Perfect Codes and the Mathieu Groups

Stephen Zhou

June 2025

Contents

1 Introduction

One of the most important mathematical achievements of the 20th century was the classification of simple finite groups.

Definition 1.0.1. *A group G is simple if G has no normal subgroups.*

This classification took tens of thousands of pages written over several decades by hundreds of authors. Some of the longest papers in math were written as part of this classification. For example, the last major gap of the classification was filled by a monstrous 1221 page paper of Ascherbacher and Smith [AS04], which proved that all the *quasithin* groups had been found. Another famous example is the 255 page proof of the Feit-Thompson Theorem [FT63], which states that groups of odd order are solvable, which proves that the only simple groups of odd order are S_p.

Theorem 1.1. *Any finite simple group is either:*

i) Contained in one of 18 *infinite families of groups, collectively known as the groups of* Lie type

ii) One of 26 *other* sporadic *groups*

The reason we care so much about simple groups is the Jordan-Holder Theorem. Informally, this theorem states that all groups are, in some sense, products of finite groups, and that this product is basically unique. So the simple groups are like the "prime numbers" of groups.

Theorem 1.2 (Jordan-Holder)**.** *Let G be a group. Then G has a composition series; that is, a sequence*

$$G = G_0 \rhd G_1 \rhd \cdots \rhd G_n = 1, \tag{1}$$

where G_i/G_{i+1} is a simple group for all i. Moreover, the composition factors *G_i/G_{i+1} are unique up to isomorphism and ordering.*

However, there is more than one way to multiply groups. That is, it's possible that two groups $G_1 \ncong G_2$ both have a normal subgroup N such that $G_1/N \cong G_2/N \cong H$ for some group H. Determining what groups G have this property for fixed N and H is known as the *extension problem*. The extension problem is very, very, very hard, and it's unlikely to ever be solved. So knowing all the finite simple groups doesn't let us construct all the finite groups, the way knowing the primes lets us build integers. However, it does tell us the possible composition factors for groups, and knowing this turns out to be crucial to studying groups.

The sporadic groups are weird. It seems logical to expect that simple groups would have a nice classification, given how abstract their definition is. But somehow, the sporadic groups are just there, by themselves, without any other groups like them. We will construct the 5 sporadic Mathieu groups $\mathcal{M}_{24}, \mathcal{M}_{23}, \mathcal{M}_{22}, \mathcal{M}_{12}, \mathcal{M}_{11}$. These are the simplest examples of sporadic groups. Along the way, we will explore a few of the bizarre coincidences and connections that make the existence of these simple groups possible.

2 Codes

We will construct $\mathcal{M}_{24}$ as the symmetries of a special error-correcting code called the *Golay code*. We will then find all the other Mathieu groups as subgroups of $\mathcal{M}_{24}$. Of course, we need to define codes first. We will use the most general definition possible.

Definition 2.0.1. *A (block) code C is a subset of $\mathbb{F}_2^n$. The elements of C are called the codewords.*

Why is this an interesting definition? Say that we are sending a message represented as a sequence of bits. The system we use to transmit the message isn't perfect, so it's possible the recipient receives a corrupted version of the message. Specifically, a 0 can become 1 or a 1 becomes a 0, but bits will never be added or deleted. The inspiration for this is how computers send bits. It's relatively easy for a computer to receive the wrong bit, but nearly impossible for it to not receive a bit at all.

An error correcting code is a way to add redundancy to our message so that the recipient can automatically fix errors in transmission. The words in the

original message get broken up into blocks of size n and sent to codewords. The mapping between uncoded words and codewords is totally arbitrary, so we don't care about it and only look at the codewords C.

Definition 2.0.2. *The hamming distance between two vectors* $\mathbf{a} = (a_1, \ldots, a_n)$ *and* $\mathbf{b} = (b_1, \ldots, b_n)$ *is the number of* i *such that* $a_i \neq b_i$.

If a code has minimal Hamming distance d, then we can correct up to $\lfloor \frac{d-1}{2} \rfloor$ errors by sending each element of $\mathbb{F}_p^n$ to the codeword closest to it in terms of Hamming distance. if there are more, we could decrypt to the wrong codeword. We can detect up to $d-1$ errors because we need at least d errors to corrupt a codeword into another codeword. Other than minimal Hamming distance, there are two other things we care about in codes: the length of the original message, k and the length of the encoded message, n. We will often say a code with these parameters is a $[n, k, d]$ code. The vast majority of codes are not very interesting: if we just pick a subset of points at random, they probably won't make a very good code. So we need to add some structure to our codes.

Definition 2.0.3. *A code* C *is* linear *if it is a subspace of* $\mathbb{F}_2^n$ *over* $\mathbb{F}_2$. *We call a matrix* G *whose rows span* C *a generator matrix.*

Example 2.0.1. *The* n*-repetition code* $C = \{(0, \ldots, 0), (1, \ldots, 1)\}$ *over* $\mathbb{F}_2^n$ *is a linear* $[n, 1, n]$ *code with generator matrix* $[1\,1 \ldots 1]$.

Of course, there are more complicated codes:

Example 2.0.2. *The* $(7, 4)$ *Hamming code is the* $[7, 4, 3]_2$ *code with generator matrix*

$$G = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{2}$$

This code is small enough that we can just write down all 2^4 elements and check their distance from each other to prove that the minimal distance between codewords is 3. Actually, we just need to check distance from 0.

Definition 2.0.4. *The weight of* $\mathbf{a} \in \mathbb{F}_2^n$, *denoted as* $w(\mathbf{a})$ *is the number of* 1 *entries in* $\mathbf{a}$.

Theorem 2.1. *Let* C *be a linear* $[n, k, d]$ *code over* $\mathbb{F}_2^n$. *Then* d *is the minimum weight the elements of* C.

Proof. Let $\mathbf{a}, \mathbf{b} \in C$ have minimal Hamming distance from each other. Notice that the Hamming distance between $\mathbf{a}$ and $\mathbf{b}$ is just the weight of $\mathbf{a} + \mathbf{b}$. Since C is linear, $\mathbf{a} + \mathbf{b} \in C$. □

Both of these examples have the very nice property that there is no "wasted space" because any element of $\mathbb{F}_2^n$ can be corrected into a codeword by changing at most $(d-1)/2$ bits. That is, every element of $\mathbb{F}_2^n$ is within $(d-1)/2$ hamming distance of a unique codeword. (Insert here) (Visualization with 3-repetition code) We should give this property a name.

Definition 2.1.1. *A $[n, k, d]$ code C is perfect if every element $\mathbb{F}_2^n$ is within $(d-1)/2$ Hamming distance of a codeword.*

The perfect codes are basically the best codes we can get. We can use a simple counting argument to find which parameters $[n, k, d]$ a perfect binary code can have.

Theorem 2.2. *If there exists a $[n, k, d]$ perfect binary code, then*

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{(d-1)/2} = 2^{n-k} \tag{3}$$

Proof. We need exactly 2^k codewords, one for each k-bit binary string. The code is perfect, so we need the Hamming spheres of radius $(d-1)/2$

There are $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{(d-1)/2}$ ways to choose at most $\frac{d-1}{2}$ bits to change, each Hamming sphere around the codewords has that many bits. Thus

$$2^k(\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{(d-1)/2}) = 2^n. \tag{4}$$

Dividing by 2^k, we get the Theorem, □

When does this happen? If $n = d$ or $n = (d-1)/2$, the condition is obviously met. These correspond to the code with one element and the n-repetition codes for odd n, respectively. What other $[n, k, d]$ work? With a bit more effort, we can find that $n = 2^r - 1, k = 2^r - r - 1, d = 1$ works. It turns out there exists a family of codes with these parameters, called the Hamming codes. We have already seen the $(7, 4)$ Hamming code. For a construction of these codes, see [CS99].

Aside from these cases, extensive computer searches have only found two more. Specifically:

$$\binom{90}{0} + \binom{90}{1} + \binom{90}{2} = 2^{12} \tag{5}$$

and

$$\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} = 2^{11}. \tag{6}$$

The $n = 90, d = 5$ case doesn't correspond to a code.

