both one-dimensional extensions of $\mathbb{Q}_p$ we have $\mathrm{Gal}(\bar{\mathbb{Q}}_{\mathfrak{p}}/\mathbb{Q}_p) = \mathrm{Gal}(\bar{\mathbb{Q}}_{\mathfrak{p}}/K_{\mathfrak{p}}) = \mathrm{Gal}(\bar{\mathbb{Q}}_{\bar{\mathfrak{p}}}/K_{\bar{\mathfrak{p}}})$. It follows that the whole chain of ramification groups is the same in both cases and therefore we have $f_{\mathfrak{p}}(\rho') = f_{\bar{\mathfrak{p}}}(\rho') = f_p(\rho)$ in the split case.

We now consider the inert case. We now aim to show that $f_p(\rho') = f_p(\rho)$. Again we will do this by showing that the whole chain of groups must in fact be equal. This time K_p is a degree two extension of $\mathbb{Q}_p$ so we do not simply have $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p) = \mathrm{Gal}(\bar{\mathbb{Q}}_p/K_p)$. However the inertia groups are still the same in each case since $K_p/\mathbb{Q}_p$ is an unramified extension. From this it follows that the whole chain must actually be the same since we can replace $\sigma \in G$ with $\sigma \in G_0$ in the definition of the ramification groups. Hence we conclude that $f_p(\rho') = f_p(\rho)$.

It therefore follows that the base change is an element of $S_2(\mathfrak{N})$ with $\mathfrak{N} = (N)$ and $\mathrm{N}(\mathfrak{N}) = N^2$.

For details of the Hecke eigenvalues, refer to [Cre]. □

Remark 5.3.2. The calculations carried out in [Cre] are for imaginary quadratic fields and weight $k = 2$. These results also hold in the case of real quadratic fields. We also note that in the case of more general weights k, the Hecke eigenvalue at an inert prime becomes $a_p(f)^2 - 2p^{k-1}$.

Notice that in the statement of Theorem 5.3.1, we avoid levels containing primes that ramify in K. The reason for this is that the level of the base changed form is not so straightforward in this case. If N contains ramified primes then the norm of $\mathfrak{N}$ is in fact smaller than N^2.

Conjecture 5.3.3. *Let $f \in S_2(\Gamma_0(N))$ with $N = p_1^{r_1}p_2^{r_2}\dots p_n^{r_n}$. Let $K = \mathbb{Q}(\sqrt{d})$ with $d > 0$ and squarefree. Let $I = \{i_1, i_2, \dots, i_m\}$ be an indexing set for the primes that ramify in K. Then the base change of f to K will be a Hilbert modular form f' defined over K with $f' \in S_2(\mathfrak{N})$ where*

$$\mathrm{N}(\mathfrak{N}) = \frac{p_1^{2r_1}p_2^{2r_2}\dots p_n^{2r_n}}{\prod_{i\in I} p_i^{r_i}} = \frac{N^2}{\prod_{i\in I} p_i^{r_i}}.$$

This conjecture has been tested in many cases and should hold. The proof however would not be as straightforward as for Theorem 5.3.1. In particular the chain of ramification groups will not be the same in each case. Therefore a more in depth argument would be required. It should be possible to prove on a case by case basis by so called local-global arguments. We also note that although these results have been stated with weight 2 classical cusp forms in mind, the results still hold in the case of weight $k > 2$. This will be important when we consider the vector valued case later as the first part of the method remains the same.

§ 5.4 Level Raising Congruences

In this section we will look at congruences between modular forms of the same type but different level. What we mean by this is that, rather than having a congruence between a non cusp form and a cusp form like the congruences we have investigated so far, these congruences will be between two cusp forms. The levels however will be analogous to the congruences we have already considered. That is, one modular form will have level N, for some N, and the other will have level Np for some prime p. We will first consider the case of classical modular forms as an introduction to the idea of level raising as introduced by Ribet in 1990. We will then move on to look at the case of Hilbert modular forms. Here we will make use of a result of Taylor in order to raise the level of a Hilbert modular form. This will be the second step in our process of producing a congruence involving a Siegel modular form.

5.4.1 Classical Modular Forms

Ribet's original paper [Rib] from 1990 only covers the case of weight 2 modular forms. Although we will only be working with weight 2 forms at first, we will be considering higher weights in the vector valued case. It is therefore worth noting that the result can be extended to any weight $k \geq 2$. The result was generalised by Diamond and can be found in another paper of Ribet. See [Rib2, §5]. The following is a restatement of Theorem 1 of [Rib].

Theorem 5.4.1. *Let $f \in S_2(\Gamma_0(N))$ be a newform. Let $p \nmid \ell N$ be a prime satisfying one or both of the identities*

$$a_p(f) \equiv \pm(p+1) \pmod{\ell}.$$

Further, assume that the reduction $\overline{\rho}_f$ modulo ℓ of the associated Galois representation is irreducible. Then there exists a newform $g \in S_2(\Gamma_0(Np))$ with $\mathrm{Tr}(\rho_f(\mathrm{Frob}_q)) \equiv \mathrm{Tr}(\rho_g(\mathrm{Frob}_q)) \pmod{\ell}$ for all q with $(Np\ell, q) = 1$.

Remark 5.4.2. The identity to be satisfied can be viewed instead as $a_p(f)^2 \equiv (p+1)^2 \pmod{\ell}$.

Remark 5.4.3. The generalisation to higher weight only changes the identity that needs to be satisfied and adds a condition on ℓ. The identity changes to $a_p(f) \equiv \pm(p^{\frac{k}{2}} + p^{\frac{k}{2}-1}) \pmod{\ell}$ and we require $2 \leq k \leq \ell + 1$.

Let us now consider a couple of examples to see how this theorem works in practice.

Example 5.4.4. Consider $f \in S_2(\Gamma_0(11))$ with

$$f = q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 - 2q^9 + \dots.$$

Suppose we choose $p = 3$. That is we wish to find $g \in S_2(\Gamma_0(33))$ with $a_p(f) \equiv a_p(g) \pmod{\ell}$ for some ℓ. The condition we must check is $\ell | a_3(f)^2 - (3+1)^2$. This gives

$$\begin{aligned} &\ell \mid \left((-1)^2 - (3+1)^2\right) \\ &\ell \mid (1 - 16) \\ &\ell \mid -15. \end{aligned}$$

Hence $\ell = 3$ or $\ell = 5$. If we consider $\ell = 5$, it turns out that the Galois representation associated to f becomes reducible modulo ℓ. Hence there is no congruence modulo 5 in this case. However we can find a congruence modulo 3. Consider the following modular forms:

$$\begin{aligned} &f \in S_2(\Gamma_0(11)), \quad f = q - 2q^2 - q^3 + 2q^4 + q^5 + 2q^6 - 2q^7 - 2q^9 + \dots, \\ &g \in S_2(\Gamma_0(33)), \quad g = q + q^2 - q^3 - q^4 - 2q^5 - q^6 + 4q^7 - 3q^8 + q^9 + \dots. \end{aligned}$$

It is clear that the Hecke eigenvalues of these two modular forms are congruent modulo 3. Note that the congruence holds for those q with $3|q$ even though the theorem doesn't cover these cases.

We will also consider an example of higher weight.

Example 5.4.5. Consider $f \in S_4(\Gamma_0(7))$ with

$$f = q - q^2 - 2q^3 - 7q^4 + 16q^5 + 2q^6 - 7q^7 + 15q^8 - 23q^9 + \dots.$$

Suppose we choose $p = 2$. That is we wish to find $g \in S_4(\Gamma_0(14))$ with $a_p(f) \equiv a_p(g) \pmod{\ell}$ for some ℓ. The condition we must check in this case is $\ell | a_2(f)^2 - (2^2+2)^2$. This gives

$$\begin{aligned} &\ell \mid \left((-1)^2 - (4+2)^2\right) \\ &\ell \mid (1 - 36) \\ &\ell \mid -35. \end{aligned}$$

Hence $\ell = 5$ or $\ell = 7$. If we consider $\ell = 7$, then we see that $\ell | N$ since $N = 7$. Hence there is no congruence modulo 7 in this case. However we can find a congruence modulo 5. Consider the following modular forms:

$$\begin{aligned} &f \in S_4(\Gamma_0(7)), \quad f = q - q^2 - 2q^3 - 7q^4 + 16q^5 + 2q^6 - 7q^7 + 15q^8 - 23q^9 + \dots, \\ &g \in S_4(\Gamma_0(14)), \quad g = q - 2q^2 + 8q^3 + 4q^4 - 14q^5 - 16q^6 - 7q^7 - 8q^8 + 37q^9 + \dots. \end{aligned}$$

It is clear that the Hecke eigenvalues of these two modular forms are congruent modulo 5 except when $2|q$. This is because $p = 2$ and so we shouldn't expect a congruence at these coefficients anyway.

Now that we have seen a couple of examples of the level raising result we move on to look at an analogous result for Hilbert modular forms that was proven by Taylor.

5.4.2 Hilbert Modular Forms

A very similar result exists for Hilbert modular forms. Here we will state the result in the case we are currently looking at, the case of parallel weight 2 Hilbert modular forms. The following is a restatement of Theorem 1 of [Ta].

Theorem 5.4.6. *Let $f \in S_2(\mathfrak{N})$ be a newform. Let $\mathfrak{p} \nmid \ell\mathfrak{N}$ be a prime ideal satisfying the identity*

$$a_{\mathfrak{p}}(f)^2 \equiv (\mathrm{N}(\mathfrak{p}) + 1)^2 \pmod{\lambda},$$

for some $\lambda|\ell$. Further assume that the reduction $\overline{\rho}_f$ modulo λ is irreducible and that $\ell \nmid \mathrm{N}(\mathfrak{p})+1$. Then there exists a newform $g \in S_2(\mathfrak{N}\mathfrak{p})$ with $\mathrm{Tr}(\rho_f(\mathrm{Frob}_{\mathfrak{q}})) \equiv \mathrm{Tr}(\rho_g(\mathrm{Frob}_{\mathfrak{q}}))$ (mod λ) for all $\mathfrak{q}$ with $(\mathfrak{N}\mathfrak{p}\ell d_K, \mathfrak{q}) = 1$, where d_K is the discriminant of K and $\lambda|\ell$.

Remark 5.4.7. The assumption that $\ell \nmid \mathrm{N}(\mathfrak{p}) + 1$ allows us to ignore the error term in Taylor's result. For details of this error term, see Section 1 of [Ta].

Remark 5.4.8. This result of course holds for more general weights. We will need this result for parallel weight $k > 2$ when working in the vector valued case later. Just like the result in the classical case, it generalises in the same way.

Recall that we need to raise the level of our Hilbert modular form by a prime p that splits in K in order to guarantee that we have a Hilbert modular form that is not Galois-invariant. Suppose we have $p = \mathfrak{p}\overline{\mathfrak{p}}$. Note that since p splits we have $\mathrm{N}(\mathfrak{p}) = p$ and $a_{\mathfrak{p}}(f') = a_p(f)$. Hence the condition required for there to exist a congruence raising the level of f' by $\mathfrak{p}$ will be the same as the condition for there to exist a congruence raising the level of f by p. This might imply that we could either base change and then level raise, or first raise the level and then take a base change. However this is not the case. Using our current method, taking a base change and then raising the level, we end up with a Hilbert modular form with level $(N)\mathfrak{p}$ that is not Galois-invariant. This follows since its Galois conjugate will have level $(N)\overline{\mathfrak{p}}$ and hence be a different form. If we however raised the level first and then took a base change, we would end up with a Hilbert modular form of level (Np) that *is* Galois-invariant. This would cause the induced four-dimensional representation to be reducible. Another incarnation of the same problem is that we require non Galois invariance in order to apply the theta lifting result that we give in the next section. Also note that in this case the level is larger. In particular we could obtain this Hilbert modular form by first base changing and then raising the level twice, once by $\mathfrak{p}$, and once by $\overline{\mathfrak{p}}$. This explains why we must first base change and then raise the level.

We will now consider an example to illustrate how the level raising theorem for Hilbert modular forms works in practice. The following data was obtained using the LMFDB's database on Hilbert modular forms [LMFDB] and we use their notation. In particular, an ideal $\mathfrak{N}$ is given a label $[N, m, \alpha]$ where N is the norm of the ideal $\mathfrak{N}$, m is the smallest rational integer in the ideal, and α is an element chosen so that $(m, \alpha) = \mathfrak{N}$.

Example 5.4.9. Consider f of level $[17, 17, 3\sqrt{2}+1]$. This Hilbert modular form is not a base change. The following list gives the norm of each prime ideal in the first column, its label in the second, and the Hecke eigenvalue in the third.

Norm	Prime	Eigenvalue
2	$[2, 2, -\sqrt{2}]$	0
7	$[7, 7, -2\sqrt{2}+1]$	-4
7	$[7, 7, -2\sqrt{2}-1]$	2
9	$[9, 3, 3]$	-2
17	$[17, 17, 3\sqrt{2}+1]$	-1
17	$[17, 17, 3\sqrt{2}-1]$	6
23	$[23, 23, \sqrt{2}+5]$	-6
23	$[23, 23, -\sqrt{2}+5]$	0
25	$[25, 5, 5]$	2
31	$[31, 31, 4\sqrt{2}+1]$	2
31	$[31, 31, -4\sqrt{2}+1]$	-4
41	$[41, 41, 2\sqrt{2}-7]$	6
41	$[41, 41, -2\sqrt{2}-7]$	-6

Suppose we choose $\mathfrak{p}$ to be a prime above 23 with label $[23, 23, \sqrt{2}+5]$. Then we have $a_{\mathfrak{p}}(f) = -6$ and $\mathrm{N}(\mathfrak{p}) = 23$. Hence, ignoring the error term, the level raising condition is

$$\begin{aligned}
&\ell \mid \left(a_{\mathfrak{p}}(f)^2 - (\mathrm{N}(\mathfrak{p})+1)^2\right) \\
&\ell \mid \left((-6)^2 - (23+1)^2\right) \\
&\ell \mid (36 - 576) \\
&\ell \mid -540 = -2^2 \times 3^3 \times 5.
\end{aligned}$$

We find g of level $[391, 391, -5\sqrt{2}+21]$ that satisfies a congruence. We now give another list with the previous information along with a new column containing the eigenvalues of g.

Norm	Prime	Eigenvalues of f	Eigenvalues of g
2	$[2, 2, -\sqrt{2}]$	0	0
7	$[7, 7, -2\sqrt{2}+1]$	-4	1
7	$[7, 7, -2\sqrt{2}-1]$	2	-3
9	$[9, 3, 3]$	-2	-2
17	$[17, 17, 3\sqrt{2}+1]$	-1	-1
17	$[17, 17, 3\sqrt{2}-1]$	6	-4
23	$[23, 23, \sqrt{2}+5]$	-6	1
23	$[23, 23, -\sqrt{2}+5]$	0	5
25	$[25, 5, 5]$	2	-3
31	$[31, 31, 4\sqrt{2}+1]$	2	-8
31	$[31, 31, -4\sqrt{2}+1]$	-4	-9
41	$[41, 41, 2\sqrt{2}-7]$	6	-4
41	$[41, 41, -2\sqrt{2}-7]$	-6	-1

It is clear that these two modular forms have Hecke eigenvalues that are congruent modulo 5 except at the prime $\mathfrak{p}$. Note that we shouldn't expect the congruence to hold at the prime $[17, 17, 3\sqrt{2}+1]$ since this is the level of f. However the congruence does hold in this case. We also shouldn't expect a congruence at the prime $[2, 2, -\sqrt{2}]$ since this divides d_K. Again, however, the congruence holds in this case.

