Oberlin

Digital Commons at Oberlin

Honors Papers Student Work

2022

Jacobi's Four Squares Theorem

Arman Yagci
Oberlin College

Follow this and additional works at: https://digitalcommons.oberlin.edu/honors

Part of the Mathematics Commons

Repository Citation

Yagci, Arman, "Jacobi's Four Squares Theorem" (2022). *Honors Papers*. 844.
https://digitalcommons.oberlin.edu/honors/844

This Thesis - Open Access is brought to you for free and open access by the Student Work at Digital Commons at Oberlin. It has been accepted for inclusion in Honors Papers by an authorized administrator of Digital Commons at Oberlin. For more information, please contact megan.mitchell@oberlin.edu.

Jacobi's Four Squares Theorem

Arman Yagci

Supervisor: Benjamin Linowitz

Abstract

Jacobi's Four Squares Theorem is a celebrated result of number theory that provides a formula for the number of ways a positive integer n can be written as a sum of four integral squares. In this paper, we prove this theorem using the theory of modular forms.

1 Introduction

Given a natural number n, there are efficient ways to determine whether n is a perfect square. If it is not, an interesting question to follow with is whether it is a sum of two squares. In general, given $m > 1$, mathematicians have wanted to determine which integers could be expressed as a sum of m integral squares. In 1640, Pierre de Fermat wrote to Marin Mersenne about his Sum of Two Squares Theorem, in which he outlined a proof method of descent. As he often did, however, Fermat did not publish the full proof. It was a century later, in 1749, that Leonhard Euler managed to provide the first published proof of the theorem by following Fermat's method of descent.

Fermat's Two Squares Theorem. *An integer $n > 1$ can be written as a sum of two squares if and only if the prime factorization of n does not contain p^k (where k is the multiplicity of p) for all primes p such that $p \equiv 3 \pmod 4$, k odd.*

Later, in 1797, Adrien-Marie Legendre published his Three Squares Theorem, which provides a classification for $m = 3$.

Legendre's Three Squares Theorem. *A natural number n can be written as a sum of three squares if and only if $n \neq 4^a(8b + 7)$ for any $a, b \geq 0$.*

It was actually a bit earlier in 1770 that Joseph-Louis Lagrange famously proved the following result based on Euler's work on sums of two squares.

Lagrange's Four Squares Theorem. *Every natural number can be expressed as a sum of four integral squares.*

This completed the picture, as for $m > 4$, one can simply add squares of zeros as necessary to express any integer as a sum of m squares. Now that it was known that given a natural number, there was at least one way of writing it as a sum of four squares, the question became: exactly how many ways are there? In this paper, we prove Jacobi's Four Squares Theorem, first proved in 1834 by Carl Jacobi, which expands on Lagrange's result by providing a formula for the number of ways an integer n can be expressed as a sum of four squares. For an in-depth history of these theorems, please refer to [Weil 1906].

Before we state Jacobi's theorem, a couple of remarks are in order. When we count the number of ways of writing n as a sum of four squares, we take order into account. For example, here are all of the eight different ways 1 can be written as a sum of four integral squares:

$$\begin{aligned} 1 &= (\pm 1)^2 + 0^2 + 0^2 + 0^2 = 0^2 + (\pm 1)^2 + 0^2 + 0^2 \\ &= 0^2 + 0^2 + (\pm 1)^2 + 0^2 = 0^2 + 0^2 + 0^2 + (\pm 1)^2 \ . \end{aligned}$$

Although the difference between the eight ways of writing 1 as a sum of four squares is rather trivial, this is not necessarily the case. As an example, two nontrivially different ways of writing 10 as a sum of four squares are $3^2 + 1^2 + 0^2 + 0^2$ and $2^2 + 2^2 + 1^2 + 1^2$. We now state the theorem:

Jacobi's Four Squares Theorem. *For any positive integer n, let a_n denote the number of ways n can be expressed as a sum of four integral squares. Then,*

$$a_n = \begin{cases} 8\sigma(n) & \text{for } n \text{ odd,} \\ 24\sigma(n_0) & \text{for } n = 2^r n_0 \text{ even, } n_0 \text{ odd.} \end{cases}$$

Here, $\sigma(n)$ gives the sum of the positive divisors of n. Although the theorem only considers values of n that are positive integers, it's trivial to see that 0 can only be expressed as $0^2+0^2+0^2+0^2$ and that there is no way to express a negative integer as a sum of squares.

Our proof of the theorem heavily utilizes the theory of modular forms. We begin Section 2 by introducing the concept of modular forms as defined for congruence subgroups of $SL_2(\mathbb{Z})$. Section 3 introduces Hecke operators defined on spaces of modular forms. Our proof of Jacobi's Four Squares Theorem begins in Section 4, and it follows the following outline: We first find a generating series for the sequence $\{a_n\}_{n\in\mathbb{Z}}$ and show that it is a modular form. Next, we find a basis for the space of modular forms containing said generating series and use that basis to find eigenforms for our Hecke operators. Finally, by comparing the Fourier coefficients of particular eigenforms, we obtain the desired formula. The reader is encouraged to skim the technical lemmas and propositions in these sections and focus on following this outline for the initial reading.

2 Modular Forms

2.1 The general linear group acts on the Riemann sphere

For any subring R of $\mathbb{R}$, the *general linear group* $GL_2(R)$ is defined as the set of 2×2 invertible matrices with entries in R. The *special linear group* $SL_2(R)$ is the subgroup of $GL_2(R)$ consisting of matrices with determinant 1.

$GL_2(\mathbb{R})$ acts on the Riemann sphere $\mathbb{C} \cup \{\infty\}$ by fractional linear transformations. (Here, imagining the point at infinity far along the imaginary axis, we take $\infty = i\infty$.) That is, given $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R})$ and $z \in \mathbb{C} \cup \{\infty\}$, we define

$$\gamma z := \frac{az+b}{cz+d}; \qquad \qquad \gamma\infty := \lim_{z \to i\infty} \gamma z = a/c.$$

We remark that for $c \neq 0$, $\gamma(-d/c) = \lim_{z \to -d/c} \gamma z = \infty$, and if $c = 0$, $\gamma\infty = \lim_{z \to i\infty} z/d = \infty$ ($d \neq 0$ as $\det(\gamma) \neq 0$).

We will denote by $\mathcal{H}$ the upper half-plane of the complex plane; that is,

$$\mathcal{H} = \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}.$$

Let $GL_2^+(\mathbb{Q})$ denote the subgroup of $GL_2(\mathbb{Q})$ consisting of matrices with positive determinant.

Lemma 2.1. *$GL_2^+(\mathbb{Q})$ preserves $\mathcal{H}$.*

Proof. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2^+(\mathbb{Q})$ and $z \in \mathcal{H}$. It suffices to show that $\operatorname{Im}(\gamma z) > 0$. We have

$$\operatorname{Im}(\gamma z) = \operatorname{Im} \frac{az+b}{cz+d} = \operatorname{Im} \frac{(az+b)(c\bar{z}+d)}{|cz+d|^2} = |cz+d|^{-2} \operatorname{Im}(adz + bc\bar{z}),$$

where

$$\operatorname{Im}(adz + bc\bar{z}) = ad \operatorname{Im}(z) + bc \operatorname{Im} \bar{z} = (ad - bc) \operatorname{Im} z.$$

Since $z \in \mathcal{H}$, $\operatorname{Im} z > 0$; and $\gamma \in GL_2^+(\mathbb{Q})$ implies $\det \gamma = ad - bc > 0$.

Hence, $\operatorname{Im}(\gamma z) = |cz+d|^{-2}(ad-bc) \operatorname{Im} z > 0$ as well. □

As $GL_2^+(\mathbb{Q}) \supseteq SL_2(\mathbb{Z})$, we note that $SL_2(\mathbb{Z})$ preserves $\mathcal{H}$ as well by Lemma 2.1.

Throughout the paper, we'll frequently encounter two matrices, which are

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

These matrices generate $SL_2(\mathbb{Z})$, and the proof is similar to our proof of Proposition 2.3. (See Exercise 1.1.1 in [Diamond and Shurman (2005)]).

2.2 Congruence Subgroups of $SL_2(\mathbb{Z})$

Definition 2.2. We define the *principal congruence subgroup of level N* as

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

where two matrices are said to be congruent modulo N if their corresponding entries are congruent modulo N.

For all $N \in \mathbb{Z}^+$, note that $\Gamma(N)$ is the kernel of the surjective homomorphism from $SL_2(\mathbb{Z})$ to $SL_2(\mathbb{Z}/N\mathbb{Z})$, induced by reduction modulo N. Hence, $\Gamma(N)$ is a normal subgroup of $SL_2(\mathbb{Z})$. Moreover, as the homomorphism embeds the cosets of $\Gamma(N)$ in $SL_2(\mathbb{Z})$ into the finite group $SL_2(\mathbb{Z}/N\mathbb{Z})$, it follows that $\Gamma(N)$ has finite index in $SL_2(\mathbb{Z})$.

A subgroup of $SL_2(\mathbb{Z})$ is called a *congruence subgroup* if it contains $\Gamma(N)$ for some N. It will then further be called a *congruence subgroup of level N*. Note that for all multiples N' of N, $\Gamma(N') \subseteq \Gamma(N)$. This implies that a subgroup that contains $\Gamma(N)$ will also contain $\Gamma(N')$. Therefore, the level of a congruence subgroup is not unique. (Athough some refer to the smallest level as "the level" of the congruence subgroup, this is not necessary for our purposes.) Moreover, it is immediate that $SL_2(\mathbb{Z})$ itself is a congruence subgroup of all levels N.

Some important congruence subgroups we will use are

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

and

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

where the entries marked by $*$ can be any integer.

Below, we find generators for $\Gamma_0(4)$, which will be particularly useful in the proof of our main theorem.

Proposition 2.3. *$-I$, T, and ST^4S generate $\Gamma_0(4)$.*

Proof. First, we calculate

$$ST^4S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 4 & -1 \end{pmatrix} .$$

It's trivial to see that $-I, T, ST^4S \in \Gamma_0(4)$, and so they must generate a subgroup of $\Gamma_0(4)$. Let $\gamma = \begin{pmatrix} a & b \\ 4c & d \end{pmatrix} \in \Gamma_0(4)$ be arbitrary. It suffices to show that γ can be written in terms of $-I, T$, and ST^4S (and their inverses).

Note that $\det(\gamma) = ad - b(4c) = 1$ implies $ad \equiv 1 \pmod 4$ so that either $a \equiv d \equiv 1 \pmod 4$ or $a \equiv d \equiv -1 \pmod 4$. In particular, this shows that $d \not\equiv 0 \pmod 2$.

We will use induction on c. As we can multiply γ by $-I$ if necessary, without loss of generality, we may assume that $c \geq 0$. If $c = 0$, then $ad = 1$ forces $a = d = \pm 1$. If $a = d = 1$, then $\gamma = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = T^b$. If $a = d = -1$, then $\gamma = \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} = -IT^{-b}$.

Now, suppose $\gamma = \begin{pmatrix} a & b \\ 4c & d \end{pmatrix}$ can be written in terms of $-I, T$, and ST^4S whenever $0 \leq c < c'$ for some $c' \in \mathbb{Z}^+$. Consider $\gamma = \begin{pmatrix} a & b \\ 4c' & d \end{pmatrix}$. As we saw that $d \not\equiv 0 \pmod 2$, we have $2kc' < |d| < 2(k+1)c'$ for some $k \in \mathbb{N} \cup \{0\}$.

If k is even and $d > 0$, let $n = -1$ and $m = -k/2$.
If k is even and $d < 0$, let $n = 1$ and $m = k/2$.

If k is odd and $d > 0$, let $n = 1$ and $m = -(k+1)/2$.
If k is odd and $d < 0$, let $n = -1$ and $m = (k+1)/2$.

