Summer 2011

Construction and Simplicity of the Large Mathieu Groups

Robert Peter Hansen
San Jose State University

CONSTRUCTION AND SIMPLICITY OF THE LARGE MATHIEU GROUPS

A Thesis

Presented to

The Faculty of the Department of Mathematics

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

R. Peter Hansen

August 2011

The Designated Thesis Committee Approves the Thesis Titled

CONSTRUCTION AND SIMPLICITY OF THE LARGE MATHIEU GROUPS

by

R. Peter Hansen

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

SAN JOSÉ STATE UNIVERSITY

August 2011

Dr. Timothy Hsu	Department of Mathematics
Dr. Roger Alperin	Department of Mathematics
Dr. Brian Peterson	Department of Mathematics

ABSTRACT

CONSTRUCTION AND SIMPLICITY OF THE LARGE MATHIEU GROUPS

by R. Peter Hansen

In this thesis, we describe the construction of the Mathieu group M_{24} given by Ernst Witt in 1938, a construction whose geometry was examined by Jacques Tits in 1964. This construction is achieved by extending the projective semilinear group $P\Gamma L_3(\mathbb{F}_4)$ and its action on the projective plane $P^2(\mathbb{F}_4)$. $P^2(\mathbb{F}_4)$ is the projective plane over the field of 4 elements, with 21 points and 21 lines, and $P\Gamma L_3(\mathbb{F}_4)$ is the largest group sending lines to lines in $P^2(\mathbb{F}_4)$. This plane has 168 6-point subsets, hexads, with the property that no 3 points of a hexad are collinear. Under the action of the subgroup $PSL_3(\mathbb{F}_4)$, the hexads in $P^2(\mathbb{F}_4)$ break into 3 orbits of equal size. These orbits are preserved and permuted by $P\Gamma L_3(\mathbb{F}_4)$, and can be viewed as 3 points, which, when added to the 21 points of $P^2(\mathbb{F}_4)$, yield a set X of 24 points. Using lines and hexads in $P^2(\mathbb{F}_4)$, we define certain 8-point subsets of X, view them as vectors in $\mathbb{F}_2^{24}$, and define the subspace they span as the Golay 24-code. We then define M_{24} as the automorphism group of the Golay 24-code and show that it acts 5-transitively on X, establishing its simplicity. We calculate the order of M_{24} and the order of two simple subgroups, M_{23} and M_{22}, the other large Mathieu groups.

ACKNOWLEDGEMENTS

I want to thank my wife, Claire. Without her constant encouragement, I could not have finished a master's degree in mathematics. She has my enduring gratitude. I would like to thank my thesis advisor, Tim Hsu. I would have been quickly lost in navigating the "Witt-Tits" construction had it not been for his guidance. I also would like to thank Brian Peterson and Roger Alperin for their courses in abstract algebra, where my interest in group theory was first sparked.

TABLE OF CONTENTS

CHAPTER

CHAPTER 1

INTRODUCTION

A simple group is one whose normal subgroups are trivial. Normal subgroups allow the definition of factor groups, a decomposition of a group. Having no such decomposition, simple groups are "atoms of symmetry" (Ronan [Ron06]). It was Evariste Galois who first noted the importance of normal subgroups in 1832, but his description remained unpublished until 1846 (Galois [Gal46]). In 1870, Camille Jordan, along with Otto Holder, determined each group has a unique decomposition as a sequence of factor groups, a "composition series" that is unique up to order and isomorphism (Jordan [Jor70]).

Groups were initially conceived of as collections of permutations possessing an intrinsic coherence. The elements permuted were variously called "symbols," "letters," and "points." It was noted that some groups could permute subsets of a given size to any other subset of that size, a feature known as transitivity. If k points can be moved to any other k points, that group is k-transitive. In 1861 and 1873, a French mathematician, Emile Mathieu, published papers that described two 5-transitive groups, one acting on 12 points, and the other on 24 points, neither of which were alternating (Mathieu [Mat61], [Mat73]). The one acting on 24 points, M_{24}, is the topic of this thesis. These groups, along with three others that are subgroups of the other two, were the first "sporadic" simple groups to be discovered. The next sporadic group was not discovered for another hundred years, in 1965 (Janko [Jan66]). Its discovery, and the proof of the odd order theorem in 1962 (Feit [FT63]), initiated a great deal of work in finite group theory, culminating with the

construction of the largest sporadic group, the Monster, in 1982 (Griess [Gri82]).

The classification theorem for finite simple groups, mostly finished by 1983, but under revision even today, completely classifies these groups (Aschbacher [Asc04]). There are 18 countably infinite families, as well as 26 groups that do not fit into any such family (Mazurov [Maz]). These 26 are the sporadic groups, and although not part of an infinite family, they are interrelated, as all but six are subgroups or subquotients of the Monster [Ron06]. The 18 infinite families include the cyclic groups of prime order, and the alternating groups acting on 5 or more points. The next 16 families are the groups of Lie type, including the projective linear groups. One of these, $PSL_3(\mathbb{F}_4)$, is used in this thesis. It is a subgroup of M_{24}, and plays a key role in proving its simplicity. This projective group is a natural subgroup of the projective semilinear group, $P\Gamma L_3(\mathbb{F}_4)$, and this larger group will be used to construct M_{24}.

There are several constructions of M_{24}, and in fact, Mathieu did not fully convince the mathematical community he had constructed two 5-transitive groups. It was not until 1935 that Ernst Witt gave a definitive construction of M_{24} (Witt [Wit38a]). In 1964, Jacques Tits published a paper that explored the geometry of Witt's construction (Tits [Tit64]). Conway and Sloane refer to this construction as the "Witt-Tits" construction, and it is the one explained in this thesis (Conway and Sloane [CS93]). Witt describes a certain combinatorial structure: 759 blocks, each an 8-element subset of 24 points X, possessing the property that any 5-element subset of X lies in exactly one block. Witt showed M_{24} could be realized as the automorphism group of this "Steiner system" (Witt [Wit38b]).

This Steiner system has a fascinating connection with modern communications. Each of the 759 blocks can be viewed as a 24-tuple over the field of two elements, and the subspace generated by these vectors is known as the Golay

24-code (Pless [Ple89]). It is a 12-dimensional subspace which allows 12 bits of "information" to be transmitted, along with 12 bits of "redundancy." These codewords have the property that, upon reception, any 3-bit corruption can be detected and corrected. The Mathieu group M_{24} can be defined as the automorphism group of the Golay 24-code.

The main work of this thesis is in showing how to define 24 points to obtain the desired Golay code and Steiner system (and from there to M_{24}). The projective plane over the field of four elements, $P^2(\mathbb{F}_4)$, has 21 points. Its "collineation group" is the projective semilinear group, $P\Gamma L_3(\mathbb{F}_4)$, the largest group sending lines to lines in $P^2(\mathbb{F}_4)$. This projective plane has 6-element subsets, hexads, with the property that no three points in a hexad are collinear. Under the action of $PSL_3(\mathbb{F}_4)$, there are three equally-sized orbits of hexads in $P^2(\mathbb{F}_4)$. These orbits are preserved (and permuted) by $P\Gamma L_3(\mathbb{F}_4)$. These three orbits give our remaining points, for 24 points X. Using lines and hexads in $P^2(\mathbb{F}_4)$, we define 8-element subsets of X, viewed as vectors in $\mathbb{F}_2^{24}$. The Golay code is the subspace generated by these vectors, and from there, we determine its automorphism group, M_{24}.

We now briefly outline this thesis. Chapter 2 contains lemmas used in the following three chapters. Chapter 3 examines group actions, develops the theory of stabilizers and orbits, and establishes a simplicity criterion based on multiple transitivity. Chapter 4 looks at linear groups (considered mainly as matrices), examining the special subgroup of matrices of determinant 1, and the factor groups which are these groups' "projective" versions. A major result of this chapter is the simplicity of $PSL_3(\mathbb{F}_4)$. Chapter 5 extends linear groups by considering vector spaces over fields with nontrivial automorphisms. Linear transformations, in conjunction with applying a nontrivial field automorphism to coordinates, preserve vector addition but not scalar multiplication. They are thus "semi-linear."

Chapters 6 and 7 consider bilinear forms and quadratic forms over fields of characteristic 2. These chapters prepare for Chapter 8, where lines and hexads in $P^2(\mathbb{F}_4)$ are considered as the zero sets of homogeneous polynomials. This approach has limited success in describing hexads, however, and thus Chapter 9 takes a finite geometry approach to yield the desired results. Chapter 10 on binary linear codes prepares for the characterization of the Golay code, and introduces the idea of a code's automorphism group.

Chapter 11 defines the Golay code and M_{24}. The Mathieu groups M_{23}, M_{22} and M_{21} are defined as the pointwise stabilizers of 1, 2 and 3 points, respectively, and M_{21} is shown to be isomorphic to $PSL_3(\mathbb{F}_4)$. This allows our simplicity criterion to be invoked, establishing the simplicity of M_{22}, M_{23}, and M_{24}.

CHAPTER 2

PREPARATORY LEMMAS

In this chapter, we prove the lemmas that will be used in Chapters 3, 4, and 5.

2.1 Normal and Characteristic Subgroups

In this section, we look at the action of conjugation on group elements and subgroups, and the conjugacy class of an element. We examine characteristic subgroups, and in particular, center and commutator subgroups. The material on these subgroups will be used in Chapter 4 on linear groups. We follow the treatment of Rotman [Rot95].

Definition 2.1.1. For $a, b \in G$, if $a = g^{-1}bg$ for some $g \in G$, then a and b are said to be *conjugate*. For $H \leq G$, we define $g^{-1}Hg = \{g^{-1}hg \mid h \in H\}$.

Lemma 2.1.2. *Conjugacy is an equivalence relation.*

Proof. Now $1 \in G$ and $a = 1^{-1}a1$ for all $a \in G$, and so $a \sim a$.

Let $a \sim b$. Thus $a = g^{-1}bg$ for some $g \in G$. Since $g^{-1} \in G$, we have

$$(g^{-1})^{-1}ag^{-1} = g(g^{-1}bg)g^{-1} = (gg^{-1})b(gg^{-1}) = b. \tag{2.1}$$

Thus $b \sim a$.

Let $a \sim b$ and $b \sim c$. Thus there are $g_1, g_2 \in G$ such that $a = g_1^{-1}bg_1$ and $b = g_2^{-1}cg_2$. This gives us

$$a = g_1^{-1}bg_1 = a = g_1^{-1}(g_2^{-1}cg_2)g_1 = (g_1^{-1}g_2^{-1})c(g_2g_1) = (g_2g_1)^{-1}c(g_2g_1). \tag{2.2}$$

Thus $a \sim c$. □

Definition 2.1.3. The *conjugacy class* of $a \in G$, denoted a^G, is the set $\{g^{-1}ag \mid g \in G\}$.

Lemma 2.1.4. *If $H \leq G$, then $g^{-1}Hg \cong H$, for all $g \in G$.*

Proof. Now $1 \in H$ and $g^{-1}1g = g^{-1}g = 1$ for all $g \in G$. Thus $1 \in g^{-1}Hg$. Let $a, b \in g^{-1}Hg$. Thus there exist $a_1, b_1 \in H$ such that $a = g^{-1}a_1g$ and $b = g^{-1}b_1g$. And we have

$$ab = (g^{-1}a_1g)(g^{-1}b_1g) = g^{-1}a_1(gg^{-1})b_1g = g^{-1}a_1b_1g. \tag{2.3}$$

Since $a_1b_1 \in H$, we have $ab = g^{-1}a_1b_1g \in g^{-1}Hg$. Finally, for $a \in g^{-1}Hg$ such that $a = g^{-1}a_1g$ with $a_1 \in H$, we have $a_1^{-1} \in H$ and thus $g^{-1}a_1^{-1}g \in g^{-1}Hg$. And this gives

$$(g^{-1}a_1g)(g^{-1}a_1^{-1}g) = g^{-1}a_1(gg^{-1})a_1^{-1}g = g^{-1}a_1a_1^{-1}g = g^{-1}1g = 1. \tag{2.4}$$

And similarly, $(g^{-1}a_1^{-1}g)(g^{-1}a_1g) = 1$. Thus $a^{-1} = g^{-1}a_1^{-1}g \in g^{-1}Hg$. Therefore, $g^{-1}Hg$ is a subgroup of G.

Now let $g \in G$ and define $\varphi : H \to g^{-1}Hg$ such that $\varphi(h) = g^{-1}hg$. Assume that $\varphi(h_1) = \varphi(h_2)$. Thus $g^{-1}h_1g = g^{-1}h_2g$, and by cancellation with g and g^{-1}, we see that $h_1 = h_2$. Now let $h \in g^{-1}Hg$. Thus there is an $h_1 \in H$ such that $h = g^{-1}h_1g$ and $\varphi(h_1) = h$. Finally, let $h_1, h_2 \in H$. We have

$$\varphi(h_1)\varphi(h_2) = (g^{-1}h_1g)(g^{-1}h_2g) = g^{-1}h_1(gg^{-1})h_2g = g^{-1}h_1h_2g = \varphi(h_1h_2). \tag{2.5}$$

Thus φ is a bijective homomorphism. □

Definition 2.1.5. If $H \leq G$, then H is a *normal* subgroup if $g^{-1}Hg = H$ for all $g \in G$. We denote this $H \lhd G$. Equivalent definitions of normality include $Hg = gH$ for all $g \in G$, and $g^{-1}Hg \leq H$ for all $g \in G$.

Lemma 2.1.6. *Let H be a subgroup of a group G. Then $N = \bigcap_{g\in G} g^{-1}Hg$ is the largest subgroup of H that is normal in G.*

Proof. So $\bigcap_{g\in G} g^{-1}Hg \leq H$ since $1^{-1}H1 = H$.

Now let $g' \in G$. So we have

$$\begin{aligned} g'^{-1}\left(\bigcap_{g\in G} g^{-1}Hg\right)g' &= \bigcap_{g\in G} g'^{-1}(g^{-1}Hg)g' = \bigcap_{g\in G}(g'^{-1}g^{-1})H(gg') \\ &= \bigcap_{g\in G}(gg')^{-1}H(gg') = \bigcap_{g\in G} g^{-1}Hg, \end{aligned} \tag{2.6}$$

since $Gg' = G$. Thus N is a normal subgroup.

Now let $N' \leq H$ be a normal subgroup. Thus

$$N' = g^{-1}N'g \leq g^{-1}Hg \tag{2.7}$$

for all $g \in G$, and so

$$N' \leq \bigcap_{g\in G} g^{-1}Hg. \tag{2.8}$$

Since any normal subgroup of H is contained in N, it is the largest normal subgroup in H. □

Definition 2.1.7. An *automorphism* of a group G is an isomorphism $\alpha : G \to G$. A subgroup H of G is called *characteristic*, and is denoted $H \operatorname{char} G$, if $\alpha(H) = H$ for every automorphism α of G.

Lemma 2.1.8. *If H is a characteristic subgroup of G, then H is normal.*

Proof. By Lemma 2.1.4, we have $g^{-1}Gg \cong G$ for all $g \in G$, making conjugation of G by $g \in G$ an automorphism of G. Since $H \operatorname{char} G$, H is invariant under automorphisms of G. Thus $g^{-1}Hg = H$ for all $g \in G$, and so $H \lhd G$. □

Definition 2.1.9. The *center* of a group G, denoted $Z(G)$, is the set $\{z \in G \mid zg = gz, g \in G\}$.

Lemma 2.1.10. *The center of a group is a characteristic subgroup.*

Proof. Since $1g = g = g1$ for all $g \in G$, $1 \in Z(G)$. Let $a, b \in Z(G)$ and let $g \in G$. We have

$$(ab)g = a(bg) = a(gb) = (ag)b = (ga)b = g(ab). \tag{2.9}$$

Thus $ab \in Z(G)$. Now let $a \in Z(G)$ and $g \in G$. We have

$$a^{-1}g = a^{-1}(g^{-1})^{-1} = (g^{-1}a)^{-1} = (ag^{-1})^{-1} = (g^{-1})^{-1}a^{-1} = ga^{-1}. \tag{2.10}$$

Thus $a^{-1} \in Z(G)$. Now let α be an automorphism of G. A typical element of $\alpha(G)$ has the form $\alpha(g)$, and so for $c \in Z(G)$, we have

$$\alpha(c)\alpha(g) = \alpha(cg) = \alpha(gc) = \alpha(g)\alpha(c). \tag{2.11}$$

Thus $\alpha(c) \in Z(\alpha(G)) = Z(G)$. Thus $\alpha(Z(G)) \leq Z(G)$. Now α^{-1} is also an automorphism of G and so by a similar argument, $\alpha^{-1}(Z(G)) \leq Z(G)$. Thus

$$Z(G) = (\alpha\alpha^{-1})(Z(G)) = \alpha(\alpha^{-1}(Z(G))) \leq \alpha(Z(G)) \tag{2.12}$$

Since α was arbitrary, $\alpha(Z(G)) = Z(G)$ for all automorphisms α of G. □

Definition 2.1.11. The *commutator subgroup* of a group G is the subgroup $G' = \langle a^{-1}b^{-1}ab \mid a, b \in G \rangle$.

Lemma 2.1.12. *The commutator subgroup G' of G is a characteristic subgroup of G. Further, if N is a normal subgroup of G, then G/N is abelian if and only if $G' \leq N$.*

Proof. Let α be an automorphism of G. For $x \in G'$,

$$x = \prod_{i=1}^{k} a_i^{-1}b_i^{-1}a_ib_i \tag{2.13}$$

where $a_i, b_i \in G$ and the $a_i b_i$ are not all necessarily distinct. Thus

$$\begin{aligned} \alpha(x) &= \alpha\left(\prod_{i=1}^{k} a_i^{-1} b_i^{-1} a_i b_i\right) = \prod_{i=1}^{k} \alpha(a_i^{-1} b_i^{-1} a_i b_i) \\ &= \prod_{i=1}^{k} \alpha(a_i^{-1})\alpha(b_i^{-1})\alpha(a_i)\alpha(b_i) \\ &= \prod_{i=1}^{k} \alpha(a_i)^{-1}\alpha(b_i)^{-1}\alpha(a_i)\alpha(b_i) \in G'. \end{aligned} \tag{2.14}$$

So $\alpha(G') \leq G'$. By a similar argument, $\alpha^{-1}(G') \leq G'$. Thus

$$G' = (\alpha\alpha^{-1})(G') = \alpha(\alpha^{-1}(G')) \leq \alpha(G'). \tag{2.15}$$

Thus $\alpha(G') = G'$ for all automorphisms α of G.

Now let $G' \leq N \lhd G$, and let $aN, bN \in G/N$. We have $b^{-1}a^{-1}ba \in G' \leq N$. Thus

$$(aN)(bN) = abN = ab(b^{-1}a^{-1}ba)N = (abb^{-1}a^{-1})(ba)N = baN = (bN)(aN) \tag{2.16}$$

and so G/N is abelian. Conversely, let G/N be abelian. So for all $aN, bN \in G/N$, $abN = (aN)(bN) = (bN)(aN) = baN$, and therefore

$$a^{-1}b^{-1}abN = N. \tag{2.17}$$

Thus $a^{-1}b^{-1}ab \in N$ for all $a, b \in G$. Since all generators of G' are contained in N, we have $G' \leq N$. □

Corollary 2.1.13. *The factor group G/G' is abelian, and if $\varphi : G \to A$ is a homomorphism such that A is an abelian group, then $G' \leq \ker\varphi$.*

Proof. For the first statement, let $N = G'$. For the second one, let $\varphi : G \to A$ be a homomorphism. By the canonical isomorphism, we have $G/\ker\varphi \cong \varphi(G) \leq A$. Since A is abelian, so is $\varphi(G)$, implying $G/\ker\varphi$ is as well, and so $G' \leq \ker\varphi$. □

Corollary 2.1.14. *If G is a non-abelian simple group, then $G' = G$.*

Proof. Let $a, b \in G$ such that $ab \neq ba$. Thus $a^{-1}b^{-1}ab \neq 1$, and $G' \neq 1$. Since G is simple, it lacks nontrivial proper subgroups. Thus $G' = G$. □

Definition 2.1.15. If $f : B \to C$ is a mapping and $A \subseteq B$, then the *restriction of f to A*, denoted $f|_A$, is $f|_A : A \to C$ such that $f|_A(a) = f(a)$ for all $a \in A$.

Lemma 2.1.16. *If K char H and $H \lhd G$, then $K \lhd G$.*

Proof. Let $a \in G$ and let $\alpha : G \to G$ be conjugation by a. Since $H \lhd G$, we have $a^{-1}Ha = H$. Thus $\alpha|_H(H) = H$, and so $\alpha|_H$ is an automorphism of H. Since K char H, $\alpha|_H(K) = K$. Thus if $k \in K$, $a^{-1}ka = \alpha(k) \in K$. □

2.2 Cosets and Products of Groups

In this section, we prove results on products of subgroups, especially where at least one subgroup is normal. We prove that $N \lhd G$ is maximal if and only if G/N is simple, for use in Section 4.3. We follow the treatment of Fraleigh [Fra03].

Lemma 2.2.1. *Let G be a group and K, H subgroups of G such that $K \leq H \leq G$. If we fix $g \in G$, then either $Kg \cap H = \varnothing$ or $Kg \subseteq H$. Thus H is a union of cosets of K (relative to $g \in G$).*

Proof. If $Kg \cap H = \varnothing$, we are done. So assume $Kg \cap H \neq \varnothing$. So there exists $h \in Kg \cap H$ with $h = k_1 g$ for some $k_1 \in K$. For $kg \in Kg$, $kk_1^{-1} \in K \leq H$, so we have

$$kg = k(k_1^{-1}k_1)g = (kk_1^{-1})(k_1 g) = (kk_1^{-1})h \in H \tag{2.18}$$

by closure. Therefore $Kg \subseteq H$. The second statement easily follows. □

Definition 2.2.2. If H and K are subgroups of a group G, define the *product of H and K* as $HK = \{hk \mid h \in H, k \in K\}$.

Lemma 2.2.3. *For a group G, if $K \leq G$ and $H \lhd G$, then $K \leq HK = KH \leq G$.*

Proof. Since $k = 1k$ for $k \in K$, $K \subseteq HK$. Now let $a = hk \in HK$. Thus $hk \in Hk$. Since H is normal in G, $Hk = kH$, and thus there exists $h' \in H$ such that $hk = kh'$. Thus $a \in KH$, and $HK \subseteq KH$. Similarly, $KH \subseteq HK$. Thus $HK = KH$.

Now let $a, b \in HK$. Thus $a = h_1k_1$ and $b = h_2k_2$ for $h_1, h_2 \in H$ and $k_1, k_2 \in K$. Now $k_1h_2 \in KH = HK$, and so there exists $h_3 \in H$ such that $k_1h_2 = h_3k_1$. Therefore

$$ab = (h_1k_1)(h_2k_2) = h_1(k_1h_2)k_2 = h_1(h_3k_1)k_2 = (h_1h_3)(k_1k_2) \in HK, \tag{2.19}$$

and thus HK is closed. Since $1 \in H \cap K$, we have $1 \cdot 1 = 1 \in HK$. Finally, let $a = h_1k_1 \in HK$. So $h_1^{-1} \in H$ and $k_1^{-1} \in K$. So $k_1^{-1}h_1^{-1} \in KH = HK$, and thus $a^{-1} \in HK$. □

Lemma 2.2.4. *If $H, K \lhd G$, then $HK \lhd G$.*

Proof. Let $H, K \lhd G$. By Lemma 2.2.3, $HK \leq G$. So for $g \in G$ we have

$$g^{-1}HKg \leq (g^{-1}Hg)(g^{-1}Kg) = HK. \tag{2.20}$$

Therefore, $HK \lhd G$. □

Lemma 2.2.5. *If $N \lhd G$ and $K \lhd H \leq G$, then $NK \lhd NH \leq G$.*

Proof. We must show $g^{-1}NKg = NK$ for all $g \in NH$. Now $NH = HN$ and $NK = KN$ by Lemma 2.2.3. Let $g \in HN$. Thus $g = hn$ for some $h \in H$ and $n \in N$, and we have

$$\begin{aligned} g^{-1}NKg = (n^{-1}h^{-1})KN(hn) &\leq (n^{-1}h^{-1})K(hn)N \\ &= n^{-1}(h^{-1}Kh)nN = n^{-1}KN = n^{-1}NK = NK. \end{aligned} \tag{2.21}$$

□

Lemma 2.2.6. *Let $\varphi : G_1 \to G_2$ be a homomorphism. If $N_1 \lhd G_1$, then $\varphi(N_1) \lhd \varphi(G_1)$. If $N_2 \lhd \varphi(G_1)$, then $\varphi^{-1}(N_2) \lhd G_1$.*

Proof. Let $n_2 \in \varphi(N_1)$ and $g_2 \in \varphi(G_1)$. Thus there exists $n_1 \in N_1$ such that $\varphi(n_1) = n_2$, and $g_1 \in G_1$ such that $\varphi(g_1) = g_2$. Now $\varphi(g_1^{-1}) = \varphi(g_1)^{-1} = g_2^{-1}$. We have

$$g_2^{-1} n_2 g_2 = \varphi(g_1^{-1})\varphi(n_1)\varphi(g_1) = \varphi(g_1^{-1} n_1 g_1) = \varphi(n_1') \tag{2.22}$$

for some $n_1' \in N_1$, because $N_1 \lhd G_1$. Thus $\varphi(N_1) \lhd \varphi(G_1)$.

Now $\varphi^{-1}(N_2)$ is the pre-image of N_2 in G_1. Let $a_1, b_1 \in \varphi^{-1}(N_2)$. We have

$$\varphi(a_1 b_1^{-1}) = \varphi(a_1)\varphi(b_1^{-1}) = \varphi(a_1)\varphi(b_1)^{-1} \in N_2 \tag{2.23}$$

and thus $a_1 b_1^{-1} \in \varphi^{-1}(N_2)$, meaning $\varphi^{-1}(N_2) \leq G_1$. We let $g_1 \in G_1$ and $n_1 \in \varphi^{-1}(N_2)$. Thus $\varphi(n_1) = n_2 \in N_2$. We have

$$\varphi(g_1^{-1} n_1 g_1) = \varphi(g_1^{-1})\varphi(n_1)\varphi(g_1) = \varphi(g_1)^{-1} n_2 \varphi(g_1) = n_2' \in N_2. \tag{2.24}$$

Thus $g_1^{-1} n_1 g_1 \in \varphi^{-1}(N_2)$, meaning $\varphi^{-1}(N_2) \lhd G$. □

Lemma 2.2.7. *N is a maximal normal subgroup of G if and only if G/N is simple.*

Proof. Let N be a maximal normal subgroup of G. Assume G/N has a normal subgroup K/N such that $N/N < K/N < G/N$. Define the map

$$\varphi : G \to G/N \qquad \text{such that} \qquad \varphi(g) = gN. \tag{2.25}$$

We have $\varphi^{-1}(N/N) < \varphi^{-1}(K/N) < \varphi^{-1}(G/N)$ yielding

$$N < K < G \tag{2.26}$$

where $K \lhd G$ by Lemma 2.2.6. This contradiction of the maximality of N shows G/N is simple. Now let G/N be simple, and assume N is not a maximal normal

subgroup. Thus there is a $K \lhd G$ such that $N < K < G$. Under φ, we have $N/N < K/N < G/N$, and so $K/N > 1$. By Lemma 2.2.6,

$$K/N = \varphi(K) \lhd \varphi(G) = G/N, \tag{2.27}$$

contradicting the simplicity of G/N. Thus N is a maximal normal subgroup. □

2.3 p-Groups

In this section, we define p-groups, state Cauchy's theorem, and look at elementary abelian p-groups. We use these results in Chapter 3, in the study of group actions. We follow the treatment of Fraleigh [Fra03] and Rotman [Rot95].

Definition 2.3.1. For $g \in G$, let $\langle g \rangle$ denote the *cyclic subgroup generated by* g, and define the *order* of g, written $|g|$, as $|g| = |\langle g \rangle|$.

Definition 2.3.2. If G is a finite group, then the *exponent of* G, denoted $\exp(G)$, is the smallest positive integer n such that $g^n = 1$ for all $g \in G$.

Lemma 2.3.3. *Let G be a finite group. Then for every $g \in G$, $|g|$ divides $\exp G$. Also, $\exp G \leq |G|$ and $\exp G$ divides $|G|$.*

Proof. Let $\exp(G) = n$ for some $n \in \mathbb{Z}^+$, and let $g \in G$. Thus $|g| = k$ for some $0 < k \leq n$, and we have $g^k = 1 = g^n$. By the division algorithm, there exist $q, r \in \mathbb{Z}$ such that $n = qk + r$, where $0 \leq r < k$. Thus

$$g^r = g^{n-qk} = g^n g^{-qk} = g^n (g^k)^{-q} = (1)(1)^{-q} = 1. \tag{2.28}$$

However, since $|g| = k$, we must have $r = 0$. Thus k is a factor of n.

Now for $g \in G$, $|g|$ divides $|G|$ by the theorem of Lagrange. Say that $|g|\ell = |G|$ for some $\ell \in \mathbb{Z}^+$. This gives

$$g^{|G|} = g^{|g|\ell} = (g^{|g|})^\ell = 1. \tag{2.29}$$

This holds true for all $g \in G$, and since $\exp G$ is the smallest exponent sending every element to the identity, $\exp G \leq |G|$. By use of the division algorithm, as in the first part of the proof, we have $\exp G$ dividing $|G|$. □

Theorem 2.3.4. *(Cauchy's Theorem) Let p be a prime. Let G be a finite group and let p divide $|G|$. Then G has an element of order p and, consequently, a subgroup of order p.* □ *[Fra03]*

Lemma 2.3.5. *Let G be a finite group and p be a prime number. Also, let $0 \leq k \leq n$, where k and n are integers. Then the following are equivalent:*

(1) $|G| = p^n$*;*

(2) Every element of G has order a power of p;

(3) $\exp(G) = p^k$*.*

Proof. Assume $|G| = p^n$ for some $n \in \mathbb{Z}^+$, and let $g \in G$. Now $\langle g \rangle \leq G$, and so $|g|$ divides $|G|$, by the theorem of Lagrange. Thus $|g| = p^\ell$ for some $0 \leq \ell \leq n$.

Now assume every element of G has order a power of p. Let p^k be the largest such prime power. Thus for an arbitrary $g \in G$, $|g| = p^\ell$, where $0 \leq \ell \leq k$. Thus

$$g^{p^k} = (g^{p^\ell})^{p^{k-\ell}} = (1)^{p^{k-\ell}} = 1, \tag{2.30}$$

where $k - \ell \geq 0$. Since p^k is the largest such prime power, there exists a $g' \in G$ such that $|g'| = p^k$. Thus no smaller exponent will yield the identity for each element of G, and so $\exp(G) = p^k$.

Now assume $\exp(G) = p^k$. By Lemma 2.3.3, we have p^k dividing $|G|$. Let q be a prime divisor of $|G|$. By Cauchy's theorem, there is a $g \in G$ of order q. By Lemma 2.3.3, q divides p^k. Thus $q = p$, and $|G| = p^n$, where $k \leq n$. □

Definition 2.3.6. Let p be a prime. A group G is a *p-group* if G meets one of the equivalent conditions of Theorem 2.3.5.

Definition 2.3.7. An *elementary abelian p-group* is a finite group G isomorphic to $\mathbb{Z}_p \times \cdots \times \mathbb{Z}_p$, that is, ${\mathbb{Z}_p}^n$ for some positive integer n.

Lemma 2.3.8. *A finite abelian p-group G is elementary if and only if it has exponent p.*

Proof. If a finite abelian p-group G is elementary, then

$$G = \mathbb{Z}_p \times \cdots \times \mathbb{Z}_p = (\mathbb{Z}/p\mathbb{Z})^n \tag{2.31}$$

for $n \in \mathbb{Z}^+$. Let $g \in G$. Thus $g = (g_1, \ldots, g_n)$ where $g_i \in \mathbb{Z}_p$ for $1 \leq i \leq n$. Since $|g_i| = 1$ or p, we have

$$|g| = \operatorname{lcm}(|g_1|, \ldots, |g_n|) = 1 \text{ or } p. \tag{2.32}$$

Further, $|g| = 1$ if and only if $g_i = 0$ for all i. Since there exists a $g \neq 0$ in G, $\exp(G) = p$.

Conversely, let a finite abelian p-group G have exponent p. Since G is finite, it is finitely-generated, and by the fundamental theorem of finitely generated abelian groups,

$$G \cong \mathbb{Z}_{{p_1}^{r_1}} \times \cdots \times \mathbb{Z}_{{p_n}^{r_n}} \times \mathbb{Z} \times \cdots \times \mathbb{Z}. \tag{2.33}$$

Since G is finite, it has no copies of $\mathbb{Z}$, and since it is a p-group, all the p_i are identical, $1 \leq i \leq n$. Thus

$$G \cong \mathbb{Z}_{p^{r_1}} \times \cdots \times \mathbb{Z}_{p^{r_n}}. \tag{2.34}$$

If we have some $r_i > 1$, then G has an element of order greater than p, and thus $\exp(G) \neq p$. Thus $r_i = 1$ for all i, and so

$$G \cong \mathbb{Z}_p \times \cdots \times \mathbb{Z}_p. \tag{2.35}$$

□

2.4 Cyclic Groups

In this section, we look at that well-behaved group, the cyclic group, examining subgroups and the orders of elements. We finish with a beautiful lemma: any finite subgroup of a field's multiplicative group is cyclic. These results will be used in Section 4.3. We follow Fraleigh [Fra03], Gallian [Gal02], and Jacobson [Jac85].

Definition 2.4.1. Let G be a group. If there is some $a \in G$ such that $\langle a \rangle = G$, then G is a *cyclic group*.

Lemma 2.4.2. *If $a \in G$ such that $|a| = n = mk$ for $m, k \geq 1$, then $|a^k| = m$.*

Proof. We have $\langle a \rangle = \{a^1, a^2, \ldots, a^n = 1\}$, such that if $i \not\equiv j \pmod{n}$, then $a^i \neq a^j$. For $1 \leq k \leq n$, this is equivalent to $ik \not\equiv jk \pmod{n}$ implying $a^{ik} \neq a^{jk}$. Thus $\langle a^k \rangle = \{a^k, a^{2k}, \ldots, a^{mk} = 1\}$ is a set of distinct elements, and $|a^k| = m$. □

Lemma 2.4.3. *If $a \in G$ such that $|a| = n$, then $\{k \in \mathbb{Z} \mid a^k = 1\} = n\mathbb{Z}$.*

Proof. Let $k \in n\mathbb{Z}$. Thus $k = n\ell$ for some $\ell \in \mathbb{Z}$, and we have

$$a^k = a^{n\ell} = (a^n)^\ell = 1^\ell = 1. \tag{2.36}$$

Thus $k \in \{k \in \mathbb{Z} \mid a^k = 1\}$. Now let $k \in \{k \in \mathbb{Z} \mid a^k = 1\}$. Now $\langle a \rangle = \{a^1, a^2, \ldots, a^n = 1\}$, such that if $i \not\equiv j \pmod{n}$, then $a^i \neq a^j$. Thus $a^k = 1 = a^n$ implies $k \equiv n \pmod{n}$. Thus there is some $\ell \in \mathbb{Z}$ such that $k = n\ell$, and $k \in n\mathbb{Z}$. □

Definition 2.4.4. The *greatest common divisor* of $n, k \in \mathbb{Z}^+$ will be denoted $\gcd(n, k)$.

Lemma 2.4.5. *Every subgroup of a cyclic group is cyclic.*

Proof. Let $G = \langle a \rangle$ for some $a \in G$, and let $C \leq G$. Now every element of C is some power of a, and we consider their exponents. There exists a minimum $k \in Z^+$ such that $a^k \in C$. Let $a^m \in C$. By the division algorithm, there exist unique $q, r \in \mathbb{Z}$ where

$$m = kq + r \qquad \text{such that} \qquad 0 \leq r < k. \tag{2.37}$$

Thus

$$a^m = a^{kq+r} = a^{kq}a^r \qquad \Rightarrow \qquad a^{-kq}a^m = a^r. \tag{2.38}$$

Now $a^{-kq} = (a^k)^{-q} \in C$, and thus $a^r \in C$. Since k is the smallest positive integer such that $a^k \in C$, we have $r = 0$. Thus $a^m = (a^k)^q$, and $\left\langle a^k \right\rangle = C$. □

Lemma 2.4.6. *Let $k \in \mathbb{Z}^+$ and $a \in G$ such that $|a| = n$. Then $\left\langle a^k \right\rangle = \left\langle a^{\gcd(n,k)} \right\rangle$ and $|a^k| = n/\gcd(n,k)$.*

Proof. Let $d = \gcd(n,k)$. Thus $k = dr$ for some $r \in \mathbb{Z}^+$. Since $a^k = (a^d)^r$, we have $a^k \in \left\langle a^d \right\rangle$ and thus $\left\langle a^k \right\rangle \leq \left\langle a^d \right\rangle$. Now there exist integers r and s such that $nr + ks = d$ [Gal02]. Thus

$$a^d = a^{nr+ks} = a^{nr}a^{ks} = (a^n)^r(a^k)^s = (a^k)^s \in \left\langle a^k \right\rangle. \tag{2.39}$$

Thus $\left\langle a^d \right\rangle \leq \left\langle a^k \right\rangle$, and so $\left\langle a^k \right\rangle = \left\langle a^d \right\rangle$.

Let d be a positive divisor of n. We have $(a^d)^{n/d} = a^n = 1$, and so $|a^d| \leq n/d$. If j is a positive integer such that $j < n/d$, then $dj < n$, and $(a^d)^j \neq 1$. Thus $|a^d| = n/d$, which gives

$$|a^k| = |\left\langle a^k \right\rangle| = |\left\langle a^{\gcd(n,k)} \right\rangle| = |a^{\gcd(n,k)}| = n/\gcd(n,k). \tag{2.40}$$

□

Definition 2.4.7. The *Euler ϕ function* is the function $\phi : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that $\phi(n) = |\{k \in \mathbb{Z}^+ \mid k \leq n \text{ and } \gcd(n,k) = 1\}|$.

Lemma 2.4.8. *Let $|a| = n$ for $a \in G$. Then $\langle a^i \rangle = \langle a^j \rangle$ if and only if $\gcd(n, i) = \gcd(n, j)$. In particular, if $\langle a \rangle = G$ is cyclic of order n, then $\langle a^k \rangle = G$ if and only if $\gcd(n, k) = 1$. In this case, $\phi(n)$ is the number of generators of G.*

Proof. By Lemma 2.4.6, $\langle a^i \rangle = \langle a^{\gcd(n,i)} \rangle$ and $\langle a^j \rangle = \langle a^{\gcd(n,j)} \rangle$. Thus we must show $\langle a^{\gcd(n,i)} \rangle = \langle a^{\gcd(n,j)} \rangle$ if and only if $\gcd(n, i) = \gcd(n, j)$. Now one direction is trivial, and so assume $\langle a^{\gcd(n,i)} \rangle = \langle a^{\gcd(n,j)} \rangle$. By Lemma 2.4.6, we have

$$n/\gcd(n, i) = |a^{\gcd(n,i)}| = |\langle a^{\gcd(n,i)} \rangle| = |\langle a^{\gcd(n,j)} \rangle| = |a^{\gcd(n,j)}| = n/\gcd(n, j). \tag{2.41}$$

This implies $\gcd(n, i) = \gcd(n, j)$.

Now if $G = \langle a \rangle$ is of order n, then $\langle a^k \rangle = \langle a^1 \rangle = G$ if and only if $\gcd(n, k) = \gcd(n, 1) = 1$. Since G is cyclic of order n, we need only consider a^k for $1 \leq k \leq n$. Thus the number of distinct generators of G is $\phi(n)$, by definition of the Euler ϕ function. □

Lemma 2.4.9. *If G is a cyclic group of order n, then there is a unique subgroup of order d for every divisor d of n.*

Proof. Let d be a divisor of n. If $\langle a \rangle = G$ and $|a| = n$, then $|a^{n/d}| = d$, by Lemma 2.4.2, and so $\langle a^{n/d} \rangle$ is a subgroup of order d. By Lemma 2.4.5, every subgroup of G is cyclic, and so let $\langle b \rangle$ be a subgroup of order d. Thus $b^d = 1$ and $b = a^j$ for some j such that $1 \leq j \leq n$. Since $a^{jd} = (a^j)^d = b^d = 1$, $jd = kn$ for some $k \in \mathbb{Z}$, by Lemma 2.4.3. Thus $j = (n/d)k$ and

$$b = a^j = a^{(n/d)k} = (a^{n/d})^k \in \langle a^{n/d} \rangle. \tag{2.42}$$

Since $\langle b \rangle \leq \langle a^{n/d} \rangle$, and both are of order d, $\langle b \rangle = \langle a^{n/d} \rangle$. Thus there is only one subgroup of order d. □

Corollary 2.4.10. *If G is a cyclic group of order n, then for every divisor d of n, there are exactly d elements in $\{g \in G \mid g^d = 1\}$.*

Proof. If $d|n$, then by Lemmas 2.4.5 and 2.4.9, G has a unique subgroup $\langle h \rangle$ of order d. Let $h_i \in \langle h \rangle$ such that $h_i = h^i$. Thus $(h_i)^d = (h^i)^d = (h^d)^i = 1^i = 1$, and so $|\{g \in G \mid g^d = 1\}| \geq d$. Now assume $g^d = 1$. Thus $|g| = k$ and $k|d$, by Lemma 2.4.3. Now $k|d$ and $d|n$ implies $k|n$. Thus G has a unique subgroup of order k, and since $\langle h \rangle$ is cyclic and $k|d$, $\langle h \rangle$ does as well. Thus they are the same subgroup, and $\langle g \rangle \leq \langle h \rangle$, yielding $|\{g \in G \mid g^d = 1\}| = d$. □

Lemma 2.4.11. *If n is a positive integer and $1 \leq d \leq n$, then $n = \sum_{d|n} \phi(d)$.*

Proof. If $C \leq G$ is a cyclic subgroup, let $\operatorname{gen} C$ denote the set of its generators. We have

$$G = \bigcup_{C \leq G} \operatorname{gen} C, \tag{2.43}$$

and this is a disjoint union over all cyclic subgroups C. If G is cyclic, then by Lemma 2.4.9, for $d|n$ we have a unique subgroup of order d. Let C_d denote this subgroup, and $\operatorname{gen} C_d$ its generators. Thus

$$n = |G| = \sum_{d|n} |\operatorname{gen} C_d| = \sum_{d|n} \phi(d), \tag{2.44}$$

where the last equality follows from Lemma 2.4.8. □

Lemma 2.4.12. *A group G of order n is cyclic if and only if for each divisor d of n, there is at most one cyclic subgroup of G having order d.*

Proof. If G is cyclic, then by Lemma 2.4.9, there is exactly one cyclic subgroup of order d for $d|n$, and we have our result. Now assume that if $d|n$, there is at most one cyclic subgroup C_d of order d. Thus $|C_d| = 0$ or d, and C_d has 0 or $\phi(d)$

generators. In Lemma 2.4.11, we observed that any group G is the disjoint union of the generators of its cyclic subgroups. We have

$$n = |G| = \sum_{d|n} |\operatorname{gen} C_d| \leq \sum_{d|n} \phi(d) = n, \tag{2.45}$$

where $|\operatorname{gen} C_d| = 0$ if there is no cyclic subgroup of order d. Thus our inequality is actually equality, and there is exactly one cyclic subgroup C_d of order d for $d|n$. In particular, there is a cyclic subgroup of order n and G is cyclic. □

Corollary 2.4.13. *Let $|G| = n$. If $|\{g \in G \mid g^d = 1\}| \leq d$ for all divisors d of n, then G is cyclic.*

Proof. Let $d|n$. If $|\langle g \rangle| = d$, then $(g^i)^d = (g^d)^i = 1^i = 1$. Thus if $h \in \langle g \rangle$, then $h^d = 1$. Conversely, if $g^d = 1$, then $|\langle g \rangle| = d$. If there were two cyclic subgroups of order d, C_d and C'_d, then $|C_d \cup C'_d| > d$. This implies $|\{g \in G \mid g^d = 1\}| > d$. Thus $|\{g \in G \mid g^d = 1\}| \leq d$ implies there is at most one cyclic subgroup of order d, and so by Lemma 2.4.12, G is cyclic. □

Lemma 2.4.14. *If F is a field, and G is a finite subgroup of $F^\times$, the multiplicative group of nonzero elements of F, then G is cyclic. In particular, if F is a finite field, then $F^\times$ is cyclic.*

Proof. If $|G| = n$ and $a \in G$ satisfies $a^d = 1$, where $d|n$, then a is a root of the polynomial $x^d - 1 \in F[x]$. Such a polynomial has at most d roots, and so

$$|\{g \in G \mid g^d = 1\}| \leq d. \tag{2.46}$$

Thus by Corollary 2.4.13, G is cyclic. If F is finite, then $F^\times$ is a finite subgroup of $F^\times$. □

2.5 Field Theory Lemmas

The results in this section prepare for Section 5.1 on field extensions and nontrivial field automorphisms. To that end, we examine the notion of a prime subfield. We follow the treatment of Fraleigh [Fra03].

Definition 2.5.1. In a ring R with unity 1, for $n \in \mathbb{Z}$ we let

$$n \cdot 1 = \begin{cases} 1 + \ldots + 1 & \text{if } n > 0, \\ -1 - \ldots - 1 & \text{if } n < 0, \\ 0 & \text{if } n = 0. \end{cases}$$

where there are $|n|$ summands.

Lemma 2.5.2. *If R is a ring with unity 1, then*

$$\varphi : \mathbb{Z} \to R \qquad \textit{such that} \qquad \varphi(n) = n \cdot 1 \tag{2.47}$$

is a ring homomorphism.

