The Spirit of Moonshine: Connections between the Mathieu Groups and Modular Forms

Jeremy Booher

jbooher@fas.harvard.edu

Adviser: Benedict Gross

March 22, 2010

Submitted to the Harvard University Department of Mathematics in partial fulfillment of the requirements for the degree of A.B. in Mathematics

Acknowledgements

I wish to thank Dick Gross, my thesis adviser, for his wisdom and patience in helping me select a manageable piece of moonshine on which to write my thesis. It would have been very easy to pick so broad a topic it would be impossible to move beyond the superficial. Thank you also for reading and improving several drafts. I also to wish to thank Noam Elkies for being an excellent source of knowledge about Steiner systems and the Mathieu groups.

Thanks also to Amanda Folsam, and Riad Masri, and Ken Ono at the Wisconsin REU in Number Theory for introducing me to the theory of modular forms and monstrous moonshine.

Thanks also to Carl Erickson and Christian Zamora Jaen for agreeing to edit the mathematical heart of my thesis, and to Jeremy Aron-Dine and Don Larsen for giving me a non-mathematicians impression of my thesis.

Finally, I thank my parents for supporting me as my path into mathematics went far from their path in life. They always were interested, they always provided food, and they always loved. Thank you also for finding Steve Munden, who showed me in high school what math is all about by teaching me to prove.

moonshine

2. a. (*n.*) Appearance without substance; something unsubstantial or unreal; (now) *esp.* foolish or fanciful talk, ideas, plans, etc. Originally †**moonshine in the water**. ... **1887** *Spectator* 3 Sept. 1173 As for all this talk about Federalism, it is moonshine. It means nothing practical at all.

2. (*adj.*) Vain, empty, foolish; worthless. *rare.* ... **1668** H. MORE Divine Dialogues I. III. xxvi. 471 They are weak, abortive, Moon-shine Conceptions.

-Oxford English Dictionary [**20**]

Contents

Introduction

A starting point in the theory of moonshine is the 1979 paper of J. H. Conway and S. P. Norton entitled "Monstrous Moonshine" [**18**]. It collects several seeming-coincidences about how the then conjectural Monster group, along with other sporadic finite simple groups like the Mathieu groups, relate to modular forms. The amazing fact was that these were not coincidences, but the beginnings of an unexpected link between the representation theory of finite groups and modular forms, a seemingly unrelated branch of number theory.

In the forty year since this original paper, many more connections between modular forms and the Monster group, the Mathieu groups, and most of the other finite simple sporadic groups have been discovered. They are collectively referred to as moonshine. The connection between the Mathieu group M_{24} and Hecke eigenforms which is the focus of this thesis was described by G. Mason in 1985 [**17**]. Significant progress was made in the 1990's, and R. Borcherds won a Fields medal in 1998 in part for his work in proving Conway and Norton's original conjectures. The proof opened up connections between number theory and representation theory with mathematical physics. Moonshine acts a bridge allowing knowledge about number theory or representation theory to illuminate aspects of the other field, and to explain apparent coincidences. (For surveys, see [**3**] and [**10**]). But despite being able to prove these connections, the central question remains unanswered: why do the Monster group and its relatives have anything to do with modular forms?

1. The History of Monstrous Moonshine

A look at the outlines of monstrous moonshine will help motivate the rest of this thesis which examines the aspects of moonshine relating to the Mathieu groups. The story, as related in [**3**], begins with the three equalities

$$\begin{aligned} 1 &= 1 \\ 196884 &= 196883 + 1 \\ 21493760 &= 21296876 + 196883 + 1 \end{aligned} \tag{1}$$

The numbers on the left of (1) come from the modular function $j(z)$. It is a function on the upper half plane $z \in H = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ that appears in complex analysis and number theory as the prototypical example of a modular function, which is a function that transforms "nicely" so that $j(z) = j(z+1) = j(-1/z)$. If we let $q = e^{2\pi i z}$, then $j(z)$ is expressible as a $q-$series with integer coefficients

$$j(z) = q^{-1} + 744 + 196884q + 21493760q^2 + \dots$$

The numbers on the right of (1) are associated to the Monster group. The Monster group is the largest of 26 sporadic simple groups. All the other simple groups, groups with only the one element subgroup and whole groups as normal subgroups, lie in infinite families

according to the classification of finite simple groups. The Monster contains

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$$

elements, a 54 digit number. Its existence is a non-trivial fact: when the original moonshine conjectures were made, mathematicians suspected its existence, and had been able to work out its character table, but could not prove it actually existed. They did know that the dimensions of the smallest irreducible representations would be 1, 196883, and 21296876.

When J. McKay first noticed these relationships, other mathematicians were sure they were simply coincidences, so the subsequent theory took on the name *moonshine*, after foolish and empty ideas.[1] McKay and Thompson suggested that the explanation for these equalities should lie in the existence of an infinite dimensional graded representation $V = \bigoplus_{n\in\mathbb{Z}} V_n$ of the Monster group such that the series

$$T_g(z) = \sum_n \mathrm{tr}(g|V_n)q^n$$

is an "interesting" function for each element g of the Monster group. Taking the identity element, the traces are simply the dimensions of the graded piece V_n, which appeared to give the coefficients of the $q-$expansion of the function $j(z)$. If one only looked at the identity element, each V_n could simply be copies of the trivial representation and still produce the j function, so the condition that $T_g(z)$ be interesting for other elements of the Monster group is crucial. Conway and Norton [**18**] proposed that all of the $T_g(z)$ are special types of modular functions called Hauptmodul, of which $j(z)$ is an example. A Hauptmodul for a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ is an isomorphism $\Gamma\backslash H \to \mathbb{C}$ normalized so that its $q-$expansion begins $q^{-1} + O(1)$. More precisely, Conway and Norton conjectured the following, which Borcherds proved:

Conjecture 0.1. *There is an infinite dimensional graded representation* $\bigoplus_{n\in\mathbb{Z}} V_n$ *of the Monster group such that for any element* g *of the Monster group, the series* $T_g(z)$ *is a Hauptmodul for a genus* 0 *subgroup of* $\mathrm{SL}_2(\mathbb{R})$.

A. O. L. Atkin, P. Fong, and S. D Smith were able to prove the existence of such a representation through computer calculations, but shed no light onto what this representation actually was. Frenkel, Lepowsky, and Meurman managed to find an explicit construction of a representation so that $T_1(z) = j(z)$, but it was not obvious that it satisfied the remaining parts of the conjecture. Borcherds' work showed that these two representations agreed.

2. The Mathieu Groups and Moonshine

In the mid 80's, G. Mason attempted to understand understand monstrous moonshine by looking for analogs of monstrous moonshine for smaller sporadic finite simple groups. The Mathieu groups are the five smallest sporadic simple groups, and were known to Mathieu and Frobenius in the 19th century. The largest of these 5, M_{24}, contains $2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23 = 244823040$ elements. It is denoted M_{24} because it arises as a permutation group for a 24 element set S.

[1] Moonshine also refers to whiskey, especially illegal whiskey, and there is alcohol in this story. When A. Ogg noticed that the prime factors of the order of the Monster group are precisely those primes that satisfy a condition in the theory of modular functions, he offered a bottle of Jack Daniels as a prize for an explanation of this fact [**14**].

Mason realized that just as elements of the Monster group correspond to Hauptmodul, elements of M_{24} correspond to Hecke eigenforms, a special type of modular form [**17**]. A building block for modular forms is the Dedekind $\eta-$function. To each element $g \in \mathrm{M}_{24}$ is associated a product of $\eta-$ functions f_g based on how g acts on S.

More precisely, Mason proceeded as follows. For $q = e^{2\pi iz}$, define the Dedekind $\eta-$function by

$$\eta(z) := q^{\frac{1}{24}} \prod_{n=1}^{\infty}(1-q^n).$$

Now $g \in \mathrm{M}_{24}$ permutes the set S, and so one can define the cycle shape of g to be $1^{r(1)}2^{r(2)}\dots 24^{r(24)}$ where g acts as a product of $r(i)$ cycles of length i. Then define

$$f_g(z) = \prod_{d=1}^{24} \eta(dz)^{r(d)}$$

Furthermore, define the weight $k(g)$ to be half of the number of cycles of g and the level $N(g)$ to be the product of the lengths of the longest and shortest cycles. The key observation will be the following relation to modular forms.

THEOREM 3.2. *For each $g \in \mathrm{M}_{24}$, $f_g(z)$ is a cusp form and a Hecke eigenform, with weight $k(g)$, level $N(g)$, and a quadratic nebentypus character. The character is trivial if $k(g)$ is even.*

Furthermore, each of the products $f_g(z)$ has a $q-$expansion

$$f_g(z) = \sum_{n=1}^{\infty} \gamma_n(g) q^n$$

where each of the γ_n is a class function on M_{24} taking on integral values. In fact, each of the γ_n are virtual characters, see Theorem 3.7. Because each of the $f_g(z)$ is a Hecke eigenform, the coefficients of the $q-$expansion are multiplicative so $\gamma_n \cdot \gamma_m = \gamma_{nm}$ for relatively prime n and m. This interesting family of characters comes from an natural infinite dimensional graded virtual representation, described in another paper by Mason [**16**], analogous to the module for the Monster.

To describe it, let M be the standard $24-$dimensional permutation representation of M_{24}. Let $\Lambda^r(M)$ denote the rth exterior power of M. Represent a partition λ of n by $(\lambda_1, \lambda_2, \dots, \lambda_n)$ where λ_k is the number of times k appears in the partition. Being a partition means that $\sum_{k=1}^{n} k\lambda_k = n$. Denote λ being a partition of n by $\lambda \lhd n$. For example, $(1,0,1,0) \lhd 4$, since $4 = 3+1$. Define

$$\sigma(\lambda) := (-1)^{\sum_{k=1}^{n} \lambda_k}$$

and define

$$M_\lambda := \bigotimes_{k=1}^{n} \Lambda^{\lambda_k}(M).$$

We will see that these are the graded components of the virtual module.

THEOREM 3.11. *Let V be the infinite dimensional graded virtual module*

$$V = \oplus V_n \quad \textit{where} \quad V_n = \sum_{\lambda \lhd (n-1)} \sigma(\lambda) M_\lambda.$$

Then V explains the family of characters γ_n in the sense that for $g \in \mathrm{M}_{24}$

$$\mathrm{tr}(g|V_n) = \gamma_n(g).$$

The bulk of this thesis will explain Mason's results summarized above. It aims to be accessible to anyone who has studied mathematics at an undergraduate level. The main prerequisites which are not reviewed are basic group theory and complex analysis. Chapter 1 is devoted to reviewing the necessary representation theory, constructing the Mathieu groups via Steiner systems, and understanding the representation theory of the Mathieu groups, especially the conjugacy classes and permutation representation. Chapter 2 defines modular forms and focuses on properties and examples of Hecke eigenforms. Chapter 3 deals with Mason's results directly, proving Theorems 3.2 and 3.11 and illustrating how the bridge of moonshine between representation theory and number theory can be used to better understand both.

CHAPTER 1

Representation Theory of M_{24}

The Mathieu groups M_{11}, M_{12}, M_{22}, M_{23}, and M_{24} are the simplest of the sporadic finite simple groups. The goal of this section is to provide the background material from representation theory, the construction of the Mathieu groups, and the relevant information about the character table and conjugacy classes of M_{24} in preparation for a discussion of moonshine.

1. Facts from Representation Theory

The motivating idea behind representation theory is to reduce the problem of understanding a group to that of understanding its image in a group of linear transformations of a finite dimensional vector space. These can be understood using the techniques of linear algebra and in turn can give group theoretic information. An excellent summary for basic representation theory is Serre's *Linear Representations of Finite Groups* [**24**].

1.1. Virtual Representations, Modules, and Characters. The only representations necessary here will be of the following type:

Definition 1.1. Let G be a finite group and V an n dimensional vector space over $\mathbb{C}$. A representation of G on V is a homomorphism $\rho : G \to \mathrm{GL}(V)$.

Recall that this data is the same as making V a $\mathbb{C}[G]$ module.

The character of a representation is the trace of the homomorphism ρ. A representation is irreducible if it has no non-trivial vector subspaces that are fixed under the action of G. The character of an irreducible representation is said to be irreducible. The following basic fact is proved in Serre [**24**]:

Proposition 1.2. *Every representation splits into a direct sum of irreducible representations. In terms of $\mathbb{C}[G]$ modules, this means that every representation is semisimple.*

The irreducible characters of G form an orthonormal basic for the the space of complex valued class functions on G with the inner product

$$(f, f') := \frac{1}{|G|} \sum_{g \in G} f(g)\overline{f'}(g) \tag{2}$$

In addition to the basic theory of ordinary representations and characters, the theory of moonshine requires a basic understanding of virtual representation and virtual characters, as discussed in the appendix of Serre [**24**]. A virtual representation (or virtual module) is nothing more than formal $\mathbb{Z}$-linear combination of $G-$representations. They are elements of the Grothendieck group for the category of finitely generated $\mathbb{C}[G]$ modules. The Grothendieck group is an Abelian group generated by $[M]$ where M is a finitely generated $\mathbb{C}[G]$ module, with the relation that $[M_2] = [M_1] + [M_3]$ for each short exact sequence

$$0 \to M_1 \to M_2 \to M_3 \to 0$$

Since all $\mathbb{C}[G]$ modules are semi-simple, they split into direct sums of simple modules (irreducible representations) and hence the Grothendieck group is precisely a free Abelian group generated by the irreducible representations of G. The Grothendieck group can be made into a ring: the product of two finitely generated $\mathbb{C}[G]$ modules is their tensor product over $\mathbb{C}[G]$.

A virtual character is simply a $\mathbb{Z}$-linear combination of the irreducible characters. To a virtual representation one can associate a virtual character by replacing each representation in the formal linear combination by its character.

1.2. Tensor Products and Exterior Products. A standard way to construct new representations is by taking tensor products. If M_1 and M_2 are $\mathbb{C}[G]$ modules, $M_1 \underset{\mathbb{C}[G]}{\otimes} M_2$ is a $\mathbb{C}[G]$ module, with the action of g defined through $g(m_1 \otimes m_2) = (gm_1) \otimes (gm_2)$. The character of the tensor product is the product of the characters. Through this, $\Lambda^r(M)$ becomes a $\mathbb{C}[G]$ module as it is a quotient of $M^{\otimes r}$ by the ideal generated by elements of the form $m_1 \otimes m_2 + m_2 \otimes m_1$, which is fixed by the action of G. The image of $v_1 \otimes v_2 \otimes \ldots \otimes v_r$ in the quotient is denoted by $v_1 \wedge v_2 \wedge \ldots \wedge v_r$. Flipping the order of two adjacent factors negates the wedge product. The character of $\Lambda^r(M)$ has a concise description.

PROPOSITION 1.3. *Let M be an n dimensional representation of G. Let $g \in G$ act on M by a linear transformation which has eigenbasis $\{e_i\}$ with eigenvalues $\{\lambda_i\}$. Then the character of $\Lambda^r(M)$ evaluated at g equals*

$$\sum_{1 \leq i_1 < i_2 < \ldots i_r \leq n} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_r}$$

In particular, this equals the rth elementary symmetric polynomial evaluated on the eigenvalues $\{\lambda_i\}$.

PROOF. This follows from picking an eigenbasis for $\Lambda^r(M)$ given by $e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_r}$ with $1 \leq i_1 < i_2 \ldots < i_r \leq n$ and by the definition of the elementary symmetric polynomials. □

In the case $r = 2$, Section 2.1 of Serre [**24**] describes the standard decomposition of $M \otimes M$ into $\mathrm{Sym}^2(M) \oplus \Lambda^2(M)$, where $\mathrm{Sym}^2(M)$ is the subspace of $M \otimes M$ generated by $m_1 \otimes m_2 + m_2 \otimes m_1$. In particular, if χ is the character associated to M then the character of $\Lambda^2(M)$ is

$$\chi^2_{alt}(g) = \frac{1}{2}(\chi(g)^2 - \chi(g^2)). \tag{3}$$

1.3. Induced Representations and Permutation Representations. Another way to construct new representations is through extending representations from subgroups to the whole group, or by restricting representations to subgroups.

DEFINITION 1.4. Let $H < G$ and let W be a representation of G. Then the restriction of W to H is given by viewing W as an $\mathbb{C}[H]$ module, and is denoted by $\mathrm{Res}^G_H(W)$.

It is obvious this is a representation of H. Going in the other direction, one can construct a representation of the whole group through extension of scalars.

DEFINITION 1.5. Let W be a representation for a subgroup $H < G$. The induced representation $\mathrm{Ind}^G_H(W)$ is defined to be $\mathbb{C}[G] \underset{\mathbb{C}[H]}{\otimes} W$.

An alternate description is in terms of left cosets. Let H be a subgroup of G with H represented on a vector space W. If R is a system of representatives for G/H, then $\mathrm{Ind}_H^G(W)$ is a direct sum $\oplus_{\sigma\in R} W_{\sigma H}$, where each W_σ is a copy of W. Writing element $g \in G$ as σh with $h \in H$ and $\sigma \in R$, g acts on $\mathrm{Ind}_H^G(W)$ by sending $w \in W_{\tau H}$ to $h(w) \in W_{\sigma\tau H}$.

It is also useful to understand the character of an induced representation.

THEOREM 1.6. *If χ is the character of a representation of $H \subset G$ acting on W, then the character of $\mathrm{Ind}_H^G(W)$ is*

$$\mathrm{Ind}_H^G(\chi)(g') := \frac{1}{|H|} \sum_{\substack{g\in G \\ g^{-1}g'g\in H}} \chi(g^{-1}g'g)$$

PROOF. This is Theorem 12 of Serre [**24**]. □

A special kind of representation is a permutation representation. Suppose a group G acts on a finite set S. Using the set S as a index set for a basis of an $|S|$ dimensional vector space V, V becomes a representation of G by defining ge_s by $e_{g(s)}$. Such a representation is a permutation representation, and the associated module is called a permutation module. The trace of a permutation representation evaluated at g is simply the number of fixed points of the action of g on S. This implies that not all representations are permutation representations: for some groups like C_n for $n \geq 3$, not all characters take on integral values.

EXAMPLE 1.7. Taking $S = G$ with the group acting on the left, the associated permutation representation is the regular representation of G. Taking S to be a single point gives the trivial representation.

EXAMPLE 1.8. If H is a subgroup G which acts on a one dimensional space trivially, the induced representation is simply the permutation representation associated with the action of G on the finite set of cosets G/H.

Although not every permutation representation is induced from a trivial representation of a subgroup, there is the following partial result.

PROPOSITION 1.9. *A transitive permutation representation is induced from the trivial representation on the stabilizer of a point H.*

PROOF. This can be proved using characters. Let χ be the character associated to the transitive permutation representation, and H the stabilizer of a point. Let g_α be coset representatives for G/H corresponding to the points α. By Theorem 1.6,

$$\begin{aligned}\mathrm{Ind}_H^G(1_H)(g) &= \sum_{\alpha,\, g_\alpha^{-1}gg_\alpha\in H} 1 \\ &= \sum_{g_\alpha^{-1}gg_\alpha 1=1} 1 = \sum_{gg_\alpha 1=g_\alpha} 1 \\ &= \sum_{g\alpha=\alpha} 1 = \chi(g)\end{aligned}$$

since the number of fixed points of a permutation representation is its trace. □

Permutation representations also naturally give embeddings of a group inside a permutation group, allowing discussion of the cycle shape of an element.

DEFINITION 1.10. Given an embedding $G \hookrightarrow S_n$, the cycle shape of $g \in G$ is said to be $1^{r(1)}2^{r(2)}\dots n^{r(n)}$ if there are $r(i)$ cycles of length i in the standard cycle notation for $g \in S_n$. The cycles of length 1 can be omitted from the description.

Note that this depends on the embedding: however, in the case of the Mathieu groups there is a natural embedding of $M_n \hookrightarrow S_n$ that is understood.

1.4. Reduction to Subgroups. One way to understand representations of a group G is to understand the representations of its subgroups. The first example of this sort of theorem is Frobenius reciprocity.

THEOREM 1.11 (Frobenius Reciprocity). *Let $H < G$, ψ be a class function of H and ϕ a class function of G. Then*

$$(\psi, \mathrm{Res}^G_H(\phi))_H = (\mathrm{Ind}^G_H(\psi), \phi)_G.$$

PROOF. This follows from the fact that class functions are linear combinations of characters, that for characters χ_1, χ_2 with representations V and W,

$$\dim \mathrm{Hom}_G(V, W) = (\chi_1, \chi_2),$$

and from the adjointness of Hom and tensor products. The full details are in section 7.2 of Serre [**24**]. □

It also useful to study whether combinations of characters induced from special subgroups are enough to produce all characters of G. Two general theorems along these lines are Artin's theorem and Brauer's theorem. The proofs are found in Chapters 9 and 10 of Serre [**24**].

THEOREM 1.12 (Artin). *Every character of a group G is a rational linear combination of characters induced from characters of cyclic subgroups of G.*

A subgroup H of G is said to be $p-$elementary if $H = A \times B$ with A cyclic of order prime to p and B a p-group. A subgroup H is said to be elementary if it is $p-$elementary for some prime.

THEOREM 1.13 (Brauer). *Every character of a group G is an integral linear combination of characters induced from characters of elementary subgroups of G.*

We will need a specific kind of reduction that applies to the symmetric groups.

PROPOSITION 1.14. *Every character of S_n is of the form $\sum_{H_i \subset S_n} a_i \mathrm{Ind}^{S_n}_{H_i}(1)$ with $a_i \in \mathbb{Z}$.*

This result depends on the very concrete description of the representation theory of S_n in terms of Young diagrams. A concise exposition, without proofs, is found in Zhao [**30**]. The proofs can be found in Chapter 7 of Fulton [**9**].

DEFINITION 1.15. A partition of n is a decomposition $n = \sum_{i=1}^m \lambda_i$ where $\lambda_i \in \mathbb{Z}$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m \geq 1$.

