INTRODUCTION TO MODULAR FORMS

ALEXANDRE ACRA

Abstract. The subject of Modular Forms is a rich sub-area of complex analysis that reveals powerful and surprising results in Number Theory. In this paper, we will introduce Modular Forms as holomorphic functions on the upper half of the complex plane $\mathbb{H} \subset \mathbb{C}$, exhibiting specific regularity conditions (of "modularity") and some asymptotic behavior at infinity in the imaginary direction.

The main angle of the paper is to illustrate and enjoy how Abstract Algebra (e.g., group theory on matrices), Linear Algebra (vector spaces and dimensions), Complex Analysis (elliptic functions and lattices on $\mathbb{C}$, Eisenstein series and Fourier analysis, holomorphicity, summation techniques and convergence of series) come together to deliver powerful results in Number Theory, such as relationships between sums of divisors of integers, and the Four Square Theorem which counts the ways in which any integer can be written as a sum of four squares.

The paper assumes sufficient background in Complex Analysis and Holomorphic and Elliptic Functions, as it is the capstone of a course on these topics, but we have attempted to make it self-contained when it comes to definitions and results from other domains, notably from Abstract and Linear Algebra. This has led to significant sections on necessary preliminaries.

We will also discuss examples and benefits of broadening the topic to include Mock Modular Forms, and will end with an introduction to Hecke Operators and discuss how they led to proofs of some conjectures posited by Ramanujan.

1. Historical Background

By the late nineteenth century, the field of Number Theory had already evolved to make advances by using non-elementary techniques, i.e., by leveraging other fields of mathematics including real analysis (e.g., Dirichlet's use of the pigeonhole principle) and the use of infinite sums and products (e.g., by Euler then Gauss).

By that time, the use of Complex Analysis towards Number Theory was also well underway, such as with the work of Klein and Jacoby on Elliptic Curves, and of course Riemann's work and insights in leveraging the ζ function that now bears his name and the corresponding Riemann Hypothesis about zeros of the ζ function and the distribution of the primes.

The field of Modular Functions had also been developed by that late nineteenth century timeframe, although it was mostly aimed at achieving results in Hyperbolic Geometry. A seminal period appeared in the early twentieth century regarding Modular Forms strongly impacting Number Theory, when Ramanujan had the idea of expressing Modular Forms through their corresponding *q-series*, also known as *q-expansions*, and more formally as

Date: May 24, 2021.

Fourier Series representation.

This intuition led him to formulate three conjectures around his so-called Δ function, subsequently designated as Ramanujan's Δ function. These conjectures led to active work for the rest of the twentieth century, including Mordell's proof of two conjectures, then Hecke's development of his framework of linear operators on Modular Forms, culminating with Deligne's proof of the third conjecture as part of proving the Weil conjecture in the 1970's.

The impact of Modular Forms has continued through the twentieth century, with a high point being Andrew Wiles' 1994 proof of Fermat's Last Theorem via his proof of the Shimura-Taniyama-Wiles Modularity Theorem which states that every elliptic curve corresponds to a Modular Form. The juxtaposition of the Wiles result with a previous result by Frey (1986) that a solution of Fermat's equation would lead to an elliptic curve, and with a subsequent result by Ribet (1986) that Frey curves do *not* correspond to Modular Forms, achieved the proof by contradiction that solutions to Fermat's Last Equation could not exist.

Exciting work continues well into the twenty-first century, notably with the connections between Mock Modular Forms (a relaxed and more inclusive class of functions) and Finite Simple Groups (with no non-trivial quotient groups). Important results are revealed by the Representation Theory *Moonshine* approach to quantifying the order (i.e., cardinality) of Sporadic Groups (Finite Simple Groups that are neither cyclic nor alternating nor Lie), such as Mathieu Groups and *The Monster*.

2. Abstract Algebra Preliminaries

As we will see when we introduce Modular Forms, a key condition to be satisfied by these is a certain form of invariance in the face of fractional linear transformations derived from a matrix group. In order to make this document self-contained, we introduce in this section the main definitions and results from Group Theory that are needed to understand the definition of modular forms. In particular, we will lead towards concepts of a Quotient Group, the Index of a Subgroup, and Groups Actions. Lastly, since we mentioned how modular forms impacted research on Finite Simple Groups in the Historical Background section, we will cover that topic as well.

Definition 2.1 (**Group**)**.** We recall that a group $(G, \cdot)$ is defined as a non-empty set G endowed with a binary operation denoted by "$\cdot$" (a dot symbol): $G \times G \to G$ such that

- *Closure*: $f, g \in G \implies f \cdot g \in G$.
- *Associativity*: $\forall f, g, h \in G$, we have $(f \cdot g) \cdot h = f \cdot (g \cdot h)$.
- *Identity*: there exists and element $e \in G$ acting as the neutral identity element for the $\cdot$ operation, i.e., $e \cdot g = g \cdot e = g$, $\forall g \in G$.
- *Invertibility*: $\forall g \in G$, there exists an inverse element denoted as $g^{-1} \in G$ such that $g^{-1} \cdot g = g \cdot g^{-1} = e$.

Remark 2.2. It is frequent to denote the binary operation by "$*$" or "$+$", and it is also frequent to omit the explicit presence of the operator symbol altogether, and to denote $f \cdot g$ simply by fg when the context is clear enough. It is also frequent to refer to the operation as multiplying, when the context is clear enough.

Theorem 2.3 (**Uniqueness of Inverse**). *If $x, y \in G$ and $x \cdot y = e$ then $y \cdot x = e$. The left inverse and right inverse are necessarily the same. In addition, there is a unique inverse for each element.*

Proof. If $x \cdot y = e$, then we left-multiply by y and we have $y \cdot (x \cdot y) = y \cdot e = y$. Now we left-multiply by y^{-1} and we have $y^{-1} \cdot (y \cdot (x \cdot y)) = y^{-1} \cdot y = e$ and we now now associativity on the left-hand side to get $y^{-1} \cdot (y \cdot (x \cdot y)) = (y^{-1} \cdot y) \cdot (x \cdot y) = e \cdot (x \cdot y) = x \cdot y$. We have therefore shown that $x \cdot y = e \implies y \cdot x = e$, so the left inverse and the right inverse are necessarily the same.
To complete the uniqueness proof, we show uniqueness of the inverse on a given side, e.g., the left side, and we are done. If $x \cdot y = x \cdot z = e$, we left-multiply by x^{-1} and we have $x^{-1} \cdot (x \cdot y) = x^{-1}(x \cdot z)$, i.e., $(x^{-1} \cdot x) \cdot y = (x^{-1} \cdot x) \cdot z$, i.e., $e \cdot y = e \cdot z$, i.e., $y = z$. □

Theorem 2.4 (**Inverse of Product**). $\forall x, y \in G$, $(x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$.

Proof. We multiply and we have the following by repeatedly using associativity of the operation

$$(x \cdot y)(y^{-1} \cdot x^{-1}) = [(x \cdot y) \cdot y^{-1}] \cdot x^{-1} = [x \cdot (y \cdot y^{-1}] \cdot x^{-1} = (x \cdot e) \cdot x^{-1} = x \cdot x^{-1} = e.$$

By virtue of Theorem 2.3, it is enough to show on one side that two elements are inverses, for them to be each other's unique inverse. □

Theorem 2.5 (**Injection Induced by Operation**). *Let $(G, \cdot)$ be a group and let $S \subseteq G$ be a subset of G. Let g be any element of G. Then the function*

$$S \to G$$
$$f_g(s) = g \cdot s$$

is injective. It is therefore bijective when its co-domain is restricted to its image (which we will define below as a coset). The same is true for the function defined by right-multiplying by g, i.e., $f_g^*(s) = s \cdot g$.

Proof. If $f_g(s) = f_g(t)$ for two elements $s, t \in S$, then $g \cdot s = g \cdot t$. We multiply both sides by g^{-1} which is guaranteed to be an element in G, and we get by invoking associativity of the group operation

$$g^{-1} \cdot (g \cdot s) = g^{-1} \cdot (g \cdot t) \implies (g^{-1} \cdot g) \cdot s = (g^{-1} \cdot g) \cdot t \implies e \cdot s = e \cdot t \implies s = t.$$

The proof is identical for the function f_g^*. □

Corollary 2.6. *If we designate by $g \cdot S$ the image set of S by f_g, i.e., the set $\{x \in G : x = g \cdot s, \text{ for some } s \in S\}$, then f_g is a bijection from S to $g \cdot S$. Similarly, if we designate by $S \cdot g$ the image set of S by f_g^*, i.e., the set $\{x \in G : x = s \cdot g, \text{ for some } s \in S\}$, then f_g^* is a bijection from S to $S \cdot g$.*

Proof. We have already shown that f_g is injective, and it is clear that it is surjective by construction of the co-domain as its image. It is therefore a bijection. A similar reasoning on f_g^* concludes that it is bijective. □

Corollary 2.7. *For any subset S of G, and any element g of G, we have*

$$|S| = |g \cdot S| = |S \cdot g|,$$

where we are referring to cardinalities of sets.

Proof. This is a direct consequence of having f_g as a bijection from S to $g \cdot S$ and a bijection from S to $S \cdot g$. □

Definition 2.8 (**Conjugates**)**.** We say that two elements $a, b \in G$ are conjugates if there exists an element g in G such that

$$b = g^{-1} \cdot a \cdot g.$$

Theorem 2.9 (**Conjugacy Equivalence Relation**)**.** *Conjugacy between elements defines an equivalence relation, and conjugate elements are within equivalence classes.*

Proof. We let the relation $x\mathcal{R}y$ be defined as $x\mathcal{R}y \iff \exists\, g \in G : y = g^{-1} \cdot x \cdot g$. We prove that $\mathcal{R}$ is an equivalence relation on G.

- Reflexive: $\forall\, x \in G, x = e^{-1} \cdot x \cdot e$ so we have established that $x\mathcal{R}x$.
- Symmetric: If $x\mathcal{R}y$, then $y = g^{-1} \cdot x \cdot g$ for some $g \in G$. We left-multiply by g and right-multiply by g^{-1} and with associativity, we get

$$x\mathcal{R}y \iff y = g^{-1} \cdot x \cdot g \iff g \cdot y \cdot g^{-1} = x \iff \left(g^{-1}\right)^{-1} \cdot y \cdot (g^{-1}) = x \iff y\mathcal{R}x.$$

- Transitive: If $x\mathcal{R}y$ and $y\mathcal{R}z$, then let $y = g^{-1} \cdot x \cdot g$ and $z = h^{-1} \cdot y \cdot h$ for some $h, g \in G$. We then have

$$z = h^{-1} \cdot y \cdot h = h^{-1} \cdot (g^{-1} \cdot x \cdot g) \cdot h = (h^{-1} \cdot g^{-1}) \cdot x \cdot (g \cdot h) = (g \cdot h)^{-1} \cdot x(g \cdot h) \implies x\mathcal{R}z,$$

where we have used associativity as well as Theorem 2.4.

□

Notation. Given a subset S of G and an element g of G, we denote by

$$g^{-1} \cdot S \cdot g = \{g^{-1} \cdot x \cdot g : x \in S\},$$

the set of the results of conjugation of all elements of S by the element g.

Definition 2.10 (**Order of an Element**)**.** The order of an element $g \in G$ is the smallest $m \in \mathbb{N}$ such that $g^m = g \cdot g \cdot \ldots \cdot g = e$.

Definition 2.11 (**Abelian Group**)**.** An Abelian group, also known as a Commutative Group is a group in which the defining "$\cdot$" binary operation is commutative, i.e.

$$\forall f, g \in G,\ f \cdot g = g \cdot f.$$

Example. The integers with the addition operation form an Abelian group $(\mathbb{Z}, +)$.

Example. The Rubik's cube with the operation of composition of actions is a group. However, it is not an Abelian group as switching the order in which two actions are performed generally leads to different results.

Definition 2.12 (**Subgroup**)**.** A non-empty subset $H \subseteq G$ of a group G is called a subgroup of G if the defining "$\cdot$" binary operation defines a group $(H, .)$ when restricted to $H \times H$. In particular, the result of operating on two members of the subgroup remains in the subgroup, the neutral element is in the subgroup, and the inverse of each element in the subgroup is also in the subgroup.

Definition 2.13 (**Proper Subgroup**)**.** A subgroup H of a group G is proper if $H \subsetneq G$.

Example. The set of even integers with the addition operation $(2\mathbb{Z}, +)$ forms a proper subgroup of $(\mathbb{Z}, +)$.

Definition 2.14 (**Left Cosets**)**.** Given a subgroup H of a group G, and an element $g \in G$, the left coset $g \cdot H$ is

$$g \cdot H = \{g \cdot h : h \in H\},$$

i.e., it is the set of the results of the operation on g with all elements of H, with g being in the left position of the binary operation.

Definition 2.15 (**Right Cosets**)**.** Given a subgroup H of a group G, and an element $g \in G$, the right coset $H \cdot g$ is

$$H \cdot g = \{h \cdot g : h \in H\},$$

i.e., it is the set of the results of the operation on all elements of H with g, with g being in the right position of the binary operation.

Remark 2.16. Clearly when $(G, \cdot)$ is Abelian, $g \cdot H = H \cdot g$, i.e., the left coset and the right coset are the same set for a given element g and a given subgroup H.

Remark 2.17. We note that $g \in g \cdot H$, $\forall g \in G$ because $e \in H$ due to the fact that H is a subgroup. Therefore, $g \cdot e \in g \cdot H$ and since $g \cdot e = g$, this implies that $g \in g \cdot H$. It is similarly true that $g \in H \cdot g$, $\forall g \in G$.

Lemma 2.18. *We have $x \in g \cdot H$ if and only if $g \in x \cdot H$ if and only if $g^{-1} \cdot x \in H$ if and only if $x^{-1} \cdot g \in H$.*

Proof. $x \in g \cdot H \iff \exists h \in H$ such that $x = gh$. Left-multiplying both sides by g^{-1} maintains the equivalence with $g^{-1} \cdot x = g^{-1} \cdot (g \cdot h) = (g^{-1} \cdot g)h = e \cdot h = h \in H$. Reversing the roles of x and g shows the rest of the equivalence being true. □

Corollary 2.19. *We note that by Lemma 2.18, a given coset may very well have several representatives and that it can be equivalently referred to using any of its representatives.*

Theorem 2.20 (**Cosets are Equivalence Classes**)**.** *Given a subgroup H, the relation $x\mathcal{R}y \iff x^{-1} \cdot y \in H$ is an equivalence relation over G. Similarly, the relation $x\mathcal{R}'y \iff x \cdot y^{-1} \in H$ is an equivalence relation.*

Proof. We prove that the relation $\mathcal{R}$ is reflexive, symmetric, and transitive.

- Reflexivity: for any $x \in G$, $x\mathcal{R}x \iff x^{-1} \cdot x \in H \iff e \in H$ which is true since H is a subgroup.
- Symmetry: for any $x, y \in G$, $x\mathcal{R}y \iff x^{-1} \cdot y \in H \implies (x^{-1} \cdot y)^{-1} \in H \iff y^{-1} \cdot x \in H \iff y\mathcal{R}x$, where we have used Theorem 2.4 for the inverse of a product.
- Transitivity: for any $x, y, z \in G$, if $x\mathcal{R}y$ and $y\mathcal{R}z$, then by definition $x^{-1} \cdot y \in H$ and $y^{-1} \cdot z \in H$. Since H is a subgroup, the product of two of its elements is also an element of H, so that $(x^{-1} \cdot y) \cdot (y^{-1} \cdot z) = ((x^{-1} \cdot y) \cdot y^{-1}) \cdot z = (x^{-1} \cdot (y \cdot y^{-1})) \cdot z = x^{-1} \cdot z \in H$, which shows that $x\mathcal{R}z$.

The proof for the right-coset equivalence relation is identical and we omit detailing it. □

Corollary 2.21 (**Cosets Partition a Group**)**.** *Since every equivalence relation on a set induces a partition of the set into equivalence classes, this is the case with left cosets of a subgroup H. It is also the case with right cosets of a subgroup H.*

Theorem 2.22 (**Conjugation of Subgroup**)**.** *If H is a subgroup of G and $g \in G$, then $g^{-1} \cdot H \cdot g$ is also a subgroup.*

Proof. We show that the identity is in $g^{-1} \cdot H \cdot g$, that it is closed under the group operation, and that the inverse of an element in $g^{-1} \cdot H \cdot g$ is also in $g^{-1} \cdot H \cdot g$.

- Identity: since $e \in H$, and therefore $g^{-1} \cdot e \cdot g \in g^{-1} \cdot H \cdot g$, and $g^{-1} \cdot e \cdot g = g^{-1} \cdot g = e$, we have established that $e \in g^{-1} \cdot H \cdot g$.
- Closure: let $x_1, x_2 \in g^{-1} \cdot H \cdot g$, then there exist $h_1, h_2 \in H$ such that $x_1 = g^{-1} \cdot h_1 \cdot g$ and $x_2 = g^{-1} \cdot h_2 \cdot g$. We write
$$x_1 \cdot x_2 = (g^{-1} \cdot h_1 \cdot g) \cdot (g^{-1} \cdot h_2 \cdot g) = g^{-1} \cdot (h_1 \cdot h_2) \cdot g,$$
and $h_1 \cdot h_2 \in H$ because H is a subgroup.
- Inverse: let $x = g^{-1} \cdot h \cdot g \in g^{-1} \cdot H \cdot g$. Then by Theorem 2.4, $x^{-1} = g^{-1} \cdot h^{-1} \cdot g$, and $h^{-1} \in H$ because H is a subgroup. Therefore, $x^{-1} \in g^{-1} \cdot H \cdot g$.

□

Theorem 2.23 (**Cosets and Conjugate Subgroups**). *Given a subgroup H of G, and an element g of G, the right coset $H \cdot g$ of the subgroup H with respect to the element g is the left coset $g \cdot (g^{-1} \cdot H \cdot g)$ of the conjugate subgroup $g^{-1} \cdot H \cdot g$ with respect to the element g, i.e.,*
$$H \cdot g = g \cdot (g^{-1} \cdot H \cdot g).$$

Proof. We have
$$\begin{aligned} x \in H \cdot g &\iff \exists\, h \in H \text{ such that } x = h \cdot g \\ &\iff \exists\, h \in H \text{ such that } x = (g \cdot g^{-1}) \cdot h \cdot g \\ &\iff \exists\, h \in H \text{ such that } x = g \cdot (g^{-1} \cdot h \cdot g) \\ &\iff \exists\, h' = (g^{-1} \cdot h \cdot g) \in g^{-1} \cdot H \cdot g \text{ such that } x = g \cdot h' \\ &\iff x \in g \cdot (g^{-1} \cdot H \cdot g) \end{aligned}$$

□

Remark 2.24. We note that in the general case the right cosets of a subgroup H are distinct from its left cosets. However, by Corollary 2.7, the cardinality of all cosets (right or left) of a given subgroup is the same as the cardinality of that subgroup, and by Theorem 2.23, the number of left cosets of a subgroup is equal to the number of its right cosets.

Definition 2.25 (**Normal Subgroup**). A subgroup N of a group G is a *normal* subgroup if and only if for any element $g \in G$, the corresponding left and the right cosets are equal, i.e., $g \cdot N = N \cdot g$. Equivalently, this means that a normal subgroup is its own conjugate subgroup, i.e.,
$$\forall\, g \in G,\, n \in N, \text{ we have } g^{-1} \cdot n \cdot g \in N.$$

Remark 2.26. We note that the condition for a subgroup being normal does not imply that an element in that subgroup commutes with all elements of the group G (see the definition for *center*). It simply implies that if $x \in g \cdot N = g \cdot n_1$ with $n_1 \in N$, then $x \in N \cdot g$, i.e., there exists some $n_2 \in N$ such that $x = n_2 \cdot g$ but we may have $n_1 \neq n_2$.

Definition 2.27 (**Center of a Group**). The center of a group G, denoted by $Z(G)$, is the set of elements of G that commute with all elements of G, i.e.,
$$Z(G) = \{z \in G \,|\, \forall\, g \in G,\, g \cdot z = z \cdot g\}.$$

Theorem 2.28 (**The Center is a Subgroup**). *The center $Z(G)$ of a group G is a subgroup of G.*

Proof.
- Identity: the identity element commutes with all elements of a group because $\forall\, g \in G$, $e \cdot g = g \cdot e = g$, so it is in the center.
- Closure: if $c_1, c_2 \in Z(G)$, then for any $g \in G$, we have
$$(c_1 \cdot c_2) \cdot g = c_1 \cdot (c_2 \cdot g) = c_1 \cdot (g \cdot c_2) = (c_1 \cdot g) \cdot c_2 = (g \cdot c_1) \cdot c_2 = g \cdot (c_1 \cdot c_2),$$
which shows that $c_1 \cdot c_2 \in Z(G)$.
- Inverse: let $c \in Z(G)$, we want to show that $c^{-1} \in G$. For any $g \in G$, we have after left-multiplying by c^{-1}, then right-multiplying by c^{-1}:
$$c \cdot g = g \cdot c \implies c^{-1} \cdot c \cdot g = c^{-1} \cdot g \cdot c \implies g \cdot c^{-1} = c^{-1} \cdot g \cdot c \cdot c^{-1} = c^{-1} \cdot g,$$
which shows that c^{-1} also commutes with all elements g of G, thus $c^{-1} \in Z(G)$.

□

Corollary 2.29 (**The Center is a Normal Subgroup**). *The center $Z(G)$ of a group G is a normal subgroup of G, i.e., it is its own conjugate subgroup, and its left coset is equal to its right coset with respect to any element of G.*

Proof. This is clear since elements of $Z(G)$ commute with all elements of G, then the left-coset and the right-coset of $Z(G)$ are the same with respect to any element $g \in G$. The condition for the center is stronger than the condition for a normal subgroup. □

Definition 2.30 (**Quotient Group**). For a normal subgroup N of a group G, we define the set of all left cosets of N (which are the same as its right cosets, because N is normal),
$$G/N = \{g \cdot N : g \in G\}.$$
We define the binary operation on elements of G/N as follows:
$$(g \cdot N) * (g' \cdot N) = (g \cdot g') \cdot N.$$

Theorem 2.31. *$(G/N, *)$ is a group.*

Proof. We first show that the binary operation is well defined and does not depend on the choice of representative elements of the cosets. Indeed, if $g \cdot N = h \cdot N$ and $g' \cdot N = h' \cdot N$, then we have (omitting the dot notation for convenience)
$$(gg')N = g(g'N) = g(h'N) = g(Nh') = (gN)h' = (hN)h' = h(Nh') = h(h'N) = (hh')N.$$
We remark that we used repeatedly the fact that N was a normal subgroup in the transformations above, notably when switching freely between a left coset and a right coset of N with respect with the same element of G.

We then note that associativity of the $*$ operation on G/N follows directly from the associativity of the $\cdot$ operation on G.

The identity element is the coset $e \cdot N = N$, as we can easily see that
$$N * (g \cdot N) = (e \cdot N) * (g \cdot N) = (e \cdot g) \cdot N = g \cdot N, \forall\, g \in G.$$
And the inverse with respect to $*$ of a coset gN is the coset $g^{-1}N$, as we have by definition of the $*$ operation
$$(g^{-1} \cdot N) * (g \cdot N) = (g^{-1} \cdot g) \cdot N = e \cdot N = N,$$
with N being the identity element. This completes the proof that $(G/N, *)$ is a group, thus justifying its designation as a quotient group. □

Remark 2.32. Given that N is normal, it is clear that we could have defined the quotient group G/N by way of right cosets, those being the same as the left cosets of N.

Remark 2.33. We note (but don't prove) that there is a converse statement to Claim 2.31, which states that if the operation $*$ above is well defined on a subgroup N, then N must be a normal subgroup.

Definition 2.34 (**Index of Subgroup**)**.** As noted in Remark 2.24, every left or right coset of a subgroup H has the same number of elements (or cardinality, if infinite). Furthermore, the number of left cosets of H is equal to the number of right cosets of H. This number is the *index* of the subgroup H in G and is denoted by $[G : H]$.

Remark 2.35. When a group is finite, then Lagrange's Theorem lets us relate the index of a subgroup H in G to the cardinalities of H and G, via the relation

$$|G| = [G : H]|H|.$$

Definition 2.36 (**Finite Index Subgroup**)**.** A finite index subgroup is a subgroup whose index is finite. The number of left cosets or of right cosets of such a subgroup is finite. We will refer to this concept when discussing certain matrix subgroups later on in the paper.

As we made reference to Finite Simple Groups in the Historical Background section of this paper, we provide the definition of simple groups.

Definition 2.37 (**Simple Group**)**.** A non-trivial group is said to be a simple group if its only normal subgroups are the trivial subgroup $(\{e\}, \cdot)$,where e is the identity element, and the group itself.

Definition 2.38 (**Finite Simple Group**)**.** A finite simple group is a simple group that has finite cardinality, i.e., it has no non-trivial normal subgroups and has finite order (which is a synonym for the cardinality of a group).

Remark 2.39. A non-simple group can be broken down into a proper normal subgroup and the corresponding quotient group. This can be recursively pursued until arriving (in the case of a finite group) at uniquely determined simple groups, by the Jordan-Hölder Theorem.

Not so for a simple group as it has no proper normal subgroups, thus cannot be broken down into normal subgroup and corresponding quotient group substructures.

Definition 2.40 (**Subgroup Generated by Set**)**.** Given a group $(G, \cdot)$ and $S \subseteq G$ a subset of G, the Subgroup Generated by S , denoted as $\langle S \rangle$ is the subgroup of all elements of G that can be expressed as the finite product of elements of S and their inverses.

Remark 2.41. This is clearly a subgroup as can be derived from its definition.

Definition 2.42 (**Generator Set**)**.** We say that a subset $S \subseteq G$ is a Generator Set of the group G if the subgroup $\langle S \rangle$ is equal to G, i.e., every element in G can be expressed as a product of elements of S and their inverses. We also say that G is generated by S, and write

$$\langle S \rangle = G.$$

Example. The additive group of integers $(\mathbb{Z}, +)$ is generated by the one-element set $\{1\}$ as every integer can be expressed as a finite sum of the element 1 or its inverse -1. Therefore, we have

$$\langle \{1\} \rangle = (\mathbb{Z}, +).$$

Definition 2.43 (**Presentation of a Group**)**.** A Presentation of a Group is a method of specifying a group via a generator set S such that $\langle S\rangle = G$ and a set of relations R between the elements of the generator set. We then say that G has a presentation

$$G = \langle S \mid R\rangle.$$

Example. The multiplicative group $(\mathbb{Z}/5\mathbb{Z})^\times$ of integers modulo 5 is generated by the element 2 as $2^2 \equiv 4 \pmod 5$, $2^3 \equiv 3 \pmod 5$, and $2^4 \equiv 1 \pmod 5$, and this element is characterized by its fourth power being equal to the identity element of the group. We therefore have

$$(\mathbb{Z}/5\mathbb{Z})^\times = \langle \{2\} \mid \{2^4 = 1\}\rangle.$$

Example. The Dihedral Group $D4$ is the group of symmetries on the square. It is called the dihedral group of order 8 (because it has 8 elements). It consists of all possible 90° rotations and reflections about the horizontal or vertical or diagonal axes of symmetry of the square, with the operation of composition of such transformations. The group $D4$ can be generated by any one 90° rotation and any one reflection about a symmetry axis. So if we designate a counter-clockwise 90° rotation as the element a, and a reflection about the vertical axis of symmetry as b, then we have the generator set

$$\langle S\rangle = \langle \{a, b\}\rangle = D_4.$$

Furthermore, and if designate the identity operation as e, then the elements a and b are characterized by the relations $a^4 = e$, $b^2 = e$, and $a \cdot b = b^c \dot a^{-1}$, so that the set of relations is

$$\mathcal{R} = \{a^4 = b^2 = e,\, a \cdot b = b \cdot a^{-1}\},$$

and therefore a presentation of the group $D4$ is the following

$$D_4 = \langle \{a,\, b\} \mid \{a^4 = b^2 = e,\, a \cdot b = b \cdot a^{-1}\}\rangle.$$

Definition 2.44 (**Finitely Generated / Related / Presented**)**.** If a group can be generated by a finite subset S, then it is said to be finitely generated. If the set of relations among the elements of the generating set is finite, then the group is said to be finitely related. If both conditions are met, then the group is said to be finitely presented.

