# Math 491 - The Mathieu Groups

Billie-Jo Powers
Supervisor: Dr Martin Cook

March 2024

**Abstract**

The Mathieu groups are the first five discovered sporadic simple groups. In this dissertation, we aim to construct the large Mathieu groups. We first introduce and establish the main properties of multiply transitive groups and projective spaces. This allows us to construct the large Mathieu groups using one-point extensions of multiply transitive groups. Finally, we view the large Mathieu groups as automorphism groups on Steiner systems.

## Contents

# 1 Introduction

## 1.1 Motivation

In 1861, Mathieu published a paper constructing the groups now known as $M_{11}$ and $M_{12}$, in 1873 he published another paper in which he constructed the large Mathieu groups $M_{22}, M_{23}$, and $M_{24}$, these were the first of the 26 groups known as the sporadic simple groups to be discovered. As the name suggests, the Mathieu groups are simple, and their label "sporadic" means that they do not belong to any of the 18 countably infinite classes of finite simple groups given by the Classification Theorem for Simple Groups [1, p.2].

In this dissertation, we will be taking the same approach to building the Mathieu groups as Biggs and White [2]. As a consequence of this, many of the results (and in some cases the proofs) are taken verbatim from the book, but as often as possible we aim to add clarification to proofs and reasoning for using each theorem (e.g. the full construction of the large Mathieu groups will be significantly longer in this dissertation).

## 1.2 Preliminary definitions

We first recap some basic definitions in group theory that will be used throughout this dissertation.

Recall that for a group $G$, the centre of $G$, denoted $Z(G)$, is the set of all elements of $G$ which commute with every other element. A group $G$ is called abelian if all of its elements commute with each other (or, equivalently, if $Z(G) = G$).

**Definition 1.1** ([3, p.69]). *Let $p$ be a prime and $G$ a group. A group whose order is a positive power of $p$ is called a $p$-group. A subgroup of $G$ whose order is a power of $p$ is called a $p$-subgroup.*

**Definition 1.2** (Sylow p-subgroup [4, p.78]). *If $p$ is a prime, then a Sylow $p$-subgroup $P$ of a group $G$ is a maximal $p$-subgroup.*

Another important concept to recall is that of a commutator.

**Definition 1.3** (Commutator [4, p.33]). *Let $a, b \in G$, the commutator of $a$ and $b$, $[a, b]$ is given by:*

$$[a, b] = aba^{-1}b^{-1}.$$

**Definition 1.4** (Commutator subgroup [4, p.33]). *The commutator subgroup (or derived subgroup) of $G$ is the subgroup of $G$ generated by all of the commutators. We denote this subgroup by $G'$.*

When $G = G'$ we call $G$ a perfect group. Further, recall that for elements $g, h \in G$, we say that $h$ is conjugate to $g$ if there is some $x \in G$ such that $h = xgx^{-1}$, and conjugacy is an equivalence relation on $G$ [3, p.5]. We can also conjugate subgroups as well as elements.

**Definition 1.5** ([4, p.44]). *If $H \leq G$ and $g \in G$, then the conjugate $gHg^{-1}$ is $\{ghg^{-1} : h \in H\}$.*

**Definition 1.6** (Normaliser [4, p.44]). *If $H \leq G$, then the normaliser of $H$ in $G$, denoted by $N_G(H)$, is*

$$N_G(H) = \{a \in G : aHa^{-1} = H\}.$$

The final part of this section will be looking at results related to group actions.

**Definition 1.7** (Group action [4, p.247]). *If $X$ is a set and $G$ is a group, then $X$ is a $G$-set if there is a function $\alpha : G \times X \to X$ (called an action) denoted by $\alpha : (g, x) \mapsto gx$, such that:*

*(i) $1x = x$ for all $x \in X$, and*

*(ii) $g(hx) = (gh)x$ for all $g, h \in G$ and $x \in X$.*

*We also say that $G$ acts on $X$. If $|X| = n$, then $n$ is called the degree of the $G$-set $X$.*

In fact, the following theorem allows us to see that $G$-sets are another way of viewing permutation representations.

**Theorem 1.8** ([4, Theorem 3.18]). *If $X$ is a $G$-set with action $\alpha$, then there is a homomorphism $\tilde{\alpha} : G \to S_X$ given by $\tilde{\alpha} : x \mapsto gx = \alpha(g, x)$. Conversely, every homomorphism $\phi : G \to S_X$ defines an action, namely, $gx = \phi(g)(x)$.*

*Proof.* Let $g \in G$ and $x \in X$, then,

$$\tilde{\alpha}(g^{-1})\tilde{\alpha}(g) : x \mapsto \tilde{\alpha}(g^{-1})(gx) = g^{-1}g(x) = x,$$

so we have that each $\tilde{\alpha}(g)$ is a permutation of $X$ with inverse $\tilde{\alpha}(g^{-1})$. The definition of an action ensures that $\tilde{\alpha}$ is a homomorphism. □

**Definition 1.9** (Faithful [4, p.248]). *A $G$-set $X$ with action $\alpha$ is faithful if $\tilde{\alpha} : G \to S_X$ is injective.*

We note that this definition means that, when $X$ is a faithful $G$-set, if $gx = x$ for all $x \in X$, then $g = e$.
For a $G$-set with action $\alpha$, Theorem 1.8 gives that the subgroup $\mathrm{Im}(\tilde{\alpha}) \leq S_x$ is a permutation group, so, if $X$ is a faithful $G$-set then $G$ can be identified with $\mathrm{Im}(\tilde{\alpha})$ and we recover the permutation group $(G, X)$, and say that the group has degree $|X|$. From this point, whenever we are working with a permutation group $(G, X)$, it is given that $X$ is a faithful $G$-set.
We now give a reminder of orbits and stabilisers for group actions which will be used throughout the following section.

**Definition 1.10** (Orbit [3, Definition 4.2] )**.** *If $G$ acts on $X$, the orbit of an element $x \in X$ is the set:*

$$\mathcal{O}_x = \{gx : g \in G\}$$

**Definition 1.11** (Stabiliser [3, p.55])**.** *If $G$ is a group acting on a set $X$, then the stabiliser of $x \in X$ in $G$ is the subset:*

$$G_x = \{g \in G : gx = x\}.$$

For a subset $Y$ of a set $X$, say $Y = \{x_1, \cdots, x_n\}$, we call the set $G_Y = \{g \in G : gx_i = x_i, \forall i = 1, \cdots, n\}$ the pointwise stabiliser of $Y$. Similarly, we call the set $G_{(Y)} = \{g \in G : gx_i \in Y, \forall x_i \in Y\}$ the set-wise stabiliser of $Y$. Finally, we recall the Orbit-Stabiliser Theorem.

**Theorem 1.12** (Orbit-Stabiliser Theorem [3, p.56])**.** *If $G$ acts on $X$ and $x \in X$, then $|\mathcal{O}_x| = |G : G_x|$.*

# 2 Transitivity

## 2.1 Multiply transitive groups

This section will establish multiply transitivity and - most importantly - extensions of multiply transitive groups. We begin with establishing the definition of transitive groups.

**Definition 2.1** (Transitive [2, Defintion 1.3.1])**.** *The permutation group $(G, X)$ is transitive if there is just one orbit in the action of $G$ on $X$.*

From this, we can see that when $(G, X)$ is transitive, the Orbit-Stabiliser Theorem (Theorem 1.12) becomes:

$$|X| = |G : G_x|.$$

**Definition 2.2** (k-transitive [2, Definition 1.3.5])**.** *Let $G$ be a group acting on a set $X$ of degree $n$. The permutation group $(G, X)$ is called $k$-transitive ($k \leq n$) if for any two ordered pairs of $k$-tuples, say $(x_1, \cdots, x_k)$ and $(y_1, \cdots, y_k)$ with distinct entries in $X$, we can find some $g \in G$ such that:*

$$g(x_i) = y_i$$

*for all $1 \leq i \leq k$.*

From this, we can see that any group which is $k$-transitive is also $l$ transitive for any $l = 1, \cdots, k$, so a group being called $k$-transitive is meant to imply that $k$ is the largest integer for which the condition is satisfied. A group is called multiply transitive if it is $k$-transitive for some $k \geq 2$.
We now establish some important properties of multiply transitive groups, which will allow us to consider extensions of multiply transitive groups.

**Lemma 2.3** ([2, Lemma 1.3.6])**.** *Let $G$ be a transitive group on $X$. Then $(G, X)$ is $k$-transitive if and only if, for any $x \in X$, $(G_x, X \setminus \{x\})$ is $(k-1)$-transitive.*

*Proof.* First, suppose $(G, X)$ is $k$-transitive, then by the above we have that it is $(k-1)$-transitive, and the set of all pairs of $(k-1)$-tuples in $X$ must contain all of the $(k-1)$-tuples in $X \setminus \{x\}$, so the result follows.
For the converse, suppose that $(G_x, X \setminus \{x\})$ is $(k-1)$-transitive. For any two ordered $k$-tuples $(x_1, \cdots, x_k)$ and $(y_1, \cdots, y_k)$, we can find $g_1, g_2 \in G$ and $h \in G_x$ which satisfy the following:

$$g_1(x_1) = x \ , \ g_2(y_1) = x$$
$$h(g_1(x_i)) = g_2(y_i) \ (2 \leq i \leq k)$$

So that $g_2^{-1}hg_1(x_1) = g_2^{-1}(x) = y_1$, and (for $i = 2, \cdots, k$), $g_2^{-1}hg_1(x_i) = g_2^{-1}g_2(y_i) = y_i$. Hence $g_2^{-1}hg_1$ is an element of $G$ which transforms the $k$-tuples as required. □

The above theorem tells us that we can determine whether a group is multiply transitive by looking at successive stabilisers $G_x$, $(G_x)_y$, and so on. Let $|X| = n$ and suppose $G$ is k-transitive on $X$, then the Orbit-Stabiliser Theorem gives the result:

$$|G| = n(n-1)(n-2)\cdots(n-k+1)|G_{x_1 x_2 \cdots x_k}|,$$

where $G_{x_1 x_2 \cdots x_n}$ is the pointwise stabiliser of $x_1, x_2, \cdots, x_k$.
If $G$ is $k$-transitive with the identity being the only permutation which fixes $k$ points, then we call $G$ sharply $k$-transitive, and the order of $G$ is exactly $n(n-1)\cdots(n-k+1)$. We are particularly interested in sharp transitivity when $k = 1$ [2, p.8].

**Definition 2.4** (Regular)**.** *A group $G$ is called regular on $X$ if $(G, X)$ is sharply 1-transitive.*

When $G$ is regular on $X$, we have $|G| = |X|$. Equivalently, we have that $G$ is regular on $X$ if for any pair $x, y \in X$, there is a unique element $g \in G$ such that $gx = y$ [5, p.78].

**Theorem 2.5** ([4, p.81])**.** *Let $(G, X)$ be a permutation group, if $H \leq G$ acts transitively on $X$, then $G = HG_x$ for each $x \in X$.*

*Proof.* Fix $x \in X$, it is clear that $G_x \leq HG_x$. Hence it remains to show that for any $g \in G$ such that $g \notin G_x$, we have $g \in HG_x$. For any $g \in G$, we have that there is some $h \in H$ with $h \cdot (g \cdot x) = x$, so $hg = g' \in G_x$ and $g = h^{-1}g' \in HG_x$. Hence the result follows. □

We are interested in constructing multiply transitive groups (the Mathieu groups), the following lemma is useful for this purpose.

**Lemma 2.6** ([2, Lemma 1.3.9])**.** *Let $(G, X)$ be a $k$-transitive group, where $k \geq 2$. Then for $g \notin G_x$, we have $G = G_x \cup G_x g G_x$.*

*Proof.* Let $h \in G \setminus G_x$, and fix $g$ as above. Since $G$ is transitive, we can find some $g_1 \in G$ which sends $g^{-1}(x)$ to $h^{-1}(x)$ and fixes $x$, so that $g_1 \in G_x$. This then gives $hg_1g^{-1}(x) = h(h^{-1}(x)) = x$, so we have:

$$hg_1g^{-1} \in G_x \implies h \in G_xgG_x.$$

Since this holds for all $h \in G \setminus G_x$, the result follows. □

**Definition 2.7** (One-point extension [2, Definition 1.5.1])**.** *Let $(G, X)$ be a transitive group. We call $(G^+, X^+)$ a one-point extension when $X^+ = X \cup \{*\}$, with $* \notin X$ and $G^+$ is transitive on $X^+$ with stabilizer $(G^+)_* = G$.*

We then have from Lemma 2.3 that if $G$ is $k$-transitive, then $G^+$ is $(k+1)$-transitive. Further, if $G$ is sharply $k$-transitive, then the one-point extension $G^+$ is sharply $(k+1)$-transitive, as any permutation fixing $k+1$ points in $G^+$ must fix $k$ points in $G$, hence the only such permutation is the identity.
Now that we have defined a one-point extension for a multiply transitive group, we want to see when it is possible to find such an extension. To this aim, we establish the following theorem.

**Theorem 2.8** (Conditions for the existence of a one-point extension [2, Theorem 1.5.2] )**.** *Let $(G, X)$ be a $k$-transitive group with $k \geq 2$, and let $X^+ = X \cup \{*\}$, where $* \notin X$. Suppose that there is some permutation $h$ of $X^+$ and an element $g \in G$ such that:*

*(i) $h$ switches $*$ and some $x \in X$, and $h$ fixes some point $y \in X$;*

*(ii) $g$ switches $x$ and $y$;*

*(iii) $(gh)^3$ and $h^2$ are elements of $G$;*

*(iv) $hG_xh = G_x$.*

*Then the group $G^+ = \langle G, h\rangle$ acts on $X^+$ as a one-point extension of $(G, X)$.*

*Proof.* From condition $(ii)$, we have that $g \notin G_x$, and so Lemma 2.6 gives us that $G = G_x \cup G_xgG_x$. The result will follow if we can show that $G^+ = \langle G, h\rangle = G \cup GhG$, as $h$ does not fix $*$, so no element in $GhG$ can fix $*$, hence $(G^+)_* = G$.
Clearly, $G \cup GhG \subseteq \langle G, h\rangle$, and if $G \cup GhG$ is a group, $\langle G, h\rangle \subseteq G \cup GhG$. Therefore, it is enough to show that $G \cup GhG$ is a group. We first show that $G \cup GhG$ is closed under composition. For this, we only need to show that $hGh \subset G \cup GhG$, since we have that, if $x \in hGh$, either $x \in G$ or $x \in GhG$, so it follows that for all $g_1, g_2 \in G$, either $g_1xg_2 \in G$ or $g_1xg_2 \in GhG$, which gives the following:

$$GhG \cdot GhG = G \cdot hGh \cdot G \subseteq G \cup GhG.$$

We know that $h^2 \in G$ fixes $x$, so that $h^2 \in G_x$, combining this with condition (iv), we have that $hG_x = G_xh$. Also, $(gh)^3 \in G$, so:

$$(ghg)hgh \in G \implies hgh \in (ghg)^{-1}G$$

and $(ghg)^{-1}G = g^{-1}h^{-1}G = g^{-1}hG$. So that $hgh \in g^{-1}hG$. These three remarks are used in the following calculation:

$$\begin{aligned} hGh &= h(G_x \cup G_x g G_x)h \\ &= hG_x h \cup hG_x g G_x h \\ &= G_x \cup G_x \cdot hgh \cdot G_x \\ &\subseteq G \cup G_x \cdot g^{-1}hG \cdot G_x \\ &\subseteq G \cup GhG, \end{aligned}$$

as required. Thus $G \cup GhG = \langle G, h \rangle$. □

In most cases, $g$ and $h$ are chosen so that both $(gh)^3$ and $h^2$ are the identity. Using the assumptions in Theorem 2.8 we can make the following useful observation. Since $h(*) = x, h(x) = *, h(y) = y, g(x) = y, g(*) = *, g(y) = x$, we have that:

$$\begin{aligned} (gh)^3(x) &= ghghgh(x) \\ &= ghgh(*) = gh(y) \\ &= x. \\ (gh)^3(y) &= ghghgh(y) \\ &= ghgh(x) = gh(*) \\ &= y. \\ (gh)^3(*) &= ghghgh(*) \\ &= ghgh(y) = gh(x) \\ &= *. \end{aligned}$$

Therefore, when we check condition (iii) in future proofs, we are only concerned with showing the result for the other elements of $G$.

