Utrecht University

Department of Mathematics

---

mathematics bachelor thesis

A Formalisation of the Mathieu-24 Group

Supervisor:
Johan Commelin

Candidate:
Edward van de Meent

July 11, 2024

Abstract

Group theory is the study of symmetries of objects. An important theorem in this field is the Classification theorem of Finite Simple Groups. This statement has been proven over 20 years ago, and yet the full proof has not yet been completely published, in part due to its size. The proof spans more than 10000 pages, and as such, it is hard to completely verify. In order formally prove that this is a true classification with a computer, one first needs to have a formalisation of the group classes, along with a proof that the groups within it are indeed finite and simple. In this thesis, a formalisation will be given of the Mathieu-24 group M_{24}, and a proof of its simplicity, in the Lean proof assistant, using its Mathlib library. Along the way, related concepts such as (Linear) Codes, Steiner systems, R-Metrics and the (Extended) Golay Code will be given a formal definition as well.

Contents

1. Introduction

Between 1955 and 2004, papers have been published to complete a proof of the statement known as the Classification theorem of Finite Simple Groups (CFSG). This was a project in which hundreds of mathematicians participated, and the result is a milestone in finite group theory. The proof itself covers more than 10000 pages, and as a result, multiple errors had snuck in along the way. The size of the proof is not conducive to intuition, understanding, nor trust. Proof assistants can help with this: proof assistants are tools which are able to formally check proofs written in a certain format. In order to make small steps towards this kind of formalization and verification of the classification theorem, we attempt to define one of the sporadic simple groups in the proof assistant Lean 4, making use of its Mathlib library. The concerning group is one of 26 sporadic simple groups, i.e. one of the finite simple groups which does not fit in one of 16 infinite families. The specific sporadic group of concern here is the group M_{24}, which is one of 5 sporadic finite simple groups discovered by the French mathematician Émile Léonard Mathieu in two papers published in 1861 and 1873. The subscript refers to the fact that the group naturally occurs as subgroup of S_{24}. Other sporadic simple groups can be obtained as a subgroup of this one, among which are the other Mathieu-groups: M_{11}, M_{12}, M_{22} and M_{23}. The used construction is the one explained in chapter 5 of *The Finite Simple Groups* by Robert A. Wilson[1]. Similarly, the structure of the proof of simplicity is from there, but with more detailed steps. We will define the group as the maps which preserve the GolayCode, and then show that that linear code contains a steiner system $S(5, 8, 24)$

2. Linear Codes

The group M_{24} can be defined as the automorphism group of a linear code called the *GolayCode*. We will devote this chapter to the construction of this structure, as well mentioning some noted features.

2.1 HexaCode

Definition: a *linear code* C is formally a combination of $(K, V, C, \|\cdot\|)$ where V is a Vectorspace over K), C is a linear subspace of V whose members are called codewords, and $\|\cdot\|$ is a norm on V. Additionally, the norm of non-zero codewords in C is strictly positive, while the norm on a general member of V is strictly less than infinity.
In the cases we are concerned with, we will only consider nontrivial fields and finite vectorspaces, meaning we can endow V with the Hammingnorm or Hammingweight, which is defined as the number of non-zero entries in V, or $\|v\| = \|\{i \in \mathbb{N} | v_i \neq 0\}\|$. In this case, we will trivially have that the norm has finite non-zero value for non-zero members of V. The norm is finite because it is bound by the dimension of V, which must be finite, and the norm is only equal to zero when every entry of V with respect to a basis is equal to 0, making it equal to 0.

The HexaCode is a particular example of a linear code. It is given as the subspace of F_4^6 as a vectorfield over $F_4 = \langle\omega\rangle$, where the subspace of F_4^6 is the span of these elements:

$$b_1 = (\omega, \omega^{-1}, \omega^{-1}, \omega, \omega^{-1}, \omega)$$
$$b_2 = (\omega^{-1}, \omega, \omega, \omega^{-1}, \omega^{-1}, \omega)$$
$$b_3 = (\omega^{-1}, \omega, \omega^{-1}, \omega, \omega, \omega^{-1})$$

Alternatively, it is the span given by these three elements:

$$\begin{aligned} a_1 &= \omega \cdot b_1 + \omega^{-1} \cdot b_2 && = (1,0,0,1,\omega^{-1},\omega) \\ a_2 &= \omega^{-1} \cdot b_1 + \omega^{-1} \cdot b_2 + b_3 && = (0,1,0,1,\omega,\omega^{-1}) \\ a_3 &= b_2 + b_3 && = (0,0,1,1,1,1) \end{aligned}$$

This last triple must be independent, as their first three indices form an identity matrix. Since they are all linear combinations of the first triple, they must be codewords, and as a result, the dimension of the subspace of codewords is at least 3. However, as the subspace is also generated by 3 codewords, the dimension must be equal to 3, and hence both of these triples must be bases. The fact that we have a basis means that we can characterise codewords as follows:

lemma 2.1.1: For all $x \in F_4^6$, x is a codeword exactly when $x = (x_1, x_2, x_3, x_1 + x_2 + x_3, \omega^{-1}x_1 + \omega x_2 + x_3, \omega x_1 + \omega^{-1}x_2 + x_3)$.

