Ila Varma

Sums of Squares, Modular Forms, and Hecke Characters

Master thesis, defended on June 18, 2010

Thesis advisor: Bas Edixhoven

Mastertrack: Algebra, Geometry, and Number Theory

Mathematisch Instituut, Universiteit Leiden

Abstract. This thesis discusses the classical problem of how to calculate $r_n(m)$, the number of ways to represent an integer m by a sum of n squares. To this day, there are very few formulas that allow for easy calculation of $r_n(m)$. Here, we focus on the case when n is even, hence we can use the theory of integral weight modular forms on $\Gamma_1(4)$ to write down formulas for the theta function $\theta_n(q)$ associated to sums of n squares. In particular, we show that for only a small finite list of n can θ_n be written as a linear combination consisting entirely of Eisenstein series and cusp forms with complex multiplication. These give rise to "elementary" formulas for $r_n(m)$, in which knowing the prime factorization of m allows for their efficient computation. This work is related to Couveignes and Edixhoven's forthcoming book and Peter Bruin's forthcoming Ph.D. thesis concerning polynomial-time algorithms for calculating the prime Fourier coefficients of modular forms.

Contents

I am extremely grateful to Prof. Bas Edixhoven for advising this master thesis as well as patiently explaining his insight on various areas of mathematics. I would not have been able to come and study at Leiden without his help, and his guidance during the past year has been truly valuable. I would also like to thank Dr. Ronald van Luijk, Dr. Lenny Taelman, and Peter Bruin for enlightening mathematical discussions, as well as Profs. Hendrik Lenstra and Richard Gill for sitting on my exam committee. The entire mathematics department at Leiden University has been very supportive and welcoming, and I am grateful for the classes and seminars I have had the opportunity to participate in here. I also deeply appreciate the mathematical and personal support from Alberto Vezzani and other friends I have met during my study in Holland. Finally, I would like to thank my parents for their continual support in my mathematical endeavors.

This master thesis and my study in Leiden was supported by the U.S. Fulbright program and the HSP Huygens program.

1 Introduction

The simple Diophantine equation

$$x_1^2 + x_2^2 + ... + x_n^2 = m$$

has been of interest to many mathematicians throughout time. From understanding the lengths of a right triangle to distances in n-dimensional space, the physical and geometric aspects of this expression are clear. However, studying *sums of squares problems* is deeply linked to almost all of number theory. Fermat started by investigating which primes can be "represented" by a sum of 2 squares, i.e. whether or not there exists an integral pair (x_1, x_2) such that $x_1^2 + x_2^2 = p$ for each prime p (see [16]). To this day, students in an elementary number theory class are quickly introduced to his theorem stated in 1640 and later proved by Euler:

an odd prime p can be written as a sum of 2 squares if and only if $p \equiv 1 \bmod 4$.

After studying the multiplicative properties of solutions to this equation, it is not hard to conclude that the integers represented by sums of squares are those with prime factorization such that primes $p \equiv 3 \pmod 4$ occur in even powers.

Fermat also studied the sums of 3 squares problems, but the following statement describing which integers can be represented was not proven until Legendre in 1798 ([16]):

an integer $m > 0$ can be written as a sum of 3 squares if and only if $m \not\equiv 7 \bmod 8$ and $4 \nmid m$.

In 1770, Lagrange proved that every natural number can be written as a sum of 4 squares (see [11]). Other mathematicians gave different proofs involving surprising tools such as quaternions and elliptic functions (see [16]). For example, Ramanujan gave a proof in 1916 involving calculation of the coefficient $r_4(m)$ of x^m in

$$(1 + 2x + 2x^4 + ...)^4 = \left(\sum_{k=-\infty}^{\infty} x^{k^2} \right)^4 \quad \in \mathbb{Z}[[x]]$$

as the number of solutions of $m = x_1^2 + x_2^2 + x_3^2 + x_4^2$ in the integers (see [26]). In this language, Lagrange's theorem amounted to proving that $r_4(m) > 0$ for all integers $m > 0$. These coefficients were further studied by Jacobi in 1829, additionally gave exact formulas for representing a natural number by a sum of 4 squares (see [19]):

$$r_4(m) = \begin{cases} 8 \sum_{d|m} d & \text{if } m \text{ is odd} \\ 24 \sum_{2 \nmid d|m} d & \text{if } m \text{ is even.} \end{cases}$$

Jacobi continued the study of sums of n squares by writing down exact formulas for the cases of $n = 6$ and $n = 8$. Writing down formulas for $r_5(m)$ and $r_7(m)$ in fact came much later due to their surprising difficulty, and it was worked on by Eisenstein, Smith, Minkowski, Mordell, Ramanujan, and Hardy ([16]). Even for $r_3(m)$, Gauss gave the simplest formula in 1801, which still involved the class number of binary quadratic forms with discriminant $-m$. We therefore focus our attention to the case when n is even.

Denote $r_n(m)$ as the coefficient of x^m in $(1 + 2x + 2x^4 + ...)^n$, i.e.

$$\sum_{m=0}^{\infty} r_n(m) x^m = \left(\sum_{k=-\infty}^{\infty} x^{k^2} \right)^n \quad \text{or equivalently,} \quad r_n(m) = \#\{\mathbf{x} \in \mathbb{Z}^n : x_1^2 + x_2^2 + ... + x_n^2 = m\}.$$

If we define $\chi : \mathbb{Z} \to \mathbb{C}$ such that

$$\chi(d) = \begin{cases} 1 & \text{if } d \equiv 1 \bmod 4 \\ -1 & \text{if } d \equiv -1 \bmod 4 \\ 0 & \text{if } d \equiv 0 \bmod 2, \end{cases}$$

then Jacobi's formulas can be written as

$$r_6(m) = 16 \cdot \sum_{d|m} \chi\left(\frac{m}{d}\right) d^2 - 4 \cdot \sum_{d|m} \chi(d) d^2 \qquad \text{and} \qquad r_8(m) = 16 \cdot (-1)^m \sum_{d|m} (-1)^d d^3.$$

When mathematicians started writing down formulas for $n > 8$, they got noticeably more complicated. Liouville in 1864 wrote the first formula for $r_{10}(m)$ in terms of summations with respect to divisors of m as well as decompositions of m into sums of 2 squares (see [22]). Glaisher noted that $r_{10}(m)$ can be equivalently written as a linear combination of three functions (see [14]):

$$\begin{aligned} E_4(m) &= \sum_{2 \nmid d|m} (-1)^{(d-1)/2} d^4 \\ E_4'(m) &= \sum_{2 \nmid d|m} (-1)^{(d-1)/2} \left(\frac{m}{d}\right)^4 \\ \psi_4(m) &= \frac{1}{4} \sum_{\substack{N(\alpha)=m \\ \alpha \in \mathbb{Z}[i]}} \alpha^4. \end{aligned}$$

While E_4 and E_4' look similar to the summations that came up in previous $r_n(m)$, $n < 10$, $\psi_4(m)$ is very distinctive, particularly in its use of $\mathbb{Z}[i]$. Liouville's original formula can be expressed as follows (see [14]):

$$r_{10}(m) = \frac{4}{5} \cdot E_4(m) + \frac{64}{5} \cdot E_4'(m) + \frac{32}{5} \cdot \psi_4(m).$$

Liouville also produced a formula for sums of 12 squares, which was again rewritten by Glaisher as

$$r_{12}(m) = \begin{cases} -8 \cdot \sum_{d|m} (-1)^{d+m/d} d^5 & \text{if } m \text{ is even} \\ 8 \cdot \sum_{d|m} d^5 + 2 \cdot \Omega(m) & \text{if } m \text{ is odd.} \end{cases}$$

Here, Ω can either be defined as coefficients of elliptic function expansions or arithmetically. For the latter, let S_m of all $\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathbb{Z}^4$ such that $x_1^2 + x_2^2 + x_3^2 + x_4^2 = m$.

$$\Omega(m) = \frac{1}{8} \cdot \sum_{\mathbf{x} \in S_m} x_1^4 + x_2^4 + x_3^4 + x_4^4 - 2x_1^2x_2^2 - 2x_1^2x_3^2 - 2x_1^2x_4^2 - 2x_2^2x_3^2 - 2x_2^2x_4^2 - 2x_3^2x_4^2.$$

There is no straightforward method to compute $\Omega(m)$ in polynomial time with respect to $\log(m)$, even when the prime factorization of m is given. In a 1916 article (see [26]), Ramanujan remarked

$$\sum_{k=1}^{\infty} \Omega(m) x^m = \eta^{12}(x^2) \quad \text{where} \quad \eta(x) = x^{1/24} \prod_{k=1}^{\infty} (1 - x^k).$$

In order to express $r_{14}(m)$, $r_{16}(m)$, and $r_{18}(m)$, Glaisher also described functions similar to Ω that were defined as coefficients of elliptic function, and Ramanujan later wrote them as coefficients of expansions of powers and products of $\eta(x)$. In addition, he included the relation

$$\sum_{k=1}^{\infty} \psi_4(m)x^m = \eta^4(x)\eta^2(x^2)\eta^4(x^4).$$

Unfortunately, this equation does not seem to simplify the computations of $\psi_4(m)$ and additionally implies an equal or higher level of difficulty in computing ψ_4 and Ω. However, even though ψ_4 is not simply a summation running over divisors, Fermat's theorem for sums of two squares allows us to understand the ψ_4 quite well. A straightforward consequence of Fermat's theorem is

$$r_2(p) = \begin{cases} 4 & \text{if } p = 2 \\ 8 & \text{if } p \equiv 1 \bmod 4 \\ 0 & \text{if } p \equiv 3 \bmod 4. \end{cases}$$

Using the fact that the norm of $\mathbb{Z}[i]$ is multiplicative, one can compute $\psi_4(m)$ by finding an element of $\mathbb{Z}[i]$ with norm p for each prime $p \mid m$ and $p \equiv 1 \bmod 4$.

The modern perspective on the sums of squares problems involves the theory of modular forms, the sequel to the elliptic functions of Jacobi and Ramanujan. Through this perspective, we will discuss the generating function for $r_n(m)$ when n is even. The fact that these formal series give rise to integral weight modular forms will allow us to precisely understand when formulas for $r_n(m)$ are straightforward and easily computable.

2 Main statements

We begin by recalling the general situation. For even $n \in \mathbb{Z}_{>0}$ and $m \in \mathbb{Z}_{\geq 0}$, we define

$$r_n(m) := \#\{\mathbf{x} \in \mathbb{Z}^n : x_1^2 + x_2^2 + ... + x_n^2 = m\}. \tag{1}$$

We wish to understand the generating function of $r_n(m)$, which we denote by θ_n. A key insight is to interpret this formal series as a complex (in fact, holomorphic) function on the open unit disk $\mathcal{D} \in \mathbb{C}$. For $q \in \mathcal{D}$, let

$$\theta_n(q) = \sum_{m=0}^{\infty} r_n(m) q^m = 1 + 2n \cdot q + 4\binom{n}{2} \cdot q^2 + 8\binom{n}{3} \cdot q^3 + \left[2^4\binom{n}{4} + 2n\right] \cdot q^4 + \dots \tag{2}$$

Equivalently, one can define $\theta_n(q)$ by using the multiplicative property of $r_n(m)$

$$\theta_n(q) = \theta_1(q)^n = \left(1 + 2q + 2q^4 + 2q^9 + ...\right)^n.$$

Furthermore, if we view $\mathcal{D}$ as the image of the upper half plane $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ under the map $z \mapsto e^{2\pi i z}$, then we may write

$$\theta_n(z) = \sum_{m=0}^{\infty} r_n(m) e^{2\pi i m z}.$$

Jacobi noted certain symmetries of θ_n; in particular, it satisfies the equations (see [25], 3.2),

$$\theta_n(-1/4z) = (2z/i)^{n/2}\theta_n(z) \qquad \theta_n(z+1) = \theta_n(z). \tag{3}$$

Note that we have assumed n is even, hence there is no need to choose a square root. The above equalities illustrate that θ_n as a function of the $\mathbb{H}$-coordinate z, the coordinate of the upper half plane, is a modular form of weight $n/2$ on the congruence subgroup $\Gamma_1(4)$ consisting of matrices $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma \equiv \left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right) \pmod 4$ (see [10], 1.2 or [25], 3.2). We are interested in analyzing when $\theta_n(z)$ has coefficients that are easily computable, thus we establish the following definition.

Definition 1. A modular form f on the congruence subgroup $\Gamma_1(N)$ of weight $k \in \mathbb{Z}$ is *elementary* if and only if f is a linear combination of Eisenstein series and cusp forms with complex multiplication as defined in Section 3.9.

Denote the space of modular forms on a congruence subgroup Γ of weight k as $\mathcal{M}_k(\Gamma)$, and let its subspace of cusp forms be $\mathcal{S}_k(\Gamma)$. The Eisenstein space $\mathcal{E}_k(\Gamma)$ is then the orthogonal subspace to $\mathcal{S}_k(\Gamma)$ with respect to the Petersson inner product defined in Section 3.5. Finally, we define the subspace $\mathcal{S}_k^{cm}(\Gamma) \subset \mathcal{S}_k(\Gamma)$ as the space of cusp forms with complex multiplication, i.e. those which are invariant under twisting by a quadratic character.

By definition, $\theta_n \in \mathcal{M}_{n/2}(\Gamma_1(4))$, is elementary if and only if θ_n is an element of the subspace $\mathcal{E}_{n/2}(\Gamma_1(4)) \oplus \mathcal{S}_{n/2}^{cm}(\Gamma_1(4))$. Note that in [33], Serre calls cusp forms *lacunary* if the density of the nonzero coefficents in the q-expansion is zero, and proves that a cusp form f is lacunary if and only if $f \in S^{cm}$. While it is false that θ_n is lacunary for any $n \geq 4$, it is true that θ_n is elementary if and only if the cuspidal part in its decomposition are lacunary, i.e. contribute to the value of $r_n(m)$ very rarely. The following theorem is our main result.

Theorem 1. *Suppose n is even. Then θ_n is elementary if and only if $n = 2, 4, 6, 8,$ or 10.*

To prove this, we will first compute the dimensions of the various subspaces introduced above:

Lemma 1. *The dimensions of $\mathcal{M}_k(\Gamma_1(4))$ and its subspace of cusp forms for arbitrary $k \in \mathbb{Z}_{>0}$ are as follows:*

$$\dim_{\mathbb{C}}(\mathcal{M}_k(\Gamma_1(4))) = \begin{cases} \frac{k+2}{2} & \text{if } k \text{ is even} \\ \frac{k+1}{2} & \text{if } k \text{ is odd} \end{cases} \quad \dim_{\mathbb{C}}(\mathcal{S}_k(\Gamma_1(4)) = \begin{cases} 0 & \text{if } k \leq 4 \\ \frac{k-4}{2} & \text{if } k \geq 3 \text{ is even} \\ \frac{k-3}{2} & \text{if } k \geq 3 \text{ is odd} \end{cases}$$

In particular, this implies that $\mathcal{E}_k(\Gamma_1(4))$ has dimension 3 for even $k > 2$ and dimension 2 for odd $k > 2$. We will prove this by using a geometric interpretation of these spaces and applying the Riemann-Roch formula. When k is odd, the genus formula arising from Riemann-Roch gets a contribution from the irregular cusp $1/2$ on $\Gamma_1(4)$.

A corollary of this statement (and the fact that θ_n is a modular form of weight $n/2$) is that if $n = 2, 4, 6$, and 8, then θ_n is elementary (consisting entirely of Eisenstein series).

Lemma 2. *The dimension of $\mathcal{S}_k^{cm}(\Gamma_1(4))$ is 1 if $k \equiv 1 \bmod 4$ for $k \geq 5$ and 0 otherwise.*

A more general theorem for all $\Gamma_1(N)$ can be found in [28]. We focus on the case of $\Gamma_1(4)$ here, proving that there exists a unique algebraic Hecke character on $\mathbb{Q}(i)$ of conductor 1 and ∞-type equal to $\#\mathcal{O}_{\mathbb{Q}(i)}^{\times} = 4$, and the only possible CM cusp forms on $\Gamma_1(4)$ arise from its powers.

Using these lemmas, we will prove that for even $n > 8$, the modular form θ_n is not a linear combination of Eisenstein series. Thus, the only possible $n > 8$ and even for which θ_n can be elementary are such that $\frac{n}{2} \equiv 1 \bmod 4$. Then we have reduced the problem to producing an elementary formula for $n = 10$, and showing that for $n > 10$ with the above property, any decomposition of θ_n must include cusp forms that do not have complex multiplication.

3 Modular forms

We introduce the general theory and notation of modular forms that will be used throughout this thesis. This material is found in [5], [10], [9], [23], [25], [29], and [41]. In order to prove Lemmas 1 and 2, we will focus on modular forms related to the congruence subgroup $\Gamma_1(4)$, the geometric interpretation of the space of modular forms, and the general theory of modular forms with complex multiplication (see [28]).

3.1 $\mathrm{SL}_2(\mathbb{Z})$ and congruence subgroups

The group $\mathrm{SL}_2(\mathbb{R})$ consisting of 2×2 matrices with determinant 1 and coefficients in $\mathbb{R}$ acts on the upper half plane of the complex numbers, denoted $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$, for any $z \in \mathbb{H}$, we define the linear fractional transformation by γ as

$$\gamma(z) = \frac{az+b}{cz+d} \in \mathbb{H}.$$

The element $-1 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ has trivial action on $\mathbb{H}$, and $\mathrm{SL}_2(\mathbb{R})/\{\pm 1\}$ acts faithfully on $\mathbb{H}$. Although some authors use $\mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$ instead, we will call $\mathrm{SL}_2(\mathbb{Z})$ the *modular group*.

