Channels: Where Disciplines Meet

Volume 6
Number 2 *Spring 2022*

Article 2

May 2022

An Overview of Monstrous Moonshine

Catherine E. Riley
Cedarville University, catherineriley@cedarville.edu

Recommended Citation

Riley, Catherine E. (2022) "An Overview of Monstrous Moonshine," *Channels: Where Disciplines Meet*: Vol. 6: No. 2, Article 2.
DOI: 10.15385/jch.2022.6.2.2
Available at: https://digitalcommons.cedarville.edu/channels/vol6/iss2/2

Abstract

The Conway-Norton monstrous moonshine conjecture set off a quest to discover the connection between the Monster and the *J*-function. The goal of this paper is to give an overview of the components of the conjecture, the conjecture itself, and some of the ideas that led to its solution. Special focus is given to Klein's *J*-function.

Keywords

Monstrous moonshine, the Monster, modular functions, the J-function, representation theory, character theory, group theory

1 Introduction

In 1979, John Conway and Simon Norton published a paper conjecturing a connection between the Monster group and the J-invariant function. It was this paper that first introduced the term "moonshine", or crazy idea, in connection with this conjecture [11]. The name stuck, and the title of "monstrous moonshine" was here to stay.

Even before Griess constructed the Monster, amazing discoveries were already being discussed. These were often tied to a conjectured 196 883-dimensional irreducible representation of the Monster. When Griess constructed this representation, he added new mysteries, such as a new algebra which he used to construct the Monster.

Group theorists wanted to know why the dimensions of the irreducible representations of the Monster were tied to the coefficients of the J-function. This connection between group theory and number theory took a collection of brilliant minds to discover.

After Conway and Norton conjectured the connection, Frenkel, Lepowsky, and Meurman helped the problem by diving into the area of vertex operator algebras. Vertex operator algebras were being used in string theory, so that connection was unexpected but intriguing. Borcherds then built on the work of Frenkel, Lepowksy, and Meurman to solve the conjecture. Although the original conjecture has a solution, there are still unanswered questions related to group theory, modular forms, and physics.

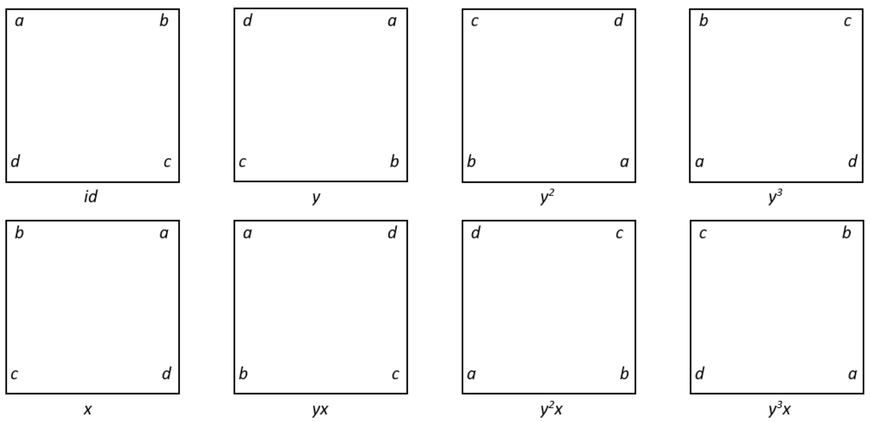
In this paper, we will work through some key elements of the monstrous moonshine conjecture and solution. We will start with an overview of group theory that will give us enough background information to understand the Monster group. After that, we will briefly discuss character theory, which will allow us to comprehend some of the basics of the argument for the monstrous moonshine conjecture. From there, we will put the Monster on hold temporarily to discuss modular functions and specifically the J-function. There we will clearly see the connection between the Monster and the J-function that intrigued so many. We will look at some of Frenkel, Lepowsky, and Meurman's work creating a vertex operator algebra, which will lead us to Borcherds' solution to the monstrous moonshine problem.

The monstrous moonshine conjecture is surprising but beautiful. The theory connects different fields in a fascinating way. Although there are still questions related to moonshine, the progress that has been made in the past forty years has been impressive. Seeing what has been accomplished should give us optimism about what has yet to be discovered.

Channels Vol. 6 No. 2 (2022): 27–50
ISSN 2474-2651

2 Overview of Group Theory

Group theory is an area of mathematics that deals with a collection of objects, called a set, along with some operation between them, called a binary operation. A set with a binary operation that together satisfy certain requirements is called a **group**. As an example of this, let us think about D_4, the group of symmetries of a square, or the ways a square can be moved without changing its appearance. D_4 has 8 elements. We define y to be a rotation by $\frac{\pi}{2}$ and x to be reflection. Hence, our eight elements are



The Elements of D_4

Composition of these functions is our binary operation. These together form a group. For a set G and a binary operation $*$, the group is denoted by $(G, *)$. If we want to perform the binary operation on two elements in the set, say $a, b \in G$, we denote that by $a * b$. Note that $a * b$ must be in G for any a and b in G. We often ignore the binary operation when denoting a group, and instead simply write the symbol for the set. In other words, instead of writing $(G, *)$ every time the group is mentioned, we will often denote the group as G. Additionally, multiplicative or additive notation is often used for the binary operation, although the operation may not be multiplication or addition.

For G to be a group, a few qualifications must be met. A key qualification is that every group must have an **identity element**, usually denoted e, such that $a * e = e * a = a$ for all a in G. Going back to our previous example with the square group, doing nothing with the square is the identity element, because if you do nothing with the square and then move it in some way it is the same as just moving it in some way without purposefully doing nothing first. This qualification is important for this discussion because it will play a role when we define simple groups.

Given a subset G' of G (such that $(G, *)$ is a group), we say that G' is a **subgroup** of G if G' is a group under $*$. If G' is a subgroup of G we denote it by $G' \leq G$.

An important concept for the discussion that will follow is the idea of a simple group. A **simple group** is a group that has no proper, nontrivial, normal subgroups [4]. By a proper subgroup, we mean a subgroup whose set is not the set of the whole group (in finite groups the set of a proper subgroup is strictly smaller in cardinality than the set of the original group). By nontrivial, we mean that the subgroup must contain more elements than just the identity element. Note that simple groups themselves must be nontrivial. In other words, the group whose set is just the identity element is not a simple group. Groups are classified as finite or infinite depending on the size of the set. A subgroup being normal is a technical condition.

We are interested in simple groups since they are the "building blocks" of larger groups. These are the groups that cannot be "factored" down into smaller groups. There are two broad categories of simple groups. They are 18 infinite families of groups and 26 sporadic groups [7].

In this paper, we will be focusing on the largest sporadic group, the **Monster** (denoted $\mathbb{M}$).

3 The Monster Group

In 1973, Fischer and Griess independently conjectured that the Monster group existed. In 1980, Griess constructed it by hand as the automorphism group of a 196884-dimensional commutative non-associative algebra [7]. This algebra is known as the Griess algebra. An automorphism is an isomorphism from a group to itself, and an isomorphism is a bijective function that has the homomorphism property. A **homomorphism** is a function, say ρ, such that $\rho(ab) = \rho(a)\rho(b)$.

