First Swig of Moonshine

An Introduction to Modular Forms, Representation Theory, and the Monster

Justin Kelm
Department of Mathematics and Statistics, Oakland University

Contents

1 Overview

The first part of the paper will begin with the history of monstrous moonshine and a general overview of its subject matter. This is Section 2.

The second part of the paper—Sections 3 through 8—will be devoted to the study of modular forms; a fluency in undergraduate complex analysis will be assumed. We will start from the bottom by introducing Eisenstein series (the most basic example of modular forms). We will proceed to characterize all modular forms by means of these series, and in particular define the j-invariant. Then, in Section 6, we directly calculate the Fourier series coefficients of the Eisenstein series. Carrying these calculations through to the j-invariant will give us the raw data behind one half of the "mysterious" connection between complex analysis and representation theory.

The third part of the paper—Sections 9 through 11—will introduce the core concepts of linear representation and character theory. We build up the theory of characters in the context of finite groups, and discuss character tables. Although we shall not have the leisure of working out the character table of the Monster group ourselves, we shall hopefully acquire a taste for the mathematics involved.

2 History and Background

The first rigorous formulation and study of the mathematical object known as a *group* began in the mid-19th century with the work of Galois and Cauchy. They arose in the context of what would come to be known as Galois theory, in the form of permutation groups which acted on the roots of rational polynomial functions. Despite the field's infancy, even Galois at his time realized the important notion that complex groups could be collapsed into simpler groups by taking the quotient of it by a normal subgroup. Those groups that could be collapsed no further (nontrivially speaking) were known as simple groups. Galois provided the first example of a family of finite noncyclic simple groups in order to conclude his proof of the unsolvability of the quintic, namely, the alternating groups A_n for $n \geq 5$. He would also introduce and prove the simplicity of the projective special linear groups $\mathrm{PSL}(2,p)$.

In the best of cases, when a group G contains a nontrivial proper normal subgroup N, we may precisely describe the more complex group structure G in terms of the relatively less complex group structures N and G/N (such is the case if G contains a subgroup complement to N—then G decomposes as a semidirect product involving N). Although even today we do not possess every possible tool in constructing larger groups out of smaller ones, this motivated mathematicians to completely classify finite simple groups. This task—which spanned tens of thousands of pages from hundreds of authors over 150 years—was finally considered complete in 2004 ([1]; although a small error was discovered and fixed in 2008). The conclusion:

Theorem. *Every finite simple group is isomorphic to one of the following:*

- *A cyclic group of prime order*
- *An alternating group of order* ≥ 5
- *A so-called group of Lie type*
- *The Tits group (sometimes considered a group of Lie type)*
- *One of the 26 "sporadic groups"*

The sporadic groups were always considered rather mysterious and exceptional objects. The largest one of these (and the final simple group proven to exist, first by Griess in 1982 [9] and later more simply by Conway in 1985 [2]) is known as the Monster group, and contains

$$2^46 \cdot 3^20 \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 =$$
$$808,017,424,794,512,875,886,459,904,961,710,757,005,754,368,000,000,000 \approx$$
$$8 \times 10^{53}$$

elements. Although the infinite families of simple groups such as the alternating groups or the groups of Lie type obviously contain examples of simple groups with more elements, the Monster group is special in that there is no known "easy" or efficient way to represent its elements (in terms of, for example, relatively small permutation or linear representations) and is the only finite simple group that is still highly resistant to calculations by computer.

The entirety of the character table (an important set of invariants associated to a finite group, explored in Section 10) of the Monster group was calculated in 1979 (one would be surprised how much mathematicians manage to prove about an object before they can even prove it exists!), whereupon it was found the smallest nontrivial complex representation of the Monster was in dimension 196,883.

This number would've lived out an otherwise innocuous life in but a handful of dusty academic papers were it not for John McKay's closer look at the first few terms of something called the (normalized) j-invariant in 1978:

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + ...$$
$$1 = 1$$
$$196884 = 1 + 196883$$
$$21493760 = 1 + 196883 + 21296876$$

where $1, 196883, 21296876...$ are the dimensions of the irreducible representations of the Monster group in increasing order. This was most startling because the j-invariant was native to modular curve theory, a field of mathematics that was completely unrelated to the theory of finite simple groups, and there was no explanation for this infinite set of seeming "coincidences." As such, the study behind this serendipitous phenomenon became known as "monstrous moonshine" [3].

An even stronger formulation of the link between the monster group and modular curve theory was finally standardized, and proven in a landmark paper by Borcherds in 1992, for which he received the Fields medal in 1998.

3 Modular Forms

Allow us to begin by considering a special symmetry, known as the *modular action on the upper-half plane.* If $H = \{z \in \mathbb{C} \mid \text{Im } z > 0\}$ denotes the upper-half plane, and $\text{SL}_2(\mathbb{Z}) = \{M \in \text{GL}_2(\mathbb{Z}) \mid \det M = 1\}$ denotes the *(integral) special linear group*, then $\text{SL}_2(\mathbb{Z})$ affords a left action on H via linear fractional transformations. That is, for $\tau \in H$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \tau = \frac{a\tau + b}{c\tau + d}$$

Left associativity of this action can be seen by a routine computation; that $M\tau$ is also in H follows from said associativity, and the fact that $\mathrm{SL}_2(\mathbb{Z})$ is generated as a group under composition by the elements $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (a fact itself following from the existence of reduced row echelon form of matrices in $\mathrm{SL}_2(\mathbb{Z})$), which correspond respectively to the operations $\tau \mapsto \tau + 1$ and $\tau \mapsto -\frac{1}{\tau}$, both of which preserve H. We see that the kernel of this action is precisely $\pm I$, so this action factors through to the quotient $\mathrm{SL}_2(\mathbb{Z})/\{\pm I\}$, hereafter referred to as the *modular group* and notated as Γ. Throughout this paper, γ is to refer to an element of Γ represented by $\pm \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Definition 3.1. For given even $k \in \mathbb{N}$, we define the *(k^{th}) factor of automorphy* to be a map j from Γ to the set of invertible holomorphic functions on H, given by $j_\gamma(\tau) = (c\tau + d)^k$. A holomorphic function f is said to be a *modular form (of weight k)* if it satisfies the *weak modularity condition*: $f(\gamma\tau) = j_\gamma(\tau)f(\tau)$ for all $\gamma \in \Gamma$ and all $\tau \in H$, and is bounded as $\tau \to i\infty$. (In the more general context, which we shall not explore in this paper, Γ may be any group acting on a complex-analytic manifold H, and the *automorphic functions* are those holomorphic functions on H satisfying a similar relationship together with a pre-defined factor of automorphy.)

Remark. The condition of boundedness as $\tau \to i\infty$ is also referred to as "holomorphy at infinity"; the justification for this will become apparent in Section 4.

It is convenient to define linear operators $[\gamma]_k$ (for all real matrices γ of positive determinant) on the space of holomorphic functions on H given by $f[\gamma]_k(\tau) = j_\gamma(\tau)^{-1}f(\gamma\tau)$, so that the modular forms are precisely the common fixed points of each operator $[\gamma]_k$. Since the correspondence $\gamma \mapsto [\gamma]_k$ is associative as a right action on the holomorphic functions on H (another routine computation), again by considering generators of Γ we deduce that a holomorphic function is a modular form if and only if $f(\tau + 1) = f(\tau)$ (i.e. is $\mathbb{Z}$-periodic) and $f(-1/\tau) = \tau^k f(\tau)$ for all $\tau \in H$.

Modular forms may perhaps be understood most simply as complex-valued functions on (the moduli space of) lattices satisfying a certain homogeneity property. By a lattice, we mean an abelian group $L \subseteq \mathbb{C}$ of rank two that generates $\mathbb{C}$ as a vector space over $\mathbb{R}$. Figure 1 shows a prototypical lattice generated by $\omega_1 = 1$ and $\omega_2 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$, denoted $[\omega_1, \omega_2]$.

We may assume that any lattice given in these terms is such that $\frac{\omega_2}{\omega_1} \in H$.

Let $\mathbb{L}$ denote the set of all lattices. There is an action of $\mathbb{C}^\times$ on $\mathbb{L}$ via $\lambda L = \{\lambda\omega \mid \omega \in L\}$. We prove the following correspondence:

Proposition 1. *Let an even $k \in \mathbb{N}$ be fixed. Consider the collection of functions $F : \mathbb{L} \to \mathbb{C}$ such that $F(\lambda L) = \lambda^{-k}F(L)$ for all $L \in \mathbb{L}$ and all $\lambda \in \mathbb{C}^\times$. Consider also the collection of functions $f : H \to \mathbb{C}$ such that $f[\gamma]_k = f$ for all $\gamma \in \Gamma$. There is a correspondence between the two $F \leftrightarrow f$ which is uniquely constrained by*

$$F([1, \omega]) = f(\omega)$$

Proof. Suppose first that we have such an F; we define $f(\omega)$ to be the value of F on the lattice