Certain basic functions on $\mathbb{H}$ such as translation, $z \mapsto z+n$ for $n \in \mathbb{Z}$, and the transformation $z \mapsto -1/z$ can be written as matrices in $\mathrm{SL}_2(\mathbb{Z})$:

$$T^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

respectively. Furthermore, it is well known (see [29], 7.1) that the modular group is generated by S and $T = T^1$.

For each $N \in \mathbb{Z}_{>0}$, let $\Gamma(N)$ denote the kernel of the reduction map

$$\varphi : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}).$$

A *congruence subgroup* Γ of $\mathrm{SL}_2(\mathbb{Z})$ is then any subgroup containing some $\Gamma(N)$. The *level* of Γ is defined as the smallest such N for which $\Gamma(N) \subset \Gamma$. We are particularly interested in the following congruence subgroups

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

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

Here $*$ denotes any element of $\mathbb{Z}$. Equivalently, one can think of $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ as acting on $(\mathbb{Z}/N\mathbb{Z}) \times (\mathbb{Z}/N\mathbb{Z})$ and $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$. Then $\Gamma_1(N)$ is the preimage under φ of the stabilizer of the vector $\binom{1}{0} \in (\mathbb{Z}/N\mathbb{Z})^2$ and similarly, $\Gamma_0(N) = \varphi^{-1}(\mathrm{Stab}\,[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}])$ where $[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}] \in \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$. When $N = 1$, $\Gamma_1(1) = \Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z})$.

3.2 Cusps

The action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$ is defined by

$$m \mapsto \gamma(m) = \frac{am+b}{cm+d}, \qquad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Here, $\gamma(\infty) = \frac{a}{c}$ and if $cm + d = 0$, $\gamma(m) = \infty$ (similarly, if $c = 0$, γ fixes ∞).

The *cusps* of a congruence subgroup Γ are the Γ-orbits of $\mathbb{P}^1(\mathbb{Q})$. It is a nontrivial fact that the set of cusps for any congruence subgroup Γ is finite. If $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is only one cusp, i.e. for any $m_1, m_2 \in \mathbb{P}^1(\mathbb{Q})$, there always exists some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma(m_1) = m_2$. If $\Gamma = \Gamma_1(4)$, one can easily check that there are 3 distinct orbits, normally represented by ∞, 0, and $\frac{1}{2}$, and we say that there are the three cusps of $\Gamma_1(4)$.

Remark. Geometrically, adding the cusps of Γ to $\mathbb{H}$ "compactifies" the upper half plane with respect to the action of Γ in the following sense. We can view $\Gamma\backslash\mathbb{H} = Y_\Gamma(\mathbb{C})$ as a *modular curve*, which can also be viewed as a Riemann surface. However, it is not compact, so we also consider the quotient of the action of Γ on $\mathbb{H}^* = \mathbb{H} \cup \mathbb{P}^1(\mathbb{Q})$ (the topology on $\mathbb{H}^*$ is defined by using the usual open sets of $\mathbb{H}$ along with the sets $\gamma(\{x + iy : y > C\} \cup \{\infty\})$ for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $C \in \mathbb{R}_{\geq 0}$). This decomposes to $X_\Gamma(\mathbb{C}) = Y_\Gamma(\mathbb{C}) \cup (\Gamma\backslash\mathbb{P}^1(\mathbb{Q}))$, i.e. our original modular curve with the cusps defined above added to it. $X_\Gamma(\mathbb{C})$ is a compact connected Riemann surface, thus by Riemann's existence theorem (see [30]), we can view and study it as a projective algebraic curve over $\mathbb{C}$. (It is also possible to define X_Γ as a compactified moduli space of elliptic curves with Γ-structure that makes sense over $\mathbb{Q}$. Then one can show that its complex points give this compactification of $\Gamma\backslash\mathbb{H}$ (see [9], II.9))

3.3 Modular functions

For any integer k, define the *weight k (right) action of* $\mathrm{SL}_2(\mathbb{Z})$ on the set of functions $f : \mathbb{H} \to \mathbb{C}$ as follows: For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and f as above, define

$$f \mid [\gamma]_k(z) = \frac{f(\gamma(z))}{(cz + d)^k}.$$

Since $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathbb{H}$ on the left, this yields a right action of $\mathrm{SL}_2(\mathbb{Z})$ on the set of all functions $f : \mathbb{H} \to \mathbb{C}$ as

$$f \mid [\gamma_1\gamma_2]_k = (f \mid [\gamma_1]_k) \mid [\gamma_2]_k.$$

Moreover, this action can be defined for any $\gamma \in \mathrm{GL}_2(\mathbb{R})$ with positive determinant, where we add multiplication by the factor of $\det(\gamma)^{k-1}$ in the right-hand side of the definition of $f \mid [\gamma]_k$.)

A *modular form* of weight $k \geq 0$ with respect to a congruence subgroup Γ is a function $f : \mathbb{H} \to \mathbb{C}$ such that for all $z \in \mathbb{H}$,

1. f is holomorphic on $\mathbb{H}$, i.e. $\lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$ exists independent of the path h may approach 0 on;
2. f is invariant under the weight k action of Γ, i.e. $f \mid [\gamma]_k = f$ for all $\gamma \in \Gamma$;
3. f is holomorphic on the cusps of Γ.

We define the third condition as follows. Note that the matrix $T^N \in \Gamma(N)$, thus there exists a smallest positive integer h such that $f(z + h) = f(z)$ for all $z \in \mathbb{H}$ if f is a candidate for a modular form of level N. In particular, f has a *Fourier expansion* (at ∞)

$$f(z) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n z/h} \quad \forall z \in \mathbb{H}.$$

Let $q^{1/h} = q^{1/h}(z) = e^{2\pi i z/h}$, which we view as a map $\mathbb{H} \to \mathbb{D}^*$, where $\mathbb{D}^*$ denotes the punctured open unit disk (i.e., with origin removed). We then say f is *holomorphic at* ∞ if the map

$F : \mathbb{D}^* \to \mathbb{C}$ defined by $F(q(z)) = f(z)$ extends to and is well-behaved at 0, i.e. if $a_n = 0$ for all $n < 0$ (one can check that this condition is independent of the choice of h).

Note that for $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, $f \mid [\alpha\gamma]_k = f \mid [\alpha]_k$ for all $\gamma \in \alpha^{-1}\Gamma\alpha$, thus the Fourier expansion at ∞ of f immediately gives one for $f \mid [\alpha]_k$. Then, f is *holomorphic at the cusps* of Γ if $f \mid [\alpha]_k$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Furthermore, f *vanishes at the cusps* if $f \mid [\alpha]_k$ is both holomorphic at ∞ and $a_0 = 0$ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. Checking if $a_0 = 0$ when $\alpha = 1$ gives the criterion for vanishing at ∞.

The complex vector space of all modular forms of weight k on Γ will be denoted by $\mathcal{M}_k(\Gamma)$. An important subspace $\mathcal{S}_k(\Gamma)$ consists of all modular forms that also vanish on all the cusps of Γ, known as the space of *cusp forms*. It has an orthogonal complement, denoted $\mathcal{E}_k(\Gamma)$, with respect to the inner product defined in Section 3.5. The space $\mathcal{E}_k(\Gamma)$ consists of modular forms called *Eisenstein series* which do not vanish at every cusp of Γ. It is well known that $\mathcal{M}_k(\Gamma) = \mathcal{S}_k(\Gamma) \oplus \mathcal{E}_k(\Gamma)$ has finite dimension over $\mathbb{C}$ (see [10]).

Remark. The matrix $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma_1(N) \subseteq \Gamma_0(N)$ for all N, thus any modular form $f \in \mathcal{M}_k(\Gamma_1(N)) \supseteq \mathcal{M}_k(\Gamma_0(N))$ has a q-expansion at ∞ of the form

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

If $f(z) \in \mathcal{S}_k(\Gamma_1(N))$ or $\mathcal{S}_k(\Gamma_1(N))$, then $a_0 = 0$ as well (but the converse implication does not hold).

3.4 Hecke operators

We first restrict to the case of level 1 modular forms. For any positive integer n, let

$$X_n = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Z}) : a \geq 1, ad = n, \text{ and } 0 \leq b < d \right\}.$$

It is not hard to see that X_n is in bijection with the set of sublattices of $\mathbb{Z}^2$ of index n (by letting the rows, (a, b) and $(0, d)$ define basis elements). Recall that the weight k action of $\gamma \in X_n$ on a function $f : \mathbb{H} \to \mathbb{C}$ is

$$(f \mid [\gamma]_k)(z) = n^{k-1} \cdot d^{-k} \cdot f\left(\frac{az + b}{d}\right).$$

The *n-th Hecke operator* of weight k, denoted $T_{n,k}$ (or T_n since the weight will always be obvious, corresponding to the weight of the modular form) is the operator on the set of functions on $\mathbb{H}$ defined by

$$T_{n,k}(f) = \sum_{\gamma \in X_n} f \mid [\gamma]_k.$$

The Hecke operators of a fixed weight k satisfy the following formulas (see [29], 7.5.2, Lemma 2):

$$\begin{aligned} T_m T_n &= T_{mn} && \text{if } \gcd(m, n) = 1, \\ T_{p^n} &= T_{p^{n-1}} T_p - p^{k-1} T_{p^{n-2}} && \text{if } p \text{ is prime.} \end{aligned}$$

So in particular, the prime power Hecke operators T_{p^n} can be written as integer-polynomials in T_p. Furthermore, Hecke operators commute, i.e. for any $n, m \in \mathbb{Z}$, $T_n T_m = T_m T_n$.

On modular forms of level 1, we can write the action of T_n explicitly when f is written as a q-expansion. If $q = e^{2\pi iz}$ then the Fourier expansion of a modular form f of weight k is written as $\sum_{n\geq 0} a_n q^n$. A Hecke operator T_n (of weight k) acts on f as

$$T_n(f) = \sum_{m\in\mathbb{Z}} \left(\sum_{d|\gcd(m,n)} d^{k-1} a_{mn/d^2} \right) q^m,$$

where the summation runs over positive d ([29], 7.5.3). If f is a modular form, then $T_n(f)$ is also a modular form of the same weight ([29], 7.5.3, Prop. 12).

For the action of Hecke operators on modular forms of higher level N, first define the *diamond operator* for all $d \nmid N$,

$$\langle d\rangle_k : f \mapsto f \mid [\sigma_d]_k \quad \text{where } \sigma_d \equiv \begin{pmatrix} \overline{d} & 0 \\ 0 & d \end{pmatrix} \bmod N,$$

where $\overline{d} \equiv d^{-1} \bmod N$, and σ_d is any element of the modular group satisfying the equation. Indeed, the action of $\langle d\rangle_k$ only depends on $d \bmod N$, not on the choice of σ_d.

The *n-th Hecke operator* of level N is then the operator on the set of functions on $\mathbb{H}$ defined by

$$T_n(f) = \sum_{\gamma\in X_n} (\langle a_\gamma\rangle f) \mid [\gamma]_k;$$

here, a_γ denotes the top left entry (i.e. the "a" entry) of the matrix γ. (Here, we are implicitly assuming that f is a weight k modular form and the diamond operator is of the same weight, thus T_n is a weight k action.) The Hecke operators of arbitrary level satisfy formulas similar to those in the level 1 case. In particular,

$$\begin{aligned} T_m T_n &= T_{mn} && \text{if } \gcd(m,n) = 1; \\ T_{p^n} &= T_{p^{n-1}} T_p - p^{k-1}\langle p\rangle T_{p^{n-2}} && \text{if } p \nmid N \text{ is prime.} \end{aligned}$$

Remarks. In our notation, we are using the fact that diamond operators and Hecke operators commute (with themselves and each other), (see [10], 5.2), e.g. it is okay to use notation $T_n f$ rather than $f \mid T_n$.

Furthermore, the action of Hecke operators and diamond operators preserve the decomposition of the space of modular forms of a given weight into the cusp forms and the Eisenstein series.

3.4.1 Operators on $\mathcal{M}_k(\Gamma_0(N))$ and $\mathcal{M}_k(\Gamma_1(N))$

As $\Gamma_1(N) \subseteq \Gamma_0(N)$, all modular forms on $\Gamma_0(N)$ are modular forms on $\Gamma_1(N)$. The converse is not true, but one can consider the action of elements of $\Gamma_0(N)$ on $f \in \mathcal{M}_k(\Gamma_1(N))$, because $\Gamma_1(N)$ is a normal subgroup of $\Gamma_0(N)$ and furthermore, $\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/N\mathbb{Z})^\times$. The diamond operators $\langle d\rangle_k$ defined earlier act on the space $\mathcal{M}_k(\Gamma_1(N))$, and they in fact represent the action of $\Gamma_0(N)/\Gamma_1(N)$ on $\mathcal{M}_k(\Gamma_1(N))$. If $\varepsilon : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$ is a Dirichlet character mod N, view it as a map from $\mathbb{Z}$ by defining $\varepsilon(p) = 0$ for primes $p \mid N$ and extending multiplicatively. We say that a modular form $f \in \mathcal{M}_k(\Gamma_1(N))$ has *Nebentypus* ε if it satisfies

$$(f \mid [\gamma]_k)(z) = \varepsilon(d_\gamma) f(z), \quad \forall \gamma \in \Gamma_0(N),$$

where d_γ denotes the bottom right entry (the "d" entry) of the matrix γ. For a fixed Nebentypus ε, these modular forms form a subspace of $\mathcal{M}_k(\Gamma_1(N))$ denoted $\mathcal{M}_k(\Gamma_0(N), \varepsilon)$. Moreover, the

space $\mathcal{M}_k(\Gamma_0(N), \varepsilon)$ can be thought of as the ε-eigenspace of the $\langle d \rangle$ for $d \in (\mathbb{Z}/N\mathbb{Z})^\times$.Thus, an equivalent definition of this space is

$$\mathcal{M}_k(\Gamma_0(N), \varepsilon) = \{f \in \mathcal{M}_k(\Gamma_1(N)) : \langle d \rangle f = \varepsilon(d) f \quad \forall d \in (\mathbb{Z}/N\mathbb{Z})^\times\}.$$

It follows that we have a decomposition of $\mathcal{M}_k(\Gamma_1(N))$ into subspaces $\mathcal{M}_k(\Gamma_0(N), \varepsilon)$, indexed by the Dirichlet characters mod N where $\mathcal{M}_k(\Gamma_0(N), 1_N) = \mathcal{M}_k(\Gamma_0(N))$ when the Nebentypus is the trivial character 1_N (see [10], 4.3). Since $-\mathbf{1} = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \Gamma_0(N)$, if the Nebentypus of $f \in \mathcal{M}_k(\Gamma_1(N))$ is ε, then

$$f \mid [-\mathbf{1}](z) = \varepsilon(-1) f(z) \Rightarrow (-1)^{-k} f(z) = \varepsilon(-1) f(z),$$

for all $z \in \mathbb{H}$. Hence we conclude that in order for $\mathcal{M}_k(\Gamma_0(N), \varepsilon)$ to be nontrivial, the Nebentypus ε must have the property that $\varepsilon(-1) = (-1)^k$. Thus, we conclude that

$$\mathcal{M}_k(\Gamma_1(N)) = \bigoplus_{\varepsilon} \mathcal{M}_k(\Gamma_0(N), \varepsilon) \quad \text{where} \quad \varepsilon(-1) = (-1)^k. \tag{4}$$

The action of Hecke operators T_n preserves the decomposition for n coprime to N, i.e. if $f \in \mathcal{M}_k(\Gamma_0(N), \varepsilon)$, then $T_n f \in \mathcal{M}_k(\Gamma_0(N), \varepsilon)$ and we can write out the q-series expansion of $T_n f$ in terms of $f(z) = \sum_{m \geq 0} a_m q^m$

$$T_n f = \sum_{m \geq 0} \left(\sum_{d \mid \gcd(n,m)} \varepsilon(d) \cdot d^{k-1} \cdot a_{mn/d^2} \right) q^m,$$

where d runs through positive divisors and ε is the Dirichlet character viewed as a map on $\mathbb{Z}$ (see [10], 5.3.1).

In the space $\mathcal{S}_k(\Gamma_1(N))$, cusp forms of a fixed Nebentypus ε form a subspace denoted $\mathcal{S}_k(\Gamma_0(N), \varepsilon)$. The diamond operators preserve cusp forms, thus the restriction of their action to the subspace of cusp forms partition the space analogously with respect to the possible Dirichlet characters ε mod N:

$$\mathcal{S}_k(\Gamma_1(N)) = \bigoplus \mathcal{S}_k(\Gamma_0(N), \varepsilon), \quad \text{where} \quad \varepsilon(-1) = (-1)^k.$$

In particular, the cuspidal ε-eigenspace $\mathcal{S}_k(\Gamma_0(N), \varepsilon) = \mathcal{M}_k(\Gamma_0(N), \varepsilon) \cap \mathcal{S}_k(\Gamma_1(N))$. Consequently, Hecke operators preserve this decomposition as well as the decomposition $\mathcal{E}_k(\Gamma_1(N)) = \bigoplus \mathcal{E}_k(\Gamma_0(N), \varepsilon)$ (see [10], 5.2).

3.5 Petersson inner product

If $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup, there is a "natural" inner product on the cuspidal space $\mathcal{S}_k(\Gamma)$ known as the *Petersson inner product*. It allows us to focus our attention on certain types forms that are eigenvectors for all Hecke and diamond operators.