2.1. **Functions on Groups.**

Definition 2.45 (**Group Homomorphism**)**.** A group homomorphism is a function from a group to another that preserves the operations of the groups, i.e., for $x, y \in G$:

$$f : (G, \cdot) \to (G', *)$$
$$f(x \cdot y) = f(x) * f(y)$$

Example. The structure $(\mathbb{R}, +)$ is a group and the structure $(\mathbb{R}^+ \setminus \{0\}, \cdot)$ is another group, and the exponential function

$$f : (\mathbb{R}, +) \to (\mathbb{R}^+ \setminus \{0\}, \cdot)$$
$$f(x) = e^x$$

is a group homomorphism between the two groups above, because we have

$$f(x + y) = e^{x+y} = e^x \cdot e^y = f(x) \cdot f(y).$$

Definition 2.46 (**Group Isomorphism**)**.** A group isomorphism is a bijective group homomorphism between two groups.

Definition 2.47 (**Group Automorphism**)**.** A group automorphism is an isomorphism from a group onto itself.

Definition 2.48 (**Kernel**)**.** Given a homomorphism $f : G \to G'$ from a group to another, we define the *kernel* of the homomorphism as the subset of elements of G whose image by f is the identity element of G', i.e.,

$$\ker(f) = \{g \in G : f(g) = e_{G'}\} \subseteq G.$$

Definition 2.49 (**Image**)**.** Given a homomorphism $f : G \to G'$ from a group to another, we define the *image* of the homomorphism as the subset of elements of G' that are images by f of some elements of G, i.e.,

$$\mathrm{im}(f) = \{g' \in G' : \exists\, g \in G \text{ such that} f(g) = g'\} \subseteq G'.$$

Theorem 2.50 (**Group of Automorphisms on a Group**)**.** *Given a group G, the set of automorphisms on G, with the operation of function composition, is a group.*

Proof. We prove that all the conditions of a group structure are met.

- Identity: the identity function on a group is an automorphism of that group, and is the identity element for the function composition operation.
- Associativity: function composition is associative in general, so it is so for automorphisms on a group.
- Closure: if f_1 and f_2 are two automorphisms on G, then they are bijective from G to G, and their composition is also a bijection from G to G, as this is a general result. We now examine the behavior of their composition with respect to homomorphic behavior. For any two elements $x, y \in G$, we have

$$(f_2 \circ f_1)(xy) = f_2[f_1(xy)] = f_2[f_1(x)f_1(y)] = f_2(f_1(x))f_2(f_1(y)) = (f_2 \circ f_1)(x) \cdot (f_2 \circ f_1)(y).$$

- Inverse: if f is an automorphism on G, then f^{-1} is a bijection from G to G because this is a general result for bijections. We now show that the inverse satisfies the homomorphism condition. For any $x, y \in G$, there exist $u, v \in G$, such that $x = f(u)$ and $y = f(v)$ because the bijection f is a surjection. We then have

$$f^{-1}(x \cdot y) = f^{-1}[f(u) \cdot f(v)] = f^{-1}[f(u \cdot v)] = u \cdot v = f^{-1}(x) \cdot f^{-1}(y),$$

  where we have used the homomorphic property of f in the second equality above.

□

Definition 2.51 (**Isomorphic Groups**)**.** Two groups $(G, \cdot)$ and $(G', *)$ are said to be isomorphic if there exists an isomorphism from one to the other (and thus vice-versa). We denote such isomorphic groups as

$$(G, \cdot) \cong (G', *)$$

Example. The exponential function is a group isomorphism from $(\mathbb{R}, +)$ to $(\mathbb{R}^+ \setminus \{0\}, \cdot)$, as it is a homomorphism and is also bijective.

Theorem 2.52 (**Properties of Isomorphisms**)**.** *We state, and they are straightforward to prove, a few important properties of isomorphisms. Let $f : (G, \cdot) \to (G', *)$ be a group isomorphism, and let e_G and $e_{G'}$ be the identity elements of G and G', respectively. We then have*

- $f(e_G) = e_{G'}$.
- *G is Abelian if and only if G' is Abelian.*

- $f(g^{-1}) = [f(g)]^{-1}$, $\forall\, g \in G$.
- $f(g^m) = [f(g)]^m$.
- *The order of $g \in G$ is equal to the order of $f(g) \in G'$.*
- *The kernel of f is e_G, i.e., e_G is the one and only element in G whose image by f is $e_{G'}$.*

Theorem 2.53 (**Equivalence Classes of Isomorphic Groups**). *The relation of being isomorphic between groups $(G, \cdot) \cong (G', *)$, as defined in Definition 2.51, is an equivalence relation.*

Proof. The relation is clearly reflexive as any group is isomorphic to itself with the isomorphism being the identity function.
It is symmetric because if f is an operation-preserving bijection from $(G, \cdot)$ to $(G', *)$, then its inverse f^{-1} is an operation-preserving bijection from $(G', *)$ to $(G, \cdot)$.
It is transitive because if $f_1 : (G, \cdot) \to (G', *)$ is an isomorphism and $f_2 : (G', *) \to (G'', \times)$ is an isomorphism, then $f_2 \circ f_1 : (G, \cdot) \to (G'', \times)$ is a composition of bijections, therefore a bijection. And it is operation-preserving because if $g, g' \in G$, then we have

$$(f_2 \circ f_1)(g \cdot g') = f_2[f_1(g \cdot g')] = f_2[f_1(g) * f_1(g')] = f_2[f_1(g)] \times f_2[f_1(g')] = (f_2 \circ f_1)(g) \times (f_2 \circ f_1)(g').$$

□

We end this section with a result for all homomorphisms, which we will show in the context of groups. The result will be useful when discussing modular forms in its form for linear transformations between vector spaces, and in fact this result has versions that apply to many algebraic structures, including groups, rings, and others. We will re-state the result in the vector space context that will be of interest to us for modular forms.

Theorem 2.54 (**Isomorphism Theorem**). *Let $(G, \cdot)$ and $(G', *)$ be two groups, and let $f : G \to G'$ be a group homomorphism. We then have the following results:*

- *The kernel* $\ker(f)$ *of f is a normal subgroup of G.*
- *The image* $\mathrm{im}(f)$ *of f is a subgroup of G'.*
- *The image* $\mathrm{im}(f)$ *of f is isomorphic to the quotient group* $G/\ker(f)$.

Proof.

- Recalling from Definition 2.25 that a normal subgroup is a subset of the group such that its elements conjugated with any element of G remain in the subset in question. We want to show that this is the case for $\ker(f)$, so we suppose $k \in \ker(f)$, and let g be any element of G. We want to show that $g^{-1} \cdot k \cdot g \in \ker(f)$. We have

$$\begin{aligned} f(g^{-1} \cdot k \cdot g) &= f(g^{-1}) * f(k) * f(g) \\ &= f(g^{-1}) * e_{G'} * f(g) \\ &= f(g^{-1} * f(g) \\ &= f(g^{-1} \cdot g) \\ &= f(e_G) \\ &= e_{G'}. \end{aligned}$$

  The kernel $\ker(f)$ is therefore a normal subgroup of G.
- Since $f(e_G) = e_{G'}$, we have established that $e_{G'} \in \mathrm{im}(f)$. We now verify that inverses of elements of $\mathrm{im}(f)$ are also in $\mathrm{im}(f)$: we let $i \in \mathrm{im}(f)$ be such an element, then

there exists $g \in G$ such that $f(g) = i$. We now have

$$\begin{aligned} e_{G'} = f(e_G) = f(g^{-1} \cdot g) = f(g^{-1}) * f(g) = f(g^{-1}) * i \\ \implies i^{-1} = f(g^{-1}) \\ \implies i^{-1} \in \text{im}(f). \end{aligned}$$

Lastly, we show closure of $\text{im}(f)$ under the group operation $*$. Let i_1 and i_2 be two elements of $\text{im}(f)$, so there are two elements $g_1, g_2 \in G$ such that $i_1 = f(g_1)$ and $i_2 = f(g_2)$. We then have

$$f(g_1 \cdot g_2) = f(g_1) * f(g_2) = i_1 * i_2,$$

so we have found an element $g_1 \cdot g_2 \in G$ such that $i_1 * i_2 = f(g_1 \cdot g_2)$, which shows closure of $\text{im}(f)$ under the operation of the group G'.
We have threfore shown that $\text{im}(f)$ is closed under the group operation, contains the identity element $e_{G'}$, and contains the inverse of any of its elements. It is therefore a subgroup of G'.

- To show that $(\text{im}(f), *) \cong (G/\ker(f), \cdot)$, we define the function

$$\phi : (G/\ker(f), \cdot) \to (\text{im}(f), *)$$

$$\phi(g) = f(g)$$

and we show that it is a well-defined function, that it is homomorphic, and that it is bijective. We first note that an element in the subgroup $G/\ker(f)$ is an equivalence class of an element g of G multiplied by all elements $k \in \ker(f)$.

The function ϕ is well-defined because if $k_1, k_2 \in \ker(f)$, then $f(k_1) = f(k_2) = e_{G'}$ and $f(g \cdot k_1) = f(g) * f(k_1) = f(g) * e_{G'} = f(g)$ and $f(g \cdot k_2) = f(g) * f(k_2) = f(g) * e_{G'} = f(g)$, therefore the function ϕ defined on a coset $g \cdot \ker(f)$ as $f(g)$ is well-defined.

The function ϕ is homomorphic because it is derived from f which is homomorphic from $(G, \cdot)$ to $(G', *)$, and since group structure is preserved by f, then group structure is also preserved by ϕ since for any two elements $g_1, g_2 \in G$, and the two corresponding cosets $g_1 \cdot \ker(f)$ and $g_2 \cdot \ker(f)$, we have $\phi[g_1 \cdot \ker(f)] = f(g_1)$ and $\phi[g_2 \cdot \ker(f)] = f(g_2)$, therefore

$$\phi[(g_1 \cdot g_2) \cdot \ker(f)] = f(g_1 \cdot g_2) = f(g_1) * f(g_2) = \phi[g_1 \cdot \ker(f)] * \phi[g_2 \cdot \ker(f)],$$

which proves that we have a homomorphism.

We now prove that ϕ is a bijection. It is clearly a surjection because by definition of $\text{im}(f)$, every element is the image by f of some element g of G, which means that it is the image by ϕ of this element's coset $g \cdot \ker(f)$ which is an element of

$G/\ker(f)$. It is also an injection because

$$\begin{aligned}
&\phi[g_1 \cdot \ker(f)] = \phi[g_2 \cdot \ker(f)] \iff f(g_1) = f(g_2) \\
&\implies f(g_1 \cdot g_2^{-1}) = f(g_1) * f(g_2^{-1}) = f(g_2) * f(g_2^{-1}) = f(g_2 \cdot g_2^{-1}) = f(e_G) = e_{G'} \\
&\implies g_1 \cdot g_2^{-1} \in \ker(f) \\
&\iff \exists k \in \ker(f) \text{ such that } g_1 \cdot g_2^{-1} = k \\
&\iff g_1 = k \cdot g_2 \\
&\iff g_1 \text{ is in the right coset } \ker(f) \cdot g_2 \\
&\implies g_1 \text{ is in the left coset } g_2 \cdot \ker(f) \text{ because } \ker(f) \text{ is a normal subgroup} \\
&\implies g_1 \cdot \ker(f) = g_2 \cdot \ker(f).
\end{aligned}$$

This completes the proof of the isomorphism

$$\operatorname{im}(f) \cong G/\ker(f).$$

□

2.2. **Group Actions.**

Definition 2.55 (**Transformations**)**.** Given a set or a geometric space S, a transformation on S is a bijective function from S to S.

Theorem 2.56 (**Transformation Group**)**.** *The set of transformations on a space forms a group with the function composition operation, with the identity element of the group being the identity function.*

Proof. Composition of functions is associative, the identity function acts as the identity for function composition, the composition of bijective functions from a set to itself results in a bijective function from that set to itself, and each bijective function from a set to itself has an inverse which is also a bijection from that set to itself. We therefore have a group. □

Definition 2.57 (**Group Action**)**.** A group action of a group G on a space S is a group homomorphism from the group G to the group of transformations of S (which we know to be a group from Theorem 2.56). A group action of a group G on an algebraic structure is a group homomorphism from the group G to the group of automorphisms of the structure (which we have shown to be a group when the structure is a group, in Theorem 2.50). More generally, a group action of a group G on a set X is a group homomorphism from G to the group of all bijections of X onto itself, i.e., to the symmetric group of X.

Definition 2.58 (**Representation**)**.** When the target structure of the group action is a finite-dimensional vector space (e.g., $\mathbb{R}^n$), the group action is called a *representation* of the group G.

Remark 2.59. We will see in the following sections that a frequent use of such a group action is to allow the identification of the group G with subgroups of the so-called linear group of invertible square matrices of some given finite dimension over a ring such as $\mathbb{Z}$ or a field such as $\mathbb{R}$.

Definition 2.60 (**Orbit of an Element**)**.** In the context of a group action, each element g of a group G induces a bijection $f_g : X \to X$ or an automorphism on the target set X. Given an element $x \in X$, each of the induced bijections maps x to some image $f_g(x)$ that is

also in X. The set of all images of the element x by all the bijections on X induced by the elements g of G is called the *orbit* of the element x. The orbit of $x \in X$ is therefore the set denoted as

$$G \cdot x = \{f_g(x) : g \in G\}.$$

Theorem 2.61 (**Orbits as Equivalence Classes**). *The relation on elements of X*

$$x \mathcal{R} y \iff G \cdot x = G \cdot y$$

is an equivalence relation.

Proof. The relation is defined by equality of sets, so it is trivially easy to see that it is reflexive, symmetric, and transitive. □

Notation. The set of all orbits of elements of X by a group action of a group G is denoted by X/G or also $G\backslash X$.

Definition 2.62 (**Fundamental Domain**). In the context of a group action (homomorphism) from a group G into the group of transformations on a set or space X, a Fundamental Domain for this action is a set D of representatives of all the orbits of all elements $x \in X$ (i.e., the set of images $f_g(x)$ of x by all the induced actions $f_g : X \to X$ by all the elements $g \in G$).
As an important note, fundamental domains are often considered in the context of X being a Topological Space (i.e., endowed with a distance and where concepts of open and closed sets and convergence of sequences are well-defined). In those cases, fundamental domains are particularly useful when they exhibit "good" properties such as being open, connected sets, for instance.
The key benefit of fundamental domains is that they enable the study of the behavior of the induced transformations $f_g : X \to X$ over their entire domain X by reducing the study to the behavior over the fundamental domain, as all such behaviors will be exhibited on points in the fundamental domain due to the fact that the fundamental domain contains a (usually unique) representative from all orbits of all elements of X.

Example. Let $(\mathbb{Z}, +)$ be the additive group of integers, and let $X = \mathbb{R}$. We consider the group action as the homomorphism that maps an element $n \in \mathbb{Z}$ to the function

$$f_n : \mathbb{R} \to \mathbb{R}$$
$$f_n(x) = n + x$$

It is easy to see that this is a homomorphism as

$$f_{n+m}(x) = (n + m) + x = n + (m + x) = n + f_m(x) = f_n(f_m(x)) = (f_n \circ f_m)(x).$$

And it is also clear that each of the functions $f_n(x)$ as defined is a bijection from $\mathbb{R}$ to $\mathbb{R}$, so we have a group action. The orbit $\mathcal{O}(x)$ of a given element $x \in \mathbb{R}$ is the set of reals

$$\mathcal{O}(x) = \{n + x : n \in \mathbb{Z}\}.$$

A fundamental domain for this group action would be the real semi-open interval

$$\mathcal{D} = [0,\ 1).$$

Indeed, every orbit $\mathcal{O}(x)$, for any $x \in \mathbb{R}$, has one and only one representative in $\mathcal{D}$ because for any real x, we have $x - \lfloor x \rfloor \in [0, 1)$ and $x - \lfloor x \rfloor \in \mathcal{O}(x)$ since it is the image of x by the function $f_{-\lfloor x \rfloor}(y) = -\lfloor x \rfloor + y$, for $y \in \mathbb{R}$.

We can see that studying the behavior of the functions $\{f_n(x) = n + x\}_{n \in \mathbb{Z}}$ on the fundamental domain $[0, 1)$ reveals their behavior over the entire original set $\mathbb{R}$.

2.3. **Rings and Fields.** We end this section with a reminder of the definition of a ring, as we will use this definition further down when discussing groups of matrices defined over rings, i.e., matrices whose entries are ring elements.

Definition 2.63 (**Ring**)**.** A ring is an Abelian group whose operation is called *addition*, with a secondary binary operation called *multiplication* that is associative, distributive over the addition operation, and that has a multiplicative identity element.

Example. The algebraic structure $(\mathbb{Z}, +, \cdot)$ of the integers endowed with addition and multiplication is a ring.

Definition 2.64 (**Field**)**.** A field is a ring in which, additionally to the ring conditions, multiplication is commutative, and every element except the additive identity element has a multiplicative inverse in the field.

Example. The structure $(\mathbb{Q}, +, \cdot)$ of the rationals with addition and multiplication is a field.

Example. The structure $(\mathbb{R}, +, \cdot)$ of real numbers with real addition and real multiplication is a field.

Definition 2.65 (**Vector Space**)**.** Given a field F, a vector space over the field F is a set V endowed with two operations that satisfy the eight axioms below. In the context of vector spaces, elements of F are called *scalars* and elements of V are called *vectors*. The first operation is vector addition $+ : V \times V \to V$ and the second operation is scalar multiplication $\cdot : F \times V \to V$. Scalar multiplication is frequently represented without the $\cdot$ (dot) symbol.

- Associativity of addition: $\forall\, u, v, w \in V$, $u + (v + w) = (u + v) + w$.
- Commutativity of addition: $\forall\, u, v \in V$, $u + v = v + u$.
- Identity element of addition: $\exists\, 0 \in V$ such that $\forall\, v \in V$, $v + 0 = 0 + v = v$. This vector 0 is called the *zero vector*.
- Inverse elements of addition: $\forall\, v \in V$, $\exists\, -v \in V$ such that $v + (-v) = 0$. This vector $-v$ is called the *additive inverse* of v.
- item Compatibility between scalar multiplication and field multiplication: $\forall\, a, b \in F$ and $\forall\, v \in V$, we have $a \cdot (b \cdot v) = (a \cdot b) \cdot v$, i.e., $a(bv) = (ab)v$.
- Identity element of scalar multiplication: with 1 denoting the multiplicative identity in the field F, we have $\forall\, v \in V$, $1v = v$.
- Distributivity of scalar multiplication with respect to vector addition: $\forall\, a \in F$ and $\forall\, u, v \in V$, we have $a(u + v) = au + av$.
- Distributivity of scalar multiplication with respect to field addition: $\forall\, a, b \in F$ and $\forall\, v \in V$, we have $(a + b)v = av + bv$.

Example. Given any field F, the set of *ordered n-tuples* $(a_1, a_2, \ldots, a_n)$ with $a_1, a_2, \ldots, a_n \in F$ is a vector space usually denoted by F^n and called a coordinate space. A common example is with $F = \mathbb{R}$ or $F = \mathbb{C}$ and $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$, respectively.

Definition 2.66 (**Linear Subspace**)**.** A non-empty subset U of a vector space V is called a linear subspace or a vector subspace (or simply subspace) of V if it is closed under vector addition and scalar multiplication, and therefore linear combination operations, and as a consequence of which it contains the 0 vector in particular.

Theorem 2.67 (**A Subspace is a Vector Space**). *A linear subspace of a vector space is itself a vector space.*

Proof. It is easy to prove all eight conditions defining a vector space for a subspace, in that the vector addition identity element is in the subspace, it has the required closure conditions, and it carries over the other conditions related to scalar multiplication with elements of the field over which these vector spaces are defined. □

Definition 2.68 (**Quotient Space**). Similarly to quotient groups being cosets attached to a subgroup of a group, if U is a subspace of a vector space V, then the quotient space V/U is defined as

$$V/U = \{v + U : v \in V\},$$

i.e., it is the set of cosets of each element of V with the subspace U. An element in that quotient space is an equivalence class of vectors whose difference is an element of the subspace U.

Theorem 2.69 (**A Quotient Space is a Subspace**). *The quotient space of a vector space by one of its subspaces is itself a linear subspace of the vector space.*

Proof. We omit the proof because it is very similar structure to the proof we gave for the equivalent result of quotient groups being subgroups of a group. □

Definition 2.70 (**Linear Combination**). For any set of scalars $\{\lambda_i\}_{i\in I}$ indexed by some set I and for any similarly indexed set of vectors $\{v_i\}_{i\in I}$, a sum of the form

$$\sum_{i\in I} \lambda_i v_i = \lambda_1 v_1 + \lambda_2 v_2 + \dots$$

is called a linear combinations of the vectors $\{v_i\}_{i\in I}$.

Definition 2.71 (**Basis**). A set of vectors $\{b_i\}_{i\in I}$, where $b_i \in V$ and I is some set of indices, is a basis of the vector space V if

$$\forall\, v \in V,\ \exists\, \{\lambda_i\}_{i\in I} \text{ with all } \lambda_i \in F, \text{ such that } v = \sum_{i\in I} \lambda_i b_i.$$

This means that a basis is a set of vectors from V such that any vector in V can be obtained by a linear combination of the vectors in the basis. The basis $\{b_i\}_{i\in I}$ is said to *span* the vector space V.

Remark 2.72. The choice of a basis allows the unique decomposition of any vector into its linear combination of the basis vectors, and it allows the representation of any vector as the ordered n-tuple of the scalars that appear in its linear combination representation. These scalars are called the *coordinates* of the vector in the basis.

Remark 2.73. It is important to note that the coordinates of a vector depend on the choice of the basis for the vector space, and are *not* invariant with respect to changes of bases.

Proposition 2.74 (**Uniqueness of Basis Decomposition**). *Once a basis $\{b_i\}_{i\in I}$ has been chosen for a vector space, the decomposition of a vector $v \in V$ as a linear combination of vectors from that basis is unique.*

Proposition 2.75 (**Dimension**). *All bases of a vector space V have the same cardinality, i.e., if $\{b_i\}_{i\in I}$ and $\{b'_j\}_{j\in J}$ are two bases of V, then $|I| = |J|$. As a result, there is a well-defined quantity called the dimension of a vector space and denoted by* $\dim(V)$ *that is*

independent of the choice of basis, and is equal to the cardinality of any basis of the vector space.

Remark 2.76. A vector space can have finite dimension when $|I| = n \in \mathbb{N} \cup \{0\}$ such as with $\mathbb{R}^3$, or it can have a countably infinite dimension when $|I| = \aleph_0 = |\mathbb{N}|$ such as with the space of all real polynomials of any degree, or it can have a dimension that is uncountable when $|I| = 2^{\aleph_0} = \mathfrak{c}$ such as with the space of all functions on real variables.

Definition 2.77 (**Linear Transformation or Map**)**.** A function $f : V \to V'$ from a vector space to another is a linear map or linear transformation if it satisfies

$$\forall \{\lambda_i \in F\}_{i\in I}, \forall \{v_i \in V\}_{i\in I}, \text{ we have } f(\sum_{i\in I} \lambda_i v_i) = \sum_{i\in I} \lambda_i f(v_i).$$

We note that this is really a homomorphism that preserves the algebraic structure of a vector space, in that the image by the function of a linear combination of vectors is equal to the corresponding linear combination of images by the function of the individual vectors.

Definition 2.78 (**Endomorphism**)**.** When $V = V'$, a linear transformation from V to itself is called an endomorphism.

Definition 2.79 (**Eigenvectors and Eigenvalues**)**.** Given an endomorphism f on a vector space V over a field F, a vector $v \in V$ is an *eigenvector* with associated *eigenvalue* $\lambda \in F$ if

$$f(v) = \lambda v.$$

Theorem 2.80 (**Eigen-subspace**)**.** *The set of vectors in V associated with a common eigenvalue λ for en endomorphism f on V is a linear subspace of V. Its dimension is called the multiplicity of the eigenvalue λ, and as a consequence there is a basis of eigenvectors for that subspace, and their number is equal to the multiplicity of λ.*

Proof. The vector 0 satisfies $f(0) = 0 = \lambda 0$, so our set satisfies the first condition of a subspace, i.e., being non-empty. Furthermore, if $u, v \in V$ are such that $f(u) = \lambda u$ and $f(v) = \lambda v$, then for any $a, b \in F$, we have

$$f(au + bv) = af(u) + bf(v) = a\lambda u + b\lambda v = \lambda(au + bv),$$

which shows that our set is closed under linear combinations. □

Definition 2.81 (**Automorphism**)**.** When an endomorphism is also bijective, i.e., it is also an isomorphism, then it is called an automorphism of the vector space V. Any eigenvalue λ of an automorphism f must satisfy $\lambda \neq 0$.

Definition 2.82 (**Linear Operator**)**.** When $V' = F$, i.e., the linear map is from a vector space V over the field F to the field F, then it is called a linear operator on V.

Definition 2.83 (**Kernel**)**.** The kernel of a linear map $f : V \to V'$ is the set $\ker(f) \subseteq V$ defined as the set of elements in V whose image by f is the 0 vector in V', i.e.,

$$\ker(f) = \{v \in V : f(v) = 0_{V'}\}.$$

Definition 2.84 (**Image**)**.** The image of a linear map $f : V \to V'$ is the set $\mathrm{im}(f) \subseteq V'$ defined as the set of elements in V' that are images by f of vectors of V, i.e.,

$$\mathrm{im}(f) = \{v' \in V' : \exists\, v \in V \text{ such that } f(v) = v'\}.$$

Theorem 2.85 (**Kernel, Image, Dimension**). *An equivalent theorem to Theorem 2.54 that we introduced in the context of group homomorphisms is the following in the context of linear maps (i.e., homomorphisms) between vector spaces. If $f : V \to V'$ is a linear map, then we have*

- $\ker(f)$ *is a linear subspace of* V.
- $\text{im}(f)$ *is a linear subspace of* V'.
- $V/\ker(f) \cong \text{im}(f)$ *and, when all dimensions are finite, we also have*

$$\dim(V) = \dim[\ker(f)] + \dim[\text{im}(f)].$$

Proof. We omit the proof because it is very similar in structure to the proof of the equivalent result for group homomorphisms that we gave for Theorem 2.54. □

Definition 2.86 (**Matrix**). A matrix with entries from a ring R or a field F is a rectangular array of scalars from the ring or field. Although infinite dimension matrices exist, we will restrict our attention to finite-dimension matrices where the number of (horizontal) rows and of (vertical) columns are both finite positive integers.

Definition 2.87 (**Matrix Dimensions**). A matrix of finite dimensions is said to be $n \times m$, with $n, m \in \mathbb{N}$, if it has n rows and m columns.