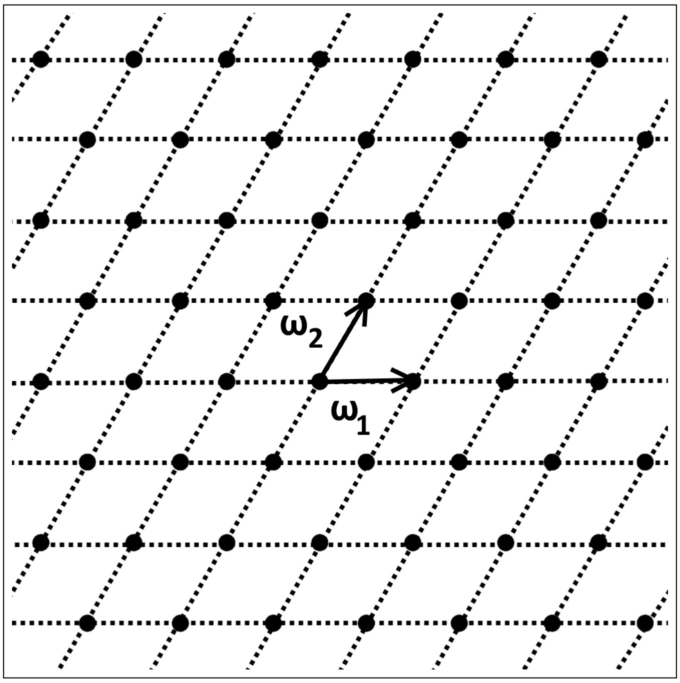

Figure 1: The lattice $[\omega_1, \omega_2]$

$[1, \omega]$. Then

$$f[\gamma]_k(\omega) = (c\omega + d)^{-k} \cdot F([1, \frac{a\omega + b}{c\omega + d}]) = F([c\omega + d, a\omega + b]) = F([1, \omega]) = f(\omega)$$

so that $f[\gamma]_k = f$ as claimed. Now, suppose we have a function f satisfying the weak modularity condition of weight k; in this case we define $F([\omega_1, \omega_2]) = \omega_1^{-k} f(\frac{\omega_2}{\omega_1})$. Our first order, of course, is to prove that this is well defined independent of ω_1, ω_2: Suppose $[\omega_1, \omega_2] = [\omega_1', \omega_2']$. Then

$$\begin{bmatrix} \omega_1' \\ \omega_2' \end{bmatrix} = \gamma \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$$

for some $\gamma \in \Gamma$ (this will follow, for example, from Proposition 4 together with the fact that if $\det \gamma = -1$ then $\frac{\omega_2'}{\omega_1'} \notin H$). As such, if $\gamma' \in \Gamma$ is the matrix γ with first its columns and then its rows interchanged (i.e. conjugated by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$), we have

$$F([\omega_1', \omega_2']) = F([a\omega_1 + b\omega_2, c\omega_1 + d\omega_2]) = (a\omega_1 + b\omega_2)^{-k} f(\frac{c\omega_1 + d\omega_2}{a\omega_1 + b\omega_2}) =$$

$$\omega_1^{-k}(b\frac{\omega_2}{\omega_1} + a)^{-k} f\Big(\frac{d\frac{\omega_2}{\omega_1} + c}{b\frac{\omega_2}{\omega_1} + a}\Big) = \omega_1^{-k} f[\gamma']_k(\frac{\omega_2}{\omega_1}) = \omega_1^{-k} f(\frac{\omega_2}{\omega_1}) = F([\omega_1, \omega_2])$$

Last but not least, the desired property $F(\lambda L) = \lambda^{-k} F(L)$ follows immediately from the definition. □

Thus, from this point forward, we shall frequently cross-interpret the two concepts without undue formalism.

Now that we understand the basics of the modular action and what modular forms are, what are some examples? If we wish to construct an example of a modular form of weight k, it seems easiest to approach it from the point of view of assigning complex values to lattices in such a way that the homogeneity property is satisfied. Therefore, examining the function that takes a given lattice and sums up the $-k^{\text{th}}$ powers of its nonzero elements is a perfectly natural idea. This leads us to the *Eisenstein series of weight k.*

Proposition 2. *Let $k > 2$ be an even integer, and consider the following function $G_k : H \to \mathbb{C}$:*

$$G_k(\tau) = \sum_{(c,d)\in\mathbb{Z}^2-(0,0)} \frac{1}{(c\tau+d)^k}$$

The series above converges absolutely for all τ in the upper half plane, and as such the summands may be permuted freely. The series also converges uniformly on compact subsets, and thus defines a holomorphic function on the upper-half plane. $G_k(\tau)$ is bounded as $\tau \to i\infty$, and G_k is also weakly modular, and so defines a modular form of weight k. The G_k are not constantly zero, and in fact $G_k(\tau) \to 2\zeta(k)$ as $\tau \to i\infty$.

Proof. Proving that the series absolutely converges for any given τ amounts to showing the same for the Weierstrass series with respect to the lattice $L = [1, \tau]$; a proof may be found in any introductory book on elliptic curves (e.g. [10], Chapter I, Section 4), so we merely outline the idea. One may contain the lattice points of L in disjoint balls of uniform size $\delta > 0$. Since each annulus of radii $R-1, R+2$ has area $O(R)$, the convergence of $\sum \frac{1}{|c\tau+d|^k}$ follows by comparison with the convergent series $\sum_{R=1}^{\infty} \frac{1}{R^{k-1}}$.

For positive real numbers A, B, let $\Omega = \{\tau \mid |\text{Re}(\tau)| \leq A, |\text{Im}(\tau)| \geq B\}$. We show $G_k(\tau)$ converges uniformly on Ω, and hence on compact subsets in general, thereby proving $G_k(\tau)$ defines a holomorphic function. We do this by showing that $\sum_{(c,d)\in\mathbb{Z}^2-(0,0)} \sigma_{c,d}$ converges on Ω, where

$$\sigma_{c,d} = \sup_{\tau\in\Omega} \frac{1}{|c\tau+d|^k}$$

When $A = 1$, this will be enough to show $G_k(\tau)$ is bounded as $\tau \to i\infty$, once we demonstrate weak modularity (since this implies $\mathbb{Z}$-periodicity). Note that this supremum is approached by bringing $c\tau$ as close to $-d$ as possible. If c, d are not both 0, it is clear that $|c\tau + d| > |\frac{cB}{2}|$, and also that $|c\tau + d| > |\frac{Bd}{A}|$, for all $\tau \in \Omega$. That is to say, there exists real $C > 0$ such that $|c\tau + d| \geq C \cdot \sup\{|c|, |d|\} \geq \frac{C}{\sqrt{2}} \cdot |c + di|$. This implies

$$\sum \sigma_{c,d} \leq \left(\frac{\sqrt{2}}{C}\right)^k \sum \frac{1}{|c+di|^k}$$

which is finite, so that $G_k(\tau)$ is indeed holomorphic on H. Finally, weak modularity of $G_k(\tau)$ follows from Proposition 1 and the fact that $G_k(\tau)$ represents the analytic form of a complex-valued function on lattices we know to satisfy the homogeneity property.

Finally, that $G_k(\tau) \to 2\zeta(k)$ as $\tau \to i\infty$ may be observed by considering the limit of the sequence $G_k(2^n i)$ as $n \to \infty$. Each term of this sequence is by definition the sum of the k^{th} powers of the reciprocals of some subset Y_n of the Gaussian integers, which we may partition into $\mathbb{Z} \setminus \{0\}$ and its complement in Y which we denote X_n. As $X_{n+1} \subseteq X_n$ and $X_n \to \varnothing$ as $n \to \infty$, this leaves us with only the sum of powers of $\mathbb{Z} \setminus \{0\}$ which is equal to $2\zeta(k)$. □

Our mentioning of the Eisenstein series is not merely to satisfy our constructivist insecurities; as we shall see in Section 5, the Eisenstein series are all that we shall need in our goal to completely characterize modular forms of any weight!

4 q-Expansions, Modular Functions, and the Fundamental Domain

It has been noted that modular forms exhibit $\mathbb{Z}$-periodicity, that is, $f(\tau + z) = f(\tau)$ for all $z \in \mathbb{Z}$. If $B' = \{z \in \mathbb{C} \mid |z| < 1\} \setminus \{0\}$, consider the covering map $q : H \to B'$ given by $q(\tau) = e^{2\pi i\tau}$.

Remark. It is customary to write merely q in referring to $q(\tau)$, and to interpret statements involving complex exponentiation q^z as $e^{2z\pi i\tau}$ (of course, this distinction is of no consequence when z is integral).

Note that at each point of B', there is a family of local holomorphic inverses that differ from each other merely by addition of some integer. Thus, given a holomorphic $\mathbb{Z}$-periodic function f on H, there is a unique holomorphic function F on B' that satisfies $f = F \circ q$. If we expand the Laurent series for F about the origin and plug in q instead, we arrive at the following series for f, referred to as its *q-expansion*:

$$f(\tau) = \sum_{m=-\infty}^{\infty} a_m q^m$$

Note that in the case of modular forms, by definition f is bounded as $\tau \to i\infty$. This is to say, F is bounded as $q \to 0$, so that by Riemann's Removable Singularities Theorem, $a_m = 0$ for all $m < 0$, i.e. F extends to a holomorphic function on all of the open unit ball B.

Definition 4.1. A function $f : H \to \mathbb{C}$ is said to be a *modular function* if it satisfies the following criteria:

1. f is meromorphic on H.
2. f is weakly modular of weight 0, i.e. $f(\gamma\tau) = f(\tau)$ for all $\gamma \in \Gamma$.
3. The corresponding $F : B' \to \mathbb{C}$ is meromorphic at $q = 0$. This is equivalent to the condition that either $f(\tau)$ be bounded as $\tau \to i\infty$ or $f(\tau) \to \infty$ as $\tau \to i\infty$.

From now on, we shall use H^* to refer to the upper-half plane together with $i\infty$, so that, for example, (1) and (3) may be subsumed under the condition "meromorphic on H^*."

Definition 4.2. If Y is a space with an afforded equivalence relation $\sim$, then a subspace $X \subseteq Y$ is called a *fundamental domain* for this relation if $x \sim y$ if and only if $x = y$ for $x, y \in X$, and the closure of X provides the full set of representatives for $\sim$.

It is well known that the region $\Omega = \{z \in \mathbb{C} \mid |z| > 1, -\frac{1}{2} < \mathrm{Re}(z) < \frac{1}{2}\}$ constitutes a fundamental domain for H under the action of Γ (see Figure 2) with only border points of equal absolute real part being equivalent under the action. This may be proven via elementary albeit tedious methods, so we take it as granted ([11], Chapter 3, Section 1).