As mentioned previously, the Monster is the largest of the sporadic groups. The order of the Monster (the number of elements in the Monster considered as a set) is

$$\begin{aligned} |\mathbb{M}| &= 2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \\ &= 808,017,424,794,512,875,886,459,904,961,710,757,005,754,368,000,000,000. \end{aligned}$$

The Monster is generated by involutions, which are elements of order 2 [7]. The order of an element of a group refers to the smallest possible power of the element that is the identity element. The order of any element in a group divides the order of the group.

3.1 Representation and Character Theory

We can consider any finite group as a finite subgroup of the group of $n \times n$ invertible matrices acting on V, where V is a finite-dimensional complex vector space. If we consider G in this way, we are looking at a (complex linear finite-dimensional) representation of G which is a pair (ρ, V), where $\rho : G \to GL(V)$ is a homomorphism (a function that preserves the group operation) [11]. Officially, we can define a **representation** of G over F (where we will think of F as $\mathbb{C} = \{a+bi : a, b \in \mathbb{R}, i^2 = -1\}$, the complex numbers) as a homomorphism ρ from G to $GL(n, F)$ for some n [9]. Note that $GL(n, F)$ is the general linear group of $n \times n$ invertible matrices with entries from F. We define the dimension (or degree) of the representation to be the dimension of V [11].

Let us think about the square group we looked at earlier. Here is an example of a homomorphism, ρ, from the square group to $GL(2, \mathbb{C})$:

g	$\rho(g)$	g	$\rho(g)$
id	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	x	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
y	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	yx	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$
y^2	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	y^2x	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
y^3	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	y^3x	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Note that if we multiply these matrices, they behave in the same way as their corresponding elements in D_4. For example, if we multiple the matrices corresponding to y and x we get $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$, which is the matrix corresponding to yx.

Groups can be partitioned into conjugacy classes. Given an element $a \in G$, we define the **conjugacy class** of a to be

$$cl(a) = \{b \in G | b = c^{-1}ac, c \in G\}.$$

Here is another example of a matrix representation of the square group that is conjugate to the one above:

g	$\sigma(g)$	g	$\sigma(g)$
id	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	x	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$
y	$\begin{pmatrix}i&0\\0&-i\end{pmatrix}$	yx	$\begin{pmatrix}0&i\\-i&0\end{pmatrix}$
y^2	$\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$	y^2x	$\begin{pmatrix}0&-1\\-1&0\end{pmatrix}$
y^3	$\begin{pmatrix}-i&0\\0&i\end{pmatrix}$	y^3x	$\begin{pmatrix}0&-i\\i&0\end{pmatrix}$.

The **trace** is the sum of the diagonal entries of a matrix, and is denoted by Tr. The trace of a matrix has the property that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, which implies that $\mathrm{Tr}(B^{-1}AB) = \mathrm{Tr}(ABB^{-1}) = \mathrm{Tr}(A)$, and so the trace is constant on each conjugacy class. Hence, we can tell that two representations are conjugate if they have the same trace for all elements. If ρ and σ are homomorphisms from G into $GL(n, \mathbb{C})$, they are conjugate if there exists a g in $GL(n, \mathbb{C})$ such that $g^{-1}\rho(h)g = \sigma(h)$ for all $g \in G$. If we represent a group by a homomorphism, we can define the **character** of (ρ, V), to be a function $\chi_{(\rho,V)} : G \to \mathbb{C}$ such that $g \mapsto \mathrm{Tr}(\rho(g))$. Note that this allows us to find the characters of these elements of D_4. By computing the trace of these matrices, we find that id has a character of 2, y^2 has a character of -2, and all of the other elements have a character of 0.

g	$\rho(g)$	$\sigma(g)$	$\mathrm{Tr}(\rho(g))) = \mathrm{Tr}(\sigma(g))) = \chi(g)$
id	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	2
y	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$	0
y^2	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	-2
y^3	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$	0
x	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	0
yx	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$	0
y^2x	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$	0
y^3x	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$	0

Comparison of ρ and σ

Note that the characters are the same for both representations, as expected.

Basically, when we are looking at representations of a group, we can think of it as looking at a homomorphism from a group into $GL(n, \mathbb{C})$. Alternatively, we can think of representations as FG-modules. Note that F is a field, G is a group, and an FG-module is similar to a vector space. Since these ways of thinking of representations are equivalent, we will present a theorem that refers to modules and then apply it to homomorphisms.

Maschke's Theorem: "Let G be a finite group, let F be $\mathbb{R}$ or $\mathbb{C}$, and let V be an FG-module. If U is an FG-submodule of V, then there is an FG-submodule W of V such that

$$V = U \oplus W\text{"} \quad [9].$$

Alternatively, "Let G be a finite group and F a field whose characteristic does not divide $|G|$. Then every $F[G]$-module is completely reducible" [8]. This theorem tells us that we can break modules down into a direct sum of submodules. In other words, if we take any one of these (complex linear finite-dimensional) modules, it will be isomorphic to a direct sum of irreducible submodules. The number of these irreducible submodules for a given group is equal to the number of conjugacy classes in the group [11].

Since representations can be thought of as modules or homomorphisms, Maschke's Theorem applies to representations thought of as homomorphisms. Modules can be broken down until we reach irreducible submodules; in other words, until we get to a direct sum of irreducible representations. Similarly, we can break a homomorphism down into a direct sum of multiple irreducible homomorphisms. Concretely, since our homomorphism sends the elements of the group to matrices, this looks like breaking down a matrix into smaller matrices along the diagonal.

As an example, let us think about A_5, the group of even permutations of a five element set. Suppose we wanted to represent A_5 in 18 dimensions and hence each element of A_5 maps to an 18 by 18 matrix. An example of what such a matrix would look like is

$$\left(\begin{array}{cccccccccccccccccc}
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare
\end{array}\right),$$

where the black squares represent elements in $\mathbb{C}$. Given that A_5 has irreducible representations of dimension 5, dimension 4, two different representations of dimension 3, and one of dimension 1, this representation can be broken down into irreducibles, perhaps a 5 by 5, two 4 by 4, a 3 by 3, and two 1 by 1 matrices (note that $5+4+4+3+1+1=18$). Hence, our irreducible matrix would look like this

$$\left(\begin{array}{cccccccccccccccccc}
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare
\end{array}\right).$$

These two representations could be conjugate to each other, since the original matrix is conjugate to some block diagonal sum of irreducible representations. We could also create a different irreducible representation, as long as we use the dimensions of irreducible representations listed above. For example, if we wanted to create a representation with one block of dimension 4, one block of dimension 3, two blocks of the other representation of dimension 3, and five blocks of dimension 1, we would get

$$\begin{pmatrix}
\blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare
\end{pmatrix},$$

where the first representation of dimension 3 is shown by the black squares and the other representation is shown by the asterisks. We could even have three blocks of dimension 5 and three blocks of dimension 1, which gives us

$$\begin{pmatrix}
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare
\end{pmatrix}.$$

Note that our original matrix would not be conjugate to all of these irreducible representations, but these are examples of what it might be conjugate to. Hence, we see how a representation of a group can be broken into irreducible representations and how the dimensions of those irreducible representations work.