Then, we have

$$\gamma T^m(-IST^4S)^{-n} = \begin{pmatrix} a & b \\ 4c' & d \end{pmatrix} \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 4n & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ 4c' & d \end{pmatrix} \begin{pmatrix} 4mn+1 & m \\ 4n & 1 \end{pmatrix},$$

which has lower left entry $4c'(4mn+1) + 4nd = 4((4mn+1)c' + nd)$. It's straightforward to check that our choice of m, n satisfies $-c' < (4mn+1)c' + nd < c'$. Multiplying γ by $-I$ if necessary, without loss of generality, this implies that $\gamma T^m(-IST^4S)^{-n}$ has lower left entry in the form of $4c$ where $0 \le c < c'$. By the induction hypothesis, we conclude that $\gamma T^m(-IST^4S)^{-n}$ can be written in terms of $-I, T$, and ST^4S, which in turn implies that γ can be written in terms of $-I, T$, and ST^4S.

We've shown that $-I$, T, and ST^4S generate $\Gamma_0(4)$. □

Although it is not a group, another useful set to define is

$$\Delta^n(N, \{1\}, \mathbb{Z}) := \left\{ \text{integer matrices } \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & n \end{pmatrix} \pmod{N} \text{ and } \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = n \right\}.$$

Below we find "coset representatives" for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$.

First, for all $N \in \mathbb{Z}$, we define

$$\alpha_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}.$$

Lemma 2.4. *For $n, N \in \mathbb{Z}^+$, let $n = ad$ where $a \in (\mathbb{Z}/N\mathbb{Z})^*$ and $d \in \mathbb{Z}^+$. For each possible choice for a, fix $\sigma_a \equiv \begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix} \pmod{N}$ in $\Gamma_0(N)$ where a^{-1} denotes an integer representing the inverse of a modulo N. Then,*

$$\Delta^n(N, \{1\}, \mathbb{Z}) = \bigcup_{\textit{disjoint}} \Gamma_1(N)\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix},$$

where the disjoint union is taken over all pairs a, b such that $b \in \{0, 1, \ldots, d-1\}$.

Moreover, if $\gcd(n, N) = 1$, then n has an inverse modulo N so that we can fix $\sigma_n \equiv \begin{pmatrix} n^{-1} & 0 \\ 0 & n \end{pmatrix}$ $\pmod{N}$ in $\Gamma_0(N)$. Let $a \in (\mathbb{Z}/N\mathbb{Z})^*$ satisfy $n = ad$, and let $b \in \{0, 1, \ldots, d-1\}$. Define $\alpha_{a,b} = \sigma_n \alpha_N \sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \alpha_N^{-1}$. Then,

$$\Delta^n(N, \{1\}, \mathbb{Z}) = \bigcup_{\textit{disjoint}} \Gamma_1(N)\alpha_{a,b}\ .$$

Proof. To prove the first equality, it suffices to show that $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ form a complete set of coset representatives for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$. To see that $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \Delta^n(N, \{1\}, \mathbb{Z})$ for all a, b, observe that $\det \left(\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \right) = \det(\sigma_a) \det \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = 1 \cdot n = n$ and

$$\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \equiv \begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} 1 & ba^{-1} \\ 0 & n \end{pmatrix} \pmod{N},$$

as required.

<u>Claim 1</u>: $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ are in distinct cosets of $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$.

<u>Proof</u>: Suppose, for the sake of contradiction, that there exist distinct $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ and $\sigma_{a'} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}$ such that $\Gamma_1(N)\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \Gamma_1(N)\sigma_{a'} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}$. Then, there exists a $\gamma \in \Gamma_1(N)$ such that

$$\gamma\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \sigma_{a'} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}. \tag{1}$$

Since $\gamma, \sigma_a, \sigma_{a'} \in SL_2(\mathbb{Z})$, we must have

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}^{-1} = \begin{pmatrix} a/a' & \frac{-ab'+ba'}{a'd'} \\ 0 & d/d' \end{pmatrix} \in SL_2(\mathbb{Z}) \tag{2}$$

as well. But then $\det \begin{pmatrix} a/a' & \frac{-ab'+ba'}{a'd'} \\ 0 & d/d' \end{pmatrix} = 1$ implies a/a' and d/d' are two integers whose product is 1. Thus, $a/a' = d/d' = 1$ since a, a', d, d' are positive integers, which forces $a = a'$ and $d = d'$.

Now, note that $b' + d \left(\frac{-ab'+ba}{ad} \right) = b$, where $\frac{-ab'+ba}{ad} = \frac{-ab'+ba'}{a'd'}$ is an integer by equation (2). Furthermore, since we have $0 \leq b, b' < d$, we must have $d \left(\frac{-ab'+ba}{ad} \right) = 0$. Hence, $b = b'$. We've shown that $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \sigma_{a'} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}$, contradicting our assumption that $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ and $\sigma_{a'} \begin{pmatrix} a' & b' \\ 0 & d' \end{pmatrix}$ were distinct. We conclude that $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ are in distinct cosets of $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$, proving the claim. $\triangle$

<u>Claim 2</u>: For any $\alpha = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \Delta^n(N, \{1\}, \mathbb{Z})$, $\alpha \in \Gamma_1(N)\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ for some $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$.

<u>Proof</u>: We'll first show that there exists a $\gamma \in SL_2(\mathbb{Z})$ such that $\gamma\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$, where $\gcd(a, N) = 1$, $a > 0$, $ad = n$, and $0 \leq b < d$.

Let $g = \frac{c'}{\gcd(a',c')}$ and $h = -\frac{a'}{\gcd(a',c')}$. Then, $ga' + hc' = 0$. Since $\gcd(g, h) = 1$, there exist $e, f \in \mathbb{Z}$ such that $eh - gf = 1$. Thus, $\gamma = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \in SL_2(\mathbb{Z})$. Consider $\gamma\alpha$ and write $\gamma\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some $a, b, c, d \in \mathbb{Z}$. Since $\det(\gamma\alpha) = \det(\gamma)\det(\alpha) = 1 \cdot n = n$ and $c = ga' + hc' = 0$, $\gamma\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$, where $d = n/a$ so that $\det(\gamma\alpha) = n$. Observe also that $a' \equiv 1 \pmod{N}$ and $c' \equiv 0 \pmod{N}$ imply $a = ea' + fc' \equiv e \pmod{N}$. As $N \mid c'$ and $\gcd(a', N) = 1$, we also have $g = \frac{c'}{\gcd(a',c')} \equiv 0 \pmod{N}$. Thus, $eh - gf = 1$ implies $eh \equiv 1 \pmod{N}$ so that $\gcd(e, N) = 1$. As $a \equiv e \pmod{N}$, $\gcd(a, N)$ must divide N and e, which forces $\gcd(a, N) = 1$.

Since $\det(\gamma\alpha) = n \neq 0$, $a \neq 0$. Suppose $a < 0$. Since $d = n/a$, $d < 0$ as well. Choose $j \in \mathbb{Z}$ such that $0 \leq -b + jd < -d$. Then, $\begin{pmatrix} -1 & j \\ 0 & -1 \end{pmatrix} \gamma\alpha = \begin{pmatrix} -a & -b + jd \\ 0 & -d \end{pmatrix}$ is a matrix with determinant n, $-a > 0$, and $0 \leq -b + jd < -d$. Similary, if $a > 0$, choose $j \in \mathbb{Z}$ such that $0 \leq b + jd < d$. Then, $\begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix} \gamma\alpha = \begin{pmatrix} a & b + jd \\ 0 & d \end{pmatrix}$ is a matrix with determinant n, $a > 0$, and $0 \leq b + jd < d$. Hence, without loss of generality, we may suppose that $\gamma\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ has determinant n, where $a > 0$ and $0 \leq b < d$.

Next, note that $\alpha \in \Delta^n(N, \{1\}, \mathbb{Z})$ implies $\alpha \equiv \begin{pmatrix} 1 & s \\ 0 & n \end{pmatrix} \pmod{N}$ for some $s \in \mathbb{Z}$. Thus, $\alpha = \gamma^{-1} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ gives us $\begin{pmatrix} 1 & s \\ 0 & n \end{pmatrix} \equiv \gamma^{-1} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \pmod{N}$. Hence, $\gamma^{-1} \equiv \begin{pmatrix} a^{-1} & t \\ 0 & a \end{pmatrix} \pmod{N}$ for some $t \in \mathbb{Z}$. Observe that

$$\gamma^{-1}\sigma_a^{-1} \equiv \begin{pmatrix} a^{-1} & t \\ 0 & a \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \equiv \begin{pmatrix} aa^{-1} & ta^{-1} \\ 0 & aa^{-1} \end{pmatrix} \equiv \begin{pmatrix} 1 & ta^{-1} \\ 0 & 1 \end{pmatrix} \pmod{N},$$

which shows that $\gamma^{-1}\sigma_a^{-1} \in \Gamma_1(N)$.
Hence, $\gamma^{-1} \in \Gamma_1(N)\sigma_a$ so that $\alpha \in \Gamma_1(N)\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$, as desired. $\triangle$

This proves the first equality. Next, to prove the second equality, it suffices to show that $\alpha_{a,b}$ form a complete set of coset representatives for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$. To see that $\alpha_{a,b} \in \Delta^n(N, \{1\}, \mathbb{Z})$ for all a, b, observe that

$$\begin{aligned} \alpha_{a,b} &\equiv \begin{pmatrix} n^{-1} & 0 \\ 0 & n \end{pmatrix} \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix} \begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} a & b \\ 0 & n/a \end{pmatrix} \begin{pmatrix} 0 & 1/N \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & * \\ -Nbna^{-1} & n \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & * \\ 0 & n \end{pmatrix} \pmod{N}. \end{aligned}$$

Furthermore, as $\sigma_n, \sigma_a \in \Gamma_0(N)$, we see that

$$\det(\alpha_{a,b}) = \det(\sigma_n)\det(\alpha_N)\det(\sigma_a)\det\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}\det(\alpha_N^{-1}) = 1 \cdot N \cdot 1 \cdot n \cdot \frac{1}{N} = n.$$

Hence, $\alpha_{a,b} \in \Delta^n(N, \{1\}, \mathbb{Z})$.

Claim 3: $\alpha_{a,b}$ are in distinct cosets of $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$.

