The Mathieu Groups

by

Megan E. Stiles

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the

Mathematics

Program

YOUNGSTOWN STATE UNIVERSITY

May, 2011

The Mathieu Groups

Megan E. Stiles

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Megan E. Stiles, Student — Date

Approvals:

Dr. Thomas Wakefield, Thesis Advisor — Date

Dr. Neil Flowers, Committee Member — Date

Dr. Eric Wingler, Committee Member — Date

Peter J. Kasvinsky, Dean of School of Graduate Studies & Research — Date

ABSTRACT

The Classification of Finite Simple Groups was a prominent goal of algebraists. The Classification Theorem was complete in 1983 and many textbooks from the 1980s include detailed proofs and explorations of many aspects of this subject. For example, J.J. Rotman devotes an entire chapter to the Mathieu groups [13].

It seems that there is still disagreement amongst mathematicians as to whether the Classification Theorem should be deemed thorough or without major error. Looking into the entire Classification Theorem would be a huge undertaking, so in this paper we are discussing only the five sporadic Mathieu groups. Looking at these small sporadic simple groups opened up a discussion of transitivity and k-transitivity.

In addition to traditional abstract algebra material, this paper explores a relationship between the five sporadic Mathieu groups and the combinatorial Steiner Systems. Included in this discussion is the relationship of M_{24} with the Binary Golay Code.

This thesis ends in a proof of the simplicity of the Mathieu Groups. The proof of the simplicity of M_{11} and M_{23} which was developed by R. Chapman in his note, *An elementary proof of the Mathieu groups* M_{11} *and* M_{23}, makes the preliminary theorems to the simplicity proof in J.J. Rotman's book look much less perfunctory [2],[13]. This raises the question of whether there could possibly be a more succinct proof of the simplicity of M_{12}, M_{22} and M_{24}.

ACKNOWLEDGEMENTS

I would like to thank my family and friends for being so encouraging and supportive as I have pursued my master's degree in mathematics. I would like to sincerely thank all the mathematics professors at Grove City College for giving me such a good foundation in my mathematics education. I am particularly grateful to Dr. Thompson for his incredibly interesting Abstract Algebra and Number Theory course.

I would also like to thank Dr. Wakefield, my thesis advisor, for his corrections and guidance during this process. Additionally, I am very grateful for his instruction in my Abstract Algebra courses at Youngstown State University. I would like to thank the rest of my Thesis Committee, Dr. Wingler, and Dr. Flowers, for their corrections and comments.

Lastly, I would like to thank my fellow classmates for their aid in working with LaTeX.

Contents

1 Background Information **1**

1.1 Some History of the Classification of Finite Simple Groups 1

1.2 Emile Mathieu 3

2 Mathematical Background **5**

2.1 Essential Definitions and Theorems 5

2.2 Group Actions and Transitivity 14

2.3 The Sylow Theorems 28

3 The Golay Code and Steiner Systems **37**

3.1 Definitions 37

3.2 Establishing a Relationship 41

4 The Mathieu Groups **48**

4.1 Group Representations 48

4.2 Orders 50

5 Simplicity **51**

5.1 Simplicity of the Mathieu Groups with Prime Degree 51

5.2 Simplicity of the Mathieu Groups with Degrees 12, 22, and 24 54

1 Background Information

1.1 Some History of the Classification of Finite Simple Groups

A group is a set of elements with a binary operation on that set in which four basic properties hold. The set must be closed under the operation, there exists an identity element, there exists an inverse for each element, and all the elements exhibit an associative property under the operation. A subgroup, H, of a group G is a subset of G which is a group under the same operation. The subgroup consisting of only the identity element is called the trivial subgroup and every subgroup H such that $H \subsetneq G$ is called a proper subgroup. Special subsets of a group are called cosets; specifically left cosets of H in G are defined as $gH = \{gh : h \in H\}$ and right cosets of H in G are $Hg = \{hg : h \in H\}$ for some $g \in G$. If for every $g \in G$ the left and right cosets are equal, the subgroup H is said to be a normal subgroup. A simple group is one which has no nontrivial proper normal subgroups.

Dealing with only the finite simple groups makes further classification of these groups more palatable. There are five different types of finite simple groups. These types include:

1. the cyclic groups of prime order;
2. the alternating groups;
3. the groups of Lie type over the Galois field of order q, where q is odd, which is denoted $GF(q)$;
4. the groups of Lie type over $GF(q)$ where q is even; and
5. the 26 sporadic groups.

In this paper, we will concentrate on some particular sporadic simple groups.

The classification of all finite simple groups began in 1861 with the discovery of two of the sporadic simple groups by French mathematician Emile Mathieu. In 1861 and 1873, Emile Mathieu published two papers revealing the first five sporadic simple groups, aptly named the Mathieu groups. It took until 1965 for the next sporadic group, J_1, to be discovered by Zvonimir Janko.

Most of the theorems involving the classification of finite simple groups were published between 1955 and 1983 [15]. At the beginning of this important period in classifying these groups, mathematicians were trying to deduce a general classification strategy. The first of these was suggested by Richard Brauer in 1954 at the International Congress of Mathematicians in Amsterdam. This proposal was:

> In a finite nonabelian simple group G, choose an involution z... and consider its centralizer $C_G(z)$.... Show that the isomorphism type of $C_G(z)$ determines the possible isomorphism types of G [15].

Between 1950 and 1965, some mathematicians (including Brauer) spent time working on perfecting the techniques of classification. This work, along with the proof of Burnside's odd order conjecture by Thompson and Feit, helped make this particular strategy more viable [15].

Consequently, a flood of articles and papers were published in the late 1960s and the 1970s by mathematicians such as Thompson, Walter, Alperin, Brauer, Gorenstein, Harada, and Aschbacher on classification of particular simple group types. In conjunction with this, many new sporadic simple groups were being discovered. Many of these groups were discovered by finding evidence of a group satisfying certain conditions rather than constructing the group itself. During this uncertain time, people

thought that maybe the sporadic groups would be an infinite set. Then there was a pivotal conference in Duluth, Minnesota in 1976 which presented key theorems that illustrated that the production of a complete classification of finite simple groups would take place in the near future [7]. This insight proved to be true as the mathematical endeavor of classifying all finite simple groups essentially ended in 1981 with Englishman Simon Norton and his proof of the uniqueness of one of the other sporadic simple groups [7].

1.2 Emile Mathieu

Emile Mathieu was born in Metz, France in 1835. He was a promising student even in his younger days, earning awards for academics and behavior; he was honored mostly in the areas of Latin and Greek studies, but also excelled in mathematics. As a teenager, Mathieu entered the École Polytechnique in Paris to study mathematics. His bachelor's degree was earned partially due to a paper he published which extended some algebraic theorems about derivatives and differentials originally published by Descartes and Budan [6].

By the age of 24, Mathieu earned the degree of Doctor of Mathematical Sciences with his thesis titled *On the Number of Values a Function Can Assume, and on the Formation of Certain Multiply Transitive Functions*. His thesis concentrated on the theory of substitutions which led to two articles which were published in Liouville's *Journal de mathématiques pures et appliquées*. In the second of these articles, the first two sporadic groups were discussed. These two and another article he published in 1862 on the solution of equations with prime degree helped attract the notice of the scientific community. As a consequence, Mathieu was placed on the list of candidates

for the Paris Academy of Sciences in the area of geometry. He was never elected into the Academy, however.

Mathieu spent much time in scientific research, but earned his living as a private tutor and working for both public and private schools in France. Most of his research was in applied mathematics, particularly mathematical physics. He returned to pure mathematics in his papers *On the theory of biquadratic remainders*, published in 1867, and *Sur la fonction cinq fois transitive de 24 quantités*, published in 1873. The latter of these articles details the last of the Mathieu groups. These things, though, were only a brief interlude in his studies in applied mathematics. He mainly concentrated on his work in mathematical physics along with analytical and celestial mechanics. Mathieu explained this in a quote found in Duhem's article (translated from French):

> Not having found the encouragement I had expected for my researches in pure mathematics, I gradually inclined toward applied mathematics, not for the sake of any gain that I might derive from them, but in the hope that the results of my investigations would more engage the interest of scientific men [6].

In 1867, Mathieu was offered a job teaching a course in mathematical physics at the Sorbonne. But, with the death of Lamé, one of the last notable applied mathematicians of his day in France, mathematical physics was going out of style [6]. Thus, in 1869, Mathieu applied for and received the position of the chair of pure mathematics at Basançon. In 1873, he transferred to Nancy and worked in the same position. After this point, he was overlooked several times for a position as a chair at the Sorbonne's Collège de France.

He died at the age of 55 in relative anonymity for a man of his talent and early

promise. He is most remembered for the Mathieu Groups and the Mathieu Functions, $C(a, q, z)$ and $S(a, q, z)$ which for $q = 0$ are:

$$C(a, 0, z) = \cos(\sqrt{a}z) \text{ and } S(a, 0, z) = \sin(\sqrt{a}z).$$

These are part of the solutions of the Mathieu differential equation:

$$\frac{d^2u}{dz^2} + (a + 16q\cos 2z)u = 0,$$

for which the general solutions are $y = k_1C(a, q, z) + k_2S(a, q, z)$ where k_1 and k_2 are constants.

According to P. Duhem in his biography of Emile Mathieu,

> After a life full of disappointments, he died at a time when the official men of science hardly had begun to suspect that somewhere in the provinces [of France], far away from the capital, there lived a mathematician whose works were an honor to his country [6].

2 Mathematical Background

2.1 Essential Definitions and Theorems

We begin the mathematical portion of this paper with a discussion of the basic definitions in group theory which will be essential when exploring the Mathieu groups. Many of the definitions will be very general, but most will be referred to in later proofs and statements.

Definition 2.1. A set G is a **group** under an operation $*$ if:

1. For all $a, b \in G$, $a * b \in G$;

2. There exists $e \in G$, such that for all $a \in G$, $e * a = a * e = a$;

3. For all $a \in G$, there exists $a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$; and

4. For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$.

Many examples of groups are very obvious to any mathematician. It is imperative, however to be careful not to assume that if a set is a group under one operation, it must be a group under any common operation.

Example 2.2. The set of integers, $\mathbb{Z}$, forms a group under the operation of addition. It is clear that the integers are closed under addition and exhibit the associative property under addition. We have a property of 0 such that $a + 0 = 0 + a = a$. We also know that each integer has a unique opposite, so for any $a \in \mathbb{Z}$, there exists $-a \in \mathbb{Z}$ such that $a + -a = 0$.

Example 2.3. The set of integers under the operation of multiplication is not a group. Note that for any $a \in \mathbb{Z}$, $1 \cdot a = a \cdot 1 = a$. So 1 is our identity for this set. But the inverse of any integer a is the reciprocal of a which is not an integer. Therefore $\mathbb{Z}$ is not a group under multiplication.

There are other commonly used groups which are not so apparent to those outside of the abstract algebra field. Many of these groups are collections of functions or other elements which are not numbers.

Definition 2.4. Let $X = \{1, 2, \ldots, n\}$ and S_X be the collection of all permutations of X. Then S_X is a group under the operation of function composition called the **symmetric group** of degree n.

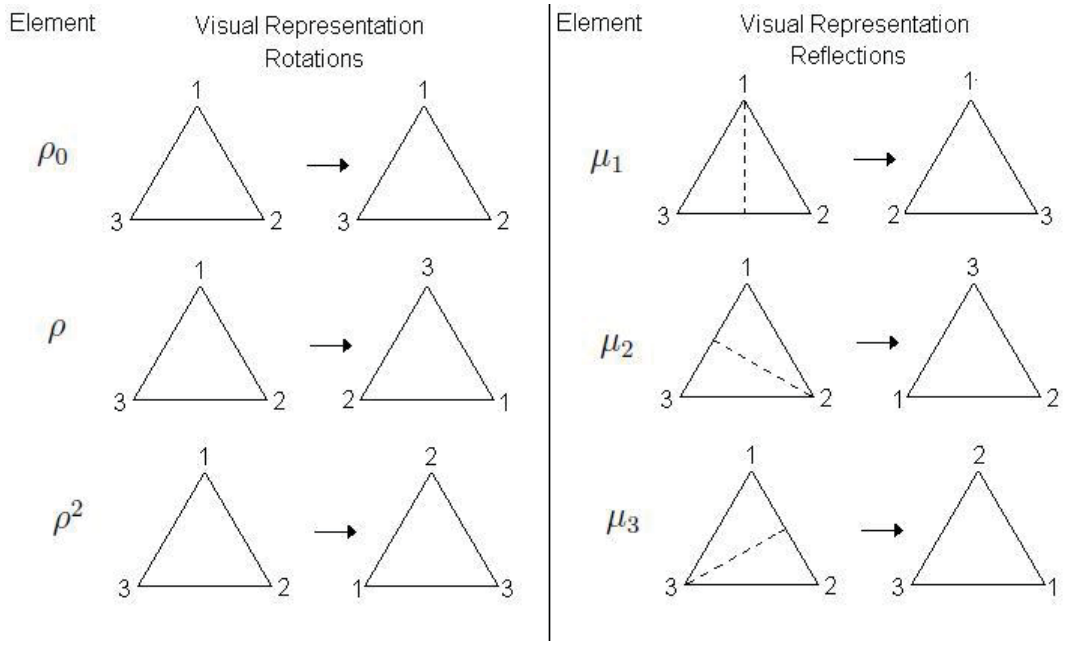

Figure 1: Visual Representation of S_3

Example 2.5. Consider the symmetric group, S_3. This is the group consisting of rotations and reflections of an equilateral triangle. The elements are $\{\rho_0, \rho, \rho^2, \mu_1, \mu_2, \mu_3\}$. These elements are also permutations of the numbers $1, 2$, and 3. See Figure 1. The operation is composition and therefore works from right to left. Table 1 shows the multiplication table of S_3. Note that $\mu_1\rho = \mu_2$ and $\rho\mu_1 = \mu_3$ so the group is non-commutative.

The next example is often introduced as an example with elements which are matrices. For the purposes of this paper, we will discuss the Klein four group in a more general sense.

Example 2.6. The Klein four group, denoted V, is the group of four elements $\{e, a, b, c\}$ where e is the identity. Refer to Table 2 for the multiplication table. This is an example of a commutative, or abelian, group.

$\circ$	ρ_0	ρ	ρ^2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ	ρ^2	μ_1	μ_2	μ_3
ρ	ρ	ρ^2	ρ_0	μ_3	μ_1	μ_2
ρ^2	ρ^2	ρ_0	ρ	μ_2	μ_3	μ_1
μ_1	μ_1	μ_2	μ_3	ρ_0	ρ	ρ^2
μ_2	μ_2	μ_3	μ_1	ρ^2	ρ_0	ρ
μ_3	μ_3	μ_1	μ_2	ρ	ρ^2	ρ_0

Table 1: Multiplication Table of S_3

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Table 2: Multiplication Table for the Klein four Group

Definition 2.7. A group G is said to be **abelian** if G is commutative.

Many times in abstract algebra it becomes useful to discuss the internal structure of a group. To discuss this, we must begin with the definition of a subgroup.