## 2.2 Primitivity

Now that we have looked at multiply transitive groups, we want to discuss another property of groups that is useful when establishing properties of the Mathieu groups.

**Definition 2.9** ([4, p.256]). *Let $(G, X)$ be a permutation group, then a block is subset $B$ of $X$ such that, for each $g \in G$, either $gB = B$ or $gB \cap B = \emptyset$. (Here, $gB = \{gx : x \in B\}$).*

Clearly, $\emptyset$, singleton sets and $X$ are blocks, we call these trivial blocks. Any other block is called non-trivial.

**Definition 2.10** (Primitive [4, p.256]). *The transitive group $(G, X)$ is primitive if it has no non-trivial blocks.*

If we can find non-trivial blocks of $X$, then we say it is imprimitive.

**Theorem 2.11** ([4, Theorem 9.12]). *Any 2-transitive group is primitive.*

*Proof.* Suppose that there is a non-trivial block $B$, then, since $|B| \geq 2$ and $B \neq X$, we can find elements $x, y, z \in X$, with $x, y \in B$ and $z \notin B$. Since $(G, X)$ is 2-transitive, we can find some $g \in G$ such that $gx = x$ and $gy = z$. Hence we have that $x \in B \cap gB$, so that $B \cap gB$ is non-empty, and $B \neq gB$, contradicting the assumption that $B$ is a block. □

Combining the above theorem with the fact that any $k$-transitive group is $i$-transitive, for $1 \leq i \leq k$, it follows that any $k$-transitive group with $k \geq 2$ is primitive.

**Theorem 2.12** ([4, Theorem 9.15]). *Let $(G, X)$ be a transitive permutation group. $(G, X)$ is primitive if and only if, for any $x \in X$, the stabiliser $G_x$ is a maximal subgroup of $G$.*

*Proof.* Assume $G_x$ is not maximal, so we can find a subgroup $H$ satisfying $G_x < H < G$, we want to show that $Hx = \{hx : h \in H\}$ is a non-trivial block (meaning that $(G, X)$ is imprimitive). If $g \in G$ and $Hx \cap gHx \neq \emptyset$, then $hx = gh'x$ for some $h, h' \in H$, and $h^{-1}gh' \in G_x \subset H$. So we have $g \in H$. Hence we have that $gHx = Hx$ and $Hx$ is a block. It remains to show that $Hx$ is non-trivial. We already have that $Hx$ is non-empty. To show that $Hx \neq X$, let $g \in G$ with $g \notin H$. If $Hx = X$, then for every $y \in X$, there is a $h \in H$ with $y = hx$; in particular $gx = hx$ for some $h \in H$. Therefore $g^{-1}h \in G_x < H$, and $g \in H$, contradicting $g \notin H$. Finally, if $Hx$ is a singleton, then $Hx = \{x\}$, so $H \leq G_x$, contradicting $G_x < H$. Therefore, $(G, X)$ is imprimitive.
For the converse, assume that every $G_x$ is a maximal subgroup and that there is a non-trivial block $B$ in $X$. Define a subgroup $H$ of $G$ by

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

Choose $x \in B$, if $gx = x$, then $x \in B \cap gB$, and so $gB = B$ (by the assumption that $B$ is a block), therefore $G_x \leq H$. Since $B$ is non-trivial, there is $y \in B$ with $y \neq x$. Since $(G, X)$ is transitive, there is a $g \in G$ with $gx = y$; hence $y \in B \cap gB$ and so $gB = B$. Thus $g \in H$ while $g \notin G_x$, i.e $G_x < H$. If $H = G$, then $gB = B$ for all $g \in G$. It then follows that for $b \in B$ and $c \in X \setminus B$, there is no $g \in G$ satisfying $g \cdot b = c$, which contradicts $X \neq B$ being a transitive $G$-set. Therefore $G_x < H < G$, contradicting the assumption that $G_x$ is maximal. □

**Lemma 2.13** ([4, Lemma 9.16]). *Let $(G, X)$ be a permutation group and $x, y \in X$. Then,*

*(i) If $H \leq G$ then $Hx \cap Hy \neq \emptyset \Rightarrow Hx = Hy$.*

*(ii) If $H \lhd G$, then the subsets $Hx$ are blocks of $X$.*

*Proof.* (i) We have that $Hy = Hx$ if and only if $y \in Hx$. If $Hx \cap Hy \neq \emptyset$, then there must be some $h, h' \in H$ with $hx = h'y$. Then we have $y = (h')^{-1}hx \in Hx$, so that $Hy = Hx$.

(ii) Let $g \in G$ and assume $gHx \cap Hx \neq \emptyset$. Since $H \lhd G$, $gHx \cap Hx = Hgx \cap Hx$. Then there are $h, h' \in H$ with $hgx = h'x$ so that $gx = h^{-1}hx \in Hx$, hence $gHx = Hx$ and $Hx$ is a block.

□

**Theorem 2.14** ([4, Theorem 9.17]). *If $(G, X)$ is primitive and $H \trianglelefteq G$, $H \neq 1$, then $H$ is transitive on $X$.*

*Proof.* Lemma 2.13 gives that $Hx$ is a block for every $x \in X$. Since $X$ is primitive, either $Hx = \emptyset$, $Hx = \{x\}$, or $Hx = X$. Clearly, $Hx \neq \emptyset$. If $Hx = \{x\}$, then $H \leq G_x$, but if $g \in G$, then the normality of $H$ gives that $H = gHg^{-1} \leq gG_xg^{-1} = G_{gx}$. By the transitivity of $(G, X)$, $H \leq \cap_{y \in X} G_y = 1$ (since $X$ is faithful), which gives a contradiction. Hence we must have $Hx = X$, so that $H$ is transitive on $X$. □

This result allows us to establish the following theorem, which in turn will be instrumental in proving the simplicity of the Mathieu groups.

**Theorem 2.15** ([2, Theorem 1.6.7]). *Let $(G, X)$ be a primitive group with $G_x$ simple, then either:*

*(i) $G$ is simple, or*

*(ii) $G$ has a normal subgroup $N$ which acts regularly on $X$.*

*Proof.* For this proof, we first note that, since $(G, X)$ is primitive, it is transitive and so, since we have assumed that $G_x$ is simple for some $x \in X$, we have $G_x \cong gG_xg^{-1} = G_{gx}$, hence $G_x$ is simple for all $x \in X$. Suppose that $G$ is not simple, so that there exists some proper non-trivial normal subgroup $N$. Given $x \in X$, consider $N \cap G_x$, this group is normal in $G_x$, and since $G_x$ is simple, it must be either 1 or the whole of $G_x$. Suppose $G_x \cap N = G_x$, then we have $G_x \leq N$. Since $N \lhd G$, we have that $N$ is transitive, and clearly $G_x$ is not transitive, so we must have $G_x < N$. Thus by Theorem 2.12, $N = G$, which contradicts the assumption that $N$ is proper, and so $N \cap G_x = 1$, hence $N$ acts regularly. □

We now look at some results which will be used when proving the simplicity of the large Mathieu groups. It is first useful to define equivalence of permutation groups.

**Definition 2.16** (Equivalent permutation groups, [2, Definition 1.7.1]). *The permutation groups $(G, X)$ and $(G', X')$ are equivalent if there is an isomorphism $g \mapsto g'$ of $G$ and $G'$, and a one-to-one correspondence $\beta : X \mapsto X'$ such that:*

$$g'(\beta x) = \beta(gx)$$

*for all $g \in G$ and $x \in X$.*

**Lemma 2.17** ([2, Lemma 1.7.2]). *Let $(G, X)$ be a transitive permutation group, and suppose there is a normal subgroup $N$ of $G$ acting regularly on $X$. Given $x \in X$, there is an action of $G_x$ on $N^* = N \setminus \{1\}$, equivalent to its action on $X \setminus \{x\}$. Furthermore, the elements of $G_x$ act as automorphisms of $N$.*

*Proof.* Let $f : X \setminus \{x\} \to N^*$ be the map defined by $f(y) = n_y$, with $n_y \in N$ being the (unique) element taking $x$ to $y$. Given $g \in G_x$, we define the corresponding permutation $g'$ of $N^*$ by:

$$g'(n_y) = n_{gy}.$$

By definition, this action is equivalent to the action of $G_x$ on $X \setminus \{x\}$. Now, $gng^{-1} \in N$ sends $x$ to $gy$, and since $N$ is regular, $n_{gy} = gn_yg^{-1}$. Thus, each $g \in G_x$ acts by conjugation, $g'(n) = gng^{-1}$, on $N^*$. If we extend the action to $N$, by letting $g'(1) = 1$, then we have an automorphism of $N$. □

From the above, we have that for any transitive permutation group $(G, X)$ to have a regular normal subgroup $N$, $N$ must admit a group of automorphisms equivalent to the action of the stabiliser $G_x$ on $X \setminus \{x\}$. The following theorem aims to show that there are only a few circumstances in which this happens. For this theorem, recall that an $n$-dimensional vector space $V$ over a field $K$ satisfies $V \cong K^n$, and let $\mathbb{Z}_p$ be the cyclic group of order $p$.

**Theorem 2.18** ([4, Lemma 9.24]). *Let $(G, X)$ be $k$-transitive for some $k \geq 2$, with degree $n$. If $G$ has a regular normal subgroup $H$, then $k \leq 4$ and,*

*(i) If $k \geq 2$, then $H \cong (\mathbb{Z}_p)^n$, for some prime $p$.*

*(ii) If $k \geq 3$, then $H \cong (\mathbb{Z}_2)^n$ or $H \cong \mathbb{Z}_3$.*

*(iii) If $k = 4$, then $H \cong (\mathbb{Z}_2)^2$.*

*Proof.* We first note that from Lemma 2.3, $(G_x, X \setminus \{x\})$ is $(k-1)$-transitive for each fixed $x \in X$. Also, as in the proof of Lemma 2.17 we have that each $g \in G_x$ acts by conjugation on $H^* = H \setminus \{1\}$ and so $(G_x, H^*)$ is $(k-1)$-transitive.

(i) Let $k \geq 2$, since $G_x$ acts by conjugation on $H^*$, we have that any $g \in G_x$ preserves the order of an element of $H^*$, and so every element must have the same order. Further, if an $g$ element has order $nm$, then $g^n$ has order $m$, hence every element is of order some prime $p$. Thus, $H$ is a $p$-group and its centre $Z(H)$ is non-trivial, again, each $g \in G_x$ must preserve the elements of $Z(H)$, and so we have $H = Z(H)$, and $H$ is abelian. Hence $H \cong (\mathbb{Z}_p)^n$ [2, p.18].

(ii) Suppose $k \geq 3$ and let $h \in H^*$, then we want to show that $\{h, h^{-1}\}$ is a block. If $h' = ghg^{-1}$, for some $g \in G$ then $(h')^{-1} = gh^{-1}g^{-1}$, so it follows that $\{h, h^{-1}\}$ and $\{h', h'^{-1}\} = g\{h, h^{-1}\}g^{-1}$ are either equal or disjoint, so $\{h, h^{-1}\}$ is indeed a block. Then, $(G_x, H^*)$ is 2-transitive, and hence (by Theorem 2.11) is primitive, so that either $\{h, h^{-1}\} = H^*$, or $\{h, h^{-1}\} = \{h\}$. In the first case, $|H| = 3$, so that $n = 3$ and $H \cong \mathbb{Z}_3$. In the second case, $h$ has order 2, and so applying part (i), we get $p = 2, H \cong (\mathbb{Z}_2)^n$.

(iii) Let $k \geq 4$, by part (ii), we can assume that $H \cong (\mathbb{Z}_2)^n$, with $n \geq 2$ (we exclude the cases $\mathbb{Z}_2$ and $\mathbb{Z}_3$ as they are too small). We have that $H$ contains a 4-group $\{1, h, k, hk\}$. Then, the stabiliser $(G_x)_h$ acts 2-transitively, and hence primitively, on $H^* \setminus \{h\}$. In this case, we have that, for $g \in (G_x)_h$, we have $g\{k, hk\}g^{-1} = \{gkg^{-1}, hgkg^{-1}\}$, so letting $k' = gkg^{-1}$, we have if $\{k, hk\} \cap \{k', hk'\} \neq \emptyset$, then either $k = k'$, in which case $hk = hk'$, or $k = hk'$, which gives $hk = h^2k' = k'$ (since all elements of $H$ have order 2 by assumption). In both cases $\{k, hk\} = \{k', hk'\}$. Hence, $\{k, hk\}$ is a block, and so $H^* \setminus \{h\} = \{k, hk\}$. Hence $H = \{1, h, k, hk\} \cong (\mathbb{Z}_2)^2$ as required. Further, this shows that $k \neq 5$, as $5 \geq 4$ would give $H \cong \mathbb{Z}_2^2$, and so it is not possible to be 5-transitive on $H^*$, which has only 3 elements.

□

**Corollary 2.19** ([2, Theorem 1.7.6] ). *Let $(G, X)$ be $k$-transitive, and suppose it has a regular normal subgroup $N$, then:*

*(i)* $k \geq 2 \Rightarrow |X| = p^n$, *$p$ prime*

*(ii)* $k \geq 3 \Rightarrow |X| = 2^n$, *or 3,*

*(iii)* $k \geq 4 \Rightarrow |X| = 4$,

*Proof.* If $(G, X)$ is $k$-transitive, $(G_x, X \setminus \{x\})$ is $(k-1)$-transitive, and so, by Lemma 2.17, $N$ admits a $(k-1)$- transitive group of automorphisms. Then, we can apply Theorem 2.18, and the result follows. □

# 3 Projective spaces

## 3.1 Finite fields and finite vector spaces

Throughout the remainder of this dissertation, we will be working over finite fields and finite vector spaces, this section aims to give a brief overview (without proof) of the relevant properties of these, beginning with finite fields.