The Monster has 194 conjugacy classes, and hence it has 194 irreducible representations. In 1978, Fischer, Livingstone, and Thorne determined the dimensions of these representations. Starting with the trivial representations, the dimensions of the first few irreducible representations are

$$(r_n)_{n=1,\cdots,194} = (1, 196\,883, 21\,296\,876, 842\,609\,326, 18\,538\,750\,076, \cdots).$$

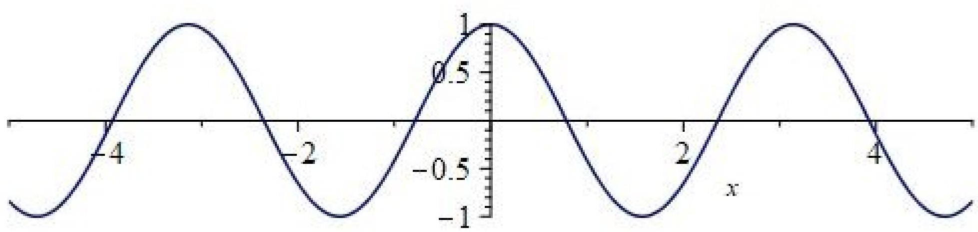
The rest of the sequence can be found in the ATLAS of Finite Groups [3]. We will now temporally leave the Monster, but this sequence of numbers will be important in the discussion that follows.

4 The J-Function

We are going to work our way to the J-function, which is a modular function that connects number theory to the Monster. Modular functions and the closely related but slightly more complicated modular forms come up in number theory in a variety of ways. It was the theory of modular forms that allowed Andrew Wiles to prove Fermat's Last Theorem. To truly understand the J-function we need to lay quite a bit of groundwork first.

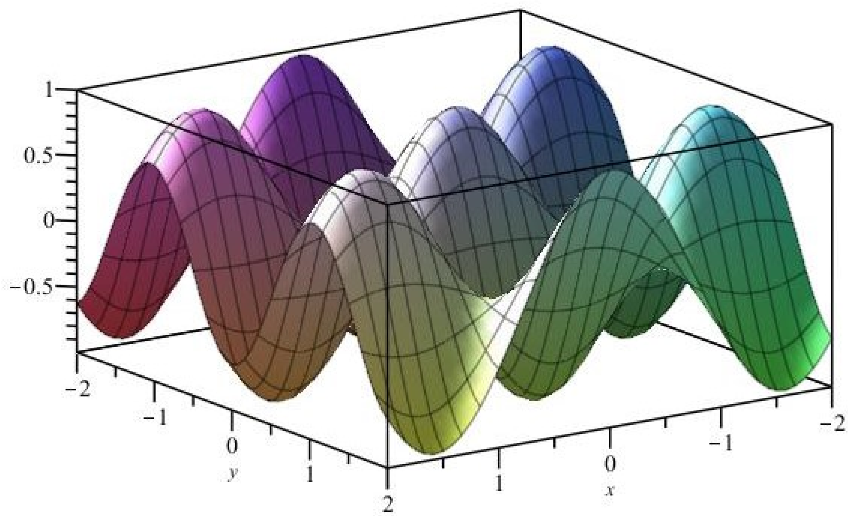
4.1 Preliminary Notation and Functions

We are going to need to understand elliptic functions. A function f is **elliptic** if it is doubly periodic and meromorphic (which means it is complex differentiable on every point of an open subset of the complex plane except for a set of isolated points called poles). **Poles** are isolated singular points [2]. A good way to recognize poles is where the denominator of the function is zero. A pole of order n in a rational function is a zero of the denominator of multiplicity n (which is not a zero of the numerator). A function is **periodic** if there is some period ω such that $f(z+\omega) = f(z)$, for z and $z+\omega$ in the domain of f. An example of a periodic function is $\cos 2x$, which has a period of π.



Graph of $y = \cos 2x$

If ω is a period of a function, so is $n\omega$ for all $n \in \mathbb{Z}$. For a function to be **doubly periodic** it must have two periods, ω_1 and ω_2 in $\mathbb{C}$, with ratio ω_2/ω_1, where this ratio is not real. Hence, the periods will go in two different directions. An example of a function that has two periods in two different directions is $\cos 2x \cdot \cos 3y$. It is not technically doubly periodic since the ratio of the periods is real, but it does have periods in two different directions. Doubly periodic functions can only be seen in four dimensions.



Graph of $f(x,y) = \cos 2x \cdot \cos 3y$

A pair of periods (ω_1, ω_2) for a function f is a **fundamental pair** if all the periods of f can be written in the form $m\omega_1 + n\omega_2$ for $m, n \in \mathbb{Z}$. We can create a parallelogram by considering $0, \omega_1, \omega_2$, and $\omega_1 + \omega_2$.

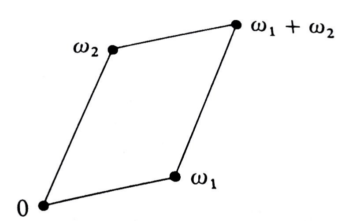
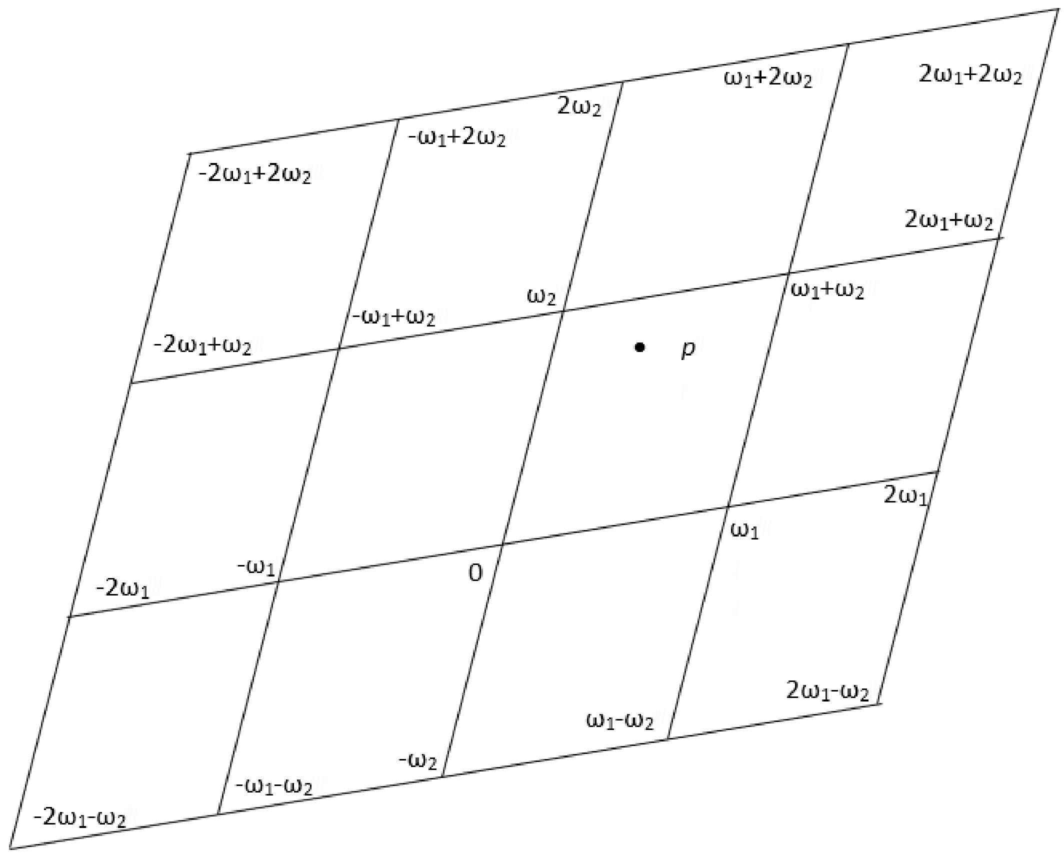


Figure 1.2.a from [1]

This parallelogram is important because it is part of a larger lattice.