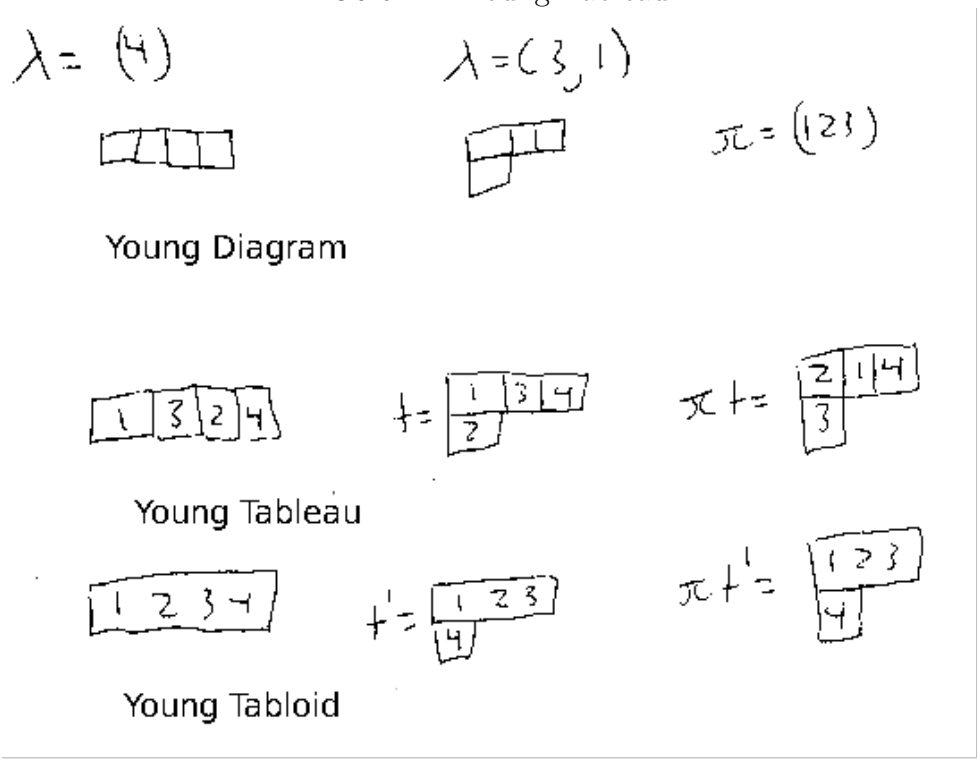
To each partition $\lambda = (\lambda_1, \ldots, \lambda_m)$, associate a Young diagram which consists of λ_i boxes in the ith row.

A Young tableau is an assignment of the integers $1, 2, \ldots, n$ to the boxes of a Young diagram of shape λ.

A Young tabloid is an equivalence class of Young tableaux under row-equivalence: two tableaux are equivalent if each row contains the same elements.

These definitions are illustrated in Figure 1. The group S_n acts on the set of Young tabloids of shape λ in the obvious manner. Denote the corresponding representation by M^λ.

FIGURE 1. Young Tableaux



LEMMA 1.16. *For a permutation λ of n, the permutation representation M^λ is transitive. In particular, it is of the form* $\mathrm{Ind}_H^{S_n}(1)$ *for $H < S_n$.*

PROOF. Transitivity is clear from the definition. Proposition 1.9 shows that it is induced from the trivial representation. □

Young tableaux can also be used to construct the irreducible representations of S_n.

DEFINITION 1.17. For a tableau t, define C_t to be the group of permutations that only permutes elements within each column of t. Define

$$e_t = \sum_{\pi \in C_t} \mathrm{sign}(\pi)\pi(t)$$

The Specht module, denoted S^λ, is the submodule of M^λ spanned by e_t where t ranges over all tableaux of shape λ.

THEOREM 1.18. *The S^λ where λ ranges over partitions of n form a complete list of irreducible representations of S_n. S^λ appears in M^λ with multiplicity one.*

PROOF. A proof is found in Section 7.2 of Fulton [**9**]. □

The last standard result we will need is about which other irreducible representations appear in M^λ. The first step is to define a partial ordering on partitions of n.

DEFINITION 1.19. For partitions $\mu = (\mu_1, \ldots, \mu_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_m)$ of n, let $\mu \geq \lambda$ if

$$\mu_1 + \ldots + \mu_i \geq \lambda_1 + \ldots + \lambda_i$$

for every i. (μ_i or λ_i are set to be 0 if $i > n$ or $i > m$ respectively.)

This ordering determines which irreducible representations appear in M^μ.

THEOREM 1.20. *For partitions μ and λ of n, M^μ contains S^λ if and only if $\lambda \geq \mu$.*

PROOF. Again, a proof is contained in Section 7.2 of Fulton [**9**]. □

This knowledge about the Specht modules allows an easy proof of Proposition 1.14. Note it suffices to prove the assertion for irreducible representations, which are exactly the S^λ by Theorem 1.18. The theorem is true for $\lambda = (n)$, for the Specht module in this case is the trivial representation. Note that this partition is maximal with respect to the partial ordering. Suppose that the assertion holds for all S^λ with $\lambda > \mu$. Then the only irreducible representations in M^μ are S^μ with multiplicity one and the S^λ for $\lambda > \mu$ by Theorem 1.20. It follows that for integers c_λ we have

$$S^\mu = M^\mu - \sum_{\lambda > \mu} c_\lambda S^\lambda.$$

But each S^λ is a integral linear combination of representations of the form $\mathrm{Ind}_{H_i}^{S_n}(1)$ by hypothesis, and M^μ itself is of this form by Lemma 1.16, so S^μ is as well.

EXAMPLE 1.21. If $n = 4$ then $\lambda = (4)$ and $\mu = (3, 1)$ are partitions of 4. M^λ is the trivial representation, since the all of the boxes are in the same row and hence the only tabloid is fixed by the action of S_n. On the other hand, M^μ is a four dimensional representation since a tabloid is uniquely determined by the element in the box in the second row, and the action of S_4 is the standard permutation action. The associated Specht module is three dimensional, and spanned by $e_1 - e_i$ for $i = 2, 3, 4$ (with the action of $\pi \in S_4$ still given by $\pi(e_i) = e_{\pi(i)}$). Thus in this case $S^{(3,1)} = M^{(3,1)} - S^4$.

2. Constructions of M_{12}

The goal of this section will be to construct M_{12}, a first example of a sporadic finite simple group. The most important feature of M_{12} is its action on a set of 12 points.

2.1. Transitive Actions. A group G acts on a set S via a homomorphism $G \xrightarrow{\sigma} \mathrm{Aut}(S)$. The map σ is almost always suppressed, so $\sigma(g)x$ is written gx.

- The action is faithful if σ is injective, i.e. only the identity element of G acts by the identity permutation on S.
- The action is transitive if for any $x, y \in S$ there exists a $g \in G$ such that $gx = y$.
- The action is n-transitive if for every pair $(x_1, \ldots, x_n)$ $(y_1, \ldots, y_n)$ of n tuples with distinct entries, there exists a $g \in G$ such that $gx_i = y_i$ for $1 \leq i \leq n$.
- The action is n-homogeneous if for every set pair of sets of n points, there exists a $g \in G$ that sends that sends one set to the other.
- The action is sharply n-transitive if it is $n-$transitive and only the identity fixes n distinct elements of S.
- If the action of G on S is sharply $n-$transitive and faithful, then

$$|G| = |S| \cdot (|S| - 1) \cdot \ldots \cdot (|S| - n + 1).$$

These standard facts are found in Chapter 9 Section 1 of Rotman [**22**].

EXAMPLE 1.22. The standard action of S_n on the set of n elements is sharply n transitive.

EXAMPLE 1.23. As detailed in Chapter 9 of Rotman [**22**], the group $\mathrm{PGL}_2(\mathbb{F}_{11})$, defined to be the projectivization of two by two matrices with entries in the finite field with 11 elements and nonzero determinant, acts on the projective line $\mathbb{P}^1_{\mathbb{F}_{11}}$ by viewing a point with homogeneous coordinates $[x_1, x_2]$ as a vector in $\mathbb{F}_{11}^2$ and acting by matrix multiplication. This can be shown to be a sharply three transitive group action. The group $\mathrm{PSL}_2(\mathbb{F}_{11})$, consisting of the image in $\mathrm{PGL}_2(\mathbb{F}_{11})$ of two by two matrices with determinant 1, also acts on $\mathbb{P}^1_{\mathbb{F}_{11}}$. It is not 3−transitive, but it is 3−homogeneous.

The most important property of group M_{12} is its action. We will show the following:

THEOREM 1.24. *There exists a sharply* 5−*transitive group* M_{12} *acting on a set of* 12 *elements. The group has order* $12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$.

There are many ways to construct the Mathieu groups. Different approaches are given in detail in Rotman [**22**] and Dixon and Mortimer [**7**]. Many different methods are summarized in the Atlas's entry on the Mathieu groups [**5**]. The method here is based around Steiner systems, and begins with the outer automorphism of S_6.

2.2. The Outer Automorphism of S_6.

DEFINITION 1.25. An inner automorphism of a group G is automorphism of G given by conjugation. An outer automorphism is an automorphism that is not inner.

EXAMPLE 1.26. The map $f : S_6 \to S_6$ sending σ to $(12)^{-1}\sigma(12)$ is an inner automorphism of S_6. It, and any other inner automorphism, must send a transposition to a transposition.

The surprising fact is that S_6 has an outer automorphism, unlike any other symmetric group. It arises through an exceptional action of S_6 on a 6 point set. The matrix group $\mathrm{PGL}_2(\mathbb{F}_5)$ acts on a 6 point set $\mathbb{P}^1_{\mathbb{F}_5}$, and hence is a subgroup of S_6. The action is faithful and sharply three transitive: for a proof see Theorem 9.48 of Rotman [**22**]. Now $|\,\mathrm{GL}_2(\mathbb{F}_5)| = 24 \cdot 20$, the number of ways to pick a nonzero vector in $\mathbb{F}_5^2$ and to pick a second vector outside the span of the first. Thus $|\,\mathrm{PGL}_2(\mathbb{F}_5)| = 24 \cdot 20/4 = 24 \cdot 5$. On the other hand, $|S_6| = 6!$, so the index of $\mathrm{PGL}_2(\mathbb{F}_5)$ in S_6 is 6. Therefore S_6 acts on the 6 cosets $S_6/\,\mathrm{PGL}_2(\mathbb{F}_5)$: for $\sigma, \tau \in S_6$, $\sigma \cdot (\tau\,\mathrm{PGL}_2(\mathbb{F}_5)) = (\sigma\tau)\,\mathrm{PGL}_2(\mathbb{F}_5)$. In particular, this gives an homomorphism ρ from S_6 to $\mathrm{Aut}(S_6/\,\mathrm{PGL}_2(\mathbb{F}_5)) \simeq S_6$.

PROPOSITION 1.27. *The map* ρ *is an outer automorphism of* S_6.

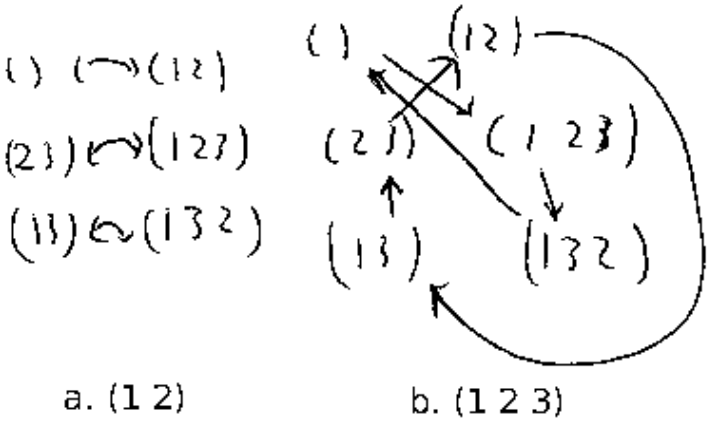
PROOF. To show ρ is an automorphism, it suffices to check it is injective. Let H be the kernel of ρ. Since A_6 is the only proper normal subgroup of S_6, H is either $\{1\}$, S_6, or A_6. However, a direct calculation can rule out the last two. A set of coset representatives for $\mathrm{PGL}_2(\mathbb{F}_5)$ in S_6 are given by $(), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)$ because any element in $\mathrm{PGL}_2(\mathbb{F}_5)$ that fixes the three points $0, 4, \infty$ is the identity. The element $(1\ 2\ 3) \in S_6$ acts as two three cycles, permuting $()$ $(1\ 2\ 3)$ and $(1\ 3\ 2)$ and permuting $(2\ 3)$, $(1\ 2)$, and $(1\ 3)$. Since $(1\ 2\ 3)$ is even and does not lie in the kernel, the kernel is trivial. Furthermore, since the image of $(1\ 2\ 3)$ is not a three cycle, the automorphism must be outer, since any inner automorphism sends a three cycle to a three cycle. □

REMARK 1.28. S_n has no outer automorphisms for $n \neq 6$. Any outer automorphism sends the conjugacy class of transpositions to a conjugacy class of order 2 with the same number of elements. For $n \neq 6$, there does not exist a different conjugacy class with the proper number of elements (Theorem 7.5 of Rotman [**22**]).

It will be useful to understand what ρ does to cycles.

LEMMA 1.29. *The automorphism ρ sends a transposition to an element of shape 2^3. It sends a three cycle to an element of shape 3^2. Furthermore, ρ sends a four cycle to a four cycle and an element of cycle shape 1 2 3 to a six cycle.*

FIGURE 2. S_6 acting by its outer automorphism



PROOF. Figure 2 shows the action of a specific transposition and a three cycle. Note that all transpositions and three cycles act with the same pattern since they are conjugates.

Both the 4 cycle and the element of shape 2 3 are a composition of a transposition τ and a three cycle σ. Since $\rho(\sigma)$ is a product of two three cycles and $\rho(\tau)$ is of shape 2^3, $\rho(\tau\sigma)$ is determined by the way $\rho(\tau)$ interacts with the two sets of points induced by $\rho(\sigma)$. If it interchanged the two three cycles, the composition would be a 6 cycle. Otherwise, the transposition interchanges two points within each three cycles, and hence the composition fixes two points. In the case of an element of order four, the last case must occur and the image must be of order four, so $\rho(\tau\sigma)$ is a four cycle. For an element of shape 2 3, the image must be of order 6 and so is a six cycle. □

2.3. The Steiner System $S(5, 6, 12)$. The exceptional outer automorphism of S_6 yields a construction of the exceptional Steiner system of type $S(5, 6, 12)$.

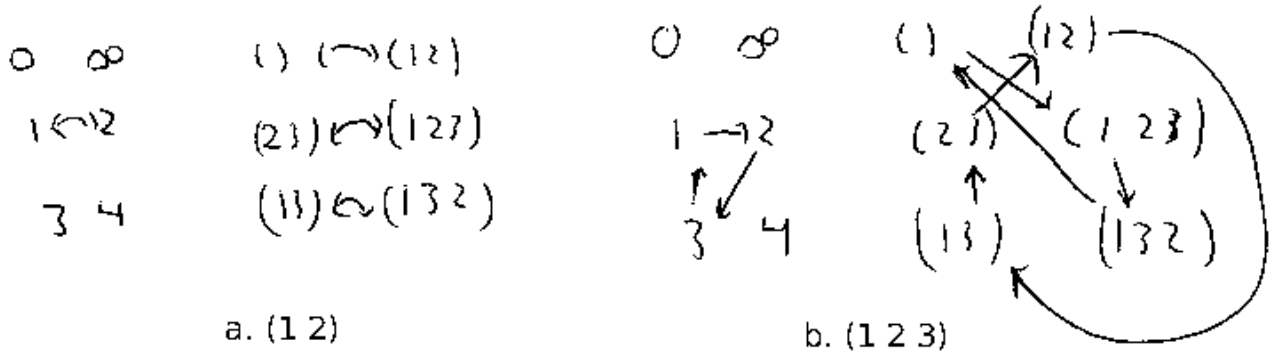
DEFINITION 1.30. Let $1 < t < k < v$ be integers. A Steiner system of type $S(t, k, v)$ consists of the pair $(S, \mathcal{B})$, with S a finite set with v elements, and $\mathcal{B}$ a family of k element subsets of S such that every t elements of S lie in a unique block $B \in \mathcal{B}$.

EXAMPLE 1.31. The set of lines in an affine or projective plane over a finite field with q elements give Steiner systems of type $S(2, q, q^2)$ and $S(2, q + 1, q^2 + q + 1)$.

The Steiner systems of type $S(5, 6, 12)$ is exceptional because it does not fall into any known infinite family. Determining whether any Steiner systems exist for a given set of parameters is an open problem.

Let S be a set of 12 points, 6 labeled $0, 1, 2, 3, 4, \infty$ and the other 6 labeled by cosets of $S_6/\operatorname{PGL}_2(\mathbb{F}_5)$. Let the points be denoted by T and the cosets by T'. Let S_6 act on T via the usual action and on T' via the outer automorphism. The action of (1 2) and (1 2 3) are pictured in Figure 3.

FIGURE 3. Action of S_6 on S



PROPOSITION 1.32. *Let $\mathcal{B}$ consist of the following subsets of S:*

1. *For a transposition $\tau \in S_6$, the two elements of T that are interchanged by action of τ, plus two of the pairs of the cosets switched by the action.*
2. *For a transposition $\tau \in S_6$, the four elements of T fixed by the action of τ, plus one of the three pairs of cosets switched.*
3. *For a three cycle $\sigma \in S_6$, the three elements of T permuted by the action of S_6, and one of the two cycles of T'.*
4. *T and T'.*

The set S along with $\mathcal{B}$ form a Steiner system of type $S(5, 6, 12)$.

The following lemmas are useful in proving this:

LEMMA 1.33. *Consider the three pairs of cosets induced by the action of* (1 2)*. No three cycle of the form* (1 2 p) *for $p \neq 1, 2 \in T$ sends a coset to the coset it is paired with. Furthermore, for any three cosets, one from each pair, there is a unique p such that the action of* (1 2 p) *permutes them.*

PROOF. It is possible to prove this purely mechanically through direct calculation and checking which permutations arise through fractional linear transformations of $\mathbb{F}_5$. Alternately, suppose $p = 4$ ($p = 3$ is obvious from Figure 3, and the other cases are identical to the case $p = 4$). Consider the action of (1 2 4)(1 2) = (1 4): since it is a transposition it pairs up the three cosets. However, if the action of (1 2 4) sent a coset to its pair under (1 2), then the action of the composition would fix a coset, a contradiction.

Furthermore, there are 8 triples of cosets induced by the four three-cycles (1 2 p). There are 8 ways to pick three cosets one from each pair, so it suffices to show no two three-cycles identify the cosets in the same way. Suppose (1 2 p) and (1 2 q) did. Then either (1 2 p)(1 2 q) or (1 2 p)(1 q 2) will fix three of the triples. However, the composition is a product of two disjoint transpositions. The outer automorphism of S_6 fixes the conjugacy class of two disjoint transpositions, so the action of the composition fixes at exactly 2 cosets, a contradiction. □

LEMMA 1.34. *There exists a permutation of S of order 2 that switches T and T' and sends blocks to blocks.*

PROOF. Such a permutation is given by

$$(1\ ())(2\ (1\ 2))(3\ (1\ 3))(4\ (1\ 2\ 3))(0\ (1\ 3\ 2))(\infty\ (2\ 3)).$$

A computer can quickly verify that it sends blocks to blocks.

A conceptual way to see this will be presented by giving a second construction of $S(5,6,12)$ involving $\mathrm{PSL}_2(\mathbb{F}_{11})$, along with an explanation of why this is presented now. □

The proof of Proposition 1.32 now proceeds by cases.

PROOF. Given 5 points, we must show a unique block contains them.

If all five points are elements of T', then the 6 points of T' form a block containing them. Inspection of the other cases shows that no other blocks contain five points.

If only one point is an element of T, then the only possible kind of block that can contain the five points is a block of type 1. There is a unique transposition $(1\ p)$ that swaps the two unselected cosets. Let the unselected cosets be H and H'. There are 5 possible transpositions of the form $(1\ p)$, and 5 cosets for H to be sent to. If $(1\ p)$ and $(1\ q)$ send H to the same coset, then $(1\ p)(1\ q)$ fixes it H. But $(1\ p)(1\ q) = (1\ q\ p)$ does not fix any coset. Thus no two transpositions $(1\ p)$ send H to the same coset, so one must send it to H'. This shows the five points are contained in a unique block.

If two points are elements of T, look at the transposition flipping the two of them. For concreteness, assume the points are 1 and 2. There are three cosets. $(1\ 2)$ pairs up the cosets. If two of the three selected cosets lie in the same pair, a block from condition 1 exists containing all 5 points. No block from condition 2 or 4 contains all 5. No block from condition 3 can contain the 5 points by Lemma 1.33 since two of the three cosets lie in the same pairing under $(1\ 2)$. Otherwise, if no two of the cosets are paired by $(1\ 2)$, there is a three cycle $(1\ 2\ p)$ which identifies the three cosets by Lemma 1.33. The block consisting of the three cosets and $1, 2$, and p works. It is clear no other block can contain the 5 points.

If 3 or more points are elements of T, then applying the map of Lemma 1.34 reduces it to one of the above cases.

This completes the proof that S and $\mathcal{B}$ form a Steiner system of type $S(5,6,12)$. □

The existence of the Steiner system of type $S(5,6,12)$ implies the existence of smaller Steiner systems.

PROPOSITION 1.35. *If $(S,\mathcal{B})$ is a Steiner system of type $S(t,k,v)$, define S' to be $S-\{p\}$ and $\mathcal{B}'$ to be $\{B-\{p\} : B \in \mathcal{B} \text{ and } p \in B\}$. Then $(S',\mathcal{B}')$ is a Steiner system of type $S(t-1,k-1,v-1)$.*

PROOF. Given $t-1$ points of S', there is a unique element of $\mathcal{B}$ containing them and the point p. This block corresponds to a block of $\mathcal{B}'$ containing the $t-1$ points. □

In particular, contracting the Steiner system of type $S(5,6,12)$ gives a Steiner system of type $S(4,5,11)$.