Definition 2.8. A **subgroup** of a group G is a subset H of G where H is a group under the same operation as G. This is denoted by $H \leq G$.

Theorem 2.9. *Let G be a group. Then $H \subseteq G$ is a subgroup if and only if the following conditions are true under the same operation as the group G:*

1. *$H \neq \emptyset$;*
2. *for all $a, b \in H$ $ab \in H$;*
3. *for every $a \in H$ there exists $a^{-1} \in H$ such that $aa^{-1} = a^{-1}a = e$, where e is the identity in G.*

Proof. Suppose H is a subgroup of G. Then these conditions are met by the Definition 2.8.

Conversely, suppose the three conditions are true. By (1), there exists some $a \in H$. Then by (3), there exists a^{-1} such that $aa^{-1} = a^{-1}a = e$. By (2), $e \in H$ is an identity element. Let $b \in H$. Then since $H \subseteq G$, $b \in G$. Therefore, since G is a group, $be = eb = b$. Similarly, for any $a, b, c \in H$, these elements are also in G and inherit the associative property of G. Therefore H is a group under the same operation as G. □

There is a relationship between subgroups and elements in a group, and this relationship forms sets. It will become apparent that these sets are used to classify groups further.

Definition 2.10. Let G be a group, H a subgroup of G, and $g \in G$. A **left coset of H in G** is the set $gH = \{gh : h \in H\}$. A **right coset of H in G** is the set $Hg = \{hg : h \in H\}$. The number of left cosets of H in G or equivalently, the number of right cosets of H in G is known as the **index of H in G**, denoted $|G : H|$. The identity in a set of cosets is $1H = H$. Note that $aH = H$ means that $a \in H$. (Often we will refer to the left cosets of a subgroup in a group as the cosets.)

Now we must look at special types of subgroups and how we use them to classify their respective groups. These next definitions will be the basic foundation of this paper.

Definition 2.11. Let G be a group and H a subgroup of G. Then H is a **normal subgroup of G**, denoted $H \trianglelefteq G$, if $Hg = gH$ for all $g \in G$.

Definition 2.12. A group is **simple** if it has no nontrivial proper normal subgroups.

Example 2.13. Let $H = \{\mu_1, \rho_0\}$. The cosets of H in S_3 are:

$$\begin{aligned}
\rho_0 H &= H & H\rho_0 &= H \\
\rho H &= \{\mu_2, \rho\} & H\rho &= \{\mu_3, \rho\} \\
\rho^2 H &= \{\mu_3, \rho^2\} & H\rho^2 &= \{\mu_2, \rho^2\} \\
\mu_1 H &= \{\rho_0, \mu_1\} & H\mu_1 &= \{\rho_0, \mu_1\} \\
\mu_2 H &= \{\rho, \mu_2\} & H\mu_2 &= \{\rho^2, \mu_2\} \\
\mu_3 H &= \{\rho^2, \mu_3\} & H\mu_3 &= \{\rho, \mu_3\}
\end{aligned}$$

The left and right cosets are only equal for some elements of S_3. Thus H is not a normal subgroup of S_3. Let $K = \{\rho_0, \rho, \rho^2\}$. The cosets of K in S_3 are:

$$\begin{aligned}
\rho_0 K &= K & K\rho_0 &= K \\
\rho K &= \{\rho, \rho^2, \rho_0\} & K\rho &= \{\rho, \rho^2\rho_0\} \\
\rho^2 K &= \{\rho^2, \rho_0, \rho\} & K\rho^2 &= \{\rho^2, \rho_0, \rho\} \\
\mu_1 K &= \{\mu_1, \mu_3, \mu_2\} & K\mu_1 &= \{\mu_1, \mu_2, \mu_3\} \\
\mu_2 K &= \{\mu_2, \mu_1, \mu_3\} & K\mu_2 &= \{\mu_2, \mu_3, \mu_1\} \\
\mu_3 K &= \{\mu_3, \mu_2, \mu_1\} & K\mu_3 &= \{\mu_3, \mu_1, \mu_2\}
\end{aligned}$$

All of the left and right cosets are equal. So K is a normal subgroup of S_3. In fact, $K = A_3$ is called the alternating group. Thus S_3 is an example of a finite group that is not simple.

Theorem 2.14. *Let G be an abelian group. Every nontrivial, proper subgroup of G is also a normal subgroup of G.*

Proof. Let G be an abelian group and H a nontrivial, proper subgroup of G. Let $g \in G$. Let $x \in gH$. Then $x = gh$ for some $h \in H$. Since G is abelian, $x = gh = hg \in Hg$. So $gH \subseteq Hg$. Since the left and right cosets have the same size, $|gH| = |Hg|$. Therefore $H \trianglelefteq G$. □

Next we will see many properties of groups and their internal structure. It is important to be able to explore various aspects of a group's basic construction.

Lagrange's Theorem. *Let G be a group and $H \leq G$. Then*

1. *$|H| \mid |G|$ and*

2. *$|G|/|H|$ is the number of distinct cosets of G in H.*

Proof. Let $g \in G$. We claim that the equivalence class of g is the left coset, gH, under the relation of coset equality, for all $a, b \in G$, $aH = bH$ or $a^{-1}b \in H$. To prove this claim, we must show the three equivalence relation properties. Let $a \in G$. Then $a^{-1}a = e \in H$. So the relation is reflexive. For $a, b \in G$, suppose $a^{-1}b \in H$. Then $aH = bH$. This implies that $b^{-1}a \in H$. Therefore the relation is also symmetric. Let $a, b, c \in G$. Then suppose $a^{-1}b \in H$ and $b^{-1}c \in H$. Then $(a^{-1}b)(b^{-1}c) \in H$ which implies that $a^{-1}ec = a^{-1}c \in H$. Thus, the relation is transitive. Hence the left cosets are equivalence classes. So these cosets partition G. So if a_iH, $i = 1, \ldots, k$ form the set of cosets of H in G, $|G| = |a_1H| + \cdots + |a_kH|$. We claim that $|H| = |a_iH|$ for all $i = 1, \ldots, k$. Let $a_j \in G$ for a fixed j, $1 \leq j \leq i$. Define a map, $\phi : H \to a_jH$ by $\phi(h) = a_jh$. Note this is one-to-one since if $\phi(h_1) = \phi(h_2)$, then $a_jh_1 = a_jh_2$. Hence, by applying a_j^{-1} to both sides of the equation, we obtain $h_1 = h_2$. To see that ϕ is onto, let $a_jh_3 \in a_jH$ for some $h_3 \in H$. Then since $\phi(h_3) = a_jh_3$, ϕ is onto. Hence our claim is proven. So since $|G| = |a_1H| + \cdots + |a_kH|$, $|G| = k|H|$. So $|H| \mid |G|$. Also, $|G|/|H| = k$ which is the number of distinct cosets of G in H. □

Definition 2.15. The **centralizer** of an element z in G is the collection of elements in G which commute with z. This set is denoted $C_G(z) = \{g \in G : gz = zg\}$.

Example 2.16. In the group, S_3, the centralizer of ρ is $C_{S_3}(\rho) = \{\rho_0, \rho, \rho^2\}$.

The centralizer is instrumental in many proofs involving the classification of finite simple groups. The centralizer of an element $a \in G$ is also a subgroup of G.

Definition 2.17. Let G be a group and H a subgroup of G. Then the **normalizer of H in G** is $N_G(H) = \{g \in G : gHg^{-1} = H\}$.

Note that the normalizer of $H \leq N_G(H) \leq G$.

Example 2.18. One subgroup of S_3 is $H = \{\mu_1, \rho_0\}$. Its normalizer is $N_G(H) = \{\rho_0, \mu_1\}$.

Definition 2.19. Let G be a group. Then the **center of G**, $Z(G)$, is the set of elements in G which commute with every other element in G.

The center of a group is also a subgroup of the group.

Example 2.20. In S_3, $Z(G) = \{\sigma \in S_3 : \sigma\tau = \tau\sigma \text{ for all } \tau \in S_3\} = \{\rho_0\}$.

Definition 2.21. Let G be a group with $a \in G$. Then the **order of a** is the smallest positive integer k such that $a^k = 1$, where 1 is the identity element in G. If no such integer exists, we say that a has infinite order.

Example 2.22. In S_3, $|\rho| = |\rho^2| = 3$, $|\mu_1| = |\mu_2| = |\mu_3| = 2$, and $|\rho_0| = 1$.

Definition 2.23. An **involution** is an element of order two.

Example 2.24. In S_3, μ_1, μ_2, and μ_3 are involutions.

Example 2.25. In V, every element except e is an involution.

As in many areas of mathematics, it is important in Algebra to discuss how two groups may be related. We use several different levels of morphisms to achieve this goal.

Definition 2.26. Let $(G, *)$ and $(G', \diamond)$ be groups under the operations $*$ and $\diamond$ respectively. Then the map $\phi : G \to G'$ is a **homomorphism** if for all $a, b \in G$, $\phi(a * b) = \phi(a) \diamond \phi(b)$.

Example 2.27. Define $\phi : \mathbb{Z} \to 2\mathbb{Z}$ by $\phi(x) = 2x$. Then ϕ is a homomorphism since for all $x, y \in \mathbb{Z}$, $\phi(x + y) = 2(x + y) = 2x + 2y = \phi(x) + \phi(y)$.

Definition 2.28. An **isomorphism** is a homomorphism which is a bijection. The groups G and G' are said to be **isomorphic** if there is an isomorphism between G and G'. This is denoted $G \cong G'$.

Example 2.29. The homomorphism defined in Example 2.27 is an isomorphism. Let $\phi(x) = \phi(y)$. Then $2x = 2y$ for $x, y \in \mathbb{Z}$. This implies that $x = y$. So ϕ is an injection. To see that ϕ is a surjection, let $c \in 2\mathbb{Z}$. Then by the definition of $2\mathbb{Z}$, there exists $z \in \mathbb{Z}$ such that $c = 2z$. Also, $\phi(z) = 2z = c$. Thus ϕ is onto. Therefore ϕ is an isomorphism.

Definition 2.30. Let G and H be groups. An **isomorphism class** (or type) is the equivalence class $\{H : H \cong G\}$.

Definition 2.31. A function $\phi : X \to X$ is a **permutation of a set X** if ϕ is a bijection.

A field is a set which is an abelian group under one operation and is commutative, has closure, inverses of nonzero elements, and an identity under another operation as well as a distributive property over both operations.

Definition 2.32. Let p be a prime and $n \in \mathbb{Z}, n \geq 1$. The fields with p^n elements are called **Galois Fields**, denoted by $GF(q)$ where $q = p^n$.

The following definitions present special types of matrix groups which will be used in specific proofs in later sections of this paper.

Definition 2.33. Let F be a field. The **General Linear Group, $GL(m, F)$**, is the set of $m \times m$ matrices, with nonzero determinants, whose entries are from F. If $F = GF(q)$, then $GL(m, F)$ may be written as $GL(m, q)$.

Definition 2.34. Let F be a field. The **Special Linear Group, $SL(m, F)$** is the set of $m \times m$ matrices, with a determinant of one, whose entries are from F. If $F = GF(q)$, then $SL(m, F)$ may be written as $SL(m, q)$.

Definition 2.35. Let F be a field. Let E denote the $m \times m$ identity matrix. The group **$Z_1(m, F)$** is the group of scalar multiples kE with $k^m = 1$. Again, if $F = GF(q)$, then $Z_1(m, F)$ may be written as $Z_1(m, q)$.

Definition 2.36. Let G be a group. Let H be a normal subgroup of G. Then the **quotient group** G/N read G mod N is the set of cosets of N in G. Since N is normal, this set is also a group.

Definition 2.37. Let F be a field. The **Projective Special Linear Group, $PSL(m, F)$**, is the group $SL(m, F)/Z_1(m, F)$. When $F = GF(q)$, then $SL(m, F)/Z_1(m, F)$ may be written as $SL(m, q)/Z_1(m, q)$.

2.2 Group Actions and Transitivity

Group actions connect the idea of general groups with permutations and allow us to study transitivity in groups. The subject of transitivity is extremely important when discussing the Mathieu groups. Emile Mathieu discovered these groups while seeking highly transitive permutation groups [7].

Definition 2.38. Let G be a group and X a set. The **left group action of G on X** is a map from $G \times X$ to X which fulfills the following conditions:

1. $e \cdot x = x$ for all $x \in X$ where e is the identity element in G.

2. $g \cdot (h \cdot x) = (gh) \cdot x$ for all $g, h \in G$ and $x \in X$.

Note: The operation $\cdot$ represents the group action and the regular multiplication represents the group operation in G.

Example 2.39. The permutation group S_3 has a very natural group action on the three vertices of the triangle, $X = \{1, 2, 3\}$. The action is the rotation or reflection as illustrated in Example 2.5. For example, $\mu_1 \cdot 1 = 1$, since μ_1 does not move the vertex 1. So $\mu_1 \cdot 2 = 3$ since μ_1 moves the vertex 2 into the 3rd vertex position. Then $\mu_1 \cdot 3 = 2$ since μ_1 moves vertex 3 into the 2nd vertex position. This is a group action because the identity, ρ_0 fixes every element of X. Also, for the second property, consider $G \backslash \{\rho_0\}$. Using the multiplication table, Table 1, $(\rho\rho) \cdot x = \rho^2 \cdot x$. By inspection, we see that $\rho \cdot (\rho \cdot 1) = \rho \cdot (2) = 3$ and $\rho^2 \cdot 1 = 3$. Checking all cases gives the conclusion that the action of S_3 on X is a group action.

Theorem 2.40. *Let G be a group acting on a set X. Then*

1. *for each $g \in G$, $\sigma_g : X \to X$ defined by $\sigma_g(x) = g \cdot x$ is a permutation of X.*

2. *the map $\phi : G \to S_X$ defined by $\phi(g) = \sigma_g$ is a group homomorphism.*

Proof.

1. To see that for each $g \in G$, $\sigma_g : X \to X$ defined by $\sigma_g(x) = g \cdot x$ is a permutation of X, or a bijection, we must check to see if σ_g is one-to-one and onto. Suppose

$\sigma_g(x) = \sigma_g(y)$, for $x, y \in X$. Then $g \cdot x = g \cdot y$. Thus $g^{-1} \cdot (g \cdot x) = g^{-1} \cdot (g \cdot y)$. Therefore, $x = e \cdot x = (g^{-1}g) \cdot x = g^{-1} \cdot (g \cdot x) = g^{-1} \cdot (g \cdot y) = (g^{-1}g) \cdot y = e \cdot y = y$ by the definition of left group action. Let $w \in X$. Thus $g^{-1} \cdot w \in X$. Note that since $\sigma_g(g^{-1} \cdot w) = g \cdot (g^{-1} \cdot w) = (gg^{-1}) \cdot w = e \cdot w = w$, σ_g is onto. Thus σ_g is an injection.

2. Let $g, h \in G$. Recall that the operation in S_X is composition. Then $\phi(gh)(x) = \sigma_{gh}(x) = (gh) \cdot x = g \cdot (h \cdot x) = \sigma_g(\sigma_h(x)) = (\sigma_g \circ \sigma_h)(x) = \phi(g) \circ \phi(h)(x)$. Therefore, ϕ is a group homomorphism.