More of the Lattice

The lattice extends indefinitely in all directions. If we understand what is happening in one parallelogram, for instance, where it has poles, we will know what is happening in all the other parallelograms in the lattice. So if this main parallelogram has a pole at p, all the other parallelograms also have a pole at the same position. Note that the lattice itself is only made up of the vertices of the parallelograms, but sometimes it is helpful to think about what is happening in the interior of the parallelograms (in other words, the points $a\omega_1 + b\omega_2$ where neither a nor b is in $\mathbb{Z}$).

The number of zeros of f over this period parallelogram is the same as the number of poles (each counted with multiplicity, which means that if we have multiple poles at the same point we count them separately) and is called the order of the function. We can show this is true by considering a cell, which is a period parallelogram shifted so that it has no poles or zeros on its boundary. If we take the integral

$$\frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}dx$$

around the boundary C of the cell, it will give us the difference between the number of zeros and poles inside the cell. But f'/f is elliptic and has the same period as f, and the contour integral of an elliptic function around the boundary of any cell is zero, since the integrals along parallel edges cancel because of periodicity. So the difference between the number of zeros and the number of poles is 0, and hence the number of poles and the number of zeros for a given elliptic function are the same [1].

We now turn to the task of constructing a nonconstant elliptic function. The minimum order of such a function is 2, and so we start by constructing an elliptic function of that order. We will prescribe the periods and find an elliptic function of that order. We will start with a function of order 3 and then integrate it to get to a function of order 2.

If we consider the series $f(z) = \sum_{\omega \in \Omega} \frac{1}{(z-\omega)^3}$ (where $\Omega = \{\omega = m\omega_1 + n\omega_2 | m, n \in \mathbb{Z}\}$ is a lattice), we have an elliptic function with two periods, ω_1 and ω_2, and a pole of order 3 at each ω. We will use this function to construct an elliptic function of order 2 called the **Weierstrass $\wp$ function**. We do this by integrating $f(z)$ term by term. To do this, we see that the integral of $(z-\omega)^{-3}$ is $-(z-\omega)^{-2}/2$, and hence we will have a summand of $-(z-\omega)^{-2}/2$ near each period. We can then multiply by -2 to simplify. Additionally, we can also consider starting our integration at the origin ($\omega = 0$), and hence we will have an initial term of $\frac{1}{z^2}$. Basically, we can consider this as splitting the sum into two parts, $\omega = 0$ and $\omega \neq 0$, and integrating those two parts separately. Hence,

$$\wp(z) = \frac{1}{z^2} + \int_0^z \sum_{\omega \neq 0} \frac{-2}{(t-\omega)^3} dt = \frac{1}{z^2} + \sum_{\omega \neq 0} \left[\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right].$$

Note that $\wp'(z) = -2 \sum_{\omega \in \Omega} \frac{1}{(z-\omega)^3}$, which is -2 times our original function.

The Weierstrass function is incredibly important in the area of elliptic functions. It turns out that given some lattice, any elliptic function for that lattice is a rational function in $\wp$ and $\wp'$ relative to that lattice [10]. This means that if you have the $\wp$ function for some lattice, you can get whatever other elliptic function you want in terms of $\wp$ and $\wp'$.

In the discussion that follows, we are going to refer to the series $G_n = \sum_{\omega \neq 0} \frac{1}{\omega^n}$, which is called the **Eisenstein series of order** n. This series is used in the differential equation $[\wp'(z)]^2 = 4\wp^3(z) - 60G_4\wp(z) - 140G_6$, which the Weierstrass $\wp$ function satisfies since we can form a linear combination of the powers of $\wp$ and $\wp'$ which eliminates the pole at $z = 0$. Therefore, we can get a constant elliptic function, since it will have no poles and an elliptic function without poles is constant. We can verify that $\wp$ satisfies this differential equation by simply plugging the relevant functions into this differential equation. Near $z = 0$

$$\wp'(z) = -\frac{2}{z^3} + 6G_4 z + 20G_6 z^3 + \cdots ,$$

which is an elliptic function of order 3. If we square it, we get an elliptic function of order 6 since

$$[\wp'(z)]^2 = \frac{4}{z^6} - \frac{24G_4}{z^2} - 80G_6 + \cdots ,$$

where $+\cdots$ indicates a power series in z which vanishes at $z = 0$. If we cube $\wp$ and multiple it by 4 we get

$$4\wp^3(z) = \frac{4}{z^6} + \frac{36G_4}{z^2} + 60G_6 + \cdots .$$

Therefore

$$[\wp'(z)]^2 - 4\wp^3(z) = -\frac{60G_4}{z^2} - 140G_6 + \cdots .$$

Hence

$$[\wp'(z)]^2 - 4\wp^3(z) + 60G_4\wp(z) = -140G_6 + \cdots .$$

The left hand side of the equation does not have a pole at $z = 0$ and hence has no poles in the period parallelogram. Therefore, it must be constant. That constant must be $-140G_6$, and hence $[\wp'(z)]^2 = 4\wp^3(z) - 60G_4\wp(z) - 140G_6$. We define **invariants** $g_2 = 60G_4$ and $g_3 = 140G_6$, which simplifies the differential equation for $\wp$ [1].

One last notation we need to cover before we get to the J-function is the **discriminant**, defined by $\Delta = g_2^3 - 27g_3^2$. This function, along with g_2, is what we use to define the J-function. Note that this is not the only way to define Δ. If we look at the Euler function $\phi(q) = \prod_{n=1}^{\infty}(1 - q^n)$, we see that if we multiply it by $q^{\frac{1}{24}}$ we get the function $\eta(q)$. If we take $(\eta(q))^{24}$ we get a function that many people call Δ. This is not quite the same as our Δ, but if we take $(\eta(q))^{24}(2\pi)^{12}$ we get our Δ function. Although this may seem like a more natural way to define Δ, our more artificial construction of Δ will help us as we define the J-function.