If $z = x + iy \in \mathbb{H}$, then the *hyperbolic measure* $d\mu(z) := \frac{dxdy}{y^2}$ on $\mathbb{H}$ is $\mathrm{SL}_2(\mathbb{Z})$-invariant. It induces a measure on $\Gamma\backslash\mathbb{H}$, which is given by a smooth volume form outside the elliptic points. In fact, the integral $\int_{\Gamma\backslash\mathbb{H}} d\mu$ converges to the volume

$$V_\Gamma = [\mathrm{SL}_2(\mathbb{Z}) : \pm\Gamma] \left(\frac{\pi}{3}\right).$$

Note that for $f \in \mathcal{S}_k(\Gamma)$, $|f(z)|^2 y^k$ is Γ-invariant and bounded on $\mathbb{H}$, hence the measure

$$d\mu_f(z) := |f(z)|^2 y^{k-2} dxdy = |f(z)|^2 y^k d\mu,$$

is Γ-invariant on $\mathbb{H}$, and furthermore, $\int_{\Gamma\backslash\mathbb{H}} d\mu_f$ converges to an element of $\mathbb{R}_{\geq 0}$ (see [10], 5.4). Thus, there is an inner product

$$\langle f, g\rangle = \frac{1}{V_\Gamma} \int_{\Gamma\backslash\mathbb{H}} f(z)\overline{g(z)} y^k d\mu, \quad f, g \in \mathcal{S}_k(\Gamma),$$

In fact, this inner product can be extended to a sesquilinear pairing $\mathcal{M}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \to \mathbb{C}$; however, it is not an inner product on $\mathcal{M}_k(\Gamma)$ as the integral does not converge in the larger space.

The set of $f \in \mathcal{M}_k(\Gamma_1(N))$ such that $\langle f, g\rangle = 0$ for all $g \in \mathcal{S}_k(\Gamma_1(N))$ is exactly $\mathcal{E}_k(\Gamma_1(N))$. (The statement also holds true for $\Gamma_0(N)$) (see [10], 5.4).

3.5.1 Eigenforms & newforms

On the space of cusp forms $\mathcal{S}_k(\Gamma_1(N))$, one can show that the diamond and Hecke operators away from the level are normal, i.e. they commute with their adjoints with respect to the Petersson inner product. From linear algebra, $\mathcal{S}_k(\Gamma_1(N))$ then has an orthogonal basis of elements which are eigenvectors simultaneously for all the operators away from the level. We define an *eigenform* as a nonzero modular form $f \in \mathcal{M}_k(\Gamma_1(N))$ with this above property, i.e. an eigenform is an eigenvector for all Hecke and diamond operators of level coprime to N. However, the eigenspaces attached to these eigenforms may not necessarily be 1-dimensional.

In general, we say an eigenform f is *normalized* if the q-expansion of f has coefficient 1 for q. Normalization is motivated by the fact that it forces an eigenform $f \in \mathcal{M}_k(\Gamma_0(N), \varepsilon)$ to have q-series coefficients described by the action of the Hecke operators ($T_n(f) = a_n f$ when $\gcd(n, N) = 1$).

As a consequence, the q-series coefficients of a normalized eigenform $\sum_{n\geq 0} a_n q^n \in \mathcal{M}_k(\Gamma_0(N), \varepsilon)$ must satisfy $a_1 = 1$ along with:

1. $a_{p^r} = a_{p^{r-1}} a_p - \varepsilon(p) p^{k-2} a_{p^{r-2}}$ for all primes $p \nmid N$ and $r \geq 2$
2. $a_{mn} = a_m a_n$ when m and n are coprime to the level, and $\gcd(m, n) = 1$.

Suppose M and N are positive integers such that $M \mid N$. For any divisor $t \mid \frac{N}{M}$, define the t-th degeneracy map of cusp forms of level M to those of level N as

$$\imath_{t,M} : \mathcal{S}_k(\Gamma_1(M)) \hookrightarrow \mathcal{S}_k(\Gamma_1(N)), \quad \text{where} \quad \imath_{t,M} : f(z) \mapsto f(tz).$$

On q-expansions, $\imath_{t,M}$ sends $\sum_{n=0}^{\infty} a_n q^n \mapsto \sum_{n=0}^{\infty} a_n q^{tn}$. This map commutes with the action of the diamond and Hecke operators coprime to N described previously (see [10], 5.6). Note also that when $t = 1$, $\imath_{t,M}$ is the identity inclusion.

The *old subspace* of $\mathcal{S}_k(\Gamma_1(N))$ is the sum of the images of all such $\imath_{t,M}$ where M runs through proper divisors of N, and t runs through all divisors of $\frac{N}{M}$ (given M). We define the *new subspace* to then be the orthogonal subspace in $\mathcal{S}_k(\Gamma_1(N))$ with respect to the Petersson inner product, so in particular, we have the following decomposition,

$$\mathcal{S}_k(\Gamma_1(N)) = \mathcal{S}_k(\Gamma_1(N))_{new} \oplus \mathcal{S}_k(\Gamma_1(N))_{old}.$$

(The names are derived from the idea that forms from the old subspace originate from lower levels, i.e. proper divisors M of N, while the forms from new subspace do not.) The Hecke and diamond operators away from the level respect the decomposition of $\mathcal{S}_k(\Gamma_1(N))$ into old and new subspaces, and furthermore, both subspaces have bases of eigenforms (see [10], 5.6). We call the normalized eigenforms for $\mathcal{S}_k(\Gamma_1(N))_{new}$, *newforms*. The set of newforms is a basis for the new subspace in $\mathcal{S}_k(\Gamma_1(N))$, and in particular, the eigenspaces in the new subspace each have dimension 1 (see [21]). Thus, since Hecke operators commute with each other, newforms are eigenforms for all Hecke operators, including those that are not coprime to the level. We can decompose $\mathcal{S}_k(\Gamma_1(N))$ as follows (due to Atkin and Lehner [1] and Li [21]):

$$\mathcal{S}_k(\Gamma_1(N)) = \bigoplus_{M|N} \bigoplus_{t|\frac{N}{M}} i_{t,M}(\mathcal{S}_k(\Gamma_1(M))_{new}).$$

Thus, for any normalized eigenform $g \in \mathcal{S}_k(\Gamma_1(N))$, there exists a unique newform $f \in \mathcal{S}_k(\Gamma_1(M))_{new}$ for some $M \mid N$ such that the coefficients of q^n for n coprime to the level in the q-series expansions of g and f coincide. This decomposition allows us to view *newforms* as eigenforms for all Hecke operators, including those that are not coprime to the level. A priori, a newform $f = \sum a_n q^n$ has the property that the eigenvalue of T_n for n coprime to N is the nth coefficient of the q-series for f. For any positive $n \in N$ with nontrivial $\gcd(n, N) = 1$, the "additional" Hecke operators T_n also satisfy $T_n f = a_n f$ (see [21]). This coincides with the earlier formulas (1 and 2 above) viewing ε as a map on $\mathbb{Z}$ where $\varepsilon(p) = 0$ if p is a prime dividing N, and extending multiplicatively. (For diamond operators $\langle d \rangle$ such that $\gcd(d, N) \neq 1$, we define $\langle d \rangle f = 0$, hence f is automatically an eigenform for such $\langle d \rangle$, with eigenvalues equal to 0.)

3.6 Geometric view

Modular forms on a congruence subgroup Γ also have a geometric interpretation, as holomorphic sections of line bundles on the corresponding modular curves introduced in Section 3.2. The main reference for this entire section is [9], II.

Let $k \in \mathbb{Z}_{\geq 0}$ and Γ a congruence subgroup satisfying the following conditions:

1. Either $k = 0$ or the image of Γ under the projection $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$ acts freely on $\mathbb{H}$.

2. If k is odd, then the cusps have unipotent stabilizer in Γ, i.e. the eigenvalues are 1. This only occurs when $-1 \notin \Gamma$. Under this assumption, if a cusp written as $\gamma(\infty)$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ has unipotent stabilizer in Γ if it is contained in $\gamma S_\infty \gamma^{-1}$ where $S_\infty = \{\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)\}$.

We call such cusps *regular*. Let X denote the modular curve $\Gamma\backslash\mathbb{H}^*$, and $Y = \Gamma\backslash\mathbb{H}$. $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathbb{C} \times \mathbb{H}$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : (\tau, z) \mapsto \left((cz + d)^k \tau,\ \frac{az + b}{cz + d} \right).$$

The quotient of the action of Γ on $\mathbb{C} \times \mathbb{H}$ has a natural projection to Y, giving $\Gamma\backslash(\mathbb{C} \times \mathbb{H})$ the structure of a complex line bundle over Y (see [9]). We extend it to a line bundle over X via the trivial action on open neighborhoods of a cusp in $\mathbb{H}^*$, defined as:

$$\gamma(\tau, z) \mapsto (\tau, \gamma(z)), \quad \text{for any } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \text{ and } z = x + iy \text{ with } y > 0.$$

(The image under $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ of the sets $\{x+iy : y > C\} \cup \{\infty\}$ extend the base of open sets of $\mathbb{H}$ to $\mathbb{H}^*$.) The sections that are defined to be generators are exactly those with q-expansions $\sum_{m=0}^{\infty} a_m q^m$ such that $a_0 \neq 0$. Denote the resulting line bundle as $\underline{\omega}_k$ and let $\psi : \underline{\omega}_k \to X$ be the projection map. Consider the sheaf $\mathcal{G}_k$ on X of holomorphic sections on $\underline{\omega}_k$. It is an invertible sheaf of $\mathcal{O}_X$-modules, where $\mathcal{O}_X$ denotes the sheaf of holomorphic functions on X (also, $\mathcal{O}_X = \mathcal{G}_0$). A modular form f of weight k on Γ defines an element of $\mathcal{G}_k(Y)$ which sends $z \mapsto (f(z), z)$. Since f is holomorphic at the cusps of Γ, we automatically get that this element extends to a holomorphic section $\phi_f : X \to \underline{\omega}_k$. In fact, $f \mapsto \phi_f$ produces a natural correspondence between spaces $\mathcal{M}_k(\Gamma)$ and $\mathcal{G}_k(X) = H^0(X, \mathcal{G}_k)$.

For the analogous interpretation of cusp forms, let $\mathcal{C}_k$ denote the subsheaf of holomorphic functions on X which vanish at the cusps, inside $\mathcal{O}_X$. We can define $\mathcal{F}_k = \mathcal{G}_k \otimes_{\mathcal{O}_X} \mathcal{C}_k$ as the invertible sheaf of $\mathcal{O}_X$-modules on X, which naturally lies in $\mathcal{G}_k$. Thus, $\mathcal{F}_k(X) = H^0(X, \mathcal{F}_k)$ lying inside $\mathcal{G}_k(X)$ corresponds to the cusp forms $\mathcal{S}_k(\Gamma)$.

3.6.1 $\Gamma_1(4)$ and its irregular cusp

Although $\Gamma_1(4)$ acts freely on $\mathbb{H}$, the cusp $1/2$ is irregular, i.e. the stabilizer of $\gamma(\infty)$ contains the element $\gamma \left(\begin{smallmatrix} -1 & * \\ 0 & -1 \end{smallmatrix}\right) \gamma^{-1}$, with eigenvalues equal to -1. Thus, the above discussion only applies to $\Gamma = \Gamma_1(4)$ when the weight k is even. For k odd, consider the normal subgroup $\Gamma(4) \trianglelefteq \Gamma_1(4)$. One can check that its 6 cusps are regular, and since its image in $\mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$ has no nontrivial elements of finite order, $\Gamma(4)$ satisfies the above conditions, in particular when k is odd (see [25], 4.2.10).

Let $Y' = \Gamma(4)\backslash\mathbb{H}$, and $X' = \Gamma(4)\backslash\mathbb{H}^*$. There is a natural projection map $\pi : X' \twoheadrightarrow X$. Following the above discussion for $\Gamma(4)$, we can produce an invertible sheaf $\mathcal{G}'_k$ of $\mathcal{O}_{X'}$-modules on X'. Furthermore, we can define an action of $\Gamma_1(4)$ on the direct image sheaf $\pi_*\mathcal{G}'_k$ which factors through the quotient $\Gamma_1(4)/\Gamma(4)$. In particular, in the natural correspondence between $\pi_*\mathcal{G}'_k(X) = \mathcal{G}'_k(X')$ and $\mathcal{M}_k(\Gamma(4))$, the action of γ on the sections coincides with the action of the operator $| \, [\gamma^{-1}]_k$ on the space of modular forms. Let $\mathcal{G}_k = (\pi_*\mathcal{G}'_k)^{\Gamma_1(4)}$ be the subsheaf consisting of sections that are invariant under the action of $\Gamma_1(4)$. It is an invertible sheaf of $\mathcal{O}_X$-modules and we can conclude,

$$\mathcal{M}_k(\Gamma_1(4)) = \mathcal{G}_k(X) = H^0(X, \mathcal{G}_k).$$

For cusp forms, we let $\mathcal{F}'_k \subset \mathcal{G}'_k$ be the invertible sheaf of $\mathcal{O}_{X'}$-modules on X' obtained by tensoring the subsheaf of holomorphic functions on X' which vanish at its cusps with $\mathcal{G}'_k$. The action of $\Gamma_1(4)$ on $\pi_*\mathcal{G}'_k$ restricts to an action on $\pi_*\mathcal{F}'_k$, thus analogously, we let $\mathcal{F}_k = (\pi_*\mathcal{F}')^{\Gamma_1(4)}$. $\mathcal{F}_k$ is an invertible subsheaf of $\mathcal{G}_k$ of $\mathcal{O}_X$-modules. If $\mathcal{C}_k$ is the sheaf of holomorphic functions which vanish at the regular cusps of $\Gamma_1(4)$, then it is also true that $\mathcal{F}_k = \mathcal{G}_k \otimes_{\mathcal{O}_X} \mathcal{C}_k$. This results in

$$\mathcal{S}_k(\Gamma_1(4)) = \mathcal{F}_k(X) = H^0(X, \mathcal{F}_k).$$

Remark. The above construction for odd k is not dependent on $\Gamma(4)$. Starting with another normal subgroup Γ' of $\Gamma_1(4)$ satisfying the regularity and freeness conditions would have resulted in sheaves that were canonically isomorphic to $\mathcal{G}_k$ and $\mathcal{F}_k$. In fact, the entire discussion when k is odd holds for any congruence subgroup Γ. One must choose a normal subgroup Γ' satisfying the two stated conditions, and if $\Gamma = \Gamma'$, the two definitions of $\mathcal{G}_k$ (and $\mathcal{F}_k$) coincide.

3.7 Proof of Lemma 1

To compute the dimension of $\mathcal{M}_k(\Gamma_1(4))$ and $\mathcal{S}_k(\Gamma_1(4))$, we can now use the Riemann-Roch formula (below is the formulation necessary for Lemma 1) (see [15]).

Theorem 2 (Riemann-Roch)**.** *Let $\mathcal{R}$ be a compact Riemann surface of genus g. If D is a divisor on $\mathcal{R}$ such that* $\deg(D) > 2g-2$*, then*

$$\dim_{\mathbb{C}} H^0(\mathcal{R}, D) = \deg(D) - g + 1.$$

First, assume $k = 2$. We can show that $\mathcal{F}_2$ can be viewed as the $\mathcal{O}_X$-sheaf of holomorphic differentials on X. For an open subset U of X, consider a differential $\omega \in \Omega^1_X(U)$. If φ denotes the natural map $\varphi : \mathbb{H} \to X$, let $\varphi^*\omega = f(z)dz$ where f is a holomorphic function on $\varphi^{-1}(U)$. We can then define a holomorphic map $U \cap Y \to \underline{\omega}_2$ which sends $z \mapsto (z, f(z))$; this is an element of $\mathcal{F}_2(U \cap Y)$ and extends uniquely to an element ϕ_ω of $\mathcal{F}_2(U)$. In fact, the map sending $\omega \mapsto \phi_\omega$ turns out to be a $\mathcal{O}_X(U)$-linear isomorphism between $\Omega^1_X(U)$ and $\mathcal{F}_2(U)$ (see [9]). One can check that this is compatible with restriction, thus we can conclude that $\Omega^1_X \cong \mathcal{F}_2$. Since $\mathcal{S}_2(\Gamma_1(4)) = H^0(X, \mathcal{F}_2)$, we furthermore get an isomorphism

$$\Omega^1_X(X) \xrightarrow{\sim} \mathcal{S}_2(\Gamma_1(4)), \qquad \omega \mapsto f(z) \quad \text{where } \varphi^*\omega = f(z)dz.$$

This also allows us to conclude that the dimension of $\mathcal{S}_2(\Gamma_1(4))$ is equal to the genus of X. (see [25]). Note that the weight 2 case does not utilize 2

Recall from Section 3.2 that $X = \Gamma_1(4)\backslash\mathbb{H}^*$ is a compact Riemann surface. To compute its genus, we use the following general fact: if a congruence subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ and $\gamma \in \mathrm{GL}_2(\mathbb{Q})$ with positive determinant satisfy

$$\Gamma \subset \gamma^{-1}\,\mathrm{SL}_2(\mathbb{Z})\gamma,$$

then the map $\tau \mapsto \gamma(\tau)$ on points $\tau \in \mathbb{H}^*$ induces a holomorphic map (see [9])

$$\Gamma\backslash\mathbb{H}^* \longrightarrow \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}^* = \mathbb{H} \cup \{\infty\} \cong \mathbb{P}^1(\mathbb{C}).$$

Viewed as a cover of the Riemann sphere, this map can have ramification over the cusp $\{\infty\}$ and i and $\zeta = e^{\pi i/3}$, the points with non-trivial stabilizer in $\mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$. On the Riemann sphere, these correspond to the *elliptic points* 0, 1728, and $\infty \in \mathbb{P}^1(\mathbb{C})$. The Riemann-Hurwitz formula then tells us that the genus of X can be calculated by

$$g(X) = 1 + \frac{[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(4)]}{24} - \frac{\nu_i}{4} - \frac{\nu_\zeta}{3} - \frac{\#\text{ cusps}}{2},$$

where $\nu_{\{i,\zeta\}}$ denotes the number of elliptic points over i and ζ (see [36], 1.6). Since $\Gamma_1(4)$ is an index 12 subgroup of $\mathrm{SL}_2(\mathbb{Z})$ with 3 cusps and no elliptic points, we can conclude that $g(X) = 0$, i.e. the genus of the modular curve $\Gamma_1(4)\backslash\mathbb{H}^*$ is 0, and thus $\mathcal{S}_2(\Gamma_1(4))$ is trivial.