We will see another example of level raising for Hilbert modular forms later when we apply our method, see Example 5.6.2. In that case we will be raising the level of a Hilbert modular form that is a base change of a classical weight 2 modular form. In a way this case may be viewed as a little harder for a congruence to exist. This is because at primes q that split in K, as say $q = \mathfrak{q}\overline{\mathfrak{q}}$, the eigenvalues of f' at $\mathfrak{q}$ and $\overline{\mathfrak{q}}$ will be the same. The level raised form however will not be a base change. So in general the Hecke eigenvalues of this form at $\mathfrak{q}$ and $\overline{\mathfrak{q}}$ will be different. Hence the differences will be non-zero but we still require them to be divisible by ℓ for there to exist a congruence. Since the differences are bounded by $2q^{1/2}$, using small q will in general force ℓ to be small. In particular, the example we check will have $\ell = 2$.

We now move on to explain the next step in our process. Once we have base changed our classical weight 2 modular form to obtain a Hilbert modular form of parallel weight 2 and then raised the level of this form, we then want to take a theta lift to obtain a Siegel paramodular cusp form.

§ 5.5 Theta Lifts of Hilbert Modular Forms

There are many different results involving theta lifts, the result we will be interested in is a result of Johnson-Leung and Roberts. This results takes a theta lift of a weight $(2, 2n+2)$ Hilbert modular form and gives a genus 2 weight $n+2$ Siegel paramodular newform. Since we are working with a parallel weight 2 Hilbert modular form (i.e. $n = 0$) we will obtain a genus 2, weight 2 Siegel paramodular newform. The result also gives us all the information we require about the Hecke eigenvalues of this Siegel form.

The following is a restatement of (part of) the main theorem of [JLR].

Theorem 5.5.1. *Let K be a real quadratic extension of $\mathbb{Q}$. Let $f \in S_{2,2n+2}(\mathfrak{N}_0)$ be a Hilbert cusp form whose associated Galois representation ρ is irreducible. Assume that ρ is not Galois-invariant. Let $M = d_K^2 N_{\mathbb{Q}}^K(\mathfrak{N}_0)$, where d_K is the discriminant of $K/\mathbb{Q}$. Then there exists a non-zero Siegel paramodular newform $F : \mathcal{H}_2 \to \mathbb{C}$ of weight $k = n+2$ and paramodular level M such that:*

For every prime p,

$$T_p F = a_p(F)F \text{ and } T_{p^2} F = b_p(F)F$$

with

$$a_p(F) = p^{k-3}\lambda_p \text{ and } b_p(F) = p^{2(k-3)}\mu_p$$

where λ_p and μ_p are determined by the Hecke eigenvalues of f as follows. If p splits, let $\mathfrak{p}$ and $\overline{\mathfrak{p}}$ be the places above p. If p does not split, let $\mathfrak{P}$ be the place above p.

(1) If $\mathrm{val}_p(M) = 0$,

$$\lambda_p = \begin{cases} p(a_{\mathfrak{p}}(f) + a_{\overline{\mathfrak{p}}}(f)) & \text{if } p \text{ is split,} \\ 0 & \text{if } p \text{ is not split.} \end{cases}$$

$$\mu_p = \begin{cases} p^2 + p a_{\mathfrak{p}}(f) a_{\overline{\mathfrak{p}}}(f) - 1 & \text{if } p \text{ is split,} \\ -(p^2 + p a_{\mathfrak{P}}(f) + 1) & \text{if } p \text{ is not split.} \end{cases}$$

(2) If $\mathrm{val}_p(M) = 1$, then p splits and $\mathrm{val}_{\mathfrak{p}}(\mathfrak{N}_0) = 1, \mathrm{val}_{\overline{\mathfrak{p}}}(\mathfrak{N}_0) = 0$, and

$$\lambda_p = p a_{\mathfrak{p}} + (p+1) a_{\overline{\mathfrak{p}}}(f), \quad \mu_p = p a_{\mathfrak{p}}(f) a_{\overline{\mathfrak{p}}}(f).$$

(3) If $\mathrm{val}_p(M) \geq 2$, then:

p inert:

$$\lambda_p = 0, \quad \mu_p = -p^2 - p a_{\mathfrak{P}}(f);$$

p ramified:

$$\lambda_p = p\lambda_{\mathfrak{P}}, \quad \mu_p = \begin{cases} 0 & \text{if } \mathrm{val}_{\mathfrak{P}}(\mathfrak{N}_0) = 0, \\ -p^2 & \text{if } \mathrm{val}_{\mathfrak{P}}(\mathfrak{N}_0) \geq 1; \end{cases}$$

p split and $\mathrm{val}_{\mathfrak{p}}(\mathfrak{N}_0) \leq \mathrm{val}_{\overline{\mathfrak{p}}}(\mathfrak{N}_0)$*:*

$$\lambda_p = p(a_{\mathfrak{p}}(f) + a_{\overline{\mathfrak{p}}}(f)), \quad \mu_p = \begin{cases} 0 & \text{if } \mathrm{val}_{\mathfrak{p}}(\mathfrak{N}_0) = 0, \\ -p^2 & \text{if } \mathrm{val}_{\mathfrak{p}}(\mathfrak{N}_0) \geq 1. \end{cases}$$

For particular ρ, λ_p and μ_p, see Proposition 4.2 of [JLR].

Since we are working with parallel weight 2 Hilbert modular forms, we are taking $n = 0$ in this theorem. Hence the theta lift will produce a Siegel paramodular cusp form of weight 2. The Hecke operator we will be interested in is T_p; this is the generalisation of the T_p operator from the classical case. We are therefore interested in the $a_p(F)$ eigenvalues. Also when checking for a congruence we would only expect to find one at primes not dividing the level. So we will be in the case where $\mathrm{val}_p(M) = 0$.

§ 5.6 The Main Theorem

We have now covered each of the steps that we will need to prove the existence of a congruence between the Hecke eigenvalues of a Siegel paramodular cusp form of weight 2 and the Hecke eigenvalues of a classical weight 2 cusp form. We now state the main theorem.

Theorem 5.6.1. *Let $f \in S_2(\Gamma_0(N))$ and let K be a real quadratic field with discriminant d_K. Suppose $(d_K, N) = 1$. Choose a prime p that splits in K. Suppose $p \nmid \ell N$, $\ell \nmid (p+1)$ and*

$$a_p(f)^2 \equiv (p+1)^2 \pmod{\lambda},$$

for some $\lambda | \ell$. Further assume that $\rho_f|_{G_K}$ is irreducible modulo λ.

Then there exists a Siegel paramodular cusp form $F \in S_2(K(N^2 d_K^2 p))$ satisfying

$$a_q(F) \equiv a_q(f)\,(1 + \chi_K(q)) \pmod{\lambda}$$

for $q \nmid N^2 d_K^2 p$, where χ_K is the quadratic character associated to K.

The proof of this result will simply require us to apply the method we have described in this section.

Proof. Let $f = a_1 q + a_2 q^2 + a_3 q^3 + \cdots \in S_2(\Gamma_0(N))$. By Theorem 5.3.1, the base change of f will be $f' \in S_2(\mathfrak{N})$ with $\mathfrak{N} = (N)$ and $\mathrm{N}(\mathfrak{N}) = N^2$. We note that at a prime q that splits as say $\mathfrak{q}\overline{\mathfrak{q}}$ we have $a_{\mathfrak{q}}(f') = a_{\overline{\mathfrak{q}}}(f') = a_q(f)$ since f' is a base change.

We now use the level raising result for Hilbert modular forms. Choose a prime p that splits in K as say $p = \mathfrak{p}\overline{\mathfrak{p}}$. Suppose $\mathfrak{p} \nmid \ell N$ and

$$a_{\mathfrak{p}}(f')^2 \equiv (\mathrm{N}(\mathfrak{p}) + 1)^2 \pmod{\lambda}.$$

Further assume that the reduction $\overline{\rho}_{f'}$ modulo λ is irreducible and $\ell \nmid \mathrm{N}(\mathfrak{p})+1$. Note that here $\mathrm{N}(\mathfrak{p}) = p$ and $a_{\mathfrak{p}}(f') = a_p(f)$, hence the congruence condition reduces to

$$a_p(f)^2 \equiv (p+1)^2 \pmod{\lambda}.$$

We also have $\ell \nmid (\mathrm{N}(\mathfrak{p})+1)$ reducing to $\ell \nmid (p+1)$.

By Theorem 5.4.6 there exists a newform $g \in S_2(\mathfrak{N}\mathfrak{p})$ with $\mathrm{Tr}(\rho_{f'}(\mathrm{Frob}_{\mathfrak{q}})) \equiv \mathrm{Tr}(\rho_g(\mathrm{Frob}_{\mathfrak{q}})) \pmod{\lambda}$ for all $\mathfrak{q}$ with $(\mathfrak{N}\mathfrak{p}\ell d_K, \mathfrak{q}) = 1$ and $\lambda | \ell$. Further since p was chosen to be a prime that splits in K, this newform will be non Galois-invariant so its Galois conjugate will be different. This ensures that for a split prime $q = \mathfrak{q}\overline{\mathfrak{q}}$ we have $a_{\mathfrak{q}}(g) \neq a_{\overline{\mathfrak{q}}}(g)$ in general.

By Theorem 5.5.1 there exists a non-zero Siegel paramodular newform $F \in S_2(K(N^2 d_K^2 p))$ whose Hecke eigenvalues are determined by those of g. We have

$$a_q(F) = \begin{cases} a_{\mathfrak{q}}(g) + a_{\overline{\mathfrak{q}}}(g) & \text{if } q \text{ is split,} \\ 0 & \text{if } q \text{ is not split.} \end{cases}$$

We now note that by the level raising congruence of Hilbert modular forms we have $a_{\mathfrak{q}}(g) \equiv a_{\mathfrak{q}}(f') \pmod{\lambda}$. Since f' is a base change, we have $a_{\mathfrak{q}}(f') = a_{\overline{\mathfrak{q}}}(f') = a_q(f)$ for primes q that split. We also note that the quadratic character χ_K associated to the quadratic field K satisfies

$$\chi_K(q) = \begin{cases} 1 & \text{if } q \text{ splits in } K, \\ -1 & \text{otherwise.} \end{cases}$$

It therefore follows that we have the congruence

$$a_q(F) \equiv a_q(f)\,(1 + \chi_K(q)) \pmod{\lambda}$$

for $q \nmid N^2 d_K^2 p$.

□

We will now consider an example.

Example 5.6.2. We begin with a classical weight 2 cusp form. Consider $f \in S_2(\Gamma_0(15))$ with

$$f = q - q^2 - q^3 - q^4 + q^5 + q^6 + 3q^8 + q^9 - q^{10} - 4q^{11} + q^{12} - 2q^{13} - q^{15} - q^{16} + 2q^{17} + \ldots.$$

Suppose we wish to take a base change to $K = \mathbb{Q}(\sqrt{2})$. We have $d_K = 8$. Hence the only prime to ramify in K is 2. Since $2 \nmid 15$, we know from Theorem 5.3.1, that there exists a Hilbert modular form $f' \in S_2(\mathfrak{N})$ that is a base change of f with $\mathrm{N}(\mathfrak{N}) = 15^2 = 225$. Checking in the LMFDB we find a Hilbert modular form with label $[225, 15, 15]$ that is a base change of f. This form has the following eigenvalues:

Norm	Prime	Eigenvalue
2	$[2, 2, -\sqrt{2}]$	-1
7	$[7, 7, -2\sqrt{2}+1]$	0
7	$[7, 7, -2\sqrt{2}-1]$	0
9	$[9, 3, 3]$	1
17	$[17, 17, 3\sqrt{2}+1]$	2
17	$[17, 17, 3\sqrt{2}-1]$	2
23	$[23, 23, \sqrt{2}+5]$	0
23	$[23, 23, -\sqrt{2}+5]$	0
25	$[25, 5, 5]$	1
31	$[31, 31, 4\sqrt{2}+1]$	0
31	$[31, 31, -4\sqrt{2}+1]$	0
41	$[41, 41, 2\sqrt{2}-7]$	10
41	$[41, 41, -2\sqrt{2}-7]$	10

Note that at the split primes 7 and 17, we see that f and f' both have the same eigenvalues as expected.

We now use Theorem 5.4.6 to raise the level of this Hilbert modular form, and in the process, we ensure that the associated Galois representation is no longer Galois-invariant. Suppose we choose $\mathfrak{p}$ to be a prime above 7 with label $[7, 7, -2\sqrt{2}+1]$. Note that this is a split prime, so the level raised form will be non Galois-invariant. We have $a_{\mathfrak{p}}(f') = 0$ and $\mathrm{N}(\mathfrak{p}) = 7$. Hence

$$\begin{aligned}
&\ell \mid \left(a_{\mathfrak{p}}(f')^2 - (\mathrm{N}(\mathfrak{p})+1)^2\right) \\
&\ell \mid \left(0^2 - (7+1)^2\right) \\
&\ell \mid (0 - 64) \\
&\ell \mid -64 = -2^6.
\end{aligned}$$

We find g of level $[1575, 105, -30\sqrt{2}+15]$ that satisfies a congruence modulo 2.

Norm	Prime	Eigenvalues of f'	Eigenvalues of g
2	$[2, 2, -\sqrt{2}]$	-1	0
7	$[7, 7, -2\sqrt{2}+1]$	0	-1
7	$[7, 7, -2\sqrt{2}-1]$	0	0
9	$[9, 3, 3]$	1	-1
17	$[17, 17, 3\sqrt{2}+1]$	2	0
17	$[17, 17, 3\sqrt{2}-1]$	2	-4
23	$[23, 23, \sqrt{2}+5]$	0	0
23	$[23, 23, -\sqrt{2}+5]$	0	8
25	$[25, 5, 5]$	1	1
31	$[31, 31, 4\sqrt{2}+1]$	0	2
31	$[31, 31, -4\sqrt{2}+1]$	0	-10
41	$[41, 41, 2\sqrt{2}-7]$	10	-2
41	$[41, 41, -2\sqrt{2}-7]$	10	-2

We can see clearly that the Hecke eigenvalues are congruent modulo 2 at all primes except 2 and 7. These primes are ones where we wouldn't expect the congruence to hold since $2|d_K$ and we chose $p = 7$. We also note that at the split primes, the Hecke eigenvalues of g are different in general. Hence the associated Galois representation ρ_g will not be Galois-invariant. We are therefore able to use Theorem 5.5.1 to produce a Siegel paramodular cusp form.

We know that $\mathrm{N}(\mathfrak{N}) = 1575 = 7 \times 3^2 \times 5^2$ and that $d_K = 8$. Hence Theorem 5.5.1 produces a Siegel paramodular newform $F : \mathcal{H}_2 \to \mathbb{C}$ of weight 2 and paramodular level $100800 = 8^2 \times 7 \times 3^2 \times 5^2$. The theorem also tells us precisely what the Hecke eigenvalues are in terms of those of the Hilbert modular form g. We are interested in the T_p eigenvalues. These are given by $T_p F = a_p(F)F$ where

$$a_p(F) = \begin{cases} p^{k-2}(a_{\mathfrak{p}}(g) + a_{\overline{\mathfrak{p}}}(g)) & \text{if } p \text{ is split,} \\ 0 & \text{if } p \text{ is not split.} \end{cases}$$

Since we have $k = 2$, the eigenvalues at split primes are just the sum of the eigenvalues $a_{\mathfrak{p}}(g)$ and $a_{\overline{\mathfrak{p}}}(g)$. The following table shows the first few eigenvalues of F along with the RHS of the congruence and the difference between the two.

p	$a_p(F)$	$a_p(f)\,(1+\chi_K(p))$	$a_p(F)-a_p(f)\,(1+\chi_K(p))$
2	0	0	0
7	-1	0	-1
9	0	0	0
17	-4	4	-8
23	8	0	8
25	0	0	0
31	-8	0	-8
41	-4	20	-24

We can see clearly from the right hand column that we have a congruence modulo 2 except at the prime 7. However as we have mentioned, this is the prime that we raised the level of the Hilbert modular form by. Hence we wouldn't expect there to be a congruence at this prime.