□

Theorem 2.41. *Let G be a group and S_X the group of permutations of a set X. For a given group homomorphism, $\phi : G \to S_X$, the map $G \times X \mapsto X$ defined by $(g, x) \mapsto g \cdot x = \phi(g)(x)$ is a group action on X.*

Proof. To show that this map is a group action on X, the conditions in Definition 2.38 must be verified.

1. Let id denote the identity function in S_X. Then $e \cdot x = \phi(e)(x) = id(x) = x$.

2. Let $g_1, g_2 \in G$. Then $g_1 \cdot (g_2 \cdot x) = \phi(g_1)(\phi(g_2)(x)) = \phi(g_1) \circ \phi(g_2)(x) = \phi(g_1 g_2)(x) = (g_1 g_2) \cdot x$ since ϕ is a homomorphism.

Therefore, the defined map is a group action on X. □

Definition 2.42. Let $\phi : G \to G'$ be a homomorphism. Then the **kernel of ϕ**, denoted Kernϕ, is the set $\{g \in G : \phi(g) = e'\}$ where e' is the identity in G'.

Theorem 2.43. *A homomorphism $\phi : G \to G'$ is an injection, or one-to-one, if and only if Kern$\phi = \{e\}$, where e is the identity element in G.*

Proof. Suppose ϕ is one-to-one. Then $x \in \text{Kern}\phi$ if and only if $\phi(x) = e'$. But $\phi(e) = e'$. So $\phi(x) = \phi(e)$. Since ϕ is an injection, $x = e$. Therefore $x \in \text{Kern}\phi$ if and only if $x = e$.

Conversely, let $\text{Kern}\phi = \{e\}$. Let $g, h \in G$ such that $\phi(g) = \phi(h)$. Then $\phi(g)(\phi(h))^{-1} = e'$, which implies that $\phi(gh^{-1}) = e'$, so $gh^{-1} \in \text{Kern}\phi$. By the assumption, $gh^{-1} = e$. Therefore $g = h$. □

Definition 2.44. Let G be a group acting on X. Then the **kernel of a group action** is the kernel of the group homomorphism, $\phi : G \to S_X$ which was defined in Theorem 2.41.

Cayley's Theorem. *Every group is isomorphic to a subgroup of a group of permutations.*

Proof. Let G be a group acting on itself by left multiplication, i.e., $G \times G \to G$, where $(g, x) \mapsto gx$. By Theorem 2.40 there exists a group homomorphism $\phi : G \to S_X$ defined by $\phi(g) = \sigma_g \in S_X$ where $\sigma_g(x) = gx$ for all $x \in G$. To see that ϕ is one-to-one, consider the $\text{Kern}\phi$. We can see that $x \in \text{Kern}\phi$

if and only if $\phi(g)(x) = id(x)$ for all $x \in G$

if and only if $\sigma_g(x) = x$ for all $x \in G$

if and only if $gx = x$ for all $x \in G$

if and only if $g = e$.

Therefore $\text{Kern}\phi = e$ and ϕ is one-to-one. This implies that $G \cong \text{Im}(\phi) \leq S_X$. □

Definition 2.45. Let G be a group and X a set. The homomorphism $\phi : G \to S_X$ associated with the action of G on X is called the **permutation representation** of the action.

Definition 2.46. Let G be a group acting on a set X. The **stabilizer of x in G** is $G_x = \{g \in G : g \cdot x = x\}$.

Example 2.47. Let S_3 act on X where $X = \{1, 2, 3\}$, the three vertices of the triangle. The stabilizer of 1 is everything in S_3 that leaves 1 fixed. In this case, $G_1 = \{\rho_0, \mu_1\}$.

Theorem 2.48. *Let G be a group acting on a set X with $x \in X$. Then $G_x \leq G$.*

Proof. Let $x \in X$. It is clear that $G_x \subseteq G$ by the definition of G_x. Note that if e is the identity in G, $e \cdot x = x$ by Definition 2.38 and so $e \in G_x$. Thus $G_x \neq \emptyset$. Let $a, b \in G_x$. Then $a \cdot x = x$ and $b \cdot x = x$. So $(ab) \cdot x = a \cdot (b \cdot x)$ by Definition 2.38. Then $a \cdot (b \cdot x) = a \cdot x = x$ since $a, b \in G_x$. Thus $(ab) \cdot x = x$ so $ab \in G_x$. Let $c \in G_x$. Then $c \cdot x = x$. Thus $c^{-1} \cdot (c \cdot x) = c^{-1} \cdot x$. Simplifying the left side of this equation using Definition 2.38, $c^{-1} \cdot (c \cdot x) = (c^{-1}c)x = ex = x$. So $x = c^{-1} \cdot x$. Thus $c^{-1} \in G_x$. Therefore by Theorem 2.9, $G_x \leq G$. □

Definition 2.49. Let G be a group acting on a set X, with $a \in X$. The **orbit of G in X containing a** is $\mathcal{O}_a = \{g \cdot a : g \in G\}$.

Example 2.50. Let S_3 act on X where $X = \{1, 2, 3\}$, the three vertices of the triangle. Recall that the elements of S_3 are $\rho_0, \rho, \rho^2, \mu, \mu_1, \mu_2$. Note that $\rho_0 \cdot 1 = 1$, $\rho \cdot 1 = 3$, and $\rho^2 \cdot 1 = 2$. So without considering the other two elements, we know that $\mathcal{O}_1 = \{1, 2, 3\} = X$.

The Orbit-Stabilizer Relation. *Let G be a group acting on a set X and $a \in X$. The size of the orbit of a, $|\mathcal{O}_a|$, is equal to the order of the index of the stabilizer of a in G. If G is finite, then $|\mathcal{O}_a| = |G|/|G_a|$.*

Proof. Note that by Definition 2.10, $|G : G_a|$ is the number of left cosets of G_a in G. To show that $|G : G_a| = |\mathcal{O}_a|$, we must show there exists a bijection between the cosets of G_a in G- call this set K- and $\mathcal{O}_a$. Let $b \in \mathcal{O}_a$. Then by Definition 2.49, there exists $g \in G$ such that $b = g \cdot a$. Define the map:

$$\psi : \mathcal{O}_a \to K$$

by $\psi(b) = gG_a$, where $b = g \cdot a \in \mathcal{O}_a$. To see that ψ is well-defined, let $x = y \in \mathcal{O}_a$. Then there exist $g_1, g_2 \in G$ such that $x = g_1 \cdot a$ and $y = g_2 \cdot a$. So $\psi(x) = g_1G_a$ and $\psi(y) = g_2G_a$. But $x = y$ implies that $g_1 \cdot a = g_2 \cdot a$. Thus $(g_2^{-1}g_1) \cdot a = g_2^{-1} \cdot (g_1 \cdot a) = g_2^{-1} \cdot (g_2 \cdot a) = (g_2^{-1}g_2) \cdot a = a$. So $g_2^{-1}g_1 \in G_a$. By Definition 2.10, $g_2^{-1}g_1G_a = G_a$ and therefore $g_1G_a = g_2G_a$. Hence ψ is well-defined. In order to show that ψ is a one-to-one function, let $\psi(c) = \psi(d) \in K$. So there exist $g, h \in G$ with $c = g \cdot a$ and $d = h \cdot a$. Since $\psi(c) = \psi(d)$, $gG_a = hG_a$. But then $h^{-1}gG_a = G_a$, which implies that $h^{-1}g \in G_a$. By Definition 2.46, $(h^{-1}g) \cdot a = a$. Therefore $h^{-1} \cdot (g \cdot a) = a$ and so $c = g \cdot a = h \cdot a = d$. Hence ψ is a one-to-one function. All that is left to prove is that ψ is onto. Let $k \in G$ such that $kG_a \in K$. Then $g \in kG_a$ if and only if $g \cdot a = k \cdot g_1 \cdot a = k \cdot a$ for some $g_1 \in G$. Thus $g \cdot a \in \mathcal{O}_a$ and $\psi(g \cdot a) = \psi(k \cdot a) = kG_a$. Hence ψ is onto. We can now conclude that there exists a bijection from $\mathcal{O}_a$ to the left cosets of G_a in G. Thus, $|G : G_a| = |\mathcal{O}_a|$. □

Example 2.51. The Orbit-Stabilizer relation is demonstrated very clearly using the action of $G = S_3$ on the vertices of a triangle. Note that $|G| = 3! = 6$, and from

Examples 2.47 and 2.50 that $|\mathcal{O}_1| = 3$ and $|G_1| = 2$. So the these results agree with the Orbit-Stabilizer relation.

Definition 2.52. The **conjugation of G on X** is the group action of G on X defined by $g \cdot x = gxg^{-1}$ for $g \in G$ and $x \in X$.

Definition 2.53. Let G be a group. Then $a, b \in G$ are said to be **conjugate in G** if there exists a $g \in G$ such that $b = gag^{-1}$. In other words, if G is acting on itself by conjugation, a and b are conjugate in G if a and b are in the same orbit.

Definition 2.54. Let G be a group acting on itself by conjugation and let $a \in G$. The **conjugacy class of a in G**, denoted $Cl(a) = \{gag^{-1} : g \in G\}$, is the set of all conjugates of a. In this case, the conjugacy class of a is also the orbit of a in this action.

Note: Let G be a group acting on itself by conjugation and let $a \in G$. Then, recalling Definition 2.15, the centralizer of a in G is $C_G(a) = \{g \in G : ga = ag\} = \{g \in G : gag^{-1} = a\}$. So from Definition 2.46 we can see that the centralizer of an element in a group acting on itself by conjugation is the same as the stabilizer.

Theorem 2.55. *Let G be a finite group acting on itself by conjugation and let $a \in G$. Then $|Cl(a)| = |G|/|C_G(a)|$.*

Proof. By the Orbit Stabilizer Relation, we know that $|\mathcal{O}_a| = |G|/|G_a|$. Definition 2.54 and the above comment implies that $|Cl(a)| = |G|/|C_G(a)|$. □

Definition 2.56. Let G be a group acting on a set X. The action is called **faithful** if $\phi : G \to S_X$ is injective, i.e. Kern$\phi = \{e\}$ where e is the identity element in G.

Example 2.57. The action of S_3 on $X = \{1, 2, 3\}$, the set of vertices of a triangle, is faithful. We can test this by looking at the kernel of the action, or the set of elements in S_3 which fix every element in X. We can see that ρ_0 fixes every element of S_3. But $\rho \cdot 1 = 2$, $\rho^2 \cdot 1 = 3$, $\mu_1 \cdot 2 = 3$, $\mu_2 \cdot 1 = 3$, and $\mu_3 \cdot 1 = 2$. So no element besides ρ_0 fixes every element of X.

Definition 2.58. Let G be a group acting on a set, X. A group G is **transitive** on X if G has only one orbit.

Example 2.59. The group action of S_3 on its vertices, $X = \{1, 2, 3\}$ is an example of a transitive action since $\rho \cdot 1 = 2$, $\rho^2 \cdot 1 = 3$, and $\mu_2 \cdot 1 = 3$. Therefore $\mathcal{O}_1 = \{1, 2, 3\} = X$.

Example 2.60. Another example of a group action involves the group D_4, which is a subgroup of the symmetric group S_4. The group D_4 consists of the rotations and reflection of a square. The elements of D_4 are $\rho_0, \rho, \rho^2, \rho^3, \tau, \rho\tau, \rho^2\tau$ and $\rho^3\tau$ and can be represented as in Figure 2. This group has a group action on $X = \{1, 2, 3, 4, s_1, s_2, s_3, s_4, d_1, d_2, c\}$ where these elements are illustrated in Figure 3. This action is not transitive since every element of D_4 of type s_i with $1 \leq i \leq 4$ is sent to another element of type s_i. So $\mathcal{O}_{s_1} = \{s_1, s_2, s_3, s_4\}$. Similarly, $\mathcal{O}_1 = \{1, 2, 3, 4\}$, $\mathcal{O}_{d_1} = \{d_1, d_2\}$, and $\mathcal{O}_c = \{c\}$.

Definition 2.61. Let k be a positive integer less than $|X|$. Then G is ***k*-transitive** on X if it acts transitively on the set of k-tuples of distinct elements of X. In other words, for $(x_1, x_2, \ldots, x_k)$ and $(y_1, y_2, \ldots, y_k)$ where $x_i, y_i \in X$, for all $i = 1, 2, \ldots, k$, there exists some $g \in G$ such that $x_i = gy_i$ for all $i = 1, 2, \ldots, k$. Often we call k-transitive actions, for $k \geq 2$, **multiply transitive actions**.

Emile Mathieu discovered that M_{12} and M_{24} are 5-transitive, or quintuply transitive and M_{11} and M_{23} are 4-transitive or quadruply transitive permutation groups. In

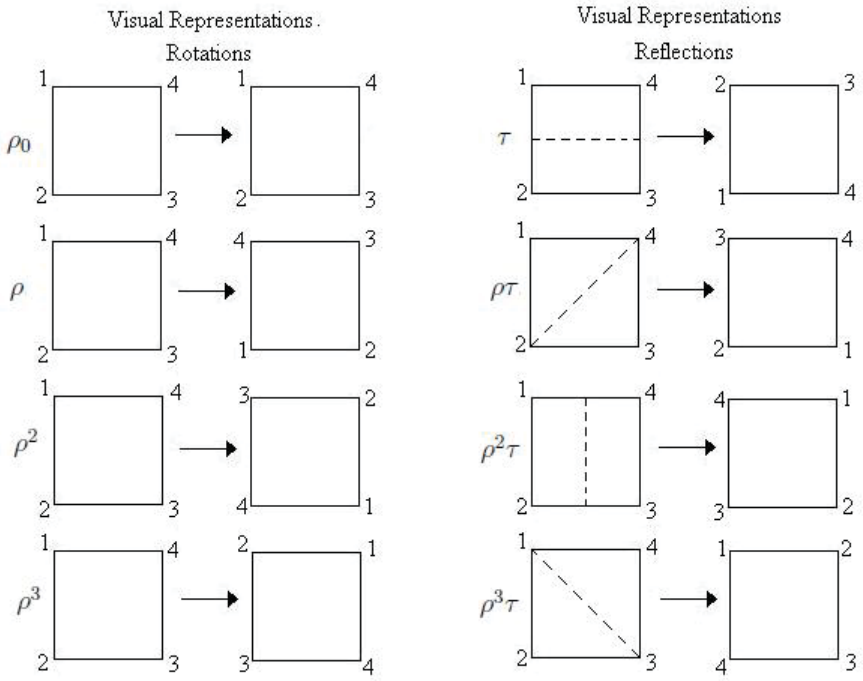

Figure 2: Visual Representation of D_4

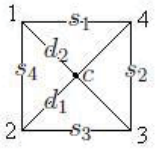

Figure 3: Set $X = \{1, 2, 3, 4, s_1, s_2, s_3, s_4, d_1, d_2, c\}$

fact, these four are the only quintuply and quadruply transitive permutation groups, aside from the alternating and symmetric groups [7].