Since $\omega = m\omega_1 + n\omega_2$, we can consider g_2, g_3, and Δ as functions of the two periods ω_1 and ω_2, and hence, $g_2 = g_2(\omega_1, \omega_2)$, $g_3 = g_3(\omega_1, \omega_2)$ and $\Delta = \Delta(\omega_1, \omega_2)$. By the Eisenstein series we see that g_2 has degree -4 (since $g_2 = 60\sum_{\omega\neq 0}(\omega)^{-4}$) and g_3 has degree -6 (since $g_3 = 140\sum_{\omega\neq 0}(\omega)^{-6}$). Both these functions are homogeneous. A function f is **homogeneous** if $f(cx) = c^n f(x)$ for some n. So

$$\begin{aligned} g_2(c\omega_1, c\omega_2) &= 60 \sum_{m\omega_1+n\omega_2\neq 0} \frac{1}{(cm\omega_1 + cn\omega_2)^4} \\ &= 60 \sum_{m\omega_1+n\omega_2\neq 0} \frac{1}{c^4(m\omega_1 + n\omega_2)^4} \\ &= \frac{60}{c^4} \sum_{m\omega_1+n\omega_2\neq 0} \frac{1}{(m\omega_1 + n\omega_2)^4} \\ &= c^{-4} g_2(\omega_1, \omega_2). \end{aligned}$$

Similarly, $g_3(c\omega_1, c\omega_2) = c^{-6} g_3(\omega_1, \omega_2)$ and $\Delta(c\omega_1, c\omega_2) = c^{-12}\Delta(\omega_1, \omega_2)$.

If we take $\tau = \omega_2/\omega_1$ and factor an ω_1^{-1} out of these functions, we see that

$$\begin{aligned} g_2(1, \tau) &= 60 \sum_{m+n\tau\neq 0} \frac{1}{(m + n\tau)^4} \\ &= 60 \sum_{m+n\frac{\omega_2}{\omega_1}\neq 0} \frac{1}{(m + n\frac{\omega_2}{\omega_1})^4} \\ &= 60 \sum_{m\omega_1+n\omega_2\neq 0} \frac{1}{\omega_1^{-4}(m\omega_1 + n\omega_2)^4} \\ &= 60 \sum_{m\omega_1+n\omega_2\neq 0} \frac{\omega_1^4}{(m\omega_1 + n\omega_2)^4} \\ &= \omega_1^4 g_2(\omega_1, \omega_2). \end{aligned}$$

Similarly, $g_3(1, \tau) = \omega_1^6 g_3(\omega_1, \omega_2)$ and $\Delta(1, \tau) = \omega_1^{12}\Delta(\omega_1, \omega_2)$. If $\tau \in \mathbb{H}$, we denote these three functions by $g_2(\tau)$, $g_3(\tau)$, and $\Delta(\tau)$, respectively, and $\Delta(\tau) = (g_2(\tau))^3 - 27(g_3(\tau))^2 \neq 0$ [1].

4.2 Definition of the J-Function

As we are about to define the J-function, let us survey for a moment the previous discussion. In trying to build a nonconstant elliptic function of order 2, we created the Weierstrass function. From there, we defined a relationship between Eisenstein series and the coefficients of a differential equation that $\wp(z)$ satisfies. This relationship gave us g_2 and g_3 which we then used to define Δ. We defined all three of those functions in terms of τ.

We now define Klein's function. This function is a quotient of g_2 and Δ. A function of the periods ω_1 and ω_2, it is homogeneous of degree 0. This means that we can multiply ω_1 and ω_2 by a constant and not change the value of the function. For $\omega_2/\omega_1 \notin \mathbb{R}$, we define

Klein's function as

$$J(\omega_1,\omega_2)=\frac{g_2^3(\omega_1,\omega_2)}{\Delta(\omega_1,\omega_2)}=\frac{\lambda^{-12}g_2^3(\omega_1,\omega_2)}{\lambda^{-12}\Delta(\omega_1,\omega_2)}=\frac{g_2^3(\lambda\omega_1,\lambda\omega_2)}{\Delta(\lambda\omega_1,\lambda\omega_2)}=J(\lambda\omega_1,\lambda\omega_2)$$

for $\lambda\neq 0$. This equality is true since both g_2^3 and Δ are homogeneous of degree -12. Because of this fact, for $\tau\in\mathbb{H}$ we see that $J(1,\tau)=J(\omega_1,\omega_2)$. We will denote this function simply as $J(\tau)$ [1].

If we define $\omega_2'=a\omega_2+b\omega_1$ and $\omega_1'=c\omega_2+d\omega_1$, where $a,b,c,d\in\mathbb{Z}$ such that $ad-bc=1$, this new pair of periods generates the same lattice Ω. Note that

$$\tau'=\frac{\omega_2'}{\omega_1'}=\frac{a\omega_2+b\omega_1}{c\omega_2+d\omega_1}=\frac{a\omega_2+b\omega_1}{c\omega_2+d\omega_1}\cdot\frac{\frac{1}{\omega_1}}{\frac{1}{\omega_1}}=\frac{a\frac{\omega_2}{\omega_1}+b}{c\frac{\omega_2}{\omega_1}+d}=\frac{a\tau+b}{c\tau+d}.$$

Functions of the form $f(\tau)=\frac{a\tau+b}{c\tau+d}$ where a,b,c,d are integers and $ad-bc=1$ are called **unimodular transformations**. The set of all these transformations form a group under composition called the modular group. Hence, "$J(\tau)$ is invariant under the transformations of the modular group" [1]. In other words, $J\left(\frac{a\tau+b}{c\tau+d}\right)=J(\tau)$. One instance of this is $\tau'=\tau+1$, which means that $J(\tau)$ is a periodic function with period of 1, since both τ and τ' generate the same lattice. Let us pause our discussion of the J-function temporarily to show that $SL(2,\mathbb{Z})$ behaves in this way. This is important for our focus on monstrous moonshine.

4.3 Modular Functions in General

To begin, we will consider the group $SL(2,\mathbb{R})$ of 2×2 matrices with entries from $\mathbb{R}$ with determinant 1. We will let $SL(2,\mathbb{R})$ act on the upper half plane $\mathbb{H}=\{\tau\in\mathbb{C}|\text{Im}(\tau)>0\}$ by

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\cdot\tau=\frac{a\tau+b}{c\tau+d}\qquad [7].$$

Technically we are dealing with $SL(2,\mathbb{R})/\{\pm I\}$ on $\mathbb{H}$. In other words, the quotient group of $SL(2,\mathbb{R})$ modded out by $\{\pm I\}$, since

$$\begin{pmatrix}\pm 1 & 0\\ 0 & \pm 1\end{pmatrix}\cdot\tau=\frac{\pm\tau+0}{0\cdot\tau\pm 1}=\tau$$

and hence I and $-I$ act as the identity element. For the purposes of this paper we will think of it in terms of $SL(2,\mathbb{R})$ [7].