Definition 2.88 (**Matrices and Linear Maps**). Given a linear map $f : V \to V'$ between two vector spaces, with $\dim(V) = n$ and $\dim(V') = m$, and given a choice of a basis for V and a basis for V', vectors in V can be represented by their coordinates in the chosen basis of V as $n \times 1$ matrices and vectors in V' can be represented by their coordinates in the chosen basis of V' as $m \times 1$ matrices, and the linear map f can be represented as an $n \times m$ matrix.

Theorem 2.89 (**Bijective Linear Maps Isomorphic to Invertible Matrices**). *The following propositions are important results in linear algebra.*

- *The set of invertible linear maps from a vector space V with $\dim(V) = n$ to itself, and with V defined over a field F, is a group with the function composition operation, and with the identity function as the identity element. This group is the group of automorphisms of the vector space V.*
- *The set of invertible square $n \times n$ matrices with entries in a field F is a group with the matrix multiplication operation, and with the identity matrix $I_{n\times n}$ as the identity element, where the identity matrix has the 1 element of the field F as its main diagonal entries and the 0 element of the field F as its off-diagonal entrie.*
- *For each choice of a basis for the space V, there is an isomorphism between the group of automorphisms of a vector space V over a field F and with $\dim(V) = n$, and the group of invertible square $n \times n$ matrices with entries in F.*
- *If an automorphism $f_{V\to V}$ corresponds to a matrix $M_{n\times n}$, then the entry m_{ij} in the i^{th} row and j^{th} column of $M_{n\times n}$ is the j^{th} coordinate (in the chosen basis) of the image by f of i^{th} basis vector of V.*

Theorem 2.90 (**Vector Space Generated by Vectors**). *Given a set of n vectors $\{v_1, \ldots, v_n\}$ with entries in a field K, all with equal number of entries greater than or equal to n, the set of all linear combinations with scalars in K of the vectors in $\{v_1, \ldots, v_n\}$ is a vector space designated as the vector space generated by these vectors. Furthermore, if the vectors are linearly independent, then the dimension of the generated vector space is equal to the cardinality of the set of generating vectors, in this case n.*

Proof. The proof is very straightforward as it is easy to see that the 0 vector can be generated by a linear combination with all scalar coefficients equal to the 0 of the field K, and closure under linear combinations is evident from the definition of the set being all linear combinations of vectors from the initial set. □

Notation (**Generation of Vector Space from Subspaces**)**.** Given a set of vectors $\{v_1, \ldots, v_n\}$ on a field K that generate a vector space by their linear combinations with scalars in K, we denote the vector space generated by the set of vectors as follows

$$Kv_1 \oplus \cdots \oplus Kv_n.$$

Similarly, if a vector space is generated by the linear combinations of a given vector v and a subspace W (i.e., by all linear combinations of the vector v and of all vectors in W), then the generated vector space is denoted as follows

$$Kv \oplus W.$$

Example. As an example, if we designate by $\mathcal{P}_2$ the vector space of polynomials with complex coefficients of degree less than or equal to 2, then it is generated by the subspace of polynomials of degree less than or equal to 1, i.e., $\mathcal{P}_1$ and the monomial X^2, so we can denote this result as

$$\mathcal{P}_2(X) = \mathbb{C}X^2 \oplus \mathcal{P}_1(X)$$

Definition 2.91 (**Direct Sum**)**.** When generating a vector space over a field K from linear combinations of all vectors in some subspaces or of a specific set of vectors, then we say that the generated vector space is the *direct sum* of the subspaces (or vectors) that generate it by all linear combinations with coefficients in K.

3. Matrix Groups

Definition 3.1 (**Matrix Group**)**.** A matrix group is a group consisting of invertible matrices over a given ring R or field F (as defined in Definition 2.63 and Definition 2.64), with the matrix multiplication operation.

Example. The group of $n \times n$ real matrices with non-zero determinant form a matrix group with the matrix multiplication operation.

Definition 3.2 (**General Linear Group**)**.** The General Linear Group of degree n over a ring R or a field F, denoted as $GL_n(R)$ or $GL_n(F)$, respectively (or sometimes also $GL(n, R)$ or $GL(n, F)$), is the set of $n \times n$ invertible matrices, whose entries are from the ring R or the field F, respectively, together with the matrix multiplication operation.

Remark 3.3. This forms a group because such matrix multiplication is associative, the identity matrix $I_{n\times n}$ is the identity element, the product of two invertible matrices is also invertible, and an invertible matrix has an inverse which is also invertible.

Remark 3.4. In the case of a matrix with entries from a field, a matrix is invertible whenever its determinant is non-zero. However, in the case of a matrix with entries from a ring, a matrix is invertible whenever its determinant is a unit in that ring, i.e., the determinant must be equal to an element x of the ring such that there exists an inverse x^{-1} in the ring, and the product $x \cdot x^{-1} = x^{-1} \cdot x = 1$.

Example. When considering $n \times n$ matrices with entries from the ring $(\mathbb{Z}, +, \cdot)$, invertibility requires the determinant to be equal to 1 or -1 because these are the only two units in $\mathbb{Z}$.

Example. On the other hand, when considering matrices with entries from the ring $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$, invertibility requires the determinant to be equal to one of the units in the multiplicative group of integers modulo n, i.e., $\mathbb{Z}_n^\times$. Such units are those integers in $\{1, \ldots, n-1\}$ that are relatively prime with n. If $n = p$ prime, then all non-zero integers in $\mathbb{Z}/p\mathbb{Z}$ are units.

Remark 3.5. For a vector space V over a field F, there is a definition of general linear group for that vector space, defined as the group of all automorphisms of V, i.e., all bijective linear transformations from V to V, together with function composition as the operation. It is denoted as $GL(V)$.

Remark 3.6. When the vector space V is finite-dimensional of, e.g., degree n, then this general linear group of the vector space can become isomorphic to the general linear group $GL_n(F)$ as defined in Definition 3.2, once a basis for the vector space V has been chosen.

Definition 3.7 (**Special Linear Group**)**.** The Special Linear Group of degree n over a ring R or a field F, respectively denoted as $SL_n(R)$ or $SL_n(F)$, is the set of $n \times n$ invertible matrices whose entries are from the ring R or the field K, respectively, and whose determinant is equal to 1.

Example. $SL_2(\mathbb{Z})$ is the group of 2×2 matrices with integer entries and whose determinant is equal to 1., with matrix multiplication as the group operation. We note that this group is sometimes denoted Γ in the context of modular forms.

Definition 3.8 (**Projective General Linear Group**)**.** The Projective General Linear Group of degree n over a ring R or a field F, respectively denoted as $PGL_n(R)$ and $PGL_n(F)$, is the quotient group of $GL_n(R)$ or $GL_n(F)$ by their centers (recalling from Definition 2.28 that the center of a group is the subgroup of elements that commute with all elements of the group).

Example. The center of $GL_n(\mathbb{R})$ is the subgroup made of $n \times n$ invertible matrices with real coefficients and that commute with all other $n \times n$ invertible real matrices. This set is in fact composed of all non-zero multiples of the identity matrix, i.e., it is the set

$$\{\lambda I_{n\times n} : \lambda \in \mathbb{R}, \lambda \neq 0\}.$$

Therefore, $PGL_n(\mathbb{R})$ is the quotient of $GL_n(\mathbb{R})$ by these multiples-of-identity matrices. Two matrices are equivalent in this context if they are non-zero real multiples of one another.

Definition 3.9 (**Projective Special Linear Group**)**.** The Projective Special Linear Group of degree n over a ring R or a field F, respectively denoted as $PSL_n(R)$ and $PSL_n(F)$, is the quotient group of $SL_n(R)$ or $SL_n(F)$ by their centers (recalling from Definition 2.28 that the center of a group is the subgroup of elements that commute with all elements of the group).

Example. The center of $SL_n(\mathbb{Z})$ is made of just the two matrices $I_{n\times n}$ and $-I_{n\times n}$ as these are the only invertible matrices with integer coefficients that commute with all invertible integer matrices. Therefore, $PSL_n(\mathbb{Z})$ is made of equivalence classes each of size 2, with a class having an integer matrix of determinant 1 and its additive inverse (i.e., "its negative").

Remark 3.10 (**Matrices in** $\Gamma = SL_2(\mathbb{Z})$)*.* In the introduction of modular forms, we will be referring to matrices of $SL_2(\mathbb{Z})$ and $PSL_2(\mathbb{Z})$. Recalling that

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

we note that such matrices can be constructed by choosing two non-zero and relatively prime integers a and b, then running Euclid's algorithm to arrive at the Bézout identity of $gcd(a,b)=1$, i.e.,

$$gcd(a,b)=1 \implies \exists\, c,d,\in \mathbb{Z} \,|\, ad-bc=1.$$

We now introduce the concept of congruence subgroups and illustrate a few of the important ones in the group $\Gamma = SL_2(\mathbb{Z})$, as these will play an important role in the development of modular forms on subgroups.

Definition 3.11 (**Congruence Subgroups**)**.** A congruence subgroup of a matrix group with integer entries is a subgroup defined by congruence conditions on the entries.

Definition 3.12 (**Reduction Modulo** n)**.** For $n \in \mathbb{N}$, there is a homomorphism

$$\pi_n : \Gamma = SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/n\mathbb{Z})$$
$$\pi_n[\begin{pmatrix} a & b \\ c & d \end{pmatrix}] = \begin{pmatrix} a \pmod n & b \pmod n \\ c \pmod n & d \pmod n \end{pmatrix}.$$

Theorem 3.13 (π_n **is a Homomorphism**)**.** *The mapping π_n introduced in Definition 3.12 is a group homomorphism from the multiplicative matrix group $\Gamma = SL_2(\mathbb{Z})$ to the multiplicative matrix group $SL_2(\mathbb{Z}/n\mathbb{Z})$.*

Proof. This is the case because modular multiplication and addition are well-defined over the integers $\mathbb{Z}$ for any modulus n, so the result of the mapping by π_n of a product of two matrices in $\Gamma = SL_2(\mathbb{Z})$ is equal to the matrix product of the mappings of each matrix. If $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ and $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$, then we have

$$\begin{aligned}
\pi_n(\gamma_1\gamma_2) &= \begin{pmatrix} a_1a_2+b_1c_2 \pmod n & a_1b_2+c_1d_2 \pmod n \\ c_1a_2+d_1c_2 \pmod n & c_1b_2+d_1d_2 \pmod n \end{pmatrix} \\
&= \begin{pmatrix} a_1 \pmod n a_2 \pmod n + b_1 \pmod n & c_2 \pmod n \pmod n & ... \\ & ... & ... \end{pmatrix} \\
&= \pi_n(\gamma_1)\pi_n(\gamma_2).
\end{aligned}$$

□

Definition 3.14 (**Principal Congruence Subgroup of Level** n)**.** The principal congruence subgroup of level n in $\Gamma = SL_2(\mathbb{Z})$ is the kernel of π_n and is denoted by $\Gamma(n)$, i.e., it is

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

so it is the set of matrices whose image by π_n (entry-wise reduction modulo n) is the congruence class of the identity matrix $I_{2\times 2}$ in $SL_2(\mathbb{Z}/n\mathbb{Z})$.

Theorem 3.15 ($\Gamma(n)$ **is a Finite Index Normal Subgroup**)**.** $\Gamma(n)$ *is a normal subgroup. Furthermore, it is a finite index subgroup in* Γ*, i.e.,*

$$[SL_2(\mathbb{Z}) : \Gamma(n)] = [\Gamma : \Gamma(n)] < \infty.$$

Proof. By the isomorphism theorem proven in Theorem 2.54, the kernel of a group homomorphism is a normal subgroup, therefore $\Gamma(n) = \ker(\pi_n)$ is a normal subgroup. In addition, and by that same theorem, the image subgroup $\text{im}(\pi_n)$ is isomorphic to the quotient group

$SL_2(\mathbb{Z}/\ker(\pi_n))$.

Since the cardinality of $\mathbb{Z}/n\mathbb{Z}$ is finite (with n elements), the cardinality of $SL_2(\mathbb{Z}/n\mathbb{Z})$ is also finite, therefore the cardinality of $\mathrm{im}(\pi_n)$ is finite since it is a subgroup of $SL_2(\mathbb{Z}/n\mathbb{Z})$. Since this image is isomorphic to the quotient group $SL_2(\mathbb{Z})/\ker(\pi_n)$, we deduce that $SL_2(\mathbb{Z})/\ker(\pi_n) = SL_2(\mathbb{Z})/\Gamma(n)$ has finite cardinality (equal to the cardinality of $\mathrm{im}(\pi_n)$). But this cardinality is the index of the subgroup $\Gamma(n)$ in $SL_2(\mathbb{Z})$. □

Remark 3.16 (**Index of** $\Gamma(n)$). It can be shown -- but we won't prove it here -- that the actual (finite) index of $\Gamma(n)$ is:

$$[SL_2(\mathbb{Z}) : \Gamma(n)] = [\Gamma : \Gamma(n)] = n^3 \cdot \prod_{\substack{p|n \\ p \text{ prime}}} \left(1 - \frac{1}{p^2}\right).$$

Definition 3.17 (**Congruence Subgroup of Level n**)**.** If H is a subgroup of $\Gamma = SL_2(\mathbb{Z})$, then it is a *congruence subgroup of level* n if H contains the principal congruence subgroup $\Gamma(n)$ and if n is the smallest integer -- therefore leading to the largest subgroup $\Gamma(n)$ -- for which $\Gamma(n) \subset H$.

Definition 3.18 (**Hecke Congruence Subgroup** $\Gamma_0(n)$)**.** The pre-image by π_n of the group of upper triangular matrices in $SL_2(\mathbb{Z}/n\mathbb{Z})$ is called the *Hecke congruence subgroup* $\Gamma_0(n)$ and is therefore

$$\Gamma_0(n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma : c \equiv 0 \pmod{n} \right\}$$

Example. An example which we will use when we explore modular forms on subgroups of $SL_2(\mathbb{Z})$ is the group of matrices with an even lower left entry, i.e.,

$$\Gamma_0(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : c \equiv 0 \pmod{2} \right\}$$

Theorem 3.19 ($\Gamma_0(n)$ **Finite Index Subgroup**)**.** $\Gamma_0(n)$ *is a finite index subgroup of* $SL_2(\mathbb{Z})$.

Proof. In short, it is a subgroup because the identity matrix $I_{2\times 2}$ is in $\Gamma_0(n)$ and it is closed under matrix multiplication and inversion. Indeed, if $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ and $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$, with $c_1 \equiv 0 \pmod{n}$ and $c_2 \equiv 0 \pmod{n}$, then the lower left entry of the matrix product $\gamma_1\gamma_2$ is $c_1a_2 + d_1c_2 \equiv 0 \pmod{n}$, which shows closure under matrix multiplication. In addition, if $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(n)$, then $c \equiv 0 \pmod{n}$. And we have the following, noting that $ad - bc = 1$ because $\gamma \in SL_2(\mathbb{Z})$

$$\gamma^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

and $c \equiv 0 \pmod{n} \implies -c \equiv 0 \pmod{n} \implies \gamma^{-1} \in \Gamma_0(n)$. $\Gamma_0(n)$ is therefore a subgroup.

To show that $\Gamma_0(n)$ has finite index, we note that $\Gamma(n) \subset \Gamma_0(n)$ because the conditions on $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ being in $\Gamma(n)$ are $a, d \equiv 1 \pmod{n}$ and $b, c \equiv 0 \pmod{n}$ which are a

superset of the condition $c \equiv 0 \pmod{n}$ which defines $\Gamma_0(n)$. Therefore the cardinality of $SL_2(\mathbb{Z})/\Gamma_0(n)$, which is the index of $\Gamma_0(n)$, is less than the cardinality of $SL_2(\mathbb{Z})/\Gamma(n)$ which is the index of $\Gamma(n)$. And we have proven in Theorem 3.15 that $\Gamma(n)$ had finite index in $SL_2(\mathbb{Z})$. □

Remark 3.20 (**Index of $\Gamma_0(n)$**). It can be shown -- but we won't prove it here -- that the index of $\Gamma_0(n)$ is

$$[SL_2(\mathbb{Z}) : \Gamma_0(n)] = [\Gamma : \Gamma_0(n)] = n \cdot \prod_{\substack{p|n \\ p \text{ prime}}} \left(1 + \frac{1}{p}\right).$$

We now move our focus back to the group $\Gamma = SL_2(\mathbb{Z})$ and we examine a generating subgroup of that group, which will simplify our work on modular forms when we get to that point in the paper.

Definition 3.21 (S **and** T **in** $SL_2(\mathbb{Z})$)**.** We introduce the following two matrices, and we will subsequently prove that they generate $SL_2(\mathbb{Z})$, and play a fundamental role in understanding modular forms.

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

Theorem 3.22. *The matrices S and T as introduced in Definition 3.21 satisfy the following properties*

$$S^2 = -I_{2\times 2}, \qquad (ST)^3 = I_{2\times 2}.$$

Proof. This can be verified by simple matrix multiplication. □

Theorem 3.23 (S **and** T **Generate** $SL_2(\mathbb{Z})$)**.** *The group $\langle S, T\rangle$ generated by S and T is all of $SL_2(\mathbb{Z})$.*

Proof. We show that every integral (i.e., with integer entries) matrix with determinant 1, i.e.,

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ with } ad - bc = 1$$

can be expressed as a product of the matrices S and T or their inverses, i.e., $\gamma \in \langle S, T\rangle$. We first examine the effects of each of S and T on the matrix γ, and we have

$$S\gamma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ -a & -b \end{pmatrix},$$

and we verify that, as expected, we have

$$\det(S\gamma) = c(-b) - (-a)d = ad - bc = 1.$$

We then have

$$T\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ c & d \end{pmatrix},$$

which by a trivial induction implies for $n \in \mathbb{Z}$,

$$T^n\gamma = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+nc & b+nd \\ c & d \end{pmatrix},$$

and we verify that, as expected, we have

$$\det(T^n\gamma) = (a+nc)d - c(b+nd) = ad - bc = 1.$$

Choosing $\gamma = I_{2\times 2}$ we see in particular that

$$T^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$$

We now show that any matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL_2(\mathbb{Z})$ can be transformed into another matrix with lower-left diagonal element equal to 0 by repeated left-multiplication by S and powers of T (which can be of either algebraic sign).

If $c = 0$, we are done. Otherwise, if $|c| > |a|$, then by left-applying S once, a gets replaced with c and c gets replaced with $-a$ such that we can bring things back to the higher-left diagonal element being greater or equal in absolute value to the lower-left diagonal element, i.e., $|a| \geq |c|$. With this in place, we can write a in the form of its division with remainder by c, so we have

$$a = cq + r, \text{ with } q, r \in \mathbb{Z} \text{ and } 0 \leq r < |c|.$$

By now left-multiplying by T^{-q}, we have a matrix $T^{-q}\gamma$ with an upper left entry $a - qc = r$ which is now smaller in absolute value than the lower left entry that remained unchanged and equal to c in $T^{-q}\gamma$. Left-multiplying by S places c in the upper left position and $-r = -(a - qc)$ in the lower left position, with $|c| > |-r|$. If $r = 0$, then we are done.

Otherwise, we can repeat the process by applying the integer division algorithm to c divided by r which gives us a remainder strictly less than $|r|$ in absolute value. This means that we have a procedure that yields (by repeated applications of S and the right powers of T) a monotonous, strictly decreasing sequence of positive integers in the lower left entry of the resulting matrix.

A strictly decreasing monotonous sequence of positive integers must eventually reach 0 in a finite number of steps, so we have proven our intermediate result that we can transform our initial matrix γ into a matrix with a lower left entry equal to 0, and this transformed matrix has a determinant of 1. But now that the lower left entry is 0, and with the entries of the matrix being integers, the only possibility is that the main diagonal elements of this matrix are either both 1 or both -1, i.e., we have transformed our initial matrix γ into a matrix of the form

$$\begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} = T^k \qquad \text{or} \qquad \begin{pmatrix} -1 & -k \\ 0 & -1 \end{pmatrix} = -T^k \text{ for some } k \in \mathbb{Z},$$

and we note that two such matrices are indistinguishable in the quotient group $PSL_2(\mathbb{Z}$ which is the quotient group of $SL_2(\mathbb{Z})$ by the subgroup $\langle I_{2\times 2}, -I_{2\times 2}\rangle$ in which a matrix and its opposite (additive inverse) are in the same equivalence class.

Therefore, by left-multiplying one more time the matrix T^{-k}, we arrive at the identity matrix $I_{2\times 2}$. If we summarize what we have done, we have found a matrix $\alpha \in \langle S, T\rangle$ (the subgroup of $SL_2(\mathbb{Z})$ generated by the matrices S and T), such that $\alpha\gamma = I_{2\times 2}$. Equivalently, we have a found a matrix $\alpha \in \langle S, T\rangle$ such that $\gamma = \alpha^{-1}$ which is itself in $\langle S, T\rangle$ because a

subgroup is closed under taking the inverse of any of its elements.

We have therefore shown that the subgroup $\langle S, T\rangle$ generates the entire group $LS_2(\mathbb{Z})$ because we have shown that for any matrix $\gamma \in SL_2(\mathbb{Z})$, we also have $\gamma \in \langle S, T\rangle$, i.e., $SL_2(\mathbb{Z}) \subseteq \langle S, T\rangle \subseteq SL_2(\mathbb{Z})$. This shows that

$$PSL_2(\mathbb{Z}) = \langle S, T\rangle.$$

□

Corollary 3.24 (**Presentation of** $SL_2(\mathbb{Z})$)**.** *A presentation of* $SL_2(\mathbb{Z})$ *is*

$$PSL_2(\mathbb{Z}) = \langle S,\, T \mid S^2 = -I_{2\times 2},\, (ST)^3 = I_{2\times 2}\rangle.$$

Corollary 3.25 (**Presentation of** $PSL_2(\mathbb{Z})$)**.** *A presentation of* $PSL_2(\mathbb{Z})$ *is*

$$PSL_2(\mathbb{Z}) = \langle S,\, T \mid S^4 = I_{2\times 2},\, (ST)^3 = I_{2\times 2}\rangle.$$

Remark 3.26. As a matter of notation, since $-I_{2\times 2}$ is indistinguishable from $I_{2\times 2}$ in the quotient group $PSL_2(\mathbb{Z})$ which is the quotient of the group $SL_2(\mathbb{Z})$ by the group generated by $I_{2\times 2}$ and $-I_{2\times 2}$, i.e., the quotient by $\langle I_{2\times 2}, -I_{2\times 2}\rangle$, it is appropriate to turn the relation $S^2 = -I_{2\times 2}$ into the relation $S^4 = I_{2\times 2}$ since there is no existence of an element $-I_{2\times 2}$ distinctly from $I_{2\times 2}$ in the quotient group $PSL_2(\mathbb{Z})$.

4. Group Action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}$

Definition 4.1 (**Fractional Linear Transformation**)**.** We define a fractional linear transformation to be an invertible transformation (FLT) on $\mathbb{C}$ which can be expressed as a ratio (fraction) of two linear functions on $\mathbb{C}$. More formally, we define a fractional linear transformation as

$$z \to \frac{az+b}{cz+d}, \text{ with } a, b, c, d \in \mathbb{C} \text{ and } ad - bc \neq 0.$$

In the context of Modular Forms, we will be interested in the FLTs where $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$. In addition, we will be interested in the effect of these transformations on the half-plane $\mathbb{H} = \{z \in \mathbb{C} : \Im(z) > 0\}$. We show our first result in the following theorem.

Theorem 4.2. *For any* $a, b, c, d \in \mathbb{Z}$*, with* $ad - bc > 0$*, the function*

$$f : \mathbb{H} \to \mathbb{H}$$

$$f(z) = \frac{az+b}{cz+d}$$

is a bijection, i.e., a transformation of $\mathbb{H}$*.*

Proof. We must prove that the image of any element of $\mathbb{H}$ is also in $\mathbb{H}$, which is equivalent to proving that the imaginary part fo the image is positive. With c, d being integers, therefore reals and invariant by complex conjugation, we first note that

$$\overline{cz+d} = c\overline{z} + d \implies (cz+d)(c\overline{z}+d) = |cz+d|^2.$$

and therefore we have

$$f(z) = \frac{az+b}{cz+d} = \frac{(az+b)(c\overline{z}+d)}{|cz+d|^2} = \frac{(ac|z|^2+bd) + (adz+bc\overline{z})}{|cz+d|^2}.$$

Since the denominator is evidently real, and the part $(ac|z|^2 + bd)$ of the numerator is also real ($|z|^2 \in \mathbb{R}$ and $a, b, c, d \in \mathbb{Z} \subset \mathbb{R}$) is also real, we now can write

$$\mathfrak{I}[f(z)] = \mathfrak{I}\left[\frac{adz + bc\overline{z}}{|cz+d|^2}\right] = \frac{\mathfrak{I}(adz - bcz)}{|cz+d|^2} = \frac{ad-bc}{|cz+d|^2}\mathfrak{I}(z).$$

We note that the only possibility for $|cz + d|$ to be zero would be for $cz = -d \in \mathbb{Z}$ which would necessarily imply that $c = 0$ because $\mathfrak{I}(z) > 0$, and this would imply that $d = 0$. But then $c = d = 0$ would prevent the assumption $ad-bc > 0$, so we are certain that $|cz+d| > 0$.

With $\mathfrak{I}(z) > 0$ by assumption, and $|cz + d|^2 > 0$, and $ad - bc > 0$ by assumption, we see that $\mathfrak{I}[f(z)] > 0$, i.e., that $f(z) \in \mathbb{H}$.

We now show that f is injective

$$\begin{aligned} f(z) = f(z') &\implies \frac{az+b}{cz+d} = \frac{az'+b}{cz'+d} \\ &\implies aczz' + adz + bcz' + bd = aczz' + adz' + bcz + bd \\ &\implies (ad-bc)z = (ad-bc)z' \\ &\implies z = z' \text{ because } ad - bc > 0. \end{aligned}$$

We finally show that f is surjective by showing that for every $z \in \mathbb{H}$, there is a $w = \frac{dz-b}{-cz+a}$ such that $f(w) = z$. This is an easy verification:

$$f(w) = \frac{a\frac{dz-b}{-cz+a} + b}{c\frac{dz-b}{-cz+a} + d} = \frac{(ad-bc)z}{ad-bc} = z.$$

□

We now show the important result that we have a group action, i.e., an isomorphism from $PSL_2(\mathbb{Z})$ to $\mathbb{H}$ by mapping a matrix to a fractional linear transformation on $\mathbb{H}$.

Theorem 4.3 (**Group Action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}$**). *The map*

$$\begin{aligned} PSL_2(\mathbb{Z}) &\to \mathrm{Aut}(\mathbb{H}) \\ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\to f_\gamma(z) = \frac{az+b}{cz+d} \end{aligned}$$

is a group action on $\mathbb{H}$, i.e., it is an isomorphism from the multiplicative group $PSL_2(\mathbb{Z})$ to the group of automorphisms on $\mathbb{H}$ with composition of functions. We say that the group $PSL_2(\mathbb{Z})$ acts on the upper half plane $\mathbb{H}$ via fractional linear transformations.

Proof. We must show that the map is a bijection, and that it is an isomorphism, i.e., that the fractional linear transformation induced by a product of matrices from $PSL_2(\mathbb{Z})$ is equal to the composition of the two fractional linear transformations induced by each of the matrices, respectively.

The proof of the surjection will be skipped, as we have developed it in Complex Analysis theory by way of mappings from the upper half plane $\mathbb{H}$ to the unit disk $\mathbb{D}$ and back.