Example 2.62. Consider the group S_4 acting on $X = \{1, 2, 3, 4\}$. The group can be viewed as the permutations of the set X. In order to simplify notation of the elements of this group, we may write them in "cycle notation." The identity element which holds everything fixed is simply 1. The element which moves 1 to 2, 2 to 3, 3 to 4, and 4 to 1 is denoted (1 2 3 4). The action simply moves an element to the one directly to its right in the cycle. The last element is moved back to the first element of the cycle. Thus the elements of S_4 are {1, (1 2 3 4), (1 3 2 4), (1 4 2 3), (1 4 3 2), (1 3 4 2), (1 2 4 3), (1 3)(2 4), (1 2)(3 4), (1 4)(2 3), (1 2 3), (1 3 2), (1 2 4), (1 4 2), (2 3 4), (2 4 3), (1 3 4), (1 4 3), (1 2), (1 3), (1 4), (2 3), (2 4), (3 4)}. Since S_4 is by definition all the possible permutations of $X = \{1, 2, 3, 4\}$, it is clear that S_4 is 4-transitive. Using this representation of S_4 allows us to see that all the possible 4-tuples of X can be rearranged into any other 4-tuple of X. This result can be generalized to S_n for any positive integer n.

Definition 2.63. Let G be a group acting on a set, X. Let k be less than or equal to $|X|$. Then G is **sharply *k*-transitive** if for any two distinct k-tuples of X, there exists exactly one element $g \in G$ which maps the first k-tuple to the second.

Example 2.64. It is also true that the action of S_4 on $X = \{1, 2, 3, 4\}$ is sharply 4-transitive.

Definition 2.65. Let G be a group acting on a set X. Then G is called **regular** if this action is sharply 1-transitive. If a group G acts on a set X and the action is faithful and regular, then only the identity in G fixes any points in X.

Example 2.66. The Klein four group, V, is a regular group under the action $V \times V \to V$ such that $g \cdot x = gx$. Referring to the multiplication table, Table 2, makes this clear.

Definition 2.67. Let G be a group acting on a set X. A **block** of the set X is a subset B of X with the following property. If $gB = \{gb : b \in B\}$, then $gB = B$ or $gB \cap B = \emptyset$ for all $g \in G$. Note that $B = \emptyset$, $B = X$, and any one-point subset of X are called **trivial blocks**. Any other block is called **nontrivial**.

Definition 2.68. Let G be a group acting on a set X. Then X is said to be **primitive** if it contains no nontrivial blocks.

The following theorems and definitions will be used to prove Theorem 2.88 and Corollary 2.89, which are used in Section 4 to prove the simplicity of three of the Mathieu groups.

Theorem 2.69. *Let G be a group acting transitively on a set X and suppose the size of X is n. Let B be a nontrivial block of X. Then the following are true:*

1. *If $g \in G$ then gB is a block.*
2. *There are elements $g_1, g_2, \ldots, g_m$ of G such that $Y = \{B, g_1B, \ldots, g_mB\}$ is a partition of X.*

Proof. 1. Assume $gB \cap hgB \neq \emptyset$ for some $h \in G$. Since B is a block, $B = g^{-1}hgB$ or $B \cap g^{-1}hgB = \emptyset$. Since $gB \cap hgB \neq \emptyset$, $B \cap g^{-1}hgB \neq \emptyset$. Along with the fact that $B = g^{-1}hgB$, this implies that $gB = hgB$. Therefore gB is a block.

2. Let $b \in B$ with $x_1 \notin B$. Then since G acts transitively on X, there exists $g_1 \in G$ such that $g_1 b = x_1$. This implies that $B \neq g_1B$ and therefore, since B is a block,

$B \cap g_1B = \emptyset$. If $B \cup g_1B = X$ then we are done. If not, then let $x_2 \notin (B \cup g_1B)$. Then there exists $g_2 \in G$ with $g_2b = x_2$. Therefore since B and g_1B are blocks, g_2B is disjoint from both B and g_1B. Using the same argument, you can see that eventually we will have a partition of X.

□

Definition 2.70. Let G be a group acting on a set X and let $H \leq G$. The **orbit of H in X containing x** is $\mathcal{O}_{Hx} = \{hx : h \in H\}$.

Theorem 2.71. *Let G be a group acting on a set X with $x, y \in X$. Let H be a subgroup of G. Suppose that $\mathcal{O}_{Hx} \cap \mathcal{O}_{Hy} \neq \emptyset$. Then $\mathcal{O}_{Hx} = \mathcal{O}_{Hy}$. If $H \trianglelefteq G$ then $\mathcal{O}_{Hx}$ for all $x \in X$ are blocks of X.*

Proof. Let G be a group acting on a set X with $x, y \in X$. Let H be a subgroup of G. Suppose that $\mathcal{O}_{Hx} \cap \mathcal{O}_{Hy} \neq \emptyset$. Note that since $\mathcal{O}_{Hx} \cap \mathcal{O}_{Hy} \neq \emptyset$, there exists an element a in both of these orbits. So $a = h_1x$ and $a = h_2y$ for some $h_1, h_2 \in H$. But then $h_1x = h_2y$. So $x = h_1^{-1}h_2y$ and $y = h_2^{-1}h_1x$. Thus $y \in \mathcal{O}_{Hx}$ and $x \in \mathcal{O}_{Hy}$. Suppose $b \in \mathcal{O}_{Hy}$. Then $b = h_3y$ for some $h_3 \in H$. Since $y \in \mathcal{O}_{Hx}$, $y = h_4x$ for some $h_4 \in H$. Therefore, $b = h_3y = h_3h_4x$. Thus $b \in \mathcal{O}_{Hx}$. So $\mathcal{O}_{Hy} \subseteq \mathcal{O}_{Hx}$. Similarly, $\mathcal{O}_{Hx} \subseteq \mathcal{O}_{Hy}$. Therefore $\mathcal{O}_{Hx} = \mathcal{O}_{Hy}$.

Suppose that $H \trianglelefteq G$ and let $g \in G$. Assume that $g\mathcal{O}_{Hx} \cap \mathcal{O}_{Hx} \neq \emptyset$. Note that if $a \in g\mathcal{O}_{Hx}$ then $a = ghx$ for some $h \in H$. Then since H is normal in G, $a = (ghg^{-1})gx = h_1(gx)$ for some $h_1 = ghg^{-1} \in H$. Since G is acting on a set X, $gx \in X$. Thus $a \in \mathcal{O}_{H(gx)}$. So $g\mathcal{O}_{Hx} \subseteq \mathcal{O}_{H(gx)}$. Let $b \in \mathcal{O}_{H(gx)}$. Then $b = h_1gx$ for some $h_1 \in H$. By the normality of H in G, $b = h_1gx = g(g^{-1}h_1g)x = gh_2x$ for some $h_2 \in H$. So $b \in g\mathcal{O}_{Hx}$. Therefore, $\mathcal{O}_{H(gx)} \subseteq g\mathcal{O}_{H(x)}$. So $\mathcal{O}_{H(gx)} = g\mathcal{O}_{Hx}$. Then by assumption $g\mathcal{O}_{Hx} \cap \mathcal{O}_{Hx} \neq \emptyset$. But then $\mathcal{O}_{H(gx)} \cap \mathcal{O}_{Hx} \neq \emptyset$. So by the first part of this proof,

$\mathcal{O}_{H(gx)} = g\mathcal{O}_x = \mathcal{O}_{Hx}$. □

Theorem 2.72. *Let G act on a set X such that this action is multiply transitive. Then X is primitive.*

Proof. Proceed by contradiction. Assume X has a nontrivial block B. So $|B| > 1$ and $B \neq X$. Let $x, y \in B$ with $x \neq y$. Let $z \notin B$. Note that $(x, y, *, \ldots, *)$ and $(x, z, *, \ldots, *)$ are k-tuples in X. Since X is k-transitive, there exists $g \in G$ such that $gx = x$ and $gy = z$ by definition of k-transitive. But then since $z \notin B$ and $z \in gB$, by the definition of block, $B \cap gB = \emptyset$. Also, since $x \in B$ and $x \in gB$, $B \cap gB \neq \emptyset$. This is a contradiction. Therefore X is primitive. □

Theorem 2.73. *Let G be a group acting on a set X transitively. Then X is primitive if and only if, for each $x \in X$, the stabilizer G_x is a maximal proper subgroup of G.*

Proof. Assume X is primitive and let $x \in X$. Proceed by contradiction. Suppose that there exists a subgroup H such that $G_x \subsetneq H \subsetneq G$. Define $B = \mathcal{O}_{Hx}$. To see that B is a block, let $g \in G$ and with $\mathcal{O}_{Hx} \cap g\mathcal{O}_{Hx} \neq \emptyset$. So there exist $h, h' \in H$ such that $hx = gh'x$. This implies that $x = h^{-1}gh'x$. So $h^{-1}gh' \in G_x \subsetneq H$. Therefore $g \in H$. Thus $gB = g\mathcal{O}_{Hx} = \mathcal{O}_{Hx} = B$. Hence B is a block. Now it remains to be shown that B is nontrivial. Now $B \neq \emptyset$ since $B \cap gB \neq \emptyset$. If $B = X$, then let $g \in G$ with $g \notin H$. Then since this is a transitive action, $gx \in X = B$, so $gx = hx$ for some $h \in H$. So $h^{-1}gx = x$. Therefore $h^{-1}g \in G_x \subsetneq H$. But then $g \in H$ which is a contradiction. So $B \neq X$. Also since $G_x \subsetneq H$, $H \neq \{1\}$, so B is not a one-point set for if $B = \{x\}$ then $\mathcal{O}_{Hx} = \{hx : h \in H\} = \{x\} \leq G_x$. So X is not primitive, which is a contradiction. Thus G_x is a maximal proper subgroup of G.

Assume that every G_x is a maximal proper subgroup. Proceed by contradiction.

Suppose there exists a nontrivial block B in X. Define the subgroup H of G as follows:

$$H = \{g \in G : gB = B\}$$

Then let $x \in B$. We claim that $G_x \subseteq H$. Let $g \in G_x$. By definition of block, either $gB = B$ or $gB \cap B = \emptyset$. But we know that $x \in B$ and $gx = x$ so then $gB = B$. Hence $g \in H$, so $G_x \subseteq H$. But B is nontrivial, so there exists $y \in B$ with $y \neq x$. Since G acts on X is transitively, there exists $g_1 \in G$ such that $g_1 x = y$. So $g_1 \notin G_x$, but $g_1 \in H$. So $G_x \subsetneq H$. By Theorem 2.69 part 2, $H = G$ if only if $B = X$. Thus $G_x \subsetneq H \subsetneq G$ which is a contradiction. So there exists no nontrivial block B in X. Hence, X is primitive. □

Definition 2.74. Let G be a group acting on a set X. Let $H \leq G$. Then the action on X is said to be ***H*-transitive** if $\mathcal{O}_{Hx} = X$ for all $x \in X$.

Theorem 2.75. *Let G be a group acting on a set X. If the action on X is faithful and primitive, $H \trianglelefteq G$, and $H \neq \{1\}$, then X is H-transitive.*

Proof. By Theorem 2.71 for all $x \in X$, $\mathcal{O}_{Hx}$ is a block. Since G is a primitive action, either $\mathcal{O}_{Hx} = X$ or $\mathcal{O}_{Hx} = \{x\}$. If $\mathcal{O}_{Hx} = \{x\}$, $h \cdot x = x$ for all $h \in H$. Then $\phi(h)(x) = x$ which implies that $h \in \text{Kern}(\phi)$ for all $h \in H$. But this action is faithful, so $\mathcal{O}_{Ha} \neq \{x\}$. Thus X is H-transitive. □

Theorem 2.76. *Let G be a group acting on a set X. Let X be faithful and primitive and suppose G_x is simple. Then either G is simple or for every $H \trianglelefteq G$, such that $H \neq \{1\}$, H acts regularly on X.*

Proof. Assume $H \neq \{1\}$ and $H \trianglelefteq G$. By Theorem 2.75, X is H-transitive. So either H acts regularly and $|H_x| = 1$ for all $x \in X$, or $H \cap G_x \neq \{1\}$ for some $x \in X$. Suppose

$H \cap G_x \neq 1$. Clearly, $H \cap G_x \trianglelefteq G_x$. Let $g \in G_x$ and $h \in H \cap G_x$. Then $ghg^{-1} \in G_x$. But since $H \trianglelefteq G$, $ghg^{-1} \in H$. Therefore $ghg^{-1} \in H \cap G_x$ and $H \cap G_x \trianglelefteq G_x$. Since G_x is simple, $H \cap G_x = G_x$ or $H \cap G_x = \{1\}$ which is not possible. Thus $G_x \subseteq H$. By Theorem 2.73, either $H = G_x$ or $H = G$. Since H is transitive, $H \neq G_x$. Therefore G is simple. □

Definition 2.77. Let G act on X and Y. A function $\phi : X \to Y$ is called a ***G*-map** if $\phi(gx) = g\phi(x)$ for all $g \in G$ and $x \in X$. If ϕ is one-to-one and onto, then ϕ is a ***G*-isomorphism** and X and Y are **isomorphic.**

Theorem 2.78. *Let G be a finite group acting on a finite set X transitively. Let $H \trianglelefteq G$ and let the action of H on X be regular. Let $x \in X$ and let G_x act on H^*=$H\backslash\{e\}$ by conjugation. Then H^* and $X\backslash\{x\}$ are G_x-isomorphic.*

Proof. Define $\phi : H^* \to X\backslash\{x\}$ by $\phi(h) = h \cdot x$ for $h \in H^*$. If for some $h, k \in H^*$, $\phi(h) = \phi(k)$, then $hx = kx$. This implies that $h^{-1}k \in H_x$. Since H is regular, $H_x = \{1\}$. Thus $h^{-1}k = 1$ and so $h = k$. Therefore ϕ is one-to-one. Since H is regular, $|\mathcal{O}_{Ha}| = |X|$ for any $a \in X$ and $|H_a| = 1$. By the Orbit-Stabilizer Theorem, $|X| = \frac{|H|}{|H_a|} = |H|$. Then $|X\backslash\{x\}| = |H^*|$. So ϕ is also onto. To see that ϕ is a G_x-map, let $g \in G_x$ and $h \in H^*$. Then since G_x acts on H^* by conjugation, $\phi(g \cdot h) = \phi(ghg^{-1}) = ghg^{-1}x = g \cdot hx$ since $g \in G_x$. Also, $g\phi(h) = ghx$. Therefore ϕ is a G_x-isomorphism and $H^* \cong X\backslash\{x\}$. □

2.3 The Sylow Theorems

The Sylow Theorems deal with groups and subgroups of orders that are primes or powers of primes. Applications of the Sylow Theorems include proofs that groups

with certain properties and orders are simple. Some of this section will be used to prove the simplicity of the Mathieu groups.

Definition 2.79. A group G is said to be **cyclic** if there exists an element $g \in G$ such that $G = \{g^n : n \in \mathbb{Z}\}$. We denote these groups $G = \langle g \rangle$ and say $\boldsymbol{G}$ **is generated by** $\boldsymbol{g}$.