Proof: Suppose, for the sake of contradiction, that there exist distinct $\alpha_{a,b}$ and $\alpha_{a',b'}$ such that $\Gamma_1(N)\alpha_{a,b} = \Gamma_1(N)\alpha_{a',b'}$. Then, there exists a $\gamma \in \Gamma_1(N)$ such that $\gamma\alpha_{a,b} = \alpha_{a',b'}$. Thus, we have

$$\gamma\left(\sigma_n\alpha_N\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix}\alpha_N^{-1}\right) = \sigma_n\alpha_N\sigma_{a'}\begin{pmatrix}a' & b'\\ 0 & d'\end{pmatrix}\alpha_N^{-1} \Rightarrow \gamma\sigma_n\alpha_N\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} = \sigma_n\alpha_N\sigma_{a'}\begin{pmatrix}a' & b'\\ 0 & d'\end{pmatrix}$$
$$\Rightarrow \sigma_n^{-1}\gamma\sigma_n\alpha_N\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} = \alpha_N\sigma_{a'}\begin{pmatrix}a' & b'\\ 0 & d'\end{pmatrix}$$

Let $\gamma' = \sigma_n^{-1}\gamma\sigma_n$. Write $\gamma = \begin{pmatrix}j & k\\ \ell & m\end{pmatrix}$ and note that

$$\gamma' = \sigma_n^{-1}\gamma\sigma_n \equiv \begin{pmatrix}n & 0\\ 0 & n^{-1}\end{pmatrix}\begin{pmatrix}j & k\\ \ell & m\end{pmatrix}\begin{pmatrix}n^{-1} & 0\\ 0 & n\end{pmatrix} = \begin{pmatrix}j & n^2k\\ \ell(n^{-1})^2 & m\end{pmatrix} \pmod N.$$

As $\sigma_n, \gamma' \in \Gamma_0(N)$, $\det(\gamma') = \det(\sigma_n^{-1})\det(\gamma)\det(\sigma_n) = 1\cdot 1\cdot 1 = 1$. Furthermore, $\gamma \in \Gamma_1(N)$ implies $j \equiv m \equiv 1 \pmod N$ and $\ell(n^{-1})^2 \equiv 0 \pmod N$. It follows that $\gamma' \in \Gamma_1(N)$. We have

$$\gamma'\alpha_N\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} = \alpha_N\sigma_{a'}\begin{pmatrix}a' & b'\\ 0 & d'\end{pmatrix} \Rightarrow \alpha_N^{-1}\gamma'\alpha_N\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} = \sigma_{a'}\begin{pmatrix}a' & b'\\ 0 & d'\end{pmatrix}.$$

Next, let $\bar\gamma = \alpha_N^{-1}\gamma'\alpha_N$. Once again, we write $\gamma' = \begin{pmatrix}j' & k'\\ \ell' & m'\end{pmatrix}$ for some $j', k', \ell', m' \in \mathbb{Z}$ and note that

$$\bar\gamma = \alpha_N^{-1}\gamma'\alpha_N = \begin{pmatrix}0 & 1/N\\ -1 & 0\end{pmatrix}\begin{pmatrix}j' & k'\\ \ell' & m'\end{pmatrix}\begin{pmatrix}0 & -1\\ N & 0\end{pmatrix} = \begin{pmatrix}m' & -\ell'/N\\ -Nk' & j'\end{pmatrix}.$$

As conjugation doesn't change the determinant, $\det\begin{pmatrix}m' & -\ell'/N\\ -Nk' & j'\end{pmatrix} = \det\begin{pmatrix}j' & k'\\ \ell' & m'\end{pmatrix} = 1$. Since $\gamma' \in \Gamma_1(N)$, we also have $m' \equiv j' \equiv 1 \pmod N$ and $-\ell'/N \in \mathbb{Z}$. Finally, $-Nk' \equiv 0 \pmod N$ shows that $\bar\gamma \in \Gamma_1(N)$. We've obtained the equation

$$\bar\gamma\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} = \sigma_{a'}\begin{pmatrix}a' & b'\\ 0 & d'\end{pmatrix}. \tag{3}$$

Now, following the proof of Claim 1 from equation (1) onwards with the substitution $\gamma = \bar\gamma$ shows that $a = a'$, $b = b'$, and $d = d'$. Hence, $\alpha_{a,b} = \alpha_{a',b'}$, which contradicts our assumption that $\alpha_{a,b}$ and $\alpha_{a',b'}$ were distinct. We conclude that $\alpha_{a,b}$ are in distinct cosets of $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$, proving the claim. △

Claim 1 and Claim 2 show that every coset of $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$ can be written in the form $\Gamma_1(N)\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix}$. Furthermore, it is clear from the definition of $\alpha_{a,b}$ that each $\alpha_{a,b}$ is uniquely determined by $\sigma_a\begin{pmatrix}a & b\\ 0 & d\end{pmatrix}$. Since Claim 3 shows that $\alpha_{a,b}$ are in distinct cosets, they must form a complete set of coset representatives. This completes the proof of the lemma. □

2.3 Weight-k operators

Definition 2.5. Let $f : \mathcal{H} \cup \mathbb{Q} \cup \{\infty\} \longrightarrow \mathbb{C} \cup \{\infty\}$ be a function, and let $k \in \mathbb{Z}$. Given $\gamma = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in GL_2^+(\mathbb{Q})$, we define the *weight-k operator* $[\gamma]_k$ by

$$f(z)|[\gamma]_k := (\det\gamma)^{k/2}(cz+d)^{-k}f(\gamma z)\ .$$

Here, Lemma 2.1 ensures that γz is in the domain of f.

Proposition 2.6. *The weight-k operator preserves addition and scalar multiplication of functions that map from $\mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$ to $\mathbb{C} \cup \{\infty\}$.*

Proof. Let $f, g : \mathcal{H} \cup \mathbb{Q} \cup \{\infty\} \longrightarrow \mathbb{C} \cup \{\infty\}$, $r \in \mathbb{C}$, $k \in \mathbb{Z}$, and $\gamma \in GL_2^+(\mathbb{Q})$. Then, by Definition 2.5, we have

$$\begin{aligned}(f+g)(z)|[\gamma]_k &= \det(\gamma)^{k/2}(cz+d)^{-k}(f+g)(\gamma z)\\ &= \det(\gamma)^{k/2}(cz+d)^{-k}(f(\gamma z)+g(\gamma z))\\ &= \det(\gamma)^{k/2}(cz+d)^{-k}f(\gamma z)+\det(\gamma)^{k/2}(cz+d)^{-k}g(\gamma z)\\ &= f(z)|[\gamma]_k + g(z)|[\gamma]_k\end{aligned}$$

and

$$(rf)(z)|[\gamma]_k = \det(\gamma)^{k/2}(cz+d)^{-k}(rf)(\gamma z) = r\det(\gamma)^{k/2}(cz+d)^{-k}f(\gamma z) = rf(z)|[\gamma]_k,$$

as desired. □

Proposition 2.7. $f|[\gamma_1\gamma_2]_k = (f|[\gamma_1]_k)|[\gamma_2]_k$ *for all* $\gamma_1, \gamma_2 \in GL_2^+(\mathbb{Q})$.

Proof. For all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2^+(\mathbb{Q})$, observe that

$$\begin{aligned}\left(\frac{\mathrm{d}\gamma z}{\mathrm{d}z}\right)^{k/2} f(\gamma z) &= \left(\frac{\mathrm{d}(\frac{az+b}{cz+d})}{\mathrm{d}z}\right)^{k/2} f(\gamma z)\\ &= \left(\frac{a(cz+d)-c(az+b)}{(cz+d)^2}\right)^{k/2} f(\gamma z)\\ &= \frac{(acz+ad-caz-cb)^{k/2}}{(cz+d)^{2k/2}} f(\gamma z)\\ &= \frac{(ad-bc)^{k/2}}{(cz+d)^k} f(\gamma z)\\ &= (\det\gamma)^{k/2}(cz+d)^{-k}f(\gamma z)\\ &= f(z)|[\gamma]_k\ .\end{aligned}$$

Let $\gamma_1, \gamma_2 \in GL_2^+(\mathbb{Q})$. Since $GL_2^+(\mathbb{Q})$ is closed under matrix multiplication, $\gamma_1\gamma_2 \in GL_2^+(\mathbb{Q})$ as well. Consider

$$f(z)|[\gamma_1\gamma_2]_k = \left(\frac{\mathrm{d}\gamma_1\gamma_2 z}{\mathrm{d}z}\right)^{k/2} f(\gamma_1\gamma_2 z).$$

By the chain rule, this equals $\left(\frac{\mathrm{d}\gamma_1(\gamma_2 z)}{\mathrm{d}(\gamma_2 z)} \cdot \frac{d(\gamma_2 z)}{dz}\right)^{k/2} f(\gamma_1\gamma_2 z) = \left(\frac{\mathrm{d}\gamma_1(\gamma_2 z)}{\mathrm{d}(\gamma_2 z)}\right)^{k/2} \left(\frac{\mathrm{d}(\gamma_2 z)}{\mathrm{d}z}\right)^{k/2} f(\gamma_1\gamma_2 z)$.

On the other hand, let $g(z) = f(z)|[\gamma_1]_k = \left(\frac{\mathrm{d}\gamma_1 z}{\mathrm{d}z}\right)^{k/2} f(\gamma_1 z)$ and observe that

$$(f(z)|[\gamma_1]_k)|[\gamma_2]_k = g(z)|[\gamma_2]_k = \left(\frac{\mathrm{d}\gamma_2 z}{\mathrm{d}z}\right)^{k/2} g(\gamma_2 z).$$

Since $g(\gamma_2 z) = \left(\frac{\mathrm{d}\gamma_1(\gamma_2 z)}{\mathrm{d}(\gamma_2 z)}\right)^{k/2} f(\gamma_1\gamma_2 z)$, we have shown that $f(z)|[\gamma_1\gamma_2]_k = (f(z)|[\gamma_1]_k)|[\gamma_2]_k$. □

2.4 Definition of A Modular Form for Congruence Subgroups

Definition 2.8. Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$ and let k be an integer. A function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is a *modular form of weight k with respect to* Γ if

(1) f is holomorphic on $\mathcal{H}$,
(2) $f|[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in SL_2(\mathbb{Z})$, and
(3) $f = f|[\gamma]_k$ for all $\gamma \in \Gamma$.

A function that satisfies the third condition is called *weight-k invariant under* Γ. The set of modular forms of weight k with respect to Γ is denoted $M_k(\Gamma)$.

Let f be a modular form. This means that $f \in M_k(\Gamma)$ where $\Gamma \subseteq SL_2(\mathbb{Z})$ is a congruence subgroup of level N for some $N \in \mathbb{Z}^+$. Then, $T^N = \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \in \Gamma(N) \subseteq \Gamma$ implies $f(z+N) = f(z)|[T^N]_k = f(z)$ so that f is N-periodic. In particular, since f is both holomorphic (i.e., $f \in C^\infty$) and periodic, it must have a Fourier expansion that converges to f. That is, one can write

$$f(z) = \sum_{n=-\infty}^{\infty} a_n q^{n/h}, \quad \text{where } q = e^{2\pi i z}$$

for some period $h \in \mathbb{Z}^+$, and the sequence of Fourier coefficients $\{a_n(f)\}_{n\in\mathbb{Z}}$ uniquely determines f.

Furthermore, when we say that a holomorphic and h-periodic function f is holomorphic at ∞ (which is true for all modular forms f by Definition 2.8), we mean that $f(z)$ is bounded as $z \to i\infty$, or, equivalently, that f has a Fourier expansion of the form

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

To see that these two definitions are equivalent, note that if f were to have a nonzero Fourier coefficient a_n for some $n < 0$, then

$$\lim_{z\to i\infty} |a_n e^{2\pi i n z/h}| = \lim_{b\to\infty} |a_n e^{2\pi |n| b/h}| = +\infty$$

would imply that f is not bounded as $z \to i\infty$. On the other hand, when $n \geq 0$, we have

$$\lim_{z\to i\infty} |a_n e^{2\pi i n z/h}| = \begin{cases} |a_0| < +\infty & \text{if } n = 0; \\ 0 & \text{if } n > 0. \end{cases}$$

Due to the polynomial growth rate of a_n (as will be seen in Lemma 2.10), $\lim_{\substack{n\to\infty \\ z\to i\infty}} a_n q^{n/h} = 0$ so that the series $\sum_{n=0}^{\infty} a_n q^{n/h}$ diverges if and only if there exists an $n \in \mathbb{Z}$ such that $\lim_{z\to i\infty} |a_n q^{n/h}| = +\infty$. As we've shown that no such $n \geq 0$ exists, we conclude that $\sum_{n=0}^{\infty} a_n q^{n/h}$ is indeed bounded as $z \to i\infty$.

Remark 2.9. Suppose $f \in M_k(\Gamma)$ where Γ contains $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then, $f(z+1) = f(z)|[T]_k = f(z)$ so that f has period 1 and hence has a Fourier expansion $f(z) = \sum_{n=0}^{\infty} a_n(f) q^n$.

To confirm conditions (1) and (2) in Definition 2.8, we combine Lemma 4.3.3 in [Miyake (2006)] and Proposition 1.2.4 in [Diamond and Shurman (2005)] into the following lemma:

Lemma 2.10. *Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$ of level N, and let $f : \mathcal{H} \longrightarrow \mathbb{C}$ be a 1-periodic function that is weight-k invariant under Γ. Then, $f \in M_k(\Gamma)$ if and only if f has a Fourier expansion of the form*

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

that satisfies $|a_n| \leq cn^r$ for all $n \in \mathbb{Z}^+$ and some positive constants c and r.