The proof of the injection amounts to equating two fractional linear transformations for all $z \in \mathbb{H}$, then asserting the equality at $z = i$, and at $z = iN$ with $N \in \mathbb{N}$ and letting $N \to \infty$ and at $z = i\varepsilon$ with $\varepsilon \in \mathbb{R}^+$ and letting $\varepsilon \to 0+$. The calculations lead us to the interesting result that the fractional linear transformations are identical if and only if they are respectively induced by a matrix γ and another matrix $\gamma' = \pm\gamma$. But then since we are considering the matrices γ in the quotient group $PSL_2(\mathbb{Z}$ (i.e., the projective special linear group as opposed to the special linear group), this is precisely the case where two elements of the same equivalence class are indistinguishable, so that $-\gamma \equiv \gamma$ in this quotient group.

We now show the homomorphism property by examining the action induced by the product of two matrices, and showing that it is equal to the composition of the induced actions by each of the matrices. Let $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ and let $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ be both from $PSL_2(\mathbb{Z})$, so we have

$$\gamma_2\gamma_1 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} = \begin{pmatrix} a_2a_1 + b_2c_1 & a_2b_1 + b_2d_1 \\ c_2a_1 + d_2c_1 & c_2b_1 + d_2d_1 \end{pmatrix},$$

so that this resulting matrix induces the fractional linear transformation

$$f_{\gamma_2\gamma_1}(z) = \frac{(a_2a_1 + b_2c_1)z + (a_2b_1 + b_2d_1)}{(c_2a_1 + d_2c_1)z + (c_2b_1 + d_2d_1)}.$$

Meanwhile, if we compose the two FLTs $(f_{\gamma_2} \circ f_{\gamma_1})$, we get

$$\begin{aligned}
(f_{\gamma_2} \circ f_{\gamma_1})(z) &= f_{\gamma_2}(f_{\gamma_1}(z)) \\
&= f_{\gamma_2}\left(\frac{a_1z + b_1}{c_1z + d_1}\right) \\
&= \frac{a_2\frac{a_1z+b_1}{c_1z+d_1} + b_2}{c_2\frac{a_1z+b_1}{c_1z+d_1} + d_2} \\
&= \frac{(a_2a_1 + b_2c_1)z + (a_2b_1 + b_2d_1)}{(c_2a_1 + d_2c_1)z + (c_2b_1 + d_2d_1)} \\
&= f_{\gamma_2\gamma_1}(z).
\end{aligned}$$

We have therefore shown the homomorphic property and this, along with the sketch of the proof of bijection, shows that we have a group isomorphism, and therefore a group action. □

Proposition 4.4 (**Actions of S and T on $\mathbb{H}$**). *The action induced on the half-plane $\mathbb{H}$ by the matrix $S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is*

$$f_S : \mathbb{H} \to \mathbb{H}$$
$$f_S(\tau) = -\frac{1}{\tau}$$

and the action on $\mathbb{H}$ induced by the matrix $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is

$$f_T : \mathbb{H} \to \mathbb{T}$$
$$f_T(\tau) = \tau + 1$$

Proof. From the general case that a matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL_2(\mathbb{Z})$ induces the fractional linear transformation

$$f_\gamma : \mathbb{H} \to \mathbb{H}$$
$$f_\gamma(\tau) = \frac{a\tau + b}{c\tau + d},$$

we easily see that the fractional linear transformation corresponding to S is

$$f_S(\tau) = \frac{0 \cdot \tau + 1}{(-1) \cdot \tau + 0} = -\frac{1}{\tau}.$$

Similarly, we have

$$f_T(\tau) = \frac{1 \cdot \tau + 1}{0 \cdot \tau + 1} = \tau + 1.$$

□

Remark 4.5 (**Interpreting the Effects of T and S Actions**). We note that the transformation induced by T (i.e., the action of T) makes it such that the orbit of a given point $\tau \in \mathbb{H}$ has translates of τ with a period of 1 in the "real" (horizontal) direction of the complex upper half plane. In particular, this transformation leaves the imaginary part of τ intact.
In addition, the transformation induced by S (i.e., the action of S) sends points in $\mathbb{H}$ that are outside the upper half of the unit circle into the upper half of the unit circle, and vice-versa.

Proof. The first statement about the effect of T is obvious, so we focus on proving the second statement, relative to S. Let $\tau = x + iy \in \mathbb{H}$. We have

$$f_S(\tau) = -\frac{1}{\tau} = -\frac{1}{x + iy} = \frac{-x + iy}{x^2 + y^2} \implies |f_S(\tau)| = \frac{1}{\sqrt{x^2 + y^2}} = \frac{1}{|\tau|}.$$

This shows that if $|\tau| < 1$, then $|f_S(\tau)| > 1$, and if $|\tau| > 1$, then $|f_S(\tau)| < 1$. □

Theorem 4.6 (**Fundamental Domain of Action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}$**). *The region of $\mathbb{H}$ defined by*

$$\mathcal{F} = \left\{ \tau \in \mathbb{H} : |\Re(\tau)| \leq \frac{1}{2}, |\tau| \geq 1 \right\}$$

is a Fundamental Domain (by Definition 2.62) of the group action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}$, recalling that a fundamental domain is a set in which the orbit by actions of $PSL_2(\mathbb{Z})$ of every element of $\mathbb{H}$ is represented.
Said otherwise, for every $\tau \in \mathbb{H}$, there is some $\gamma \in PSL_2(\mathbb{Z})$ such that $f_\gamma(\tau) \in \mathcal{F}$. And since we have proven that $\langle S, T \rangle = PSL_2(\mathbb{Z})$, i.e., the subgroup generated by S and T generates all of $PSL_2(\mathbb{Z})$, a corollary is that there exists some composition of actions of S, T, their powers, and the powers of their inverses, which brings the image of τ by that composition into $\mathcal{F}$.

Proof. For any $\tau \in \mathbb{H}$, if τ is not already in the vertical strip $\{z \in \mathbb{H} : |\Re(z)| \leq \frac{1}{2}\}$, we can first have an image of τ in that vertical strip as follows.

Let $f_{T^{-\lfloor \Re(\tau) \rfloor}}$ be the action that iterates T^{-1} a number of times equal to the floor (i.e., integer part) of the real part of τ. This action produces an image of τ that is in the vertical strip $\{z \in \mathbb{H} : 0 \leq \Re(z) < 1\}$. If this image has a real part less than or equal to $\frac{1}{2}$, then we are done

because this image is in the vertical strip $\{z \in \mathbb{H} : 0 \leq \Re(z) \leq \frac{1}{2}\} \subset \{z \in \mathbb{Z} : |\Re(z) \leq \frac{1}{2}\}$. Otherwise, the real part of this image is greater than $\frac{1}{2}$, so by applying the action of T^{-1} one more time, the resulting image will be in $\{z \in \mathbb{Z} : -\frac{1}{2} < \Re(z) < 0\} \subset \{z \in \mathbb{Z} : |\Re(z) \leq \frac{1}{2}\}$.

So we can assume that for every $\tau \in \mathbb{H}$, its orbit has a representative in the vertical strip $\{z \in \mathbb{Z} : |\Re(z) \leq \frac{1}{2}\}$. We now seek to show that there is a representative of the orbit in $\mathcal{F}$. If the representative of the orbit that is in the vertical strip $\{z \in \mathbb{Z} : |\Re(z) \leq \frac{1}{2}\}$ also satisfies $|z| \geq 1$, then we are done as the combination of these two conditions define $\mathcal{F}$.

Otherwise, this representative of the orbit of τ that we have on hand is in the upper half of the unit disk and satisfies $|\Re(z)| \leq \frac{1}{2}$, so it is of the form $\rho e^{i\theta}$ with $\frac{\pi}{3} \leq \theta \leq \frac{2\pi}{3}$ and $\rho < 1$.

We then apply to it the action of S which transforms z to $-\frac{1}{z}$, so that the new image becomes $-\frac{1}{\rho e^{i\theta}} = -\frac{e^{-i\theta}}{\rho} = \frac{e^{i(\pi-\theta)}}{\rho}$, with $\frac{1}{\rho} > 1$ and $\frac{\pi}{3} \leq \pi - \theta \leq \frac{2\pi}{3}$, i.e., with a modulus greater than 1 and a real part less than or equal to $\frac{1}{2}$ in absolute value. This is the definition of $\mathcal{F}$, so we have a representative of the orbit of τ in $\mathcal{F}$. □

5. Modular Forms Introduction

The main idea behind modular forms is that they are holomorphic functions over the upper half of the complex plane that exhibit a certain invariance under composition with fractional linear transformations (modularity conditions), in addition to having certain regularity regularity conditions. We formally introduce them now.

Definition 5.1 (**Modular Form**)**.** For $k \in \mathbb{Z}$, a Modular Form of weight k for $SL_2(\mathbb{Z})$ is a function $f : \mathbb{H} \to \mathbb{C}$ satisfying the following three conditions:

- f is holomorphic on $\mathbb{H}$.
- Modularity: $f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau + d)^k f(\tau)$ for *all matrices* $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ and all $\tau \in \mathbb{H}$.
- $f(\tau)$ is bounded as $\tau \to i\infty$, i.e., as $\tau = \alpha + i\beta$ and $\beta \to \infty$.

Remark 5.2. We note that the modularity condition amounts to an infinite set of conditions since the equality must be true for all matrices in $SL_2(\mathbb{Z})$. However, we have shown in Theorem 3.23 that the two matrices $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ generate the entire multiplicative group $SL_2(\mathbb{Z})$. We also recall from Theorem 4.3 that we have a group action from $PSL_2(\mathbb{Z})$ to $\mathbb{H}$ by way of induced fractional linear transformations on $\mathbb{H}$.

The modularity condition applied to these two group-generating matrices yields the following two conditions

$$\text{Condition on } S : f(-\frac{1}{\tau}) = \tau^k f(\tau),$$
$$\text{Condition on } T : f(\tau + 1) = f(\tau).$$

Remark 5.3 (**Defining with** $SL_2(\mathbb{Z})$ **vs** $PSL_2(\mathbb{Z})$). It is worth noting a subtle difference between defining the modular forms for the group $SL_2(\mathbb{Z})$ as opposed to doing the same for the group $PSL_2(\mathbb{Z})$. The group $PSL_2(\mathbb{Z})$ is the quotient group of $SL_2(\mathbb{Z})$ by the subgroup

$\langle I_{2\times 2}, -I_{2\times 2}\rangle$ generated by the identity matrix and its additive inverse.

The advantage of defining modular forms for the main group $SL_2(\mathbb{Z})$ is that we get an easy proof that there are no modular forms of odd weight, and this is straightforward to prove (see Theorem 5.5 below) by singling out the modularity condition on the matrix $-I_{2\times 2}$. A disadvantage, however, is that the matrix $I_{2\times 2}$ and the matrix $-I_{2\times 2}$ are distinguishable in $SL_2(\mathbb{Z})$ even though their induced fractional linear transformations are identical as transformations on $\mathbb{H}$. This is generally true for any pair of opposite matrices $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $-\gamma = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$ due to the fact that $\frac{a\tau+b}{c\tau+d} = \frac{(-a)\tau+(-b)}{(-c)\tau+(-d)}$ for all $\tau \in \mathbb{H}$. This choice of matrix group gives us a non-injective mapping from the matrix group to the group of frational linear transformations that is induced by it.

On the other hand, the advantage of using the quotient group is that there is an isomorphic relationship (i.e., both bijective and homomorphic map), and therefore a group action, from $PSL_2(\mathbb{Z})$ (with matrix left-multiplication) to the fractional linear transforms on $\mathbb{H}$ (with function composition). A matrix $\gamma \in PSL_2(\mathbb{Z})$ is in fact an equivalence class of a matrix from $SL_2(\mathbb{Z})$ together with its additive inverse, and this uniquely corresponds to a fractional linear transformation on $\mathbb{H}$ induced by it.

We now verify that the modularity condition is preserved under matrix multiplication and matrix inversion, and this lets us reduce the modularity condition to being met on the two generating matrices of $SL_2(\mathbb{Z})$, i.e., S and T.

Theorem 5.4 (**Modularity Satisfied for Products, Inverses**). *Let* $\gamma_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ *and* $\gamma_2 = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ *be two matrices in* $SL_2(\mathbb{Z})$ *and let* f_{γ_1} *and* f_{γ_2} *be the induced fractional linear transformations, respectively. If a function* $g : \mathbb{H} \to \mathbb{C}$ *satisfies the modularity condition with weight* k *for* f_{γ_1} *and* f_{γ_2}*, then it satisfies the modularity condition with weight* k *for* $f_{\gamma_1} \circ f_{\gamma_2}$ *which is induced by* $\gamma_1\gamma_2$ *and for* $f_{\gamma_1^{-1}}$ *which is induced by* γ_1^{-1}*.*

Proof. We start by proving the result for the action induced by the matrix product $\gamma_1\gamma_2$. We have

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix},$$

so that satisfying the modularity condition with weight k for the action of the product matrix $\gamma_1\gamma_2$ would mean satisfying

$$g(f_{\gamma_1\gamma_2})(\tau) = [(c_1a_2 + d_1c_2)\tau + (c_1b_2 + d_1d_2)]^k g(\tau). \quad (!)$$

We know from the group action being an isomorphism, and from the assumption that g satisfies the modularity condition with weight k for γ_1 and γ_2, that we have

$$\begin{aligned} f_{\gamma_1\gamma_2}(\tau) &= (f_{\gamma_1} \circ f_{\gamma_2})(\tau) \\ &= f_{\gamma_1}(f_{\gamma_2}(\tau)) \\ &= (c_1 f_{\gamma_2}(\tau) + d_1)^k (c_2\tau + d_2)^k g(\tau) \\ &= \left(c_1 \frac{a_2\tau + b_2}{c_2\tau + d_2} + d_1\right)^k (c_2\tau + d_2)^k g(\tau) \\ &= (c_1 a_2 \tau + c_1 b_2 + d_1 c_2 \tau + d_1 d_2)^k g(\tau) \\ &= [(c_1 a_2 + d_1 c_2)\tau + (c_1 b_2 + d_1 d_2)]^k g(\tau). \end{aligned}$$

We see that the last expression in the series of equalities above matches exactly the expression (!) for satisfying the modularity condition with weight k for the fractional linear transformation induced by $\gamma_1\gamma_2$.

We now show the result for the inverse, and we start by noting that the inverse of γ_1 is

$$\gamma_1^{-1} = \begin{pmatrix} d_1 & -b_1 \\ -c_1 & a_1 \end{pmatrix}.$$

We apply the modularity condition for f_{γ_1} to $f_{\gamma_1^{-1}}(\tau)$ and we get

$$g(\tau) = g(f_{\gamma_1}(f_{\gamma_1^{-1}}(\tau))) = (c_1 f_{\gamma_1^{-1}}(\tau) + d_1)^k g(f_{\gamma_1^{-1}}(\tau)),$$

which implies by dividing both sides that

$$g(f_{\gamma_1^{-1}}(\tau)) = \frac{1}{(c_1 f_{\gamma_1^{-1}}(\tau) + d_1)^k} g(\tau). \quad (!!)$$

We now calculate the expression in the denominator and we have

$$(c_1 f_{\gamma_1^{-1}}(\tau) + d_1)^k = \left(c_1 \frac{d_1\tau - b_1}{-c_1\tau + a_1} + d_1\right)^k = \left(\frac{a_1 d_1 - b_1 c_1}{-c_1\tau + a_1}\right)^k = \frac{1}{(-c_1\tau + a_1)^k}.$$

We plug this expression back into (!!) and we get

$$g(f_{\gamma_1^{-1}}(\tau)) = (-c_1\tau + a_1)^k g(\tau).$$

But this is exactly the modularity condition with weight k for the action of the matrix γ_1^{-1}, so we have completed the proof of our theorem. □

Theorem 5.5 (**No Non-Zero Odd Weight Modular Form**). *There are no non-trivial modular forms with odd weight for $SL_2(\mathbb{Z})$.*

Proof. We note that since the modularity condition must be true for all matrices in $SL_2(\mathbb{Z})$, it must also be true for the matrix $-I_{2\times 2} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. But the modularity condition then becomes

$$f(\tau) = (-1)^k f(\tau).$$

This is only possible if k is even, or if the function $f(\tau)$ is the constant zero function, which is a trivial case. □

Remark 5.6. We note that it is possible to define modular forms on other groups than $PSL_2(\mathbb{Z})$ or $SL_2(\mathbb{Z})$, and in particular it is possible to define modular forms on subgroups of $SL_2(\mathbb{Z})$ or of $PSL_2(\mathbb{Z})$. In those cases, it is possible to have modular forms with odd weight, as long as these subgroups do not include the matrix $-I_{2\times 2}$. We will such examples further in the paper.

Given that the matrices S and T generate the groups $SL_2(\mathbb{Z})$ and $PSL_2(\mathbb{Z})$, and given that the modularity condition is preserved under matrix multiplication and inverse, we can state an alternative formulation of the definition of modular forms. We also present an alternative formulation to the third condition in the alternative definition below

Definition 5.7 (**Alternative Definition**)**.** For $k \in \mathbb{Z}$, a Modular Form of weight k for $SL_2(\mathbb{Z})$ is a function $f : \mathbb{H} \to \mathbb{C}$ satisfying the following three conditions:

- f is holomorphic on $\mathbb{H}$.
- Modularity: $f(\tau + 1) = f(\tau)$ and $f(-\frac{1}{\tau}) = \tau^k f(\tau)$ for all $\tau \in \mathbb{H}$.
- $f(\tau)$ converges to a limit as $\tau \to i\infty$, i.e., as $\tau = \alpha + i\beta$ and $\beta \to \infty$.

Theorem 5.8 (**Vector Spaces of Modular Forms**)**.** *The set of modular forms for $SL_2(\mathbb{Z})$ and of weight $k \in \mathbb{N}$ is a vector space on the field $\mathbb{C}$.*

Proof. It is clear that for any choice of weight k, the constant function equal to 0 on all of $\mathbb{H}$ satisfies all the conditions of a modular form, as it is holomorphic, it is bounded as $\tau \to i\infty$, and the modularity conditions are satisfied. These sets are therefore non-empty.

We now prove closure under linear combinations: if f and g are two modular forms of weight k, then it is easy to see that any of their linear combinations would be holomorphic (as the two functions are), and any linear combination is also bounded as $\tau \to i\infty$ (as the two functions are). As to the modularity conditions, we can see that if $\lambda f + \mu g$ is a linear combination of f and g (with $\lambda, \mu \in \mathbb{C}$), then

$$f(-\frac{1}{\tau}) = \tau^k f(\tau) \text{ and } g(-\frac{1}{\tau}) = \tau^k g(\tau) \implies (\lambda f + \mu f)(-\frac{1}{\tau}) = \tau^k[\lambda f(\tau) + \mu g(\tau)],$$

and

$$(\lambda f + \mu g)(\tau + 1) = (\lambda f + \mu g)(\tau).$$

Therefore each set of modular functions of weight k is either a trivial vector space with only 0 as its element, or a vector space on the field $\mathbb{C}$. □

We cite and prove one additional result that will be important in developing the results on dimensionality of the vector spaces of modular forms of given weight.

Theorem 5.9 (**Product of Modular Forms**)**.** *Let f be a modular form on $\mathbb{H}$ of weight k for $SL_2(\mathbb{Z})$ and g be a modular form on $\mathbb{H}$ of weight l for $SL_2(\mathbb{Z})$. Then the point-wise product function*

$$fg : \mathbb{H} \to \mathbb{C}$$

$$(fg)(\tau) = f(\tau)g(\tau)$$

is a modular form on $\mathbb{H}$ of weight $k + l$ for $SL_2(\mathbb{Z})$.

Proof. It is clear that the product of two holomorphic functions on $\mathbb{H}$ is also holomorphic on $\mathbb{H}$. It is also clear that if $f(\tau)$ and $g(\tau)$ are both bounded as $\tau \to i\infty$, then $(fg)(\tau) = f(\tau)g(\tau)$ is bounded as $\tau \to i\infty$.

Moving on to the modularity conditions, we have

$$(fg)(\tau+1) = f(\tau+1)g(\tau+1) = f(\tau)g(\tau) = (fg)(\tau),$$

and we have

$$(fg)\left(-\frac{1}{\tau}\right) = f\left(-\frac{1}{\tau}\right) g\left(-\frac{1}{\tau}\right) = \tau^k f(\tau)\tau^l g(\tau) = \tau^{k+l}(fg)(\tau).$$

This completes the proof that (fg) is a modular form of weight $k+l$ for $SL_2(\mathbb{Z})$. □

Corollary 5.10 (**Ratio of Modular Forms**). *Let f be a modular form of weight k for $SL_2(\mathbb{Z})$, and let g be a modular form of weight l for $SL_2(\mathbb{Z})$, such that g does not vanish anywhere on $\mathbb{H}$ (no zeros). If the point-wise ration function $\left(\frac{f}{g}\right)(\tau) = \frac{f(\tau)}{g(\tau)}$ is bounded as $\tau \to i\infty$, then the point-wise ratio function $\left(\frac{f}{g}\right)$ is a modular form of weight $k-l$ for $SL_2(\mathbb{Z})$.*

Proof. The proof is similar throughout with the proof of Theorem 5.9 so we omit laying out all of its details again for this corollary. The important difference from the previous theorem about products is the added condition that the ratio must be bounded as $\tau \to i\infty$ in order to satisfy the modular form conditions, and this boundedness condition cannot be always assumed for the ratio. □

We now proceed to introduce actual examples of modular form, by defining the Eisenstein Series. In the next section, we are assuming familiarity with lattices on the complex plane and more generally with the main concepts of complex analysis, including elliptic functions and holomorphic functions.

6. Eisenstein Series

Definition 6.1 (**Eisenstein Series**). Let $k \in \mathbb{N}$ be even and $k \geq 4$. We define the Eisenstein Series of weight k as the function expressed in elliptic function form

$$G_k : \mathbb{H} \to \mathbb{C}$$
$$G_k(\tau) = \sum_{\substack{(m,n)\in\mathbb{Z}\times\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(m\tau+n)^k}.$$

Theorem 6.2 (**Eisenstein Series as Modular Form**). *The Eisenstein series of weight $k \geq 4$ is a modular form of weight k for the group $SL_2(\mathbb{Z})$.*

Proof. We must first show that the function is well-defined (i.e., convergent). The function is in fact absolutely convergent and this can be shown by way of a Lemma that proves that for any $\tau \in \mathbb{H}$, there is a $\delta_\tau \in (0,1)$, such that $|m\tau+n| \geq \delta_\tau |mi+n|$ for all $m,n \in \mathbb{Z}$. (In fact, we can find a uniform value δ that does not depend on τ if we consider τ in a vertical strip of $\mathbb{H}$ with finite width for the real part and a strictly positive lower bound for the imaginary part, i.e., $\{z \in \mathbb{H} : |\Re(z)| \leq \alpha, \Im(\tau) \geq \beta\}$, for some $\alpha > 0, \beta > 0$.)

The lemma lets us bracket the terms of the Eisenstein series

$$0 < \frac{1}{|m\tau + n|^k} \leq \frac{1}{\delta_\tau^k |mi + n|^k} = \frac{1}{\delta_\tau^k (m^2 + n^2)^{\frac{k}{2}}}.$$

The exponent $\frac{k}{2}$ on n and m is greater than or equal to 2, so absolute convergence of $G_k(\tau)$ is a result of convergence of the the series $\sum_{\substack{(m,n)\in\mathbb{Z}\times\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(m^2+n^2)^{\frac{k}{2}}}$ when $k \geq 4 > 2$. This is an instance in 2 dimensions of a broader result on convergence of lattice sums in any number of dimensions.

We must them prove that the Eisenstein series is holomorphic on $\mathbb{H}$. We know that each of its terms is clearly holomorphic as it has no poles or singularities in $\mathbb{H}$. We recall a result from complex analysis about limits of holomorphic functions: if a sequence $\{h_n\}$ of holomorphic functions on a common open set $\Omega \subseteq \mathbb{C}$ converges uniformly on all compact subsets of Ω, then the pointwise limit $h(z) = \lim_{n\to\infty} h_n(z)$ is also holomorphic on Ω. The region Ω that lets us invoke this result for the Eisenstein series is a vertical strip of the sort $\{z \in \mathbb{H} : |\Re(z)| \leq \alpha, \Im(\tau) \geq \beta\}$, for some $\alpha > 0, \beta > 0$ that we invoked in parentheses above, and applying the Weierstra M-test (proof not detailed here).

We now prove that G_k satisfies two modularity conditions. For the first condition, we have

$$G_k(\tau + 1) = \sum_{(m,n)\neq(0,0)} \frac{1}{(m(\tau+1)+n)^k} = \sum_{(m,n)\neq(0,0)} \frac{1}{(m\tau + (m+n))^k}.$$

If we reindex the double sum by setting $m' = m$ and $n' = m+n$, then the condition $(m,n) \neq (0,0)$ is equivalent to $(m', n'-m') \neq (0,0)$ which is also equivalent to $(m',n') \neq (0,0)$, so our double sum becomes

$$G_k(\tau+1) = \sum_{(m',n')\neq(0,0)} \frac{1}{(m'\tau + n')^k} = G_k(\tau).$$

For the second condition, we have

$$G_k(-\frac{1}{\tau}) = \sum_{(m,n)\neq(0,0)} \frac{1}{[(-\frac{m}{\tau}) + n]^k} = \sum_{(m,n)\neq(0,0)} \frac{\tau^k}{(n\tau - m)^k} = \tau^k \sum_{(m,n)\neq(0,0)} \frac{1}{(n\tau - m)^k}.$$

If we reindex the double sum by setting $m' = n$ and $n' = -m$, then the condition $(m,n) \neq (0,0)$ is equivalent to $(-n', m') \neq (0,0)$ which is equivalent to $(m',n') \neq (0,0)$, so our double sum becomes

$$G_k(-\frac{1}{\tau}) = \tau^k \sum_{(m',n')\neq(0,0)} \frac{1}{(m'\tau + n')^k} = \tau^k G_k(\tau).$$

The last condition to verify is bounded behavior of $G_k(\tau)$ as $\tau \to i\infty$. Since we are interested in the limit at infinity in the imaginary direction, we can assume that $\Im(\tau) \geq L > 0$ for some $L > 0$ (say $L = 1$), and with the periodicity $G_k(\tau+1) = G_k(\tau)$, we can assume that $|\Re(\tau)| \leq \frac{1}{2}$. We are therefore in the situation of the vertical half-strip with bounded width that we described in the earlier part of this proof when sketching proof of absolute convergence of $G_k(\tau)$. So there is some $\delta > 0$ such that $|m\tau + n| \geq \delta |mi + n|$ for all τ in this strip and all $m, n \in \mathbb{Z}$.

Since $G_k(\tau)$ is absolutely convergent, we can rearrange its terms without altering the result of the sum, so we can separate the terms $m = 0$ from the terms $m \neq 0$ and write

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

where we were able to convert sums over non-zero integers into twice the sums over the natural numbers because k is even therefore $(-m\tau - n)^k = (m\tau + n)^k$. We then note that the first sum $2\sum_{n\in\mathbb{N}} \frac{1}{n^k}$ does not depend on τ so it is a constant, and therefore bounded as $\tau \to i\infty$. As to the second term $2\sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{Z}} \frac{1}{(m\tau+n)^k}$, we note that it is bounded in absolute value by $2\sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{Z}} \frac{1}{|m\tau+n|^k} \geq \frac{1}{\delta^k}\sum \frac{1}{|mi+n|^k}$ and the right-hand side converges independently of τ, so the entire sum is bounded as $\tau \to i\infty$. □

We now introduce the q-series, also known as the Fourier series representation of $G_k(\tau)$, as this representation will be very fruitful in deriving results around spaces of modular forms, their dimensions, relations between modular forms, as well as results in Number Theory.