Remark 5.6.3. Notice that this example had $\ell = 2$. It therefore does not satisfy the condition that $\ell \nmid (p+1)$ that is required in order for the error term of Taylor's level raising result to be ignored. The method however still works and illustrates the general method.

§ 5.7 Comparison with the Bloch-Kato Formula

We have proved the existence of a congruence between the Hecke eigenvalues of a classical cusp form of level N and a Siegel paramodular newform of level $N^2 d_K^2 p$ in Theorem 5.6.1. We might be interested in how we could link this congruence with the Bloch-Kato conjecture. In particular we might expect such a congruence to allow us to construct a non-zero element in a Bloch-Kato Selmer group. This would in turn hopefully give the divisibility of an incomplete L-value. The statement of Theorem 5.6.1 however, has no mention of the divisibility of an Euler factor arising from a partial L-value but we do have a divisibility criterion. We might be interested in whether this divisibility criterion can be linked with a particular L-value. In our main result (Theorem 3.0.1) the modulus came from a divisor of an Euler factor arising from a Dirichlet L-function. This came about because the Dirichlet L-function appeared as the constant term of the Eisenstein series appearing in the congruence. In this case, we have no such analogy as we are now working with two cusp forms. We will see however, that there is a particular L-function with an Euler factor that will be very familiar. The modulus of the congruence will then be a divisor of this Euler factor.

5.7.1 Reduction of the Four-Dimensional Representation

Just as in Section 3.3, we consider the reduction of a Galois representation modulo λ. In this case it will be the four-dimensional Galois representation associated to the Siegel paramodular newform.

Let $k, p, \ell, \lambda, N, d_K, f$ and F be as in Theorem 5.6.1. Assume also that $\ell \neq p$, and let $L = \mathbb{Q}(\{a_n\})$. As in Section 3.3, there exists a continuous representation attached to f (note that we have a different f here) given by:

$$\rho_f = \rho_{f,\lambda} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(L_\lambda),$$

unramified outside $N\ell$, such that if $q \nmid N\ell$ is a prime, and Frob_q is an arithmetic Frobenius element, then

$$\mathrm{Tr}(\rho_f(\mathrm{Frob}_q^{-1})) = a_q(f), \quad \det(\rho_f(\mathrm{Frob}_q^{-1})) = q^{k-1}.$$

We may conjugate so that ρ_f takes values in $\mathrm{GL}_2(\mathcal{O}_\lambda)$ and reduce modulo λ to get a continuous representation

$$\overline{\rho}_f = \overline{\rho}_{f,\lambda} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{F}_\lambda).$$

We assume that this reduced representation is irreducible as this will appear as a composition factor of our four-dimensional representation.

We can also do the same with the Siegel paramodular cusp form F. As in Proposition 5.2.7, we have an attached representation:

$$\rho_F = \rho_{F,\lambda} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_4(L_\lambda),$$

unramified outside $Nd_K p\ell$, such that if $q \nmid Nd_K p\ell$ is a prime, and Frob_q is an arithmetic Frobenius element, then

$$\mathrm{Tr}(\rho_F(\mathrm{Frob}_q^{-1})) = a_q(F), \quad \det(\rho_F(\mathrm{Frob}_q^{-1})) = q^{4k-6}$$

We may assume that this representation is irreducible. We may again conjugate so that this representation takes values in $\mathrm{GL}_4(\mathcal{O}_\lambda)$ and reduce modulo λ to obtain a continuous representation

$$\overline{\rho}_F = \overline{\rho}_{F,\lambda} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_4(\mathbb{F}_\lambda).$$

This depends in general on the choice of invariant $\mathcal{O}_\lambda$-lattice but the composition factors are well-defined. This representation is reducible with composition factors $\overline{\rho}_f$ and $\overline{\rho}_f(\chi_K)$ where χ_K is the quadratic character associated to K. This follows because of the Cebotarev density theorem, the Brauer-Nesbitt theorem and the existence of the congruence in Theorem 5.6.1 as we have $a_p(F) \equiv a_p(f)\,(1 + \chi_K) \pmod{\lambda}$.

Without loss of generality we may choose our invariant $\mathcal{O}_\lambda$-lattice in such a way that $\overline{\rho}_F$ has the form

$$\overline{\rho}_F \sim \begin{bmatrix} \overline{\rho}_f(\chi_K) & * \\ 0 & \overline{\rho}_f \end{bmatrix}.$$

Moreover, an argument of Ribet [Rib3, Proposition 2.1] says that we can also choose our invariant $\mathcal{O}_\lambda$-lattice such that $\overline{\rho}_F$ is realised on a space V such that

$$0 \longrightarrow W(\chi_K) \xrightarrow{\iota} V \xrightarrow{\pi} W \longrightarrow 0$$

is a non-split extension of $\mathbb{F}_\lambda[\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})]$-modules. Note that here W is the space of the representation $\overline{\rho}_f$. Choose a map $s : W \to V$ that is $\mathbb{F}_\lambda$ linear and $x \in W$. As in Section 3.3, for $g \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ consider $g(s(g^{-1}(x))) - s(x)$. As before we have $g(s(g^{-1}(x))) - s(x) \in \ker(\pi)$. Since the sequence is exact, we therefore see that $g(s(g^{-1}(x))) - s(x) \in \mathrm{im}(\iota)$. Then we define $C : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Hom}(\overline{\rho}_f, \overline{\rho}_f(\chi_K))$ such that $C(g)(x) := \iota^{-1}\left(g(s(g^{-1}(x))) - s(x)\right)$. As in Section 3.3, this is a cocycle. The same argument can be used to show this. Choosing a different map s here would result in a different cocycle but it would only differ by a coboundary. We therefore see that we get a unique class $c := [C] \in H^1(G_\mathbb{Q}, \mathrm{Hom}(\overline{\rho}_f, \overline{\rho}_f(\chi_K)))$ independent of the choice of x. Again this is non-zero since the extension is non-split.

Now we have a pairing $\rho_f \times \rho_f \to L_\lambda(1-k)$. This pairing is similar to the Weil pairing for elliptic curves. In particular, it satisfies the same kind of properties, i.e., it is bilinear, Galois-equivariant, non-degenerate and skew-symmetric. We see that $\rho_f \simeq \mathrm{Hom}(\rho_f, L_\lambda(1-k)) \simeq \rho_f^*(1-k)$ where ρ_f^* represents the dual representation. From this it follows that $\rho_f^* \simeq \rho_f(k-1)$. We therefore see that

$$\mathrm{Hom}(\overline{\rho}_f, \overline{\rho}_f(\chi_K)) \simeq \overline{\rho}_f^* \otimes \overline{\rho}_f(\chi_K) \simeq (\overline{\rho}_f \otimes \overline{\rho}_f)(k-1)(\chi_K).$$

In fact, part 2 of Proposition 5.2.7 tells us that ρ_F actually lands inside $\mathrm{GSp}_4(L_\lambda)$. We see that the class c actually lands in $\mathrm{Sym}^2(\overline{\rho}_f)(k-1)(\chi_K)$. Hence $c \in H^1(G_\mathbb{Q}, \mathrm{Sym}^2(\overline{\rho}_f)(k-1)(\chi_K))$. Similarly to Section 3.3, let $V_\lambda = \mathrm{Sym}^2(W)(k-1)(\chi_K)$, $M_\lambda = \mathrm{Sym}^2(\mathcal{M}_\lambda)(k-1)(\chi_K)$, where $\mathcal{M}_\lambda$ is an invariant $\mathcal{O}_\lambda$-lattice in W, and let $A_\lambda = V_\lambda/M_\lambda$. Also let $A[\lambda]$ be the kernel of multiplication by λ in A_λ. Let $i : A[\lambda] \to A_\lambda$ be the inclusion and let $d := i_*(c) \in H^1(G_\mathbb{Q}, A_\lambda)$. In a similar manner to Section 3.3 we will show that i_* is injective and therefore $d \neq 0$. As before we have the short exact sequence

$$0 \longrightarrow A[\lambda] \xrightarrow{i} A_\lambda \xrightarrow{\text{“}\lambda\text{”}} A_\lambda \longrightarrow 0.$$

Again, “λ” means multiplication by some uniformiser for λ. This sequence gives rise to a long exact sequence in cohomology of which we consider the following piece:

$$H^0(G_\mathbb{Q}, A_\lambda) \xrightarrow{\delta} H^1(G_\mathbb{Q}, A[\lambda]) \xrightarrow{i_*} H^1(G_\mathbb{Q}, A_\lambda).$$

Again knowing when i_* is injective is the same as knowing when the image of δ is trivial since the sequence is exact. This is when $H^0(G_\mathbb{Q}, A_\lambda)$ is trivial. Note that $H^0(G_\mathbb{Q}, A_\lambda)$ consists of those elements of A_λ fixed by $G_\mathbb{Q}$. Suppose that $(d_K, N) = 1, \ell \nmid d_K$ and consider $q | d_K$. We will restrict to the local Galois group and then take inertia in there. That is we have $I_q \subset G_{\mathbb{Q}_q} \subset G_\mathbb{Q}$. Now I_q acts trivially on $\mathrm{Sym}^2(W)(k-1)$ since $\mathrm{Sym}^2(W)$ is only ramified at ℓ and primes dividing N, and the $(k-1)$ twist is by the ℓ-adic character χ_ℓ only ramified at ℓ. However I_q does not act trivially on the character (χ_K). Take $\sigma \in I_q$ with $\chi_K(\sigma) = -1$. Then $\sigma(v) = -v \quad \forall\, v \in V_\lambda$. Hence σ is an element of $G_\mathbb{Q}$ which does not fix any non-zero element of A_λ. It follows that $H^0(G_\mathbb{Q}, A_\lambda)$ is trivial. Hence i_* is injective and therefore $d \neq 0$.

We define the Bloch-Kato Selmer group in the same way as in Section 3.3.

Proposition 5.7.1. *Let* $\Sigma = \{q : q|N\}$. *Then* $d \in H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, A_\lambda)$.

We omit the proof here but note that it follows in exactly the same way as Proposition 3.3.3. That is, the necessary local conditions all hold and so $d \in H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, A_\lambda)$. Note that we now require $\ell > 2j + 3k - 2$ in order for the local condition at ℓ to be satisfied. This follows from Lemma 7.2 of [D2]. Note that here we have $j = 0$.

We now consider the λ-part of the Bloch-Kato conjecture. This will give (conjecturally) a condition on there being a non-zero element of $H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))$. Let $L_{\Sigma\cup\{p\}}(\mathrm{Sym}^2(f)(\chi_K), k)$ be the symmetric square L-function with Euler factors at primes $q \in \Sigma \cup \{p\}$ omitted.

Conjecture 5.7.2 (Case of λ-part of Bloch-Kato)**.**

$$\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(\mathrm{Sym}^2(f)(\chi_K), k)}{\Omega}\right)$$
$$= \mathrm{ord}_\lambda\left(\frac{\mathrm{Tam}^0_\lambda(\mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k)(\chi_K))\#H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))}{\#H^0(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))}\right).$$

Here Ω is an appropriately chosen period. We omit the definition of the Tamagawa factor $\mathrm{Tam}^0_\lambda(\mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k)(\chi_K))$, but note that (assuming $\ell > 2k-1$, see [D2, §5](recall that $\lambda|\ell$)), its triviality is a direct consequence of [BlKa, Theorem 4.1(iii)]. Also note that $H^0(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K)) = H^0(G_{\mathbb{Q}}, A_\lambda)$, which we have shown to be trivial. Since this factor and the Tamagawa factor are both trivial, we see that divisibility of $\frac{L_{\Sigma\cup\{p\}}(\mathrm{Sym}^2(f)(\chi_K),k)}{\Omega}$ by λ is equivalent to divisibility of $\#H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))$ by λ. In the same way as Section 3.3 if we could show that $\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(\mathrm{Sym}^2(f)(\chi_K),k)}{\Omega}\right) > 0$, this would imply that $\mathrm{ord}_\lambda\left(\#H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))\right) > 0$. This would then imply there was a non-zero element in the Bloch-Kato Selmer group.

Since we have already constructed a non-zero element in $H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))$, we see that the Bloch-Kato conjecture predicts a λ appearing in $\frac{L_{\Sigma\cup\{p\}}(\mathrm{Sym}^2(f)(\chi_K),k)}{\Omega}$. Also, a similar Bloch-Kato formula holds if we drop the $\cup\{p\}$, however we wouldn't be able to construct a non-zero element in the Bloch-Kato Selmer group in this case. This suggests that we should expect λ to appear in the p-part of the L-value. This gives some evidence for the Bloch-Kato conjecture in this case. We would also like to go the other way. That is, if $\mathrm{ord}_\lambda\left(\frac{L_{\Sigma\cup\{p\}}(\mathrm{Sym}^2(f)(\chi_K),k)}{\Omega}\right) > 0$, with λ appearing in the p-part, we would like λ to divide $\#H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))$.

As usual, since we are raising the level by p, we would like to deduce divisibility of the Euler factor at p arising from the L-value. In fact, in this case since we chose a p that splits in K, we in fact don't actually need the χ_K here since $\chi_K(p) = 1$. Since we

expect the λ to appear in the p-part we consider $L_p(\mathrm{Sym}^2(f), k)$. We therefore expect that if we have $\lambda | L_p(\mathrm{Sym}^2(f), k)$ then, since the other factors cause no cancellation, we must also have $\lambda | H^1_{\Sigma\cup\{p\}}(G_{\mathbb{Q}}, \mathrm{Sym}^2(W/\mathcal{M}_\lambda)(k-1)(\chi_K))$. We also note that the critical values in this case are the odd integers $1, \ldots, k-1$ and the even integers $k, \ldots, 2k-2$, which in this case gives 1 and 2 as critical values since $k = 2$.

Remark 5.7.3. Congruences such as those studied in [D2] and [U] give evidence for Conjecture 5.7.2 in the case of higher weights. Here we are dealing with the low weight case of $k = 2$. In the higher weight cases there exist congruences between Siegel cusp forms and Klingen Eisenstein series. In the case of weight 2 however, the weight is too low for such a congruence to exist and we instead end up with a congruence between a Siegel cusp form and a classical cusp form.

Now that we have linked our congruence with the Bloch-Kato conjecture, we see that we should expect divisibility of an Euler factor arising from a symmetric square L-function to give a congruence as in Theorem 5.6.1. We will see in the next section that this is exactly the case.