Proof. Now $(n \cdot 1) \in R$, and so φ is well-defined. We have

$$\varphi(n) + \varphi(m) = (n \cdot 1) + (m \cdot 1) = (n + m) \cdot 1 = \varphi(n + m). \tag{2.48}$$

and

$$\varphi(n)\varphi(m) = (n \cdot 1)(m \cdot 1) = nm \cdot 1 = \varphi(nm) \tag{2.49}$$

for $n, m \in \mathbb{Z}$, so φ is a ring homomorphism. □

Lemma 2.5.3. *If a ring R with unity has characteristic $n > 0$, then it has a subring isomorphic to $\mathbb{Z}_n$. If R has characteristic 0, then it has a subring isomorphic to $\mathbb{Z}$. Such subrings are contained in any subring S of R that contains 1.*

Proof. By Lemma 2.5.2, $\varphi(\mathbb{Z}) \leq R$. Thus $\ker\varphi$ is an ideal of $\mathbb{Z}$, whose ideals are of the form $k\mathbb{Z}$ for $k \in \mathbb{Z}$. If R is of characteristic $n > 0$, then $\ker\varphi = n\mathbb{Z}$. By the first homomorphism theorem for rings, the subring generated by 1 is

$$\mathbb{Z}_n \cong \mathbb{Z}/n\mathbb{Z} \cong \varphi(\mathbb{Z}) \leq R. \tag{2.50}$$

If R is of characteristic 0, then $n \cdot 1 \neq 0$ for all $n \neq 0$. Thus $\ker\varphi = 0$, and so the subring generated by 1 is

$$\mathbb{Z} \cong \mathbb{Z}/0 \cong \varphi(\mathbb{Z}) \leq R. \tag{2.51}$$

If a subring S of R contains 1, then S contains the subring generated by 1. □

Lemma 2.5.4. *A field F is either of prime characteristic p and contains a subfield isomorphic to $\mathbb{Z}_p$, or is of characteristic 0 and contains a subfield isomorphic to $\mathbb{Q}$. Such subfields are contained in any subfield of F.*

Proof. If F is of characteristic $n > 0$, then by Lemma 2.5.3, F has a subring isomorphic to $\mathbb{Z}_n$. Assume $n \neq p$, for p a prime. Let $n = ab$ such that $1 < a, b < n$. Thus

$$(a \cdot 1)(b \cdot 1) = \varphi(a)\varphi(b) = \varphi(ab) = n \cdot 1 = 0 \tag{2.52}$$

such that $(a \cdot 1), (b \cdot 1) \neq 0$. Since these are zero-divisors, they have no multiplicative inverse, implying F is not a field. Thus $n = p$, a prime. Now $\mathbb{Z}_p$ is a field, and so F has a subfield isomorphic to $\mathbb{Z}_p$.

If the characteristic of F is 0, then by Lemma 2.5.3, F has a subring S isomorphic to $\mathbb{Z}$. We can construct a field of fractions T from S, such that $T = \{ab^{-1} \mid a, b \in S, b \neq 0\} \leq F$. This field of fractions is isomorphic to $\mathbb{Q}$. Thus if K is a subfield of F, since $1 \in K$, K also contains the subfield generated by 1. □

2.6 Vector Space Lemmas

In this section, we prove some basic linear algebra facts, in preparation for Chapter 4 on linear groups and Chapter 6 on bilinear forms. We establish the isomorphism between an n-dimensional vector space, over a field F, with F^n. We look at the matrix of a linear transformation, relative to a fixed basis. We also examine the matrix manipulations corresponding to the mapping of vectors and the composition of transformations. We follow the treatment of Friedberg, Insel, and Spence [FIS03].

Remark 2.6.1. We use row spaces and right linear transformations throughout the thesis. The text of Friedberg, Insel, and Spence uses column spaces and left linear transformations, but the results still hold *mutatis mutandis*.

Definition 2.6.2. Let X be a set and F be a field. We define $\mathcal{F}(X)$ as the set of all functions from X to F, that is, $\{f \mid f : X \to F\}$.

Lemma 2.6.3. *$\mathcal{F}(X)$ is a vector space over F, where for $f, g \in \mathcal{F}(X)$ and $\alpha \in F$, $f + g$ is defined as*

$$(f + g)(x) = f(x) + g(x) \text{ for all } x \in X, \tag{2.53}$$

and αf is defined as

$$(\alpha f)(x) = \alpha f(x) \in F \text{ for all } \alpha \in F \text{ and } x \in X. \tag{2.54}$$

Proof. First, $f(x), g(x) \in F$ implies $f(x) + g(x) \in F$, and $\alpha, f(x) \in F$ implies $\alpha f(x) \in F$, because $+$ and $\cdot$ are both defined in F. Thus $f + g, \alpha f \in \mathcal{F}(x)$. For all $x \in X$, we have

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x), \tag{2.55}$$

because the additive group of F is abelian. Thus $f+g=g+f$ for all $f,g\in\mathcal{F}(X)$. Now let $h\in\mathcal{F}(X)$. For all $x\in X$, we have

$$\begin{aligned}((f+g)+h)(x)&=(f+g)(x)+h(x)=(f(x)+g(x))+h(x)\\&=f(x)+(g(x)+h(x))=f(x)+(g+h)(x)=(f+(g+h))(x),\end{aligned}\tag{2.56}$$

because the additive group of F obeys the associative property. Thus $(f+g)+h=f+(g+h)$ for all $f,g,h\in\mathcal{F}(X)$. Define $0:X\to F$ such that $0(x)=0$. Thus $0\in\mathcal{F}(X)$, and if $f\in\mathcal{F}(X)$, then for all $x\in X$,

$$(f+0)(x)=f(x)+0(x)=f(x)+0=f(x).\tag{2.57}$$

Thus $f+0=f$ for all $f\in\mathcal{F}(X)$. If $f\in\mathcal{F}(X)$, define $-f$ such that $(-f)(x)=-f(x)$ for all $x\in X$. Since $-f(x)\in F$, $-f\in\mathcal{F}(x)$. For all $x\in X$, we have

$$(f+(-f))(x)=f(x)+(-f)(x)=f(x)+(-f(x))=0=0(x).\tag{2.58}$$

Thus $f+(-f)=0$ for all $f\in\mathcal{F}(X)$. Now for $1\in F$ and $f\in\mathcal{F}(X)$, we have

$$(1f)(x)=1f(x)=f(x)\tag{2.59}$$

for all $x\in X$. Thus $1f=f$ for all $f\in\mathcal{F}(X)$. Now let $\alpha,\beta\in F$ and $f\in\mathcal{F}(X)$. For all $x\in X$, we have

$$((\alpha\beta)f)(x)=(\alpha\beta)f(x)=\alpha(\beta f(x))=\alpha((\beta f)(x))=(\alpha(\beta f))(x),\tag{2.60}$$

because $\cdot$ obeys the associative law in F. Thus $(\alpha\beta)f=\alpha(\beta f)$ for all $\alpha,\beta\in F$ and $f\in\mathcal{F}(X)$. We again let $\alpha,\beta\in F$ and $f\in\mathcal{F}(X)$. For all $x\in X$, we have

$$\begin{aligned}((\alpha+\beta)f)(x)&=(\alpha+\beta)f(x)=\alpha f(x)+\beta f(x)\\&=(\alpha f)x+(\beta f)(x)=(\alpha f+\beta f)(x),\end{aligned}\tag{2.61}$$

because F obeys the right-distributive law. Thus $(\alpha+\beta)f = \alpha f + \beta f$ for all $\alpha, \beta \in F$ and $f \in \mathcal{F}(X)$. Similarly, let $\alpha \in F$ and $f, g \in \mathcal{F}(X)$. For all $x \in X$, we have

$$(\alpha(f+g))(x) = \alpha(f+g)(x) = \alpha(f(x)+g(x)) = \alpha f(x) + \alpha g(x)$$
$$= (\alpha f)(x) + (\alpha g)(x) = (\alpha f + \alpha g)(x), \quad (2.62)$$

because F obeys the left-distributive law. Thus $\alpha(f+g) = \alpha f + \alpha g$ for all $\alpha \in F$ and $f, g \in \mathcal{F}(X)$. Thus $\mathcal{F}(X)$ satisfies the axioms of a vector space. □

Definition 2.6.4. We define F^n as the set of $1 \times n$ row vectors with entries from a field F. If $x \in F^n$, then x_j denotes the jth entry of x, for $1 \leq j \leq n$. If we define for $x, y \in F^n$ and $a \in F$, that

$$x + y = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n) \in F^n \quad (2.63)$$

and

$$ax = a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n) \in F^n, \quad (2.64)$$

then F^n is an n-dimensional vector space over F. (In fact, $F^n \cong \mathcal{F}(\{1, \ldots, n\})$.)

Lemma 2.6.5. *If V is an n-dimensional vector space over F, then V is isomorphic to F^n.*

Proof. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V. Therefore, the ordered set of scalars from F, $(x_1, \ldots, x_n)$, determines exactly one vector $x = x_1v_1 + \ldots + x_nv_n \in V$. Define $\phi_\beta : V \to F^n$ such that $\phi_\beta(x) = (x_1, \ldots, x_n)$. By the uniqueness of the set of scalars associated with $x \in V$, ϕ_β is well-defined. Assume $\phi_\beta(x) = \phi_\beta(y)$. Then $(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$, which implies $x_i = y_i$ for $1 \leq i \leq n$. Again by the uniqueness of scalars, we have $x = y$. Thus ϕ_β is one-to-one. Now let

$(x_1, \ldots, x_n) \in F^n$. Now $x_1v_1 + \ldots + x_nv_n = x \in V$, and thus $\phi_\beta(x) = (x_1, \ldots, x_n)$ and ϕ_β is onto. For $x, y \in V$, we have

$$\begin{aligned} x + y &= (x_1v_1 + \ldots + x_nv_n) + (y_1v_1 + \ldots + y_nv_n) \\ &= (x_1v_1 + y_1v_1) + \ldots + (x_nv_n + y_nv_n) = (x_1 + y_1)v_1 + \ldots + (x_n + y_n)v_n. \end{aligned} \quad (2.65)$$

Thus

$$\phi_\beta(x)+\phi_\beta(y) = (x_1, \ldots, x_n)+(y_1, \ldots, y_n) = (x_1+y_1, \ldots, x_n+y_n) = \phi_\beta(x+y). \quad (2.66)$$

Now for $x \in V$ and $a \in F$, we have

$$ax = a(x_1v_1+\ldots+x_nv_n) = a(x_1v_1)+\ldots+a(x_nv_n) = (ax_1)v_1+\ldots+(ax_n)v_n. \quad (2.67)$$

Thus

$$a\phi_\beta(x) = a(x_1, \ldots, x_n) = (ax_1, \ldots, ax_n) = \phi_\beta(ax). \quad (2.68)$$

Therefore, ϕ_β is linear. □

Definition 2.6.6. For ϕ_β as in Lemma 2.6.5, we denote the image $\phi_\beta(x)$ as $[x]_\beta$.

Definition 2.6.7. Let V be an n-dimensional vector space over F, and let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V. If $T : V \to V$ is a linear transformation such that $x \mapsto xT$, we define the *representing matrix* of T with respect to β as the $n \times n$ matrix $[T]_\beta$ which has $[v_iT]_\beta$ as its ith row, for $1 \leq i \leq n$.

Definition 2.6.8. If $A \in M_n(F)$ and $x \in F^n$, then define $xA \in F^n$ such that the jth coordinate of xA, $(xA)_j$, is $x_1A_{1j} + \ldots + x_nA_{nj}$, for $1 \leq j \leq n$. Thus if A_i is the ith row of A,

$$xA = x_1A_1 + \ldots + x_nA_n. \quad (2.69)$$

Lemma 2.6.9. *Let V be n-dimensional over F, $\beta = \{v_1, \ldots, v_n\}$ a basis, and T a linear transformation. We have $[xT]_\beta = [x]_\beta[T]_\beta$.*

Proof. Now $x = x_1v_1 + \ldots + x_nv_n$, and by the linearity of T,

$$xT = (x_1v_1 + \ldots + x_nv_n)T = (x_1v_1)T + \ldots + (x_nv_n)T = x_1(v_1T) + \ldots + x_n(v_nT). \quad (2.70)$$

Thus

$$\begin{aligned} [xT]_\beta &= [x_1(v_1T) + \ldots + x_n(v_nT)]_\beta \\ &= [x_1(v_1T)]_\beta + \ldots + [x_n(v_nT)]_\beta \\ &= x_1[v_1T]_\beta + \ldots + x_n[v_nT]_\beta \\ &= [x]_\beta[T]_\beta. \end{aligned} \quad (2.71)$$

□

Definition 2.6.10. Let $T : V \to V$ be linear, where V is n-dimensional over F. Let $\beta = \{v_1, \ldots, v_n\}$ and $\gamma = \{w_1, \ldots, w_n\}$ be bases for V. Define $[T]^\gamma_\beta \in M_n(F)$ such that $[v_iT]_\gamma$ is the ith row of $[T]^\gamma_\beta$, for $1 \leq i \leq n$.

The next two results are from Friedberg, Insel, and Spence [FIS03].

Lemma 2.6.11. *Let $T : V \to V$ be a linear transformation, and let β, γ be bases for V. Then*

$$[v]_\beta[T]^\gamma_\beta = [vT]_\gamma \quad (2.72)$$

for all $v \in V$. □

Lemma 2.6.12. *Let $S, T : V \to V$ be linear transformations, and let α, β, γ be bases for V. Then*

$$[S]^\gamma_\beta[T]^\beta_\alpha = [ST]^\gamma_\alpha. \quad (2.73)$$

□

Definition 2.6.13. Let V be n-dimensional over F, and let β, γ be bases for V. We call $Q = [I]^\gamma_\beta$ the *change of coordinate* matrix from bases β to γ.

Lemma 2.6.14. *Let Q be the change of coordinate matrix from bases β to γ. Then Q is invertible, and $[v]_\beta Q = [v]_\gamma$ for all $v \in V$.*

Proof. Now $Q = [I]^\gamma_\beta$, and we have $[I]^\beta_\gamma \in M_n(F)$. Now by Lemma 2.6.12,

$$[I]^\gamma_\beta [I]^\beta_\gamma = [I]^\gamma_\gamma = I = [I]^\beta_\beta = [I]^\beta_\gamma [I]^\gamma_\beta \tag{2.74}$$

and thus $[I]^\beta_\gamma = Q^{-1}$. Now let $v \in V$. Thus $[v]_\beta \in F^n$ and by Lemma 2.6.11

$$[v]_\beta Q = [v]_\beta [I]^\gamma_\beta = [vI]_\gamma = [v]_\gamma. \tag{2.75}$$

□

CHAPTER 3

GROUP ACTIONS AND MULTIPLE TRANSITIVITY

In this chapter, we develop a criterion for group simplicity following from multiple transitivity (as well as from the order of the set being acted upon). We follow the treatment of Rotman [Rot95].

3.1 Group Actions

In this section, we define (right) group actions and the homomorphism of a group action.

Definition 3.1.1. Let X be a set and G a group. A *group action* of G on X is a map $X \times G \to X$ satisfying:

(1) $x1 = x$

(2) $x(g_1 g_2) = (xg_1)g_2$

for all $x \in X$ and all $g_1, g_2 \in G$. To say G *acts on* X means to fix a particular action of G on X. If G acts on X, X is called a *G-set*, and the elements of X are called *points*. Note that if G acts on X and $H \leq G$, then H also acts on X.

Remark 3.1.2. All group actions described in this thesis will be *right* actions, with right permutations written as right operators.

Example 3.1.3. S_n acts on $\{1, \ldots, n\}$, and a group G acts on itself under conjugation by $g \in G$.

Theorem 3.1.4. *Let G act on X, and let S_X denote the symmetric group on X. For each $g \in G$, the function $\rho_g : X \to X$ such that $x\rho_g = xg$ for all $x \in X$ is a permutation of X, and the map $\rho : G \to S_X$ defined by $\rho(g) = \rho_g$ is a homomorphism.*

Proof. The product $\rho_a\rho_b$ is function composition, so $x(\rho_a\rho_b) = (x\rho_a)\rho_b$ for $x \in X$. Let $g \in G$. By definition, $g^{-1} \in G$. For $x \in X$, we have

$$x(\rho_g\rho_{g^{-1}}) = (x\rho_g)\rho_{g^{-1}} = (xg)\rho_{g^{-1}} = (xg)g^{-1} = x(gg^{-1}) = x1 = x, \tag{3.1}$$

and similarly, $x(\rho_{g^{-1}}\rho_g) = x$. Since this is true for all $x \in X$,

$$\rho_g\rho_{g^{-1}} = 1_{S_X} = \rho_{g^{-1}}\rho_g \tag{3.2}$$

for all $g \in G$. Since ρ_g has an inverse, $\rho_g : X \to X$ is a bijection and thus a permutation of X.

Now define $\rho : G \to S_X$ where $\rho(g) = \rho_g$, and let $g_1, g_2 \in G$. So we have

$$x\rho_{g_1g_2} = x(g_1g_2) = (xg_1)g_2 = (x\rho_{g_1})g_2 = (x\rho_{g_1})\rho_{g_2} = x\rho_{g_1}\rho_{g_2} \tag{3.3}$$

Since $x\rho_{g_1g_2} = x\rho_{g_1}\rho_{g_2}$ for all $x \in X$, we have $\rho_{g_1g_2} = \rho_{g_1}\rho_{g_2}$ for all $\rho_{g_1}, \rho_{g_2} \in S_X$. Thus

$$\rho(g_1g_2) = \rho_{g_1g_2} = \rho_{g_1}\rho_{g_2} = \rho(g_1)\rho(g_2) \tag{3.4}$$

for all $g_1, g_2 \in G$. □

Definition 3.1.5. The homomorphism ρ in Theorem 3.1.4 is the *homomorphism of the action of G on X*.

3.2 Transitivity, Orbits, and Stabilizers

In this section, we define transitivity of a group action. We describe stabilizer subgroups and the orbit of a subgroup, and we bring these concepts together in the

orbit-stabilizer theorem. We also distinguish pointwise and setwise stabilizers. We define the kernel of an action and faithful actions. Finally, we describe the conditions under which a factor group induces an action. Then we discuss the example of p-groups acting on themselves by conjugation.

Theorem 3.2.1. *Let G act on X. For $x_1, x_2 \in X$, let $x_1 \sim x_2$ if and only if there is a $g \in G$ such that $x_1 g = x_2$. Then $\sim$ is an equivalence relation.*

Proof. By definition, $1 \in G$ and $x1 = x$ for all $x \in X$. Thus $x \sim x$, and $\sim$ is reflexive. Now let $x_1 \sim x_2$. Thus there is a $g \in G$ such that $x_1 g = x_2$. So $g^{-1} \in G$ and we have

$$x_1 = x_1 1 = x_1(gg^{-1}) = (x_1 g)g^{-1} = x_2 g^{-1}. \tag{3.5}$$

Thus $x_2 \sim x_1$, and $\sim$ is symmetric. Now let $x_1 \sim x_2$ and $x_2 \sim x_3$. So there exist $g, h \in G$ such that $x_1 g = x_2$ and $x_2 h = x_3$. So we have

$$x_1(gh) = (x_1 g)h = x_2 h = x_3 \tag{3.6}$$

Thus $x_1 \sim x_3$, and $\sim$ is transitive. □

Definition 3.2.2. We call the equivalence class $xG = \{xg \mid g \in G\}$ the *orbit* of x under G.

Definition 3.2.3. We say G acts *transitively* on X if for each $x_1, x_2 \in X$, there exists some $g \in G$ such that $x_1 g = x_2$.

Lemma 3.2.4. *If G acts on X, then the following are equivalent:*

(1) G acts transitively on X.

(2) There exists an $x_0 \in X$ such that for all $x \in X$, there exists a $g \in G$ where $xg = x_0$.

(3) There exists an $x_0 \in X$ such that for all $x \in X$, there exists a $g \in G$ where $x_0 g = x$.

Proof. Assume G act transitively on X. Fix $x_0 \in X$, and let $x \in X$. By definition, there exists a $g \in G$ such that $xg = x_0$.

Now let there be an $x_0 \in X$ where for all $x \in X$, there exists a $g \in G$ such that $xg = x_0$. If $x \in X$, we have

$$x = x1 = x(gg^{-1}) = (xg)g^{-1} = x_0 g^{-1}. \tag{3.7}$$

Thus $g^{-1} \in G$ gives $x_0 g^{-1} = x$.

Assume there is an $x_0 \in X$, where for all $x \in X$, there exists a $g \in G$ such that $x_0 g = x$. Let $x_1, x_2 \in X$. Thus there exist $g_1, g_2 \in G$, such that $x_0 g_1 = x_1$ and $x_0 g_2 = x_2$. By the same reasoning as above, $x_1 g_1^{-1} = x_0$. Therefore,

$$x_1(g_1^{-1} g_2) = (x g_1^{-1}) g_2 = x_0 g_2 = x_2, \tag{3.8}$$

and since $g_1^{-1} g_2 \in G$, we see that G acts transitively on X. □

Lemma 3.2.5. *If $\{X_i \mid i \in I\}$ is a partition of a G-set X such that G acts transitively on each X_i, then the X_i are the orbits of X under G.*

Proof. We must show that each X_i is an orbit. If $x_i \in X_i$, then $x_i G \subseteq X_i$, because X_i is a G-set. Now if $y, x_i \in X_i$, then since X_i is transitive, there exists $g \in G$ such that $y = x_i g$. Thus $y \in x_i G$ and so $X_i \subseteq x_i G$. □

Lemma 3.2.6. *G acts transitively on X if and only if X has only one orbit under the action of G.*

Proof. Let G act transitively on X. Fix $x_0 \in X$ and let $x \in X$. Thus there exists a $g \in G$ such that $x_0 g = x$. Thus $x \in xG$, and so $X \subseteq xG$. As the reverse inclusion is trivial, $xG = X$, and X has one orbit under G.

Conversely, let X have only 1 orbit under the action of G. Thus $xG = X$ for all $x \in X$. By Lemma 3.2.4, G acts transitively on X. □

Theorem 3.2.7. *The stabilizer of* x, $\mathrm{Stab}(x) = \{g \in G \mid xg = x\}$, *is a subgroup of* G.

Proof. Let $g_1, g_2 \in \mathrm{Stab}(x)$. Thus

$$x(g_1g_2) = (xg_1)g_2 = xg_2 = x \tag{3.9}$$

and so $g_1g_2 \in \mathrm{Stab}(x)$. Also $x1 = x$, so $1 \in \mathrm{Stab}(x)$. Now let $g \in \mathrm{Stab}(x)$. So $g^{-1} \in G$ and

$$xg^{-1} = (xg)g^{-1} = x(gg^{-1}) = x1 = x \tag{3.10}$$

and so $g^{-1} \in \mathrm{Stab}(x)$. □

Definition 3.2.8. If G acts on X and $x_1, \ldots, x_t \in X$, then the *pointwise stabilizer* of these points is $\mathrm{Stab}(x_1, \ldots, x_t) = \{g \in G \mid x_ig = x_i, 1 \leq i \leq t\}$. We have

$$\mathrm{Stab}(x_1, \ldots, x_t) = \bigcap_{i=1}^{t} \mathrm{Stab}(x_i). \tag{3.11}$$

Definition 3.2.9. If G acts on X and $Y = \{x_1, \ldots, x_t\} \subseteq X$, then the *setwise stabilizer* of Y is $\mathrm{Stab}(Y) = \{g \in G \mid Yg = Y\}$. Thus $g \in \mathrm{Stab}(Y)$ permutes the points of Y, while $g \in \mathrm{Stab}(x_1, \ldots, x_t)$ fixes Y pointwise.

Lemma 3.2.10. *We have*

$$\mathrm{Stab}(x_1, \ldots, x_i) \leq \mathrm{Stab}(\{x_1, \ldots, x_i\}) \tag{3.12}$$

and

$$\mathrm{Stab}(x_1, \ldots, x_i, x_{i+1}) \leq \mathrm{Stab}(x_1, \ldots, x_i). \tag{3.13}$$

□

Definition 3.2.11. If G acts on X, then $\ker\rho$ is the *kernel of the action*, where ρ is the homomorphism of the action. Thus $\ker\rho$ is the subgroup of G whose elements fix X pointwise.

Theorem 3.2.12. *We have* $\ker\rho = \bigcap_{x\in X} \mathrm{Stab}(x)$.

Proof. If $h \in \bigcap_{x\in X} \mathrm{Stab}(x)$, then $xh = x = x1$ for all $x \in X$. Thus

$$\rho(h) = \rho_h = \rho_1 = 1_{S_X} \tag{3.14}$$

and so $h \in \ker\rho$. Conversely, if $h \in \ker\rho$, then $\rho(h) = \rho_h = 1_{S_X} = \rho_1$ and

$$xh = x\rho_h = x\rho_1 = x1 = x \tag{3.15}$$

for all $x \in X$, and so $h \in \bigcap_{x\in X} \mathrm{Stab}(x)$. □

Theorem 3.2.13. *Let G act on X and let $x, y \in X$. If $y = xh$ for $h \in G$, then* $\mathrm{Stab}(y) = h^{-1}\,\mathrm{Stab}(x)h$.

Proof. If $g \in \mathrm{Stab}(x)$, then

$$y(h^{-1}gh) = (xh)(h^{-1}gh) = x(hh^{-1})gh = x(gh) = (xg)h = xh = y. \tag{3.16}$$

Thus $h^{-1}gh \in \mathrm{Stab}(y)$, and we have $h^{-1}\,\mathrm{Stab}(x)h \subseteq \mathrm{Stab}(y)$. Conversely, if $g \in \mathrm{Stab}(y)$, then

$$x(hgh^{-1}) = (xh)gh^{-1} = y(gh^{-1}) = (yg)h^{-1} = yh^{-1} = x. \tag{3.17}$$

Thus $hgh^{-1} \in \mathrm{Stab}(x)$, implying $g \in h^{-1}\,\mathrm{Stab}(x)h$. So $\mathrm{Stab}(y) \subseteq h^{-1}\,\mathrm{Stab}(x)h$, and thus $\mathrm{Stab}(y) = h^{-1}\,\mathrm{Stab}(x)h$ for $y = xh$. □

Theorem 3.2.14. *Let G act transitively on X and fix $x \in X$. If $\ker\rho$ is the kernel of the action, then*

$$\ker\rho = \bigcap_{g\in G} \mathrm{Stab}(xg) = \bigcap_{g\in G} g^{-1}\,\mathrm{Stab}(x)g. \tag{3.18}$$

Proof. By Lemma 3.2.6, we know $xG = X$. Thus if $y \in X$, there exists a $g \in G$ such that $y = xg$. Thus

$$\bigcap_{x \in X} \mathrm{Stab}(x) = \bigcap_{g \in G} \mathrm{Stab}(xg) \tag{3.19}$$

and so by Theorem 3.2.12, we have the first equality. By Theorem 3.2.13 we have the second equality. □

Definition 3.2.15. A group G acts *faithfully* on X if the kernel of the action is trivial. When $\ker \rho = 1$, ρ is one-to-one and G is isomorphic to a subgroup of S_X.

Lemma 3.2.16. *If G acts on X and $N \lhd G$, then G/N has a well-defined action on X by coset representatives if and only if $N \leq \ker \rho$.*

Proof. Let G/N have a well-defined action on X such that $x(gN) = xg$ for all $x \in X$ and $gN \in G/N$. Since it is well-defined, $g_1N = g_2N$ implies $xg_1 = xg_2$, for all $g_1, g_2 \in G$ and all $x \in X$. If $n \in N$, then $nN = 1N = N$. Thus $xn = x1 = x$ for all $x \in X$, and $n \in \ker \rho$. Therefore, $N \leq \ker \rho$.

Now let $N \leq \ker \rho$ such that $N \lhd G$. Thus $xn = x$ for all $n \in N$, all $x \in X$. Let $g_1N = g_2N$. Since $N \lhd G$, this implies $Ng_1 = Ng_2$. But this is equivalent to $N(g_1g_2^{-1}) = N$, and thus $g_1g_2^{-1} \in N \leq \ker \rho$. If $x \in X$, then $x(g_1g_2^{-1}) = x$, and since G acts on X,

$$xg_2 = (x(g_1g_2^{-1}))g_2 = x((g_1g_2^{-1})g_2) = x(g_1(g_2^{-1}g_2)) = x(g_11) = xg_1. \tag{3.20}$$

Thus using coset representatives is well-defined. Now let $x \in X$. Since N is the identity of G/N and

$$xN = x(1N) = x1 = x, \tag{3.21}$$

Condition (1) is satisfied. Now let $g_1N, g_2N \in G/N$ and $x \in X$. We have

$$x(g_1Ng_2N) = x(g_1g_2N) = x(g_1g_2) = (xg_1)g_2 = (xg_1)g_2N = (xg_1N)g_2N, \tag{3.22}$$

and thus Condition (2) is satisfied. Therefore, G/N acts on X by coset representatives. □

Definition 3.2.17. If G acts on X, $N \lhd G$, and $N \leq \ker \rho$, then the action of G/N on X by coset representatives (from Lemma 3.2.16) is the *induced* action of G/N on X.

Corollary 3.2.18. *If G acts on X and ρ is the homomorphism of the action, then the induced action of $G/\ker \rho$ on X is faithful.*

Proof. By Lemma 3.2.16, $G/\ker \rho$ acts on X by coset representatives. Now the identity element of $G/\ker \rho$ is $\ker \rho$, the kernel of the inherited action of G on X. By definition, $G/\ker \rho$ acts faithfully on X. □

Lemma 3.2.19. *If G acts on X, then $\mathrm{Stab}(x)$ acts on $X \setminus \{x\}$. Further, if $\mathrm{Stab}(x)$ acts faithfully on $X \setminus \{x\}$, then G acts faithfully on X.*

Proof. Assume $yg = x$, for $x \neq y \in X$ and $g \in \mathrm{Stab}(x)$. This implies

$$y = y1 = y(gg^{-1}) = (yg)g^{-1} = xg^{-1}, \tag{3.23}$$

which contradicts that $g^{-1} \in \mathrm{Stab}(x)$. Thus $yg \in X \setminus \{x\}$ for all $y \in X \setminus \{x\}$. Thus let $\mathrm{Stab}(x)$ act on $X \setminus \{x\}$, such that for $g \in \mathrm{Stab}(x) \leq G$ and $y \in X \setminus \{x\} \subseteq X$, yg has the same value as in the original action. Conditions (1) and (2) are thus satisfied.

Now let $\mathrm{Stab}(x)$ act faithfully on $X \setminus \{x\}$, and assume G does not act faithfully on X. Then there is a $g \in G$ such that $xg = x$ for all $x \in X$, such that $g \neq 1$. But this implies $1 \neq g \in \mathrm{Stab}(x)$, and $yg = y$ for all $y \in X \setminus \{x\}$, a contradiction. □

Theorem 3.2.20. *(Orbit-Stabilizer) Let G act on X. Fix $x \in X$, and let* $\mathrm{Stab}(x)\backslash G$ *denote the right cosets of* $\mathrm{Stab}(x)$. *Define* $\psi : \mathrm{Stab}(x)\backslash G \to xG$ *such that* $\psi(\mathrm{Stab}(x)g) = xg$. *Then ψ is a bijection and* $|xG| = [G : \mathrm{Stab}(x)]$. *If G is finite, then $|xG|$ is a divisor of $|G|$.*

Proof. Let $\mathrm{Stab}(x)g_1 = \mathrm{Stab}(x)g_2$ for $g_1, g_2 \in G$. Thus $\mathrm{Stab}(x) = \mathrm{Stab}(x)(g_2 g_1^{-1})$ and so $g_2 g_1^{-1} \in \mathrm{Stab}(x)$. Thus

$$xg_1 = (xg_2g_1^{-1})g_1 = xg_2(g_1^{-1}g_1) = xg_2, \tag{3.24}$$

and ψ is well-defined.

Let $\psi(\mathrm{Stab}(x)g_1) = \psi(\mathrm{Stab}(x)g_2)$. Thus $xg_1 = xg_2$. So we have

$$x = x1 = x(g_1g_1^{-1}) = (xg_1)g_1^{-1} = (xg_2)g_1^{-1} = x(g_2g_1^{-1}), \tag{3.25}$$

and so $g_2g_1^{-1} \in \mathrm{Stab}(x)$. Thus $g_2 \in \mathrm{Stab}(x)g_1$ and so $\mathrm{Stab}(x)g_1 = \mathrm{Stab}(x)g_2$, and ψ is one-to-one. Let $y \in xG$. So $y = xg_1$ for some $g_1 \in G$. Thus $\psi(\mathrm{Stab}(x)g_1) = xg_1 = y$, and ψ is onto.

Since ψ is a bijection, the sets $\mathrm{Stab}(x)\backslash G$ and xG have equal cardinality. Thus $|xG| = [G : \mathrm{Stab}(x)]$. If G is finite, then $|xG| = [G : \mathrm{Stab}(x)] = |G|/|\mathrm{Stab}(x)|$, and so $|xG|$ divides $|G|$. □

Corollary 3.2.21. *If G acts transitively on X and $|X| = n$, then* $|X| = [G : \mathrm{Stab}(x)]$ *and* $|G| = n|\mathrm{Stab}(x)|$.

Proof. Fix $x \in X$. From Theorem 3.2.20, $|xG| = [G : \mathrm{Stab}(x)]$. Since G acts transitively, $xG = X$, by Lemma 3.2.6, and we have $|X| = [G : \mathrm{Stab}(x)]$. Thus

$$n = |X| = [G : \mathrm{Stab}(x)] = |G|/|\mathrm{Stab}(x)|. \tag{3.26}$$

□

We give an example of the utility of the orbit-stabilizer theorem by considering the action of p-groups on themselves by conjugation.

Definition 3.2.22. The *centralizer* of $a \in G$ is the set $C(a) = \{g \in G \mid ga = ag\}$. Thus $Z(G) \subseteq C(a)$. Since this can be written $\{g \in G \mid g^{-1}ag = a\}$, it is the stabilizer of a under the action of conjugation. Thus the orbit-stabilizer theorem says that for a^G, the *conjugacy class* of a, $|a^G| = |G : C(a)|$ and $|a^G|$ divides $|G|$, for all $a \in G$.

Corollary 3.2.23. *Any finite group G of order a power of a prime p has a center $Z(G) \neq 1$.*

Proof. Let G act on itself by conjugation. By Lemma 3.2.5 and 3.2.20, G is partitioned into orbits such that the order of each orbit divides the order of G. If $a \in G$ is in a one-point orbit, then $g^{-1}ag = a$ for all $g \in G$, meaning $ag = ga$ for all $g \in G$. Thus $a \in Z(G)$. Now let $\{h_i\}$ be a set of representatives from the orbits of more than one point. Thus

$$|G| = |Z(G)| + \sum_i |h_i^G|. \tag{3.27}$$

Since $|h_i^G|$ divides $|G| = p^k$ and $|h_i^G| > 1$, we know p divides $|h_i^G|$, for all i. Since p divides $|G|$, this implies that p divides $|Z(G)|$, requiring $Z(G)$ to have at least p elements, since $1 \in Z(G)$. □

For convenience, we repeat the definition of elementary abelian p-groups.

Definition 3.2.24. An *elementary abelian p-group* is a finite group G isomorphic to $\mathbb{Z}_p \times \cdots \times \mathbb{Z}_p$, that is, ${\mathbb{Z}_p}^n$ for some positive integer n.

Lemma 3.2.25. *If for finite $H \lhd G$ all the non-identity elements of H are conjugate in G, then H is an elementary abelian p-group.*

Proof. Let h_1, h_2 be elements of finite $H \lhd G$ such that $h_1, h_2 \neq 1$. Thus there exists a $g \in G$ such that $g^{-1}h_1 g = h_2$. Let $|h_2| = n$. We have

$$1 = {h_2}^n = (g^{-1}h_1 g)^n = g^{-1}{h_1}^n g \tag{3.28}$$

implying ${h_1}^n = 1$. Thus $|h_1| \leq |h_2|$. By a similar argument, if $|h_1| = n'$, then ${h_2}^{n'} = 1$. Thus $|h_2| \leq |h_1|$. Since $|h_1| = |h_2|$, we conclude that every non-identity element of H is of the same order n.

Let $1 \neq h \in H$. Thus $|h| = |h^i|$ for $1 \leq i < n = |h|$, meaning $\langle h \rangle = \langle h^i \rangle$ for $1 \leq i < n$. Thus $\gcd(n, i) = 1$ for such i, by Lemma 2.4.8. But this only happens if $n = p$, where p is a prime. Thus $\exp(H) = p$, and H is a p-group by Lemma 2.3.5.

By Corollary 3.2.23 there is a $1 \neq z \in Z(H)$, and if $1 \neq h \in H$, there exists a $g \in H$ such that $h = g^{-1}zg$. By Lemma 2.1.10, $Z(H)$ char H, and so $Z(H) \lhd H$. Thus $h = g^{-1}zg \in Z(H)$. We conclude that $Z(H) = H$ and H is abelian. By Lemma 2.3.8, H is an elementary abelian p-group. □

3.3 Multiple Transitivity and Sharp Transitivity

In this section, we describe multiple transitivity and sharply k-transitive group actions.

Definition 3.3.1. Let G act on X, where $|X| = n$ and let $k \leq n$. Then G acts *k-transitively* if for any two k-tuples of distinct points of X, $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$, there exists a $g \in G$ such that $x_i g = y_i$ for $1 \leq i \leq k$. If G acts k-transitively, we say G is *k-transitive*.

Lemma 3.3.2. *If G acts on X, then the following are equivalent:*

(1) G acts k-transitively on X.

(2) There exist distinct $x_1, \ldots, x_k \in X$, where for all ordered subsets of size k in X, such as $\{y_1, \ldots, y_k\}$, there is a $g \in G$ such that $y_i g = x_i$, $1 \leq i \leq k$.

(3) There exist distinct $x_1, \ldots, x_k \in X$, where for all ordered subsets of size k in X, such as $\{y_1, \ldots, y_k\}$, there is a $g \in G$ such that $x_i g = y_i$, $1 \leq i \leq k$.

Proof. Apply Lemma 3.2.4 to sets of distinct k-tuples. □

Lemma 3.3.3. *Let G act transitively on X, where $|X| = n$, and let $2 \leq k \leq n$. Then the following are equivalent:*

(1) G acts k-transitively on X

(2) For all $x \in X$, $\mathrm{Stab}(x)$ acts $(k-1)$-transitively on $X \setminus \{x\}$.

(3) There is an $x \in X$ such that $\mathrm{Stab}(x)$ acts $(k-1)$-transitively on $X \setminus \{x\}$.

Proof. Assume G acts k-transitively. Let $x \in X$ and let $(a_1, \ldots, a_{k-1})$ and $(b_1, \ldots, b_{k-1})$ be two $(k-1)$-tuples of $X \setminus \{x\}$, each with distinct entries. There exists some $g \in G$ such that

$$a_i g = b_i \tag{3.29}$$

for $1 \leq i \leq k-1$, and also $xg = x$. Thus $g \in \mathrm{Stab}(x)$, and $\mathrm{Stab}(x)$ acts $(k-1)$-transitively on $X \setminus \{x\}$.

The second condition implies the third condition *a fortiori*.

Assume that $\mathrm{Stab}(x_1)$ acts $(k-1)$-transitively on $X \setminus \{x_1\}$. Let $(a_1, \ldots, a_k)$ be a k-tuple of X with distinct entries, and choose $x_2, \ldots, x_k \in X \setminus \{x_1\}$. Since G acts transitively on X, there exists a $g_1 \in G$ such that $a_1 g_1 = x_1$. Thus

$$(a_1, \ldots, a_k)g_1 = (x_1, b_2 \ldots, b_k), \tag{3.30}$$

where $b_i = a_i g_1$ for $2 \leq i \leq k$. By assumption, there is a $g_2 \in \mathrm{Stab}(x_1)$ such that $b_i g_2 = x_i$ for $2 \leq i \leq k$. Thus

$$(x_1, b_2 \ldots, b_k) g_2 = (x_1, \ldots, x_k). \tag{3.31}$$

Thus if $g = g_1 g_2$, we have

$$(a_1, \ldots, a_k) g = (x_1, \ldots, x_k), \tag{3.32}$$

and G acts k-transitively on X, by Lemma 3.2.4. □

Theorem 3.3.4. *If G acts k-transitively on X, where $|X| = n$, then*

$$|G| = n(n-1)(n-2)\cdots(n-k+1)|\,\mathrm{Stab}(x_1, \ldots, x_k)| \tag{3.33}$$

for every choice of k distinct points $x_1, \ldots, x_k \in X$.

Proof. If $x_1 \in X$ then

$$|G| = n|\,\mathrm{Stab}(x_1)|. \tag{3.34}$$

Since $\mathrm{Stab}(x_1)$ acts $(k-1)$-transitively on $X \setminus \{x_1\}$, we have

$$|\,\mathrm{Stab}(x_1| = (n-1)|\,\mathrm{Stab}(x_1, x_2)|. \tag{3.35}$$

Since $\mathrm{Stab}(x_1, x_2)$ acts $(k-2)$-transitively on $X \setminus \{x_1, x_2\}$,

$$|\,\mathrm{Stab}(x_1, x_2)| = (n-2)|\,\mathrm{Stab}(x_1, x_2, x_3)|. \tag{3.36}$$

We proceed in a like manner until we have

$$|\,\mathrm{Stab}(x_1, \ldots, x_{k-1})| = (n-(k-1))|\,\mathrm{Stab}(x_1, \ldots, x_k)|. \tag{3.37}$$

Substitution yields our desired result. □

Definition 3.3.5. If G acts k-transitively on X, then it acts *sharply* k-transitively if only the identity fixes k distinct points of X, i.e, $\mathrm{Stab}(x_1, \ldots, x_k) = 1$, where $x_1, \ldots, x_k$ are k distinct points of X.

Lemma 3.3.6. *If G acts sharply k-transitively on X, where $|X| = n$ and $k \leq n$, then G acts faithfully on X.*

Proof. Let $a \in G$ fix all points of X. Since $k \leq n$, a also fixes k distinct points in X. Since X is sharply k-transitive, this implies $a = 1$. Thus G acts faithfully on X. □

Theorem 3.3.7. *If G acts faithfully and k-transitively on X, for $|X| = n$ and $k \geq 1$, then the following are equivalent:*

(1) G acts sharply k-transitively on X.

(2) If $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ are k-tuples of distinct points in X, then there is a unique $g \in G$ such that $a_i g = b_i$, for $1 \leq i \leq k$.

(3) $|G| = n(n-1)\cdots(n-k+1)$.

(4) The pointwise stabilizer of any k points in X is trivial.

(5) If $k \geq 2$, then the above conditions are equivalent to: For every $x \in X$, $\mathrm{Stab}(x)$ acts sharply $(k-1)$-transitively on $X \setminus \{x\}$.

Proof. Let G act faithfully and k-transitively on X, where $|X| = n$. Assume the first statement. Since G acts k-transitively, there exists a $g_1 \in G$ such that $a_i g_1 = b_i$ for all i. If $g_2 \in G$ satisfies these same k equations, then $a_i = b_i g_2^{-1}$, and so

$$a_i(g_1 g_2^{-1}) = (a_i g_1) g_2^{-1} = b_i g_2^{-1} = a_i \tag{3.38}$$

for $1 \leq i \leq k$. But only $1 \in G$ fixes k distinct points of X. Thus $g_1 g_2^{-1} = 1$ and $g_1 = g_2$. Thus $g \in G$ is unique and the second statement follows.

Assume the second statement. From Theorem 3.3.4 we have

$$|G| = n(n-1)(n-2)\cdots(n-k+1)|\operatorname{Stab}(x_1,\ldots,x_k)|. \tag{3.39}$$

So $\operatorname{Stab}(x_1,\ldots,x_k)$ is the subgroup of G fixing k distinct points of X. Since $x_i 1 = x_i$ for $1 \leq i \leq k$, $1 \in G$ is the unique element sending $(x_1,\ldots,x_k)$ to itself. Therefore $|\operatorname{Stab}(x_1,\ldots,x_k)| = 1$, and the third statement follows.

Assume the third statement. From Theorem 3.3.4 we derive the equations

$$n(n-1)(n-2)\cdots(n-k+1)|\operatorname{Stab}(x_1,\ldots,x_k)| = |G| = n(n-1)\cdots(n-k+1). \tag{3.40}$$

Thus $|\operatorname{Stab}(x_1,\ldots,x_k)| = 1$, and since $1 \in \operatorname{Stab}(x_1,\ldots,x_k)$, the fourth statement follows.

Assume the fourth statement. The stabilizer of any k-tuple of distinct points in X being trivial is equivalent to only $1 \in G$ fixing k distinct points of X. Thus G acts sharply k-transitively, the first statement.

Now let $k \geq 2$, and assume the first statement. From Lemma 3.3.3, we know that for a fixed $x \in X$, $\operatorname{Stab}(x)$ acts $(k-1)$-transitively on $X \setminus \{x\}$. Let $x \in X$ and let $(a_1,\ldots,a_{k-1})$ be a $(k-1)$-tuple of distinct points of $X \setminus \{x\}$. By assumption, $1 \in G$ is the unique element fixing $(a_1,\ldots,a_{k-1},x)$. Thus 1 is the only element of $\operatorname{Stab}(x)$ fixing $(a_1,\ldots,a_{k-1})$. Thus $\operatorname{Stab}(x)$ acts sharply $(k-1)$-transitively on $X \setminus \{x\}$.

Now suppose the fifth statement. By Lemma 3.3.3, G acts k-transitively on X, and we must show this is a sharp action. Let $(a_1,\ldots,a_k)$ be a k-tuple of distinct points of X. Let $g \in G$ fix this k-tuple. Thus $g \in \operatorname{Stab}(a_i)$ for $1 \leq i \leq k$. But $\operatorname{Stab}(a_i)$ has a sharp $(k-1)$-transitive action on $X \setminus \{a_i\}$, and thus $1 \in \operatorname{Stab}(a_i)$ is the unique element of $\operatorname{Stab}(a_i)$ fixing $(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_k)$. Thus $g = 1$. Since 1 is the only element in G fixing our k-tuple, G acts sharply k-transitively on X. $\square$

3.4 Primitive Actions

In this section, we examine primitive group actions, which, informally, occupy a space between 1-transitive and 2-transitive actions. These results will be used in the final section of this chapter.

Definition 3.4.1. If G acts on X, then a *block* is a subset B of X such that, for each $g \in G$, either $Bg = B$ or $Bg \cap B = \varnothing$ (where $Bg = \{xg \mid x \in B\}$). The *trivial* blocks are $\varnothing$, X, and one-point subsets.

Definition 3.4.2. If G acts transitively on X, then G acts *primitively* on X if all blocks of X are trivial. In such a case, we say X is a *primitive* G-set. Otherwise, we say G acts *imprimitively.*

Lemma 3.4.3. *Let G act transitively on X. Then G acts primitively if and only if* $\mathrm{Stab}(x)$ *is a maximal subgroup for each $x \in X$.*

Proof. Assume $\mathrm{Stab}(x)$ is not maximal for some $x \in X$. Thus there exists $H \leq G$ such that $\mathrm{Stab}(x) < H < G$. To show imprimitivity, we must show xH is a nontrivial block. If $g \in G$ and $(xH)g \cap xH = \varnothing$, we are done. So assume there is a $g \in G$ where $(xH)g \cap xH \neq \varnothing$. Thus there are $h_1, h_2 \in H$ such that $xh_1g = xh_2$. This implies $x(h_1gh_2^{-1}) = x$. Thus $h_1gh_2^{-1} \in \mathrm{Stab}(x) < H$, and so $g \in H$. We have

$$(xH)g = x(Hg) = xH, \tag{3.41}$$

and $xH \subseteq X$ is a block. Since H acts on X, there is a bijection between the cosets of $\mathrm{Stab}(x)$ in H and the points in xH, by Theorem 3.2.20. Now it follows from our assumption that $\mathrm{Stab}(x)$ has more than 1 coset in H, and H has more than 1 coset in G. Therefore, xH is nontrivial and G acts imprimitively.