Let $F$ be a finite field, then we must have $|F| = p^n$, where $p$ is a prime number and $n$ is a positive integer. Furthermore, given any prime power $p^n$, there is a unique field (up to isomorphism) of size $p^n$ [6, p.60]. We call this the Galois field, denoted $GF(p^n)$.

The Galois field $GF(p^n)$ can be constructed in the following manner [2, p.26]:

1. Identify the sequence $(a_0, a_1, \cdots, a_{r-1})$, with the polynomial $a_0 + a_1t + \cdots + a_{r-1}t^{r-1}$ in the ring of polynomials $\mathbb{Z}_p[t]$;

2. Choose a polynomial $f(t)$ of degree $r$ which is irreducible in $\mathbb{Z}_p[t]$;

3. Define multiplication of two sequences by multiplying the corresponding polynomials in $\mathbb{Z}$ and then reducing modulo $f(t)$.

The following example gives a finite field that we will be working over in later sections.

**Example 3.1.** *To construct $GF(2^2)$, consider the polynomial $t^2+t+1$, which is irreducible over $\mathbb{Z}_2 = \{0,1\}$. The elements of $GF(4)$ may be given as follows:*

| *element* | *0* | $t$ | $t^2$ | $t^3$ |
|---|---|---|---|---|
| *reduction* | *0* | $t$ | $t+1$ | $1$ |

*In particular, we have that for any element $x \in GF(4), x^3 = 1$, and $x^2+x+1 = 0$.*

We now discuss some properties of finite vector spaces, letting $V = V(n,q)$ denote a vector space of dimension $n$ over the field $F = GF(q)$. The vector space $V$ will have $q^n$ elements. Recall that the general linear group $GL(V)$ is the group of all linear automorphisms of $V$, i.e permutations $l$ of $V$ which satisfy:

$$l(x+y) = l(x) + l(y) \quad \text{for all } x, y \in V$$
$$l(\lambda x) = \lambda l(x) \quad \text{for all } x \in V \lambda \in GF(q).$$

Similarly, the special linear group $SL(V)$ is the group of all linear automorphisms of $V$ with determinant 1.
It will first be useful to establish some properties of $GL(V)$ and $SL(V)$.

**Lemma 3.2** ([2, Lemma 2.3.3])**.** *$SL(V)$ is a normal subgroup of $GL(V)$, and if $V = V(n,q)$ then the index $|GL(V) : SL(V)| = q-1$.*

*Proof.* $GL(V)$ is the group of all linear automorphisms of $V$, and so is the group of automorphisms with non-zero determinant. Hence, we have that the determinant function maps $GL(V)$ into $F^* = F \setminus \{0\}$. We then have the following property of the determinant:

$$\det(AB) = \det(A)\det(B).$$

We also have that det is surjective, as for any $x \in F$, we can find a matrix $X \in GL(V)$ with determinant $x$, by letting $X$ have first entry $x$, all other diagonal entries being 1, and off diagonal entries being 0. These properties show us that the determinant is a homomorphism onto the multiplicative group $F^*$. Furthermore, the kernel of this homomorphism (the elements which are mapped to 1) is $SL(V)$. Hence, $SL(V)$ is a normal subgroup of $GL(V)$, and its index is $|\mathbb{F}^*| = q-1$. □

**Theorem 3.3** ([2, Theorem 2.3.4])**.** *The orders of $GL(V)$ and $SL(V)$ are given by:*

$$|GL(n,q)| = q^{n(n-1)/2}\prod_{i=1}^{n}(q^i - 1);$$
$$|SL(n,q)| = q^{n(n-1)/2}\prod_{i=2}^{n}(q^i - 1).$$

*Proof.* To show the result for $|GL(n,q)|$, we see that for each pair of ordered bases $\{e_1,\cdots,e_n\},\{f_1,\cdots,f_n\}$ of $V(n,q)$ there is a unique linear automorphism $l$ taking $e_i$ to $f_i$ (for $i=1,\cdots,n$). From this, we have that $|GL(n,q)|$ is equal to the number of ordered bases of $V(n,q)$. The first member of an ordered basis can be any element of $V$ except 0, and so there are $q^n-1$ choices for the first element. The second element cannot be linearly dependant on the first, so there are $q^n-q$ choices for the second element. Continuing in this fashion we have:

$$\begin{aligned}|GL(n,q)| &= (q^n-1)(q^n-q)\cdots(q^n-q^{n-2})(q^n-q^{n-1})\\ &= q^{1+2+\cdots+(n-1)}(q^n-1)(q^{n-1}-1)\cdots(q^2-1)(q-1)\\ &= q^{n(n-1)/2}\prod_{i=1}^{n}(q^i-1),\end{aligned}$$

which gives the desired result.
For $SL(n,q)$, we apply Lemma 3.2 and the result follows directly. □

## 3.2 Projective planes

We begin by introducing the idea of a projective plane:

**Definition 3.4** ([2, p.24])**.** *A finite projective plane is a pair $(P,\mathcal{L})$ of finite sets and an incidence relation between them which satisfy the following three axioms:*

- *Any pair of distinct elements of $P$ is incident to exactly one element $l\in\mathcal{L}$.*
- *Each pair of distinct elements of $\mathcal{L}$ is incident to a single element of $P$.*
- *There are at least four members of $P$ having the property that no three of them are incident to a single member of $\mathcal{L}$.*

We refer to the elements of $P$ as points and the elements of $\mathcal{L}$ as lines.

**Definition 3.5** (Projective geometry [2, Definition 2.5.1])**.** *Let $V$ be a vector space of dimension $n$ over some field $\mathbb{K}$. The statement "$x=\lambda y$ for some $\lambda\in\mathbb{K}^*$" defines an equivalence relation on $V^*=V\setminus\{0\}$, and the equivalence classes are the points of the projective geometry $PG(V)$.*

We will denote the equivalence class of $x\in V^*$ by $[x]$. A subspace $[U]$ of $PG(V)$ is the image of $U\leq V$ under the map $x\mapsto[x]$. The convention is to say that if $U$ has dimension $k$, then the equivalence class $[U]$ has projective dimension $k-1$, and so for $V=V(n,q)$, we write $PG(V)=PG(n-1,q)$. In cases where we can view $PG(V)$ as a projective plane, the points are given by $[x]$, for $x\in V^*$ and the lines are given by:

$$\mathcal{L}=\{[U]:U\text{ is a 2-dimensional subspace of }V\}.$$

We have that a point $[x]\in P$ is incident with $[U]\in\mathcal{L}$ if $x\in U$. In general, $PG(2,q)$ is a projective plane. To see this, we note that if $[x]$ and $[y]$ are distinct points of

$PG(2,q)$, then $x, y \in V(3,q)$ and there is no $\lambda \in GF(q)^*$ which satisfies $x = \lambda y$. Here $U = \langle x, y \rangle$ is the unique 2-dimensional subspace of $V(3,q)$ containing $x$ and $y$, so $[U]$ is the unique element of $\mathcal{L}$ incident to $[x]$ and $[y]$, so the first axiom is satisfied. For the second axiom, let $[U], [W] \in \mathcal{L}$ with $[U] \neq [W]$. We have $U \cap W \leq U$, so that $\dim(U \cap W) \leq \dim U = 2$. If $U \cap W = U$, then $U \leq W$ and $\dim W = \dim U = 2$, so $W = U$. Hence we must have $\dim(U \cap W) \leq 1$. If $U \cap W = \{0\}$ then $U + W$ would have dimension 4, which is impossible in a 3-dimensional vector space. Hence $\dim(U \cap W) = 1$, so $U \cap W = \langle v \rangle$, and $[v] \in P$ is the unique element of $P$ incident to $[U]$ and $[W]$ in $\mathcal{L}$. For the final axiom, let $\{v_1, v_2, v_3\}$ be a basis of $V(3,q)$, then $[v_1], [v_2], [v_3]$ and $[v_1 + v_2 + v_3]$ have the property that no three of them are incident to a single member of $\mathcal{L}$.

**Example 3.6** ([2, p.37]). *Let $V = V(3,2)$, so $V$ has 8 points, and $V^*$ has 7 points. Since $V$ is a vector space over the field $\mathbb{K} = GF(2)$, we have $\mathbb{K}^*$ has only one element, and so every element of $V^*$ is in its own equivalence class. Hence $PG(2,2)$ has 7 elements. To give an explicit representation of $PG(2,2)$ we can choose coordinates $(x_0, x_1, x_2)$ for points $x \in V^*$, and denote $[x]$ in $PG(2,2)$ by $[x_0, x_1, x_2]$. If we let $U$ be the subspace whose equation is $x_0 + x_1 + x_2 = 0$, then this gives rise to a line $[U]$ containing the points $[1,1,0], [1,0,1]$ and $[0,1,1]$, there are 7 such lines,written in the same format as above, these lines are:*

$$\begin{aligned}\mathcal{L} = \{&[\langle(1,0,0),(0,0,1)\rangle], [\langle(1,0,0),(0,1,0)\rangle], [\langle(0,1,0),(0,0,1)\rangle],\\ &[\langle(1,1,0),(0,1,1)\rangle], [\langle(1,0,0),(0,1,1)\rangle], [\langle(0,0,1),(1,1,0)\rangle],\\ &[\langle(0,1,0),(1,0,1)\rangle]\},\end{aligned}$$

*and so we can construct the below figure 3.1. In fact, we can see from figure 3.1 that this is a projective plane, as all of the conditions of 3.4 are satisfied. The projective plane given by $PG(2,2)$ is called the Fano plane.*

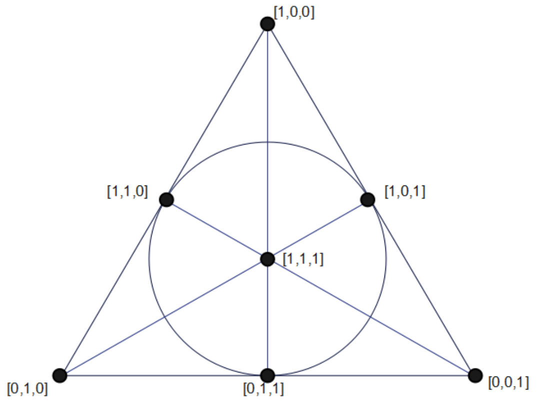


Figure 3.1: The Fano Plane; PG(2,2) [2, p.38]

In this dissertation, we are interested in permutations of projective planes. In particular, we will want to look at permutations which preserve the structure of the projective plane. By "preserve the structure", we mean that for a line $l \in \mathcal{L}$, the permutation $\pi$ satisfies $\pi(l) \in \mathcal{L}$.

## 3.3 Transvections

In this section, we continue to work with $V$ as a vector space over the field $F = GF(q)$, with the assumption that $n \geq 2$ (where $n = \dim V$). We also note that a function $\phi : V \to F$ is called a linear functional if it is a $F$-linear transformation (i.e $\phi(x + \lambda y) = \phi(x) + \lambda\phi(y)$ for $x, y \in V, \lambda \in F$). The main aim of this section is to discuss transvections, which are useful for studying certain matrix groups. We first define a hyperplane.

**Definition 3.7** (Hyperplane [4, p.228]). *If $V$ is an $m$-dimensional vector space over a field $F$, then a hyperplane $H$ in $V$ is a subspace of dimension $m-1$.*

In particular, if $u : V \to F$ is a non-trivial linear functional, the set

$$U = \{x \in V : u(x) = 0\}$$

is a hyperplane, and $u(x) = 0$ is an equation for it. We also have that for any $g \in GL(V)$, $g(U)$ is a hyperplane with equation $u'(x) = 0$ where $u' = u \circ g^{-1}$. If $u(x) = 0$ and $t(x) = 0$ are two equations for the same hyperplane, then there is some $\lambda \in F^* = F \setminus \{0\}$ satisfying $u(x) = \lambda t(x)$ for all $x \in V$ [2, p.30]. To see this, let $\{v_1, v_2, \cdots, v_m\}$ be a basis for $V$, with $\{v_2, \cdots, v_m\}$ a basis for $U$. Then we can write each $x \in V$ as $x = \lambda_1 v_1 + \lambda_2 v_2 + \cdots \lambda_m v_m$ for $\lambda_1, \cdots, \lambda_m \in F$, and we have $u(x) = \lambda_1 u(v_1)$, $t(x) = \lambda_1 t(v_1)$. So then $v_1 \notin U$ implies that $u(v_1) \neq 0$ and $t(v_1) \neq 0$, which means $\lambda = u(v_1)t(v_1)^{-1} \in F^*$ satisfies $u(x) = \lambda t(x)$.

**Definition 3.8** (Transvection [2, Definition 2.4.1]). *A linear automorphism $\tau$ in $GL(V)$ is a transvection, with direction $d \in V^*$, if $\tau$ fixes $d$ and $\tau(x) - x$ is a scalar multiple of $d$ (dependent on $x$) for all $x \in V$.*

Transvections can be viewed as maps moving every point in a direction parallel to $d$. An easy example of a transvection is the identity (which is a transvection for any direction $d$). We also note that a transvection with direction $d$ is a transvection with direction $\lambda d$, for all $\lambda \in F^*$.

**Theorem 3.9** ([2, Theorem 2.4.2]). *If $\tau$ is a transvection with direction $d$, then there is a hyperplane containing $d$ each of whose points is fixed by $\tau$.*

*Proof.* When $\tau$ is the identity, the result is clear, so let $\tau$ be a non-identity transvection.
Let $\tau(x) - x = u(x)d$ for each $x \in V$. Since $\tau$ is linear, $u$ is a non-trivial linear functional on $V$. Let $U$ be the hyperplane with equation $u(x) = 0$; then $\tau(x) = x$ whenever $x \in U$ and $u(d) = 0$, as required. □

The above theorem allows us to write any transvection in the form $\tau = \tau_{u,d}$, where

$$\tau_{u,d}(x) = x + u(x)d, \tag{3.1}$$

with $u$ a linear functional, $0 \neq d \in V$ and $u(d) = 0$.

**Lemma 3.10** ([2, Lemma 2.4.3])**.** *Given $g \in GL(V)$, and using the notation for transvections given in* (3.1)*, we have:*

*(i)* $\tau_{u,d}\tau_{v,d} = \tau_{u+v,d}$*;*

*(ii)* $g\tau_{u,d}g^{-1} = \tau_{u',d'}$*, where* $u' = u \circ g^{-1}$ *and* $d' = g(d)$*.*

*Proof.* We proceed by direct calculation, using (3.1), and noting that $u(v(x)x) = v(x)u(x)$ since $v(x)$ is a scalar.