5.7.2 The Symmetric Square L-Function

Suppose we have a normalised cuspidal Hecke eigenform $f = \sum_{n=1}^{\infty} a_n q^n \in S_k(\Gamma_0(N))$. Associated to this f is the usual L-function

$$\begin{aligned} L_f(s) = \sum_{n=1}^{\infty} a_n n^{-s} &= \prod_p \left(1 - a_p p^{-s} + p^{k-1-2s}\right)^{-1} \\ &= \prod_p \left[(1-\alpha_p p^{-s})(1-\beta_p p^{-s})\right]^{-1}. \end{aligned}$$

From this we can deduce that $\alpha_p + \beta_p = a_p$ and $\alpha_p \beta_p = p^{k-1}$. In a similar way we can define the symmetric square L-function associated to f. This is given by

$$L(\mathrm{Sym}^2(f), s) = \prod_p \left[(1-\alpha_p^2 p^{-s})(1-\beta_p^2 p^{-s})(1-\alpha_p\beta_p p^{-s})\right]^{-1}.$$

We will now expand out this expression and simplify using the expressions deduced for α_p and β_p. We first note that $\alpha_p^2 + \beta_p^2 = (\alpha_p + \beta_p)^2 - 2\alpha_p\beta_p = a_p^2 - 2p^{k-1}$. We have

$$\begin{aligned} L(\mathrm{Sym}^2(f), s) &= \prod_p \left[(1-\alpha_p^2 p^{-s})(1-\beta_p^2 p^{-s})(1-\alpha_p\beta_p p^{-s})\right]^{-1} \\ &= \prod_p \left[(1-\alpha_p^2 p^{-s})(1-\beta_p^2 p^{-s})(1-p^{k-1-s})\right]^{-1} \\ &= \prod_p \left[(1-(\alpha_p^2+\beta_p^2)p^{-s} + \alpha_p^2\beta_p^2 p^{-2s})(1-p^{k-1-s})\right]^{-1} \\ &= \prod_p \left[(1-a_p^2 p^{-s} + 2p^{k-1-s} + p^{2k-2-2s})(1-p^{k-1-s})\right]^{-1}. \end{aligned}$$

We now concentrate on the first factor in this expression for a particular choice of p. Multiplying through by p^s to clear denominators we get

$$\begin{aligned} p^s(1 - a_p^2 p^{-s} + 2p^{k-1-s} + p^{2k-2-2s}) &= p^s - a_p^2 + 2p^{k-1} + p^{2k-2-s} \\ &= -\left[a_p^2 - (p^{2k-2-s} + 2p^{k-1} + p^s)\right]. \end{aligned}$$

Note that in our case we have $f \in S_2(\Gamma_0(N))$. Hence when $k = 2$ this factor reduces to

$$-\left[a_p^2 - (p^{2-s} + 2p + p^s)\right].$$

Note now that if we take $s = 2$ we get the following:

$$-\left[a_p^2 - (1 + 2p + p^2)\right] = -\left[a_p^2 - (p+1)^2\right].$$

Recall that it makes sense to consider $s = 2$ by our argument in the previous section. Notice how if we assume there is an ℓ dividing this factor, we have exactly the condition for there to exist a level raising congruence. The modulus of the congruence in Theorem 5.6.1 is then a prime $\lambda|\ell$. We could therefore say that we expect a congruence to exist when we divide an Euler factor coming from the symmetric square L-function associated to the cusp form f.

5.7.3 Galois Deformations

There is actually another way that we could construct a non-zero element in a Bloch-Kato Selmer group using the information we know about the congruence in Theorem 5.6.1. Recall that in our case the condition for the existence of a level raising congruence of Hilbert modular forms is the same as the condition for there to be such a congruence between two classical cusp forms. This follows since we are working with a quadratic field and $\mathrm{N}(\mathfrak{p}) = p$ where $p = \mathfrak{p}\overline{\mathfrak{p}}$ is a split prime in the quadratic field K. However in order for our method to work we need to level raise after base changing to a Hilbert modular form. Since the condition is the same in both cases however, we can consider the case where we first level raise and see what happens when we consider the Bloch-Kato formula. The argument we will use is similar to those used in the theory of Galois deformations.

Take $f \in S_2(\Gamma_0(N))$. Assume this satisfies a level raising congruence with some $g \in S_2(\Gamma_0(Np))$. That is we have $a_p(f) \equiv a_p(g) \pmod{\lambda}$ for some $\lambda|\ell$ where $\ell|(a_p(f)^2 - (p+1)^2)$. Then we have $\overline{\rho}_f \simeq \overline{\rho}_g$ where these are the associated Galois representations reduced modulo λ. We assume that these representations are irreducible. Since we know that a congruence exists we may choose a maximal $r \geq 1$ such that ρ_f and ρ_g are the same modulo λ^r. We then know that ρ_f and ρ_g differ modulo λ^{r+1}. We have

$$\rho_g(\sigma) \equiv \rho_f(\sigma)\,(I + \lambda^r \theta_g(\sigma)) \pmod{\lambda}^{r+1},$$

where $\theta_g(\sigma) \in M_2(\mathbb{F}_\lambda) = \mathrm{Hom}(\mathbb{F}_\lambda^2, \mathbb{F}_\lambda^2)$. Here $\theta_g(\sigma)$ actually lies in the space of trace 0 matrices since $\det(\overline{\rho}_f) = \det(\overline{\rho}_g)$. We now consider $\rho_g(\sigma\tau)$ and show that we can produce a non-trivial cocycle. We have

$$\begin{aligned}\rho_g(\sigma\tau) &\equiv \rho_f(\sigma\tau)\left(I + \lambda^r\theta_g(\sigma\tau)\right) \pmod{\lambda}^{r+1} \\ &\equiv \rho_f(\sigma)\rho_f(\tau)\left(I + \lambda^r\theta_g(\sigma\tau)\right) \pmod{\lambda}^{r+1}\end{aligned}$$

But we also have

$$\begin{aligned}\rho_g(\sigma\tau) &= \rho_g(\sigma)\rho_g(\tau) \\ &\equiv \rho_f(\sigma)\left(I + \lambda^r\theta_g(\sigma)\right)\rho_f(\tau)\left(I + \lambda^r\theta_g(\tau)\right)\end{aligned}$$

We deduce that

$$\theta_g(\sigma\tau) = \overline{\rho}_f(\tau)^{-1}\theta_g(\sigma)\overline{\rho}_f(\tau)\theta_g(\tau),$$

and so

$$\theta_g(\sigma\tau) = \tau(\theta_g(\sigma))\theta_g(\tau).$$

Here the action of τ is the adjoint $\overline{\rho}_f$ action $\mathrm{Ad}_{\overline{\rho}_f}$. This action is really the Ad_0 action by the above argument. This gives us a non-trivial cocycle which then leads to a non-zero element lying in some Bloch-Kato Selmer group by a similar argument to those we used in Sections 3.3 and 5.7.

Although this argument was for a congruence between two classical cusp forms $f \in S_k(\Gamma_0(N))$ and $g \in S_k(\Gamma_0(Np))$ with $a_p(f) \equiv a_p(g) \pmod{\lambda}$, we still actually end up with the same conditions arising in the Bloch-Kato formula as we would for the congruence in Theorem 5.6.1. This arises because of the level raising condition being the same in both cases and the fact that it is analogue of the Euler factor of an L-function.

§ 5.8 The Vector Valued Case

Now that we have dealt with the scalar valued case, we move on to the more general case of vector valued Siegel modular forms. This will mean that we can start with a classical modular form $f \in S_k(\Gamma_0(N))$ with $k > 2$. Running through our process, we will base change to a Hilbert modular form of parallel weight k, which we will then level raise to get a non Galois-invariant Hilbert modular form. This is the point at which we need a different approach. We will no longer be able to make use of the theta lifting result of [JLR] since the only parallel weight that is covered by the result is parallel weight 2. As we know, this leads to a scalar valued Siegel modular form. We will however be able to slightly modify the work of [JLR] in order to obtain a vector valued Siegel modular form. In order to do this we will need to consider L-parameters and L-packets. Before we do this however, we will explain the construction of the Siegel modular form.

5.8.1 Constructing the Siegel Modular Form

In this section we will give an outline of the construction of the Siegel modular form. There will be many new concepts introduced in this section, most of which will be properly defined or explained in later sections. The method we will use is the same as that of Johnson-Leung and Roberts from Section 3 of [JLR]. Let K be a real quadratic field and π_0 a cuspidal, irreducible, automorphic representation of $\mathrm{GL}_2(\mathbb{A}_K)$ with trivial central character, where $\mathbb{A}_K$ is the adeles of K. Note that this will be the automorphic representation associated to our level raised Hilbert modular form. In particular we will have taken a base change of our classical weight k cusp form to a Hilbert cusp form over the field K. For every place v of $\mathbb{Q}$ we define $\pi_{0,v} = \otimes_{w|v}\pi_{0,w}$. In particular, for a prime v that is inert in K, $\pi_{0,v} = \pi_{0,(w)}$, for a prime v that splits in K, $\pi_{0,v} = \pi_{0,w_1} \otimes \pi_{0,w_2}$ and for a prime v that ramifies in K, $\pi_{0,v} = \pi_{0,w}$.

Let $\varphi(\pi_{0,v}) : W_{\mathbb{Q}_v} \to \mathrm{GSp}_4(\mathbb{C})$ be the L-parameter associated to $\pi_{0,v}$. Here $W_{\mathbb{Q}_v}$ is the Weil group of $\mathbb{Q}_v$. This is a dense subgroup of $\mathrm{Gal}(\bar{\mathbb{Q}}_v/\mathbb{Q}_v)$, and consists of all elements whose image in the Galois group of the residue field is an integral power of the Frobenius automorphism. Note that these L-parameters, along with the L-packets, will be defined in Section 5.8.2. By the remark in Section 2.2 of [Bla], the representation $\pi_{0,w}$ is tempered for all finite places w of K. Let $\Pi(\varphi(\pi_{0,v}))$ be the L-packet of tempered, irreducible, admissible representations of $\mathrm{GSp}_4(\mathbb{Q}_v)$ with trivial central character associated to $\varphi(\pi_{0,v})$ as in [Rob]. For finite $v = p$, the packet $\Pi(\varphi(\pi_{0,p}))$ coincides with the packet associated to $\varphi(\pi_{0,p})$ in [GT]. This packet may contain several representations but it contains a unique generic representation π_p of $\mathrm{GSp}_4(\mathbb{Q}_p)$. We will show in Section 5.8.2 that $\Pi(\varphi(\pi_{0,\infty}))$ contains the lowest weight representation of $\mathrm{GSp}_4(\mathbb{R})$ denoted by $\pi_\lambda[0]$. This is a particular type of discrete series representation where λ is a vector since we are in the vector valued case.

We note that the Langlands correspondence between the L-parameters and the representations of $\mathrm{GSp}_4(\mathbb{R})$ means that the L-parameter, in a sense, labels the representation. Now, $\mathrm{GSp}_4(\mathbb{R})$ is not a compact group, however $\mathrm{Sp}_4(\mathbb{R}) \subset \mathrm{GSp}_4(\mathbb{R})$ is a semisimple group whose maximal compact subgroup is a connected compact Lie group. In order to understand the representations of $\mathrm{GSp}_4(\mathbb{R})$, we may restrict to the subgroup $\mathrm{Sp}_4(\mathbb{R})$ and then further restrict to its maximal compact subgroup. The Lie algebra of $\mathrm{Sp}_4(\mathbb{R})$ is given by

$$\begin{aligned}\mathfrak{g} &= \{X \in M_4(\mathbb{R}) | XJ + JX^T = 0\} \\ &= \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in M_4(\mathbb{R}) |\, B = B^T, C = C^T, A = -D^T \right\}.\end{aligned}$$

The standard maximal compact subgroup K_∞ of $\mathrm{Sp}_4(\mathbb{R})$ is

$$K_\infty = \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in \mathrm{GL}_4(\mathbb{R}) |\, A^TA + B^TB = \mathbf{1}, A^TB = B^TA \right\}.$$

We have $K_\infty \simeq U(2)$ via $\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \mapsto A + iB$.

If we consider the restriction of our representation to the maximal compact subgroup $K_\infty \simeq U(2)$, we obtain a set of finite dimensional pieces. This restriction makes our calculations easier as the representation theory of compact Lie groups is completely understood and we do not lose any information about the larger group $\mathrm{GSp}_4(\mathbb{R})$. If we consider the maximal torus T inside K_∞, we can determine the representations of K_∞, from those of T. In general, such tori are isomorphic to several copies of the circle S^1 and therefore the representations of T can be determined from the representations of the circle. Since T is commutative, Schur's lemma says that each irreducible representation ρ of T is one-dimensional:

$$\rho : T \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times.$$

It must also map into $S^1 \subset \mathbb{C}$ since T is compact. We let $\mathfrak{t}$ be the Lie algebra of T and we let points $h \in T$ be denoted

$$h = e^H,\ H \in \mathfrak{t}.$$

In these coordinates, ρ has the form

$$\rho(e^H) = e^{i\lambda(H)}$$

for some linear functional λ on $\mathfrak{t}$. As it stands the linear functionals do not give a well-defined map of T into S^1. This is because the exponential map $H \to e^H$ is not injective. We let Γ denote the kernel of the exponential map:

$$\Gamma = \{H \in \mathfrak{t} |\, e^{2\pi H} = \mathrm{Id}\},$$

where Id is the identity element of $\mathfrak{t}$. Then for the linear functional λ to give a well-defined map ρ, it must satisfy

$$\lambda(H) \in \mathbb{Z}, H \in \Gamma.$$

Such a linear functional is called an analytically integral element.

The irreducible representations of K_∞ break up as a direct sum of irreducible representations of T. As we have seen, these representations are described by linear functionals λ. Let Σ be a finite-dimensional irreducible representation of K_∞. If a given functional λ appears in the restriction of Σ to T, we call λ a weight of Σ. These weights can then be determined by considering a root system for K_∞. When we restrict a representation such as Σ to T we have an ordering on the weights of the irreducible representations in the direct sum. This ordering gives rise to a highest weight for each irreducible piece. The representation with lowest highest weight is said to have minimal K_∞-type. The irreducible representations of $K_\infty \simeq U(2)$ are parametrised by elements in the weight lattice $\Lambda = \mathbb{Z}e_1 + \mathbb{Z}e_2$, modulo the action of the real Weyl group (to be defined later). Since the Weyl group acts by permuting the e_i, the irreducible representations of K_∞ are are in 1-1 correspondence with the weight

$$\lambda = r_1 e_1 + r_2 e_2,\ r_1, r_2 \in \mathbb{Z}, r_1 \geq r_2.$$

We denote the finite-dimensional irreducible representation of K_∞ corresponding to this λ by τ_λ. For us, $\lambda = (k-1, 0)$; so we have $r_1 = k-1$ and $r_2 = 0$. We will show this in Section 5.8.2. This will mean that we actually have a limit of discrete series representation. There is an irreducible, admissible representation π_λ of $\mathrm{Sp}_4(\mathbb{R})$ which is exactly this limit of discrete series representation. This representation has minimal K_∞-type τ_λ for our choice of λ. In other words, the piece with lowest highest weight of the restriction of π_λ to K_∞ is τ_λ. We also note that π_λ is itself a piece of the restriction of a larger limit of discrete series representation of $\mathrm{GSp}_4(\mathbb{R})$. This representation is denoted $\pi_\lambda[0]$. We have $\pi_\lambda[0] \cong \pi_\lambda \oplus \hat{\pi}_\lambda$. We set $\pi_\infty = \pi_\lambda[0]$.

5.8.1.1 Siegel to Automorphic

Our next step will be to produce an automorphic representation from which we can extract the Siegel modular form that we are interested in as a particular vector. Before we do this however we will describe the general process in the opposite direction. That is, we will show how given a Siegel modular form, we can produce an automorphic form. We will then be able to essentially run the argument backwards to determine our Siegel modular form. Here we will describe the process for general Siegel modular forms of genus n, but we note that for us $n = 2$.