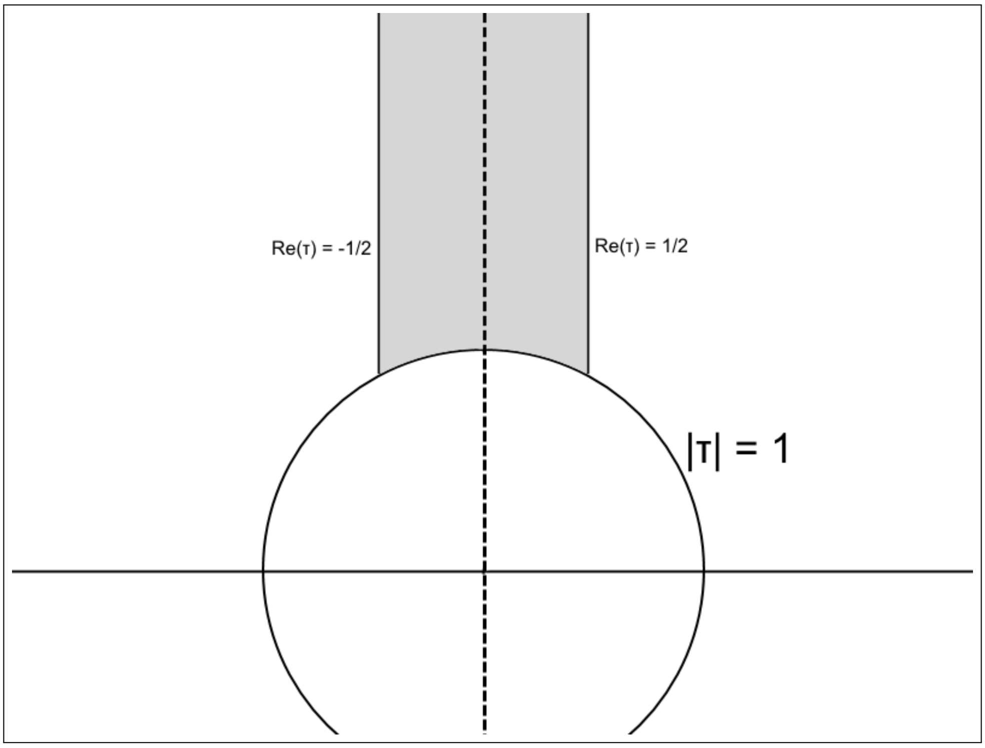

Figure 2: The fundamental domain of the modular action

From a geometrical point of view, we may view modular functions as being the meromorphic functions on the compact Riemann surface obtained by stitching opposite ends of the border of Ω together and adding an appropriate compactifying point at $i\infty$. Naturally, this implies that the only modular *forms* of weight 0 are constant, since any holomorphic function on a compact Riemann surface is constant.

5 The Graded Algebra of Modular Forms

Let nonzero f be weakly modular of weight k, meromorphic on H^*, and not constantly zero. Let $\rho = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $-\overline{\rho} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ be the bottom corners of the fundamental domain of the modular action. We shall integrate $\frac{f'(z)}{f(z)}$ over the contour indicated in Figure 3 (as R approaches ∞ and the radii of the detour arcs around the zeros/poles of f approaches 0) using two different methods which we shall compare in order to obtain useful information relating the order of the zeros of f.

Firstly, $\frac{f'(z)}{f(z)}$ will have a simple pole at and only at each zero/pole z of f, with residue the corresponding order $v_z(p)$ of the zero/pole. As such, the residue theorem tells us that the value of the (limit of) the integral is

$$-2\pi i \sum_{z\neq\rho,i} v_z(f)$$

where $z \in H$ varies over the zeros/poles of f unique up to the modular action, excluding representatives for i and ρ.

Secondly, we compute the integral a bit more directly. It is clear that $\mathbb{Z}$-periodicity of $\frac{f'(z)}{f(z)}$ will

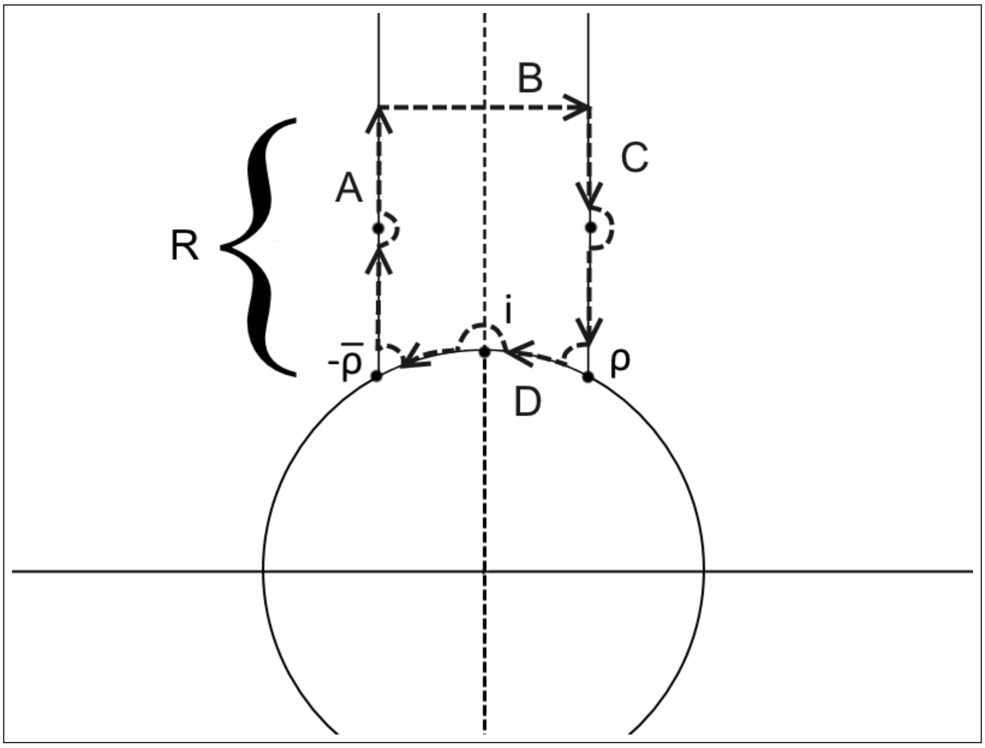

Figure 3: Integrating along the fundamental domain

entail that the integral over paths A and C (neither of which contain the detour arcs around ρ and $-\overline{\rho}$) cancel out. As well, integrating over the contour B, after applying the q-change of variables, is $2\pi i$ times the order of the zero/pole at $q = 0$, here notated by $v_\infty(f)$. Continuing on, we split the contour D (not including the detour arcs around ρ, $-\overline{\rho}$, or i) into its right and left halves, say E and F, respectively, and work with the complex change of variables $z \to -\frac{1}{z}$ sending E to F (and inverting orientation):

$$\oint_F \frac{f'(z)}{f(z)}\, dz = -\oint_E \frac{f'(-\frac{1}{z})}{z^2 f(-\frac{1}{z})}\, dz \tag{1}$$

Now, if we take the equality

$$f(-\frac{1}{z}) = z^k f(z)$$

and differentiate both sides, we get

$$\frac{f'(-\frac{1}{z})}{z^2} = kz^{k-1}f(z) + z^k f'(z) \implies f'(-\frac{1}{z}) = kz^{k+1}f(z) + z^{k+2}f'(z)$$

so that we can plug this expression for $f'(-\frac{1}{z})$ into (1) to get

$$-\oint_E \frac{f'(-\frac{1}{z})}{z^2 f(-\frac{1}{z})}\, dz = -\oint_E \frac{k}{z} + \frac{f'(z)}{f(z)}\, dz$$

Thus,

$$\oint_D \frac{f'(z)}{f(z)}\,dz = \oint_E \frac{f'(z)}{f(z)}\,dz + \oint_F \frac{f'(z)}{f(z)}\,dz = 2\pi i \cdot \frac{k}{12}$$

Finally, $\frac{f'(z)}{f(z)}$ composed with a suitable translation will look like

$$\frac{m}{z} + \text{holomorphic term}$$

at a zero/pole of f of degree m. Since the integral of a holomorphic term along a vanishingly small arc will itself vanish, we conclude that the integral of the detour arc around i is πi, and the sum of the integrals of the detour arcs around ρ and $-\overline{\rho}$ is $\frac{2\pi i}{3}$.

Comparing the results of our two different methods of computing the integral yields the following theorem.

Theorem 1. *Let f be weakly modular of weight k, meromorphic on H^*, and not constantly zero. Let $\rho = \frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $-\overline{\rho} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ be the bottom corners of the fundamental domain of the modular action. If $v_z(f)$ denotes the order of the zero/pole of f at z, then*

$$4v_\rho(f) + 6v_i(f) + 12v_\infty(f) + 12\sum_{z\neq\rho,i} v_z(f) = k$$

where $z \in H$ varies over the zeros/poles of f unique up to the modular action, excluding representatives for i and ρ.

Remark. Those familiar with Riemann surface theory and the branched mapping principle (if $f : X \to Y$ is a proper nonconstant holomorphic map between Riemann surfaces, then up to multiplicities, f takes any value the same number of times; thus, any meromorphic function on a compact Riemann surface has as many zeros as poles) might notice that when $k = 0$ (i.e. when f is a modular function), this is merely the statement that the analytic quotient map $H^* \to H^*/\Gamma$ (where the latter is endowed with an appropriate complex structure) ramifies with index 3 over ρ and index 2 over i.

Believe it or not, this formula together with our Eisenstein series are all that we need to construct any modular form imaginable!

Theorem 2. *Let M_k denote the complex vector space of modular forms of weight k, and let M denote the $\mathbb{C}$-algebra that is generated by M_k for all natural k. Then $M = \oplus M_k$ is a graded $\mathbb{C}$-algebra, and if G_4 and G_6 represent the Eisenstein series of weights 4 and 6 respectively, then $\mathbb{C}[x, y] \to M$ given by $x \mapsto G_4$ and $y \mapsto G_6$ is an isomorphism of $\mathbb{C}$-algebras.*

Proof. Note that $M_n M_m \subseteq M_{n+m}$, so that to verify $M = \oplus M_k$ it only remains to show that nontrivial finite sums of modular functions of varying weights is nonzero (we shall in fact have to prove something slightly stronger). Suppose this is not the case, and that we have a set $f_1, f_2, ..., f_n$ of nonzero forms of weights $k_1 < k_2 < ... < k_n$ respectively satisfying a linear dependence not necessarily over $\mathbb{C}$, but merely over $\mathbb{C}(\tau)$ (the set of complex rational functions):

$$p_1 f_1 + p_2 f_2 + ... + p_n f_n = 0$$

We may assume $p_1 = 1$. Now, it is clear that no nonconstant rational function is $\mathbb{Z}$-periodic, for otherwise roots and poles would not be preserved. If one of p_i is not constant, then applying the transformation $\tau \mapsto \tau + 1$ to both sides of this expression and subtracting this new dependence from the original would give us a nontrivial $\mathbb{C}(\tau)$-linear dependence among the $f_2, ..., f_n$. If the p_i are constant for all i, then apply the transformation $\tau \mapsto -\frac{1}{\tau}$ to both sides, divide by τ^{k_1}, and again subtract one dependence from the other to obtain a nontrivial $\mathbb{C}(\tau)$-linear dependence among the $f_2, ..., f_n$. Induction now applies to verify $M = \oplus M_k$.