Remark 2.11. Lemma 4.3.3 in [Miyake (2006)] considers Fourier series of the form $f = \sum_{n=1}^{\infty} a_n(f)q^n$. On the other hand, if we let $g = \sum_{n=0}^{\infty} a_n(g)q^n$ such that $g = a_0(g) + f$, Proposition 1.2.4 in [Diamond and Shurman (2005)] looks at g written in the form $g = \sum_{n=0}^{\infty} a_n'(g)q^{n/N}$. Suppose $|a_n(f)| \leq cn^r$ for all $n \in \mathbb{Z}^+$ and some positive constants c, r. Then, since $a_n(f) = a_n(g)$ for all $n \in \mathbb{Z}^+$, we also have $|a_n(g)| \leq cn^r$. Observe that

$$g = \sum_{n=0}^{\infty} a_n(g)q^n = \sum_{\substack{n \geq 0 \\ N|n}} a_{n/N}(g)q^{n/N} = \sum_{n=0}^{\infty} a_n'(g)q^{n/N},$$

where

$$|a_n'(g)| = \begin{cases} 0 \leq cn^r & \text{if } N \nmid n; \\ |a_{n/N}(g)| \leq c(n/N)^r \leq cn^r & \text{if } N \mid n. \end{cases}$$

Hence, for all $n \in \mathbb{Z}^+$, $|a_n'(g)| \leq cn^r$ as well. Also noting that g is holomorphic at ∞ by definition, this enables us to combine the two results as we did. The "if and only if" nature of the statement is due to only considering 1-periodic functions and noting the remark right after the statement of Proposition 1.2.4 in [Diamond and Shurman (2005)] that the converse of the proposition is also true.

Proposition 2.12. *For all congruence subgroups Γ of $SL_2(\mathbb{Z})$, $M_k(\Gamma)$ is a complex vector space.*

Proof. We verify the vector space axioms.
First, we show closure under addition and scalar multiplication. Let $f, g \in M_k(\Gamma)$, $c \in \mathbb{C}$, and $\gamma \in \Gamma$. Then, by the algebra of holomorphic functions, $cf+g$ satisfies conditions (1) and (2) in Definition 2.8. As for the third condition, by Proposition 2.6, we have $(cf+g)|[\gamma]_k = c \cdot f|[\gamma]_k + g|[\gamma]_k = cf + g$, as desired.
The additive identity is the zero function, and the rest of the axioms follow from the properties of functions. For example, commutativity can be seen from $(f + g)|[\gamma]_k = f + g = g + f = (g + f)|[\gamma]_k$. □

For odd k, if $-I \in \Gamma$, note that $f(z) = f(z)|[-I]_k = -f(z)$ forces that f is the zero function; that is, $\dim(M_k(\Gamma)) = 0$. In general, condition (2) in Definition 2.8 ensures that $M_k(\Gamma)$ has finite dimension. Explicit dimension formulas for even k and odd k can be found in Theorem 3.5.1 and Theorem 3.6.1 in [Diamond and Shurman (2005)], respectively.

The following proposition reduces condition (3) in Definition 2.8 to weight-k invariance under a finite set.

Proposition 2.13. *Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Then, f is weight-k invariant under Γ if and only if f is weight-k invariant under the generators of Γ. Therefore, it suffices to check weight-k invariance under the generators of Γ to confirm condition* (3) *in Definition* 2.8.

Proof. Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Since Γ contains the finite index subgroup $\Gamma(N)$ for some N by definition, Γ has finite index as well. As a finite index subgroup of the finitely generated group $SL_2(\mathbb{Z})$, it follows that Γ is finitely generated. Let $\gamma_1, \gamma_2, \ldots, \gamma_n$ be the generators of Γ. If f is weight-k invariant under Γ, then clearly f is weight-k invariant under $\{\gamma_1, \gamma_2, \ldots, \gamma_n\} \subseteq \Gamma$.
Next, suppose f is weight-k invariant under $\{\gamma_1, \gamma_2, \ldots, \gamma_n\}$. By Proposition 2.7, observe that for $1 \leq i \leq n$, we have

$$f|[\gamma_i^{-1}]_k = (f|[\gamma_i])\,|[\gamma_i^{-1}]_k = f|[\gamma_i\gamma_i^{-1}]_k = f|[I_2]_k = f$$

as well. Define $\gamma_{n+1} = \gamma_1^{-1}, \gamma_{n+2} = \gamma_2^{-1}, \ldots, \gamma_{2n} = \gamma_n^{-1}$. Then, our observation shows that f is weight-k invariant under $\{\gamma_1, \gamma_2, \ldots, \gamma_{2n}\}$.

For all $\gamma \in \Gamma$, we can write $\gamma = \gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_m}$ for some $m \in \mathbb{Z}^+$, where $a_1, \ldots, a_m \in \{1, 2, \ldots, 2n\}$, possibly with repetition.

To show that $f|[\gamma]_k = f$, we will use induction on m. When $m = 1$, we have $f|[\gamma]_k = f|[\gamma_{a_1}]_k = f$, since $\gamma_{a_1} \in \{\gamma_1, \gamma_2, \ldots, \gamma_{2n}\}$. Now, suppose $f|[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_m}]_k = f$ for some $m \in \mathbb{Z}^+$. By Proposition 2.7, we have

$$f|[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_m}\gamma_{a_{m+1}}]_k = (f|[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_m}]_k)|[\gamma_{a_{m+1}}]_k,$$

which equals $f|[\gamma_{a_{m+1}}]_k$ by our inductive hypothesis. Furthermore, since $\gamma_{a_{m+1}} \in \{\gamma_1, \gamma_2, \ldots, \gamma_{2n}\}$, we see that $f|[\gamma_{a_1}\gamma_{a_2}\cdots\gamma_{a_m}\gamma_{a_{m+1}}]_k = f|[\gamma_{a_{m+1}}]_k = f$, as desired.

We've shown that f is weight-k invariant under Γ if and only if f is weight-k invariant under the generators of Γ. □

Proposition 2.14. *Let Γ, Γ' be congruence subgroups of $SL_2(\mathbb{Z})$. If $\Gamma' \subseteq \Gamma$, then $M_k(\Gamma) \subseteq M_k(\Gamma')$.*

Proof. Suppose Γ, Γ' are congruence subgroups of $SL_2(\mathbb{Z})$ such that $\Gamma' \subseteq \Gamma$. Let $f \in M_k(\Gamma)$. Then, f satisfies conditions (1) and (2) in Definition 2.8. Furthermore, $f = f|[\gamma]_k$ for all $\gamma \in \Gamma$. Since $\Gamma' \subseteq \Gamma$, this implies $f = f|[\gamma']_k$ for all $\gamma' \in \Gamma'$. Hence, $f \in M_k(\Gamma')$. We conclude that $M_k(\Gamma) \subseteq M_k(\Gamma')$. □

Lemma 2.15. *The weight-k operator $[\alpha_N]_k$ preserves $M_k(\Gamma_0(N))$.*

Proof. Let $f(z) \in M_k(\Gamma_0(N))$ and let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$. As $z \in \mathcal{H}$, z is never 0 or $i\infty$ so that f being holomorphic implies $f(z)|[\alpha_N]_k = \det(\alpha_N)^{k/2}(Nz)^{-k}f(\alpha_N z) = N^{-k/2}z^{-k}f\left(\frac{-1}{Nz}\right)$ is also holomorphic. Note that as $z \to i\infty$, we have

$$f(z)|[\alpha_N]_k = N^{-k/2}z^{-k}f\left(\frac{-1}{Nz}\right) \sim N^{-k/2}\left(z^{-k}f(-1/z)\right) = N^{-k/2}f(z)|[S]_k\ .$$

Recall that $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in SL_2(\mathbb{Z})$ and that $SL_2(\mathbb{Z})$ is closed under matrix multiplication. Then, as $z \to i\infty$, since $f(z)|[S\alpha]_k = (f(z)|[S]_k)|[\alpha]_k$ is bounded by condition (2) in Definition 2.8 for all $\alpha \in SL_2(\mathbb{Z})$, it follows that $(f(z)|[\alpha_N]_k)|[\alpha]_k \sim N^{-k/2}(f(z)|[S]_k)|[\alpha]_k$ is also bounded.

It now suffices to show that $(f|[\alpha_N]_k)|[\gamma]_k = f|[\alpha_N]_k$. Note that

$$\alpha_N \gamma \alpha_N^{-1} = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1/N \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} d & -c/N \\ -Nb & a \end{pmatrix}.$$

Because $\gamma \in \Gamma_0(N)$, we have $c \equiv 0 \pmod{N}$ so that $-c/N \in \mathbb{Z}$. Also, as conjugation doesn't change the determinant, $\det(\alpha_N \gamma \alpha_N^{-1}) = \det(\gamma) = 1$. Finally, it's clear that $-Nb \equiv 0 \pmod{N}$ and that $a, d \in \mathbb{Z}$. Hence, $\alpha_N \gamma \alpha_N^{-1} \in \Gamma_0(N)$. In particular, this implies $f = f|[\alpha_N \gamma \alpha_N^{-1}]_k$ since $f \in M_k(\Gamma_0(N))$. Then, by Proposition 2.7, we have

$$f|[\alpha_N]_k = (f|[\alpha_N \gamma \alpha_N^{-1}]_k)|[\alpha_N]_k = (f|[\alpha_N]_k)|[\gamma]_k,$$

as desired. □

3 Hecke Operators

3.1 The T_n Operator

Definition 3.1. For $f \in M_k(\Gamma_1(N))$ and $n \in \mathbb{Z}^+$, we define the T_n *Hecke operator* as

$$T_n f := n^{(k/2)-1} \sum f|[\alpha_j]_k,$$

where α_j runs through a set of coset representatives for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$.

We must show that T_n is well-defined, i.e., it does not depend on our choice of coset representatives.
Let $\Gamma_1(N)\alpha_j = \Gamma_1(N)\beta_j$ for some $\alpha_j, \beta_j \in \Delta^n(N, \{1\}, \mathbb{Z})$. Then, $\alpha_j = \gamma\beta_j$ for some $\gamma \in \Gamma_1(N)$. As $f \in M_k(\Gamma_1(N))$, by Proposition 2.7 and Definition 2.8, we have

$$f|[\alpha_j]_k = f|[\gamma\beta_j]_k = (f|[\gamma]_k)|[\beta_j]_k = f|[\beta_j]_k,$$

which implies

$$n^{(k/2)-1} \sum f|[\alpha_j]_k = n^{(k/2)-1} \sum f|[\beta_j]_k,$$

as required.

At the end of the section, we'll show that T_n preserves $M_k(\Gamma_0(N))$. For now, we prove some results about commutativity and Fourier coefficients.

Proposition 3.2. *Let $f \in M_k(\Gamma_0(N))$. Then, $T_n(T_m f) = T_m(T_n f)$ for all $n, m \in \mathbb{Z}^+$ with* $\gcd(n, m) = 1$.