proof: The right hand side of the equation can be written as follows: $(x_1, x_2, x_3, x_1 + x_2 + x_3, \omega^{-1}x_1 + \omega x_2 + x_3, \omega x_1 + \omega^{-1}x_2 + x_3) = x_1 \cdot a_1 + x_2 \cdot a_2 + x_3 \cdot a_3$. If x is a codeword, there is a unique way to write x as a linear combination of (a_1, a_2, a_3), because it is a basis. That means there is a unique triple $y_1, y_2, y_3 \in F_4$ such that $(x_1, x_2, x_3, x_4, x_5, x_6) = y_1 \cdot a_1 + y_2 \cdot a_2 + y_3 \cdot a_3 = (y_1, y_2, y_3, y_1 + y_2 + y_3, \omega^{-1}y_1 + \omega y_2 + y_3, \omega y_1 + \omega^{-1}y_2 + y_3)$. Then by extensionality, we must have that $x_1 = y_1$, $x_2 = y_2$ and $x_3 = y_3$, and so the equation must hold. For the other direction, it suffices to see that the right hand side is always a codeword as it is a linear combination of a basis of the codewords. □

lemma 2.1.2 The weight of a nonzero codeword in the HexaCode is 4 or 6.

proof: Notice that we can simplify the equation of lemma 2.1.1 by associating with an element $v \in GolayCode$ a polynomial $f(x) = v_1 * x^2 + v_2 * x + v_3$, allowing us to say that $v = (v_1, v_2, f(0), f(1), f(\omega), f(\omega^{-1}))$. Now, let us consider the following cases: If $v_1 = v_2 = 0$ and $v \neq 0$, f is constant

and non-zero, making the weight equal to 4. If $v_1 = 0$ but $v_2 \neq 0$, f is a linear polynomial, with one root exactly at $-v_3/v_2$, making the weight of the codeword again equal to 4. If $v_1 \neq 0$ but $v_2 = 0$, it is the case that there is again a single solution, given by $v_3^2 * v_1$, giving v weight 4. Lastly, if both v_1 and v_2 are non-zero, because f is a degree 2 polynomial, there are at most two roots, meaning the weight of v is at least 4, and at most 6. There cannot be exactly one root in this case, as then $f(x) = v_1(x - \alpha)^2$, making b equal to zero, which is a contradiction. This allows us to conclude that the weight cannot be 5 in this case, which concludes our proof. □

2.2 The Extended Binary Golay Code

The Extended Binary Golay Code, or just GolayCode for short, is given by a subspace of $(\mathbb{Z}/2)^{6\times|F_4|} = (\mathbb{Z}/2)^{24}$, typically described as a 4×6 matrix where the rows are indexed by elements of F_4, enumerated from top to bottom as $(0, 1, \omega, \omega^{-1})$, and the columns are indexed by numbers 1 through 6. We identify these matrices with the set of indexes where they are non-zero (which is more commonly known as its *support*), creating a bijection between $(\mathbb{Z}/2)^{6\times|F_4|}$ and $\mathcal{P}(6 \times F_4)$. This identification turns addition into symmetric difference, while scalar multiplication is intersection with either or the full set $6 \times F_4$, depending on if the scalar is 0 or 1 respectively. The codewords of the GolayCode can be defined in several equivalent ways.

Definition: $m \in (\mathbb{Z}/2)^{6\times|F_4|}$ is a codeword of the GolayCode exactly when the following two properties hold:

- for every $i \in \mathbb{N}_{\geq}$ with $1 \leq i \leq 6$, $\sum_{x\in F_4} m_{(i,x)} = \sum_{1\leq j\leq 6} m_{(j,0)}$, or in words, the parity of each column is equal to the parity of the first row.
- there is some *codeword* $y \in F_4^6$ of the HexaCode such that for every $1 \leq i \leq 6$, $\sum_{x\in F_4}(m_{(i,x)} \cdot x) = y_i$, or in words, the per-column sum of indices where the value of m is non-zero (and therefore 1) is a codeword in the HexaCode. note that the multiplication $(M_{(i,x)} \cdot x)$ here is a *scalar multiplication* between an element of $\mathbb{Z}/2$ and an element of F_4, given by the interpretation of F_4 as a field extension of F_2, or $\mathbb{Z}/2$.

It is relatively easy to verify that this definition indeed gives a linear subspace. The first property is obviously linear. The second property is linear too, which is more obvious when one considers taking the sum per column as evaluating the sum $\sum_{x \in \mathbb{F}_4} v_{(i,x)} \cdot x$, as well as the fact that $(a + b) \cdot x = a \cdot x + b \cdot x$ holds.