Let (ρ, V) be a finite-dimensional rational representation of $\mathrm{GL}_n(\mathbb{C})$. As defined in Section 5.2, a vector valued Siegel modular form of genus n and weight ρ is a holomorphic function $F : \mathcal{H}_n \to V$ such that

$$F(\gamma\langle Z\rangle) = \rho(j(\gamma, Z))F(Z) \text{ for all } \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma, Z \in \mathcal{H}_n,$$

where Γ is some congruence subgroup of $\mathrm{Sp}_{2n}(\mathbb{R})$ and $j(\gamma, Z) = (CZ + D)$.

Assuming ρ is irreducible, we may associate a function Φ on $\mathrm{GSp}_{2n}(\mathbb{A})$ to the modular form F. Let $m \in \mathbb{R}$ be the number such that $\rho(s) = s^{2m}\mathrm{id}_V$, for each scalar matrix $s = \mathrm{diag}(s, \ldots, s) \in \mathrm{GL}_n(\mathbb{C}), s > 0$.

Remark 5.8.1. In the scalar valued case $\rho = \det^k$. We therefore have $\rho(s) = \det(s)^k = (s^n)^k = s^{nk}$. Hence m must satisfy $s^{nk} = s^{2m}$. It follows that in this case $m = nk/2$. Hence in the genus 2 case we have $m = k$.

Remark 5.8.2. We also note that the definition of the slash operator in [AS] differs from the definition in this thesis. This therefore affects the choice of m. We will point out the difference later when we define our Siegel modular form. The general theory however remains the same.

Let $G = \mathrm{GSp}(2n)$. We note that using strong approximation (see [Kne]) we have

$$G(\mathbb{A}) = G(\mathbb{Q})G(\mathbb{R})^+ \prod_{p<\infty} K_p,$$

where $G(\mathbb{R})^+$ denotes the elements of $G(\mathbb{R})$ with positive multiplier, and K_p is $G(\mathbb{Z}_p)$ for all but finitely many p and $G(K(p^n))$ otherwise, where $K(p^n)$ is the local paramodular group. We note that here K_p could be any open compact subgroup of $G(\mathbb{Q}_p)$ as explained in [SS], but the local paramodular group is the one that we need for our case. We write an element $g \in G(\mathbb{A})$ as

$$g = g_{\mathbb{Q}} g_\infty k_0 \text{ with } g_{\mathbb{Q}} \in G(\mathbb{Q}), g_\infty \in G(\mathbb{R})^+, k_0 \in K_0, \tag{5.1}$$

where $K_0 = \prod_{p<\infty} K_p$. We may now define

$$\tilde{\Phi}(g) = \mu(g_\infty)^m \rho(CI + D)^{-1} F(g\langle I \rangle),$$

with $g = g_{\mathbb{Q}} g_\infty k_0$ as in (5.1) and $g_\infty = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

The factor $\mu(g_\infty)^m$ ensures that $\tilde{\Phi}$ descends to a function on $\mathrm{PGSp}_{2n}(\mathbb{A})$. Further, $\tilde{\Phi}(\gamma g k_0) = \tilde{\Phi}(g)$ for all $\gamma \in G(\mathbb{Q})$ and $k_0 \in K_0$, and

$$\tilde{\Phi}(g k_\infty) = \rho(k_\infty)^{-1} \tilde{\Phi}(g) \text{ for all } k_\infty \in K_\infty \simeq U(n).$$

As it stands, the function $\tilde{\Phi}$ is vector valued, but we would like to obtain a scalar valued function. We let L be any non-zero linear form on V, the space of ρ, and define

$$\Phi(g) = L(\tilde{\Phi}(g)),\ g \in G(\mathbb{A}).$$

It can be shown that the map $F \mapsto \Phi$ is a norm-preserving map of Hilbert spaces from $S_\rho(\Gamma)$ to $L^2(Z(\mathbb{A})G(\mathbb{Q})\backslash G(\mathbb{A}))$ whose image is contained in the space of cuspidal functions. That is Φ can be realised as being inside the space of functions $L^2_0(Z(\mathbb{A})G(\mathbb{Q})\backslash G(\mathbb{A}))$. We then consider the subspace of this L^2_0 space spanned by all the right translates of Φ. Note that this means the choice of linear form L above is arbitrary. We let π be any irreducible constituent of this unitary representation. Then π is an automorphic representation of $G(\mathbb{A})$ which is trivial on $Z(\mathbb{A})$. We may therefore consider π as an automorphic representation of $\mathrm{PGSp}_{2n}(\mathbb{A})$. Let

$$\pi = \bigotimes_v \pi_v$$

be the decomposition of π into local representations. For finite places, these are irreducible representations π_p of the local groups $G_p = G(\mathbb{Q}_p)$. Note that this is a restricted tensor product. Because Φ is right K_p-invariant at each finite place p, the representation π_p is spherical (to be defined later) for every such p. As we will see, these representations take a particular form that tell us about the related Satake parameters. As explained in [AS, §4.5, p.196], π_∞ is the lowest weight representation of $G(\mathbb{R})$ with minimal K_∞-type τ_λ for some weight λ. The representation τ_λ contains a vector v_0 that is annihilated by the compact positive root vectors. This v_0 is a highest weight vector; further, we have $\Phi = v_0$.

5.8.1.2 Automorphic to Siegel

We have now discussed the process of going from a cuspidal Siegel modular form to an automorphic representation. We now wish to go in the other direction. We note that with some careful choices we can do this. We define π to be the restricted tensor product

$$\pi = \bigotimes_v \pi_v,$$

where π_v is the unique generic representation inside $\Pi(\varphi_{0,v}))$ for finite v and π_∞ is the lowest weight representation $\pi_\lambda[0]$. Note that these are all representations of $\mathrm{GSp}_4(\mathbb{R})$. What we have done here is create a tensor product of local representations for each place v of $\mathbb{Q}$. It isn't immediately obvious that this should create an automorphic representation. However, by Theorem 8.6 of [Rob], π is a cuspidal, irreducible, admissible, automorphic representation of $\mathrm{GSp}_4(\mathbb{A}_\mathbb{Q})$ with trivial central character. We note that for our case we take $F = \mathbb{Q}$ and $E = K$, then part (1) gives us the result.

Now that we have the automorphic representation, we need to identify the correct vector inside this representation. For each prime p of $\mathbb{Q}$, let Φ_p be the local paramodular newform in π_p, and let $\Phi_\infty \in \pi_\infty = \pi_\lambda[0]$ be the highest weight vector v_0 lying in τ_λ. By [RS2], Φ_p exists and is unique up to scalars; we may assume that for almost all p, Φ_p is the unramified vector used to define the tensor product. Note that the L-packet is unramified for almost all primes and contains spherical representations (to be defined in Section 5.8.4). We set

$$\Phi = \bigotimes_v \Phi_v.$$

This tensor product is now the Siegel modular form viewed as an automorphic form, so we have $\Phi : \mathrm{GSp}_4(\mathbb{A}) \to \mathbb{C}$. This Φ is a scalar valued function that we can view as lying inside the space of cuspidal functions $L^2_0(Z(\mathbb{A})G(\mathbb{Q})\backslash G(\mathbb{A}))$. We now wish to move from the scalar valued Φ to a vector valued function $\tilde{\Phi}$. As explained in the discussion preceding Lemma 6.2 of [D2], this can be achieved in the vector valued case. The results of the scalar valued case need some slight modifications. We do not go into the details here. Given that we now have $\tilde{\Phi}$ as defined in Section 5.8.1.1, we will define the Siegel modular form in classical notation. Let (ρ, V) be the representation $\mathrm{Sym}^{k-2}(\mathbb{C}^2)\otimes\det^2$. We will see later in Section 5.8.2 why this is the correct representation. Define $F : \mathcal{H}_2 \to V$ by $F(Z) = \mu(g)^{-(k-3)}\rho(j(g,I))\tilde{\Phi}(g_\infty)$ where $g \in \mathrm{GSp}_4(\mathbb{R})^+$ is such that $g\langle I\rangle = Z$ and $I = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \in \mathcal{H}_2$.

Remark 5.8.3. We note that in our definition of F we have the factor $\mu(g)^{-m}$ where $m = k-3$. This is in order to match our definition of slash operator so that we have appropriate cancellation and therefore trivial central character. If we had instead used the normalisation used in [AS] we would have ended up with $m = (k+2)/2$. Note that for the representation $\mathrm{Sym}^j(\mathbb{C}^2)\det^\kappa$ we have $m = j + \kappa - 3$ in our normalisation and $m = (j + 2\kappa)/2$ in the normalisation of [AS].

By a generalisation of Lemma 7 of [AS] to the vector valued case, F is holomorphic. Since we started with an automorphic form whose local pieces comprised of paramodular newforms and have run through the argument of Section 5.8.1.1 backwards, it follows that $F \in S_\rho^{\text{new}}(K(M))$. Again, we will see in Section 5.8.2 that M is the same as in the scalar case; that is $M = N^2 d_K^2 p$.

Remark 5.8.4. We note that we could have chosen any representation at the finite places and the restricted tensor product would still have given us an automorphic representation. Recall that the L-packet $\Pi(\varphi(\pi_{0,p}))$ may contain several representations but contains a unique generic representation for each prime p. Our choice of the unique generic representation is what leads us to a Siegel paramodular newform.

Now that we have shown how we will construct the Siegel newform, we go into more details on the local conditions. That is, we now consider the L-parameters and L-packets that were critical in this construction. This will explain how we determined both the level and weight of the Siegel paramodular form. They will also help us determine the Hecke eigenvalues of F. We will see exactly how in Section 5.8.4.

5.8.2 L-parameters and L-packets

As we have seen, the way to obtain the correct Siegel modular form together with its Hecke eigenvalues is by considering the automorphic representation obtained by the process outlined above. The Siegel modular form that we want is then associated to a particular vector in this automorphic representation. We will need to consider the local components at each finite place and at the infinite places. Each of these pieces has an associated L-parameter. We then package together certain isomorphism classes of these representations to form an L-packet. From this packet we can then extract the necessary information. The L-parameters at the finite places will tell us about the Hecke eigenvalues of the Siegel modular form, whereas the L-parameters at the infinite places will tell us about the weight of the Siegel modular form. We also note that the level of the modular form is determined by the local behaviour at finite places.

In constructing the representation of $\mathrm{GSp}_4(\mathbb{R})$ we can use both the work in Section 4 of [JLR] and the examples on pages 206-207 of [Mor]. In order to determine the weight of our Siegel modular form we must consider the infinite places. Here we are dealing with discrete series representations; or rather, limit of discrete series representations as it will turn out in our case.

Define two-dimensional representations $\varphi_{\mu,N} : W_{\mathbb{R}} \to \mathrm{GL}_2(\mathbb{C}) \quad (\mu \in \mathbb{C}, N \in \mathbb{Z}_{\geq 0})$ as follows:

$$\varphi_{\mu,N}\left(re^{i\theta}\right) = \begin{pmatrix} r^{2\mu-N}e^{-iN\theta} & 0 \\ 0 & r^{2\mu-N}e^{iN\theta} \end{pmatrix}, \quad \varphi_{\mu,N}(j) = \begin{pmatrix} 0 & (-1)^N \\ 1 & 0 \end{pmatrix}.$$

Here $W_{\mathbb{R}}$ is the Weil group of $\mathbb{R}$, defined by $W_{\mathbb{R}} = \mathbb{C}^\times \sqcup \mathbb{C}^\times \cdot j \, (jz = \overline{z}j, \, j^2 = -1)$. The representation $\varphi_{\mu,N}$ is irreducible if $N > 0$, which for us, will always be the case.

Our main work will involve applying our case to two examples from [Mor, p 207]. In order to do this however we first consider a simpler example to introduce some notation that will be necessary in both these examples and in later work.

This is part of [Mor, §1, Ex1]. Consider $G = \mathrm{SL}_2(\mathbb{R})$. Here we may fix an Iwasawa decomposition $G = NAK$ as follows:

$$N := \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} | x \in \mathbb{R} \right\}, \quad A := \left\{ \begin{pmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{pmatrix} | y > 0 \right\},$$

$$K := \left\{ r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} | \theta \in \mathbb{R} \right\}.$$

As a $\mathbb{C}$-basis of the Lie algebra $\mathfrak{sl}_2(\mathbb{C})$, we take

$$H := \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}, \quad X_\pm := \frac{1}{2} \begin{pmatrix} 1 & \pm\sqrt{-1} \\ \pm\sqrt{-1} & -1 \end{pmatrix}.$$

Then $\{H, X_+, X_-\}$ is an $\mathfrak{sl}_2$-triplet. For each $k \in \mathbb{C}$, consider the Verma module

$$M(k) = U(\mathfrak{g}) \otimes_{U(\overline{\mathfrak{b}})} \mathbb{C}_k \quad \text{with} \quad \overline{\mathfrak{b}} = \mathbb{C} \cdot H \oplus \mathbb{C} \cdot X_-.$$

Here $\mathbb{C}_k = \mathbb{C}_{v_k}$ is a one-dimensional $\overline{\mathfrak{b}}$-module characterised by $H \cdot v_k = k v_k$ and $X_- \cdot v_k = 0$. It is well known that $M(k)$ is unitarisable if $k \in \mathbb{Z}_{>0}$. In other words, there exists an irreducible unitary representation D_k^+ of G such that its underlying $(\mathfrak{g}, K)$-module is $M(k)$. We denote by D_k^- the contragredient representation of D_k^+. The representation $D_k^\pm (k \geq 2)$ is called the discrete series representation with Blattner parameter $\pm k$ (to be defined later). These representations are limit of discrete series representations.

We now consider the two examples from [Mor, p 207]. We note that in the following $\hat{G}$ denotes the Langlands dual of the reductive algebraic group G.

Example 5.8.5 ($G = \mathrm{GL}_2(\mathbb{R})$)**.** In this case, $\hat{G} \simeq \mathrm{GL}_2(\mathbb{C})$. We denote by $\pi = D_k[c] (k \geq 1, c \in \mathbb{C})$ the irreducible admissible representation of $\mathrm{GL}_2(\mathbb{R})$ characterised by $\pi|_{\mathrm{SL}_2(\mathbb{R})} \simeq D_k^+ \oplus D_k^-$ and $\pi(zI_2) = z^c \times \mathrm{id}(z > 0)$. Then we have

$$\Pi_{\varphi_{\mu,N}}(\mathrm{GL}_2(\mathbb{R})) = \{D_k[c]\}, \quad \text{with } \mu = (c + k - 1)/2, N = k - 1.$$

Example 5.8.6 ($G = \mathrm{GSp}_4(\mathbb{R})$)**.** In this case, $\hat{G} \simeq \mathrm{GSp}_4(\mathbb{C})$. Let $\pi_\lambda[c]$ be the representation of $\mathrm{GSp}_4(\mathbb{R})$ with $\pi_\lambda[c]|_{\mathrm{Sp}_4(\mathbb{R})} \simeq \pi_\lambda \oplus \hat{\pi}_\lambda$ and $\pi(zI_4) = z^c \times \mathrm{id}(z > 0)$. Note that π_λ is a particular discrete series representation of $\mathrm{Sp}_4(\mathbb{R})$. For $\lambda_1 > \lambda_2 > 0$, we consider the Langlands parameter $\varphi : W_\mathbb{R} \to \hat{G} = \mathrm{GSp}_4(\mathbb{C})$ defined by

$$\varphi(w) = \begin{pmatrix} a_1 & 0 & b_1 & 0 \\ 0 & a_2 & 0 & b_2 \\ c_1 & 0 & d_1 & 0 \\ 0 & c_2 & 0 & d_2 \end{pmatrix} \quad \text{with} \quad \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} = \varphi_{\mu_i, N_i}(w) \quad (i = 1, 2),$$

where we set

$$\mu_1 = (c+\lambda_1-\lambda_2)/2, N_1 = \lambda_1-\lambda_2; \mu_2 = (c+\lambda_1+\lambda_2)/2, N_2 = \lambda_1+\lambda_2.$$

Then the corresponding L-packet consists of two elements:

$$\Pi_\varphi\left(\mathrm{GSp}_4(\mathbb{R})\right) = \{\pi_{(\lambda_1,\lambda_2)}[c], \pi_{(\lambda_1,-\lambda_2)}[c]\}.$$

In our case we have a parallel weight k Hilbert modular form which at each infinite place, has an associated irreducible admissible representation of $\mathrm{GL}_2(\mathbb{R})$. For both places, this is $D_k[c]$, where $c = 0$ since we have trivial central character. Using example 5.8.5 we see the associated L-parameter has $\mu = \frac{k-1}{2}$ and $N = k-1$. In order to obtain the four-dimensional representation that we require we will be inducing the two-dimensional representation associated to our Hilbert modular form. Strictly speaking this is by automorphic induction from $\mathrm{GL}_2(\mathbb{A}_K)$ to $\mathrm{GSp}_4(\mathbb{A}_\mathbb{Q})$ and amounts to combining the two representations at the infinite places of K to get a four-dimensional representation. Hence using example 5.8.6 we will have two copies of the representation $\varphi_{\mu,N}$ with $\mu = \frac{k-1}{2}$ and $N = k-1$.