Proof. Let $f \in M_k(\Gamma_0(N)) \subseteq M_k(\Gamma_1(N))$. By Definition 3.1 and Proposition 2.7, we have

$$\begin{aligned}(T_nf)(z) &= n^{k/2-1} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} f(z)|[\sigma_a \begin{pmatrix} a & b \\ 0 & d\end{pmatrix}]_k \\ &= n^{k/2-1} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} (f(z)|[\sigma_a]_k)|[\begin{pmatrix} a & b \\ 0 & d\end{pmatrix}]_k \ .\end{aligned}$$

Since $\sigma_a \in \Gamma_0(N)$ and $f \in M_k(\Gamma_0(N))$, $f(z)|[\sigma_a]_k = f(z)$ so that

$$\begin{aligned}(T_nf)(z) &= n^{k/2-1} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} f(z)|[\begin{pmatrix} a & b \\ 0 & d\end{pmatrix}]_k \\ &= n^{k/2-1} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} n^{k/2}d^{-k} f\left(\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right) \\ &= n^{k/2-1}n^{k/2} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} \left(\frac{n}{a}\right)^{-k} f\left(\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right) \\ &= \frac{1}{n} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} a^k f\left(\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right).\end{aligned}$$

Thus, we have

$$\begin{aligned}(T_m(T_nf))(z) &= \frac{1}{m} \sum_{\substack{\gcd(a',N)=1\\ a'>0,\ a'd'=m\\ 0\le b'<d'}} \left((a')^k \frac{1}{n} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} a^k f\left(\begin{pmatrix} a' & b' \\ 0 & d'\end{pmatrix}\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right)\right) \\ &= \frac{1}{mn} \sum_{\substack{\gcd(a',N)=1\\ a'>0,\ a'd'=m\\ 0\le b'<d'}} \sum_{\substack{\gcd(a,N)=1\\ a>0,\ ad=n\\ 0\le b<d}} (aa')^k f\left(\begin{pmatrix} aa' & a'b+b'd \\ 0 & dd'\end{pmatrix} z\right).\end{aligned}$$

Here, without loss of generality, we remark that the restriction $0 \le b < d$ in the sums can be replaced by the restriction of b being an element of a (fixed) complete residue system modulo d. To see this, note that for all $q \in \mathbb{Z}$, as $T^q = \begin{pmatrix} 1 & q \\ 0 & 1\end{pmatrix} \in \Gamma_0(N)$, $f(T^qz) = f(z)|[T^q]_k = f(z)$. Thus,

$$f\left(\begin{pmatrix} a & b+qd \\ 0 & d\end{pmatrix} z\right) = f\left(T^q\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right) = f\left(\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right).$$

Claim: As a and a' run through the positive divisors of n and m such that $\gcd(a,N) = \gcd(a',N) = 1$, the product aa' runs through the positive divisors of mn. Furthermore, as b and

b' run through the elements of complete residue systems modulo d and d', respectively, the linear combination $a'b + b'd$ runs through the elements of a complete residue system modulo dd'.

Proof: Let

$$A = \{a : \gcd(a, N) = 1, a > 0, a \mid n\},$$
$$B = \{a' : \gcd(a', N) = 1, a' > 0, a' \mid m\},$$
$$C = \{s : \gcd(s, N) = 1, s > 0, s \mid mn\},$$
$$D = \{aa' : a \in A \text{ and } a' \in B\}.$$

Fix $a \in A$ and $a' \in B$; let $d = n/a$, $d' = m/a'$. Define

$$E = \{a'b + b'd : 0 \le b < d \text{ and } 0 \le b' < d'\}.$$

Our claim is equivalent to the assertion that $C = D$ and E is a complete residue system modulo dd'.

To see that $D \subseteq C$, let $aa' \in D$. Then, $\gcd(a, N) = \gcd(a', N) = 1$ implies $\gcd(aa', N) = 1$; $a, a' > 0$ implies $aa' > 0$; and $a \mid n$ and $a' \mid m$ imply $aa' \mid mn$. It follows that $aa' \in C$, which implies $D \subseteq C$.

To see that $C \subseteq D$, let $s \in C$. Furthermore, let $a = \gcd(s, n)$, and $a' = \gcd(s, m)$. Then, $a \mid n$, $a' \mid m$, and $a, a' > 0$.
We were given that $\gcd(n, m) = 1$. Since we also have $a, a' \mid s$ and $\gcd(a, a') = \gcd(s, n, m) = \gcd(s, \gcd(n, m)) = \gcd(s, 1) = 1$, we conclude that $aa' \mid s$. By Bezout's identity, we can write $a = \gcd(s, n) = sx + ny$ and $a' = \gcd(s, m) = sx' + my'$ for some $x, x', y, y' \in \mathbb{Z}$. Thus, $aa' = (sx + ny)(sx' + my') = s^2xx' + sxmy' + sx'ny + mnyy' = s(sxx' + xmy' + x'ny + \frac{mn}{s}yy')$ where $\frac{mn}{s} \in \mathbb{Z}$ as $s \in C$. This implies that $s \mid aa'$, and together with $aa' \mid s$, we conclude that $s = \pm aa'$. Since $s \in C$ implies $s > 0$, and we also have $a, a' > 0$, we can further conclude that $s = aa'$.
Next, without loss of generality, suppose $\gcd(a, N) > 1$. Then, $\gcd(s, N) = \gcd(aa', N) \ge \gcd(a, N) > 1$ contradicts $s \in C$. Hence, $\gcd(a, N) = \gcd(a', N) = 1$.
We've shown that $c \in D$, which implies $C \subseteq D$. Together with our earlier result $D \subseteq C$, we conclude that $C = D$, as desired.

Next, we show that E is a complete residue system modulo dd'. Assume $a'b+b'd = a'\beta+\beta'd$ for some $0 \le b, \beta < d$ and $0 \le b', \beta' < d'$. Then, $a'(b - \beta) = d(\beta' - b')$ implies $b - \beta = t \cdot \operatorname{lcm}(a', d)/a'$ for some $t \in \mathbb{Z}$. But since $\gcd(m, n) = 1 \Rightarrow \gcd(m/d', n/a) = \gcd(a', d) = 1$, we must have $\operatorname{lcm}(a', d) = a'd$ so that $b - \beta = ta'd/a' = td$. The restriction $0 \le b, \beta < d$ forces $t = 0$ so that $b = \beta$.
With the substitution $\beta = b$, we now have $a'b+b'd = a'b+\beta'd$, which implies $b' = \beta'$. We've shown that each pair (b, b') corresponds to a unique element of E. Hence, E has a total of dd' elements.
It remains to show that the elements of E are pairwise incongruent modulo dd'. Assume $a'b+b'd \equiv a'\beta + \beta'd \pmod{dd'}$ for some $0 \le b, \beta < d$ and $0 \le b', \beta' < d'$. Then, as $d \mid dd'$, we also have $a'b + b'd \equiv a'\beta + \beta'd \pmod{d}$, which implies $a'b \equiv a'\beta \pmod{d}$. Because $\gcd(a', d) = 1$, we can divide both sides by a' to obtain $b \equiv \beta \pmod{d}$. As $0 \le b, \beta < d$, this means $b = \beta$.
With the substitution $\beta = b$, we now have $a'b + b'd \equiv a'b + \beta'd \pmod{dd'}$, which implies $b'd \equiv \beta'd \pmod{dd'}$. As $d \ne 0$, we can divide both sides by d to obtain $b' \equiv \beta' \pmod{dd'/\gcd(d, dd')}$, that

is, $b' \equiv \beta' \pmod{d'}$. As $0 \leq b', \beta' < d'$, this means $b' = \beta'$. We've shown that the elements of E are pairwise incongruent.
Overall, we conclude that E is a complete residue system modulo dd', which completes the proof of our claim. △

The claim that we proved implies that

$$(T_m(T_n f))(z) = \frac{1}{mn} \sum_{\substack{\gcd(aa',N)=1 \\ aa'>0,\ (aa')(dd')=mn \\ 0 \leq a'b+b'd < dd'}} (aa')^k f\left(\begin{pmatrix} aa' & a'b+b'd \\ 0 & dd' \end{pmatrix} z\right)$$

$$= (T_{mn}f)(z) = (T_{nm}f)(z) = (T_n(T_m f))(z),$$

proving the proposition. □

The following lemma will be useful in the proof of Proposition 3.4.

Lemma 3.3. *Let $f \in M_k(\Gamma_1(N))$, and let p be a prime such that $p \mid N$. Then,*

$$T_p f = p^{(k/2)-1} \sum_{j=0}^{p-1} f|[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}]_k.$$

Proof. Consider the coset representatives $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ that we found in Lemma 2.4. By Definition 3.1, we have

$$T_p f = p^{(k/2)-1} \sum f|[\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}]_k.$$

Recall the definitions in Lemma 2.4 and substitute $n = p$. Since p is prime, $a > 0$, and $a \mid p$, either $a = 1$ or $a = p$. Suppose $a = p$. Since $p \mid N$, $\gcd(p, N) > 1$, which contradicts the fact that a was defined to be invertible modulo N. Hence, $a = 1$. So $\sigma_a \in \Gamma_0(N)$ such that $\sigma_a \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (mod N). We can fix σ_a to be the identity matrix. Then, from $p = ad$, we have $d = p$, which in turn implies $b \in \{0, 1, 2, \ldots, p-1\}$. We conclude that

$$T_p f = p^{(k/2)-1} \sum_{j=0}^{p-1} f|[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}]_k,$$

as desired. □

Proposition 3.4. *Let $f \in M_k(\Gamma_1(N))$, and let p be a prime divisor of N. Then, $T_p f$ has a Fourier expansion given by*

$$T_p f(z) = \sum_{n=0}^{\infty} a_{np}(f) q^n.$$

In other words, $a_n(T_p f) = a_{np}(f)$.

Proof. By Lemma 3.3, we have

$$\begin{aligned} T_p f(z) &= p^{(k/2)-1} \sum_{j=0}^{p-1} f(z)|[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}]_k \\ &= p^{(k/2)-1} \sum_{j=0}^{p-1} p^{k/2}(0z+p)^{-k} f\left(\frac{z+j}{p}\right) \\ &= p^{(k/2)-1} p^{-k/2} \sum_{j=0}^{p-1} f\left(\frac{z+j}{p}\right) \\ &= \frac{1}{p}\sum_{j=0}^{p-1} f\left(\frac{z+j}{p}\right). \end{aligned}$$

As $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_1(N)$, Remark 2.9 implies that f has a Fourier expansion $f(z) = \sum_{n=0}^{\infty} a_n(f)q^n$, where $q = e^{2\pi i z}$. Then, expressing f as a Fourier series, $T_p f(z)$ equals

$$\frac{1}{p}\sum_{j=0}^{p-1}\left(\sum_{n=0}^{\infty} a_n(f)e^{2\pi i n(z+j)/p}\right) = \frac{1}{p}\sum_{j=0}^{p-1}\left(\sum_{n=0}^{\infty} a_n(f)e^{2\pi i nz/p}e^{2\pi i nj/p}\right) = \frac{1}{p}\sum_{n=0}^{\infty}\left(a_n(f)e^{2\pi i nz/p}\sum_{j=0}^{p-1} e^{2\pi i nj/p}\right).$$

Note that if $p \mid n$, $e^{2\pi i nj/p} = 1$ for all j so that $\sum_{j=0}^{p-1} e^{2\pi i nj/p} = p$. On the other hand, if $p \nmid n$, then $e^{2\pi i n/p} \neq 1$ so that $\sum_{j=0}^{p-1} e^{2\pi i nj/p} = \frac{1-(e^{2\pi i n/p})^p}{1-e^{2\pi i n/p}} = 0$.
Thus,

$$\begin{aligned} T_p f(z) &= \frac{1}{p}\sum_{n=0}^{\infty}\left(a_n(f)e^{2\pi i nz/p}\sum_{j=0}^{p-1} e^{2\pi i nj/p}\right) \\ &= \sum_{\substack{n\geq 0 \\ p|n}} a_n(f)e^{2\pi i nz/p} \\ &= \sum_{n=0}^{\infty} a_{np}(f)q^n. \end{aligned}$$

We've shown that $a_n(T_p f) = a_{np}(f)$. □

Concerning the proof of our main theorem, we only need the following corollary.