Now, we shall first show that the aforementioned map $\mathbb{C}[x, y] \to M$ is surjective. We shall proceed by induction on k to show that M_k is within the image of this map. In each case suppose f is a nonzero modular form of weight k.

Case $k = 0$**:** Since modular forms have no poles on H^*, the variables in the formula of Theorem 1 must be nonnegative integers. When $k = 0$ necessarily all these must be 0, so that f has no zeros. Subtracting an appropriate constant to give it a zero will force the result to be constantly zero, from which we deduce f is constant. $M_0 = \mathbb{C}$.

Case $k = 2$**:** There is no way to satisfy Theorem 1. $M_2 = \varnothing$.

Case $k = 4$**:** The only possibility is that f has a unique and simple zero at ρ (and its Γ-equivalents). For some constant c, we observe $f - cG_4$ will have another zero, and be forced to be constantly zero. $M_4 = \mathbb{C}G_4$.

Case $k = 6$**:** The only possibility is that f has a unique and simple zero at i. For some constant c, we observe $f - cG_6$ will have another zero, and be forced to be constantly zero. $M_6 = \mathbb{C}G_6$.

Case $k = 8$**:** The only possibility is that f has a unique and double zero at ρ. For some constant c, we observe $f - cG_4^2$ will have another zero, and be forced to be constantly zero. $M_8 = \mathbb{C}G_4^2$.

Case $k = 10$**:** The only possibility is that f have unique and simple zeros at ρ and i. For some constant c, we observe $f - cG_4G_6$ will have another zero, and be forced to be constantly zero. $M_8 = \mathbb{C}G_4G_6$.

Case $k \geq 12$**:** Note that G_4 and G_6 have unique and simple zeros at ρ and i, respectively, so that for any other point τ, we may choose an appropriate constant c so that $G_4^3 - cG_6^2$ has a (necessarily unique and simple) zero at τ. We now have, for every point $\tau \in H^*$, a modular form G, generated by G_4 and G_6, with a unique and simple zero at τ. If τ is chosen to be a zero of f, we may write $\frac{f}{G}$ to get a modular form of smaller weight; we conclude the case analysis by an appeal to induction and multiplying both sides by G.

All that remains is to prove that the map $\mathbb{C}[x, y] \to M$ is injective, i.e. that G_4 and G_6 are algebraically independent over $\mathbb{C}$. But this should be clear in light of what we have already worked out: If $c_1G_4^r + c_2G_4^{r-3}G_6^2 +$ etc. is a weight-balanced expression that is constantly zero (recall we have already proven that sums of nonzero weight-balanced expressions of distinct weights will be nonzero), then $c_1G_4^r(i) = 0$, so $c_1 = 0$; we may thus divide by G_6 and appeal to induction. □

In the case that $\tau = i\infty$ in the last case, this makes G an example of a *cusp form.* Ater multiplying G by a constant so that $a_1 = 1$ in its q-expansion (technically we don't yet know this is possible), we arrive at Δ, known as the *modular discriminant* (the term "discriminant" comes from its appearance in elliptic curve theory). Now, we are in a position to define the elusive j-invariant: $j = \frac{G_4^3}{8\zeta(4)^3\Delta}$. Note that j is a modular function with a q-expansion beginning

$$j(\tau) = \frac{1}{q} + ...$$

Similar in utility to the Eisenstein series, j generates the space of modular functions:

Proposition 3. *Every modular function can be represented as a (complex) rational function of j.*

Proof. Note that by Theorem 1, j induces a bijection (in fact, an analytic isomorphism) from H^*/Γ to the Riemann sphere. Thus, for any nonconstant modular function g with some prescribed zero/pole at τ, we may divide/multiply it by $j - c$ for some constant c with $j(c) = \tau$ and proceed by induction until g is constant. □

We can almost taste the moonshine. We merely need a clever way of computing the Fourier coefficients of G_4 and G_6.

6 Cotangent Cleverness

Consider the function $f(\tau) = \pi \cot(\pi\tau)$. It is a $\mathbb{Z}$-periodic meromorphic function on $\mathbb{C}$ with unique and simple poles at each integer, each with residue 1. It is also bounded in each open set of uniform distance away from each integer. But these exact same properties are also shared by the function

$$g(\tau) = \frac{1}{\tau} + \sum_{d=1}^{\infty} \frac{1}{\tau + d} + \frac{1}{\tau - d}$$

It follows that the function $f - g$ is a bounded entire function, and thus constant by Liouville's theorem. But in a Laurent series expansion about the origin, both of f and g have zero constant term. It follows that $f = g$.

As the astute reader might already sense, the $(k-1)^{\text{th}}$ derivative of g can be used to construct the Eisenstein series. Before we perform this, then, we calculate the q-expansion of g, or equivalently, of f:

$$\pi\cot(\pi\tau) = \pi\frac{\cos(\pi\tau)}{\sin(\pi\tau)} = \pi\frac{i(q^{1/2} + q^{-1/2})}{q^{1/2} - q^{1/2}} = \pi\frac{i(q+1)}{q-1} =$$

$$-\pi i(q+1)(1 + q + q^2 + ...) = \pi i - 2\pi i \sum_{m=0}^{\infty} q^m$$

Now, when $k > 2$ is an even integer, we take the $(k-1)^{\text{th}}$ derivative of both sides of the equality $\frac{1}{\tau} + \sum \frac{1}{\tau+d} + \frac{1}{\tau-d} = \pi i - 2\pi i \sum q^m$ to obtain

$$(-1)^{k-1}(k-1)! \sum_{d\in\mathbb{Z}} \frac{1}{\tau + d}^{k} = -(2\pi i)^k \sum_{m=0}^{\infty} m^{k-1} q^m$$

Since k is even, $(-1)^{k-1} = -1$; multiply through appropriately to get

$$\sum_{d\in\mathbb{Z}} \frac{1}{(\tau+d)^k} = \frac{(2\pi i)^k}{(k-1)!} \sum_{m=0}^{\infty} m^{k-1} q^m$$

Sum this equality with the substitutions $c\tau$ for $c \in \mathbb{Z} \setminus \{0\}$, and also add the equality $\sum_{d\in\mathbb{Z}\setminus\{0\}} \frac{1}{d^k} = 2\zeta(k)$ to arrive at a formula for $G_k(\tau)$. Since k is even, the sums will simplify to

$$G_k(\tau) = 2\zeta(k) + 2\sum_{c=1}^{\infty}\sum_{d\in\mathbb{Z}} \frac{1}{(c\tau+d)^k} = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!} \sum_{c=1}^{\infty}\sum_{m=0}^{\infty} m^{k-1} q^{cm} =$$

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

where $\sigma_k(N)$ denotes the sum of the k^{th} powers of the positive divisors of a number N.

Recall $\zeta(4) = \frac{\pi^4}{90}$ and $\zeta(6) = \frac{\pi^6}{945}$. As such, the first few values of the q-expansions of G_4 and G_6 are

$$G_4(\tau) = \frac{\pi^4}{45}(1 + 240\sum_{m=1}^{\infty}\sigma_3(m)q^m) = \frac{\pi^4}{45}(1 + 240q + 2160q^2 + 6720q^3 + ...)$$

$$G_6(\tau) = \frac{2\pi^6}{945}(1 - 504\sum_{m=1}^{\infty}\sigma_5(m)q^m) = \frac{2\pi^6}{945}(1 - 504q - 16672q^2 - 122976q^3 - ...)$$

In fact, after normalizing $G_k(\tau)$ so that its constant term is equal to 1, the resulting q-coefficients will always be straight integers, due to the relation $\zeta(2n) = (-1)^{n+1}\frac{B_{2n}(2\pi)^{2n}}{2(2n)!}$.

After performing the necessary arithmetic with the Fourier series of G_4 and G_6, we may rotely compute the Fourier series for the j-invariant:

$$j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + ...$$

7 A Brief Sideline on Integral Matrices

Before we continue, it is important to understand a fact or two about (integer-valued) matrices and their corresponding action on lattices. If as before we pass to the moduli space of lattices by associating τ with $[1, \tau]$, then the modular action $\tau \mapsto \frac{a\tau+b}{c\tau+d}$ induces an action on lattices $[1, \tau] \mapsto \frac{1}{c\tau+d}[c\tau + d, a\tau + b] = \frac{1}{c\tau+d}[1, \tau]$ (those familiar with elliptic curve theory will know that this latter lattice is analytically isomorphic to the original), but this is not quite what we're talking about. Instead, similar to the case for vector spaces, we consider integral matrices as endomorphisms of free $\mathbb{Z}$-modules and their resultant images.

Proposition 4. *Let L be a finitely generated free $\mathbb{Z}$-module with a prescribed basis $e_1, ..., e_n$. Then if M is a nonsingular $n \times n$ matrix with integer entries and we interpret M as a linear transformation $M : L \to L$, then $[L : M(L)] = |\det M|$, i.e. $L/M(L)$ is an abelian group of order $|\det(M)|$.*

Proof. By repeated application of the Euclidean algorithm, put M into row echelon form M' by interchanging, adding, and subtracting rows. Then $M(L) = M'(L)$ and $|\det M| = |\det M'|$; write $\det M' = \prod d_i$, where the d_i are the diagonal entries of M'. It is clear that every element of the quotient $L/M'(L)$ may be written in a reduced form with respect to the generators $M'(e_i)$ where the i^{th} coefficient c_i satisfies $0 \le c_i < |d_i|$; as well, by upper-triangularity of M', this reduced form will be unique. As such, $L/M(L) = L/M'(L)$ is an abelian group of order $|\prod d_i| = |\det M'| = |\det M|$. □

With L as in the above proposition, it is a well-known fact (cf. [4], Chapter 12) that every subgroup of L will be free of rank $\le n$. As such, every subgroup of L will be the image of some M. Or, put in different terms, the orbit of L under the ring action of $\mathrm{GL}_n(\mathbb{Z})$ on the subgroup lattice of L is the entire subgroup lattice.