Now assume $\mathrm{Stab}(x)$ is a maximal subgroup for each $x \in X$. Using proof by contradiction, we assume there exists a nontrivial block B. Let

$H = \{h \in G \mid Bh = B\}$, the setwise stabilizer of B. Let $x \in B$, and let $g \in \text{Stab}(x)$. Thus $xg = x$, and so $x \in Bg \cap B$. Since B is a block, $Bg = B$. Thus $g \in H$, and $\text{Stab}(x) \leq H$. Since B is not trivial, there is some $y \in B$ such that $x \neq y$, and since G acts transitively, there is some $g \in G$ such that $xg = y$. Thus $y \in Bg \cap B$, and since B is a block, $Bg = B$. Thus $g \in H$ and $g \notin \text{Stab}(x)$. Thus $\text{Stab}(x) < H$. Now B is nontrivial and so $B \neq X$. If $H = G$, then $Bg = B$ for all $g \in G$. But this contradicts G acting transitively on X. Thus $\text{Stab}(x) < H < G$, and we contradict the maximality of $\text{Stab}(x)$. Therefore, all blocks B are trivial and G acts primitively. □

Lemma 3.4.4. *If G acts 2-transitively on X, then G acts primitively.*

Proof. Using proof by contradiction, assume G acts imprimitively, that is, X has a nontrivial block B. Thus there are distinct $x, y, z \in X$ with $x, y \in B$ and $z \notin B$. Since G acts 2-transitively, there exists a $g \in G$ such that $xg = x$ and $yg = z$. Thus $x \in Bg \cap B$ and so $Bg \cap B \neq \varnothing$. We also have $z \in Bg$ and so $Bg \neq B$. But this contradicts the definition of a block, and thus B must have been trivial. □

3.5 Simplicity of Multiply-Transitive Groups

In this section, we develop our desired simplicity criterion (Theorem 3.5.11).

Lemma 3.5.1. *If G acts primitively on a set X and $H \lhd G$ is not contained in the kernel of the action, then H acts transitively on X.*

Proof. Now H also acts on X, and so it partitions X into orbits xH. Since $H \lhd G$, if $x \in X$ is fixed and $g \in G$, then

$$(xH)g = x(Hg) = x(gH) = (xg)H. \tag{3.42}$$

Thus orbit xH is taken to orbit $(xg)H$ by $g \in G$. Since these orbits are a partition of X, for all $x \in X$ and all $g \in G$, then either $(xH)g = xH$ or $(xH)g \cap xH = \varnothing$. Thus the H-orbits are blocks of X. Since G acts primitively on X, these orbits are either X or single points of X. Since H is not contained in the kernel, there is an $h \in H$ such that $yh \neq y$ for some $y \in X$, and so $yH \neq \{y\}$. Thus $yH = X$ for some $y \in X$, and H acts transitively on X, by Lemma 3.2.6. □

Corollary 3.5.2. *If G acts primitively and faithfully on a set X and $H \lhd G$ such that $H \neq 1$, then H acts transitively on X.* □

Lemma 3.5.3. *If G acts on X and $H \leq G$ acts transitively on X, then $G = \mathrm{Stab}(x)H$.*

Proof. Fix $x \in X$, and let $g \in G$. Thus there exists $h \in H$ such that $xh = y = xg$. Thus

$$x = x1 = x(hh^{-1}) = (xh)h^{-1} = (xg)h^{-1} = x(gh^{-1}) \tag{3.43}$$

and so $gh^{-1} \in \mathrm{Stab}(x)$. Thus

$$g = g1 = g(h^{-1}h) = (gh^{-1})h \in \mathrm{Stab}(x)H. \tag{3.44}$$

Thus $G \subseteq \mathrm{Stab}(x)H$. The reverse containment is trivial, and so $G = \mathrm{Stab}(x)H$. □

Definition 3.5.4. If G acts sharply 1-transitively on X, we say X is a *regular G-set*.

Definition 3.5.5. Let G act on X and $H \lhd G$. If X is a regular H-set, then H is a *regular normal subgroup.*

Theorem 3.5.6. *Let G act 2-transitively and faithfully on X, with $\mathrm{Stab}(x)$ a simple group for some $x \in X$. If $H \lhd G$ for $1 < H < G$, then X is a regular H-set, that is, H is a regular normal subgroup.*

Proof. Let $H \lhd G$ and $1 < H < G$. By Lemma 3.4.4, G acts primitively on X, and thus by Corollary 3.5.2, H acts transitively on X. Fix $x \in X$, where $\mathrm{Stab}(x)$ is a simple group. Now $\mathrm{Stab}(x) \cap H \lhd \mathrm{Stab}(x)$ by the second isomorphism theorem. Since $\mathrm{Stab}(x)$ is simple, $\mathrm{Stab}(x) \cap H = \mathrm{Stab}(x)$ or $\mathrm{Stab}(x) \cap H = 1$.

If $H \cap \mathrm{Stab}(x) = \mathrm{Stab}(x)$, then $\mathrm{Stab}(x) \leq H < G$. By Lemmas 3.4.3 and 3.4.4, $\mathrm{Stab}(x)$ is a maximal subgroup of G, and so $\mathrm{Stab}(x) = H$. But since $xg = x$ for all $g \in \mathrm{Stab}(x)$, $\mathrm{Stab}(x)$ is not transitive on X, contradicting the transitivity of H on X.

So $H \cap \mathrm{Stab}(x) = 1$, and thus 1 is the only element of H fixing $x \in X$. By definition, H acts sharply transitively on X. □

Definition 3.5.7. If G acts on X and Y, then a function $\varphi : X \to Y$ is a *G-map* if $\varphi(xg) = \varphi(x)g$ for all $x \in X$ and $g \in G$. If φ is also a bijection, then φ is called a *G-isomorphism*. Two sets X and Y are *G-isomorphic*, denoted $X \cong Y$, if there exists a G-isomorphism $\varphi : X \to Y$.

Definition 3.5.8. If G is a group, then $G^{\#} = G \setminus \{1\}$.

Lemma 3.5.9. *Let G act transitively on X and let H be a regular normal subgroup of G. Choose $x \in X$ and let $\mathrm{Stab}(x)$ act on $H^{\#}$ by conjugation. Then the $\mathrm{Stab}(x)$-sets $H^{\#}$ and $X \setminus \{x\}$ are $\mathrm{Stab}(x)$-isomorphic.*

Proof. Fix $x \in X$ and define

$$\varphi : H^{\#} \to X \setminus \{x\} \qquad \text{such that} \qquad \varphi(h) = xh. \tag{3.45}$$

As a regular normal subgroup, H acts sharply transitively on X, and so only $1 \in H$ fixes a given point. Thus $xh \neq x$ for all $h \in H^{\#}$, and φ is well-defined. Let $\varphi(h_1) = \varphi(h_2)$. Now $xh_1 = xh_2$ implies $xh_1h_2^{-1} = x$, and so $h_1h_2^{-1} \in H_x = 1$. This

gives $h_1 h_2^{-1} = 1$, or $h_1 = h_2$, making φ one-to-one. Now $|H| = |X|$, by Theorem 3.3.7, and so $|H^{\#}| = |X \setminus \{x\}|$. Since these are finite sets, φ is onto. Thus φ is a bijection.

We also have, for $g \in \mathrm{Stab}(x)$ and $h \in H^{\#}$,

$$\varphi(g^{-1}hg) = x(g^{-1}hg) = (xg^{-1})(hg) = x(hg) = (xh)g = (\varphi(h))g, \tag{3.46}$$

because $g^{-1} \in \mathrm{Stab}(x)$. Therefore, φ is a $\mathrm{Stab}(x)$-isomorphism. □

Theorem 3.5.10. *Let $k \geq 2$, and let G act faithfully and k-transitively on X, where $|X| = n \geq k$. If G has a regular normal subgroup H, then $k \leq 4$. Moreover:*

(1) H is an elementary abelian p-group for some prime p, and n is a power of p.

(2) If $k \geq 3$, then either $H \cong \mathbb{Z}_3$ and $n = 3$, or H is an elementary abelian 2-group and n is a power of 2.

(3) If $k = 4$, then $H \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ and $n = 4$.

Proof. Let H be a regular normal subgroup of G. By Lemma 3.3.3, for fixed $x \in X$, $\mathrm{Stab}(x)$ acts $(k-1)$-transitively on $X \setminus \{x\}$. Thus by Lemma 3.5.9, $\mathrm{Stab}(x)$ acts $(k-1)$-transitively on $H^{\#}$ by conjugation.

(1) Since $k \geq 2$, $\mathrm{Stab}(x)$ acts transitively on $H^{\#}$, and all elements of $H^{\#}$ are conjugate. By Lemma 3.2.25, H is an elementary abelian p-group, of order p^r for some $r \in \mathbb{Z}^+$. Now $|H| = |X|$, by Theorem 3.3.7(iii), and so $n = |X| = |H| = p^r$.

(2) Let $h \in H^{\#}$ and $B = \{h, h^{-1}\}$. Now G acts on $H^{\#}$ by conjugation and if $g \in G$, we have

$$g^{-1}Bg = g^{-1}\{h, h^{-1}\}g = \{g^{-1}hg, (g^{-1}hg)^{-1}\} = \{h', h'^{-1}\}. \tag{3.47}$$

Now $h', h'^{-1} \in H^{\#}$, since $H \lhd G$ and because $h \neq 1$ implies $g^{-1}hg \neq 1$. Since $g^{-1}Bg = B$ or $g^{-1}Bg \cap B = \varnothing$, B is a block in $H^{\#}$.

Since $k \geq 3$, $\mathrm{Stab}(x)$ acts 2-transitively on $H^{\#}$, and $H^{\#}$ is a primitive $\mathrm{Stab}(x)$-set. Since blocks of $H^{\#}$ are trivial, either $\{h, h^{-1}\} = H^{\#}$ or $\{h, h^{-1}\} = \{h\}$. In the first case, $|H| = 3$, $H \cong \mathbb{Z}_3$ and $n = 3$. In the second case, h is an involution, and so the prime p in part (1) must be 2. Thus H is an elementary abelian 2-group, and $n = 2^r$ for some $r \in \mathbb{Z}^+$.

(3) Let $k \geq 4$. Thus $|X| \geq 4$, and since $|X \setminus \{x\}| = |H^{\#}|$, we have $|H^{\#}| \geq 3$. Thus $H \neq \mathbb{Z}_3$, and we are in the second case of part (2), where H is an elementary abelian 2-group. Since $H \neq \mathbb{Z}_2$, it must contain a copy of $\mathbb{Z}_2 \times \mathbb{Z}_2$, say $\{1, h, k, hk\}$. Now the stabilizer subgroup

$$\mathrm{Stab}(x, h) = \{g \in G \mid xg = x \text{ and } g^{-1}hg = h\} \tag{3.48}$$

acts 2-transitively on $H^{\#} \setminus \{h\}$, and hence primitively. Consider $B = \{k, hk\}$ and $g \in \mathrm{Stab}(x, h)$. If $g^{-1}kg = k$, then

$$g^{-1}hkg = (g^{-1}hg)(g^{-1}kg) = hk \tag{3.49}$$

and $g^{-1}Bg = B$. If $g^{-1}kg = hk$, then

$$g^{-1}hkg = (g^{-1}hg)(g^{-1}kg) = h(hk) = h^2k = k. \tag{3.50}$$

and $g^{-1}Bg = B$. Similarly, if conjugation sends neither point to the other, then $g^{-1}Bg \cap B = \varnothing$. Thus $B = \{k, hk\}$ is a block in $H^{\#} \setminus \{h\}$. But since B is trivial, we must have $H^{\#} \setminus \{h\} = \{k, hk\}$. As a result, $H = \{1, h, k, hk\} = \mathbb{Z}_2 \times \mathbb{Z}_2$ and $n = 4$.

We finally have our first assertion because $k \leq n \leq 4$. □

Theorem 3.5.11. *Let G act faithfully and k-transitively on X, $k \geq 2$, and assume that* $\mathrm{Stab}(x)$ *is simple for some $x \in X$.*

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

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

(3) If $k \geq 2$ and $|X|$ is not a prime power, then G is simple.

Proof. Assume G has a proper nontrivial normal subgroup H. By Theorems 3.5.6 and 3.5.10, H is a regular normal subgroup and $k \leq 4$.

(1) If $k \geq 4$, then $k = 4$ and by Theorem 3.5.10, $H \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ and $|X| = 4$. Now $\rho(G) \leq S_4$ under the homomorphism of the action. The only 4-transitive subgroup of S_4 is S_4, and the stabilizer of a point in S_4 is S_3. But S_3 is not simple since it has A_3 as a normal subgroup, which contradicts the simplicity of $\mathrm{Stab}(x)$. Thus no such H exists, and G is simple.

(2) If $k \geq 3$ and $|X|$ is not a power of 2, then by Theorem 3.5.10, we must have $H \cong \mathbb{Z}_3$ and $n = 3$. Thus $\rho(G) \leq S_3$. The only 3-transitive subgroup of S_3 is S_3, and the stabilizer of a point in S_3 is S_2, with $S_2 \cong \mathbb{Z}_2$ simple. Thus either $G \cong S_3$ or G is simple.

(3) If $k \geq 2$ and $|X|$ is not a prime power, then a contradiction is immediate, since Theorem 3.5.10 gives n as a prime power. Thus no such H exists, and G is simple. □

CHAPTER 4

LINEAR GROUPS AND THE SIMPLICITY OF $PSL_3(\mathbb{F}_4)$

In this chapter, we establish that if $n \geq 3$, or $n = 2$ and $|F| > 3$, then $PSL_n(F)$ is a simple group. We use this fact in Chapter 11, as $PSL_3(\mathbb{F}_4)$ is a subgroup of the large Mathieu groups, and its simplicity is used to establish theirs. We follow the treatment of Jacobson [Jac85].

4.1 Linear Groups

In this section, we define linear groups, and make use of the identification between matrices and linear transformations (given a fixed base). We show that the special linear group is generated by elementary matrices, the transvections. We establish some facts about the commutator subgroups of linear groups, and in preparation for defining projective linear groups, we describe the center subgroup.

Definition 4.1.1. The *general linear group*, denoted $GL_n(F)$, is the group of bijective linear transformations of an n-dimensional vector space V over a field F. Recall that, given the correspondence between linear transformations and $n \times n$ matrices, relative to a fixed basis for V, we can identify $GL_n(F)$ with the group of invertible $n \times n$ matrices (Friedberg, Insel, Spence [FIS03] and Section 2.6).

Definition 4.1.2. We let $\mathbb{F}_q$ denote the unique finite field of order q.

Definition 4.1.3. The *special linear group*, denoted $SL_n(F)$, is the subgroup of $GL_n(F)$ whose matrices have determinant 1. This group can also be defined as the

kernel of the determinant homomorphism, det: $GL_n(F) \to F^\times$, where $F^\times$ is the multiplicative group $F \setminus \{0\}$.

Lemma 4.1.4. *We have $SL_n(F) \lhd GL_n(F)$.*

Proof. Since $SL_n(F)$ is the kernel of a homomorphism of $GL_n(F)$, this follows from the fundamental homomorphism theorem. □

Definition 4.1.5. We define E_{ij} to be the $n \times n$ matrix with an entry of 1 in the ij-position, and 0 elsewhere. By inspection, we see $E_{ik}E_{kj} = E_{ij}$ and $E_{ik}E_{lj} = 0$ for $k \neq l$.

Definition 4.1.6. A *transvection* is a matrix $T_{ij}(b) = I + bE_{ij}$ such that $b \in F^\times$ and $i \neq j$. The matrix $AT_{ij}(b)$ changes A by adding b times the ith column to the jth column. The matrix $T_{ij}(b)A$ changes A by adding b times the jth row to the ith row.

Lemma 4.1.7. *For all $b \in F^\times$ and $i \neq j$, $T_{ij}(b) \in SL_n(F)$ and $T_{ij}(b)^{-1} = T_{ij}(-b)$.*

Proof. Since $T_{ij}(b)$ is an upper or lower triangular matrix, its determinant is the product of the diagonal entries, which are all 1. Thus $T_{ij}(b) \in SL_n(F)$. Now $-b \in F$, and so for $i \neq j$, we have

$$T_{ij}(b)T_{ij}(-b) = (I + bE_{ij})(I - bE_{ij}) = I + bE_{ij} - bE_{ij} + 0 = I, \tag{4.1}$$

and a similar result holds for $T_{ij}(-b)T_{ij}(b)$. Thus $T_{ij}(b)^{-1} = T_{ij}(-b)$. □

Definition 4.1.8. We define $P_{ij} = I - E_{ij} + E_{ji} - E_{ii} - E_{jj}$, and call them *elementary signed permutation matrices*. For example, if $n = 2$, then $P_{21} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. The matrix AP_{ij} interchanges the negated ith column of A with the jth column of A. The matrix $P_{ij}A$ interchanges the negated jth row of A with the ith row of A.

Lemma 4.1.9. *A P_{ij} matrix is invertible, and $P_{ij}^{-1} = P_{ji}$.*

Proof. Let I be the usual identity matrix, and assume without loss of generality that $1 \leq i < j \leq n$. We have

$$P_{ij}P_{ji} = [e_1 \dots e_j \dots -e_i \dots e_n]P_{ji} = [e_1 \dots -(-e_i) \dots e_j \dots e_n] = I \tag{4.2}$$

and

$$P_{ji}P_{ij} = [e_1 \dots -e_j \dots e_i \dots e_n]P_{ji} = [e_1 \dots e_i \dots -(-e_j) \dots e_n] = I. \tag{4.3}$$

□

Lemma 4.1.10. *Each P_{ij} is a product of transvections, and is thus in $SL_n(F)$.*

Proof. We have

$$\begin{aligned} T_{ji}(1)T_{ij}(-1)T_{ji}(1) &= (I + E_{ji})(I - E_{ij})(I + E_{ji}) \\ &= (I + E_{ji} - E_{ij} - E_{jj})(I + E_{ji}) \\ &= I + E_{ji} - E_{ij} - E_{jj} + E_{ji} + 0 - E_{ii} - E_{ji} \\ &= I - E_{ij} + E_{ji} - E_{ii} - E_{jj} \\ &= P_{ij}. \end{aligned} \tag{4.4}$$

The second statement follows from Lemma 4.1.7, and the fact that the determinant is multiplicative (Friedberg, Insel, and Spence [FIS03]). □

Definition 4.1.11. We define $U_i(d) = I + (d^{-1} - 1)E_{ii} + (d - 1)E_{(i+1)(i+1)}$, where $d \in F^\times$ and $1 \leq i \leq n - 1$. For example, if $n = 2$, then $U_1(d) = \begin{bmatrix} d^{-1} & 0 \\ 0 & d \end{bmatrix}$.

Lemma 4.1.12. *For $n = 2$, $U_1(d)$ is a product of transvections. The same result holds for $U_i(d)$, where $n > 2$.*

Proof. We have

$$\left(\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ -1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\right)\left(\begin{bmatrix}1 & -d\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ d^{-1} & 1\end{bmatrix}\begin{bmatrix}1 & -d\\ 0 & 1\end{bmatrix}\right)$$
$$= \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}0 & -d\\ d^{-1} & 0\end{bmatrix} = \begin{bmatrix}d^{-1} & 0\\ 0 & d\end{bmatrix}. \quad (4.5)$$

The same result holds if each of the above transvections are identically-placed blocks on the diagonals of $n \times n$ identity matrices. □

Theorem 4.1.13. *$SL_n(F)$ is generated by the matrices $T_{ij}(b)$, where $b \in F^\times$ and $i \neq j$.*

Proof. We must show that if $A \in SL_n(F)$, then there exist matrices P, Q such that P, Q are the products of transvections and $PAQ = I$ (as then, $A = P^{-1}Q^{-1}$). Let $A \in SL_n(F)$. By Gaussian elimination, we multiply A with $T_{ij}(b)$ and P_{ij} matrices to row reduce it, and may now assume $A = \operatorname{diag}\{d_1, d_2, \ldots, d_n\} \in SL_n(F)$.

Since $\det A = 1$, we have $(d_1 d_2 \cdots d_n) = 1$. Thus each d_i is invertible. If we right-multiply our diagonal matrix by $U_1(d_1^{-1})$, we get $\operatorname{diag}\{1, (d_1 d_2), d_3, \ldots, d_n\}$. If we right-multiply this by $U_2((d_1 d_2)^{-1})$, we get $\operatorname{diag}\{1, 1, (d_1 d_2 d_3), d_4, \ldots, d_n\}$. Continuing this way, we obtain $\operatorname{diag}\{1, \ldots, 1, (d_1 d_2 \cdots d_n)\} = I$. When we multiply the transvections we have used, we obtain $PAQ = I$. □

Theorem 4.1.14. *If $n \geq 3$, or if $n = 2$ and $|F| > 3$, then $SL_n(F)$ is identical with its commutator subgroup $SL_n(F)'$, for $n \geq 2$, where the commutator subgroup $G' = \langle a^{-1}b^{-1}ab \mid a, b \in G\rangle \leq G$ is as in Definition 2.1.11.*

Proof. Let $G = SL_n(F)$. If we can show that the generators $T_{ij}(b)$ are contained in the commutator subgroup G', we are done. If $n \geq 3$, then let $i \neq j$ and $k \neq i, j$. We

have:

$$\begin{aligned} T_{ik}(b)T_{kj}(1)T_{ik}(b)^{-1}T_{kj}(1)^{-1} &= T_{ik}(b)T_{kj}(1)T_{ik}(-b)T_{kj}(-1) \\ &= (I+bE_{ik})(I+E_{kj})(I-bE_{ik})(I-E_{kj}) \\ &= (I+bE_{ik}+E_{kj}+bE_{ij})(I-bE_{ik}-E_{kj}+bE_{ij}) \\ &= I+bE_{ik}+E_{kj}+bE_{ij}-bE_{ik}+0+0+0 \\ &\quad -E_{kj}-bE_{ij}+0+0+bE_{ij}+0+0+0 \\ &= I+bE_{ij} \\ &= T_{ij}(b), \end{aligned} \tag{4.6}$$

and thus $T_{ij}(b) \in G'$.

Let $n = 2$. Now $U_1(d)^{-1} = \begin{bmatrix} d^{-1} & 0 \\ 0 & d \end{bmatrix}^{-1} = \begin{bmatrix} d & 0 \\ 0 & d^{-1} \end{bmatrix} = U_1(d^{-1})$, and we have:

$$\begin{bmatrix} d & 0 \\ 0 & d^{-1} \end{bmatrix} \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d^{-1} & 0 \\ 0 & d \end{bmatrix} \begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & c(d^2-1) \\ 0 & 1 \end{bmatrix} \in G'. \tag{4.7}$$

Let $b \in F^\times$, which is a cyclic multiplicative group. If $|F| > 3$, then $|F^\times| > 2$, and if $\langle d \rangle = F^\times$, then $d \neq 1$ and $d^2 \neq 1$. Thus $d^2 - 1 \neq 0$ and $(d^2-1)^{-1} \in F^\times$, as is $c = b(d^2-1)^{-1}$. Referring to our matrix above, we see $T_{12}(b) \in G'$, and a similar result holds for $T_{21}(b)$. □

Lemma 4.1.15. *If $n \geq 3$, or if $n = 2$ and $|F| > 3$, then $SL_n(F) = GL_n(F)'$.*

Proof. Let $\det : GL_n(F) \to F^\times$ be the determinant homomorphism. If $A = \operatorname{diag}\{\alpha, 1, \ldots, 1\}$ for $\alpha \in F^\times$, then $\det A = \alpha$ and $A \in GL_n(F)$. Thus the homomorphism is onto. By the first isomorphism theorem,

$$GL_n(F)/\ker(\det) = GL_n(F)/SL_n(F) \cong F^\times. \tag{4.8}$$

Since $F^\times$ is abelian, so is the factor group, and thus by Lemma 2.1.12, $GL_n(F)' \leq SL_n(F)$. From Theorem 4.1.14, we have

$$SL_n(F) = SL_n(F)' \leq GL_n(F)', \tag{4.9}$$

and thus $SL_n(F) = GL_n(F)'$. □

Theorem 4.1.16. *We have*

$$Z(GL_n(F)) = \{\alpha I \mid \alpha \in F^\times\},$$

the set of scalar matrices. Also,

$$Z(SL_n(F)) = \{\alpha I \mid \alpha \in F^\times, \alpha^n = 1\} = Z(GL_n(F)) \cap SL_n(F).$$

Proof. Let $D \in GL_n(F)$ commute with all elements of $GL_n(F)$. Thus

$$DT_{ij}(1) = T_{ij}(1)D \tag{4.10}$$

for $1 \leq i, j \leq n$, where $i \neq j$. Now $DT_{ij}(1)$ is an altered D, such that D's ith column has been added to its jth column, and $T_{ij}(1)D$ is an altered D, such that D's jth row has been added to its ith row. By hypothesis, the two are identical.

The (ii)-entry of our altered matrix is

$$D_{ii} = D_{ii} + D_{ji}. \tag{4.11}$$

We conclude $D_{ji} = 0$. Since i and j were arbitrary, other than $i \neq j$, all off-diagonal entries of D are 0. The (ij)-entry of our altered matrix is

$$D_{ij} + D_{ii} = D_{ij} + D_{jj}. \tag{4.12}$$

We conclude $D_{ii} = D_{jj}$. Since i and j were arbitrary, other than $i \neq j$, all diagonal entries of D are identical. Since $D \in GL_n(F)$, $\det D \neq 0$. Now $\det D = (D_{ii})^n = 0$ if and only if $D_{ii} = 0$. Thus $D_{ii} = \alpha \in F^\times$. If $D \in SL_n(F)$, then $\det D = 1$, and so $(D_{ii})^n = \alpha^n = 1$. □

Definition 4.1.17. The *projective general linear group* is the factor group of the general linear group modulo its center. We have:

$$PGL_n(F) = GL_n(F)/Z(GL_n(F)). \tag{4.13}$$

Definition 4.1.18. The *projective special linear group* is the factor group of the special linear group modulo its center. We have:

$$PSL_n(F) = SL_n(F)/Z(SL_n(F)). \tag{4.14}$$

4.2 Action of Linear Groups on Projective Space

In this section, we define an action of $GL_n(F)$ (and its subgroups) on projective $(n-1)$-space. We show that projective linear groups act faithfully on projective space. We show that $SL_n(F)$ is 2-transitive, and prove a technical lemma (Lemma 4.2.9) for use in the final section of the chapter.

Definition 4.2.1. The *row vector space* $V = F^n$ is defined as $\{(a_1, \ldots, a_n) \mid a_i \in F\}$, the $1 \times n$ row vectors with entries from F, and we define $V^{\#} = F^n \setminus \{0\}$.

Lemma 4.2.2. *For $v, w \in V^{\#}$, define $v \sim w$ if there is an $\alpha \in F^{\times}$ such that $w = \alpha v$. Then $\sim$ is an equivalence relation on $V^{\#}$.*

Proof. We have $1 \in F^{\times}$ and $v = 1v$ for all $v \in F^{n\#}$, so $\sim$ is reflexive. Let $v \sim w$. So $w = \alpha v$ for some $\alpha \in F^{\times}$. Now $\alpha^{-1} \in F^{\times}$ and

$$v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1}w, \tag{4.15}$$

so $\sim$ is reflexive. Now let $v \sim w$ and $w \sim x$. So $w = \alpha v$ and $x = \beta w$ for some $\alpha, \beta \in F^{\times}$. Now $\beta\alpha \in F^{\times}$ and

$$x = \beta w = \beta(\alpha v) = (\beta\alpha)v, \tag{4.16}$$

so $\sim$ is transitive. □

Definition 4.2.3. Let $v \in V^{\#}$, and define $[v] = \{\alpha v \mid \alpha \in F^{\times}\}$, the equivalence class of v. We define *projective* $(n-1)$*-space* as $P^{n-1}(F) = \{[v] \mid v \in V^{\#}\}$.

Lemma 4.2.4. *$GL_n(F)$ acts on $P^{n-1}(F)$ by $[v]A = [vA]$ for $A \in GL_n(F)$ and $[v] \in P^{n-1}(F)$.*

Proof. Suppose $[v] = [w]$ and let $A \in GL_n(F)$. Then $v = \alpha w$ for $v, w \in V^{\#}$ and $\alpha \in F^{\times}$. Since $v, w \neq 0$ and $A \in GL_n(F)$, we have $vA, wA \neq 0$, and so $vA, wA \in V^{\#}$. Further,

$$vA = (\alpha w)A = \alpha(wA) \tag{4.17}$$

and so $[vA] = [wA]$. Finally,

(1) $[v]I = [vI] = [v]$ and

(2) $[v](AB) = [v(AB)] = [(vA)B] = [vA]B = ([v]A)B$

for all $v \in V^{\#}$ and all $A, B \in GL_n(F)$. □

Lemma 4.2.5. *The kernel of the action in Lemma 4.2.4 is the center of $GL_n(F)$, that is, the set of scalar matrices.*

Proof. Let $A \in Z(GL_n(F))$. Thus $A = \alpha I$ for $\alpha \in F^{\times}$, by Theorem 4.1.16. If $v \in V^{\#}$, we have

$$[v]A = [vA] = [v(\alpha I)] = [\alpha(vI)] = [\alpha v] = [v], \tag{4.18}$$

and A is in the kernel of the action.

Now let $(e_1, \ldots, e_n)$ be the standard basis for F^n, and let A be in the kernel of the action. Since $vA = \alpha v$ for $v \in V^{\#}$ and $\alpha \in F^{\times}$, denote the eigenvalue α with

respect to v as α_v. Since e_iA is the ith row of A, and $e_iA = \alpha_{e_i}e_i$, the ith row of A has $\alpha_{e_i} \in F^\times$ in the ith column, and 0 elsewhere. For $i \neq j$, we have that

$$\alpha_{(e_i+e_j)}e_i+\alpha_{(e_i+e_j)}e_j = \alpha_{(e_i+e_j)}(e_i+e_j) = (e_i+e_j)A = e_iA+e_jA = \alpha_{e_i}e_i+\alpha_{e_j}e_j. \quad (4.19)$$

Since $\{e_i, e_j\}$ is linearly independent, we conclude $\alpha_{e_i} = \alpha_{(e_i+e_j)} = \alpha_{e_j}$. Therefore, $A = \alpha I$ where $\alpha = \alpha_{e_i} = \alpha_{e_j} \neq 0$, and thus $A \in Z(GL_n(F))$. □

Corollary 4.2.6. *$PGL_n(F)$ acts faithfully on $P^{n-1}(F)$.*

Proof. By Lemma 4.2.5, the kernel of the action of $GL_n(F)$ is $Z(GL_n(F))$. By Corollary 3.2.18 and Definition 4.1.17, $PGL_n(F) = GL_n(F)/Z(GL_n(F))$ acts faithfully on $P^{n-1}(F)$. □

Corollary 4.2.7. *$PSL_n(F)$ acts faithfully on $P^{n-1}(F)$.*

Proof. Now $SL_n(F) \leq GL_n(F)$ acts on $P^{n-1}(F)$, where the action of $GL_n(F)$ is restricted to the elements of $SL_n(F)$. By Lemma 4.2.5, the kernel of the action of $GL_n(F)$ is $Z(GL_n(F))$. Since $SL_n(F)$ has the same action, the kernel of its action is $Z(GL_n(F)) \cap SL_n(F) = Z(SL_n(F))$. By Corollary 3.2.18 and Definition 4.1.18, $PSL_n(F) = SL_n(F)/Z(SL_n(F))$ acts faithfully on $P^{n-1}(F)$. □

Theorem 4.2.8. *For $n \geq 2$, $SL_n(F)$ acts 2-transitively on $P^{n-1}(F)$.*

Proof. Let $n \geq 2$. By Lemma 3.3.2, it suffices to show that for a distinct pair $[v_1], [v_2]$ in $P^{n-1}(F)$, there exists an $A \in SL_n(F)$ such that $[v_1]A = [e_1]$ and $[v_2]A = [e_2]$. Since $[v_1] \neq [v_2]$, $v_1 \neq \alpha v_2$ for all $\alpha \in F^\times$. Thus $\{v_1, v_2\}$ is linearly independent, and we can extend it to a basis $(v_1, v_2, \ldots, v_n)$ for F^n.

By linear algebra, there exists an $\widetilde{A} \in GL_n(F)$ such that $v_i\widetilde{A} = e_i$ for $1 \leq i \leq n$ [FIS03]. Now $\det \widetilde{A} = \alpha$ for some $\alpha \in F^\times$. Let $A_0 = \mathrm{diag}\{\alpha^{-1}, 1, \ldots, 1\}$.

Consider $A = \widetilde{A}A_0 \in GL_n(F)$. We have

$$\det A = \det \widetilde{A}A_0 = (\det \widetilde{A})(\det \operatorname{diag}\{\alpha^{-1}, 1, \ldots, 1\}) = \alpha\alpha^{-1} = 1, \tag{4.20}$$

implying $A \in SL_n(F)$. Since $v_1A = \alpha^{-1}e_1$ and $v_2A = e_2$, we have $[v_1]A = [e_1]$ and $[v_2]A = [e_2]$. Thus $SL_n(F)$ acts 2-transitively on $P^{n-1}(F)$. □

Lemma 4.2.9. *For $n \geq 2$, let $(e_1, \ldots, e_n)$ be the standard basis for F^n, and let $G = SL_n(F)$. Then $\operatorname{Stab}_G([e_1])$ contains a normal abelian subgroup $A(e_1)$ whose conjugates in G generate G.*

Proof. Let $T \in \operatorname{Stab}_G([e_1]) \leq G$. Now $[e_1]T = [e_1]$, and so $e_1T = \alpha e_1$ for some $\alpha \in F^\times$, implying the first row of T is αe_1. Thus the $(n-1) \times (n-1)$ submatrix formed by deleting the first row and first column of T is of determinant α^{-1}, and is in $GL_{n-1}(F)$. Define a map φ from $\operatorname{Stab}_G([e_1])$ to $GL_{n-1}(F)$ where $T \in \operatorname{Stab}_G([e_1])$ is mapped to such a submatrix. Thus

$$T = \begin{bmatrix} \alpha & 0 & \cdots & 0 \\ r_{21} & & & \\ \vdots & & \varphi(T) & \\ r_{n1} & & & \end{bmatrix}. \tag{4.21}$$

If $S, T \in \operatorname{Stab}_G([e_1])$, we see by inspection that $\varphi(ST) = \varphi(S)\varphi(T)$, and φ is a homomorphism.

If $K \in \ker \varphi$, then $\varphi(K) = I_{n-1}$, and K is of the form:

$$K = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ k_{21} & & & \\ \vdots & & I_{n-1} & \\ k_{n1} & & & \end{bmatrix}. \tag{4.22}$$

We define $A(e_1) = \ker \varphi$. If $K, L \in A(e_1)$, we have

$$\begin{aligned} KL &= \begin{bmatrix} 1 & \cdots & 0 \\ k_{21} & & \\ \vdots & & I_{n-1} \\ k_{n1} & & \end{bmatrix} \begin{bmatrix} 1 & \cdots & 0 \\ l_{21} & & \\ \vdots & & I_{n-1} \\ l_{n1} & & \end{bmatrix} \\ &= \begin{bmatrix} 1 & \cdots & 0 \\ k_{21}+l_{21} & & \\ \vdots & & I_{n-1} \\ k_{n1}+l_{n1} & & \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ l_{21}+k_{21} & & \\ \vdots & & I_{n-1} \\ l_{n1}+k_{n1} & & \end{bmatrix} \\ &= \begin{bmatrix} 1 & \cdots & 0 \\ l_{21} & & \\ \vdots & & I_{n-1} \\ l_{n1} & & \end{bmatrix} \begin{bmatrix} 1 & \cdots & 0 \\ k_{21} & & \\ \vdots & & I_{n-1} \\ k_{n1} & & \end{bmatrix} = LK, \end{aligned} \tag{4.23}$$

and $A(e_1)$ is an abelian subgroup of $\mathrm{Stab}_G([e_1])$.

It suffices to show all transvections are the products of the conjugations of elements of $A(e_1)$. In this case, since transvections generate G by Theorem 4.1.13, the conjugates of the elements of $A(e_1)$ generate G. Now the transvections $T_{i1}(b)$ are in $A(e_1)$ for $2 \leq i \leq n$. Since the P_{ij} matrices are in G, we have

$$\begin{aligned} P_{i1}T_{i1}(b)P_{i1}^{-1} &= (P_{i1}T_{i1}(b))P_{1i} = (P_{i1} + bE_{11})P_{1i} \\ &= P_{i1}P_{1i} + bE_{11}P_{1i} = I + (-b)E_{1i} = T_{1i}(-b) \end{aligned} \tag{4.24}$$

as a conjugate of $T_{i1}(b)$, for $2 \leq i \leq n$. So if $n = 2$, $T_{21}(b) \in A(e_1)$ and $T_{12}(-b)$ are conjugate, and we are done. If $n \geq 3$, we know from (4.6) that

$$T_{ij}(b) = T_{ik}(b)T_{kj}(1)T_{ik}^{-1}(b)T_{kj}^{-1}(1) \tag{4.25}$$

and thus

$$T_{ij}(b) = T_{i1}(b)T_{1j}(1)T_{i1}^{-1}(b)T_{1j}^{-1}(1) = T_{i1}(b)T_{1j}(1)T_{i1}(-b)T_{1j}(-1), \tag{4.26}$$

making $T_{ij}(b)$ the product of conjugates of elements of $A(e_1)$. □

4.3 Simplicity of $PSL_n(F)$

In this section, we develop a second simplicity criterion and use it to prove the simplicity of $PSL_n(F)$ (given minimal restrictions on n and the order of F). In particular, we conclude $PSL_3(\mathbb{F}_4)$ is simple. We also develop formulas for calculating the orders of various linear groups.

Theorem 4.3.1. *Let G act on X and let K be the kernel of the action. If*

(1) G acts primitively on X; and

(2) $G = G'$, the commutator subgroup of G; and

(3) There exists an $x \in X$ such that $\mathrm{Stab}(x)$ contains a normal abelian subgroup, $A(x)$, with the property that G is generated by the conjugates $g^{-1}A(x)g$, for $g \in G$;

then it follows that G/K is simple.

Proof. Let N be a nontrivial normal subgroup of G/K. By Lemma 2.2.6, there exists an $H \lhd G$ such that $H/K = N$, and $K < H$. Then by Lemma 3.5.1, H acts transitively on X. Let $x \in X$ satisfy Condition (3). By Lemmas 3.5.3 and 2.2.3, we know that

$$G = \mathrm{Stab}(x)H = H\,\mathrm{Stab}(x). \tag{4.27}$$

Now by Lemma 2.2.5, $HA(x) \lhd H\,\mathrm{Stab}(x) = G$, and so for $a \in A(X)$ and $g \in G$, $g^{-1}ag \in HA(x)$. Since all conjugations of the elements of $A(x)$ are in $HA(x)$, and

by Condition (3), G is generated by these conjugates, $G = HA(x) = A(x)H$. By the second isomorphism theorem,

$$G/H = A(x)H/H \cong A(x)/(A(x) \cap H), \tag{4.28}$$

and $A(x)/(A(x) \cap H)$ is abelian because $A(x)$ is. Hence by Lemma 2.1.12, $G = G' \leq H$. Thus $H = G$, and $N = G/K$. Thus G/K has no nontrivial proper normal subgroups, and so is simple. □

Theorem 4.3.2. *If $n \geq 3$, or if $n = 2$ and $|F| > 3$, then $PSL_n(F)$ is simple.*

Proof. Let $n \geq 2$. Now $SL_n(F)$ acts 2-transitively on $P^{n-1}(F)$, and the kernel of its action is $Z(SL_n(F))$, by Corollary 4.2.7 and Theorem 4.2.8. This implies $SL_n(F)$ acts primitively on $P^{n-1}(F)$, by Lemma 3.4.4. Also, except for $n = 2$ and $|F| = 2$ or 3, $SL_n(F) = SL_n(F)'$, by Theorem 4.1.14. We have $A(e_1) < \mathrm{Stab}(e_1) < SL_n(F)$, where $A(e_1)$ is the normal abelian subgroup whose conjugates generate $SL_n(F)$, by Lemma 4.2.9. We conclude that $SL_n(F)/Z(SL_n(F)) = PSL_n(F)$ is simple (minus the exceptions), by Theorem 4.3.1. □

Theorem 4.3.3. *Let $\mathbb{F}_q$ be a finite field. We have:*

- $|GL_n(\mathbb{F}_q)| = \prod_{i=0}^{n-1} q^n - q^i = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$.
- $|SL_n(\mathbb{F}_q)| = \left(\prod_{i=0}^{n-1} q^n - q^i\right)/(q-1) = (q^n - 1) \cdots (q^n - q^{n-2})q^{n-1}$.
- $|PGL_n(\mathbb{F}_q)| = \left(\prod_{i=0}^{n-1} q^n - q^i\right)/(q-1) = (q^n - 1) \cdots (q^n - q^{n-2})q^{n-1}$.
- $|PSL_n(\mathbb{F}_q)| = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-2})q^{n-1}/d$, *where* $d = \gcd(n, q-1)$.

Proof. An $n \times n$ matrix has q^n choices for each row. Consider an arbitrary matrix in $GL_n(\mathbb{F}_q)$. Such a matrix has n linearly independent rows. In particular, we cannot have a row of all 0's, and thus there are $q^n - 1$ choices for the first row. The second

row cannot be a scalar multiple of the first row, and since our field is of order q, there are q scalar multiples. Thus there are $q^n - q$ choices for the second row. The third row cannot be in the span of the first two, and there are q^2 ways this could happen. Thus there are $q^n - q^2$ choices for the second row. Continuing in this way, we reach the nth row, which cannot be in the span of the preceding $n-1$ rows. Thus there are $q^n - q^{n-1}$ choices for the nth row. Multiplying the available choices for each row, we arrive at the first equality.

By Definition 4.1.3, $SL_n(F) = \ker(\det)$, where $\det : GL_n(F) \to F^\times$ is the determinant homomorphism. Since det is onto, we have, by the fundamental homomorphism theorem,

$$GL_n(F)/SL_n(F) \cong F^\times. \tag{4.29}$$

Since $\mathbb{F}_q$ is finite,

$$|GL_n(\mathbb{F}_q)|/|SL_n(\mathbb{F}_q)| = |F^\times| = q - 1, \tag{4.30}$$

and so $|SL_n(\mathbb{F}_q)| = |GL_n(\mathbb{F}_q)|/(q-1)$.

By Definition 4.1.17, $PGL_n(\mathbb{F}_q) = GL_n(\mathbb{F}_q)/Z(GL_n(\mathbb{F}_q))$. Thus

$$|PGL_n(\mathbb{F}_q)| = |GL_n(\mathbb{F}_q)|/|Z(GL_n(\mathbb{F}_q))|. \tag{4.31}$$

Now $Z(GL_n(\mathbb{F}_q)) = F^\times I$ by Theorem 4.1.16, and so $|Z(GL_n(\mathbb{F}_q))| = |F^\times| = q-1$. Thus $|PGL_n(\mathbb{F}_q)| = |GL_n(\mathbb{F}_q)|/(q-1)$.

Now $PSL_n(\mathbb{F}_q) = SL_n(\mathbb{F}_q)/Z(SL_n(\mathbb{F}_q))$, and so

$$|PSL_n(\mathbb{F}_q)| = |SL_n(\mathbb{F}_q)|/|Z(SL_n(\mathbb{F}_q))|. \tag{4.32}$$

By Theorem 4.1.16, $Z(SL_n(\mathbb{F}_q)) = \{fI \mid f \in F^\times, f^n = 1\}$. We have $F^\times$ as a cyclic group of order $q-1$, by Lemma 2.4.14. Thus if $d|(q-1)$, then $F^\times$ has exactly one cyclic subgroup of order d, by Lemmas 2.4.9 and 2.4.5. Therefore, by Lemma 2.4.10, $|\{f \in F^\times \mid f^d = 1\}| = d$.

Now $f^{q-1} = 1$ for all $f \in F^\times$, and

$$f^n = 1 = f^{q-1} \iff f^{\gcd(n,q-1)} = 1. \tag{4.33}$$

Therefore, $Z(SL_n(\mathbb{F}_q)) = \{fI \mid f \in F^\times, f^d = 1, d = \gcd(n, q-1)\}$. There are d such elements, and the third equality follows. □

Corollary 4.3.4. *$PSL_2(\mathbb{F}_2)$ and $PSL_2(\mathbb{F}_3)$ are not simple.*

Proof. Now $\gcd(2, 2-1) = 1$, and so $|PSL_2(\mathbb{F}_2)| = (2^2-1)(2)/1 = 6$. We also have $\gcd(2, 3-1) = 2$, and thus $|PSL_2(\mathbb{F}_3)| = (3^2-1)(3)/2 = (8)(3)/2 = 12$. The smallest simple group of non-prime order is A_5, which is of order 60. Thus neither of these groups are simple. □

Example 4.3.5. We have

$$|GL_3(\mathbb{F}_4)| = (4^3-1)(4^3-4)(4^3-4^2) = (63)(60)(48) = 181440 \tag{4.34}$$

and

$$|SL_3(\mathbb{F}_4)| = |PGL_3(4)| = 181440/3 = 60480. \tag{4.35}$$

Finally, the order of $PSL_3(\mathbb{F}_4)$, is

$$60480/3 = 20160 = 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 = 8!/2 \tag{4.36}$$

because $\gcd(3, 4-1) = 3$.

It is an interesting fact that there are two non-isomorphic simple groups of order $8!/2$, namely $PSL_3(\mathbb{F}_4)$ and A_8 [Rot95].

CHAPTER 5

SEMILINEAR GROUPS

In this chapter, we extend linear groups to semilinear groups by considering vector spaces over fields with nontrivial automorphisms. Our goal is to describe the projective semilinear group $P\Gamma L_3(\mathbb{F}_4)$, and its subgroups $PGL_3(\mathbb{F}_4)$, $P\Sigma L_3(\mathbb{F}_4)$, and $PSL_3(\mathbb{F}_4)$. As we shall see, $P\Gamma L_3(\mathbb{F}_4)$ is a maximal subgroup of the largest Mathieu group, M_{24}, and has an important role in defining an action on 3 of the 24 points acted on by M_{24}.

5.1 Field Extensions and Automorphism Groups

In this section, we describe field extensions, which give rise to nontrivial field automorphisms. We construct the field of four elements, and describe its nontrivial automorphism. We note that since the cube of every element of $\mathbb{F}_4^\times$ is 1, $PSL_3(\mathbb{F}_4)$ is a normal subgroup of $PGL_3(\mathbb{F}_4)$. We follow the treatment of Jacobson [Jac85].