As it turns out, this definition allows us to calculate a basis of the codewords of the code. Suppose we know the following values of a codeword of the GolayCode:

$$\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & ? \\ \blacksquare & \blacksquare & \blacksquare & ? & ? & ? \\ \blacksquare & \blacksquare & \blacksquare & ? & ? & ? \\ \blacksquare & ? & ? & ? & ? & ? \end{pmatrix}$$

Then, we know the parity of the columns and the first row, as the first column is completely filled in. From this, we can determine the values at $(2, \omega^{-1})$ and $(3, \omega^{-1})$, as well as $(6, 0)$. This means we know the following values:

$$\begin{pmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare & ? & ? & ? \\ \blacksquare & \blacksquare & \blacksquare & ? & ? & ? \\ \blacksquare & \blacksquare & \blacksquare & ? & ? & ? \end{pmatrix}$$

Because we completely know the values of the first three columns, we can determine the first three values of the corresponding codeword in the HexaCode, which according to lemma 3.1.1, completely determines the codeword. At this point, we want to be able to determine the value of a column by knowing its parity, the value at the first index, and the sum of the indices where it is nonzero. We can do this via a clever trick, using (ω, ω^{-1}) as a basis for F_4:

lemma 2.2.1: For all $m \in (\mathbb{Z}/2)^{6 \times |F_4|}$ and all $1 \leq i \leq 6$, if $\sum_{x \in F_4} m_{(i,x)} = p$ and $\sum_{x \in F_4} (m_{(i,x)} \cdot x) = a \cdot \omega + b \cdot \omega^{-1}$,

then $m_i = (m_{(i,0)}, m_{(i,0)} + a + b + p, m_{(i,0)} + b + p, m_{(i,0)} + a + p)$.
Proof: Firstly, lets rewrite the sum-notation:

$$\sum_{x \in F_4} m_{(i,x)} = p = m_{(i,0)} + m_{(i,1)} + m_{(i,\omega)} + m_{(i,\omega^{-1})}$$

$$\begin{aligned}\sum_{x \in F_4} (m_{(i,x)} \cdot x) &= a \cdot \omega + b \cdot \omega^{-1} \\ &= m_{(i,0)} \cdot 0 + m_{(i,1)} \cdot (\omega + \omega^{-1}) + m_{(i,\omega)} \cdot \omega + m_{(i,\omega^{-1})} \cdot \omega^{-1} \\ &= (m_{(i,1)} + m_{(i,\omega)}) \cdot \omega + (m_{(i,1)} + m_{(i,\omega^{-1})}) \cdot \omega^{-1}.\end{aligned}$$

Since (ω, ω^{-1}) forms a basis, from the second equation we can conclude that $a = m_{(i,1)} + m_{(i,\omega)}$ and $m_{(i,1)} + m_{(i,\omega^{-1})}$. Substituting these equations for p, a and b, results in the following:

$$\begin{pmatrix} m_{(i,0)} \\ m_{(i,0)} + a + b + p \\ m_{(i,0)} + b + p \\ m_{(i,0)} + a + p \end{pmatrix} = \begin{pmatrix} m_{(i,0)} \\ m_{(i,1)} + m_{(i,\omega)} + m_{(i,\omega^{-1})} + a + b \\ m_{(i,1)} + m_{(i,\omega)} + m_{(i,\omega^{-1})} + b \\ m_{(i,1)} + m_{(i,\omega)} + m_{(i,\omega^{-1})} + a \end{pmatrix}$$

$$= \begin{pmatrix} m_{(i,0)} \\ m_{(i,\omega^{-1})} + b \\ m_{(i,1)} + m_{(i,\omega)} + m_{(i,\omega^{-1})} + b \\ m_{(i,\omega^{-1})} \end{pmatrix} = \begin{pmatrix} m_{(i,0)} \\ m_{(i,1)} \\ m_{(i,\omega)} \\ m_{(i,\omega^{-1})} \end{pmatrix} = m_i$$

□

Using this lemma, we can finally fill in the rest of the unknown values in m, to conclude that the values at these indices completely determine the codeword. Now, choosing a single one of the starting indexes to be 1 and setting

the rest at 0, results in the following 12 (linearly independent) elements:

$$b_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}; \quad b_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}; \quad b_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix}$$

$$b_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}; \quad b_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}; \quad b_6 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

$$b_7 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{pmatrix}; \quad b_8 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}; \quad b_9 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$b_{10} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}; \quad b_{11} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}; \quad b_{12} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Because these starting indices completely determine the value of a codeword, and we have established 12 linearly independent codewords which completely span the space of these starting indices, we have found a basis. This basis allows us to give a second definition equivalent of the codewords of the GolayCode, namely the subspace that is the span of these basis elements.

Codewords in the GolayCode are quite spaced out, as the distance between any two different codewords is always at least 8:

Lemma 2.2.2 If $m \in$ GolayCode and $m \neq 0$, then $8 \leq \|m\|$.

The following proof is a more detailed version of one written by Robin J. Chapman[2].