If we let G be a discrete subgroup of $SL(2,\mathbb{R})$, we can regard $G\setminus\mathbb{H}$ (a quotient topology where all points in one equivalence class get mapped to one point) as a complex curve. We can think of $\mathbb{H}$ as a topological space. Since we have a quotient topology, this curve inherits some complex structure from $\mathbb{H}$. We define the genus of G as the genus of $G \setminus \mathbb{H}$. The group $G = SL(2, \mathbb{Z})$ can be interpreted as the modular group of the torus. Of the groups we are interested in, most of them are commensurable with $SL(2, \mathbb{Z})$. This means that $G \cap SL(2, \mathbb{Z})$ has finite index in both G and $SL(2, \mathbb{Z})$. The index of a subgroup G' of a group G is defined as $\frac{|G|}{|G'|}$, which is always an integer by Lagrange's Theorem. One family of subgroups of $SL(2, \mathbb{Z})$ we are interested in is

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) | N \text{ divides } c \right\}$$

where N will determine the genus [7].

This leads us to two definitions, from the article "*Monstrous Moonshine: The First Twenty-Five Years*", that will be influential in our discussion of moonshine.

Definition 1: "We call a discrete subgroup G of $SL(2, \mathbb{R})$ a *moonshine-type modular group* if it contains some $\Gamma_0(N)$, and also obeys the condition that

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \in G \text{ if and only if } t \in \mathbb{Z}\text{"} \quad [7].$$

A few notes about moonshine modular groups: first, they are commensurable with $SL(2, \mathbb{Z})$. Second, if we take any meromorphic function $f : G\backslash\mathbb{H} \to \mathbb{C}$, where G is a moonshine modular group, it will have a Fourier expansion. We will discuss Fourier expansions later. For now, we simply note that this expansion will be of the form $f(\tau) = \sum_{n=-\infty}^{\infty} a_n q^n$, where $a_n \in \mathbb{C}$ and $q = e^{2\pi i \tau}$ [7]. A **meromorphic function** is a function that is complex differentiable on all but a discrete subset of its domain, and those points must be poles. The second definition we are interested in points us to the J-function.

Definition 2: "Let G be any subgroup of $SL(2, \mathbb{R})$ commensurable with $SL(2, \mathbb{Z})$. By a *modular function* f for G we mean a meromorphic function $f : \mathbb{H} \to \mathbb{C}$, such that

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = f(\tau) \qquad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$$

and such that, for any $A \in SL(2, \mathbb{Z})$, the function $f(A.\tau)$ has Fourier expansion of the form $\sum_{n=-\infty}^{\infty} b_n q^{n/N}$ for some N and b_n (both depending on A), and where $b_n = 0$ for all but finitely many negative n" [7].

In this definition, $b_n \in \mathbb{C}$ and $N \in \mathbb{Z}^+$. Note that this definition limits how many terms with negative exponents we can have, which is exactly the case with the J-function.

4.4 Fourier Expansion of $J(\tau)$

We will now follow the proof presented in *Modular Functions and Dirichlet Series in Number Theory* [1] to prove that $J(\tau)$ has a Fourier expansion. Note that a Fourier expansion is similar to a Taylor series, except that it is written in terms of sines and cosines. We can use Euler's formula to represent this in terms of exponentials, since $e^{ix} = \cos x + i \sin x$.

Theorem: If $\tau \in \mathbb{H}$, $J(\tau)$ can be represented by an absolutely convergent Fourier series

$$J(\tau) = \sum_{n=-\infty}^{\infty} a(n) e^{2\pi i n \tau}.$$

Proof: Let $x = e^{2\pi i \tau}$. Then $\mathbb{H}$ maps into the punctured unit disk $D = \{x \mid 0 < |x| < 1\}$.

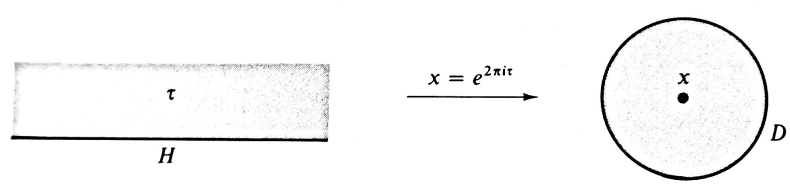


Figure 1.5

Figure 1.5 from [1]

Each τ maps onto a unique point x in D, but x is the image of infinitely many points in $\mathbb{H}$. If τ and τ' both map onto x, this means that $e^{2\pi i \tau} = e^{2\pi i \tau'}$. This means that τ and τ' differ by an integer, since by Euler's formula we have $\cos(2\pi\tau) + i\sin(2\pi\tau) = \cos(2\pi\tau') + i\sin(2\pi\tau')$, which is true if τ and τ' differ by an integer as $\cos(x)$ and $\sin(x)$ both have a period of 2π. For $x \in D$, let $f(x) = J(\tau)$, where τ is any point that maps to x (this function is well defined since J is periodic with period 1). We see that f is analytic in D because

$$f'(x) = \frac{d}{dx} J(\tau) = \frac{d}{d\tau} J(\tau) \frac{d\tau}{dx} = (J'(\tau)) / \left(\frac{dx}{d\tau}\right) = \frac{J'(\tau)}{2\pi i e^{2\pi i \tau}},$$

and $J(\tau)$ is analytic. Therefore, $f'(x)$ exists at every x. Hence, f has a Laurent expansion about 0,

$$f(x) = \sum_{n=-\infty}^{\infty} a(n) x^n,$$

which is absolutely convergent for each $x \in D$. Substituting $e^{2\pi i\tau}$ for x we see that the theorem is proved [1].

QED

A **Laurent expansion** for a function g can be defined as

$$g(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} \quad [2].$$

Since we are dealing with functions with poles of finite order, there cannot be more terms with a negative degree than the order of the poles. So most of the b_n terms are zero in this expansion.

So $J(\tau)$ has a Fourier expansion, but we must now show that it has integer coefficients. It can be shown that for $\tau \in \mathbb{H}$ we have the Fourier expansions

$$g_2(\tau) = \frac{4\pi^4}{3}\left[1 + 240\sum_{k=1}^{\infty} \sigma_3(k)e^{2\pi ik\tau}\right]$$

and

$$\Delta(\tau) = (2\pi)^{12}\sum_{n=1}^{\infty} \tau(n)e^{2\pi in\tau},$$

where $\sigma_3(k) = \sum_{d|k} d^3$ and $\tau(n)$ is Ramanujan's tau function, where the coefficients are integers and $\tau(1) = 1$ and $\tau(2) = -24$. In this definition of Δ it is slightly more clear that $\Delta = (2\pi)^{12}(\eta)^{24}$. Using these formulas we can find the Fourier expansion for $J(\tau)$. Note that both of these formulas have integer coefficients if we ignore the factors with powers of π. Again, we will follow a proof from [1].

Theorem: If $\tau \in \mathbb{H}$, we have the Fourier expansion

$$12^3 J(\tau) = e^{-2\pi i\tau} + 744 + \sum_{n=1}^{\infty} c(n)e^{2\pi in\tau},$$

where the $c(n)$ are integers.