For arbitrary even k, note that $\underline{\omega}_k \cong \underline{\omega}_1^{\otimes k}$ naturally, thus it is also true that $\mathcal{G}_k \cong \mathcal{G}_1^{\otimes k}$, tensoring over $\mathcal{O}_X$. Furthermore, the isomorphism $\Omega^1_X \cong \mathcal{F}_2$ for $k = 2$ induces

$$\mathcal{F}_k \cong \mathcal{G}_{k-2} \otimes_{\mathcal{O}_X} \Omega^1_X$$

for all even k. We can also calculate the degree of these sheaves $\mathcal{G}_k$ and $\mathcal{F}_k$ as

$$\begin{aligned} \deg(\mathcal{G}_k) &= (g-1)k + (\#\text{ of cusps}) \cdot \frac{k}{2} = \frac{k}{2}, \\ \deg(\mathcal{F}_k) &= (g-1)k + (\#\text{ of cusps})(\frac{k}{2} - 1) = \frac{k}{2} - 3. \end{aligned}$$

Since the genus is 0, $\mathcal{G}_k$ always satisfies $\deg(\mathcal{G}_k) > 2g-2$ since k is non-negative. However, $\deg(\mathcal{F}_k) > 2g-2$ only when $k > 2$. For these cases, the Riemann-Roch formula gives

$$\begin{aligned}\dim_{\mathbb{C}} \mathcal{M}_k(\Gamma_1(4)) &= 1 - g(X) + \deg(\mathcal{G}_k) = \frac{k}{2} + 1, \\ \dim_{\mathbb{C}} \mathcal{S}_k(\Gamma_1(4)) &= 1 - g(X) + \deg(\mathcal{F}_k) = \frac{k}{2} - 2.\end{aligned}$$

When k is odd, there is still a natural map $\mathcal{G}_1^{\otimes k} \to \mathcal{G}_k$ arising from $\underline{\omega}_1^{\otimes k} \cong \underline{\omega}_k$. However, it is not necessarily an isomorphism, specifically at the cusps of $\Gamma_1(4)$. Let $\mathcal{D}_k$ be the sheaf of holomorphic functions with zeroes of order at least $k/2$. Then we have an isomorphism

$$\mathcal{G}_1^{\otimes k} \cong \mathcal{G}_k \otimes_{\mathcal{O}_X} \mathcal{D}_k$$

which can be checked by computing on the stalks of the cusps. In particular, the irregular cusp takes away from the degree of $\mathcal{G}_k$ as computed before, and in fact

$$\begin{aligned}\deg(\mathcal{G}_k) &= (g-1)k + (\#\text{ of reg. cusps}) \cdot \frac{k}{2} + (\#\text{ of irreg. cusps}) \cdot \frac{k-1}{2} \\ \deg(\mathcal{S}_k) &= (g-1)k + (\#\text{ of reg. cusps}) \cdot \left(\frac{k}{2} - 1\right) + (\#\text{ of irreg. cusps}) \cdot \frac{k-1}{2}\end{aligned}$$

(for details, see [36], 2.4 & 2.6 or [25], 2.5). The degrees for both $\mathcal{G}_k$ and $\mathcal{S}_k$ are strictly greater than $-2 = 2g-2$ when $k > 2$, thus the Riemann-Roch formula says

$$\dim_{\mathbb{C}} \mathcal{M}_k(\Gamma_1(4)) = 1 - g + \deg(\mathcal{G}_k) = \frac{k+1}{2} \tag{5}$$

$$\dim_{\mathbb{C}} \mathcal{S}_k(\Gamma_1(4)) = 1 - g + \deg(\mathcal{F}_k) = \frac{k-3}{2} \tag{6}$$

It follows from a similar argument (and the fact that the number of regular cusps is greater than $2g-2$) that $\dim_{\mathbb{C}} \mathcal{M}_1(\Gamma_1(4)) = \frac{(\#\text{ of reg. cusps})}{2} = 1$ and $\dim_{\mathbb{C}} \mathcal{S}_1(\Gamma_1(4)) = 0$. For details, see 2.5 of [25]. □

3.8 *L*-functions and the Mellin transform

To a modular form $f(z) = \sum_{m=0}^{\infty} a_m e^{2\pi i m z}$, one can attach a the Dirichlet L-series,

$$L(s, f) = \sum_{m=1}^{\infty} a_m m^{-s},$$

and vice versa. However, this correspondence between L-series and modular forms is more than a formal relationship between series. One can obtain $L(s, f)$ from $f(z)$ by means of the Mellin transformation (see [**?**]).

$$\int_0^{\infty} f(iy) y^{s-1} dy = \Gamma(s)(2\pi)^{-s} L(s, f) = \Lambda(s, f).$$

Here, $\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1} dt$ is the usual gamma function associated to the Riemann zeta function $\zeta(s)$. Furthermore, to obtain $f(z)$ from a Dirichlet L-series $L(s) = \sum_{m=1}^{\infty} a_m m^{-s}$, we use the inverse Mellin transform

$$f(iy) = \frac{1}{(2\pi i)} \int \Lambda(s, f) x^{-s} ds,$$

where the integral is taken on a vertical line in the right half of the complex plane. If $f(z)$ is an Eisenstein series, the constant coefficient is constructed by looking at the residue of Λ at $s = 0$. Otherwise, if $f(z)$ is cuspidal, $L(s)$ is absolutely convergent. More precisely, we have the following theorem.

Theorem 3. *Let N be a positive integer and ε a Dirichlet character defined mod N. Let r be a positive integer coprime to N and and χ a primitive Dirichlet character defined mod r. For $f(z) = \sum_{m=0}^{\infty} a_m e^{2\pi i m z} \in \mathcal{M}_k(\Gamma_0(N), \varepsilon)$, let*

$$L_\chi(s, f) = \sum_{m=1}^{\infty} \chi(m) a_m m^{-s} \quad \textit{and} \quad \Lambda_\chi(s, f) = \left(\frac{r\sqrt{N}}{2\pi}\right)^s \Gamma(s) L_\chi(s, f).$$

Then $L_\chi(s, f)$ can be holomorphically continued to the whole s-plane. Moreover, it satisfies a functional equation

$$\Lambda_\chi(s, f) = (-1)^{k/2} \varepsilon(r) \chi(N) \left(\sum_{j=0}^{r-1} \chi(j) e^{2\pi i j/r}\right)^2 r^{-1} \Lambda_{\overline{\chi}}(k - s, f \mid [\gamma]_k),$$

where $\gamma = \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$. If $f(z) \in \mathcal{S}_k(\Gamma_0(N), \varepsilon)$, then $L_\chi(s, f)$ is absolutely convergent for $\mathrm{Re}(s) > 1 + (k/2)$.

This theorem describing the correspondence between $f(z)$ and $L(s)$ is due to Hecke (see [36], Thm. 3.66). Weil furthermore showed the converse also holds, i.e. if the functional equation holds for $L_\chi(s) = \sum_{n=1}^{\infty} \chi(n) a_n n^{-s}$ for "sufficiently many" characters χ, then the associated $f(z)$ belongs to $\mathcal{M}_k(\Gamma_0(N), \varepsilon)$ for some N and ε depending on the functional equation for the associated $\Lambda(s)$. Moreover, if $L(s)$ is absolutely convergent for $s = k/2 - \epsilon$ for some $\epsilon > 0$, then $f(z)$ is a cusp form (see [25], Thm. 4.3.15).

3.9 Modular forms with complex multiplication

For a normalized eigenform $f \in \mathcal{S}_k(\Gamma_1(N))$, let K_f be the field over $\mathbb{Q}$ generated by the coefficients in its q-series expansion. Using the property that the rational subspace of $\mathcal{S}_k(\Gamma_1(N))$ generates the entire space over $\mathbb{C}$ and it is stable under the action of operators, one can show that K_f is a number field. Furthermore, K_f contains the image of the Nebentypus of f (see [28]). (This follows from the fact that two eigenforms of possibly different weight and level and Nebentypus coincide everywhere (away from their levels) if the prime coefficients a_p in their q-series expansion coincide on a set of primes of density 1. In particular, they are of the same weight, all of the coefficients of q^n with n coprime to the levels are equal, and the images of the Nebentypus are equal for all integers coprime to both levels (see [8], 6.3).)

For any eigenform f, the structure of K_f depends on its Nebentypus. More precisely, K_f is either a *field with complex multiplication* or a totally real field, and it is real if and only if the Nebentypus ε factors through $\{\pm 1\} \subseteq \mathbb{C}^\times$ and

$$\varepsilon(p) a_p = a_p \quad \forall \text{ primes } p \nmid N,$$

where as usual a_p denotes the coefficient of q^p in the Fourier expansion of f (see [28]). Recall that a field with complex multiplication, also called a CM field, is an imaginary quadratic extension of a totally real field.

For a newform $f \in \mathcal{S}_k(\Gamma_1(N))$ of Nebentypus ε, we can twist by a Dirichlet character φ mod D as follows:

$$f \otimes \varphi = \sum_{n=1}^{\infty} \varphi(n) a_n q^n \in \mathcal{S}_k(\Gamma_0(ND^2), \varepsilon\varphi^2).$$

Moreover, $f \otimes \varphi$ is an eigenform as the action of the Hecke operator T_p for $p \nmid ND$ has eigenvalue $\varphi(p)a_p$.

We say that a form f has *complex multiplication* (or CM) by φ if $f \otimes \varphi = f$. Note that one must check that $\varphi(p)a_p = a_p$ (or equivalently, either $\varphi(p) = 1$ or $a_p = 0$) for a set of primes of density 1 in order to conclude that $f(z) = \sum_{n \geq 1} a_n q^n$ has CM by φ. Furthermore, this implies that $\varepsilon\varphi^2 = \varepsilon$, so φ must be a quadratic character.

Remark. Using $\Gamma_1(N)$, we can define the notion of a CM cusp form on all Γ, namely in the direct limit over all levels.

4 Galois representations

The theory of Galois representations stems from the study of the absolute Galois group of $\mathbb{Q}$ and number fields. Originating as a generalization to class field theory and the Kronecker-Weber theorem, such representations have proven to be useful in a variety of subjects. The material discussed here focuses on 2-dimensional ℓ-adic representations, which are closely related to the theory of modular forms (see Theorem 4). The main references for this material include [5], [6], [9], [36], and [42]. In particular, the relationship between CM cusp forms and certain ℓ-adic representations as described in [28] allow us to prove Lemma 2.

4.1 Basic theory and notation

Fix an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$ and consider the profinite group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ with its natural topology. Let K be a topological field. An n-dimensional representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is a homomorphism

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_n(K),$$

which is continuous with respect to the topology of K. When the field $K = \mathbb{C}$, such ρ are called *Artin representations*, and continuity is equivalent to requiring that the representation ρ factors through $\mathrm{Gal}(F/\mathbb{Q})$ where F is a finite and Galois extension over $\mathbb{Q}$. When K is a finite extension of $\mathbb{Q}_\ell$, for some prime ℓ, they are called the *ℓ-adic representations*. The case when $n = 1$ is described by class field theory (particularly the Kronecker-Weber theorem), and the 2-dimensional case is particularly connected to the theory of modular forms on $\Gamma_1(N)$.

We can extend the usual theory of ramification at primes to infinite extensions as follows. For any prime p, choose a place $\mathfrak{p}$ of $\overline{\mathbb{Q}}$ over p in order to fix a decomposition subgroup $D_\mathfrak{p}$ of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, which acts as a stabilizer. Let $I_\mathfrak{p}$ denote the inertia subgroup inside $D_\mathfrak{p}$, arising from the following exact sequence

$$1 \longrightarrow I_\mathfrak{p} \longrightarrow D_\mathfrak{p} \overset{\mathrm{mod}\,\mathfrak{p}}{\longrightarrow} \mathrm{Gal}(\overline{\mathbb{F}_p}/\mathbb{F}_p) \longrightarrow 1.$$

The automorphism $x \mapsto x^p$ (topologically) generates this Galois group of residue fields, and any element of $D_\mathfrak{p}$ in its preimage is a *Frobenius element* for the prime p (for all choices of $D_\mathfrak{p}$). We say that a representation of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is *unramified at p* if all Frobenius elements for p lie in the same conjugacy class in the image (and the conjugacy class is therefore well-defined). Equivalently, a map $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_n(K)$ for a field K must vanish on any (and therefore all) inertia subgroups $I_\mathfrak{p}$.

In general, we will implicitly assume that an ℓ-adic representation, along with being continuous, is unramified at all but finitely many primes. Unlike Artin representations, this is not always true, and furthermore ℓ-adic representations need not be semi-simple. However, any semi-simple ℓ-adic representation is completely determined by its trace (see [3], 8.12.1).

Example: The ℓ-adic cyclotomic character. Consider the unique map defined by

$$\chi_\ell : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathbb{Z}_\ell^\times \hookrightarrow \mathbb{Q}_\ell^\times, \qquad \sigma \mapsto \chi_\ell(\sigma) \quad \text{s.t. } \sigma(\zeta) = \zeta^{\chi(\sigma)} \quad \forall \zeta \in \overline{\mathbb{Q}}^\times[\ell^\infty],$$

i.e., ζ is a ℓ-power torsion element of $\overline{\mathbb{Q}}^\times$. In particular, it factors through $\mathrm{Gal}(\mathbb{Q}(\zeta_\ell^\infty)/\mathbb{Q}) = \bigcup_n \mathbb{Q}(\zeta_{\ell^n})$. Furthermore, χ_ℓ is unramified at all primes $p \neq \ell$, and for such p, $\chi_\ell(\mathrm{Frob}_p) = p$. Since it is a 1-dimensional ℓ-adic representation, we call it a character.

Note that the image of χ_ℓ lies in $\mathbb{Z}_\ell^\times$. In fact, it is true that any continuous ℓ-adic representation has image in $\mathrm{GL}_n(\mathcal{O}_K)$ after suitable conjugation (see [6], 3). Additionally, Artin

representations can be viewed as (finite) ℓ-adic representations by fixing an isomorphism of fields $\overline{\mathbb{Q}_\ell} \cong \mathbb{C}$.

After fixing an isomorphism $\overline{\mathbb{Q}}_l \xrightarrow{\sim} \mathbb{C}$, to any ℓ-adic representation ρ, we can attach an L-function $L(s,\rho)$ by taking the product of the L-factors $L_p(s) = \det(1-\rho(\mathrm{Frob}_p)p^{-s})^{-1}$ over all primes p:

$$L(s,\rho) = \prod_p \frac{1}{\det(1-\rho(\mathrm{Frob}_p)p^{-s})}.$$

Semi-simple ℓ-adic representations are completely determined by their L-functions.

4.2 ℓ-adic Representations in connection with cuspidal eigenforms

Let $f \in \mathcal{S}_k(\Gamma_1(N))$ be an eigenform with Nebentypus ε, and denote K_f as the number field generated by the q-series coefficients a_n of f in its q-series expansion (Recall that a_p for $n = p$ prime also satisfy $T_p f = a_p f$).

Theorem 4. *For any prime ℓ, there exists an ℓ-adic representation*

$$\rho_\ell : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(K_f \otimes_{\mathbb{Q}} \mathbb{Q}_\ell)$$

such that for any prime $p \nmid \ell N$:

1. *ρ_ℓ is unramified at p;*
2. *For all choices of Frob_p, $\rho_\ell(\mathrm{Frob}_p)$ has trace equal to a_p;*
3. *Furthermore, $\rho_\ell(\mathrm{Frob}_p)$ has determinant equal to $\varepsilon(p)p^{k-1}$.*

This was first proven for the "classical" case of $k = 2$, using the Jacobian variety of $X_1(N)$ (this follows from results of Eichler, Shimura, and Igusa [5]). Since $X_1(4)$ has genus 0, the Jacobian $J_1(4)$ has dimension 0, i.e. $J_1(4) = 0$, which corresponds to the fact that there are no weight 2 cusp forms on $\Gamma_1(4)$. Serre conjectured the statement for $k > 2$ and it was later proved by Deligne using étale cohomology for the space of forms of weight k on $\Gamma_1(N)$ (see [7]). The final positive case, $k = 1$ was presented by Deligne and Serre using results from the higher weight cases (see [8]).

Note that the $\mathbb{Q}_\ell$-algebra $K_f \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ is the product of the completions of K_f at primes lying over ℓ, hence we have the decomposition

$$\mathrm{GL}_2(K_f \otimes_{\mathbb{Q}} \mathbb{Q}_\ell) = \prod_{\lambda|\ell} \mathrm{GL}_2(K_\lambda).$$

Thus, we can define λ-adic representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ by composing an ℓ-adic representation ρ_ℓ with the projection $\mathrm{GL}_2(K_f \otimes_{\mathbb{Q}} \mathbb{Q}_\ell) \to \mathrm{GL}_2(K_\lambda)$ to get

$$\rho_\lambda : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(K_\lambda) \qquad \text{and} \qquad \rho_\ell = \bigoplus_{\lambda|\ell} \rho_\lambda.$$

We can then consider the λ-adic representation attached to a newform f, which is in fact a Galois representation over a field.

It is also true that the ℓ-adic representation attached to f whose existence is guaranteed by the previous theorem is unique. This follows from the fact that for each $\lambda \mid \ell$, ρ_λ is simple, thus ρ_ℓ is semi-simple with trace determined by the Fourier coefficients of f (see [28], 2).