Proof: Suppose that the columns of m have odd parity. Since then the

weight of each column is at least 1, we just need to prove that there must be at least one column with more than 1 non-zero entry. Suppose to the contrary. In that case the projection to the HexaCode is given precisely by the unique non-zero index for that column. Since codewords in the HexaCode have weight either 0,4, or 6, that means there must be an even number of zeroes in the projection to the HexaCode of m. However, that would mean that the parity of the top row is even, which is in contradiction with the assumption that the columns have odd parity. Therefore we must have at least one column with more than 1 non-zero entry, making that column have 3 entries, meaning that the weight of m is at least 8. Now suppose that the columns (and top row) have even parity. If the projection to the HexaCode is non-zero, then we know that there are at least 4 non-zero entries in the projection. Given that the parity of the column is even, this can only occur when the weight of the column is 2, meaning that m has at least 4 columns with weight 2, making the weight of m at least 8. If, on the other hand, the projection to the HexaCode is zero, we know that the weight of a column is either 0 or 4. It cannot be the case that one or less columns have weight zero, because that would mean that $m = 0$ in the first case, or that the parity of the top row is odd in the second case. Therefore, we must have that there are at least two columns with weight 4, meaning that m must have weight at least 8. □

Due to this fact, it is always possible to detect up to 7 errors, and correct up to 3 errors. Furthermore, a simple calculation reveals that if one were to take codewords of the Extended Binary GolayCode, and drop a single dimension, every non-codeword has a unique closest codeword, making it a perfect code. This variant is also sometimes referred to as the GolayCode, which is why in those contexts the code described earlier in this chapter is referred to as "extended".

3. Code Automorphisms

3.1 Semilinear Automorphisms

Definition: a *Semilinear Automorphism of a Linear Code* is a map $\phi : V \to V$, such that the following properties hold:

- There is some Field automorphism σ such that for all $k \in K$ and all $v \in V$, φ conjugates scalar multiplication via σ, i.e $\varphi(k \cdot v) = \sigma(k) \cdot \varphi(v)$,
- It maps addition, i.e. for all $x, y \in V$, $\varphi(x + y) = \varphi(x) + \varphi(y)$,
- It preserves distance, i.e. for all $x, y \in V$, $d(x, y) = d(\varphi(x), \varphi(y))$
- It exclusively and always maps codewords to codewords, i.e. for all $x \in V$, x is a codeword exactly when $\varphi(x)$ is one.
- It has an inverse, or equivalently, it is bijective.

The first two conditions make the map semilinear, the third and left-to-right version of the fourth condition make it a map between codes, and the last condition together with the right-to-left implication of the fourth condition ensure that the inverse function also satisfies all these conditions, making it an Automorphism. As it turns out, these semilinear automorphisms can be composed and have inverses, and the identity map is also always a semilinear Automorphism. In other words, semilinear Automorphisms form a Group. In particular, when assuming a basis and using the HammingNorm associated with that basis, it gives a subgroup of $((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$, where $\phi((f, \sigma), d) = f^{-1} \circ d \circ \sigma$, and we identify elements of $(K^\times)^n$ with their indexing function.

Lemma 3.1.1: The group $((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$ has a faithful group-action on V.

Proof: Define the action via $((d, (f, \sigma)) \cdot x) = f \circ (d \cdot x) \circ \sigma^{-1}$, where we use $d \cdot x$ to denote elementwise multiplication of these vectors. Now, let

$(d,(f,\sigma),(d',(f',\sigma')) \in ((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$. Then we have that the following holds for all $x \in V$:

$$
\begin{aligned}
&(d,(f,\sigma) \cdot ((d',(f',\sigma')) \cdot x) \\
&= (d,(f,\sigma) \cdot (f' \circ (d' \cdot x) \circ {\sigma'}^{-1}) \\
&= f \circ (d \cdot (f' \circ (d' \cdot x) \circ {\sigma'}^{-1})) \circ \sigma^{-1} \\
&= f \circ \left((f' \circ {f'}^{-1} \circ d \circ \sigma' \circ {\sigma'}^{-1}) \cdot (f' \circ (d' \cdot x) \circ {\sigma'}^{-1}) \right) \circ \sigma^{-1} \\
&= f \circ f' \circ \left(({f'}^{-1} \circ d \circ \sigma') \cdot (d' \cdot x) \right) \circ {\sigma'}^{-1} \circ \sigma^{-1} \\
&= (f \circ f') \circ \left(\left(({f'}^{-1} \circ d \circ \sigma') * d' \right) \cdot x) \right) \circ (\sigma \circ \sigma')^{-1} \\
&= \left(\left(\left({f'}^{-1} \circ d \circ \sigma' \right) * d' \right), \left(\left(f \circ f' \right), \left(\sigma \circ \sigma' \right) \right) \right) \cdot x \\
&= \left(\left(\phi \left((f', \sigma'), d \right) * d' \right), \left(\left(f \circ f' \right), \left(\sigma \circ \sigma' \right) \right) \right) \cdot x \\
&= ((d,(f,\sigma) * (d',(f',\sigma'))) \cdot x
\end{aligned}
$$

Therefore, this action is indeed a group-action. To see that this action is Faithful, one can focus on appropriate choices of elements of $x \in V$. Suppose that $f \circ (d * x) \circ \sigma^{-1} = f' \circ (d' * x) \circ {\sigma'}^{-1}$ for all $x \in V$. Then one can verify that $\sigma(i) = \sigma'(i)$ by noting that when one chooses to evaluate at $x = b_i$, the value of $f \circ (d * x) \circ \sigma^{-1}$ at index $\sigma(i)$ is nonzero (as it is equal to $f(d_i * 1)$). Because f' is a ring-homomorphism, and d' has only units as entries by definition, this means that b_i at index ${\sigma'}^{-1}(\sigma(i))$ is non-zero, which means that $\sigma(i) = \sigma'(i)$.