Theorem 2.80. *Let G be a finite group, and let $g_1, g_2, \ldots, g_r$ be representatives of the noncentral conjugacy classes. Then $|G| = |Z(G)| + \sum_{i=1}^{r} |G : C_G(g_i)|$.*

Proof. Note that the conjugacy classes partition G. Let $x \in Z(G)$. Then $Cl(x) = \{gxg^{-1} : g \in G\} = \{xgg^{-1} : g \in G\} = \{x\}$, since $x \in Z(G)$. Therefore each element in the center is contained in its own conjugacy class. So $G = Z(G) \cup (\cup Cl(g_i))$ where $Cl(g_i)$ are disjoint and $Z(G)$ is disjoint from the conjugacy classes. Therefore $|G| = |Z(G)| + \sum_{i=1}^{r} |Cl(g_i)| = |Z(G)| + \sum_{i=1}^{r} |G : C_G(g_i)|$ by Theorem 2.55. □

Lemma 2.81. *Let G be a finite group with no nontrivial proper subgroups. Then G is cyclic.*

Proof. Since G is a group, if $G = \{1\}$, then $G = \langle 1 \rangle$ is cyclic. Suppose $G \neq \langle 1 \rangle$. Let $g \in G$. Since G is finite, $|g| = k$ for some $k \in \mathbb{Z}$; in other words, there exists an integer k such that $g^k = 1$. If $1 < k < |G|$, then G would have a nontrivial proper subgroup generated by g. Thus $|G| = k$. Then $G = \langle g \rangle$ and thus G is cyclic. □

Theorem 2.82. *Let G be a finite abelian group and p a prime dividing the order of G. Then G has an element of order p.*

Proof. Proceed by induction on $|G|$. If $|G| = 1$ there is nothing to prove. If $|G| = 2$ or $|G| = 3$ then G is cyclic and has an element of order 2 or 3 respectively. Assume the

statement is true for all abelian groups of order less than $|G|$. If $|G|$ has no nontrivial proper subgroups, then G is cyclic and therefore has an element of order p. Let H be a nontrivial, proper subgroup of G. If $p \mid |H|$, then by the induction hypothesis, H contains an element of order p and so G contains an element of order p. Assume $p \nmid |H|$. Since G is abelian, $H \trianglelefteq G$ so we may consider the abelian group G/H. Now $|G/H| = |G|/|H| < |G|$ and $p \mid |G/H|$. Therefore by the induction hypothesis, G/H contains an element of order p. So there exists an element $bH \in G/H$ of order p. Thus $(bH)^p = H$. This implies that $b^p \in H$. Let $c = b^{|H|}$. Note that $c^p = b^{|H|p} = 1$. Since $|c| \mid p$, either $|c| = p$ or $|c| = 1$. If $|c| = 1$, then c is the identity element. So $H = eH = b^{|H|}H = (bH)^{|H|}$. But $|bH| = p$ so p must divide the order of H. This is a contradiction. Therefore $|c| = p$ and G has an element of order p. □

Cauchy's Theorem. *Let p be a prime and G a finite group. If p divides the order of G, then G has an element of order p.*

Proof. Proceed by induction on $|G|$. If $|G| = p$ then by Lagrange's Theorem, the only possible subgroups of G would be G itself and the trivial subgroup. Therefore by Lemma 2.81, G is cyclic and has an element of order p. Assume that if H is a group with $|H| < |G|$, and $p \mid |H|$, then H has an element of order p. If p divides the order of any proper subgroup of G, by the inductive hypothesis, the subgroup has an element of order p and so G has an element of order p. Suppose no proper subgroup of G has order divisible by p. Let $g_1, g_2, \ldots, g_r$ be representatives of the noncentral conjugacy classes of G. Then by Theorem 2.80, $|G| = |Z(G)| + \sum_{i=1}^{r} |G : C_G(g_i)|$. Since each g_i is noncentral, it follows that $C_G(g_i) \lneq G$. So $p \nmid |C_G(g_i)|$. Since $p \mid |G|$, $p \mid |G : C_G(g_i)|$ for all $i = 1, 2, \ldots r$ and therefore $p \mid \sum_{i=1}^{r} |G : C_G(g_i)|$. Thus $p \mid |Z(G)|$. Hence, $Z(G)$ must not be a proper subgroup of G. Therefore $Z(G) = G$ and G is abelian. By

Theorem 2.82, the result follows. □

The Correspondence Theorem. *Let $K \trianglelefteq G$ and define $\phi : G \to G/K$ by $\phi(g) = gK$. Then ϕ defines a one - to - one correspondence between the set of subgroups of G containing K and the set of all subgroups of G/K.*
If $K \subseteq S \subseteq G$, and the subgroup of G/K is denoted by S^* then

1. *$S^* = S/K = \phi(S)$,*
2. *$T \subseteq S$ if and only if $T^* \subseteq S^*$, and then $|S : T| = |S^* : T^*|$,*
3. *$T \trianglelefteq S$ if and only if $T^* \trianglelefteq S^*$ and then $S/T \cong S^*/T^*$.*

Proof. We will prove parts one and two since the third part is not pertinent for our results. To see that ϕ is one-to-one, let S and T be subgroups of G that both contain K and assume $S/K = T/K$. We wish to show that $S = T$. Let $s \in S$. Then $sK = tK$ for some $t \in T$. Then $t^{-1}sK = K$, so $t^{-1}s \in K$ and thus $s = kt$ for some $k \in K$. Therefore $s \in K \subseteq T$. Similarly, $s^{-1}t \in K$, so $t = sk'$ for some $k' \in K \subseteq S$. So $S = T$.

1. Let A be a subgroup of G/K. Define $S = \phi^{-1}(A)$. Note that the preimage of a subgroup of the codomain is a subgroup of the domain. Also, S must contain K by definition. Lastly, $S/K = \phi(S) = \phi(\phi^{-1}(A)) = A$, since ϕ itself is onto.

2. Let $T \subseteq S$. Then suppose $tK \in T/K$. Then since $t \in S$, $tK \in S/K$. So $T/K \subseteq S/K$. Let $t \in T$. then $tK = sK$ for some $s \in S$. So using the same argument as in the proof the map is one-to-one, $T \subseteq S$. To see that $|S : T| = |S^* : T^*|$, define the map $\psi : S/T \to S^*/T^*$ by $\psi(sT) = \phi(s)T^*$. Let $\psi(aT) = \psi(bT)$. Then $\phi(a)T^* = \phi(b)T^*$ and so $aKT^* = bKT^*$. Thus $aT = bT$. So ψ is one-to-one. To see that ψ is onto, let $cT^* \in S^*/T^*$. Then since ϕ is onto, there exists

some $d \in S$ with $\phi(d) = c$. Hence, $cT^* = \phi(d)T^*$. But then $\psi(dT) = \phi(d)T^*$. Thus ψ is onto. So $|S : T| = |S^* : T^*|$.

□

Definition 2.83. A finite ***p*-group** is a group whose order is p^n for some $n \geq 1$.

Definition 2.84. A finite ***p*-subgroup** of a group G is any subgroup of G that is also a p-group.

Definition 2.85. Let G be a group with $|G| = m \cdot p^n$ where $n \geq 1$ and $p \nmid m$. A **Sylow *p*-subgroup** is any p-subgroup of G whose order is p^n. The set of Sylow p-subgroups of G is denoted $\mathrm{Syl}_p(G)$. We write P is a Sylow p-subgroup of G as $P \in \mathrm{Syl}_p(G)$.

Lemma 2.86. *Let P be a finite p-group acting on a finite set A. Let $A_0 = \{a \in A : g \cdot a = a \text{ for all } g \in P\} \subseteq A$. Then $|A_0| \equiv |A| (\mathrm{mod}\ p)$.*

Proof. Note that $A = A_0 \cup \mathcal{O}_{a_1} \cup \cdots \cup \mathcal{O}_{a_k}$ is a series of disjoint unions where the a_i are representatives of orbits with size greater than 1, since A_0 is the disjoint union of orbits of A about a. The Orbit-Stabilizer Relation implies that $p \mid |\mathcal{O}_{a_i}|$ for all $i = 1, 2, \ldots, k$ since P is a p-group. Therefore $p|(|A| - |A_0|)$. □

Lemma 2.87. *Let G be a group such that $|G| = p^n m$ where $p \nmid m$. Let H be a p-subgroup of a finite group G. Then $|N_G(H)| \equiv |G : H| (\mathrm{mod}\ p)$.*

Proof. Note that $|H| = p^i$ for some $i = 1, 2, \ldots n$. So $|G : H| = mp^{n-i}$. Also, $|N_G(H)| \mid |G|$ and $|H| \mid |N_G(H)|$ by Lagrange's Theorem. So, $|N_G(H)|$ is mp^j for some $i \leq j$. But then $|N_G(H)| = mp^j \equiv mp^{n-i} (\mathrm{mod}\ p)$. Hence $|N_G(H)| \equiv |G : H| (\mathrm{mod}\ p)$. □

The Sylow Theorems. *Let p be a prime, G a finite group and suppose $p \mid |G|$.*

1. *Let $|G| = p^n m, n \geq 1, p \nmid m$. Then for each i, $1 \leq i \leq n$, there is a subgroup of G of order p^i. Every subgroup of the order p^i is a normal subgroup of a subgroup of order p^{i+1}, $1 \leq i \leq n-1$. (In other words, G contains a Sylow p-subgroup.)*

2. *If $P \in Syl_p(G)$ and Q is any other p-subgroup of G then $Q \subseteq kPk^{-1}$ for some $k \in G$.*

3. *Let $|G| = p^n m, n \geq 1, p \nmid m$ and let n_p be the number of Sylow p-subgroups in G. Then:*

   (a) $n_p \equiv 1 \pmod{p}$, and

   (b) $n_p \mid |G|$.

Proof.

1. Proceed by induction on i. If $i = 1$, G has a subgroup of order p by Cauchy's Theorem. Suppose G has a subgroup of order p^i for some i, $1 \leq i \leq n-1$. Then $|G : H| = p^{n-i}m$. Therefore $p \mid |G : H|$ which implies that $p \mid |N_G(H) : H|$ (by Lemma 2.87). But $H \leq N_G(H)$ so consider $N_G(H)/H$. By Cauchy's Theorem, $N_G(H)/H$ has an element of order p. Hence $N_G(H)/H$ has a subgroup of order p. By the Correspondence Theorem, there exists a subgroup H_1 of $N_G(H)/H$ such that $H \leq H_1 \leq N_G(H)$ and $|H_1 : H| = p$. So $H \trianglelefteq H_1$ and $\frac{H_1}{H}$ has order p. Therefore, $|H_1| = p^{i+1}$. Thus the result is proven.

2. Let $P \in \mathrm{Syl}_p(G)$. Let Q act on the set $A = G/P$. Then $|A| = |G|/|P| = \frac{p^n m}{p^n} = m$ and $p \nmid m$. Also consider the set $A_0 = \{a \in A : g \cdot a = a \text{ for all } g \in Q\}$. Note that $|A_0| \equiv |A| (\mathrm{mod}\ p)$ by Lemma 2.86. In particular $|A_0| \neq 0$. A typical

element in A is gP. So $gP \in A_0$

if and only if $q \cdot (gP) = gP$ for all $q \in Q$

if and only if $(qg)P = gP$ for all $q \in Q$

if and only if $g^{-1}qgP = P$ for all $q \in Q$

if and only if $g^{-1}qg \in P$ for all $q \in Q$

if and only if $g^{-1}Qg \subseteq P$.

But $|g^{-1}Qg| = |Q| = |P|$ so $g^{-1}Qg = P$. Therefore P and Q are conjugates in G.

3. Let n_p be the number of Sylow p-subgroups of G. Let $P \in \mathrm{Syl}_p(G)$. By (2) the set of Sylow p-subgroups of G are the conjugates of P in G. Since each Sylow p-subgroup is a conjugate of P in G, if Q is a Sylow p-subgroup, then $Q = kPk^{-1}$ for some $k \in G$. Now $N_G(P) = \{g \in G : gPg^{-1} = P\}$, and $|G : N_G(P)|$ is the number of elements of G divided by the number of elements in $N_G(P)$. So $|G : N_G(P)|$ gives the number of elements in G with gPg^{-1} not equal to P. So, $|G : N_G(P)|$ gives the number of subgroups conjugate to P. Thus $|G : N_G(P)| = n_p$. So then $n_p \mid |G|$. Let P act on a set A of all Sylow p-subgroups of G such that $P \times A \to A$ is defined by $g \cdot a = gag^{-1}$. Recall the set $A_0 = \{a \in A : g \cdot a = a \text{ for all } g \in P\}$. So $|A| = n_p$ and $n_p \equiv |A_0| (\mathrm{mod}\ p)$. Let $R \in \mathrm{Syl}_p(G)$ and suppose $R \in A_0$. Then $g \cdot R = R$ for all $g \in P$. This is true if and only if $gRg^{-1} = R$ for all $g \in P$. So $g \in N_G(R)$ for all $g \in P$. So $P \leq N_G(R)$. So R and P are Sylow p subgroups of $N_G(R)$. Therefore by part (2), P and R are conjugate in $N_G(R)$. But, $N_G(R) = \{g \in G : gRg^{-1} = R\}$,

so R can only be conjugate with itself in its normalizer. Therefore $P = R$. So $A_0 = \{P\}$. Hence $n_p \equiv 1 (\text{mod } p)$.

□

Theorem 2.88. *Let G act on a set X. Let X be k-transitive for $k \geq 3$ with $|X| = n$. If G has a regular normal subgroup H then $k \leq 4$, and*

1. *If $k = 3$ then $H \cong \mathbb{Z}_3$ and $n = 3$ or H is an abelian Sylow 2-group and $n = 2^m$.*

2. *If $k = 4$ then $H \cong V$ and $n = 4$.*

Proof. If X is k-transitive with $k \geq 3$, then for some fixed $x \in X$ G_x acts $(k-1)$-transitively on $X \backslash \{x\}$. To see this is true, consider $(x_1, \ldots, x_{k-1})$ and $(y_1, \ldots, y_{k-1})$ where $x_i, y_i \in X \backslash \{x\}$ and these $(k-1)$-tuples are distinct. Consider distinct k-tuples $(x_1, \ldots, x_{k-1}, x)$ and $(y_1, \ldots, y_{k-1}, x)$. Since G acts k-transitively on X, there exists $g \in G$ such that $gx_i = y_i$ for all $i = 1, \ldots, k-1$ and $gx = x$ so $g \in G_x$. Thus G_x acts $k-1$-transitively on $X \backslash \{x\}$. Also, by Theorem 2.78, G_x acts $(k-1)$-transitively on $H^* = H \backslash \{e\}$, where this action is by conjugation.

1. Assume that $k = 3$. Therefore H^* is primitive by Theorem 2.72 with respect to G_x. Note that $|H| \leq 3$. Let $h \in H^*$. To see that $B = \{h, h^{-1}\}$ is a block, let $g \in G_x$. Assume $gB \cap B \neq \emptyset$. Suppose $h \in gB$. Thus either $h = g \cdot h$ or $h = g \cdot h^{-1}$.

   If $h = g \cdot h$, then $h = ghg^{-1}$ and hence $h^{-1} = (ghg^{-1})^{-1} = gh^{-1}g^{-1} = g \cdot h^{-1}$.