4.3 ℓ-adic representations in connection with Eisenstein series

Although we have mostly discussed the action of operators on cusp forms, they are in fact well-defined operators on all of $\mathcal{M}_k(\Gamma_1(4))$ (this is due to the fact that the Petersson inner product mentioned in Section 3.5.1 can be extended to $\mathcal{M}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \longrightarrow \mathbb{C}$), see Section 3.5 or [10], 5.4). The actions on the entire space extend as follows:

$$\langle d \rangle f = \varepsilon(d) f \quad \text{for } \gcd(d, N) = 1 \qquad \text{and} \qquad T_p f = c_p \cdot f \quad \text{for } p \nmid N.$$

Here, ε is the Nebentypus as usual, but c_p are eigenvalues that are not necessarily equal to the coefficients a_p unless f is a normalized eigenform (Eisenstein or cuspidal). If f is an Eisenstein series, we can construct a representation ρ_ℓ on the number field generated by the image of ε and c_p. A classical result of Hecke (see [17], p. 690) tells us that there exist Dirichlet characters ε_1 and ε_2 with the property that the product of their conductors divides N such that

$$\varepsilon_1 \cdot \varepsilon_2 = \varepsilon \qquad \text{and} \qquad c_p = \varepsilon_1(p) + \varepsilon_2(p) p^{k-1},$$

for all $p \nmid N$. One can show that ε_1 and ε_2 have images contained in K_f, thus we can regard them as characters of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ over $K \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$. We can then attach to f the (reducible) ℓ-adic representation

$$\rho_\ell = \varepsilon_1 \oplus \varepsilon_2 \chi_\ell^{k-1}.$$

For primes $p \nmid \ell N$, ρ_ℓ is unramified as $\varepsilon_1, \varepsilon_2$, and χ_ℓ are, and the trace and determinant of $\rho_\ell(\mathrm{Frob}_p)$ are analogous to the cusp form case:

$$\mathrm{tr}(\rho_\ell(\mathrm{Frob}_p)) = \varepsilon_1(p) + \varepsilon_2(p) p^{k-1} = c_p \qquad \det(\rho_\ell(\mathrm{Frob}_p)) = \varepsilon(p) p^{k-1}.$$

With this construction, Theorem 4 holds for all eigenforms.

5 The space $S_k^{cm}(\Gamma_1(N))$

Definition 2. The *space of CM cusp forms* denoted $\mathcal{S}_k^{cm}(\Gamma_0(N), \varepsilon)$ is the subspace of $\mathcal{S}_k(\Gamma_0(N), \varepsilon)$ generated by cusp forms with complex multiplication. The corresponding subspace of CM cusp forms of $\mathcal{S}_k(\Gamma_1(N))$ is $\mathcal{S}_k^{cm}(\Gamma_1(N)) = \bigoplus_\varepsilon \mathcal{S}_k^{cm}(\Gamma_0(N), \varepsilon)$.

We will produce eigenforms $f_{K,\psi}(r \cdot z)$ below attached to Hecke characters ψ which also have complex multiplication. Later, Theorem 6 will imply that such forms are a basis for the space of CM cusp forms.

5.1 Hecke characters of imaginary quadratic fields

Let $K \subseteq \mathbb{C}$ be an imaginary quadratic field of discriminant $-d$ and ring of integers $\mathcal{O}_K$. Denote χ_K as the quadratic character of conductor d defined in terms of the Kronecker symbol:

$$\chi_K(p) = \left(\frac{-d}{p}\right) \quad \text{if } p \text{ is prime and } p \nmid 2d$$

Let $t \in \mathbb{N}$ and $\mathfrak{f}$ be a nonzero ideal of $\mathcal{O}_K$. An algebraic Hecke character of K of ∞-type t and of conductor $\mathfrak{f}$ is a homomorphism

$$\psi : \left\{ \begin{matrix} \text{fractional ideals of } K \\ \text{which are prime to } \mathfrak{f} \end{matrix} \right\} \longrightarrow \mathbb{C}^\times \quad \text{s.t.} \quad \psi((\alpha)) = \alpha^t \quad \text{if } \alpha \in K^\times \text{ and } \alpha \equiv 1 \bmod^\times \mathfrak{f},$$

where $\mathfrak{f}$ should be minimal in the following sense: if ψ can be defined modulo $\mathfrak{f}'$, then $\mathfrak{f} \mid \mathfrak{f}'$. Consider the homomorphism $\omega_\psi : \mathbb{Z} \to \mathbb{C}^\times$ such that

$$\omega_\psi : a \longmapsto \frac{\psi((a))}{a^t} \qquad \forall a \in \mathbb{Z} \text{ s.t. } a \text{ is coprime to } \mathfrak{f} \text{ (or } \mathrm{Nm}(\mathfrak{f})).$$

We call ω_ψ the character attached to ψ; it is a Dirichlet character mod $\mathrm{Nm}(\mathfrak{f})$.

5.2 Hecke characters, idelically

Hecke characters can also be viewed as continuous homomorphisms on the idèle group of any finite extension of $\mathbb{Q}$. More precisely, let K be a number field with ring of integers $\mathcal{O}_K$. The *adèles* of K are defined as

$$\mathbb{A}_K = \left(\prod_{\mathfrak{p} \subseteq \mathcal{O}_K} K_\mathfrak{p} \right)' \times (\mathbb{R} \otimes_\mathbb{Q} K),$$

where the first factor is a restricted product: for any element $x = (x_\mathfrak{p})_\mathfrak{p} x_\infty \in \mathbb{A}_K$, for almost all $\mathfrak{p}$, $x_\mathfrak{p} \in \mathcal{O}_{K,\mathfrak{p}}$. We can think of $\mathbb{R} \otimes K$ as the product of completions of the infinite places, i.e.,

$$K_\mathbb{R} = \mathbb{R} \otimes_\mathbb{Q} K = \prod_{v \mid \infty} K_v.$$

The *idèles* of K are then

$$\mathbb{A}_K^\times = \{x \in \mathbb{A}_K : \text{for almost all } \mathfrak{p}, |x|_\mathfrak{p} = 1 \text{ and for all } v, x_v \neq 0\}.$$

Automatically, $\mathbb{A}_K$ has the restricted product topology, i.e. it is induced by the product topology on the open subgroup $\prod \mathcal{O}_{K,\mathfrak{p}} \times K_\mathbb{R}$; however, the idèles have the topology induced from the map

$$\mathbb{A}_K^\times \hookrightarrow \mathbb{A}_K \times \mathbb{A}_K, \quad x \mapsto (x, x^{-1}).$$

Note that the image $\{(x, y) \in \mathbb{A}_K \times \mathbb{A}_K : xy = 1\}$ is closed, and the induced topology makes $\mathbb{A}_K^\times$ locally compact.

An *(idelic) algebraic Hecke character* of K is a continuous character $\psi : \mathbb{A}_K^\times \to \mathbb{C}^\times$ which is trivial on the diagonally embedded $K^\times$, and on the (connected component of the origin of the) Archimedean factor,

$$t : (K_\mathbb{R}^\times)^\circ \to \mathbb{C}^\times, \quad z \mapsto \prod_{\sigma : K \hookrightarrow \mathbb{C}} \sigma(z)^{t(\psi,\sigma)},$$

where each $t(\psi, \sigma) \in \mathbb{Z}$.

We have not yet defined the conductor of ψ. Given an integer $m \geq 0$ and a finite prime $\mathfrak{p} \subset \mathcal{O}_K$, define $U_{\mathfrak{p},m} := \{u \in \mathcal{O}_{K,\mathfrak{p}}^\times : v_\mathfrak{p}(1-u) \geq m\}$. For an infinite place $v \mid \infty$, define U_v as the connected component of $K_v^\times$ containing the origin. Then the subgroup attached to $\mathfrak{m}$, a *modulus* $\mathfrak{m}$ is

$$U_\mathfrak{m} = \prod_\mathfrak{p} U_{\mathfrak{p},m(\mathfrak{p})} \times \prod_{\substack{m(v)>0 \\ v}} U_v \quad \text{where} \quad \mathfrak{m} = \prod_\mathfrak{p} \mathfrak{p}^{m(\mathfrak{p})} \cdot \prod_v v^{m(v)},$$

where only a finite number of $m(\mathfrak{p})$ are nonzero. An algebraic character ψ has modulus $\mathfrak{m}$ if ψ is trivial on $U_\mathfrak{m}$. Note that every Hecke character has a modulus since the restriction of ψ to the finite idéles has open kernel and each $U_\mathfrak{m}$ is also open. The conductor of ψ is then the minimal modulus.

For K an imaginary quadratic field, there is only one pair of complex embeddings $(\sigma, \overline{\sigma})$ and no real embeddings, thus for an algebraic Hecke character ψ on K, $t \in \mathbb{Z}^2$. However, when studying Hecke characters in connection with the theory of classical modular forms, we are only interested in Hecke characters ψ such that $t(\overline{\sigma}) = 0$. Thus, the map t sends $z \mapsto z^{t(\sigma)}$, and $t(\sigma)$ then coincides with the notion of ∞-type in the classical description of Hecke characters.

5.3 Cusp forms attached to Hecke characters

For a Hecke character ψ of conductor $\mathfrak{f}$ with ∞-type t on an imaginary quadratic field K, define the q-series

$$f_{K,\psi}(z) = \sum_{\substack{\mathfrak{a} \text{ integral} \\ \text{coprime to } \mathfrak{f}}} \psi(\mathfrak{a}) \cdot q^{\mathrm{Nm}(\mathfrak{a})} \quad (q = e^{2\pi i z}, \mathrm{Im}(z) > 0) \tag{7}$$

Theorem 5 (Hecke [17], Shimura [34] & [35])**.** *The q-series $f_{K,\psi}(z)$ is a newform of weight $t+1$ and level $d \cdot \mathrm{Nm}(\mathfrak{f})$ in the space*

$$f_{K,\psi} \in \mathcal{S}_{t+1}(\Gamma_0(d \cdot \mathrm{Nm}(\mathfrak{f})), \chi_K \cdot \omega_\psi)) \subset \mathcal{S}_{t+1}(\Gamma_1(d \cdot \mathrm{Nm}(\mathfrak{f}))).$$

In addition, distinct cusp forms $f_{K,\psi}$ arise from distinct pairs (K, ψ).

Corollary 5.1. *For any positive integer r,*

$$f_{K,\psi}(r \cdot z) = \sum_{\mathfrak{a}} \psi(\mathfrak{a}) \cdot q^{r \cdot \mathrm{Nm}(\mathfrak{a})} \in \mathcal{S}_{t+1}(\Gamma_0(N), \varepsilon) \quad \Longleftrightarrow \quad r \cdot d \cdot \mathrm{Nm}(\mathfrak{f}) \mid N \text{ and } \chi_K \cdot \omega_\psi = \chi$$

if ψ has ∞-type t. The second equality takes place while viewing χ_K, ω_ψ and χ as characters on $\mathbb{Z}$. Thus, for all primes $p \nmid N$, $\chi_K(p) \cdot \omega_\psi(p) = \chi(p)$.

Note that the cusp forms $f_{K,\psi}(r \cdot z)$ have CM by χ_K. (By construction, the coefficient of q^p in the q-series expansion of $f_{K,\psi}(r \cdot z)$ is 0 if no ideal of k has norm equal to p. Since $\varphi(p) = -1$ exactly when this holds for p, $a_p = \varphi(p)a_p$, the result holds.) However, it is not at all obvious that these are all the cusp forms with complex multiplication in $\mathcal{S}_k(\Gamma_1(N))$. Using the theory of Galois representations, Ribet proves that a newform f has CM by an imaginary quadratic field K if and only if it arises from a Hecke character on K (see Theorem 6 below).

5.4 λ-adic representations in connection with CM cusp forms

If $f \in \mathcal{S}_k(\Gamma_0(N), \varepsilon)$ is a newform, there are special properties of the ℓ-adic representation that depend on whether f has complex multiplication. If K_f denotes the field of eigenvalues of f, let ℓ be a prime in $\mathbb{Q}$ with λ lying over ℓ in K_f. For all newforms f, recall from the end of Section 4.2 that ρ_λ is irreducible over K_λ. It is furthermore true that the image of ρ_λ is not abelian. (If $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ denotes complex conjugation, note that $\rho_\lambda(\sigma)$ has eigenvalues ± 1, which are distinct units in K_λ. Let S denote the subgroup of matrices in $\mathrm{GL}_2(K_\lambda)$ which commute with $\rho_\lambda(\sigma)$. Note that S is abelian and diagonalizable, but since ρ_λ is irreducible, $\mathrm{Im}(\rho_\lambda)$ is not contained inside S, so in particular there exists elements in the image which do not commute with $\rho_\lambda(\sigma)$.)

Additionally, the restriction of a semi-simple representation of a group to a subgroup with finite index is again semi-simple, thus for any open subgroup $H \subseteq \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, $\rho_\lambda|_H$ is semi-simple. Consider the composition of ρ_λ with the projection $\mathrm{GL}_2(K_\lambda) \twoheadrightarrow \mathrm{PGL}_2(K_\lambda)$, the quotient group of $\mathrm{GL}_2(K_\lambda)$ by the center $K_\lambda^\times$. Then, the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ inside $\mathrm{PGL}_2(K_\lambda)$ can be shown to have infinite image (see [28], 4.3). Using these observations along with the Cebotarev Density Theorem, one can deduce the following (see [28]):

Proposition 1 (Ribet)**.** *One of the following is true.*

1. *For each open subgroup H of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, $\rho_\lambda(H)$ is irreducible and non-abelian.*
2. *There exists an open subgroup H_2 of index 2 such that for each open subgroup H of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, $\rho_\lambda(H)$ is abelian if and only if $H \subseteq H_2$.*

If H_2 exists, then let K be the fixed field corresponding to $H_2 \subset \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. K is quadratic and unramified outside of ℓN, and f has complex multiplication by the character of K.

Conversely, if f has CM by a character φ and K is quadratic field associated to φ, then the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/K)$ under ρ_λ is abelian.

From Theorem 5 and Proposition 1, we can conclude that the space $\mathcal{S}_k^{cm}(\Gamma_1(N))$ is generated by the eigenforms $f_{K,\psi}(r \cdot z)$ with K, ψ and r satisfying the conditions in Corollary 5.1. The precise statement is as follows.

Theorem 6 (Ribet, see [28])**.** *If there exists an open subgroup $H \subset \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ of index 2 such that the image of H under ρ_λ is abelian, then for all primes λ' of K, $\rho_{\lambda'}(H)$ is abelian. Furthermore, the fixed field of H is an imaginary quadratic field which is unramified at all primes away from N, the level of f. Finally, f is obtained from an algebraic Hecke character ψ of K as described in Corollary 5.1.*

5.5 Proof of Lemma 2

Assume $N = 4$. First note that there are two Dirichlet characters on $(\mathbb{Z}/4\mathbb{Z})^\times$, which we will denote ε_+ and ε_- depending on where they send the only nontrivial element, -1 mod 4.

$$\varepsilon_+ : \pm 1 \longmapsto 1 \quad \text{and} \quad \varepsilon_- : \pm 1 \longmapsto \pm 1.$$

These can also be viewed as maps from $\mathbb{Z} \to \mathbb{C}$ by sending any even integer to 0. Furthermore, note that the decomposition of

$$\mathcal{S}_k(\Gamma_1(N)) = \bigoplus_{\varepsilon} \mathcal{S}_k(\Gamma_0(N), \varepsilon)$$

occurs over all characters modulo N such that $\varepsilon(-1) = -1$ if k is odd and $\varepsilon(-1) = 1$ if k is even due to (4). Thus,

$$\mathcal{S}_k(\Gamma_1(4)) = \begin{cases} \mathcal{S}_k(\Gamma_0(4), \varepsilon_+) & \text{if } k \text{ is even} \\ \mathcal{S}_k(\Gamma_0(4), \varepsilon_-) & \text{if } k \text{ is odd.} \end{cases}$$

Let ε denote the correct Dirichlet character such that $\mathcal{S}_k(\Gamma_1(4)) = \mathcal{S}_k(\Gamma_0(4), \varepsilon)$ depending on whether k is even or odd. We first want to understand the CM subspace $\mathcal{S}_k^{cm}(\Gamma_0(4), \varepsilon)$. By definition, the imaginary quadratic field K must have discriminant $-d$ such that $d \mid N$, thus for $N = 4$, $d = 4$, i.e. $K = \mathbb{Q}(i)$ (By Stickelberger's theorem, there is no field of discriminant -2 and $d = 1$ is associated to the field $\mathbb{Q}$, which is not quadratic.) Note that χ_K is a nontrivial Dirichlet character of conductor 4, thus $\chi_K = \varepsilon_-$.

Since the discriminant of K is -4, we are left with finding all Hecke characters ψ on K of conductor 1 such that $\psi = \varepsilon_+$ if the ∞-type t of ψ is even, and $\psi = \varepsilon_-$ if t is odd. By definition, if $\alpha \in K^\times$ then since $(\alpha) = (u\alpha)$ for all $u \in \mathcal{O}_K^\times$,

$$\psi((\alpha)) = \alpha^t = (u\alpha)^t = \psi((u\alpha)).$$

Since $\mathcal{O}_K^\times = \langle i \rangle$, we can conclude that $4 \mid t$, since if $u = \pm i$, $i^t = 1$ if and only if $t \equiv 0 \bmod 4$. In particular, there are no Hecke characters of odd ∞-type, hence $\mathcal{S}_k^{cm}(\Gamma_0(4), \varepsilon)$ is trivial for even k. In addition, $\mathcal{S}_k^{cm}(\Gamma_0(4), \varepsilon)$ can only be non-trivial when $k \equiv 1 \pmod 4$ and $k > 1$.