If we are interested only in subgroups of L of a given finite index m, then it is necessary to consider which matrices/linear transformations yield identical images. Since $\pm\mathrm{SL}_n(\mathbb{Z})$ consists of precisely the matrices that correspond to linear isomorphisms of free $\mathbb{Z}$-modules of rank n, the answer is that it is the size of the set of integral matrices of determinant $\pm m$ quotiented out by the action of $\pm\mathrm{SL}_n(\mathbb{Z})$ under left multiplication. Since $\pm\mathrm{SL}_n(\mathbb{Z})$ contains (and is in fact generated by) all the matrices that correspond to row addition/subtraction/interchanging, the row echelon form considered in the proof of the above proposition gives us a way to obtain a class representative of a matrix under this action. In fact, if a matrix is put into lower-triangular form with positive diagonal entries such that each lower entry $c_{i,j}$ satisfies $0 \leq c_{i,j} < d_j$, then this form will be unique, for the entries are uniquely determined by $M(L)$: With respect to the lexicographic ordering $e_n > e_{n-1} > ... > e_1$, row i of our reduced matrix represents the smallest element of $M(L)$ subject to nonnegative entries and a nonzero i^{th} coefficient. The number of such reduced forms is easy to compute, and the result is stated below:

Theorem 3. *Let L be a finitely generated free $\mathbb{Z}$-module of rank n. Then the number of subgroups of L of finite index m is given by*

$$\sum_{(q_1,...,q_n)} \prod_i q_i^{i-1}$$

where $(q_1, ..., q_n)$ runs over all positive n-tuples with $\prod q_i = m$.

Corollary 1. *The number of sublattices of a given lattice in $\mathbb{C}$ of finite index m is given by $\sigma(m)$, the sum of the divisors of m.*

8 Hecke Operators

Given a modular form of given weight, how can we apply some interesting and natural transformation to yield another modular form of the same weight? If this transformation is required to be linear $M_k \to M_k$, then there is a good chance that complex scalar multiplication is the only way to go about it, since M_k for small k (specifically, $k = 0, 4, 6, 8, 10, 14$) is one-dimensional. But this doesn't make the endeavor fruitless; in fact, a prima facie "interesting" transformation that turns out to be "uninteresting" means that we have discovered some simple way to describe the not-so-simple. This is what happens when one considers Hecke operators and the corresponding Hecke eigenforms.

To define some action on the analytic f, we shall again turn to its interpretation as a complex-valued function on lattices.

Definition 8.1. Let $\mathfrak{L}$ denote the free abelian group on the generators $\mathbb{L}$. We define the *Hecke operator $T(n)$ of order n* to be the endomorphism $T(n) : \mathfrak{L} \to \mathfrak{L}$ induced by

$$T(n)(L) = \sum_{[L:L']=n} L'$$

Any complex-valued function on lattices extends uniquely to a morphism $\mathfrak{L} \to \mathbb{C}$, so extend the action of $T(n)$ to complex-valude functions on lattices by $T(n)F(L) = F(T(n)L) = \sum_{[L:L']=n} F(L')$. Once again extend the action of $T(n)$ to modular functions by $T(n)f(\tau) = T(n)F([1, \tau])$.

Remark. At this point it's necessary to disambiguate $nL \in \mathfrak{L}$ as signifying either $\sum_{i=1}^{n} L$, or $\{n\omega \mid \omega \in L\}$ as in the preceding sections. From now on, we shall use nL to mean exclusively the former, and $R(\lambda)L$ for $\lambda \in \mathbb{C}^{\times}$ will mean the latter.

Definition 8.2. The algebra of operators generated by the $T(n)$ and $R(m)$ for all n, m with composition for multiplication is called the Hecke algebra, notated by $\mathfrak{H}$.

It is easy to verify that if F satisfies the homogeneity property, then so does $T(n)F$, since the operator $R(\lambda) : L \to R(\lambda)L$ is an isomorphism and thus provides a bijection between sublattices of index n. As well, it will soon become clear that if f is a modular form, then $T(n)f$ is holomorphic on H^*, and is thus a modular form.

Note that, by the results of the preceding section, we can write

$$T(n)L = \sum_{\gamma} \gamma L$$

where the sum runs over the integral matrices

$$\gamma = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$$

with $a, d > 0$, $ad = n$, and $0 \leq c < a$. We may thus compute

$$T(n)f(\tau) = \sum_{\gamma} F(\gamma[1, \tau]) = \sum F([a, c + d\tau]) = \sum F((a)[1, \frac{c + d\tau}{a}]) =$$

$$\sum a^{-k} F([1, \frac{c + d\tau}{a}]) = \sum a^{-k} f(\frac{c + d\tau}{a})$$

From this vantage, we may explicitly compute the effect of $T(n)$ on the q-expansion of f. If as before

$$f(\tau) = \sum_{m=0}^{\infty} a_m q^m$$

then

$$T(n)f(\tau) = \sum_{a|n} \sum_{c=0}^{a-1} \sum_{m=0}^{\infty} a_m e^{2\pi i (\frac{c+d\tau}{a})m} = \sum_{m=0}^{\infty} \sum_{a|n} a^{-k} a_m e^{2\pi i \frac{dm}{a}\tau} \sum_{c=0}^{a-1} e^{2\pi icm/a}$$

Focus on the expression $\sum_{c=0}^{a-1} e^{2\pi icm/a}$; note that multiplying it by $e^{2\pi im/a}$ will not affect the sum. If $a \nmid m$, then $e^{2\pi im/a} \neq 1$, so this is to say $\sum_{c=0}^{a-1} e^{2\pi icm/a} = 0$ in this case. If instead $a \mid m$, then $\sum_{c=0}^{a-1} e^{2\pi i \frac{c}{a} m} = \sum_{c=0}^{a-1} 1 = a$. So we continue:

$$T(n)f(\tau) = \sum_{m=0}^{\infty} \sum_{a|(n,m)} a^{-k+1} a_m q^{dm/a} = \sum_{m'=0}^{\infty} \sum_{a|n} a^{-k+1} a_{m'a} q^{dm'} = \sum_{m''=0}^{\infty} \sum_{d|(n,m'')} \left(\frac{n}{d}\right)^{-k+1} a_{m''n/d^2} q^{m''}$$

This is finally in a form we want. We clean up the notation and summarize the result in the following theorem.

Theorem 4. *Let f be a modular form of weight k, and let $T(n)$ be the Hecke operator of order n.*

If f has q-expansion $\sum a_m q^m$, then $T(n)f$ is a modular form with q-expansion $\sum b_m q^m$, where

$$b_m = n^{-k+1} \sum_{d|(n,m)} d^{k-1} a_{mn/d^2}$$

If $\sigma_k(N)$ denotes the sum of the k^{th} powers of the positive divisors of a number N, then in particular

$$b_0 = n^{-k+1}\sigma_{k-1}(n)a_0 \qquad b_1 = n^{-k+1}a_n$$

Proposition 5. *Every Eisenstein series G_k is a Hecke eigenform of every order, i.e. for every natural n, there exists a constant $\lambda \in \mathbb{C}$ such that $T(n)G_k = \lambda G_k$.*

Proof. The proof of the proposition is split into three parts. First, we show that the Hecke operator assignment is a multiplicative function, i.e. $T(nm) = T(n)T(m)$ when n and m are relatively prime. Second, we show that in the algebra of Hecke operators (with composition as multiplication) $T(p^r)$ is in the subalgebra generated by $T(p)$ for each prime p and all r. Third, we show that every Eisenstein series is a Hecke eigenform of every prime order. This will be enough to verify the statement.

The first step is the easiest. Let L' be a subblattice of L of index mn, and consider the abelian group L/L'; it contains unique subgroups of index m and n, which is to say that there are unique sublattices L_n and L_m of of L of index n and m respectively containing L'. Thus, the list of all sublattices of index n in another sublattice of index m in L is the same as the list of all sublattices of index mn in L, including (nonexistent) multiplicities. This carries over to Hecke operators to verify $T(nm) = T(n)T(m)$.

Now, consider $T(p^r)$ for $n \geq 2$. We verify

$$T(p^r) = T(p^{r-1})T(p) - pR(p)T(p^{r-2})$$

We must show that the multiplicities of the sublattices associated to each operator are the same. To this end, let $L' \subseteq L$ be of index p^n, and suppose $L' \subseteq R(p)L$. Then since $R(p)L$ is contained in every sublattice of index p (of which there are $\sigma(p) = p+1$), and since $R(n)$ and $T(m)$ commute for all n, m ($R(n) : L \to L$ is an embedding), both sides of the equation assign the multiplicity 1 to L'. So suppose $L' \nsubseteq R(p)L$. Then since the intersection of any two distinct sublattices of index p is $R(p)L$, we again find that both sides of the equation assign the multiplicity 1 to L'.

Finally, interpret G_k as a complex-valued function on lattices.

$$T(p)G_k(L) = \sum_{[L:L']=p} G_k(L')$$

Note that the L' form a cover of L with multiplicity $p+1$ on elements of $R(p)L$ and multiplicity 1 elsewhere. By definition of the Eisenstein series, we may thus equate

$$\sum_{[L:L']=p} G_k(L') = G_k(L) + pG_k(pL) = (1 + p^{1-k})G_k(L)$$

□

Although we will not have a chance to apply our theory of Hecke operators further to modular form theory, they are a basic tool for any interested reader moving forward. One interesting result that can already be obtained from our study of Hecke operators, though, is number-theoretic: If f is a Hecke eigenform of all orders with Fourier coefficients a_n such that $a_1 = 1$, then the third equation of Theorem 4 together with our previous observation that $T(nm) = T(n)T(m)$ for relatively prime n and m implies Fourier coefficient assignment is actually multiplicative: $a_{nm} = a_n a_m$ if n and m are relatively prime. This is precisely how Mordell [12] in 1917 proved the Ramanujan tau function $\tau(n)$ is multiplicative, where $\tau(n)$ is the n^{th} Fourier coefficient of the modular discriminant form Δ (Δ must a Hecke eigenform of all orders as it is the unique cusp form of weight 12 up to scalar multiplication), among other interesting properties.

9 Representation Theory

It's no secret that groups like to act on things. It's also no secret that mathematicians understand linear algebra better than almost any other subject. It is as such that the following investigation of (finite) groups acting on vector spaces is motivated.

Definition 9.1. Let G be a finite group, let F be a field, and let V be a vector space over F of finite rank. Then a *representation* of G over F is a group morphism $\varphi : G \to \mathrm{GL}(V)$. Alternatively, G is said to *act on the space* V, namely by $g \cdot v = \varphi(g)(v)$. The degree of the representation is defined to be the rank of the V over F.