(i)

$$\begin{aligned}\tau_{u,d}\tau_{v,d}(x) &= \tau_{u,d}(x + v(x)d)\\ &= x + v(x)d + u(x + v(x)d)d\\ &= x + v(x)d + u(x)d + u(v(x)d)d\\ &= x + v(x)d + u(x)d + v(x)u(d)d\\ &= x + (u(x) + v(x))d = \tau_{u+v,d}(x).\end{aligned}$$

(ii)

$$\begin{aligned}g\tau_{u,d}g^{-1}(x) &= g(g^{-1}(x) + u(g^{-1}(x))d)\\ &= x + u(g^{-1}(x))g(d) = \tau_{u',d'}(x).\end{aligned}$$

□

**Theorem 3.11** ([2, Theorem 2.4.4])**.** *Let $\mathcal{T}$ denote the set of all transvections in $GL(V)$ and $\mathcal{T}^* = \mathcal{T} \setminus \{1\}$, then we have:*

*(i) If* $\dim V \geq 2$*, then $\mathcal{T}^*$ is a complete conjugacy class in $GL(V)$;*

*(ii)* $\mathcal{T} \subseteq SL(v)$*;*

*(iii) if* $\dim V \geq 3$*, then $\mathcal{T}^*$ is a complete conjugacy class in $SL(V)$.*

*Proof.* (i) By Lemma 3.10, the conjugate of a transvection is a transvection. For the converse, suppose that $\tau = \tau_{u,d}$ and $\tau' = \tau_{u',d'}$, are two transvections. Choose bases $\{d, v_1, \cdots, v_m\}$ and $\{d', w_1, \cdots, w_m\}$ for the hyperplanes $U$ and $U'$ with equations $u(x) = 0$ and $u'(x) = 0$ respectively. We can then select $v \notin U, w \notin U'$ satisfying $u(v) = u'(w) = 1$. The sets $\{d, v_1, \cdots, v_m, v\}$ and $\{d', w_1, \cdots, w_m, w\}$ are both bases for the whole space $V$, and consequently we

can find $h \in GL(V)$ such that $h(d) = d', h(v_i) = w_i$ and $h(v) = w$. We claim that $\tau' = h\tau h^{-1}$.
Since $h(d) = d'$, both transvections have the same direction, $d'$, and we have that $u'(x) = 0, (u \circ h^{-1})(x) = 0$ are both equations for the hyperplane $U'$, so that there is some $\lambda \neq 0$ such that $u'(x) = \lambda(u \circ h^{-1})(x)$ for all $x \in V$. If we let $x = w$, then $u'(w) = 1$, and $\lambda(u \circ h^{-1})(w) = \lambda u(v) = \lambda$, so $\lambda = 1$. Hence $\tau'$ and $h\tau h^{-1}$ have the same formula, so $\tau' = h\tau' h^{-1}$.

(ii) Since (by (i)) all transvections are conjugate in $GL(V)$, they all have the same non-zero determinant $\delta$. So, $\delta^2 = \det(\tau_{u,d})\det(\tau_{u,d}) = \det(\tau_{u+v}, d) = \delta$, which implies $\delta = 1$.

(iii) Let $\dim V \geq 3$, and $\tau$ and $\tau'$ as in part (i) of this proof. Let the determinant of $h$ (also as found in (i)) be $\mu \neq 1$. Replace the basis vector $v_1$ by $\mu v_1$ (since $\dim V \geq 3$, the dimension of a hyperplane is at least 2, so $d, v_1 \in U$), and let $h \in GL(V)$ be as in (i). Then, we define the map $\phi$ with $\phi(\mu v_1) = v_1$, and $\phi(y) = y$ for all $y \in \{d, v_2, \cdots, v_m, v\}$, and let $h^* = h\phi \in GL(V)$, with $\tau' = h^*\tau {h^*}^{-1}$. Further, $\det h^* = \det h \det \phi = \mu\mu^{-1} = 1$; so we have that $\tau$ and $\tau^*$ are conjugate in $SL(V)$.

□

**Theorem 3.12** ([2, Theorem 2.4.5]). *The set $\mathcal{T}(d)$ of transvections with direction $d$ is an abelian normal subgroup of the stabiliser of $d$ in the action of $SL(V)$ on $V^*$. The groups $\mathcal{T}(d)$ $(d \in V^*)$ are all conjugate in $SL(V)$.*

*Proof.* Theorem 3.11 gives $T \subseteq SL(V)$. We then have that $\mathcal{T}(d)$ is an abelian group from Lemma 3.10 (i), and Lemma 3.10(ii) gives that if $g$ is an element of $SL(V)$ fixing $d$, then $g\mathcal{T}(d)g^{-1} = \mathcal{T}(d)$. Hence $\mathcal{T}(d)$ is an abelian normal subgroup of the stabiliser of $d$. Let $d_1$ and $d_2$ be any two elements of $V^*$, then let $\{d_1, u_1, \cdots, u_m\}$ and $\{d_2, u'_1, \cdots, u'_m\}$ be bases for $V$. Then we can find $\phi \in GL(V)$ with $\phi(d_1) = d_2$ and $\phi(u_i) = u'_i$ (for $1 \leq i \leq m$), and $\psi \in GL(V)$ with $\psi(d_2) = \det(\phi)^{-1}d_2$ and $\psi(u'_i) = u'_i$ (for $1 \leq i \leq m$). It then follows that $\psi\phi \in SL(V)$ with $\psi\phi(d_1) = (\det(\phi))^{-1}d_2$, hence we have an element of $SL(V)$ taking $d_1$ to a scalar multiple of $d_2$, and $\mathcal{T}(d_1)$ and $\mathcal{T}(d_2)$ are conjugate under this element. □

One of the main aims of this section is to show that the set of transvections generates $SL(V)$, for the proof of this, we need to introduce some more notation. Suppose that $y$ is a non-zero element of $V$, and let $\langle y \rangle = \{\lambda y : \lambda \in F\}$. We denote the quotient vector space $V/\langle y \rangle$ by $\overline{V}$, the elements of this set will be cosets of the form $\overline{x} = \{x' \in V : x' - x \in \langle y \rangle\} = x + \langle y \rangle$.

**Lemma 3.13** ([2, Lemma 2.4.6]). *Taking $y, \langle y \rangle, \overline{V}$ be as above. Given $g \in SL(V)$ such that $gy = y$, put $\overline{g}(\overline{x}) = \overline{g}(x + \langle y \rangle) = g(x) + \langle y \rangle = \overline{gx}$. Then:*

*(i) $\overline{g}$ is a well defined element of $SL(\overline{V})$;*

*(ii) if we are given a transvection $\sigma \in SL(\overline{V})$, there is a transvection $\tau \in SL(V)$ such that $\overline{\tau} = \sigma$.*

*Proof.* (i) We first show that $\overline{g}$ is well defined as a linear automorphism of $\overline{V}$. Let $x, u \in V$ with $\overline{x} = \overline{u}$, then $x - u \in \langle y \rangle$, so that $g(x) - g(u) = g(x-u) \in g(\langle y \rangle)$. We also have that $g(\langle y \rangle) = \langle g(y) \rangle = \langle y \rangle$. Hence $\overline{g(x)} = \overline{g(u)}$ and $\overline{g}$ is well defined.

It remains to show that $\overline{g} \in SL(\overline{V})$. Let $\{y, v_2, \cdots, v_n\}$ be a basis for $V$, and let $M$ be the corresponding matrix representing $g$; we are given that $\det M = 1$ (since $g \in SL(V)$). Since $gy = y$, the first column of $M$ is the transpose of $(1, 0, \cdots, 0)$, so that $\det \overline{M} = 1$, where $\overline{M}$ is obtained by deleting the first row and column of $M$. We then note that if $g(v_i) = m_{1i}y + m_{2i}v_2 + \cdots m_{ni}v_n$, where $m_{1i}, \cdots, m_{ni} \in F$, then $\overline{g}(\overline{v_i}) = \overline{g(v_i)} = m_{1i}\overline{y} + m_{2i}\overline{v_2} + \cdots + m_{ni}\overline{v_n} = m_{2i}\overline{v_2} + \cdots m_{ni}\overline{v_n}$, hence, $\overline{M}$ is the matrix representing $\overline{g}$ with respect to the basis $\{\overline{v_2}, \cdots, \overline{v_n}\}$ of $\overline{V}$, and so $\overline{g} \in SL(\overline{V})$.

(ii) If a transvection $\sigma$ in $SL(\overline{V})$ has direction $\overline{d} \in \overline{V}$, and its hyperplane has the equation $\overline{u}(\overline{x}) = 0$, then choose a representative $d \in V$ for $\overline{d}$, and define a linear functional $u$ on $V$ by:

$$u(y) = 0, \ u(x) = \overline{u}(\overline{x}) \ (x \notin Y).$$

We can confirm that this is indeed a linear functional by the following:

$$\begin{aligned} u(x+a) &= \overline{u}(\overline{x+a}) = \overline{u}(\overline{x} + \overline{a}) \\ &= \overline{u}(\overline{x}) + \overline{u}(\overline{a}) = u(x) + u(a), \\ u(\lambda x) &= \overline{u}(\overline{\lambda x}) = \overline{u}(\lambda \overline{x}) = \lambda \overline{u}(\overline{x}) = \lambda u(x). \end{aligned}$$

The transvection $\tau$ on $V$ given by $\tau = \tau_{u,d}$ satisfies $\overline{\tau} = \sigma$, since,

$$\begin{aligned} \overline{\tau}(\overline{x}) &= \overline{\tau(x)} \\ &= \overline{x + u(x)d} \\ &= \overline{x} + \overline{u}(\overline{x})\overline{d} \\ &= \sigma(\overline{x}). \end{aligned}$$

□

**Theorem 3.14** ([2, Theorem 2.4.7]). *The set $\mathcal{T}$ of transvections generates $SL(V)$.*

*Proof.* We want to show that every element of $SL(V)$ can be written as a product of transvections, we proceed by induction on $n = \dim V$. When $n = 1$, $\mathcal{T} = \{1_V\} = SL(V)$, since $V = \langle d \rangle$ and a transvection maps $d$ to $d$, and hence fixes $V$ pointwise, and so the result holds.
Let $n \geq 2$ and $g \in SL(V)$. There are three cases to consider, the first being the case where $g$ fixes some non-zero $y \in V$. In this case, we can apply Lemma 3.13(i) to find a corresponding $\overline{g} \in SL(\overline{V})$, where $\dim \overline{V} = n - 1$. We can then (by induction) assume that $\overline{g}$ is a product of transvections $\overline{g} = \sigma_1 \sigma_2 \cdots \sigma_k$.

We can now apply Lemma 3.13(ii) to find transvections $\tau_1, \cdots, \tau_k \in SL(V)$ satisfying $\overline{\tau_i} = \sigma_i$ for all $i \in 1, \cdots, k$. Let $h = \tau_1\tau_2 \cdots \tau_k$, then we have:

$$\begin{aligned} \overline{hx} &= \overline{h}(\overline{x}) \\ &= \overline{\tau_1\tau_2 \cdots \tau_k}(\overline{x}) \\ &= \overline{\tau_1}\ \overline{\tau_2} \cdots \overline{\tau_k}(\overline{x}) \\ &= \sigma_1 \cdots \sigma_k(\overline{x}) = \overline{g}(\overline{x}) \\ &= \overline{g(x)}. \end{aligned}$$

From this, we have that, for all $x \in V$, $h(x) - g(x)$ is a scalar multiple of $y$, and since $g$ fixes $y$, we have that $g^{-1}h(x) - x$ is a scalar multiple of $y$. Hence $g^{-1}h$ is a transvection with direction $y$. Since $h$ is a product of transvections, so is $g$, and so we have the desired result in this case.
The next case we consider is when $g$ takes $y$ to some linearly independent vector $g(y)$. Then $y - g(y)$ and $g(y)$ are linearly independent, and we can find some linear functional $t$ on $V$ satisfying $t(y-g(y)) = 0$ and $t(g(y)) = 1$. Let $\theta$ be the transvection $\tau_{t,y-g(y)}$ (so that $\theta(x) = x + t(x)(y - g(y))$), then we have:

$$\begin{aligned} \theta g(y) = \theta(g(y)) &= g(y) + t(g(y))(y - g(y)) \\ &= y, \end{aligned}$$

so $\theta g$ fixes $y$, and we can apply the previous argument to see that $\theta g$ is a product of transvections. Hence $g$ is a product of transvections, as required.
The final case we need to consider is the case where $g$ takes $y$ to some linearly dependant (but not equal) vector $g(y)$. We choose some $b \in V$ which is not linearly dependent on $y$, and a linear functional $w$ satisfying $w(g(y)) = 1$ and $w(b) = 0$. Let $\phi = \tau_{w,b}$, (so that $\phi(x) = x + w(x)b$), so that we have:

$$\begin{aligned} \phi g(y) = \phi(g(y)) &= gy + w(g(y))b \\ &= gy + b, \end{aligned}$$

which is linearly independent to $y$, and so we can apply the above case, and again we find that $g$ is a product of transvections. Hence the induction step is complete. □

**Theorem 3.15** ([2, Theorem 2.4.8]). *$SL(n, q)$ coincides with it commutator subgroup when $n \geq 2$ and $(n, q) \neq (2, 2)$, or $(2, 3)$.*

*Proof.* By Theorem 3.14, it is sufficient to show that any transvection can be written as a commutator $[a, b] = aba^{-1}b^{-1}$, with $a, b \in SL(n, q)$.
First, suppose $n \geq 3$, and $\tau \in \mathcal{T}(d)$ (so $\tau$ is a transvection with direction $d$). Let $\sigma \in \mathcal{T}(d)$ be such that $\sigma \neq \tau^{-1}$ so $\tau\sigma \in \mathcal{T}(d)$. Then, by Theorem 3.11(iii), there is some $g \in SL(n, q)$ such that $\tau\sigma = g\sigma g^{-1}$, that is $\tau = [g, \sigma]$.
For the case $n = 2$, we use the matrix representation. If $\tau$ is a given transvection in $SL(2, q)$ with matrix $\begin{pmatrix} 1 & \gamma \\ 0 & 1 \end{pmatrix}$ then we can consider the following identity:

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} 1 & -\beta \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \beta(\alpha^2 - 1) \\ 0 & 1 \end{pmatrix},$$

which is valid over any field. This identity shows that we only need to find $\alpha, \beta \in GF(q)$ such that $\gamma = \beta(\alpha^2 - 1)$. When $q \neq 2, 3$ we can take $\alpha$ to be a primitive element of $GF(q)$ (so that $\alpha^2 \neq 1$) and set $\beta = \gamma(\alpha^2 - 1)^{-1}$. □

## 3.4 PGL and PSL

This section aims to introduce the groups $PGL(V)$ and $PSL(V)$ and give some important results, particularly concerning $PSL(V)$, as we use a specific example of these groups to form the large Mathieu groups. Before we can define these groups, we need the following theorem concerning the centre of $GL(V)$.