Definition 5.1.1. A *field automorphism* is a bijective ring homomorphism from a field E to itself. If σ is an automorphism of E, then the mapping of $a \in E$ by σ will be written as a^σ, where σ is a right operator.

Lemma 5.1.2. *The automorphisms of a field E form a group,* $\mathrm{Aut}(E)$.

Proof. Let the operation be function composition. This is an associative operation, and the composition of two mappings is a mapping. The identity automorphism, 1, is the mapping that sends every element to itself. If σ is an automorphism of E, it is an isomorphism, and thus has an inverse, σ^{-1}, which is also an isomorphism. $\square$

Definition 5.1.3. A field E is an *extension* of a field F if $F \leq E$, and this *extension field* is denoted E/F.

Definition 5.1.4. If E is a field, then the intersection of all its subfields is called the *prime subfield* P of E. Now P is the smallest subfield of E, and from Lemma 2.5.4, we have that if E is of characteristic p, then P is isomorphic to $\mathbb{Z}_p$, and if E is of characteristic 0, then P is isomorphic to $\mathbb{Q}$.

Lemma 5.1.5. *Every element of* $\mathrm{Aut}(E)$ *fixes the prime subfield* P*.*

Proof. Let $\sigma \in \mathrm{Aut}(E)$, where E has P as its prime subfield. Let $a \in P$. If $\mathrm{char}\, E = p$, then $P \cong \mathbb{Z}_p$ which, as an additive group, is generated by $1 \in P$. Let $a = a' \cdot 1$ for $a' \in \mathbb{Z}$. Thus

$$a^\sigma = (a' \cdot 1)^\sigma = a' \cdot 1^\sigma = a' \cdot 1 = a, \tag{5.1}$$

and so P is fixed by σ. If $\mathrm{char}\, E = 0$, then $P \cong \mathbb{Q}$. Thus $a = rs^{-1}$, where $r = r' \cdot 1$ and $s = s' \cdot 1$ for $r', s' \in \mathbb{Z}$. Thus

$$\begin{aligned} a^\sigma = (rs^{-1})^\sigma = r^\sigma (s^{-1})^\sigma = r^\sigma (s^\sigma)^{-1} = (r' \cdot 1)^\sigma ((s' \cdot 1)^\sigma)^{-1} \\ = (r' \cdot 1^\sigma)(s' \cdot 1^\sigma)^{-1} = (r' \cdot 1)(s' \cdot 1)^{-1} = rs^{-1} = a, \end{aligned} \tag{5.2}$$

and so P is fixed by σ. □

Definition 5.1.6. If E/F is an extension field and $r \in E$, then $F(r)$ is the intersection of all subfields of E containing F and r. Thus it is the smallest subfield of E containing F and r. We *adjoin* r to F to form $F(r)$. We define $F(r_1, r_2) = (F(r_1))(r_2)$, and thus recursively define $F(r_1, \ldots, r_k)$, for $k \in \mathbb{Z}^+$.

Definition 5.1.7. If every element in a field E is the root of a nonzero polynomial with coefficients from F, then E is an *algebraic extension* of F.

Definition 5.1.8. Let F be a field, and let $f(x)$ be a monic polynomial in $F[x]$. If

$$f(x) = (x - r_1)(x - r_2)\dots(x - r_n) \in E[x] \tag{5.3}$$

and

$$E = F(r_1, r_2, \dots, r_n), \tag{5.4}$$

then E/F is a *splitting field* over F of $f(x)$.

Definition 5.1.9. If $f(x)$ splits over E with distinct roots, then $f(x)$ is *separable*, and the splitting field E/F, formed by adjoining the roots of $f(x)$, is called a *Galois extension* of F.

Definition 5.1.10. The *Galois group*, $\mathrm{Gal}(E/F)$, is the automorphism group of a Galois extension field E/F, where each automorphism fixes every element of the *base field* F.

Lemma 5.1.11. *If the Galois extension field E/P has prime subfield P, then* $\mathrm{Aut}(E) = \mathrm{Gal}(E/P)$.

Proof. If $\sigma \in \mathrm{Aut}(E)$, then by Lemma 5.1.5, σ fixes every element of P. Thus $\sigma \in \mathrm{Gal}(E/P)$. If $\sigma \in \mathrm{Gal}(E/P)$, then it is an automorphism of E fixing every element in P. Thus $\sigma \in \mathrm{Aut}(E)$. □

Theorem 5.1.12. *For every prime p and positive integer k, there exists a unique finite field of order p^k. This field has as its prime subfield the finite field $\mathbb{Z}_p$. A finite field of order q is denoted $\mathbb{F}_q$, and is the splitting field of the polynomial* $x^q - x \in \mathbb{Z}_p[x]$.

Proof. See Jacobson [Jac85]. □

Example 5.1.13. Consider $f(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$. Since $f(0) = 1 = f(1)$, $f(x)$ has no linear factors, and since it is of degree 2, $f(x)$ is irreducible in $\mathbb{Z}_2[x]$. Let $\langle f(x) \rangle$ be the ideal in $Z_2[x]$ generated by $f(x)$. Now

$$E = \mathbb{Z}_2[x] / \langle f(x) \rangle \tag{5.5}$$

is a field, because $f(x)$ is irreducible. The zero and unity elements of E are $0 = 0 + \langle f(x) \rangle$ and $1 = 1 + \langle f(x) \rangle$, respectively. Let

$$\omega = x + \langle f(x) \rangle \qquad \text{and} \qquad \bar{\omega} = \omega + 1 = (x + 1) + \langle f(x) \rangle . \tag{5.6}$$

These elements are the two roots of $f(x)$ in E, since

$$\omega^2 + \omega + 1 = (x^2 + x + 1) + \langle f(x) \rangle = f(x) + \langle f(x) \rangle = 0 + \langle f(x) \rangle = 0 \tag{5.7}$$

and

$$\bar{\omega}^2 + \bar{\omega} + 1 = (x^2 + 2x + 1) + (x + 1) + 1 + \langle f(x) \rangle = (x^2 + x + 1) + \langle f(x) \rangle = 0. \tag{5.8}$$

Thus our field is a splitting field of $f(x)$. Since the two roots differ by 1, we get both by adjoining ω or $\bar{\omega}$ to our base field. Thus

$$\mathbb{Z}_2[x] / \langle f(x) \rangle = \mathbb{Z}_2(\omega). \tag{5.9}$$

Since the two roots are distinct, our field is a Galois extension field. Now

$$f(x) = x^2 + x + 1 = 0 \qquad \Rightarrow \qquad x^2 = -x - 1 = x + 1. \tag{5.10}$$

Thus $x^2 + \langle f(x) \rangle = (x + 1) + \langle f(x) \rangle$. Since our field is a vector space with basis $\{1, \omega\}$, the elements in our field are

$$\mathbb{Z}_2(\omega) = \{0 + 0\,\omega, 1 + 0\,\omega, 0 + \omega, 1 + \omega\} = \{0, 1, \omega, \bar{\omega}\}. \tag{5.11}$$

$+$	0	1	ω	$\bar{\omega}$
0	0	1	ω	$\bar{\omega}$
1	1	0	$\bar{\omega}$	ω
ω	ω	$\bar{\omega}$	0	1
$\bar{\omega}$	$\bar{\omega}$	ω	1	0

$\cdot$	0	1	ω	$\bar{\omega}$
0	0	0	0	0
1	0	1	ω	$\bar{\omega}$
ω	0	ω	$\bar{\omega}$	1
$\bar{\omega}$	0	$\bar{\omega}$	1	ω

Table 5.1: Operation Tables for $\mathbb{F}_4$

The addition and multiplication tables for this field, denoted $\mathbb{F}_4$ from now on, are shown in Table 5.1 below.

Thus the additive group of $\mathbb{F}_4$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$, while the multiplicative group $\mathbb{F}_4^\times$ is isomorphic to $\mathbb{Z}_3$, with $\omega^3 = 1 = \bar{\omega}^3$.

Now $\mathbb{F}_4$ has two automorphisms:

$$1 : \omega \mapsto \omega \qquad \text{and} \qquad \sigma : \omega \mapsto \bar{\omega}. \tag{5.12}$$

The identity automorphism, 1, fixes every element in $\mathbb{F}_4$, while σ acts like complex conjugation, fixing the prime subfield $\mathbb{F}_2$, but exchanging ω and $\bar{\omega}$. Thus

$$\mathrm{Gal}(\mathbb{F}_4/\mathbb{F}_2) = \{1, \sigma\} \cong \mathbb{Z}_2. \tag{5.13}$$

Using right actions, we denote the mapping of elements by σ as

$$0^\sigma = 0, \qquad 1^\sigma = 1, \qquad \omega^\sigma = \bar{\omega}, \qquad \bar{\omega}^\sigma = \omega. \tag{5.14}$$

Theorem 5.1.14. *We have $PSL_3(\mathbb{F}_4) \lhd PGL_3(\mathbb{F}_4)$, and the determinant of $[A] \in PGL_3(\mathbb{F}_4)$ is well-defined.*

Proof. Now $Z(GL_3(\mathbb{F}_4)) = \{fI \mid f \in \mathbb{F}_4^\times\}$ and $Z(SL_3(\mathbb{F}_4)) = \{fI \mid f \in \mathbb{F}_4^\times, f^3 = 1\}$ from Theorem 4.1.16. Since $\omega^3 = 1 = \bar{\omega}^3$, this implies $Z(GL_3(\mathbb{F}_4)) = Z(SL_3(\mathbb{F}_4)) = \{I, \omega I, \bar{\omega} I\} = \mathbb{F}_4^\times I$. Since $SL_3(\mathbb{F}_4) \lhd GL_3(\mathbb{F}_4)$, by Lemma 4.1.4,

$$SL_3(\mathbb{F}_4)/\mathbb{F}_4^\times I \lhd GL_3(\mathbb{F}_4)/\mathbb{F}_4^\times I. \tag{5.15}$$

Let $[A] \in PGL_3(\mathbb{F}_4)$. Thus $[A] = A(\mathbb{F}_4^\times I) = \{A, \omega A, \bar{\omega} A\}$, for some $A \in GL_3(\mathbb{F}_4)$. We have

$$\det(\omega A) = \det(\omega I)(A) = \det(\omega I) \det A = \omega^3 \det A = \det A. \tag{5.16}$$

Similarly, $\det(\bar{\omega} A) = \bar{\omega}^3 \det A = \det A$. □

Remark 5.1.15. We usually denote $[A] \in PGL_3(\mathbb{F}_4)$ as $A \in PGL_3(\mathbb{F}_4)$, even though $[A] \in PGL_3(\mathbb{F}_4)$ is the coset containing $A \in GL_3(\mathbb{F}_4)$.

5.2 Semilinear Groups

In this section, we describe how linear groups are extended to semilinear groups, and prove several facts about semilinear groups. Our definitions are from Gruenberg [GW77].

Definition 5.2.1. Let V, W be two vectors spaces over a field K and let θ be a field automorphism of K. A transformation $T : V \to W$ is *θ-semilinear* if, for all $x, y \in V$ and $a \in K$,

$$(x + y)T = xT + yT \qquad \text{and} \qquad (ax)T = a^\theta xT. \tag{5.17}$$

If T is a *θ-semilinear* transformation for some $\theta \in \mathrm{Aut}(K)$, we say T is a *semilinear transformation*.

Remark 5.2.2. For the rest of this chapter, let E be a Galois extension field with F as its prime subfield. Thus $\mathrm{Aut}(E) = \mathrm{Gal}(E/F)$ by Lemma 5.1.11. We also let $V = E^n$, for $n \in \mathbb{Z}^+$.

Theorem 5.2.3. *Let $V = E^n$ be a vector space over a Galois extension field E/F. The set of all invertible semilinear transformations from V to V forms a group, the general semilinear group, denoted $\Gamma L(V)$.*

Proof. Let $\phi, \theta \in \mathrm{Gal}(E/F)$, and let $S, T \in \Gamma L(V)$ such that S is ϕ-semilinear and T is θ-semilinear. Let $x, y \in V$ and $a \in E/F$. We have

$$(x+y)ST + ((x+y)S)T = (xS + yS)T = (xS)T + (yS)T = xST + yST \tag{5.18}$$

and

$$(ax)ST = ((ax)S)T = (a^{\phi}(xS))T = (a^{\phi})^{\theta}((xS)T) = a^{\phi\theta}(xST). \tag{5.19}$$

Since $\mathrm{Gal}(E/F)$ is a group, $\phi\theta \in \mathrm{Gal}(E/F)$. Thus ST is $\phi\theta$-semilinear, and $ST \in \Gamma L(V)$.

Now composition of functions is associative. Also, define the identity transformation 1 such that $x1 = x$ for all $x \in V$. Then since

$$(x+y)1 = x + y = x1 + y1 \qquad \text{and} \qquad (ax)1 = ax = a^1 x1, \tag{5.20}$$

where $1 \in \mathrm{Gal}(E/F)$, we have $1 \in \Gamma L(V)$, and by the usual properties of the identity, 1 is the identity element of $\Gamma L(V)$.

Let $T \in \Gamma L(V)$, where T is θ-semilinear. Since the semilinear transformations in $\Gamma L(V)$ are invertible, there exists a $T^{-1} : V \to V$ such that $TT^{-1} = 1 = T^{-1}T$. Now $(xT^{-1})T = x(T^{-1}T) = x1 = x$, and we have

$$(xT^{-1} + yT^{-1})T = (xT^{-1})T + (yT^{-1})T = x + y = ((x+y)T^{-1})T. \tag{5.21}$$

Multiplying both sides on the right by T^{-1}, we have $xT^{-1} + yT^{-1} = (x+y)T^{-1}$. Since automorphisms are invertible, $\theta^{-1} \in \mathrm{Gal}(E/F)$. Let $a \in E/F$. Thus

$$(a^{\theta^{-1}}(xT^{-1}))T = (a^{\theta^{-1}})^{\theta}((xT^{-1})T) = a^1(x1) = ax \tag{5.22}$$

Multiplying both sides on the right by T^{-1}, we have $a^{\theta^{-1}}(xT^{-1}) = (ax)T^{-1}$. We conclude T^{-1} is θ^{-1}-semilinear. $\square$

Theorem 5.2.4. *Given a vector space $V = E^n$ over a Galois extension field E/F, $GL(V) \lhd \Gamma L(V)$.*

Proof. Let $S \in GL(V)$, and let $x \in V$ and $a \in K$. Now

$$(ax)S = axS = a^1 xS, \tag{5.23}$$

so linear maps are 1-semilinear maps. Thus $S \in \Gamma L(V)$, and $GL(V) \leq \Gamma L(V)$.

Now let $T \in \Gamma L(V)$. If T is θ-semilinear, then T^{-1} is θ^{-1}-semilinear, by the proof of Theorem 5.2.3. We must show $T^{-1}ST \in GL(V)$. Since elements of $\Gamma L(V)$ preserve vector addition, we need only determine how this conjugate acts on scalar multiples. Let $x \in V$ and $a \in K$. We have

$$\begin{aligned} (ax)T^{-1}ST &= ((ax)T^{-1})ST = (a^{\theta^{-1}}(xT^{-1}))ST = ((a^{\theta^{-1}})^1(xT^{-1}S))T \\ &= (a^{\theta^{-1}})^{\theta}(xT^{-1}ST) = a^1(xT^{-1}ST) = a(xT^{-1}ST). \end{aligned} \tag{5.24}$$

Thus $T^{-1}ST \in GL(V)$, and we conclude $GL(V) \lhd \Gamma L(V)$. □

Lemma 5.2.5. *Let $V = E^n$ be a vector space over a Galois extension field, E/F. Define $f : \mathrm{Gal}(E/F) \to \Gamma L(V)$ such that*

$$(x_1, \ldots, x_n)f(\sigma) = (x_1^\sigma, \ldots, x_n^\sigma). \tag{5.25}$$

Then f is a one-to-one homomorphism.

Proof. Given a basis β, we express $x, y \in V$ as $x = (x_1, \ldots x_n)$ and $y = (y_1, \ldots, y_n)$, where $x_i, y_i \in E/F$ for $1 \leq i \leq n$. Let $\sigma \in \mathrm{Gal}(E/F)$. Since σ is an automorphism of E/F, $(x_i + y_i)^\sigma = x_i^\sigma + y_i^\sigma$ for all i. Thus

$$\begin{aligned} (x+y)f(\sigma) &= ((x_1, \ldots, x_n) + (y_1, \ldots, y_n))f(\sigma) = (x_1 + y_1, \ldots, x_n + y_n)f(\sigma) \\ &= ((x_1 + y_1)^\sigma, \ldots, (x_n + y_n)^\sigma) = (x_1^\sigma + y_1^\sigma, \ldots, x_n^\sigma + y_n^\sigma) \\ &= (x_1^\sigma, \ldots, x_n^\sigma) + (y_1^\sigma, \ldots, y_n^\sigma) = (x_1, \ldots, x_n)f(\sigma) + (y_1, \ldots, y_n)f(\sigma) \\ &= xf(\sigma) + yf(\sigma). \end{aligned} \tag{5.26}$$

Since σ is an automorphism of E/F, for $a \in E/F$ we have $(ax_i)^\sigma = a^\sigma x_i^\sigma$. Thus

$$\begin{aligned}(ax)f(\sigma) &= (a(x_1,\ldots,x_n))f(\sigma) = (ax_1,\ldots,ax_n)f(\sigma)\\ &= ((ax_1)^\sigma,\ldots,(ax_n)^\sigma) = (a^\sigma x_1^\sigma,\ldots,a^\sigma x_n^\sigma)\\ &= a^\sigma(x_1^\sigma,\ldots,x_n^\sigma) = a^\sigma(xf(\sigma)).\end{aligned} \tag{5.27}$$

Therefore, $f(\sigma) \in \Gamma L(V)$, and f is well-defined. Now let $\sigma, \phi \in \mathrm{Gal}(E/F)$ and let $x \in V$. We have

$$\begin{aligned}x(f(\sigma)f(\phi)) &= (xf(\sigma))f(\phi) = ((x_1,\ldots,x_n)f(\sigma))f(\phi) = (x_1^\sigma,\ldots,x_n^\sigma)f(\phi)\\ &= ((x_1^\sigma)^\phi,\ldots,(x_n^\sigma)^\phi) = (x_1^{\sigma\phi},\ldots,x_n^{\sigma\phi}) = (x_1\ldots x_n)f(\sigma\phi)\\ &= xf(\sigma\phi).\end{aligned} \tag{5.28}$$

Thus $f(\sigma)f(\phi) = f(\sigma\phi)$, and f is a homomorphism.

Now suppose $f(\sigma_1) = f(\sigma_2)$. For $k \in E/F$ we have

$$(k^{\sigma_1},0,\ldots,0) = (ke_1)f(\sigma_1) = (ke_1)f(\sigma_2) = (k^{\sigma_2},0,\ldots,0), \tag{5.29}$$

implying $k^{\sigma_1} = k^{\sigma_2}$. Since k is arbitrary, $\sigma_1 = \sigma_2$, and f is one-to-one. □

Theorem 5.2.6. *Given a vector space $V = E^n$ over a Galois extension field E/F, we have*

$$\Gamma L(V) = GL(V)f(\mathrm{Gal}(E/F)). \tag{5.30}$$

Proof. From Lemma 5.2.5, we have $f(\mathrm{Gal}(E/F)) \leq \Gamma L(V)$, where $\mathrm{Gal}(E/F)$ and its image have the same cardinality. By Lemma 2.2.3 and Theorem 5.2.4, we have

$$f(\mathrm{Gal}(E/F)) \leq GL(V)f(\mathrm{Gal}(E/F)) \leq \Gamma L(V). \tag{5.31}$$

Let $T \in \Gamma L(V)$, where T is σ-semilinear. Now $f(\sigma^{-1}) \in \Gamma L(V)$, and thus $Tf(\sigma^{-1}) \in \Gamma L(V)$. Let $a \in E/F$ and $x \in V$. We have

$$\begin{aligned}(ax)(Tf(\sigma^{-1})) &= ((ax)T)f(\sigma^{-1}) = (a^\sigma(xT))f(\sigma^{-1})\\ &= (a^\sigma)^{\sigma^{-1}}((xT)f(\sigma^{-1})) = a^1(x(Tf(\sigma^{-1}))).\end{aligned} \tag{5.32}$$

Thus $Tf(\sigma^{-1}) \in GL(V)$. Since f is a homomorphism,

$$(Tf(\sigma^{-1}))f(\sigma) = T(f(\sigma^{-1})f(\sigma)) = T(f(\sigma^{-1}\sigma)) = Tf(1) = T1 = T. \tag{5.33}$$

Thus every element in $\Gamma L(V)$ is the product of an element of $GL(V)$ with an element of $f(\mathrm{Gal}(E/F))$. □

Corollary 5.2.7. *If $f(\sigma)T_1 \in \Gamma L(V)$, such that $T_1 \in GL(V)$ and $f(\sigma) \in \Gamma L(V)$, then there exists $T_2 \in GL(V)$ such that $f(\sigma)T_1 = T_2 f(\sigma)$.*

Proof. Now $\Gamma L(V) = GL(V)f(\mathrm{Gal}(E/F))$ and $GL(V) \lhd \Gamma L(V)$, by Theorems 5.2.6 and 5.2.4. For $\sigma \in \mathrm{Gal}(E/F)$ we have $f(\sigma^{-1}) \in \Gamma L(V)$. Thus

$$f(\sigma^{-1})^{-1}T_1 f(\sigma^{-1}) = T_2 \in GL(V), \tag{5.34}$$

and so

$$f(\sigma^{-1})^{-1}T_1 = T_2 f(\sigma^{-1})^{-1} \iff f(\sigma)T_1 = T_2 f(\sigma). \tag{5.35}$$

□

Definition 5.2.8. Let K be a subgroup of a group G. A subgroup H of G is a *complement* of K in G if $KH = G$ and $K \cap H = 1$.

Definition 5.2.9. A group G is a *semidirect product* of K by Q, denoted by $G = K \rtimes Q$, if $K \lhd G$ and K has a complement $Q_1 \cong Q$.

Corollary 5.2.10. *Given a vector space $V = E^n$ over a Galois extension field E/F, we have $\Gamma L(V) = GL(V) \rtimes \mathrm{Gal}(E/F)$.*

Proof. From Theorem 5.2.4,

$$GL(V) \lhd \Gamma L(V), \tag{5.36}$$

and from Theorem 5.2.6,

$$\Gamma L(V) = GL(V)f(\mathrm{Gal}(E/F)). \tag{5.37}$$

Let $T \in GL(V) \cap f(\mathrm{Gal}(E/F))$. Now for $\theta \in \mathrm{Gal}(E/F)$, $f(\theta)$ is θ-semilinear. Since $T \in GL(V)$ is 1-semilinear, and f is one-to-one, $T = f(1) = 1$ for $1 \in \mathrm{Gal}(E/F)$. Thus $GL(V) \cap f(\mathrm{Gal}(E/F)) = 1$, and $f(\mathrm{Gal}(E/F))$ is a complement of $GL(V)$. From Lemma 5.2.5,

$$f(\mathrm{Gal}(E/F) \cong \mathrm{Gal}(E/F). \tag{5.38}$$

We conclude $\Gamma L(V) = GL(V) \rtimes \mathrm{Gal}(E/F)$. □

Corollary 5.2.11. *Given a vector space $V = E^n$ over a finite field E/F, we have $|\Gamma L(V)| = |GL(V)||\,\mathrm{Gal}(E/F)|$.*

Proof. Define

$$\varphi : GL(V) \times f(\mathrm{Gal}(E/F)) \to \Gamma L(V) \qquad \text{such that} \qquad \varphi(S, f(\sigma)) = Sf(\sigma). \tag{5.39}$$

From Theorem 5.2.6, this mapping is onto. Now let $\varphi(S_1, f(\sigma_1)) = \varphi(S_2, f(\sigma_2))$. Thus $S_1 f(\sigma_1) = S_2 f(\sigma_2)$. Now $S_i f(\sigma_i)$ is σ_i-semilinear, so $f(\sigma_1) = f(\sigma_2)$, and since f is one-to-one, $\sigma_1 = \sigma_2$. Thus $S_1 f(\sigma_1) = S_2 f(\sigma_1)$. Multiplying both sides on the right by $f(\sigma_1)^{-1}$, we see $S_1 = S_2$. Thus φ is one-to-one, and we have

$$|\Gamma L(V)| = |GL(V)||f(\mathrm{Gal}(E/F))| = |GL(V)||\,\mathrm{Gal}(E/F)|, \tag{5.40}$$

where the last equality follows from Lemma 5.2.5. □

Definition 5.2.12. Let $V = E^n$ be a vector space over a Galois extension field, E/F. For $S \in GL(V)$ and $\sigma \in \mathrm{Gal}(E/F)$, define $S^\sigma = f(\sigma)^{-1} S f(\sigma)$. If $GL(V)$ is viewed as a group of $n \times n$ invertible matrices, for $A \in GL(V)$, define $A^\sigma = (a_{ij})^\sigma = (a_{ij}^\sigma)$.

Theorem 5.2.13. *Let $V = E^n$ be a vector space over the Galois extension field E/F. If $S \in GL(V)$, and the matrix of S is A, then $S^\sigma \in GL(V)$, and the matrix of S^σ is A^σ.*

Proof. From Theorem 5.2.4, we have $GL(V) \lhd \Gamma L(V)$. Thus for $S \in GL(V)$ and $f(\sigma) \in f(\mathrm{Gal}(E/F)) \leq \Gamma L(V)$, we have $f(\sigma)^{-1} S f(\sigma) \in GL(V)$. Let $\beta = \{e_1, \ldots, e_n\}$ be the standard ordered basis. Using right actions, $[e_i S]_\beta$ is the ith row of matrix A. But then, since $1^{\sigma^{-1}} = 1$, we have

$$[e_i S^\sigma]_\beta = [e_i f(\sigma^{-1}) S f(\sigma)]_\beta = [(e_i f(\sigma^{-1})) S f(\sigma)]_\beta = [(e_i)^{\sigma^{-1}} S f(\sigma)]_\beta$$
$$= [e_i S f(\sigma)]_\beta = [(e_i S) f(\sigma)]_\beta = [(e_i S)^\sigma]_\beta. \quad (5.41)$$

Therefore, the matrix of S^σ is A^σ. □

Corollary 5.2.14. *We have* $\det A^\sigma = (\det A)^\sigma$.

Proof. Now for $A \in GL(E/F)$, where E/F is n-dimensional and $\sigma \in \mathrm{Gal}(E/F)$, we have

$$\det A = \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^{n} a_{i,\pi(i)} \quad (5.42)$$

Since σ is an automorphism of E/F, it preserves addition and multiplication in E/F. Thus

$$\begin{aligned} \det A^\sigma &= \sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^{n} a_{i,\pi(i)}^\sigma = \sum_{\pi \in S_n} \operatorname{sgn} \pi \left(\prod_{i=1}^{n} a_{i,\pi(i)} \right)^\sigma \\ &= \left(\sum_{\pi \in S_n} \operatorname{sgn} \pi \prod_{i=1}^{n} a_{i,\pi(i)} \right)^\sigma = (\det A)^\sigma. \quad (5.43) \end{aligned}$$

□

Theorem 5.2.15. *Given a vector space $V = E^n$ over a Galois extension field E/F, we have $SL(V) \lhd \Gamma L(V)$.*

Proof. Let $R \in SL(V)$, and let $T \in \Gamma L(V)$. From Theorem 5.2.6, we know that $T = S f(\sigma)$ where $S \in GL(V)$ and $f(\sigma) \in f(\mathrm{Gal}(E/F))$. We also know

$SL(V) \lhd GL(V)$, since $SL(V)$ is the kernel of the determinant homomorphism $\det : GL_n(F) \to F^\times$. Thus we have

$$\begin{aligned} T^{-1}RT &= (Sf(\sigma))^{-1}R(Sf(\sigma)) = (f(\sigma)^{-1}S^{-1})R(Sf(\sigma)) \\ &= f(\sigma)^{-1}(S^{-1}RS)f(\sigma) = f(\sigma)^{-1}\widehat{R}f(\sigma) = \widehat{R}^\sigma. \end{aligned} \tag{5.44}$$

Now $\widehat{R} = S^{-1}RS \in SL(V)$, since $S \in GL(V)$ and $SL(V) \lhd GL(V)$. Using Corollary 5.2.14, we have

$$\det \widehat{R}^\sigma = (\det \widehat{R})^\sigma = 1^\sigma = 1. \tag{5.45}$$

Thus $T^{-1}RT = \widehat{R}^\sigma \in SL(V)$, and $SL(V) \lhd \Gamma L(V)$. □

Theorem 5.2.16. *The product $SL(V)f(\mathrm{Gal}(E/F))$ is a subgroup of $\Gamma L(V)$, called the special semilinear group, denoted $\Sigma L(V)$. Thus $\Sigma L(V) = SL(V) \rtimes \mathrm{Gal}(E/F)$.*

Proof. Since $SL(V) \lhd \Gamma L(V)$ from Theorem 5.2.15, we use Lemma 2.2.3 to obtain:

$$f(\mathrm{Gal}(E/F)) \leq (SL(V))f(\mathrm{Gal}(E/F)) \leq \Gamma L(V). \tag{5.46}$$

Now $SL(V) \lhd \Gamma L(V)$ and $f(\mathrm{Gal}(E/F)) \cap SL(V) = 1$, and so the second statement follows. □

Remark 5.2.17. We gather together our results in this section on the subgroups of $\Gamma L(V)$. We have

$$SL(V) \lhd \Gamma L(V) \qquad \text{and} \qquad \begin{array}{ccc} \Sigma L(V) & \leq & \Gamma L(V) \\ \triangledown & & \triangledown \\ SL(V) & \lhd & GL(V) \end{array}. \tag{5.47}$$

Example 5.2.18. As in Example 5.1.13, let our field be $\mathbb{F}_4 = \{0, 1, \omega, \bar{\omega}\}$. Thus $\mathrm{Gal}(\mathbb{F}_4/\mathbb{F}_2) = \{1, \sigma\}$, where σ maps $\omega \mapsto \bar{\omega}$ and $\bar{\omega} \mapsto \omega$. Thus we have

$$SL_3(\mathbb{F}_4) \lhd \Gamma L_3(\mathbb{F}_4) \qquad \text{and} \qquad \begin{array}{ccc} \Sigma L_3(\mathbb{F}_4) & \leq & \Gamma L_3(\mathbb{F}_4) \\ \triangledown & & \triangledown \\ SL_3(\mathbb{F}_4) & \lhd & GL_3(\mathbb{F}_4) \end{array}. \tag{5.48}$$

We wish to form factor groups modulo $SL_3(\mathbb{F}_4)$. By Theorem 4.3.3, $GL_3(\mathbb{F}_4)$ is three times bigger than $SL_3(\mathbb{F}_4)$, because there are three nonzero determinants that elements in $GL_3(\mathbb{F}_4)$ possess, while 1 is the determinant of every element in $SL_3(\mathbb{F}_4)$. By Corollary 5.2.11, $\Gamma L_3(\mathbb{F}_4)$ is twice as big as $GL_3(\mathbb{F}_4)$, because it is extended by the nontrivial automorphism σ. Thus $\Gamma L_3(\mathbb{F}_4)$ is six times bigger than $SL_3(\mathbb{F}_4)$. Similarly, $\Sigma L_3(\mathbb{F}_4)$ is twice as big as $SL_3(\mathbb{F}_4)$, since it is extended by the nontrivial automorphism σ.

There are two groups of order 6, $\mathbb{Z}_6$ and S_3. We will show $\Gamma L_3(\mathbb{F}_4)/SL_3(\mathbb{F}_4)$ is not abelian, and therefore, $\Gamma L_3(\mathbb{F}_4)/SL_3(\mathbb{F}_4) \cong S_3$. Recall that (Lemma 2.1.12) if $N \lhd G$ and $G' \not\leq N$, then G/N is not abelian. Taking $G = \Gamma L_3(\mathbb{F}_4)$ and $N = SL_3(\mathbb{F}_4)$, we have $A, f(\sigma) \in G$ where

$$A = \begin{bmatrix} \omega & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.49}$$

Thus $A^{-1}f(\sigma)^{-1}Af(\sigma) \in G'$, and

$$A^{-1}f(\sigma)^{-1}Af(\sigma) = A^{-1}\left(f(\sigma)^{-1}Af(\sigma)\right) = A^{-1}A^{\sigma} \tag{5.50}$$

by Theorem 5.2.13. Thus

$$A^{-1}f(\sigma)^{-1}Af(\sigma) = \begin{bmatrix} \bar{\omega} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{\omega} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \omega & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.51}$$

whose determinant is $\omega \neq 1$. Thus $G' \not\leq N$, and $\Gamma L_3(\mathbb{F}_4)/SL_3(\mathbb{F}_4) \cong S_3$. Now $\langle\sigma\rangle \cong \mathbb{Z}_2$ and $\langle\omega\rangle \cong \mathbb{Z}_3$. Modulo $SL_3(\mathbb{F}_4)$, we have

$$1 \lhd S_3 \qquad \text{and} \qquad \begin{array}{ccc} \langle\sigma\rangle & \leq & S_3 \\ \triangledown & & \triangledown \\ 1 & \lhd & \langle\omega\rangle \end{array}. \tag{5.52}$$

Since $\langle\sigma\rangle$, which is isomorphic to a transposition subgroup, is not normal in S_3, this example shows $\Sigma L(V)$ is not a normal subgroup of $\Gamma L(V)$.

5.3 Action of Semilinear Groups on Projective Space

In this section, we define the action of the semilinear group (and its projective form) on projective space.

Lemma 5.3.1. *The semilinear group $\Gamma L_n(E)$ acts on $P^{n-1}(E)$, such that $[v]T = [vT]$, for $T \in \Gamma L_n(E)$ and $[v] \in P^{n-1}(E)$. The special semilinear group $\Sigma L_n(E)$ acts similarly on $P^{n-1}(E)$.*

Proof. Recall $V^{\#} = V \setminus \{0\}$. Let $T \in \Gamma L_n(E)$ be σ-semilinear. If $[v] \in P^{n-1}(E)$, then $v \in E^{n\#}$. Since $v \neq 0$ and T is invertible, $vT \neq 0$. Thus $vT \in E^{n\#}$, meaning $[vT] \in P^{n-1}(E)$. Now suppose $[v] = [w]$. Then $v = \alpha w$ for $v, w \in E^{n\#}$ and $\alpha \in E^{\times}$. Then

$$vT = (\alpha w)T = \alpha^{\sigma}(wT). \tag{5.53}$$

Since $\alpha^{\sigma} \in E^{\times}$, $[vT] = [wT]$. Finally,

(1) $[v]I = [vI] = [v]$ and

(2) $[v]ST = [v(ST)] = [(vS)T] = [vS]T = ([v]S)T$

for all $v \in E^{n\#}$ and all $S, T \in \Gamma L_n(E)$. The case of $\Sigma L_n(E)$ follows by a restriction on the action. □

Lemma 5.3.2. *We have $Z(GL_n(E)) \lhd \Gamma L_n(E))$ and $Z(SL_n(E)) \lhd \Sigma L_n(E))$.*

Proof. Now $Z(GL_n(E)) = \{\alpha I \mid \alpha \in E^{\times}\}$ by Theorem 4.1.16. Let $\alpha I \in Z(GL_n(E))$

and $Sf(\sigma) \in \Gamma L_n(E)$. We have

$$\left(f(\sigma)^{-1}S^{-1}\right)(\alpha I)\left(Sf(\sigma)\right) = f(\sigma)^{-1}\left(S^{-1}(\alpha I)\,S\right)f(\sigma) = f(\sigma)^{-1}(\alpha I)f(\sigma) = \alpha^{\sigma} I \in Z(GL_n(E)). \quad (5.54)$$

We have $Z(SL_n(E)) = \{\alpha I \mid \alpha \in E^{\times}, \alpha^n = 1\}$, also by Theorem 4.1.16. Let $\alpha I \in Z(SL_n(E))$ and $Sf(\sigma) \in \Sigma L_n(E)$. We have

$$\left(f(\sigma)^{-1}S^{-1}\right)(\alpha I)\left(Sf(\sigma)\right) = f(\sigma)^{-1}\left(S^{-1}(\alpha I)\,S\right)f(\sigma) = f(\sigma)^{-1}(\alpha I)f(\sigma) = \alpha^{\sigma} I \in Z(GL_n(E)). \quad (5.55)$$

Now $(\alpha^{\sigma})^n = (\alpha^n)^{\sigma} = 1^{\sigma} = 1$, and so $\alpha^{\sigma} \in Z(SL_n(E))$. □

Definition 5.3.3. The *projective semilinear group* is:

$$P\Gamma L_n(E) = \Gamma L_n(E)/Z(GL_n(E)). \quad (5.56)$$

Definition 5.3.4. The *projective special semilinear group* is:

$$P\Sigma L_n(E) = \Sigma L_n(E)/Z(SL_n(E)). \quad (5.57)$$

Lemma 5.3.5. *The kernels of the actions of* $\Gamma L_n(E)$ *and* $GL_n(E)$ *on* $P^{n-1}(E)$ *are identical, as are those of* $\Sigma L_n(E)$ *and* $SL_n(E)$.

Proof. Now $Z(GL_n(E)) = \{\alpha I \mid \alpha \in E^{\times}\}$ is the kernel of the action of $GL_n(E)$ on $P^{n-1}(E)$ by Lemma 4.2.5, and it suffices to show that the kernel of the action of $\Gamma L_n(E)$ on $P^{n-1}(E)$ is contained in $Z(GL_n(E))$.

Let $Sf(\phi)$ be in the kernel of the action of $\Gamma L_n(E)$ on $P^{n-1}(E)$, where $S \in GL_n(E)$ and $\phi \in \mathrm{Gal}(E/F)$. Thus for $1 \leq i \leq n$, and for some $\alpha \in E^{\times}$,

$$\alpha e_i = e_i(Sf(\phi)) = (e_i S)f(\phi) = (e_i S)^{\phi}, \quad (5.58)$$

implying the ith row of S is $\alpha^{\phi^{-1}} e_i$, where $\beta = \alpha^{\phi^{-1}} \in E^\times$. Thus S is the scalar matrix βI.

Let $(\lambda, 1, \ldots, 1) \in E^n$, where $\lambda \in E^\times$. We have

$$(\lambda, 1, \ldots, 1)Sf(\phi) = (\lambda\beta, \beta, \ldots, \beta)f(\phi) = ((\lambda\beta)^\phi, \beta^\phi, \ldots, \beta^\phi)$$
$$= (\beta^\phi \lambda^\phi, \beta^\phi, \ldots, \beta^\phi) = \beta^\phi(\lambda^\phi, 1, \ldots, 1). \quad (5.59)$$

Since $(\lambda, 1, \ldots, 1)Sf(\phi)$ is a scalar multiple of $(\lambda, 1, \ldots, 1)$, we have $\lambda = \lambda^\phi$, and since $\lambda \in E^\times$ is arbitrary, $\phi = 1$. Thus $Sf(\phi) = Sf(1) = S$, a scalar matrix, and the kernel of the action of $\Gamma L_n(E)$ on $P^{n-1}(E)$ is $Z(GL_n(E))$.

The argument for the identity between the kernels of $\Sigma L_n(E)$ and $SL_n(E)$ is similar. □

Corollary 5.3.6. *$P\Gamma L_n(E)$ and $P\Sigma L_n(E)$ act faithfully on $P^{n-1}(E)$.*

Proof. By Lemmas 3.2.16 and 5.3.2, an induced action of $P\Gamma L_n(E)$ on $P^{n-1}(E)$ is well-defined. Since $Z(GL_n(E))$ is the kernel of the action of $\Gamma L_n(E)$ on $P^{n-1}(E)$, we have a faithful action by Corollary 3.2.18.

The case of $P\Sigma L_n(E)$ is similar. □

Example 5.3.7. We have $\mathrm{Gal}(\mathbb{F}_4/\mathbb{F}_2) = \{1, \sigma\}$ such that $\omega^\sigma = \bar{\omega}$ and $\bar{\omega}^\sigma = \omega$. Thus

$$|\Gamma L_3(\mathbb{F}_4)| = |GL_3(\mathbb{F}_4)||\,\mathrm{Gal}(\mathbb{F}_4/\mathbb{F}_2)| = (181440)(2) = 362880 \quad (5.60)$$

and

$$|\Sigma L_3(\mathbb{F}_4)| = |SL_3(\mathbb{F}_4)||\,\mathrm{Gal}(\mathbb{F}_4/\mathbb{F}_2)| = (60480)(2) = 120960. \quad (5.61)$$

Now $\omega^3 = \bar{\omega}^3 = 1$, and so $\det \omega I = \det \bar{\omega} I = 1$. Thus $Z(GL_3(\mathbb{F}_4)) = Z(SL_3(\mathbb{F}_4)$, and $|Z(GL_3(\mathbb{F}_4))| = |Z(SL_3(\mathbb{F}_4))| = 3$. We have

$$|P\Gamma L_3(\mathbb{F}_4)| = |\Gamma L_3(\mathbb{F}_4)|/|Z(GL_3(\mathbb{F}_4))| = (362880)/(3) = 120960 \quad (5.62)$$

and

$$|P\Sigma L_3(\mathbb{F}_4)| = |\Sigma L_3(\mathbb{F}_4)||Z(SL_3(\mathbb{F}_4))| = (120960)/(3) = 40320. \tag{5.63}$$

Since the centers are identical, we have the following analogue of diagram (5.48):

$$PSL_3(\mathbb{F}_4) \lhd P\Gamma L_3(\mathbb{F}_4) \qquad \text{and} \qquad \begin{array}{ccc} P\Sigma L_3(\mathbb{F}_4) & \leq & P\Gamma L_3(\mathbb{F}_4) \\ \triangledown & & \triangledown \\ PSL_3(\mathbb{F}_4) & \lhd & PGL_3(\mathbb{F}_4) \end{array} . \tag{5.64}$$

CHAPTER 6

BILINEAR FORMS

In this chapter, we develop the theory of bilinear forms, in preparation for Chapter 7, where we use them to define quadratic forms over fields of characteristic 2. We narrow our focus from bilinear forms, to symmetric and alternate forms, and finally to alternate forms. The sections on orthogonality contain the most important results, which are later used in Chapter 10. In that chapter, a nondegenerate symmetric bilinear form, the dot product, plays an important role in proving facts about binary linear codes. We follow the treatment of Grove [Gro02].

6.1 Bilinear Forms

In this section, we define bilinear forms on a vector space V (of dimension n, over a field E). We show that the set of these forms is a subspace of all the maps $V \times V \to E$. Given a fixed basis for V, we identify this subspace with $M_n(E)$, the $n \times n$ matrices over E. We define a mapping $V \times V \to E$ for these matrices.

Definition 6.1.1. Let V denote a vector space of dimension n over a field E. A *bilinear form* is a map $B : V \times V \to E$ that is linear in each variable as the other is held fixed. Thus

$$\begin{aligned} B(x+y,z) &= B(x,y) + B(y,z), \\ B(ax,y) &= aB(x,y), \\ B(x,y+z) &= B(x,y) + B(x,z), \quad \text{and} \\ B(x,ay) &= aB(x,y), \end{aligned} \tag{6.1}$$

for all $x, y, z \in V$ and all $a \in E$.

Example 6.1.2. Let E^n be composed of row vectors, and let $A \in M_n(E)$. Define $B : E^n \times E^n \to E$ such that $B(x, y) = xAy^t$. Now xAy^t is a 1×1 matrix, and so $B(x, y) \in E$. Matrix multiplication distributes over matrix addition, and if we let $x, y, z \in E^n$ and $a \in E$, we have

$$B(x + y, z) = (x + y)Az^t = ((x + y)A)z^t = (xA + yA)z^t$$
$$= xAz^t + yAz^t = B(x, z) + B(y, z) \quad (6.2)$$

and

$$B(ax, y) = (ax)Ay^t = ((ax)A)y^t = (a(xA))y^t = a(xAy^t) = aB(x, y). \quad (6.3)$$

Linearity in the second argument is similarly verified. Thus B is a bilinear form. If $A = I$, then the bilinear form is the standard *dot product*.

Definition 6.1.3. We denote the set of all bilinear forms defined on a vector space V as $\mathcal{B}(V)$.

Theorem 6.1.4. *$\mathcal{B}(V)$ is a subspace of $\mathcal{F}(V \times V) = \{f : V \times V \to E\}$.*

Proof. Let $V \cong E^n$. Now $\mathcal{F}(V \times V) = \{f \mid f : V \times V \to E\}$ is a vector space, by Lemma 2.6.3, and $\mathcal{B}(V) \subseteq \mathcal{F}(V \times V)$. We must show $\mathcal{B}(V)$ is a subspace of $\mathcal{F}(V \times V)$.

Let $B_1, B_2 \in \mathcal{B}(V)$. We have

$$(B_1+B_2)(x+y, z) = B_1(x+y, z)+B_2(x+y, z) = B_1(x, z)+B_1(y, z)+B_2(x, z)+B_2(y, z)$$
$$= B_1(x, z) + B_2(x, z) + B_1(y, z) + B_2(y, z) = (B_1 + B_2)(x, z) + (B_1 + B_2)(y, z) \quad (6.4)$$

and

$$(B_1 + B_2)(ax, y) = B_1(ax, y) + B_2(ax, y) = aB_1(x, y) + aB_2(x, y)$$
$$= a(B_1(x, y) + B_2(x, y)) = a(B_1 + B_2)(x, y). \quad (6.5)$$

Similarly, the second argument is also linear. Thus $B_1 + B_2 \in \mathcal{B}(V)$.

Now let $k \in E$ and $B \in \mathcal{B}(V)$. We have

$$kB(x+y,z) = k(B(x+y,z)) = k(B(x,z)+B(y,z)) = \\ k(B(x,z)) + k(B(y,z)) = kB(x,z) + kB(y,z) \quad (6.6)$$

for all $x, y, z \in V$, and if $a \in E$, then

$$kB(ax,y) = k(B(ax,y)) = k(aB(x,y)) = \\ ka(B(x,y)) = a(k(B(x,y))) = a(kB)(x,y). \quad (6.7)$$

Similarly, the second argument is also linear. Thus $kB \in \mathcal{B}(V)$.