Theorem 2.3 ([CS99])**.** *There is no $[90, 11, 5]_2$ code.*

Proof. Assume that C is a $[90, 11, 5]$ code. We have just shown that C must be perfect. Notice that the Hamming distance in $\mathbb{F}_2^n$ is unaffected by translation. So we can assume that $0 \in C$. Since C is perfect, it must contain no other codewords of weight less then 5. The elements of weight 3 in $\mathbb{F}_2^{90}$ must be within Hamming distance 2 of a codeword. This codeword must be of weight 5. Consider the vectors of weight 3 that begin with two 1's. We will write these as $(1, 1, 0, \ldots, 0) + \mathbf{e}_i$, where $\mathbf{e}_i$ has zeroes in all places except for the ith. These must be within two bits of a codeword, which must be of form

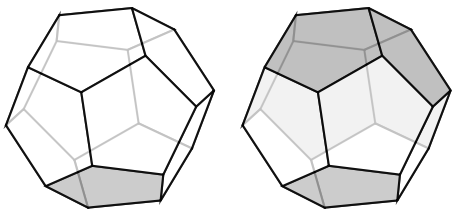
$(1, 1, 0, \ldots, 0) + \mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k$. All the different choice of i, j, k must be distinct because the minimal Hamming distance is 5. But 3 does not divide 88, so this is impossible for all i. Thus there is no $[90, 11, 5]$ Hamming code. □

So now the only remaining case is $n = 23, d = 8$. The code corresponding to this is the Golay code G_{23}. Moreover, the Golay code turns out to be the only $[23, 12, 7]$ code up to permuting the coordinates. Since the Golay code happens to be unique, we could just construct it by going through all 2^{23} elements of $\mathbb{F}_2^{23}$ in some order, and choosing every element that differs in 3 or more places from the previously chosen elements. But that doesn't tell us anything about the code. So instead of doing that, we will construct the closely related *extended Golay code* G_{24}, which has block size $n = 24$. G_{24} has the property that removing any one of its coordinates gives us G_{23}, up to permuting coordinates.

There are many constructions of the Golay code. The following is taken from [Fre]. A standard way to transmit codes is to send the original message, then follow it up with some redundant bits that can be used to correct errors. The bits making up the original message are called the *information bits*, and the added bits are called the *padding bits*. For example, we send two padding bits identical to the information bit in the 3-repetition code. In the case of the extended Golay code, we need 12 information bits and 12 padding bits.

Take a look at a dodecahedron. It has 12 faces. We will construct the Golay code by putting an information bit and a padding bit on the $12 \cdot 2 = 24$ facts of two dodecahedrons. Number the faces of the first dodecahedron from 1 to 12 like below. On the second dodecahedron, just add 12 to each face. This numbering is completely arbitrary, but we need to choose a particular numbering to find a generating matrix for G_{24}.

Given an element $\mathbf{a} \in \mathbb{F}_2^{12}$, we will put one bits on each face corresponding to the element of $\mathbb{F}_2^{24}$ the word is are coded as. For the ith face, the information bit on the first dodecahedron will just be the ith entry of $\mathbf{a}$. The padding bit on a face f of the second dodecahedron is determined as the sum of the bits on the first dodecahedron not adjacent to the face in the same position as f, including that face itself. For example, if we want a first data bit to be 1, and all the others 0, we place bits on the dodecahedron as below, where the shaded faces are 0 and the others 1.



In general, if we want one data bit to be a 1 and the rest to be 0, we get rotations of the diagram above. This gives us the following generating matrix G for G_{24}.

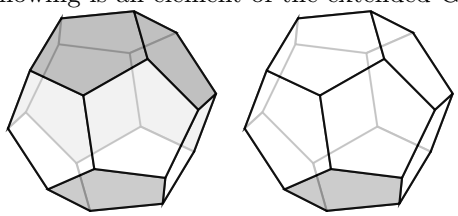
$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1
\end{pmatrix}$$

We will now prove G_{24} is a $[24, 12, 8]$ code. It's clear that $n = 24$. Because the first twelve bits form the identity matrix, the subspace this matrix generates is 12 dimensional, so $k = 12$. It remains to show that $d = 8$. Since the padding bits are just a linear sum of data bits, the extended Golay code is linear. So we just need to prove that the minimum number of ones in an element of the Golay code is 8 . We can just do casework on the number of data bits that are 1.

Theorem 2.4. *The minimal Hamming distance between two elements of the extended Golay code is* 8*.*

Lemma 2.4.1. *If* $\mathbf{a} = (a_1, \ldots, a_{12}, \ldots, a_{24}) \in G_{24}$*, then* $\mathbf{a}' = (a_{13}, a_{14}, \ldots, a_1, a_2, \ldots, a_{12}) \in G_{24}$*.*

Proof. This means that swapping the two dodecahedrons is an automorphism of the Golay code We just need to prove this for all the generators given by the rows of A. Since all the generators are the same up to rotation, we only need to prove this for a single generator, which is equivalent to showing that the following is an element of the extended Golay code.



We can just compute that this is true. □

Proof. We have to do a lot of casework. If both the padding and data bits have more than 4 ones, then the weight is at least 8. So we can assume one of them has less than 4 ones. The Lemma lets us swap data and padding bits, so assume data bits have less than 4 ones. We can just do casework on the number of ones.

If there is only one one, then we have a row of A, which is of weight 8.

If we have two ones, we need to consider how rows of A intersect. Put the bits on a dodecahedron as before.

If the two bits are antipodal, then we get a word of weight 12.

Else, it's easy to see that we get a word of weight 8.

If we have three ones, there are three cases. Either two rings are antipodal, in which case we get a codeword of weight 8, the 3 padding bits share a single 1 bit, or none do. Either way, we can check to see that the codewords have weight at least 8.

□

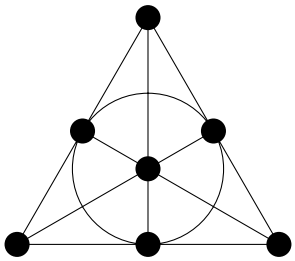
The codewords of G_{24} of weight 8 are called *octads*. The set of octads is called $\mathscr{C}_8$. Now we have proven the extended Golay code G_{24} is an $[24, 12, 8]$ code. We claim that removing any one bit gives us a perfect $[23, 12, 7]$ code. The code clearly still is linear. Since $d = 8$, no elements of the extended code differ in exactly one place. So there are still 2^{12} elements of the code. Thus we have a perfect $[23, 12, 7]$ code. We haven't proven that these $[23, 12, 7]$ codes are all the same though. So for now, say that G_{23} is G_{24} with the last bit removed.

3 Steiner Systems

The octads of G_{24} have a structure that helps explain why the Mathieu groups are so special.

Definition 3.0.1 ([Cur25])**.** *A Steiner system $S(t, k, n)$ is a set with n elements S, along with a set of k-element subsets of S, called blocks, such that every t-element subset of S is contained in exactly 1 block.*

For example, the *Fano Plane* is a $S(2, 3, 7)$ Steiner system, with every line, including the circle, a block.



We don't know many Steiner systems. Specifically, we have no examples of $S(t, k, n)$ Steiner systems with $t > 5$, and only a small number of $S(5, k, n)$ Steiner systems are known. It was recently shown that $S(t, k, n)$ Steiner systems existed for all t, but the proof is nonconstructive. So it should be surprising that $\mathscr{C}_8$ has the structure of a Steiner system.

Definition 3.0.2. *An octad is a codeword of the extended Golay code with weight 8.*

Theorem 3.1. *For each octad of the extended Golay code, choose an 8-element subset of $\{1, 2, \ldots, 24\}$ by choosing i to be in the subset if and only if the ith digit of the corresponding octad is 1. These subsets form a Steiner system $S(5, 8, 24)$.*

Proof. We just need to show that every vector $\mathbf{v} = (v_1, v_2, \ldots, v_{24}) \in \mathbb{F}_2^{24}$ of weight 5 has Hamming distance 3 to exactly one octad.