Proof: In this proof we will write I for any power series in x with integer coefficients. Hence, letting $x = e^{2\pi i\tau}$ as before, we have

$$g_2^3(\tau) = \frac{64\pi^{12}}{27}(1 + 240x + I)^3 = \frac{64\pi^{12}}{27}(1 + 720x + I)$$

(where the coefficient on x comes from the multinomial theorem), and

$$\Delta(\tau) = \frac{64\pi^{12}}{27}[12^3x(1 - 24x + I)].$$

This means that
$$J(\tau) = \frac{g_2^3(\tau)}{\Delta(\tau)} = \frac{1+720x+I}{12^3 x(1-24x+I)} = \frac{1+720x+I}{12^3 x(1-24x+I)} \cdot \frac{1+24x+I}{1+24x+I} = \frac{1}{12^3 x}(1+720x+I)(1+24x+I),$$
which implies that
$$12^3 J(\tau) = \frac{1}{x} + 744 + \sum_{n=1}^{\infty} c(n)x^n.$$
In this formula, the $c(n)$ are integers [1].

QED

Thus, we have our expansion of the J-function. The coefficients have been calculated for $n \leq 100$. The first few are

$c(0) = 744$,

$c(1) = 196\,884$,

$c(2) = 21\,493\,760$, and

$c(3) = 864\,299\,970$.

In some versions of this formula, the 744 is subtracted from both sides, and the result is called $J(\tau)$. Letting $q = e^{2\pi i \tau}$, that gives us $12^3 J(\tau) - 744 = q^{-1} + 196\,884q + 21\,493\,760q^2 + 864\,299\,970q^3 + \cdots$. From now on we will refer to this version of the formula as $J(\tau)$.

Going back to modular functions in general for a moment, for a group G, we can expand its unique modular function $J_G(\tau)$ as
$$J_G(\tau) = q^{-1} + \sum_{n=1}^{\infty} a_n q^n.$$
"This function J_G is called the (normalised) *Hauptmodul* for the genus 0 group G" [7]. If we let $G = SL(2, \mathbb{Z})$ we get $J(\tau) = q^{-1} + 196\,884q + 21\,493\,760q^2 + 864\,299\,970q^3 + \cdots$ [7].

At this point, we return to the Monster. Recall that the first few dimensions of irreducible representations of the Monster are
$$(r_n)_{n=1,\cdots,194} = (1, 196\,883, 21\,296\,876, 842\,609\,326, 18\,538\,750\,076, \cdots).$$

We now notice that
$$\begin{aligned} 196\,884 &= 1 + 196\,883 \\ 21\,493\,760 &= 1 + 196\,883 + 21\,296\,876 \\ 864\,299\,970 &= 2 \cdot 1 + 2 \cdot 196\,883 + 21\,296\,876 + 842\,609\,326, \end{aligned}$$
where the numbers on the left hand side of the equation are coefficients from $J(\tau)$ and the numbers on the right are from the dimensions of the irreducible representations of the Monster. So we see that there is a connection between modular functions and a sporadic finite simple group [7].

5 Monstrous Moonshine

Based on the connection between the coefficients of the J-function and the dimensions of the irreducible representations of the Monster, John Thompson made the following conjecture:

Conjecture 1: "There is a (somehow) natural infinite-dimensional graded representation of the monster group $(\rho_\natural, V^\natural = \oplus_{i\geq -1} V_i^\natural)$, such that each graded part $V_i^\natural$ is finite dimensional, and such that:

$$\tilde{J}(\tau) = \sum_{i\geq -1} \dim(V_i^\natural) q^i.$$

Equivalently, the elements of $\mathbb{M}$ act naturally as (infinite) matrices on an infinite dimensional graded vector space (these matrices are block-diagonal with every block of finite size), and the graded-dimension of $V^\natural$ is the q-expansion of the normalized J-function" [11]. Note that in this definition, $\tilde{J}(\tau)$ is the same as what we have been referring to as $J(\tau)$, namely, the normalized J-function.

This conjecture is quite intense, so some explanation is in order. Similar to the representations we were considering earlier, this conjecture says that there is an infinite-dimensional representation of the Monster where $\rho_\natural$ is our homomorphism and $V^\natural$ is our vector space. The vector space we are considering is graded, which means that the vector space is a direct sum of subspaces which are indexed by some set. A direct sum is similar to a direct product except that all but finitely many terms must be 0. Direct sums are used in commutative groups where we can consider our binary operation using additive notation [4]. The notation for a direct sum is $\oplus$. In this conjecture, each of the subspaces has finite dimension.

Thompson's conjecture says that there is an infinite dimensional graded module for the Monster whose graded submodule corresponding to i has dimension the cofficient of q^i in the J-function. The conjecture can also be thought of in terms of matrices. If we consider the elements of $\mathbb{M}$ as matrices that act on an infinite dimensional graded vector space, the graded dimension is the q-expansion of $J(\tau)$.

If we look at representations of groups, we are simply looking at the trace of $\rho(e)$, the identity element. For this reason, Thompson suggested that the series

$$T_{[g]} = \sum_{i\geq -1} \mathrm{Tr}(\rho_{\mathbb{M}}(g)_{|V_i^\natural}) q^i = \frac{1}{q} + \sum_{n=0}^{\infty} H_n([g]) q^n$$

should be studied. These series (which are now called McKay-Thompson series) exist for each conjugacy class. So there is a separate formula for each conjugacy class in $\mathbb{M}$.

Using the character table of the Monster, in 1979 Conway and Norton added to Thompson's conjecture. This new conjecture, Conway-Norton's Monstrous Moonshine, can be stated as follows:

Conjecture: Conway-Norton's Monstrous Moonshine: "There is a (somehow) natural infinite-dimensional graded representation $(\rho_{\mathbb{M}}, V^\natural = \oplus V_i^\natural)$ of the monster group, with finite dimensional graded parts $V_i^\natural$, and such that for each conjugacy class $[g]$ of the monster, the series $T_{[g]}$ is the q-expansion of the normalised Hauptmodul of some subgroup $\Gamma_{[g]}$ of $PSL(2, \mathbb{R})$ commensurable with $PSL(2, \mathbb{Z})$" [11].

This definition leads back to our previous discussion of modular functions. Conway and Norton also computed $\mathrm{Tr}(\rho_{\mathbb{M}}(g)_{|V_i^\natural})$ for all $g \in \mathbb{M}$ and $i \leq 10$, as well as verified that the functions from G to $\mathbb{C}$ defined by $g \mapsto \mathrm{Tr}(\rho_{\mathbb{M}}(g)_{|V_i^\natural})$ are characters of some representations of $\mathbb{M}$ [11].

Using an IBM 370/158 computer, Atkin, Fong, and Smith showed in 1979 that the H_n class function are truly characters of $\mathbb{M}$. They did this by reducing the infinite number of verifications down to a finite one by using results like Brauer's characterization of characters [11].