Now $0 \in \mathcal{F}(V \times V)$ is the function such that $0(x,y) = 0$ for all $x, y \in V$. Let $x, y, z \in V$ and $a \in E$. We have

$$0(x+y,z) = 0 = 0+0 = 0(x,z) + 0(y,z) \quad (6.8)$$

and

$$0(ax,y) = 0 = a0 = a0(x,y). \quad (6.9)$$

Similarly, the second argument is also linear. Thus $0 \in \mathcal{B}(V)$, and we conclude $\mathcal{B}(V)$ is a subspace of $\mathcal{F}(V \times V)$. □

Theorem 6.1.5. *A bilinear form B is determined by its values on a basis β of V.*

Proof. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V over E. Thus for $x, y \in V$, we have $x = \sum_i x_i v_i$ and $y = \sum_j y_j v_j$, where $x_i, y_j \in E$. Using the bilinearity of B, we have

$$B(x,y) = B\left(\sum_i x_i v_i, \sum_j y_j v_j\right) = \sum_{i,j} x_i y_j B(v_i, v_j) \quad (6.10)$$

which is a sum of values of B on $\{v_1, \ldots, v_n\}$. □

Definition 6.1.6. Given a bilinear form B on V, with basis $\beta = \{v_1, \ldots, v_n\}$, we define the *representing matrix* of B with respect to β as

$$[B]_\beta = [b_{ij}] \in M_n(E), \text{ where } b_{ij} = B(v_i, v_j) \text{ for } 1 \leq i, j \leq n. \tag{6.11}$$

Theorem 6.1.7. *Let V be n-dimensional over E and fix a basis β. Then*

$$\psi_\beta : \mathcal{B}(V) \to M_n(E) \qquad \textit{such that} \qquad \psi_\beta(B) = [B]_\beta \tag{6.12}$$

is a vector space isomorphism.

Proof. Let $\beta = \{v_1, \ldots, v_n\}$, and let $B_1, B_2 \in \mathcal{B}(V)$. Now for $1 \leq i, j \leq n$, we have

$$([B_1+B_2]_\beta)_{ij} = (B_1+B_2)(v_i, v_j) = B_1(v_i, v_j)+B_2(v_i, v_j) = ([B_1]_\beta)_{ij}+([B_2]_\beta)_{ij} \tag{6.13}$$

and so

$$\psi_\beta(B_1 + B_2) = [B_1 + B_2]_\beta = [B_1]_\beta + [B_2]_\beta = \psi_\beta(B_1) + \psi_\beta(B_2). \tag{6.14}$$

Let $a \in E$ and $B \in \mathcal{B}(V)$. For $1 \leq i, j \leq n$, we have

$$([aB]_\beta)_{ij} = (aB)(v_i, v_j) = a(B(v_i, v_j)) = a([B]_\beta)_{ij}, \tag{6.15}$$

and so

$$\psi_\beta(aB) = [aB]_\beta = a[B]_\beta = a\psi_\beta(B). \tag{6.16}$$

Therefore, ψ_β is linear.

Now let $\psi_\beta(B_1) = \psi_\beta(B_2)$. Thus $[B_1]_\beta = [B_2]_\beta$, implying $B_1(v_i, v_j) = B_2(v_i, v_j)$ for $1 \leq i, j \leq n$. By Theorem 6.1.5, a bilinear form is determined by its values on a basis, and thus $B_1 = B_2$, and ψ_β is one-to-one. Finally, let $A \in M_n(E)$. Let $B : V \times V \to E$ be the mapping

$$B(x, y) = [x]_\beta A[y]_\beta^t \tag{6.17}$$

for $x, y \in V$. From Example 6.1.2, we know $B \in \mathcal{B}(V)$. For $1 \leq i, j \leq n$, we have

$$B(v_i, v_j) = [v_i]_\beta A [v_j]_\beta^t = e_i A e_j^t = A_{ij}. \tag{6.18}$$

Thus $[B]_\beta = A$, and ψ_β is onto. □

Corollary 6.1.8. *Let V be n-dimensional over E and fix a basis β. Then if $B \in \mathcal{B}(V)$ and $A \in M_n(E)$, then $[B]_\beta = A$ if and only if, for all $x, y \in V$,*

$$B(x, y) = [x]_\beta A {[y]_\beta}^t. \tag{6.19}$$

Proof. Let $\beta = \{v_1, \ldots, v_n\}$, and let $[B]_\beta = A$. Thus $B(v_i, v_j) = A_{ij}$ for $1 \leq i, j \leq n$, and for $x, y \in V$,

$$B(x, y) = B\left(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right) = \sum_{i,j=1}^{n} x_i y_j B(v_i, v_j) = \sum_{i,j=1}^{n} x_i a_{ij} y_j = [x]_\beta A [y]_\beta^t \tag{6.20}$$

using the bilinearity of B. The converse was shown in Theorem 6.1.7. □

6.2 Congruence Classes

In this section, we look at the congruence classes (orbits) of bilinear forms under the action of the general linear group. We then describe analogous results for the matrix representation of these forms.

Definition 6.2.1. For B a bilinear form on V and $T \in GL(V)$, define

$$B^T : V \times V \to E \qquad \text{such that} \qquad B^T(v, w) = B(vT^{-1}, wT^{-1}) \tag{6.21}$$

for all $v, w \in V$.

Theorem 6.2.2. *Definition 6.2.1 defines a right action of $GL(V)$ on $\mathcal{B}(V)$, the set of all bilinear forms defined on V.*

Proof. Let $T \in GL(V)$ and $B \in \mathcal{B}(V)$. We have $B^T(v,w) = B(vT^{-1}, wT^{-1})$ for all $v, w \in V$. Let $x, y, z \in V$ and $a \in E$. Using the linearity of T^{-1} and the bilinearity of B, we have

$$\begin{aligned} B^T(x+y,z) &= B((x+y)T^{-1}, zT^{-1}) = B(xT^{-1} + yT^{-1}, zT^{-1}) \\ &= B(xT^{-1}, zT^{-1}) + B(yT^{-1}, zT^{-1}) = B^T(x,z) + B^T(y,z) \end{aligned} \quad (6.22)$$

and

$$\begin{aligned} B^T(ax,y) &= B((ax)T^{-1}, yT^{-1}) = B(a(xT^{-1}), yT^{-1}) \\ &= aB(xT^{-1}, yT^{-1}) = aB^T(x,y). \end{aligned} \quad (6.23)$$

Thus the first argument of B^T is linear as the second argument is held fixed. The proof of the second argument's linearity is similar. Thus $B^T \in \mathcal{B}(V)$. Let $I \in GL(V)$ be the identity transformation. For $B \in \mathcal{B}(V)$ and $x, y \in V$, we have

$$B^I(x,y) = B(xI^{-1}, yI^{-1}) = B(xI, yI) = B(x,y) \quad (6.24)$$

and thus $B^I = B$. Now let $S, T \in GL(V)$, $B \in \mathcal{B}(V)$, and $x, y \in V$. We have

$$\begin{aligned} (B^S)^T(x,y) &= B^S(xT^{-1}, yT^{-1}) = B((xT^{-1})S^{-1}, (yT^{-1})S^{-1}) \\ &= B(x(T^{-1}S^{-1}), y(T^{-1}S^{-1})) = B(x(ST)^{-1}, y(ST)^{-1}) = B^{ST}(x,y) \end{aligned} \quad (6.25)$$

and thus $(B^S)^T = B^{ST}$. □

Theorem 6.2.3. *Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V, and let $B \in \mathcal{B}(V)$ and $T \in GL(V)$. Then*

$$[B^T]_\beta = [T^{-1}]_\beta [B]_\beta [T^{-1}]_\beta{}^t. \quad (6.26)$$

Proof. By Corollary 6.1.8, we have

$$[B^T(v,w)]_\beta = [v]_\beta [B^T]_\beta [w]_\beta{}^t \quad (6.27)$$

for all $v, w \in V$. We also have

$$\begin{aligned}
[B^T(v,w)]_\beta &= [B(vT^{-1}, wT^{-1})]_\beta = [vT^{-1}]_\beta [B]_\beta [wT^{-1}]_\beta{}^t \\
&= \left([v]_\beta [T^{-1}]_\beta\right) [B]_\beta \left([w]_\beta [T^{-1}]_\beta\right)^t \\
&= [v]_\beta \left([T^{-1}]_\beta [B]_\beta [T^{-1}]_\beta{}^t\right) [w]_\beta{}^t \qquad (6.28)
\end{aligned}$$

for all $v, w \in V$. We conclude

$$[B^T]_\beta = [T^{-1}]_\beta [B]_\beta [T^{-1}]_\beta{}^t. \qquad (6.29)$$

□

Corollary 6.2.4. *We have that $GL_n(E)$ acts on $M_n(E)$, such that if $P \in GL_n(E)$ and $A \in M_n(E)$, then $A^P = P^{-1}A(P^{-1})^t$.*

Proof. By Theorem 6.1.7, when a basis β of V is fixed, $\mathcal{B}(V)$ is isomorphic to $M_n(E)$. Thus we use the action of Theorem 7.3.2, which has the desired form for matrices, as shown in Theorem 6.2.3. □

Definition 6.2.5. Relative to a basis β, $GL_n(E)$ partitions $M_n(E)$ into orbits, called *congruence classes*. The congruence class of $A \in M_n(E)$ is $\{P^{-1}A(P^{-1})^t \mid P \in GL_n(E)\}$. If $B = P^{-1}A(P^{-1})^t$ for some $P \in GL_n(E)$, we say B is *congruent* to A.

Theorem 6.2.6. *Two matrices are congruent if and only if they represent the same bilinear form with respect to different bases.*

Proof. Let $[B]_\beta, [B]_\gamma$ represent the bilinear form B with respect to bases β and γ. Let Q be the change of coordinate matrix from γ to β, as in Definition 2.6.13. Thus Q^{-1} is the change of coordinate matrix from β to γ, by Lemma 2.6.14. We have, by

Corollary 6.1.8,

$$[x]_\beta[B]_\beta[y]_\beta^t = B(x,y) = [x]_\gamma[B]_\gamma[y]_\gamma^t$$
$$= ([x]_\beta Q^{-1})[B]_\gamma([y]_\beta Q^{-1})^t = [x]_\beta(Q^{-1}[B]_\gamma(Q^{-1})^t)[y]_\beta^t, \quad (6.30)$$

for all $x, y \in V$. Thus $[B]_\beta = Q^{-1}[B]_\gamma(Q^{-1})^t$, and the two matrices are congruent.

Let $B \in \mathcal{B}(V)$, with $\beta = \{v_1, \ldots, v_n\}$ as a basis for V. Let $[B]_\beta$ be congruent to $A \in M_n(E)$. Thus there is a $Q \in GL_n(E)$ such that $A = Q^{-1}[B]_\beta(Q^{-1})^t$. Define

$$w_i = \sum_{j=1}^{n}(Q^{-1})_{ij}\, v_j, \quad (6.31)$$

and let $\gamma = \{w_1, \ldots, w_n\}$. Since $[w_i]_\beta$ is the ith row of Q^{-1}, and Q^{-1} is invertible, we have γ as a basis for V, and therefore, Q^{-1} is the change of coordinate matrix from bases γ to β. This gives

$$[x]_\gamma A[y]_\gamma^t = [x]_\gamma(Q^{-1}[B]_\beta(Q^{-1})^t)[y]_\gamma^t = ([x]_\gamma Q^{-1})[B]_\beta([y]_\gamma Q^{-1})^t$$
$$= [x]_\beta[B]_\beta[y]_\beta^t = B(x,y) \quad (6.32)$$

for all $x, y \in V$. Thus by Corollary 6.1.8, $Q^{-1}[B]_\beta(Q^{-1})^t = A = [B]_\gamma$, and our congruent matrices represent the same bilinear form relative to different bases. □

Theorem 6.2.7. *Let $[B]_\beta, [B]_\gamma$ be two matrix representations of a bilinear form B with respect to bases β, γ. Then* $\det[B]_\beta = 0$ *if and only* $\det[B]_\gamma = 0$.

Proof. By Theorem 6.2.6, we know that there exists an invertible $P \in M_n(E)$ such that $[B]_\beta = P^{-1}[B]_\gamma(P^{-1})^t$. Thus

$$\det[B]_\beta = \det P^{-1}[B]_\gamma(P^{-1})^t = (\det P^{-1})(\det[B]_\gamma)(\det(P^{-1})^t) = (\det P^{-1})^2 \det[B]_\gamma. \quad (6.33)$$

Since $\det P^{-1} \neq 0$, our result follows. □

Definition 6.2.8. Let B and B' be bilinear forms defined on E^n. We say that B and B' are *equivalent under an invertible linear substitution* if

$$\begin{aligned} B'((x_1,\ldots,x_n),(y_1,\ldots,y_n)) \\ &= B((x_1a_{11}+\ldots+x_na_{n1},\ldots,x_1a_{1n}+\ldots+x_na_{nn}), \\ &\qquad (y_1a_{11}+\ldots+y_na_{n1},\ldots,y_1a_{1n}+\ldots+y_na_{nn})) \end{aligned} \quad (6.34)$$

where $[a_{ij}] = A \in M_n(E)$ is invertible. Thus $B'(x,y) = B(xA, yA)$.

Theorem 6.2.9. *Let $V = E^n$ and B, B' be bilinear forms on V. Then the following are equivalent:*

(1) There exists a basis $\{w_1,\ldots,w_n\}$ for V such that $B(e_i,e_j) = B'(w_i,w_j)$.

(2) There exist bases $\{v_1,\ldots,v_n\}$ and $\{w_1,\ldots,w_n\}$ for V such that $B(v_i,v_j) = B'(w_i,w_j)$.

(3) B and B' are in the same $GL(V)$ orbit.

(4) There exists an invertible matrix $P \in M_n(E)$ such that $[B'] = P^{-1}[B](P^{-1})^t$.

(5) B and B' are equivalent under an invertible linear substitution.

Proof. $(1a \Rightarrow 1b)$ Take $\{v_1,\ldots,v_n\} = \{e_1,\ldots,e_n\}$, such that $v_i = e_i$ for $1 \le i \le n$.

$(1b \Rightarrow 2)$ Assume there exist bases $\{v_1,\ldots,v_n\}$ and $\{w_1,\ldots,w_n\}$ such that $B(v_i,v_j) = B'(w_i,w_j)$. Let $T \in GL(V)$ be the unique transformation such that $v_iT = w_i$ for all i. Thus $B'(w_i,w_j) = B(v_i,v_j) = B(w_iT^{-1}, w_jT^{-1})$. Since B' is determined by its values on the basis vectors $\{w_i\}$, we have $B'(x,y) = B(xT^{-1}, yT^{-1})$ for all $x, y \in V$. Thus $B'(x,y) = B(xT^{-1}, yT^{-1}) = B^T(x,y)$, and B and B' are in the same $GL(V)$ orbit.

$(2 \Rightarrow 3)$ Assume that B and B' are in the same $GL(V)$ orbit. Thus $B' = B^T$ for some $T \in GL(V)$. By Theorem 6.2.3, for a given basis β, we have

$$[B']_\beta = [B^T]_\beta = [T^{-1}]_\beta [B]_\beta [T^{-1}]_\beta{}^t = P^{-1}[B]_\beta (P^{-1})^t, \tag{6.35}$$

where the invertibility of $P^{-1} = [T^{-1}]_\beta$ follows from the invertibility of T^{-1}.

$(3 \Rightarrow 4)$ Assume there exists an invertible matrix $P^{-1} \in M_n(E)$ such that $[B']_\beta = P^{-1}[B]_\beta (P^{-1})^t$, with respect to basis β. We have

$$\begin{aligned} B'([x]_\beta, [y]_\beta) &= [x]_\beta [B']_\beta [y]_\beta^t = [x]_\beta (P^{-1}[B]_\beta (P^{-1})^t)[y]_\beta^t \\ &= ([x]_\beta P^{-1})[B]_\beta ((P^{-1})^t [y]_\beta^t) = ([x]_\beta P^{-1})[B]_\beta ([y]_\beta P^{-1})^t = B([x]_\beta P^{-1}, [y]_\beta P^{-1}) \end{aligned} \tag{6.36}$$

where $P^{-1} \in M_n(E)$ is invertible. Thus B and B' are equivalent under an invertible linear substitution.

$(4 \Rightarrow 1a)$ Let B and B' be equivalent under an invertible linear substitution. Thus there exists an invertible $A \in M_n(E)$ such that
$B'([x]_\beta, [y]_\beta) = B([x]_\beta A, [y]_\beta A)$, where β is a basis for V. We have

$$B'([x]_\beta A^{-1}, [y]_\beta A^{-1}) = B(([x]_\beta A^{-1})A, ([y]_\beta A^{-1})A) = B([x]_\beta, [y]_\beta). \tag{6.37}$$

Let $\beta = \{e_1, \ldots, e_n\}$, the standard basis for E^n. Let $[e_i]_\beta A^{-1} = [v_i]_\beta$, the ith row of A^{-1}. Thus

$$B(e_i, e_j) = B([e_i]_\beta, [e_j]_\beta) = B'([e_i]_\beta A^{-1}, [e_j]_\beta A^{-1}) = B'([v_i]_\beta, [v_j]_\beta) = B'(v_i, v_j) \tag{6.38}$$

Since A^{-1} is invertible, its rows are linearly independent vectors. Thus $\{v_1, \ldots, v_n\}$ is a basis for V. $\square$

6.3 Symmetric and Alternate Forms

In this section, we define symmetric and alternate forms and describe their matrices. Each of these special types constitutes a subspace of the bilinear forms.

Definition 6.3.1. If $B(x, y) = B(y, x)$ for all $x, y \in V$, then B is a *symmetric bilinear form.*

Theorem 6.3.2. *A bilinear form is symmetric if and only if its associated matrix is symmetric.*

Proof. Let B be a symmetric bilinear form on a vector space V with basis β. And so $b_{ij} = B(v_i, v_j) = B(v_j, v_i) = b_{ji}$, and thus the associated matrix is symmetric. Conversely, if the matrix representation of B is symmetric, we know that B is symmetric on the basis vectors. Since each vector in V can be expressed as a linear combination of the vectors of β, we use the bilinearity of B to obtain

$$B(x, y) = B\left(\sum_{i=1}^{n} a_i v_i, \sum_{j=1}^{n} b_j v_j\right) = \sum_{i,j=1}^{n} a_i b_j B(v_i, v_j)$$
$$= \sum_{i,j=1}^{n} b_j a_i B(v_j, v_i) = B\left(\sum_{j=1}^{n} b_j v_j, \sum_{i=1}^{n} a_i v_i\right) = B(y, x) \quad (6.39)$$

for all $x, y \in V$. □

Theorem 6.3.3. *The symmetric bilinear forms are a subspace of bilinear forms.*

Proof. Let $x, y \in V$ and $k \in E$, and let B_1, B_2, B be symmetric bilinear forms. We have

$$(B_1 + B_2)(x, y) = B_1(x, y) + B_2(x, y) = B_1(y, x) + B_2(y, x) = (B_1 + B_2)(y, x) \quad (6.40)$$

and

$$(kB)(x, y) = k(B(x, y)) = k(B(y, x)) = (kB)(y, x). \quad (6.41)$$

Also, $0(x, y) = 0 = 0(y, x)$, and thus $0 \in \mathcal{B}(V)$ is symmetric. □

Definition 6.3.4. If $B(x,x) = 0$ for all $x \in V$, then B is an *alternate bilinear form.*

Theorem 6.3.5. *If B is an alternate form, then $B(x,y) = -B(y,x)$ for all $x, y \in V$.*

Proof. We have

$$0 = B(x+y, x+y) = B(x,x)+B(y,y)+B(x,y)+B(y,x) = B(x,y)+B(y,x), \quad (6.42)$$

implying $B(x,y) = -B(y,x)$. □

Remark 6.3.6. Because of the above property, alternate forms are often called *skew-symmetric* forms. However, if E is of characteristic 2, an alternate form B gives $B(x,y) = -B(y,x) = B(y,x)$ for all $x, y \in V$, and thus B is also symmetric.

Definition 6.3.7. We define $A \in M_n(E)$ as a *skew-symmetric* matrix if $A^t = -A$.

Theorem 6.3.8. *A bilinear form B is alternate if and only if its matrix, with respect to a basis β, is skew-symmetric and has all 0's on the diagonal.*

Proof. Let B be an alternate form on V over E, with an associated basis $\beta = \{v_1, \ldots, v_n\}$. Now $[B]_\beta = [b_{ij}] = B(v_i, v_j)$ for $1 \leq i, j \leq n$. Thus

$$[B]_\beta{}^t = [b_{ij}]^t = [b_{ji}] = B(v_j, v_i) = -B(v_i, v_j) = -[b_{ij}] = -[B]_\beta, \quad (6.43)$$

by Theorem 6.3.5. By definition, $B(v_i, v_i) = 0$ for $1 \leq i \leq n$. Thus $[B]_\beta$ has all 0's on the diagonal.

Now let $[B]_\beta{}^t = -[B]_\beta$, where $[B]_\beta$ has all 0's on the diagonal. Thus $B(v_i, v_i) = 0$ for all i. For $x \in V$, we have

$$B(x,x) = B\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n x_j v_j\right) = \sum_{i,j=1}^n x_i x_j B(v_i, v_j) = \sum_{i,j=1}^n b_{ij} x_i x_j$$

$$= \sum_{i<j} (b_{ij} x_i x_j + b_{ji} x_j x_i) + \sum_{i=j} x_i^2 b_{ii} = \sum_{i<j} (b_{ij} x_i x_j - b_{ij} x_i x_j) + \sum_{i=j} x_i^2 b_{ii} = 0 + 0 = 0 \quad (6.44)$$

Thus B is an alternate form. □

Theorem 6.3.9. *The alternate bilinear forms are a subspace of bilinear forms.*

Proof. Let $x, y \in V$ and $k \in E$, and let B_1, B_2, B be alternate bilinear forms. We have

$$(B_1 + B_2)(x, x) = B_1(x, x) + B_2(x, x) = 0 + 0 = 0 \tag{6.45}$$

and

$$(kB)(x, x) = k(B(x, x)) = k(0) = 0. \tag{6.46}$$

Now $0(x, x) = 0$ for all $x \in V$, and thus $0 \in \mathcal{B}(V)$ is alternate. □

Remark 6.3.10. Let $A, P \in M_n(E)$ with P invertible. It follows from Theorem 6.2.6 that A is symmetric (or, alternate) if and only if $P^{-1}A(P^{-1})^t$ is symmetric (or, alternate). Thus congruence preserves symmetric and alternate forms.

6.4 Reflexive Forms and Dual Spaces

In this section, we restrict ourselves to symmetric and alternate forms, known as reflexive forms. For these forms, $(v, w) \mapsto 0$ implies $(w, v) \mapsto 0$, and this reflexivity allows us to define orthogonal subspaces. In preparation for this definition, we consider dual spaces and the linear functionals that compose them.

Definition 6.4.1. Let V be an n-dimensional vector space over field E, with B a bilinear form on V. If $B(v, w) = 0$ for $v, w \in V$, then v is *orthogonal* to w, denoted $v \perp w$.

Definition 6.4.2. If orthogonality is a reflexive relation, that is, $B(v, w) = 0$ if and only if $B(w, v) = 0$, we say B is a *reflexive* form.

Theorem 6.4.3. *Symmetric and alternate bilinear forms are reflexive.*

Proof. A symmetric form is clearly reflexive. If B is alternate, then $B(v,w) = -B(w,v)$ by Theorem 6.3.5. If $B(v,w) = 0$, then $B(w,v) = -0 = 0$, and conversely. Thus an alternate form is reflexive. □

Remark 6.4.4. For the rest of Chapter 6, all bilinear forms considered will be reflexive.

Definition 6.4.5. Let V be an n-dimensional vector space over field E. A *linear functional* on V is a linear function $f : V \to E$. We have $f \in V^* = \mathrm{Hom}_E(V, E)$, where V^* is the *dual space* of V. We denote such functionals as $v^* \in V^*$.

Theorem 6.4.6. *If V is an n-dimensional vector space over E, then V^* is also an n-dimensional vector space over E.*

Proof. Let $x^*, y^* \in V^*$ and $a \in E$. Define $x^* + y^*$ such that $(x^* + y^*)(v) = x^*(v) + y^*(v)$ for all $v \in V$, and ax^* such that $(ax^*)(v) = a(x^*(v))$. It is routine to show $x^* + y^*$ and ax^* are in V^*. There is also $0^* \in V^*$ such that $0^*(v) = 0 \in E$ for all $v \in V$. The other axioms are easily verified, and rely on the field properties of E. Thus V^* is a vector space over E.

Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V. Define $\beta^* = \{v_1^*, \ldots, v_n^*\} \subseteq V^*$ such that $v_i^*(v_j) = \delta_{ij}$ for all i and j, where δ is the Kronecker delta, defined such that

$$\delta_{ij} = 1, \text{ if } i = j \qquad \text{and} \qquad \delta_{ij} = 0, \text{ if } i \neq j. \tag{6.47}$$

We extend this by linearity to a mapping of all $v \in V$. If $x^* \in V^*$ and $x^*(v_i) = a_i \in E$, then

$$\left(\sum_j x^*(v_j) v_j^* \right)(v_i) = \left(\sum_j a_j v_j^* \right)(v_i) = a_i = x^*(v_i). \tag{6.48}$$

Now our linear map x^* is determined by where it maps the basis β, and thus

$$x^* = \sum_i a_i v_i^* \tag{6.49}$$

and β^* generates V^*. If $\sum_i a_i v_i^* = 0^*$, the 0-functional, then

$$a_j = \sum_i a_i v_i^*(v_j) = 0^*(v_j) = 0 \tag{6.50}$$

for all j, and thus β^* is an independent set. Thus β^* is a basis for V^*. □

Definition 6.4.7. The basis $\beta^* = \{v_1^*, \ldots, v_n^*\}$ for V^* is the *dual basis* of V.

Definition 6.4.8. For vector space V with bilinear form B, define $L : V \to V^*$ and $R : V \to V^*$ such that

$$\begin{aligned} L(v) &= L_v, \text{ where } L_v(w) = B(v, w) \text{ for all } w \in V, \\ R(v) &= R_v, \text{ where } R_v(w) = B(w, v) \text{ for all } v \in V. \end{aligned} \tag{6.51}$$

The bilinearity of B ensures that L and R are well-defined.

Lemma 6.4.9. *The maps L and R are linear.*

Proof. Let $v, w \in V$. We have $L(v + w) = L_{v+w}$, where, for all $x \in V$,

$$L_{v+w}(x) = B(v+w, x) = B(v, x) + B(w, x) = L_v(x) + L_w(x) = (L_v + L_w)(x). \tag{6.52}$$

Thus $L(v + w) = L(v) + L(w)$. For $a \in E$ and $v \in V$, we have $L(av) = L_{av}$, where, for all $x \in V$,

$$L_{av}(x) = B(av, x) = aB(v, x) = aL_v(x) = (aL_v)(x). \tag{6.53}$$

Thus $L(av) = aL(v)$. The argument for the linearity of R is similar. □

Lemma 6.4.10. *$L = R$ if and only if B is symmetric, and* $\ker L = \ker R$ *if and only if B is reflexive.*

Proof. Now $L = R$ if and only if $L(v) = R(v)$ for all $v \in V$, if and only if $L_v(w) = R_v(w)$ for all $w \in W$, if and only if $B(v, w) = B(w, v)$ for all $v, w \in V$.

Thus $L = R$ is equivalent to B being symmetric. Similarly, $\ker L = \ker R$ is equivalent to the statement "$B(v, w) = 0$ if and only if $B(w, v) = 0$." By Definition 6.4.2, this is reflexivity. □

Definition 6.4.11. For a reflexive form B on V, we define the *radical of* V relative to B as

$$\operatorname{rad}_B V = \{v \in V \mid B(v, w) = 0, w \in V\} = \{v \in V \mid B(w, v) = 0, w \in V\}. \quad (6.54)$$

Thus $\operatorname{rad}_B = \ker L = \ker R$.

Definition 6.4.12. If B is a bilinear form on V with basis β, B is *nondegenerate* if $\det[B]_\beta \neq 0$. By Theorem 6.2.7, a nonzero determinant does not depend on the basis β.

Lemma 6.4.13. *The standard dot product, defined as $x \cdot y = xIy^t = xy^t$ for $x, y \in E^n$, is a nondegenerate symmetric form for a vector space V over a field E.*

Proof. Now $E^n \times E^n \to E$ such that $(x, y) \mapsto xIy^t$ is a bilinear form, as in Example 6.1.2. Since $xIy^t = yIx^t$, it is symmetric, and since $\det I = 1 \neq 0$, it is nondegenerate. □

Theorem 6.4.14. *A reflexive bilinear form B on V is nondegenerate if and only if* $\operatorname{rad}_B V = 0$.

Proof. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V, with $[B]_\beta$ the matrix representation of B with respect to β. Let $x = \sum_{i=1}^n a_i v_i$ for $a_i \in E$. Suppose $x \in \operatorname{rad}_B V$. Then for all j,

$$0 = B(x, v_j) = B\left(\sum_i a_i v_i, v_j\right) = \sum_i a_i B(v_i, v_j) = \sum_i a_i b_{ij}, \quad (6.55)$$

which follows if and only if the row vector $[x]_\beta = (a_1, \ldots, a_n)$ satisfies $[x]_\beta[B]_\beta = 0$. If B is nondegenerate, then $\det[B]_\beta \neq 0$. This implies $[B]_\beta$ is invertible, implying $[x]_\beta = 0$. Thus $\operatorname{rad}_B V = 0$.

If $\operatorname{rad}_B V = 0$, then only $[x]_\beta = (0, \ldots, 0)$ satisfies $[x]_\beta[B]_\beta = 0$. Thus the null space of $[B]_\beta$ is of dimension 0, and $[B]_\beta$ is invertible, implying $\det[B]_\beta \neq 0$. By definition, B is nondegenerate. □

Corollary 6.4.15. *A reflexive bilinear form B on V is nondegenerate if and only if*

$$V^* = L(V) = R(V), \tag{6.56}$$

that is, if $v^* \in V^*$, there exist $x, y \in V$ such that

$$B(x, w) = L_x(w) = v^*(w) = R_y(w) = B(w, y) \tag{6.57}$$

for all $w \in V$.

Proof. Let reflexive B be nondegenerate. By Theorem 6.4.14, this implies

$$\operatorname{rad}_B V = \ker L = \ker R = 0. \tag{6.58}$$

Using the L function, we have by the dimension theorem,

$$n = \dim V = \dim \ker L + \dim L(V) = 0 + \dim L(V) = \dim L(V) \tag{6.59}$$

and thus $\dim L(V) = n$. Since $L(V) \leq V^*$, and $\dim V = n = \dim V^*$, we conclude $L(V) = V^*$.

Conversely, let $V^* = L(V) = R(V)$. Using the dimension theorem, and the fact that $\dim V = n = \dim V^*$, we have $\dim \ker L = \dim \ker R = 0$. We conclude $\operatorname{rad}_B V = 0$, which by Theorem 6.4.14, implies B is nondegenerate. □

6.5 Orthogonal Complements

In this section, we define orthogonal subspaces relative to reflexive forms. If a reflexive form B is nondegenerate (has a matrix of nonzero determinant), then the dimension of the orthogonal subspace $W^\perp$ is the codimension of $W \leq V$, and $W^{\perp\perp} = W$. We also define the radical of a vector subspace $W \leq V$ relative to a reflexive form B.

Definition 6.5.1. Let B be a reflexive bilinear form on V. For $S \subseteq V$, define

$$S^\perp = \{v \in V \mid B(v, w) = 0,\ w \in S\} = \{v \in V \mid B(w, v) = 0,\ w \in S\}. \tag{6.60}$$

Thus $v \in S^\perp$ if and only if $L_v|_S = 0$ if and only if $R_v|_S = 0$. We denote $\{x\}^\perp$ as $x^\perp$.

Theorem 6.5.2. *Let B be a reflexive form on V. If S and T are subsets of V, then $S \subseteq T$ implies $T^\perp \subseteq S^\perp$.*

Proof. Let $x \in T^\perp$. Thus $B(x, w) = 0$ for all $w \in T$. Now let $s \in S \subseteq T$, implying $s \in T$. Thus $B(x, s) = 0$, and since $s \in S$ is arbitrary, $x \in S^\perp$. Thus $T^\perp \subseteq S^\perp$. □

Theorem 6.5.3. *Let B be a reflexive form on V, and let S be a subset of V. We have $S \subseteq S^{\perp\perp}$.*

Proof. Let $s \in S$ and $x \in S^\perp$. Thus $B(s, x) = 0$ and so
$s \in \{v \in V \mid B(v, x) = 0, x \in S^\perp\} = S^{\perp\perp}$. □

Theorem 6.5.4. *If $S \subseteq V$, then $S^\perp$ is a subspace of V.*

Proof. We have

$$B(0, w) = B(v - v, w) = B(v, w) - B(v, w) = 0 \tag{6.61}$$

for all $w \in S$. Thus $0 \in S^\perp$. Now let $x, y \in S^\perp$. We have

$$B(x + y, w) = B(x, w) + B(y, w) = 0 + 0 = 0 \tag{6.62}$$

for all $w \in S$. Thus $x + y \in S^{\perp}$. Now let $x \in S^{\perp}$ and $a \in E$. We have

$$B(ax, w) = aB(x, w) = a(0) = 0 \tag{6.63}$$

for all $w \in S$. Thus $ax \in S^{\perp}$. We conclude $S^{\perp} \leq V$. □

Definition 6.5.5. If W is a subspace of V, then $W^{\perp}$ is the *orthogonal complement* of W.

Lemma 6.5.6. *Let B be a nondegenerate form on V, and W a subspace of V. If $w^* \in W^*$, there exist $x, y \in V$ such that $w^* = L_x|_W = R_y|_W$.*

Proof. We extend a basis $\{v_1, \ldots, v_k\}$ for W to a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V. Now extend $w^* \in W^*$ to $\widetilde{w}^* \in V^*$ such that $\widetilde{w}^*|_W = w^*$ and $\widetilde{w}^*(v_i) = 0$ for $i > k$. By Corollary 6.4.15, there exist $x, y \in V$ such that $\widetilde{w}^* = L_x = R_y$, and therefore, $w^* = L_x|_W = R_y|_W$. □

Theorem 6.5.7. *If B is a reflexive nondegenerate form on V, and W is a subspace of V, then*

$$\dim W^{\perp} = \dim V - \dim W. \tag{6.64}$$

Proof. By Lemma 6.5.6, the map $L|_W : V \to W^*$, where $L|_W(v) = L_v|_W$, is onto W^*. We have $\ker L|_W = W^{\perp}$, by definition of $W^{\perp}$. By the dimension theorem:

$$\dim V = \dim W^* + \dim W^{\perp} = \dim W + \dim W^{\perp}. \tag{6.65}$$

Rearranging terms gives our result. □

Corollary 6.5.8. *If B is a reflexive nondegenerate form on V, and W is a subspace of V, then $W^{\perp\perp} = W$.*

Proof. By Theorems 6.5.3 and 6.5.4, $W \leq W^{\perp\perp}$. By Theorem 6.5.7,

$$\dim W^{\perp\perp} = \dim V - \dim W^{\perp} = \dim V - (\dim V - \dim W) = \dim W. \tag{6.66}$$

Since W and $W^{\perp\perp}$ are of the same dimension, and $W \leq W^{\perp\perp}$, $W = W^{\perp\perp}$. □

Theorem 6.5.9. *Let B be a reflexive form on V over E. If $x \in V$, then $x^{\perp} = \langle x \rangle^{\perp}$.*

Proof. We have $\{x\} \subseteq \langle x \rangle$, and thus by Theorem 6.5.2, $\langle x \rangle^{\perp} \subseteq x^{\perp}$. So let $v \in x^{\perp}$. Thus $B(v, x) = 0$. Now $\langle x \rangle = \{ax \mid a \in E\}$. If $ax \in \langle x \rangle$, then

$$B(v, ax) = aB(v, x) = a(0) = 0. \tag{6.67}$$

Thus $v \in \langle x \rangle^{\perp}$, and therefore, $x^{\perp} \subseteq \langle x \rangle^{\perp}$. Our conclusion follows. □

Theorem 6.5.10. *Let B be a reflexive form on V. If W_1 and W_2 are subspaces of V, then $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$.*

Proof. Now $(W_1 + W_2)^{\perp} = \{v \in V \mid B(v, w_1 + w_2) = 0,\ w_1 + w_2 \in W_1 + W_2\}$, and W_1 and W_2 are each subspaces of $W_1 + W_2$ (setting w_1 or w_2 equal to 0). Thus by Theorem 6.5.2, we have

$$(W_1 + W_2)^{\perp} \subseteq W_1^{\perp} \qquad \text{and} \qquad (W_1 + W_2)^{\perp} \subseteq W_2^{\perp}. \tag{6.68}$$

Therefore,

$$(W_1 + W_2)^{\perp} \subseteq W_1^{\perp} \cap W_2^{\perp}. \tag{6.69}$$

Now let $v \in W_1^{\perp} \cap W_2^{\perp}$. Thus $B(v, w_1) = 0 = B(v, w_2)$ for all $w_1 \in W_1$, $w_2 \in W_2$. If $x \in W_1 + W_2$, then $x = w_1 + w_2$ for $w_1 \in W_1$ and $w_2 \in W_2$. In this case, we have

$$B(v, x) = B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2) = 0 + 0 = 0. \tag{6.70}$$

Since $x \in W_1 + W_2$ was arbitrary, we conclude $v \in (W_1 + W_2)^{\perp}$. Thus

$$W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1 + W_2)^{\perp}. \tag{6.71}$$

□

Definition 6.5.11. Let B be a reflexive form on V. If W is a subspace of V, define the *radical of* W as

$$\operatorname{rad}_B W = W \cap W^\perp = \{v \in W \mid B(v, w) = 0, w \in W\}. \tag{6.72}$$

Note that $\operatorname{rad}_B V = V \cap V^\perp = V^\perp = \{w \in V \mid B(v, w) = 0, v \in V\}$ agrees with Definition 6.4.11.

Definition 6.5.12. Given a reflexive form defined on V, we say W is a *nondegenerate subspace* of V if $\operatorname{rad}_B W = W \cap W^\perp = 0$.

Corollary 6.5.13. *Let B be a reflexive form on V, and let $B|_W$ be the restriction of B to a subspace W of V. We have W nondegenerate if and only if $B|_W$ is a nondegenerate form on W.*

Proof. This follows immediately from Theorem 6.4.14. □

Theorem 6.5.14. *Suppose B is a reflexive bilinear form on V, and W is a nondegenerate subspace of V. Then $V = W \oplus W^\perp$.*

Proof. Let V be n-dimensional, and let W be of dimension $k \leq n$. By Theorem 6.5.7, $W^\perp$ is of dimension $n - k$. Since W is nondegenerate, $W \cap W^\perp = 0$. This implies $W + W^\perp = W \oplus W^\perp$, by definition of direct sum from Friedberg, Insel and Spence [FIS03]. In general, for subspaces W_1 and W_2 of V, we have

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2), \tag{6.73}$$

also from Friedberg, Insel and Spence [FIS03]. Thus

$$\dim(W + W^\perp) = \dim W + \dim W^\perp - \dim(W \cap W^\perp) = k + (n - k) + 0 = n. \tag{6.74}$$

Now $W + W^\perp \leq V$, which is of dimension n, and thus $W \oplus W^\perp = V$. □

Definition 6.5.15. Let W_1 and W_2 be subspaces of V, and let B be a reflexive form defined on V. If $W_1 \cap W_2 = 0$, and $B(w_1, w_2) = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$, then we denote the *orthogonal direct sum* of W_1 and W_2 as $W_1 \perp W_2$.

Corollary 6.5.16. *If B is a reflexive form on V, and W_1 and W_2 are subspaces of V such that $W_1 \perp W_2$, then*

$$\operatorname{rad}_B(W_1 \perp W_2) = \operatorname{rad}_B W_1 \oplus \operatorname{rad}_B W_2. \tag{6.75}$$

Proof. Let $x \in \operatorname{rad}_B(W_1 \perp W_2)$. By Definition 6.5.11, $x \in (W_1 \perp W_2) \cap (W_1 \perp W_2)^\perp$, implying $x \in (W_1 \oplus W_2) \cap (W_1 \oplus W_2)^\perp$. Thus $x = w_1' + w_2'$ for $w_l' \in W_1$ and $w_2' \in W_2$, such that for all $w_1 + w_2 \in W_1 \oplus W_2$,

$$B(w_1 + w_2, w_1' + w_2') = 0. \tag{6.76}$$

Now $W_1 \leq W_1 \oplus W_2$, and so for $w_1 \in W_1$,

$$0 = B(w_1, w_1' + w_2') = B(w_1, w_1') + B(w_1, w_2') = B(w_1, w_1') + 0, \tag{6.77}$$

because $W_1 \perp W_2$. Thus $B(w_1, w_1') = 0$, and $w_1' \in W_1 \cap W_1^\perp = \operatorname{rad}_B W_1$. Similarly, $w_2' \in \operatorname{rad}_B W_2$. Thus $w_1' + w_2' \in \operatorname{rad}_B W_1 + \operatorname{rad}_B W_2$. Now $\operatorname{rad}_B W_1 \leq W_1$ and $\operatorname{rad}_B W_2 \leq W_2$. Thus if $W_1 \cap W_2 = 0$, then $\operatorname{rad}_B W_1 \cap \operatorname{rad}_B W_2 = 0$ as well. Thus $w_1' + w_2' \in \operatorname{rad}_B W_1 \oplus \operatorname{rad}_B W_2$, and $\operatorname{rad}_B(W_1 \perp W_2) \subseteq \operatorname{rad}_B W_1 \oplus \operatorname{rad}_B W_2$.

Now let $x \in \operatorname{rad}_B W_1 \oplus \operatorname{rad}_B W_2$. Thus $x = w_1' + w_2'$ such that $w_1' \in \operatorname{rad}_B W_1$ and $w_2' \in \operatorname{rad}_B W_2$. Thus $w_1' \in W_1 \cap W_1^\perp$ and $w_2' \in W_2 \cap W_2^\perp$. Thus $w_1' + w_2' \in W_1 + W_2$ and $w_1' + w_2' \in W_1^\perp + W_2^\perp$. Because $W_1 \perp W_2$, for all $w_1 + w_2 \in W_1 + W_2$ we have

$$B(w_1+w_2, w_1'+w_2') = B(w_1, w_1')+B(w_1, w_2')+B(w_2, w_1')+B(w_2, w_2') = 0+0+0+0 = 0. \tag{6.78}$$

Thus $w_1' + w_2' \in (W_1 + W_2) \cap (W_1 + W_2)^\perp = \operatorname{rad}_B(W_1 + W_2) = \operatorname{rad}_B(W_1 \perp W_2)$. Therefore, $\operatorname{rad}_B W_1 \oplus \operatorname{rad}_B W_2 \subseteq \operatorname{rad}_B(W_1 \perp W_2)$, and equality follows. □

6.6 Classification of Alternate Forms

In this section, we give an orthogonal decomposition for a vector space V, relative to an alternate form B. We decompose V into 2-dimensional subspaces, called hyperbolic planes, with the radical of V as the remaining subspace. Thus our alternate form B is of even rank, and if B is nondegenerate, V is of even dimension.

Lemma 6.6.1. *Let B be an alternate form on V, and let $u, v \in V$. If $B(u, v) \neq 0$, then $\{u, v\}$ is linearly independent.*

Proof. Using the contrapositive, we let $\{u, v\}$ be linearly dependent. Thus, without loss of generality, $v = au$ for some $a \in E$. Thus

$$B(u, v) = B(u, au) = aB(u, u) = a(0) = 0. \tag{6.79}$$

□

Definition 6.6.2. If B is an alternate form over V, and if $B(u, v) = 1$, then (u, v) is called a *hyperbolic pair* and the subspace $W = \langle u, v \rangle$ is called a *hyperbolic plane*.

Remark 6.6.3. If $B(u, v) = a \neq 0$, then we know by Lemma 6.6.1 that $\{u, v\}$ is an independent set. If we set $u_1 = u$ and $v_1 = a^{-1}v$, then

$$B(u_1, v_1) = B(u, a^{-1}v) = a^{-1}B(u, v) = a^{-1}a = 1, \tag{6.80}$$

and we may normalize (u, v) to obtain (u_1, v_1) as a hyperbolic pair, such that $\langle u_1, v_1 \rangle = \langle u, v \rangle$.

Lemma 6.6.4. *Let B be an alternate form on V, and let W be a hyperbolic plane in V. Then if $\beta = \{u, v\}$ is a hyperbolic pair generating W, the restriction of B to W has representing matrix* $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ *relative to β.*

Proof. Since $\beta = \{u, v\}$ is an independent set, we can can take it as a basis for W. Restrict B to W. Thus $[B|_W]_\beta = [b_{ij}]$ is a 2×2 matrix. We have

$$b_{11} = B(u, u) = 0 = B(v, v) = b_{22} \tag{6.81}$$

and

$$b_{12} = B(u, v) = 1 \qquad \Rightarrow \qquad -1 = B(v, u) = b_{21}. \tag{6.82}$$

□

Theorem 6.6.5. *If B is an alternate form on V, then*

$$V = W_1 \perp W_2 \perp \ldots \perp W_r \perp \operatorname{rad} V \tag{6.83}$$

a direct sum of mutually orthogonal subspaces with each W_i a hyperbolic plane. Thus V has a basis

$$\{u_1, v_1, u_2, v_2, \ldots, u_r, v_r, x_1, \ldots, x_{n-2r}\} \tag{6.84}$$

under which the representing matrix $[B]_\beta$ has block diagonal form

$$\begin{bmatrix} M & & & 0 \\ & \ddots & & \\ & & M & \\ 0 & & & 0_{n-2r} \end{bmatrix}, \tag{6.85}$$

where $M = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Proof. If $B \equiv 0$, then for $v \in V$ we have $B(v, w) = 0$ for all $w \in V$. Thus $V^\perp = V$, and $\operatorname{rad} V = V \cap V^\perp = V$, so we are done.

Assume $B \not\equiv 0$. Therefore, there exist $u, v \in V$ such that $B(u, v) \neq 0$. Thus $\{u, v\}$ is linearly independent, by Lemma 6.6.1. We normalize (u, v) to obtain a

hyperbolic pair (u_1, v_1), and let $W_1 = \langle u_1, v_1 \rangle$. By Lemma 6.6.4, we know the matrix of B restricted to W_1, relative to base $\beta_1 = \{u_1, v_1\}$, is M. By Corollary 6.5.13, since

$$\det[B|_{W_1}]_{\beta_1} = \det M = \det \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = 1 \neq 0, \tag{6.86}$$

W_1 is a nondegenerate subspace of V. Thus by Theorem 6.5.14,

$$V = W_1 \perp W_1^{\perp}, \tag{6.87}$$

and

$$\operatorname{rad}_B V = V \cap V^{\perp} = V^{\perp} = (W_1 \perp W_1^{\perp})^{\perp} = W_1^{\perp} \cap W_1^{\perp\perp} = \operatorname{rad}_B W_1^{\perp}, \tag{6.88}$$

where the fourth equality follows from Theorem 6.5.10.