First we show that there exists such an octad. Without loss of generality, assume that the first coordinate of $\mathbf{v}$ is 0. Consider $\mathbf{v}' = (v_2, v_2, \ldots, v_{23}) \in \mathbb{F}_2^{23}$. Since the $[23, 12, 7]$ Golay code G_{23} is perfect, $\mathbf{v}'$ is within 3 Hamming distance of a codeword of G_{23}, say $\mathbf{c}$. The weight of $\mathbf{c}$ must be 7 or 8. Since G_{23} is found by removing the first entry of G_{24}, we can add that entry back to get a codeword $\mathbf{c}' \in G_{24}$. Since the minimal weight of a word in G_{24} is 8, and the weight of a codeword must be even, the first entry of $\mathbf{c}'$ is 1 if $\mathbf{c}$ has weight 7, and 0 if $\mathbf{c}$ has weight 8.

Now we prove that octad must be unique. If $\mathbf{v}$ was contained in both the octads corresponding to codewords $\mathbf{a}$ and $\mathbf{b}$, then $\mathbf{a}$ and $\mathbf{b}$ can differ in at most 6 places, which isn't possible because G_{24} is a $[24, 12, 8]$ code. □

We sometimes will refer to the specific $S(5, 8, 24)$ Steiner system we found here as just $S(5, 8, 24)$. This is fine because the $S(5, 8, 24)$ Steiner turns out to be unique, which we will prove later. When our proofs apply specifically to this Steiner system, and not $S(5, 8, 24)$ Steiner systems in general, we will abuse notation a bit and call it $\mathscr{C}_8$.

We found a generating matrix for G_{24} consisting of all octads in (insert). Thus any permutation that sends octads of G_{24} to each other also maps G_{24} to itself. And since permuting coordinates doesn't change the weight of any codeword, any permutation that sends G_{24} to itself must also map octads to each other. So $\mathcal{M}_{24}$ is also the group of automorphisms of the $S(5, 8, 24)$ Steiner system of octads.

We need to investigate the structure of Steiner systems a bit more to prove some facts about G_{24}.

Theorem 3.2. *In an $S(t, k, n)$ Steiner system, any set of $i < t$ elements is contained within*

$$\frac{\binom{n-i}{t-i}}{\binom{k-i}{t-i}} \tag{7}$$

blocks.

Proof. By the definition of $S(t, k, n)$, each t element set is contained in a unique block of k elements. There are $\binom{n-i}{t-i}$ subsets of t elements that contain a particular i element subset. Each of these corresponds to a single $k$element block. A particular block has $\binom{k-i}{t-i}$ subsets that contain a certain i element subset, so we need to divide by that. □

Using this theorem, we can find the *Todd triangle* corresponding to any $S(t, k, n)$. These Pascal's triangle-like structures tell us how many blocks intersect x element subsets in certain ways. Specifically, for any fixed subset $x_1, \ldots, x_{l-1}$, the mth entry of the lth row from the top tells us how many blocks intersect in exactly $x_1, x_2, \ldots, x_{m-1}$ with $x_1, \ldots, x_{l-1}$. After the $t+1$th line, we have to assume that there exists a block containing $x_1, \ldots, x_{l-1}$, so everything

doesn't become 0. We will show how to do this for $S(5,8,24)$. Theorem 3.2 lets us find the last row: $759 = \frac{\binom{24}{5}}{\binom{8}{5}}, 253 = \frac{\binom{23}{4}}{\binom{7}{4}}$ and so on, until we reach the sixth row. Then we fill in ones, since we assume there is a block containing all these elements. Now we can fill in the rest of the rows. The first entry of the second row from the top is the number of rows that don't contain some x_1, which we find as $759-253 = 506$. The second entry of the third row is the number of rows that contain x_1 but not x_2. The number of rows that contain both x_1 and x_2 is 77, and the number of rows that contain x_1 is 253, so we fill in $253 - 77 = 176$. Continuing like this, we get the whole triangle as shown below.

										759												
								506				253										
							330			176			77									
						210			120			56			21							
					130			80			40			16			5					
				78			52			28			12			4			1			
			46			32			20			8			4			0			1	
		30			16			16			4			4			0			0			1
	30			0			16			0			4			0			0			0			1

We mostly care about the last row. All the entries corresponding to octads intersecting in an odd number of spaces are zero. So two octads always intersect at an even number of points. This tells us a fact about the extended Golay code.

Theorem 3.3. *The weight of any codeword of G_{24} is divisible by 4.*

Proof. Notice that $w(\mathbf{a}+\mathbf{b}) = w(\mathbf{a}) + w(\mathbf{b}) - 2(\mathbf{a}\cdot\mathbf{b})$ over $\mathbb{F}_2^n$. Assume that $\mathbf{a},\mathbf{b} \in G_{24}$. Since $\mathbf{a}\cdot\mathbf{b}$ corresponds to the intersection of $\mathbf{a}$ and $\mathbf{b}$, it must be even, so $w(\mathbf{a}+\mathbf{b}) \equiv w(\mathbf{a})+w(\mathbf{b}) \pmod 4$. Since G_{24} is generated by codewords with weight $8 \equiv 0 \pmod 4$, the weights of all codewords are divisible by 4. □

This implies that the only possible weights of elements of G_{24} are $0, 8, 12, 16, 24$. G_{24} cannot have a word of weight 4, because the minimal weight of an element is d. Since $(1,1,\ldots,1) \in G_{24}$, there cannot be any words of weight 20 either, as summing that word with $(1,1,\ldots,1)$ would give a word of weight 4. In general, a word of weight i corresponds to a word of weight $24-i$. We can find the *weight distribution* of the code from this.

Theorem 3.4 ([Cur25])**.** *Let $\mathscr{C}_i$ be the set of words of weight i in G_{24}. Then*
i) $|\mathscr{C}_1| = 1$
ii) $|\mathscr{C}_8| = 759$
iii) $|\mathscr{C}_{12}| = 2576$
iv) $|\mathscr{C}_{16}| = 759$
v) $|\mathscr{C}_{24}| = 1$

Proof. The union of the $\mathscr{C}_i$ listed above make up the whole Golay code, which has 4096 elements. There obviously is only one element each of weight 1 and 24, so i) and v) have been proven. We have already proven that $\mathscr{C}_8$ has 759 elements, so $\mathscr{C}_{16}$ does as well. Thus $\mathscr{C}_{12}$ has $4096 - 2\cdot 759 - 2 = 2576$ elements, □

We defined $\mathcal{M}_{24}$ as the automorphism group of an $S(5,8,24)$ Steiner system. We naturally can think of $\mathcal{M}_{24}$ as acting on the set of 24 coordinate positions. We will define the Mathieu groups $\mathcal{M}_{24-i}$ as the stabilizers of i coordinates, and prove that $\mathcal{M}_{24-i}$ is also the automorphism group of another Steiner system.

We must be careful about the definition of a stabilizer of a set of points. There are two kinds: standard stabilizers, and pointwise stabilizers.

Definition 3.4.1. *Let G act on X. The* stabilizer *of $S \subseteq X$ is the set of all $g \in G$ such that for all $x \in X$, $x \cdot g \in X$.*

Definition 3.4.2. *Let G act on X. The* pointwise stabilizer *of $S \subseteq X$ is the set of all $g \in G$ such that for all $x \in X$, $x \cdot g = x$.*

$\mathcal{M}_{23}$ and $\mathcal{M}_{22}$ have the simplest definitions:

Definition 3.4.3. *$\mathcal{M}_{23}$ is the stabilizer of a point in $\mathcal{M}_{24}$ acting on $X = \{1, 2, \ldots, 24\}$.*

Definition 3.4.4. *$\mathcal{M}_{22}$ is the pointwise stabilizer of any two points in $\mathcal{M}_{24}$ acting on $X = \{1, 2, \ldots, 24\}$.*

Of course, we haven't proven that different choices of points will lead to isomorphic $\mathcal{M}_{23}$ and $\mathcal{M}_{22}$. We will prove this after we show $S(5,8,24)$ is unique. But even without proving everything is well that $\mathcal{M}_{23}$ and $\mathcal{M}_{22}$ are also the automorphism groups of a $S(4,7,23)$ and $S(3,6,22)$ Steiner system. $\mathcal{M}_{23}$ is the automorphism group of G_{24} with an element removed, that is, G_{23}. Alternatively, we could take all the octads of $\mathscr{C}_8$, remove those not containing a certain point, and remove that point from the octads that do contain it. These octads form a $S(4,7,23)$ Steiner system. Viewed as vectors, its easy to see that these blocks generate G_{23}, so any automorphism of these blocks is also an automorphism of G_{23}. Thus,

Theorem 3.5. *$\mathcal{M}_{23}$ is the automorphism group of an $S(4,7,23)$ Steiner system.*

Similarly, we can prove that

Theorem 3.6. *$\mathcal{M}_{22}$ is the automorphism group of an $S(3,6,22)$ Steiner system.*

$\mathcal{M}_{24}, \mathcal{M}_{23}, \mathcal{M}_{22}$ are collectively known as the *large Mathieu groups.*

The definitions of the *small Mathieu groups* $\mathcal{M}_{12}$ and $\mathcal{M}_{11}$ is a bit more complicated. A *dodecad* is a word of weight 12.