**Theorem 3.16** ([2, Theorem 2.4.9]). *The centre of $GL(V)$ consists of the scalar transformations $x \mapsto \lambda x (\lambda \in F^*)$; the centre of $SL(V)$ consists of those scalar transformations for which $\lambda^n = 1$, where $n = \dim V$.*

*Proof.* The scalar transformations clearly belong to the center of $GL(V)$. So we only need to show the converse inclusion. Let $g$ be an element which commutes with every element of $GL(V)$, so that for any $x \in V^*$, g commutes with a transvection $\tau$ which has direction $x$. Since $\tau = g\tau g^{-1}$ and $g\tau g^{-1}$ is a transvection with direction $g(x)$, there must be some $\lambda_x \in F^*$ satisfying $g(x) = \lambda_x x$. This is true for any $x \in V^*$, and since $g$ is linear we have, for $x, y \in V^*$,

$$\lambda_{x+y}(x + y) = \lambda_x x + \lambda_y y.$$

If $x$ and $y$ are linearly independent, we have $\lambda_x = \lambda_y = \lambda_{x+y}$. Otherwise, we can choose a $z$ independent of both $x$ and $y$, so that $\lambda_x = \lambda_y = \lambda_z$. Hence $\lambda_x$ is constant and $g$ is a scalar transformation.
Since, by Theorem 3.11, transvections belong to $SL(V)$, the same argument applies to $SL(V)$, the additional requirement that the determinant is 1 means that we must have $\lambda^n = 1$. □

We are interested in how the action of the general linear group is affected by passing from $V$ to $PG(V)$. Let $g \in GL(V)$, we can define a permutation $\hat{g}$ of $PG(V)$ by the rule:

$$\hat{g}[v] = [g(v)] \quad (v \in V^*).$$

We will call this the induced permutation. Note that this definition is independent of the representative of $[v]$ since if $[v] = [v']$, then $v = \lambda v'$, and $g(v) = \lambda g(v')$, and so $[g(v)] = [g(v')]$. However the assignment $g \mapsto \hat{g}$ is not a faithful representation of $GL(V)$ on $PG(V)$, since some non-identity automorphisms may induce the identity on $PG(V)$.

**Lemma 3.17** ([2, Lemma 2.5.2]). *Given $g \in GL(V)$, the induced permutation $\hat{g}$ is the identity of $PG(V)$ if and only if $g$ is a scalar transformation.*

*Proof.* Suppose that $g$ is a scalar transformation, then it is clear (from the definition of the projective geometry) that $[g(v)] = [v]$, and so the induced permutation $\hat{g}$ is

the identity.
Conversely, suppose that $\hat{g}[v] = [g(v)] = [v]$ for all $v \in V^*$, let $\{v_1, \cdots, v_n\}$ be basis of $V$. Then for any $i, j \in 1, \cdots, n$, we have $\hat{g}([v_i]) = [v_i], \hat{g}([v_j]) = [v_j]$ and $\hat{g}([v_i + v_j]) = [v_i + v_j]$, so that $g(v_i) = \lambda_i v_i, g(v_j) = \lambda_j v_j$, and $g(v_i + v_j) = \lambda(v_i + v_j)$. Then, by the linearity of $g$, we have:

$$\begin{aligned}\lambda(v_i + v_j) = g(v_i + v_j) = g(v_i) + g(v_j) = \lambda_i v_i + \lambda_j v_j \\ \implies (\lambda - \lambda_i)v_i + (\lambda - \lambda_j)v_j = 0.\end{aligned}$$

Since $v_i$ and $v_j$ are linearly independent, we must have $\lambda = \lambda_i = \lambda_j$. Hence, for any element $v = \mu_1 v_1 + \cdots \mu_n v_n \in V$ we have:

$$\begin{aligned}g(v) &= g(\mu_1 v_1 + \cdots + \mu_n v_n) \\ &= \mu_1 g(v_1) + \cdots + \mu_n g(v_n) \\ &= \mu_1(\lambda v_1) + \cdots + \mu_n(\lambda v_n) \\ &= \lambda(\mu_1 v_1 + \cdots + \mu_n v_n) = \lambda v.\end{aligned}$$

Hence $g$ is a scalar transformation, as required. □

Theorem 3.16 tells us that the scalar elements of $GL(V)$ and $SL(V)$ comprise the centres of these groups. Hence, Lemma 3.17 tells us that the centre of $GL(V)$ (respectively $SL(V)$) is the kernel of its permutation representation on $PG(V)$. The main consequence of this which we are interested in is that the quotient group by its center is therefore a group which acts on $PG(V)$. We define these groups as follows.

**Definition 3.18** ([2, Definition 2.5.3]). *The projective general linear group $PGL(V)$ is defined by:*

$$PGL(V) = GL(V)/Z(GL(V)).$$

*Similarly, the projective special linear group $PSL(V)$ is defined by:*

$$PSL(V) = SL(V)/Z(SL(V)).$$

Throughout, when $V = V(n, q)$, we will denote the projective groups as $PGL(n, q)$ and $PSL(n, q)$. Note that $PGL(n, q)$ (resp $PSL(n, q)$) acts on $PG(n - 1, q)$.

**Theorem 3.19** ([2, Theorem 2.5.4]). *$PGL(n, q)$ and $PSL(n, q)$ both act 2-transitively on the points of $PG(n - 1, q)$.*

*Proof.* Given $g \in GL(n, q)$, let $[g]$ denote its coset in $PGL(n, q)$ so that the action of $PGL(n, q)$ on $PG(n - 1, q)$ can be defined by the rule $[g][v] = [g(v)]$ $(v \in V^*(n, q))$. Then, let $([x], [y])$ and $([x'], [y'])$ be two ordered pairs of distinct points in $PG(n - 1, q)$. The ordered pairs $(x, y)$ and $(x', y')$ are then both linearly independent pairs in $V(n, q)$, so we can choose them as the initial members of two ordered bases. Then we we will be able to find some $g \in GL(n, q)$ such that $g(x) = x'$ and $g(y) = y'$. The corresponding element $[g]$ takes $[x]$ to $[x']$ and $[y]$ to $[y']$. Hence we have that $PGL(n, q)$ is 2-transitive on $PG(n - 1, q)$.

To show that $PSL(n,q)$ is 2-transitive, we want to find some $[h] \in PSL(n,q)$ with the same properties as $[g]$. Suppose $\det(g) = \lambda \neq 1$, and let $h \in GL(n,q)$ be the element taking $x$ to $\lambda^{-1}x'$ and acting as $g$ on every other member of the basis (including sending $y$ to $y'$). Then $\det(h) = 1$, so that $[h] \in PSL(n,q)$ and $[h][x] = [\lambda^{-1}x'] = [x']$, $[h][y] = [y']$. So $PSL(n,q)$ acts 2-transitively on $PG(n-1,q)$. □

**Lemma 3.20** ([2, Lemma 2.5.5]). *Let $(G,\Pi)$ denote the permutation group $PSL(n,q)$ acting on the set $PG(n-1,q)$ and suppose $\pi \in \Pi$. The stabiliser $G_\pi$ contains an abelian normal subgroup $H_\pi$; the groups $H_\pi$ are all conjugate in $G$ and they generate $G$.*

*Proof.* Suppose that $p \in V(n,q)$ is chosen so that $[p] = \pi$; then the set of transvections with direction $p$ is independent of the chosen $p$. Let $H_\pi$ be the image of $T(p)$ under the morphism $SL(n,q) \to PSL(n,q)$. Theorem 3.12 gives that $H_\pi$ is an abelian normal subgroup of $G_\pi$, and that every $H_\pi$ is conjugate in $G$. Similarly, Theorem 3.14 gives us that the $H_\pi$ generate $G$. □

**Lemma 3.21** ([2, Lemma 2.5.6]). *$PSL(n,q)$ coincides with its commutator subgroup, when $n \geq 2$ and $(n,q) \neq (2,2)$, or $(2,3)$.*

*Proof.* This follows from Theorem 3.15, since $PSL(n,q)$ is a quotient of $SL(n,q)$. □

**Theorem 3.22** ([2, Theorem 2.5.7]). *The group $PSL(n,q)$ is simple, provided $n \geq 2$ and $(n,q) \neq (2,2)$ or $(2,3)$.*

*Proof.* We use the same notation as in Lemma 3.20, letting $(G,\Pi)$ denote $PSL(n,q)$ acting on $PG(n-1,q)$. Suppose that $N \neq 1$ is a normal subgroup of $G$. Then, since $G$ acts 2-transitively (and hence primitively) Theorem 2.14 gives us that $N$ must be transitive on $\Pi$.

Fix $\pi \in \Pi$, and let $H = H_\pi$ be as in Lemma 3.20. We consider the set $NH \subseteq G$. Since $N$ is normal, $NH = HN$ and $NH$ is a subgroup of $G$. We want to show that $NH = G$. Let $k \in H_\sigma$, for some $\sigma \in \Pi$, and choose $n \in N$ with $n(\pi) = \sigma$. Then $n^{-1}\pi n \in H = H_\pi$, and so $k \in NHN = NH$. We have from Lemma 3.20 that the groups $H_\sigma$ generate $G$, and so $NH = G$.

We now want to show that any commutator in $G$ is an element of $N$. Let $g_1, g_2 \in G$, since $G = NH$, we can write $g_1 \in Nh_1$ and $g_2 \in Nh_2$ for some $h_1, h_2 \in H$, so $g_1 g_2 g_1^{-1} \in Nh_1 Nh_2 Nh_1^{-1}$. From Lemma 3.20 we have that $H$ is abelian, and by assumption $N$ is normal, hence we have:

$$\begin{aligned}
Nh_1Nh_2Nh_1^{-1} &= Nh_1h_2h_1^{-1}N \\
&= Nh_2N \\
&= Nh_2 \\
&= Ng_2.
\end{aligned}$$

Hence $g_1g_2g_1^{-1} \in Ng_2$, and so $[g_1, g_2] \in N$. Since, by Lemma 3.21 $G$ coincides with its commutator subgroup, we must have $G = N$, and so $G$ is simple. □

**Lemma 3.23** ([2, Lemma 2.6.1]). *The number of points in $PG(n, q)$ is $\frac{(q^{n+1}-1)}{q-1}$.*

*Proof.* The points of $PG(n, q)$ are equivalence classes $[x]$ of points in $x \in V = V(n+1, q)$. Fix a basis for $V$, and let the the coordinates for a typical point $x$ be given by $(x_1, x_2, \cdots, x_n, x_{n+1})$. If $x_{n+1} \neq 0$, then the point $(y_1, y_2, \cdots, y_n, 1)$ with $y_i = x_i(x_{n+1})^{-1}$ (for each $i = 1, \cdots, n+1$), is a representative of $[x]$ in $PG(n, q)$. In fact, it is the unique representative with final coordinate 1. Since each $y_i$ (for $i = 1, \cdots, n$) can be any element of $GF(q)$, there are $q^n$ such elements in $PG(n, q)$. Similarly, if $x_{n+1} = 0$ and $x_n \neq 0$, then we can find the unique representative of $[x]$ of the form $(y_1, y_2, \cdots, y_{n-1}, 1, 0)$, so that there are $q^{n-1}$ such points. Continuing in this way, we have that there are $q^n + q^{n-1} + \cdots q + 1 = (q^{n+1} - 1)/(q - 1)$ points of $PG(n, q)$ as required. □

**Theorem 3.24** ([2, Theorem 2.6.3]). *For all $n \geq 2$ we have:*

*(i)* $|PGL(n, q)| = q^{n(n-1)/2} \prod_{i=2}^{n}(q^i - 1)$*;*

*(ii)* $|PSL(n, q)| = (q-1, n)^{-1}|SL(n, q)|$*, where* $(,)$ *denotes the greatest common divisor.*

*Proof.* (i) By definition, $PGL(n, q)$ is the quotient of $GL(n, q)$ by its centre, and from 3.16, the centre consists of $q - 1$ elements. Hence applying the formula for $GL(n, q)$ given in Theorem 3.3 we get the result.

(ii) By definition $PSL(n, q)$ is the quotient of $SL(n, q)$ by its centre, and again by Theorem 3.16 consists of the scalar transformations with $\lambda^n = 1$. Since the non-zero elements of $GF(q)$ form, under multiplication, a cyclic group of order $q - 1$, there are $(q - 1, n)$ such transformations.

□

In particular, we have that $|PSL(3, 4)| = 20160$.

# 4 Designs

In this section, we will be looking at purely combinatorial structures on a set. We want to find a family of subsets which satisfy some particular conditions. We will be interested in determining when these families exist and if we can find a transitive group of permutations on them. To this aim, we introduce designs.

**Definition 4.1** (t-design [7, Definition 3.16]). *Let $v \geq k \geq t \geq 1$ and $\lambda \geq 1$ be natural numbers. A $t$-design $(X, \mathcal{B})$ with parameters $v, k, \lambda$ consists of a family $\mathcal{B}$ of subsets of $X$ such that:*

*(i)* $|X| = v$*;*

*(ii)* $|B| = k$*, for all* $B \in \mathcal{B}$*;*

*(iii) each $t$-subset of $X$ is contained in exactly $\lambda$ sets from $\mathcal{B}$.*

*We call the elements of $X$ points, and the elements of $\mathcal{B}$ blocks.*

A $t$-design with $t \geq 2$ and $\lambda = 1$ is called a Steiner system, we will use the notation $S(t, k, v)$.

**Example 4.2.** *The Fano plane, $PG(2, 2)$ as introduced in example 3.6 is a $2-(7, 3, 1)$ design. We can re-label the points given in example 3.6 as follows; let $[1, 0, 0] = 1, [1, 1, 0] = 2, [1, 0, 1] = 3, [0, 1, 0] = 4, [0, 1, 1] = 5, [1, 1, 1] = 6, [0, 0, 0] = 7$, so that the points of the Fano plane can be written as the set $X = \{1, 2, 3, 4, 5, 6, 7\}$. We can now explicitly write the blocks of $X$ when viewed as a $2-(7, 3, 1)$ design as follows:*

$$\begin{array}{cccc} 1\ 2\ 4 & 2\ 3\ 5 & 3\ 4\ 6 & 4\ 5\ 7 \\ 1\ 5\ 6 & 2\ 6\ 7 & 1\ 3\ 7. & \end{array}$$

**Theorem 4.3** ([2, Theorem 3.2.2])**.** *A $t$-design $(X, \mathcal{B})$ is also an $s$-design, for any $0 \leq s \leq t$. If the parameters of the $t$-design are $t-(v, k, \lambda)$, then its parameters as an $s$-design are $s-(v, k, \lambda_s)$, where,*

$$\lambda_s = \lambda \cdot \frac{(v-s)(v-s-1)\cdots(v-t+1)}{(k-s)(k-s-1)\cdots(k-t+1)}.$$

*Proof.* We proceed by induction on $t-s$, when $t-s=0$ the result clearly holds. Suppose that the result holds for $s = i+1$ for $0 \leq i \leq t-1$, so that each $(i+1)$-subset of $X$ occurs as a subset of exactly $\lambda_{i+1}$ blocks, so we want to show that the result holds for $s = i$. Let I be any $i$-subset of $X$ and consider the pairs $(x, \beta) \in X \times \mathcal{B}$ satisfying the conditions:

$$x \in X \setminus \mathrm{I}\ ,\ \mathrm{I} \cup \{x\} \subseteq \beta.$$

We then note that we can find two formulas for the number of pairs $(x, \beta)$. First, we can select an $x \in X \setminus I$, of which there are $(v-1)$ choices, and then select $\beta \in \mathcal{B}$ with $I \cup \{x\} \subseteq \beta$, of which there are $\lambda_{i+1}$ choices, so we have $(v-i)\lambda_{i+1}$ pairs $(x, \beta)$. Alternatively, we can first select $\beta_0 \in \mathcal{B}$ with $I \subseteq \beta_0$, of which we say there are $\lambda_i(I)$ choices, and then select an $x \in \beta_0 \setminus I$, of which there are $(k-i)$ choices. This approach gives us that there are $(k-i)\lambda_i(\mathrm{I})$ choices for $(x, \beta)$, since we have two different formulas for the same number, we then have the following:

$$(v-i)\lambda_{i+1} = (k-i)\lambda_i(\mathrm{I}). \tag{4.1}$$

Where $\lambda_i(\mathrm{I})$ is the number of blocks containing I. We can then use (4.1) to see:

$$\lambda_i(\mathrm{I}) = \frac{v-i}{k-i}\lambda_{i+1}, \tag{4.2}$$

so that $\lambda_i(\mathrm{I})$ is independent of I, and so we have an $i$-design. Using repeated application of (4.2), we then have the following:

$$\begin{aligned}
\lambda_s &= \frac{v-s}{k-s}\lambda_{s+1} \\
&= \frac{(v-s)(v-s-1)}{(k-s)(k-s-1)}\lambda_{s+2} \\
&\vdots \\
&= \frac{(v-s)(v-s-1)\cdots(v-(t-1))}{(k-s)(k-s-1)\cdots(k-(t-1))}\lambda_t.
\end{aligned}$$

Since $\lambda_t = \lambda$, we have the required formula, so the induction step is complete. □

The proof of Theorem 4.3 allows us to establish two useful formulas when working with designs. From this point on, we will use the following notation when working with designs; the number of blocks containing any given point (previously denoted $\lambda_1$) will be given by $r$; and $|\mathcal{B}|$ will be denoted by $b$. So we have that the equation (4.1) can be written as:

$$(v-i)\lambda_{i+1} = (k-i)\lambda_i \quad (0 \leq i \leq t-1), \tag{4.3}$$

and, in particular, when $i = 0$ we get the equation:

$$vr = bk. \tag{4.4}$$

Throughout the remainder of this dissertation, we will be interested in particular permutations of designs known as automorphisms of designs.