But now, since $\dim W_1^{\perp} = \dim V - 2$, the result follows by induction on n. □

Corollary 6.6.6. *An alternate form B has even rank, and if B is nondegenerate, then the dimension of V is even.*

Proof. We select a basis β that exhibits V as the orthogonal direct product of r hyperbolic planes M and $\operatorname{rad} V$. We have $\operatorname{rank} M = 2$ and $\operatorname{rank} B|_{\operatorname{rad}_B V} = 0$. Thus $\operatorname{rank} B = 2r$. Let B be nondegenerate. Thus $\operatorname{rad}_B V = 0$, and $\{u_1, v_1, \ldots, u_r, v_r\}$ is a basis for V, implying $\dim V = 2r$. □

CHAPTER 7

QUADRATIC FORMS IN CHARACTERISTIC 2

In fields not of characteristic 2, quadratic forms and symmetric bilinear forms mutually determine one other. In a field of characteristic 2, this is not the case. We now examine quadratic forms over such fields, and identify quadratic forms with homogeneous quadratic polynomials. We describe regular quadratic forms, and use them for defining quadratic spaces. The final result is a particularly simple polynomial over a vector space of dimension 3, used in Chapter 8 to determine a regular conic in the projective plane $P^2(\mathbb{F}_q)$, where $\mathbb{F}_q$ is of characteristic 2. This chapter follows the treatment of Grove [Gro02].

7.1 Quadratic Forms

In this section, we define quadratic forms on a vector space $V = E^n$, where E is of characteristic 2. These forms are a subspace of all maps $V \to E$. A quadratic form Q has an associated alternate form B, which can be thought of as the "corrective term" in the near-additivity of a quadratic form. The choice of quadratic form Q determines a unique alternate bilinear form B, while a given alternate form B may have more than one quadratic form associated with it.

Definition 7.1.1. Let $V \cong E^n$, where E is of characteristic 2. A *quadratic form* is a map $Q : V \to E$ that satisfies

(1) $Q(av) = a^2 Q(v)$, for all $v \in V$ and $a \in E$; and

(2) $Q(u + v) = Q(u) + Q(v) + B(u, v)$,

where B is a bilinear form defined on V.

Theorem 7.1.2. *We have* $Q(0) = 0$.

Proof.

$$Q(0) = Q(0v) = 0^2 Q(v) = 0. \tag{7.1}$$

□

Theorem 7.1.3. *A quadratic form* Q *determines an alternate bilinear form* B.

Proof. Rearranging (2) above, we have

$$B(u, v) = Q(u + v) + Q(u) + Q(v). \tag{7.2}$$

Thus B is determined by Q. If $x \in V$, then

$$0 = Q(0) = Q(2x) = Q(x + x) = 2Q(x) + B(x, x) = B(x, x). \tag{7.3}$$

Since $B(x, x) = 0$ for $x \in V$, B is alternate. □

Example 7.1.4. In characteristic 2, the choice of alternate form B does not determine the quadratic form Q. For example, let V be of dimension 1, and let $B \equiv 0$. Define $Q_1(x) = 0$ for all $x \in V = E$. Let $a \in E$ and $x, y \in V$. We have

$$Q_1(ax) = 0 = a^2 0 = a^2 Q_1(x) \tag{7.4}$$

and

$$Q_1(x + y) = 0 = 0 + 0 + 0 = Q_1(x) + Q_1(y) + B(x, y). \tag{7.5}$$

Thus Q_1 is a quadratic form. Now define $Q_2(x) = x^2$ for $x \in V = E$. Let $a \in E$ and $x, y \in V$. We have

$$Q_2(ax) = (ax)^2 = a^2 x^2 = a^2 Q_2(x) \tag{7.6}$$

and

$$Q_2(x+y) = (x+y)^2 = x^2+y^2+2xy = x^2+y^2+0 = Q_2(x)+Q_2(y)+B(x,y). \quad (7.7)$$

Thus Q_2 is a quadratic form, and distinct quadratic forms Q_1, Q_2 are associated with the same alternate form $B \equiv 0$.

Definition 7.1.5. We denote the set of all quadratic forms Q defined on a vector space V as $\mathcal{Q}(V)$.

Theorem 7.1.6. *$\mathcal{Q}(V)$ is a subspace of $\mathcal{F}(V) = \{f : V \to E\}$.*

Proof. Let $V \cong E^n$. By Lemma 2.6.3, $\mathcal{F}(V)$ is a vector space. Now $\mathcal{Q}(V) \subseteq \mathcal{F}(V)$, and we must show it is a subspace. Let $Q_1, Q_2 \in \mathcal{Q}(V)$, with $B_1, B_2 \in \mathcal{B}(V)$ as their associated alternate forms. Let $x \in V$ and $a \in E$. We have

$$\begin{aligned}(Q_1+Q_2)(ax) = Q_1(ax)+Q_2(ax) &= a^2Q_1(x)+a^2Q_2(x) \\ &= a^2(Q_1(x)+Q_2(x)) = a^2(Q_1+Q_2)(x). \quad (7.8)\end{aligned}$$

and

$$\begin{aligned}(Q_1+Q_2)(x+y) &= Q_1(x+y)+Q_2(x+y) \\ &= Q_1(x)+Q_1(y)+B_1(x,y)+Q_2(x)+Q_2(y)+B_2(x,y) \\ &= Q_1(x)+Q_2(x)+Q_1(y)+Q_2(y)+B_1(x,y)+B_2(x,y) \\ &= (Q_1+Q_2)(x)+(Q_1+Q_2)(y)+(B_1+B_2)(x,y), \quad (7.9)\end{aligned}$$

where $B_1+B_2 \in \mathcal{B}(V)$, by Theorem 6.1.4. Thus $Q_1+Q_2 \in \mathcal{Q}(V)$.

Now let $k \in E$ and $Q \in \mathcal{Q}(V)$, with $B \in \mathcal{B}(V)$ its associated alternate form. For $a \in E$, we have

$$(kQ)(ax) = k(Q(ax)) = k(a^2Q(x)) = a^2(k(Q(x))) = a^2(kQ)(x) \quad (7.10)$$

and

$$\begin{aligned}(kQ)(x+y) &= k(Q(x+y)) = k(Q(x)+Q(y)+B(x,y)) \\ &= k(Q(x))+k(Q(y))+k(B(x,y)) \\ &= (kQ)(x)+(kQ)(y)+(kB)(x,y) \end{aligned} \tag{7.11}$$

where $kB \in \mathcal{B}(V)$, by Theorem 6.1.4. Thus $kQ \in \mathcal{Q}(V)$. Example 7.1.4 showed that $Q \equiv 0$ is a quadratic form, and so $0 \in \mathcal{Q}(V)$. Thus $\mathcal{Q}(V) \leq \mathcal{F}(V)$. □

7.2 Quadratic Forms and Homogeneous Polynomials

In this section, we identify quadratic forms with the set of quadratic homogenous polynomials, by use of the polarization identity.

Definition 7.2.1. Let $E[x_1, \ldots, x_n]$ be the ring of polynomials with coefficients from a field E. A *homogeneous polynomial* is a polynomial $f \in E[x_1, \ldots, x_n]$ where each term of f has the same total degree. If this degree is 2, then f is a *quadratic* homogeneous polynomial. A term of the form $x_i x_j$ where $i \neq j$ is called a *mixed* term of f, while one of the form x_i^2 is called an *unmixed* term of f.

Lemma 7.2.2. *Let $V = E^n$. For $1 \leq i \leq n$ and $1 \leq j < k \leq n$, we define $f, g : V \to E$ such that for $x = (x_1, \ldots, x_n) \in V$, we have $f(x) = x_j^2$ and $g(x) = x_j x_k$. Then $f, g \in \mathcal{Q}(V)$.*

Proof. Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be in E^n and $a \in E$. Thus

$$f(ax) = f(ax_1, \ldots, ax_n) = (ax_j)^2 = a^2 x_j^2 = a^2 f(x), \tag{7.12}$$

and

$$f(x+y) = f(x_1+y_1, \ldots, x_n+y_n) = (x_j+y_j)^2 = x_j^2+y_j^2+2x_j y_j = f(x)+f(y). \tag{7.13}$$

Now $B \equiv 0$ is an alternate bilinear form. Thus

$$f(x+y) = f(x) + f(y) + 0 = f(x) + f(y) + B(x,y). \tag{7.14}$$

Therefore, $f \in \mathcal{Q}(V)$.

We also have

$$g(ax) = g(ax_1, \ldots, ax_n) = (ax_j)(ax_k) = a^2(x_j x_k) = a^2 g(x) \tag{7.15}$$

and

$$\begin{aligned} g(x+y) &= g(x_1+y_1, \ldots, x_n+y_n) = (x_j+y_j)(x_k+y_k) \\ &= x_j x_k + y_j y_k + x_j y_k + x_k y_j = g(x) + g(y) + x_j y_k + x_k y_j. \end{aligned} \tag{7.16}$$

Let B be the bilinear form on V such that $B(e_j, e_k) = 1 = B(e_k, e_j)$, and is 0 on any other combination of basis vectors $\{e_1, \ldots, e_n\}$. We have

$$B(x,y) = B((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{r,s} x_r y_s B(e_r, e_s) = x_j y_k + x_k y_j, \tag{7.17}$$

and so

$$g(x+y) = g(x) + g(y) + x_j y_k + x_k y_j = g(x) + g(y) + B(x,y). \tag{7.18}$$

Therefore, $g \in \mathcal{Q}(V)$. □

Corollary 7.2.3. *A homogeneous quadratic polynomial $f \in E[x_1, \ldots, x_n]$ determines a quadratic form Q on $V = E^n$.*

Proof. This follows from Theorem 7.1.6 and Lemma 7.2.2. □

Theorem 7.2.4. *(Polarization) Let Q be a quadratic form on a vector space V over a field E of characteristic 2. Let V have basis $\beta = \{e_1, \ldots, e_n\}$ and let B be the alternate form determined by Q. If $x = \sum_i x_i e_i$ for $x \in V$, then*

$$Q(x) = \sum_i x_i^2 Q(e_i) + \sum_{i<j} x_i x_j B(e_i, e_j). \tag{7.19}$$

Proof. By induction on the dimension n of V. Let $n = 1$, and $x = x_1 e_1 \in V$. We have

$$Q(x) = Q(x_1 e_1) = x_1^2 Q(e_1) + 0 = \sum_i x_i^2 Q(e_i) + \sum_{i<j} x_i x_j B(e_i, e_j). \tag{7.20}$$

Now assume equality holds for all vector spaces of dimension less than n. Let V be a vector space of dimension n, and let $W = \{v \in V \mid v_n = 0\}$. Thus W is a subspace of dimension $n-1$. Now $V = W \oplus \{ke_n | k \in E\}$. So for $x \in V$, $x = x' + x_n e_n$, where $x' \in W$. Thus for $x' = x_1 e_1 + \ldots + x_{n-1} e_{n-1}$, we have

$$\begin{aligned} Q(x) &= Q(x' + x_n e_n) = Q(x') + Q(x_n e_n) + B(x', x_n e_n) \\ &= Q\left(\sum_{i<n} x_i e_i\right) + Q(x_n e_n) + B\left(\sum_{i<n} x_i e_i, x_n e_n\right) \\ &= \left(\sum_{i<n} x_i^2 Q(e_i) + \sum_{i<j<n} x_i x_j B(e_i, e_j)\right) + x_n^2 Q(e_n) + \sum_{i<n} x_i x_n B(e_i, e_n) \\ &= \sum_i x_i^2 Q(e_i) + \sum_{i<j} x_i x_j B(e_i, e_j). \end{aligned} \tag{7.21}$$

□

Corollary 7.2.5. *A quadratic form Q on V is determined by its values on a basis β and the values of its associated alternate form B on β.* □

Corollary 7.2.6. *Let Q be a quadratic form on E^n. Then $Q(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)$, where f is a homogeneous quadratic polynomial.*

Proof. By the polarization identity,

$$Q(x_1, \ldots, x_n) = \sum_i x_i^2 Q(e_i) + \sum_{i<j} x_i x_j B(e_i, e_j). \tag{7.22}$$

Let $a_i = Q(e_i) \in E$ and $b_{ij} = B(e_i, e_j) \in E$. Thus

$$Q(x_1, \ldots, x_n) = \sum_i a_i x_i^2 + \sum_{i<j} b_{ij} x_i x_j = f(x_1, \ldots, x_n). \tag{7.23}$$

□

7.3 Congruence Classes

In this section, we define an action of the general linear group on quadratic forms, and consider the orbits of quadratic forms (and associated alternate forms) under this action.

Definition 7.3.1. For Q a quadratic form on V and $T \in GL(V)$, define

$$Q^T : V \to E \qquad \text{such that} \qquad Q^T(v) = Q(vT^{-1}) \tag{7.24}$$

for all $v \in V$.

Theorem 7.3.2. *Definition 7.3.1 defines a right action of $GL(V)$ on $\mathcal{Q}(V)$.*

Proof. Let $T \in GL(V)$ and $Q \in \mathcal{Q}(V)$. Let $x, y \in V$ and $a \in E$. Using the linearity of T^{-1}, we have

$$Q^T(ax) = Q((ax)T^{-1}) = Q(a(xT^{-1})) = a^2Q(xT^{-1}) = a^2Q^T(x) \tag{7.25}$$

and

$$\begin{aligned} Q^T(x+y) &= Q((x+y)T^{-1}) = Q(xT^{-1} + yT^{-1}) \\ &= Q(xT^{-1}) + Q(yT^{-1}) + B(xT^{-1}, yT^{-1}) = Q^T(x) + Q^T(y) + B^T(x,y). \end{aligned} \tag{7.26}$$

Thus $Q^T \in \mathcal{Q}(V)$. Let $I \in GL(V)$ be the identity transformation. For $Q \in \mathcal{Q}(V)$ and $x, y \in V$, we have

$$Q^I(x) = Q(xI^{-1}) = Q(xI) = Q(x) \tag{7.27}$$

and thus $Q^I = Q$. Now let $S, T \in GL(V)$, $Q \in \mathcal{Q}(V)$, and $x \in V$. We have

$$(Q^S)^T(x) = Q^S(xT^{-1}) = Q((xT^{-1})S^{-1}) = Q(x(T^{-1}S^{-1})) = Q(x(ST)^{-1}) = Q^{ST}(x) \tag{7.28}$$

and thus $(Q^S)^T = Q^{ST}$. □

Corollary 7.3.3. *Let $T \in GL(V)$. If Q has associated alternate form B, then Q^T has associated alternate form B^T, and* $(\operatorname{rad}_B V)T = \operatorname{rad}_{B^T} V$.

Proof. Let Q determine the alternate form B. From the proof of Theorem 7.3.2, we have

$$Q^T(x+y) = Q^T(x) + Q^T(y) + B^T(x,y). \tag{7.29}$$

Thus Q^T determines the alternate form B^T.

Let $r \in \operatorname{rad} V_B$ and $s \in V$. We have

$$B^T(rT, s) = B((rT)T^{-1}, sT^{-1}) = B(r, sT^{-1}) = 0. \tag{7.30}$$

Since $T \in GL(V)$, s and sT^{-1} in V are equally arbitrary, and so $rT \in \operatorname{rad}_{B^T} V$. Thus $(\operatorname{rad}_B V)T \subseteq \operatorname{rad}_{B^T} V$. Now let $r \in \operatorname{rad}_{B^T} V$ and $s \in V$. We have

$$B(rT^{-1}, sT^{-1}) = B^T(r, s) = 0. \tag{7.31}$$

Since s and sT^{-1} in V are equally arbitrary, we have $rT^{-1} \in \operatorname{rad}_B V$ and $(rT^{-1})T = r \in (\operatorname{rad}_B V)T$, implying $\operatorname{rad}_{B^T} V \subseteq (\operatorname{rad}_B V)T$. We conclude that $(\operatorname{rad}_B V)T = \operatorname{rad}_{B^T} V$. □

Definition 7.3.4. Let Q and Q' be bilinear forms defined on E^n with a basis β. We say that Q and Q' are *equivalent under an invertible linear substitution* if

$$Q'(x_1, \ldots, x_n) = Q(a_{11}x_1 + \ldots + a_{n1}x_n, \ldots, a_{1n}x_1 + \ldots + a_{nn}x_n) = Q((x_1, \ldots, x_n)A), \tag{7.32}$$

where $[a_{ij}] = A \in M_n(E)$ is invertible.

Theorem 7.3.5. *Let $V = E^n$. If Q and Q' are quadratic forms on V, with associated alternate forms B and B', then the following are equivalent:*

(1) There exists a basis $\{w_1, \ldots, w_n\}$ *such that* $Q(e_i) = Q'(w_i)$*, and* $B(e_i, e_j) = B'(w_i, w_j)$.

(2) There exist bases $\{v_1, \ldots, v_n\}$ *and* $\{w_1, \ldots, w_n\}$ *such that* $Q(v_i) = Q'(w_i)$*, and* $B(v_i, v_j) = B'(w_i, w_j)$.

(3) Q *and* Q' *are in the same* $GL(V)$ *orbit.*

(4) Q *and* Q' *are equivalent under an invertible linear substitution.*

Proof. Assume there exists a basis $\{w_1, \ldots, w_n\}$ such that $Q(e_i) = Q'(w_i)$ and $B(e_i, e_j) = B'(w_i, w_j)$. Take $v_i = e_i$ for $1 \leq i \leq n$. Thus $Q(v_i) = Q'(w_i)$ and $B(v_i, v_j) = B'(w_i, w_j)$.

Assume there exist bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ such that $Q(v_i) = Q'(w_i)$ and $B(v_i, v_j) = B'(w_i, w_j)$. Let $T \in GL(V)$ be the unique map such that $v_i T = w_i$, for all i. Thus

$$Q'(w_i) = Q(v_i) = Q(w_i T^{-1}) = Q^T(w_i) \tag{7.33}$$

and

$$B'(w_i, w_j) = B(v_i, v_j) = B(w_i T^{-1}, w_j T^{-1}) = B^T(w_i, w_j). \tag{7.34}$$

So by Theorem 7.2.4 we obtain $Q'(x) = Q^T(x)$ for all $x \in V$. Thus Q and Q' are in the same $GL(V)$ orbit.

Assume Q and Q' are in the same $GL(V)$ orbit. Thus $Q' = Q^T$ for some $T \in GL(V)$, and so $Q'(x) = Q^T(x) = Q(xT^{-1})$ for all $x \in E^n$. For the standard basis $\beta = \{e_1, \ldots, e_n\}$, we have $[T^{-1}]_\beta = A \in M_n(E)$, which is invertible because T^{-1} is invertible. For $x \in V$, we have $x = \sum_i x_i e_i$, and we consider Q as acting on coordinates such that $Q(x) = Q(x_1 e_1 + \ldots + x_n e_n) = Q(x_1, \ldots, x_n) = Q([x]_\beta)$. Thus

$$Q'([x]_\beta) = Q'(x) = Q(xT^{-1}) = Q([xT^{-1}]_\beta) = Q([x]_\beta [T^{-1}]_\beta) = Q([x]_\beta A). \tag{7.35}$$

Therefore, Q and Q' are equivalent under an invertible linear substitution.

Assume Q and Q' are equivalent under an invertible linear substitution. Thus there exists an invertible $A \in M_n(E)$ such that $Q'([x]_\beta) = Q([x]_\beta A)$ for all $[x]_\beta \in E^n$, where $\beta = \{e_1, \ldots, e_n\}$ is the standard basis. We have

$$Q'([x]_\beta A^{-1}) = Q(([x]_\beta A^{-1})A) = Q([x]_\beta). \tag{7.36}$$

Let $w_i = e_i A^{-1}$, the ith row of A^{-1}. Thus

$$Q(e_i) = Q([e_i]_\beta) = Q'([e_i]_\beta A^{-1}) = Q'([w_i]_\beta) = Q'(w_i). \tag{7.37}$$

Since Q and Q' are equivalent, their associated bilinear forms are as well. Thus

$$B(e_i, e_j) = B([e_i]_\beta, [e_j]_\beta) = B'([e_i]_\beta A^{-1}, [e_j]_\beta A^{-1}) = B'([w_i]_\beta, [w_j]_\beta) = B'(w_i, w_j). \tag{7.38}$$

Since A^{-1} is invertible, its rows are linearly independent vectors. Thus $\{w_1, \ldots, w_n\}$ is a basis for V. □

7.4 Classification of Regular Quadratic Forms

In this section, we define regular quadratic forms and quadratic spaces. A regular form has the property that $0 \in V$ is the only vector from the radical of V (relative to B) mapped to 0 by Q, implying the radical of V is of dimension 0 or 1. We also introduce additional terms: V is nondefective if it has a nondegenerate alternate form B, and a subspace is singular if it has a nonzero vector mapped to 0. We then describe an orthogonal decomposition of a quadratic space V into 2-dimensional hyperbolic planes, allowing Q to be written in one of three simple ways (one way if n is odd, and two ways if n is even). For $n = 3$, this polynomial will be used in Section 8.5 for finding the points of a regular conic.

Remark 7.4.1. For the remainder of this chapter, assume E is a finite field of characteristic 2, and that $V = E^n$ has a quadratic form Q defined on it, where Q has the associated bilinear form B.

Lemma 7.4.2. *If $\theta : E \to E$ is defined by $a^\theta = a^2$, then $\theta \in \mathrm{Aut}(E)$.*

Proof. Let $a, b \in E$. Thus

$$(ab)^\theta = (ab)^2 = (ab)(ab) = a^2b^2 = a^\theta b^\theta. \tag{7.39}$$

We also have

$$(a+b)^\theta = (a+b)^2 = (a+b)(a+b) = a^2 + b^2 + 2ab = a^2 + b^2 = a^\theta + b^\theta. \tag{7.40}$$

Thus θ is an endomorphism of E. Finally, suppose $a^\theta = b^\theta$. Thus $a^2 = b^2$, and so

$$0 = a^2 + b^2 = (a+b)^2. \tag{7.41}$$

Since fields have no zero-divisors, $a + b = 0$, implying $a = b$. Thus θ is one-to-one. Since E is finite, this implies θ is onto as well. Thus $\theta \in \mathrm{Aut}(E)$. □

Lemma 7.4.3. *If Q is restricted to $\mathrm{rad}_B V$, then $Q : \mathrm{rad}_B V \to E$ is a semilinear transformation.*

Proof. From Lemma 7.4.2, we have $\theta \in \mathrm{Aut}(E)$, where $a^\theta = a^2$. Let $x, y \in \mathrm{rad}_B V$ and let $a, b \in E$. We have

$$Q(ax + by) = a^2Q(x) + b^2Q(y) + abB(x, y) = a^\theta Q(x) + b^\theta Q(y). \tag{7.42}$$

Thus Q restricted to $\mathrm{rad}_B V$ is θ-semilinear. □

Definition 7.4.4. We say Q is *regular* on V if $\ker(Q|_{\mathrm{rad}_B V}) = 0$, where

$$\ker(Q|_{\mathrm{rad}_B V}) = \{v \in \mathrm{rad}_B V \mid Q(v) = 0\}. \tag{7.43}$$

Definition 7.4.5. If Q is regular on V, then (V, Q) is a *quadratic space*. If (V, Q) is a quadratic space, and Q is regular on $W \leq V$, then (W, Q) is a *quadratic subspace* of V.

Lemma 7.4.6. *If Q is regular on V, then* $\dim \operatorname{rad}_B V = 0$ *or* 1.

Proof. Let Q be regular. If $\dim \operatorname{rad}_B V = 0$, we are done. Assume $\dim \operatorname{rad}_B V \geq 1$. From Lemma 7.4.3, $Q|_{\operatorname{rad}_B V} : \operatorname{rad}_B V \to E$ is semilinear. Since $Q|_{\operatorname{rad}_B V}$ is one-to-one and $\dim E = 1$, our mapping is onto, making it an isomorphism. Thus $\dim \operatorname{rad}_B V = \dim E = 1$. □

Definition 7.4.7. We say V is *nondefective* if $\det[B]_\beta \neq 0$, where β is a basis for V. This is equivalent to $\operatorname{rad}_B V = 0$. If $\det[B]_\beta = 0$, we say V is *defective*. This is equivalent to $\operatorname{rad}_B V \neq 0$.

Definition 7.4.8. If $0 \neq v \in V$ and $Q(v) = 0$, we say v is a *singular vector*. If a subspace W of V has a singular vector, we say it is a *singular subspace*; otherwise, it is a *nonsingular subspace*.

Definition 7.4.9. If (V, Q, B) is a quadratic space and $u, v \in V$ are vectors such that $Q(u) = 0 = Q(v)$ and $B(u, v) = 1$, then (u, v) is a *hyperbolic pair* and the subspace $H = \langle u, v \rangle$ is a *hyperbolic plane*. Note that by Lemma 6.6.1, $\{u, v\}$ is linearly independent.

Lemma 7.4.10. *If (V, Q) is a quadratic space, and $H \leq V$ is a hyperbolic plane, then (H, Q) is a nondefective quadratic subspace of V.*

Proof. Since H is a hyperbolic plane, there exist linearly independent $u, v \in V$ such that $H = \langle u, v \rangle$, where $B(u, v) = 1$ and $Q(u) = 0 = Q(v)$. Let

$x \in \operatorname{rad}_B H = H \cap H^\perp$. Thus for some $a, b \in E$, we have $x = au + bv$ and $B(x, \alpha u + \beta v) = 0$ for all $\alpha, \beta \in E$. We have

$$0 = B(au + bv, \alpha u + \beta v) = a\alpha B(u,u) + b\beta B(v,v) + a\beta B(u,v) + b\alpha B(v,u)$$
$$= (a\beta + b\alpha)B(u,v) = a\beta + b\alpha. \quad (7.44)$$

Taking $\alpha = 0$ and $\beta = 1$, we see that $a = 0$, and taking $\beta = 0$ and $\alpha = 1$, we see that $b = 0$. Thus $x = 0u + 0v = 0$, and $\operatorname{rad}_B H = 0$, making H nondefective. Now $Q(0) = 0$, and so by definition, Q is regular on H, making H a quadratic subspace of V. □

Remark 7.4.11. For the rest of this section, we assume that V is a quadratic space of dimension at least 2.

Theorem 7.4.12. *If $u \in V \setminus \operatorname{rad}_B V$ is singular, then there exists a $v \in V$ such that (u, v) is a hyperbolic pair.*

Proof. Let $u \in V \setminus \operatorname{rad}_B V$ be singular. Since $u \notin \operatorname{rad}_B V$, there exists a $w \in V$ such that $B(u, w) \neq 0$. By normalizing w, we may assume that $B(u, w) = 1$. Let $v = Q(w)u + w \in V$. We have

$$Q(v) = Q(Q(w)u + w) = (Q(w))^2 Q(u) + Q(w) + B(Q(w)u, w)$$
$$= (Q(w))^2(0) + Q(w) + Q(w)B(u,w) = 2Q(w) = 0 \quad (7.45)$$

and

$$B(u, v) = B(u, Q(w)u + w) = Q(w)B(u,u) + B(u,w) = 0 + 1 = 1. \quad (7.46)$$

Thus (u, v) is a hyperbolic pair. □

Theorem 7.4.13. *If V is a defective quadratic space, and $\dim V \geq 2$, then V is singular and contains a hyperbolic plane.*

Proof. By Lemma 7.4.6, $\dim \operatorname{rad}_B V \leq 1$, and since B is defective, $\operatorname{rad}_B V \neq 0$. Thus $\dim \operatorname{rad}_B V = 1$. Now $V \neq \operatorname{rad}_B V$ because $\dim V \geq 2$. Let $0 \neq y \in \operatorname{rad}_B V$ and $x \in V \setminus \operatorname{rad}_B V$. Since Q is regular, $Q(y) \neq 0$. Thus $Q(x)/Q(y) = b \in E$. Since $E = E^2$ by Lemma 7.4.2, $b = a^2$ for some $a \in E$. Thus

$$Q(x) = a^2 Q(y). \tag{7.47}$$

Let $u = x + ay$. Since $y \in \operatorname{rad}_B V$, and $u - ay = x \notin \operatorname{rad}_B V$, $u \notin \operatorname{rad}_B V$. Since $y \in \operatorname{rad}_B V$, we have

$$\begin{aligned} Q(u) = Q(x + ay) &= Q(x) + a^2 Q(y) + B(x, ay) \\ &= Q(x) + Q(x) + aB(x, y) = 2Q(x) + a0 = 0 \end{aligned} \tag{7.48}$$

and thus u is a singular vector. By Theorem 7.4.12, V contains a hyperbolic plane. □

Theorem 7.4.14. *If V is a nondefective quadratic space, and $\dim V \geq 3$, then V is singular and contains a hyperbolic plane.*

Proof. Let $0 \neq x \in V$. If $Q(x) = 0$, we are done. Assume $Q(x) \neq 0$. By Theorem 6.5.4, $x^\perp$ is a subspace of V, and by Theorem 6.5.9 $x^\perp = \langle x \rangle^\perp$. Now $\dim \langle x \rangle^\perp = \dim V - \dim \langle x \rangle \geq 3 - 1 = 2$ by Theorem 6.5.7. Thus let $y \in x^\perp \setminus \langle x \rangle$. We have $Q(y)/Q(x) = b \in E$. Since $E = E^2$ by Lemma 7.4.2, $b = a^2$ for some $a \in E$. Thus

$$Q(y) = a^2 Q(x). \tag{7.49}$$

Let $u = ax + y$. Now $y \in x^\perp \setminus \langle x \rangle$, implying $y \neq ax \in \langle x \rangle$, and therefore, $u \neq 0$. Since $y \in x^\perp$, $B(x, y) = 0$. We have

$$Q(u) = Q(ax + y) = a^2 Q(x) + Q(y) + B(ax, y) = 2a^2 Q(x) + 0 = 0. \tag{7.50}$$

Thus u is a singular vector. By Theorem 7.4.12, V contains a hyperbolic plane. □

Theorem 7.4.15. *If V is a 2-dimensional nondefective and nonsingular quadratic space, there is a basis $\{v_1, v_2\}$ for V and an irreducible polynomial $x^2 + x + c \in E[x]$ such that*

$$Q(a_1v_1 + a_2v_2) = a_1^2 + a_1a_2 + ca_2^2 \tag{7.51}$$

for all $a_i \in E$. Conversely, suppose $x^2 + x + c \in E[x]$ is irreducible and $\{v_1, v_2\}$ is a basis for V, and define $Q : V \to E$ by $Q(a_1v_1 + a_2v_2) = a_1^2 + a_1a_2 + ca_2^2$, for all $a_i \in E$. Then Q is a quadratic form for which V is nondefective and nonsingular.

Proof. On the one hand, since V is nondefective, there exist $w_1, w_2 \in V$ such that, after normalizing, $B(w_1, w_2) = 1$. Thus

$$Q(a_1w_1 + a_2w_2) = a_1^2Q(w_1) + a_1a_2 + a_2^2Q(w_2). \tag{7.52}$$

Now $Q(w_1) \neq 0$, since V is not singular, and since $E^2 = E$ by Lemma 7.4.2, $Q(w_1) = b^2$, where $b \in E^\times$. Since $b \neq 0$, let $v_1 = b^{-1}w_1$ and $v_2 = bw_2$, and let $c = b^2Q(w_2)$. Thus $\{v_1, v_2\}$ is also a basis, and

$$\begin{aligned} Q(a_1v_1 + a_2v_2) &= Q((a_1b^{-1})w_1 + (a_2b)w_2) \\ &= a_1^2b^{-2}Q(w_1) + (a_1b^{-1})(a_2b) + a_2^2b^2Q(w_2) \\ &= a_1^2(b^{-2}b^2) + a_1a_2 + a_2^2b^2(b^{-2}c) \\ &= a_1^2 + a_1a_2 + ca_2^2. \end{aligned} \tag{7.53}$$

If $a \in E$, then $Q(av_1 + v_2) = a^2 + a + c$. Since $\{v_1, v_2\}$ is a basis for V, $av_1 \neq v_2$ for all $a \in E$. Thus $av_1 + v_2 \neq 0$ for all $a \in E$. Since V is nonsingular,

$$a^2 + a + c = Q(av_1 + v_2) \neq 0 \tag{7.54}$$

for all $a \in E$. Since $x^2 + x + c \in E[x]$ could only have linear factors, the polynomial is irreducible.

Conversely, let $x^2 + x + c \in E[x]$ be irreducible, and let $\{v_1, v_2\}$ be a basis for V. Define $Q : V \to E$ such that

$$Q(a_1v_1 + a_2v_2) = a_1^2 + a_1a_2 + ca_2^2 \tag{7.55}$$

for all $a_i \in E$. By Corollary 7.2.3, this is a quadratic form on V. We have

$$Q(v_1) = 1^2 + (1)(0) + c(0^2) = 1 \text{ and } Q(v_2) = 0^2 + (0)(1) + c(1^2) = c. \tag{7.56}$$

Comparing our definition of Q and Property (2) of quadratic forms, we have

$$c = 1^2 + (1)(1) + c(1^2) = Q(v_1 + v_2) = Q(v_1) + Q(v_2) + B(v_1, v_2) = 1 + c + B(v_1, v_2). \tag{7.57}$$

We conclude $1 = B(v_1, v_2) = B(v_2, v_1)$. Let $x \in \operatorname{rad}_B V$. Thus $x = d_1v_1 + d_2v_2$ for $d_1, d_2 \in E$. Now $B(x, v) = 0$ for all $v \in V$. Thus

$$0 = B(x, v_2) = B(d_1v_1 + d_2v_2, v_2) = d_1B(v_1, v_2) + d_2B(v_2, v_2) = d_1 + 0 \tag{7.58}$$

and $d_1 = 0$. We also have

$$0 = B(x, v_1) = B(d_1v_1 + d_2v_2, v_1) = d_1B(v_1, v_1) + d_2B(v_2, v_1) = 0 + d_2 \tag{7.59}$$

and $d_2 = 0$. Thus $x = 0$, and so $\operatorname{rad}_B V = 0$, making V nondefective. Also,

$$\ker(Q|_{\operatorname{rad}_B V}) = \ker Q|_0 = 0 \tag{7.60}$$

and so Q is regular, making V a quadratic space. Let $0 \neq y = a_1v_1 + a_2v_2 \in V$ be singular. Thus

$$0 = Q(y) = Q(a_1v_1 + a_2v_2) = a_1^2 + a_1a_2 + ca_2^2 \tag{7.61}$$

Now $a_2 = 0$ implies $a_1 = 0$. Since $y \neq 0$, $a_2 \neq 0$. Dividing by a_2^2 we obtain

$$0 = a_1^2/a_2^2 + a_1a_2/a_2^2 + ca_2^2/a_2^2 = (a_1/a_2)^2 + a_1/a_2 + c \tag{7.62}$$

But our polynomial is irreducible, and so $a_1/a_2 \in E$ cannot satisfy this equation. Thus no such y exists, and V is nonsingular. □

Lemma 7.4.16. *If (V, Q) is a quadratic space, and if (W, Q) is a nondefective quadratic subspace of V, then $V = W \perp W^{\perp}$, where $(W^{\perp}, Q)$ is a quadratic subspace of V.*

Proof. Let W be a nondefective quadratic subspace. Since $\operatorname{rad}_B W = 0$, W is a nondegenerate subspace relative to B, and thus $V = W \oplus W^{\perp} = W \perp W^{\perp}$, by Theorem 6.5.14. Now, by Corollary 6.5.16,

$$\operatorname{rad}_B V = \operatorname{rad}_B(W \perp W^{\perp}) = \operatorname{rad}_B W \oplus \operatorname{rad}_B W^{\perp} = \{0\} \oplus \operatorname{rad}_B W^{\perp} = \operatorname{rad}_B W^{\perp}. \tag{7.63}$$

Therefore, $0 = \ker(Q|_{\operatorname{rad}_B V}) = \ker(Q|_{\operatorname{rad}_B W^{\perp}})$, and so Q is regular on $W^{\perp}$, making $W^{\perp}$ a quadratic subspace of V. □

Lemma 7.4.17. *If $W_1 \perp W_2$, then $Q(w_1 + w_2) = Q(w_1) + Q(w_2)$ for all $w_1 \in W_1$ and $w_2 \in W_2$.*

Proof. Let $w_1 \in W_1$ and $w_2 \in W_2$. Thus

$$Q(w_1 + w_2) = Q(w_1) + Q(w_2) + B(w_1, w_2) = Q(w_1) + Q(w_2) + 0. \tag{7.64}$$

□

Theorem 7.4.18. *Let V be a quadratic space of dimension n over E, a finite field of characteristic 2. Then V has a basis $\beta = \{v_1, \ldots, v_n\}$ such that Q takes one of the following three forms:*

1. If $n = 2m + 1$ is odd, then

$$Q(\sum_i x_i v_i) = \sum_{i=1}^{m} x_{2i-1}x_{2i} + x_{2m+1}^2. \tag{7.65}$$

2. If $n = 2m$ is even, then

$$Q(\sum_i x_i v_i) = \sum_{i=1}^{m} x_{2i-1}x_{2i} \tag{7.66}$$

or

$$Q(\sum_i x_i v_i) = (\sum_{i=1}^{m-1} x_{2i-1}x_{2i}) + (x_{2m-1}^2 + x_{2m-1}x_{2m} + cx_{2m}^2), \tag{7.67}$$

where $x^2 + x + c \in E[x]$ *is irreducible.*

Proof. Let $n = 1$. Then $V = E$, and is generated by any $0 \neq v_1 \in E$. Since B is alternate, $B(v_1, v_1) = 0$, and thus $B \equiv 0$, which means $\operatorname{rad}_B V = V = E$. Since Q is regular on E, $Q(v_1) = a \neq 0$. Since $E = E^2$ by Lemma 7.4.2, $a = b^2$ for some $b \in E$, and we select $b^{-1}v_1 \neq 0$ for our basis vector. Thus

$$Q(x) = Q(x_1(b^{-1}v_1)) = x_1^2((b^{-1})^2 Q(v_1)) = x_1^2(a^{-1}a) = x_1^2, \tag{7.68}$$

and thus for $n = 1$, V has a basis such that Q is of form (7.65).

If $n = 2$, then V is either singular or nonsingular and either defective or nondefective. If V is singular and defective, then $\operatorname{rad}_B V = V$ by Theorem 6.6.5. Since $\dim V = 2$, this contradicts that Q is a regular quadratic form. If V is singular and nondefective, then by Theorem 7.4.12, there exist $v_1, v_2 \in V$ such that (v_1, v_2) is a hyperbolic pair, and $V = \langle v_1, v_2 \rangle$ is a hyperbolic plane. Thus

$$Q(x) = Q(x_1v_1 + x_2v_2) = x_1^2 Q(v_1) + x_2^2 Q(v_2) + x_1x_2B(v_1, v_2) = x_1x_2, \tag{7.69}$$

and there is a basis for V such that Q has form (7.66). If V is nonsingular, then by Theorem 7.4.13, V is nondefective, and so by Theorem 7.4.15, there exist $v_1, v_2 \in V$ such that $V = \langle v_1, v_2 \rangle$ and

$$Q(x_1v_1 + x_2v_2) = x_1^2 + x_1x_2 + cx_2^2, \tag{7.70}$$

where $x^2 + x + c \in E[x]$ is irreducible. In this case, there is a basis for V such that Q has form (7.67).

Proceeding by induction on n, suppose $n \geq 3$. Thus V has a hyperbolic plane H_1, by Theorem 7.4.13 or Theorem 7.4.14. Now H_1 is a nondefective quadratic

subspace by Lemma 7.4.10, and so by Lemma 7.4.16, $V = H_1 \perp H_1^{\perp}$, where $H_1^{\perp}$ is a quadratic subspace of dimension $n-2$. Then by Lemma 7.4.17,

$$Q(V) = Q(H_1) + Q(H_1^{\perp}). \tag{7.71}$$

Thus

$$Q(\sum_i x_i v_i) = Q(x_1 v_1 + x_2 v_2) + Q(\sum_{i>2} x_i v_i), \tag{7.72}$$

and the result follows by induction. □

Corollary 7.4.19. *Let Q be a regular quadratic form on an n-dimensional vector space over a finite field E of characteristic 2. Then Q, acting on coordinate vectors with respect to an appropriate basis β, has the form:*

$$n = 1 : Q(a_1) = a_1^2.$$

$$n = 2 : Q(a_1, a_2) = a_1 a_2, \quad \textit{or}$$

$$Q(a_1, a_2) = a_1^2 + a_1 a_2 + c a_2^2, \ \textit{where } x^2 + x + c \in E[x] \textit{ is irreducible.}$$

$$n = 3 : Q(a_1, a_2, a_3) = a_1 a_2 + a_3^2.$$

If (X, Y, Z) is an arbitrary coordinate vector in E^3, then any regular quadratic form Q is equivalent to the form $Q(X, Y, Z) = YZ + X^2$.

Proof. The first statement follows directly from Theorem 7.4.18. If (X, Y, Z) is an arbitrary coordinate vector in E^3, then it follows that any quadratic form Q is of the form $Q(X, Y, Z) = XY + Z^2$, for an appropriate basis β. By the change of basis exchanging X and Z, this is equivalent to $Q(X, Y, Z) = YZ + X^2$. □

CHAPTER 8

CURVES IN THE PROJECTIVE PLANE

In this chapter, we define projective space and then consider the projective plane. We take an algebraic geometry approach to the projective plane, considering curves as the zero sets of homogeneous polynomials. If the polynomial is linear, the curve is a line, and if the polynomial is quadratic, the curve is a conic. We define an action of $P\Gamma L_n(E)$ on curves by the mapping of zero sets of homogeneous polynomials. We show that regular conics in $P^2(\mathbb{F}_q)$ are sets of $q+1$ points, which can be extended to a set of $q+2$ points, a hyperconic, by adjoining a certain point. We examine lines in $P^2(\mathbb{F}_q)$, and determine the intersection properties between a line and a hyperconic. A more thorough investigation of the intersection properties of hyperconics must wait until Chapter 9, where a finite geometry approach is taken. This chapter follows the treatment of Rotman [Rot95] and Fulton [Ful08].

8.1 Projective Space

In this section, we define projective space $P^{n-1}(E)$ by taking the vector space E^n and considering the nonzero multiples of a nonzero vector as equivalent. Thus lines are reduced to points, and in general, subspaces of dimension k are of dimension $(k-1)$ in projective space. We calculate the order of various projective spaces over finite fields, and introduce homogeneous and non-homogeneous coordinates.

For convenience, we first repeat some definitions from Section 4.2.

Definition 8.1.1. If E is a field, then *affine n-space* is $E^n = \{(v_1, \ldots, v_n) | v_i \in E\}$. We also define $E^{n\#} = E^n \setminus \{0\}$.

Definition 8.1.2. Let $v \in E^{n\#}$, and define $[v] = \{\alpha v \mid \alpha \in E^\times\}$. We then define *projective n-space* as $P^n(E) = \{[v] \mid v \in E^{n+1\#}\}$.

Definition 8.1.3. Let $\langle w \rangle$ be a 1-dimensional subspace of E^{n+1}. Let $Gr_1(E^{n+1})$ be the collection of all such subspaces. Define $\pi : Gr_1(E^{n+1}) \to P^n(E)$ such that $\langle w \rangle \mapsto [w]$.

Lemma 8.1.4. *The mapping π is well-defined and bijective.*

Proof. Let $w \in E^{n\#}$. If $\alpha \in E^\times$ then $[w] = [\alpha w]$. Thus

$$\pi \langle w \rangle = [w] = [\alpha w] = \pi \langle \alpha w \rangle , \tag{8.1}$$

and π is well-defined. Let $\pi \langle w_1 \rangle = \pi \langle w_2 \rangle$. Thus $[w_1] = [w_2]$, implying $w_1 = \alpha w_2$, where $w_1, w_2 \neq 0$ and $\alpha \in E^\times$. Thus $\langle w_1 \rangle = \langle w_2 \rangle$, and π is one-to-one. Now let $[x] \in P^n(E)$. Thus $0 \neq x \in E^{n+1}$, and $\pi \langle x \rangle = [x]$. Thus π is onto. □

Definition 8.1.5. If $n = 2$, $P^2(E)$ is called a *projective plane*. If $n = 2$ and $|E| = q$, then q is the *order* of the plane. If $n = 1$, then $P^1(E)$ is called a *projective line*.

Definition 8.1.6. Let π be the mapping in Definition 8.1.3. A *point* in $P^n(E)$ is the image under π of a 1-dimensional subspace in E^{n+1}. A *line* in $P^n(E)$ is the image under π of a 2-dimensional subspace in E^{n+1}.

Theorem 8.1.7. *If E is of order q, then $|P^n(E)| = q^n + q^{n-1} + \ldots + q + 1$.*

Proof. We have

$$P^n(E) = \{[v] | v \in E^{n+1\#}\} = E^{n+1\#}/E^\times. \tag{8.2}$$

Thus

$$|P^n(E)| = |E^{n+1\#}|/|E^\times| = (q^{n+1}-1)/(q-1) = q^n + q^{n-1} + \ldots + q + 1. \qquad (8.3)$$

□

Corollary 8.1.8. *If E is of order q, a projective plane has q^2+q+1 points and a projective line has $q+1$ points.* □

Definition 8.1.9. Now $v = (v_1, \ldots, v_{n+1}) \in E^{n+1\#}$ determines $[v] \in P^n(E)$. The *homogeneous* coordinates of $[v]$ are $[v_1 : \ldots : v_{n+1}]$.

The value of a nonzero coordinate v_i in $[v]$ is not well-defined. However, $v_i = 0$ if and only if $\alpha v_i = 0$ for $\alpha \in E^\times$, and so the value of $v_i = 0$ is well-defined. Also, if $v_i \neq 0$, then v_j/v_i is well-defined, because $(\alpha v_j)/(\alpha v_i) = v_j/v_i$ for $\alpha \in E^\times$.

Definition 8.1.10. Define the set $U_i = \{[v_1 : \ldots : v_{n+1}] \in P^n(E) \mid v_i \neq 0\}$. Now $[v] \in U_i$ can be written uniquely as

$$[v] = [v_1/v_i : \ldots : v_i/v_i : \ldots : v_n/v_i] = [w_1 : \ldots : 1 : \ldots : w_n]. \qquad (8.4)$$

The coordinates $(w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_n)$ are referred to as the *nonhomogeneous* coordinates of $[v]$, with respect to U_i. We have $U_i \cong E^n$. Note that U_i has a natural bijection with E^n that identifies points of U_i with points of E^n (in the usual sense) and lines of U_i with lines of E^n (in the usual sense).

Definition 8.1.11. We define the *hyperplane at infinity* as

$$H_\infty = P^n(E) - U_{n+1} = \{[v_1 : \ldots : v_{n+1}] \mid v_{n+1} = 0\}. \qquad (8.5)$$

Now $[v_1 : \ldots : v_n]$ corresponds with $[v_1 : \ldots : v_n : 0]$, and thus $H_\infty \cong P^{n-1}(E)$. Since

$$P^n(E) = U_{n+1} \cup H_\infty, \qquad (8.6)$$

projective n-space is the union of affine n-space with the hyperplane at infinity.