Corollary 3.4.1. *If $f \in M_2(\Gamma_0(4))$, then $a_n(T_2 f) = a_{2n}(f)$.*

Proof. Note that $\Gamma_1(4) \subseteq \Gamma_0(4)$ implies $M_2(\Gamma_0(4)) \subseteq M_2(\Gamma_1(4))$ by Proposition 2.14. The result follows from Proposition 3.4 because 2 is a prime divisor of 4. □

Proposition 3.5. *Let $f \in M_k(\Gamma_0(N))$ and $n \in \mathbb{Z}^+$. Considering the Fourier expansions $f = \sum_{m\in\mathbb{Z}} a_m(f)q^m$ and $T_nf = \sum_{m\in\mathbb{Z}} a_m(T_nf)q^m$, for all $m \in \mathbb{Z}$, we have*

$$a_m(T_nf) = \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|\gcd(m,n)}} a^{k-1}a_{nm/a^2}(f).$$

In particular, $a_m(T_nf) = a_{mn}(f)$ whenever $\gcd(m,n) = 1$.

Proof. As in the proof of Proposition 3.2, for $f \in M_k(\Gamma_0(N))$, we have

$$(T_nf)(z) = \frac{1}{n} \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ ad=n \\ 0\leq b<d}} a^k f\left(\begin{pmatrix} a & b \\ 0 & d\end{pmatrix} z\right) = \frac{1}{n} \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ ad=n \\ 0\leq b<d}} a^k f\left(\frac{az+b}{d}\right).$$

Since $T \in \Gamma_0(N)$, Remark 2.9 shows that expressing f as a Fourier series gives

$$\begin{aligned}(T_nf)(z) = \frac{1}{n} \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ ad=n \\ 0\leq b<d}} a^k \sum_{n'=0}^{\infty} \left(a_{n'}(f)e^{2\pi in'(az+b)/d}\right) &= \frac{1}{n} \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ ad=n \\ 0\leq b<d}} a^k \sum_{n'=0}^{\infty} \left(a_{n'}(f)e^{2\pi in'az/d}e^{2\pi in'b/d}\right) \\ &= \frac{1}{n} \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ ad=n \\ n'\in\mathbb{Z}}} \left(a^k a_{n'}(f)e^{2\pi in'az/d} \sum_{b=0}^{d-1} e^{2\pi in'b/d}\right).\end{aligned}$$

Note that if $d \mid n'$, $e^{2\pi in'b/d} = 1$ for all b so that $\sum_{b=0}^{d-1} e^{2\pi in'b/d} = d$. On the other hand, if $d \nmid n'$, then $e^{2\pi in'/d} \neq 1$ so that $\sum_{b=0}^{d-1} e^{2\pi in'b/d} = \sum_{b=0}^{d-1}(e^{2\pi in'/d})^b = \frac{1-(e^{2\pi in'/d})^d}{1-e^{2\pi in'/d}} = 0$.

Hence, we may only consider the case $d|n'$. Writing $n' = dm'$, we have

$$(T_nf)(z) = \frac{1}{n} \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ ad=n \\ m'\in\mathbb{Z}}} \left(da^k a_{dm'}(f)e^{2\pi im'az}\right).$$

Substituting $d = n/a$, this equals

$$\sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|n \\ m'\in\mathbb{Z}}} \left(a^{k-1}a_{nm'/a}(f)q^{m'a}\right).$$

We make a final change of variables $m = am'$ to obtain

$$(T_nf)(z) = \sum_{\substack{\gcd(a,N)=1 \\ a|n,\ a|m \\ a>0,\ m\in\mathbb{Z}}} \left(a^{k-1}a_{nm/a^2}(f)q^m\right) = \sum_{m=0}^{\infty}\left(q^m \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|\gcd(m,n)}} a^{k-1}a_{nm/a^2}(f)\right).$$

Hence, for all $m \in \mathbb{Z}$, we have

$$a_m(T_n f) = \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|\gcd(m,n)}} a^{k-1} a_{nm/a^2}(f).$$

(When $m < 0$, $a_m(T_n f) = a_{nm/a^2}(f) = 0$ as both $T_n f$ and f are modular forms.)

In particular, when $\gcd(m,n) = 1$, $\{a \in \mathbb{Z}^+ : a \mid \gcd(m,n)\} = \{1\}$ implies

$$a_m(T_n f) = 1^{k-1} a_{nm/1^2}(f) = a_{mn}(f). \qquad \square$$

Proposition 3.6. *The Hecke operator T_n preserves $M_k(\Gamma_0(N))$.*

Proof. First, we'll show that $T_n f$ is weight-k invariant under $\Gamma_0(N)$.
Let α be a coset representative for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$. For any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, note that $\det \gamma = ad - bc \equiv 1 \pmod N$ and $c \equiv 0 \pmod N$ imply $ad \equiv 1 \pmod N$. Thus, we have

$$\gamma^{-1}\alpha\gamma \equiv \begin{pmatrix} d & -b \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & * \\ 0 & n \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \equiv \begin{pmatrix} d & * \\ 0 & an \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \equiv \begin{pmatrix} ad & * \\ 0 & adn \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & n \end{pmatrix} \pmod N.$$

Since we also have $\det(\gamma^{-1}\alpha\gamma) = \det(\gamma^{-1}) \cdot \det(\alpha) \cdot \det(\gamma) = 1 \cdot n \cdot 1 = n$, we conclude that $\gamma^{-1}\alpha\gamma \in \Delta^n(N, \{1\}, \mathbb{Z})$ for all $\gamma \in \Gamma_0(N)$. Noticing that $\Gamma_1(N) = \Delta^1(N, \{1\}, \mathbb{Z})$, these calculations also show that $\Gamma_1(N)$ is normal in $\Gamma_0(N)$.

Suppose $\gamma^{-1}\alpha\gamma\Gamma_1(N) = \gamma^{-1}\alpha'\gamma\Gamma_1(N)$ for some $\alpha' \in \Delta^n(N, \{1\}, \mathbb{Z})$. Then, there exists a $\gamma_1 \in \Gamma_1(N)$ such that $\alpha\gamma\gamma_1 = \alpha'\gamma$. Multiplying on the right by γ^{-1}, we get $\alpha\gamma\gamma_1\gamma^{-1} = \alpha'$. Since $\Gamma_1(N)$ is normal in $\Gamma_0(N)$, $\gamma\gamma_1\gamma^{-1} \in \Gamma_1(N)$ so that $\alpha\Gamma_1(N) = \alpha'\Gamma_1(N)$. This shows that α' is another representative for the same coset as α. So if $\{\alpha_j\}$ is a set of coset representatives for $\Gamma_1(N)$, then $\{\gamma^{-1}\alpha_j\gamma\}$ is a set of representatives as well. Thus, by Proposition 2.6 and Proposition 2.7,

$$(T_n f)|[\gamma]_k = \left(n^{(k/2)-1} \sum f|[\alpha_j]_k\right) |[\gamma]_k = n^{(k/2)-1} \sum f|[\gamma^{-1}\alpha_j\gamma]_k = T_n f,$$

as desired.

Next, by Proposition 3.5, we can write $T_n f = \sum_{m=0}^{\infty} a_m(T_n f) q^m$ where

$$a_m(T_n f) = \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|\gcd(m,n)}} a^{k-1} a_{nm/a^2}(f).$$

As $n \in \mathbb{Z}^+$ is fixed and a goes through the elements of a subset of the set of divisors of n, we always have $1 \leq a \leq n$ so that $a^{k-1} \leq n^{k-1}$. Furthermore, as $T \in \Gamma_0(N)$, f has period 1 by Remark 2.9 so that Lemma 2.10 implies $|a_{nm/a^2}(f)| \leq c(nm/a^2)^r = \frac{cn^r}{a^{2r}} m^r \leq (cn^r)m^r$ for some positive constants c and r. Let $N \in \mathbb{Z}^+$ denote the number of terms in the sum. Then, if we let C be a new constant such that $C = Nn^{k-1}cn^r > 0$, we see that

$$a_m(T_n f) = \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|\gcd(m,n)}} a^{k-1} a_{nm/a^2}(f) \leq \sum_{\substack{\gcd(a,N)=1 \\ a>0,\ a|\gcd(m,n)}} n^{k-1}(cn^r)m^r = Nn^{k-1}cn^r m^r = Cm^r,$$

which shows that $T_n f \in M_k(\Gamma_0(N))$ by Lemma 2.10. □

The following is an immediate application of Proposition 2.6 and Proposition 3.6.

Corollary 3.6.1. *The Hecke operator T_n is a linear operator from $M_k(\Gamma_0(N))$ to $M_k(\Gamma_0(N))$.*

Proposition 3.7. *For $f \in M_k(\Gamma_0(N))$, T_n commutes with $\alpha_N = \left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$ whenever* $\gcd(n, N) = 1$.

Proof. Let $f \in M_k(\Gamma_0(N)) \subseteq M_k(\Gamma_1(N))$. Lemma 2.4 tells us that $\alpha_{a,b}$ (which is defined only when $\gcd(n, N) = 1$) form a complete set of coset representatives for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$. Then, by our definition of $T_n f$ and Proposition 2.7, we have

$$(T_n f)|[\alpha_N]_k = n^{(k/2)-1} \sum_{a,b} (f|[\alpha_{a,b}]_k)|[\alpha_N]_k = n^{(k/2)-1} \sum_{a,b} f|[\alpha_{a,b}\alpha_N]_k$$

$$= n^{(k/2)-1} \sum_{a,b} f|[\sigma_n \alpha_N \sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}]_k = n^{(k/2)-1} \sum_{a,b} f|[\alpha_N \sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}]_k$$

since $f|[\sigma_n]_k = f$ as $\sigma_n \in \Gamma_0(N)$.

On the other hand, Lemma 2.4 also tells us that $\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ form a complete set of coset representatives for $\Gamma_1(N)$ in $\Delta^n(N, \{1\}, \mathbb{Z})$. Note that $f|[\alpha_N]_k \in M_k(\Gamma_0(N)) \subseteq M_k(\Gamma_1(N))$ by Lemma 2.15. Thus, we have

$$T_n(f|[\alpha_N]_k) = n^{(k/2)-1} \sum_{a,b} (f|[\alpha_N]_k)|[\sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}]_k = \sum_{a,b} f|[\alpha_N \sigma_a \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}]_k = (T_n f)|[\alpha_N]_k \ ,$$

as desired. □

4 Main Theorem

In this section, we will prove our main theorem:

Main Theorem (Jacobi's Four Squares Theorem). *For any positive integer n, let a_n denote the number of ways n can be expressed as a sum of four integral squares. Then,*

$$a_n = \begin{cases} 8\sigma(n) & \textit{for } n \textit{ odd,} \\ 24\sigma(n_0) & \textit{for } n = 2^r n_0 \textit{ even, } n_0 \textit{ odd.} \end{cases}$$

As an immediate corollary of our main theorem, we obtain Lagrange's Theorem:

Corollary (Lagrange's Four Squares Theorem). *Every nonnegative integer can be expressed as a sum of four integral squares.*

Our strategy is to consider a modular form that is the generating series for the number of ways an integer n can be expressed as a sum of four squares.

We define the *theta-function*:

$$\Theta := \sum_{n\in\mathbb{Z}} q^{n^2} \quad \text{for } z \in \mathcal{H}, \quad q = e^{2\pi i z},$$

and observe that

$$\Theta^4 = \sum_{n_1,n_2,n_3,n_4\in\mathbb{Z}} q^{{n_1}^2} q^{{n_2}^2} q^{{n_3}^2} q^{{n_4}^2} = \sum_{n_1,n_2,n_3,n_4\in\mathbb{Z}} q^{{n_1}^2+{n_2}^2+{n_3}^2+{n_4}^2} = \sum_{n\in\mathbb{Z}} a_n q^n,$$

where a_n is the number of ways n can be written as a sum of four squares. Hence, it will suffice to confirm our formula for the n^{th} Fourier coefficient of Θ^4.

Proposition 4.1. $\Theta^4 \in M_2(\Gamma_0(4))$.