Definition 3.6.1. *$\mathcal{M}_{12}$ is the stabilizer of a dodecad D in G_{24}.*

Definition 3.6.2. *$\mathcal{M}_{11}$ is the stabilizer of a point of $\mathcal{M}_{12}$ acting on 12 points.*

Theorem 3.7. *$\mathcal{M}_{12}$ is the automorphism group of a $S(5,6,12)$ Steiner system.*

Proof. Let $X = \{1, 2, \ldots, 24\}$. For any 5 points in X/D, there exists exactly one octad O_1 of G_{24} containing those points. We claim that $|X/D \cap O_1| = 6$. For if $|X/D \cap O_1| = 5, 7, 8$, that would imply that $|X/D \oplus O_1| = 10, 6, 4$,

respectively. Thus the $X/D \cap O_1$ are the blocks of a $S(5,6,12)$ Steiner system. An automorphism fixing D must fix X/D as well, and thus the blocks $X/D \cap O_1$. Conversely, its easy to see that the choices of O_1 generate G_{24} as a vector space, so any permutation of the $X/D \cap O_1$ also fixes G_{24}. □

Similarly to the proof of Theorem 3.5, we can now show that.

Theorem 3.8. *$\mathcal{M}_{11}$ is the automorphism group of a $S(4,5,11)$ Steiner system.*

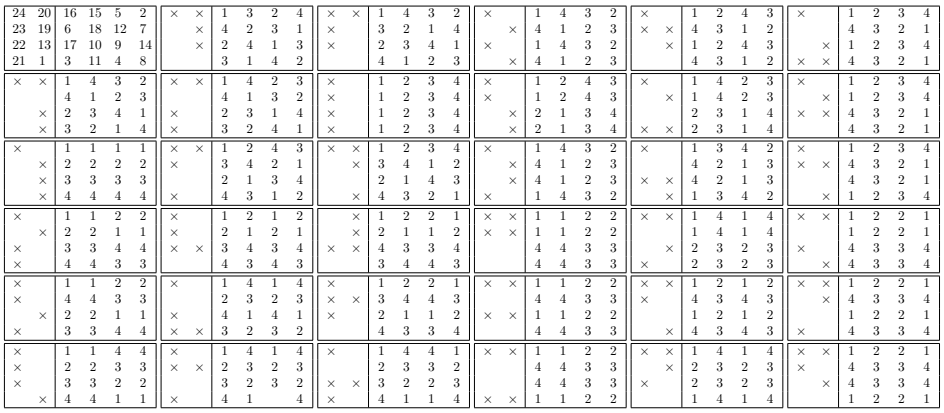
In conclusion, we have proven that all the Mathieu groups are automorphism groups of Steiner systems.

Theorem 3.9. *i) $\mathcal{M}_{23}$ is the automorphism group of an $S(5,8,24)$ Steiner system. ii) $\mathcal{M}_{23}$ is the automorphism group of an $S(4,7,23)$ Steiner system. iii) $\mathcal{M}_{22}$ is the automorphism group of an $S(3,6,22)$ Steiner system. iv) $\mathcal{M}_{12}$ is the automorphism group of a $S(5,6,12)$ Steiner system. v) $\mathcal{M}_{11}$ is the automorphism group of a $S(4,5,11)$ Steiner system.*

4 The MOG

So far, we don't have a nice way of deciding if a set of 8 elements forms an octad of $\mathcal{M}_{24}$. Historically, not being able to calculate efficiently in $\mathcal{M}_{24}$ this has been a problem. When Mathieu discovered his groups in the mid 19th century, he was unable to prove that they were distinct from alternating groups. The existence of these groups was contested over the next eighty years, until Carmichael found another construction of them in 1931. But still, the standard way of finding octads was just to check this list of them. (insert) So even checking if a permutation was in $\mathcal{M}_{24}$ would require 759 separate checks. Nowadays, a computer can do all these calculations near instantly, but computers weren't widely available in the 1960's, when people started getting interested in the Mathieu groups again. Obviously, this made doing anything with $\mathcal{M}_{24}$ very difficult.

This all changed when Curtis discovered the *Miracle Octad Generator*, or MOG, in 1976. The MOG is basically a diagram of $S(5,8,24)$ that takes advantage of its symmetries to avoid just listing out all 759 octads. The diagram consists of the 36 4×6 blocks shown below. We will construct this diagram and explain how to find octads with it. The first block shows us the correspondence between indices of $\mathbb{F}_2^{24}$ and spaces in the blocks. 35 of these blocks let us find in a way we will explain later.



We need to take a look at the structure of $S(5, 8, 24)$. Call any set of four points a *tetrad*. Given a tetrad T_1 contained within an octad U, the triangle tells us there are exactly 4 other octads that intersect that octad in exactly that tetrad. The intersection of these 4 octads with the sixteen elements not in U divides them evenly into 4 sets of *special tetrads* of 4 elements each. For if any of these special tetrads intersected with each other, their corresponding octads would intersect at more than 5 points, which contradicts the fact that any 5 points are contained in a unique octad. We call the chosen tetrad, its complement in U, and the four special tetrads a *sextet*.

We care about sextets because they let us find lots of octads.

Theorem 4.1. *Let U be an octad, and T_1 a tetrad in U with complement T_2. Let $T_3 \ldots, T_6$ be the corresponding special tetrads. Then $T_i \cup T_j, i \neq j$ is always an octad.*

Lemma 4.1.1. *Let U_1 and U_2 be octads in the $S(5, 8, 24)$ Steiner system defined as the weight 8 words of the Golay code such that $|U_1 \cap U_2| = 4$. Then the symmetric difference $U_1 \oplus U_2 = U_1 \cup U_2 - U_1 \cap U_2$ is also an octad.*

Proof. If the octads U_1 and U_2 correspond to Golay codewords $\mathbf{v_1}$ and $\mathbf{v}_2$, then $U_1 \oplus U_2$ corresponds to their sum $\mathbf{v}_1 + \mathbf{v}_2$. This sum has weight $|U_1| + |U_2| - 2|U_1 \cap U_2| = 8 + 8 - 8 = 8$, so it also corresponds to an octad. □

Proof. By the definition of special tetrads, $T_1 \cup T_i$ is an octad for all $i \neq 1$. By taking the symmetric difference of $T_1 \cup T_i$ and $T_1 \cup T_j$, we get that $T_i \cup T_j$ is also an octad, since T_i and T_j are disjoint. □

This shows that complementary choices of the first tetrad $T_1 \in U$ give us the same octads. There are $\binom{8}{4}/2$ ways to divide an octad into two complementary tetrads. So there are a total of 140 different special tetrads for a particular octad. These special tetrads have a nice structure.

Theorem 4.2. *The set of special tetrads for a particular octad U forms a Steiner system $S(3, 4, 16)$.*

Proof. Say that two partitions of U into disjoint tetrads, $U = A_1 \cup A_2$ and $U = B_1 \cup B_2$, result in special tetrads differing in exactly one place, say T_1 for $A_1 \cup A_2$ and T_2 for $B_1 \cup B_2$ with $|T_1 \cap T_2| = 3$. There must exist i, j such that $|A_i \cap B_j| \geq 2$. Then the octads $T_1 \cup A_i$ and $T_2 \cup B_j$ intersect in more than 5 points, which is impossible. □

We are now ready to explain what the MOG is. The first diagram has the 24 coordinates arranged into 6 columns, each of which is a tetrad. This grid is divided into three 2×4 octads, called $\Lambda_1, \Lambda_2, \Lambda_3$ in that order.