Definition 9.2. Consider the vector space W of rank $|G|$ over F, with some basis e_g indexed by $g \in G$. Then there is an action of G on W induced by $h \cdot e_g = e_{hg}$ for each $h, g \in G$, affording a representation φ. As elements of $\mathrm{GL}(V)$, the $\varphi(h)$ are linearly independent over F, since only the null linear combination of them will send e_1 to 0. We call the F-algebra generated by the elements $\varphi(h)$ the *group ring* FG, and it is customary to write the elements as linear sums of elements of G over F, omitting φ. For example,

$$3 + \frac{4}{3}r - r^3 + s$$

is an element of $\mathbb{Q}D_8$.

If we consider G as being contained in FG and inheriting its multiplicative structure, then given a representation $\varphi : G \to \mathrm{GL}(V)$, there is a unique morphism of F-algebras $\varphi : FG \to \mathrm{GL}(V)$ extending φ. Conversely, every morphism of F-algebras φ restricts to a representation of G over F. In perhaps more familiar terms, a morphism of F-algebras $\varphi : FG \to \mathrm{GL}(V)$ is just the definition of an FG-module structure on V. It is as such that we note the following correspondences:

$$\left\{ V \text{ an } FG\text{-module} \right\} \longleftrightarrow \left\{ \begin{array}{c} V \text{ a vector space over } F \\ \text{and} \\ \varphi : G \to \mathrm{GL}(V) \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} V \text{ a vector space over } F \\ \text{and} \\ G \text{ acts on } V \end{array} \right\}$$

Example 1. If $V = F$, we may define a trivial action of G on V via $gv = v$. This is known as the *principal representation* of G, and (believe it or not) carries some interest in the character theory presented in the next section.

Example 2. FG acting on itself by left multiplication gives it the structure of an FG-module. The associated representation is known as the *regular representation* of G.

Proposition 6. *Let G be a finite group with two representations $\varphi : G \to GL(V)$, $\psi : G \to GL(W)$ of the same degree over the same field F. The following three conditions are equivalent:*

(1) Ring/module theoretic: V and W are isomorphic as FG-modules.

(2) Matrix computational: There exists some nonsingular matrix A that simultaneously conjugates $\varphi(g)$ into $\psi(g)$ for all $g \in G$.

(3) Category theoretic: There exists some linear isomorphism $T : V \to W$ allowing the following diagram to commute for all $g \in G$:

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "\varphi(g)"'] & W \arrow[d, "\psi(g)"] \\
V \arrow[r, "T"] & W
\end{tikzcd}$$

In such a case, the representations φ and ψ are said to be equivalent.

Proof. (1) $\implies$ (2): Suppose $A : V \to W$ is an FG-module isomorphism. Then A may be considered as a linear isomorphism of vector spaces over F such that $A(g \cdot v) = g \cdot A(v)$ for all $v \in V$ and $g \in G$. That is to say as matrices $A \cdot \varphi(g) = \psi(g) \cdot A$, from which (2) follows.

(2) $\implies$ (3): This is simply (2) with T in place of A and converted into the language of linear transformations.

(3) $\implies$ (1): We see T must also be a morphism of FG-modules, as $T(g \cdot v) = T(\varphi(g)(v) = \psi(g)(Tv) = g \cdot T(v)$, and hence also an isomorphism. □

Suppose FG is a group ring with V an FG-module. The module/representation/group action is said to be *faithful* if the associated representation $\varphi : G \to \mathrm{GL}(V)$ is injective; as such, the group structure is fully demonstrated in its action on V. Suppose further that $W_1 \subset V$ is a proper FG-submodule (this is equivalent to the condition that $gW_1 \subseteq W_1$ for all $g \in G$, i.e. W_1 is stabilized under the action of G). The associated action of G on W_1 is in a certain sense simpler than that on V, although it may lose its faithfulness. In ideal circumstances, we should wish that V decomposes as a direct FG-module sum involving W_1:

$$V = W_1 \oplus W_2$$

in which case the action of G on V is completely described by the corresponding actions on W_1 and W_2, and $\varphi = \psi_1 \times \psi_2$ for the corresponding representations ψ_1 and ψ_2. The following theorem demonstrates that this is always the case given nice F and G, and is fundamental to the representation theory of finite groups.

Theorem 5. (Maschke's Theorem) *Let G be a finite group of order n, and let F be a field of characteristic not dividing n. Then every FG-module is* injective*, i.e. for every pair of FG-modules W_1 and V such that $W_1 \subseteq V$, there is another FG-submodule W_2 of V such that $V = W_1 \oplus W_2$.*

Proof. Since W_1 and V are in particular vector spaces over F, we may assume that there is a direct sum decomposition of V involving W_1 in terms of vector spaces. By means of this decomposition, let $\pi : V \to W_1$ be an F-linear projection, i.e. a linear transformation such that $\pi(w) = w$ for all $w \in W_1$. Then the idea is to construct from π an FG-module projection τ of V onto W_1; the theorem will then follow immediately with $W_2 = \ker \tau$.

Given a linear transformation $\varphi : M \to N$ of vector spaces over F, we may define $g\varphi$ and φg in the obvious way, and also define $g * \varphi = g\varphi g^{-1}$. These are seen to be actions of G on the set $\mathrm{Hom}(M, N)$, which uniquely extend in order to give $\mathrm{Hom}(M, N)$ FG-module structures. By similar reasoning as in Proposition 6, it is a necessary and sufficient condition that $g * \varphi = \varphi$ for all g in order for φ to be a morphism of FG-modules. As such, we shall "average" π over G in order to construct τ (this is where char $F \nmid n$ is necessary):

$$\tau = (\frac{1}{n} \sum_{g \in G} g) * \pi$$

Since $\frac{1}{n} \sum g$ absorbs multiplication by elements of G, and $*$ is associative, τ is by the above a morphism of FG-modules. Since as well $\tau(w) = \frac{1}{n} \sum g\pi(g^{-1}w) = \frac{1}{n} \sum gg^{-1}w = w$, we have τ is a projection, completing the proof. □

Example 3. Without the condition char $F \nmid n$, Maschke's Theorem does not hold in general. For example, let F be a field of characteristic p, and let P be a nontrivial p-group. Then the one-dimensional subspace of FP generated by $\sum_{g \in P} g$ is also an FP-submodule, yet we shall show FP does not decompose as a direct sum of *any* two proper submodules (this proof may be skipped without any hindrance to the rest of this paper's exposition). First, a lemma:

Lemma 1. *Let G be a group (not necessarily finite), let F be a field, and let V be an FG-module. If $N \trianglelefteq G$ is a normal subgroup, then the set of elements $W \subseteq V$ that are fixed by N is an FG-submodule, and there is a natural $F(G/N)$-module action on W. The FG-submodules and $F(G/N)$-submodules of W are the same. Furthermore, if the module V is itself FG affording the regular representation on G, then the induced module is isomorphic to the regular representation of G/N.*

Proof. It is clear W is a subspace, and if $n \in N$, $g \in G$, and $w \in W$, then $ng \cdot w = gg^{-1}ng \cdot w = g(g^{-1}ng) \cdot w = g \cdot w$, so that W is G-stable. There is a natural action of G/N on W, thereby turning W into a $F(G/N)$-module under this action. We see an F-subspace of W is G-stable if and only if it is G/N-stable, so its FG-submodules and $F(G/N)$-submodules are the same.

Suppose V is the module given by the regular representation. For each distinct left coset xN define

$$\alpha_{xN} = \sum_{g \in xN} g$$

as an element of FG. It is clear, then, that $\alpha_{xN} \in W$ and the α_{xN} are linearly independent over F. Conversely, any element of FG fixed by each element of N must retain equivalent coefficients on elements in the same left coset of N, hence be an F-linear combination of the α_{xN}. As such, the α_{xN} form an F-basis for W. Define an F-linear isomorphism $\varphi : W \to F(G/N)$ given by $\varphi(\alpha_{xN}) = \overline{x}$. This is in fact an $F(G/N)$-module isomorphism, since $\varphi(\overline{g}\alpha_{xN}) = \varphi(\alpha_{gxN}) = \overline{gx} = \overline{g}\varphi(\alpha_{xN})$. □

We proceed by induction on the order of P. The base case will obviously be satisfied when $|P| = 1$. Suppose $FP = V_1 \oplus V_2$ as FP-modules. If P is nontrivial, let $x \in Z(P)$ be of order p. Then on any FP-submodule of FP, we have $(x - 1)^p = x^p - 1 = 0$ as FP-module transformations, hence $x - 1$ has nontrivial kernel in each of V_1, V_2; this is to say the subspaces of elements fixed by $\langle\ x\ \rangle$ in each of V_1, V_2 are nontrivial, call them W_1, W_2 respectively. Then if W is the FP-submodule of elements of FP fixed by $\langle\ x\ \rangle$, we in fact have $W = W_1 \oplus W_2$. This nontrivial direct sum expression as FP-modules translates to the same as $F(P/\langle\ x\ \rangle)$-modules. But now $W \cong F(P/\langle\ x\ \rangle)$ may be written as a nontrivial direct sum, a contradiction. □

A module is said to be *irreducible* if it contains no nonzero proper submodules (and *reducible* otherwise), and a representation is said to be *irreducible* (or *simple*) if the FG-module affording it is irreducible. By repeated application of Maschke's Theorem, we may write any finite-degree representation of G as a direct product of irreducible representations; as for the afforded module, we say that it is *completely reducible* (confusing terminology note: completely redicible does not imply reducible, for irreducible modules are trivially completely reducible!). In the next section, we shall find that this decomposition is unique. Before that, however, we will establish Wedderburn's Theorem, which tells us among other things that there are a finite number of distinct irreducible representations.

Theorem 6. (Wedderburn's Theorem) *Let R be a ring with identity. The following five conditions are equivalent:*

(1) Every R-module is injective.

(2) Every R-module is completely reducible.

(3) For every R-module M and proper submodule $N \subset M$, there exists an irreducible submodule $L \subseteq M$ such that $N \cap L = 0$.

(4) Considering R as a left R-module, we have the following decomposition of R into irreducible submodules (i.e. left ideals):

$$R = Re_1 \oplus Re_2 \oplus \ldots \oplus Re_n$$

where the e_i are orthogonal idempotents summing to 1.

(5) As a ring, R is isomorphic to the m-fold direct product of $n_i \times n_i$ matrix rings of over division rings Δ_i. Up to permutation of the factors, this decomposition is unique, with m, n_i, and Δ_i being uniquely determined by R.