Proof. First, we show weight-2 invariance under $\Gamma_0(4)$. By Proposition 2.13, it suffices to check weight-2 invariance under the generators $-I$, T, and ST^4S of $\Gamma_0(4)$ that we found in Proposition 2.3.

To show invariance under $[-I]_2$, we check that $\Theta^4(z)|[-I]_2 = (-1)^{-2}\Theta^4(z) = \Theta^4(z)$.

Next, to show invariance under $[T]_2$, we have

$$\Theta^4(z)|[T]_2 = (1)^{-2}\Theta^4(z+1) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n(z+1)} = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n z} e^{2\pi i n} = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n z} = \Theta^4(z).$$

Finally, to show invariance under $[ST^4S]_2$, we begin by observing that $ST^4S = \frac{1}{4}\alpha_4 T \alpha_4$, which follows directly from matrix multiplication. Then,

$$\Theta^4(z)|[ST^4S]_2 = \Theta^4(z)|[\frac{1}{4}\alpha_4 T\alpha_4]_2 = (\Theta^4(z)|[\frac{1}{4}I]_2)|[\alpha_4 T\alpha_4]_2$$

by Proposition 2.7. Note that $\Theta^4(z)|[\frac{1}{4}I]_2 = \det(\frac{1}{4}I)^{2/2}(\frac{1}{4})^{-2}\Theta^4(\frac{z/4}{1/4}) = \Theta^4(z)$. Thus, we have

$$\Theta^4(z)|[ST^4S]_2 = \Theta^4(z)|[\alpha_4 T\alpha_4]_2 = ((\Theta^4(z)|[\alpha_4]_2)|[T]_2)|[\alpha_4]_2.$$

Equation (3.5) on page 124 in [Koblitz (1993)] tells us that

$$\Theta^2|[\alpha_4]_1 = -i\Theta^2. \tag{4}$$

Squaring the left side of equation (4), we get

$$(\Theta^2|[\alpha_4]_1)^2 = (\det(\alpha_4)^{1/2}(4z)^{-1}\Theta^2(\alpha_4 z))^2 = \det(\alpha_4)(4z)^{-2}\Theta^4(\alpha_4 z) = \Theta^4|[\alpha_4]_2.$$

Squaring the right side of equation (4), we get $(-i\Theta^2)^2 = -\Theta^4$. Thus, we have obtained the identity

$$\Theta^4|[\alpha_4]_2 = -\Theta^4. \tag{5}$$

This means that Θ^4 is an eigenform (with associated eigenvalue -1) of $[\alpha_4]$, which will be a useful piece of information going forward.

Back to our problem at hand, we have

$$\Theta^4(z)|[ST^4S]_2 = ((\Theta^4(z)|[\alpha_4]_2)|[T]_2)|[\alpha_4]_2 = ((-\Theta^4)|[T]_2)|[\alpha_4]_2$$

by equation (5). Since we've shown that Θ^4 is invariant under $[T]_2$, this equals $-\Theta^4|[\alpha_4]_2$, which is again Θ^4 by equation (5). We've shown that $\Theta^4(z)|[ST^4S]_2 = \Theta^4(z)$, as desired.

Next, observe that as negative integers cannot be expressed as a sum of four squares, we can write

$$\Theta^4 = \sum_{n\in\mathbb{Z}} a_n q^n = \sum_{n=0}^{\infty} a_n q^n.$$

Given $n \in \mathbb{Z}^+$, for n to be expressed as a sum of four squares $a^2 + b^2 + c^2 + d^2$, it must be that $-n < a, b, c, d \leq n$ (in fact, we could say $-\sqrt{n} \leq a, b, c, d \leq \sqrt{n}$). There are $(2n)^4$ possible ways to select 4 numbers a, b, c, d in this range, possibly with repetition, and taking order into account. As these $(2n)^4$ possible selections need not all satisfy $a^2 + b^2 + c^2 + d^2 = n$, we see that $|a_n| \leq 16n^4$. By Lemma 2.10, we conclude that $\Theta^4 \in M_2(\Gamma_0(4))$. □

We'd like to know more about the space $M_2(\Gamma_0(4))$ that contains our generating series Θ^4. We now find another function in $M_2(\Gamma_0(4))$, which will turn out to form a basis for the space together with Θ^4.

Proposition 4.2. *The function F given by $F := \sum_{odd\ n>0} \sigma(n)q^n$ lies in $M_2(\Gamma_0(4))$.*

Proof. We begin by checking weight-2 invariance under the generators $-I$, T, and ST^4S of $\Gamma_0(4)$.

Invariance under $[-I]_2$ is trivial as $F(z)|[-I]_2 = (-1)^{-2}F(z) = F(z)$. Invariance under $[T]_2$ is also easily seen as

$$F(z)|[T]_2 = F(z+1) = \sum_{\text{odd } n>0} \sigma(n)e^{2\pi i n(z+1)} = \sum_{\text{odd } n>0} \sigma(n)e^{2\pi i n z}e^{2\pi i n} = \sum_{\text{odd } n>0} \sigma(n)e^{2\pi i n z} = F(z).$$

To show invariance under $[ST^4S]_2$, we need a few lemmas.

We define the *Eisenstein series of weight 2* as follows:

$$E_2(z) := 1 - 24\sum_{n=1}^{\infty} \sigma(n)q^n.$$

Lemma 4.3. $F(z) = -\frac{1}{24}(E_2(z) - 3E_2(2z) + 2E_2(4z))$

Proof. By our definition of $E_2(z)$, we have

$$\begin{aligned}
E_2(z) - 3E_2(2z) + 2E_2(4z) &= 1 - 24\left(\sum_{n=1}^{\infty}\sigma(n)e^{2\pi inz}\right) - 3\left(1 - 24\sum_{n=1}^{\infty}\sigma(n)e^{4\pi inz}\right) \\
&\quad + 2\left(1 - 24\sum_{n=1}^{\infty}\sigma(n)e^{8\pi inz}\right) \\
&= -24\left(\left(\sum_{n=1}^{\infty}\sigma(n)e^{2\pi inz}\right) - 3\left(\sum_{n=1}^{\infty}\sigma(n)e^{4\pi inz}\right) + 2\sum_{n=1}^{\infty}\sigma(n)e^{8\pi inz}\right) \\
&= -24\left(\left(\sum_{n=1}^{\infty}\sigma(n)q^n\right) - 3\left(\sum_{n>0,\ 2|n}\sigma(n/2)q^n\right) + 2\sum_{n>0,\ 4|n}\sigma(n/4)q^n\right) .
\end{aligned}$$

Thus, the coefficient of q^n is given by

$$a_n = \begin{cases} -24(\sigma(n) - 3\sigma(n/2) + 2\sigma(n/4)) & \text{if } 4 \mid n \ , \\ -24(\sigma(n) - 3\sigma(n/2) & \text{if } 2 \mid n \text{ and } 4 \nmid n \ , \\ -24\sigma(n) & \text{if } 2 \nmid n \ . \end{cases}$$

If $4 \mid n$, write $n = 2^t r$ where $2 \nmid r$ and $t > 1$. Then,

$$a_n = -24(\sigma(2^t r) - 3\sigma(2^{t-1}r) + 2\sigma(2^{t-2}r)).$$

As σ is multiplicative and $\gcd(2, r) = 1$, this equals

$$\begin{aligned}
-24\sigma(r)(\sigma(2^t) - 3\sigma(2^{t-1}) + 2\sigma(2^{t-2}) &= -24\sigma(r)\left(\sum_{j=0}^{t}2^j - 3\sum_{j=0}^{t-1}2^j + 2\sum_{j=0}^{t-2}2^j\right) \\
= -24\sigma(r)\left(\left(3\sum_{j=0}^{t-2}2^j - 3\sum_{j=0}^{t-2}2^j\right) + 2^{t-1} + 2^t - 3(2^{t-1})\right) &= 0.
\end{aligned}$$

If $2 \mid n$ and $4 \nmid n$, write $n = 2r$ where $2 \nmid r$. Then,

$$a_n = -24(\sigma(2r) - 3\sigma(r)) = -24\sigma(r)(\sigma(2) - 3\sigma(1)) = -24\sigma(r)(3 - 3) = 0.$$

We conclude that $F(z) = \sum_{\text{odd } n>0}\sigma(n)q^n = -\frac{1}{24}(E_2(z) - 3E_2(2z) + 2E_2(4z))$. □

Lemma 4.4. *Let $g(z) = f(nz)$ for some integer n, and let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$. Then, $g(z)|[\gamma]_k = f(nz)|[\left(\begin{smallmatrix} a & nb \\ c/n & d \end{smallmatrix}\right)]_k$.*

Proof. We have

$$\begin{aligned}
f(nz)|[\left(\begin{smallmatrix} a & nb \\ c/n & d \end{smallmatrix}\right)]_k &= \left(\frac{c}{n}(nz) + d\right)^{-k} f\left(\frac{anz + nb}{\frac{c}{n}(nz) + d}\right) \\
&= (cz + d)^{-k} f\left(n \cdot \frac{az + b}{cz + d}\right) \\
&= (cz + d)^{-k} g\left(\frac{az + b}{cz + d}\right) \\
&= g(z)|[\gamma]_k.
\end{aligned}$$

□

Lemma 4.5. *For all $a \in \mathbb{Z}$, $E_2(ST^{-a}Sz) = (az+1)^2 E_2(z) - \frac{6ai}{\pi}(az+1)$.*

Proof. By Proposition 7 in [Koblitz (1993)], we have

$$E_2(ST^{-a}Sz) = E_2(S(-a-1/z)) = E_2\left(\frac{-1}{-a-1/z}\right) = (a+1/z)^2 E_2(-a-1/z) + \frac{6}{\pi i}(-a-1/z)\ .$$

Since E_2 has period 1, $E_2(-a-1/z) = E_2(-1/z) = E_2(Sz)$, so

$$\begin{aligned} E_2(ST^{-a}Sz) &= (a+1/z)^2\left(z^2 E_2(z) + \frac{6z}{\pi i}\right) + \frac{6}{\pi i}(-a-1/z) \\ &= (az+1)^2 E_2(z) + \frac{6(az+1)^2}{\pi i z} - \frac{6(az+1)}{\pi i z} \\ &= (az+1)^2 E_2(z) + \frac{6(az+1)(az)}{\pi i z} \\ &= (az+1)^2 E_2(z) - \frac{6ai}{\pi}(az+1). \end{aligned}$$

□

By Lemma 4.3 and Lemma 4.4, as well as noting that $ST^4S = \left(\begin{smallmatrix} -1 & 0 \\ 4 & -1 \end{smallmatrix}\right)$, we have

$$\begin{aligned} -24F(z)|[ST^4S]_2 &= E_2(z)|[ST^4S]_2 - 3E_2(2z)|[\left(\begin{smallmatrix} -1 & 2(0) \\ 4/2 & -1 \end{smallmatrix}\right)]_2 + 2E_2(4z)|[\left(\begin{smallmatrix} -1 & 4(0) \\ 4/4 & -1 \end{smallmatrix}\right)]_2 \\ &= E_2(z)|[ST^4S]_2 - 3E_2(2z)|[ST^2S]_2 + 2E_2(4z)|[STS]_2. \end{aligned}$$

By Lemma 4.5, this is equal to

$$\begin{aligned} &\frac{1}{(4z-1)^2}\left((-4z+1)^2E_2(z) + \frac{24i}{\pi}(-4z+1)\right) - 3\frac{1}{(4z-1)^2}\left((-4z+1)^2E_2(2z) + \frac{12i}{\pi}(-4z+1)\right) \\ &+ 2\frac{1}{(4z-1)^2}\left((-4z+1)^2E_2(4z) + \frac{6i}{\pi}(-4z+1)\right) \\ &= E_2(z) - \frac{24i}{\pi}\frac{1}{4z-1} - 3\left(E_2(2z) - \frac{12i}{\pi}\frac{1}{4z-1}\right) + 2\left(E_2(4z) - \frac{6i}{\pi}\frac{1}{4z-1}\right) \\ &= E_2(z) - 3E_2(2z) + 2E_2(4z) - \frac{24i}{\pi(4z-1)} + \frac{36i}{\pi(4z-1)} - \frac{12i}{\pi(4z-1)} \\ &= -24F(z), \end{aligned}$$

as desired.