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

$=$

Λ_1	Λ_2	Λ_3

.

We call Λ_1 the *special brick*, and $\Lambda_2 \cup \Lambda_3$ the *square*. The special brick is separated from the square by a line in each diagram. For each of the 35 ways to split Λ_1 into an octad, we show how the 4 special tetrads, which we label as $1, 2, 3, 4$ partition $\Lambda_2 \cup \Lambda_3$. Given any 5 points, 4 of which are in Λ_1, the MOG tells us how to complete it. But what if there aren't three points in Λ_1?

This is the miraculous part of the miracle octad generator. It turns out that the MOG diagram stays the same even if we swap bricks around. That is, $\mathcal{M}_{24}$ contains a subgroup isomorphic to S_3 that swaps the positions of the bricks $\Lambda_1, \Lambda_2, \Lambda_3$ in the MOG without changing the relative position of elements.

The easiest way to see this is just by writing the generators of this S_3. The permutation swapping Λ_1 and Λ_3 and the permutation Λ_2 and Λ_3 generate this copy of S_3. It's clear from the symmetry of the MOG diagram that swapping Λ_2 and Λ_3 inside the square is an automorphism. (insert)

The permutation swapping Λ_1 and Λ_2 is (insert). We can just check all 12 of the octads corresponding to the generators of the Golay code to see that they get mapped to other octads under σ. Then every other octads are mapped to each other as well, since they correspond to sums of generators in $\mathbb{F}_{24}$. Alternatively, we can just calculate every single octad from those octads given by the MOG.

Theorem 4.3. *In any $S(5, 8, 24)$ Steiner system. The octads that intersect a specific octad U in exactly four points uniquely determine all other octads of the system.*

Proof. Consider all the octads containing any 3 points x_1, x_2, x_3 of an octad U_0. Octads intersect in an even number of points, so pairs of these octads intersect in exactly 4 points. Todd's triangle shows that there are 21 of these. Since these 21 octads intersect each other in 4 points, the $\binom{21}{2} = 210$ symmetric differences of these octads are also octads. These 210 octads are also all distinct. Say that $U_1 \oplus U_2 = U_3 \oplus U_4$, where all U_i contain x_1, x_2, x_3. There are 5 points in U_1 other than the x_i. At most two of these points can be shared with U_2. So there are at least 3 points of U_1 that are also in $U_1 \oplus U_2$. Since $U_1 \oplus U_2 = U_3 \oplus U_4$, one of U_3 or U_4 contains at least two of those points. Say that U_3 does. Then

$U_1 = U_3$ since they intersect in at least 5 points. It follows that $U_2 = U_4$. Todd's triangle again shows that these 210 octads are all the octads not containing the x_i. But for any octad $U_5 \neq U_0$, there exist three $x_i \in U_0$ such that U_5 does not contain any x_i, since octads intersect in at most 5 points. □

The copy of S_3 in $\mathcal{M}_{24}$ lets us find the octads that intersect in 4 points with any Λ_i. Since octads intersect each other in either 0, 2, or 4 points, any octad U must intersect with at least one of the Λ_i in four points. We call this Λ_i the *heavy brick*. By permuting the bricks, we can send the heavy brick to the special brick Λ_1, and then the MOG shows the octad. So the MOG diagram shows us all the octads of $S(5, 8, 24)$.

5 Multiple Transitivity

Recall how we defined the other Mathieu groups as stabilizer subgroups of $\mathcal{M}_{24}$ acting on $X = \{1, 2, \ldots, 24\}$. The reason these groups are well defined is that $\mathcal{M}_{24}$ is 5-transitive.

Definition 5.0.1. *A group G acting on a set X is k-transitive if for any two sets of k elements $a_1, a_2, \ldots, a_k \in X$ and $b_1, b_2, \ldots, b_k \in X$, there exists a g such that $a_i \cdot g = b_i$.*

Definition 5.0.2. *A group G acting on X is sharply k-transitive if for any two sets of k elements $a_1, a_2, \ldots, a_k \in X$ and $b_1, b_2, \ldots, b_k \in X$, there exists exactly one g such that $a_i \cdot g = b_i$.*

Example 5.0.1. *S_n acts sharply n-transitively on $\{1, 2, \ldots, n\}$, and A_n acts $n - 2$ transitively for $n \geq 5$.*

The proof that $S(5, 8, 24)$ is unique will also proves that $\mathcal{M}_{24}$ is 5-transitive on $X = \{1, \ldots, 24\}$. This is interesting because highly transitive groups are rare. In fact, it turns out that the only k-transitive groups for $k > 5$ are the symmetric groups S_n and the alternating groups A_n. [ONa75] gives a proof of this fact as a consequence of the classification of simple finite groups. Unfortunately, no direct proof of this fact is known.

The order of $\mathcal{M}_{24}$ is central to proving it's simplicity. Luckily, the orbit stabilizer theorem tells us a lot of information about the order of k-transitive groups.

Theorem 5.1. *If G acts on X k-transitively, then Stab(x) acting on $X - \{x\}$ is $k - 1$ transitive, and $|Stab(x)| = \frac{|G|}{|X|}$ for any $x \in X$.*

Proof. Since G is transitive, $\text{orb}(x) = X$. By the orbit stabilizer theorem,

$$|G| = |X||\text{Stab}(x)| \tag{8}$$

which is equivalent to $|\text{Stab}(x)| = \frac{|G|}{|X|}$.

Let $a_1, a_2 \ldots, a_{k-1}$ and $b_1, b_2, \ldots, b_{k-1}$ be any $k-1$ points of $X - \{x\}$. Since G acts k-transitively on X, there exists $g \in G$ such that $x \cdot g = x$ and $a_i \cdot g = b_i$. By definition, $g \in \text{Stab}(x)$, so $\text{Stab}(x)$ acts $k-1$-transitively on $X - \{x\}$. □

Theorem 5.2. *If G acting on X is k-transitive and $|X| = n$, then $|G| = n(n-1)\ldots(n-k-1)|Stab(a_1, \ldots a_k)|$, where $a_1, \ldots, a_k$ are any k elements of X.*

Proof. We proceed by induction on k. The base case of $k = 1$ is proven by the previous theorem. Assume that $|G| = n(n-1)\ldots(n-k-2)|\text{Stab}(x_1, \ldots x_{k-1})|$ if G acts on X $k-1$-transitively, where $|X| = n$.

Let H act k-transitively on X. By Theorem (insert), $\text{Stab}(x_k)$ acts $k-1$ transitively on $X - \{x_k\}$, and $|G| = n|\text{Stab}(x_k)|$. But $|\text{Stab}(x_k)| = (n-1)(n-2)\ldots(n-k-2)|\text{Stab}(x_1, \ldots, x_k)|$ by induction, so $|G| = n(n-1)\ldots(n-k-1)|\text{Stab}(a_1, \ldots a_k)|$. □

Corollary 5.2.1. *If G acts on X sharply k-transitively, and $|X| = n$, then $|G| = n(n-1)\ldots(n-k-1)$.*

Proof. By definition, if G acts sharply k-transitive on X, then there is only one $g \in G$ such that $x_i \cdot g = x_i$ for any k elements of G $x_1, x_2, \ldots, x_k$. Specifically, this element must be the identity. So $|\text{Stab}(x_1, \ldots, x_k)| = 1$, which implies $|G| = n(n-1)\ldots(n-k-1)$ by the previous theorem. □

6 S(5,8,24) is Unique

We are finally ready to prove that $S(5, 8, 24)$ is unique. Along the way, we will find the order of $\mathcal{M}_{24}$ and prove that it is 5-transitive. If we can derive the MOG as a diagram of the octads of an arbitrary $S(5, 8, 24)$, possibly with the first diagram relabeled, then we will have shown that there is only one $S(5, 8, 24)$.

First, notice that everything we have proven about the *particular* $S(5, 8, 24)$ Steiner system $\mathscr{C}_8$ is also true for *any* $S(5, 8, 24)$ Steiner system, except for (insert), which relied on the linearity of the Golay code, and the construction of the MOG. Fortunately, it's not too hard to prove that we can still take symmetric differences in an arbitrary $S(5, 8, 24)$.