Next, through similar arguments, it can be concluded that $d_i = d'_i$, by choosing $x = d_i^{-1} \cdot b_i$ and noting that the value of $(d_i,(f,\sigma)) \cdot x$ at index $\sigma(i)$ must be 1.

Finally, we will be able to conclude that $f(a) = f'(a)$ for all $a \in K$ by choosing $x = d_i \cdot (a \cdot b_i)$ and noting that the value of $(d,(f,\sigma)) \cdot x$ at index $\sigma(i)$ is equal to $f(a)$. In conclusion, by extensionality we have that if for all $x \in V$, $(d,(f,\sigma)) \cdot x = (d',(f',\sigma')) \cdot x$, then $(d,(f,\sigma)) = (d',(f',\sigma'))$, which means that the action is faithful.

Let φ be some semilinear automorphism of the code $(K, V, C, \|\cdot\|$. Then, since scalar multiplication in a (nontrivial) vectorspace is faithful, there is a unique field-automorphism f_φ such that $\varphi(k \cdot x) = f(k) \cdot \varphi(x)$ for all $k \in K$ and $x \in V$. Then, note that the weight of $\varphi(b_i)$ must be 1, as the weight is the distance to 0 and φ preserves distance. Therefore, $\varphi(b_i)$ must be a scalar multiple of some basis element. Define σ_φ and d_φ such that $\varphi(b_i) = f_\varphi(d_\varphi(i) * b_{(}\sigma_\varphi(i)))$. Then, by induction over the basis, we can say that for all $x \in V$, $\varphi(x) = f_\varphi \circ (d_\varphi \circ x \circ \sigma_\varphi^{-1})$.

Theorem 3.1.2: The group of semilinear code automorphisms of some code form a subgroup of $((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$.

Proof: First some auxillary definitions. Let φ be some semilinear automorphism of V. Then, since scalar multiplication on a (nontrivial) vectorspace is faithful, there is a unique field-automorphism denoted f_φ such that $\varphi(k \cdot x) = f(k) \cdot \varphi(x)$ for all $k \in K$ and $x \in V$. Then, note that the weight of $\varphi(b_i)$ must be 1, as the weight is the distance to 0 and φ preserves distance. Therefore, $\varphi(b_i)$ must be a scalar multiple of some basis element. Define $\sigma_\varphi \in S_n$ and $d_\varphi \in (K^\times)^n$ such that $\varphi(b_i) = f_\varphi(d_\varphi(i) * b_{(}\sigma_\varphi(i)))$. Then, by induction over the basis, it can be proven that for all $x \in V$, $\varphi(x) = f_\varphi \circ (d_\varphi \circ x \circ \sigma_\varphi^{-1})$.

Define g to be the mapping from Code Automorphisms on V to the group $((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$ given by $\varphi \mapsto (d_\varphi, (f_\varphi, \sigma_\varphi))$. Then it is the case that $\varphi(x) = f_\varphi \circ (d_\varphi \cdot x) \circ \sigma_\varphi^{-1} = g(\varphi) \cdot x$. Then, because the action of $((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$ on V is faithful and because of the extensionality of semilinear automorphisms, we conclude that g is injective. Similarly, we can conclude that g is a group-homomorphism because $g(\varphi * \varphi') \cdot x = (\varphi * \varphi')(x) = \varphi(\varphi'(x)) = g(\varphi) \cdot (g(\varphi') \cdot x) = (g(\varphi) * g(\varphi')) \cdot x$. In conclusion, there exists an injective group homomorphism g from the group of Semilinear code automorphisms to the group $((K^\times)^n \rtimes_\phi ((\mathrm{Aut}(K)) \times S_n)^{\mathrm{mop}})^{\mathrm{mop}}$, meaning the first is isomorphic to a subgroup of the second.

3.2 Automorphisms of the GolayCode

Firstly, due to the fact that the field over which the GolayCode is defined is $\mathbb{Z}_2$, follows from lemma 3.1.2 that M_{24} (which is the group of semilinear automorphisms of the GolayCode) is a subgroup of S_{24}, because both $(\mathbb{Z}_2)^\times$ and $Aut(\mathbb{Z}_2)$ are trivial.

lemma 3.2.1: The action of M_{24} on $6 \times \mathbb{F}_4$ is faithful.
Proof: This action is inherited from S_{24}, and its action is faithful, therefore so is the action of M_{24}.

As it turns out, Semilinear Code Automorphisms of the Hexacode also have an action on $6 \times \mathbb{F}_4$ given by $\varphi, (i, x) \mapsto (\sigma_{\varphi i}, f_\varphi(d_{\varphi i} * x))$, which then induces a map on $\mathbb{Z}_2^{6\times\mathbb{F}_4}$ given by $m_{(i,x)} = \varphi(m)_{(\sigma_{\varphi i}, f_\varphi(d_{\varphi i} * x))}$

Theorem 3.2.2 This action defines a semilinear automorphism of the GolayCode.
Proof: Let m be a codeword of the GolayCode, and φ be a semilinear code automorphism of the HexaCode. It is obvious that the image of a column $m_{(i,)}$ is another column $m_{(\sigma_{\varphi i},)}$. Because this mapping is bijective, it also preserves the parity of the columns. Furthermore, the mapping $x \mapsto f_\varphi(d_{\varphi i} * x)$ is an additive homomorphism. From this we can conclude two things; Firstly that the top row gets mapped to the top row (preserving its parity), and secondly that the per-column sum also gets mapped with this mapping. This means that the associated codeword in the hexacode gets mapped such that $(\varphi \cdot v)_{\sigma_{\varphi i}} = f_\varphi(d_{\varphi i} * v_i)$ holds, meaning that the new corresponding vector is precisely the old one mapped by φ, meaning it is once again a codeword. In conclusion, this action does map codewords. Since it is defined by a permutation on indices, the other properties readily follow, meaning that this indeed is a semilinear code automorphism for the GolayCode.