We note that

$$\begin{aligned}\varphi_{\frac{k-1}{2},k-1}\left(re^{i\theta}\right) &= \begin{pmatrix} r^{(k-1)-(k-1)}e^{-i(k-1)\theta} & 0 \\ 0 & r^{(k-1)-(k-1)}e^{i(k-1)\theta} \end{pmatrix} \\ &= \begin{pmatrix} e^{i(1-k)\theta} & 0 \\ 0 & e^{i(k-1)\theta} \end{pmatrix}.\end{aligned}$$

Alternatively we can write

$$\varphi_{\frac{k-1}{2},k-1}(z) = \begin{pmatrix} (z/\overline{z})^{\frac{1-k}{2}} & 0 \\ 0 & (z/\overline{z})^{\frac{k-1}{2}} \end{pmatrix},$$

with $z = re^{i\theta}$. Following example 5.8.6, we therefore have the L-parameter $\varphi : W_\mathbb{R} \to \mathrm{GSp}_4(\mathbb{C})$ given by

$$\varphi(z) = \begin{pmatrix} (z/\overline{z})^{\frac{1-k}{2}} & 0 & 0 & 0 \\ 0 & (z/\overline{z})^{\frac{1-k}{2}} & 0 & 0 \\ 0 & 0 & (z/\overline{z})^{\frac{k-1}{2}} & 0 \\ 0 & 0 & 0 & (z/\overline{z})^{\frac{k-1}{2}} \end{pmatrix}.$$

We note that in the example 5.8.6, we have $\mu_1 = \frac{c+\lambda_1-\lambda_2}{2}, N_1 = \lambda_1-\lambda_2, \mu_2 = \frac{c+\lambda_1+\lambda_2}{2}$, and $N_2 = \lambda_1+\lambda_2$. Since we know that $c = 0$, $\mu_1 = \mu_2 = \frac{k-1}{2}$ and $N_1 = N_2 = k-1$, we see that $\lambda_1 = k-1$ and $\lambda_2 = 0$. Although example 5.8.6 deals with the case $\lambda_1 > \lambda_2 > 0$ this process is actually still applicable. We are working in the case of limit of discrete series representations. See Section 5.8.3 for more details. Hence the L-packet consists of a single element:

$$\Pi_\varphi(\mathrm{GSp}_4(\mathbb{R})) = \{\pi_{(k-1,0)}[0]\}.$$

Now that we know what the L-parameters and L-packets look like in our case, we need to consider Harish-Chandra parameters and Blattner parameters in order to determine the weight that our Siegel modular form will have. In 1965-1966 Harish-Chandra classified the discrete series representations of connected semisimple groups G. It turns out that such a group G has discrete series representations if and only if the rank of G is equal to the rank of some maximal compact subgroup K. The Harish-Chandra parameter is then a way to measure this. We also have some associated parameters called Blattner parameters. If λ denotes the Harish-Chandra parameter, then the Blattner parameter is given by

$$\Lambda = \lambda + \delta_{\mathrm{nc}} - \delta_{\mathrm{c}},$$

where δ_{nc} (respectively δ_{c}) is half of the sum of the non-compact (respectively compact) positive roots. For more details see [Mor, Theorem 2.2].

In our case we have the Harish-Chandra parameter $\lambda = (\lambda_1, \lambda_2) = (k-1, 0)$. This tells us that the Blattner parameter Λ is given by $\Lambda = \lambda + (1, 2) = (k, 2)$. This follows since $\lambda \in \Xi^{3,0}$ as defined in [Mor, p.205]. We note that we could have $\lambda \in \Xi^{2,1}$ but this would lead to a contradiction as we would end up with a Siegel modular form with scalar weight 0. As we mentioned earlier, the representation π_∞ is the one of minimal K_∞-type τ_λ. This representation is then associated with the representation of highest weight $(j + \kappa, \kappa)$, this is the representation $\mathrm{Sym}^j(\mathbb{C}^2) \otimes \det^\kappa$. In our case we have $(k, 2) = (j + \kappa, \kappa)$. Hence it follows that the scalar part of the weight is given by $\kappa = 2$ and then $j = k - 2$. Therefore the representation of highest weight is $\mathrm{Sym}^{k-2}(\mathbb{C}^2) \otimes \det^2$.

We note that the L-parameters at the finite places are constructed in a very similar manner. These are constructed in the same way as in Section 4 of [JLR]. Due to the special nature of our case, we end up with the split and non-split cases behaving in the same way. In fact we end up with a four dimensional representation that has the same form as the ones at the infinite places in both cases. In order to consider the L-parameters at finite places we consider our quadratic field K tensored with $\mathbb{Q}_p$ for each prime p. We end up with the following

$$K \otimes \mathbb{Q}_p = \begin{cases} \mathbb{Q}_p \oplus \mathbb{Q}_p & \text{if } p \text{ splits in } K, \\ K_{\mathfrak{p}} & \text{otherwise.} \end{cases}$$

We note that in the notation of [JLR] we have $F = \mathbb{Q}_p$ and $E = K \otimes \mathbb{Q}_p$. For ease of notation we will use F and E in the following.

The split case is parametrised by pairs (π_1, π_2) where π_1 and π_2 are irreducible, admissible representations of $\mathrm{GL}_2(F)$ having the same central character. In our case, this will be trivial. The non-split case is parametrised by triples (E, π_0, η) where E is a quadratic extension of F, π_0 is an irreducible, admissible representation of $\mathrm{GL}_2(E)$ with Galois-invariant central character ω_{π_0}, and η a character of $F^\times$ such that $\omega_{\pi_0} = \eta \circ \mathrm{N}_F^E$.

To define the parameter $\varphi(\pi_1, \pi_2) : W_F \to \mathrm{GSp}_4(\mathbb{C})$ associated to a pair (π_1, π_2), let $\varphi_1 : W_F \to \mathrm{GL}_2(\mathbb{C})$ and $\varphi_2 : W_F \to \mathrm{GL}_2(\mathbb{C})$ be the L-parameters of π_1 and π_2, respectively. We define

$$\varphi(\pi_1, \pi_2)(x) = \begin{pmatrix} a_1 & 0 & b_1 & 0 \\ 0 & a_2 & 0 & b_2 \\ c_1 & 0 & d_1 & 0 \\ 0 & c_2 & 0 & d_2 \end{pmatrix} \tag{5.2}$$

$$\text{for } \varphi_1(x) = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \quad \varphi_2(x) = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \text{ and } x \in W_F.$$

The definition of the L-parameter in the non-split case is a little more involved. For more details see Section 4 of [JLR]. We note here however that we end up with a very similar matrix. For $y \in W_E$ and g_0 a representative for the non-trivial coset of $W_E \backslash W_F$,

$$\varphi(E, \pi_0, \eta)(y) = \begin{pmatrix} a & 0 & \eta(g_0)^{-1} b & 0 \\ 0 & a' & 0 & b' \\ \eta(g_0) c & 0 & d & 0 \\ 0 & c' & 0 & d' \end{pmatrix} \tag{5.3}$$

$$\text{for } \varphi_0(y) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{and } \varphi_0(g_0 y g_0^{-1}) = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

Since the character η will be trivial in our case, we see that we end up with the same kind of interleaved matrix.

Since the level of the Siegel modular form that we end up with is determined by what happens at the finite places, we see that we still get a Siegel paramodular form of level $N^2 d_K^2 p$. Here N is the level of the classical cusp form we start with, d_K is the discriminant of K and p is the prime we raise the level of the Hilbert modular form by. This follows because the local components at finite places are constructed in exactly the same way as in [JLR]. The only change we have in the vector valued case is with the local component at the infinite place, hence the change in weight.

5.8.3 Limit of Discrete Series Representation

As was mentioned in the previous section, we aren't working exactly in the case of Example 5.8.6 from [Mor]. This is because we end up with $\lambda_2 = 0$. We therefore are not working with a typical discrete series representation, rather a limit of discrete series representation. Here we aim to give a rough description of such series. As we know, Harish-Chandra classified the discrete series representations in 1965-1966. Recall from Section 5.8.1, that in order to study the representations of $\mathrm{GSp}_4(\mathbb{R})$, we instead consider the subgroup $G = \mathrm{Sp}_4(\mathbb{R})$ which is a connected semisimple group. Then K_∞ is the maximal compact subgroup and we let T denote the maximal torus in

K_∞. As described earlier the $\mathbb{Z}$-lattice generated by the e_i is called the weight lattice. If we let L be the weight lattice, then we have a discrete series representation π_λ for every vector λ of

$$L + \rho,$$

which is not orthogonal to any root of G, where ρ is the Weyl vector (half sum of the positive roots) of G. If we consider the hyperplane perpendicular to each root in our system and consider the group generated by the set of reflections about these hyperplanes, we get the Weyl group. The complement of the set of hyperplanes is disconnected and each connected component is called a Weyl chamber. In the notation of [Mor], the Weyl chambers are the $\Xi^{p,q}$. Elements that lie inside these chambers, but not on the walls of the chambers, parametrise the discrete series representations. All discrete series representations occur in this way. Two vectors v correspond to the same discrete series representation if and only if they are conjugate under the action of the Weyl group of K.

The elements that lie on the walls of Weyl chambers are known as limit of discrete series representations. These representations still behave very much like discrete series representations, for example the character formulas still make sense but the parameter that is plugged into such formulas has been allowed to move into the wall of a Weyl chamber. Two limit of discrete series representations give the same representation if they are conjugate under the action of the Weyl group of K. Recall that in our case we have the Harish-Chandra parameter $\lambda = (k-1, 0)$ with $\lambda \in \Xi^{3,0}$. Now $\Xi^{3,0} := \{\lambda = (\lambda_1, \lambda_2) \in \widehat{T} | \lambda_1 > \lambda_2 > 0\}$, where $\widehat{T}$ is the character group of T. Since we have $\lambda_2 = 0$ we see that our Harish-Chandra parameter doesn't quite satisfy the correct inequalities. This amounts to our parameter lying in the wall of the Weyl chamber since we actually have $\lambda_1 > \lambda_2 = 0$.

Since we are working with a slightly unusual case we might wonder whether we can still carry out the process of moving from an automorphic representation to a Siegel modular form as described in Section 5.8.1.2. Fortunately, this is still the case. Recall that we are working at the infinite place here. The main result that is important to us is Theorem 3.1 from [Mor]. We now state this result.

Theorem 5.8.7. *Suppose that* $\Lambda_1 \geq \Lambda_2 \geq 3$. *If we set* $\lambda = (\Lambda_1 - 1, \Lambda_2 - 2)$, *then we have*

$$\mathrm{Hom}_{\mathfrak{g},K}(\pi_\lambda, \mathcal{A}(\Gamma\backslash G)) \simeq M_\Lambda^{\mathrm{holo}}(\Gamma\backslash G) \simeq M_\Lambda(\Gamma\backslash X).$$

This result is essentially saying that from the representation in our L-packet we can associate to it a modular form. Here $\mathcal{A}(\Gamma\backslash G)$ is the space of automorphic forms on G with respect to Γ, with Γ a congruence subgroup of $\mathrm{Sp}_4(\mathbb{R})$, and we identify $X = G/K$ with the Siegel upper half space $\mathcal{H}_2$. This result is therefore saying that the infinite dimensional representation π_λ has an associated automorphic form. We then have an associated vector valued Siegel modular form.

Now in our case we have $\Lambda_2 = 2$ so we do not satisfy the conditions of the theorem. However, by remark (ii) on page 210 of [Mor] the theorem still holds if we replace π_λ

with the generalised Verma module $M(\Lambda)$. This is essentially a method of dealing with the fact that we are in the case of limit of discrete series representations. In this case we end up with

$$\mathrm{Hom}_{\mathfrak{g},K}(M(\Lambda), \mathcal{A}^{\mathrm{cusp}}(\Gamma\backslash G)) \simeq \mathrm{Hom}_{\mathfrak{g},K}(L(\Lambda), \mathcal{A}^{\mathrm{cusp}}(\Gamma\backslash G)) \simeq S_\Lambda^{\mathrm{holo}}(\Gamma\backslash G) \simeq S_\Lambda(\Gamma\backslash X).$$

Here $S_\Lambda^{\mathrm{holo}}(\Gamma\backslash G)$ (respectively $S_\Lambda(\Gamma\backslash X)$) is the space of cusp forms in $M_\Lambda^{\mathrm{holo}}(\Gamma\backslash G)$ (respectively $M_\Lambda(\Gamma\backslash X)$), and $L(\Lambda)$ is isomorphic to the limit of discrete series representation.

5.8.4 The Satake Isomorphism and Satake Parameters

Our process so far comprises of starting with a classical modular form $f \in S_k(\Gamma_0(N))$ and base changing to a Hilbert modular form f' say. We then raise the level of this modular form to get another Hilbert modular form g say. We then induce the associated Galois representation in order to obtain a four-dimensional representation whose associated modular form is a Siegel paramodular newform. We now wish to know, given the L-parameters and L-packets that we have obtained, what the Hecke eigenvalues of this Siegel paramodular newform are. This will be determined via Satake parameters and the Satake isomorphism. There are many different sources that cover the Satake isomorphism. We will stick mostly with the treatment given in [AS], however for a more thorough treatment, the reader might like to consult [Gross].