Theorem 6.1. *For any two blocks A, B in any Steiner system $S(5, 8, 24)$, if $|A \cap B| = 4$, then $A \oplus B$ is also a block.*

Proof. Let $A = \{a_1, \ldots, a_8\}$ and $B = \{a_1, a_2, a_3, a_4, b_5, b_6, b_7, b_8\}$. For the sake of contradiction, assume that $\{a_5, a_6, a_7, a_8, b_5, b_6, b_7, b_8\}$ is not an octad. Consider the unique octad U_1 that contains the points a_5, a_6, a_7, a_8, b_5. Since two octads intersect in an even number of points, U_1 must contain another point of B. This point cannot also be in A, since 4 of the 5 points are already in A, and octads intersect in at most 4 points. So without loss of generality we can assume that U_1 also contains b_6, so $U_1 \supset \{a_5, a_6, a_7, a_8, b_5, b_6\}$.

Similarly, we can assume the octad U_2 containing a_5, a_6, a_7, a_8, b_7 contains b_8, or that $U_2 \supset \{a_5, a_6, a_7, a_8, b_7, b_8\}$. Now consider the octad U_3 that contains a_5, a_6, a_7, b_5, b_7. This octad must contain another point of A, which cannot be a_8. Say it is a_1. U_3 must intersect with B in exactly one more element. U_3 already contains four a_i's, so the new element must be a b_i. We have already found three elements of U_3 that intersect with U_1 and U_2, so this new element must not be in U_2 or U_3. But since $U_2 \cup U_3$ contain all the b_i, we have reached a contradiction, and $\{a_5, a_6, a_7, a_8, b_5, b_6, b_7, b_8\}$ is an octad. □

Before we prove $S(5, 8, 24)$ is unique, we need to understand how sextets intersect with octads.

Theorem 6.2. *An octet intersects with the tetrads of a sextet in one of the following three ways:*

i) In four places with two tetrads

ii) In 2 places with 4 sextets.

iii) In 3 places with one tetrads, and 1 place with the 5 others

Proof. Octets intersect in an even number of points. If two tetrads T_1, T_2 of a sextet intersect an octad U with different parities, $T_1 \cup T_2$ intersects U in an odd number of points, since all tetrads are disjoint. So all tetrads of a sextet must intersect U with the same parity. If the tetrads intersect evenly, then the octad must intersect either as i) or ii). If the tetrads intersect oddly, there is at most one tetrad intersecting in three places with the octad, since the union of two tetrads that intersect the octad in three places would intersect the octad in 6 places, which is impossible. There are only 6 octads, so there must exist at least one tetrad intersecting the octad in 3 places if the tetrads intersect oddly. So iii) is the only remaining choice. □

Now we are ready to prove the main theorem of this section.

Theorem 6.3.

i) $S(5, 8, 24)$ *is unique up to relabeling.*

ii) $\mathcal{M}_{24}$ *acts* 5*-transitively on* $X = \{1, 2, 3 \ldots, 24\}$.

iii) $|\mathcal{M}_{24}| = 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 3 \cdot 16 = 244,823,040$

Proof. Say that we have an arbitrary $S(5, 8, 24)$ Steiner system $\mathscr{C}_8'$. Let $\mathcal{M}_{24}'$ be the subgroup of S_{24} consisting of all automorphisms of $\mathscr{C}_8'$. We know that the MOG diagram determines all the octads of $\mathscr{C}_8$. We will prove that the MOG also shows the octads of $\mathscr{C}_8'$, possibly with a different first diagram, so $\mathscr{C}_8$ and $\mathscr{C}_8'$ are isomorphic. The way we will do is to prove that some of the diagrams are true by computation, and then derive the rest of the diagrams from them using Theorem 6.1.

Let Y be the set of all ordered 7-tuples $(x_1, x_2, x_3, x_4, x_5, x_6, y)$, where the x_i are elements of an octad Λ_1 in $\mathscr{C}_8$, and $y \notin \Lambda_1$. Along with proving $S(5, 8, 24)$ is unique, we will also prove that $\mathcal{M}_{24}'$ acts sharply 1-transitively on Y. Since there are $24 \cdot 23 \cdot 22 \cdot 21 \cdot 20$ ways to choose the first 5 x_i's, 3 ways to choose the

sixth, and 16 ways to choose $y \notin \Lambda_1$, Corollary 5.2.1 implies ii). And since any 5 points are contained in some octad, this also implies iii).

For the first diagram, arrange the 7 points like below. The 21 $\cdot$ marks are undetermined points. We choose to arrange them such that the columns are tetrads of the hexad we get from choosing x_1, x_2, x_3, x_4 as a tetrad of Λ_1.

x_1	x_5	y	$\cdot$	$\cdot$	$\cdot$
x_2	x_6	$\cdot$	$\cdot$	$\cdot$	$\cdot$
x_3	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$
x_4	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$

Because we assumed the columns form tetrads, the MOG diagram D_1 works.

$$D_1 = \begin{array}{|c|cccc|}\hline \times & 1 & 1 & 1 & 1 \\ \times & 2 & 2 & 2 & 2 \\ \times & 3 & 3 & 3 & 3 \\ \times & 4 & 4 & 4 & 4 \\ \hline \end{array}$$

Now consider the MOG diagram corresponding to the partition $\begin{array}{|ccc|}\hline & & \times \\ \times & & \\ \times & & \\ \times & & \\ \hline \end{array}$.
The $\times$ are the places containing elements of the MOG.

The octad U_1 determined by the points $\{x_2, x_3, x_4, x_5, y\}$ must intersect each column of the square in exactly one point. For if it intersected, say, column i in 3 places as well, U_1 would intersect the union of columns 1 and i in more than 5 places, which isn't allowed because any two columns form a octad. So by rearranging the points in the last three columns, we can assume that the top row forms a special tetrad. We can repeat this for the rest of the columns to get the diagram

$$D_2 = \begin{array}{|c|cccc|}\hline \times & 1 & 1 & 1 & 1 \\ \times & 2 & 2 & 2 & 2 \\ \times & 3 & 3 & 3 & 3 \\ \times & 4 & 4 & 4 & 4 \\ \hline \end{array}$$

Since all the permutation we have used preserve columns, D_1 stays accurate as well. Now consider the partition $\begin{array}{|ccc|}\hline & & \times \\ \times & & \\ \times & & \\ \times & & \\ \hline \end{array}$. We now have found all the octads disjoint from Λ_1.

Similarly to before, the octad containing $\{x_1, x_3, x_4, x_5, y\}$ must intersect with each of the square's columns in exactly one place. By permuting the last 3 columns, we can assume that $\begin{array}{|c|cccc|}\hline \times & \times & & & \\ \times & & \times & & \\ \times & & & \times & \\ \times & & & & \times \\ \hline \end{array}$ is a tetrad. Since

$\begin{array}{|c|cccc|}\hline \times & \times & & & \times \\ \times & & & \times & \times \\ \times & & & \times & \times \\ \times & \times & & & \times \\ \hline \end{array}$ is also a octad, we can take the symmetric difference of these to find the octad $\begin{array}{|c|cccc|}\hline \times & & & & \times \\ \times & & & \times & \\ \times & & \times & & \\ \times & \times & & & \\ \hline \end{array}$. If we continue like this, we find that the octads containing $\begin{array}{|c|c|}\hline & \times \\ \times & \\ \times & \\ \times & \\ \hline \end{array}$ are exactly those indicated in

$$D_3 = \begin{array}{|c|cccc|}\hline \times & 1 & 2 & 4 & 3 \\ \times & 2 & 1 & 3 & 4 \\ \times & 4 & 3 & 1 & 2 \\ \times & 3 & 4 & 2 & 1 \\ \hline \end{array}.$$