DEFINITION 1.36. An automorphism of a Steiner system is a permutation of its points which sends each block to a block. Two Steiner systems are said to be isomorphic if there exists a bijection between their sets that sends blocks to blocks.

The Mathieu group M_{12} is defined to be the automorphism group of $S(5,6,12)$.

The Mathieu group M_{11} is defined to be the automorphism group of $S(4, 5, 11)$.

This definition of M_{12} clearly show the action on a 12 point set. However, it is far from clear that the action is sharply 5−transitive, or even that there are many elements at all in M_{12}. One permutation that obviously lies in M_{12} is the permutation of order 2 from Lemma 1.34 which has cycle shape 2^6. A second class arise from the action of S_6.

PROPOSITION 1.37. *The permutations that arise from S_6 acting on S are automorphisms of the Steiner system and hence lie in M_{12}.*

PROOF. Since S_6 is generated by the transpositions, it suffices to check that the action of a transposition preserves the block structure. It suffices to check this for the transposition $(1\ 2)$. This transposition's action clearly fixes T and T'. It is also clear that it preserves the blocks from conditions 1 and 2 arising from $(1\ 2)$.

Now consider a block arising from the transposition $(1\ p)$, where $p \neq 1, 2,$. The coset $\sigma \operatorname{PGL}_2(\mathbb{F}_5)$ is paired with $(1\ p)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$. A block that consists of $1, p$ and two pairs of T' is sent by $(1\ 2)$ to $2, p$, and four points of T'. They are paired, for $\sigma \operatorname{PGL}_2(\mathbb{F}_5)$ and $(1\ p)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$ are sent to

$$(1\ 2)\sigma \operatorname{PGL}_2(\mathbb{F}_5) \quad \text{and} \quad (1\ 2)(1\ p)\sigma \operatorname{PGL}_2(\mathbb{F}_5) = (2\ p)(1\ 2)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$$

which are paired by $(2\ p)$. Thus the blocks arising from condition 1 and 2 for $(1\ p)$ are preserved.

A transposition $(p\ q)$ with $p, q \neq 1, 2$ commutes with $(1\ 2)$, and hence the transposition $(1\ 2)$ respects the pairing induced by $(p\ q)$. This implies the blocks from condition 1 and 2 are preserved.

A three cycle $(1\ 2\ p)$ creates two blocks by condition 3. By Lemma 1.33, the action of $(1\ 2)$ flips these two blocks, and so preserves the block structure.

A three cycle $(p\ q\ r)$ with $p, q, r \neq 1, 2$ also creates two blocks. It commutes with $(1\ 2)$, and hence the action of $(1\ 2)$ preserves the two three cycles $(p\ q\ r)$ induces on T'. Since $(1\ 2)$ fixes the three points $p, q, r \in T$, this means it sends blocks of this type to blocks.

A three cycle $(1\ p\ q)$ with $p, q \neq 1, 2$ creates two blocks. The transposition $(1\ 2)$ sends the points $1, p, q$ to $2, p, q$. $(1\ 2)$ sends the three cosets $\sigma \operatorname{PGL}_2(\mathbb{F}_5)$, $(1\ p\ q)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$, and $(1\ q\ p)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$ to $(1\ 2)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$, $(1\ 2)(1\ p\ q)\sigma \operatorname{PGL}_2(\mathbb{F}_5) = (2\ p\ q)(1\ 2)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$, and $(2\ q\ p)(1\ 2)\sigma \operatorname{PGL}_2(\mathbb{F}_5)$. These three cosets, along with the points $2, p, q$, form a block through condition 3. Thus $(1\ p\ q)$ preserves the block structure of the Steiner system.

This verifies that $(1\ 2)$ preserves the block structure of $S(5, 6, 12)$. The same proof works for any transposition, so the permutations induced by the action of S_6 on S lie in the Mathieu group. □

There are many alternate constructions of a Steiner system $S(5, 6, 12)$ that easily give other pieces of information about M_{12}. The choice of construction is irrelevant, because:

PROPOSITION 1.38. *All Steiner systems of type $S(5, 6, 12)$ are isomorphic. In particular, M_{12} is well defined.*

This proof is adapted from the proof found in Chapter IV Section 2 of Beth, Jungnickel, and Lenz [**2**]. The key idea is to contract the Steiner system three times, ending up with a Steiner system of type $S(2, 3, 9)$, which is an affine plane. By understanding the geometry of this plane one shows that $S(5, 6, 12)$ is uniquely determined. The first step is to show that Steiner systems of type $S(2, 3, 9)$ are unique up to isomorphism.

FIGURE 4. Leech triangle for M_{12}

					132						
				66		66					
			30		36		30				
		12		18		18		12			
	4		8		10		8		4		
1		3		5		5		3		1	
1	0		3		2		3		0		1

LEMMA 1.39. *All Steiner systems of type $S(2,3,9)$ are isomorphic to a two dimensional affine plane over $\mathbb{F}_3$.*

PROOF. Pick a block B of a Steiner system of type $S(2,3,9)$. Let p be a point not in this block. p induces a pairing on the remaining 6 points based on which pairs of points are in a block with p. Three pairs include a point of B and a point not in B, while the remaining pair comes from a block B' consisting of 3 points not in B. Let q be a point not in B or B'. Again, there is a block B'' disjoint from B containing q. It cannot intersect B' in two points since there is a unique block through any two points. Suppose it intersects B' in a single point p'. Then pairing the points outside of B based on the point p', the $B' - \{p'\}$ and $B'' - \{p'\}$ are paired by hypothesis. This means that the sixth point outside of B must be paired with a point of B, but then that two points of B must be paired. This violates the assumption that through any two points there is a unique block. Thus there are blocks disjoint B' and B'' that partition the set of 9 blocks.

Referring to the blocks as lines, this says there are four parallel classes of lines in this Steiner system, corresponding to the 4 lines through a fixed point p. (There are $12 = \binom{9}{2}/\binom{3}{2}$ lines total.) Pick one parallel class, and write down the lines in three rows. Rearrange within each row so that each column is a line. Given any two points not in the same row or column, there is a unique point that is in the third row and column. This is the only point that can be on the line joining these two points, since any two points must determine a unique line. This matches the description of the lines in an affine plane. □

To prove Proposition 1.38, the first step is to use the Leech triangle to extract information about the intersection of blocks based solely on the combinatorial properties of $S(5,6,12)$. Fix a block S, subsets $S_i \subset S$ with $|S_i| = i$ and $S_i \subset S_{i+1}$. The Leech triangle is defined so that the ith number in the nth row is the number of blocks B such that $S_i = S_n \cap B$. This turns out to be easy to calculate. The fact that the number of blocks containing S_i is the number of ways to pick $5-i$ more points (which is $\binom{12-i}{5-i}/\binom{6-i}{5-i}$) gives the right hand side of the triangle. Note that this is independent of the choice of S_i. The number of blocks meeting S_n in exactly S_i is the number of blocks meeting S_{n+1} in S_i plus the number of blocks meeting S_{n+1} is $S_i \cup (S_{n+1} \backslash S_n)$, which is the same as the number of blocks meeting S_{n+1} in S_{i+1}. Then this relation along with the rightmost column completely determines the Leech triangle, and is independent of the choices made. The whole triangle is shown in Figure 4. In particular, the triangle tells that the since no blocks intersect in 1 point the complement of a block must be a block.

Now let $D = (S, \mathcal{B})$ be a Steiner system of type $S(5,6,12)$, and $D' = (S', \mathcal{B}')$ be the contraction at three points ∞_1, ∞_2 and ∞_3. It is of type $S(2,3,9)$, so is an affine plane by Lemma 1.39. The 132 blocks of D are of the following types:

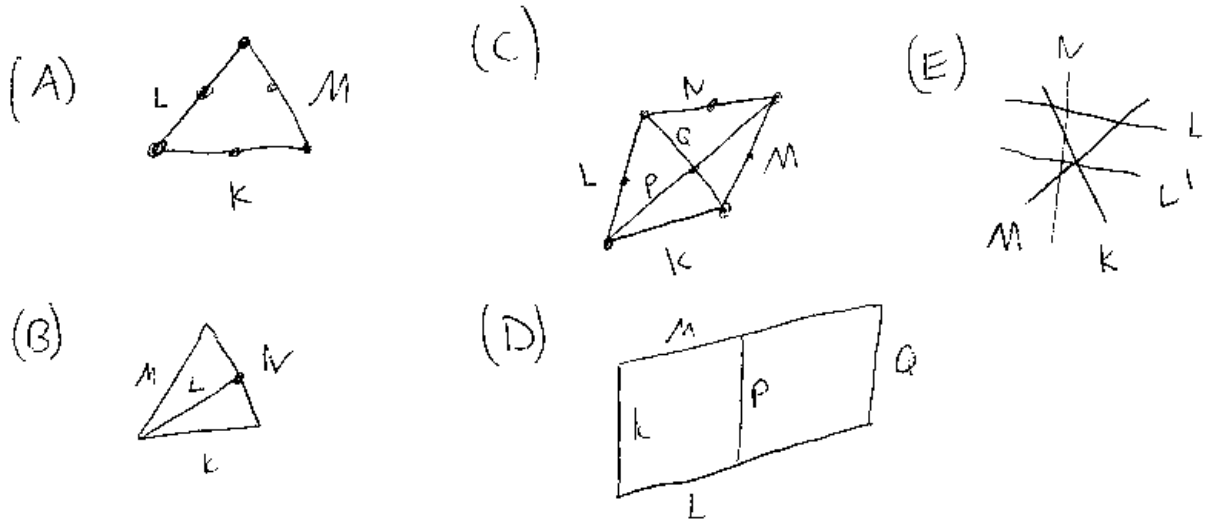
(1) 12 blocks of the form $L \cup \{\infty_1, \infty_2, \infty_3\}$ where L is a line of D'.
(2) 12 blocks that the union of two parallel lines of D', the complement of the blocks in 1.
(3) Blocks of the form $A \cup \{\infty_i, \infty_j\}$ where $i \neq j$ and A is a four element subset of S' with no three points colinear.
(4) Complements of blocks in 3, which are sets $\{\infty_i\} \cup L \cup M$ where L and M are lines of D' that intersect in a point.

Since there 132 blocks total and the number of blocks arising from 3 and 4 are equal, there 54 of each of these types. Call the union of two non-parallel lines in D' a counteroval.

Observe that the counterovals O are subdivided into three classes A_i based on which ∞_i to add so that $A \cup \{\infty_i\} \in \mathcal{B}$. There are 18 for each choice of ∞_i. If two counterovals are in the same class, then they cannot intersect in exactly four points, for then two blocks of D would intersect in 5 points, violating the Leech triangle.

The next step is to show that there is a unique equivalence relation on the set of 54 counterovals divided into three equivalence classes with the property that two equivalent counterovals never intersect in exactly four points. This proceeds in a sequence of steps, illustrated in Figure 5.

FIGURE 5. Proof of Uniqueness of $S(5, 6, 12)$



(A) If K, L, and M are sides of a triangle, then since $K \cup L$ meets $K \cup M$ in 4 points $K \cup L \not\simeq K \cup M$.

(B) If K, L, M, and N are mutually non-parallel with K, L, and M having a common point not on N, then $K \cup M \simeq L \cup N \not\simeq K \cup N \simeq L \cup M \not\simeq K \cup L \simeq M \cup N$. To prove this, note that by A, we know that $L \cup N \not\simeq K \cup N, M \cup N$ and that $K \cup M \not\simeq K \cup N, M \cup N$. Since there are only three equivalence classes, this forces $K \cup M \simeq L \cup N$. The other statements follow by permuting the lines.

(C) Let K, L, M, N be four distinct lines of D' with $K \nparallel L \parallel M \nparallel N \parallel K$ (geometrically, a parallelogram). Then the diagonals P and Q intersect in a point s. Using condition (B), $K \cup L \simeq P \cup Q$ and $M \cup N \simeq P \cup Q$. Since $\simeq$ is an equivalence relation, $K \cup L \simeq M \cup N$.

(D) Let $K \nparallel L \parallel M$. Letting P and Q be the lines parallel to K as pictured in the figure. Then using (C), $K \cup L \simeq M \cup P \simeq L \cup Q \simeq K \cup M$.

(E) If K, L, M, N are mutually non-parallel, then there is a line $L' \parallel L$ such that K, L', and M have a point in common. Using (B) and (D), it follows that $K \cup L \simeq K \cup L' \simeq M \cup N$.

These conditions show that if $K \cap L \neq \emptyset$ and $M \cap N \neq \emptyset$ then

$$K \cup L \simeq M \cup N \quad \text{if} \quad K \parallel M \text{ and } L \parallel N, \text{ or } K \parallel N \text{ and } L \parallel M, \text{ or } K, L \nparallel M, N$$

This condition is only if as well, because otherwise all counterovals would be equivalent. Thus there is a unique equivalence relation with the property that equivalent counterovals cannot intersect in four points.

This tells us that the three classes of counterovals A_i are uniquely determined by D up to re-indexing. In particular, the blocks of D are uniquely determined, up to re-indexing, because the blocks of D are determined by conditions 1 and 4 since the complement of a block is a block. Thus there is a unique Steiner system of type $S(5, 6, 12)$ up to isomorphism.

An alternate construction of the Steiner system is through the squares modulo 11, which makes it clear that M_{12} is transitive on the blocks of the Steiner system.

THEOREM 1.40. *The group* $\mathrm{PSL}_2(\mathbb{F}_{11})$ *acts on* $\mathbb{P}^1_{\mathbb{F}_{11}}$. *Let* $T = \{\infty, 1, 3, 4, 5, 9\}$, *the quadratic residues modulo* 11 *and* ∞. *Define* $\mathcal{B} := \{L(T) : L \in \mathrm{PSL}_2(\mathbb{F}_{11})\}$. *Then* $(\mathbb{P}^1_{\mathbb{F}_{11}}, \mathcal{B})$ *is a* $S(5, 6, 12)$ *Steiner system.*

PROOF. This proof is inspired by a construction from Beth, Jungnickel, and Lenz [**2**], Chapter IV Section 1.2. It is clear that for any three points there are the same number of blocks containing them, for the action of $\mathrm{PSL}_2(\mathbb{F}_{11})$ is 3−homogeneous (Example 1.23). Note that T is fixed by the five mappings $x \to sx$, where s is a non-zero square. These mappings are also in $\mathrm{PSL}_2(\mathbb{F}_{11})$. This means that the order of the stabilizer of T is $5m$ for some $m \in \mathbb{Z}$. The number of blocks is $|\mathrm{PSL}_2(\mathbb{F}_{11})|/5m = \frac{132}{m}$. The number of blocks through the three points $0, 1, \infty$ is equal to the number of blocks times the number of ways to pick 3 points in the block, divided by the number of ways to pick three points of $\mathbb{P}^1_{\mathbb{F}_{11}}$, so there are $\frac{12}{m}$ blocks containing $0, 1, \infty$. Twelve such blocks are listed in Table 1, which implies $m = 1$.

TABLE 1. Blocks containing $0, 1, \infty$.

A	A(T)	A	A(T)
$\begin{pmatrix} 0 & 1 \\ 10 & 4 \end{pmatrix}$	$0, 1, \infty, 2, 4, 10$	$\begin{pmatrix} 0 & 1 \\ 10 & 5 \end{pmatrix}$	$0, 1, \infty, 3, 6, 8$
$\begin{pmatrix} 1 & 2 \\ 3 & 7 \end{pmatrix}$	$0, 1, \infty, 4, 8, 9$	$\begin{pmatrix} 1 & 8 \\ 0 & 1 \end{pmatrix}$	$0, 1, \infty, 2, 6, 9$
$\begin{pmatrix} 1 & 8 \\ 10 & 4 \end{pmatrix}$	$0, 1, \infty, 3, 9, 10$	$\begin{pmatrix} 1 & 8 \\ 6 & 5 \end{pmatrix}$	$0, 1, \infty, 2, 7, 8$
$\begin{pmatrix} 1 & 10 \\ 8 & 4 \end{pmatrix}$	$0, 1, \infty, 3, 4, 7$	$\begin{pmatrix} 1 & 10 \\ 5 & 7 \end{pmatrix}$	$0, 1, \infty, 5, 7, 9$
$\begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix}$	$0, 1, \infty, 5, 8, 10$	$\begin{pmatrix} 1 & 7 \\ 1 & 8 \end{pmatrix}$	$0, 1, \infty, 6, 7, 10$
$\begin{pmatrix} 1 & 7 \\ 2 & 4 \end{pmatrix}$	$0, 1, \infty, 4, 5, 6$	$\begin{pmatrix} 1 & 6 \\ 1 & 7 \end{pmatrix}$	$0, 1, \infty, 2, 3, 5$

These twelves blocks can be realized as the rows, columns, and generalized diagonals of

$$\begin{pmatrix} 2 & 3 & 5 \\ 6 & 7 & 10 \\ 9 & 4 & 8 \end{pmatrix}.$$

But these 12 sets form a Steiner system of type $S(2,3,9)$ on these nine points (an affine geometry). Thus through any 2 points besides $0, 1, \infty$, there is a unique six element set containing $0, 1, \infty$ and those two points. Because $\mathrm{PSL}_2(\mathbb{F}_{11})$ is 3−homogeneous, any three points can be moved to $\{0, 1, \infty\}$ so this is a $S(5, 6, 12)$ Steiner system. □

COROLLARY 1.41. M_{12} *is transitive on the blocks of* $S(5, 6, 12)$.

This provides enough information to prove Theorem 1.24.

PROOF. From the second construction of the Steiner system, it is clear that $\mathrm{PSL}_2(\mathbb{F}_{11})$ is a subgroup of M_{12} and that M_{12} acts transitively on blocks. Thus the block T can be sent to the unique block containing $0, 1, 2, 3$ and 4. Then composing with the appropriate element of M_{12} arising from S_6 by Proposition 1.37, the points $0, 1, 2, 3, 4$ can be permuted so the composition sends them to the five given points in the proper order, showing the action to be five transitive. If any permutation fixes 5 elements, by transitivity assume they are five elements of T'. Since the permutation preserves the block structure, the sixth point of T' must be fixed as well. Now, suppose the permutation sends p_1 to p_2 in T. Since the block containing p_1, p_2, and 4 points of T' must be fixed, p_2 is sent to p_1. For any third point p_3 of T, the block containing p_1, p_3, and four points of T' must be sent to a block. This implies that p_3 is sent to the last point of the block containing those four points of T' and p_2. However, this means the permutation cannot fix the block containing p_1, p_2, p_3, and three points of T', but the three points of T', p_1, and p_2 are kept within the block. This contradiction shows that M_{12} acts sharply 5−transitively on S. In particular, $|\,M_{12}\,| = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$. □

The original motivation for looking at the Mathieu groups was their relation to the Monster. It is worth noting that:

THEOREM 1.42. M_{12} *is a simple group.*

PROOF. This is proven as Theorem 9.59 of Rotman [**22**]. □

2.4. Alternate Constructions. There are many other constructions of M_{12}, many of them sketched briefly in the atlas of finite groups. One alternative is to construct M_{24} first, in which case M_{12} is a 12 point stabilizer. This is the approach taken in Griess [**11**]. Alternatively, the Steiner system $S(5, 6, 12)$ and their automorphism groups can be constructed starting with $S(2, 3, 9)$, an affine plane. In Dixon and Mortimer [**7**], this affine is plane is shown to extend uniquely to a $S(3, 4, 10)$, which in turn extends uniquely to a $S(4, 5, 11)$ and $S(5, 6, 12)$. The perspective is flipped in Rotman [**22**], where the automorphism groups of the Steiner systems are constructed without reference to the Steiner systems, starting with a subgroup of $\mathrm{PGL}_2(\mathbb{F}_9)$. On the other hand, Mathieu's original definitions simply gave M_{12} in terms of generators (Chapter 6 Section 8 of Dixon and Mortimer [**7**]).

3. Constructions of M_{24}

This section will briefly discuss constructions of M_{24}, the relative of M_{12} acting on 24 points, but in much less detail than the constructions of M_{12}. The basic fact is that:

THEOREM 1.43. *There exists a simple* 5 *transitive group* M_{24} *acting on a set of* 24 *points. It has order* 244823040.

As for M_{12}, there are multiple methods of constructing M_{24}. The group M_{24} can be constructed analogously to M_{12} using an outer automorphism. The key fact is M_{12}'s outer automorphism.

THEOREM 1.44. M_{12} *has an outer automorphism. It interchanges the conjugacy classes of orders* 4 *and interchanges the classes of order* 8.

PROOF. There is a proof in Rotman [**22**]. □

Using this outer automorphism, M_{12} can be made to act on a set of 24 points, and the cycle shapes of elements in M_{12} can be used to construct a Steiner system of type $S(5, 8, 24)$. The group M_{24} is then the automorphism group of this Steiner system. The process stops here, since M_{24} has no outer automorphisms [**5**].

However, this isn't quite the way to go about constructing M_{24}, because the simple proof of the existence of this outer automorphism, given in Rotman [**22**], depends on the existence of M_{24}. Instead, as presented after Lemma 8.5 of Cameron [**4**], one looks at the graph whose nodes are pairs of complementary blocks with nodes connected if the blocks intersect in 3 points. Using a uniqueness result on graphs, pairs of complementary blocks are put in bijection with pairs of points drawn from another 12 element set. A $S(5, 6, 12)$ Steiner system can be constructed on this second set. These can be pieced together to get a $S(5, 8, 24)$ Steiner system.