The first intuition in the following result which leads to the Fourier representation is that if a function is periodic in the sense that $f(\tau+1) = f(\tau)$, then it is reminiscent of the function $e^{2\pi i\tau}$ which also satisfies this periodicity.

Theorem 6.3 (**First q Representation Result**). *Let $f : \mathbb{H} \to \mathbb{C}$ be holomorphic, and satisfying $f(\tau+1) = f(\tau)$ for all $\tau \in \mathbb{H}$, and with f bounded as $\tau \to i\infty$. Then there is a sequence $a_n \in \mathbb{C}$ such that for all $\tau \in \mathbb{H}$*

$$f(\tau) = \sum_{n\in\mathbb{Z}^+} a_n e^{2\pi i n\tau}.$$

In addition, $f(\tau)$ has a limit as $\tau \to i\infty$.

Proof. We will only sketch the proof to convey the idea of how the coefficients a_n are derived. If we let $q = e^{2\pi i\tau}$, with $\tau = x + iy \in \mathbb{H}$ (i.e., $y > 0$), then $q = e^{-2\pi y}e^{2\pi ix}$, and we have $0 < |q| = e^{-2\pi y} < 1$ because $y > 0$, therefore the image of $\mathbb{H}$ by the function $q(\tau)$ is the punctured unit disk $\mathbb{D}^* = \{q \in \mathbb{C} : 0 < |q| < 1\}$. Also, we have

$$e^{2\pi i\tau} = e^{2\pi i\tau'} \text{ if and ony if } \tau' = \tau + n, \text{ with } n \in \mathbb{Z},$$

so if we define the function

$$\tilde{f}(q) = f(\tau) \text{ for } q = e^{2\pi i\tau},$$

then this function $\tilde{f}$ is well-defined on the punctured unit disk.

Since f is bounded as $\tau \to i\infty$ which corresponds to $y \to \infty$ and to $q \to 0$, then $\tilde{f}(q)$ which is holomorphic on $\mathbb{D}^*$ can be analytically continued at $q = 0$ (by the Riemann singularity removal theorem), so that $\tilde{f}$ becomes holomorphic on the unit disk $\mathbb{D}$. We can therefore write $\tilde{f}$ as a power series around the point $q = 0$, so let this power series be

$$\tilde{f}(q) = \sum_{n=0}^{\infty} a_n q^n = \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau}, \text{ with } a_0 = \tilde{f}(0) = f(i\infty).$$

□

Since the result above is true even before considering the other weight-k modularity condition, it is evidently also true whenever f is a modular form because a modular form certainly satisfies the conditions of Theorem 6.3. We therefore given the definition in the context of a modular form.

Definition 6.4 (**q-expansion and Fourier Coefficients**)**.** The *q-expansion* of a modular form $f(\tau)$ is the series

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

The coefficients a_n are the *Fourier coefficients* of f.

Remark 6.5. It is frequent to see the abuse of notation in the writing $f(q)$ when writing the q-expansion for $f(\tau)$, instead of always distinguishing the variable change with a $\tilde{f}(q)$ notation.

Remark 6.6 (**Second Modularity Condition Hidden from View**). We note that while the first modularity condition $f(\tau+1) = f(\tau)$ is well encoded in the q-expansion of a modular form, the second condition $f(-\frac{1}{\tau}) = \tau^k f(\tau)$ is not visible at all in the q-expansion. In fact, by looking at a q-expansion, it is not apparent at all whether it is the q-expansion of a modular form or not.

We now state and sketch the proof of an explicit expression for the q-expansion of the Eisenstein series. On the way to doing so, we state two beautiful definitions and results from Fourier analysis, namely the introduction of the Fourier transform of a function and the Poisson summation formula that relates a function and its Fourier transform.

Definition 6.7 (**Fourier Transform**)**.** For an absolutely integrable function $f : \mathbb{R} \to \mathbb{C}$, i.e., such that $\int_{-\infty}^{\infty} |f(x)\, dx < \infty$, we define its *Fourier Transform* as the function (which is a continuous function)

$$\hat{f} : \mathbb{R} \to \mathbb{C}$$
$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{2\pi i x \xi}\, dx.$$

Lemma 6.8 (**Poisson Summation Formula**)**.** *If a function $f : \mathbb{R} \to \mathbb{C}$ as well as its Fourier transform $\hat{f} : \mathbb{R} \to \mathbb{C}$ are both continuous and absolutely integrable, then the Poisson summation formula gives the remarkable result*

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n).$$

Remark 6.9. We will not provide the proof of the Poisson summation formula but we note that it is a very powerful result (which we will use in the next theorem), that equates the sum from sampling one function over all the integers to the sum from sampling its Fourier transform over all the integers.

Lemma 6.10. *For $z \in \mathbb{H}$ and $k \geq 3$, the following equality holds as a result of the Poisson summation formula*

$$\sum_{n\in\mathbb{Z}} \frac{1}{(z+n)^k} = \frac{(-2\pi i)^k}{(k-1)!} \sum_{n\in\mathbb{N}} n^{k-1} e^{2\pi i n z}.$$

Proof. We will not give the detailed proofs but mention that the right-hand side is obtained as a result of calculating the Fourier transform of the function $f_z(x) = \frac{1}{(z+x)^k}$ using the Cauchy residue theorem on a proper contour, then summing the samples of $f_z(x)$ and of its Fourier transform over all the integers. □

Theorem 6.11. *For even $k \geq 4$, the q-expansion of $G_k(\tau)$ is*

$$G_k(\tau) = \tilde{G}_k(q) = 2\zeta(k) + \frac{2(2\pi i)^k}{(k-1)!}\sum_{n\in\mathbb{N}} \sigma_{k-1}(n)q^n,$$

where $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$ is the sum of the k^{th} powers of the divisors of n.

Proof. We only sketch the proof in its main outline, and we start from an expression of $G_k(\tau)$ that we had derived at the end of the proof of Theorem 6.2 on Eisenstein series being modular forms. We have

$$G_k(\tau) = 2\sum_{n\in\mathbb{N}} \frac{1}{n^k} + 2\sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{Z}} \frac{1}{(m\tau+n)^k} = 2\zeta(k) + 2\sum_{m\in\mathbb{N}}\left(\sum_{n\in\mathbb{Z}} \frac{1}{(m\tau+n)^k}\right).$$

By Lemma 6.10, the rightmost term in parentheses in the equation above can be expressed as

$$\sum_{n\in\mathbb{Z}} \frac{1}{(m\tau+n)^k} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{n\in\mathbb{N}} n^{k-1}e^{2\pi inm\tau},$$

where we replaced the parameter z from Lemma 6.10 with $m\tau$.
We therefore have

$$G_k(\tau) = \tilde{G}_k(q) = 2\zeta(k) + \frac{2(-2\pi i)^k}{(k-1)!}\sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} n^{k-1}e^{2\pi inm\tau}.$$

We can now reindex the double sum over the natural numbers with a single sum over the values that the product $r = nm$ of the indices can take, and regroup terms so that a term in the series has exponent $e^{2\pi ir\tau}$ with $r = nm$. It becomes clear that this term will be multiplied by a sum of terms that came from n^{k-1} where $n \mid nm$, i.e., a sum over $d \mid r$ of terms d^{k-1}. More specifically, we have

$$\begin{aligned} G_k(\tau) = \tilde{G}_k(q) &= 2\zeta(k) + \frac{2(-2\pi i)^k}{(k-1)!}\sum_{r\in\mathbb{N}}\left(\sum_{\substack{d|r\\ d\in\mathbb{N}}} d^{k-1}\right)e^{2\pi ir\tau} \\ &= 2\zeta(k) + \frac{2(-2\pi i)^k}{(k-1)!}\sum_{r\in\mathbb{N}} \sigma_{k-1}(r)\left(e^{2\pi i\tau}\right)^r \\ &= 2\zeta(k) + \frac{2(-2\pi i)^k}{(k-1)!}\sum_{r\in\mathbb{N}} \sigma_{k-1}(r)q^r. \end{aligned}$$

□

Remark 6.12. Euler had already found a formula for $\zeta(k)$ when $k \geq 2$ is even, by way of Bernoulli numbers, and it is

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

where B_k is the k^{th} Bernoulli number, i.e., it is one of the rationals appearing in the power series

$$\frac{x}{e^x-1}=\sum_{k\geq 0}\frac{B_k}{k!}x^k=1-\frac{1}{2}x+\frac{1}{12}x^2-\frac{1}{720}x^4+\dots$$

The Bernoulli numbers are 0 for odd $k>1$ and early values of Bernoulli numbers are the following:

$k:$	0	1	2	4	6	8	10	12
$B_k:$	1	$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{30}$	$\frac{1}{42}$	$-\frac{1}{30}$	$\frac{5}{66}$	$-\frac{691}{2730}$

We therefore have another expression for the Eisenstein series, using the Bernoulli numbers, which we state in the next theorem.

Theorem 6.13 (**Eisenstein Series and Bernoulli Numbers**). *We have the following*

$$\begin{aligned}G_k(\tau)=\tilde{G}_k(q)&=2\zeta(k)-\frac{4k\zeta(k)}{B_k}\sum_{n\in\mathbb{N}}\sigma_{k-1}(n)q^n\\&=2\zeta(k)\left[1-\frac{2k}{B_k}\sum_{n\in\mathbb{N}}\sigma_{k-1}(n)q^n\right].\end{aligned}$$

Since it is often convenient to have the constant term equal to 1, we add a definition for the *normalized Eisenstein series of weight k*

Definition 6.14 (**Normalized Eisenstein Series**). Foe even $k\geq 4$, we define the normalized Eisenstein series of weight k to be

$$E_k(\tau)=\tilde{E}_k(q)=\frac{\tilde{G}_k(q)}{2\zeta(k)}=1-\frac{2k}{B_k}\sum_{n\in\mathbb{N}}\sigma_{k-1}(n)q^n.$$

Example. Some examples of starting terms of q-expansions of standardized Eisenstein series are:

- $E_4(\tau)=\tilde{E}_4(q)=1+240\sum_{n\in\mathbb{N}}\sigma_3(n)q^n=1+240q+2160q^2+6720q^3+\dots$
- $E_6(\tau)=\tilde{E}_6(q)=1-504\sum_{n\in\mathbb{N}}\sigma_5(n)q^n=1-504q-16632q^2-122976q^3-\dots$
- $E_8(\tau)=\tilde{E}_8(q)=1+480\sum_{n\in\mathbb{N}}\sigma_7(n)q^n=1+480q+61920q^2+1050240q^3+\dots$
- $E_{10}(\tau)=\tilde{E}_{10}(q)=1-264\sum_{n\in\mathbb{N}}\sigma_9(n)q^n=1-264q-135432q^2-5196576q^3-\dots$
- $E_{12}(\tau)=\tilde{E}_{12}(q)=1+\frac{65520}{691}\sum_{n\in\mathbb{N}}\sigma_{11}(n)q^n=1+\frac{65520}{691}q+\frac{134250480}{691}q^2+\frac{11606736960}{691}q^3+\dots$
- $E_{14}(\tau)=\tilde{E}_{14}(q)=1-24\sum_{n\in\mathbb{N}}\sigma_{13}(n)q^n=1-24q-196632q^2-38263776q^3-\dots$

Remark 6.15. We note that for some q-expansions of standardized Eisenstein series, the a_n coefficients are integers. This is the case for $k=4,6,8,10,14$, and is consistent with the fact that the ratio $\frac{2k}{B_k}$ is an integer for these values of k. This is not the case for all values of k, however, as can be seen with the case $k=12$.

We have seen in Theorem 5.9 the result that the product (in the sense of pointwise product of functions) of a modular form of weight k with a modular form of weight l is in fact a modular form of weight $k+l$. We will see in the next section an illustration of this result in the form of some normalized Eisenstein series being the product of two others of lower weight, but this result is spurious and depends in very interesting ways on the dimensions

of the vector spaces of modular forms of a given weight, which is the subject of the following section after next.

Having introduced Eisenstein series, and derived results for their q-expansions and associated Fourier coefficients, we end this section with a question relative to whether the expression we have derived for the q-expansion of normalized Eisenstein series for $k \geq 4$ could also work for $k = 2$ and give us a modular form for $SL_2(\mathbb{Z})$ of weight 2. We show that "we can come close" to having one, but we will have a more definitive result about the existence (or lack) of modular forms of weight 2 for $SL_2(\mathbb{Z})$ in the next section.

Lemma 6.16. *For all $\tau \in \mathbb{H}$, we have the equality*

$$\frac{1}{\Im\left(-\frac{1}{\tau}\right)} = \tau^2 \frac{1}{\Im(\tau)} - 2i\tau.$$

Proof. Let $\tau = x + iy$, $x, y, \in \mathbb{R}$, so that $y = \Im(\tau)$. We have

$$-\frac{1}{\tau} = -\frac{1}{x+iy} = -\frac{x-iy}{x^2+y^2} = \frac{-x+iy}{x^2+y^2}.$$

This implies that

$$\Im\left(-\frac{1}{\tau}\right) = \frac{y}{x^2+y^2},$$

and

$$\frac{1}{\Im\left(-\frac{1}{\tau}\right)} = \frac{x^2+y^2}{y}.$$

We now calculate the right-hand side of our desired equality and we have

$$\begin{aligned}\tau^2 \frac{1}{\Im(\tau)} - 2i\tau &= \frac{x^2 - y^2 + 2ixy}{y} - 2i(x+iy) \\ &= \frac{x^2 - y^2 + 2ixy - 2ixy + 2y^2}{y} \\ &= \frac{x^2+y^2}{y}.\end{aligned}$$

We have therefore shown equality. □

Corollary 6.17. *A corollary of the result above is that*

$$\frac{3}{\pi} \cdot \frac{1}{\Im\left(-\frac{1}{\tau}\right)} = \tau^2 \left(\frac{3}{\pi} \cdot \frac{1}{\Im(\tau)}\right) - \frac{6i}{\pi}\tau.$$

Proof. All we need is multiply the equality from Lemma 6.16 by $\frac{3}{\pi} \neq 0$ on both sides. □

Corollary 6.18. *The function*

$$\begin{aligned}\phi : \mathbb{H} &\to \mathbb{C} \\ \phi(\tau) &= \frac{3}{\pi} \cdot \frac{1}{\Im(\tau)}\end{aligned}$$

satisfies the condition

$$\phi\left(-\frac{1}{\tau}\right) = \tau^2 \phi(\tau) - \frac{6i}{\pi}\tau.$$

In other words, this function almost exhibits the second modularity condition of weight 2 *for* $SL_2(\mathbb{Z})$*, except for the extra term* $-\frac{6i}{\pi}\tau$*.*

Proof. This is a direct consequence of the definition of the second modularity condition with weight 2 and of the result of Corollary 6.17. □

Definition 6.19 (**Eisenstein Series of Weight 2**)**.** From the general expressions that we have previously derived for Eisenstein series and normalized Eisenstein series for even weights $k \geq 4$, we extend these expressions to the case $k = 2$ and we will then assess what properties these Eisenstein series still exhibit from the relevant standpoint to modular forms. We define

$$G_2(\tau) = \tilde{G}_2(q) = 2\zeta(2) + \frac{2(2\pi i)^2}{(2-1)!}\sum_{n\in\mathbb{N}}\sigma_1(n)q^n = \frac{\pi^2}{3} - 8\pi^2\sum_{n\in\mathbb{N}}\sigma_1(n)q^n,$$

and

$$E_2(\tau) = \tilde{E}_2(q) = \frac{G_2(\tau)}{2\zeta(2)} = \frac{\tilde{G}_2(q)}{2\zeta(2)} = 1 - 24\sum_{n\in\mathbb{N}}\sigma_1(n)q^n.$$

Theorem 6.20 (**Properties of** $G_2(\tau)$ **and** $E_2(\tau)$)**.** *We state, without fully proving, the following facts about these Eisenstein series of weight* 2 *for* $SL_2(\mathbb{Z})$*: these functions satisfy all of the conditions of being modular forms for* $SL_2(\mathbb{Z})$ *with weight* $k = 2$ *except for the second modularity condition.*

Proof. The series $\tilde{G}_2(q)$ converges for all $q \in \mathbb{D}$ because $\sigma_1(n) \leq \sum_{m=1}^{n} m = \frac{n(n-1)}{2} \leq \frac{n^2}{2}$, therefore $\sum_{n\in\mathbb{N}}\sigma_1(n)q^n$ is dominated as a non-negative series by the series with terms n^2q^n whose radius of convergence is the unit disk $\mathbb{D}$. It is also holomorphic in q as any power series is within its radius of convergence. By the inverse transformation of $q = e^{2\pi i\tau}$, i.e., by choice of a proper logarithm branch for τ as a function of q, the functions $G_2(\tau)$ and $E_2(\tau)$ are also convergent and holomorphic in τ on $\mathbb{H}$.

Also, $G_2(\tau) \to \frac{\pi^2}{3}$ and $E_2(\tau) \to 1$ as $\tau \to i\infty$, because with $\tau = x + iy$ and $y > 0$, we have

$$q^n = e^{2\pi in\tau} = e^{-2\pi ny}e^{2\pi inx},$$

and this implies

$$|\sum_{n\in\mathbb{N}}\sigma_1(n)q^n| \leq \sum_{n\in\mathbb{N}}|\sigma_1(n)|e^{-2\pi ny} \leq e^{-2\pi y}\sum_{n\in\mathbb{N}}n^2e^{-2\pi(n-1)y} \leq Me^{-2\pi y},$$

where M can be a constant that uniformly bounds the series $\sum_{n\in\mathbb{N}}n^2e^{-2\pi(n-1)y}$ for all values of y above a certain positive threshold (say, $y > 1$), which is the case when we're examining behavior as $y \to \infty$, i.e., as $\tau \to i\infty$. Therefore the series $\sum_{n\in\mathbb{N}}n^2e^{-2\pi(n-1)y}$ is dominated in absolute terms by a constant (that does not depend on y) times the term $e^{-2\pi y}$ which goes to 0 as $y \to \infty$.

In addition, $G_2(\tau + 1) = G_2(\tau)$ and correspondingly $E_2(\tau + 1) = E_2(\tau)$, because the terms of their q-expansion series $q^n = e^{2\pi in\tau}$ clearly satisfy $e^{2\pi in(\tau+1)} = e^{2\pi in\tau} \cdot e^{2\pi i} = e^{2\pi in\tau}$.

So the last question is does $G_2(\tau)$ satisfy the second modularity condition with $k = 2$ for $SL_2(\mathbb{Z})$, i.e., is it the case that $G_2\left(-\frac{1}{\tau}\right) = \tau^2G_2(\tau)$? The answer is in fact "no", but there is an alternative result, which we state without proving in the next theorem. □

Theorem 6.21 (**Not Quite Modular G_2 and E_2**). *The functions $G_2(\tau)$ and $E_2(\tau)$ satisfy the following equivalent equalities. For all $\tau \in \mathbb{H}$, we have*

$$G_2\left(-\frac{1}{\tau}\right) = \tau^2 G_2(\tau) - 2\pi i\tau,$$

and

$$E_2\left(-\frac{1}{\tau}\right) = \tau^2 E_2(\tau) - \frac{6i}{\pi}\tau.$$

Remark 6.22. We notice that the function $E_2(\tau)$ satisfies the exact same functional equality as the function $\phi(\tau) = \frac{3}{\pi} \cdot \frac{1}{\mathfrak{J}(\tau)}$ that we introduced in Corollary 6.18, so we ask what happens if we consider the function that is equal to their difference.

Definition 6.23 ($E_2^*(\tau)$). We define the function on $\mathbb{H}$

$$E_2^*(\tau) = E_2(\tau) - \phi(\tau) = E_2(\tau) - \frac{3}{\pi}\frac{1}{\mathfrak{J}(\tau)}.$$

Proposition 6.24 ($E_2^*(\tau)$ **Not Holomorphic**). *The function $E_2^*(\tau)$ is bounded as $\tau \to i\infty$, and it satisfies the two modularity conditions for being a modular form of weight 2 for $LS_2(\mathbb{Z})$. However, it is not holomorphic.*

Proof. It is easy to verify that the function satisfies the two modularity conditions. Also, it has a limit of 1 as $\tau \to i\infty$ because $E_2(\tau) \to 1$ and $\frac{3}{\pi}\frac{1}{\mathfrak{J}(\tau)} \to 0$ as $\tau \to i\infty$. However, it is not holomorphic because the term $\frac{1}{\mathfrak{J}(\tau)}$ is real, and a function that only takes values in $\mathbb{R}$ cannot be holomorphic by the *Open Mapping Theorem* which states that a holomorphic function must map an open set to an open set (and there is no open set in $\mathbb{C}$ that can be a subset of $\mathbb{R}$ as an open set in $\mathbb{C}$ must contain some open ball and $\mathbb{R}$ doesn't). □

7. Mock Modular Forms and Modular Forms on Subgroups

We have found in the previous section that we could construct a function that met all the conditions of being a modular form on $SL_2(\mathbb{Z})$ of weight 2, except that it was not a holomorphic function on $\mathbb{H}$. We introduce a definition to represent the holomorphic part of such functions, which turn out to have a very useful set of applications, in particular in the direction of quantifying the cardinality of some of the Finite Simple Groups (introduced in Definition 2.38), and of the dimensions of the vector spaces towards which they have isomorphisms, i.e., the dimensions of their representations (as introduced in Definition 2.58).

Definition 7.1 (**Mock Modular Form**). When a function on $\mathbb{H}$ satisfies the conditions of modular forms, except for the condition of being holomorphic, then if the function has a holomorphic part (e.g., a subset of its terms), then the holomorphic part of such a function is called a *mock modular form*.

Example. The function $E_2^*(\tau)$ introduced in Definition 6.23 and discussed as being non-holomorphic in Proposition 6.24 gives us a mock modular form by taking the holomorphic part which is $E_2(\tau)$ and letting go of the non-holomorphic part which is the term $\frac{3}{\pi}\frac{1}{\mathfrak{J}(\tau)}$.

We know define a few concepts that set the stage for defining modular forms on finite-index subgroups of $SL_2(\mathbb{Z})$.

Definition 7.2 (**Compactification**)**.** The compactification of a topological set (a set armed with a distance, so that limits and convergence have a meaning) is the process of adding to that set the points that it needs to become a compact set. We recall that a compact set is a set such that from any cover of that set made of open sets, there is a cover of that compact set by a finite subset of the open sets that gave the cover.

Example. As an example that is relevant to our context, we have described in Theorem 4.6 the Fundamental Domain $\mathcal{F}$ of the group action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$. This set has closed boundaries on three of its sides (an arc of the unit circle and two vertical lines), but it is not compact in the direction $i\infty$. So we compactify it by taking its union with $\{\infty\}$.

Definition 7.3 (**Cusp**)**.** A cusp is an element in the set that is added to compactify a given set. Referring to Definition 7.2, we say that the set is compactified by adding a set of cusps to it, such that the union of the original set with the set of cusps becomes a compact set.

Example. The only cusp of the fundamental domain of the group action of $SL_2(\mathbb{Z})$ (or of $PSL_2(\mathbb{Z})$) on $\mathbb{H}$ is the point at $i\infty$, i.e., $\tau = i\infty$. In the q-representation, it is the point $q = 0$, which corresponds to the $\tau = i\infty$.

Definition 7.4 (**Cusp of Group Action**)**.** When we have a group action from a group G on a space S, we define the cusp of the group action to be the cusp that compactifies the orbits of the elements of S by the action of the elements of G.

Example. The cusp of the group action of $\Gamma = SL_2(\mathbb{Z})$ on $\mathbb{H}$ is the set of points needed to compactify the orbits of all the points in $\mathbb{H}$. These orbits each have one representative in the Fundamental Domain, but they each also have an infinity of points within $\mathbb{H}$ under the action of elements of $\Gamma = SL_2(\mathbb{Z})$. We have the following theorem.

Theorem 7.5 (**Cusps of Group Actions of** $\Gamma = SL_2(\mathbb{Z})$ **on** $\mathbb{H}$)**.** *The cusps of the action by* Γ *on the points in* $\mathbb{H}$ *is made of all rational points as well as the point at infinity. It is denoted as*

$$\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}.$$

Remark 7.6. While we don't give the proof of Theorem 7.5, we note the importance of cusps to the upcoming definition of modular forms for subgroups of $SL_2(\mathbb{Z})$, especially due to the fact that their fundamental domains will have additional cusps to the single cusp of the fundamental domain of $\Gamma = SL_2(\mathbb{Z})$.

We saw in Theorem 6.21 that the Eisenstein series E_2 was not quite a modular form because it failed to satisfy the modularity condition under all transformations of $\Gamma = SL_2(\mathbb{Z})$. We also saw at the start of this section one approach to relax the conditions by defining a *mock modular form* which we extracted as the holomorphic part of E_2^* which satisfied the modularity conditions with weight 2 but was not holomorphic.

In the next definition, we will explore a different direction of relaxing the conditions of a modular form, by requiring the modularity conditions to be met for all transformations by matrices from a subgroup of $SL_2(\mathbb{Z})$ instead of the entire $\Gamma = SL_2(\mathbb{Z})$. This is motivated by the following observation.

Theorem 7.7. *The function* $F(\tau) = 2E_2(2\tau) - E_2(\tau)$ *is holomorphic, bounded at* $i\infty$*, and it also satisfies the modularity condition with weight* 2 *for the Hecke congruence subgroup*

$\Gamma_0(2)$ of $\Gamma = SL_2(\mathbb{Z})$, which -- recalling Definition 3.18 -- is the subgroup of $SL_2(\mathbb{Z})$ with an even lower left entry, i.e.,

$$\Gamma_0(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : c \equiv 0 \pmod 2 \right\}.$$

Remark 7.8. We note that the matrix $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is *not* an element of $\Gamma_0(2)$ because its lower left entry is odd. This is the intuitive explanation to why the function $F(\tau)$ is able to satisfy the modularity condition of weight 2 for this particular subgroup of $SL_2(\mathbb{Z})$.

Theorem 7.9 (**Finite Index 3 of $\Gamma_0(2)$**)**.** *The subgroup $\Gamma_0(2)$ has finite index of 3, and is generated by three matrices*

$$\Gamma_0(2) = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right\rangle.$$

We now proceed to define modular forms on subgroups. We will see that we will place a higher requirement than simply translating all conditions of a modular form for $SL_2(\mathbb{Z})$ to similar conditions for a subgroup of $SL_2(\mathbb{Z})$. In particular, we will introduce a condition of being *holomorphic at the cusps*.

Definition 7.10 (**Modular Form for Finite Index Subgroup**)**.** A modular form of weight k for a finite index subgroup $G \subseteq SL_2(\mathbb{Z})$ is a function $f : \mathbb{H} \to \mathbb{C}$ satisfying the following three conditions:

- f is holomorphic on $\mathbb{H}$.
- Modularity of weight k for the finite-index subgroup G:

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k f(\tau) \text{ for all } \tau \in \mathbb{H} \text{ and all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G.$$

- Holomorphicity at the cusps:

$$\frac{1}{(c\tau + d)^k} f\left(\frac{a\tau + b}{c\tau + d}\right) \text{ is bounded as } \tau \to i\infty \text{ for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}).$$

Remark 7.11. We note that the third condition of holomorphicity at the cusps requires the product $(c\tau + d)^{-k} f\left(\frac{a\tau+b}{c\tau+d}\right)$ to be bounded at $\tau \to i\infty$ not only for actions of matrices in the finite index subgroup G, but in fact for actions of all matrices in $SL_2(\mathbb{Z})$. It is the modularity condition with weight k which is relaxed to only apply for actions of the finite index subgroup G.