Remark 8.1.12. Points in $P^n(E) - H_\infty$, where the last coordinate is nonzero, are referred to as *affine points*. They are often represented in nonhomogeneous coordinates, with every coordinate divided by the last, which is then dropped.

Example 8.1.13. $[X : Y : Z]$ is a point in homogeneous coordinates in $P^2(E)$. For $Z \neq 0$,

$$[X : Y : Z] = [X/Z : Y/Z : Z/Z] = [x : y : 1], \tag{8.7}$$

where $X/Z = x$ and $Y/Z = y$. The values of x, y are unique for $Z = 1$, and we have (x, y) as the affine coordinates of our point. The hyperplane at infinity is here the *line at infinity*, L_∞, since the points $(X, Y, 0) \in E^3$ are projected from a 2-dimensional subspace in E^3 to a 1-dimensional subset of $P^2(E)$.

8.2 Action of $\Gamma L_n(E)$ on Zero Sets of Homogeneous Polynomials

In this section, we consider the zero sets of homogeneous polynomials. Specifically, we consider the action of $\Gamma L_n(E)$ on homogeneous polynomials. While a transformed polynomial may no longer be a polynomial, its zero set remains well-defined, and there exists a homogeneous polynomial possessing the same zero set. Thus $\Gamma L_n(E)$ sends zero sets of homogeneous polynomials to zero sets of homogeneous polynomials.

Definition 8.2.1. A *homogeneous polynomial* is a polynomial in $E[x_1, \ldots, x_n]$ where each term is of the same total degree. If that degree is 1, it is a *linear* homogeneous polynomial, while if it is 2, it is a *quadratic* homogeneous polynomial.

Lemma 8.2.2. *If f is homogeneous of degree d, and if $\alpha \in E$ and $x \in E^n$, then $f(\alpha x) = \alpha^d f(x)$.*

Proof. We have

$$f(x) = a_1({x_{11}}^{e_{11}} \cdots {x_{1n}}^{e_{1n}}) + \ldots + a_k({x_{k1}}^{e_{k1}} \cdots {x_{kn}}^{e_{kn}}) = \sum_{1 \leq i \leq k} a_i \prod_{1 \leq j \leq n} {x_{ij}}^{e_{ij}}, \quad (8.8)$$

where x_{ij} is the jth variable in the ith term, where $e_{ij} \in \mathbb{N}$, and for a given i, $e_{i1} + \ldots + e_{in} = d$. If $x = (x_1, \ldots, x_n)$, then $\alpha x = (\alpha x_1, \ldots, \alpha x_n)$. Thus

$$f(\alpha x) = \sum_i a_i \prod_j (\alpha x_{ij})^{e_{ij}} = \sum_i a_i \alpha^d \left(\prod_j {x_{ij}}^{e_{ij}} \right) = \alpha^d \left(\sum_i a_i \prod_j {x_{ij}}^{e_{ij}} \right) = \alpha^d f(x). \quad (8.9)$$

□

Definition 8.2.3. Let $f : E^n \to E$. The *zero set* of f (or, the *variety* of f) is:

$$Z(f) = \{[x] \in P^{n-1}(E) \mid f(x) = 0\}. \quad (8.10)$$

Lemma 8.2.4. *The zero set $Z(f)$, where f is a homogeneous polynomial, is well-defined.*

Proof. Let f be a homogeneous polynomial of degree d, and let $f(x) = 0$, where $x \neq 0$ for $x \in E^{n+1}$. If $a \in E^\times$, then

$$f(ax) = a^d f(x) = a^d 0 = 0, \quad (8.11)$$

by Lemma 8.2.2. Conversely, let $f(ax) = 0$, such that $ax \neq 0$. Thus $a \in E^\times$ and $x \neq 0$. We have

$$0 = f(ax) = a^d f(x), \quad (8.12)$$

and since $a \neq 0$, $a^d \neq 0$, implying $f(x) = 0$. Therefore, $[x] \in Z(f)$. □

Definition 8.2.5. A *curve* in $P^{n-1}(E)$ is the zero set of a homogeneous polynomial $f \in E[x_1, \ldots, x_n]$. If f is linear, then $Z(f)$ is called a *line*. If f is quadratic, then $Z(f)$ is called a *conic*.

Theorem 8.2.6. *The definitions of lines in $P^2(E)$ given in Definitions 8.1.6 and 8.2.5 agree.*

Proof. Let W be a 2-dimensional subspace of E^3. Thus $\pi(W)$ is a projective line by Definition 8.1.6. Now W is a plane through the origin, and by Corollary 6.5.8, W is the set of $x \in E^3$ such that $a \cdot x = 0$, where a is a nonzero vector orthogonal to W and $\cdot$ is the dot product. Let $f = a_1x_1 + a_2x_2 + a_3x_3$. We see $\pi(W) = Z(f)$.

Conversely, let $Z(f)$ be the zero set of the linear homogeneous polynomial $f = a_1x_1 + a_2x_2 + a_3x_3$. Since the coefficients of f are not all 0, $a = (a_1, a_2, a_3) \neq 0$. Thus $Z(f)$ is the set of all $[y]$ such that $a \cdot y = 0$. By Theorem 6.5.7 the set of all y satisfying this equation is a 2-dimensional subspace W, and $Z(f) = \pi(W)$. □

Definition 8.2.7. By Lemma 5.3.1, $\Gamma L_n(E)$ has a well-defined action on $P^{n-1}(E)$, and we define

$$Z(f)^T = \{[x]T \in P^{n-1}(E) \mid f(x) = 0\}. \tag{8.13}$$

Definition 8.2.8. If $f \in E[x_1, \ldots, x_n]$ is homogeneous of degree d, and $T \in \Gamma L_n(E)$, we define

$$f^T(x) = f\left(xT^{-1}\right), \tag{8.14}$$

for all $x \in E^n$. Note that if T has a nontrivial field automorphism, then f^T is not a polynomial.

Lemma 8.2.9. *For $T \in \Gamma L_n(E)$ and $f \in E[x_1, \ldots, x_n]$, the zero set of f^T is well-defined.*

Proof. Let $T \in \Gamma L_n(E)$. Now $T^{-1} \in \Gamma L_n(E)$, and by definition, there is a $\theta \in \mathrm{Gal}(E)$ such that for $x \in E^n$ and $a \in E$, $(ax)T^{-1} = a^\theta\,(xT^{-1})$. We assume $f^T(x) = 0$ for some $x \neq 0$, and we let $a \in E^\times$. We have

$$f^T(ax) = f\left((ax)\,T^{-1}\right) = f\left(a^\theta\left(xT^{-1}\right)\right) = \left(a^\theta\right)^d f\left(xT^{-1}\right) = \left(a^\theta\right)^d(0) = 0. \tag{8.15}$$

Conversely, let $f^T(ax) = 0$ such that $ax \neq 0$. Thus $a \in E^\times$ and $x \neq 0$. We have

$$0 = f^T(ax) = \left(a^\theta\right)^d f^T(x). \tag{8.16}$$

Since $a \neq 0$, $a^\theta \neq 0$, and $\left(a^\theta\right)^d \neq 0$, implying $f^T(x) = 0$. Thus $[x] \in Z\left(f^T\right)$. □

Lemma 8.2.10. *If f is homogeneous and $T \in \Gamma L_n(E)$, then $Z(f)^T = Z\left(f^T\right)$.*

Proof. Under the action of $\Gamma L_n(E)$ on $P^{n-1}(E)$ in Lemma 5.3.1, we have $[x]T = [xT]$. If $[x] \in Z(f^T)$, then $f^T(x) = f\left(xT^{-1}\right) = 0$. Thus $[xT^{-1}] \in Z(f)$, and

$$[xT^{-1}]T = [x(TT^{-1})] = [x] \in Z(f)^T. \tag{8.17}$$

If $[xT] = [x]T \in Z(f)^T$, then

$$f^T(xT) = f((xT)T^{-1}) = f(x) = 0. \tag{8.18}$$

Thus $[xT] = [x]T \in Z\left(f^T\right)$. □

Definition 8.2.11. If $f \in E[x_1, \ldots, x_n]$ such that

$$f(x) = \sum_{1 \leq i \leq k} a_i \prod_{1 \leq j \leq n} {x_{ij}}^{e_{ij}}, \tag{8.19}$$

and $\theta \in \mathrm{Gal}(E)$, then define $f_\theta \in E[x_1, \ldots, x_n]$ such that

$$f_\theta(x) = \sum_{1 \leq i \leq k} a_i^\theta \prod_{1 \leq j \leq n} {x_{ij}}^{e_{ij}}. \tag{8.20}$$

Thus f_θ is the polynomial resulting from the application of θ to the coefficients of f.

Remark 8.2.12. Since $\mathrm{Gal}(E)$ is isomorphic to a subgroup of $\Gamma L_n(E)$, we embed $\mathrm{Gal}(E)$ in $\Gamma L_n(E)$ by the action

$$(x_1, \ldots, x_n)\theta = (x_1^\theta, \ldots, x_n^\theta), \tag{8.21}$$

where $\theta \in \mathrm{Gal}(E)$. This was the same action denoted by $f(\theta)$ in Lemma 5.2.5.

Lemma 8.2.13. *If f is homogeneous of degree d, and $T = T_1\theta \in \Gamma L_n(E)$, where $T_1 \in GL_n(E)$ and $\theta \in \mathrm{Gal}(E)$, then*

$$Z\left(f^T\right) = Z\left(\left(f^{T_1}\right)_\theta\right), \tag{8.22}$$

and $\left(f^{T_1}\right)_\theta$ is homogeneous of degree d.

Proof. If $T \in GL_n(E)$, and $x = (x_1, \ldots, x_n) \neq 0$, then

$$xT^{-1} = ((b_{11}x_1 + \ldots + b_{n1}x_n), \ldots, (b_{1n}x_1 + \ldots + b_{nn}x_n)) \neq 0, \tag{8.23}$$

for some $b_{ij} \in E$. If the first term of $f(x)$ is

$$a_1 \prod_{1\leq j\leq n} {x_j}^{e_j}, \text{ such that } \sum_j e^j = d, \tag{8.24}$$

then the first term of $f\left(xT^{-1}\right)$ is

$$a_1 \prod_{1\leq j\leq n} (b_{1j}x_1 + \ldots + b_{nj}x_n)^{e_j}, \text{ such that } \sum_j e^j = d. \tag{8.25}$$

The first factor in this product is the exponent $(b_{11}x_1 + \ldots + b_{n1}x_n)^{e_1}$, which is homogeneous of degree e_1. Thus after multiplying out our product, we have a homogeneous polynomial of degree d. After summing all terms of $f(xT^{-1})$, we continue to have a homogeneous polynomial of degree d. Since $f\left(xT^{-1}\right) = f^T(x)$, f^T is a homogeneous polynomial.

We have

$$f^\theta(x) = f\left(x\theta^{-1}\right) = f\left(x_1^{\theta^{-1}}, \ldots, x_n^{\theta^{-1}}\right), \tag{8.26}$$

and so for

$$f(x) = \sum_{1\leq i\leq k} a_i \prod_{1\leq j\leq n} {x_{ij}}^{e_{ij}}, \tag{8.27}$$

we have

$$f^\theta(x) = \sum_{1\leq i\leq k} a_i \prod_{1\leq j\leq n} \left(x_{ij}^{\theta^{-1}}\right)^{e_{ij}}. \tag{8.28}$$

If $f^\theta(x) = 0$ for some $0 \neq x \in E^n$, then

$$0 = \sum_{1 \leq i \leq k} a_i \prod_{1 \leq j \leq n} \left(x_{ij}^{\theta^{-1}} \right)^{e_{ij}} \tag{8.29}$$

if and only if

$$0 = 0^\theta = \left(\sum_{1 \leq i \leq k} a_i \prod_{1 \leq j \leq n} (x_{ij}^{\theta^{-1}})^{e_{ij}} \right)^\theta = \sum_{1 \leq i \leq k} a_i^\theta \prod_{1 \leq j \leq n} \left(x_{ij}^{\theta^{-1}\theta} \right)^{e_{ij}} = \sum_{1 \leq i \leq k} a_i^\theta \prod_{1 \leq j \leq n} {x_{ij}}^{e_{ij}}. \tag{8.30}$$

Thus $f_\theta(x) = 0$ if and only if $f^\theta(x) = 0$, and f_θ is a homogeneous polynomial with the same zero set as f^θ.

If $T = T_1\theta \in \Gamma L_n(E)$, where T_1 is linear, then f^{T_1} is a homogeneous polynomial, as is $\left(f^{T_1}\right)_\theta$. This last polynomial has the same zero set as $\left(f^{T_1}\right)^\theta$. Now

$$\left(f^{T_1}\right)^\theta (x) = f^{T_1}\left(x\theta^{-1}\right) = f\left(\left(x\theta^{-1}\right)T_1^{-1}\right) = f\left(x\left(\theta^{-1}T_1^{-1}\right)\right) = f\left(xT^{-1}\right) = f^T(x), \tag{8.31}$$

and so $\left(f^{T_1}\right)^\theta = f^T$. Thus $Z\left(f^T\right) = Z\left(\left(f^{T_1}\right)_\theta\right)$, and $\left(f^{T_1}\right)_\theta$ is a homogeneous polynomial of degree d. □

Theorem 8.2.14. *$\Gamma L_n(E)$ acts on the zero sets of homogeneous polynomials.*

Proof. Let $Z(f)$ be the zero set of a homogeneous $f \in E[x_1, \ldots, x_n]$ and let $T \in \Gamma L_n(E)$. Now by Theorem 5.2.6, we have $T = T_1\theta$ where $T_1 \in GL_n(E)$ and $\theta \in \mathrm{Gal}(E)$. Combining Lemmas 8.2.10 and 8.2.13, we obtain:

$$Z(f)^T = Z\left(f^T\right) = Z\left(f^{T_1\theta}\right) = Z\left(\left(f^{T_1}\right)_\theta\right), \tag{8.32}$$

where $(f^{T_1})_\theta$ is a homogeneous polynomial. Thus the zero set of a homogeneous polynomial is taken to the zero set of a homogeneous polynomial by $T \in \Gamma L_n(E)$. Now $I \in \Gamma L_n(E)$, and $f^I(x) = f(xI^{-1}) = f(xI) = f(x)$. Thus $f^I = f$, and so

$$Z(f)^I = Z(f^I) = Z(f). \tag{8.33}$$

Now let $S, T \in \Gamma L_n(E)$, and let $x \in E^n$. We have

$$f^{ST}(x) = f\left(x\left(ST\right)^{-1}\right) = f\left(x\left(T^{-1}S^{-1}\right)\right)$$
$$= f\left(\left(xT^{-1}\right)S^{-1}\right) = f^S\left(xT^{-1}\right) = \left(f^S\right)^T(x), \quad (8.34)$$

and thus $f^{ST} = \left(f^S\right)^T$, implying $Z(f)^{ST} = Z\left(f^{ST}\right) = Z\left(\left(f^S\right)^T\right) = \left(Z\left(f\right)^S\right)^T$. □

Corollary 8.2.15. *$P\Gamma L_n(E)$ acts on the zero sets of homogeneous polynomials.*

Proof. Let f be homogeneous of degree d, and let $\beta I \in \Gamma L_n(E)$ such that $\beta \in E^\times$. Now

$$f^{\beta I}(x) = f(x(\beta I)^{-1}) = f(x(\beta^{-1}I)) = f(\beta^{-1}x) = \beta^{-d}f(x). \quad (8.35)$$

Since $\beta^{-d} \neq 0$, $f(x) = 0$ if and only if $\beta^{-d}f(x) = 0$, and so $Z(f) = Z(f^{\beta I})$ by Lemma 8.2.9. Thus $Z(f)^{\beta I} = Z(f^{\beta I}) = Z(f)$, and therefore $\{\beta I \mid \beta \in E^\times\}$ is in the kernel of the action defined in Theorem 8.2.14. By Lemma 3.2.16, the induced action of $P\Gamma L_n(E)$ on zero sets of homogeneous polynomials is well-defined. □

Corollary 8.2.16. *The action of elements of $P\Gamma L_3(E)$ preserves lines in $P^2(E)$.*

Proof. By Corollary 8.2.15, elements of $P\Gamma L_3(E)$ send zero sets of linear homogeneous polynomials to zero sets of linear homogeneous polynomials. By Theorem 8.2.6, lines are the zero sets of linear homogeneous polynomials. □

8.3 Duality

Projective planes have the property that what is said about lines can also be said about points, and vice versa. In this section, we look at this duality and at the homogeneous coordinates for duals.

Definition 8.3.1. If there is a transformation that maps points to lines and lines to points, while preserving incidence, then points and lines are *dual* to one other. In

such a situation, the two terms can switched so that what holds for one also holds for the other.

Definition 8.3.2. Let $S \subseteq P^2(E)$ such that S is a point or line in $P^2(E)$. Thus $S = \pi(W)$, the projection of a subspace $W \leq E^3$ of dimension 1 or 2, where π is the bijection in Definition 8.1.3. Define a mapping $\varphi : P^2(E) \to P^2(E)$ such that

$$\varphi(S) = \pi((\pi^{-1}S)^\perp) = \pi(W^\perp). \tag{8.36}$$

Theorem 8.3.3. *Points and lines are dual in the projective plane $P^2(E)$.*

Proof. Now the dot product is a nondegenerate symmetric bilinear form (by Lemma 6.4.13), and so by Theorem 6.5.7,

$$\dim W^\perp = \dim E^3 - \dim W = 3 - \dim W. \tag{8.37}$$

If S is a point and W has dimension 1, then $W^\perp$ has dimension 2 and $\pi(W^\perp)$ is a line. Similarly, if S is a line and W has dimension 2, then $W^\perp$ has dimension 1 and $\pi(W^\perp)$ is a point.

Since the dot product is a nondegenerate symmetric form, we have $W^{\perp\perp} = W$ by Corollary 6.5.8. Thus

$$\varphi(\varphi(S)) = \varphi(\pi(W^\perp)) = \pi((\pi^{-1}(\pi(W^\perp)))^\perp) = \\ \pi((\pi^{-1}\pi(W^\perp))^\perp) = \pi(W^{\perp\perp}) = \pi(W) = S, \tag{8.38}$$

and φ is a bijection.

Let line ℓ be incident with point p. Since ℓ is a set of points, $p \in \ell$. Thus if $\pi^{-1}(\ell) = W_1$ and $\pi^{-1}(p) = W_2$, then $W_2 < W_1$. By Theorem 6.5.2, $W_1^\perp < W_2^\perp$. Thus if

$$\varphi(\ell) = \pi(W_1^\perp) = \ell^\perp \qquad \text{and} \qquad \varphi(p) = \pi(W_2^\perp) = p^\perp \tag{8.39}$$

are the dual point of ℓ, and the dual line of p, respectively, then $\ell^\perp \in p^\perp$. Therefore, point $\ell^\perp$ is incident with line $p^\perp$, and φ preserves incidence. □

Corollary 8.3.4. *If $|E| = q$, then $P^2(E)$ has $q^2 + q + 1$ lines, and $q + 1$ lines in $P^2(E)$ contain a given point.* □

Lemma 8.3.5. *If $\ell : aX + bY + cZ = 0$, where $a, b, c \in E$ are not all 0, then the solution set of ℓ and $\alpha\ell : (\alpha a)X + (\alpha b)Y + (\alpha c)Z = 0$, for $\alpha \in E^\times$, are identical. The projection of this set is a line in $P^2(E)$.*

Proof. Let $(x_1, y_1, z_1) \in E^3$ satisfy ℓ. Then

$$\alpha(ax_1) + \alpha(by_1) + \alpha(cz_1) = \alpha(ax_1 + by_1 + cz_1) = \alpha(0) = 0. \tag{8.40}$$

Similarly,

$$a(\beta x_1) + b(\beta y_1) + c(\beta z_1) = \beta(ax_1 + by_1 + cz_1) = \beta(0) = 0. \tag{8.41}$$

Thus $(\beta x_1, \beta y_1, \beta z_1)$ also satisfies ℓ (and thus $\alpha\ell$). Now ℓ is a plane through the origin in E^3, a 2-dimensional subspace. By projecting, we have a line in $P^2(E)$. □

Definition 8.3.6. By Theorem 8.3.3, points and lines are dual in $P^2(E)$, and by Lemma 8.3.5, lines can be expressed as $aX + bY + cZ = 0$. Now every point, $[a : b : c]$, in $P^2(E)$ gives rise to a unique such equation. In fact,

$$[a : b : c]^\perp = \{[X : Y : Z] \in P^2(E) \mid aX + bY + cZ = 0\} \tag{8.42}$$

and

$$\{[X : Y : Z] \in P^2(E) \mid aX + bY + cZ = 0\}^\perp = [a : b : c]. \tag{8.43}$$

8.4 Lines

In this section, we verify that any two lines in the projective plane intersect in a single point. We also classify lines in the projective plane into three types, in preparation for a proof in the next section.

Theorem 8.4.1. *Two distinct lines in $P^2(E)$ intersect in exactly one point.*

Proof. Let ℓ_1, ℓ_2 be distinct lines in $P^2(E)$. With the addition of the origin, they are 2-dimensional subspaces W_1, W_2 in E^3. Since ℓ_1, ℓ_2 are distinct, so are W_1, W_2. We have

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2). \tag{8.44}$$

Since W_1, W_2 are distinct, $2 < \dim(W_1 + W_2) \leq 3$, and thus $\dim(W_1 + W_2) = 3$. Therefore,

$$\dim(W_1 \cap W_2) = \dim W_1 + \dim W_2 - \dim(W_1 + W_2) = 2 + 2 - 3 = 1. \tag{8.45}$$

Since $W_1 \cap W_2$ is 1-dimensional in E^3, when the origin is removed, it is a point in $P^2(E)$, the point of intersection of ℓ_1, ℓ_2. □

Theorem 8.4.2. *Lines in $P^2(E)$ have three forms:*

(i) $\{[x : mx + b : 1] \mid m, b \in E\} \cup \{[1 : m : 0]\}$*, with m, b fixed as x varies.*

(ii) $\{[b : y : 1] \mid b \in E\} \cup \{[0 : 1 : 0]\}$*, with b fixed as y varies.*

(iii) $\{[1 : m : 0] \mid m \in E\} \cup \{[0 : 1 : 0]\}$*, where m varies.*

Proof. Lines in $P^2(E)$ are the projections of 2-dimensional subspaces of E^3, which by Theorem 8.2.6, are defined by equations of the form $a_1X + a_2Y + a_3Z = 0$, such that the $a_i \in E$ are not all 0.

Case 1: $a_2 \neq 0$. Solving for Y, we have

$$Y = -\frac{a_1}{a_2}X - \frac{a_3}{a_2}Z = mX + bZ, \tag{8.46}$$

which is satisfied by points of the form $(X, mX + bZ, Z)$. When this subspace is projected, we have the line satisfied by the points (in homogeneous coordinates) $[X : mX + bZ : Z]$. When $Z \neq 0$, these points have the unique form

$$[X : mX + bZ : Z] = [X/Z : (mX + bZ)/Z : 1] = [x : mx + b : 1]. \tag{8.47}$$

This is $(x, mx + b)$ in nonhomogeneous coordinates, equivalent to $y = mx + b$ in the affine plane. However, if $Z = 0$, we have the point $[X : mX + b(0) : 0] = [X : mX : 0]$. Since $[0 : 0 : 0] \notin P^2(E)$, $X \neq 0$ and we have

$$[X/X : mX/X : 0] = [1 : m : 0] \tag{8.48}$$

as a point on our line.

Case 2: $a_2 = 0, a_1 \neq 0$. Solving for X, we have

$$X = -\frac{a_1}{a_3}Z = bZ, \tag{8.49}$$

which is satisfied by points of the form (bZ, Y, Z). When this subspace is projected, we have the line satisfied by the points (in homogeneous coordinates) $[bZ : Y : Z]$. When $Z \neq 0$, these points have the unique form

$$[bZ : Y : Z] = [bZ/Z : Y/Z : 1] = [b : y : 1]. \tag{8.50}$$

This is (b, y) in nonhomogeneous coordinates, equivalent to $x = b$ in the affine plane. However, if $Z = 0$, we have the point on our line $[b(0) : Y : 0] = [0 : Y : 0]$. Since $[0 : 0 : 0] \notin P^2(E)$, we see $Y \neq 0$ and we have

$$[0/Y : Y/Y : 0/Y] = [0 : 1 : 0] \tag{8.51}$$

as a point on our line.

Case 3: $a_1 = a_2 = 0$. We have $Z = 0$, which is satisfied by points of the form $(X, Y, 0)$. When this subspace is projected, we have the line satisfied by the points

$[X : Y : 0]$, which is L_∞. If $X \neq 0$, then $[X : Y : 0] = [1 : Y/X : 0]$. Since $Y/X = (Y/Z)/(X/Z) = y/x$, we have

$$[1 : Y/X : 0] = [1 : y/x : 0] = [1 : m : 0] \tag{8.52}$$

for $m \in E$. If $X = 0$, we have

$$[0 : Y : 0] = [0/Y : Y/Y : 0/Y] = [0 : 1 : 0] \tag{8.53}$$

since $Y \neq 0$. □

Remark 8.4.3. Lines of the form $\{[x : mx + b : 1]\} \cup \{[1 : m : 0]\}$ will often be denoted by $y = mx + b$, and those of the form $\{[b : y : 1]\} \cup \{[0 : 1 : 0]\}$ by $x = b$.

Example 8.4.4. Let $E = \mathbb{F}_4$. Thus $|P^2(\mathbb{F}_4)| = 4^2 + 4 + 1 = 21$. Each projective line in $P^2(\mathbb{F}_4)$ has $4 + 1 = 5$ points. By Theorem 8.4.2, we have 3 types of lines in $P^2(\mathbb{F}_4)$. Lines of the form $y = mx + b$ have 4 choices for m, and 4 choices for b, for 16 lines; lines of the form $x = b$ have 4 choices for b, for 4 lines; and there is 1 line at infinity, for 21 lines. We see that each line of type (i) and (ii) has 4 affine points, plus 1 point at infinity, while the line at infinity has 4 affine slopes, plus the slope ∞.

8.5 Regular Conics and Hyperconics

In this section, we classify a *regular conic* in $P^2(\mathbb{F}_q)$, $\mathbb{F}_q$ of characteristic 2, by use of the regular quadratic form $YZ + X^2$ (from the end of Chapter 7). We determine this polynomial's zero set of $q + 1$ points. It turns out that there is a unique point in $P^2(\mathbb{F}_q)$ not on any of this conic's chords, and the addition of this point gives a *hyperconic* of $q + 2$ points. Any regular quadratic form (in 3 variables, over a finite field of characteristic 2) can be transformed to the above form, which generalizes the result to all regular conics and hyperconics in $P^2(\mathbb{F}_q)$. Finally, we show that a line and hyperconic intersect in zero or two points.

Remark 8.5.1. We will omit colons in homogeneous coordinates if the value of each coordinate is clear.

Definition 8.5.2. A *conic* C in $P^2(E)$ is $C = \{[XYZ] \in P^2(E) \mid Q[XYZ] = 0\}$, the zero set of a quadratic homogeneous polynomial Q. If Q is a regular quadratic form, then C is a *regular conic*.

Theorem 8.5.3. *The action of elements of $P\Gamma L_3(E)$ preserves conics and regular conics in $P^2(E)$.*

Proof. Let C be a conic, the zero set of the quadratic homogeneous polynomial Q. By Theorem 7.3.2, elements of $GL_3(E)$ act on Q, and thus elements of $PGL_3(E)$ preserve conics in $P^2(E)$.

Let $\theta \in \Gamma L_3(E)$ apply a field automorphism to vector coordinates in E^3 (thus θ corresponds to an element in $\mathrm{Gal}(E/F)$). By Lemmas 8.2.10 and 8.2.13, there is a homogeneous quadratic polynomial Q_θ (θ applied to the coefficients of Q) that has the zero set $Z(Q)^\theta$ (θ applied to the zero set of Q). We have

$$Z(Q)^\theta = Z(Q^\theta) = Z(Q_\theta). \tag{8.54}$$

Thus θ sends C (the zero set of Q) to C' (the zero set of Q_θ).

By Theorem 8.2.14, $\Gamma L_3(E)$ acts on the zero sets of homogeneous quadratic polynomials, and Corollary 8.2.15 extends this action to $P\Gamma L_3(E)$. Thus $P\Gamma L_3(E)$ preserves conics in $P^2(E)$. □

Remark 8.5.4. For the rest of this section, $E = \mathbb{F}_q$ where $q = 2^n$.

Theorem 8.5.5. *For the regular quadratic form $Q(X, Y, Z) = YZ + X^2$ of Corollary 7.4.19, we have the conic $C_0 = \{[x : x^2 : 1] \mid x \in \mathbb{F}_q\} \cup \{[010]\}$ of $q + 1$ points.*

Proof. First, assume $Z \neq 0$. Thus $YZ + X^2 = 0$ implies

$$YZ/Z^2 = X^2/Z^2 \qquad \Rightarrow \qquad y = x^2, \tag{8.55}$$

and $\{[x : x^2 : 1] \mid x \in \mathbb{F}_q\}$ are the q points satisfying this equation. If $Z = 0$, then $C_0 : X^2 = 0$ has $[0a0]$ as its one solution, with $a \in \mathbb{F}_q$. We take $[010]$ as this point. □

Corollary 8.5.6. *For any regular quadratic form Q, the associated regular conic C has $q + 1$ points in $P^2(\mathbb{F}_q)$.*

Proof. By Corollary 7.4.19, there is a basis for $V = \mathbb{F}_q^3$ (where $\mathbb{F}_q$ is of characteristic 2) such that Q becomes $Q(X, Y, Z) = YZ + X^2$. Thus a regular conic C defined on $\mathbb{F}_q^3$ can be put in the desired form under the action of $GL_3(\mathbb{F}_q)$, and $q + 1$ points will be taken to $q + 1$ points. □

Definition 8.5.7. The line determined by any two distinct points on a regular conic C is a *chord* of the conic.

Theorem 8.5.8. *The point* $[100]$ *is the only point of $P^2(\mathbb{F}_q)$ not on any chord of the conic C_0 of Theorem 8.5.5.*

Proof. By Theorem 8.5.5, $C_0 = \{[x : x^2 : 1] \mid x \in \mathbb{F}_q\} \cup \{[010]\}$ has q affine points and $[010]$ on L_∞. The q chords of C_0 passing through $[010]$ have the form $x = b$, $b \in \mathbb{F}_q$, and consist of the points $[b : y : 1]$, $y \in \mathbb{F}_q$. Therefore, the q^2 affine points of $P^2(\mathbb{F}_q)$ and $[010]$ are on these chords.

Consider the chords determined by the other q points in C_0, which have the form (x, x^2) in affine coordinates. Now $(0, 0)$ is one of these points. Since $\mathbb{F}_q^2 = \mathbb{F}_q$, if $\alpha \in \mathbb{F}_q^\times$, then $[\sqrt{\alpha} : \alpha : 1] \in C_0$. Thus the chords determined by $(0, 0)$ and the other affine points of C_0 are precisely the lines with equations $y = \alpha x$ for $\alpha \in \mathbb{F}_q^\times$. This

implies that the $q-1$ points of the form $[1\alpha 0]$, for $\alpha \neq 0$, are on these chords. Thus every point of $P^2(\mathbb{F}_q)$ except $[100]$ is on a chord of C_0.

In order for $[100]$ to be on a chord, there must be an affine chord with slope 0. Because $\mathbb{F}_q^2 = \mathbb{F}_q$, the q affine points of C_0 have distinct values in the y coordinate, and thus no chord determined by them has slope 0. □

Corollary 8.5.9. *A regular conic C has exactly one point in $P^2(\mathbb{F}_q)$ not on any of its chords.*

Proof. The proof is similar to that of Corollary 8.5.6. □

Definition 8.5.10. The point in Corollary 8.5.9 is the *nucleus* of the conic. The union of a regular conic and its nucleus is a *hyperconic*.

Theorem 8.5.11. *A line and a hyperconic in $P^2(\mathbb{F}_q)$ intersect in 0 or 2 points.*

Proof. By Corollary 7.4.19 and Theorem 8.5.8, we can let our hyperconic be $C_0 \cup \{[100]\}$. Thus we have the q affine points (x, x^2), and $[010], [100]$ at infinity. By Theorem 8.4.2, we know lines in $P^2(\mathbb{F}_q)$ have one of 3 forms.

Case 1*a*: $\{[x : b : 1] \mid b \in \mathbb{F}_q\} \cup \{[1 : 0 : 0]\}$. On L_∞, we have $[100]$ as a point of intersection. In the affine plane, our lines are of the form $y = b$, $b \in \mathbb{F}_q$. Since our affine points have distinct y values, and squaring is an automorphism of $\mathbb{F}_q$ by Lemma 7.4.2, there is 1 other point of intersection for each $b \in \mathbb{F}_q$, namely, $(\sqrt{b}, b)$. In this case, we have 2 points of intersection.

Case 1*b*: $\{[x : mx + b : 1] \mid m \in \mathbb{F}_q^\times, b \in \mathbb{F}_q\} \cup \{[1 : m : 0] \mid m \in \mathbb{F}_q^\times\}$. On L_∞, this line has no intersection with our hyperconic. Now C_0 is the solution set of $YZ = X^2$, which has the affine equation $y = x^2$. Thus in the affine plane, our line is of the form $y = mx + b$, $m \neq 0$, and by substituting we obtain $x^2 = mx + b$, or $x^2 + mx + b = 0$. If this polynomial is irreducible, we have 0 points of intersection.

If it is reducible, assume there is a double root r. We have

$$x^2 + mx + b = (x + r)(x + r) = x^2 + 2rx + r^2 = x^2 + r^2. \tag{8.56}$$

But this implies $m = 0$, contrary to our assumption. Thus, if it is reducible,

$$x^2 + mx + b = (x + r_1)(x + r_2) = x^2 + (r_1 + r_2)x + r_1 r_2 = 0, \tag{8.57}$$

such that $r_1 \neq r_2$. Now r_1, r_2 are the 2 solutions, and so $y = (r_1 + r_2)x + r_1 r_2$ and $y = x^2$ have $[r_1 : r_1^2 : 1]$ and $[r_2 : r_2^2 : 1]$ as their 2 points of intersection. Thus when there is intersection, $[x : mx + b : 1]$ is of the form $[x : x^2 : 1]$. In this case, there are 0 or 2 points of intersection.

Case 2: $\{[b : y : 1] \mid b \in \mathbb{F}_q\} \cup \{[0 : 1 : 0]\}$. On L_∞, we have $[010]$ as a point of intersection. In the affine plane, our lines are of the form $x = b$, $b \in \mathbb{F}_q$. Since our affine points have distinct x values, there is 1 other point of intersection for each $b \in \mathbb{F}_q$. In this case, we have 2 points of intersection.

Case 3: $\{[1 : m : 0] \mid m \in \mathbb{F}_q\} \cup \{[0 : 1 : 0]\}$. We have $[010]$ and $[100]$ as our 2 points of intersection. □

CHAPTER 9

HEXADS IN $P^2(\mathbb{F}_4)$

In this chapter, we shift to a finite geometry approach, with combinatorial arguments and the use of incidence properties. We identify hyperconics with hexads, 6-point subsets of $P^2(\mathbb{F}_4)$ where no three points are collinear. We consider the action of $\Gamma L_3(\mathbb{F}_4)$ on these hexads. We show that there are 168 hexads in $P^2(\mathbb{F}_4)$, falling into three orbits of 56 under the action of $PSL_3(\mathbb{F}_4)$, and upon which $P\Gamma L_3(\mathbb{F}_4)$ acts transitively. Further, hexads are in the same $PSL_3(\mathbb{F}_4)$-orbit if and only if they intersect in an even number of points. Finally, we show that $P\Gamma L_3(\mathbb{F}_4)$ sends hexad orbits to hexad orbits, preserving the set of three orbits. These results enable us to consider the hexad orbits as 3 points acted on by $P\Gamma L_3(\mathbb{F}_4)$, and we adjoin them to the 21 points of $P^2(\mathbb{F}_4)$, a step taken in Chapter 11. We will also use these results to prove that in the Golay code (a certain subspace of $\mathbb{F}_2^{24}$), the number of 1's in every nonzero vector is divisible by 4. This chapter follows the work of Hughes [HP85] and Edge [Edg65].

9.1 k-Arcs in $P^2(\mathbb{F}_q)$

In this section, we introduce the notion of k-arc, a subset of $P^2(\mathbb{F}_q)$ such that no three points are collinear, and terminology for various values of k is given. We give a method for constructing tetrads from triads, show that the projective general linear group acts sharply-transitively on ordered tetrads, and calculate the number of tetrads in $P^2(\mathbb{F}_4)$. We follow Hirschfeld in our definitions [Hir79].

Remark 9.1.1. We use the abbreviated notation $[XYZ]$ for $[X : Y : Z] \in P^2(\mathbb{F}_q)$.

Definition 9.1.2. A *k-arc* is a set of k points in $P^2(\mathbb{F}_q)$ such that no 3 are collinear.

If $k < 3$, for any set of k points, this is vacuously true. If a point is removed from a k-arc, the remaining points form a $(k-1)$-arc. For $k = 2, 3, 4, 5, 6$, we respectively call k-arcs *duads, triads, tetrads, pentads*, and *hexads.*

Lemma 9.1.3. *If $\{v_i\} \subseteq \mathbb{F}_q^3$, for $1 \leq i \leq k$, then the following are equivalent:*

(1) For any $a_i \neq 0$ in $\mathbb{F}_q$, every 3-element subset of $\{a_1v_1, \ldots, a_kv_k\}$ is linearly independent.

(2) Every 3-element subset of $\{v_1, \ldots, v_k\}$ is linearly independent.

(3) $\{[v_1], \ldots, [v_k]\}$ is a k-arc.

Proof. $(1 \Rightarrow 2)$ Take all $a_i = 1$.

$(2 \Rightarrow 3)$ Choose $\{v_1, v_2, v_3\}$ as a representative subset. Thus $v_i \neq 0$ for $i = 1, 2, 3$, and $[v_1], [v_2], [v_3]$ are the corresponding points in $P^2(\mathbb{F}_q)$. Since any 2 of these vectors form an independent set, their span is a 2-dimensional subspace of $\mathbb{F}_q^2$, corresponding to a line in $P^2(\mathbb{F}_q)$. Since the third vector is not in this subspace, $[v_1], [v_2], [v_3]$ are not collinear. Thus $\{[v_1], \ldots, [v_k]\}$ is a k-arc.

$(3 \Rightarrow 1)$ Let $[v_1], [v_2], [v_3]$ be elements of a k-arc, and take $a_i \neq 0$ for $i = 1, 2, 3$. Since they are not collinear, any one of them is not on the line defined by the other two, and so $[v_3]$ is not on the line defined by $[v_1]$ and $[v_2]$. This line corresponds to $\text{Span}\{a_1v_1, a_2v_2\}$, and so $a_3v_3 \notin \text{Span}\{a_1v_1, a_2v_2\}$. Thus $\text{Span}\{a_1v_1, a_2v_2, a_3v_3\}$ is a 3-dimensional subspace of $\mathbb{F}_q^3$, and $\{a_1v_1, a_2v_2, a_3v_3\}$ is linearly independent. □

Lemma 9.1.4. *The group $PGL_3(\mathbb{F}_q)$ acts transitively on ordered triads in $P^2(\mathbb{F}_q)$.*

Proof. Now $\beta = \{e_1, e_2, e_3\}$ is a basis for $\mathbb{F}_q^3$, and so $([e_1], [e_2], [e_3])$ is an ordered triad. We view elements of $PGL_3(\mathbb{F}_q)$ as equivalence classes of matrices. Let

$([v_1], [v_2], [v_3])$ be another ordered triad. The matrix

$$S = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{9.1}$$

maps e_i to v_i, $i = 1, 2, 3$. Since $\{v_1, v_2, v_3\}$ is independent, S is invertible, and thus is in an equivalence class $[S] \in PGL_3(\mathbb{F}_q)$ mapping $([e_1], [e_2], [e_3])$ to $([v_1], [v_2], [v_3])$. The result follows by Lemma 3.2.4. □

Lemma 9.1.5. *The following are equivalent:*

(1) $\{[v_1], [v_2], [v_3], [x]\}$ *is a tetrad.*

(2) $\{v_1, v_2, v_3\}$ *is an independent set, and* $x = a_1v_1 + a_2v_2 + a_3v_3$ *such that no* $a_i = 0$.

Proof. Let $\{[v_1], [v_2], [v_3], [x]\}$ be a tetrad. Thus $\{[v_1], [v_2], [v_3]\}$ is a triad, and so (v_1, v_2, v_3) is an independent set by Lemma 9.1.3. Thus it is a basis for $\mathbb{F}_q^3$, and so $x = a_1v_1 + a_2v_2 + a_3v_3$, where $a_i \in \mathbb{F}_q$. Assume one of these coefficients is 0, say $a_3 = 0$. Thus $x = a_1v_1 + a_2v_2$, and so $\{v_1, v_2, x\} = \{v_1, v_2, a_1v_1 + a_2v_2\}$ is not an independent set, and

$$\text{Span}\{v_1, v_2, a_1v_1 + a_2v_2\} = \text{Span}\{v_1, v_2\}, \tag{9.2}$$

making $[v_1], [v_2], [x]$ collinear, contrary to assumption. Thus none of our coefficients is 0.

Conversely, let $\{v_1, v_2, v_3\}$ be independent, and let $x = a_1v_1 + a_2v_2 + a_3v_3$, such that no $a_i = 0$. Thus $x \neq 0$, and $[x] \in P^2(\mathbb{F}_q)$. Now $\{v_1, v_2, a_1v_1 + a_2v_2 + a_3v_3\}$ is independent if and only if $\{v_1, v_2, v_3\}$ is independent, which it is by assumption. Similarly, if we omit v_1 or v_2 from $\{v_1, v_2, v_3, x\}$, the resulting set of 3 elements is independent. Thus $\{[v_1], [v_2], [v_3], [x]\}$ is a tetrad. □

Corollary 9.1.6. *We have* $q_0 = \{[100], [010], [001], [111]\}$ *as a tetrad in* $P^2(\mathbb{F}_q)$.

Proof. Clearly, all 4 points are in $P^2(\mathbb{F}_q)$, and $\{[100], [010], [001]\}$ form an independent set. Since $(1,1,1) = (1,0,0) + (0,1,0) + (0,0,1)$, we have q_0 as a tetrad by Lemma 9.1.5. □

Theorem 9.1.7. $PGL_3(\mathbb{F}_q)$ *acts sharply transitively on ordered tetrads in* $P^2(\mathbb{F}_q)$.

Proof. Let $q_1 = ([v_1], [v_2], [v_3], [v_4])$ and $q_2 = ([w_1], [w_2], [w_3], [w_4])$ be ordered tetrads in $P^2(\mathbb{F}_q)$. By Lemma 9.1.5,

$$v_4 = a_1v_1 + a_2v_2 + a_3v_3 \qquad \text{and} \qquad w_4 = b_1w_1 + b_2w_2 + b_3w_3, \tag{9.3}$$

such that no a_i or b_i is 0. Since $\{v_1, v_2, v_3\}$ and $\{w_1, w_2, w_3\}$ are independent, $\{a_1v_1, a_2v_2, a_3v_3\}$ and $\{b_1w_1, b_2w_2, b_3w_3\}$ are as well. As in Lemma 9.1.4, there exists a $T \in GL_3(\mathbb{F}_q)$ such that $(a_iv_i)T = b_iw_i$ for $i = 1, 2, 3$. Since

$$(a_1v_1 + a_2v_2 + a_3v_3)T = (a_1v_1)T + (a_2v_2)T + (a_3v_3)T = b_1w_1 + b_2w_2 + b_3w_3, \tag{9.4}$$

we have $v_4T = w_4$, and T sends $\{a_1v_1, a_2v_2, a_3v_3, v_4\}$ pointwise to $\{b_1w_1, b_2w_2, b_3w_3, w_4\}$. Thus $[v_i]T = [v_iT] = [w_i]$, for $1 \leq i \leq 4$, and $G = PGL_3(\mathbb{F}_q)$ acts transitively on ordered tetrads in $P^2(\mathbb{F}_q)$.

We let $q_0 = \{[100], [010], [001], [111]\}$ be an ordered tetrad, and calculate $\mathrm{Stab}_G(q_0)$, the pointwise stabilizer. Now this stabilizer subgroup is

$$\mathrm{Stab}_G([100]) \cap \mathrm{Stab}_G([010]) \cap \mathrm{Stab}_G([001]) \cap \mathrm{Stab}_G([111]). \tag{9.5}$$

We have $T \in \mathrm{Stab}_G([100])$ exactly if $(1,0,0)T = a_1(1,0,0)$, where $a_1 \in \mathbb{F}_q^\times$, and similarly for $\mathrm{Stab}_G([010])$ and $\mathrm{Stab}_G([001])$. Thus as a matrix,

$$T = \begin{bmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{bmatrix}. \tag{9.6}$$

Now if $T \in \mathrm{Stab}_G([111])$, then

$$a(1,0,0) + a(0,1,0) + a(0,0,1) = a(1,1,1) = (1,1,1)T$$
$$= (1,0,0)T + (0,1,0)T + (0,0,1)T = a_1(1,0,0) + a_2(0,1,0) + a_3(0,0,1). \quad (9.7)$$

Since $\{e_1, e_2, e_3\}$ are independent, we conclude $a_1 = a_2 = a_3$, and

$$\mathrm{Stab}_G(q_0) = \{aI \mid a \in \mathbb{F}_q^\times\}, \quad (9.8)$$

which is trivial in $PGL_3(\mathbb{F}_q)$. Therefore, $PGL_3(\mathbb{F}_q)$ acts sharply transitively on ordered tetrads in $P^2(\mathbb{F}_q)$. □

Corollary 9.1.8. *There are* 60480 *ordered tetrads in* $P^2(\mathbb{F}_4)$.

Proof. Since $PGL_3(\mathbb{F}_4)$ acts sharply transitively on ordered tetrads, $|PGL_3(\mathbb{F}_4)|$ equals the number of such tetrads, by Theorem 3.3.7. Since $|PGL_3(\mathbb{F}_4)| = 60480$, by Theorem 4.3.3, this is the number of ordered tetrads. □

Corollary 9.1.9. *There are* 2520 *tetrads in* $P^2(\mathbb{F}_4)$.