Alternately, the blocks of a Steiner system of type $S(5, 8, 24)$ can be constructed from the action of $\mathrm{PSL}_2(\mathbb{F}_{23})$ on the projective line, much like the second construction of $S(5, 6, 12)$. Let H be the subgroup of $\mathrm{PSL}_2(\mathbb{F}_{23})$ generated by the matrices

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & 1 \\ 1 & -3 \end{pmatrix}$$

The orbit of ∞ under H is

$$B := \{\infty, 0, 1, 3, 12, 15, 21, 22\}.$$

Define $\mathcal{B} := \{g(B) : g \in \mathrm{PSL}_2(\mathbb{F}_{23})\}$. It turns out the stabilizer of B under the action of $\mathrm{PSL}_2(\mathbb{F}_{23})$ has 8 elements, and there are 21 blocks in $\mathcal{B}$ passing through $\{0, 1, \infty\}$. The other five elements of these blocks form 21 five element sets. A lengthy but simple check shows that for any two of the points in $\mathbb{P}^1(\mathbb{F}_{23})$ besides $0, 1, \infty$, there is a unique five element set containing those points. This means that the blocks $\mathcal{B}$ form a Steiner system on $\mathbb{P}^1(\mathbb{F}_{23})$ of type $S(5, 8, 24)$. There are more details in Chapter IV Section 1 of Beth, Jungnickel, and Lenz [**2**].

The blocks of the Steiner system $S(5, 8, 24)$ can also be constructed by extending a Steiner system of type $S(2, 5, 21)$ which arises as the geometry of a two dimensional projective plane over $\mathbb{F}_4$. It is shown in Dixon and Mortimer [**7**] how to extend this Steiner system three times to obtain a system of type $S(5, 8, 24)$. In Rotman [**22**], the automorphism groups of the Steiner systems are constructed through group extensions, starting with $\mathrm{PSL}_3(\mathbb{F}_4)$ acting on the projective plane. An alternate construction involves the Golay code as presented in Griess [**11**]. All of these constructions give the same Steiner system and hence the same group M_{24} by a uniqueness result similar to that for $S(5, 6, 12)$.

PROPOSITION 1.45. *All Steiner systems of type* $S(5, 8, 24)$ *are isomorphic.*

PROOF. This is proven in Dixon and Mortimer [**7**]. Alternately, it follows from a uniqueness theorem for the binary Golay code [**11**]. □

There are also smaller Mathieu groups occurring as stabilizers of M_{24}.

DEFINITION 1.46. The Mathieu group M_{23} is the stabilizer of a point in M_{24}.
The Mathieu group M_{22} is the stabilizer of a point in M_{23}.

Finally, M_{22}, M_{23}, and M_{24} can be shown to be simple groups.

4. Conjugacy Classes and Character Table for M_{12} and M_{24}

The theory of moonshine for the Mathieu groups involves the cycle-shapes, and more generally the irreducible representations, of the Mathieu groups. The above constructions of M_{24} and M_{12} provide enough information to find all conjugacy classes and compute the character tables, although the computations are long. Any one construction of course provides enough information in principal, and a computer algebra system such as Sage [**23**] can calculate the conjugacy classes and character tables in a routine manner, with significant computational effort. The benefit of multiple descriptions is that each provides an easy look at one particular aspect of the Mathieu groups. Since these are sporadic simple groups, it is understandable that there is no single description that illuminates all of the relevant properties.

In this section, the conjugacy classes for M_{12} are determined in full detail, as well as some of the irreducible characters of M_{12}. This illustrates the techniques that would be used to determine the conjugacy classes of M_{24} and the full character tables for M_{12} and M_{24}. The full character table for M_{12} and M_{24} are listed in the atlas [**5**].

4.1. Conjugacy Classes in M_{12}. There are two classes of permutations that obviously lie in M_{12}: those arising from the action of S_6 on the set S of 12 points (Proposition 1.37, about the first description) and those in $\mathrm{PSL}_2(\mathbb{F}_{11}) \subset M_{12}$ arising from the second description.

The action of S_6 illuminates conjugacy classes of the following cycle-shapes:

- $1^4\ 2^4$, arising from a two cycle in S_6. Given an element g_2 of cycle shape $1^4\ 2^4$, there exists 4 blocks such that g_2 acts on the block with cycle shape 1^42 and the complementary block with cycle shape 2^3. There is a unique block containing the two points flipped by each transposition of g_2 and three of the points fixed by g_2. Since g_2 preserves the Steiner system, this block must contain all four points fixed by g_2. Thus every element of shape $1^4\ 2^4$ corresponds to the four transpositions. There are 132 blocks, and since M_{12} is transitive on blocks by Corollary 1.41 it suffices to look at the block T where there are 15 transpositions. This gives a total of $132 \cdot 15 = 1980$ elements of shape of $1^4\ 2^4$. But each arises 4 times, so there are 495 distinct elements of cycle-shape 2^4 in M_{12}.
- $1^3\ 3^3$, arising from a three cycle 3 in S_6. Given such a $g_3 \in M_{12}$, an argument like the one above shows that any three cycle, plus the points fixed, is a block. Any block and any three cycle in that block induces a permutation of cycle shape $1^3\ 3^3$, each of which appears 3 times. Thus there are $132 \cdot 40/3 = 1760$ such elements.
- $1^4\ 4^2$, arising from a four cycle in S_6. Any cycle of shape $1^4\ 4^2$ preserves 4 blocks: each block contains one of the four cycles along with two of the fixed points. Furthermore, by Lemma 1.29, a four cycle acts with shape 4^2 on S. For each of the

132 blocks and each of the 90 four cycles in them, this creates an element of shape $1^4\ 4^2$. Each is repeated 4 times, so there are 2970 elements of cycle shape $1^4\ 4^2$.

- $2^2\ 4^2$, arising from an element of cycle shape $2\,4$ in S_6. By Lemma 1.29, an element of shape $2\,4$ cannot be sent to a four cycle, so must be sent to an element of shape $2\,4$. The counting argument is essentially the same as for the $1^4\ 4^2$ case. There are 2970 such elements.
- $1^2\ 5^2$, arising from a 5 cycle in S_6. Each element of this shape fixes 2 blocks, and there are 144 five cycles in each of 132 blocks, for a total of 9504 elements.
- 1 2 3 6, arising from an element of shape 1 2 3 in S_6. By Lemma 1.29, a 6 cycle acts with shape 1 2 3.. Given an element of this shape, there is only one block containing the 6 cycle. So for every block, and all of the 120 six cycles in it, there are distinct elements of shape 1 2 3 6 in M_{12}. Thus there are 15840 elements of this type.

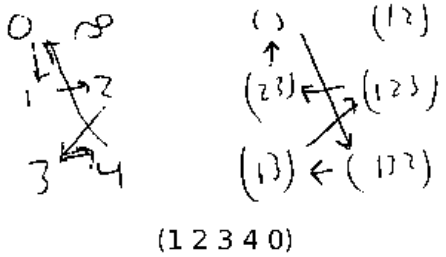
The subgroup $\mathrm{PSL}_2(\mathbb{F}_{11})$ gives further conjugacy classes. To find the size of these classes we want to find the centralizer, since the size of the conjugacy class is the order of M_{12} divided by the order of the centralizer. However, which permutations lie in the centralizer depends on the whole structure of M_{12}, not just the subgroup $\mathrm{PSL}_2(\mathbb{F}_{11})$. To compute the order of the centralizer thus requires checking mechanically whether certain permutations preserve the block structure. This makes the remaining observations inherently computational.

- Eleven cycles, arising from the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ which has order 11 by inspection. The centralizer of this element in M_{12} are just its powers, so there are $95040/11 = 8640$ such elements in the conjugacy class. However, if σ is an eleven cycle then σ and σ^2 are not conjugate. This can be verified by simply writing down the blocks and checking that any permutation that would make the two conjugate does not preserve the block structure.
- Elements of cycle shape 2^6, which can be produced from the map $z \to \frac{a}{z}$ where a is a non-residue modulo 11. Having a computer compute the centralizer, there are 396 such elements.
- Elements of cycle shape 3^4, arising from the matrix $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, which clearly has order 3 and fixes no point of $\mathbb{P}^1(\mathbb{F}_{11})$. Thus it must act with shape 3^4. There are 2640 such elements.
- Elements of cycle shape 6^2, arising for example from the matrix $\begin{pmatrix} 1 & 1 \\ 7 & 0 \end{pmatrix}$. Anything that centralizes this element must either be a power or must flip the two six cycles. There are 7920 such elements.

By combining elements from the two descriptions, the remaining conjugacy classes can be found. The results are again inherently computational, since they rely on the fact that the two constructions of a Steiner system of type $S(5, 6, 12)$ are isomorphic and need to implicitly or explicitly identify the two.

- Elements of cycle shape 10 2. These arise from a combination of the elements with shapes 5^2 and 2^6. Lemma 1.34 gives an explicit permutation τ of shape 2^6, while an element of shape $1^2\ 5^2$ arising from the action of S_6 is shown in Figure 6. The product $\tau\sigma$ can be directly computed to be of shape 10 2. The centralizer is only the 10 powers of this element, so the conjugacy class contains 9504 elements.

FIGURE 6. Permutation of Order 5 in M_{12}



- Elements of cycle shape 8 4. They arise as a product of an element of shape 2^4 and one of shape 2 3 6. However, those cannot both arise simply from the action of S_6. Instead, at least one must be conjugated by a permutation arising from $\mathrm{PSL}_2(\mathbb{F}_{11})$ that doesn't fix the block T. This requires either writing down an identification between the two Steiner systems, or finding an element of this shape directly, and verifying it preserves the Steiner blocks constructed through the outer automorphism of S_6. Such a permutation is given by

  $$(1\ ()\ \infty\ 0\ 2\ (1\ 2)\ 4\ 3)((1\ 2\ 3)\ (1\ 3)\ (1\ 3\ 2)\ (2\ 3)).$$

  A direct calculation shows it preserves blocks. The centralizer is only the 8 powers of this element, so the conjugacy class contains 11880 elements.
- Elements of shape 8 2. Composing the element of shape 8 4 given above with the element of shape 2^6 given in Lemma 1.34 gives the element

  $$(0\ (1\ 2)\ (1\ 2\ 3)\ 3\ ()\ (2\ 3)\ 4\ (1\ 3))(\infty\ (1\ 3\ 2))$$

  of shape 8 2. The centralizer is only the 8 powers, so the conjugacy class contains 11880 elements.

The number of elements identified so far is 95040, the order of M_{12}, so the above arguments have found all of the conjugacy classes. The information is summarized in Table 2. Note there are two conjugacy classes of shape 11.

TABLE 2. Cycle Shapes of M_{12}

Cycle Shape	Centralizer Order	Elements	Cycle Shape	Centralizer Order	Elements
1^{12}	95040	1	$1^4\ 2^4$	192	495
2^6	240	396	$1^3\ 3^3$	54	1760
3^4	36	2640	$2^2\ 4^2$	32	2970
$1^4\ 4^2$	32	2970	$1^2\ 5^2$	10	9504
1 2 3 6	6	15840	6^2	12	7920
$1^2\ 2\ 8$	8	11880	4 8	8	11880
$1^2\ 10$	10	9504	1 11	11	17280

4.2. Some irreducible characters of M_{12}. Recall that a character χ of an irreducible representation satisfies $(\chi, \chi) = 1$ and such an equality implies that χ comes from an irreducible representation. The goal here is to illustrate some techniques for constructing irreducible characters of M_{12} and M_{24}, but not to find the complete character table.

The trivial character χ_1 is always irreducible. Furthermore, the knowledge of the cycle shapes of M_{12} determines a character χ_{12} through the permutation representation of M_{12} on 12 points. The trace is simply the number of fixed points of the permutation. It can be written $\chi_{12} = \chi_1 + \chi_{11}$, where χ_{11} is irreducible. The assertion can be checked by verifying $(\chi_{11}, \chi_{11}) = 1$.

One method of constructing further irreducible characters is by looking at the exterior powers of a known irreducible character. $\Lambda^2(\chi_{11})$ has character $\chi_{55}(g) := \frac{1}{2}(\chi_{11}(g)^2 - \chi_{11}(g^2))$. This character turns out to be irreducible, as $(\chi_{55}, \chi_{55}) = 1$. In general, the hope is to project the new character onto the subspace of class functions spanned by the known irreducible characters and have the projection be a new irreducible character. This does not always work: trying this again on χ_{55} gives a character that decomposes as $2\chi_{55} + \chi'$, where $(\chi', \chi') = 25$ and χ' is orthogonal to χ_1, χ_{11}, and χ_{55}. The known irreducible characters are listed in Table 3.

TABLE 3. Some irreducible characters of M_{12}

	1	2^4	2^6	3^3	3^4	4^2	$4^2\,2^2$	5^2	2 3 6	6^2	8 2	8 4	10 2	11	11
χ_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
χ_{11}	11	3	-1	2	-1	3	-1	1	0	-1	1	-1	-1	0	0
χ_{55}	55	-1	-5	1	1	3	-1	0	-1	1	-1	1	0	0	0

A second place to look is in the symmetric square. The character $\chi_{66} := \mathrm{Sym}^2(\chi_{11})$ comes from a 66 dimensional representation, but is not irreducible. It decomposes as $\chi_{66} = \chi_1 + \chi_{11} + \chi_{54}$, and χ_{54} turns out to be irreducible. All these assertions can be checked using the inner product. χ_{54} turns out to be orthogonal to χ', so it gives no further help at decomposing χ'.

Since Theorem 1.44 says M_{12} has an outer automorphism, any irreducible character can be composed with the outer automorphism to yield an irreducible character which may be new. For example, the irreducible characters χ_{11} and χ_{55} give new characters χ'_{11} and χ'_{55} after switching the values on the conjugacy classes of order 4 and 8. These new characters are listed in Table 4.

TABLE 4. Some more irreducible character of M_{12}

	1	2^4	2^6	3^3	3^4	4^2	$4^2\,2^2$	5^2	2 3 6	6^2	8 2	8 4	10 2	11	11
χ'_{11}	11	3	-1	2	-1	-1	3	1	0	-1	-1	1	-1	0	0
χ_{54}	54	6	6	0	0	2	2	-1	0	0	0	0	1	-1	-1
χ'_{55}	55	-1	-5	1	1	-1	3	0	-1	1	1	-1	0	0	0

An alternate approach is to try to induce characters from a subgroup of M_{12}. A copy of S_6 is a subgroup of M_{12}, where it is embedded via the action of S_6 on the twelve points of S. Taking the trivial character on S_6, denote the induced character $\mathrm{Ind}_{S_6}^{M_{12}}(1)$ by χ_{132}. Theorem 1.6 gives the formula

$$\chi_{132}(g') = \frac{1}{720} \sum_{\substack{g \in M_{12} \\ g^{-1}g'g \in S_6}} 1.$$

Taking $g' = 1$ shows that the dimension of the representation associated to χ_{132} is in fact 132 dimensional. Table 5 shows the cycle shapes of elements in this copy of S_6.

TABLE 5. Cycle Shapes of S_6 in M_{12}.

Cycle Shape	1^{12}	$1^4\ 2^4$	$1^3\ 3^3$	$1^4\ 4^2$	$4^2 2^2$	$1^2\ 5^2$	1 2 3 6
Number of Elements	1	75	80	90	90	144	240

To calculate the induced character on other conjugacy classes, let $g \in M_{12}$ and let n_g denote the number of elements in the conjugacy class of g intersect S_6. Then g is conjugate to each of the n_g elements in m_g ways, where m_g is the order of the centralizer of g in M_{12}. Then there are $n_g m_g$ elements of M_{12} which conjugate g into S_6, and hence $\chi_{132}(g) = n_g m_g$. This character decomposes as $\chi_{132} = \chi_1 + \chi_{11} + \chi'_{11} + \chi_{54} + \psi_{55}$, where ψ_{55} is a third irreducible character of dimension 55. Its values are shown in Table 6.

TABLE 6. The Character ψ_{55} of M_{12}

	1	2^4	2^6	3^3	3^4	4^2	$4^2\ 2^2$	5^2	2 3 6	6^2	8 2	8 4	10 2	11	11
ψ_{55}	55	7	-5	1	1	-1	-1	0	1	1	-1	-1	0	0	0

This process can be continued, calculating characters associated to various exterior and symmetric powers, and inducing characters from characters of subgroups like S_6 and $\mathrm{PSL}_2(\mathbb{F}_{11})$. The point of this is that finding the character table of M_{12} requires finesse. There is no way to predict how to obtain a new character that contains only one new irreducible character without substantial trial and error: calculating complicated character tables is more of an art than a science. The complete results for M_{12} are listed in the atlas [**5**].

4.3. Conjugacy Classes and Character Table for M_{24}. The conjugacy classes of M_{24}, and some knowledge about its character table, are necessary to understand moonshine for M_{24}. The same techniques as in the previous section, along with plenty of calculation, give the conjugacy classes of M_{24} and its character table. But since this information is included in the atlas and we have already exhibited the techniques for calculating it, there is little point in working for it. A list of the conjugacy classes is reproduced here as Table 7. There are two conjugacy classes of cycle shapes $1^3\ 7^3$, 1 2 7 14, 1 3 5 15, 3 21 and 1 23.

TABLE 7. Cycle Shapes in M_{24}, from [**5**].

Cycle Shape	Centralizer Order	Cycle Shape	Centralizer Order
1^{24}	244823040	$1^3\ 7^3$	42
$1^8\ 2^8$	21504	$1^2\ 2\ 4\ 8^2$	16
2^{12}	7680	$2^2\ 10^2$	20
$1^6\ 3^6$	1080	$1^2\ 11^2$	11
3^8	504	12^2	12
$1^4\ 2^2\ 4^4$	128	2 4 6 12	12
$2^4\ 4^4$	384	1 2 7 14	14
4^6	96	1 3 5 15	15
$1^4\ 5^4$	60	3 21	21
6^4	24	1 23	23
$1^2\ 2^2\ 3^2\ 6^2$	24		

CHAPTER 2

Modular Forms and Hecke Operators

This chapter presents the definitions and essential facts about modular forms for congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$, along with the theory of Hecke operators. Two special ways to construct modular forms, $\eta-$products and modular forms with complex multiplication, are described to prepare for a discussion of moonshine for the Mathieu group. An elegant introductory account of this subject is given by Serre in Chapter 7 of Serre [**25**]. General references for this subject include Iwaniec [**12**] and Koblitz [**13**].

1. Basic Properties of Modular Forms

Loosely speaking, modular forms are complex valued functions on the upper half plane that transform "correctly" under an action of $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ on the upper half plane. The first step is to describe this action.

1.1. The Action of $\mathrm{SL}_2(\mathbb{Z})$ on the Upper Half Plane. Remember that the group $\mathrm{GL}_2(\mathbb{R})$ is the group of invertible two by two matrices with real entries. The group $\mathrm{GL}_2^+(\mathbb{R}) := \{\gamma \in \mathrm{GL}_2(\mathbb{R}) : \det(\gamma) > 0\}$ acts on the upper half plane through linear fractional transformations.

Definition 2.1. For $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R})$ and z in the upper half plane

$$H := \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\} = \{[z,1] \in \mathbb{P}^1_{\mathbb{C}} : \mathrm{Im}(z) > 0\}$$

we define the action of γ on z by

$$\gamma z := \frac{az+b}{cz+d} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} [z,1].$$

With the usual conventions viewing ∞ as the point $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathrm{GL}_2^+(\mathbb{R})$ acts on $H \cup \infty$ with γ sending ∞ to $\frac{a}{c}$ and $\frac{-d}{c}$ to ∞.

Because $\mathrm{Im}(\gamma z)$ and $\mathrm{Im}(z)$ have the same sign, the upper half plane is stable under $\mathrm{GL}_2^+(\mathbb{R})$ so this forms a well defined group action. The subgroup $\mathrm{SL}_2(\mathbb{Z})$ of integer matrices with determinant 1 acts through its inclusion in $\mathrm{GL}_2^+(\mathbb{R})$. The kernel of the map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Aut}(H)$ contains the scalar matrices, so the action descends to an action of $\mathrm{PSL}_2(\mathbb{Z})$ on H. A fundamental domain for the group action is the region of the upper half plane pictured in Figure 1. More precisely, if we let

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

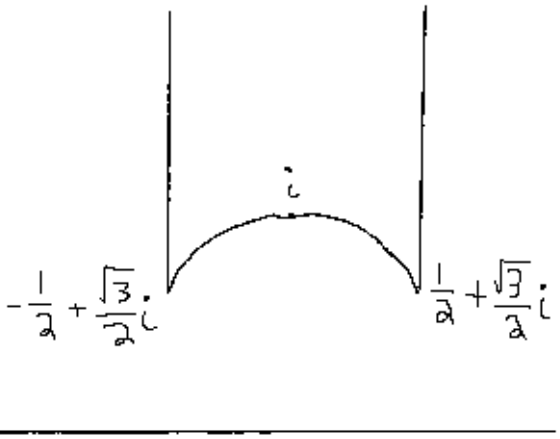
then the following theorem is true.

THEOREM 2.2. *Let D be the subset of H with $|z| \geq 1$ and $|\operatorname{Re}(z)| \leq \frac{1}{2}$. D is a fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ in the sense that*

- *For every $z \in H$, there exists a $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma z \in D$.*
- *If two distinct points z, z' of D are congruent modulo $\mathrm{SL}_2(\mathbb{Z})$, then $\operatorname{Re}(z) = \pm\frac{1}{2}$ and $z = z' \pm 1$ or $|z| = 1$ and $z' = -\frac{1}{z}$.*

Furthermore, $\mathrm{PSL}_2(\mathbb{Z})$ is generated by S and T.

FIGURE 1. Fundamental Domain for $\mathrm{SL}_2(\mathbb{Z})$



PROOF. The proof is not hard, and is Theorems 1 and 2 in Chapter 7 of Serre [**25**]. □

These groups also act on complex functions function defined on the upper half plane.

DEFINITION 2.3. If $\gamma \in \mathrm{GL}_2^+(\mathbb{R})$, k is an integer and f a function $H = \{x + iy : x, y \in \mathbb{R}, y > 0\} \to \mathbb{C}$, define

$$(f|_k\gamma)(z) := \det(\gamma)^{k/2}(cz + d)^{-k} f(\gamma \cdot z). \tag{4}$$

EXAMPLE 2.4. If $k = 0$, then the functions fixed by the action of $\mathrm{SL}_2(\mathbb{Z})$ are called modular functions. Being fixed translates into the condition that for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$

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

The j function associated with the Monster group is an example. Details are found in Apostol [**1**].

1.2. Modular Forms for Congruence Subgroups. It is important to look at functions which are invariant under the action of subgroups of $\mathrm{SL}_2(\mathbb{Z})$ through the $|_k$ action. The following congruence subgroups are the most important for understanding moonshine.