We now prove that there is a unique Hecke character with the properties described above for each ∞-type $t \in \mathbb{Z}$ such that $4 \mid t$. Let t be fixed such that $t \equiv 0 \pmod 4$. Then $\chi_K \cdot \omega_\psi = \varepsilon_-$ for any ψ, hence the attached character ω_ψ must be trivial, or abusing notation, $\omega_\psi = \varepsilon_+$. Thus, since $\mathcal{O}_K = \mathbb{Z}[i]$ is a PID, the algebraic Hecke characters on K are exactly those defined as

$$\psi_t : (\alpha) \longmapsto \alpha^t \quad \forall \alpha \in K^\times \qquad \text{where } t \in \mathbb{Z} \text{ and } 4 \mid t.$$

Thus, using Theorem 5, we conclude that the dimension of $S_k^{cm}(\Gamma_0(4), \varepsilon) \subseteq S_k(\Gamma_1(4))$ is 1 if $k \equiv 1 \bmod 4$ and $k \geq 5$ and is 0 otherwise. □

6 Construction of bases for $\mathcal{E}_k(\Gamma_1(4)) \oplus \mathcal{S}_k^{cm}(\Gamma_1(4))$

We describe how to construct "natural" generators for the space of elementary modular forms on $\Gamma_1(4)$. We use the fact that each form can be decomposed into its Eisenstein and cuspidal part, thus we treat the Eisenstein and CM cuspidal space separately.

6.1 Spaces of Eisenstein series

Suppose ε_1 and ε_2 are primitive Dirichlet characters modulo N_1 and N_2, and let $K_{\varepsilon_1,\varepsilon_2}$ be the number field containing their image. For positive k, denote the holomorphic function on $\mathbb{H}$

$$E_k^{\varepsilon_1,\varepsilon_2}(z) = a_0 + \sum_{m \geq 1} \left(\sum_{d \mid m} \varepsilon_1(d) \cdot \varepsilon_2(m/d) \cdot d^{k-1} \right) q^m \in K_{\varepsilon_1,\varepsilon_2}[[q]], \qquad q = e^{2\pi i z}$$

where d ranges over positive divisors, and a_0 is defined below. This q-series arises from (correctly) normalizing and rewriting the summation

$$G_k^{\varepsilon_1,\varepsilon_2}(z) = \sum_{\substack{\text{nonzero} \\ (m_1,m_2) \in \mathbb{Z}^2}} \frac{\varepsilon_1(m_1)\varepsilon_2(m_2)}{(m_1 z + m_2)^k}.$$

The value of a_0 is nonzero only when $\varepsilon_1 = 1_{N_1}$, and it is related to the value of the Dirichlet L-function at $L(1-k, \varepsilon_2)$. Explicitly, we can write a_0 in terms of *generalized Bernoulli numbers* attached to a character ε of conductor N. These are defined by satisfying the following identity of infinite series:

$$\sum_{j=1}^{N} \frac{\varepsilon(j) \cdot x \cdot e^{jx}}{e^{Nx} - 1} = \sum_{m=0}^{\infty} B_k^\varepsilon \cdot \frac{x^m}{m!}.$$

When ε_1 is the trivial character, $a_0 = -\dfrac{B_k^{\varepsilon_2}}{2k}$; otherwise, as noted above $a_0 = 0$ (see [41]). We then have the following description if Eisenstein eigenforms (see [25], Ch. 7 or [41], 5.3).

Theorem 7. *Let ε_1 and ε_2 be primitive of conductors N_1 and N_2 and let k and t be positive integers.*

1. *Suppose $k \neq 2$. Then the q-series $E_k^{\varepsilon_1,\varepsilon_2}(q^t)$ is a modular form of weight k on $\Gamma_1(N_1 N_2 t)$. Furthermore, it has Nebentypus $\varepsilon = \varepsilon_1 \cdot \varepsilon_2$.*

2. *Furthermore, when $k \neq 2$, the set of Eisenstein series $E_k^{\varepsilon_1,\varepsilon_2}$ such that $\varepsilon_1 \cdot \varepsilon_2 = \varepsilon$ and $N_1 N_2 t \mid N$ form a basis for the Eisenstein eigenspace $\mathcal{E}_k(\Gamma_0(N), \varepsilon)$. Furthermore, they are (normalized) eigenforms.*

3. *When $k = 2$, $E_2(q) = E_2^{\mathbf{1},\mathbf{1}}(q) - tE_2^{\mathbf{1},\mathbf{1}}(q^t)$ is a (normalized) eigenform of weight 2 in $\mathcal{M}_2(\Gamma_1(t))$ of trivial Nebentypus.*

The action of Hecke operators on these eigenforms can thus be written explicitly (see [41], 5.3). As usual, the eigenvalue of T_p is the coefficient of q^p:

$$T_p \; E_k^{\varepsilon_1,\varepsilon_2(s)} = \begin{cases} (\varepsilon_1(p) + \varepsilon_2(p)p^{k-1}) \cdot E_k^{\varepsilon_1,\varepsilon_2}(q) & \text{if } k \neq 2, \\ (1+p) \cdot E_2(q) & \text{if } k = 2 \text{ and } p \nmid t \\ 1 \cdot E_2(q) & \text{if } k = 2 \text{ and } p \mid t. \end{cases}$$

6.2 Eisenstein series via Galois representations

Assume that $k \geq 3$. We can produce the same basis for this space $\mathcal{E}_k(\Gamma_1(4))$ as above using an analogous approach to the construction of cusp forms involving Galois representations.

Case 1. First, assume the weight k is even. Lemma 1 implies that the dimension of the space $\mathcal{E}_k(\Gamma_1(4))$ is 3 when $k > 2$. Recall from Section 5.5 that there are two Dirichlet characters with conductor dividing 4, the trivial character $\varepsilon_+ = 1_4$ with image $\{1\}$, and $\varepsilon_- = \chi_4$ which sends $\pm 1 \bmod 4 \mapsto \pm 1$. From 4, we know that

$$\mathcal{E}_k(\Gamma_1(4)) = \mathcal{E}_k(\Gamma_0(4), 1_4)$$

since $\mathcal{E}_k(\Gamma_0(4), \chi_4)$ is trivial when k is even. Thus, we want to produce three linearly independent Eisenstein series of trivial Nebentypus.

Let ℓ be a prime. For $f \in \mathcal{E}_k(\Gamma_1(4))$, recall from Section 4.3 that the 2-dimensional reducible ℓ-adic representation will have the form $\rho_l = \varepsilon_1 \oplus \varepsilon_2 \chi_\ell^{k-1}$, where the Dirichlet characters $\varepsilon_1 \cdot \varepsilon_2 = 1_4$. Since f must have level 4, both characters are trivial (since there are no primitive characters of conductor 2). Thus, consider the ℓ-adic representation

$$\rho = 1 \oplus \chi_\ell^{k-1} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell).$$

The L-series attached to this representation is

$$\begin{aligned} L(s, 1 \oplus \chi_\ell^{k-1}) &= \prod_p \frac{1}{1 - p^{-s}} \cdot \prod_p \frac{1}{1 - p^{k-1}p^{-s}} \\ &= \sum_{m_1 \geq 1} \frac{1}{m_1^s} \cdot \sum_{m_2 \geq 1} \frac{m_2^{k-1}}{m_2^s} \\ &= \sum_{m_1, m_2 \geq 1} \left(\sum_{d \mid m_1 m_2} d^{k-1} \right) \frac{1}{m_1 m_2^s} \\ &= \sum_{m \geq 1} \left(\sum_{d \mid m} d^{k-1} \right) \frac{1}{m^s}. \end{aligned}$$

Applying an inverse Mellin Transform gives us the q-series

$$E(z) = a_0 + \sum_{m \geq 1} \left(\sum_{d|m} d^{k-1} \right) q^m. \tag{8}$$

In order to calculate a_0, we consider the functional equation of $L(s, \rho)$. Since the level $N = 4$, let

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

where $\Gamma(s)$ denotes the usual gamma function. If $\Lambda(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$, then $\Lambda(s, \rho)$ satisfies the following functional equation:

$$\Lambda(s)\Lambda(s-k+1) = \frac{2\pi^{k/2}}{\mu(s)}\Lambda(s, \rho), \quad \text{where } \mu(s) = \left(\frac{s-k+1}{2}\right)\left(\frac{s-k+1}{2}+1\right)...\left(\frac{s-1}{2}\right),$$

(see [25], 4.7). Thus, $\Lambda(s, \rho)$ is holomorphic for all s except $s = 0$ and k. At $s = 0$, one can calculate that it has residue $-a_0 = -\zeta(1-k)/2$, in terms of the Riemann zeta function. Thus,

$$a_0 = -\frac{B_k}{2k}$$

where $B_k = B_k^{\mathbf{1}}$ is the kth Bernoulli number associated to the trivial character as defined in Section 6.1.

From Weil's converse of Theorem 3, $E(z)$ lies inside $\mathcal{E}_k(\mathrm{SL}_2(\mathbb{Z})) \subseteq \mathcal{E}_k(\Gamma_1(4))$ since the conductor of the trivial character is 1 (see Section 3.8, [25], §4.7, and [23], Ch. 9). However, this implies that the q-series $E(q^2)$ and $E(q^4)$ are also in $\mathcal{E}_k(\Gamma_1(4))$, and furthermore, they are clearly linearly independent. These are all eigenforms, and since the dimension of $\mathcal{E}_k(\Gamma_1(4))$ is three, $\mathcal{B} = \{E(q), E(q^2), E(q^4)\}$ is a basis.

Case 2. Now assume k is odd. Here, Lemma 1 implies that the dimension of the space $\mathcal{E}_k(\Gamma_1(4))$ is 2 for $k > 1$. Furthermore, $\mathcal{E}_k(\Gamma_1(4)) = \mathcal{E}_k(\Gamma_0(4), \chi_4)$ since now $1_4(-1) = 1 \neq (-1)^k$, making $\mathcal{E}_k(\Gamma_0(4), 1_4)$ trivial by (4). Thus, we want to produce 2 linearly independent Eisenstein series with Nebentypus χ_4.

As above, let ℓ be a prime. An Eisenstein series f with the above properties will have an ℓ-adic representation $\varepsilon_1 \oplus \varepsilon_2 \chi_\ell^{k-1}$ where $\varepsilon_1 \cdot \varepsilon_2 = \chi_4$, thus one of the two characters must be χ_4 while the other must be trivial. This gives two possible 2-dimensional reducible Galois representations

$$\begin{aligned} \rho_1 := \chi_4 \oplus \chi_\ell^{k-1} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) &\longrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell), \\ \rho_2 := 1 \oplus \chi_4 \chi_\ell^{k-1} : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) &\longrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell). \end{aligned}$$

The L-series attached to these representations are

$$\begin{aligned} L(s, \rho_1) &= \prod_p \frac{1}{1 - \chi_4(p)p^{-s}} \cdot \prod_p \frac{1}{1 - p^{k-1}p^{-s}} \\ &= \sum_{m_1 \geq 1} \frac{\chi_4(m_1)}{m_1^s} \cdot \sum_{m_2 \geq 1} \frac{m_2^{k-1}}{m_2^s} \\ &= \sum_{m_1, m_2 \geq 1} \left(\sum_{d | m_1 m_2} \chi_4\left(\frac{m_1 m_2}{d}\right) d^{k-1} \right) \frac{1}{m_1 m_2^s} \\ &= \sum_{m \geq 1} \left(\sum_{d|m} \chi_4\left(\frac{m}{d}\right) d^{k-1} \right) \frac{1}{m^s}, \end{aligned}$$

$$\begin{aligned} L(s,\rho_2) &= \prod_p \frac{1}{1-p^{-s}} \cdot \prod_p \frac{1}{1-\chi_4(p)p^{k-1}p^{-s}} \\ &= \sum_{m_1\geq 1} \frac{1}{m_1^s} \cdot \sum_{m_2\geq 1} \frac{\chi_4(m_2)m_2^{k-1}}{m_2^s} \\ &= \sum_{m_1,m_2\geq 1} \left(\sum_{d|m_1m_2} \chi_4(d)d^{k-1} \right) \frac{1}{m_1m_2^s} \\ &= \sum_{m\geq 1} \left(\sum_{d|m} \chi_4(d)d^{k-1} \right) \frac{1}{m^s}. \end{aligned}$$

Applying an inverse Mellin Transform to the above L-functions produces two q-series

$$E_1(q) = a_{0,1} + \sum_{m\geq 1} \left(\sum_{d|m} \chi_4\left(\frac{m}{d}\right) d^{k-1} \right) q^m, \tag{9}$$

$$E_2(q) = a_{0,2} + \sum_{m\geq 1} \left(\sum_{d|m} \chi_4(d)d^{k-1} \right) q^m. \tag{10}$$

As in the previous case, we calculate $a_{0,1}$ and $a_{0,2}$ by considering the functional equations for $L(s,\rho_1)$ and $L(s,\rho_2)$. As before, let $\Lambda(s,\rho_i) = \pi^{-s}\Gamma(s)L(s,\rho_i)$ for $i \in \{1,2\}$. If $\Lambda(s,\chi_4) = \left(\frac{\pi}{4}\right)^{-s/2}\Gamma(\frac{s+1}{2})L(s,\chi_4)$ and again $\Lambda(s) = \pi^{-s/2}\Gamma(\frac{s}{2})\zeta(s)$, then $\Lambda(s,\rho_i)$ satisfy the following functional equations:

$$\Lambda(s,\chi_4)\Lambda(s-k+1) = \frac{2\pi^{k/2}}{\mu_1(s)}\Lambda(s,\rho_1), \quad \text{and} \quad \Lambda(s)\Lambda(s-k+1,\chi_4) = \frac{2^{2-k}\pi^{k/2}}{\mu_2(s)}\Lambda(s,\rho_2),$$

where $\mu_1(s) = \left(\frac{s-k+1}{2}\right)\left(\frac{s-k+1}{2}+1\right)...\left(\frac{s}{2}-1\right)$ and $\mu_2(s) = \left(\frac{s-k}{2}+1\right)\left(\frac{s-k}{2}+2\right)...\left(\frac{s-1}{2}\right)$ (see [25], 4.7). Both $\Lambda(s,\rho_i)$ are holomorphic for all s except $s = 0$ and k. At $s = 0$, one can calculate that $\Lambda(s,\rho_1)$ has residue $-a_{0,1} = -L(1-k,\chi_4)/2$, in terms of the L-series associated to the Dirichlet character χ_4, and $\Lambda(s,\rho_2) = -\zeta(1-k)/2 = 0$ has residue $-a_{0,2} = 0$. Thus,

$$a_{0,1} = -\frac{B_k^{\chi_4}}{2k}, \quad \text{and} \quad a_{0,2} = 0.$$

From Theorem 3 and its converse, both $E_1(q)$ and $E_2(q)$ are distinct eigenforms on $\Gamma_1(4)$ of weight k, hence the two series are linearly independent (see Section 3.8, [25], §4.7, and [23], Ch. 9). By construction, $E_1(q), E_2(q) \in \mathcal{E}_k(\Gamma_1(4))$, thus we can take $\mathcal{B} = \{E_1(q), E_2(q)\}$ as a basis for the Eisenstein space.

Remarks. The cases for $k = 1$ and $k = 2$ have not been discussed here. Those are treated in Section 8 while finding elementary formulas for θ_2 and θ_4.

One can easily see that the Eisenstein series produced here coincide exactly with those in the previous section.

6.3 CM cusp forms via L-functions of Hecke characters

Although Theorem 5 gives the q-expansion of the CM cusp form on $\Gamma_1(4)$, we give an alternative construction which demonstrates the relationship to (potentially abelian) Galois representations (see [13]).

By Lemma 2, $\mathcal{S}_k^{cm}(\Gamma_1(4))$ is nontrivial if and only if $k \equiv 1 \pmod 4$ and $k > 1$. Restrict ourselves to such k, and note that the dimension of the CM subspace is equal to 1 (we showed this by proving there exists exactly one Hecke character satisfying the requirements of Corollary 5.1). Let ψ denote the algebraic Hecke character unramified away from 2 on $K = \mathbb{Q}(i)$ with ∞-type $k-1$. Viewing it idelically, ψ acts on $\mathbb{C}^\times$ by sending $z \mapsto z^{-k+1}$, and on primes $\pi \in \mathcal{O}_K$, ψ sends $\pi \mapsto \pi^{k-1}$. The L-series attached to this character is

$$\begin{aligned} L(s,\psi) &= \left(1 + (1+i)^{k-1}2^{-s}\right)^{-1} \prod_{p \equiv 3 \bmod 4} \frac{1}{1-p^{k-1}p^{-2s}} \prod_{p \equiv 1 \bmod 4} \frac{1}{(1-\pi^{k-1}p^{-s})\left(1-\overline{\pi}^{k-1}p^{-s}\right)} \\ &= \frac{1}{4}\sum_{m \geq 1} \left(\sum_{\substack{d \in \mathbb{Z}[i] \\ \mathrm{Nm}(d)=m}} d^{k-1} \right) \frac{1}{m^s}. \end{aligned}$$

Applying an inverse Mellin transform as described in Section 3.8 gives the normalized q-series

$$\begin{aligned} C(q) &= \frac{1}{4}\sum_{m \geq 1} \left(\sum_{\substack{d \in \mathbb{Z}[i] \\ \mathrm{Nm}(d)=m}} d^{k-1} \right) q^m \qquad (11) \\ &= q + (-4)^{\frac{k-1}{4}} q^2 + 2^{k-1}q^4 + \ldots \end{aligned}$$

By Theorem 3, this is a cusp form on $\Gamma_1(4)$ of weight k, and by construction, it has complex multiplication by the Dirichlet character χ_4 attached to $\mathbb{Q}(i)$. Thus $C(q) \in \mathcal{S}_k^{cm}(\Gamma_1(4))$ and in fact, generates the space.