Under any of these conditions, R is said to be semisimple with minimum condition.

Proof. (2) $\implies$ (1): Suppose $N \subseteq M$ are R-modules. Consider the collection $\mathfrak{L}$ of submodules of M that have trivial intersection with N, partially ordered by inclusion. Since every chain in $\mathfrak{L}$ has an upper bound (namely, by taking the union of the chain), by Zorn's Lemma, we may choose a maximal member $L \in \mathfrak{L}$. Suppose $N \oplus L \neq M$; then completely reducing M/L, there is an irreducible submodule $\overline{L'}$ with trivial intersection with $\overline{N}$. If L' is the complete preimage of this submodule, then $L \subset L'$ and L' has trivial intersection with N, contradicting the maximality of L.

(3) $\implies$ (2): Let M be an R-module. Consider the collection $\mathfrak{N}$ of submodules of M which have a direct sum decomposition, partially ordered by $N_1 \leq N_2$ if N_1 has a direct sum complement in N_2. Since every chain in $\mathfrak{N}$ has an upper bound (namely, by taking the union of the chain), by Zorn's Lemma we may choose a maximal member $N \in \mathfrak{N}$. If $N \neq M$, then choose an irreducible submodule $L \subseteq M$ with $L \cap N = 0$; then $L \oplus N$ contradicts the maximality of N.

(4) $\implies$ (3): Choose an element $m \in M \setminus N$; then for some i, we have $e_i m \neq 0$. Then $Re_i m$ will be irreducible, as Re_i is an irreducible left ideal in R, and necessarily have empty intersection with N (lest $Re_i m \cap N$ be a nonzero proper submodule of $R_e im$).

(5) $\implies$ (4): We may assume without loss of generality that R is but a single matrix ring over a division ring; then if e_i are the matrices with 1 in the (i, i)-entry and 0 elsewhere, Re_i will be the submodule of R consisting of matrices with 0 in every column besides the i^{th}. As can be readily checked, every nonzero element of Re_i generates Re_i as a submodule, hence Re_i is irreducible. Clearly $R = Re_1 \oplus Re_2 \oplus \ldots \oplus Re_n$, and the e_i will be orthogonal idempotents summing to 1.

(1) $\implies$ (5): This requires the most effort. First, we show that R satisfies the descending chain condition (DCC) on left ideals (since every R-module is injective, it will also demonstrate the ascending chain condition [ACC], since decreasing left ideals leads to increasing direct sum complements, and vice versa). Suppose there existed an infinite decreasing chain of left ideals of R:

$$R = M_0 \supset M_1 \supset M_2 \supset M_3 \supset ...$$

Then, if $N = \cap M_i$, and $M_i = N_{i+1} \oplus M_i$, we may decompose R as an infinite direct sum $R = N \oplus (\bigoplus N_i)$. But this is impossible for a ring with identity, for then the identity could be written as a sum involving terms from only a finite number of the direct summands, which summands would then generate the whole ring as a left ideal. Therefore, R satisfies the DCC, and so too does every quotient of R.

By the DCC, we may obtain a minimal nonzero two-sided ideal R_1 of R. Let R' be its direct sum complement in R. Suppose R' were not a two-sided ideal; then the right-sided closure of R' would be a two-sided ideal containing both R_1 and R', hence be all of R, hence contain the identity, so we could write $1 = s_1r_1 + s_2r_2 + ... + s_mr_m$ for some $s_i \in R'$ and $r_i \in R$. Let $q \in R_1$ be nonzero. Then $qs_i \in R_1 \cap R' = 0$, yet $q \cdot 1 = q(s_1r_1 + s_2r_2 + ... + s_mr_m) = 0$, a contradiction. Thus R_1 and R' are two-sided ideals. They are seen to be orthogonal, so we may write $R = R_1 \times R'$ as rings. Continue this process for R', and by the DCC we will arrive at a finite direct product decomposition of R into minimal two-sided ideals:

$$R = R_1 \times R_2 \times ... \times R_m$$

Since R contains an identity, each R_i contains an identity z_i. Let L_i be an irreducible left ideal of R_i (note that it is irreducible both over R and over R_i); then the right-sided closure of L_i generates R_i, so write $z_i = s_{i,1}r_1 + s_{i,2}r_2 + ... + s_{i,n_i}r_{n_i}$ with $s_{i,j} \in L_i$, $r_j \in R_i$, and with minimal n_i. Note that each L_ir_j is a left ideal isomorphic as a left R-module to L_i. As well, if $L_ir_j \cap (L_ir_1 + ... + L_ir_{j-1} + L_ir_{j+1} + ... + L_ir_{n_i}) \neq 0$, then we could write another expression for e_i which would contradict the minimality of n_i. Therefore, $R_i \cong L_i^{n_i}$ as left R-modules.

For each $r \in R_i$, consider the left R_i-module endomorphism on R_i given by right multiplication by r. This is an embedding of R_i^{opp} (same element and addition as R_i, with multiplication given by $x \cdot y = yx$, where the latter is computed in R_i) into the ring $\text{End}_{R_i}(R_i)$. Since such endomorphisms are uniquely determined by their action on the identity, this is in fact an isomorphism. But $\text{End}_{R_i}(R_i) = \text{Hom}_{R_i}(L_i^{n_i}, L_i^{n_i}) \cong M_{n_i}(Q_i)$, where $Q_i = \text{End}_{R_i}(L_i)$ is a division ring since L_i is an irreducible module (a simple but satisfying result known as *Schur's Lemma*). There is an isomorphism between $M_n(Q_i)^{\text{opp}}$ and $M_n(Q_i^{\text{opp}})$ given by taking the transpose of a matrix. We conclude

$$R \cong M_{n_1}(\Delta_1) \times M_{n_2}(\Delta_2) \times ... \times M_{n_m}(\Delta_m)$$

where Δ_i is the division ring Q_i^{opp}.

All that remains to be shown is the uniqueness of the Δ_i and n_i. Suppose $R \cong \prod_i^m M_{n_i}(\Delta_i) \cong \prod_i^{m'} M_{n_i'}(\Delta_i')$. Then the factor matrix rings in each direct product are the complete set of minimal two-sided ideals, so it suffices to show that if $S \cong M_n(\Delta) \cong M_{n'}(\Delta')$ then $\Delta \cong \Delta'$ and $n = n'$. There is only one irreducible left-sided ideal up to isomorphism in each ring: the set of matrices with zeros everywhere outside the first column, denoted in each ring by L and L'. Since L is irreducible, an S-module endomorphism of L is uniquely determined by its action on $e_{1,1}$, the matrix with a 1 in the $(1,1)$ entry and zeros elsewher. On the other hand, since $e_{1,1}$ is idempotent in S, any endomorphism must map $e_{1,1}$ to a Δ-multiple of itself. This is enough to show that $\text{End}_S(L) \cong \Delta^{\text{opp}}$, as right scalar multiplication by Δ is an S-module endomorphism. By similar reasoning, $\text{End}_S(L') \cong \Delta'^{\text{opp}}$, and now $\Delta \cong \Delta'$. Finally, $n = n'$ since this is the number of minimal left ideals of S. □

We apply our results to the group ring FG, since Maschke's Theorem implies the condition (1).

Corollary 2. *Let G be a finite group, and let F be an algebraically closed field of characteristic not dividing the order of G. Then FG is isomorphic to the m-fold direct product of matrix rings of degree n_i over division rings Δ_i, where*

(1) The Δ_i are in fact all isomorphic to the field F.

(2) $\sum n_i^2 = |G|$.

(3) m is equal to the number of conjugacy classes of G.

(4) There are m irreducible modules/representations of FG up to isomorphism/equivalence.

Proof. (1) The center of FG in the Wedderburn decomposition is precisely the product of the rings of scalar matrices in each matrix ring; as such, the field $F \subseteq FG$ is contained in this product. For each i, by projecting onto the i^{th} coordinate if necessary, the division ring Δ_i may be regarded as an F-algebra. Since this makes Δ_i an F-subalgebra of the finite rank F-algebra FG, Δ_i is also of finite rank over F. But if F is algebraically closed, then this is only possible if $\Delta_i \cong F$.

(2) On the one hand, we have $\sum n_i^2$ is the rank of the Wedderburn decomposition of FG over F. On the other, $|G|$ is the rank of FG over F (with basis $\{g\}_{g \in G}$). Clearly, these coincide.

(3) Consider the center of FG. For each i, let z_i denote the element of the Wedderburn decomposition of FG with the identity matrix in $M_{n_i}(F)$ and zero matrices elsewhere. On the one hand, it has a basis as a vector space over F the m elements z_i. On the other hand, an element x of FG is central if and only if $gxg^{-1} = x$ for all $g\ inG$. It is as such that we see the center has for basis $\sum_{g \in C} g$, for each conjugacy class C of G. The result follows.

(4) Let M be an irreducible left FG-module with nonzero element $m \in M$. Note that FG is additively generated by the minimal left ideals from each Wedderburn component. Write them as FGe_i for orthogonal idempotents $e_i \in FG$ summing to the identity. Then $e_i m \neq 0$ for some e_i, so since M is irreducible, $FGe_i m = M$ and M is isomorphic to FGe_i. Now slightly abusing notation, let the e_i be such that the FGe_i are the m minimal left ideals that are isomorphically distinct within each Wedderburn component. Suppose there were an φ isomorphism between FGe_i and FGe_j as FG-modules for $i \neq j$. Then we would have $\varphi(e_i) = \varphi(e_i e_i) = e_i \varphi(e_i) = 0$ since e_i annihilates FGe_j, a contradiction. Thus, the FGe_i are up to isomorphism the m unique irreducible FG-modules. □

In what follows, as is most typical in the literature, we shall only concern ourselves with $F = \mathbb{C}$. In the next section, we explore the basic methods of describing and calculating the invariants of the irreducible complex representations of a finite group.

10 Character Theory

Given a representation $\varphi : G \to \mathrm{GL}_n(V)$ over $\mathbb{C}$, we define the *(group) character* of φ to be the function $\psi : G \to \mathbb{C}$ given by the trace of the representation: $\psi(g) = \mathrm{tr}\ \varphi(g)$. Note that the trace of a matrix/linear transformation is preserved under conjugation/change of basis, so a group character is invariant under equivalency of representations. As well, for the same reason, a group character is an example of a *class function*, that is, a function $\mathcal{C} \to F$ where $\mathcal{C}$ is the set of conjugacy classes of G.