Corollary: From the construction used, it is obvious that the group of semilinear automorphisms of the hexacode is a subgroup of the group of semilinear automorphisms of the GolayCode.

Theorem 3.2.3: The codewords of the Hexacode also induce semilinear code automorphisms of the GolayCode, such that for a codeword v, the

mapping is such that $(v \cdot m)_{(i,x+v_i)} = m_{(i,x)}$.

Proof: Firstly, because this is again a permutation on the indices, it suffices to show that this mapping preserves codewords in the GolayCode. Secondly, note that it suffices to show that this is true for scalar multiples of basis vectors of the Hexacode, due to the fact that the composition of the maps given for some $u, v \in$ HexaCode corresponds to the map given by $u + v$. Furthermore, due to the fact that (ω, ω^{-1}) is a basis for $\mathbb{F}_4$, it again suffices to show that only the maps corresponding to those scalars multiplied with the basis words of the Hexacode map codewords of the Golaycode. First of all, it is relatively easy to verify that this is the case for the codewords of the hexacode which are ω times a basisvector or a basisvector itself, on the basis of the GolayCode. From there, since the mapping generated is a permutation of the indices, it is an additive homomorphism, meaning that ω times a basisvector as well as basisvectors themselves preserve all codewords of the GolayCode. Next, since $(\omega, 1)$ is a basis for $\mathbb{F}_4$, and mapping the sum of two vectors in the Hexacode corresponds to composing their maps, we know that all scalar multiples of basisvectors preserve all codewords of the HexaCode. By applying the same argument, all codewords of the Hexacode do this. In conclusion, the mapping given by $(v \cdot m)_{(i,x+v_i)} = m_{(i,x)}$ also gives a semilinear automorphism of the GolayCode. □

Lemma 3.2.4: The intersection of semilinear automorphisms of the GolayCode which are induced by automorphisms of the Hexacode and those which are induced by codewords of the HexaCode, is precisely the trivial subgroup. **Proof:** Any nontrivial semilinear automorphism induced by one of the Hexacode preserves the top row. If an automorphism induced by a codeword in the hexacode preserves the top row, it must preserve the entire set, and therefore be trivial. □

3.3 Sextets

In order to prove that M_{24} is simple, we will use Iwasawa's criterion.

Iwasawa's criterion: If G is a nontrivial perfect group, and has a faithful primitive action on Ω, and if G is generated by the conjugates of an abelian

normal subgroup of the stabilizer (of any element of Ω), then G is simple.

In order to make use of this criterion, we will first of all define a set Ω and an action of M_{24} on this set.

Consider the cosets of the GolayCode in $\mathbb{Z}_2^{6\times\mathbb{F}_4}$.

Theorem 3.3.1: If u, v are vectors in $\mathbb{Z}_2^{6\times\mathbb{F}_4}$, and $\|u\| \leq 3$ and $\|v\| \leq 4$ which are members of the same coset of the GolayCode, then $u = v$.

Proof: By the triangle inequality, we must then have that $\|u + v\| \leq \|u\| + \|v\| \leq 8$. Since addition and subtraction are inherited from $\mathbb{F}_4$, they are identical. This means that $u - v = u + v$, and therefore $\|u - v\| \leq 7$. However, because they are members of the same coset, we must have that $u - v$ is a codeword the GolayCode. Then, due to lemma 2.2.2, we must have that $u - v = 0$, meaning that $u = v$. □

Corollary: Any vector of weight 3 or less uniquely defines a coset, which additionally does not contain any weight 4 vectors. This means that we can account for $1 + \binom{1}{24} + \binom{2}{24} + \binom{3}{24} = 1 + 24 + 276 + 2024 = 2325$ cosets out of $2^{24}/2^{12} = 4096$, leaving 1771 unidentified cosets.

Lemma 3.3.2: If u, v are two different vectors in the same coset, and both have weight 4, then they cannot overlap.
Proof: This is easy to see: It is the case that $\|u + v\| = \|u\| + \|v\| - 2\|u * v\|$ for all u and v. however, since they are different, $u + v$ is not 0, but because they overlap, we also have that $0 < \|u * v\|$. This is in clear contradiction with the fact that $u + v$ is a codeword in the GolayCode, as those have weigth of at least 8 if they are non-zero. □

Corollary: Any coset can contain at most 6 vectors of weight 4, due to the fact that they don't overlap and the maximum total weight is 24.