Remark 7.12. The fundamental domain of the group action of a finite index subgroup of $SL_2(\mathbb{Z})$ is a superset of the fundamental domain of the group action of the entire $SL_2(\mathbb{Z})$.

For instance, the fundamental domain of $\Gamma_0(2)$ consists of $\mathcal{F} \cup f_S(\mathcal{F}) \cup f_{ST}(\mathcal{F})$, where $f_S(\mathcal{F})$ and $f_{ST}(\mathcal{F})$ refer to the respective actions of the matrices S and ST on the original fundamental domain $\mathcal{F}$ of $SL_2(\mathbb{Z})$. In particular, this fundamental domain reaches into the unit circle (which $\mathcal{F}$ did not) and has additional cusps at 0, besides the cusp at ∞ that was already a cusp for $\mathcal{F}$.

Theorem 7.13. *The function $F_1(\tau) = 2E_2(2\tau) - E_2(\tau)$ is a modular form of weight 2 for the finite index subgroup $\Gamma_0(2)$ of $SL_2(\mathbb{Z})$.*

Remark 7.14. We don't provide the proof of this theorem but one of its remarkable consequences is a path to the proof of the four-square theorem, which we examine in the next section.

8. Four-Square Theorem

In 1770, Joseph-Louis Lagrange proved the four-square theorem which states that every natural number can be written as the sum of four integer squares. The proof given by Lagrange, while valid, was non-constructive. In 1834, Carl Gustav Jakob Jacobi gave a constructive proof with a formula for the number of ways that a natural number could be written as the sum of four squares. We will state the theorem, and proceed to sketch its proof using modular forms.

Remark 8.1. In the course of the proof sketch, we will use in anticipation a result that will only be discussed in the following section on dimensions of the vector spaces of modular forms of a given weight.

Theorem 8.2 (**Four-Square Theorem**). *For $n \in \mathbb{Z}^+$, let $r_4(n)$ designate the number of ways that n can be written as the sum of four integer squares, where order matters, i.e.,*

$$r_4(n) = |\{(a,b,c,d) \in \mathbb{Z}^4 : a^2+b^2+c^2+d^4 = n\}|.$$

Then we have

$$\begin{cases} r_4(n) = 8\sigma_1(n), & \text{if } n \text{ is odd.} \\ r_4(n) = 24 \sum_{\substack{d|n \\ d \text{ odd}}} d, & \text{if } n \text{ is even.} \end{cases}$$

Another way to write the same result is the following:

$$r_4(n) = 8 \sum_{m : 4 \nmid m | n} m.$$

Proof. The rest of this section is a sketch of the proof of the theorem, so we will first define the Jacobi Θ function, then proceed to outline the proof. □

Example. We have

$$r_4(0) = 1,$$
$$r_4(1) = 8,$$
$$r_4(2) = 24.$$

Definition 8.3 (**Jacobi Theta Function**). We define the Jacobi Θ function as

$$\Theta : \mathbb{H} \to \mathbb{C}$$
$$\Theta(\tau) = \sum_{m \in \mathbb{Z}} e^{2\pi i m^2 \tau}, \text{ i.e.}$$
$$\tilde{\Theta}(q) = \sum_{m \in \mathbb{Z}} q^{m^2}, \text{ with } q = e^{2\pi i \tau}.$$

Another way to also write $\tilde{\Theta}$ by regrouping opposite terms in $\mathbb{Z}$ is the following:

$$\tilde{\Theta}(q) = 1 + 2\sum_{m \in \mathbb{N}} q^{m^2} = 1 + 2q + 2q^4 + 2q^9 + 2q^{16} + \dots$$

We now consider the function Θ raised to the 4^{th} power, i.e., $\Theta^4 = \Theta \cdot \Theta \cdot \Theta \cdot \Theta$, and we have the following result.

Theorem 8.4.

$$\tilde{\Theta}^4(q) = \sum_{n\in\mathbb{Z}^+} r_4(n)q^n.$$

Proof. We write out $\tilde{\Theta}$ as the product of its four factors

$$\tilde{\Theta}^4(q) = \left(\sum_{m\in\mathbb{Z}} q^{m^2}\right)\left(\sum_{m\in\mathbb{Z}} q^{m^2}\right)\left(\sum_{m\in\mathbb{Z}} q^{m^2}\right)\left(\sum_{m\in\mathbb{Z}} q^{m^2}\right),$$

so the coefficient of q^n for $\tilde{\Theta}^4$ is exactly the number of ways that 4 integer squares in $\mathbb{Z}$ can add up to the exponent n, since the exponents in each of the four sum-factors are squares of integers in $\mathbb{Z}$.

We conclude that the Fourier coefficients of $\Theta^4(\tau)$, i.e., the coefficients of $\tilde{\Theta}^4(q)$ are the quantities $r_4(n)$ as we defined them in Theorem 8.2. □

Corollary 8.5. *We have*

$$\tilde{\Theta}^4(q) = 1 + 8q + 24q^2 + \dots$$

Claim 8.6 (Θ^4 **Modular Form for** $\Gamma_0(4)$)**.** *The fourth power of the Jacobi* Θ *function is a modular form of weight* 2 *for the finite index subgroup* $\Gamma_0(4)$ *of* $SL_2(\mathbb{Z})$*, i.e., the subgroup of matrices whose lower left entry is a multiple of* 4*.*

Proof. We will not give the proof other than mentioning that the proof is derived from the application of the Poisson summation formula. □

Remark 8.7. It can be said in a certain sense (of the fourth power being a modular form of weight 2) that the Θ function is "modular of weight $\frac{1}{2}$" for the subgroup $\Gamma_0(4)$.

We cited in Theorem 7.13 the result that $F_1(\tau) = 2E_2(2\tau) - E_2(\tau)$ is a modular form of weight 2 for the finite index subgroup $\Gamma_0(2)$. Since $\Gamma_0(4) \subset \Gamma_0(2)$, this implies that $F_1(\tau)$ is also a modular form of weight 2 for the finite index subgroup $\Gamma_0(4)$. This is because a function that satisfies the modularity condition for all matrices in a larger subgroup will also satisfy these conditions for the smaller subgroup which is a subset of the larger subgroup.

Another fact that we cite without proving it is that the function

$$F_2(\tau) = 2E_2(4\tau) - E_2(2\tau) = F_1(2\tau)$$

is in fact a modular form of weight 2 for the finite index subgroup $\Gamma_0(4)$ of $SL_2(\mathbb{Z})$, although it is not a modular form of weight 2 for the finite index subgroup $\Gamma_0(2)$. We note that the function $F_2(\tau)$ being the same as $F_1(2\tau)$, its q-expansion $\tilde{F_2}(q)$ is the same as the q-expansion $\tilde{F_1}(q)$ except for the series being expressed in powers of q^2 instead of powers of q.

Theorem 8.8 (**Expansions of** F_1 **and** F_2)**.** *We have the following q-expansions for the functions* $F_1(\tau) = 2E_2(2\tau) - E_2(\tau)$ *and* $F_2(\tau) = 2E_2(4\tau) - E_2(2\tau)$*:*

$$F_1(\tau) = \tilde{F_1}(q) = 1 + 24\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d\ odd}} d\right)q^n = 1 + 24q + \dots$$

$$F_2(\tau) = \tilde{F}_2(q) = 1 + 24 \sum_{n \in \mathbb{N}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^{2n} = 1 + 24q^2 + \dots$$

Proof. We recall from Definition 6.19 that

$$E_2(\tau) = \tilde{E}_2(q) = 1 - 24 \sum_{n \in \mathbb{N}} \sigma_1(n) q^n.$$

Therefore, we have

$$\begin{aligned} F_1(\tau) = \tilde{F}_1(q) &= 2 \left(1 - 24 \sum_{n \in \mathbb{N}} \sigma_1(n) q^{2n} \right) - \left(1 - 24 \sum_{n \in \mathbb{N}} \sigma_1(n) q^n \right) \\ &= 1 + 24 \sum_{n \in \mathbb{N}} \sigma_1(n) q^n - 2 \cdot 24 \sum_{n \in \mathbb{N}} \sigma_1(n) q^{2n} \\ &= 1 + 24 \sum_{\substack{n \in \mathbb{N} \\ n \text{ odd}}} \sigma_1(n) q^n + 24 \sum_{n \in \mathbb{N}} [\sigma_1(2n) - 2\sigma_1(n)] q^{2n} \end{aligned}$$

Now looking at the term $24 \sum_{\substack{n \in \mathbb{N} \\ n \text{ odd}}} \sigma_1(n) q^n$ on the right hand side of the last equality, the coefficient $\sigma_1(n)$ when n is odd is the sum of the divisors of n, and these are all odd because n is odd.

And looking at the term $24 \sum_{n \in \mathbb{N}} [\sigma_1(2n) - 2\sigma_1(n)] q^{2n}$, we note that the difference $\sigma_1(2n) - 2\sigma_1(n)$ only leaves the odd divisors of $2n$ because the even divisors of $2n$ get canceled by twice a divisor of n, term per term.

Our expression for F_1 therefore becomes

$$\begin{aligned} F_1(\tau) = \tilde{F}_1(q) &= 1 + 24 \sum_{\substack{n \in \mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 24 \sum_{n \in \mathbb{N}} \left(\sum_{\substack{d|2n \\ d \text{ odd}}} d \right) q^{2n} \\ &= 1 + 24 \sum_{\substack{n \in \mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 24 \sum_{\substack{n \in \mathbb{N} \\ n \text{ even}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n \\ &= 1 + 24 \sum_{n \in \mathbb{N}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n. \end{aligned}$$

This proves the desired result for F_1. The proof of the similar result for F_2 is identical, except for the substitution of q^2 instead of q in the series, so the result is easily shown to be as stated in this theorem. □

Theorem 8.9 (F_1, F_2 **Not Linearly Dependent**). *The funtions F_1 and F_2 are linearly independent, i.e., they are not scalar multiples of one another.*

Proof. The Fourier coefficient for the single power of q is non-zero in $\tilde{F_1}(q)$ but it is zero in $\tilde{F_2}(q)$. Just as with polynomials, infinite series cannot be linearly dependent if they non-zero terms in one are zero terms in the other. □

Claim 8.10 (**Space of Weight 2 Modular Forms for $\Gamma_0(4)$**). *The modular forms of weight 2 for the subgroup $\Gamma_0(4)$ form a vector space of dimension 2. As a result of showing the linear independence of F_1 and F_2 which were both also claimed to be modular forms of weight 2 for $\Gamma_0(4)$, the functions F_1 and F_2 form a basis of this vector space.*

Proof. It is fairly easy to see that modular forms of a given weight and for a given subgroup of $SL_2(\mathbb{Z}$ form a vector space over the scalar field $\mathbb{C}$, as the 0 function is a modular form for all weights and all subgroups of $SL_2(\mathbb{Z})$, and linear combinations with coefficients in $\mathbb{C}$ of similar weight modular forms are modular forms of the same weight, and all for the same subgroup.

The next section will be dedicated to showing the dimensions of the spaces of modular forms, so we will skip the part of the proof relative to the dimension being 2.

And we have proved in Theorem 8.9 that F_1 and F_2 are linearly independent. Since the space is of dimension 2, two elements of the space that are linearly independent form a basis of that space. □

We now proceed to the final step of the proof of the Four-Square Theorem, building on the last result that F_1 and F_2 form a basis for the two-dimensional space of modular forms of weight 2 for $\Gamma_0(4)$, and on the stated claim (not proven here) that Θ^4 is itself a modular form of weight 2 for the subgroup $\Gamma_0(4)$.

Theorem 8.11. *We have the following relationship between Θ^4 and F_1 and F_2*

$$\Theta^4(\tau) = \frac{1}{3}F_1(\tau) + \frac{2}{3}F_2(\tau).$$

Proof. We have shown previously that

$$\tilde{\Theta}^4(q) = 1 + 8q + 24q^2 + \dots$$

and that

$$\tilde{F_1}(q) = 1 + 24q + \dots$$

and that

$$\tilde{F_2}(q) = 1 + 24q^2 + \dots$$

and by virtue of F_1 and F_2 forming a basis for our two-dimensional vector space of interest, we can express the vector $\tilde{\Theta}^4$ as a linear combination of the two vectors of the basis, so there exist $\alpha, \beta \in \mathbb{C}$ such that

$$\tilde{\Theta}^4 = \alpha F_1 + \beta F_2.$$

Looking at the constant coefficients of all three q-expansions, we have

$$\alpha + \beta = 1,$$

and looking the coefficient of q, we have

$$8 = 24\alpha.$$

Our solution is therefore

$$\alpha = \frac{1}{3}, \qquad \beta = \frac{2}{3}.$$

□

This last result delivers the proof of the Four-Square Theorem, as it now suffices to regroup the terms of the sum

$$\tilde{\Theta}^4(q) = \frac{1}{3}\left(1 + 24\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^n\right) + \frac{2}{3}\left(1 + 24\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^{2n}\right),$$

and to set the coefficient of q^n equal to $r_4(n)$. We have

$$\begin{aligned}
\tilde{\Theta}^4(q) &= \frac{1}{3}\left(1 + 24\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^n\right) + \frac{2}{3}\left(1 + 24\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^{2n}\right) \\
&= 1 + 8\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^n + 16\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^{2n} \\
&= 1 + 8\sum_{\substack{n\in\mathbb{N}\\ n \text{ odd}}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^n + 8\sum_{\substack{n\in\mathbb{N}\\ n \text{ even}}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^n + 16\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^{2n} \\
&= 1 + 8\sum_{\substack{n\in\mathbb{N}\\ n \text{ odd}}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^n + 8\sum_{n'\in\mathbb{N}}\left(\sum_{\substack{d|2n'\\ d \text{ odd}}} d\right) q^{2n'} + 16\sum_{n\in\mathbb{N}}\left(\sum_{\substack{d|n\\ d \text{ odd}}} d\right) q^{2n},
\end{aligned}$$

where we have reindexed the middle term with the sum over the all the even natural numbers via $n = 2n'$ with $n' \in \mathbb{N}$.

We now note that the odd divisors of n are the same as the odd divisors of $2n$ because for any of these odd divisors d we have $\gcd(d, 2) = 1$, and we use the well-known result that if an integer divides the product of two others while being relatively prime with one of the factors of the product, then it must divide the other factor of that product.

So in the next step of simplification of the right-hand side above, we change the condition on the rightmost term $d \mid n$ into the equivalent (as argued above) condition $d \mid 2n$ due to the fact that we're conditioning this by d being odd. We also simply rename the index n'

in the middle term on the right-hand side as n, and we get the following:

$$\begin{aligned}
\tilde{\Theta}^4(q) &= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 8 \sum_{n'\in\mathbb{N}} \left(\sum_{\substack{d|2n' \\ d \text{ odd}}} d \right) q^{2n'} + 16 \sum_{n\in\mathbb{N}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^{2n} \\
&= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 8 \sum_{n'\in\mathbb{N}} \left(\sum_{\substack{d|2n' \\ d \text{ odd}}} d \right) q^{2n'} + 16 \sum_{n\in\mathbb{N}} \left(\sum_{\substack{d|2n \\ d \text{ odd}}} d \right) q^{2n} \\
&= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 8 \sum_{n\in\mathbb{N}} \left(\sum_{\substack{d|2n \\ d \text{ odd}}} d \right) q^{2n} + 16 \sum_{n\in\mathbb{N}} \left(\sum_{\substack{d|2n \\ d \text{ odd}}} d \right) q^{2n} \\
&= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 24 \sum_{n\in\mathbb{N}} \left(\sum_{\substack{d|2n \\ d \text{ odd}}} d \right) q^{2n}
\end{aligned}$$

We lastly note that when n is odd, all of its divisors are also odd, so we can remove the extra condition "d odd" that qualifies the sum of its divisors in the right-hand side of the expression above. We also reindex the rightmost term with $n' = 2n$ and summing over n' even natural number, then we reindex it back to n even natural number, so our expression becomes

$$\begin{aligned}
\tilde{\Theta}^4(q) &= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n + 24 \sum_{n\in\mathbb{N}} \left(\sum_{\substack{d|2n \\ d \text{ odd}}} d \right) q^{2n} \\
&= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{d|n} d \right) q^n + 24 \sum_{\substack{n'\in\mathbb{N} \\ n' \text{ even}}} \left(\sum_{\substack{d|n' \\ d \text{ odd}}} d \right) q^{n'} \\
&= 1 + 8 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} \left(\sum_{d|n} d \right) q^n + 24 \sum_{\substack{n\in\mathbb{N} \\ n \text{ even}}} \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n \\
&= 1 \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} 8\sigma_1(n) q^n + \sum_{\substack{n\in\mathbb{N} \\ n \text{ even}}} 24 \left(\sum_{\substack{d|n \\ d \text{ odd}}} d \right) q^n \\
&= 1 + \sum_{\substack{n\in\mathbb{N} \\ n \text{ odd}}} r_4(n) q^n + \sum_{\substack{n\in\mathbb{N} \\ n \text{ even}}} r_4(n) q^n
\end{aligned}$$

Identifying term-wise the last two lines on the right-hand side of the last expression above, we obtain

$$r_4(n) = \begin{cases} 8\sigma_1(n) & \text{if } n \text{ is odd} \\ 24 \sum_{\substack{d|n \\ d \text{ odd}}} d & \text{if } n \text{ is even} \end{cases},$$

and this completes the proof of the Four-Square Theorem!

9. Dimensions of Spaces of Modular Forms

This section gives a few relative to the structure of modular forms of even weight as vector spaces over the field $\mathbb{C}$, each of finite dimensions. In fact, the set of conditions that define modular forms (notably those relative to being bounded at infinity and being holomorphic) are key to having the vector spaces keep finite dimensions.

In this section we also develop a few results for the dimensions of each space of modular forms, via a calculation of the first few dimensions followed by a recursion on the dimensions of modular forms of higher weights from those of lower weights.

We will derive results on bases of some spaces of modular forms, by examining the Eisenstein series as well as by introducing the *discriminant* function Δ which is a discriminant of an elliptic curve associated with the solution the Weierstra equation, and which has the property of not vanishing anywhere on $\mathbb{H}$ and having no constant Fourier coefficient.

We start by stating a theorem that the set of modular forms with even weight k is a vector space.

Theorem 9.1 (M_k **Vector Space of Weight** k **Modular Forms**)**.** *The modular forms of weight k with respect to $SL_2(\mathbb{Z})$ form a vector space over the field $\mathbb{C}$. We designate this vector space by M_k.*

Proof. Most of the axioms defining a vector space are inherited from the fact that functions on $\mathbb{H}$ are themselves a vector space. The two key conditions to check are the fact that the 0 vector (i.e., the constant 0 function on $\mathbb{H}$) is in M_k, and this is the case, and that modular forms of weight k for $SL_2(\mathbb{Z})$ are closed under linear combinations with coefficients in $\mathbb{C}$, which is also the case. □

We now proceed to prove that the vector spaces M_k for $k < 0$ are all reduced to the trivial vector space $\{0\}$. To do so, we start by introducing a lemma about an invariant of modular forms to $SL_2(\mathbb{Z})$.

Lemma 9.2 ($SL_2(\mathbb{Z})$ **Invariance**)**.** *For any even integer k, and any modular form f of weight k for $SL_2(\mathbb{Z})$, the expression*

$$|f(\tau)| \cdot [\Im(\tau)]^{\frac{k}{2}}$$

is $SL_2(\mathbb{Z})$ invariant, i.e., the expression for τ is equal to the expression for the image of τ by the action induced by a matrix in $SL_2(\mathbb{Z})$.

Proof. We have shown in the proof of Theorem 4.2 the equality

$$\Im\left(\frac{az+b}{cz+d}\right) = \frac{ad-bc}{|cz+d|^2}\Im(z),$$

which, for an action induced by $SL_2(\mathbb{Z})$ implies

$$\Im\left(\frac{a\tau+b}{c\tau+d}\right) = \frac{1}{|c\tau+d|^2}\Im(\tau).$$

Raising this equality to the power $\frac{k}{2}$ (with k even) gives us

$$\left[\Im\left(\frac{a\tau+b}{c\tau+d}\right)\right]^k = \frac{1}{|c\tau+d|^k}[\Im(\tau)]^{\frac{k}{2}}.$$

By the modularity of f with weight k, we also have

$$f\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^k f(\tau) \implies \left|f\left(\frac{a\tau+b}{c\tau+d}\right)\right| = |c\tau+d|^k|f(\tau)|.$$

Multiplying our two equalities together, we get

$$\left|f\left(\frac{a\tau+b}{c\tau+d}\right)\right| \cdot \left[\Im\left(\frac{a\tau+b}{c\tau+d}\right)\right]^k = |c\tau+d|^k|f(\tau)| \cdot \frac{1}{|c\tau+d|^k}[\Im(\tau)]^{\frac{k}{2}} = |f(\tau)| \cdot [\Im(\tau)]^{\frac{k}{2}},$$

which shows the invariance of the expression $|f(\tau)| \cdot [\Im(\tau)]^{\frac{k}{2}}$ under group actions from $SL_2(\mathbb{Z})$. □

Theorem 9.3 ($M_k = \{0\}$ **for** $k < 0$)**.** *The vector space of modular forms of weight $k < 0$ for the group $SL_2(\mathbb{Z})$ is the trivial space with just the zero function as an element.*

Proof. By Lemma 9.2, the function on $\mathbb{H}$ (with real values) $|f(\tau)| \cdot [\Im(\tau)]^{\frac{k}{2}}$ is invariant by actions induced by $SL_2(\mathbb{Z})$. It therefore attains all of its values for $\tau \in \mathcal{F}$, the fundamental domain of the group actions of $SL_2(\mathbb{Z})$, and we recall that this fundamental domain is bordered by an arc of the unit circle at its bottom end, and by the two verticals lines $\{\tau \in \mathbb{H} : \Re(\tau) = \pm\frac{1}{2}\}$ at its lateral ends.

We split $\mathcal{F}$ by a horizontal line $\{z \in \mathbb{H} : \Im(z) = I\}$, for a real value $I > 1$ which will shortly determine in this proof. As $\tau \to i\infty$ with $\tau \in \mathcal{F}$, $|f(\tau)|$ is bounded as this is one of the conditions of being a modular form, and $\Im(\tau) > 1$ with $k < 0$ implies that

$$[\Im(\tau)]^{\frac{k}{2}} \to 0 \text{ as } \tau \to i\infty.$$

Therefore, the product

$$|f(\tau)| \cdot [\Im(\tau)]^{\frac{k}{2}} \to 0 \text{ as } \tau \to i\infty.$$

We can therefore now pick a value for I such that

$$\Im(\tau) > I \implies |f(\tau)| \cdot [\Im(\tau)]^{\frac{k}{2}} \leq 1.$$

And on the remaining part of $\mathcal{F}$ where $\Im(\tau) \leq I$, we have a compact set (delimited by the horizontal line at I, the two vertical lines, and the arc of the unit circle). Noting that the real-valued function $|f(\tau)[\Im(\tau)]^{\frac{k}{2}}$ is continuous, it must be bounded on a compact set.

We therefore have an upper bound $M > 0$ such that

$$|f(\tau)|[\Im(\tau)]^{\frac{k}{2}} = |f(x+iy)|y^{\frac{k}{2}} \leq M \implies |f(x+iy)| \leq My^{-\frac{k}{2}},$$

for all $x + iy \in \mathcal{F}$, and by $SL_2(\mathbb{Z})$ invariance also for all $x + iy \in \mathbb{H}$. We now write the modular form f using its q-expansion with $q = e^{2\pi i\tau} = e^{-2\pi y}e^{2\pi ix}$, which is

$$f(x+iy) = \sum_{n\in\mathbb{Z}^+} a_n q^n = \sum_{n\in\mathbb{Z}^+} a_n e^{-2\pi ny}e^{2\pi inx}.$$

Multiplying both sides by $e^{-2\pi imx}$ for some $m \in \mathbb{Z}^+$, we get

$$f(x+iy)e^{-2\pi imx} = \sum_{n\in\mathbb{Z}^+} a_n e^{-2\pi ny} e^{2\pi i(n-m)x}.$$

We now fix $y > 0$, and we integrate the two sides of the last equality above over the range x going from 0 to 1. We note that thanks to the term $e^{-2\pi ny}$ the series on the right-hand side is absolutely convergent, so we can bring the integral into the infinite series term-wise, so we have

$$\int_0^1 f(x+iy)e^{-2\pi imx}\,dx = \sum_{n\in\mathbb{Z}^+} a_n e^{-2\pi ny} \int_0^1 e^{2\pi i(n-m)x}\,dx.$$

The integral on the right is equal to 0 for $n \neq m$, and is equal to 1 for $n = m$, so we only have one term from the series contributing, which results in the equality

$$\int_0^1 f(x+iy)e^{-2\pi imx}\,dx = a_m e^{-2\pi my} \iff a_m = e^{2\pi my}\int_0^1 f(x+iy)e^{-2\pi imx}\,dx.$$

We therefore have, taking absolute values,

$$|a_m| = \left|e^{2\pi my}\int_0^1 f(x+iy)e^{-2\pi imx}\,dx\right| \leq e^{2\pi my}\int_0^1 |f(x+iy)|\,dx.$$

Recalling our calculated bound $|f(x+iy)| \leq My^{-\frac{k}{2}}$, our inequality above becomes

$$|a_m| \leq e^{2\pi my}My^{-\frac{k}{2}}\int_0^1 dx = \frac{Me^{2\pi my}}{y^{\frac{k}{2}}},$$

and this inequality holds for any arbitrary value of $y > 0$. If we let $y \to 0+$, the numerator of the last fraction goes to $M > 0$ and the denominator goes to ∞ because $k < 0$, therefore the entire fraction goes to 0. We therefore have $|a_m| = 0 \implies a_m = 0$. But this was true for any arbitrary choice of $m \in \mathbb{Z}^+$, so we have just shown that every Fourier coefficient of f is 0, so the function f must be the constant 0 function. □

We will now seek to leverage this result that $M_k = \{0\}$ for all $k < 0$, to derive dimensions of vector spaces M_k for even $k > 0$. To do this, we will leverage the existence of a modular form $\Delta \in M_{12}$ (modular form of weight 12) which vanishes nowhere on $\mathbb{H}$ and with no constant term in the q-expansion, i.e., its constant Fourier term is $a_0 = 0$.

To construct the modular form Δ, we will need a second Poisson summation result, which we state together with the first result in one theorem (which we don't prove).

Theorem 9.4 (**Two Poisson Summations**). *For a function $f : \mathbb{R} \to \mathbb{C}$ such that both f and its Fourier transform $\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{2\pi ix\xi}\,dx$ are continuous and absolutely integrable, we have the following two summation equalities:*

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} \hat{f}(n),$$

and

$$\sum_{\substack{n\in\mathbb{Z}\\ n\ odd}} (-1)^{\frac{(n-1)}{2}} f(n) = \frac{i}{2}\sum_{\substack{n\in\mathbb{Z}\\ n\ odd}} (-1)^{\frac{(n-1)}{2}} \hat{f}\left(\frac{n}{4}\right).$$

We will apply the second Poisson summation to the function $f(x) = xe^{-\pi ax^2}$, with $a > 0$, whose Fourier transform is $\hat{f}(\xi) = (-\frac{i\xi}{a^{3/2}})e^{-\pi\xi^2/a}$, noting that both functions are continuous and absolutely integrable. The result we obtain is stated in the following theorem, which we do not prove.