**Definition 4.4** (Automorphisms of designs [2, Definition 3.4.1])**.** *An automorphism of a design $(X, \mathcal{B})$ is a permutation $\pi$ of $X$ such that $\beta \in \mathcal{B}$ if and only if $\pi(\beta) \in \mathcal{B}$.*

We can see that the automorphisms of $(X, \mathcal{B})$ form a group which acts on $X$. Since an automorphism takes blocks to blocks, the group also has a permutation representation on the set $\mathcal{B}$.

**Theorem 4.5** ([2, Theorem 3.4.3])**.** *Let $(G, X)$ be a $t$-transitive permutation group ($t \geq 2$), and suppose $\beta$ is a subset of $X$, with $|\beta| = k$, $|X| = v$, and $t < k < v-1$. Then the set*

$$\mathcal{B} = \{g(\beta) : g \in G\}$$

*is the set of blocks of a $t$-design $(X, \mathcal{B})$, and $G$ is a group of automorphisms acting transitively on $\mathcal{B}$.*

*Proof.* Let $S$ and $T$ be any two $t$-subsets of $X$, then there is some $h$ in $G$ satisfying $h(S) = T$. If $S$ belongs to a block $g(\beta) \in \mathcal{B}$, then $T$ belongs to the block $hg(\beta)$. So $S$ and $T$ occur as subsets of a block the same number of times. □

We now want to find the parameters of the block constructed in this theorem; $t, v$ and $k$ are already given, so it remains to find $\lambda$. For this calculation, we use the equations (4.3) and (4.4), noticing that $b$ is the length of the orbit of $\beta$ under the action of $G$. So it follows from the Orbit-Stabiliser Theorem that $b = |G : G_{(\beta)}|$, with $G_{(\beta)}$ the set-wise stabiliser of $\beta$, hence,

$$\lambda = \lambda_t = \frac{k(k-1)\cdots(k-t+1)}{v(v-1)\cdots(v-t+1)}|G : G_{(\beta)}|.$$

We will be interested in extending t-designs. To this aim we introduce the concept of a contraction of a t-design. Let $D = (X, \mathcal{B})$ be a t-design with parameters $t-(v,k,\lambda)$, and let $x \in X$, then the family of sets,

$$\mathcal{B}_x = \{\beta \setminus \{x\} : x \in \beta, \beta \in \mathcal{B}\}$$

gives a block design on $X \setminus \{x\}$. To see this, let $\{x_1, \cdots, x_{t-1}\} \subseteq X \setminus \{x\}$, then there are $\lambda$ blocks $\beta$ containing $\{x_1, \cdots x_{t-1}\} \cup \{x\}$ and hence there are $\lambda$ blocks $\beta \setminus \{x\} \in \mathcal{B}_x$ containing $\{x_1, \cdots x_{t-1}\}$. This design will have parameters $(t-1)-(v-1, k-1, \lambda)$. The design $D_x = (X \setminus \{x\}, \mathcal{B}_x)$ is called a contraction of the design $D$. We also notice that $b_x$ (the number of blocks of $D_x$) is equal to the number of blocks of $D$ containing $x$ (which is $r$) and so, applying equation (4.4), we get:

$$b_x = \frac{bk}{v}. \tag{4.5}$$

**Example 4.6** ([2, p.67]). *Let us consider the* $3-(8,4,1)$ *design, where we take the set* $X = \{0,1,2,3,4,5,6,7\}$*, then the design has 14 blocks given as follows:*

| | |
|---|---|
| 0 1 2 4 | 3 5 6 7 |
| 0 2 3 5 | 1 4 6 7 |
| 0 3 4 6 | 1 2 5 7 |
| 0 4 5 7 | 1 2 3 6 |
| 0 1 5 6 | 2 3 4 7 |
| 0 2 6 7 | 1 3 4 5 |
| 0 1 3 7 | 2 4 5 6. |

*Letting* $x = 0$*, we see that the block design on* $X \setminus \{x\}$ *is the* $2-(7,3,1)$ *discussed in example 4.2.*

Since we are more interested in extending designs, we want to try and reverse this process, which we can do in the following manner.

**Definition 4.7** (Extendable designs [2, Definition 3.5.2]). *The design* $D^+ = (X^+, \mathcal{B}^+)$ *is an extension of* $D = (X, \mathcal{B})$ *if* $X^+ = X \cup \{z\}$*, for some point* $z \notin X$*, and the contraction* $(D^+)_z$ *is just* $D$*.* $D$ *is said to be extendable if it has some extension* $D^+$*.*

**Lemma 4.8** ([2, Lemma 3.5.3]). *In order for a* $t-(v,k,\lambda)$ *design with* $b$ *blocks to be extendable* $(k+1)$ *should divide* $b(v+1)$*.*

*Proof.* Suppose that the design is extendable, then the extension would give a $(t+1)-(v+1,k+1,\lambda)$ design with $b^+$ blocks, and using (4.5) we have:

$$b = \frac{b^+(k+1)}{v+1}.$$

So, to ensure that $b^+$ is an integer we need $(k+1)$ to divide $b(v+1)$. □

This lemma allows us to prove the following theorem, which will be useful when constructing the large Mathieu groups.

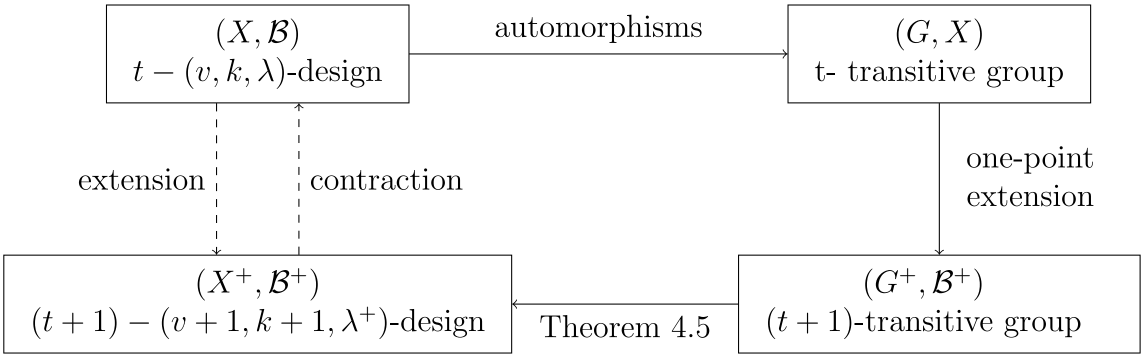
**Theorem 4.9** ([2, Theorem 3.5.4])**.** *Let $q$ be a prime power. A necessary condition that a design with parameters $2-(q^2+q+1, q+1, 1)$ (in particular $PG(2,q)$) should be extendable is that $q = 2$ or $4$.*

*Proof.* If we have an extendable $2-(q^2+q+1, q+1, 1)$ design, then the extension would have parameters $3-(q^2+q+2, q+2, 1)$, and so the condition in Lemma 4.8 says that $q+2$ should divide $(q^2+q+2)(q^2+q+1)$. Then

$$\frac{(q^2+q+2)(q^2+q+1)}{q+2} = (q^3+4q-5) + \frac{12}{q+2},$$

so that $q+2$ divides 12, since $q$ is a prime power, we must have $q = 2$ or $4$. □

As shown in Example 4.6, $PG(2,2)$ is extendable, and in fact $PG(2,4)$ is extendable too, we will discuss this at length in the next section. We can construct extensions of designs by simultaneously extending the automorphism group of the design, using the one-point extension discussed in Theorem 2.8.
The below flow diagram shows the procedure, where we begin with the $t-(v,k,\lambda)$ design $(X,\mathcal{B})$, and we find the $t$-transitive group $(G,X)$ acting transitively on $\mathcal{B}$. Then we are interested in the case where $G$ admits a one-point extension $(G^+,X^+)$, where $X^+ = X \cup \{z\}$. We can then apply Theorem 4.5, taking the basic block $\beta = \beta_0 \cup \{z\}$, for some $\beta_0 \in \mathcal{B}$, to construct a $(t+1)-(v+1,k+1,\lambda^+)$ design $(X^+,\mathcal{B}^+)$. It is not always the case that this design is an extension of our original one, as, for example, we may have $\lambda \neq \lambda^+$ [2, p.69].



The following theorem gives us the conditions for which the procedure above does result in the new design being an extension of the original.

**Theorem 4.10** ([2, Theorem 3.5.5])**.** *Let $(X, \mathcal{B})$ be a $t-(v, k, \lambda)$ design whose automorphism group $G$ is transitive on blocks, and suppose that $(G, X)$ is $t$-transitive and has a one-point extension. The design $(X^+, \mathcal{B}^+)$ constructed by the technique shown in the above flow diagram is an extension of $(X, \mathcal{B})$ if and only if $\lambda^+ = \lambda$.*

*Proof.* For the forward direction, suppose $(X^+, \mathcal{B}^+)$ is an extension of $(X, \mathcal{B})$, then it follows that $(X, \mathcal{B})$ is a contraction of $(X^+, \mathcal{B}^+)$ and so we have $\lambda^+ = \lambda$.
For the converse, let $\lambda^+ = \lambda$. Then, applying Theorem 4.3 to $(X^+, \mathcal{B}^+)$ we have:

$$\begin{aligned} r^+ = \lambda_1^+ &= \lambda_2^+ \cdot \frac{((v+1)-1)}{(k+1)-1} \\ &\vdots \\ &= \lambda_{t+1}^+ \cdot \frac{((v+1)-1)\cdots((v+1)-t)}{((k+1)-1)\cdots((k+1)-t)} \\ &= \lambda^+ \cdot \frac{v(v-1)\cdots(v-t+1)}{k(k-1)\cdots(k-t+1)}. \end{aligned}$$

Then, applying Theorem 4.3 to $(X, \mathcal{B})$ gives

$$r = \lambda \cdot \frac{(v-1)\cdots(v-t+1)}{(k-1)\cdots(k-t+1)},$$

so that,

$$\frac{r^+}{r} = \frac{\lambda^+}{\lambda} \cdot \frac{v}{k}.$$

Then, since $\lambda^+ = \lambda$ (and recalling equation (4.4)),we can rearrange this to give:

$$r^+ = \frac{vr}{k} = \frac{bk}{k} = b.$$

Now, following the notation given in Definition 4.7, let $z \notin X$, we have that $z$ is contained in $r^+$ blocks of $\mathcal{B}^+$. Therefore, the contraction of $(X^+, \mathcal{B}^+)$ with respect to $z$ must have $r^+ = b$ blocks. We want to show that these are exactly the blocks in $\mathcal{B}$.
Suppose the basic block used in the construction of $(X^+, \mathcal{B}^+)$ is $\beta = \beta_0 \cup \{z\}$, for some $\beta_0 \in \mathcal{B}$ and let $\gamma$ be any member of $\mathcal{B}$. Since $G$ is transitive on $\mathcal{B}$, we can find $g \in G$ such that $\gamma = g(\beta_0)$. Thus $\gamma \cup \{z\} = g(\beta)$, and $\gamma \cup \{z\}$ is in $\mathcal{B}^+$, since $g \in G \leq G^+$. Contracting $\mathcal{B}^+$ with respect to $z$ gives the block $\gamma$. □

# 5 The Mathieu groups

In this section, we will be constructing the large Mathieu groups, $M_{22}, M_{23}$ and $M_{24}$, using one-point extensions of permutation groups.
Throughout, we denote the points of $PG(2, 4)$ by equivalence classes of coordinate triples $[u, v, w]$, with $u, v, w \in GF(4)$. With this in mind, we recall from Example

3.1 the following properties of $GF(4)$, we have $GF(4)^* = \{1, \alpha, \alpha^2\}$ so that for any $x \in GF(4)^*$, we have $x^3 = 1$, and $x^4 = x$. Note that, from Example 3.1, we also have that for $x \in GF(4)^*$, $x^2 + x + 1 = 0$.
We can then choose a primitive element $t \in GF(4)$ and define three permutations of $PG(2, 4)$ as follows:

$$\begin{aligned} f_1[u, v, w] &= [u^2 + vw, v^2, w^2], \\ f_2[u, v, w] &= [u^2, v^2, w^2 t], \\ f_3[u, v, w] &= [u^2, v^2, w^2]. \end{aligned}$$

Since $u, v, w \in GF(4)$, we have the permutation $\alpha : u \mapsto u^2$ (which is the permutation $(t \;\; t^2)$) satisfies:

$$\begin{aligned} \alpha(x + y) = (x + y)^2 = x^2 + y^2 = \alpha(x) + \alpha(y), \\ \alpha(xy) = (xy)^2 = \alpha(x)\alpha(y). \end{aligned}$$

Hence, $\alpha$ is an automorphism and so the three permutations $f_1, f_2, f_3$ above are all permutations of elements of $PG(2, 4)$. Further, each of the permutations are invertible as:

$$\begin{aligned} (f_1)^2[u, v, w] &= f_1[u^2 + vw, v^2, w^2] = [(u^2 + vw)^2 + v^2w^2, v^4, w^4] \\ &= [u^4 + 2u^2vw + v^2w^2 + v^2w^2, v, w] \\ &= [u, v, w], \\ (f_2)^2[u, v, w] &= f_2[u^2, v^2, w^2t] = [u^4, v^4, w^4t^3] \\ &= [u, v, w], \\ (f_3)^2[u, v, w] &= f_3[u^2, v^2, w^2] = [u^4, v^4, w^4] \\ &= [u, v, w]. \end{aligned}$$

Hence, we have $f_i^{-1} = f_i$ for each $i = 1, 2, 3$. These permutations will be used in the construction of the large Mathieu groups $M_{22}, M_{23}$ and $M_{24}$. We begin with the construction of $M_{22}$.