What permutations are we still allowed to use? That is, what remaining permutations are there that permute the special tetrads of all of D_1, D_2, D_3? We can just check all the $(4!)^2$ permutations that fix D_1 and D_2 (preferably with a computer) to find that the subgroup that also fixes D_3 is generated by

$$a = \begin{array}{|c|cccc|}\hline \times & 1 & 2 & 4 & 3 \\ \times & 2 & 1 & 3 & 4 \\ \times & 4 & 3 & 1 & 2 \\ \times & 3 & 4 & 2 & 1 \\ \hline \end{array}, b = \begin{array}{|c|cccc|}\hline \times & 1 & 2 & 4 & 3 \\ \times & 2 & 1 & 3 & 4 \\ \times & 4 & 3 & 1 & 2 \\ \times & 3 & 4 & 2 & 1 \\ \hline \end{array}, c = \begin{array}{|c|cccc|}\hline \times & 1 & 2 & 4 & 3 \\ \times & 2 & 1 & 3 & 4 \\ \times & 4 & 3 & 1 & 2 \\ \times & 3 & 4 & 2 & 1 \\ \hline \end{array}.$$

Now look at the partition $\begin{array}{|cc|}\hline \times & \times \\ \times & \\ \times & \\ \hline \end{array}$. The octad containing $\{x_1, x_2, x_3, x_5, y\}$ cannot contain points labeled with 1 in D_1, D_2 or D_3. This forces it to either be $\begin{array}{|c|cccc|}\hline \times & & & & \times \\ \times & & & \times & \\ \times & & \times & & \\ \times & \times & & & \\ \hline \end{array}$ or $\begin{array}{|c|cccc|}\hline \times & & & & \times \\ \times & & & \times & \\ \times & & \times & & \\ \times & \times & & & \\ \hline \end{array}$. Since b swaps these, we can assume it is the first. Just like we did with D_3, we can use other octads disjoint from Λ_1 to find that the diagram

$$D_4 = \begin{array}{|cc|cccc|}\hline \times & \times & 1 & 4 & 3 & 2 \\ \times & & 3 & 2 & 1 & 4 \\ \times & & 2 & 3 & 4 & 2 \\ & & 4 & 1 & 2 & 3 \\ \hline \end{array}$$

is valid. Now we only have a and c left to permute with. With the same method,

we conclude that

$$D_5 = \begin{array}{|cc|cccc|} \hline \times & \times & 1 & 4 & 2 & 3 \\ \times & & 3 & 4 & 2 & 1 \\ & & 2 & 1 & 3 & 4 \\ \times & & 4 & 3 & 2 & 1 \\ \hline \end{array}$$

is also a valid diagram. But this time, D_5 is determined entirely and we don't have to use any permutations.

We also get that

$$D_6 = \begin{array}{|cc|cccc|} \hline \times & \times & 1 & 1 & 2 & 2 \\ \times & \times & 1 & 1 & 2 & 2 \\ & & 3 & 3 & 4 & 4 \\ & & 3 & 3 & 4 & 4 \\ \hline \end{array}$$

displays a sextet by using c.

Now using a,we get that

$$D_7 = \begin{array}{|cc|cccc|} \hline \times & \times & 1 & 1 & 3 & 3 \\ & & 2 & 2 & 4 & 4 \\ \times & \times & 1 & 1 & 3 & 3 \\ & & 2 & 2 & 4 & 4 \\ \hline \end{array}$$

also displays a sextet.

Now we have used al the permutations fixing the sextets, so we need all the other diagrams to be determined by these. Its easy to check that they are. For example, we can sum D_7 and D_6 to get

$$\begin{array}{|cc|cccc|} \hline \times & \times & 1 & 1 & 2 & 2 \\ \times & \times & 1 & 1 & 2 & 2 \\ & & 3 & 3 & 4 & 4 \\ & & 3 & 3 & 4 & 4 \\ \hline \end{array} + \begin{array}{|cc|cccc|} \hline & & 1 & 1 & 3 & 3 \\ & & 2 & 2 & 4 & 4 \\ \times & \times & 1 & 1 & 3 & 3 \\ & & 2 & 2 & 4 & 4 \\ \hline \end{array} = \begin{array}{|cc|cccc|} \hline \times & \times & 1 & 1 & 2 & 2 \\ & & 4 & 4 & 3 & 3 \\ & & 4 & 4 & 3 & 3 \\ \times & \times & 1 & 1 & 2 & 2 \\ \hline \end{array}.$$

Thus we have derived the MOG and $S(5, 8, 24)$ is unique. Since we have used up all the permutations fixing D_1, D_2, D_3, $\mathcal{M}_{24}$ acts sharply transitively on $\{x_1, x_2, x_3, x_4, x_5, y\}$ because the stabilizer has □

Since $\mathcal{M}_{24}$ is 5-transitive, the large Mathieu groups are well defined. The orders and transitivity of the large Mathieu groups can be easily found with Theorem (insert).

Theorem 6.4. *$\mathcal{M}_{23}$ acts 4-transitively on 23 elements and $|\mathcal{M}_{23}| = 23 \cdot 22 \cdot 21 \cdot 20 \cdot 3 \cdot 16 = 10200960$.*

Theorem 6.5. *$\mathcal{M}_{22}$ acts 4-transitively on 23 elements and $|\mathcal{M}_{22}| = 22 \cdot 21 \cdot 20 \cdot 3 \cdot 16 = 443520$.*

To prove the small Mathieu groups are well defined, we will show $\mathcal{M}_{24}$ acts transitively on dodecads, and that $\mathcal{M}_{12}$ acts sharply 5-transitively on the remaining 12 points.

Theorem 6.6 ([Cur25]). *$\mathcal{M}_{24}$ acts transitively on dodecads.*

Lemma 6.6.1. *Let $D \in \mathscr{C}_{12}$. Then there exists two octads $U_1, U_2 \in \mathscr{C}_8$ such that, as sets, $|U_1 \cap U_2| = 2$ and $D = U_1 \oplus U_2 = U_1 \cup U_2 - U_1 \cap U_2$.*

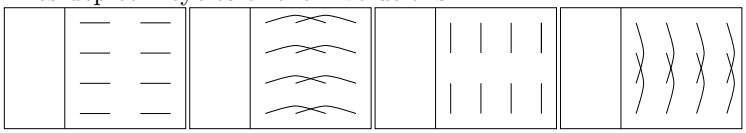
Proof. If $|U_1 \cap U_2| = 2$, it follows that $|U_1 \oplus U_2| = 12$, since it corresponds to the sum of U_1 and U_2 as vectors, so every $U_1 \oplus U_2$ correspond to a dodecad. We can just count the number of ways to make a dodecad with Todd's triangle. There are 759 different octads U_1, and $30\binom{8}{2}$ octads U_2 that intersect that octad in exactly two points by Todd's triangle. $\binom{12}{2}$ of these choices lead to the same dodecad, so $U_1 \oplus U_2$ can be

$$\frac{758 \cdot 30 \cdot \binom{8}{2}}{\binom{12}{2}} = 2576 \tag{9}$$

dodecads, which is exactly the number of dodecads in $\mathscr{C}_{12}$. □

Proof of Theorem 6.6. Consider an arbitrary dodecad $D = U_1 \cup U_2$, where U_1 and U_2 are octads such that $|U_1 \cap U_2| = 2$. Let $|U_1 \cap U_2| = \{x_1, x_2\}$. Since $\mathcal{M}_{24}$ is 5-transitive and 5 points define an octad, there exists $\sigma \in \mathcal{M}_{24}$ sending x_1 and x_2 to the top two points of Λ_1 in the MOG and U_1 to Λ_1. By Todd's triangle, there are 16 choices for the image U_2 that intersects x_1, x_2. This is small enough that we can just find all the possible U_2 and find elements of $\mathcal{M}_{24}$ mapping any choice to another. For example,

We can easily check with the MOG that these following elements are all in $\mathcal{M}_{24}$, and that they fix no choice of U_2. All these elements have order 2, and the lines depict 2-cycles of the involutions.