Theorem 9.5. *We have the following equality between series as a result of applying the second Poisson formula to the function $f(x) = xe^{-\pi zx^2}$ for $a > 0$:*

$$\sum_{\substack{n\in\mathbb{N}\\ n\ odd}} (-1)^{\frac{(n-1)}{2}} n e^{-\pi a n^2/4} = \frac{1}{a^{3/2}} \sum_{\substack{n\in\mathbb{N}\\ n\ odd}} (-1)^{\frac{(n-1)}{2}} n e^{-\pi n^2/4a}.$$

From this equality, we define a function inspired by the left hand side replacing the parameter a with $-i\tau$.

Definition 9.6. We define the function θ (not to be confused with the Jacobi Θ function discussed earlier), as

$$\theta : \mathbb{H} \to \mathbb{C}$$

$$\theta(\tau) = \sum_{\substack{n\in\mathbb{Z}\\ n\ \text{odd}}} (-1)^{\frac{(n-1)}{2}} n e^{\pi i n^2 \tau/4},$$

which can also be written as

$$\theta(x+iy) = \sum_{\substack{n\in\mathbb{Z}\\ n\ \text{odd}}} (-1)^{\frac{(n-1)}{2}} n e^{-\pi n^2 y/4} e^{\pi i n^2 x/4},$$

Theorem 9.7 (**Properties of θ**). *The function θ defined in Definition 9.6 is holomorphic on $\mathbb{H}$ and it satisfies $\theta(\tau) \to 0$ as $\tau \to i\infty$. In addition, θ satisfies the property*

$$\theta\left(-\frac{1}{iy}\right) = y^{3/2}\theta(iy).$$

Proof. We only prove the equality, starting from the definition of θ and leveraging the second Poisson summation that we obtained in Theorem 9.5. We look at the value of θ for $\tau = iy$ on the imaginary axis, and we have

$$\begin{aligned}
\theta(iy) &= \sum_{\substack{n\in\mathbb{Z}\\ n\ \text{odd}}} (-1)^{\frac{(n-1)}{2}} n e^{-\pi n^2 y/4} \\
&= \frac{1}{y^{3/2}} \sum_{\substack{n\in\mathbb{N}\\ n\ \text{odd}}} (-1)^{\frac{(n-1)}{2}} n e^{-\pi n^2/4y} \\
&= \frac{1}{y^{3/2}} \theta\left(\frac{i}{y}\right) \\
&= \frac{1}{y^{3/2}} \theta\left(-\frac{1}{iy}\right)
\end{aligned}$$

We therefore have, by swapping $\frac{1}{y}$ instead of y,

$$\theta\left(-\frac{1}{iy}\right) = y^{3/2}\theta(iy).$$

□

We are now ready to define our function Δ, which we do by raising the function θ to the 8^{th} power which turns the 3/2 exponent in the expression of Theorem 9.7 into a 12 exponent and makes it a modular form of weight 12. We will cite its properties without proving them.

Definition 9.8 (**The Δ Function**)**.** We define the function

$$\Delta(\tau) = [\theta(\tau)]^8.$$

Theorem 9.9 (**Δ Modular Form in M_{12}**)**.** *The function Δ is holomorphic, it converges to 0 as $\tau \to i\infty$, and it is modular of weight 12 for $SL_2(\mathbb{Z})$. In addition, Δ is non-vanishing over $\mathbb{H}$, its constant Fourier coefficient is 0. Lastly, its Fourier coefficient a_1 for the monomial q is non-zero, i.e., its q-expansion $\tilde{\Delta}(q)$ has a simple zero at $q = 0$, which is equivalent to Δ having a simple zero as $\tau \to i\infty$.*

With the existence of such a function Δ in hand, we can now proceed to deriving results on the dimensions of the vector spaces M_k for even $k > 0$.

Theorem 9.10 (M_0 **Vector Space of Dimension** 1)**.** *The vector space M_0 of modular forms of weight 0 has dimension 1.*

Proof. It is clear that the constant function equal to 1 everywhere on $\mathbb{H}$ satisfies all the conditions of being in M_0. Now we let f be another modular form in M_0, and let $a_0 \in \mathbb{C}$ be the constant term of it q-expansion $\tilde{f}$, i.e., $a_0 = \tilde{f}(0)$. Then the function

$$\frac{f - a_0 \cdot 1}{\Delta}$$

is a modular form in M_{-12} by Corollary 5.10 about ratios of modular forms with the denominator function vanishing nowhere on $\mathbb{H}$. And we have shown in Theorem 9.3 that for $k < 0$, $M_k = \{0\}$, which is clearly also true for $k = -12$. We therefore have

$$\frac{f - a_0 \cdot 1}{\Delta} = 0 \implies f = a_0 \cdot 1 = a_0.$$

The function f is therefore the constant function equal to a_0 everywhere on $\mathbb{H}$, which is a scalar multiple of the constant function 1 on $\mathbb{H}$. We have therefore shown that $M_0 = \mathbb{C} \cdot 1$, i.e., it is a vector space of dimension 1 with the constant function 1 being a basis for the space. □

Theorem 9.11 (M_4, M_6, M_8, M_{10} **Vector Spaces of Dimension** 1)**.** *The vector spaces M_4, M_6, M_8, M_{10} are all of dimension 1. Since we have shown the existence of Eisenstein series E_k for even $k \geq 4$, each of these vector spaces M_k for $k \in \{4, 6, 8, 10\}$ has an Eisenstein series E_k as its basis.*

Proof. Let $f \in M_k$, and let $a_0 \in \mathbb{C}$ be $a_0 = \tilde{f}(0) = \lim_{\tau \to i\infty} f(\tau)$. We then consider the function

$$\frac{f - a_0 E_k}{\Delta}$$

which we now show satisfies all the conditions of Corollary 5.10. Since Δ does not vanish n $\mathbb{H}$, it is clear that the ratio $\frac{f - a_0 E_k}{\Delta}$ is holomorphic on $\mathbb{H}$. It is also the case that this ratio satisfies the modularity conditions of weight -12.

However, we must check the important condition of the ratio being bounded as $tau \to i\infty$ or equivalently the q expansion of the ratio being bounded as $q \to 0$. The q-expansion $(\tilde{f} - a_0\tilde{E_k})(q)$ has a zero constant term by construction, since the constant term of $\tilde{E_k}$ is equal to 1 by construction of these normalized Eisenstein series.

Crucially, the q-expansion $\tilde{\Delta}(q)$ has a *simple* zero at $q = 0$, such that the ratio under consideration $\left(\frac{\tilde{f}-a_0\tilde{E_k}}{\tilde{\Delta}}\right)(q)$ not only has no constant term, but also the series in the numerator has first non-zero term q to an exponent greater than or equal to 1 whereas the series in the denominator has first non-zero term q (to the power 1) because we stated that $\tilde{\Delta}$ has a simple zero at $q = 0$. The ratio $\left(\frac{\tilde{f}-a_0\tilde{E_k}}{\tilde{\Delta}}\right)(q)$ therefore has a finite limit at $q = 0$, which equivalently means that the ratio $\frac{f-a_0E_k}{\Delta}$ has a finite limit as $\tau \to i\infty$. This shows that the boundedness condition of being a modular form is satisfied.

By Corollary 5.10, the ratio $\frac{f-a_0E_k}{\Delta}$ is therefore a modular form in the vector space M_{-12} of modular forms of weight -12 and which is the trivial space equal to the constant zero function $\{0\}$, so this ratio is the constant 0 function everywhere on $\mathbb{H}$. This implies that

$$f = a_0E_k$$

or equivalently that

$$M_k = \mathbb{C} \cdot E_k, \text{ for } k = 4, 6, 8, 10.$$

□

Theorem 9.12 (M_2 **Vector Space of Dimension** 0)**.** *The space M_2 of modular forms of weight 2 has dimension zero, i.e., it is reduced to a single element being the constant zero function $\{0\}$ on $\mathbb{H}$.*

Proof. Let $f \in M_{12}$, then the modularity condition with weight 2 implies $f(-\frac{1}{\tau}) = \tau^2 f(\tau)$. We apply the equality to $\tau = i$, and we get

$$f(-\frac{1}{i}) = i^2 f(i) = -f(i),$$

but since $-\frac{1}{i} = i$, the equality above becomes

$$f(i) = f(-\frac{1}{i}) = i^2 f(i) = -f(i) \implies f(i) = 0.$$

By Theorem 5.9, the function $f^2 = f \cdot f$ is a modular form in M_4, and we have shown in Theorem 9.11 that $M_4 = \mathbb{C} \cdot E_4$. We can therefore express f^2 as

$$f^2 = \alpha E_4, \text{ for some } \alpha \in \mathbb{C}.$$

Recalling that

$$E_4(\tau) = \tilde{E_4}(q) = 1 + 240 \sum_{n\in\mathbb{N}} \sigma_3(n) q^n,$$

we derive the equality

$$f^2(\tau) = \tilde{f}^2(q) = \alpha \left(1 + 240 \sum_{n \in \mathbb{N}} \sigma_3(n) q^n\right).$$

We now set $\tau = i$, i.e., $q = e^{2\pi i \cdot i} = e^{-2\pi}$, and we get

$$0 = f^2(i) = \tilde{f}^2(e^{-2\pi}) = \alpha \left(1 + 240 \sum_{n \in \mathbb{N}} \sigma_3(n) e^{-2\pi n}\right).$$

Since the parenthesized term on the right-hand side above is strictly positive, and the left--hand side is 0, we must have $\alpha = 0$. The function f must therefore be the constant 0 function on $\mathbb{H}$. This proves that $M_2 = \{0\}$ and its dimension is 0. □

Recapitulating the results achieved so far from the previous three theorems, we have the following for spaces of modular forms and their dimensions:

k :	0	2	4	6	8	10
$\dim(M_k)$:	1	0	1	1	1	1

We now state and prove a very powerful result on a recursion that links the dimensionalities of M_k for all even $k \geq 0$.

Theorem 9.13 (M_k **Finite Dimensional and** $\dim(M_k) = 1 + \dim(M_{k-12})$**.** *For all even $k \geq 0$, the vector space M_k has finite dimension. Furthermore, we have the recurrence relation*

$$\dim(M_k) = 1 + \dim(M_{k-12}), \text{ for all even } k \geq 0.$$

Proof. □

Theorem 9.14 (**Explicit Dimension of** M_k)**.** *For even $k \geq 0$, the dimension of the vector space M_k is given by*

$$\dim(M_k) = \begin{cases} \lfloor k/12 \rfloor + 1, & \text{if } k \not\equiv 2 \pmod{12}, \\ \lfloor k/12 \rfloor, & \text{if } k \equiv 2 \pmod{12}. \end{cases}$$

Proof. We have shown in Theorem 9.10, Theorem 9.11, and Theorem 9.12 the result to be true for $k = 0, 2, 4, 6, 8, 10$, so we now focus on the cases of even $k \geq 12$. We consider a modular form $f \in M_k$, and we let a_0 be its constant Fourier coefficient.

As with the previous cases, we consider the function $g = \frac{f - a_0 E_k}{\Delta}$ which is a modular form of weight $k - 12$ per Corollary 5.10, i.e., $g \in M_{k-12}$ and $f = a_0 E_k + g\Delta$.

We now define the function ψ as follows:

$$\psi : \mathbb{C} \oplus M_{k-12} \to M_k$$
$$\psi \left[\begin{pmatrix} \alpha \\ g \end{pmatrix} \right] = \alpha E_k + g\Delta,$$

where we recall that an element in the direct sum $\mathbb{C} \oplus M_{k-12}$ is a vector with its first coordinate being a complex scalar and its second coordinate being a complex scalar multiple of

a modular form $g \in M_{k-12}$.

We now argue that ψ is a linear mapping, i.e., it is a homomorphism of vector spaces over the field $\mathbb{C}$, and it is a bijection. As a result, we will prove that ψ is an isomorphism between its domain and its co-domain.

- ψ is a linear mapping: we verify the $\mathbb{C}$-linear mapping property of ψ by verifying its linearity for multiplication of a vector by a scalar in $\mathbb{C}$ and for summation of vectors in $\mathbb{C} \oplus E_{k-12}$. We have

$$\psi\left[\lambda\begin{pmatrix}\alpha\\ g\end{pmatrix}\right] = \psi\left[\begin{pmatrix}\lambda\alpha\\ \lambda g\end{pmatrix}\right] = (\lambda\alpha)E_k + (\lambda g)\Delta = \lambda(\alpha E_k + g\Delta) = \lambda\psi\left[\begin{pmatrix}\alpha\\ g\end{pmatrix}\right],$$

  for all $\lambda \in \mathbb{C}$ and $\begin{pmatrix}\alpha\\ g\end{pmatrix} \in \mathbb{C} \oplus M_{k-12}$. And we have

$$\psi\left[\begin{pmatrix}\alpha\\ g\end{pmatrix} + \begin{pmatrix}\beta\\ h\end{pmatrix}\right] = \alpha E_k + g\Delta + \beta E_k + h\Delta = (\alpha+\beta)E_k + (g+h)\Delta = \psi\left[\begin{pmatrix}\alpha+\beta\\ g+h\end{pmatrix}\right],$$

  for all $\begin{pmatrix}\alpha\\ g\end{pmatrix} \in \mathbb{C} \oplus M_{k-12}$ and $\begin{pmatrix}\beta\\ h\end{pmatrix} \in \mathbb{C} \oplus M_{k-12}$.
  So ψ is a linear mapping.

- ψ is surjective: for any modular form $h \in M_k$ with constant Fourier coefficient a_0, we know that we can express it as $h = a_0 E_k + \left(\frac{f - a_0 E_k}{\Delta}\right)\Delta$, where $a_0 \in \mathbb{C}$ and $\frac{f-a_0E_k}{\Delta} \in M_{k-12}$. This shows that h is the image by ψ of a vector $\begin{pmatrix} a_0 \\ \frac{f-a_0E_k}{\Delta}\end{pmatrix} \in \mathbb{C} \oplus Mk - 12$. Since h was an arbitrary modular form in M_k, this shows that ψ is surjective.

- ψ is injective: to show that a linear map is injective, we can either show as we would for any injective function that if two elements have the same image, then the two elements are the same, or we can take advantage of the fact that ψ is a linear mapping, in which case injectivity can be shown by showing that its kernel is $\{0\}$, i.e., only the 0 vector has an image of 0. We do the latter:

  If $\begin{pmatrix}\alpha\\ g\end{pmatrix} \in \mathbb{C} \oplus M_{k-12}$ is such that $\psi\left[\begin{pmatrix}\alpha\\ g\end{pmatrix}\right] = 0$, then this implies that $\alpha E_k + g\Delta = 0$ as a constant 0 function on $\mathbb{H}$. As a result, the Fourier series for $\alpha E_k + g\Delta = 0$ should have all coefficients equal to 0. We recall that the normalized Eisenstein series E_k has a constant Fourier coefficient equal to 1 and that the Δ function has a constant Fourier coefficient equal to 0 (no constant term in its q-expansion). We must therefore have

$$\alpha \cdot 1 + 0 = 0 \implies \alpha = 0.$$

  So we are now left with

$$0 \cdot E_k + g\Delta = g\Delta = 0.$$

  But Δ is the function that vanishes nowhere on $\mathbb{H}$, and its product by the function

g produces the constant 0 function on $\mathbb{H}$. Therefore the function g must be the constant 0 function on $\mathbb{H}$. We have therefore shown that

$$\psi\left[\begin{pmatrix}\alpha\\ g\end{pmatrix}\right] = 0 \implies \begin{pmatrix}\alpha\\ g\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix},$$

which is equivalent to showing that

$$\ker(\psi) = \{0\} = \begin{pmatrix}0\\ 0\end{pmatrix} \in \mathbb{C} \oplus E_{k-12}.$$

We have therefore shown that the vector spaces are isomorphic, i.e.,

$$\mathbb{C} \oplus M_{k-12} \cong M_k.$$

This implies that

$$1 + \dim(M_{k-12}) = \dim(\mathbb{C} \oplus M_{k-12}) = \dim(M_k),$$

and this proves that M_k with even $k \geq 12$ is finite dimensional, and that its dimension is

$$\dim(M_k) = 1 + \dim(M_{k-12}).$$

Lastly, we have the sequence of dimensions $\{\dim(M_k)\}$ for even $k \geq 0$ which is fully described by the recursion $\dim(M_k) = 1 + \dim(M_{k-12})$ and by the first six terms of the sequence being $1, 0, 1, 1, 1, 1$.

And we have the sequence described for even $k \geq 0$ by $\begin{cases}\lfloor k/12 \rfloor + 1, & \text{if } k \not\equiv 2 \pmod{12} \\ \lfloor k/12 \rfloor, & \text{if } k \equiv 2 \pmod{12}\end{cases}$, and this sequence starts with the same first six terms as $\{\dim(M_k)\}$, and satisfies the same recursion. The two sequences are therefore equal. This achieves the proof that for even $k \geq 0$,

$$\dim(M_k) = \begin{cases}\lfloor k/12 \rfloor + 1, & \text{if } k \not\equiv 2 \pmod{12}, \\ \lfloor k/12 \rfloor, & \text{if } k \equiv 2 \pmod{12}.\end{cases}$$

□

Remark 9.15 (**Dimensions of some** M_k). We leverage the recursion and the first values that we have already calculated for dimensions of M_k to show a few more values in the sequence.

k	0	2	4	6	8	10	12	14	16	18	20	22	24
$\dim(M_k)$	1	0	1	1	1	1	2	1	2	2	2	3	2

Remark 9.16. Another remark is that given the low dimensionality of many of these vector spaces, it becomes quite easy to relate two or more modular forms of same weight, by comparing just one or a couple of their Fourier coefficients.

For instance, when a space M_k is of dimension 1, then any modular form of weight k must be a scalar multiple of any other modular form of weight k. A simple comparison of their constant Fourier coefficients yields the scalar factor relating the first form to the second.

Similarly, if a space M_k is of dimension 2, then any two linearly independent pair of modular forms of weight k would form a basis for the space. Determination of linear independence is

often very easy and can be achieved by simple comparison of the first few Fourier coefficients. In this case too, a third modular form can be expressed as a linear combination of the first two that form a basis, and the scalar coefficients of the linear combination can be determined by equating the q-expansions forms and solving a small system of linear equations.

In fact, we can state a theorem (which we won't prove) to this effect. The important insight to derive from this theorem is that since modular forms of given weights fall within low-dimensional vector spaces, yet each such modular form has an infinite number of Fourier coefficients, there must be a large amount of redundancy among these Fourier coefficients.

Theorem 9.17 (**Form Determined by First R Fourier Coefficients**). *For each even $k \geq 0$, there is an $R \in \mathbb{Z}^+$, such that the first R Fourier coefficients of any weight k modular form for $SL_2(\mathbb{Z})$ completely determine the modular form.*

Example. A first example takes advantage of the fact that $\dim(M_8) = 1$, so all modular forms of weight 8 for $SL_2(\mathbb{Z})$ are (complex) scalar multiples of each other. We know that $E_8 \in M_8$, so E_8 is a basis for it too, and the constant coefficient of E_8 is equal to 1 (by construction of the normalized Eisenstein series).

We also know by Theorem 5.9 that since $E_4 \in M_4$, then $E_4^2 \in M_8$. Therefore, there must be some complex constant α such that $E_4^2 = \alpha E_8$. But both modular forms have, by constructions, constant coefficients equal to 1. Therefore $\alpha = 1$, and we derived the result

$$E_4^2 = E_8.$$

Example. Similarly, from $\dim(M_{10}) = 1$, we know that E_{10} is a basis for M_{10}. We also know from Theorem 5.9 that $E_4E_6 \in M_{10}$, so we know that E_4E_6 is a complex scalar multiple of E_{10}. But all three modular forms have a constant Fourier coefficient equal to 1, so the constant is 1, and we have derived another result

$$E_4E_6 = E_{10}.$$

Example. With $\dim(M_{12}) = 2$, and with the knowledge from Theorem 5.9 that $E_4^3 \in M_{12}$ and $E_6^2 \in M_{12}$, we can verify linear independence of E_4^3 and E_6^2 by a quick look at their q-expansions which easily shows that they are not scalar multiples of one another. We therefore have a basis for M_{12} composed of E_4^3 and E_6^2. Any other modular form of weight 12 for $SL_2(\mathbb{Z})$ must therefore be expressible as a complex linear combination of E_4^3 and E_6^2.

Since $E_{12} \in M_{12}$, E_{12} must be a complex linear combination of the vectors E_4^3 and E_6^2 of the basis of M_{12}. Looking at the constant and degree-1 Fourier coefficients of the three functions gives us a system of two linear equations in two unknowns, and its solution reveals another relation

$$E_{12} = \frac{441}{691}E_4^3 + \frac{250}{691}E_6^2.$$

We end this section with a result that illustrates the central role played by Eisenstein series in providing bases for all vector spaces M_k of modular forms for $SL_2(\mathbb{Z})$. In particular,

we will show that just two specific Eisenstein series, raised to the right powers and properly multiplied, deliver a basis for each M_k for even $k \geq 0$.

Lemma 9.18. *Every even integer $k \geq 4$ can be written in the form*

$$k = 4m + 6n, \text{ for some } m, n \in \mathbb{Z}^+.$$

Proof. Let $k \geq 4$ be an even integer , then either $k \equiv 0 \pmod 4$ or $k \equiv 2 \pmod 4$.

If $k \equiv 0 \pmod 4$, and with $k \geq 4$, then there exists $m \in \mathbb{N}$ such that $k = 4m = 4m + 0 \cdot 6$, so we have found a pair of non-negative integers $(m \geq 1, n = 0)$ such that $k = 4m + 6n$.

If $k \equiv 2 \pmod 4$, and with $k \geq 4$, then there exists some $q \in \mathbb{N}$ (i.e., $q \geq 1$) such that $k = 4q + 2$. If $k = 6$, then we are done with $6 = 0 \cdot 4 + 1 \cdot 6$.

Otherwise, with $k \geq 10$ and $k \equiv 2 \pmod 4$, we have the following with $q \geq 2$:

$$k = 4q + 2 = 4(q - 1)4 + 6 \cdot 1,$$

so that $(m = q - 1 \geq 1, n = 1 \geq 1)$ is a solution.

□

Remark 9.19. We note that the pair of solutions (m, n) may well not be unique -- in fact there will typically be several solutions as the values of k get larger.

Corollary 9.20 ($E_4^m E_6^n \in M_k$)**.** *For even k with $k \geq 4$, we let (m, n) be a pair of non-negative integers such that $k = 4m + 6n$. Then $E_4^m E_6^n \in M_k$ and has constant Fourier coefficient equal to 1.*

Proof. First off, it is easy to see that $E_4^m E_6^n$ has constant Fourier coefficient equal to 1 since both E_4 and E_6 have constant Fourier coefficients equal to 1 by construction of the normalized Eisenstein series.

We also know by iteratively invoking Theorem 5.9 that $E_4^m \in M_{4m}$ and $E_6^n \in M_{6n}$ and therefore the product $E_4^m E_6^n \in M_{4m+6n} = M_k$. □

Theorem 9.21 ($E_4^m E_6^n$ **Basis for** M_k)**.** *For $k \geq 0$, the set*

$$\{E_4^m E_6^n : m, n \in \mathbb{Z}^+, 4m + 6n = k\}$$

is a basis for M_k.

Proof. Let $S(k)$ be the number of solutions in non-negative integers (m, n) to the equality $4m + 6n = k$. We check the first seven values of $S(k)$, and we have:

- $k = 0$: $0 = 4 \cdot 0 + 6 \cdot 0 \implies S(0) = 1$.
- $k = 2$: $2 \neq 4m + 6n, \forall m, n \in \mathbb{Z}^+ \implies S(2) = 0$.
- $k = 4$: $4 = 4 \cdot 1 + 6 \cdot 0 \implies S(4) = 1$.
- $k = 6$: $6 = 4 \cdot 0 + 6 \cdot 1 \implies S(6) = 1$.
- $k = 8$: $8 = 4 \cdot 2 + 6 \cdot 0 \implies S(8) = 1$.
- $k = 10$: $10 = 4 \cdot 1 + 6 \cdot 1 \implies S(10) = 1$.
- $k = 12$: $12 = 4 \cdot 3 + 6 \cdot 0 = 4 \cdot 0 + 6 \cdot 2 \implies S(12) = 2$.

We see that $S(k) = \dim(M_k)$ for all $k : 0 \leq k \leq 12$.
We now sketch the proof that $S(k)$ satisfies the following induction for $k \geq 12$, and we will see that the result comes from the fact that $12 = \text{lcm}(4, 6)$:

$$S(k) = 1 + S(k - 12).$$

We note that the number of solutions $S(k)$ to $4m + 6n = k$ is the same as the number of solutions to $2m+3n = \frac{k}{2}$, since we are focusing on even values of k, where the modular form spaces are not trivially equal to $\{0\}$ as is the case for k odd. We let $h = \frac{k}{2}$ for simplification of notation.

The number of non-negative integer solutions (m, n) to $2m + 3n = h$, which we denote by $S'(h)$, is given by the coefficient of the term x^h in the generating function

$$\frac{1}{1-x^2} \cdot \frac{1}{1-x^3} = (1 + x^2 + x^4 + x^6 \cdots + x^{2i} + \dots)(1 + x^3 + x^6 + \cdots + x^{3j} + \dots)$$

We first note that there is only a single positive integer value h which cannot be realized with non-negative (m, n) and $2m+3n = h$ is 1, and this is a result of the "Chicken McNugget Theorem" that states that the largest such number is $(2)(3) - 2 - 3 = 1$. So the product of the two series above produces every power of x^h for $h \in \mathbb{N} \setminus \{1\}$.

We now observe that for $N \in \mathbb{N}$ such that $6(N - 1) \leq h < 6N$, we have

$$S'(h) = 1 + S'(h - 6) = \cdots = (N - 1) + S'[h - 6(N - 1)].$$

Indeed, we see from the expansion of the product of the two series above that the powers of x less than 6 (with the exception of x^1 which cannot be realized) can only be realized in a single way, e.g., $x^2 \cdot x^3 = x^5$ is the only way to realize x^5, and this corresponds to the equation $2 + 3 = 5$, i.e., to $(m, n) = (1, 1)$. Then we arrive at x^6 which -- because it is the LCM of 2 and 3 -- appears in both of the series, so we get 2 ways to realize it, one from the x^6 which came from $(x^3)^2$ in the second series multiplying 1 from the first series, and one from the x^6 which came from $(x^2)^3$ in the first series multiplying 1 from the second series. This corresponds to $3 \cdot 2 + 0 = 0 + 2 \cdot 3 = 6$.

This pattern continues for $6 \leq h < 12$, e.g., the term x^7 can be obtained in one way, i.e., as x^4 from the first series multiplying x^3 from the second series and this corresponds to $2 \cdot 2 + 1 \cdot 3 = 7$. And this is one more than the number of ways of realizing $x^{7-6} = x$, which was 0. On the other hand, x^9 can be realized in two ways, which is one more than for $x^{9-6} = x^3$, and this is done by the product of 1 from the series with x^9 from the second, plus the product of x^6 from the first series with x^3 from the second.

And when we arrive at x^{12}, which is double the LCM of 2 and 3, and therefore appears in both series, we have 3 ways of realizing x^{12}: 1 from the first series times x^{12} from the second, or x^{12} from the first series times 1 from the second, or x^6 from the first series times x^6 from the second. These correspond, respectively, to the three solutions to our integer linear equation $2m + 3n = 12$ being $2 \cdot 6 + 0 = 12$, $0 + 3 \cdot 4 = 12$, and $2 \cdot 3 + 3 \cdot 2 = 12$. And

this is 1 more than the 2 ways there were of realizing $x^{12-6} = x^6$.