We first set up some notation. The following is from [AS, §2 & §3]. Let $G = \mathrm{GSp}_{2n}$. Here we work with general n but note that we will be interested in the case $n = 2$. An element t of the standard maximal torus T is often written in the form

$$t = \mathrm{diag}(u_1, \ldots, u_n, u_1^{-1}u_0, \ldots, u_n^{-1}u_0),\ u_i \in \mathrm{GL}(1); \tag{5.4}$$

then $u_0 = \mu(t)$, the multiplier homomorphism. We fix the following characters of the maximal torus $T \subset G$. If $t \in T$ is written in the form (5.4), then let

$$e_i(t) = u_i,\ i = 0, 1, \ldots, n.$$

These characters are a basis for the character lattice of G,

$$X = \mathbb{Z}e_0 \oplus \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_n.$$

We also fix the following cocharacters of T:

$$\begin{aligned}
f_0(u) &= \mathrm{diag}(\underbrace{1, \ldots, 1}_{n}, \underbrace{u, \ldots, u}_{n}),\\
f_1(u) &= \mathrm{diag}(\underbrace{u, 1, \ldots, 1}_{n}, \underbrace{u^{-1}, 1, \ldots, 1}_{n}),\\
&\vdots\\
f_n(u) &= \mathrm{diag}(\underbrace{1, \ldots, 1, u}_{n}, \underbrace{1, \ldots, 1, u^{-1}}_{n}).
\end{aligned}$$

These elements are a $\mathbb{Z}$-basis for the cocharacter lattice of G,

$$X^\vee = \mathbb{Z}f_0 \oplus \mathbb{Z}f_1 \oplus \cdots \oplus \mathbb{Z}f_n.$$

With the natural pairing $\langle , \rangle : X \times X^\vee \to \mathbb{Z}$, we have

$$\langle e_i, f_j \rangle = \delta_{ij}.$$

We now choose the following set of simple roots:

$$\alpha_1(t) = u_{n-1}^{-1}u_n, \ldots, \alpha_{n-1}(t) = u_1^{-1}u_2, \alpha_n(t) = u_1^2 u_0^{-1};$$

here t is written in the form (5.4). In other words,

$$\alpha_1 = e_n - e_{n-1}, \ldots, \alpha_{n-1} = e_2 - e_1, \alpha_n = 2e_1 - e_0.$$

The corresponding coroots are

$$\alpha_1^\vee = f_n - f_{n-1}, \ldots, \alpha_{n-1}^\vee = f_2 - f_1, \hat{\alpha}_n = f_1.$$

If we let $R = \{\alpha_1, \ldots, \alpha_n\} \subset X, R^\vee = \{\alpha_1^\vee, \ldots, \alpha_n^\vee\} \subset \hat{X}$, then

$$(X, R, X^\vee, R^\vee)$$

is the root datum of G. The Cartan matrix is

$$\langle \alpha_i, \alpha_j^\vee \rangle = \begin{pmatrix} 2 & -1 & & & & & \\ -1 & 2 & -1 & & & & \\ & -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 & \\ & & & & -1 & 2 & -1 \\ & & & & & -2 & 2 \end{pmatrix}.$$

We have the Borel subgroup $B = TN$, where N consists of matrices of the form $\begin{pmatrix} A & 0 \\ 0 & {}^tA^{-1} \end{pmatrix}\begin{pmatrix} \mathbf{1} & B \\ 0 & \mathbf{1} \end{pmatrix}$, with $A \in \mathrm{GL}(n)$ lower triangular unipotent and B symmetric. The torus T acts on the Lie-algebra $\mathfrak{n}$ of N by the adjoint representation Ad. The modular factor $\delta_B(t) = \det(\mathrm{Ad}_{\mathfrak{n}}(t))$ is given by

$$\delta_B(t) = u_0^{-n(n+1)/2} u_1^2 u_2^4 \ldots u_n^{2n}.$$

We note that for our purposes, we are dealing with $G = \mathrm{GSp}_4(\mathbb{R})$ so $n = 2$. We are now in a position to consider the Satake isomorphism.

In the following F will denote a $\mathfrak{p}$-adic field, $\mathcal{O}$ will be its ring of integers, $\omega \in \mathcal{O}$ a generator of the maximal ideal and $q = |\mathcal{O}/\omega\mathcal{O}|$ the size of the residue field. Let $G = \mathrm{GSp}_4(F)$, let $K = \mathrm{GSp}_4(\mathcal{O})$, and let $T = T(F)$ be the maximal torus of G.

Let $\mathcal{H}(G,K)$ be the unramified Hecke algebra of G, consisting of locally constant, compactly supported functions $f : G \to \mathbb{C}$ which are left and right K-invariant. The product in $\mathcal{H}(G,K)$ is given by convolution

$$(f * g)(x) = \int_G f(y)g(y^{-1}x)\mathrm{d}y.$$

Here we have used the definition from [Gross, §2] since it is slightly easier to work with. We note that here $\mathrm{d}y$ is a particular (in fact unique) Haar measure on G giving K volume 1. We note that functions $f \in \mathcal{H}(G,K)$ are constant on double cosets KxK since f is left and right K-invariant. It also follows, since f has compact support, that f must be a linear combination of the characteristic functions $\mathrm{char}(KxK)$ of double cosets. It follows that these characteristic functions form a basis for $\mathcal{H}(G,K)$.

We also consider the Hecke algebra $\mathcal{H}(T, T(\mathcal{O}))$. Note that this Hecke algebra consists of locally constant, compactly supported functions $f : T \to \mathbb{C}$ which are left and right $T(\mathcal{O})$-invariant. Recall that $T \subset B = TN \subset G$ and $T(\mathcal{O}) = T \cap K$. Again, it follows that this Hecke algebra must be generated by characteristic functions. As in [Gross, Proposition 2.6], we see that T has a Cartan decomposition

$$T = \coprod_i f_i(\omega)T(\mathcal{O}),$$

where ω is a uniformising element (prime) of F. We therefore have the following special elements in this Hecke algebra:

$$\begin{aligned}
X_0 &:= \mathrm{char}\left(\mathrm{diag}(\mathcal{O}^*, \ldots, \mathcal{O}^*, \omega\mathcal{O}^*, \ldots, \omega\mathcal{O}^*)\right), \\
X_1 &:= \mathrm{char}\left(\mathrm{diag}(\omega\mathcal{O}^*, \mathcal{O}^*, \ldots, \mathcal{O}^*, \omega^{-1}\mathcal{O}^*, \mathcal{O}^*, \ldots, \mathcal{O}^*)\right), \\
&\vdots \\
X_n &:= \mathrm{char}\left(\mathrm{diag}(\mathcal{O}^*, \ldots, \mathcal{O}^*, \omega\mathcal{O}^*, \mathcal{O}^*, \ldots, \mathcal{O}^*, \omega^{-1}\mathcal{O}^*)\right),
\end{aligned}$$

where "char" stands for characteristic function. By considering the product in $\mathcal{H}(T, T(\mathcal{O}))$, we see that

$$\begin{aligned}
X_i * X_j(x) &= \int_T X_i(y)X_j(y^{-1}x)\mathrm{d}y \\
&= \int_{f_i(\omega)T(\mathcal{O})} X_j(y^{-1}x)\mathrm{d}y \\
&= \int_{x^{-1}f_i(\omega)T(\mathcal{O})} X_j(y^{-1})\mathrm{d}y \\
&= \int_{xf_i^{-1}(\omega)T(\mathcal{O})} X_j(y)\mathrm{d}y \\
&= \int_{xf_i^{-1}(\omega)T(\mathcal{O})\cap f_j(\omega)T(\mathcal{O})} \mathbf{1}\mathrm{d}y \\
&= X_{i,j}(x),
\end{aligned}$$

where $X_{i,j}$ has ω in the i-th and j-th positions (and ω^{-1} in $(n+i)$-th and $(n+j)$-th positions). It follows that we have

$$X_0^k = \mathrm{char}\left(\mathrm{diag}(\mathcal{O}^*, \ldots, \mathcal{O}^*, \omega^k\mathcal{O}^*, \ldots, \omega^k\mathcal{O}^*)\right), \quad k \in \mathbb{Z},$$

and similarly for the other X_i. This follows since $X_{i,i} = X_i^2$. It is fairly clear that the $\mathbb{C}$-algebra generated by these X_i together with their inverses is therefore the Hecke algebra $\mathcal{H}(T, T(\mathcal{O}))$. That is

$$\mathcal{H}(T, T(\mathcal{O})) = \mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}, \ldots, X_n^{\pm 1}].$$

For an element $f \in \mathcal{H}(G, K)$, the Satake transform is defined by

$$(Sf)(t) = |\delta_B(t)|^{1/2} \int_N f(tn)\mathrm{d}n = |\delta_B(t)|^{-1/2} \int_N f(nt)\mathrm{d}n.$$

We note that the element Sf is an element of $\mathcal{H}(T, T(\mathcal{O}))$, and S actually defines an isomorphism

$$S : \mathcal{H}(G, K) \xrightarrow{\sim} \mathcal{H}(T, T(\mathcal{O}))^W,$$

where W denotes the Weyl group of G with respect to T. This group is defined as the quotient of the normaliser of the torus $N(T)$ by the centraliser of the torus $Z(T)$. Note that

$$N(T) := \{x \in G |\, xtx^{-1} \in T \quad \text{for all } t \in T\},$$

and

$$Z(T) := \{x \in G |\, xtx^{-1} = t \quad \text{for all } t \in T\}.$$

We now consider spherical representations. An irreducible admissible representation of G is called spherical if it contains a non-zero vector fixed by K. Let $\chi_0, \ldots, \chi_n$ be unramified characters of F^*. Note that an unramified character of F^* is one which is trivial on $\mathcal{O}^*$. These characters then define an unramified character on the Borel subgroup $B = TN$ which is trivial on N and which, on T, is given by

$$t \mapsto \chi_0(u_0)\chi_1(u_1)\ldots\chi_n(u_n),$$

with $t = \mathrm{diag}(u_1, \ldots, u_n, u_1^{-1}u_0, \ldots, u_n^{-1}u_0)$ and $u_i \in \mathrm{GL}_1(\mathbb{R})$ the parameters in the standard maximal torus T. If we use normalised induction (see [1] for a definition) to G, we obtain a representation with a unique spherical constituent. This representation is denoted by

$$\pi(\chi_0, \chi_1, \ldots, \chi_n).$$

The isomorphism class of this representation only depends on the unramified characters modulo the action of the Weyl group. It is known that each spherical representation is obtained in this way. We therefore have a bijection between unramified characters of T modulo the action of the Weyl group, and isomorphism classes of spherical representations of G.

Each of the unramified characters χ_i of F^* are determined by their value on a prime element $\omega \in F$. This value is known as a Satake parameter and may be any non-zero complex number. We denote these Satake parameters by $b_i := \chi_i(\omega)$. We therefore see that the character of the Borel subgroup is determined by the vector $(b_0, b_1, \ldots, b_n) \in (\mathbb{C}^*)^{n+1}$. The Weyl group acts on this complex torus, and we see that the unramified representations of G are parametrised by the orbit space $(\mathbb{C}^*)^{n+1}/W$. It follows that the Satake isomorphism

$$S : \mathcal{H}(G, K) \xrightarrow{\sim} \mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^W$$

allows us to identify the Hecke algebra with the coordinate ring of $(\mathbb{C}^*)^{n+1}$. Recall that this follows since $\mathcal{H}(T, T(\mathcal{O}))^W = \mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^W$. Each point $(b_0, \ldots, b_n) \in (\mathbb{C}^*)^{n+1}$ determines a character, i.e., an algebra homomorphism $\mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^W \to \mathbb{C}$, by mapping X_i to b_i. Via the Satake isomorphism, this also defines a character of $\mathcal{H}(G, K)$. This character is simply the action of $\mathcal{H}(G, K)$ on the one-dimensional space of spherical vectors in $\pi(\chi_0, \ldots, \chi_n)$. We therefore end up with the following commutative diagram, in which all the maps are bijections:

$$\begin{array}{ccc} \{\text{spherical representations}\} & \longrightarrow & \mathrm{Hom}_{\mathrm{Alg}}(\mathcal{H}(G, K), \mathbb{C}) \\ \uparrow & & \downarrow \\ \{\text{unramified characters}\}/W & \longleftarrow & (\mathbb{C}^*)^{n+1} \end{array}$$

The left arrow here is by normalised induction and taking the unique spherical constituent. The top arrow is the action of $\mathcal{H}(G, K)$ on the space of spherical vectors. The right arrow comes form the identification $\mathcal{H}(G, K) \simeq \mathbb{C}[X_0^{\pm 1}, X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^W$. Finally the bottom arrow assigns to the Satake parameters $(b_0, \ldots, b_n)$ the unramified character with $\chi_i(\omega) = b_i$.

Remark 5.8.8. Although it is not immediately obvious that this diagram should commute, it turns out that it is a consequence of the way the process works as we move from step to step.

5.8.5 The General Theorem

Since the Satake parameters are closely linked with the Hecke algebra, it is clear that they contain information on the behaviour of the Hecke operators acting on Siegel modular forms. In particular if we take a given Hecke operator, we see from the above diagram, that there is an associated vector in $(\mathbb{C}^*)^{n+1}$ containing the Satake parameters, from which we will be able to read off the Hecke eigenvalues of the Siegel modular form that we construct.

Recall that for each finite prime q, the generic representations π_q were spherical representations. By the above, and the discussion in the introduction of [Rob], it follows that we can determine the Satake parameters from the L-parameters $\varphi(\pi_{0,q})$ by evaluation at a Frobenius element at q. This process will give us a matrix with the Satake parameters along the diagonal. We will at this point need a way of determining the Hecke eigenvalues from this matrix. For this, we turn to [Gross, §6]. Here we see that each element of the Hecke algebra has an associated semisimple conjugacy class, namely the matrix containing the Satake parameters. It turns out that plugging this semisimple conjugacy class into a particular representation of $\hat{G}$ determined by the choice of Hecke operator, and taking the trace will, up to some normalisation factors, give us the Hecke eigenvalues we want.

In our case the semisimple conjugacy classes will be exactly the interleaved matrices ((5.2) and (5.3)) defining the L-parameters $\varphi(\pi_{0,q})$ evaluated at a Frobenius element at q. Since we are working with trivial central character, we can view our representation of $\mathrm{GSp}_4(F)$ as a representation of $\overline{G} = \mathrm{PGSp}_4(F)$. We may actually use the example with $G = SO(5)$ from [Gross, p. 233] as $\mathrm{PGSp}(4) \cong \mathrm{SO}(5)$. The dual group in this case is $\mathrm{Sp}_4(\mathbb{C})$. Here we see that the standard representation is the one corresponding to the T_q Hecke operator. In the example of Gross, we see that by plugging the semisimple conjugacy class, which we denote by s, into the standard representation $V = \mathbb{C}^4$ and taking the trace, we get the following:

$$q^{3/2}\mathrm{Tr}(s|V) = q^{\frac{3-j-2k}{2}} a_q(F),$$

where $a_q(F)$ is the Hecke eigenvalue of the T_q operator applied to F. We therefore multiply by $q^{\frac{j+2k-3}{2}}$ in order to get the Hecke eigenvalue. This choice of normalisation is in order to make the eigenvalues algebraic integers. The exponent $\frac{j+2k-3}{2}$ of q is half the weight of a particular cohomology. We will therefore simply be able to read off the Hecke eigenvalues from this trace.

In order to actually calculate these eigenvalues we use our knowledge of the GL_2 case. Recall that for each finite place q, we defined $\pi_{0,q} = \otimes_{w|v}\pi_{0,w}$. Since the L-parameters for $\mathrm{GSp}_4(F)$ come from the ones for $\mathrm{GL}_2(F)$ associated to the Hilbert modular form, it follows that we may use the known theory of the GL_2 case to determine what is happening in this case. For details of this, the reader might like to consult an as yet unfinished book by Harder [Har2], or the case of $G = \mathrm{GL}_n$ from [Gross, §6]. In the $\mathrm{GL}_2(F)$ cases, evaluation of the L-parameters $\varphi(\pi_{0,w_1})$ and $\varphi(\pi_{0,w_2})$ at a Frobenius element gives matrices whose trace gives the Hecke eigenvalues of the Hilbert form at each prime above q. This then tells us that when we plug the semisimple conjugacy class associated to $\varphi(\pi_{0,q})$ into the standard representation we simply get the sum of these Hecke eigenvalues. It follows that we have

$$a_q(F) = \begin{cases} a_{\mathfrak{q}}(g) + a_{\overline{\mathfrak{q}}}(g) & \text{if } q \text{ is split,} \\ 0 & \text{if } q \text{ is not split,} \end{cases}$$

where g is the level raised Hilbert modular form. Hence, just as in the scalar valued case, we have a congruence between the Hecke eigenvalues of f and those of F.