We can write $F(z) = \sum_{n=0}^{\infty} a_n q^n$, where $|a_n| = a_n = 0 \le 1 \cdot n^2$ for all even n, and

$$|a_n| = a_n = \sigma(n) \le 1 + 2 + \cdots + n = \frac{n(n+1)}{2} = \frac{n^2+n}{2} \le \frac{2n^2}{2} = 1 \cdot n^2$$

for all odd $n \in \mathbb{Z}^+$. By Lemma 2.10, we conclude that $F \in M_2(\Gamma_0(4))$. □

Proposition 4.6. *The set $\{\Theta^4, F\}$ is a basis for $M_2(\Gamma_0(4))$.*

Proof. A computation in SAGE shows that $M_2(\Gamma_0(4))$ is two-dimensional [SageMath]. It therefore suffices to show that $\Theta^4 \in M_2(\Gamma_0(4))$ and $F \in M_2(\Gamma_0(4))$ are linearly independent. Suppose $a\Theta^4(z) = bF(z)$ for some $a, b \in \mathbb{C}$. As $z \to i\infty$, note that $a\Theta^4(z) = a(a_0(\Theta^4)) = a(1) = a$ and $bF(z) = b(a_0(F)) = b(0) = 0$ give $a = 0$. Since $F(z)$ is not the zero function, $b = 0$ as well. Hence, Θ^4 and F are linearly independent, and we conclude that $\{\Theta^4, F\}$ is a basis for $M_2(\Gamma_0(4))$. □

Remark 4.7. The dimension of $M_2(\Gamma_0(4))$ can also be calculated by using Theorem 3.5.1 in [Diamond and Shurman (2005)], though this formula is rather complicated and requires the introduction of new terminology.

Proposition 4.8. *$T_nF = \sigma(n)F$ for all odd $n > 0$.*

Proof. As $F \in M_2(\Gamma_0(4))$, Corollary 3.4.1 shows that $a_n(T_2F) = a_{2n}(F)$. Since $a_{2n}(F) = 0$ for all n by the definition of F, we conclude that $T_2F = 0$; that is, F is an eigenform for T_2 with associated eigenvalue 0.

Next, as T_2 is a linear operator from $M_2(\Gamma_0(4))$ to $M_2(\Gamma_0(4))$ by Corollary 3.6.1, we can write

$$T_2(16F + \Theta^4) = 16T_2F + T_2\Theta^4 = 0 + T_2\Theta^4 \ .$$

We know that F and Θ^4 span $M_2(\Gamma_0(4))$ by Proposition 4.6. Since $T_2\Theta^4 \in M_2(\Gamma_0(4))$, we can write $T_2\Theta^4 = a\Theta^4 + bF$ for some $a, b \in \mathbb{C}$. And since $\Theta^4 \in M_2(\Gamma_0(4))$, we have $a_0(T_2\Theta^4) = a_0(\Theta^4)$ by Corollary 3.4.1, whereas $a_0(F) = 0$. Hence, $a = 1$.
Next, we compare the a_1^{th} coefficients to see that $a_1(T_2\Theta^4) = a_2(\Theta^4) = 24$, whereas $a_1(\Theta^4) = 8$ and $a_1(F) = 1$. Thus, we have $8a + 1b = 8 + b = 24$, implying $b = 16$.
We have shown that

$$T_2(16F + \Theta^4) = T_2\Theta^4 = 16F + \Theta^4 \ . \tag{6}$$

Hence, $16F + \Theta^4$ is an eigenform for T_2 with associated eigenvalue 1.

For odd n, Proposition 3.2 shows that T_n commutes with T_2 as $\gcd(2, n) = 1$. This means $T_nT_2f = T_2T_nf$ for all $f \in M_2(\Gamma_0(4))$. Let f be an eigenform of T_2 so that $T_2f = \lambda_2 f$ for some $\lambda_2 \in \mathbb{C}$. Then,

$$T_2(T_nf) = T_n(T_2f) = T_n(\lambda_2 f) = \lambda_2(T_nf) \ .$$

In other words, if f is an eigenform of T_2, so is T_nf, and with the same associated eigenvalue.
Since $M_2(\Gamma_0(4))$ is two-dimensional and T_2 has two distinct eigenvalues 0 and 1, it follows that T_2 has two one-dimensional eigenspaces associated with the eigenvalues 0 and 1. Hence, the eigenspace of T_2 associated with λ_2 is one-dimensional, meaning $T_nf = \lambda_n f$ for some $\lambda_n \in \mathbb{C}$. We've shown that any eigenform of T_2 is also an eigenform of T_n, where n is odd.

In particular, since F is an eigenform for T_2, it follows that F is an eigenform for all T_n when n is odd. Then, for a given odd n, $T_nF = \lambda_nF$ for some constant λ_n. By Proposition 3.5, $a_1(T_nF) = a_n(F)$. We also know that $a_1(T_nF) = a_1(\lambda_nF) = \lambda_n a_1(F)$, so $a_n(F) = \lambda_n a_1(F)$. This means

$$\sigma(n) = \lambda_n\sigma(1) = \lambda_n \ .$$

□

Proposition 4.9. *$T_n\Theta^4 = \sigma(n)\Theta^4$ for odd $n > 0$.*

Proof. Let $n > 0$ be odd. We'll first show that Θ^4 is an eigenform for T_n. Equation (5) shows that Θ^4 is an eigenform of $[\alpha_4]$ with associated eigenvalue -1. Using the identity $F(z)|[\alpha_4]_2 = -\frac{1}{16}\Theta^4 + F$ (see Exercise 15(c) in [Koblitz (1993)]), we also have

$$(\Theta^4 - 32F)|[\alpha_4]_2 = \Theta^4|[\alpha_4]_2 - 32F|[\alpha_4]_2 = -\Theta^4 - 32\left(-\frac{1}{16}\Theta^4 + F\right) = \Theta^4 - 32F \ ,$$

showing that $\Theta^4 - 32F$ is another eigenform of $[\alpha_4]$ with associated eigenvalue 1.

By Proposition 3.7, $\gcd(n, 4) = 1$ implies T_n commutes with $[\alpha_4]$. As in the proof of Proposition 4.8, this means that if f is an eigenvector of $[\alpha_4]$, so is $T_n f$, and with the same associated eigenvalue. Furthermore, $[\alpha_4]$ having two distinct eigenvalues implies that any eigenform of $[\alpha_4]$ is also an eigenform of T_n. In particular, Θ^4 must be an eigenform of T_n.

Next, equation (6) shows that $\Theta^4 + 16F$ is an eigenform for T_2. Hence, by our proof of Proposition 4.8, it's also an eigenform for T_n for all odd n. Thus,

$$T_n(\Theta^4 + 16F) = \lambda(\Theta^4 + 16F)$$

for some $\lambda \in \mathbb{C}$. Then,

$$T_n\Theta^4 + 16T_nF = T_n\Theta^4 + 16\sigma(n)F = \lambda\Theta^4 + 16\lambda F$$

$$\Rightarrow \lambda\Theta^4 - T_n\Theta^4 = 16\sigma(n)F - 16\lambda F \ .$$

Since Θ^4 is an eigenform for T_n, $T_n\Theta^4 = \lambda'\Theta^4$ for some $\lambda' \in \mathbb{R}$ so that we have

$$(\lambda - \lambda')\Theta^4 = 16(\sigma(n) - \lambda)F \ .$$

As Θ^4 and F are linearly independent by Proposition 4.6, we must have $\lambda = \lambda'$ and $\sigma(n) = \lambda$. It follows that $\lambda = \lambda' = \sigma(n)$ and that $T_n\Theta^4 = \sigma(n)\Theta^4$. □

4.1 Proof of the Main Theorem

We now have all the pieces needed to complete the proof of our Main Theorem.

By Proposition 3.5, $a_n(\Theta^4) = a_1(T_n\Theta^4) = a_1(\sigma(n)\Theta^4) = \sigma(n)a_1(\Theta^4) = 8\sigma(n)$ when n is odd.

Next, note that when $n > 0$ is even, $a_n(\Theta^4 + 16F) = a_n(\Theta^4) + 16a_n(F) = a_n(\Theta^4)$ since $a_n(F) = 0$ for all even n by the definition of F. On the other hand, since $a_n(F) = \sigma(n)$ for odd $n > 0$, we get $a_n(\Theta^4 + 16F) = a_n(\Theta^4) + 16\sigma(n)$ when n is odd. Thus,

$$a_n(\Theta^4 + 16F) = \begin{cases} a_n(\Theta^4) & \text{if } n \text{ is even,} \\ a_n(\Theta^4) + 16\sigma(n) & \text{if } n \text{ is odd.} \end{cases} \tag{7}$$

Recall that $T_2\Theta^4 = \Theta^4 + 16F$ by equation (6), and compare this to

$$a_n(T_2\Theta^4) = a_{2n}(\Theta^4) \ . \tag{8}$$

Claim: $a_n(\Theta^4) = 24\sigma(n_0)$ for $n = 2^r n_0$ with $r \in \mathbb{Z}^+$ and n_0 odd.

Proof: We will use induction on r. When $r = 1$, equation (8) gives us $a_{2^1 n_0}(\Theta^4) = a_{n_0}(T_2\Theta^4)$. Since $T_2\Theta^4 = \Theta^4 + 16F$, we have $a_{n_0}(T_2\Theta^4) = a_{n_0}(\Theta^4 + 16F)$. From equation (7), we see that this equals $a_{n_0}(\Theta^4) + 16\sigma(n_0)$ when n_0 is odd. Finally, as we have shown that $a_{n_0}(\Theta^4) = 8\sigma(n_0)$, we have

$$a_{2n_0}(\Theta^4) = 8\sigma(n_0) + 16\sigma(n_0) = 24\sigma(n_0)\ .$$

Now, suppose the claim is true for some $r \in \mathbb{Z}^+$, and consider $a_{2^{r+1}n_0}(\Theta^4)$. Following the arguments in our base case, we have

$$a_{2^{r+1}n_0}(\Theta^4) = a_{2^r n_0}(T_2\Theta^4) = a_{2^r n_0}(\Theta^4 + 16F) = a_{2^r n_0}(\Theta^4)\ ,$$

which equals $24\sigma(n_0)$ by our inductive hypothesis, as desired. △

We've obtained the following formula for $a_n(\Theta^4)$, which is the number of ways n can be expressed as a sum of four integral squares:

$$a_n = \begin{cases} 8\sigma(n) & \text{for } n \text{ odd,} \\ 24\sigma(n_0) & \text{for } n = 2^r n_0 \text{ even, } n_0 \text{ odd.} \end{cases}$$

References

[Diamond and Shurman (2005)] F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Mathematics **228**, Springer, New York, 2005.

[Koblitz (1993)] N. Koblitz, *Introduction to elliptic curves and modular forms*, 2nd ed., Graduate Texts in Mathematics **41**, Springer, New York, 1990. MR 90j:11001 Zb1 0697.10023

[Miyake 2006] T. Miyake, *Modular forms*, 2nd ed., Springer, Berlin, 2006. Translated from the 1976 Japanese original by Yoshitaka Maeda. MR 2006g:11084 Zbl 1159.11014

[SageMath] SageMath, the Sage Mathematics Software System (Version 9.5), The Sage Developers, 2022, https://www.sagemath.org.

[Weil 1983] A. Weil, *Number Theory*, Birkhäuser, Boston, 1983.