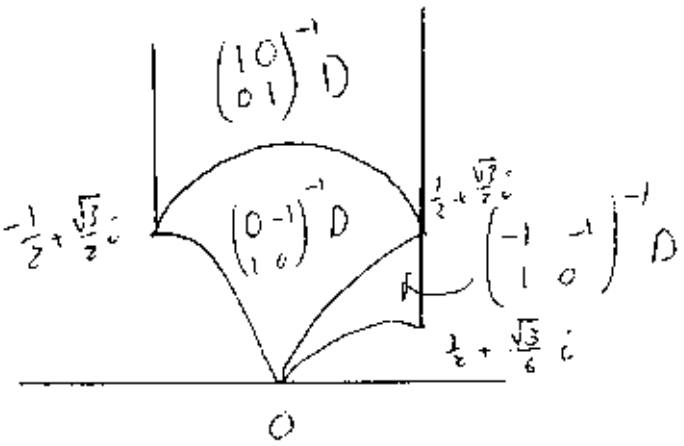
DEFINITION 2.5. Let N be a positive integer. The following are known as congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$, of level N:

$$\Gamma_0(N) := \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \mod N\}$$
$$\Gamma_1(N) := \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \mod N, \quad a \equiv d \equiv 1 \mod N\}$$
$$\Gamma(N) := \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : b \equiv c \equiv 0 \mod N, \quad a \equiv d \equiv 1 \mod N\}$$

By virtue of being subgroups of $\mathrm{SL}_2(\mathbb{Z})$, these act on the upper half plane via fractional linear transformations, and on the space of functions on the upper half plane by the $|_k$ operators.

Just as in the case for $\mathrm{SL}_2(\mathbb{Z})$, there exists a fundamental domain for this group action. Since all of these subgroups are of finite index, the fundamental domain can be chosen so that it is a finite union of translates (under $\mathrm{SL}_2(\mathbb{Z})$'s action) of the fundamental domain D. The fundamental domain for $\Gamma(2)$ is pictured in Figure 2. For more details, see Chapter III section 1 of Koblitz [**13**].

FIGURE 2. Fundamental Domain for $\Gamma_0(2)$



DEFINITION 2.6. Let Γ be a congruence subgroup. Γ acts on $H \cup \mathbb{Q} \cup \infty$ by viewing $\mathbb{Q} \cup \infty$ as $\mathbb{P}^1_{\mathbb{Q}} \subset \mathbb{P}^1_{\mathbb{C}}$. In particular, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} [r, s] = [ar + bs, cr + ds]$. The orbits of $\mathbb{Q} \cup \infty$ under this action are called the cusps of Γ.

If Γ is a congruence subgroup of level 1 so $\Gamma = \mathrm{SL}_2(\mathbb{Z})$), there is only one cusp which is customarily represented by ∞. For any pair of relatively prime integers p, q, there exist $r, s \in \mathbb{Z}$ such that $rp + sq = 1$. Then the matrix $\begin{pmatrix} r & s \\ -p & q \end{pmatrix}$ has determinant 1 and sends $\frac{q}{p}$ to ∞. Thus every rational number is equivalent to ∞, so there is only one cusp.

The cusps for $\Gamma_0(p)$, with p a prime, are also important to understand. There are two equivalence classes, represented by $[1, 0] = \infty$ and $[0, 1] = 0$. A matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(p)$ sends

the point 0 to $\frac{b}{d}$. Given relatively prime b and d, if d is not divisible by p then there is a solution in integers to $dx - bpy = 1$. Thus 0 is in the same cusp as all rational numbers whose denominator is not a multiple of p. A matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ sends ∞ to $\frac{a}{c}$, whose denominator is a multiple of p. Any such relatively prime a, c can be obtained, for again there are solutions to $ax - cy = 1$. These two equivalence classes are disjoint, for $\frac{r}{s}$ is sent to a fraction with denominator $cr + ds$: if s is a multiple of p, then since $c \equiv 0 \pmod{p}$ the new denominator is a multiple of p as well. This argument generalizes to $\Gamma_0(N)$. The number of cusps, along with representatives, are found in Section 2.4 of Iwaniec [**12**].

PROPOSITION 2.7. *The number of cusps of* $\Gamma_0(N)$ *is* $\sum_{d|N} \varphi((d, \frac{N}{d}))$ *where* φ *is Euler's totient function.*

Looking at Figure 2, the origin of the word cusp becomes clearer. 0 is visibly a cusp of the boundary of the fundamental domain. If we picked a different fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ that approached the real axis instead of ∞, a single cusp would appear there as well. We now turn to the definition of modular forms.

DEFINITION 2.8. Let Γ be a congruence subgroup of level N. A weakly meromorphic modular form of weight k is a meromorphic function on H invariant under Γ acting by $|_k$.

EXAMPLE 2.9. The matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ lies in $\Gamma_0(N)$, so any weakly meromorphic modular form must satisfy

$$f(z) = f(\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} z) = (-1)^k f(z).$$

This forces f to be the zero function when k is odd. There can be forms of odd weight for other congruence subgroups.

EXAMPLE 2.10. If $k = 2$, then then the functions fixed by the action of $\mathrm{SL}_2(\mathbb{Z})$ corresponds to the condition that

$$f(\gamma z) = (cz + d)^2 f(z).$$

If $f(z)dz$ is an differential form on H invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$, elementary calculus shows that f satisfies the same condition.

EXAMPLE 2.11. We will prove later in Remark 2.40 that the function

$$\Delta(z) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} \tag{5}$$

where $q = e^{2\pi i z}$ is a weakly meromorphic modular form of weight 12 (in fact, it is holomorphic). If we expand the product, one obtains a $q-$expansion for Δ:

$$\Delta(z) = q - 24q^2 + 252q^3 + \ldots = \sum_{n=1}^{\infty} \tau(n) q^n$$

where $\tau(n)$ is called the Ramanujan τ function. See Serre [**25**] Chapter 7 section 4.5.

The fact that Δ can be written as a power series in $q = e^{2\pi i z}$ is no accident. Note that $e^{2\pi i z}$ is a holomorphic function mapping the upper half plane onto the punctured disk $\Omega = \{q \in \mathbb{C}^\times : |q| < 1\}$. If $e^{2\pi i z} = e^{2\pi i z'}$, then $z - z' \in \mathbb{Z}$. Let f be a meromorphic

function like Δ that satisfies $f(z+1) = f(z)$. Then there is a meromorphic function $\tilde{f}$ on the punctured disk $\Omega = \{z \in \mathbb{C} : 0 < |z| < 1\}$ such that $f(z) = \tilde{f}(e^{2\pi iz})$. The point ∞ corresponds to origin. The Laurent expansion (if it exists) for $\tilde{f}$ gives the $q-$series expansion (the Fourier expansion) for f. For f to be meromorphic or holomorphic at the cusp infinity means that $\tilde{f}$ is meromorphic or holomorphic at 0. (This is exactly the definition of being meromorphic or holomorphic obtained by viewing $\mathrm{SL}_2(\mathbb{Z})\backslash H \cup \{\infty\}$ as a Riemann surface. However, it requires care to make this precise, especially for other congruence subgroups, so this remains motivation only. This approach is taken in Shimura [**26**].)

A general congruence subgroup Γ of level N will not necessarily contain $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, so a weakly meromorphic modular form may not satisfy $f(z+1) = f(z)$. However, it will contain $\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}$, in which case it satisfies $f(z+N) = f(z)$ and hence has a Fourier expansion in powers of $q_N := e^{2\pi iz/N}$. However, we will only consider those cases when f has an integral Fourier expansion. Furthermore, ∞ need not be the only cusp. To talk about the other cusps, one can simply transform them to ∞ using an element of $\mathrm{SL}_2(\mathbb{Z})$ not in Γ and proceed as before.

There are several important types of modular forms:

Definition 2.12. Let Γ be a congruence subgroup of level N and k a positive integer.

(1) A meromorphic modular form of weight k for Γ is a weakly meromorphic modular form of weight k such that for each cusp and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ that sends ∞ to the cusp, the function $f|_k\gamma$ is meromorphic at infinity.
(2) Holomorphic modular forms are defined to be meromorphic modular forms holomorphic on the upper half plane and at the cusps. The term modular form without qualification usually refers to holomorphic modular forms. The vector space of modular forms of weight k for Γ is denoted by $M_k(\Gamma)$.
(3) Cusp forms are modular forms that vanish at all cusps and are denoted by $S_k(\Gamma)$.

Remark 2.13. Showing this concept is well defined, independent of the choice of γ is straightforward and contained in Koblitz [**13**] Chapter III Section 3.

The most basic case occurs when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. A modular form of weight k then is simply a holomorphic function on H that satisfies the two transformational laws

$$f(z+1) = f(z) \quad \text{and} \quad f(-\frac{1}{z}) = z^k f(z) \tag{6}$$

By Theorem 2.2, these two conditions suffice to check that f is invariant under $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, the function must be holomorphic at ∞, the only cusp. Since $\mathrm{SL}_2(\mathbb{Z})$ contains the matrix $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, k must be even.

Being a modular form for $\Gamma_1(N)$ is often too weak a property. It is more interesting to look at the subspace of $M_k(\Gamma_1(N))$ that transform in a constrained way with respect to $\Gamma_0(N)$.

Definition 2.14. Let N be a positive integer and χ a Dirichlet character modulo N. A modular form f for $\Gamma_1(N)$ is in $M_k(\Gamma_0(N), \chi)$ if $f|_k\gamma = \chi(d)f$ where $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$. $S_k(\Gamma_0(N), \chi)$ is defined in the same way for cusp forms.

1.3. Modular Forms as Functions on Lattices. It many cases, especially when discussing Hecke operators, it is profitable and more natural to think of modular forms as functions on lattices. This is approach is developed in Serre [**25**] and more generally in Chapter 3, Section 5 of Koblitz [**13**]. A lattice Λ of $\mathbb{C}$ is simply the $\mathbb{Z}-$linear span of two linearly independent complex numbers ω_1 and ω_2. By convention, we always order the generators so $\mathrm{Im}(\omega_1/\omega_2) > 0$. The lattice spanned by ω_1' and ω_2' equals Λ if and only if ω_1' and ω_2' are a $\mathbb{Z}$-linear combination of ω_1 and ω_2 and vice verse. Thus the set of lattices corresponds to pairs (ω_1, ω_2) such that $\mathrm{Im}(\omega_1/\omega_2) > 0$ modulo the action of $\mathrm{SL}_2(\mathbb{Z})$. If we further identify Λ and $z\Lambda$ for $z \in \mathbb{C}^\times$, the map $(\omega_1, \omega_2) \to \tau = \omega_1/\omega_2$ gives a bijection between equivalence classes of lattices and $\tau \in H$ modulo the action of $\mathrm{SL}_2(\mathbb{Z})$ by fractional linear transformation. Thus classes of lattices are in bijection with the fundamental domain D for $\mathrm{SL}_2(\mathbb{Z})$.

For each type of congruence subgroup, there are corresponding modular points.

- For $\mathrm{SL}_2(\mathbb{Z})$, a modular point is a lattice.
- For $\Gamma_0(N)$, a modular point is a pair (Λ, C) where Λ is a lattice and C is a cyclic subgroup of order N inside $\mathbb{C}/\Lambda$.
- For $\Gamma_1(N)$, a modular point is a pair (Λ, t) where Λ is a lattice and t is a point of order N in $\mathbb{C}/\Lambda$.

If P is a modular point and $c \in \mathbb{C}^\times$, then $c \cdot P$ is the modular point obtained by scaling the lattice, subgroup, and vector by c. For example, for $\Gamma_0(N)$, $cP = c(\Lambda, C) = (c\Lambda, c \cdot C)$.

Given a column vector $\omega = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \in \mathbb{C}^2$ with ω_1/ω_2 in the upper half plane, the natural way to associate a modular point is as follows:

- For $\mathrm{SL}_2(\mathbb{Z})$, the modular point is $P_\omega := \Lambda_\omega = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$.
- For $\Gamma_0(N)$, the modular point is $P_\omega := (\Lambda_\omega, \mathbb{Z}\omega_2/N)$.
- For $\Gamma_1(N)$, the modular point is $P_\omega := (\Lambda_\omega, \omega_2/N)$.

A complex valued function F defined on modular points is of weight k if

$$F(c\Lambda) = c^{-k}F(\Lambda).$$

Given such a function, one can define $\tilde{F}$, a function on $\mathbb{C}^2$, by setting $\tilde{F}(\omega) := F(P_\omega)$. Furthermore, one can obtain a function on the upper half plane by taking $f(z) := \tilde{F}(\begin{pmatrix} z \\ 1 \end{pmatrix})$.

PROPOSITION 2.15. *Let $k \in \mathbb{Z}$ and Γ be a congruence subgroup of level N. The association of F to $\tilde{F}$ to f gives a one-to-one correspondence between the following sets of functions:*

- *Functions on modular points of weight k.*
- *Functions on column vectors which are invariant under the action of Γ and satisfy $\tilde{F}(\lambda\omega) = \lambda^{-k}\tilde{F}(\omega)$.*
- *Functions on the upper half plane invariant under the action of Γ through $|_k$.*

PROOF. Each type of modular point needs to be done separately. For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, this is more or less immediate. A function on lattices is simply a function on pairs of vectors that serve as a basis for the lattice. The conditions on $\tilde{F}$ listed are precisely the conditions to be a weight k function on the set of lattices. Identifying lattices to the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$ gives the third identification. More details and the other cases are in Koblitz [**13**], Chapter 3 Proposition 31. □

EXAMPLE 2.16. For a lattice Λ and an even integer $k > 2$,

$$\sum_{0 \neq z \in \Lambda} z^{-k}$$

is an absolutely convergent series and is easily seen to be a modular function on lattices of weight k. Taking Λ to be the lattice spanned by 1 and z gives the definition of Eisenstein series, a standard example of modular forms of weight k for $\mathrm{SL}_2(\mathbb{Z})$.

Thus modular forms can also be viewed as functions on lattices (equivalently, as functions on the space of elliptic curves). The lattice interpretations will make the definition of the Hecke operators more natural.

1.4. Dimension Calculations. The space of modular forms of weight k forms a vector space over $\mathbb{C}$. It will be useful to understand the dimension of this vector space.

The simplest examples are for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. If there is a modular form f_0 of weight k that is not a cusp form, any modular form g can be written as a linear combination of f_0 and a cusp form, for $g - \frac{g(\infty)}{f(\infty)} f$ has $q-$expansion with no constant term. Thus the dimension of $M_k(\mathrm{SL}_2(\mathbb{Z}))$ is one more than the dimension of $S_k(\mathrm{SL}_2(\mathbb{Z}))$.

The transformation properties of a meromorphic modular form give considerable information about its zeroes and poles. Denote the order of vanishing of f at a point p by $v_p(f)$. If $p = \infty$, define it to be the order of vanishing of $\tilde{f}$ at the origin. Let w_p be a weighting factor that is 1 except when $p = \rho = \frac{-1+\sqrt{-3}}{2}$ or $p = i$, in which case it is 3 or 2 respectively. (The factor w_p arises based on the extra elements of $\mathrm{SL}_2(\mathbb{Z})$ that fix the points ρ and i.)

THEOREM 2.17. *If f is a meromorphic modular form of weight k, not identically zero, then*

$$v_\infty(f) + \sum_{p \in D} \frac{v_p(f)}{w_p} = \frac{k}{12} \tag{7}$$

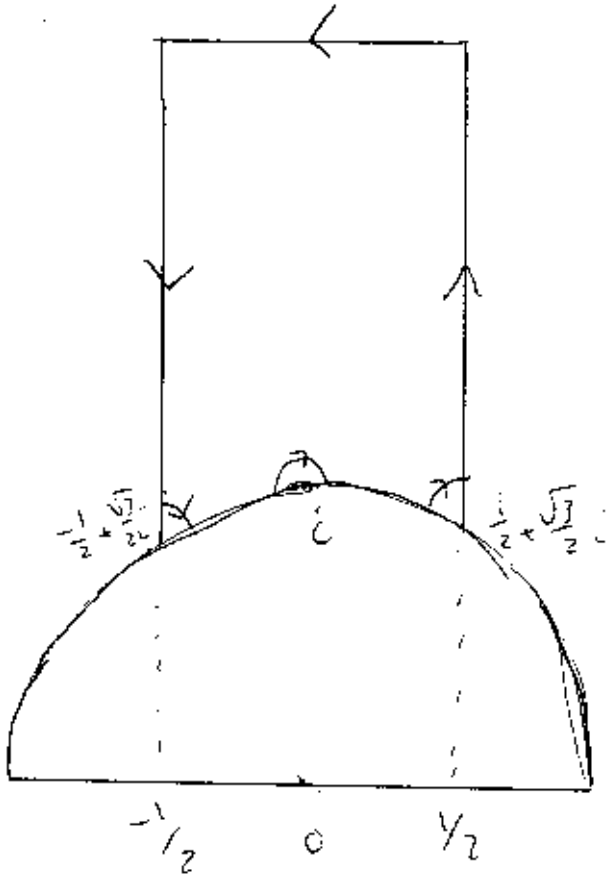
PROOF. The full details are contained in Serre [**25**] Chapter 7 Section 3. The idea is to use Cauchy's theorem from complex analysis, which relates the left hand side to a contour integral around the fundamental domain. More precisely,

$$\frac{1}{2\pi i} \int_C \frac{df}{f} = \sum_{p \in D, p \neq i, \rho} v_p(f)$$

where C is the contour pictured in Figure 3. Evaluating the contour integral along each piece of the boundary gives the desired result as the radius of the circular arcs goes to 0. If there are zeroes or poles on the boundary, modify the contour to go around them in such a way that the arcs match up under the action of $\mathrm{SL}_2(\mathbb{Z})$. □

For $\mathrm{SL}_2(\mathbb{Z})$, this theorem and elementary observations give enough information to completely determine dimensions of the space of modular forms. For example, if $k = 12$ then any cusp form must have a simple zero at ∞ and no others. $\Delta(z)$ has this property, so the ratio of any cusp form of weight 12 and Δ is a modular function (of weight 0), and hence a holomorphic function with no zeroes on the upper half plane union infinity. By basic complex analysis, it must be a constant function, and hence all cusp forms of weight 12 are multiples of $\Delta(z)$. The space of all modular forms of weight 12 is therefore 2 dimensional.

FIGURE 3. Contour of Integration for Theorem 2.17



For other congruence subgroups, dimension calculations are significantly more complicated. The following result, which can be obtained through evaluating an integral around the fundamental domain for general congruence subgroups, is the simplest generalization.

THEOREM 2.18. *If f is a nonzero meromorphic modular form of weight k for a congruence subgroup $\Gamma_0(N)$, and $X_0(N)$ is a fundamental domain for $\Gamma_0(N)$ (including the cusps), then*

$$\sum_{p \in X_0(N)} \frac{v_p(f)}{w_p} = \frac{\mu k}{12}. \tag{8}$$

Here μ is the index of $\Gamma_0(N)$ in $\Gamma_0(1)$, which equals $N \prod_{p|N} (1 + p^{-1})$.

PROOF. A proof is found in Rankin [**21**] as Theorem 4.14, which also fully explains the conventions on the weighting factor w_p. □

COROLLARY 2.19. *If $f, g \in M_k(\Gamma_0(N), \chi)$ and the first $\frac{\mu k}{12} + 1$ terms of the $q-$expansions agree, then $f = g$.*

PROOF. The difference vanishes to order at least $\frac{\mu k}{12}+1$ at ∞. Since $w_\infty = 1$, the equality in Theorem 2.18 cannot hold so $f - g$ is the 0 form. □

This corollary is crucial because it allows computations with modular forms to be reduced to finite computations in linear algebra. To check whether an equality holds between two expressions involving modular forms, it suffices to verify that enough coefficients of the $q-$expansions agree.

However, this method does not easily give exact formulas for the dimension of the space of modular forms for a congruence subgroup. An alternative approach is through understanding the upper half plane modulo $\Gamma_0(N)$ as a Riemann surface as is done in Shimura [**26**] and using the Riemann-Roch theorem. This computational result is taken from Ono [**19**].

DEFINITION 2.20. Suppose k is an integer and χ is a Dirichlet character modulo N for which $\chi(-1) = (-1)^k$. If $p|N$, let $r_p := v_p(N)$ and s_p the valuation of the conductor of χ.

Define the integer $\lambda(r_p, s_p, p)$ by

$$\lambda(r_p, s_p, p) := \begin{cases} p^{r'} + p^{r'-1} & \text{if } 2s_p \leq r_p = 2r' \\ 2p^{r'} & \text{if } 2s_p \leq r_p = 2r'+1 \\ 2p^{r_p - s_p} & \text{if } 2s_p > r_p \end{cases}$$

Define the rational numbers ν_k and μ_k by

$$\nu_k := \begin{cases} 0 & \text{if } k \text{ is odd} \\ \frac{-1}{4} & \text{if } k \equiv 2 \mod 4 \\ \frac{1}{4} & \text{if } k \equiv 0 \mod 4 \end{cases}$$

$$\mu_k := \begin{cases} 0 & \text{if } k \equiv 1 \mod 3 \\ -\frac{1}{3} & \text{if } k \equiv 1 \mod 3 \\ \frac{1}{3} & \text{if } k \equiv 0 \mod 3 \end{cases}$$

THEOREM 2.21. *Using the above notation and the convention that* $\prod_{p|N}$ *runs over prime divisors, and is* 1 *if the product is empty, it follows that*

$$\dim_{\mathbb{C}}(S_k(\Gamma_0(N), \chi)) - \dim_{\mathbb{C}}(M_{2-k}(\Gamma_0(N), \chi)) = \frac{(k-1)N}{12} \prod_{p|N}(1 + p^{-1})$$
$$-\frac{1}{2}\prod_{p|N} \lambda(r_p, s_p, p) + \nu_k \left(\sum_{\substack{x \in \mathbb{Z}/N\mathbb{Z} \\ x^2+1 \equiv 0 \mod N}} \chi(x) \right) + \mu_k \left(\sum_{\substack{x \in \mathbb{Z}/N\mathbb{Z} \\ x^2+x+1 \equiv 0 \mod N}} \chi(x) \right)$$

REMARK 2.22. If $k > 2$, then $M_{2-k}(\Gamma_0(N), \chi)$ is 0 dimensional. If $k < 0$, $S_k(\Gamma_0(N), \chi)$ is 0 dimensional. If $k = 2$, then $M_0(\Gamma_0(N), \chi)$ is 0 dimensional unless $\chi = 1$, in which case it is one dimensional: the only entire functions on the upper half plane are constants, and unless $\chi = 1$ the only constant satisfying the transformational law is 0.