Frenkel, Lepowsky, and Meurman worked to nail down precisely what was happening between the J-function and the Monster. They built a vertex operator algebra that led to the Monster by modifying the definition of vertex algebra given by Borcherds. This ended up being related to constructions in string theory. As Frenkel, Lepowsky, and Meurman say in their book, *Vertex Operator Algebras and the Monster* [6] (which was a follow-up to their ground-breaking article, "*A Natural Representation of the Fischer-Griess Monster with the Modular Function J as Character*" [5]),

> Combining this result [that $V^\natural$ has the structure of a vertex operator algebra] with the rest of our earlier work, including triality, we are able to prove the following:
>
> (A) The $\mathbb{Z}$-graded space $V^\natural$ carries an explicitly defined vertex operator algebra structure with graded dimension $J(\tau)$ and rank 24, and which acts irreducibly on itself.
>
> (B) The Monster acts faithfully and homogeneously on the $\mathbb{Z}$-graded space $V^\natural$, preserving the vertex operator algebra structure as in (25) (with $v \in V^\natural$) and fixing the elements $\mathbf{1}$ and ω [6].

Note that equation (25) mentioned in the quote above is $gY(v,z)g^{-1} = Y(g \cdot v, z)$, for $g \in \mathbb{M}$, where z can be interpreted as a nonzero complex variable, and $Y(v,z)$ is a vertex operator.

We are now prepared to unpack this quote. Triality is a relationship between three vector spaces. The authors had discussed triality in their previous paper (for more information on that paper see [5] in the References page). If we look at $V^\natural$, we find it was built as a vertex operator algebra. This structure has a graded dimension of $J(\tau)$ and rank 24. To say that $V^\natural$ acts irreducibly on itself ties into our discussion earlier of irreducible representations. Algebras are groups, and hence there is a way for them to act on themselves. We send an element of the algebra to an endomorphism of the algebra. Some of those actions are irreducible representations.

The second part of the theorem says that the Monster acts on $V^\natural$, and preserves the algebra structure. To say that the Monster acts faithfully on $V^\natural$ means that the identity element of the Monster is the only element that leaves every element in $V^\natural$ fixed [4]. That the Monster acts homogeneously means that the image of any homogeneous submodule is completely contained in another homogeneous submodule.

Frenkel, Lepowsky, and Meurman also defined the Monster as "the group M of linear automorphisms of the moonshine module $V^\natural$ generated by C and σ" [6]. The group C was a group defined to act naturally on $V^\natural$. It was "the centralizer of an involution in the Monster" [6]. The symbol σ represents an automorphism of $V^\natural$.

Finally, in 1992, Borcherds used Frenkel, Lepowsky, and Meurman's vertex operator algebra and a result from string theory to construct an infinite dimensional Lie algebra, which he called the monster Lie algebra. By building a theory of mathematical objects called generalized Kac-Moody algebras and deriving some identities for those (the denominator and twisted denominator identities), he was able to show that these identities apply to the monster Lie algebra and the coefficients of the McKay-Thompson series. Because of that, the proof of Conway and Norton's conjecture could be done using only a finite number of verifications. This was helped by the fact that Conway and Norton had already done some verifications. Hence, Borcherds proved the monstrous moonshine correspondence [11].

In this way, the connection between the J-function and the Monster was made clear. Vertex operator algebras now play a role in string theory, a type of theoretical physics [11]. For two fields that may seem so disjointed at first glance, namely, theoretical physics and group theory, it is fascinating that the Monster seems to be a connection between them.

6 Conclusion

Sometimes things that seem very disjointed come together to morph into something beautiful. The monstrous moonshine conjecture is one of those things. Bringing together number theory and algebra, and concepts from math and concepts from physics, moonshine gives us the start of a unified theory of how these things fit together.

The monstrous moonshine conjecture really starts with the Monster. This huge simple group caught the interest of many in the past half century. Due to efforts to study and understand the Monster, the connection between it and modular functions was found.

In every area of study, you need people who are not content with just noticing interesting facts. You need people who are willing to ask “why”, and then put in the work to find the answer. In the area of monstrous moonshine, those people were Conway, Norton, Frenkel, Lepowsky, Meurman, Thompson, McKay and Borcherds, among others. Because of their dedication to finding out why the coefficients of the J-function and the dimensions of the irreducible representations of the Monster are so closely connected, we have this beautiful theory that brings together people and ideas from so many different areas.

Far from being the last topic to be explored in this area, moonshine has opened our eyes to how many connections have yet to be discovered. There are many other related theories to moonshine. Subquotients of the Monster have their own moonshine theories. The umbral moonshine correspondence was connected to the Mathieu sporadic group M_{24}. There are questions of correspondences between sporadic groups and some Lie algebras [11].

It will be exciting to see where these theories go heading forward. Connections between group theory and physics could lead to all sorts of new theories. Although moonshine has radically impacted algebra, it has had a far less intense impact on number theory [7]. Perhaps this connection will be explored more from the number theory perspective in the years to come.

In conclusion, moonshine connected fields that seemed too different. It explained “coincidences” in a beautiful way. In the future, hopefully the moonshine conjecture will serve as an encouragement to mathematicians and scientists working on obscure or supposedly unexplainable theories. What today may seem like a crazy idea may tomorrow be a huge breakthrough in multiple areas of mathematics.

References

[1] Tom M. Apostol. *Modular Functions and Dirichlet Series in Number Theory*. Graduate Texts in Mathematics. Springer-Verlag New York, 1990. ISBN: 978-0-387-97127-8.

[2] Ruel V. Churchill and James Ward Brown. *Complex Variables and Applications*. 4th ed. McGraw-Hill, Inc, 1984. ISBN: 0-07-010873-0.

[3] J. H. Conway et al. *Atlas of Finite Groups*. Clarendon Press, 1985.

[4] John B. Fraleigh. *A First Course in Abstract Algebra*. 7th ed. Pearson Education, 2003. ISBN: 0-201-76390-7.

[5] I. B. Frenkel, J. Lepowsky, and A. Meurman. "A Natural Representation of the Fischer-Griess Monster with the Modular Function J as Character". In: *Proceedings of the National Academy of Sciences* 81.10 (1984), pp. 3256–3260. ISSN: 0027-8424. DOI: 10.1073/pnas.81.10.3256. eprint: https://www.pnas.org/content/81/10/3256.full.pdf. URL: https://www.pnas.org/content/81/10/3256.

[6] Igor B. Frenkel, James Lepowsky, and Arne Meurman. *Vertex Operator Algebras and the Monster*. Pure and Applied Mathematics. Academic Press, 1988. ISBN: 0-12-267065-5.

[7] T. Gannon. "Monstrous Moonshine: The First Twenty-Five Years." In: *Bulletin of the London Mathematical Society* 38.1 (2006), pp. 1–33. ISSN: 14692120. URL: https://cedarville.ohionet.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-31744448388&site=eds-live.

[8] I. Martin Isaacs. *Character Theory of Finite Groups*. AMS Chelsea Publishing, 1976. ISBN: 0-8218-4229-3.

[9] Gordon James and Martin Liebeck. *Representations and Characters of Groups*. 2nd ed. Cambridge University Press, 2001. ISBN: 0-521-00392-X.

[10] Neal Koblitz. *Introduction to Elliptic Curves and Modular Forms*. 2nd ed. Graduate Texts in Mathematics. Springer-Verlag New York, 1993. ISBN: 0-387-97966-2.

[11] Valdo Tatitscheff. *A Short Introduction to Monstrous Moonshine*. 2019. DOI: 10.48550/ARXIV.1902.03118. URL: https://arxiv.org/abs/1902.03118.