**Theorem 5.1** ([2, Theorem 3.6.1])**.** *The permutation group $PSL(3, 4)$ acting on $PG(2, 4)$ has a one-point extension.*

*Proof.* Following the notation of Theorem 2.8, let $G = PSL(3, 4)$ and $X = PG(2, 4)$, and $* = \infty$. Define $h$ as the permutation which switches $[1, 0, 0]$ and $\infty$, and acts as $f_1$ on the rest of $X$. Let $g$ be the permutation defined by $g[u, v, w] = [v, u, w]$, which written in matrix form is:

$$g = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which has determinant $-1 = 1 \in GF(4)$, so that $g \in G$. The elements $x$ and $y$ considered in Theorem 2.8 are taken as $[1, 0, 0]$ and $[0, 1, 0]$. We have defined $h$ and $g$ to satisfy conditions $(i)$ and $(ii)$ of the theorem, so it only remains to show that

conditions $(iii)$ and $(iv)$ hold.
To check condition $(iii)$, we have that $h^2$ fixes $[1,0,0]$ and $\infty$, and $f_1^{-1} = f_1$, hence $h^2 = 1 \in G$, as required. To check that $(gh)^3 \in G$, we can calculate, for $[u,v,w] \neq [1,0,0], [0,1,0]$:

$$\begin{aligned}
(gh)[u,v,w] &= g[u^2+vw, v^2, w^2] \\
&= [v^2, u^2+vw, w^2], \\
(gh)^2[u,v,w] &= (gh)[v^2, u^2+vw, w^2] \\
&= [u+v^2w^2, v+(u^2+vw)w^2, w], \\
(gh)^3[u,v,w] &= (gh)[u+v^2w^2, v+(u^2+vw)w^2, w] \\
&= [v^2+(u+v^2w^2)w, u^2+vw+u^2w^3, w^2] \\
&= [uw+v^2(1+w^3), vw+u^2(1+w^3), w^2].
\end{aligned}$$

If $w \neq 0$, then $w^3 = 1$ (again, since we are working with elements of $GF(4)$), so that $(gh)^3[u,v,w] = [uw, vw, w^2] = [u,v,w]$. If $w = 0$ and $uv \neq 0$, then we have $(gh)^3[u,v,w] = [v^2, u^2, 0] = [u,v,0]$, and so we have that $(gh)^3 = 1$. From the remark given after Theorem 2.8, we have that $(gh)^3$ fixes $x, y$ and $\infty$. Hence we have $(gh)^3 = 1 \in G$, and so condition $(iii)$ is satisfied.
For condition $(iv)$, let $p \in G_x$, so that $p$ is an element of $G$ fixing $[1,0,0]$, which will have representative matrix:

$$P = \begin{pmatrix} 1 & a & l \\ 0 & b & m \\ 0 & c & n \end{pmatrix}, \ \det P = bn - cm = 1.$$

So we have $p([x,y,z]) = [x+ay+lz, by+mz, cy+nz]$. We are interested in $hph$, we have that:

$$\begin{aligned}
hph([x,y,z]) &= hp([x^2+yz, y^2, z^2]) \\
&= h([x^2+yz+ay^2+lz^2, by^2+mz^2, cy^2+nz^2]) \\
&= [x^4+y^2z^2+a^2y^4+l^2z^4+(by^2+mz^2)(cy^2+nz^2), b^2y^4+m^2z^4, c^2y^4+n^2z^4] \\
&= [x+y^2z^2+a^2y+l^2z+bcy^4+mcy^2z^2+mnz^4+bny^2z^2, b^2y+mz^2, c^2y+n^2z] \\
&= [x+(a^2+bc)y+(l^2+mn)z+(1+bn+mc)z^2y^2, b^2y+m^2z, c^2y+n^2z] \\
&= [x+(a^2+bc)y+(l^2+mn)z, b^2y+m^2z, c^2y+n^2z].
\end{aligned}$$

Hence, $hph$ can be represented by the matrix:

$$\begin{pmatrix} 1 & a^2+bc & l^2+mn \\ 0 & b^2 & m^2 \\ 0 & c^2 & n^2 \end{pmatrix},$$

which fixes $[1,0,0]$ and belongs to $PSL(3,4)$ since $b^2n^2 - c^2m^2 = (bn-cm)^2 = 1$. Hence, $hph \in G_x$, since this holds for any element of $G_x$, we have $hG_xh \leq G_x$. Then, since both $hG_xh$ and $G_x$ are finite sets with the same order we have $G_x = hG_xh$, so condition (iv) is satisfied. Hence, all the conditions of Theorem 2.8 are satisfied and $\langle G, h\rangle$ is a one-point extension of $G = PSL(3,4)$. □

From Theorem 3.19, we know that $PSL(2,4)$ is 2-transitive on the 21 points of $PG(2,4)$, so it follows that the one-point extension is 3-transitive on 22 points, and this group is called the Mathieu group $M_{22}$. The order of $M_{22}$ is $22 \cdot |PSL(3,4)| = 443,520$.

We can now look at the design associated with $M_{22}$, following the method given in Chapter 4, beginning with the $2-(21,5,1)$ design $PG(2,4)$, which has the group $PSL(3,4)$ acting 2-transitively on its points. Since $M_{22}$ is a one-point extension of $PSL(3,4)$, we can use the procedure given in the flow diagram to construct the following design.

**Theorem 5.2** ([2, Theorem 3.6.4] ). *$PG(2,4)$ is extendable, giving a $3-(22,6,1)$ design, on which $M_{22}$ acts as a group of automorphisms.*

*Proof.* Following the notation of Theorem 4.5, let $\beta$ be the union of $\infty$ with the line $\ell$ in $PG(2,4)$ whose equation is $w = 0$. Let $\mathcal{B}^+$ denote the set of blocks generated by the action of $M_{22}$ on $\beta$ and let $b^+ = |\mathcal{B}^+|$.

The set-wise stabiliser of $\beta$ in $M_{22}$ is transitive on the six points of $\beta$, since it contains $h$ (as in the proof of Theorem 5.1) and the set-wise stabiliser of $\ell$ in $PSL(3,4)$ (which is transitive on the five points of $\ell$). Thus we may calculate, using the abbreviations $H = M_{22}$, $G = PSL(3,4)$:

$$\begin{aligned} b^+ &= |H : H_{(\beta)}| \\ &= \frac{|H : G_{(\ell)}|}{|H_{(\beta)} : G_{(\ell)}|} \\ &= \frac{|H : G| \cdot |G : G_{(\ell)}|}{|H_{(\beta)}(\infty)|} \\ &= \frac{22 \cdot 21}{6} \\ &= 77. \end{aligned}$$

So we have, using equation (4.4), $r^+ = 21$. We can then use equation (4.1) to find $\lambda_2^+ = 5, \lambda^+ = 1$ and, by Theorem 4.10, the new design is an extension of the original one. □

We can now show that it is possible to extend $M_{22}$.

**Theorem 5.3.** *The Mathieu group $M_{22}$ has a one point extension $M_{23}$.*

*Proof.* We proceed in the same manner as in Theorem 5.1, except now we denote our added element as $\infty'$, then define $h'$ to switch $\infty$ and $\infty'$ and act as $f_2$ on the rest of the set (which in this case is $PG(2,4)$), and let $g' = h$ (as in the proof of Theorem 5.1). In this case we are taking the element $x = \infty$ and $y = [1,0,0]$. We need to check that the conditions of Theorem 2.8 are satisfied. Again, conditions (i) and (ii) are satisfied by how we have defined $h'$ and $g'$.

We now check condition (iii); first we check $(h')^2$; clearly $(h')^2(\infty) = \infty$ and $(h')^2(\infty') = \infty'$. Then for any other element $[u,v,w]$ we have,

$$(h')^2[u,v,w] = [u,v,wt^3] = [u,v,w].$$

So $(h')^2$ is the identity map. We now check $(g'h')^3$, as above, we need only verify the result for $[u,v,w] \in PG(2,4)$ with $[u,v,w] \neq [1,0,0]$, so we have:

$$\begin{aligned}
(g'h')[u,v,w] &= g'[u^2, v^2, w^2t] \\
&= [u + v^2w^2t, v, wt^2], \\
(g'h')^2[u,v,w] &= (g'h')[u + v^2w^2t, v, wt^2] \\
&= [u + v^2w^2t + v^2w^2t^2, v, wt], \\
(g'h')^3[u,v,w] &= (g'h')[u + v^2w^2t + v^2w^2t^2, v, wt] \\
&= [u + v^2w^2(t^2+t+1), v, w] \\
&= [u,v,w].
\end{aligned}$$

Hence, $(g'h')^3$ acts as the identity on every element as required.
We now check condition (iv), let $p' \in (M_{22})_\infty = PSL(3,4)$. We want to show that $h'p'h' \in (M_{22})_\infty$. We first have that:

$$\begin{aligned}
h'p'h'(\infty) = h'p'(\infty') = h'(\infty') = \infty, \\
h'p'h'(\infty') = h'p'(\infty) = h'(\infty) = \infty'.
\end{aligned}$$

For the remaining elements, we again look at the representative matrix for $p'$:

$$P' = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & j \end{pmatrix}, \ \det P' = a(ej - hf) - d(bj - hc) + d(bf - ce) = 1.$$

So that $p'([x,y,z]) = [ax + dy + gz, bx + ey + hz, cx + fy + jz$. We are interested in $h'p'h'$, which gives:

$$\begin{aligned}
h'p'h'([x,y,z]) &= h'p'([x^2, y^2, z^2t]) \\
&= h'([ax^2 + dy^2 + gz^2t, bx^2 + ey^2 + hz^2t, cx^2 + fy^2 + jz^2t]) \\
&= [a^2x^4 + d^2y^4 + gz^4t^2, b^2x^4 + e^2y^4 + h^2z^4t^2, t(c^2x^4 + f^2y^4 + j^2z^4t^2)] \\
&= [a^2x + d^2y + g^2t^2z, b^2x + e^2y + h^2t^2z, c^2tx + f^2ty + j^2z].
\end{aligned}$$

So we have that $h'p'h'$ has representative matrix:

$$M = \begin{pmatrix} a^2 & d^2 & g^2t^2 \\ b^2 & e^2 & h^2t^2 \\ c^2t & f^2t & j^2 \end{pmatrix}.$$

This matrix has determinant:

$$\begin{aligned}
\det M &= a^2(e^2j^2 - h^2f^2) - d^2(b^2j^2 - h^2c^2) + g^2t^2(b^2f^2t - e^2c^2t) \\
&= (a(ej - hf) - d(bj - hc) + d(bf - ce))^2 \\
&= 1.
\end{aligned}$$

So that $h'p'h' \in (M_{22})_\infty$. Since this holds for any element of $(M_{22})_\infty$, condition (iv) is satisfied. Hence, all the conditions of Theorem 2.8 are satisfied and $M_{23} = \langle M_{22}, h' \rangle$ is a one point extension of $M_{22}$. □

Since $M_{22}$ is 3-transitive on 22 points, the Mathieu group $M_{23}$ is 4-transitive on 23 points. The order of $M_{23}$ is then $23 \cdot |M_{22}| = 10,200,960$. Again, we can now consider the design associated with $M_{23}$.

**Theorem 5.4.** *The* $3-(22,6,1)$ *design is extendable, giving a* $4-(23,7,1)$ *design on which* $M_{23}$ *acts as a group of automorphisms.*

*Proof.* We proceed in the same way as above, this time taking $\beta' = \beta \cup \{\infty'\} = \ell \cup \{\infty, \infty'\}$. Let $(\mathcal{B}')^+$ denote the set of blocks generated by the action of $M_{23}$ on $\beta'$, and label $(b')^+ = |(\mathcal{B}'^+)|$. The set-wise stabiliser of $\beta'$ in $M_{23}$ is transitive on the 7 points of $\beta'$, since it contains $h'$ (which switches $\infty$ and $\infty'$) and the set-wise stabiliser of $\beta$ in $M_{22}$ (which is transitive on the six points of $\beta$).
Again, we can calculate as follows, with the abbreviations $K = M_{23}, H = M_{22}$:

$$\begin{aligned}(b')^+ &= |K : K_{(\beta')}| \\ &= \frac{|K|}{|K_{(\beta')}|} \\ &= \frac{(\frac{|K|}{|H|})(\frac{|H|}{|G|})(\frac{|G|}{|G_{(\ell)}|})}{(\frac{|K_{(\beta')}|}{|H_{(\beta')}|})(\frac{|H_{(\beta')}|}{|G_{(\ell)}|})} \\ &= \frac{23 \cdot 22 \cdot 21}{6 \cdot 7} \\ &= 253.\end{aligned}$$

From this we can then calculate $r^+ = 77, \lambda_2^+ = 21, \lambda_3^+ = 5, \lambda_4^+ = \lambda = 1$. □

We now look at a one-point extension of $M_{23}$.