2. Hecke Operators

2.1. Definition and Elementary Properties. The Hecke operators are a family of linear transformations that preserve spaces of modular forms. However, their definition makes more sense when defined as a transformation of modular points.[1] Let $\mathcal{L}$ denote the vector space of formal, finite linear combinations of modular points. We are most interested in the case $\Gamma = \Gamma_1(N)$, in which case a modular point is a lattice Λ and a point of order N in $\mathbb{C}/\Lambda$. The Hecke operators average a modular point over lattices of index n.

Definition 2.23. For a modular point (Λ, t) and positive integer n, define a linear transformation $T(n) : \mathcal{L} \to \mathcal{L}$ by

$$T(n)(e_{(\Lambda,t)}) = \frac{1}{n} \sum_{\Lambda'} e_{\Lambda',t} \tag{9}$$

where Λ' runs over all lattices of index n containing Λ such that t still has order N in $\mathbb{C}/\Lambda'$. This sum is finite, so the map is well defined, because any such lattice Λ' must lie in $\frac{1}{n}\Lambda$.

Furthermore, if n is relatively prime to N, define $S(n) : \mathcal{L} \to \mathcal{L}$ by

$$S(n)e_{(\Lambda,t)} = \frac{1}{n^2} e_{\frac{1}{n}\Lambda,t}. \tag{10}$$

The relatively prime condition in the second definition ensures t still has order N in $\mathbb{C}/\frac{1}{n}\Lambda$. It is relatively straightforward to understand the commutativity of these operators using this definition in terms of modular points.

Proposition 2.24. *For positive integers n_1, n_2, n and m:*

1. *$S(n_1)S(n_2) = S(n_1 n_2)$ and $S(n)T(m) = T(m)S(n)$.*
2. *If $(m, n) = 1$, then $T(m)T(n) = T(n)T(m) = T(mn)$.*
3. *If p is a prime factor of N, then $T(p^l) = (T(p))^l$*
4. *If p is relatively prime to N, then for $l \geq 2$*

$$T(p^l) = T(p^{l-1})T(p) - pT(p^{l-2})S(p)$$

Proof. The first statement is immediate. The remaining three all follow from understanding the subgroup structure of Abelian groups. In the second, we need to show $T(nm)e_{(\Lambda,t)} = T(m)T(n)e_{(\Lambda,t)}$ by understandings the subgroups of $\frac{1}{mn}\Lambda/\Lambda$ of size mn. Such a subgroup corresponds to a lattice Λ' containing Λ with index mn. In order for the order of t to still be N in $\mathbb{C}/\Lambda'$, the intersection with $\mathbb{Z}t \subset \mathbb{C}/\Lambda$ must be trivial. Since n and m are relatively prime, for every lattice Λ' there is an intermediate lattice Λ'' between Λ' and Λ of indexes n and m respectively. This corresponds to an intermediate subgroup $S'' \subset S' \subset \frac{1}{mn}\Lambda/\Lambda$. Conversely, suppose $S'' = \Lambda''/\Lambda \subset \frac{1}{n}\Lambda/\Lambda$ of order n and $S' = \Lambda'/\Lambda'' \subset \frac{1}{m}\Lambda''/\Lambda''$ is a subgroup of order m. If both have trivial intersection with $\mathbb{Z}t$, then Λ'/Λ is a subgroup of order mn with trivial intersection with $\mathbb{Z}t$. Thus the modular points occurring in $T(mn)(e_{(\Lambda,t)})$ are the same as those appearing in $T(m)T(n)(e_{(\Lambda,t)})$.

The thirds and fourth are similar: they are found as Proposition 32 of Chapter 3 of Koblitz [**13**]. □

Sometime it is useful to have an explicit description of the lattices Λ' that contain Λ with index n.

[1] For an alternate treatment using double cosets, see Chapter 6 of Iwaniec [**12**].

LEMMA 2.25. *Integer matrices of the form* $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ *with* $ad = n$, $a \geq 1$, *and* $0 \leq b < d$ *are in bijection with lattices* Λ' *contained in* Λ *with index* n. *If* ω_1 *and* ω_2 *are a basis for the lattice, then the correspondence sends*

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \to \left(\frac{a}{n}\omega_1 + \frac{b}{n}\omega_2\right)\mathbb{Z} + \frac{d}{n}\omega_2\mathbb{Z}.$$

PROOF. This is an elementary fact about lattices. It is proven in Section 7.5 of Serre [**25**]. □

There is a third type of linear transformation of $\mathcal{L}$. For d relatively prime to N, define the map $[d]$ to send $e_{(\Lambda,t)}$ to $e_{(\Lambda,dt)}$. Since d is relatively prime to N the order of dt is N in $\mathbb{C}/\Lambda$. Furthermore, the map only depends on d modulo N. Note that $[d]$ commutes with $T(n)$ and $S(m)$.

The maps $T(n)$, $S(m)$, and $[d]$ are maps from $\mathcal{L} \to \mathcal{L}$: they also give linear maps on the vector space of complex valued functions on modular points. They preserve modular forms.

PROPOSITION 2.26. *With the notation as above,* $[d]$, $T(n)$, *and* $S(m)$ *preserve* $M_k(\Gamma_1(N)$ *and* $S_k(\Gamma_1(N))$. *Furthermore, if* χ *is a Dirichlet character modulo* N, *then* $f \in M_k(\Gamma_1(N))$ *is in* $M_k(\Gamma_0(N), \chi)$ *if and only if* $[d]F = \chi(d)F$.

PROOF. The transformation laws for modular forms correspond to the corresponding functions on modular points being of weight k: that $[d]$, $T(n)$, and $S(m)$ preserve the weight is simple. Being holomorphic, and the behavior at the cusps, is more involved, and included in the proof of Proposition 33 in Chapter 3 of Koblitz [**13**]. □

Furthermore, since $[d]$ commutes with $T(n)$ and $S(m)$, these two operators preserve the eignenspaces of $[d]$, which correspond to $M_k(\Gamma_0(N), \chi)$ (likewise $S_k(\Gamma_0(N), \chi)$) for a Dirichlet character $\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times$. The eigenspace has associated eigenvalue $\chi(d)$.

COROLLARY 2.27. *The operators* $T(n)$ *and* $S(m)$ *preserve the spaces* $M_k(\Gamma_0(N), \chi)$ *and* $S_k(\Gamma_0(N), \chi)$.

2.2. Hecke Operators in Terms of q-Expansions. Finally, we can finally look at the Hecke operators in terms of the q expansion of a modular form. This is sometimes used to defining them, but is unmotivated.

PROPOSITION 2.28. *Let* $f(z) \in M_k(\Gamma_0(N), \chi)$, *with* $f(z) = \sum_{n=0}^{\infty} a_n q^n$. *If* p *is prime, then using the convention* $\chi(p) = 0$ *if* $p|N$ *and* $a_{n/p} = 0$ *if* $p \not| n$,

$$T(p)(f) = \sum_{n=0}^{\infty} (a_{pn} + \chi(p)p^{k-1}a_{n/p})q^n$$

More generally, it follows that

$$T(m)(f) = \sum_{n=0}^{\infty} \Big(\sum_{d|(m,n)} \chi(d)d^{k-1}a_{mn/d^2} \Big) q^n$$

PROOF. In the case that p is a prime, by definition $T(p)f(z) = \dfrac{1}{p}\displaystyle\sum_{\Lambda'} F(\Lambda', \frac{1}{N})$. The lattices Λ' are those containing $\Lambda_z = \mathbb{Z}z + \mathbb{Z}1$ with index p such that $\frac{1}{N}$ has order N modulo

Λ'. All such lattices contain $\frac{1}{p}\Lambda_z$ which is spanned by $\frac{z}{p}$ and $\frac{1}{p}$. Any such Λ' must be generated by Λ_z along with a point $\frac{a_1 z + a_2}{p}$, where a_1 and a_2 matter only modulo p. In particular, there is a distinct lattice Λ' for each of the $p+1$ pairs $[a_1, a_2]$ in the one dimensional projective space over $\mathbb{F}_p$. If $p \not| N$, all of the lattices still have $\frac{1}{N}$ as a point of order N. The lattice corresponding to $[1, j]$ is $L_{\frac{z+j}{p}}$. The lattice corresponding to $[0, 1]$ is $\frac{1}{p}L_{pz}$, as it is generated by L_z and $\frac{1}{p}$. In this case

$$T(p)f(z) = \frac{1}{p}\sum_{j=0}^{p-1} f(\frac{z+j}{p}) + \frac{1}{p}F(\frac{1}{p}L_{pz}, \frac{1}{N}).$$

But $\frac{1}{p}F(\frac{1}{p}L_{pz}, \frac{1}{N}) = p^{k-1}F(L_{pz}, \frac{p}{N})$. Using the operator $[p]$ and Proposition 2.26, this is $p^{k-1}\chi(p)f(pz)$. Thus we have that

$$T(p)f(z) = \frac{1}{p}\sum_{j=0}^{p-1} f(\frac{z+j}{p}) + \chi(p)p^{k-1}f(pz).$$

Collecting terms of the $q-$expansion, the coefficient of q^n is $a_{pn} + \chi(p)p^{k-1}a_{n/p}$.

If $p|N$, then the lattice corresponding to $[0, 1]$, generated by $\frac{1}{p}$ and L_z, must be thrown out. This removes the last term, but $\chi(p) = 0$ when $p|N$. Thus the same formula holds.

The proof for $T(n)$ is Proposition 39 of Koblitz [**13**], and follows from the case $T(p)$. □

We can now talk about Hecke eigenforms. Since $T(m)$ and $T(n)$ commute for relatively prime n and m, they preserve the other's eigenspaces. Thus it makes sense to look for modular forms that are eigenvectors for all of the $T(n)$.

Definition 2.29. A Hecke eigenform is a (nonzero) modular form $f \in M_k(\Gamma_0(N), \chi)$ that is an eigenvector for all of the $T(n)$, in other words for all n

$$T(n)f = \lambda_n f \tag{11}$$

for some $\lambda_n \in \mathbb{C}$. It is called normalized if the coefficient of q is 1.

In particular, any nonzero modular form in a one dimensional space of modular forms must be a Hecke eigenform, for the Hecke operators preserve that space of modular forms. For example, since $S_{12}(\Gamma_0(1), 1)$ is one dimensional, Δ must be a Hecke eigenform.

We have a large amount of information about coefficients of a $q-$expansion for Hecke eigenforms.

Proposition 2.30. *If $f(z) \in M_k(\Gamma_0(N), \chi)$ is a Hecke eigenform with eigenvalues λ_n, and $f(z) = \sum_{n=0}^{\infty} a_n q^n$, then $a_m = \lambda_m a_1$. In addition, $a_m a_n = a_{mn}$ for relatively prime m and n.*

Proof. Looking at the coefficient of q in $T(m)f$ using Proposition 2.28, it is a_m. On the other hand, the coefficient of q in $\lambda_m f$ is $\lambda_m a_1$. But since for $(n, m) = 1$, $T(n)T(m)f = T(mn)f$, it follows that $\lambda_n\lambda_m = \lambda_{nm}$ and hence the coefficients a_n are multiplicative functions. □

For example, this implies that because $\Delta(z) = \sum_{n=1}^{\infty} \tau(n)q^n$ has $\tau(1) = 1$, $\tau(n) = \lambda_n$ and hence τ is a multiplicative function. Applying Proposition 2.28 gives information about what happens when n and m are not relatively prime as well.

COROLLARY 2.31. *If $f(z) \in M_k(\Gamma_0(N), \chi)$ is a normalized Hecke eigenform of the form $f(z) = \sum_{n=0}^{\infty} a_n q^n$, then*

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

2.3. The Fricke Involution. Another useful transformation is the Fricke involution. It can be defined in terms of the action of a matrix in $\mathrm{GL}_2^+(\mathbb{Z})$ or in terms of modular points. This material is discussed in Section 6.7 of Iwaniec [**12**].

DEFINITION 2.32. Let N be an integer. The Fricke involution W_N acts on complex valued functions of weight k on the upper half plane via $\begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$. Explicitly,

$$W_N f = N^{-k/2} z^{-k} f\left(\frac{-1}{Nz}\right).$$

It is immediate that $W_N^2 = (-1)^k$, hence this is called an involution. Note that W_N exchanges the cusp at 0 and the cusp at ∞. Furthermore, W_N normalizes $\Gamma_0(N)$ in the sense that for $\gamma \in \Gamma_0(N)$, if

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{then} \quad \gamma' = W_N \gamma W_N^{-1} = \begin{pmatrix} d & -c/N \\ -bN & a \end{pmatrix}.$$

Then for $f \in S_K(\Gamma_0(N), \chi)$ and $\gamma \in \Gamma_0(N)$,

$$(f|_k W_N)|_k \gamma = f|_k \gamma' W_N = \chi(\gamma') f|_k W_N$$

where $\chi(\gamma') = \overline{\chi(\gamma)}$. Composing W_N with the conjugation operator defined by $f \to \overline{f}(-\overline{z})$ gives $\overline{W}_N$ which preserves the space of cusp forms.

Alternately, the Fricke involution can be described in terms of modular points for $\Gamma_1(N)$. Consider the map defined by

$$W_N : (\omega_1\mathbb{Z} + \omega_2\mathbb{Z}, \frac{a}{N}\omega_2) \to (\omega_1\mathbb{Z} + \frac{\omega_2}{N}\mathbb{Z}, \frac{a'}{N}\omega_1) \tag{12}$$

where $\omega_1/\omega_2 \in H$ and a' is the inverse of a modulo N.

LEMMA 2.33. *The map $W_N : \mathcal{L} \to \mathcal{L}$ defined above is well defined. If $F : \mathcal{L} \to \mathbb{C}$ is the function on modular points associated to $f \in S_k(\Gamma_0(N), \chi)$, $W_N f$ corresponds to $N^{-k/2} F \circ W_N$.*

PROOF. This is well defined since given a modular point (Λ, t) Nt is a multiple of a complex number ω_2 (unique up to sign) that can be used to form a $\mathbb{Z}$-basis for Λ. Then the other basis vector ω_1 is determined up to sign and addition of ω_2. It is clear that adding ω_2 doesn't effect the image, and that because ω_1/ω_2 is required to be in the upper half plane the choice of sign doesn't matter either.

Now take the modular point P_z with $\omega_1 = z \in H$, $\omega_2 = 1$, and $a = 1$. By definition, $f(z)$ is $F(P_z)$. $W_N(P_z)$ is defined to be $(z\mathbb{Z} + \frac{1}{N}\mathbb{Z}, \frac{z}{N})$. Factoring out a z,

$$F(W_N(P_z)) = z^{-k} F(\mathbb{Z} + \frac{1}{Nz}\mathbb{Z}, \frac{1}{N})$$

The lattice is spanned by 1 and $\frac{-1}{Nz}$, with the negative sign introduced to make $\frac{-1}{Nz}$ in the upper half plane. But by definition

$$F(\mathbb{Z}+\frac{-1}{Nz}\mathbb{Z},\frac{1}{N})=F(P_{\frac{-1}{Nz}})=f(\frac{-1}{Nz})$$

which agrees with Definition 2.32. □

REMARK 2.34. W_N commutes with scalar multiplication, while for $f\in S_k(\Gamma_0(N),\chi)$ $F([d](\Lambda,t))=\chi(d)F(\Lambda,t)$. Thus since W_N sends $S_K(\Gamma_0(N),\chi)$ to $S_k(\Gamma_0(N),\overline{\chi})$, it follows that

$$W_NF((\omega_1\mathbb{Z}+\omega_2\mathbb{Z},\frac{a}{N}\omega_2))=\chi(a)F(W_N(\omega_1\mathbb{Z}+\omega_2\mathbb{Z},\frac{1}{N}\omega_2))=F(\omega_1\mathbb{Z}+\frac{1}{N}\omega_2\mathbb{Z},\frac{a'}{N})$$

which explains the strange inverse modulo N in the definition of W_N.

This alternate description makes it clear that the Fricke involution almost commutes with the Hecke operators.

THEOREM 2.35. *A Hecke eigenform $f\in S_k(\Gamma_0(N),\chi)$ is also an eigenfunction of the involution operator W_N.*

PROOF. Using slightly more of the theory of modular forms than has been discussed previously, this assertion can be reduced to checking that $W_NT_n=\chi(n)T_nW_N$ for n relatively prime to N on the space of cusp forms $S_k(\Gamma_0(N),\chi)$. By Lemma 6.25 of Iwaniec [**12**], if two Hecke eigenforms have the same eigenvalues for all T_n with $(n,N)=1$, then one is a multiple of the other. Then if $\lambda(n)$ are the eigenvalues of f with respect to T_n we have

$$T_n\overline{W_N}f=\chi(n)\overline{W_N}T_nf=\chi(n)\overline{\lambda(n)}\overline{W_N}f=\lambda(n)\overline{W_N}f$$

using the fact that the conjugation operation commutes with the Hecke operators and that $\chi(n)\overline{\lambda(n)}$ is the eigenvalue $\lambda(n)$.

To prove that $W_NT_n=\chi(n)T_nW_N$ for n relatively prime to N, we use Lemma 2.33 and Definition 2.23 and the language of modular points. Let $P=(\Lambda,t)=(\omega_1\mathbb{Z}+\omega_2\mathbb{Z},\frac{s}{N}\omega_2)$, and F be the function on modular points associated to a modular form $f\in S_K(\Gamma_0(N),\chi)$. Since n and N are relatively prime, t will still be of order N in $\mathbb{C}/\Lambda'$ for any lattice Λ' that contains Λ with index n. Thus we have that

$$nW_NT_n(F(P))=W_N\left(\sum_{\Lambda'}F(\Lambda',\frac{s}{N}\omega_2)\right)$$

where Λ' runs over lattices containing Λ with index N. By Lemma 2.25, all such lattices are of the form $\left(\frac{a}{n}\omega_1+\frac{b}{n}\omega_2\right)\mathbb{Z}+\frac{d}{n}\omega_2\mathbb{Z}$ for $\begin{pmatrix}a&b\\0&d\end{pmatrix}$ an integer matrix such that $ad=n$, $a\geq 1$, and $0\leq b<d$. Thus this equals

$$\chi(\frac{n}{d})\sum_{a,b,d}F\left((\frac{a}{n}\omega_1+\frac{b}{n}\omega_2)\mathbb{Z}+\frac{d}{Nn}\omega_2\mathbb{Z},\frac{s'}{N}(\frac{a}{n}\omega_1+\frac{b}{n}\omega_2)\right)$$

Pulling the $\overline{\chi}(d)$ inside using $\overline{\chi}(d)F(\Lambda,t)=F(\Lambda,dt)$ for $F\in S_k(\Gamma_0(N),\overline{\chi})$ gives

$$\chi(n)\sum_{a,b,d}F\left((\frac{a}{n}\omega_1+\frac{b}{n}\omega_2)\mathbb{Z}+\frac{d}{Nn}\omega_2\mathbb{Z},\frac{s'}{N}(\frac{ad}{n}\omega_1+\frac{bd}{n}\omega_2)\right)$$

On the other hand, applying W_n first gives

$$nT_nW_NF(P) = nT_nF(\omega_1\mathbb{Z} + \frac{\omega_2}{N}\mathbb{Z}, \frac{s'\omega_1}{N}) = \sum_{a,b',d}\left(F(\frac{a}{n}\omega_1 + \frac{b'}{Nn}\omega_2)\mathbb{Z} + \frac{d}{nN}\omega_2\mathbb{Z}, \frac{s'\omega_1}{N}\right)$$

Identifying b with the b' such that $bN = b' \mod d$ identifies the lattices in the two sums. Since $ad = n$ and $\frac{bds'}{nN}\omega_2$ is in the lattice, the two modular points agree as well. Thus $\chi(n)T_nW_N = W_NT_n$ which completes the proof. □

The Fricke involution is useful because it exchanges the cusp at infinity and the cusp represented by 0. It will explain a strange symmetry of the group M_{24}.

3. Modular Forms with Complex Multiplication

In general it requires special care to show that any particular modular form is a Hecke eigenform. However, for a class of modular forms that arise from imaginary quadratic fields, called modular forms with complex multiplication (CM) it is easy to tell when they are Hecke eigenforms.

Let $K = \mathbb{Q}\left(\sqrt{-D}\right)$ be an imaginary quadratic field with discriminant $-D$. Let $\mathcal{O}_K$ be its ring of integers, $\mathfrak{m}$ a nontrivial ideal and $I_\mathfrak{m}$ be the group of fractional ideals relatively prime to $\mathfrak{m}$.

Definition 2.36. A Hecke character is a group homomorphism $\phi : I_\mathfrak{m} \to \mathbb{C}^\times$ such that for all $\alpha \in K^\times$ with $\alpha \equiv 1 \pmod{\mathfrak{m}}$, ϕ satisfies

$$\phi(\alpha\mathcal{O}_K) = \alpha^{k-1},$$

for some $k \in \mathbb{Z}$ with $k \geq 2$.

Define a Dirichlet character χ_ϕ for n relatively prime to $\mathfrak{m}$ by

$$\chi_\phi(n) := \phi((n))/n^{k-1}$$

which has modulus $N\mathfrak{m}$.