We now argue that in order to prove our desired result for all h, it is sufficient to prove it for the multiples of $6 = \text{lcm}(2, 3)$ values of h and for these multiples plus 1, i.e., it is sufficient to prove it for $h \equiv 0 \pmod 6$ and $h \equiv 1 \pmod 6$.

This is because if we have a solution (m, n) that satisfies $2m + 3n = h$ with $h \equiv 0 \pmod 6$, then it *uniquely* induces a solution for $h + 2$ by way of $(m + 1, n)$, and a solution for $h + 3$ by way of $(m, n + 1)$, and a solution for $h + 4$ by way of $(m + 2, n)$, and a solution for $h + 5$ by way of $(m + 1, n + 1)$. Due to this $1 : 1$ relationship between solutions, the number of solutions (m, n) to $2m + 3n = h$ with $h \equiv 0 \pmod 6$ is equal to the number of solutions for $h + 2$, $h + 3$, $h + 4$, and $h + 5$.

This leaves out just $h \equiv 1 \pmod 6$, but for this case we have a $1 : 1$ association between the solutions to $2m + 3n = 6(N + 1) + 1$ and the number of solutions to $2m + 3n = 6N$. Indeed, there is a *unique* way to go from $6N$ to $6(N + 1) + 1 = 6N + 7$, and it is the following: (m, n) is a solution for $2m + 3n = 6N$ if and only if $(m + 2, n + 1)$ is a solution for $2m + 3n = 6N + 7 = 6(N + 1) + 1$, since $2(m + 2) + 3(n + 1) = 2m + 3n + 7 = h + 7$.

So we are down to only needing to prove the result that if $h = 6N$, then

$$S'(h) = 1 + S'(h - 6) = \cdots = N + S'(h - 6N) = N + 1.$$

To prove this equality, one way to proceed is by considering this as a "stars and bars" problem where the contributions of the multiple of 2 and of the multiple of 3 must each be a multiple of 6, and there are $N + 1$ multiples of 6 when considering the set $\{0, 6, \ldots, 6N\}$, so the number of choices for two non-negative numbers adding up to $6N$ while each is a non-negative multiple of 6 is equal to $\binom{n+k-1}{k-1}$, where $k = 2$ and $n = N$, i.e., it is

$$\binom{N + 2 - 1}{2 - 1} = \binom{N + 1}{1} = N + 1.$$

So we have now proven the result for all values of h by proving it for the multiples of 6, and we recall that this result is equivalent to the result that if $12(N - 1) \leq k < 12N$, then

$$S(k) = 1 + S(k - 12) = \cdots = (N - 1) + S[k - 12(N - 1)].$$

Since all we needed was the first equality $S(k) = 1 + S(k - 12)$, we have obtained our desired result.

We have therefore shown that the two sequences $S(k)$ and $\dim(M_k)$ satisfy the same recurrence relation and start with the same initial values, so they are equal. We have therefore shown that we have the right number of elements in the set to form a basis of M_k.

We must now verify that these are linearly independent vectors in M_k, or else they could generate only a lower dimension strict subspace of M_k. So we suppose

$$\sum_{\substack{4m+6n=k \\ m,n\in\mathbb{Z}^+}} \alpha_{m,n} E_4^m(\tau) E_6^n(\tau) = 0, \text{ for all } \tau \in \mathbb{H}.$$

If there is a term in the sum with the exponent n of E_6 equal to 0, then setting $\tau = i$ forces all the other terms with $n \geq 1$ to be zero because $E_6(i) = 0$. The last assertion is because E_6 satisfies the modularity condition of weight 6, and this implies, since $-\frac{1}{i} = i$, then

$$E_6(i) = E_6(-\frac{1}{i}) = i^6 E_6(i) = -E_6(i) \implies E_6(i) = 0.$$

So this leaves us with that single term $\alpha_{m,0} E_4^m(i) = 0$. But $E_4(i) \neq 0$, so this forces $\alpha_{m,0} = 0$. We can therefore assume that all terms have an exponent $n \geq 1$ on E_6.

Since the function E_6 is not identically 0, we can divide by it and we now have a new linear relation but with the sum of the exponents equal to $k - 6$ instead of k. This gives a path to a proof by strong induction in which we suppose the elements of the set are linearly independent for lower weights, and we prove the result for a higher weight.

We have therefore identified a set of $\dim(E_k)$ linearly independent vectors in E_k, so we have a basis of E_k. □

Definition 9.22 (**Eisenstein Basis**)**.** The basis $\{E_4^m E_6^n : m, n \in \mathbb{Z}^+, 4m + 6n = k\}$ is called the *Eisenstein basis* of the vector space M_k of modular forms of weight k for $SL_2(\mathbb{Z})$.

10. Hecke Operators

To give some historical perspective to this section, we start by mentioning that the Δ function that we discussed in the previous section arose during the study of elliptic curves by the likes of Jacobi and Klein in the 19^{th} century. Indeed, for any $\tau \in \mathbb{H}$, the Weierstra equation gives an elliptic polynomial, and the discriminant of this polynomial is $\Delta(\tau)$ for the τ that we started from.

In 1916, Ramanujan had the idea of exploring this Δ function under the aspect of its q-expansion and the associated Fourier coefficients. By some unfortunate collision of notation, it has been customary to refer to the Fourier coefficients of $\Delta(\tau)$, i.e., to the coefficients of the power series $\tilde{\Delta}(q)$ as $\tau(n)$, where we normally would generically call them a_n. This is despite $\tau(n)$ here being an arithmetic function (i.e., a function on natural numbers n) instead of a complex number in $\mathbb{H}$.

In order to avoid confusion from this notational collision, we choose a different notation and will refer to these Fourier coefficients of Δ as $\delta(n)$ instead of $\tau(n)$, breaking with the most prevalent convention denoting these coefficients. We must insist, however, that the literature refers to these as $\tau(n)$, and in fact this arithmetic function is known as *Ramanujan's Tau Function.*

Ramanujan observed for n up to 30 (and conjectured the generality of) three properties

of these Fourier coefficients. Shortly after Ramanujan published his conjectures, Mordell proved the first two, then Hecke re-proved these two conjectures after developing a new framework of linear operators on modular forms. It is this framework of Hecke Operators that we will introduce in this section.

We mention for good historical measure that the third Ramamujan conjecture was only settled in the 1970's by Deligne, as part of settling the Weil conjecture about the Riemann Hypothesis on Finite Fields (and which earned a Fields Medal!).

We will start by introducing Ramanujan's conjectures, then will proceed to describe the Hecke Operators and how proofs can be derived from them.

Remark 10.1 (**First Explicit Δ Coefficients**). Just for purposes of illustration, we show here the first few Fourier coefficients of the Δ function. We have

$$\Delta(\tau) = \tilde{\Delta}(q) = q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 - 6048q^6 + \dots$$

so that what was conventionally denoted as $\tau(n)$ and we will denote instead as $\delta(n)$ start as $\delta(0) = 0$, $\delta(1) = 1$, $\delta(2) = -24$, $\delta(3) = 252$, $\delta(4) = -1472$, and so on.

Conjecture 10.2 (**Ramanujan's Conjectures**)**.** *Since they have all been proved by now, these conjectures can now be more accurately designated as theorems, but we describe them as conjectures only for purposes of their historical context and the state they were in at the time the theory of Hecke Operators was developed.*

- *Claim #1: the arithmetic function $\delta(n)$ is multiplicative, i.e.,*

$$\gcd(m, n) = 1 \implies \delta(m \cdot n) = \delta(m) \cdot \delta(n).$$

  We can see an illustration of this with

$$\delta(6) = -6048 = (-24)(252) = \delta(2)\delta(3), \text{ and } \gcd(2, 3) = 1,$$

  and we note that this implies that it is sufficient to determine the values of $\delta(p)$ for p prime, in order to derive the values for all composite natural numbers.

- *Claim #2: For p prime and for $r \in \mathbb{N}$, there is a 3-term recursion in the exponent r for the values of $\delta()$ for powers of the prime p, as follows:*

$$\delta(p^{r+1}) = \delta(p) \cdot \delta(p^r) - p^{11} \cdot \delta(p^{r-1}).$$

- *Claim #3: We have the following inequalities as bounds on the absolute values of the $\delta(n)$ coefficients:*

$$|\delta(p)| \leq 2 \cdot p^{\frac{11}{2}} = 2 \cdot p^{5.5}, \text{ for } p \text{ prime},$$

and equivalently,

$$|\delta(n)| \leq d(n) \cdot n^{\frac{11}{2}} = d(n) \cdot n^{5.5},$$

where $d(n)$ is the usual arithmetic function counting the number of divisors of n.

Remark 10.3. We relay the historical anecdote that Hecke was not only able to prove the first two conjectures above, but he was also able to prove that $|\delta(p)| \leq 2 \cdot p^6$, as partial result towards proving the third conjecture.

Definition 10.4 (**Hecke Linear Operators**)**.** We first note that, as linear operators, the Hecke operators act on modular forms and map them to other modular form of same weight, i.e., they are linear maps from M_k to M_k.

We also highlight that the Hecke Operators are parameterized by prime numbers p, so we define the Hecke Operator T_p as

$$T_p : M_k \to M_k$$

$$(T_p f)(\tau) = p^{k-1} f(p\tau) + \frac{1}{p} \sum_{b=0}^{p-1} f\left(\frac{\tau + b}{p}\right), \text{ for all } \tau \in \mathbb{H}.$$

So a Hecke Operator maps a modular form $f(\tau) \in M_k$ to a linear combination of terms $f(p\tau), f\left(\frac{\tau}{p}\right), f\left(\frac{\tau+1}{p}\right), \ldots, f\left(\frac{\tau+p-1}{p}\right)$.

Theorem 10.5 (**Heck Operators Map from M_k to M_k**)**.** *The image by a Hecke Operator T_p of a modular form $f \in M_k$ of weight k for $SL_2(\mathbb{Z})$ is itself a modular form of weight k for $SL_2(\mathbb{Z})$, i.e., we have*

$$f \in M_k \implies (T_p f) \in M_k.$$

Proof. We show that the conditions of being a modular form of weight k are satisfied by $(T_p f)$ if they are satisfied by f in the first place.

- Since $(T_p f)$ is a finite sum of modular forms which are all holomorphic, then $(T_p f)$ is holomorphic.
- Also, since $(T_p f)$ is a finite sum of terms which are all bounded as $\tau \to i\infty$, then so is $(T_p f)$.
- We now examine invariance to integer translations in the direction of the real axis, and we have

  $$(T_p f)(\tau + 1) = p^{k-1} f(p\tau + p) + \frac{1}{p} \sum_{b=0}^{p-1} f\left(\frac{\tau + 1 + b}{p}\right).$$

  For the first term on the right-hand side, we have

  $$p^{k-1} f(p\tau + p) = p^{k-1} f(p\tau),$$

  because f itself is modular so it is invariant to any integer translation of τ's real part.

  As for the sum that makes the second term on the right-hand side, we have a sum

that now runs over $f\left(\frac{\tau+1}{p}\right), f\left(\frac{\tau+2}{p}\right), \ldots, f\left(\frac{\tau+p-1}{p}\right), f\left(\frac{\tau}{p}\right)$, so it is simply a cyclic permutation of the terms that were originally in the sum defining $(T_p f)(\tau)$.

We have therefore shown that $(T_p f)(\tau+1) = (T_p f)(\tau)$, which is the condition of invariance to translation of the real part by an integer (i.e., invariance to action of matrix $T \in SL_2(\mathbb{Z})$).

- We now examine the modularity condition of weight k, i.e., we determine the effect of swapping $\left(-\frac{1}{\tau}\right)$ for τ in the expression of $(T_p f)$. We have

$$\begin{aligned}
(T_p f)\left(-\frac{1}{\tau}\right) &= p^{k-1} f\left(-\frac{p}{\tau}\right) + \frac{1}{p}\sum_{b=0}^{p-1} f\left(\frac{-\frac{1}{\tau}+b}{p}\right) \\
&= p^{k-1} f\left(-\frac{1}{\left(\frac{\tau}{p}\right)}\right) + \frac{1}{p} f\left(-\frac{1}{p\tau}\right) + \frac{1}{p}\sum_{b=1}^{p-1} f\left(\frac{-\frac{1}{\tau}+b}{p}\right) \\
&= p^{k-1}\left(\frac{\tau}{p}\right)^k f\left(\frac{\tau}{p}\right) + \frac{1}{p}(p\tau)^k f(p\tau) + \frac{1}{p}\sum_{b=1}^{p-1} f\left(\frac{-\frac{1}{\tau}+b}{p}\right) \\
&= \tau^k\left[\frac{1}{p} f\left(\frac{\tau}{p}\right)\right] + \tau^k\left[p^{k-1} f(p\tau)\right] + \frac{1}{p}\sum_{b=1}^{p-1} f\left(\frac{-\frac{1}{\tau}+b}{p}\right)
\end{aligned}$$

We note that if we break out the first term corresponding to $b = 0$ of the rightmost sum in the expression of $(T_p f)(\tau)$, we can write it as

$$(T_p f)(\tau) = p^{k-1} f(p\tau) + \frac{1}{p} f\left(\frac{\tau}{p}\right) + \frac{1}{p}\sum_{b=1}^{p-1} f\left(\frac{\tau+b}{p}\right).$$

So we see that the sum of the first two terms in the last expression for $(T_p f)\left(-\frac{1}{\tau}\right)$ above is equal to τ^k times the sum of the first two terms in the expression of $(T_p f)(\tau)$ once we have rewritten by breaking out the term corresponding to $b = 0$ from the rightmost sum.

We now compare the remaining terms from $b = 1, \ldots, p-1$ in the rightmost sums, and we will exploit the following property: since p is a prime, all integers $1, \ldots, p-1$ are units in the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^\times$, in the sense that they all have inverses because they are all relatively prime to p. In particular, for every $b \in \{1, \ldots, p-1\}$, there is another $b' \in \{1, \ldots, p-1\}$, such that $bb' \equiv -1 \pmod{p}$.

This ability to group all remaining terms in the sum over $b = 1, \ldots, p-1$ into pairs b, b' that are multiplicative inverses $\pmod{p}$ of one another, is the key to the manipulation of the expressions towards showing that the remaining sum in the expression of $(T_p f)\left(-\frac{1}{\tau}\right)$ is equal to τ^k times the corresponding sum in the expression of $(T_p f)(\tau)$.

In other words, the change $\tau \to -\frac{1}{\tau}$ results in a permutation of the terms for a

given b and for its inverse $b' = -\frac{1}{b} \pmod{p}$ in the sum $\frac{1}{p}\sum_{b=1}^{p-1} f\left(\frac{\tau+b}{p}\right)$, and this leads to the sums of all terms becoming a factor of τ^k of one another.

This achieves the goal of proving the modularity condition with weight k for $(T_p f)$, and by consequence the proof that $f \in M_k \implies (T_p f) \in M_k$.

□

Theorem 10.6 (**Properties of Hecke Operators**). *We highlight the following important properties of the Hecke Operators.*

- *They commute, i.e.,* $T_m T_n = T_n T_m$.
- *They are multiplicative, i.e.,* $T_{mn} = T_m T_n$ *when* $\gcd(m, n) = 1$.
- *They satisfy the recurrence*

$$T_{p^{r+1}} = T_p T_{p^r} - p^{k-1} T_{p^{r-1}} \text{ for prime } p, \text{ and } r \in \mathbb{N}$$

Theorem 10.7 (q-**Expansion of** $T_p f$). *If* $f(\tau) = \tilde{f}(q) = \sum_{n\in\mathbb{Z}^+} a_n q^n$, *then we have*

$$(T_p f) = \sum_{n\in\mathbb{Z}^+} a_{pn} q^n + \sum_{n\in\mathbb{Z}^+} p^{k-1} a_n q^{pn}.$$

We now look at how to leverage the properties of the Hecke Operators, notably those in Theorem 10.6 and Theorem 10.7 to sketch proofs of the first and second Ramanujan conjectures.

Theorem 10.8 (**The First Ramanujan Conjecture is a Theorem**). *The Ramanujan tau function is multiplicative. We will use the* $\delta(n)$ *instead of* $\tau(n)$ *for the same reasons as above, but we are talking about the Ramanujan tau nevertheless.*

Proof. We start from the observation that M_{12} which is two-dimensional is realized as the direct sum of the scalar multiples of the Eisenstein series E_{12} and the scalar multiples of the Δ function (which, we recall, has a 0 constant Fourier coefficient and a 1 for the Fourier coefficient of degree 1), i.e., we have

$$M_{12} = \mathbb{C} \cdot E_{12} \oplus \mathbb{C} \cdot \Delta.$$

From Theorem 10.7, and focusing on constant Fourier terms, we see that if a_0 is the constant Fourier term for a modular form f, then the constant Fourier term for $(T_p f)$ is

$$a_{p\cdot 0} + p^{k-1} a_0 = a_0 + p^{k-1} a_0 = (1 + p^{k-1}) a_0$$

Recalling that Δ has a 0 constant term, we know that the subspace $\mathbb{C} \cdot \Delta$ of scalar multiples of Δ all modular forms have constant term $a_0 = 0$. This implies that the constant Fourier term for their images by Hecke operators are all $(1 + p^{k-1}) \cdot 0 = 0$. This means that the T_p Hecke operator sends the $\mathbb{C} \cdot \Delta$ subspace of M_{12} to itself. But this is a one-dimensional subspace with Δ as an eigenvector. We therefore have the result

$$T_p \Delta = \lambda_p \Delta, \text{ for some } \lambda_p \in \mathbb{C}.$$

We now turn to the coefficient of q (first degree monomial) in $(T_p f)$ as given to us by Theorem 10.7, and we have a term a_p coming from $\sum_{n\in\mathbb{Z}^+} a_{pn}q^n$ with $n = 1$, and there is no term with q to the power 1 in the sum $\sum_{n\in\mathbb{Z}^+} p^{k-1}a_n q^{pn}$ because all terms have a q^p factor or other exponents of q that are all multiples of p. We therefore have, noting that the Fourier coefficient of q in the q-expansion of Δ is equal to 1

$$\delta(p) = \lambda_p \cdot 1 = \lambda_p.$$

Therefore, $\delta(p)$ is the eigenvalue associated with the subspace $\mathbb{C}\cdot\Delta \subset M_{12}$ for the linear operator T_p. So we have shown that

$$T_p\Delta = \delta(p)\Delta,$$

where $\delta(p)$ is the way we are designating the Ramanujan *tau* function evaluated at the prime integer p. But we know from Theorem 10.6 that T_p is multiplicative, i.e., if $\gcd(m,n) = 1$, then $T_{mn} = T_m T_n$. This means that if $\gcd(m,n) = 1$, then $\delta(mn) = \delta(m)\delta(n)$, i.e., the Ramanujan *tau* function is also multiplicative. This proves the first Ramanujan conjecture. □

Theorem 10.9 (**The Second Ramanujan Conjecture s a Theorem**). *For prime p and exponent $r \in \mathbb{N}$, we have the following recursion for the Ramanujan tau function*

$$\delta(p^{r+1}) = \delta(p)\cdot\delta(p^r) - p^{11}\cdot\delta(p^{r-1}).$$

Proof. We first prove the result for $r = 1$, and we want to show that

$$\delta(p^2) = \delta(p)\cdot\delta(p) - p^{11}\cdot\delta(1) = [\delta(p)]^2 - p^{11},$$

because $\delta(1)$ is the Fourier coefficient for degree 1 of the q-expansion for Δ, which we know to be equal to 1. We now examine the coefficient of q^p from the q-expansion given in Theorem 10.7, and we have:

- One component comes from $a_{pn}q^n$ for $n = p$, therefore a_{p^2}, i.e., $\delta(p^2)$ in this case of the function being Δ.

- One component comes from $n = 1$ in $p^{k-1}a_n q^{pn}$, so in this case of M_{12} we have $k = 12 \implies k-1 = 11$ and with $n = 1$, this becomes $p^{11}\delta(1)$.

Since we have shown in Theorem 10.8 that $T_p\Delta = \delta(p)\Delta$, we have that the coefficient of q^p in $(T_p\Delta)$ which is $\delta(p)\Delta$ is therefore $\delta(p)$ multiplied by the coefficient of q^p in the q-expansion of Δ, but that is $\delta(p)$ too, by definition. So the product is equal to $[\delta(p)]^2$, and it is equal to the sum of the two components that we have identified above, so we have

$$[\delta(p)]^2 = \delta(p^2) + p^{11}\cdot\delta(1),$$

or, equivalently,

$$\delta(p^2) = [\delta(p)]^2 - p^{11}\cdot\delta(1).$$

So we have proven the stated claim for $r = 1$.

The proof for the recursion proceeds similarly by equating the coefficients of the terms $q^{p^2}, q^{p^3}, \ldots$, and this delivers the proof of the second Ramanujan conjecture.

□

Remark 10.10. On a final historical note, it is interesting to realize that the reason Ramanujan was studying these series is that he was interested in analogs of the Riemann zeta function by studying the Dirichlet series associated with his *tau* coefficients, i.e., he was interested in

$$L(s, \Delta) = \sum_{n \in \mathbb{N}} \frac{\delta(n)}{n^s}.$$

It turns out that the recursions which Ramanujan had identified imply an Euler product similar to the case for the Riemann zeta function, and in the case of Δ the product is the following:

$$L(s, \Delta) = \sum_{n \in \mathbb{N}} \frac{\delta(n)}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - \frac{\delta(p)}{p^s} + \frac{p^{11}}{p^{2s}}}$$

This series converges for $\Re(s) > 6.5$ but it can be analytically continued to all of $\mathbb{C}$. The functional equation resulting from the last equality above related s to $k - s$ which is $12 - s$ in this case of M_{12}, and we note that the midpoint of s and $12 - s$ is 6.

Expecting the Riemann Hypothesis to be true, all the non-trivial zeros would be on the line $\Re(s) = 6$. However, for modular forms that are *not* eigenvectors of the Hecke Operator, there is no Euler product for their Dirichlet L-series, and consequently they can have zeros in the half-plane where their L-series converges, rather than only on a line of constant real value.

11. Summary

The goal of this paper was to introduce modular forms to an audience that is already familiar with Complex Analysis, with a view to illustrating the impact of the field of modular forms on number theory. In particular, we have found it interesting to demonstrate the low dimensionality of many relevant vector spaces of modular forms. Following this path has allowed us to relate some key modular forms either by scalar proportionality factors or by simple linear combinations, from which number theory results could be derived, such as the Four Squares Theorem.

We have also introduced the concept of Hecke Operators which act linearly on vector spaces of modular forms, and which have been very instrumental in proving results such as a few famous Ramanujan Conjectures. We also mentioned, although in passing, the notion of mock modular forms which open very fruitful paths towards characterizing orders of the Sporadic Finite Simple Groups (such as Mathieu Groups and The Monster) and of dimensionality of their (vector space) representations. In an attempt to control the length of this expository

paper, we had to refrain from detailing this path via mock modular forms, despite its very stimulating rewards.

Lastly, we have attempted to make the paper self-contained when it comes to its reliance on domains external to Complex Analysis, including Group Theory, Linear Algebra, Matrix Groups, and Group Actions. This has lengthened the paper by a significant extent, but we hope to have made the paper self-contained as a result.

References

[BOYA] Luis J. Boya (Departamento de Física Teórica, Universidad de Zaragoza, Spain)
Introduction to Sporadic Groups
https://arxiv.org/pdf/1101.3055.pdf

[CHAP] Richard Chapling (Trinity College, Dublin, Ireland)
Elliptic Functions
https://rc476.user.srcf.net/furthercomplexmethods/ellfun.pdf

[CHEN] Miranda Cheng (Institute of Physics, University of Amsterdam, Amsterdam, The Netherlands)
Mock Modular Forms are Everywhere
https://www.youtube.com/watch?v=dOY_MzmS0Zk

[COHE] Henri Cohen (Institut de Mathématiques de Bordeaux, Bordeaux, France)
An Introduction to Modular Forms
https://hal.inria.fr/hal-01883058/document

[CONR] Keith Conrad, Amanda Folsom, Alvaro Lozano-Robledo, Liang Xiao (University of Connecticut), (University of Connecticut Summer School in Number Theory - August 2016)
CTNT 2016 - Introduction to Modular Forms
https://ctnt-summer.math.uconn.edu/wp-content/uploads/sites/1632/2016/02/CTNTmodularforms.pdf
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=1
https://www.youtube.com/watch?v=mPycIhlPbZo&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=2
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=3
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=4
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=5
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=6
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=7
https://www.youtube.com/watch?v=LolxzYwN1TQ&list=PLJUSzeW191Qx_rdAS8sd4nTNlSyLt97Q4&index=8

[CRAN] Sander Mack-Crane (University of California, Berkeley)
Congruence Subgroups
https://math.berkeley.edu/~sander/speaking/24June2015%20UMS%20Talk.pdf

[ELKI] Noam Elkies (Harvard University)
Rational Lattices and their Theta Functions
http://people.math.harvard.edu/~elkies/M272.19/

[FREY] Theo Johnson-Freyd (Canada/USA Math Camp 2019)
Sporadic Groups and Where to Find Them
http://categorified.net/Mathcamp-SporadicGroupsClass.pdf

[FOLS] Amanda Folsom (Amherst University)
Perspectives on Mock Modular Forms (Journal of Number Theory 176 (2017) 500-540
https://afolsom.people.amherst.edu/JNT_Perspectives_Folsom_publ.pdf
[HIRS] Daniel Hirsbrunner (U. of Washington)
Elliptic Functions with a View toward Elliptic Curves
https://sites.math.washington.edu/~reu/papers/2013/daniel/Elliptic%20Functions%20Notes.pdf
[KURI] Robert Kurinczuk (Imperial College, London, UK)
Modular Forms
http://wwwf.imperial.ac.uk/~dhelm/M4P58/ModularForms2.pdf
[LANDE] Aaron Landesman (Stanford University)
Modular Forms and the Four Squares Theorem
https://web.stanford.edu/~aaronlan/assets/landesman_junior_paper.pdf
[MILN] J.S. Milne (University of Michigan)
Modular Forms
https://www.jmilne.org/math/CourseNotes/MF110.pdf
[NEWT] James Newton (King's College, London, UK)
Modular Forms
https://nms.kcl.ac.uk/james.newton/lec1.pdf
[ONOK] Ken Ono (Emory University)
Modular Forms and Representation Theory
https://www.youtube.com/watch?v=NNa_njwB5l8
[RIBE] Ken Ribet (U.C. Berkeley) and William Stein (University of Washington)
Lectures on Modular Forms and Hecke Operators
https://wstein.org/books/ribet-stein/main.pdf
[SARN] Peter Sarnak (Princeton University)
The Unreasonable Effectiveness of Modular Forms
https://www.youtube.com/watch?v=jFtu4asyolk
[SCHU] Dan Schultz (University of Illinois)
Notes on Modular Forms https://faculty.math.illinois.edu/~schult25/ModFormNotes.pdf
[STEI] Elias M. Stein, Rami Shakarchi
Complex Analysis, Princeton Lectures in Analysis II, Princeton University Press (2003), ISBN-13: 978-0-691-11385-2, Ch.9, Ch.10.
[TELE] Constantin Teleman (U.C. Berkeley)
Representation Theory
https://math.berkeley.edu/~teleman/math/RepThry.pdf
[VENK] T.N. Venkataramana (Tata Institute of Fundamental Research, Mumbai, India)
Classical Modular Forms
http://users.ictp.it/~pub_off/lectures/Ins021/Venkataramana/Venkataramana.pdf
[ZAGI] Don Zagier (Séminaire Bourbaki 2008, France)
Ramanujan's Mock Theta Functions and Their Applications
http://numdam.org/article/AST_2009__326__143_0.pdf