**Theorem 5.5.** *The Mathieu group* $M_{23}$ *has a one point extension* $M_{24}$

*Proof.* We follow the same process as above, this time letting $* = \infty''$. We now take $h''$ to be the permutation which switches $\infty'$ and $\infty''$, fixes $\infty$ and acts as $f_3$ on the rest of the set (in this case $PG(2,4)$). We also take $g'' = h'$ (from the proof of Theorem 5.3). Following the notation of Theorem 2.8, our $x$ in this case is $\infty'$ and $y$ is $\infty$.
To verify the conditions of Theorem 2.8, we see that (i) and (ii) are satisfied by how we have defined $g''$ and $h''$. To check (iii), we first consider $(h'')^2$, it is clear that $(h'')^2$ acts as the identity on $\infty, \infty'$ and $\infty''$, and for $[u,v,w] \in PG(2,4)$ we have:

$$(h'')^2[u,v,w] = [u^4, v^4, w^4] = [u,v,w]$$

so $(h'')^2$ is the identity, as required. We now consider $(g''h'')^3$, again, this acts as the identity for $\infty, \infty'$ and $\infty''$. For $[u,v,w] \in PG(2,4)$ we have:

$$\begin{aligned}(g''h'')[u,v,w] &= g''[u^2,v^2,w^2] = [u,v,wt], \\ (g''h'')^2[u,v,w] &= (g''h'')[u,v,wt] = [u,v,wt^2], \\ (g''h'')^3[u,v,w] &= [u,v,wt^3] = [u,v,w].\end{aligned}$$

Hence, both $(h'')^2$ and $(g''h'')^3$ act as the identity on every element in the new group, so (iii) is satisfied.
To verify (iv), let $p'' \in (M_{23})_{\infty'} = M_{22}$ (the stabiliser of $\infty'$ in $M_{23}$). We want to show that $h''p''h'' \in (M_{23})_{\infty'}$. Since $p'' \in M_{22}$, we can write $p'' = p_1p_2\cdots p_m$, where each $p_i$ is either equal to $p' : [u, v, w] \mapsto [au + dv + gw, bu + ev + hw, cu + fv + jw]$ (as in the proof of theorem 5.3) or $h = (\infty\ [1, 0, 0])f_1$ (as in the proof of Theorem 5.1). We then have, in the case that $p_i = h$:

$$\begin{aligned} h''hh''(\infty) &= h''h(\infty) = h''([1,0,0]) \\ &= [1,0,0], \\ h''hh''([1,0,0]) &= h''h([1,0,0]) = h''(\infty) \\ &= \infty, \end{aligned}$$

then, for $[u, v, w] \neq [1, 0, 0]$ we have:

$$\begin{aligned} h''hh''([u,v,w]) &= f_3f_1f_3([u,v,w]) = f_3f_1([u^2,v^2,w^2]) \\ &= f_3([u + v^2w^2, v, w]) \\ &= [u^2 + vw, v^2, w^2] = h([u,v,w]). \end{aligned}$$

Hence $h''hh'' = h$. For the other case ($p_i = p'$), we have:

$$\begin{aligned} h''p'h''([u,v,w]) &= h''p'([u^2,v^2,w^2]) \\ &= h''([au^2 + dv^2 + gw^2, bu^2 + ev^2 + hw^2, cu^2 + fv^2 + jw^2]) \\ &= [a^2u + d^2v + g^2w, b^2u + e^2v + h^2z, c^2u + f^2v + j^2w]. \end{aligned}$$

Which is associated to the matrix:

$$M = \begin{pmatrix} a^2 & d^2 & g^2 \\ b^2 & e^2 & h^2 \\ c^2 & f^2 & j^2 \end{pmatrix}.$$

The determinant of this matrix is then

$$\begin{aligned} \det M &= a^2(e^2j^2 - f^2h^2) - d^2(b^2j^2 - c^2h^2) + g^2b^2f^2 - c^2e^2) \\ &= (a(ej - hf) - d(bj - hc) + d(bf - ce))^2 \\ &= 1. \end{aligned}$$

So that $h''p'h'' \in PSL(3,4) \subseteq (M_{23})_{\infty'}$. Overall, this gives $h''p''h'' = h''p_1p_2\cdots p_mh'' = h''p_1h''h''p_2h''\cdots h''p_mh''$, where each $h''p_ih''$ either lies in $PSL(3,4)$ or equals $h$, so $h''p''h'' \in \langle PSL(3,4), h\rangle = (M_{23})_{\infty'}$. Hence condition (iv) is satisfied, and so all of the conditions in Theorem 2.8 are met, so $M_{24} = \langle M_{23}, h''\rangle$ is a one-point extension of $M_{23}$. □

Since $M_{23}$ is 4-transitive on 23 points, we have $M_{24}$ is a 5-transitive group acting on 24 points. The order of $M_{24} = 24 \cdot |M_{23}| = 244,823,040$, and we again can consider the design associated to $M_{24}$.

**Theorem 5.6.** *The $4-(23,7,1)$ design is extendable, giving a $5-(24,8,1)$ design on which $M_{24}$ acts as a group of automorphisms.*

*Proof.* We again proceed in the same manner as before, this time taking $\beta''$ to be the union of $\infty''$ and $\beta'$ (where $\beta'$ is taken as in the proof of Theorem 5.4). Let $(\mathcal{B}'')^+$ denote the set of blocks generated by the action of $M_{24}$ on $\beta''$, and label $(b'')^+ = |(\mathcal{B}'')^+|$.
The set-wise stabiliser of $\beta''$ in $M_{24}$ is transitive on the 8 points of $\beta''$, since it contains $h''$ (which switches $\infty''$ and $\infty'$) and the set-wise stabiliser of $\beta'$ in $M_{23}$ (which is transitive on the seven points of $\beta'$).
We can then calculate as follows, using the abbreviations $I = M_{23}, K = M_{24}$:

$$\begin{aligned}(b'')^+ &= |I : I_{(\beta'')}| \\ &= \frac{|I|}{|I_{(\beta'')}|} \\ &= \frac{(\frac{|I|}{|H|})(\frac{|K|}{|H|})(\frac{|H|}{|G|})(\frac{|G|}{|G_{(\ell)}|})}{(\frac{|I_{(\beta'')}|}{|K_{(\beta')}|})(\frac{|K_{(\beta')}|}{|H_{(\beta')}|})(\frac{|H_{(\beta')}|}{|G_{(\ell)}|})} \\ &= \frac{24\cdot 23\cdot 22\cdot 21}{6\cdot 7\cdot 8} \\ &= 759.\end{aligned}$$

From this, we can then calculate (using the same equations as before): $r^+ = 253, \lambda_2 = 77, \lambda_3 = 21, \lambda_4 = 5, \lambda = 1$. □

**Theorem 5.7** ([2, Theorem 3.6.3])**.** *The large Mathieu groups are simple.*

*Proof.* First looking at $M_{22}$, we have that the stabiliser of any point in $M_{22}$ is the simple group $PSL(3,4)$. Since $M_{22}$ is 3-transitive on 22 points, we have that it is also 2-transitive, and so Theorem 2.11 tells us that $M_{22}$ is primitive. We can then apply Theorem 2.19(ii) to see that $M_{22}$ can have no regular normal subgroup, and so Theorem 2.15 gives that $M_{22}$ is simple.
For $M_{23}$, we use the same argument, this time with the fact that the stabiliser of a point in $M_{23}$ is the simple group $M_{22}$. In this case we apply Theorem 2.19(iii) (since $M_{23}$ is 4-transitive) to see that $M_{23}$ has no regular normal subgroups, and then we again conclude from Theorem 2.15 that $M_{23}$ is simple.
Since $M_{24}$ is 5-transitive, Theorem 2.18 tells us that it has no regular normal subgroups, and hence (as above), $M_{24}$ is simple. □

Now that we have constructed the large Mathieu groups, a natural question to ask is whether we can extend $M_{24}$ to give another Mathieu group. We first need to establish some more notation and results.
Let $(G, X)$ be a permutation group with $U \leq G$, then we set [4, p.296]:

$$\mathcal{F}(U) = \{x \in X : gx = x\ \forall g \in U\}.$$

Using this notation we have the following lemma.

**Lemma 5.8** ([4, Lemma 9.65])**.** *For a permutation group $(G, X)$, with $U \leq G$,*

$$\mathcal{F}(gUg^{-1}) = g\mathcal{F}(U) \text{ for all } g \in G.$$

*Proof.* Let $x \in X$ and $g \in G$, if $x \in \mathcal{F}(gUg^{-1})$, then $gug^{-1}(x) = x$ for all $u \in U$, so $ug^{-1}(x) = g^{-1}(x)$, hence $(g^{-1}(x)) \in \mathcal{F}(U)$. It follows that $x \in g\mathcal{F}(U)$, so the result holds. □

**Theorem 5.9** ([4, Theorem 9.66])**.** *Let $(G, X)$ be $t$-transitive, with $t \geq 2$, $H$ be the stabiliser of $t$ points (say $x_1 \cdots, x_t$) in $X$ and $U$ be a Sylow $p$-subgroup of $H$ for some prime $p$, then:*

*(i) $N_G(U)$ acts $t$-transitively on $\mathcal{F}(U)$,*

*(ii) If $k = |\mathcal{F}(U)| > t$ and $U$ is a non-trivial normal subgroup of $H$, then $(X, \mathcal{B})$ is a Steiner system of type $S(t, k, v)$, where $|X| = v$ and*

$$\mathcal{B} = \{g\mathcal{F}(U) : g \in G\}.$$

*Proof.* (i) Firstly, note that if $g \in N_g(U)$, then $U = gUg^{-1}$, and $\mathcal{F}(U) = \mathcal{F}(gUg^{-1}) = g\mathcal{F}(U)$. Now $\{x_1, \cdots, x_t\} \subseteq \mathcal{F}(U)$ since $U \leq H$, which fixes $x_1, \cdots x_t$, hence $k = |\mathcal{F}(U)| \geq t$. If $y_1, \cdots, y_t$ are distinct elements of $\mathcal{F}(U)$, then since $G$ is $t$-transitive, we can find some $g \in G$ with $gy_i = x_i$ for all $i = 1, \cdots, t$. If $u \in U$, then $gug^{-1}x_i = guy_i = gy_i = x_i$, hence $gUg^{-1} \leq H$. We then have that there is some $h \in H$ which satisfies $gUg^{-1} = hUh^{-1}$, so that $h^{-1}g \in N_G(U)$ and $h^{-1}gx_i = h^{-1}x_i = x_i$ for all $i$.

(ii) The hypothesis gives $1 < t < k \leq v$. If $k = v$, then $\mathcal{F}(U) = X$, and since we always assume that $G$ acts faithfully on $X$, we have a contradiction, since $U \neq 1$, so $k \neq v$. We also have that for any $g \in G$, $k = |\mathcal{F}(U)| = |\mathcal{F}(gUg^{-1})|$. If $y_1, \cdots, y_t$ are distinct elements of $X$, then there is some $g \in G$ with $gx_i = y_i$ for all $i$, so $\{y_1, \cdots, y_t\} \subset g\mathcal{F}(U)$. It remains to show that $g\mathcal{F}(U)$ is the unique block containing $\{y_1, \cdots, y_t\}$. Suppose $\{y_1, \cdots, y_t\} \subset h\mathcal{F}(U)$, then there are $z_1, \cdots, z_t \in \mathcal{F}(U)$ such that $y_i = hz_i$ for each $i$. By (i), there is $\sigma \in N_G(U)$ with $z_i = \sigma x_i$, for all $i = 1, \cdots, t$, hence $gx_i = y_i = h\sigma x_i$. We then have that $g^{-1}h\sigma$ fixes $x_i$ for all $i$, and so $g^{-1}h\sigma \in H$. Then, since $U \lhd H$, $H \leq N_G(U)$, so $g^{-1}h\sigma \in N_G(U)$ and $g^{-1}h \in N_G(U)$. Hence, $gUg^{-1} = hUh^{-1}$, and so $g\mathcal{F}(U) = \mathcal{F}(gUg^{-1}) = \mathcal{F}(hUh^{-1}) = h\mathcal{F}(U)$ as required.

□

The final result we need in order to show that $M_{24}$ can not be extended is the following lemma taken from Rotman, we omit the proof, as it uses particular quotient groups not previously discussed in this dissertation. We first need the following definition.

**Definition 5.10** (Elementary abelian p-group [4, p.42])**.** *If $p$ is a prime, then an elementary abelian $p$ group is a finite group $G$ isomorphic to $\mathbb{Z}_p \times \cdots \times \mathbb{Z}_p$.*

**Lemma 5.11** ([4, Lemma 9.67]). *Let $H \leq M_{24}$ be the stabiliser of the five points:*

$$\infty, \infty', \infty'', [1,0,0] \text{ and } [0,1,0]$$

*(i) $H$ is a group of order 48 with a normal elementary abelian Sylow 2-subgroup $U$ of order 16.*

*(ii) $\mathcal{F}(U) = \ell \cup \{\infty, \infty', \infty''\}$, and so $|\mathcal{F}(U)| = 8$.*

*(iii) Only the identity of $M_{24}$ fixes more than 8 points.*

We can now give our final result, that $M_{24}$ does not have an extension.

**Theorem 5.12** ([4, Theorem 9.68]). *The Mathieu Group $M_{24}$ does not have a transitive extension.*

*Proof.* We proceed by contradiction. Suppose that there is a transitive extension of $M_{24}$, which we label $M_{25}$, and label the added element $\infty'''$. By construction, we would have $M_{24} \subset M_{25}$, and hence $M_{25}$ also contains the set-wise stabiliser of $\beta'' = \ell \cup \{\infty\} \cup \{\infty'\} \cup \{\infty''\}$. When considered as a subgroup of $M_{25}$, this stabiliser must also fix $\infty'''$, so $G_1 = (M_{24})_{\beta''}$ stabilises $\beta''' = \beta'' \cup \{\infty'''\}$, and hence we have $G_1 \subseteq G_2 = (M_{25})_{\beta'''}$.
We now have two cases to consider. For the first case, suppose that $G_2$ contains an element which does not fix $\infty'''$. Here, we have that $\infty'''$ must be sent to some other element of $\beta'''$ (which must also be an element of $\beta''$), and since we know that elements of $G_1$ (contained in $G_2$) can map anything in $\beta''$ to any other element in $\beta''$, $G_2$ is transitive on $\beta'''$. Then, let $\mathcal{B}'''$ be the set of blocks generated by the action of $M_{25}$ on $\beta'''$, following the same process as in the proof of Theorem 5.6, and letting $M_{25} = J$ we then have:

$$\begin{aligned} |\mathcal{B}'''| &= \frac{|J|}{|J_{(\beta''')}|} \\ &= \frac{(\frac{|J|}{|I|})(\frac{|I|}{|K|})(\frac{|K|}{|H|})(\frac{|H|}{|G|})(\frac{|G|}{|G_{(\ell)}|})}{(\frac{|J_{(\beta''')}|}{|I_{(\beta'')}|})(\frac{|I_{(\beta'')}|}{|K_{(\beta')}|})(\frac{|K_{(\beta')}|}{|H_{(\beta)}|})(\frac{|H_{(\beta)}|}{|G_{(\ell)}|})} \\ &= \frac{25 \cdot 24 \cdot 23 \cdot 22 \cdot 21}{9 \cdot 8 \cdot 7 \cdot 6} \\ &= \frac{6325}{3} \notin \mathbb{Z}. \end{aligned}$$

Hence $M_{25}$ does not exist.
For the second case, suppose that every element of $G_2$ fixes $\infty'''$, so that $G_2 = G_1$. We can then apply Lemma 5.11, noting that $G_1 = H$, to see that $G_2 = G_1$ has a normal subgroup (which stabilises $\beta''' = \ell \cup \{\infty, \infty', \infty'', \infty'''\}$) with order $16 = 2^4$. Hence we can apply Theorem 5.9, taking $X = PG(2,4) \cup \{\infty, \infty', \infty'', \infty'''\}, G = M_{25}, H = G_2, t = 6$ and $U$ to be the Sylow 2-subgroup of $G_2$, which is the pointwise stabiliser of $\beta'''$. By (ii) of the theorem we can construct a $S(t,k,v)$ Steiner system, with $v = |X| = 25$, $t = 6$, and $k = |\mathcal{F}(U)| = |\beta'''| = 9$. As we have already seen in the first case, no such design exists, and therefore there is no extension of $M_{24}$. □

# References

[1] Robert A. Wilson. *The Finite Simple Groups.* Springer, 2009.

[2] N.L Biggs and A.T White. *Permutation Groups and Combinatorial Structures.* Cambridge University Press, 2nd edition, 1996.

[3] Daniel M. Elton. Math 321: Groups and symmetry, 2022.

[4] Joseph J. Rotman. *An Introduction to the Theory of Groups.* Springer-Verlag, 4th edition, 1995.

[5] Hans Kurzweil and Bernd Stellmacher. *The Theory of Finite Groups: An Introduction.* Springer-Verlag, New York, 1st edition, 2006.

[6] Nadia Mazza. Math 322: Commutative algebra, 2022.

[7] Bernd Schulze. Math 327: Combinatorics, 2022.