7 Proof of Theorem 1

We want to prove that θ_n is elementary, i.e. $\theta_n \in \mathcal{E}_{n/2}(\Gamma_1(4)) \oplus \mathcal{S}^{cm}_{n/2}(\Gamma_1(4))$ if and only if $n = 2, 4, 6, 8$, and 10. Lemma 1 implies that for $n = 2, 4, 6, 8$, θ_n is a linear combination of Eisenstein series since there are no cusp forms on $\Gamma_1(4)$ of weight $k \leq 4$. Furthermore, Lemma 2 implies that the "first" CM cusp form on $\Gamma_1(4)$ is of weight 5, and in fact, combining the two lemmas implies that $\mathcal{S}_5(\Gamma_1(4)) = \mathcal{S}^{cm}_5(\Gamma_1(4))$, thus θ_{10} must be elementary as well. It remains to show that for even $n > 10$, θ_n is not elementary, i.e. one must calculate the coefficients of a non-CM cusp form in order to calculate the representation of integers by sums of n squares.

First assume that $n \equiv 0 \pmod 4$, i.e. θ_n as a modular form has even weight $k = \frac{n}{2}$. By Lemma 2, there are no CM cusp forms, so consider the previously constructed basis $\mathcal{B}$ of Eisenstein series from (8) for $\mathcal{E}_k(\Gamma_1(4))$

$$\begin{aligned} E(q) &= a_0 + \sum_{m\geq 1} \left(\sum_{d|m} d^{n/2-1} \right) q^m \\ E(q^2) &= a_0 + \sum_{m\geq 1} \left(\sum_{d|\frac{m}{2}} d^{n/2-1} \right) q^m \\ E(q^4) &= a_0 + \sum_{m\geq 1} \left(\sum_{d|\frac{m}{4}} d^{n/2-1} \right) q^m \end{aligned}$$

If $\theta_n \in \mathcal{E}_k(\Gamma_1(4))$, then it can be written as a linear combination of the above series. Consider the determinant of the matrix of coefficients of the 4 q-series:

$$\det(M) = \begin{vmatrix} 2n & 4\binom{n}{2} & 8\binom{n}{3} & 16\binom{n}{4} + 2n \\ 1 & 1 + 2^{n/2-1} & 1 + 3^{n/2-1} & 1 + 2^{n/2-1} + 4^{n/2-1} \\ 0 & 1 & 0 & 1 + 2^{n/2-1} \\ 0 & 0 & 0 & 1 \end{vmatrix} \begin{matrix} (\textit{coefficients of } \theta_n) \\ (\textit{coefficients of } E(q)) \\ (\textit{coefficients of } E(q^2)) \\ (\textit{coefficients of } E(q^4)) \end{matrix}$$

The determinant of M is zero if and only if there is a linear dependence amongst the coefficients, i.e. if the coefficients of q, q^2, q^3, and q^4 of θ_n can be written in terms of coefficients of the forms in $\mathcal{B}$. Solving for the determinant gives

$$\det(M) = -\left(2n \cdot \left(1 + 3^{\frac{n}{2}-1} \right) - 8\binom{n}{3} \right) = -2n - 2n \cdot 3^{\frac{n}{2}-1} + 8\frac{n(n-1)(n-2)}{6}$$

Note that the negative exponential term takes over the growth of the function, and for $n > 8$, $\det(M)$ as a function on n is monotonically decreasing. The determinant is 0 when $n = 4, 8$, thus $\det(M)$ is nonzero for all $n > 8$, i.e. $\theta_n \notin \mathcal{E}_{n/2}(\Gamma_1(4))$ for $n \equiv 0 \pmod 4$. By Lemma 1, this implies that θ_n is not elementary for $n > 8$, $n \equiv 0 \pmod 4$.

Now assume $n \equiv 2 \pmod 4$. First note that when $n \equiv 6 \pmod 8$, θ_n has odd weight, but there is no CM subspace of $\mathcal{M}_{n/2}(\Gamma_1(4))$ since $\frac{n}{2} \equiv 3 \bmod 4$. Thus, we first check that $\theta_n \notin \mathcal{E}_{n/2}(\Gamma_1(4))$ for all $n \equiv 2 \bmod 4$ and $n > 6$, which will reduce the problem to whether there is a contribution by a CM cusp form when $n \equiv 2 \bmod 8$.

When $k = n/2$ is odd, $\mathcal{E}_{n/2}(\Gamma_1(4))$ has a basis of eigenforms from (9) and (10)

$$E_1(q) = a_{0,1} + \sum_{m\geq 1} \left(\sum_{d|m} \chi_4\left(\frac{m}{d}\right) d^{n/2-1} \right) q^m, \qquad E_2(q) = a_{0,2} + \sum_{m\geq 1} \left(\sum_{d|m} \chi_4(d) d^{n/2-1} \right) q^m$$

To show $\theta_n \notin \mathcal{E}_{n/2}(\Gamma_1(4))$, we check if the following matrix of coefficients for q, q^2, and q^3 has non-zero determinant:

$$\det(N) = \begin{vmatrix} 2n & 4\binom{n}{2} & 8\binom{n}{3} \\ 1 & 2^{n/2-1} & -1+3^{n/2-1} \\ 1 & 1 & 1-3^{n/2-1} \end{vmatrix} \begin{matrix} (\textit{coefficients of } \theta_n) \\ (\textit{coefficients of } E_1(q)) \\ (\textit{coefficients of } E_2(q)) \end{matrix}$$

Solving for the determinant,

$$\begin{aligned} \det(N) &= 2n\left(2^{n/2-1}+\right)\left(1-3^{n/2-1}\right) - 8\binom{n}{2}\left(1-3^{n/2-1}\right) + 8\binom{n}{3}\left(-2^{n/2-1}\right) \\ &= \left(-4n^2+6n\right) + 2^{n/2-1}\cdot\left(-\frac{4}{3}n^3+4n^2-\frac{2}{3}n\right) + 3^{n/2-1}\cdot\left(4n^2-6n\right) + 6^{n/2-1}\cdot(-2n) \end{aligned}$$

Note that for $n > 6$, the positive portion of $\det(N)$ is $6n + 4n^2 \cdot 2^{n/2-1} + 4n^2 \cdot 3^{n/2-1}$. However, viewing this as a function on n, the growth of $-2n \cdot 6^{n/2-1}$ is much faster, hence $\det(N)$ is monotonically decreasing. For $n = 6$, $\det(N) = 0$, thus since $n = 10$ gives negative determinant, monotonicity implies that $\det(N) < 0$ for all $n > 6$. We conclude that for $n \equiv 2 \bmod 4$ and $n > 6$, θ_n is not a linear combination of Eisenstein series. By Lemma 2, this implies that for all positive $n \equiv 6 \bmod 8$ except $n = 6$, θ_n is not elementary.

Thus, θ_n can be elementary only when $n \equiv 2 \bmod 8$. Since the weight $\frac{n}{2} \equiv 1 \bmod 4$, the space $\mathcal{S}^{cm}_{n/2}(\Gamma_1(4))$ is nontrivial. It is generated by the newform attached to the Hecke character of $\mathbb{Q}(i)$ of ∞-type $\frac{n}{2}-1$ (see (11))

$$\begin{aligned} C(q) &= \frac{1}{4}\sum_{m\geq 1}\left(\sum_{\substack{d\in\mathbb{Z}[i] \\ \mathrm{Nm}(d)=m}} d^{\frac{n}{2}-1}\right)q^m \\ &= q + (-4)^{\frac{n-2}{8}}q^2 + 2^{\frac{n-2}{2}}q^4 + \ldots \end{aligned}$$

Thus, θ_n is not elementary if the determinant of the matrix of coefficients for q, q^2, q^3 and q^4,

$$\det(N') = \begin{vmatrix} 2n & 4\binom{n}{2} & 8\binom{n}{3} & 16\binom{n}{4}+2n \\ 1 & 2^{n/2-1} & -1+3^{n/2-1} & 4^{n/2-1} \\ 1 & 1 & 1-3^{n/2-1} & 1 \\ 1 & (-4)^{\frac{n-2}{8}} & 0 & 2^{\frac{n-2}{2}} \end{vmatrix} \begin{matrix} (\textit{coefficients of } \theta_n) \\ (\textit{coefficients of } E_1(q)) \\ (\textit{coefficients of } E_2(q)) \\ (\textit{coefficients of } C(q)) \end{matrix}$$

has non-zero determinant.

A straightforward calculation demonstrates the fact that $\det(N')$ as a function on n is monotonically decreasing with growth on the order of $\mathcal{O}(12^{n/2-1})$. For $n = 10$, $\det(N') = 0$, but for $n = 18$, $\det(N') = -439,038,812,160$. Thus, we can conclude that $\det(N')$ is nonzero for all $n > 10$ such that $n \equiv 2 \bmod 8$. This proves the theorem. □

7.1 Another proof.

We provide another proof which makes use of the constants of the Eisenstein series and is motivated by the following lemma.

Lemma 3. *For any $f \in \mathcal{S}^{cm}_k(\Gamma_1(4))$ with $k > 1$, $T_p(f) = 0$ for any prime $p \equiv 3 \pmod 4$.*

Proof. Recall that Theorem 6 and Corollary 5.1 tells us that when $k \equiv 1 \pmod 4$, $\mathcal{S}_k^{cm}(\Gamma_1(4))$ is generated over $\mathbb{C}$ by forms $f_{\mathbb{Q}(i),\psi}(z)$ defined by (7). In order for $f_{\mathbb{Q}(i),\psi}(z)$ to have CM by the quadratic character $\chi_{\mathbb{Q}(i)}$, note that $a_p = 0$ for all inert primes p, thus for any prime $p \equiv 3 \pmod 4$, $T_p(f)$ must necessarily vanish. □

Thus, if we show that the cuspidal part of θ_n which vanishes on all three cusps of $\Gamma_1(4)$ has nonzero coefficient of q^p such that $p \equiv 3 \pmod 4$, then θ_n is not elementary. Since $\theta_n \in \mathcal{M}_{n/2}(\Gamma_1(4)) = \mathcal{E}_{n/2}(\Gamma_1(4)) \oplus \mathcal{S}_{n/2}(\Gamma_1(4))$, write θ_n as a linear combination of an Eisenstein series and cusp form, i.e. $\theta_n(z) = f_{n/2}(z) + s_{n/2}(z)$.

Lemma 4. *If $s_{n/2}(z)$ denotes the cuspidal part of $\theta_n(z)$, then $a_3(s_{n/2}) \neq 0$ if $n > 10$.*

Proof. The fundamental facts we use are formulas for the odd coefficients of Fourier expansions for the Eisenstein part

$$f_{n/2}(z) = 1 + \sum_{m=1}^{\infty} b_m q^m \quad \in \mathcal{E}_{n/2}(\Gamma_1(4))$$

due to Siegel (see [39]) and Shimura (see [38] and [37]). Let E_k denote the kth Euler number and B_j denote the jth Bernoulli number, which are defined by the following identities

$$\frac{2}{e^t + e^{-t}} = \sum_{k=0}^{\infty} E_k \frac{t^k}{k!}, \quad \text{and} \quad \frac{t}{e^t - 1} = \sum_{j=0}^{\infty} B_j \frac{t^j}{j!}.$$

The magnitudes of these numbers are related to values of L-series associated to the primitive Dirichlet characters of conductor 1 and 4:

$$B_j = \frac{2 \cdot j!}{(2\pi i)^j} \cdot \zeta(j) \quad \text{if } j > 0 \text{ even, and}$$

$$E_k = \frac{2^{2k+3} \cdot k!}{(2\pi i)^{k+1}} \cdot L(k+1, \chi_4) \quad \text{if } k > 0 \text{ even,}$$

where we let χ_4 denote the Dirichlet character of conductor 4 as usual. For odd m, the coefficients of the Eisenstein series $f_{n/2}$ are

$$b_m = \begin{cases} \dfrac{4}{|E_{n/2-1}|} \cdot \left(\chi_4(m) \cdot 2^{n/2-1} + \chi_4\left(\dfrac{n}{2}\right)\right) \cdot \displaystyle\sum_{d|m} \chi_4(d) d^{n/2-1} & \text{if } n \equiv 2 \pmod 4 \text{ and } n > 2 \\ \dfrac{n}{(2^{n/2}-1)|B_{n/2}|} \cdot \displaystyle\sum_{d|m} d^{n/2-1} & \text{if } n \equiv 0 \pmod 4 \end{cases}$$

From (2), recall that $r_n(3) = 8\binom{n}{3} = \frac{4n(n-1)(n-2)}{3}$. We can calculate the third coefficient of the q-series expansion of $s_{n/2}(z) = \theta_n(z) - f_{n/2}(z)$ using this equality:

$$s_{n/2}(z) = \sum_{m=1}^{\infty} a_m q^m = \sum_{m=0}^{\infty} r_n(m) q^m - \sum_{m=0}^{\infty} b_m q^m \quad \in \mathcal{S}_{n/2}(\Gamma_1(4).$$

Thus, taking the difference of b_3 and $r_n(3)$ gives the coefficients of $s_{n/2}(z)$:

$$a_3 = \begin{cases} \dfrac{4n(n-1)(n-2)}{3} + \left(\dfrac{4}{|E_{n/2-1}|}\right) \cdot \left(2^{n/2-1} - \chi_4\left(\dfrac{n}{2}\right)\right) \cdot \left(1 - 3^{n/2-1}\right) & \text{if } n \equiv 2 \pmod 4 \\ \dfrac{4n(n-1)(n-2)}{3} + \left(\dfrac{n}{|B_{n/2}|}\right) \cdot \left(\dfrac{1 + 3^{n/2-1}}{1 - 2^{n/2-1}}\right) & \text{if } n \equiv 0 \pmod 4 \end{cases}$$

In particular, we notice that if a_3 is viewed as a function on n, it can have at most six zeroes. (This is due to the fact that the values of $|B_j|$ and $|E_k|$ for even j, k have bounds (see [20]),

$$4\sqrt{\frac{\pi j}{2}}\left(\frac{j}{2\pi e}\right)^j < |B_j| < 5\sqrt{\frac{\pi j}{2}}\left(\frac{j}{2\pi e}\right)^j, \quad \text{and} \quad 8\sqrt{\frac{k}{2\pi}}\left(\frac{2k}{\pi e}\right)^k < |E_k|.$$

Some asymptotic calculations then demonstrate that a_3 is monotonically increasing for $n > 18$ and 20.)

Four of these zeroes occur when $n = 4, 6, 8, 10$, and calculating the values of a_3 for $n = 12, 14, 16, 18$ as seen in the table below along with the monotonicity show that no other zeroes occur.

$\mathbf{n}$	$\mathbf{b_3}$	$\mathbf{r_n(3)}$	$\mathbf{a_3 = r_n(3) - b_3}$
4	32	32	0
6	160	160	0
8	448	448	0
10	960	960	0
12	1952	1760	-192
14	189280/61	2912	-11648/61
16	70016/17	4480	6144/17
18	1338240/277	6528	470016/277
20	157472/31	9128	125248/31

Thus, the cuspidal part of θ_n has a nonzero coefficient for q^3 in the Fourier expansion. From Section 3.4, the first q-series coefficient of $T_3(s_{n/2})$ is a_3 which is nonzero for $n > 10$. □

We have shown that the cuspidal part of θ_n does not lie in the CM subspace of $\mathcal{S}_{n/2}(\Gamma_1(4))$ for $n > 10$, hence θ_n cannot be elementary for these cases. Theorem 1 then follows from the above two lemmas once we show that θ_n is elementary for $n = 2, 4, 6, 8, 10$. Note that this proof does not rely on the dimension arguments of Lemmas 1 and 2.

8 Elementary formulas for small n

We produce the classical nice formulas for $n = 2, 4, 6, 8$, and 10 in the manner analogous to the constructions in the proof of Theorem 1. These are originally due to Jacobi ($n = 2, 4, 6, 8$) and Liouville ($n = 10$).