Theorem 2.37. *With the notation above, define*

$$\Phi(z) := \sum_\mathfrak{a} \phi(\mathfrak{a})q^{N(\mathfrak{a})} = \sum_{n=1}^{\infty} a(n)q^n, \tag{13}$$

where $N(\mathfrak{a})$ denotes the norm of the ideal $\mathfrak{a}$ and the first sum is over integral ideals $\mathfrak{a} \subset \mathcal{O}_K$ that are prime to $\mathfrak{m}$. Then $\Phi(z)$ is a cusp form in $S_k(\Gamma_0(|D| \cdot N(\mathfrak{m})), \left(\frac{-D}{\cdot}\right)\chi_\phi)$. If ϕ is a primitive character, then f is a Hecke eigenform.

Proof. This is part of Section 12.3 of Iwaniec [**12**]. □

To define a Hecke character, it suffices to define it on a minimal set of generators for the class group. Assume ± 1 are the only units in $\mathbb{Q}\left(\sqrt{-D}\right)$. Let $\pi_1, \ldots, \pi_m$ be a minimal set of ideals whose ideal classes generate the class group. Let the order of π_i in the class group be n_i. Given $\mathfrak{m}$, we must have $\phi((\alpha)) = \alpha^{k-1}$ if $\alpha \equiv 1 \pmod{\mathfrak{m}}$. Let χ be a character on $\mathcal{O}_K/\mathfrak{m}$ extended to $\mathcal{O}_K$ that satisfies $\chi(-1) = (-1)^{k-1}$. First define ϕ on principal ideals by $\phi((\alpha)) = \chi(\alpha)\alpha^{k-1}$. Since the only units in $\mathcal{O}_K$ are ± 1, this definition is independent of the choice of generator, α. To extend ϕ to non-principal ideals, and thus obtain a Hecke character, it suffices to define it on $\pi_1, \ldots, \pi_m$ and extend multiplicatively. By the above

assumptions, π_i is non-principal and $\pi_i^{n_i} = (\alpha)$ for some $\alpha \in K^\times$. Thus $\phi(\pi_i)$ must be one of the n_i^{th} roots of $\phi((\alpha)) = \alpha^{k-1}\chi(\alpha)$. Fixing $\phi(\pi_i)$ for each i gives the Hecke character.

Modular forms with complex multiplication are easy to work with, as the coefficients of their $q-$expansion have explicit descriptions. Furthermore, one simple way to show a modular form is a Hecke eigenform is to show it has complex multiplication.

4. Eta Products

4.1. The Eta Function and its Transformation Laws. The Dedekind eta function, although not a modular form in the sense defined above, is still very important in constructing examples of modular forms.

DEFINITION 2.38. For $z \in H$, let $q = e^{2\pi i z}$. Define the Dedekind eta function by

$$\eta(z) := q^{1/24}\prod_{n=1}^{\infty}(1-q^n). \tag{14}$$

Like modular forms, the Dedekind eta function satisfies simple transformational laws. In the following, use $\sqrt{\ }$ to denote the branch of the square root function with a branch cut along the negative real axis that is positive for the positive reals.

THEOREM 2.39. *Let* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ *with* $c > 0$. *The Dedekind eta function satisfies*

$$\begin{aligned} \eta(z+1) &= e^{\pi i z/12}\eta(z) \\ \eta(-1/z) &= \sqrt{z/i}\eta(z) \\ \eta(\gamma\cdot z) &= \epsilon(\gamma)\sqrt{-i(cz+d)}\eta(z) \end{aligned}$$

where

$$\epsilon(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \exp\left(-i\pi\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix})\right)$$

and, if $(a,6)=1$, *the number* $\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix})$ *satisfies*

$$\alpha(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) \equiv \frac{1}{12}a(c-b-3) - \frac{1}{2}(1-\left(\frac{c}{a}\right)) \mod 2$$

PROOF. Note that $\left(\frac{c}{a}\right)$ in the last line is a Legendre symbol. The third statement includes the first two. A proof of the general transformation rule is given after Theorem 3.4 of Apostol [**1**]. This precise formulation comes from Ligozat [**15**], Section 3.1. □

REMARK 2.40. The function $\Delta(z)$ defined in Example 2.11 is just the 24th power of $\eta(z)$. Raising the first and second transformation laws to the 24th power shows that $\Delta(z)$ satisfies the transformation laws for a weight 12 modular form for the elements S and T which generate $\mathrm{SL}_2(\mathbb{Z})$.

Moonshine for M_{24} involves products of eta functions. Given a function $r : \mathbb{Z} \to \mathbb{Z}$ with r identically zero outside a finite set, consider the product

$$f_r(z) := \prod_{d\in\mathbb{Z}}\eta(dz)^{r(d)} \tag{15}$$

We are interested in when this will be a well-behaved modular form.

THEOREM 2.41. *Let N be a positive integer. For d a divisor of N, denote N/d by d'. Let r be a function taking on non-negative values on divisors of N and 0 elsewhere. Suppose $k = \frac{1}{2}\sum_{d|N} r(d) \in \mathbb{Z}$. If*

$$\sum_{d|N} r(d)d' \equiv 0 \mod 24 \quad \text{and} \quad \sum_{d|N} r(d)d \equiv 0 \mod 24 \tag{16}$$

then $f(z) := \prod_{d|N} \eta(dz)^{r(d)}$ satisfies

$$f(\begin{pmatrix} a & b \\ c & d \end{pmatrix} z) = \chi(d)(cz+d)^k f(z)$$

for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, where χ is a quadratic (possibly trivial) Dirichlet character.

REMARK 2.42. There is an analogous result where r takes on integer values, resulting in a quotient of eta functions.

REMARK 2.43. The two conditions (16) are saying that $f(z)$ and $W_N f(z)$ have integral Fourier expansions as the $q^{\frac{1}{24}}$ in the eta product have combined. These conditions are certainly necessary for f to be in $M_k(\Gamma_0(N), \chi)$.

PROOF. The proof generalizes proposition 3.1.1 in Ligozat [**15**]. Let $\gamma = \begin{pmatrix} a & b \\ Nc & d \end{pmatrix}$ be an element of $\Gamma_0(N)$. Then

$$\eta(\delta U z) = \eta(U_\delta \delta z) \quad \text{where} \quad U_\delta = \begin{pmatrix} a & b\delta \\ c\delta' & d \end{pmatrix}$$

by the definition of the action on H. Thus

$$f(Uz) = (-i(Ncz+d))^{\frac{1}{2}\sum_{\delta|N} r_\delta} \cdot f(z) \cdot \prod_{\delta|N} \epsilon(U_\delta)^{r_\delta} \tag{17}$$

using the transformation law in Theorem 2.39. To calculate the third factor, we can assume that $(a, 6) = 1$ and $c > 0$. Because such matrices generate $\Gamma_0(N)$ (Ligozat [**15**] section 3.1), this suffices. However,

$$\prod_{\delta|N} \epsilon(U_\delta)^{r(\delta)} = \exp(-i\pi\lambda)$$

where λ is the sum $\sum_{\delta|N} r(\delta)\alpha(U_\delta)$. By Theorem 2.39,

$$\alpha(U_\delta) \equiv \frac{a}{12}((c\delta' - b\delta - 3)) - \frac{1}{2}(1 - \left(\frac{c\delta'}{a}\right)) \mod 2$$

and hence

$$\lambda \equiv \frac{ac}{12}(\sum_{\delta|N} r(\delta)\delta') - \frac{ab}{12}(\sum_{\delta|N} r(\delta)\delta) - \frac{a}{4}\sum_{\delta|N} r(\delta) - \frac{1}{2}\sum_{\delta|N}(1 - \left(\frac{c\delta'}{a}\right))r(\delta) \mod 2$$

By hypothesis, this simplifies as

$$\lambda \equiv 0 - 0 - \frac{ak}{2} - \frac{1}{2}\sum_{\delta|N} r(\delta)(1 - \left(\frac{c\delta'}{a}\right)) \mod 2.$$

Because $\exp\left(\frac{1}{2}\pi i(1 - \left(\frac{c\delta'}{a}\right))\right) = \left(\frac{c\delta'}{a}\right)$, it follows that

$$\prod_{\delta|N} \epsilon(U_\delta)^{r(\delta)} = \exp\left(-i\pi\lambda\right) = \exp\left(-i\pi ak/2\right)\left(\frac{c^{2k}}{a}\right)\prod_{\delta|N}\left(\frac{\delta'}{a}\right)^{r(\delta)}$$

The first term in 17 simplifies as $(-i)^k(Ncz+d)^k$, so

$$f(Uz) = \exp\left(-i\pi k/2\right)(Ncz+d)^k f(z)\exp(-ia\pi k/2)\prod_{\delta|N}\left(\frac{\delta'}{a}\right)^{r(\delta)}$$

But $\chi(d) := \exp(-\frac{i\pi k}{2}(1+a))\prod_{\delta|N}\left(\frac{\delta'}{a}\right)^{r(\delta)}$ is a quadratic character as required, as $ad = 1 \mod N$.

□

4.2. Eta Products that are Hecke Eigenforms. When looking at moonshine for the Mathieu groups, the interesting eta products are also Hecke eigenforms. The simplest way to ensure this is to have the eta product lie in a one dimensional space of cusp forms. Because the Hecke operators will preserve this space, any cusp form is automatically a Hecke eigenform. Mason uses this idea to find eta products of even weight that are Hecke eigenforms [**17**].

If a nonzero eta product is to be a cusp form and Hecke eigenform, it must vanish to order 1 at the cusp at infinity. If not, then the coefficient of q must be 0, in which case all the coefficients in the $q-$expansion will be 0 by Proposition 2.30. Since the Fricke involution preserves this one dimensional space of eigenforms, it must vanish to order 1 at the cusp 0 as well. In particular, this means that

$$\sum_{d|N} r(d)d = 24 \quad \text{and} \quad \sum_{d|N} r(d)d' = 24.$$

Let N be a positive integer, k the weight, and ϵ a quadratic character modulo N (possibly the trivial character). Setting

$$\mu := |\Gamma_0(1) : \Gamma_0(N)| = N\prod_{p|N}(1+p^{-1}),$$

Theorem 2.18 shows that

$$\sum_{z\in X_0(N)} \mathrm{ord}_z(f) = \frac{\mu k}{12}$$

Furthermore, we know there are $c(N) = \sum_{d|N}\phi(d, \frac{N}{d})$ cusps. Suppose f is a cusp form and the order of vanishing at each cusp is an integer as happens when ϵ is trivial and $k \geq 2$ is even [**17**]: then if

$$c(N) = \frac{\mu k}{12}$$

then f must vanish to order 1 at each cusp and vanish nowhere else. Hence the space of all cusp forms is one dimensional.

Given these conditions, it is straightforward to calculate all eta products that are Hecke eigenforms for this reason. They are listed in Table 1: the notation $1^2 11^2$ means the eta product $\eta(z)^2\eta(11z)^2$.

v

TABLE 1. Eta Products that are Hecke Eigenforms, even weight and trivial character

N	k	f
1	12	1^{24}
2	8	$1^8 2^8$
3	6	$1^6 3^6$
4	6	2^{12}
5	4	$1^4 5^4$
6	4	$1^2 2^2 3^2 6^2$
8	4	$2^4 4^4$
9	4	3^8

N	k	f
11	2	$1^2 11^2$
14	2	1 2 7 14
15	2	1 3 5 15
20	2	$2^2 10^2$
24	2	2 4 6 12
36	2	6^4
27	2	$3^2 9^2$
32	2	$4^2 8^2$

To illustrate, suppose $N = p^r$. Then the number of cusps is for $\Gamma_0(N)$ is

$$\sum_{n=0}^{r} \varphi(p^{\max(n,r-n)}) \leq 2(1 + p + \ldots + p^{[r/2]})$$

$$\leq 2(\frac{p^{[r/2]+1} - 1}{p-1}) \leq 4p^{[r/2]}$$

On the hand, $\mu = p^r(1 + p^{-1}) = p^r + p^{r-1}$ and k must be at least 2. But

$$4p^{[r/2]} < \frac{p^r + p^{r-1}}{6}$$

whenever $r \geq 10$ (and p is at least 2) or whenever $p \geq 24$ (and r is at least 1). Thus for equality to possibly hold $r < 10$ and $p < 24$. This leaves a finite number of cases to check.

For example, one case equality holds is when $p = 11$ and $r = 1$. By Theorem 17, $f(z) = \eta(z)^2\eta(11z)^2$ lies in $M_2(\Gamma_0(11), 1)$. $f(z)^{12} = \Delta(z)\Delta(11z)$, which is known to be in $S_{24}(\Gamma_0(11))$. Furthermore, it is manifestly nonzero on the upper half plane. At the cusp infinity, $f(z)^{12}$ has a $q-$expansion $q^{12}\prod_{n=1}^{\infty}(1-q^n)^{24}(1-q^{11n})^{24}$, so has a zero of order 12 at ∞. The action of $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ sends the cusp 0 to ∞. Note that

$$f(z)^{12}|_{24}S = z^{-24}\Delta(-11/z)\Delta(-1/z) = 11^{-12}\Delta(z)\Delta(z/11)$$

has a zero of order 12 at the cusp based on its q-expansion.

By the dimension formula, there is a nonzero element $g \in S_2(\Gamma_0(11), 1)$ with a zero of order 1 at each cusp. Then $(g/f)^{12}$ is a modular function for $\Gamma_0(11)$ which is nonzero on the upper half plane. Furthermore, it is non-vanishing at the two cusps, which implies that it is a nonzero entire function and hence constant. Thus f is a multiple of g, so $\eta(z)^2\eta(11z)^2$ is in $S_2(\Gamma_0(11))$ and hence is a Hecke eigenform.

Comparing Table 1 with Table 7 gives the first hint of the moonshine that will be discussed in the next chapter.

CHAPTER 3

Moonshine For M_{24}

This chapter will present the moonshine theory for the Mathieu group M_{24}, relating eta products and the cycle shapes of elements of M_{24}. By analogy with monstrous moonshine, there is an infinite dimensional graded module that explains an infinite family of virtual characters that are multiplicative functions of their index. These ideas go back to G. Mason's papers [**16**] and [**17**].

1. M_{24} and Hecke Eigenforms

The first hint of a connection between the Mathieu group M_{24} and modular forms appears in Tables 1 and 7. Mason noticed in [**17**] that for all but the last two modular forms in Table 1, the shape of the eta product appears as the cycle shapes of elements in M_{24}. Furthermore, the level and weight of the modular form can be predicted from the cycle shape.

Definition 3.1. For an element $g \in \mathrm{M}_{24}$, of cycle shape $1^{r(1)} \dots 24^{r(24)}$, define

$$f_g(z) = \prod_{d=1}^{24} \eta(dz)^{r(d)}.$$

Define $k(g)$ to be half of the number of cycles of g and $N(g)$ to be the product of the lengths of the longest and shortest cycles.

All of the elements of M_{24} with $k(g)$ even correspond to eta products in Table 1. The other conjugacy classes in M_{24} similarly correspond to modular forms.

Theorem 3.2. *For each $g \in \mathrm{M}_{24}$, $f_g(z)$ is a cusp form and a Hecke eigenform, with weight $k(g)$, level $N(g)$, and nebentypus character ϵ_g. If the weight $k(g)$ is even, $\epsilon_g = 1$, otherwise ϵ_g is a quadratic character.*

Proof. This is already done following Mason's argument for elements of even weight: they are cusp forms and Hecke eigenforms because the space of cusp forms the eta product land in are one dimensional. For the elements of odd weight, routine calculation with Theorem 2.41 again shows that they lie in $S_{k(g)}(\Gamma_0(N(g)), \epsilon_g)$, where ϵ_g is a quadratic character. To show that these are Hecke eigenforms, for $k(g) > 1$ the dimension formula Theorem 2.21 also shows that the space of cusp forms they lie in is one dimensional.

For $k = 1$, the spaces may no longer be one dimensional. However, they are still Hecke eigenforms. One strategy is to check directly, but a nicer approach taken in [**8**] is to notice that the three eta products $\eta(z)\eta(23z)$, $\eta(3z)\eta(21z)$, $\eta(12z)^2$ agree for a sufficient number of terms in their $q-$expansion with modular forms with complex multiplication, which are known to be Hecke eigenforms by Theorem 2.37. The specific number fields and information about the Hecke characters are indicated in Table 1. The exact number of coefficients to compare is given in Corollary 2.19.

□

TABLE 1. A few Hecke Characters giving eta products.

η-product	Number Field	$\mathfrak{m}$	Order of the Character
$\eta(z)\eta(23z)$	$\mathbb{Q}(\sqrt{-23})$	1	2
$\eta(3z)\eta(21z)$	$\mathbb{Q}(\sqrt{-7})$	(3)	4
$\eta(12z)^2$	$\mathbb{Q}(\sqrt{-1})$	(6)	4

REMARK 3.3. In Dummit, Kisilevsky, and McKay [**8**], all eta products that are also cusp forms and Hecke eigenforms are found. By the reasoning in Section 4.2, the order of vanishing at ∞ must be 1, which means that the in the product

$$\eta(z)^{r(1)}\eta(2z)^{r(2)}\dots\eta(tz)^{r(t)}$$

it must hold that $\sum_d d\cdot r(d) = 24$. A computer can find all of the partitions of 24 that do not correspond to multiplicative eta products, and then the 30 remaining can be proven to be eigenforms using the techniques presented here.

REMARK 3.4. As with all instances of moonshine, the reason for the appearance of the group M_{24} is murky. It may not even have anything to do with M_{24} being a sporadic simple group and connected to the monster: instead, it may be a result of the 24 dimensional representation of M_{24}. Voskresenskaya [**28**] shows that for all groups of order 24, the eta products associated to the cycle shapes from the regular representation are multiplicative eta products. On the other hand, Dummit, Isilevsky and McKay [**8**] showed that of the conjugacy classes of S_{24}, only 30 give multiplicative eta products.

2. Representations and Multiplicative eta products

The original observations about monstrous moonshine relating coefficients in the q-expansion of the $j-$function to the irreducible representations of the monster are generalized and explained (at least, partially explained) through the existence of an infinite dimensional graded module, as alluded to in the introduction. There is a similar representation lurking for M_{24}.

2.1. A Family of Multiplicative Virtual Characters. Expanding the eta products associated to elements of M_{24}, the coefficients give class functions for M_{24}.

DEFINITION 3.5. For $g \in M_{24}$, write

$$f_g(z) = \sum_{n=1}^{\infty} a_g(n)q^n.$$

For any fixed positive integer n, define

$$\gamma_n(g) := a_g(n)$$

It is clear that γ_n is a class function. Furthermore, because all of the eta products associated with elements of M_{24} are Hecke eigenforms, it is clear that

COROLLARY 3.6. *For positive integers n and m with $(n,m) = 1$, it is true that*

$$\gamma_n\gamma_m = \gamma_{mn}$$

In other words, the γ_n form a multiplicative family of class functions.

The γ_n cannot be characters for M_{24} since the dimension of a representation cannot be negative, and $\Delta(z) = q - 24q^2 + 252q^3 - 1462q^4 + \dots$ corresponds to the identity element. However, the γ_n are virtual characters. The following proof is due to Mason [**16**].

THEOREM 3.7. *For all positive integers n and all irreducible characters χ of M_{24}, $(\gamma_n, \chi) \in \mathbb{Z}$. In other words, γ_n is a virtual character.*

For any specific n, this can be verified by simply evaluating the inner product with full knowledge of the character table from the atlas. More generally, to check that $(\gamma_n, 1) \in \mathbb{Z}$ for all n at once would be a significant step. This requires that $|\,\mathrm{M}_{24}\,|$ divides $\sum_{g\in \mathrm{M}_{24}} \gamma_n(g)$ for all n. This can be interpreted as the coefficient of q^n in the expression

$$\frac{1}{|\,\mathrm{M}_{24}\,|} \sum_{g\in \mathrm{M}_{24}} f_g(z) = \frac{1}{|\,\mathrm{M}_{24}\,|} \sum_{n=1}^{\infty} \left(\sum_{g\in \mathrm{M}_{24}} a_n(g) \right) q^n. \tag{18}$$

If the average of $f_g(z)$ over M_{24} lies in $\mathbb{Z}[[q]]$, then part of the theorem is proven. Such a statement is plausible, since there are similar combinatorial statements like Burnside's lemma (the average over a finite group of the number of points fixed by a group action is equal to the number of orbits). In fact, this suggests that such a statement might not be unique to M_{24}, but hold for any group acting on a finite set.

THEOREM 3.8. *Let G be a finite group with a fixed permutation representation. For $g \in G$, of cycle shape $1^{r(1)}2^{r(2)}\dots N^{r(N)}$, define*

$$f_g(z) = \prod_{d=1}^{N} \eta(dz)^{r(d)} = q^r \sum_{n=1}^{\infty} \gamma_n(g) q^n.$$

Then the average

$$\zeta(G, \eta) := \frac{1}{|G|} \sum_{g\in G} f_g(z)$$

has integer coefficients. In particular, for any n

$$(\gamma_n, 1_G)_G \in \mathbb{Z}$$

Assuming this theorem, the specific result for M_{24} follows. First, M_{24} embeds in S_{24}, and the characters γ_n for M_{24} and S_{24} agree on M_{24}. If γ_n is a $\mathbb{Z}-$linear combination of irreducible characters of S_{24}, then since any character of S_{24} restricts to a character of M_{24} γ_n is a $\mathbb{Z}-$linear combination of irreducible characters of M_{24}. Thus it suffices to check that $(\gamma_n, \chi) \in \mathbb{Z}$ for all irreducible characters of S_n. But all such characters are $\mathbb{Z}-$linear combinations of transitive permutation character by Proposition 1.14, so $\chi = \sum_i a_i \operatorname{Ind}_{H_i}^{S_n}(1_{H_i})$ where the H_i are subgroups of S_{24}. By Frobenius reciprocity, $(\gamma_n, \operatorname{Ind}_{H_i}^{S_n}(1_{H_i}))_{S_n} = (\gamma_n, 1_{H_i})_{H_i}$ which is an integer by Theorem 3.8. This implies $(\gamma_n, \chi) \in \mathbb{Z}$, and hence that γ_n is a virtual character of M_{24}.