   If $h = g \cdot h^{-1}$, then $h = gh^{-1}g^{-1}$ and so $h^{-1} = (gh^{-1}g^{-1})^{-1} = ghg^{-1} = g \cdot h$.

   Thus if $h \in gB$ then $h^{-1} \in gB$. Similarly, if $h^{-1} \in gB$ then $h \in gB$. This implies that $B \subseteq gB$. Since $|H^*|$ is at most 2, $|B| = |gB|$. Thus $B = gB$. Hence B is a block.

Since this is a primitive action, either $H^* = B = \{h, h^{-1}\}$ or $B = \{h\}$. If $|H^*| = 2$ then $H \cong \mathbb{Z}_3$ and $n = 3$. If $|H^*| = 1$ then $|h| = |h^{-1}| = 2$ but then H is a Sylow 2-subgroup which has order 2 and $n = |H| = 2^m$.

2. Assume $k = 4$. So $k - 1 = 3$ and $|H^*| \geq 3$. Recall that V is regular. Since H is regular, and $|H| \geq 4$, some copy of V is contained in H. Denote this $\{e, h_1, h_2, h_3\}$. Since G_x acts 3-transitively on H^*, we can see that G_{x,h_1} (which is the stabilizer of h_1 in the action of G_x on H^*) acts 2-transitively on $H^* \backslash \{h_1\}$. This action is also primitive by Theorem 2.72. Consider $B = \{h_2, h_3\}$. Since the action is 2-transitive, for the 2-tuple, (h_2, h_3), there exists $g \in G_{x,h_1}$ such that $gh_2 = h_3$. So then $gB = B$. If there were another $g' \in G$ with $g' \neq g^{-1}$, then $gB \cap B = \emptyset$. So B is a block. Since $H^* \backslash \{h_1\}$ is primitive, $B = H^* \backslash \{h_1\}$. Therefore $|H| = 4$ and $H \cong V$.

Also, since $n = 4$, and by definition $k \leq n$, $k \leq 4$. □

Theorem 2.89. *Let G be a group acting on X such that the action is faithful and k-transitive where $k \geq 2$. Assume that G_x is simple for some $x \in X$.*

1. *If $k \geq 4$, then G is simple.*

2. *If $k = 3$ and $|X|$ is not a power of 2, either $G \cong S_3$, or G is simple.*

Proof. 1. By Theorem 2.76, G is either simple or G contains a regular normal subgroup H. If H exists then Theorem 2.88 states that $k \leq 4$ and if $k = 4$, $H \cong V$ and $|X| = 4$. So $H \leq S_4$. But then this is a subgroup of S_4 which is 4-transitive (Lemma 2.62). So $H \cong S_4$. The stabilizer H_x for $x \in X$ is S_3. But S_3 is not simple. So G must be simple.

2. Note that by Theorem 2.88 (1), either $G \cong S_3$ or G has a regular normal subgroup $H \cong \mathbb{Z}_3$ and $|X| = 3$ or H is a Sylow 2-subgroup and $|X| = n$ is a power of 2. Assume $G \ncong S_3$. So $H \cong \mathbb{Z}_3$. Therefore $|X| = 3$. The stabilizer of a point in S_3 is $S_2 \cong \mathbb{Z}_2$, which is simple. So S_3 is the exception.

□

3 The Golay Code and Steiner Systems

Some more background is needed in order to present one representation of the Mathieu Groups. This section will explore the binary Golay code and how it relates to the Steiner Systems. A Steiner System is simply a set of subsets which are chosen using certain criteria. The binary Golay code is equivalent to one particular Steiner System which is used to define the largest of the Mathieu Groups.

3.1 Definitions

Definition 3.1. Let X be a finite set of elements. A **binary linear code**, $\mathcal{C}$ based on X, is a subspace of the power set, $P(X)$. This subspace is over the field F_2 where addition is defined by $A + B = (A \cup B) \backslash (A \cap B)$ for all $A, B \in P(X)$. Multiplication is defined by $A \cdot B = |A \cap B| (\text{mod } 2)$ for all $A, B \in P(X)$.[3]

Definition 3.2. Let $\mathcal{C}$ be a binary linear code based on X. The **length of $\mathcal{C}$** is $|X|$.

Definition 3.3. A binary linear code $\mathcal{C}$ is **even** if the cardinality of every nonempty subset in $\mathcal{C}$ is even.

Definition 3.4. A binary linear code $\mathcal{C}$ is **doubly even** if the cardinality of every nonempty subset in $\mathcal{C}$ is divisible by four.

Definition 3.5. The **minimal weight** of a binary linear code $\mathcal{C}$ is the cardinality of the smallest nonempty subset in $\mathcal{C}$.

Definition 3.6. Let $\mathcal{C}$ be the binary linear code over X. The **dual** of $\mathcal{C}$ is $\mathcal{C}^\dagger = \{A \in P(X) : |A \cap B| \equiv 0 \pmod 2 \ \forall B \in \mathcal{C}\}$.

Definition 3.7. A binary linear code $\mathcal{C}$ is **self-dual** if $\mathcal{C} = \mathcal{C}^\dagger$.

Definition 3.8. Let $1 < k < m < n$ be integers. A **Steiner system**, of type $\mathcal{S}(k, m, n)$ is an ordered pair $(X, \mathcal{B})$ where X is a set with n elements and $\mathcal{B}$ is a family of subsets of X; each subset $B \in \mathcal{B}$ has m elements, and every set of k distinct elements from X lies in a unique block, $B \in \mathcal{B}$.
In other words, this system is a set of $\frac{\binom{n}{k}}{\binom{m}{k}}$ subsets of X, each of size m with the property that every set of distinct k elements is contained in one and only one of these subsets. Furthermore, let $Y \subseteq X$ with $|Y| = t \leq k$ and let $X' = X \backslash Y$ and

$$\mathcal{S}' = \{A \backslash Y : A \in \mathcal{S}, Y \subseteq A\}.$$

Then $\mathcal{S}'$ is a set of $(m-t)$-element subsets of the $(n-t)$-element set X'. If B is a $(k-t)$-element subset of an element of $\mathcal{S}'$ then $Y \cup B$ is contained in a unique $A \in \mathcal{S}$. Then $B \subseteq A \backslash Y \in \mathcal{S}'$. So $\mathcal{S}'$ is an $\mathcal{S}(k-t, m-t, n-t)$ Steiner system on X'. Thus $|\mathcal{S}'| = \frac{\binom{n-t}{k-t}}{\binom{m-t}{k-t}}$

Example 3.9. This is an elementary example of a Steiner System. Let X be the set of 9 points in the vector space F^2, of dimension 2, over the field of 3 elements. For the sake of this example, consider $F = \{0, 1, 2\}$. So

$$X = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2)\}.$$

Subset Name	Subset
$B_{(0,0)(0,1)}$	$\{(0,0),(0,1),(0,2)\}$
$B_{(0,0)(1,0)}$	$\{(0,0),(1,0),(2,0)\}$
$B_{(0,0)(1,1)}$	$\{(0,0),(1,1),(2,2)\}$
$B_{(0,0)(1,2)}$	$\{(0,0),(1,2),(2,1)\}$
$B_{(0,1)(1,0)}$	$\{(0,1),(1,1),(2,1)\}$
$B_{(0,1)(1,2)}$	$\{(0,1),(1,0),(2,2)\}$
$B_{(0,1)(2,2)}$	$\{(0,1),(2,0),(1,2)\}$
$B_{(1,0)(0,1)}$	$\{(1,0),(1,1),(1,2)\}$
$B_{(1,0)(1,1)}$	$\{(1,0),(2,1),(0,2)\}$
$B_{(1,1)(1,2)}$	$\{(1,1),(2,0),(0,2)\}$
$B_{(1,2)(1,0)}$	$\{(1,2),(2,2),(0,2)\}$
$B_{(2,0)(0,1)}$	$\{(2,0),(2,1),(2,2)\}$

Table 3: Unique Subsets from Example 3.9

Let $\mathcal{B} = \{x + \alpha y : \alpha \in F\}$ for $x, y \in X$ where each subset $B \in \mathcal{B}$ is disjoint from all other subsets in $\mathcal{B}$. So then for each pair of elements a and b of X, the subset $B_{ab} = \{a + \alpha b\} = \{a, a + b, a + 2b\}$. In this case, there are 9 possible choices for the 2 points determining each subset B_{ab}. For example, for the two elements $(1, 2)$ and $(1, 0)$, $B_{(1,2),(1,0)} = \{(1,2)+0(1,0),(1,2)+1(1,0),(1,2)+2(1,0)\} = \{(1,2),(1+1,2+0),(1+2,2+0)\} = \{(1,2),(2,2),(0,2)\}$. There are 81 possible subsets of this type. The unique subsets are seen in Table 3. As can be expected, several other combinations of elements give the same subsets. The list of equivalent subsets is shown in Table 4. The sets which have one element are not included in $\mathcal{B}$. These are listed in Table 5. So there are only 12 unique subsets out of 81 possible subsets. Computationally, you can arrive at this number by considering the restriction of having 3 unique points in each subset B_{ab} and having a combination of two points from the vector space determining this set, hence we divide by $\binom{3}{2}$. Therefore there are $\frac{\binom{9}{2}}{\binom{3}{2}} = 12$ subsets of X of the form $B_{ab} = \{a + \alpha b\}$. We call this system a $\mathcal{S}(2, 3, 9)$ Steiner System. [14]

Subset	Equivalent Subsets
$B_{(0,0)(0,1)}$	$B_{(0,0)(0,2)} = B_{(0,1)(0,2)} = B_{(0,2)(0,1)} = B_{(0,2)(0,2)} = B_{(0,1)(0,1)}$
$B_{(0,0)(1,0)}$	$B_{(0,0)(2,0)} = B_{(1,0)(1,0)} = B_{(2,0)(1,0)} = B_{(1,0)(2,0)} = B_{(2,0)(2,0)}$
$B_{(0,0)(1,1)}$	$B_{(0,0)(2,2)} = B_{(1,1)(1,1)} = B_{(1,1)(2,2)} = B_{(2,2)(1,1)} = B_{(2,2)(2,2)}$
$B_{(0,0)(1,2)}$	$B_{(0,1)(2,1)} = B_{(1,2)(1,2)} = B_{(1,2)(2,1)} = B_{(2,1)(1,2)} = B_{(2,1)(2,1)}$
$B_{(0,1)(1,0)}$	$B_{(0,1)(2,0)} = B_{(1,1)(1,0)} = B_{(1,1)(2,0)} = B_{(2,1)(1,0)} = B_{(2,1)(2,0)}$
$B_{(0,1)(1,2)}$	$B_{(0,1)(2,1)} = B_{(1,0)(1,2)} = B_{(2,0)(2,2)} = B_{(2,2)(1,2)} = B_{(2,2)(2,1)} = B_{(1,0)(2,1)}$
$B_{(0,1)(2,2)}$	$B_{(0,1),(1,1)} = B_{(1,2)(1,1)} = B_{(1,2)(2,2)} = B_{(2,0)(1,1)}$
$B_{(1,0)(0,1)}$	$B_{(1,0)(0,2)} = B_{(1,1)(0,1)} = B_{(1,1)(0,2)} = B_{(1,2)(0,1)} = B_{(1,2)(0,2)}$
$B_{(1,0)(1,1)}$	$B_{(1,0)(2,2)} = B_{(2,1)(1,1)} = B_{(2,1)(2,2)} = B_{(0,2)(1,1)} = B_{(0,2)(2,2)}$
$B_{(1,1)(1,2)}$	$B_{(1,1)(2,1)} = B_{(2,0)(1,2)} = B_{(0,2),(1,2)} = B_{(0,2)(2,1)} = B_{(2,0)(2,1)}$
$B_{(1,2)(1,0)}$	$B_{(1,2)(2,0)} = B_{(2,2)(2,0)} = B_{(2,2)(1,0)} = B_{(0,2)(1,0)} = B_{(0,2)(2,0)}$
$B_{(2,0)(0,1)}$	$B_{(2,0)(0,2)} = B_{(2,1)(0,1)} = B_{(2,1)(0,2)} = B_{(2,2)(0,1)} = B_{(2,2)(0,2)}$

Table 4: Equivalent Subsets from Example 3.9

Subset Name	Subset
$B_{(0,0)(0,0)}$	$\{(0,0)\}$
$B_{(0,1)(0,0)}$	$\{(0,1)\}$
$B_{(0,2)(0,0)}$	$\{(0,2)\}$
$B_{(1,0)(0,0)}$	$\{(1,0)\}$
$B_{(1,1)(0,0)}$	$\{(1,1)\}$
$B_{(1,2)(0,0)}$	$\{(1,2)\}$
$B_{(2,0)(0,0)}$	$\{(2,0)\}$
$B_{(2,1)(0,0)}$	$\{(2,1)\}$
$B_{(2,2)(0,0)}$	$\{(2,2)\}$

Table 5: One Element Sets from Example 3.9

Definition 3.10. The **intersection triangle** of an $\mathcal{S}(k, m, n)$ Steiner System is defined by the relation

$$N_{t,t}=\begin{cases} \frac{\binom{n-t}{k-t}}{\binom{m-t}{k-t}} \text{ for } 0 \leq t \leq k \\ 1 \text{ for } k \leq t \leq m \end{cases}$$

and for $0 \leq j < t \leq m$, the recursive relation $N_{j,t} = N_{j,t-1} - N_{j+1,t}$.

Definition 3.11. A **binary Golay Code** is a binary linear code $\mathcal{C}$ over a set X with $|X| = 24$. In addition, the dimension of $\mathcal{C}$, which is the number of distinct sets spanning $\mathcal{C}$, must be at least 12 and the minimum weight at least 8.

3.2 Establishing a Relationship

Theorem 3.12. *Let X be a set with $|X| = 24$. Then if $\mathcal{C} \subseteq P(X) = V$ is a binary Golay code,*

1. *$\mathcal{C}$ has dimension 12,*
2. *$\mathcal{C}$ has minimum weight 8,*
3. *The sets of cardinality 8 in $\mathcal{C}$ form an $\mathcal{S}(5, 8, 24)$ Steiner System,*
4. *The words of cardinality 8 span $\mathcal{C}$.*

Proof. 1. Let $\omega \in X$. Let $T_\omega = \{A \in V : |A| \leq 4, \text{ and } |A| = 4 \text{ implies } \omega \in A\}$. The number of sets of order less than 4 is easily computed using combinations. The number of sets of order 4 is computed by finding the number of ways 23

sets can be ordered into three spaces since ω is held fixed. Thus

$$\begin{aligned}|T_\omega| &= \binom{24}{0} + \binom{24}{1} + \binom{24}{2} + \binom{24}{3} + \binom{23}{3} \\ &= 1 + 24 + 276 + 2024 + 1771 \\ &= 4096 \\ &= 2^{12}.\end{aligned}$$

The sum of the sizes of any two elements in T_ω must be less than 8, since any two elements of size 4 have one element, ω, in common. Therefore the sum of two distinct elements of T_ω is not contained in $\mathcal{C}$. Thus, the cosets of $A + \mathcal{C}$ for $A \in T_\omega$ are distinct. So the set of cosets $V/\mathcal{C}$ contains at least these 2^{12} cosets. But $|V| = |P(24)| = 2^{24}$ and $|\mathcal{C}| \geq 2^{12}$. Therefore $|V/\mathcal{C}| \leq \frac{2^{24}}{2^{12}} = 2^{12}$. Hence $|V/\mathcal{C}| = 2^{12}$. So $\mathcal{C}$ has dimension 12 and $A + \mathcal{C}$ as described above is the set of cosets of $\mathcal{C}$ in V.