Theorem 3.3.3: Every coset containing a vector of weight 4 contains 6 such vectors, and additionally, these account for the remaining 1771 cosets.
Proof: There are 10626 vectors of weight 4. Since there are at most 6 such vectors in a single coset, at least $10626/6 = 1771$ cosets contain such a vector. There are only 1771 cosets available, because the other 2325 were al-

ready taken by vectors with weight 3 or less. This means that it is a "tight fit", meaning each of the 1771 cosets must contain the maximum of 6 vectors of weight 4. □

Lemma: 3.3.4: Each combination of 5 indices points uniquely determine a codeword of weight 8 in the GolayCode
Proof: Consider a vector with 4 out of 5 indices non-zero. Then in the coset in which it is contained, there is also a vector of weight 4 which doesn't overlap with the original four, and has value 1 at the last index. then the sum of those two vectors is a codeword in the hexacode with weight 4. However, if there were multiple codewords of weight 8 which contain these 5 indices, then the weight of their intersection is at least 5, meaning that their sum (which also must be a codeword) has at most weight 6, which is a contradiction. □

Corollary: This means that the codewords of weight 8 form a so-called Steiner-system $(5, 8, 24)$, where you can pick any 5 out of 24 values, and those uniquely determine a set of 8 values containing those five. This system turns out to be unique up to permutation of the values, which allows you to give another definition of the GolayCode as those codewords which can be written as some symmetric difference of the sets of size 8 the Steinersystem (5,8,24).

Definition: the cosets of codewords with weight 4 define what are called Sextets, which are partitions of $6 * \mathbb{F}_4$ into 6 sets of size 4.

This is going to be the set Ω, such that the action of M_{24} is primitive on it. We will now analyse the orbits under the stabilizer of the Sextet given by the columns.

Lemma 3.3.5: The action of semilinear automorphisms of the Hexacode, preserve the sextet given by the columns.
Proof: This is easily seen, as the column i gets mapped to the column $\sigma_{\varphi i}$ for an automorphism φ. □

Lemma 3.3.6: The action of a codeword of the Hexacode preserves the sextet given by the columns

Proof: Again simple, as this action preserves the columns. □

Theorem 3.3.7: The action of the stabilizer of the column-sextet on the sextets preserves the per-column weight up to permutation.
Proof: Intersection is preserved by the semilinear automorphisms. Let us denote the vector with ones precisely in column i by col_i. Then the weight of some vector v in column i is equal to $\|v * \mathrm{col}_i\|$. Furthermore, then it must be the case that $\|v * \mathrm{col}_i\| = \|(\varphi \cdot (v * \mathrm{col}_i)\| = \|\varphi \cdot v * \varphi \cdot \mathrm{col}_i\| = \|\varphi \cdot v * \mathrm{col}_{\sigma_{\varphi i}}\|$, meaning that the weights get permuted precisely by σ_φ. □

Lemma 3.3.8: Every permutation of columns is achievable using only automorphisms generated by those of the HexaCode.
Proof: It suffices to show that we can make the permutations $(1,2,3,4,5)$ and $(5,6)$. The first is given by the automorphism $x \mapsto (\omega^{-1}x_5, \omega x_1, x_2, \omega^{-1}x_3, \omega x_4, x_6)$, while the second one is given by $x \mapsto (x_1^2, x_2^2, x_3^2, x_4^2, x_6^2, x_5^2)$. It is simple to check that these preserve the Hexacode, and it is obvious that these do indeed give the permutations required.

theorem 3.3.9: The orbit of the following Sextet under the stabilizer of the column-sextet is precisely the set of sextets with vectors with weight distribution 2^2, and has size 90

1	1	2	2	3	3
1	1	2	2	3	3
4	4	5	5	6	6
4	4	5	5	6	6

Proof: First, note that the orbit must be contained in this set, because of theorem 3.3.7. Now it suffices to show that the action of M_{24} is transitive on the vectors with such weight distribution, by showing that each such vectors can be normalized into the vector given by (1), using only members of the stabilizer. First of all, we can find a map which ensures that we can normalize vectors with this weight distribution to have the same per-column weight as (1), by using theorem 3.3.8. Then, we can apply an automorphism induced by choosing a member of the Hexacode such that the resulting vector has 1 as value at $(1,0)$ and $(2,0)$, because we can freely

choose the first three values of codewords of the hexacode. Then, we can apply an automorphism of the hexacode which multiplies by a scalar in order to ensure that the value at index $(1,1)$ is 1, while keeping the earlier set values. Finally, if we're not done yet, there are two cases: either $(2,\omega)$ or $2,\omega^{-1}$ has value 1. in the first case, apply the second permutation given in the proof of 3.3.8. in the second, don't. Then finally apply the hexacodeword $(0,1,0,1,\omega,\omega^{-1})$. This gives a procedure to generate a member of the stabilizer which maps any vector with weight distribution 2^2 to (1). Since there are 6 such vectors in the Sextet, each of the sextets in its orbits does. Since there are $\binom{2}{6} * \binom{2}{4}^2 = 540$ such vectors, there must be exactly $540/6 = 90$ cosets in this orbit. □

Through similar constructions it turns out that we can show that the remaining non-trivial orbits are those of the following Sextets:

1	2	3	3	3	3
2	1	4	4	4	4
2	1	5	5	5	5
2	1	6	6	6	6
1	1	1	2	5	6
1	2	2	2	4	3
3	5	6	4	3	6
4	6	5	3	5	4

These orbits can be characterised as sextets containing the weight patterns $3 \cdot 1$ and $2 \cdot 1 \cdot 1$ respectively. Their respective sizes can be calculated to be 240 and 1440, which in total accounts for all 1771 sextets, meaning that indeed these are all orbits.

theorem 3.3.10: The action of M_{24} on the Sextets is transitive.