These elements generate a subgroup of $\mathcal{M}_{24}$ isomorphic to $(\mathbb{Z}/2\mathbb{Z})^4$ of order 16. There are 16 possible So by the Orbit-Stabilizer Theorem, there is only one orbit on the dodecads. □

Theorem 6.7. *$\mathcal{M}_{12}$ is sharply 5-transitive on Ω/D, and $|\mathcal{M}_{12}| = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95040$.*

Proof. There are 2576 dodecads, so $\mathcal{M}_{12} = |\mathcal{M}_{24}|/2576 = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8$. By Theorem (insert), $\mathcal{M}_{12}$ must be sharply 5 transitive if it is 5-transitive at all. Let D be an arbitrary dodecad and $\Omega = \{1, 2, \ldots, 24\}$. We will show that the stabilizer of D acts 5-transitively on X/D. Notice that the stabilizer of D is exactly the stabilizer of X/D. Given two sets of 5 points a_1, a_2, a_3, a_4, a_5 and b_1, b_2, b_3, b_4, b_5 in X/D, there exists a permutation $\pi \in \mathcal{M}_{24}$ such that $a_i \cdot b = y_i$ for $i \leq 5$. Let the unique octads containing these 5 points be A and B, and say the other three points are a_6, a_7, a_8 and b_6, b_7, b_8, respectively. A and B each must intersect Ω/D in exactly 6 points. A and B are already known to intersect X/D in 5 points, and there must be a sixth point for the weight of

their symmetric difference to be divisible by 4. Assume that the sixth point of the intersection with Ω/D is a_6 and b_6 respectively.

By (the proof of) Theorem 6.3, we can assume π maps x_6 to y_6, since we proved that $\mathcal{M}_{24}$ is transitive on any six points within an octad and one point without. Since B intersects D in 6 points, there exists another octad B' with $B \cap B' = \{b_7, b_8\}$ such that $B \oplus B' = X/D$. By the previous Theorems proof, there exists a subgroup of $\mathcal{M}_{24}$ fixing B isomorphic to $\mathbb{Z}/2\mathbb{Z}$, which acts transitively on all the octads intersecting B in $\{b_7, b_8\}$. So one of these elements, say τ, maps B to B'. Then $\pi \cdot \tau$ fixes D, and sends a_i to b_i. Thus $\mathcal{M}_{12}$ acts 5-transitively. $\square$

7 Simplicity

Now we will prove the Mathieu groups are simple. The normal way to show this is through the following theorem.

Theorem 7.1 ([Rot95])**.** *Let G be a group acting faithfully and k-transitively on X. Assume there exists some $x \in X$ such that Stab(x) is simple. Then:*

i) If $k \geq 4$, then G is simple.

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

The proof of this theorem requires quite advanced group theory, so we'll skip it. But it reduces proving the Mathieu groups are simple to proving that their stabilizer is simple. Actually, we only need to prove that the stabilizer of $\mathcal{M}_{22}$ and $\mathcal{M}_{11}$ is simple.

Theorem 7.2 ([Rot95])**.** *The stabilizer of a point in $\mathcal{M}_{22}$ is the simple group $PSL_3(4)$.*

Unfortunately, explaining what $\mathrm{PSL}_3(4)$ means, or proving its the stabilizer of $\mathcal{M}_{22}$ is beyond the scope of this paper. But it relates to the projective lines, an example of which is the famous Fano plane shown previously. Another construction of the Golay codes and $S(5, 8, 24)$ uses these lines. For more on the relation between them, see [Cur25].

There is a much simpler proof that $\mathcal{M}_{23}$ and $\mathcal{M}_{11}$ are simple discovered by Chapman. [Cha95] When combined with Theorem 7.4, it proves that all the Mathieu groups are simple. Recall the Sylow theorems. These will be all we need.

Definition 7.2.1. *A p-group is a groups whose order is divisible by p.*

Definition 7.2.2. *A* Sylow p-subgroup *is a maximal p-subgroup.*

Theorem 7.3 ((insert))**.** *1. Let G be a group such that $|G| = p^n m$, where $m \nmid p$. Then there exists a sylow p-subgroup of order p^n.'*

2. All sylow p-subgroups are conjugates for fixed p.

3. Let n_p be the number of Sylow p-subgroups. Then $n_p | m$, $n_p \equiv 1 \pmod{p}$, and $n_p = |G : N_G(P)|$, where P is any Sylow p-subgroup.

Also define $r_p = |N_G(P) : P|$, so that we have $|G| = |P||N_G(P) : P||G : N_G(P)| = pn_pr_p$. Now we are ready to prove $\mathcal{M}_{23}$ and $\mathcal{M}_{11}$ are simple.

Theorem 7.4. *Let G be a subgroup of S_p acting transitively on $X = \{1, 2, \ldots, p\}$. If $|G| = pnr$, where $n > 1, n \equiv 1 \pmod{p}$, and $r < p$ prime, then G must be simple.*

Lemma 7.4.1. *Let G be a subgroup of S_p that acts transitively on $X = \{1, 2, \ldots, p\}$ with sylow p-group P. If $n_p > 1$, then $r_p > 1$.*

Proof. Say that $n_p = r_G = 1$. By the second sylow theorem, we can assume P is generated by $(12 \ldots p)$. There are $n_p(p-1) = n - n_G$ elements of order p. These elements can fix no points of X, a set with p elements. But the stabilizer G_i for $i \leq p$ must have n_p elements, so every stabilizer must be the same. The only way this can happen is if $n_p = 1$, a contradiction. □

Proof. It is clear that $r_p = r, n_p = n$. Let H be a nontrivial normal subgroup. Since G is transitive, all the orbits of H on $X = \{1, 2, \ldots, p\}$ have the same size. Since H is nontrivial, this size cannot be 1, and since p is prime, the size must be p, meaning H is also transitive. So H must contain some sylow p-subgroup, which is also a sylow p-subgroup of G. H is a normal group, closed under conjugation, and thus contains all p-subgroups. So H has the same number of sylow p-subgroups as G, so the order of H is pm_pt. We need $t < r$. Since r is prime and $t < 1$ by the lemma, $t = r$ and $G = H$. □

This theorem can easily be seen to apply to $\mathcal{M}_{23}$ and $\mathcal{M}_{11}$.

Since we have proven that $\mathcal{M}_{23}$ and $\mathcal{M}_{11}$ are simple, Theorem 7.4 proves that $\mathcal{M}_{24}$ and $\mathcal{M}_{12}$ are also simple, but we need to assume (insert) to show that $\mathcal{M}_{22}$ is simple.

References

[FT63] Walter Feit and John G. Thompson. "Solvability of groups of odd order". English. In: *Pac. J. Math.* 13 (1963), pp. 775–1029. ISSN: 1945-5844. DOI: 10.2140/pjm.1963.13.775.

[ONa75] Michael E. O'Nan. "Normal structure of the one-point stabilizer of a doubly-transitive permutation group. II". English. In: *Trans. Am. Math. Soc.* 214 (1975), pp. 43–74. ISSN: 0002-9947. DOI: 10.2307/1997095.

[Cha95] Robin J. Chapman. "An elementary proof of the simplicity of the Mathieu groups M_{11} and M_{23}". English. In: *Am. Math. Mon.* 102.6 (1995), pp. 544–545. ISSN: 0002-9890. DOI: 10.2307/2974771. URL: hdl.handle.net/10871/14832.

[Rot95] Joseph J. Rotman. *An introduction to the theory of groups.* English. 4th ed. Vol. 148. Grad. Texts Math. New York, NY: Springer-Verlag, 1995. ISBN: 0-387-94285-8.

[CS99] J. H. Conway and N. J. A. Sloane. *Sphere packings, lattices and groups. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.* English. 3rd ed. Vol. 290. Grundlehren Math. Wiss. New York, NY: Springer, 1999. ISBN: 0-387-98585-9.

[AS04] Michael Aschbacher and Stephen D. Smith. *The classification of quasithin groups. II: Main theorems: the classification of simple QTKE-groups.* English. Vol. 112. Math. Surv. Monogr. Providence, RI: American Mathematical Society (AMS), 2004. ISBN: 0-8218-3411-8.

[Cur25] Robert T. Curtis. *The art of working with the Mathieu Group M_{24}.* English. Vol. 232. Camb. Tracts Math. Cambridge: Cambridge University Press, 2025. ISBN: 978-1-00-940567-6; 978-1-00-940568-3. DOI: 10.1017/9781009405683.

[Fre] Peter Freyd. *The Golay Codes.* https://www2.math.upenn.edu/~pjf/golay.pdf.