2. Let $B \in V$ such that $|B| = 4$ with $\omega \notin B$. Then $B \notin T_\omega$, But $B \in A + \mathcal{C}$ for some $A \in T_\omega$. Now $A + B \in \mathcal{C}$ and hence has a minimum weight of 8 or greater. Since $|A| \leq 4$ and $|B| = 4$, the minimum weight of $A + B$ must be 8.

3. Recall that $A+B$ has weight 8. Since $\omega \in A$, $A+B$ contains $\{\omega\} \cup B$. Note that $|\{\omega\} \cup B| = 5$. So each subset of X containing ω which has order 5 is contained in at least one 8-element set in $\mathcal{C}$. Since ω is arbitrary, each 5-element subset of X is contained in at least one 8-element set of $\mathcal{C}$. It remains to show that a 5-element set is contained in only one unique 8-element set of $\mathcal{C}$. Suppose A, B are distinct sets of weight 8 in $\mathcal{C}$ each containing the same 5-element subset of X. But when $|A \cap B| \geq 5$, $0 < |A + B| \leq 6$. This is a contradiction since the

minimum weight of $\mathcal{C}$ is 8. In conclusion, the weight 8 words of $\mathcal{C}$ form the set $\mathcal{S}(5, 8, 24)$.

4. Let $\mathcal{C}'$ be the subspace of $\mathcal{C}$ generated by sets of order 8 in $\mathcal{C}$. Let $A \in V$ with $|A| \geq 5$. Then there exists a set $B \in \mathcal{C}$ such that $|B| = 8$ and $|A \cap B| \geq 5$. Note that $A + B \in V/\mathcal{C}' \subseteq V/\mathcal{C}$. Note that $|A + B| < |A|$. So using that argument for all the sets of this type, $|A + B| \leq 4$. Choose any set $B' \in V/\mathcal{C}'$ with $B' \notin S_\omega$, with $|B| \leq 4$. Then by part 2 of this proof, there exists $A' \in S_\omega$ with $A' + B' \in \mathcal{C}$ and $|A' + B'| = 8$. So $A' + B' \in \mathcal{C}'$ and therefore $V/\mathcal{C}' \subseteq S_\omega$. So $|V/\mathcal{C}'| \leq |S_\omega| = |V/\mathcal{C}|$. But since $\mathcal{C}' \subseteq \mathcal{C}$, $|V/\mathcal{C}| \leq |V/\mathcal{C}'|$. As $|V/\mathcal{C}| \leq |V/\mathcal{C}'|$, and $|V/\mathcal{C}'| \leq |V/\mathcal{C}|$, $|\mathcal{C}| = |\mathcal{C}'|$ and $\mathcal{C} = \mathcal{C}'$ [3].

□

Definition 3.13. Suppose $\mathcal{C}$ be the binary code of length n. Then the **weight enumerator** of this binary code is

$$W_{\mathcal{C}}(X, Y) = (X + Y)^n$$

where X, Y are elements of $\mathcal{C}$.

Theorem 3.14. *Let $\mathcal{S}$ be an $\mathcal{S}(5, 8, 24)$ Steiner system on X and let $\mathcal{C} \subseteq P(X)$ be the code spanned by the $A \in \mathcal{S}$. Then*

1. *$\mathcal{C}$ is self-dual,*
2. *$\mathcal{C}$ is a binary Golay code,*
3. *the words of weight 8 in $\mathcal{C}$ are the elements of $\mathcal{S}$,*
4. *the weight enumerator of $\mathcal{C}$ is $1 + 759Y^8 + 2576Y^{12} + 759Y^{16} + Y^{24}$.*

Proof. 1. In order for $\mathcal{C}$ to be self-dual, for any $A, B \in \mathcal{S}$, $|A \cap B| \equiv 0 (\text{mod } 2)$, or $|A \cap B|$ is even. So consider the number of elements possible in this intersection. Fix $A \in \mathcal{S}$. Let T be a subset of A. Note that the number of elements of $\mathcal{S}$ containing T depend on $|T| = t$. So if $0 \leq t < 5$ then the system behaves as in Definition 3.8 and so the number of elements of $\mathcal{S}$ containing T is $N_t = \frac{\binom{24-t}{5-t}}{\binom{8-t}{5-t}}$. So if $5 \leq t \leq 8$, then by Definition 3.8 every set of 5 distinct elements lies in only one A. Therefore $N_t = 1$. Hence, $N_5 = N_6 = N_7 = N_8 = 1$, and,

$$N_0 = \frac{\binom{24}{5}}{\binom{8}{5}} = 759$$
$$N_1 = \frac{\binom{23}{4}}{\binom{7}{4}} = 253$$
$$N_2 = \frac{\binom{22}{3}}{\binom{6}{3}} = 77$$
$$N_3 = \frac{\binom{21}{2}}{\binom{5}{2}} = 21$$
$$N_4 = \frac{\binom{20}{1}}{\binom{4}{1}} = 5.$$

Define the function $M_{j,k}$ by $M_{j,k} = N_k$ for $k = j$ and for $j \neq k$, $M_{j,k} = M_{j,k-1} - M_{j+1,k}$. To see how many sets we have which when intersected with a set of size k give you a set of size j, use Definition 3.10. The intersection triangle will give us $M_{j,k}$ where $M_{j,k}$ is the $(k+1)$th row and the $(j+1)$th column. See Table 6.

Claim: For each $A \in \mathcal{S}$ and $C \subseteq D \subseteq A$ with $|C| = j$ and $|D| = k$, $M_{j,k}$ is the number of $B \in \mathcal{S}$ with $B \cap D = C$. So it is the number of sets in the Steiner system which when intersected with a set of size k give an intersection of size j. If $j = k$, this is clear since it follows from the definition of subsets

	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$
$k=0$	759								
$k=1$	506	253							
$k=2$	330	176	77						
$k=3$	210	120	56	21					
$k=4$	130	80	40	16	5				
$k=5$	78	52	28	12	4	1			
$k=6$	46	32	20	8	4	0	1		
$k=7$	30	16	16	4	4	0	0	1	
$k=8$	30	0	16	0	4	0	0	0	1

Table 6: Intersection Triangle for $M_{j,k}$

in the Steiner System. Suppose $j < k$ and proceed by induction on $k - j$. Enumerate $C = \{a_1, a_2, \ldots, a_j\}$ and $D = \{a_1, a_2, \ldots, a_k\}$. It suffices to prove that $B \cap D = C$ if and only if $B \cap (D \backslash \{a_k\}) = C$ and $B \cap D \neq C \cup \{a_k\}$ where $B \cap (D \backslash \{a_k\}) = C$ represents $M_{j,k-1}$ and $B \cap D \neq C \cup \{a_k\}$ represents $M_{j+1,k}$. For the base case, consider when $k - j = 1$. Suppose that $B \cap D = C$. Then $B \cap (D \backslash \{a_k\}) \subseteq B \cap D = C$. So $B \cap (D \backslash \{a_k\}) \subseteq C$. Let $x \in C$. So $x \in B \cap D$. Thus $x \in B$. Note that $C = D \backslash \{a_k\}$. Therefore $x \in D \backslash \{a_k\}$. Hence $x \in B \cap D \backslash \{a_k\}$. So $C \subseteq B \cap D \backslash \{a_k\}$. Therefore $C = B \cap D \backslash \{a_k\}$.

Suppose $C = B \cap D \backslash \{a_k\}$ and $B \cap D \neq C \cup \{a_k\}$. Then since $B \cap D \backslash \{a_k\} \subseteq B \cap D$, $C \subseteq B \cap D$. Suppose that $B \cap D \not\subseteq C$. So there exists $y \in B \cap D$ such that $y \notin C$. By hypothesis $y = a_k$. But then $B \cap D = C \cup \{a_k\}$. This is a contradiction. Thus $C = B \cap D$.

Suppose this is true for $k - j \leq n - 1$. Suppose $k - j = n$. Then let $C = B \cap D$. Then $B \cap D \backslash \{a_k\} \subseteq B \cap D = C$. Suppose that $C \neq B \cap D \backslash \{a_k\}$. Then there exists $x \in C$ such that $x \notin B \cap D \backslash \{a_k\}$. So $x = a_k$. But $C = \{a_i\}$ for $1 \leq i \leq j < k$. Thus, $C = B \cap D \backslash \{a_k\}$. Suppose $B \cap D = C \cup \{a_k\}$. Then $C \cup \{a_k\} = B \cap D = C$.

But then $\{a_k\} \subseteq C$. This is not true. Therefore $B \cap D \neq C \cup \{a_k\}$.

Suppose $B \cap D \backslash \{a_k\} = C$ and $B \cap D \neq C \cup \{a_k\}$. Then $C = B \cap D \backslash \{a_k\} \subseteq B \cap D$. Suppose that $C \neq B \cap D$. Then there exists $x \in B \cap D$ such that $x \notin C$. But since $C = B \cap D \backslash \{a_k\}$, $x = a_k$. But then $C \cup \{a_k\} = B \cap D$ which contradicts the hypothesis. Thus $C = B \cup D$. Therefore the number of $B \in \mathcal{S}$ is $M_{j,k-1} - M_{j+1,k} = M_{j,k}$. Considering only the sets in $\mathcal{S}$, we must look at those which intersect sets of size eight. Since $M_{j,8} = 0$ for all odd j, the intersection of each pair of sets in $\mathcal{S}$ must have even cardinality. Therefore $\mathcal{C}$ is spanned by mutually orthogonal elements of $P(X)$ and thus $\mathcal{C}$ is self-orthogonal. Since the orders of the spanning set are divisible by 4, then $\mathcal{C}$ is also doubly even.

Let $\omega \in X$ and suppose S_ω is defined as in Theorem 3.12. Using the same method as in the proof of Theorem 3.12, the cosets $V/\mathcal{C} \subseteq \mathcal{S}_\omega$. So $|V/\mathcal{C}| \leq |\mathcal{S}_\omega| = 2^{12}$. So $\mathcal{C} \geq 2^{12}$. But $\mathcal{C}$ is self-orthogonal, so $|C| \leq |V|^{1/2} = 2^{12}$. Therefore $|\mathcal{C}| = 2^{12}$. Hence $\mathcal{C} = \mathcal{C}^\dagger$, its dual.

2. Recall that $\mathcal{C}$ contains all sets of even order and the code in which each entry is 1 is contained in $\mathcal{C}$. So using Definition 3.13, the Weight Enumerator of $\mathcal{C}$ is $W_\mathcal{C}(1, Y) = (1+Y)^{24}$. But the coefficients A_i for i not divisible by 4 are all zero. So $W_\mathcal{C}(Y, 1) = 1 + A_4 Y^4 + A_8 Y^8 + A_{12} Y^{12} + A_{16} Y^{16} + A_{20} Y^{20} + A_{24} Y^{24}$ where the coefficients are $A_k = \frac{\binom{24}{5}}{\binom{k}{5}}$. In order for this to be the desired weight enumerator, we must show that $A_4 = 0$. Recall the definition of S_ω from Theorem 3.12. In Theorem 3.12, it was shown that $|S_\omega| = 2^{12} = |V \backslash \mathcal{C}|$ and so all the elements of S_ω are congruent to one distinct element modulo $\mathcal{C}$. Suppose $A \in \mathcal{C}$ had weight 4. Then $A = B + C$ for some $C \in \mathcal{C}$. Hence $|C| = |B| = 2$. But then B, C are distinct elements of S_ω which are congruent modulo $\mathcal{C}$. This is a contradiction of

our earlier statement. Therefore $A_4 = 0$. The weight enumerator now displays that the minimum weight of $\mathcal{C}$ is 8. Thus $\mathcal{C}$ is a Binary Golay Code.

3. By Theorem 3.12 the supports, or the nonzero elements, of the words of weight 8 form the Steiner System $\mathcal{S}$.

4. Note that $A_8 = \frac{\binom{24}{5}}{\binom{8}{5}} = 759$. Using similar calculations, we find that $A_{12} = 2576$, $A_{16} = 759$, and $A_{24} = 1$. Therefore the weight enumerator is $W_{\mathcal{C}}(Y) = 1 + 759Y^8 + 2576Y^{12} + 759Y^{16} + Y^{24}$ [3]

□

It is interesting to note the connection between the binary Golay code and the Steiner System $\mathcal{S}(5, 8, 24)$. A binary Golay code uses the most basic of computer languages. Therefore the Mathieu group M_{24} can essentially be represented in a binary code. This would make computations and calculations on a computer much more palatable. The construction of M_{24} using the binary Golay code requires the use of the Miracle Octad Generator, which is a computer program created by R.T. Curtis [4]. This construction can be found in R. Chapman's "Construction of the Golay code: a survey" [3].

Continuing on, we are now able to define the groups which are other representations of the Mathieu groups.

Definition 3.15. Let $(X, \mathcal{B})$ be a Steiner System. An **automorphism of a Steiner System** $(X, \mathcal{B})$ is an isomorphism $\phi : X \to X$ such that $B \in \mathcal{B}$ implies that $\phi(B) \in \mathcal{B}$.

Theorem 3.16. *Let $(X, \mathcal{B})$ be a Steiner System of type $\mathcal{S}(k, m, n)$. Then $Aut(\mathcal{S}(k, m, n))$ is the Automorphism group of the Steiner System.*

Proof. Let $\mathrm{Aut}(X, \mathcal{B})$ denote the set of automorphisms of this Steiner System. Since the identity function $1 : X \to X$ such that $1(a) = a$ for all $a \in X$ is in $\mathrm{Aut}(X, \mathcal{B})$, $\mathrm{Aut}(X, \mathcal{B})$ is nonempty. By definition, the automorphisms of $(X, \mathcal{B})$ are permutations of X and so $\mathrm{Aut}(X, \mathcal{B})$ is a subset of S_X. Let ϕ_1 and ϕ_2 be functions in $\mathrm{Aut}(X, \mathcal{B})$. Since the composition of two automorphisms is still a permutation of X, $\phi_1 \circ \phi_2$ is still an automorphism. Let $B \in \mathcal{B}$. Then $\phi_1 \circ \phi_2(B) = \phi_1(B_1)$ where $B_1 = \phi_2(B) \in \mathcal{B}$. So then $\phi_1 \circ \phi_2(B) = \phi_1(B_1) \in \mathcal{B}$. So $\mathrm{Aut}(X, \mathcal{B})$ is closed under composition. All that needs to be shown is that for any $B \in \mathcal{B}$, $\phi^{-1}(B) \in \mathcal{B}$. Note that since $\phi^{-1} \in S_X$, and $1 = \phi^{-1}\phi$ but also if $m+1$ is the order of ϕ, $\phi^m\phi = \phi^{-1}\phi$ which implies that $\phi^{-1} = \phi^m$ for some integer m. Therefore ϕ^{-1} exhibits the aforementioned property. □

4 The Mathieu Groups

4.1 Group Representations

There are many representations of the Mathieu groups. This paper will present two of these descriptions. First, there is the permutation representation. Consider

the following permutations:

$$A = (1,2,3,4,5,6,7,8,9,10,11),$$

$$B = (5,6,4,10)(11,8,3,7),$$

$$C = (1,12)(2,11)(3,6)(4,8)(5,9)(7,10),$$

$$D = (1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23),$$

$$E = (3,17,10,7,9)(5,4,13,14,19)(11,12,23,8,18)(21,16,15,20,22),$$

$$F = (1,24)(2,23)(3,12)(4,16)(5,18)(6,10)(7,20)(8,14)(9,21)(11,17)(13,22)(19,15),$$

$$G = (1,2,3,4,5,6,7,8,9,10,11)(12,13,14,15,16,17,18,19,20,21,22),$$

$$H = (1,4,5,9,3)(2,8,10,7,6)(12,15,16,20,14)(13,19,21,18,17),$$

$$I = (11,22)(1,21)(2,10,8,6)(12,14,16,20)(4,17,3,13)(5,19,9,18).$$

Definition 4.1. The **Mathieu groups** are $M_{11} = \langle A, B\rangle$, $M_{12} = \langle A, B, C\rangle$, $M_{22} = \langle G, H, I\rangle$, $M_{23} = \langle D, E\rangle$, and $M_{24} = \langle D, E, F\rangle$.