To prove Theorem 3.8, we use a generalization of Burnside's lemma known as Pólya's Theorem. Let G be a finite group of permutations of the set D, with $|D| = N$, and let $y_1 \dots y_n$ be indeterminates. The cycle index is defined to be

$$\zeta(G) = \frac{1}{|G|} \sum_{g\in G} \prod_{i=1}^{N} y_i^{r(i)} \tag{19}$$

where $g \in G$ has cycle shape $1^{r(1)} \dots N^{r(N)}$. For a power series $f(t) \in \mathbb{Z}[[t]]$, define

$$\zeta(G, f) = \frac{1}{G} \sum_{g \in G} \prod_{i=1}^{N} f(t^i)^{r(i)} \tag{20}$$

which is an evaluation of the cycle index at the power series.

Furthermore, for another finite set R and a weight function $w : R \to A$ where A is a commutative ring, define $\mathbb{F}$ to be the set of all functions $D \to R$ and $\mathcal{O}$ to be the set of orbits of G acting on $\mathbb{F}$ via $(gf)(d) = f(gd)$. Define the weight of a function to be

$$w(f) = \prod_{d \in D} w(f(d)). \tag{21}$$

This is constant on classes of $\mathcal{O}$. Define another variant of the cycle index $\zeta(G, w)$ by replacing each y_i by $\sum_{r \in R} w(r)^i$.

THEOREM 3.9 (Pólya). *With the notation as above, we have that*

$$\zeta(G, w) = \sum_{F \in \mathcal{O}} w(F) \tag{22}$$

This has the same spirit of Burnside's theorem, for it says that a weighted sum of the number of orbits equals an average of weights. A proof will be given in Section 2.2. Pólya's Theorem is the key ingredient in the proof of Theorem 3.8.

PROOF. To show that $\zeta(G, \eta)$ has integer coefficients, it suffices to show that the coefficients are algebraic integers, since they are certainly rational numbers. For any fixed positive integer d, to show that the first d coefficients are integers it suffices to consider the expression

$$f_d(q) := \frac{1}{|G|} \sum_{g \in G} f_g^d(z) \tag{23}$$

where f_g^d is a product obtained by replacing the infinite product for eta with the finite product involving only terms up to degree $d - 1$. Suppose there exist an integer b and elements $g_k(q) \in \mathbb{C}[q]$, $1 \le k \le b$, such that $g_k(q)$ has algebraic integer coefficients for $1 \le k \le b$ and $\sum_{k=1}^{b} g_k(q)^e = f_d(q^e)$ for $1 \le e \le N$. Define a function $w : \{1, \dots, b\} \to \mathbb{C}[z]$ by $w(k) = g_k$. Then using Pólya's Theorem, since the coefficients of $\sum_{F \in \mathcal{O}} w(F)$ are all algebraic integers it follows that the coefficients of $\zeta(G, w)$ are also algebraic integers. However, since $\sum_{k=1}^{b} g_k(q)^e = f_d(q^e)$,

$$\begin{aligned} \zeta(G, w) &= \frac{1}{|G|} \sum_{g \in G} \prod_{i=1}^{N} (\sum_{k=1}^{b} g_k(q)^i)^{r(i)} \\ &= \frac{1}{G} \sum_{G} \prod_{i=1}^{N} f_d(t^i)^{r(i)} \\ &= \zeta(G, f) \end{aligned}$$

as desired. This holds for any d, so all the coefficients are rational numbers and algebraic integers and hence lie in $\mathbb{Z}$.

It remains to construct the g_k. If such g_k exist for each monomial $c_i q^i$ in a polynomial, the collection of all such g_k work for the whole polynomial. For a monomial $c_i q^i$ with $c_i > 0$,

take $b = c_i$ and $g_1 = \dots g_b = q^i$. If $c_i < 0$, take a prime $p > N$ and let λ be a primitive pth root of $-c_i$. Take $b = p-1$ and $g_k(q) = \lambda^k q^i$. Furthermore, for $1 \leq e \leq N$,

$$\sum_{k=1}^{p-1} g_k(t)^e = (\sum_{k=1}^{p-1} \lambda^{ke}) q^{e \cdot i} = c_i q^{e \cdot i}$$

as desired. Thus such g_k always exist and the proof is complete. □

REMARK 3.10. The phenomena of the coefficients of eta products giving virtual characters is not special to M_{24}. In his paper on frame-shapes [**16**], Mason shows that for any association of elements of a finite group to eta products based on cycle shapes of any permutation representation, the coefficient of q^n in the $q-$expansion is a virtual character. There is also nothing special about the eta function: a similar argument works for any power series in q. The proof is essentially the same.

2.2. A Proof of Pólya's Theorem. The following proof of Pólya's Theorem uses the notation from the previous section, and follows the proof presented in Bruijn [**6**].

Burnside's lemma says that the number of orbits $|\mathcal{O}|$ is the average over G of the number of functions fixed by a permutation $g \in G$. Furthermore, the number of orbits with weight ω is given by

$$\frac{1}{|G|} \sum_{g \in G} s_\omega(g)$$

where $s_\omega(g)$ is the number of functions of weight ω fixed by g. Summing over all possible weights gives all orbits, so

$$\sum_{f \in \mathcal{O}} w(f) = \sum_{\omega} \omega \frac{1}{|G|} \sum_{g \in G} s_\omega(g) = \frac{1}{|G|} \sum_{g \in G} \sum_{\omega} \omega s_\omega(g)$$

If g has cycle shape $1^{l(1)} 2^{l(2)} \dots n^{l(n)}$, then if g fixes a function the function must be constant on every cycle of g. The permutation divides D into a collection of disjoint components, on which the function must be constant. Labeling these components by $D_1 \dots D_k$, we will show that

$$\sum_{f=gf} w(f) = \prod_{i=1}^{k} \sum_{r \in R} w(r)^{|D_i|}.$$

To prove this, consider expanding the product: picking a term in the expansion is the same as picking a mapping of $\{1, 2, \dots, k\}$ to R, and the value of the term in the expansion is the weight of the function interpreted as a function from D to R, constant on each component. Therefore it follows that

$$\sum_{\omega} \omega s_\omega(g) = \sum_{f \in \mathbb{F}, gf=f} w(f) = \prod_{i=1}^{k} (\sum_{r \in R} w(r)^i)^{l(i)}$$

Averaging over $g \in G$ gives

$$\zeta(G, f) = \frac{1}{G} \sum_{g \in G} \prod_{i=1}^{k} (\sum_{r \in R} w(r)^i)^{l(i)} = \frac{1}{|G|} \sum_{g \in G} \sum_{\omega} \omega s_\omega(g) = \sum_{f \in \mathcal{O}} w(f).$$

This completes the proof.

3. The Moonshine Module

The family of virtual characters associated to M_{24} through eta products suggests that there should be a "natural" infinite dimensional virtual graded module V such that

$$V = \bigoplus_{n\geq 1} V_n \quad \text{and} \quad \operatorname{tr}(g|V_n) = \gamma_n(g)$$

The existence of such a virtual module is equivalent to the fact that the γ_n are virtual characters. The content of this assertion is that the virtual module V should have an independent description.

This is similar to the theory of moonshine for the monster group discussed in the introduction, in which it was possible to prove the existence of a module satisfying the moonshine conjectures without constructing it. There was a conjectured construction, but it was far from clear that it satisfied the moonshine conjectures. Proving the equivalence, which shed a bit of light on why the moonshine conjectures are true, was the heart of Borcherd's work.

Happily, the module associated to M_{24} is easier to construct.

Let M be the standard 24−dimensional permutation representation of M_{24}. Represent a partition λ of n by $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ of n where λ_k is the number of times k appears in the partition. Being a partition means that $\sum_{k=1}^{n} k\lambda_k = n$. Denote λ being a partition of n by $\lambda \lhd n$. Define

$$\sigma(\lambda) := (-1)^{\sum_{k=1}^{n} \lambda_k}$$

and define

$$M_\lambda := \bigotimes_{k=1}^{n} \Lambda^{\lambda_k}(M)$$

where the product is taken in the Grothendieck ring of finite dimensional representations of M_{24}.

THEOREM 3.11. *Let V be the infinite dimensional graded virtual module*

$$V = \oplus V_n \quad \textit{where} \quad V_n = \sum_{\lambda \lhd (n-1)} \sigma(\lambda) M_\lambda. \tag{24}$$

V explains the family of characters γ_n in the sense that for $g \in M_{24}$

$$\operatorname{tr}(g|V_n) = \gamma_n(g).$$

It will be convenient to think about graded modules as being represented as a power series, with the coefficient of q^n being the nth graded piece. The ring formed is isomorphic to $R[[q]]$, where R is the Grothendieck ring of $\mathbb{C}[G]$ modules. This ring is naturally isomorphic to $(\mathbb{C}[G])\,[[q]]$, the power series ring over the ring of class functions. The isomorphism is given by associating to each virtual module its virtual character.

Suppose that for a power series $h \in \mathbb{Z}[[q]]$, the characters defined for $g \in M_{24}$ of cycle shape $g = 1^{r(1)} \ldots (24)^{r(24)}$ by

$$h_g(q) := \prod_{i=1}^{N} h(q^i)^{r(i)} = \sum_{n\geq 0} \gamma_n^h(g) q^n \tag{25}$$

are virtual characters. Then h determines an element of $R[[q]]$. Denote it by $\imath(h)$. It is straightforward to verify that $\imath(h_1)\imath(h_2) = \imath(h_1 \cdot h_2)$. This will let us prove the theorem.

PROOF. Consider the case of $h(q) = 1 - q$. We know that

$$h_g(q) = \prod_{i=1}^{N} (1 - q^i)^{r(i)}$$

The factor $(1 - t^i)$ is the characteristic polynomial of an i cycle acting on an i dimensional vector space by permutation, so the right side is simply $\prod_\lambda (1 - \lambda q)$ where λ runs through the eigenvalues (with multiplicity) of the action of g. Therefore the coefficient of q^n is simply $(-1)^n$ times the nth elementary symmetric polynomial evaluated on the eigenvalues of g. However, we know that the character of $\Lambda^n(M)$ evaluated at g is the nth elementary symmetric polynomial evaluated on the eigenvalues of g by Proposition 1.3. Thus it follows that

$$\imath(1 - q) = \sum_{n=0}^{24} (-1)^n \Lambda^n(M) q^n$$

Similarly, it follows that

$$\imath(1 - q^k) = \sum_{n=0}^{24} (-1)^n \Lambda^n(M) q^{tk}.$$

But then

$$\imath(\prod_{k=1}^{\infty} (1 - q^k)) = \prod_{k=1}^{\infty} (\sum_{n=0}^{24} (-1)^n \Lambda^n(M) q^{nk}) = \sum_{k=0}^{\infty} \left(\sum_{\lambda \lhd k} \sigma(\lambda) M_\lambda \right) q^k$$

where it makes sense to talk of the infinite product because the coefficient of q^n depends only on $\imath(1 - q^k)$ for $k \leq n$. The characters associated to $\imath(\prod_{k=1}^{\infty}(1 - q^k))$ are by definition

$$\sum_{n=1}^{\infty} \gamma_{n-1} q^n$$

with the index offset because the $q^{1/24}$ is not included in the infinite product. This shows that $\mathrm{tr}(g|V_n) = \gamma_n(g)$. □

REMARK 3.12. This provides an independent proof of the fact that the γ_n are virtual characters of M_{24} by explicitly finding the virtual module producing them. Mason does this in more generality, showing that for any power series h with constant term 1 and any representation of a group, the class functions γ_n^h defined in terms of the frame shape $g = \prod_{i=1}^{N} i^{r(i)}$ by

$$h_g(q) = \prod_{i=1}^{N} h(q^i)^{r(i)} = \sum_{n \geq 0} \gamma_n^h(g) q^n$$

can be expressed in terms of symmetric and exterior powers of the representation of the group. In this generality, the characters are only rational characters, not virtual characters [**16**].

The fact that all of the f_g are Hecke eigenforms shows that this family of representations has the unusual property of being multiplicative.

COROLLARY 3.13. *For n and m relatively prime integers,*

$$V_n \otimes V_m \simeq V_{nm}$$

PROOF. The characters of both sides are $\gamma_n \gamma_m$ and γ_{nm} by definition. By Corollary 3.6, these are equal, which implies the representations are equal. □

4. Number Theory Explaining M_{24}

Having established this correspondence between the representation theory of M_{24} and modular forms, the natural next step is to attempt to use number theory to understand M_{24} and its representation theory. One of the original "accidents", presented in Conway and Norton [**18**] and apparent in Table 7, that led to investigations of moonshine is that for $g \in M_{24}$ of cycle shape $1^{r(1)} \dots 24^{r(24)}$, these exists an N such that $r(m) = r(N/m)$ for $m \leq N$. The bridge between M_{24} and the theory of modular forms explains this symmetry.

Recall that the Fricke involution W_N is given by the matrix $\begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$. Although W_N is not in $\Gamma_0(N)$, the Fricke involution does preserve the space of cusp forms. In particular, any Hecke eigenform is a eigenfunction for W_N. Now suppose that

$$f(z) = \prod_d \eta(dz)^{r(d)}$$

is a Hecke eigenform in $S_k(\Gamma_0(N), \chi)$. In particular, this means that all of the d with $r(d) \neq 0$ divide N and the hypotheses of Theorem 2.41 are satisfied.

Applying the Fricke involution,

$$\begin{aligned} W_n f = N^{-k/2} z^{-k} f(\frac{-1}{Nz}) &= N^{-k/2} z^{-k} \prod_d \eta(-\frac{d}{Nz})^{r(d)} \\ &= N^{-k/2} z^{-k} \prod_d \sqrt{(N/d)(z/i)}^{r(d)} \eta(\frac{Nz}{d})^{r(d)} \\ &= c \prod_d \eta(dz)^{r(\frac{N}{d})} \end{aligned}$$

using Theorem 2.39. In order for this to be a scalar multiple of $f(z)$, $r(d)$ must equal $r(N/d)$ when $r(d) \neq 0$. Using the correspondence between cycle shapes in M_{24} and Hecke eigenforms, the cycle shapes must be symmetric as well.

5. Representation Theory Explaining Number Theory

It is also possible to use the description of this virtual module to analyze Ramanujan's τ function. Recall that it is defined to be the coefficients of Δ:

$$\Delta(z) := \eta(z)^{24} = \sum_{n=1}^{\infty} \tau(n) q^n \tag{26}$$

Some general information about $\tau(n)$ is contained in Chapter 7, section 4.5 of Serre [**25**]. One surprising fact about $\tau(n)$ is that it satisfies congruences modulo powers of the primes $2, 3, 5, 7, 23$, and 691. For example, Wilton proves that

$$\tau(n) \equiv \begin{cases} 0 & \mod 23 \text{ if } \left(\frac{n}{23}\right) = -1 \\ 2 & \mod 23 \text{ if } n = u^2 + 23v^2 \text{ for } u, v \in \mathbb{Z}_{\neq 0} \\ -1 & \mod 23 \text{ otherwise} \end{cases}. \tag{27}$$

by a relatively elementary argument [**29**]. A list of all of these congruences (along with an explanation along a totally different line) is found in Swinnerton-Dyer [**27**].

The formula for V_n in Theorem 3.11 suggests there should be a congruence for $\tau(n)$ modulo 23. The coefficients of $\tau(n)$ arise from the identity element of M_{24}, and the dimension

of $\Lambda^r(M)$ is $\binom{24}{r}$, which is congruent to 0 modulo 23 for $r \neq 0, 1, 23, 24$. Since so many terms disappear, it is natural to look for a congruence modulo 23. Since for $r = 0, 1, 23, 24$ we have $\dim(\Lambda^r(M)) = 1 \mod 23$,

$$\tau(n) = \sum_{\lambda \lhd' n-1} \sigma(\lambda) \mod 23 \tag{28}$$

where the partition runs only over λ with all of the $\lambda_i = 0, 1, 23, 24$.

Define $a(n)$ to be the number of partitions of n into an even number of distinct parts minus the number of of partitions of n into an odd number of distinct parts. The pentagonal number theorem, appearing as equation (3.1) of Wilton [**29**], states that

$$\prod_n (1 - q^n) = \sum_n a(n) q^n = \sum_{n \in \mathbb{Z}} (-1)^n q^{\frac{1}{2}n(3n+1)}. \tag{29}$$

A partition of $n - 1$ into parts where each part occurs $0, 1, 23$ or 24 times is the same as selecting numbers that sum to h, taking 23 of each of them, and then picking a partition of $n - 1 - 23h$ into distinct parts. Since 23 is odd, the number of parts arising from selecting $23h$ elements is the number of elements picked that sum to h. Thus (28) gives that

$$\tau(n) = a(n-1) + a(1)a(n-24) + \ldots + a(h)a(n-1-23h) \mod 23$$

where $h = [\frac{n-1}{23}]$. We know that $a(n-1-23i) = 0$ by the pentagonal number theorem unless $n - 1 - 23i = \frac{1}{2}m(3m \pm 1)$ for some $m \in \mathbb{Z}$. Simplifying this equation modulo 23,

$$\begin{aligned} n - 1 &= \frac{1}{2}m(3m \pm 1) \mod 23 \\ n - 1 &= 12(3m^2 \pm m) \mod 23 \\ n &= (6m \pm 1)^2 \mod 23 \end{aligned}$$

Thus if n is not a quadratic residue modulo 23, $\tau(n) \equiv 0 \mod 23$.

This method of obtaining the congruence modulo 23 is no more and no less than a disguised version of the proof Wilton gives using generating functions. The fact that the only partitions that contribute to $\tau(n) \mod 23$ are those that use each integer $0, 1, 23$ or 24 times corresponds to decomposing the generating function

$$\Delta(q) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = q \prod_{n=1} (1 - q^n) \prod_{n=1}^{\infty} (1 - q^{23n}) \mod 23$$

The coefficients of $\prod_{n=1}^{\infty}(1 - q^n)$ are precisely the $a(n)$, and Wilton invokes the pentagonal number theorem for the same purpose. The point of using the M_{24} module V is that it provides an explanation for why there *should* be a congruence modulo 23: because almost all of the dimensions of the exterior powers of M vanish modulo 23.

The spirit of moonshine is that there is a connection between the representations of the finite sporadic simple groups connected with the Monster group and certain types of modular forms. For M_{24}, conjugacy classes correspond to eta products based on cycle shape. For the Monster group, its module gives a correspondence between conjugacy classes and certain Hauptmodul. This connection allows number theory and representation theory to interact, explaining the balanced cycle shapes of M_{24} in terms of Fricke involutions and allowing the description of the graded virtual module V in terms of the exterior algebra to explain congruences of the τ function. Despite the connection illuminating these issues, there is still no satisfactory answer to why these groups are involved, and why moonshine is monstrous.

Bibliography

1. Tom Apostol, *Modular functions and dirichlet series in number theory*, Springer-Verlag, 1989.
2. Thomas Beth, Dieter Jungnickel, and Hanfried Lenz, *Design theory*, Cambridge University Press, 1999.
3. Richard E. Borcherds, *What is moonshine?*, Proceedings of the International Congress of Mathematicians (1998), 607–615.
4. P J. Cameron and J. H. van Lint, *Graphs, codes, and designs*, Cambridge University Press, 1980.
5. J. H. Conway, *Atlas of finite groups: maximal subgroups and ordinary characters for simple groups*, Oxford University Press, 1985.
6. N. G. de Bruijn, *Applied combinatorial mathematics*, ch. Pólya's Theory of Counting, John Wiley and Sons, 1964.
7. John D. Dixon and Brian Mortimer, *Permuation groups*, Springer-Verlag, 1996.
8. D. Dummit, H. Kisilevsky, and J. McKay, *Multiplicative products of η functions*, Finite Groups - Coming of Age, American Mathematical Society, 1985.
9. William Fulton, *Young tableaux*, Cambridge University Press, 1997.
10. Terry Gannon, *Monstrous moonshine: The first 25 years*, Bulletin of the London Mathematical Society (2006), 1–33.
11. Robert L. Griess Jr., *Twelve sporadic groups*, Springer, 1998.
12. Henryk Iwaniec, *Topics in classical automorphic forms*, American Mathematics Society, 1997.
13. Neal Koblitz, *Introduction to elliptic curves and modular forms*, Springer, 1993.
14. James Lepowsky, *The mathematical work of the 1998 fields medalists: The work of richard e. borcherds*, Notices of the AMS (1999), 17–19.
15. Gérard Ligozat, *Courbes modulaires de genre 1*, Bulletin de la Société Mathématique de France, Société Mathématique de France, 1975.
16. Geoffrey Mason, *Frame-shapes and rational characters of finite groups*, Journal of Algebra (1984), 237–246.
17. ———, *M24 and certain automorphic forms*, Finite Groups - Coming of Age, American Mathematical Society, 1985.
18. J. H. Conway S. P. Norton, *Monstrous moonshine*, Bull. of the London Math. Soc. (1979), 308–339.
19. Ken Ono, *The web of modularity: Arithmetic of the coefficients of modular forms and q-series*, American Mathematical Society, 2004.
20. *Oxford english dictionary*, online, http://dictionary.oed.com.
21. Robert Rankin, *Modular forms and functions*, Cambridge University Press, 1977.
22. Joseph Rotman, *An introduction to the theory of groups*, Springer-Verlag, 1995.
23. *SAGE mathematics software, version 3.2.2*, http://www.sagemath.org.
24. Jean-Pierre Serre, *Linear representations of finite groups*, Springer, 1977.
25. ———, *A course in arithmetic*, Springer, 1996.
26. Goro Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton University Press, 1971.
27. H. P. F. Swinnerton-Dyer, *On ℓ-adic representations and congruences for coefficients of modular forms*, Modular Forms of One Complex Variable III, Springer Berlin, 1973, pp. 1–55.
28. G. V. Voskresenskaya, *Modular forms and regular representations of groups of order 24*, Mathematical Notes (1996), 216–218.
29. J. R. Wilton, *Congruence Properties of Ramanujan's Function $\tau(n)$*, Proc. London Math. Soc. **s2-31** (1930), no. 1, 1–10.
30. Yufei Zhou, *Young tableau and the representation theory of the symmetric group*, Harvard College Math Review (2008), 33–45.