8.1 Sum of 2 squares

By Lemma 1, we know that the dimension over $\mathbb{C}$ of weight 1 modular forms on $\Gamma_1(4)$ is 1, and the basis element is Eisenstein. We produce this generator by considering the 2-dimensional reducible Galois representation unramified away from 2. To have the correct trace and determinant, as described by Section 4.3, consider the representation

$$\rho = 1 \oplus \chi_4 : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Q}_l)$$

where χ_4 as usual can be viewed as the Dirichlet character $(\mathbb{Z}/4\mathbb{Z})^\times \to \mathbb{C}^\times$ extended to a map $(\mathbb{Z}/4\mathbb{Z}) \to \mathbb{C}$ by mapping $0, 2 \mapsto 0$. Note that this is the same representation as in the construction of basis elements during the proof of Theorem 1 since here, $\frac{n}{2} - 1 = 0$. The L-series attached to this Galois representation is

$$\begin{aligned} L(s,\rho) &= L(s,1)\cdot L(s,\chi_4) = \prod_p \frac{1}{(1-p^{-s})} \prod_p \frac{1}{(1-\chi_4(p)p^{-s})} \\ &= \sum_{m_1\geq 1} \frac{1}{m_1^s} \sum_{m_2 \geq 1} \frac{\chi_4(m_2)}{m_2^s} = \sum_{m\geq 1} \left(\sum_{d|m} \chi_4(d) \right) \frac{1}{m^s}. \end{aligned}$$

Applying the inverse Mellin transform gives the q-series $G(q) = a_{0,G} + \sum_{m\geq 1} \left(\sum_{d|m} \chi_4(d) \right) q^m$, analogous to the construction in the proof of the theorem. Theorem 3 and Lemma 1 imply that $\mathcal{M}_1(\Gamma_1(4)) = \mathbb{C} \cdot G(q)$, thus we can write $\theta_2(q) = a \cdot G(q)$ where $a \in \mathbb{C}$. Calculating the first few $r_2(m)$ thengives us

$$\theta_2(q) = 1 + \sum_{m=1}^{\infty} 4 \left(\sum_{d|m} \chi_4(d) \right) q^m. \tag{12}$$

8.2 Sum of 4 squares

When $n = 4$, the dimension of $\mathcal{M}_2(\Gamma_1(4)) = \mathcal{E}_2(\Gamma_1(4))$ is 2 by Lemma 1. We produce a 2-dimension reducible Galois representation of conductor 1 where the two characters have ∞-type 0 and 1 respectively and is unramified at 2. Analogous to the situation for $n \equiv 0 \bmod 4$ but $n > 8$, consider the system of ℓ-adic representations $1 \oplus \chi_\ell : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\mathbb{Q}_\ell)$ where χ_ℓ as usual denotes the ℓ-adic cyclotomic character. Note however that the L-function attached to χ_ℓ originally has a contribution from the prime 2. Under an inverse Mellin tranform, the Galois representation has conductor 1, hence $L(s, 1 \oplus \chi_\ell)$ would give a modular form of weight 2 on $\mathrm{SL}_2(\mathbb{Z})$, which does not exist. A priori, we know there exist modular forms of weight 2 with level 2, hence, we "correct" the L-function to have no contribution from the prime 2, which increases the conductor to 2. (This construction is equivalent to the special case of Theorem 7

for Eisenstein series of weight 2. Thus, the L-series we attach to this representation is

$$L^*(s, 1 \oplus \chi_\ell) \quad = \quad \prod_p \frac{1}{1-p^{-s}} \cdot \prod_{p>2} \frac{1}{1-p^{1-s}} = \sum_{m_1 \geq 1} \frac{1}{m_1^s} \cdot \sum_{\substack{m_2 \text{ odd} \\ m_2 \geq 1}} \frac{m_2}{m_2^s} = \sum_{m \geq 1} \left(\sum_{\substack{d \text{ odd} \\ d|m}} d \right) \frac{1}{m^s}.$$

Mellin transform results in the following q-series and its "square"

$$G'(q) = a_{0,G'} + \sum_{m \geq 1} \left(\sum_{\substack{d \text{ odd} \\ d|m}} d \right) q^m, \qquad G'(q^2) = a_{0,G'} + \sum_{m \geq 1} \left(\sum_{\substack{d \text{ odd} \\ d|m}} d \right) q^{2m}.$$

Note that $G'(q) \in \mathcal{M}_2(\Gamma_1(2)) \subseteq \mathcal{M}_2(\Gamma_1(4))$ thus, $G'(q^2) \in \mathcal{M}_2(\Gamma_1(4))$ and $\{G'(q), G'(q^2)\}$ forms a basis for this space. We write $\theta_4(q) = a \cdot G'(q) + b \cdot G'(q^2)$ and solve for the complex constants in order to produce a formula consisting of a linear combination of Eisenstein series. Calculating the first few $r_4(m)$ and the first few coefficients of the two Eisenstein series, we conclude that $a = 8$ and $b = 16$, thus

$$\theta_4(q) = 1 + \sum_{m=1}^{\infty} \left[8 \left(\sum_{\substack{d \text{ odd} \\ d|m}} d \right) + 16 \left(\sum_{\substack{d \text{ odd} \\ d|\frac{m}{2}}} d \right) \right] q^m. \tag{13}$$

8.3 Sum of 6 Squares

When $n = 6$, we note that the determinant of the matrix N is zero, hence there is a linear dependence between the element $\theta_6(q)$ and the basis $\mathcal{B} = \{E_1(q), E_2(q)\}$ from (9) and (10) of the space $\mathcal{M}_3(\Gamma_1(4))$ (by Lemma 1, the space of cusp forms is trivial). Thus, we solve for constants $a, b \in \mathbb{C}$ in the equation

$$\theta_6(q) = a \cdot E_1(q) + b \cdot E_2(q).$$

The first few coefficients force $a = 16$ and $b = -4$, hence we can conclude that

$$\theta_6(q) = 1 + \sum_{m=1}^{\infty} \left[16 \left(\sum_{d|m} \chi\left(\frac{m}{d}\right) d^2 \right) - 4 \left(\sum_{d|m} \chi(d) d^2 \right) \right] q^m. \tag{14}$$

8.4 Sum of 8 Squares

When $n = 8$, $\theta_8(q) \in \mathcal{M}_4(\Gamma_1(4))$ and from the proof of Theorem 1, we have constructed a basis $\mathcal{B} = \{E(q), E(q^2), E(q^4)\}$ from (8) for this space. Thus, there exist constants $a, b, c \in \mathbb{C}$ such that

$$\theta_8(q) = a \cdot E(q) + b \cdot E(q^2) + c \cdot E(q^4).$$

Comparing the first few coefficients of each of these q-series, we find that $a = 16$, $b = -32$, and $c = 256$, thus we can conclude that

$$\theta_8(q) = 1 + \sum_{m=1}^{\infty} \left[16 \left(\sum_{d|m} d^3 \right) - 32 \left(\sum_{d|\frac{m}{2}} d^3 \right) + 256 \left(\sum_{d|\frac{m}{4}} d^3 \right) \right] q^m. \tag{15}$$

8.5 Sum of 10 Squares

Finally when $n = 10$, there is a non-trivial subspace of cusp forms equal to $\mathcal{S}_5^{cm}(\Gamma_1(4))$. By the construction in the proof of Theorem 1 for the case of $n \equiv 2 \bmod 8$ but $n > 10$, we claim that there exist constants $a, b, c \in \mathbb{C}$ such that

$$\theta_{10}(q) = a \cdot E_1(q) + b \cdot E_2(q) + c \cdot C(q)$$

using (9), (10), and (11) since $\theta_{10}(q) \in \mathcal{M}_5(\Gamma_1(4))$. Using the matrix of coefficients, we find that $a = \frac{4}{5}$, $b = \frac{64}{5}$, and $c = \frac{32}{5}$, hence

$$\theta_{10}(q) = 1 + \sum_{m=1}^{\infty} \left[\frac{4}{5} \left(\sum_{d|m} \chi(d) d^4 \right) + \frac{64}{5} \left(\sum_{d|m} \chi\left(\frac{m}{d}\right) d^4 \right) + \frac{8}{5} \left(\sum_{\substack{d \in \mathbb{Z}[i] \\ \mathrm{Nm}(d)=m}} d^4 \right) \right] q^m. \qquad (16)$$

9 Motivation for definition of "elementary" modular forms

The motivation behind the definition of an "elementary formula" is computational: an elementary modular form should have Fourier coefficients which can be computed in a straightforward manner in polynomial time in $\log(m)$ (for the coefficient of q^m). One can broaden this by assuming factorization of m while calculating the coefficient of q^m as is the case for small n, where the set of divisors of m was needed to calculate $r_n(m)$. We have provided the nice formulas for the exceptional cases of elementary θ_n, and these give way for efficiently computing $r_n(m)$ for $n = 2, 4, 6, 8$, and 10 (factorization of m is necessary). To demonstrate that modular forms on $\Gamma_1(4)$ that involve a non-CM cusp form are not "nice", we consider the case of $n = 12$.

9.1 $n = 12$

Note that the subspace of $\mathcal{M}_6(\Gamma_1(4))$ consisting of Eisenstein series has dimension 3 and the subspace of cusp forms has dimension 1 by Lemma 1. Furthermore, by Lemma 2, $\mathcal{S}_6^{cm}(\Gamma_1(4))$ is trivial, and we can in fact use a well-known cusp form for the basis:

$$\sqrt{\Delta}(2z) = \eta^{12}(2z) := q\prod_{n=1}^{\infty}(1-q^{2n})^{12} = q\left[1 - 12q^2 + 54q^4 - 88q^6 - 99q^8 + ...\right]$$

$\eta(2z)$ is a cusp form on $\Gamma_1(2)$ of weight 1/2, hence $\sqrt{\Delta}(2z) \in \mathcal{M}_6(\Gamma_1(4))$ and it is clearly a cusp form but does not have complex multiplication.

From above and (8), take $\{E(q), E(q^2), E(q^4), \sqrt{\Delta}(q^2)\}$ as a basis for $\mathcal{M}_6(\Gamma_1(4))$. We then calculate the constants $a, b, c, d \in \mathbb{C}$ such that

$$\theta_{12}(q) = a \cdot E(q) + b \cdot E(q^2) + c \cdot E(q^4) + d \cdot \sqrt{\Delta}(q^2).$$

Note that by Theorem 1, c_3 must be nonzero. Using the first few coefficients of the q-series in the basis,

$$\begin{aligned} E(q) &= a_0 + q + 33 \cdot q^2 + 244 \cdot q^3 + 1057 \cdot q^4 + ... \\ E(q^2) &= a_0 + q^2 + 33 \cdot q^4 + ... \\ E(q^4) &= a_0 + q^4 + ... \\ \sqrt{\Delta}(q^2) &= q - 12 \cdot q^3 + ... \end{aligned}$$

Since we know that the first few coefficients $r_{12}(m)$ are

$$\theta_{12}(q) = 1 + 24 \cdot q + 264 \cdot q^2 + 1760 \cdot q^3 + 7944 \cdot q^4 + ...,$$

thus $a = 8$, $b = 0$, and $c = -512$, $d = 16$.

Here, we can conclude that it is necessary to calculate the coefficients of η^{12} in order to calculate $\theta_{12}(n)$. Furthermore, the converse is true.

From Serre's point of view, η^{12} is not lacunary, thus not only is η^{12} not a CM cusp form, it also has a positive density of non-zero coefficients in its q-series (see [33]). The formulas for $n = 2, 4, 6, 8$, and 10 illustrate that $r_n(m)$ can be calculated efficiently if the prime factorization of m is known (thus for primes p,$r_n(p)$ can be computed easily). In the case of $n = 12$, this is not enough information as there is no analogous description of the coefficients of $\eta^{12}(2z)$ in terms of divisors.

Remark. Recently, Bas Edixhoven, Jean-Marc Couveignes, and Robin de Jong have proven that if $f = \sum a_n q^n$ is a modular form on $\mathrm{SL}_2(\mathbb{Z})$, then the coefficients a_p for p prime can be

calculated in time polynomial in the weight k and $\log(p)$ (assuming GRH). This implies that the prime Fourier coefficients $\tau(p)$ of Δ can be calculated in polynomial time with respect to $\log(p)$. Furthermore, Peter Bruin in his forthcoming PhD thesis [4] will give a probabilistic algorithm in time polynomial to k and $\log(p)$ which under the assumption of GRH, computes the Fourier coefficient of q^p of eigenforms of level $2N$ where N is squarefree. This includes calculating $r_n(m)$ for all even n as discussed here, whether or not θ_n has an elementary formula or not. In particular, showing that θ_n is not elementary, and therefore does not have a nice formula, outside of the small finite set $n \in \{2, 4, 6, 8, 10\}$ demonstrates the usefulness and necessity for such an algorithm. Other than understanding classical arithmetic functions, these results are useful in computing eigenvalues of Hecke operators or equivalently, coefficients of eigenforms (see [12]).

References

[1] A. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[2] Johan Bosman. *Explicit computations with modular Galois representations*. PhD in Mathematics, Faculty of Science, Leiden University, Leiden, Netherlands, 2008.

[3] N. Bourbaki. *Éléments de mathématique, Fasc. XXIII*. Hermann, Paris, 1973. Livre II: Algèbre. Chapitre 8: Modules et anneaux semi-simples, Nouveau tirage de l'édition de 1958, Actualités Scientifiques et Industrielles, No. 1261.

[4] Peter Bruin. *Modular curves, Arakelov theory, algorithmic applications*. PhD in Mathematics, Faculty of Science, Leiden University, Leiden, Netherlands, 2010.

[5] Brian Conrad. *Modular forms and the Ramanujan conjecture*. Cambridge University Press, New York, 2010.

[6] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. In *Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993)*, pages 2–140. Int. Press, Cambridge, MA, 1997.

[7] Pierre Deligne. Formes modulaires et représentations l-adiques. *Séminaire N. Bourbaki*, (355):139–172, 1968–1969.

[8] Pierre Deligne and Jean-Pierre Serre. Formes modulaires de poids 1. *Ann. Sci. École Norm. Sup. (4)*, 7:507–530 (1975), 1974.

[9] Fred Diamond and John Im. Modular forms and modular curves. In *Seminar on Fermat's Last Theorem (Toronto, ON, 1993–1994)*, volume 17 of *CMS Conf. Proc.*, pages 39–133. Amer. Math. Soc., Providence, RI, 1995.

[10] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[11] L. E. Dickson. *History of the Theory of Numbers, Volume 2: Diophantine Analysis*. Dover, New York, 2005.

[12] Bas Edixhoven. On the computation of the coefficients of a modular form. In *Algorithmic number theory*, volume 4076 of *Lecture Notes in Comput. Sci.*, pages 30–39. Springer, Berlin, 2006.

[13] Jean-Marc Fontaine and Barry Mazur. Geometric Galois representations. In *Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993)*, Ser. Number Theory, I, pages 41–78. Int. Press, Cambridge, MA, 1995.

[14] J. W. L. Glaisher. On the numbers of representations of a number as a sum of $2r$ squares, where $2r$ does not exceed eighteen. *Proc. London Math. Soc.*, 5:479–490, 1907.

[15] Phillip Griffiths and Joseph Harris. *Principles of algebraic geometry*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1994. Reprint of the 1978 original.

[16] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman.

[17] Erich Hecke. *Mathematische Werke*. Vandenhoeck & Ruprecht, Göttingen, third edition, 1983. With introductory material by B. Schoeneberg, C. L. Siegel and J. Nielsen.

[18] A. Hurwitz. *Vorlesungen uber die Zahlentheorie der Quaternionen*. Berlin, 1919.

[19] C. G. J. Jacobi. *Fundamenta Nova Theoriae Functionum Ellipticarum*. Konigsberg, Germany, 1829.

[20] David J. Leeming. The real zeros of the Bernoulli polynomials. *J. Approx. Theory*, 58(2):124–150, 1989.

[21] Wen Ch'ing Winnie Li. Newforms and functional equations. *Math. Ann.*, 212:285–315, 1975.

[22] J. Liouville. Extrait d'une lettre adressée a m. besge (concerning the representation of the double of an odd number as the sum of 12 squares). *J. de math. pure et appl.*, 9:296–298, 1864.

[23] James S. Milne. Modular functions and modular forms (v1.20), 2009. Available at www.jmilne.org/math/.

[24] Stephen C. Milne. Infinite families of exact sums of squares formulas, Jacobi elliptic functions, continued fractions, and Schur functions. *Ramanujan J.*, 6(1):7–149, 2002.

[25] Toshitsune Miyake. *Modular forms*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, english edition, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda.

[26] S. Ramanujan. On certain arithmetical functions [Trans. Cambridge Philos. Soc. **22** (1916), no. 9, 159–184]. In *Collected papers of Srinivasa Ramanujan*, pages 136–162. AMS Chelsea Publ., Providence, RI, 2000.

[27] R. A. Rankin. On the representation of a number as the sum of any number of squares, and in particular of twenty. *Acta Arith*, 7:399–407, 1961/1962.

[28] Kenneth A. Ribet. Galois representations attached to eigenforms with Nebentypus. In *Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976)*, pages 17–51. Lecture Notes in Math., Vol. 601. Springer, Berlin, 1977.

[29] J.-P. Serre. *A course in arithmetic*. Springer-Verlag, New York, 1973. Translated from the French, Graduate Texts in Mathematics, No. 7.

[30] Jean-Pierre Serre. Géométrie algébrique et géométrie analytique. *Ann. Inst. Fourier, Grenoble*, 6:1–42, 1955–1956.

[31] Jean-Pierre Serre. Une interprétation des congruences relatives à la fonction τ de Ramanujan. In *Séminaire Delange-Pisot-Poitou: 1967/68, Théorie des Nombres, Fasc. 1, Exp. 14*, page 17. Secrétariat mathématique, Paris, 1969.

[32] Jean-Pierre Serre. Quelques applications du théorème de densité de Chebotarev. *Inst. Hautes Études Sci. Publ. Math.*, (54):323–401, 1981.

[33] Jean-Pierre Serre. Sur la lacunarité des puissances de η. *Glasgow Math. J.*, 27:203–221, 1985.

[34] Goro Shimura. On elliptic curves with complex multiplication as factors of the Jacobians of modular function fields. *Nagoya Math. J.*, 43:199–208, 1971.

[35] Goro Shimura. Class fields over real quadratic fields and Hecke operators. *Ann. of Math. (2)*, 95:130–190, 1972.

[36] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*, volume 11 of *Publications of the Mathematical Society of Japan*. Princeton University Press, Princeton, NJ, 1994. Reprint of the 1971 original, Kanô Memorial Lectures, 1.

[37] Goro Shimura. The number of representations of an integer by a quadratic form. *Duke Math. J.*, 100(1):59–92, 1999.

[38] Goro Shimura. The representation of integers as sums of squares. *Amer. J. Math.*, 124(5):1059–1081, 2002.

[39] Carl Ludwig Siegel. Über die analytische Theorie der quadratischen Formen. *Ann. of Math. (2)*, 36(3):527–606, 1935.

[40] H. J. S. Smith. *Report on the Theory of Numbers*. Chelsea Publishing Company, New York, 1986.

[41] William Stein. *Modular forms, a computational approach*, volume 79 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2007. With an appendix by Paul E. Gunnells.

[42] Richard Taylor. Galois representations. *Ann. Fac. Sci. Toulouse Math. (6)*, 13(1):73–119, 2004.