Another representation often used to describe the Mathieu groups is using Steiner systems.

Theorem 4.2. *There exist Steiner systems $\mathcal{S}(5,6,12)$ and $\mathcal{S}(5,8,24)$ such that these systems are unique and $Aut(\mathcal{S}(5,8,24)) = M_{24}$ and $Aut(\mathcal{S}(5,6,12)) = M_{12}$. The one point stabilizers of M_{24} and M_{12} are M_{23} and M_{11} respectively.*

Theorem 4.3. *There exist Steiner systems $\mathcal{S}(4,5,11)$, $\mathcal{S}(4,7,23)$, and $\mathcal{S}(3,6,22)$ such that these systems are unique and $Aut(\mathcal{S}(4,5,11)) = M_{11}$, $Aut(\mathcal{S}(4,7,23)) = M_{23}$, and $Aut(\mathcal{S}(3,6,22)) = Aut(M_{22})$. The one point stabilizers of M_{11}, M_{23} and*

M_{22} are M_{10}, M_{22}, and $PSL(3,4)$ respectively.

Note that the construction of the Steiner systems can begin using the Golay code as referenced before [3]. Also, there is a Steiner system construction for each Mathieu group in J. Rotman's *An Introduction to the Theory of Groups* [13].

Lemma 4.4. *Note that $M_{11} \leq S_{11}$, $M_{12} \leq S_{12}$, $M_{23} \leq S_{23}$, and $M_{24} \leq S_{24}$.*

Proof. By Theorem 3.16 and since all the Mathieu groups are subsets of the symmetric groups of the same degree, $M_{11} \leq S_{11}$, $M_{12} \leq S_{12}$, $M_{23} \leq S_{23}$, and $M_{24} \leq S_{24}$. □

4.2 Orders

Mathieu discovered the groups M_{12} and M_{24} with degrees of 12 and 24 and orders of $8 \cdot 9 \cdot 10 \cdot 11 \cdot 12$ and $3 \cdot 16 \cdot 20 \cdot 21 \cdot 22 \cdot 23 \cdot 24$ respectively. Note that M_{12} and M_{24} are 5-transitive on any finite set. Thus, in M_{12} for example, it is assumed that this permutation group acts on the set of 12 points and the orbit of any one of the 12 points is the entire set. From this, he found that the one point stabilizer of M_{12} is the group M_{11}, the one point stabilizer of M_{24} is M_{23}. Now M_{23} is 4-transitive and the one point stabilizer of M_{23} is M_{22}. Note that by the Orbit-Stabilizer Relation, $|M_{11}| = \frac{|M_{12}|}{12} = 8 \cdot 9 \cdot 10 \cdot 11$. Similarly, $|M_{23}| = 3 \cdot 16 \cdot 20 \cdot 21 \cdot 22 \cdot 23$ and $|M_{22}| = 3 \cdot 16 \cdot 20 \cdot 21 \cdot 22$. Note that we say n is the degree of M_n for $n = 11, 12, 22, 23,$ or 24.

5 Simplicity

5.1 Simplicity of the Mathieu Groups with Prime Degree

The proof of the simplicity of four of the Mathieu groups was provided by Robin Chapman in a note appearing in *The American Mathematical Monthly* in 1995 (544-545) using only Sylow's Theorems and information about permutation groups . According to Chapman, until then, the only proofs of the simplicity of the Mathieu groups which were found in textbooks required many more theorems and tools [2]. We will begin by concentrating on M_{11} and M_{23}.

Let G be a subgroup of S_p where p is a prime number.

Lemma 5.1. *Let $G \leq S_p$. If G is transitive then p divides $|G|$.*

Proof. Suppose G acts on the set $X = \{1, 2, \ldots, p\}$ transitively. Then let $a \in X$. Since G acts transitively, it has only one orbit. So $p = |\mathcal{O}_a| = \frac{|G|}{|G_a|}$. Therefore $p \mid |G|$. □

Lemma 5.2. *Let G be a transitive subgroup of S_p. Then G has a cyclic Sylow p-subgroup.*

Proof. Since G is transitive, $p \mid |G|$. Also, $|G| \mid p!$ which implies $|G| = pk$ where $p \nmid k$. Thus by Sylow's first theorem, G contains a Sylow p-subgroup, P, whose order is p. By Lagrange's Theorem, the only possible subgroups of any of these Sylow p-subgroups have orders of 1 or p. Thus P has no nontrivial proper subgroups. By Theorem 2.81 P is cyclic. □

Lemma 5.3. *Assume G is transitive and $P = \langle (12 \ldots p) \rangle$ is a cyclic Sylow p-subgroup of G. Let $|G| = m$, n_G be the number of Sylow p-subgroups of G and $N_G(P)$ be the normalizer of P in G with $|N_G(P) : P| = r_G$. Then the order of the group G is pr_Gn_G.*

Proof. Clearly, $m = |G| = |P||N_G(P) : P||G : N_G(P)| = pr_G n_G$. □

Lemma 5.4. *The index of the normalizer of $P \in Syl_p(G)$ in G, r_G, is the least positive residue of $\frac{m}{p}$(mod p).*

Proof. Recall that by Sylow's third theorem, $n_G \equiv 1(\text{mod } p)$. In addition to this, $P \leq N_G(P) \leq N_{S_P}(P)$. Now $N_{S_P}(P)$ is the group of all maps

$$\psi : \mathbb{Z}_p \to \mathbb{Z}_p \text{ such that } x \mapsto ax + b \text{ (mod } p)$$

where $p \nmid a$. Therefore, $|N_G(P)| = p(p-1)$ and it follows that r_G divides $p-1$. But since $\frac{m}{p} = r_G n_G$, and $n_G \equiv 1(\text{mod } p)$, this implies that $r_G \equiv \frac{m}{p}(\text{mod } p)$. Since $r_G \mid (p-1)$, r_G is the least positive residue of $\frac{m}{p}(\text{mod } p)$.

□

Lemma 5.5. *Let $G \leq S_p$ be transitive and let $n_G > 1$. Then $r_G > 1$.*

Proof. Assume $n_G > 1$, and proceeding by contradiction, let $r_G = 1$. So there are $n_G(p-1) = m - n_G$ elements of order p in G. Excluding the elements of order p, which are p-tuples and have no fixed points on $\{1, 2, \ldots p\}$, there are at most n_G elements with fixed points. Every stabilizer G_i of $i \in \{1, 2, \ldots p\}$ in G has n_G elements having at least one fixed point. Now G is transitive, so $|\mathcal{O}_i| = p$ for all $i \in \{1, 2, \ldots p\}$. Thus by the Orbit Stabilizer Relation, $|G_i| = \frac{m}{p} = r_G n_G = n_G$. But since there are at most n_G elements in these stabilizers, they must be equal. But the identity is the only element that holds every other element fixed. Thus each stabilizer is trivial. Therefore $n_G = 1$, which is a contradiction.[2] □

Theorem 5.6. *Let G be a transitive subgroup of S_p where p is a prime number. Suppose that $|G| = pnr$ where $n > 1, n \equiv 1(\text{mod } p)$, $r < p$ and r is prime. Then G is simple.*

Proof. By Lemma 5.3, $r = r_G$ and $n = n_G$. Suppose H is a nontrivial normal subgroup of G. Let H act on the set $X = \{1, 2, ..., p\}$. Thus $\mathcal{O}_{Ha} = \{a \in X : x = ha, h \in H\}$. Now G permutes the orbits of H on X since these orbits are blocks and blocks partition X by Theorem 2.71. By the same Theorem, since G is transitive and H is nontrivial, the orbits of H are all the same size and $|\mathcal{O}_{Ha}| > 1$. This implies that $|\mathcal{O}_{Ha}| = p$. Thus H is also transitive. By Lemma 5.2 there exists a $P' \in \mathrm{Syl}_p(G)$ with $P' \leq H$. By the second Sylow Theorem, Sylow p-subgroups are conjugate with each other, so H contains all the Sylow p-subgroups of G. So $|H| = pn_H t = pnt$ and, by Lagrange's Theorem, $pnt \mid pnr$; this implies that $t \mid r$. But r is a prime number and by Lemma 5.5, $t > 1$, so $t = r$. Therefore $|H| = |G|$, making $H = G$. Consequently, G has no nontrivial proper normal subgroups and, therefore, G is simple. □

Theorem 5.7. *The Mathieu groups M_{11} and M_{23} are simple.*

Proof. Now, $M_{11} \leq S_{11}$ and $m_{M_{11}} = |M_{11}| = 8 \cdot 9 \cdot 10 \cdot 11 = 7920$. Also, by Lemma 5.4, $\frac{m_{M_{11}}}{11} = 720 \equiv 5 \pmod{11}$. This implies that $r_{M_{11}} = 5$ and therefore $n_{M_{11}} = 144 > 1$. Thus $n_{M_{11}} \equiv 1 \pmod{11}$, $r_{M_{11}} < 11$, and $r_{M_{11}}$ is prime. By Theorem 5.6 M_{11} is simple. Also, $M_{23} \leq S_{23}$ and $m_{M_{23}} = |M_{23}| = 3 \cdot 16 \cdot 20 \cdot 21 \cdot 22 \cdot 23 = 10200960$. Thus by Lemma 5.4, $\frac{m_{M_{23}}}{23} = 443520 \equiv 11 \pmod{23}$. Therefore, $r_{M_{23}} = 11$ and $n_{M_{23}} = 40320 > 1$. In addition, $n_{M_{23}} \equiv 1 \pmod{23}$, $r_{M_{23}} < 23$, and $r_{M_{23}}$ is prime. By Theorem 5.6 M_{23} is simple. □

In order to prove that M_{11} is simple using the methods employed in J. Rotman's group theory text, we must prove Burnside's Theorem, which is a Theorem in advanced group theory [13]. This makes the proof found in Robin Chapman's paper attractive.

5.2 Simplicity of the Mathieu Groups with Degrees 12, 22, and 24

Following is the proof of the simplicity of M_{12}, M_{22}, and M_{24}.

Theorem 5.8. *The Mathieu groups M_{12}, M_{22} and M_{24} are simple.*

Proof. Now, M_{12} is a faithful, 5-transitive group whose stabilizer at any point is the simple group, M_{11}. Therefore, by Theorem 2.89, M_{12} is simple. The group M_{22} is a faithful 3-transitive group whose stabilizer at any point is the simple group $\mathrm{PSL}(3,4)$. This is the fourth type of finite simple group [9]. The degree of M_{22} is 22, which is not a power of two. Therefore, by Theorem 2.89, M_{22} is simple [13]. Also, M_{24} is 5-transitive whose stabilizer at any point is the simple group M_{23}. Thus by Theorem 2.89, M_{24} is simple. □

Therefore the Mathieu groups are all finite simple groups. As is apparent, the tools necessary to prove that the Mathieu groups M_{11} and M_{23} are simple include only Sylow's Theorems. In order to prove the simplicity of the other three Mathieu groups, we must prove many things about transitivity, primitivity, and multiple transitivity in group actions. This proof uses a long list of theorems and lemmas. It would be a remarkable feat to construct a proof of similar to R. Chapman's in elegance and clarity for the Mathieu groups M_{12}, M_{22}, and M_{24}.

References

[1] Adler, Stuart. Classical papers in group theory : The Mathieu groups. Thesis. University of East Anglia, 2006. Accessed 18 June 2010. <http://www.uea.ac.uk/ crd06wpu/project26.pdf>.

[2] Chapman, Robin, An elementary proof of the Mathieu groups M_{11} and M_{23}, *The American Mathematical Monthly* **102** (1995), 544-45.

[3] Chapman, Robin, Constructions of the Golay codes: a survey, University of Exeter, EX4 4QE Lecture. Exeter, UK. 23 September 1997.

[4] Curtis, R. T., A new combinatorial approach to M_{24}, *Math. Proc. Camb. Phil. Soc.* **79** (1976), 25-42.

[5] Cuypers, Hans, The Mathieu Groups and their Geometries. Lecture Notes. Eindhoven University of Technology, 1994. <http://www.win.tue.nl/$\tilde{\text{h}}$ansc/>.

[6] Duhem, P, Emile Mathieu, His Life and Works, *New York Mathematical Society* **1** (1892), 156-168. Accessed 25 June 2010. Project Euclid. <http://projecteuclid.org/euclid.bams/1183407338>.

[7] Gorenstein, Daniel, *Finite Simple Groups: An Introduction to Their Classification.* New York: Plenum Press, 1982.

[8] Gorenstein, Daniel, *The Classification of Finite Simple Groups*, volume 1. New York: Plenum Press, 1983.

[9] Hurley, J.F. and Rudvalis, A., Finite simple groups, *The American Mathematical Monthly* **84** (1977), 693-714.

[10] Ivanov, A.A, *Geometry of Sporadic Groups I: Petersen and Tilde Geometries.* Cambridge: Cambridge University Press, 1999.

[11] Ivanov, A.A. and Shpectorov, S.V., *Geometry of Sporadic Groups II: Representations and Amalgams.* Cambridge: Cambridge University Press, 1999.

[12] Papantonopoulou, Aigli, *Algebra: Pure and Applied.* Upper Saddle River, NJ: Prentice Hall, 2002.

[13] Rotman, Joseph J. *An Introduction to the Theory of Groups.* 3rd Ed. Dubuque, IA: Wm. C. Brown Publishers, 1988.

[14] Rowland, Todd and Weisstein, Eric W., Steiner System. From *MathWorld*–A Wolfram Web Resource. Accessed 22 September 2010. <http://mathworld.wolfram.com/SteinerSystem.html>.

[15] Solomon, Ron., On finite simple groups and their classification, *Notices of the AMS* **42** (1995), 231-239.