Proof: For this it suffices to show that there is some element which fuses the

orbits; this is one such element:

$$m \mapsto \begin{bmatrix} m_{(1,0)} & m_{(2,0)} & m_{(3,0)} & m_{(4,0)} & m_{(6,1)} & m_{(5,1)} \\ m_{(1,1)} & m_{(2,1)} & m_{(3,1)} & m_{(4,1)} & m_{(6,0)} & m_{(5,0)} \\ m_{(2,\omega)} & m_{(1,\omega)} & m_{(4,\omega)} & m_{(3,\omega)} & m_{(5,\omega^{-1})} & m_{(6,\omega^{-1})} \\ m_{(2,\omega^{-1})} & m_{(1,\omega^{-1})} & m_{(4,\omega^{-1})} & m_{(3,\omega^{-1})} & m_{(5,\omega)} & m_{(6,\omega)} \end{bmatrix}$$

□

Theorem 3.3.11 The action of M_{24} is Primitive.

Proof: The size of any block must be the sum of orbits of the stabilizer of one of its members. Furthermore, it must also divide 1771 since its conjugates have the same size, and tile the entire set. The orbits have sizes 1,90,240,1440, and the only sums of those values that divide 1771 are either 1 or 1771, meaning that they are trivial. □

Now we have shown that the action is primitive. To be able to use Iwasawa's criterion, it now remains to be shown that there is a normal abelian subgroup whose conjugates generate the entire group.

Theorem 3.3.12: The subgroup of semilinear automorphisms generated by codewords of the Hexacode is normal in the stabilizer of the column-sextet.

Proof: It is obvious that the subgroup is abelian, because it is isomorphic to the Hexacode. Furthermore, because they do not permute the cosets of the partition at all, it can be easily seen that this subgroup is normal. □

Theorem 3.3.13: The semilinear automorphisms generated by codewords of the Hexacode are in the commutator subgroup.

Proof: It suffices to show that this is the case for scalar multiples of basis elements of the Hexacode. This leaves us with 6 cases, since $\mathbb{F}_4$ is two-dimensional over $\mathbb{Z}_2$ with basis $(1, \omega)$, and the Hexacode is 3-dimensional over $\mathbb{F}_4$ with basis (a_1, a_2, a_3). Let us define f to be the element of M_{24} that corresponds with the automorphism of the Hexacode which maps according to $x \mapsto (x_1^2, x_2^2, x_3^2, x_4^2, x_6^2, x_5^2)$. Then, we have the following:

- a_1 is the commutator of f and $\omega \cdot a_1$
- $\omega \cdot a_1$ is the commutator of $\omega^{-1} \cdot f$ and a_1
- a_2 is the commutator of f and $\omega \cdot a_2$
- $\omega \cdot a_2$ is the commutator of $\omega^{-1} \cdot f$ and a_2
- a_3 is the commutator of f and $\omega \cdot a_3$
- $\omega \cdot a_3$ is the commutator of $\omega^{-1} \cdot f$ and a_3

This proves the lemma. □

Theorem 3.3.14: The group M_{24} is Simple

Proof: For this, it suffices to show that conjugates of the stabilizer of a sextet generate the entire set. This is the case; conjugates of the second element given in theorem 3.3.8, together with the conjugates of the element given in theorem 3.3.10 generate the entire group. Both of these elements are in the normal abelian part of the stabilizer of the sextet given in 3.3.9. This means that conjugates of a stabilizer generate the entire group. However, as conjugating fixes the commutator subgroup, these conjugates are all in the derived subgroup, meaning the derived subgroup of M_{24} is also the full group. At this point we can apply Iwasawa's criterion and conclude that M_{24} is simple. □

Appendix

A. Appendix A

The code generated in the writing of this thesis can be found at the following link:
https://github.com/leanprover-community/mathlib4/tree/blizzard_inc/mathieu12
This is a branch of the community project Mathlib, meaning that most of the code written was contributed by someone else. Most of the relevant code written for this thesis can be found under *Mathlib.InformationTheory.Code*, where the definitions of the Hexacode and GolayCode can be found, as well as the definitions of their semilinear automorphisms. Additionally, some definitions for general metrics were formalized under *Mathlib.Topology.GMetric* and *Mathlib.Topology.GPseudoMetric*.

Bibliography

[1] R. A. Wilson, *The Finite Simple Groups*. 2009, ISBN: 978-1-84800-988-2. DOI: 10.1007/978-1-84800-988-2.

[2] R. J. Chapman, "Constructions of the golay codes: A survey," 2000. [Online]. Available: https://api.semanticscholar.org/CorpusID:16382704.