We have therefore proven the following theorem.

Theorem 5.8.9. *Let $f \in S_k(\Gamma_0(N))$ with $k \geq 2$ and let K be a real quadratic field with discriminant d_K. Suppose $(d_K, N) = 1$. Choose a prime p that splits in K. Suppose $p \nmid \ell N$, $\ell \nmid (p+1)$ and*

$$a_p(f)^2 \equiv \left(p^{\frac{k}{2}} + p^{\frac{k}{2}-1}\right)^2 \pmod{\lambda},$$

for some $\lambda | \ell$. Further assume that $\rho_f|_{G_K}$ is irreducible modulo λ. Let $\rho = \mathrm{Sym}^{k-2}(\mathbb{C}^2) \otimes \det^2$.

Then there exists a Siegel paramodular cusp form $F \in S_\rho(K(N^2 d_K^2 p))$ satisfying

$$a_q(F) \equiv a_q(f)\,(1 + \chi_K(q)) \pmod{\lambda}$$

for $q \nmid N^2 d_K^2 p$, where χ_K is the quadratic character associated to K.

Notice how we now have the more general level raising condition $a_p(f)^2 \equiv \left(p^{\frac{k}{2}} + p^{\frac{k}{2}-1}\right)^2 \pmod{\lambda}$. We might wonder if this can also be viewed as an Euler factor arising from the Symmetric square L-function like in the scalar valued case. It turns out that this is indeed the case. Recall from Section 5.7.2 that we ended up with the factor

$$-\left[a_p^2 - (p^{2k-2-s} + 2p^{k-1} + p^s)\right].$$

We also have

$$\left(p^{\frac{k}{2}} + p^{\frac{k}{2}-1}\right)^2 = p^k + 2p^{k-1} + p^{k-2}.$$

It is clear that if we take $s = k$ in the Euler factor we get

$$p^{2k-2-s} + 2p^{k-1} + p^s = p^{k-2} + 2p^{k-1} + p^k.$$

We therefore see that the Euler factor matches the level raising condition in the vector valued case as well. This is as would be expected as the argument of Section 5.7 still holds for weight $k > 2$. Note also that this agrees with our choice of $s = 2$ in the scalar valued case.

References

[1-2-3] J. H. Bruinier, G. van der Geer, G. Harder, D. Zagier, *The 1-2-3 of Modular Forms*, Springer-Verlag, Berlin, 2008.

[A] A. Andrianov *Introduction to Siegel Modular Forms and Dirichlet Series*, Springer, Universitext, 2009.

[AS] M. Asgari, R. Schmidt, *Siegel modular forms and representations*, Manuscripta Math. **104** (2001) 173–200.

[BD] N. Billerey, L. V. Dieulefait, *Explicit large image theorems for modular forms*, J. Lond. Math. Soc. (2), 89(2):499523, 2014.

[BM] N. Billerey, R. Menares, *On the Modularity of Reducible mod l Galois Representations*

Math. Res. Lett. **23** (2016), no. 1, 15–41.

[BM2] N. Billerey, R. Menares, *Strong Modularity of Reducible Galois Representations*

Trans. Amer. Math. Soc. **370** (2018), no. 2, 967–986.

[Bla] D. Blasius, *Hilbert modular forms and the Ramanujan conjecture*, in: Noncommutative Geometry and Number Theory, in: Aspects Math., vol. E37, Vieweg, Wiesbaden, (2006) 35–56.

[BlKa] S. Bloch, K. Kato, *L-Functions and Tamagawa Numbers of Motives* The Grothendieck Festschrift Volume I, 333–400, Progress in Mathematics, **86**, Birkhäuser, Boston, 1990.

[Br] J. Brown, *Saito-Kurokawa Lifts and Applications to the Bloch-Kato Conjecture* Compositio Mathematica. London Mathematical Society, **143**(2), 290–322.

[BrKr] A. Brumer, K. Kramer, *Paramodular Abelian Varieties of Odd Conductor* Trans. Amer. Math. Soc. **366** (2014), no. 5, 2463–2516.

[Bu] K. Buzzard, *Computing weight one modular forms over* $\mathbb{C}$ *and* $\bar{\mathbb{F}}_p$ Computations with Modular Forms, vol **6** (2014), 129–146.

[Bu2] K. Buzzard, *Notes on Siegel Modular Forms*, April 2012

`http://wwwf.imperial.ac.uk/~buzzard/maths/research/notes/siegel_modular_forms_notes.pdf`.

[Cara] H. Carayol, *Sur les représentations l-adiques associées aux formes modulaires de Hilbert*, Annales scientifiques de l'É.N.S. 4^e série, tome **19**, n° 3 (1986), 409–468.

[Carl] L. Carlitz, *Arithmetic properties of generalized Bernoulli numbers*, J. Reine Angew. Math. **202** (1959) 174–182

[Cas] W. Casselman, *On Some Results of Atkin and Lehner*, Math. Ann. **201**, (1973), 301–314.

[Ch] N. Childress, *Class Field Theory*, Springer Universitext, 2009.

[Cog] J. W. Cogdell, *On Artin L-Functions*

`https://people.math.osu.edu/cogdell.1/artin-www.pdf`.

[Cox] D. A. Cox, *Primes of the Form $x^2 + ny^2$*, Wiley, 1989.

[Cre] J. E. Cremona, *Abelian Varieties with Extra Twist, Cusp Forms, and Elliptic Curves Over Imaginary Quadratic Fields*, J. London Math. Soc. (2) **45**, (1992), no. 3, 404–416.

[DeSe] P. Deligne, J. -P. Serre, *Formes modulaires de poids 1*, Ann. Sci. Ec. Norm. Sup. **7** (1974), 507–530.

[DemVoi] L. Dembélé, J. Voight, *Explicit Methods for Hilbert Modular Forms* Hilbert modular forms and Galois deformations, 135–198, Adv. Courses Math. CRM Barcelona, Birkhäuser/Springer, Basel, 2013.

[DFG] F. Diamond, M. Flach, L. Guo, *The Tamagawa Number Conjecture of Adjoint Motives of Modular Forms*, Ann. Sci. École Norm. Sup. (4) **37** (2004), 663–727.

[DiSh] F. Diamond, J. Shurman, *A First Course in Modular Forms*, Springer, 2005.

[Do] I. V. Dolgachev, *McKay correspondence*

`http://www.math.lsa.umich.edu/~idolga/McKaybook.pdf`.

[D] N. Dummigan, *Eisenstein Primes, Critical Values and Global Torsion* Pacific J. Math **233** (2007), 291–308.

[D2] N. Dummigan, *Symmetric Square L-Functions and Shaferavich-Tate Groups, II*, Int. J. Number Theory **5** (2009), 1321–1345.

[DF] N. Dummigan, D. Fretwell, *Ramanujan-Style Congruences of Local Origin* J. Number Theory **143** (2014), 248–261.

[E] B. Edixhoven, *Serre's Conjecture*, in Modular Forms and Fermats Last Theorem, (G. Cornell, J. H. Silverman, G. Stevens, eds.), Springer-Verlag, New York, 1997, 209–242.

[Frei] E. Freitag, *Hilbert Modular Forms*, Springer-Verlag, Berlin, 1990.

[Fret] D. Fretwell, *An Overview of Global Class Field Theory With Applications (in Terms of Ideals)*

`https://www.academia.edu/4373839/An_overview_of_global_class_field_theory_with_applications_in_terms_of_ideals_`.

[Fret2] D. Fretwell, *Level p Paramodular Congruences of Harder Type*, 2015 University of Sheffield PhD Thesis

`http://etheses.whiterose.ac.uk/9683/1/Thesis.pdf`.

[GT] W. T. Gan, S. Takeda, *The local Langlands conjecture for* GSp(4), Ann. of Math. **173** (2011) 1841–1882.

[Geer] G. van der Geer, *Hilbert Modular Surfaces*, Springer-Verlag, Berlin, 1988.

[Goren] E. Z. Goren, *Lectures on Hilbert Modular Varieties and Modular Forms*, CRM Monograph Series, vol. 14, Amer. Math. Soc. Providence, RI, 2002.

[Gross] B. H. Gross, *On the Satake Isomorphism*, in Galois Representations in Arithmetic Algebraic Geometry, (A. J. Scholl and R. L. Taylor, eds.), London Mathematical Society, Lecture Note Series 254, Cambridge University Press 1998.

[Har] G. Harder, *Secondary Operations in the Cohomology of Harish-Chandra Modules*

`http://www.math.uni-bonn.de/people/harder/Manuscripts/Eisenstein/SecOPs.pdf`.

[Har2] G. Harder, *Cohomology of Arithmetic Groups, Chapter III*

`https://www.math.uni-bielefeld.de/documenta/vol-06/12.pdf`.

[HoHu] P. N. Hoffman, J. F. Humphreys, *Projective Representations of the Symmetric Groups: Q-Functions and Shifted Tableaux*, Oxford Science Publications, 1992.

[HK] A. Huber, G. Kings, *Bloch-Kato Conjecture and Main Conjecture of Iwasawa Theory for Dirichlet Characters*

Duke Math. J., Volume **119**, Number 3 (2003), 393–464.

[J] F. Jarvis, *Algebraic Number Theory*, Springer, 2014.

[JLR] J. Johnson-Leung, B. Roberts, *Siegel Modular Forms of Degree* 2 *Attached to Hilbert Modular Forms* J. Number Theory **132** (2012), no. 4, 543–564.

[Ka] N. M. Katz, *p-adic properties of modular schemes and modular forms*, Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), Lecture Notes in Mathematics, **350**, Berlin, New-York: Springer-Verlag, 69–190.

[KhWi] C. Khare, J. P. Wintenberger, *Serre's Modularity Conjecture (I)*

`http://www.math.ucla.edu/~shekhar/papers/results.pdf`.

[Kne] M. Kneser, *Strong approximation*, Proc. Symp. Pure Math. **9**, (1966), 187–196.

[Ko] W. Kohnen, *A Short Course on Siegel Modular Forms*, POSTECH lecture series, July 2007

`https://www.mathi.uni-heidelberg.de/~winfried/siegel2.pdf`.

[L] R. P. Langlands, *Base Change for* GL(2)

`http://sunsite.ubc.ca/DigitalMathArchive/Langlands/pdf/book-ps.pdf`.

[LMFDB] The L-Functions and Modular Forms Database (LMFDB), Hilbert Modular Forms

`http://www.lmfdb.org/ModularForm/GL2/TotallyReal/`.

[M] B. Mazur, *Modular Curves and the Eisenstein Ideal*

Publ. Math. IHES **47** (1977), 33–186.

[Mil] J. S. Milne, *Class Field Theory*

`http://www.jmilne.org/math/CourseNotes/CFT.pdf`.

[MFFLT] Various Authors, *Modular Forms and Fermat's Last Theorem*, Springer-Verlag New York, Inc. 1997.

[Mor] T. Moriyama, *Representations of* $\mathrm{GSp}(4,\mathbb{R})$ *with Emphasis on Discrete Series*, in Automorphic forms on GSp(4), (M. Furusawa, ed.), Proceedings of the 9th Autumn Workshop on Number Theory, Hakuba, Japan, November 6-10, 2006.

[N] J. Neukirch, *Algebraic Number Theory*, Springer-Verlag, 1999.

[O] T. Ozawa, *Constant Terms of Eisenstein Series Over A Totally Real Field* Int. J. Number Theory **13** (2017) 309–324.

[Rib] K. A. Ribet, *Raising the Levels of Modular Representations*

Séminaire de Théorie des Nombres, Paris 1987–88, 259–271, Progr. Math., **81**, Birkhäuser Boston, Boston, MA, 1990.

[Rib2] K. A. Ribet, *Report on mod ℓ Representations of* $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ Motives (Seattle, WA, 1991), 639–676, Proc. Sympos. Pure Math., **55**, Part 2, Amer. Math. Soc., Providence, RI, 1994.

[Rib3] K. A. Ribet, *A modular construction of unramified p-extensions of $\mathbb{Q}(\mu_p)$*, Invent. Math. **34** (1976), 151–162.

[Rob] B. Roberts, *Global L-packets for* GSp(2) *and theta lifts*, Doc. Math. **6** (2001) 24–314.

[RS1] B. Roberts, R. Schmidt, *On modular forms for the paramodular group*, Automorphic Forms and Zeta Functions, Proceedings of the Conference in Memory of Tsuneo Arakawa, World Scientific, 2006.

[RS2] B. Roberts, R. Schmidt, *Local Newforms for* GSp(4), Lecture Notes in Math., vol. 1918, Springer, Berlin, Heidelberg, New York, 2007.

[SS] A. Saha, R. Schmidt, *Yoshida lifts and simultaneous non-vanishing of dihedral twists of modular L-functions*

J. Lond. Math. Soc. (2013), 88 (1): 251–270.

[Sch] G. J. Schaeffer, *The Hecke stability method and ethereal forms*, 2012 UC Berkeley PhD thesis.

[Schur] I. Schur, *Ü ber die Darstellung der symmetrischen und der alternierenden Gruppen durch gebrochene lineare Substitutionen*, J. Reine Angew. Math. **139**, 155–250.

[S] J. -P. Serre, *Modular forms of weight one and Galois representations*, Algebraic Number Fields, édité par A. Fröhlich, Acad. Press (1977), 193–268.

[Sha] R. Sharifi, *Group and Galois Cohomology*

`http://math.ucla.edu/~sharifi/groupcoh.pdf`.

[Sil] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Second Edition, Springer, 1986.

[Sil2] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag, 1994.

[StTa] I. Stewart, D. Tall, *Algebraic Number Theory and Fermat's Last Theorem - 3rd edition*, A.K.Peters, 2002.

[Ta] R. Taylor, *On Galois Representations Attached to Hilbert Modular Forms*, Springer-Verlag, Invent. math. **98**, (1989) 265–280.

[Tr] D. Trotabas, *Modular Forms of Weight One: Galois Representations and Dimension*

`https://arxiv.org/pdf/0906.4579v1.pdf`.

[Tu] J. Tunnell, *Artin's Conjecture for Representations of Octahedral Type*

Bull. Amer. Math. Soc. (N.S.), Volume 5, Number 2 (1981), 173–175.

[U] E. Urban, *Groupes de Selmer et Fonctions L p-adiques pour les Représentations Modulaires Adjointes*, 2006, `http://www.math.columbia.edu/~urban/eurp/ADJMC.pdf`

[W] R. Weissauer, *Four Dimensional Galois Representations*, Astérisque **302** (2005), 67–150.

[W2] R. Weissauer, *Existence of Whittaker models related to four dimensional symplectic Galois representations*

`https://arxiv.org/pdf/math/0703218.pdf`.

[1] Wikipedia, *Induced representation*

`https://en.wikipedia.org/wiki/Induced_representation`.

[2] Groupprops subwiki, *Linear representation theory of special linear group: SL(2,3)*

`http://groupprops.subwiki.org/wiki/Linear_representation_theory_of_special_linear_group:SL(2,3)`.

[3] Groupprops subwiki, *Linear representation theory of general linear group: GL(2,3)*

`http://groupprops.subwiki.org/wiki/Linear_representation_theory_of_general_linear_group:GL(2,3)`.

[4] Groupprops subwiki, *Linear representation theory of special linear group: SL(2,5)*

`http://groupprops.subwiki.org/wiki/Linear_representation_theory_of_special_linear_group:SL(2,5)`.