Consider the characters ψ_i afforded by the m irreducible representations φ_i of G up to equivalence. Then if z_i is the identity of the i^{th} matrix ring in the Wedderburn decomposition of $\mathbb{C}G$, we see $\psi_i(z_j) = n_i$ if and only if $i = j$ and $\psi_i(z_j) = 0$ otherwise, so that the characters ψ_i are linearly independent as class functions. This is important for the following reason: Suppose we had a $\mathbb{C}G$-module V that decomposed into irreducibles W_i in two different ways:

$$V \cong a_1W_1 \oplus a_2W_2 \oplus ... \oplus a_mW_m \cong b_1W_1 \oplus b_2W_2 \oplus ... \oplus b_mW_m$$

where $nW_i = \oplus_{j=1}^{n} W_i$. Then since the character of a direct sum is just the sum of the characters of the summands, by linear independence of the irreducible characters we would have $a_i = b_i$ for all i, and the character ψ afforded by V would equal a unique integral sum of the irreducible characters. We summarize the derived results in the following proposition.

Proposition 7. *The representation of a group is uniquely determined up to equivalency by its character, and the characters of a group are given by integral linear combinations of the irreducible characters of the group.*

Already we have a modest toolbox for computing so-called *character tables*, which display the values of the irreducible characters of a group. For example, let us calculate that of S_4:

S_4	1	(1 2)	(1 2 3)	(1 2)(3 4)	(1 2 3 4)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2				
χ_4	3				
χ_5	3				

Here we have filled in two irreducible characters, that of the trivial representation (cf. Example 1), and that of the representation which multiplies a one-dimensional complex vector by the sign of the permutation. Since $\chi_i(1) = n_i$, by Corollary 2(2) we calculate the degrees of the remaining representations by solving the essentially unique integral solution to $x^2+y^2+z^2 = |S_4|-1^2-1^2 = 22$, being $(x, y, z) = (2, 3, 3)$.

Besides the trivial and regular representations of S_4, we can naturally conjure up a 4-dimensional representation of S_4 given by $\varphi(\sigma)(e_i) = e_{\sigma(i)}$ for $i = 1, 2, 3, 4$. If χ is the corresponding character, then $\chi(\sigma)$ is the number of fixed points of σ and is given below:

S_4	1	(1 2)	(1 2 3)	(1 2)(3 4)	(1 2 3 4)
χ	4	2	1	0	0

Clearly, the only 1-dimensional complex subspace that is stable under this action is that generated by $e_1 + e_2 + e_3 + e_4$, on which S_4 acts trivially. Thus, χ decomposes as the sum of the trivial character χ_1 and a degree-3 character given by $\chi - \chi_1$. We call it χ_4:

S_4	1	(1 2)	(1 2 3)	(1 2)(3 4)	(1 2 3 4)
χ_4	3	1	0	-1	-1

Suppose φ is a representation of degree n, and χ is a representation of degree 1. χ can be regarded as a multiplicative embedding into $\mathbb{C}$; thus, the product $\chi \cdot \varphi$ is also a representation acting on the same space as φ, and the submodules of the corresponding $\mathbb{C}G$-module are the same as that of φ. In particular, if φ is irreducible, then so is $\chi \cdot \varphi$. In terms of characters, we have:

Lemma 2. *If ψ is an irreducible character, and χ is a character of degree 1, then their product $\chi \cdot \psi$ is also an irreducible character.*

Thus, our other irreducible degree-3 character is given by $\chi_5 = \chi_2 \cdot \chi_4$. Finally, we note that the regular character is given by $\rho = n_1\chi_1 + n_2\chi_2 + n_3\chi_3 + n_4\chi_4 + n_5\chi_5 = \chi_1 + \chi_2 + 2\chi_3 + 3\chi_4 + 3\chi_5$; thus we solve for χ_3 and fill in the rest of the table:

S_4	1	(1 2)	(1 2 3)	(1 2)(3 4)	(1 2 3 4)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	0	-1	2	0
χ_4	3	1	0	-1	-1
χ_5	3	-1	0	-1	1

As we have seen, both constructive and nonconstructive techniques can be used to compute the character table of a given group—we were able to determine the values of χ_3 despite not constructing any explicit action of S_4 on a two-dimensional complex vector space (strictly speaking, with some effort one could recover it from the Wedderburn decomposition and knowledge of every other representation, which is what implicitly happened when we used the regular character to deduce χ_3). The scope of the tools that we *could* further explore—such as the orthogonality relations, quotient characters, and induced characters—becomes almost unbounded in complexity, but since the actual calculation of the character table for the Monster group is beyond the means of this paper, we shall content ourselves with but a taste of the methods and theory.

11 Conclusion and Further Reading

The preceding twenty-odd pages were devoted to a very elementary exploration of two paths of research aimed toward an understanding of the phenomenon of monstrous moonshine. Essentially, our material can be summarized as follows: "There is a theory of complex-analytic functions satisfying certain symmetries according to the action of the modular group Γ on $\mathbb{H}^*$, in which context we may particularize the j-invariant; independent of all this, there is a theory on the linear representation of finite groups in which context we may particularize the Monster group. Data between these two particularizations are related in unexpected ways."

We have been a bit more rigorous in our demonstration of the first theory and particularization than the second—we explicitly constructed the j-invariant and derived its data that were of interest to us, whereas we chose not to truly construct the Monster group nor derive its data. This decision was made in the interest of brevity, difficulty, and moonshine-theoretic relevance ("There is still no expanation of why the Monster exists that does not involve many pages of obscure calculations" [7]); nonetheless, those interested in a rigorous treatment of the Monster are recommended to turn to either Griess's original construction, or Conway's simplification thereof; those interested in a rigorous treatment of its data (i.e. representation theory) are recommended to study the details of the computer program utilized by Fischer, Livingstone, and Thorne [6] to calculate the character table of the Monster. Unfortunately (yet all the more intriguingly), nothing in these purely algebraic treatments gives an obvious link with the j-invariant.

There is yet more to moonhsine, however, than we have let on. In fact, neither the j-invariant nor the Monster group represent the extent of the deep connection between the theory of modular forms and the theory of the representation of finite groups.

Complex-analytic generalizations: As we know, the degrees of the irreducible representations φ_i of the Monster are no more than the corresponding characters applied to the identity: deg $\varphi_i = \chi_i(1)$. If we instead examine the character values of nonidentity elements of the Monster, we also get interesting numbers; this is the motivation behind the McKay Thompson series. It turns out that when one changes the Monster element to which we apply the group characters, it corresponds to a change in which congruence subgroup Γ' we take from Γ (thus far we have simply taken $\Gamma' = \Gamma$), of which we then take the function field (that is, the field of meromorphic functions on the Riemann surface given by the quotient of the action of Γ on $\mathbb{H}^*$, called a *modular curve*), which has a unique Hauptmodul (a normalized generating function for the function field; the j-invariant is the Hauptmodul of the modular curve $\mathbb{H}^*/\Gamma$, which is our Proposition 3 from Section 5), whose Fourier coefficients perfectly embody the change. In other words, there is an association from conjugacy classes of the Monster to certain Hauptmoduln via McKay Thompson series. Though there are 194 conjugacy classes/irreducible characters of the Monster (a surprisingly small number, considering the group's massive order), there are only 171 distinct McKay-Thompson series. [8]

Algebraic generalizations: Moonshine isn't limited to the Monster group, or at least it is no longer. There are data from the character tables of many of the "happy" simple groups (those 20 of the 26 sporadic simple groups arising as quotients of subgroups of the Monster; the rest are called pariahs) which originally suggested similar moonshine should occur for them, and methods inspired by Borcherds's landmark proof have confirmed this. [7] There is also evidence of the Rudvalis group—a pariah—being subject to similar phenomena. [5]

Lie theory: If K is the smaller of the two conjugacy classes of involutions of the Monster, then the product of any two of these involutions will lie in one of merely nine conjugacy classes, with orders 1, 2, 2, 3, 3, 4, 4, 5, 6. These correspond with the vertex labels of the E_8 Dynkin diagram. With some slight massaging, we can also obtain similar links between the Baby Monster and the F_4 diagram, and the Fischer-24 group and the G_2 diagram. [8]

The proof of the moonshine conjectures depends on several coincidences. Even the existence of the monster seems to be a fluke in any of the known constructions: these all depend on long, strange calculations that just happen to work for no obvious reason, and would not have been done if the monster had not already been suspected to exist. Then the dimension of the Leech lattice just happens to be just 2 less than the critical dimension 26 of string theory, which is just what is needed for the no-ghost theorem to be used to construct the monster Lie algebra. The monster Lie algebra just happens to have a Weyl vector, which is extremely unusual for algebras constructed like this, and means that its simple roots can be described explicitly. [7]

References

[1] M. Aschbacher, "The status of the classification of the finite simple groups", Notices of the American Mathematical Society, 2004. 51 (7). pp. 736–740

[2] J. Conway, "A simple construction for the Fischer–Griess monster group", Inventiones Mathematicae, 1985. 79 (3): 513–540.

[3] J. Conway and S. Norton, "Monstrous Moonshine", Bull. London Math, 1979. Soc. 11 (3): 308–339

[4] D. Dummit and R. Foote, Abstract Algebra, New Jersey: Wiley, 2004

[5] J. Duncan, "Moonshine for Rudvalis's sporadic group I", arXiv:math/0609449

[6] B. Fischer, D. Livingstone, M. P. Thorne, "The characters of the 'Monster' simple group", Birmingham, 1978

[7] T. Gannon, "Moonshine beyond the monster: The bridge connecting algebra, modular forms and physics", Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, Massachusetts, 2006

[8] T. Gannon, "Monstrous Moonshine: The first twenty-five years", Bull. London Math, 2006. Soc. 38 (1): 1–33.

[9] R. Griess, "The friendly giant", Inventiones Mathematicae, 1982. 69 (1): 1–102.

[10] N. Koblitz, Introduction to Elliptic Curves and Modular Forms, Springer-Verlag, 1984

[11] S. Lang, Introduction to Algebraic and Abelian Functions, Addison Wesley, 1972

[12] L. Mordell, "On Mr. Ramanujan's empirical expansions of modular functions", Proceedings of the Cambridge Philosophical Society, 1917. 19: 117–124.