Proof. Let q be an unordered tetrad. It has 4! possible orderings, implying a $24:1$ ratio between ordered and unordered tetrads, thus yielding $60480/24 = 2520$ tetrads. □

9.2 Orbits of Hexads

In this section, we narrow our focus to the projective plane $P^2(\mathbb{F}_4)$. We determine that every tetrad is contained in exactly one hexad, and thus $PGL_3(\mathbb{F}_4)$ acts transitively on hexads. We identify the hyperconics of Chapter 8 and the hexads of this chapter. We show that $P^2(\mathbb{F}_4)$, under the action of $PSL_3(\mathbb{F}_4)$, has three hexad orbits, each of size 56, and that a hexad's setwise stabilizer is A_6 and is contained in $PSL_3(\mathbb{F}_4)$.

Many arguments in the rest of the chapter follow a similar pattern: We verify a fact for a convenient k-arc ($k = 4$ or 6), and then the transitivity of $PGL_3(\mathbb{F}_4)$ on such k-arcs extends the result to all k-arcs.

Remark 9.2.1. In the following sections of this chapter, our projective plane is exclusively $P^2(\mathbb{F}_4)$. We adopt the following convention in our representation of points: maximize the number of 1's, and then maximize the number of ω's. Thus $[11\bar{\omega}]$ is preferred to $[\omega\omega 1]$, and $[10\omega]$ is preferred to $[\bar{\omega}01]$. Thus

$$P^2(\mathbb{F}_4) = \left\{ \begin{array}{cccccc} [010], & [100], & [001], & [101], & [\omega 01], & [10\omega], \\ & [110], & [011], & [111], & [\omega 11], & [\bar{\omega}11], \\ & [1\omega 0], & [0\omega 1], & [1\omega 1], & [11\bar{\omega}], & [\bar{\omega}\omega 1], \\ & [\omega 10], & [01\omega], & [1\bar{\omega}1], & [\omega\bar{\omega}1], & [11\omega] \end{array} \right\}. \tag{9.9}$$

Lemma 9.2.2. *Let $q_0 = \{[100], [010], [001], [111]\}$ be a tetrad in $P^2(\mathbb{F}_4)$. Then:*

(1) q_0 is contained in two pentads in $P^2(\mathbb{F}_4)$.

(2) q_0 is contained in one hexad in $P^2(\mathbb{F}_4)$, $h_0 = q_0 \cup \{[\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}$.

Proof. By Lemma 9.1.5, any point of $P^2(\mathbb{F}_4)$ with a 0 coordinate will not extend $\{[100], [010], [001]\}$ to a tetrad. Thus, to extend q_0 to a pentad, we exclude the 9 such points in $P^2(\mathbb{F}_4)$.

Let $x = (1, 1, 1)$ and consider $y \in \mathbb{F}_q^3$ with two identical entries(but no 0 entries). For example, let

$$y = (a_1, a_1, a_2) = a_1 x + (a_2 - a_1)e_3. \tag{9.10}$$

Thus $\{e_3, x, y\}$ is a dependent set, and y does not extend q_0 to a pentad. Since a vector with two identical entries is a linear combination of x and one of e_1, e_2, e_3, $\{[e_1], [e_2], [e_3], [x]\}$ is not extended to a pentad by the 6 points that have two

identical coordinates (and no coordinates of 0). This leaves $[\omega\bar{\omega}1]$ and $[\bar{\omega}\omega1]$. With a series of routine calculations, we verify the addition of one is a pentad, while the addition of both is a hexad. Thus $h_0 = \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega1]\}$ is a hexad in $P^2(\mathbb{F}_4)$. □

Corollary 9.2.3. *Each tetrad or pentad in $P^2(\mathbb{F}_4)$ is in only one hexad.*

Proof. From Lemma 9.2.2, the tetrad q_0, and the two pentads formed by extending it, are in exactly one hexad. By Theorem 9.1.7, $PGL_3(\mathbb{F}_4)$ acts transitively on tetrads in $P^2(\mathbb{F}_4)$, and thus each tetrad in $P^2(\mathbb{F}_4)$ is in just one hexad. Since every pentad in $P^2(\mathbb{F}_4)$ can be constructed from some tetrad (by removing a point and adding it back, if need be), and then completed to a unique hexad, each pentad is also in a unique hexad. □

Corollary 9.2.4. *$PGL_3(\mathbb{F}_4)$ acts transitively on hexads in $P^2(\mathbb{F}_4)$.*

Proof. Let h_1 and h_2 be distinct hexads in $P^2(\mathbb{F}_4)$. Delete two points from each to make them tetrads. By Theorem 9.1.7, there exists an element of $PGL_3(\mathbb{F}_4)$ that takes one tetrad to the other. By Theorem 9.2.3, each of these tetrads are in a unique hexad, that is, h_1 and h_2. Thus $PGL_3(\mathbb{F}_4)$ takes any hexad in $P^2(\mathbb{F}_4)$ to any other. □

Corollary 9.2.5. *The set of hexads in $P^2(\mathbb{F}_4)$ is identical to the set of hyperconics.*

Proof. By Theorem 8.5.11, a line and hyperconic C in $P^2(\mathbb{F}_4)$ intersect in 0 or 2 points. Since at most two points of C lie on the same line, no three of C's points are collinear. Since C has 6 points, it is a hexad.

Conversely, we observe that the hexad h_0 in Lemma 9.2.2 is the hyperconic C_0 of Theorems 8.5.5 and 8.5.8. By Corollary 9.2.4, $PGL_3(\mathbb{F}_4)$ acts transitively on

hexads in $P^2(\mathbb{F}_4)$. Since any hexad can be taken to h_0, this implies every hexad in $P^2(\mathbb{F}_4)$ is also a hyperconic. □

Theorem 9.2.6. *There are* 168 *hexads in* $P^2(\mathbb{F}_4)$*, and each hexad contains* 360 *ordered tetrads.*

Proof. We have $\binom{6}{4}$ tetrads in each hexad, and since each tetrad has 4! possible orderings, there are $(15)(24) = 360$ ordered tetrads in each hexad. Since each tetrad lies in a unique hexad, and there are 60480 ordered tetrads by Corollary 9.1.8, we have $60480/360 = 168$ total hexads. □

Theorem 9.2.7. *Under the action of* $PSL_3(\mathbb{F}_4)$*,* $P^2(\mathbb{F}_4)$ *has* 3 *orbits of hexads, all of size* 56*.*

Proof. Let $N = PSL_3(\mathbb{F}_4)$ and $G = PGL_3(\mathbb{F}_4)$. If h is a hexad in $P^2(\mathbb{F}_4)$, it contains 360 ordered tetrads, and each of these tetrads is contained only in the hexad h, by Theorem 9.2.6. Since G acts sharply transitively on ordered tetrads, it has 360 elements sending h setwise to itself, and these form a subgroup, $\mathrm{Stab}_G(h)$. Thus if $H = \mathrm{Stab}_G(h)$, $|H| = 360$. Since H is of order 6!/2 and acts 4-transitively on the 6 points of h, we conclude $H \cong A_6$, and thus H is simple.

Consider $H \cap N$. Now $N \lhd G$ by Lemma 4.1.4. By the second isomorphism theorem, $(H \cap N) \lhd H$. Since H is simple, $H \cap N = 1$ or H. Assume $H \cap N = 1$. Since 1 is the only element of N stabilizing h, the orbit of h under N has size

$$|N|/|\mathrm{Stab}_N(h)| = 20160/1 = 20160 > 168, \tag{9.11}$$

by Theorem 3.2.20. This contradiction of Theorem 9.2.6 shows that $H \cap N = H$, and so $H \leq N$. Thus $|\mathrm{Stab}_N(h)| = 360$, and the orbit of h under N has size

$$[N : \mathrm{Stab}_N(h)] = |N|/|\mathrm{Stab}_N(h)| = 20160/360 = 56. \tag{9.12}$$

By Theorem 9.2.6, there are 168 hexads. Thus h is in an orbit with one-third of the hexads. □

Corollary 9.2.8. *Let $G = PGL_3(\mathbb{F}_4)$ and $N = PSL_3(\mathbb{F}_4)$. If h is a hexad in $P^2(\mathbb{F}_4)$, then*

$$\mathrm{Stab}_G(h) = \mathrm{Stab}_N(h) \cong A_6.$$

□

9.3 Hexagrams

To show hexads are in the same $PSL_3(\mathbb{F}_4)$-orbit if and only if they intersect evenly, it is helpful to have a list of hexads disjoint from a given hexad h, and by transitivity, we may choose a convenient hexad h_0. To find these disjoint hexads, listed in Section 9.4, we first define joins and diagonal points, which are structural features of hexads. We consider the projective duals of hexads, called hexagrams, which are sets of six lines with no three intersecting at the same point. Every hexad has a corresponding hexagram, and we determine the hexagram H_0 corresponding to the hexad h_0, and then calculate the points of intersection between its sides. These points will allow us to determine a list of hexads disjoint from h_0, a step taken in the next section.

Definition 9.3.1. Let S be a k-arc. The line determined by 2 distinct points of S is a *join* of S. For example, if $p_1, p_2 \in S$ then $\overline{p_1p_2}$ is the join determined by them. If two joins are determined by disjoint pairs of points, for example, $\overline{p_1p_2}$ and $\overline{p_3p_4}$, they are *opposite* joins. If two joins share a point in S, for example, $\overline{p_1p_2}$ and $\overline{p_1p_3}$, they are *adjacent* joins.

Definition 9.3.2. Let q be a tetrad or hexad. The intersection of opposite joins of q is a *diagonal point* of q.

Lemma 9.3.3. *Let $q = \{p_1, p_2, p_3, p_4\}$ be a tetrad subset of k-arc S. Each partition of q into two duads determines a unique diagonal point d, a point not in S. Thus q has 3 distinct diagonal points.*

Proof. Since our given tetrad is unordered, we may reorder the points $\{p_1, p_2, p_3, p_4\}$ as we please. So without loss of generality, the partition $p_1, p_2 : p_3, p_4$ determines the opposite joins $\overline{p_1p_2}$ and $\overline{p_3p_4}$, whose intersection is diagonal point d_1. Similarly, $p_1, p_3 : p_2, p_4$ determines joins $\overline{p_1p_3}$ and $\overline{p_2p_4}$, whose intersection is diagonal point d_2. If $d_1 = d_2$, then $\overline{p_1p_2}$, $\overline{p_1p_3}$, $\overline{p_3p_4}$, and $\overline{p_2p_4}$ share a common point. But the first two joins only share p_1, and the second two only share p_4, which are distinct points. This contradiction shows $d_1 \neq d_2$. If d_1 were in S, then p_1, p_2, d_1 would be collinear, contrary to the definition of a k-arc. Since our 4 points have $\binom{4}{2}/2! = 3$ partitions into two duads, there are 3 distinct diagonal points. □

Lemma 9.3.4. *Let $q = \{p_1, p_2, p_3, p_4\}$ be a tetrad. If $d \in P^2(\mathbb{F}_4) - q$, then at most two joins of q intersect at d.*

Proof. Let $d \in P^2(\mathbb{F}_4) - q$ such that two joins of q intersect at d. If the joins are not opposite, then the intersection d is a point on q, contrary to assumption. Thus these joins are opposite, and d is a diagonal point. Assume d is the intersection of $\overline{p_1p_2}$ and $\overline{p_3p_4}$, and a third join of q also intersects d. Since this third join is determined by two points in q, it intersects $\overline{p_1p_2}$ in p_1 or p_2. But two distinct lines intersect in exactly one point. This contradiction shows that no more than two joins intersect in $d \in P^2(\mathbb{F}_4) - q$. □

Corollary 9.3.5. *Each join ℓ of a tetrad q intersects exactly one diagonal point d.*

Proof. Let $q = \{p_1, p_2, p_3, p_4\}$ be a tetrad. Now $\overline{p_1p_2}$ intersects d_1, where d_1 is the intersection of $\overline{p_1p_2}$ and $\overline{p_3p_4}$. We also assume it intersects d_2, the intersection of $\overline{p_1p_3}$ and $\overline{p_2p_4}$. Thus $\overline{p_1p_2}$, $\overline{p_1p_3}$, and $\overline{p_2p_4}$ intersect at d_2, contradicting Lemma 9.3.4. □

Theorem 9.3.6. *Let q be a tetrad. There are two points in $P^2(\mathbb{F}_4)$ not on any join of q, and the addition of these two points extends q to a hexad h.*

Proof. Each tetrad q has $\binom{4}{2} = 6$ joins. Let ℓ be a join of q. Since each line in $P^2(\mathbb{F}_4)$ has exactly 5 points, ℓ is composed of 2 points from q, a diagonal point not in q, and 2 other points of $P^2(\mathbb{F}_4)$. Since pairs of adjacent joins intersect only at points in q, and pairs of opposite joins intersect only at diagonal points, these 2 other points are uniquely associated with ℓ. Thus there are

$$21 - (4 + 3 + 6 \cdot 2) = 2 \tag{9.13}$$

points not on any join of q. Now each tetrad in $P^2(\mathbb{F}_4)$ is contained in exactly one hexad by Corollary 9.2.3, and these 2 points are the only ones that can extend q to a hexad, since they do not lie on any join of q. Thus their addition achieves this. □

Lemma 9.3.7. *Every line intersecting a hexad is a join of the hexad.*

Proof. Let p be a point of hexad h. Since h has 5 other points, there are 5 joins of h containing p, where these joins are distinct because no 3 points of h are collinear. By duality in $P^2(\mathbb{F}_4)$, there are exactly 5 lines of $P^2(\mathbb{F}_4)$ intersecting at point p. Thus each of these lines is a join of h. □

Remark 9.3.8. Lemma 9.3.7 gives an alternate proof of Theorem 8.5.11, that each line of $P^2(\mathbb{F}_4)$ intersects a hyperconic (i.e., a hexad) in 0 or 2 points.

Lemma 9.3.9. *If h is a hexad and $d \in P^2(\mathbb{F}_4)$ such that $d \notin h$, then at most 3 joins of h intersect d.*

Proof. Each pair of joins of h is either adjacent or opposite. If joins are adjacent, then their point of intersection is a point in h. Thus intersection of d by multiple joins of h must come from pairs of opposite joins. Let $h = \{p_1, \ldots, p_6\}$, and assume

without loss of generality that $\overline{p_1p_2}$ and $\overline{p_3p_4}$ intersect at d. Now $\overline{p_5p_6}$ is the only join opposite to both $\overline{p_1p_2}$ and $\overline{p_3p_4}$, and thus is the only other join of h that can intersect d. □

Lemma 9.3.10. *Let q be a tetrad. The 3 diagonal points of q are collinear, and this line is completed by the 2 points that extend tetrad q to a hexad h. Thus exactly 3 joins of a hexad h intersect at each diagonal point of h, and each diagonal point of h corresponds to a partition of h into three duads.*

Proof. Let $q_0 = \{p_1 = [100], p_2 = [010], p_3 = [001], p_4 = [111]\}$. By Lemma 9.2.2, q_0 is a tetrad. There are 3 ways to partition q_0 into two duads, and each partition determines opposite joins, and thus a diagonal point. Now $\overline{p_1p_2}$ and $\overline{p_3p_4}$ intersect in [110], and so [110] is a diagonal point. Similarly, $\overline{p_2p_3}$ and $\overline{p_1p_4}$ intersect in [011], and this is a second diagonal point. Finally, $\overline{p_1p_3}$ and $\overline{p_2p_4}$ intersect in [101], and this is the third diagonal point.

Since $(1,1,0)+(0,1,1) = (1,0,1)$, we see that the 3 points are collinear. Since $\omega(1,1,0)+(0,1,1) = (\omega,\bar{\omega},1)$ and $\bar{\omega}(1,1,0)+(0,1,1) = (\bar{\omega},\omega,1)$, the remaining points on this line are $p_5 = [\omega\bar{\omega}1]$ and $p_6 = [\bar{\omega}\omega1]$, the points that by Lemma 9.2.2 extend tetrad q_0 to the hexad h_0. Since $\overline{p_5p_6}$ contains the 3 diagonal points of q_0, $\overline{p_5p_6}$ is a join of h_0 intersecting each diagonal point of q_0. Thus at least 3 joins of h_0 intersect each diagonal point of q_0. By Lemma 9.3.9, at most 3 joins of h_0 intersect each diagonal point of q_0.

Each of the above diagonal points corresponds to a partition of h_0 into three duads that includes the duad $\{p_5, p_6\}$. Now $\mathrm{Stab}(h_0) \cong A_6$, by Corollary 9.2.8, and since this is 4-transitive on h_0, we have similar results if any duad in h_0 is held fixed and the remaining 4 points are partitioned. Since $PGL_3(\mathbb{F}_4)$ acts transitively on tetrads and hexads by Theorem 9.1.7 and Corollary 9.2.4, our result holds true for

every tetrad and hexad in $P^2(\mathbb{F}_4)$. □

Corollary 9.3.11. *Let h be a hexad. Then each of the 15 points in $P^2(\mathbb{F}_4) - h$ is a unique point of intersection for 3 opposite joins of h.*

Proof. Each partition of h into three duads determines a unique diagonal point, the point of intersection of 3 joins of h, as in Lemma 9.3.10. There are $\binom{6}{2}\binom{4}{2}/3! = 15$ such partitions. Since each diagonal point is distinct, and none are points of h, they must be the points of $P^2(\mathbb{F}_4) - h$. □

Definition 9.3.12. A set of 6 lines in $P^2(\mathbb{F}_4)$ such that no 3 intersect at the same point is called a *hexagram*. The lines determining a hexagram are called its *sides*.

Lemma 9.3.13. *If h is a hexad, then there are 6 lines in $P^2(\mathbb{F}_4)$ that contain no point of h, and they form a hexagram H.*

Proof. There are $\binom{6}{2} = 15$ joins of h. Since no line of $P^2(\mathbb{F}_4)$ intersects h in just one point, by Lemma 9.3.7, there are $21 - 15 = 6$ lines that do not intersect h. Let $H = \{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6\}$ be the set of these lines. Assume 3 of the lines of H intersect in a single point p. Since p is one of the 15 points of Corollary 9.3.11, it is a point of intersection for 3 opposite joins of h. Since only 5 lines pass through p, one of the lines of H must thus be a join of h, implying this line contains two points of h. This contradicts that H is the set of lines containing no point of h. Thus H is a hexagram. □

Lemma 9.3.14. *If $H = \{\ell_1, \ldots, \ell_6\}$ is a hexagram, then the intersection of any two sides is unique, for 15 points of intersection.*

Proof. Let the intersection of ℓ_a, ℓ_b be denoted by p_{ab}. Assume $p_{ab} = p_{ac}$. Then ℓ_a, ℓ_b, ℓ_c intersect in p_{ab}, contradicting that H is a hexagram. Similarly, if $p_{ab} = p_{cd}$,

then $\ell_a, \ell_b, \ell_c, \ell_d$ intersect in p_{ab}. Thus each of the possible $\binom{6}{2} = 15$ intersections is distinct. □

Lemma 9.3.15. *Let $H = \{\ell_1, \ldots, \ell_6\}$ be a hexagram. Any partition of H into two sets of three lines determines* 6 *points of intersection between these lines, and these* 6 *points form a unique hexad, for* 10 *distinct hexads. Any two such hexads share exactly* 2 *points.*

Proof. Partition H as $\{\ell_1, \ell_2, \ell_3\}$ and $\{\ell_4, \ell_5, \ell_6\}$, which have intersections $\{p_{12}, p_{13}, p_{23}\}$ and $\{p_{45}, p_{46}, p_{56}\}$, where p_{ij} is the intersection of ℓ_i and ℓ_j. The line determined by any two points of a triad is indicated by the shared subscript, as for example, p_{12} and p_{13} determine ℓ_1. Assume p_{12}, p_{13}, p_{23} are collinear. Since any two of these points determine the same line, this implies $\ell_1 = \ell_2 = \ell_3$, contradicting the assumption that ℓ_1, ℓ_2, ℓ_3 are distinct sides of H. Now choose one point from one triad and two points from the other, as for example, p_{12}, p_{45}, p_{46}. If they are collinear, then p_{12} lies on ℓ_4, implying ℓ_1, ℓ_2, ℓ_4 intersect at p_{12}. This contradicts the definition of H. These two ways of choosing 3 points from our 2 triads exhaust the possibilities. Thus any 3 points of $\{p_{12}, p_{13}, p_{23}, p_{45}, p_{46}, p_{56}\}$ are non-collinear, making this set a hexad.

There are $\binom{6}{3}/2! = 10$ partitions of H into two sets of three lines, implying each set of 6 intersection points is distinct. By Lemma 9.3.14, every intersection point p_{ij} is unique, and therefore, so is each such hexad, for 10 distinct hexads. Now any other partition of H occurs by exchanging 2 lines of our original triads. Thus $\{\ell_1, \ell_5, \ell_3\}$ and $\{\ell_4, \ell_2, \ell_6\}$ is another partition, determining hexad $\{p_{13}, p_{15}, p_{35}, p_{24}, p_{26}, p_{46}\}$, which shares p_{13}, p_{46} with the first one. Since every other partition occurs in this way, there are 2 shared points between any two of the 10 hexads. □

Remark 9.3.16. Recall that $[abc]^\perp = \{[XYZ] \in P^2(\mathbb{F}_4) \mid aX + bY + cZ = 0\}$, the dual line of point $[abc] \in P^2(\mathbb{F}_4)$, as in Definition 8.3.6. When we refer to the "coefficients" of a line, we refer to $a, b, c \in \mathbb{F}_4$.

Lemma 9.3.17. *The hexad* $h_0 = \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}$ *has the corresponding hexagram:*

$$H_0 = \begin{Bmatrix} \ell_1 = [\omega 11]^\perp, & \ell_2 = [\bar{\omega}11]^\perp, \\ \ell_3 = [1\omega 1]^\perp, & \ell_4 = [1\bar{\omega}1]^\perp, \\ \ell_5 = [11\omega]^\perp, & \ell_6 = [11\bar{\omega}]^\perp \end{Bmatrix} \tag{9.14}$$

Proof. We claim H_0 is the set of 6 lines in $P^2(\mathbb{F}_4)$ containing no points of h_0. If a line in $P^2(\mathbb{F}_4)$ has a coefficient of 0, then one of e_1, e_2, e_3 is in its zero set. If a line in $P^2(\mathbb{F}_4)$ has coefficients adding to 0, then $e_1 + e_2 + e_3$ is in its zero set. If a line in $P^2(\mathbb{F}_4)$ has three identical coefficients, then $(\omega, \bar{\omega}, 1)$ is in its zero set. Thus the lines in H_0 are those having no coefficients of 0, such that two coefficients are identical and the third is different. This leaves the 6 lines above. □

Corollary 9.3.18. *We have the following points of intersection between the sides of* H_0*:*

$$p_{12} = [011],\ p_{13} = [11\bar{\omega}],\ p_{14} = [1\omega 0],\ p_{15} = [1\bar{\omega}1],\ p_{16} = [10\omega],$$
$$p_{23} = [\omega 10],\ p_{24} = [11\omega],\ p_{25} = [\omega 01],\ p_{26} = [1\omega 1],\ p_{34} = [101],$$
$$p_{35} = [\bar{\omega}11],\ p_{36} = [01\omega],\ p_{45} = [0\omega 1],\ p_{46} = [\omega 11],\ p_{56} = [110].$$

Proof. We have

$$\ell_1 = \{[011], [1\bar{\omega}1], [11\bar{\omega}], [10\omega], [1\omega 0]\}, \qquad \ell_2 = \{[011], [1\omega 1], [11\omega], [\omega 01], [\omega 10]\},$$
$$\ell_3 = \{[101], [11\bar{\omega}], [\bar{\omega}11], [\omega 10], [01\omega]\}, \qquad \ell_4 = \{[101], [11\omega], [\omega 11], [1\omega 0], [0\omega 1]\},$$
$$\ell_5 = \{[110], [\bar{\omega}11], [1\bar{\omega}1], [0\omega 1], [\omega 01]\}, \qquad \ell_6 = \{[110], [\omega 11], [1\omega 1], [01\omega], [10\omega]\}.$$

□

9.4 Even Intersections and Hexad Orbits

In this section, we first show there are 56 hexads intersecting a given hexad in an even number of points, the same size as that hexad's $PSL_3(\mathbb{F}_4)$-orbit. Then we show that hexads that intersect evenly are in the same $PSL_3(\mathbb{F}_4)$-orbit. Next we prove that intersecting evenly is a transitive relation, with various cases to be shown. Once the proof of an equivalence relation is completed, we have identified $PSL_3(\mathbb{F}_4)$-orbits with equivalence classes. Finally, we show $P\Gamma L_3(\mathbb{F}_4)$ sends hexad orbits to hexad orbits, preserving these three orbits.

As part of this section, we list the 10 hexads disjoint from our favorite hexad h_0, and show that they fall into three orbits under a cyclic permutation of the coordinates, which helps in proving an important lemma.

Lemma 9.4.1. *Let $\{p_1, p_2, p_3\}$ be a triad in $P^2(\mathbb{F}_4)$. There are*

- 48 *hexads containing p_1;*
- 12 *hexads containing p_1 and p_2; and*
- 3 *hexads containing p_1, p_2, p_3.*

Proof. If a hexad has point p_1, we consider the ordered tetrads containing p_1 as their first point. We have 20 choices for the second point. For the third point, we cannot choose a point on the line determined by the first and second points, for $21 - 5 = 16$ choices. Similarly, for the fourth point, we cannot choose any of the 12 points on the 3 lines determined by our first three points, and so we have $21 - 12 = 9$ choices. Conversely, within each hexad containing p_1, we have $5 \cdot 4 \cdot 3$ ordered tetrads containing p_1 as their first point, for

$$(20 \cdot 16 \cdot 9)/(5 \cdot 4 \cdot 3) = 48 \tag{9.15}$$

hexads containing p_1.

If a hexad has points p_1, p_2, we consider the ordered tetrads containing p_1, p_2 as their first and second points. We have 16 choices for our third point, and 9 for the fourth. Conversely, for each such hexad we have $4 \cdot 3$ ordered tetrads containing p_1, p_2 as their first and second points, for

$$(16 \cdot 9)/(4 \cdot 3) = 12 \tag{9.16}$$

hexads containing p_1, p_2.

If a hexad has points p_1, p_2, p_3, we consider the ordered tetrads containing p_1, p_2, p_3 as their first, second, and third points. We have 9 choices for the fourth point. Conversely, for each such hexad we have 3 ordered tetrads containing p_1, p_2, p_3 as their first, second and third points, for

$$9/3 = 3 \tag{9.17}$$

hexads containing p_1, p_2, p_3. □

Lemma 9.4.2. *Let $h = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ be a hexad in $P^2(\mathbb{F}_4)$. There are*

- 2 *hexads h' such that $h' \cap h = \{p_1, p_2, p_3\}$;*
- 3 *hexads h' such that $h' \cap h = \{p_1, p_2\}$; and*
- 12 *hexads h' such that $h' \cap h = \{p_1\}$.*

Proof. By Lemma 9.4.1, we have 3 hexads containing p_1, p_2, p_3. Excluding h, this leaves exactly 2 hexads h' intersecting h in p_1, p_2, p_3.

By Lemma 9.4.1, we have 12 hexads containing p_1 and p_2. We wish to exclude h from this, as well as hexads intersecting h in $\{p_1, p_2, p_i\}$, where $3 \leq i \leq 6$. By the reasoning in the first part, for each i there are 2 hexads of which this is true. Thus there are exactly $12 - (4 \cdot 2 + 1) = 3$ hexads h' intersecting h in p_1, p_2.

By Lemma 9.4.1, we have 48 hexads containing p_1. We wish to exclude h from this number, plus hexads intersecting h in $\{p_1, p_i\}$, where $2 \leq i \leq 6$, as well as hexads intersecting h in the points $\{p_1, p_i, p_j\}$, where $2 \leq i < j \leq 6$. By the reasoning in the second part, for each i there are 3 hexads intersecting h in $\{p_1, p_i\}$, for $5 \cdot 3 = 15$ hexads. By the reasoning in the first part, for each pair i, j there are 2 hexads intersecting h in $\{p_1, p_i, p_j\}$, for $\binom{5}{2} \cdot 2 = 20$ hexads. Thus there are exactly $48 - (15 + 20 + 1) = 12$ hexads h' intersecting h in p_1. □

Theorem 9.4.3. *Let $h = \{p_1, p_2, p_3, p_4, p_5, p_6\}$ be a hexad in $P^2(\mathbb{F}_4)$. Of the 168 hexads in $P^2(\mathbb{F}_4)$, there are:*

- *1 hexad intersecting h in 4 or more vertices;*
- *40 hexads intersecting h in 3 vertices;*
- *45 hexads intersecting h in 2 vertices;*
- *72 hexads intersecting h in 1 vertex; and*
- *10 hexads intersecting h in 0 vertices.*

Proof. Since a tetrad lies in only 1 hexad by Corollary 9.2.3, only h itself intersects h in 4 or more vertices.

From Lemma 9.4.2, there are 2 hexads intersecting h in any 3 vertices, and thus there are $\binom{6}{3} \cdot 2 = 40$ such hexads.

From Lemma 9.4.2, there are 3 hexads intersecting h in any 2 vertices, and thus there are $\binom{6}{2} \cdot 3 = 45$ such hexads.

From Lemma 9.4.2, there are 12 hexads intersecting h in any 1 vertex, and thus there are $\binom{6}{1} \cdot 12 = 72$ such hexads.

Since these exhaust the possibilities, there are $168 - (1 + 40 + 45 + 72) = 10$ hexads intersecting h in 0 vertices. □

The rest of this chapter will show that the 56 hexads intersecting hexad h in an even number of vertices ($0, 2$, or 6) form an equivalence class, identical to the orbit of h under $PSL_3(\mathbb{F}_4)$. Let $\sim$ denote the relation on hexads of intersecting in an even number of points. It is obvious $\sim$ is reflexive and symmetric. To show transitivity, we must show $h_1 \sim h_0$ and $h_0 \sim h_2$ implies $h_1 \sim h_2$. Since symmetry holds, this is equivalent to showing $h_0 \sim h_1$ and $h_0 \sim h_2$ implies $h_1 \sim h_2$. There are three cases to be checked: $|h_0 \cap h_1| = 0$ and $|h_0 \cap h_2| = 0$ (Theorem 9.4.4), $|h_0 \cap h_1| = 0$ and $|h_0 \cap h_2| = 2$ (Theorem 9.4.10), and $|h_0 \cap h_1| = 2$ and $|h_0 \cap h_2| = 2$ (Corollary 9.4.6 and Theorem 9.4.12).

Theorem 9.4.4. *If h_0, h_1, h_2 are hexads such that $|h_0 \cap h_1| = 0$ and $|h_0 \cap h_2| = 0$, then $|h_1 \cap h_2| = 2$.*

Proof. By Theorem 9.4.3, there are 10 hexads intersecting h_0 in 0 points, and we let h_1, h_2 be two of them. Let H_0 be the hexagram corresponding to h_0, as in Lemma 9.3.13. We consider the hexads formed by partitioning the sides of H_0 into two triads, as in Lemma 9.3.15. Each hexad formed in this way has its points selected from the sides of H_0. Since the sides of H_0 are skew to h_0, these hexads are all disjoint from h_0. By Lemma 9.3.15, there are 10 such hexads, with any two sharing exactly 2 points. Thus each of the 10 hexads that is disjoint from h_0 is formed in this manner, and we have our desired intersection property. □

Theorem 9.4.5. *Let $h_0 = \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}$, and let h_1, h_2, h_3 be the 3 hexads intersecting h_0 in exactly $p_5 = [\omega\bar{\omega}1]$ and $p_6 = [\bar{\omega}\omega 1]$. Let*

$q_i = h_i \setminus \{p_5, p_6\}$, *for* $0 \leq i \leq 3$. *Then:*

$$\begin{aligned} q_0 &= \{[100], [010], [001], [111]\}, \\ q_1 &= \{[\bar{\omega}11], [01\omega], [0\omega1], [\omega11]\}, \\ q_2 &= \{[1\bar{\omega}1], [\omega01], [10\omega], [1\omega1]\}, \\ q_3 &= \{[11\bar{\omega}], [1\omega0], [\omega10], [11\omega]\}. \end{aligned} \tag{9.18}$$

Thus:

(1) For $0 \leq i \leq 3$, $q_i \cap q_j = \varnothing$.

(2) The 5 *points not on* $q_0 \cup q_1 \cup q_2 \cup q_3$ *are the line* $\overline{p_5 p_6}$.

(3) For $0 \leq i \leq 3$, *the diagonal points of the* q_i *are the points on the line* $\overline{p_5 p_6}$ *that are not* p_5 *or* p_6.

Further, all the q_i *(and thus the* h_i*) are in the same* $PSL_3(\mathbb{F}_4)$ *orbit.*

Proof. By Lemma 9.4.2, there are exactly 3 hexads sharing p_5, p_6 with h_0, which we denote h_1, h_2, h_3. Consider q_0, q_1, q_2, q_3 above, of which q_0 is a tetrad by Corollary 9.1.6. If

$$A_1 = \begin{bmatrix} \bar{\omega} & 1 & 1 \\ 0 & 1 & \omega \\ 0 & \omega & 1 \end{bmatrix}, \tag{9.19}$$

then

$$[100]A_1 = [\bar{\omega}11],\ [010]A_1 = [01\omega],\ [001]A_1 = [0\omega1],\ [111]A_1 = [\omega11], \tag{9.20}$$

$$[\omega\bar{\omega}1]A_1 = [\bar{\omega}\omega1],\ [\bar{\omega}\omega1]A_1 = [\omega\bar{\omega}1], \tag{9.21}$$

and

$$[011]A_1 = [011],\ [101]A_1 = [110],\ [110]A_1 = [101]. \tag{9.22}$$

We also have $\det A_1 = \bar{\omega}(1+\bar{\omega}) = 1$. Thus $A_1 \in SL_3(\mathbb{F}_4)$ sends q_0 to q_1, and h_0 to h_1. Elements of $PGL_3(\mathbb{F}_4)$ preserve lines in $P^2(\mathbb{F}_4)$. Since tetrads and hexads are defined in terms of lines, elements of $PGL_3(\mathbb{F}_4)$ preserve tetrads and hexads. Thus q_1 is a tetrad and h_1 is a hexad. Since A_1 preserves tetrads, it also preserves their diagonal points. We see q_1 has the same diagonal points as q_0.

Now q_2 and q_3 differ from q_1 by a cyclic permutation of the coordinates. If $A_1 = [v_1 v_2 v_3]$, where v_i is the ith column of A_1, let $A_2 = [v_3 v_1 v_2]$ and $A_3 = [v_2 v_3 v_1]$. Thus A_2 sends q_0 to q_2, and A_3 sends q_0 to q_3. Permutations of the columns of a matrix change the determinant by at most a sign, and in characteristic 2 we have $-1 = 1$. Thus $A_2, A_3 \in SL_3(\mathbb{F}_4)$, implying that q_0, q_1, q_2, q_3 are tetrads in the same $PSL_3(\mathbb{F}_4)$ orbit. Now

$$[1\omega\bar{\omega}] = [\bar{\omega}1\omega] = [\omega\bar{\omega}1] = p_5, \quad [1\bar{\omega}\omega] = [\omega 1\bar{\omega}] = [\bar{\omega}\omega 1] = p_6, \tag{9.23}$$

and so p_5, p_6 are preserved by cyclic permutation of the vertices. Thus h_0 is sent to h_2 by A_2, and h_0 is sent to h_3 by A_3, and h_0, h_1, h_2, h_3 are hexads in the same $PSL_3(\mathbb{F}_4)$ orbit.

By inspection of these sets, we see that for $i \neq j$,

$$h_i \cap h_j = \{p_5, p_6\} \text{ and } q_i \cap q_j = \varnothing. \tag{9.24}$$

Now the set of 3 diagonal points is preserved by cyclic permutations of the columns of A_1. Thus A_2 and A_3 send those points to themselves, so q_0, q_1, q_2, q_3 share the same diagonal points, the points completing $\overline{p_5 p_6}$. Finally, q_0, q_1, q_2, q_3 are pairwise disjoint, and so their union accounts for 16 points of $P^2(\mathbb{F}_4)$. Thus the 5 points of $\overline{p_5 p_6}$ make up the remainder. □

Corollary 9.4.6. *If h_0, h_1, h_2 are hexads such that $h_0 \cap h_1 = h_0 \cap h_2 = \{x, y\}$, where $x, y \in P^2(\mathbb{F}_4)$, then $|h_1 \cap h_2| = 2$. Further, these 3 hexads are in the same $PSL_3(\mathbb{F}_4)$ orbit.*

Proof. Since $PGL_3(\mathbb{F}_4)$ acts transitively on hexads in $P^2(\mathbb{F}_4)$, we choose h_0 as in Theorem 9.4.5, without loss of generality. Since the stabilizer of h_0 in $PGL_3(\mathbb{F}_4)$ is isomorphic to A_6, it is 4-transitive on the points of h_0. Thus, without loss of generality, $\{x, y\} = \{p_5, p_6\}$, and we apply Theorem 9.4.5. □

Corollary 9.4.7. *If h_0, h_1 are hexads such that $|h_0 \cap h_1| = 2$, then h_0, h_1 are in the same $PSL_3(\mathbb{F}_4)$ orbit.*

Proof. Now $PGL_3(\mathbb{F}_4)$ acts transitively on hexads, and the stabilizer of h_0 in $PGL_3(\mathbb{F}_4)$ is isomorphic to A_6. Thus without loss of generality, we may choose h_0, h_1 and $h_0 \cap h_1 = \{[\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}$ as in Theorem 9.4.5. □

Lemma 9.4.8. *Let $h_0 = \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}$. The following hexads are disjoint from h_0:*

$$
\begin{aligned}
h'_1 &= \{[011], [110], [\omega 10], [0\omega 1], [11\bar{\omega}], [\omega 11]\} \;\; \ell_1, \ell_2, \ell_3 : \ell_4, \ell_5, \ell_6. \\
h'_2 &= \{[101], [011], [0\omega 1], [10\omega], [\bar{\omega}11], [1\omega 1]\} \;\; \ell_1, \ell_2, \ell_6 : \ell_3, \ell_4, \ell_5. \\
h'_3 &= \{[110], [101], [10\omega], [\omega 10], [1\bar{\omega}1], [11\omega]\} \;\; \ell_1, \ell_5, \ell_6 : \ell_2, \ell_3, \ell_4. \\
h'_4 &= \{[110], [011], [01\omega], [1\omega 0], [\bar{\omega}11], [11\omega]\} \;\; \ell_1, \ell_2, \ell_4 : \ell_3, \ell_5, \ell_6. \\
h'_5 &= \{[011], [101], [\omega 01], [01\omega], [1\bar{\omega}1], [\omega 11]\} \;\; \ell_1, \ell_2, \ell_5 : \ell_3, \ell_4, \ell_6. \\
h'_6 &= \{[101], [110], [1\omega 0], [\omega 01], [11\bar{\omega}], [1\omega 1]\} \;\; \ell_1, \ell_3, \ell_4 : \ell_2, \ell_5, \ell_6. \\
h'_7 &= \{[\omega 11], [\bar{\omega}11], [1\omega 0], [10\omega], [\omega 10], [\omega 01]\} \;\; \ell_1, \ell_4, \ell_6 : \ell_2, \ell_3, \ell_5. \\
h'_8 &= \{[1\omega 1], [1\bar{\omega}1], [01\omega], [\omega 10], [0\omega 1], [1\omega 0]\} \;\; \ell_1, \ell_4, \ell_5 : \ell_2, \ell_3, \ell_6. \\
h'_9 &= \{[11\omega], [11\bar{\omega}], [\omega 01], [0\omega 1], [10\omega], [01\omega]\} \;\; \ell_1, \ell_3, \ell_6 : \ell_2, \ell_4, \ell_5. \\
h'_{10} &= \{[11\omega], [\omega 11], [1\omega 1], [11\bar{\omega}], [\bar{\omega}11], [1\bar{\omega}1]\} \;\; \ell_1, \ell_3, \ell_5 : \ell_2, \ell_4, \ell_6.
\end{aligned}
$$

These are listed with the corresponding partition of H_0 into two triads.

Proof. This follows from Lemma 9.3.15 and Corollary 9.3.18. □

Lemma 9.4.9. *Let $G = PGL_3(\mathbb{F}_4)$, and let*

$$h_0 = \{p_1 = [100], p_2 = [010], p_3 = [001], p_4 = [111], p_5 = [\omega\bar{\omega}1], p_6 = [\bar{\omega}\omega 1]\}. \quad (9.25)$$

There is a subgroup of $\mathrm{Stab}_G(h_0) \cap \mathrm{Stab}_G(\{p_5, p_6\})$ *that stabilizes the set of* 3 *hexads intersecting* h_0 *in* $\{p_5, p_6\}$*, and the set of* 10 *hexads disjoint from* h_0*. This is the group of permutations of the* 3 *coordinates in* $\mathbb{F}_4^3$*.*

Proof. Now coordinate permutations are elements of $GL_3(\mathbb{F}_4)$ as they result from permutation matrices acting on $P^2(\mathbb{F}_4)$. Applying permutations (123) or (132) to the coordinates moves each of p_1, p_2, p_3 and fixes p_4, p_5, p_6. Applying permutations $(12), (13)$ or (23) to the coordinates moves two of p_1, p_2, p_3, while fixing p_4 and switching p_5, p_6. Both types induce even permutations of the points of h_0 and stabilize the sets $\{p_4\}$ and $\{p_5, p_6\}$. Thus they generate a subgroup of $\mathrm{Stab}(h_0) \cap \mathrm{Stab}(\{p_5, p_6\})$ isomorphic to S_3. Such a subgroup must preserve the set of 3 hexads sharing p_5, p_6 with h_0, as well as the set of 10 hexads disjoint from h_0. □

Theorem 9.4.10. *Let h, h', h'' be hexads such that $|h \cap h'| = 2$ and $|h \cap h''| = 0$. Then $|h' \cap h''| = 0$ or 2. Also, there exists a hexad h''' such that $|h''' \cap h| = 2$ and $|h''' \cap h''| = 2$.*

Proof. Since $PGL_3(\mathbb{F}_4)$ acts transitively on hexads, without loss of generality we choose

$$h_0 = \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}. \quad (9.26)$$

Since $\mathrm{Stab}(h_0)$ is isomorphic to A_6, which is 4-transitive, without loss of generality we choose $[\omega\bar{\omega}1], [\bar{\omega}\omega 1]$ as the 2 points of intersection with other hexads. By

Theorem 9.4.5, the 3 hexads intersecting h_0 in these points are

$$\begin{aligned} h_1 &= \{[\bar{\omega}11], [01\omega], [0\omega1], [\omega11], [\omega\bar{\omega}1], [\bar{\omega}\omega1]\}, \\ h_2 &= \{[1\bar{\omega}1], [\omega01], [10\omega], [1\omega1], [\omega\bar{\omega}1], [\bar{\omega}\omega1]\}, \\ h_3 &= \{[11\bar{\omega}], [1\omega0], [\omega10], [11\omega], [\omega\bar{\omega}1], [\bar{\omega}\omega1]\}. \end{aligned} \tag{9.27}$$

By Lemma 9.4.9, the group of permutations of the 3 coordinates of $\mathbb{F}_4^3$ is a subgroup of $\mathrm{Stab}(h_0) \cap \mathrm{Stab}(\{[\omega\bar{\omega}1], [\bar{\omega}\omega1]\})$ preserving $\{h_1, h_2, h_3\}$ and the 10 hexads disjoint from h_0. Thus if k is one of these disjoint hexads and π is an element of this subgroup, then $k\pi$ has the same intersection properties as k with hexads h_1, h_2, h_3. Thus we need only determine the orbits of the 10 hexads under this group, and test representatives. Replacing the notation of h_i' to k_i (for $1 \leq i \leq 10$) from Lemma 9.4.8, we see that

$$\{k_1, k_2, k_3\},\ \{k_4, k_5, k_6\},\ \{k_7, k_8, k_9\},\ \{k_{10}\} \tag{9.28}$$

are the orbits under a 3-cycle from this group. A transposition of the first 2 coordinates sends k_1 to k_4, and so we have the orbits

$$\{k_1, k_2, k_3, k_4, k_5, k_6\},\ \{k_7, k_8, k_9\},\ \{k_{10}\} \tag{9.29}$$

under the action of this subgroup. Thus we take the following as representatives:

$$\begin{aligned} k_1 &= \{[011], [110], [\omega10], [0\omega1], [11\bar{\omega}], [\omega11]\}, \\ k_7 &= \{[\omega11], [\bar{\omega}11], [1\omega0], [10\omega], [\omega10], [\omega01]\}, \\ k_{10} &= \{[11\omega], [\omega11], [1\omega1], [11\bar{\omega}], [\bar{\omega}11], [1\bar{\omega}1]\}. \end{aligned}$$

Now k_1 intersects h_1 and h_3 in 2 points and h_2 in 0 points, and k_7 and k_{10} each intersect h_1, h_2 and h_3 in 2 points. If we take $h = h_0$, $h' = h_2$, $h'' = k_1$ or k_7, then the first assertion follows.

As for the second assertion, for all three h_i, and for the k_j, $7 \leq j \leq 10$, $|h_i \cap k_j| = 2$. If $1 \leq j \leq 6$, then k_j intersects two of the three h_i in 2 points. Thus we can always find an h_i satisfying our second assertion. For example, if we take $h = h_0$ and $h'' = k_1$ or k_7, then $h''' = h_1$ satisfies our intersection requirement. □

Lemma 9.4.11. *Let*

$$h_0 = \{p_1 = [100], p_2 = [010], p_3 = [001], p_4 = [111], p_5 = [\omega\bar{\omega}1], p_6 = [\bar{\omega}\omega 1]\}. \quad (9.30)$$

Let h_1, h_2, h_3 be the 3 hexads intersecting h_0 in $\{p_5, p_6\}$, and h'_1, h'_2, h'_3 the 3 hexads intersecting h_0 in $\{p_1, p_2\}$. Then $|h_i \cap h'_j| = 0$ or 2. Let h''_1, h''_2, h''_3 be the 3 hexads intersecting h_0 in $\{p_1, p_5\}$. Then $|h_i \cap h''_j| = 2$, for $1 \leq i, j \leq 3$.

Proof. From Theorem 9.4.5, we have:

$$\begin{aligned} h_1 &= \{[\bar{\omega}11], [01\omega], [0\omega 1], [\omega 11], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}, \\ h_2 &= \{[1\bar{\omega}1], [\omega 01], [10\omega], [1\omega 1], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}, \\ h_3 &= \{[11\bar{\omega}], [1\omega 0], [\omega 10], [11\omega], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}. \end{aligned} \quad (9.31)$$

Consider the matrix

$$A' = \begin{bmatrix} \omega & \bar{\omega} & 1 \\ \bar{\omega} & \omega & 1 \\ 0 & 0 & 1 \end{bmatrix}. \quad (9.32)$$

Since $A' \in PGL_3(\mathbb{F}_4)$, A' preserves hexads. We also have $p_5 A' = p_1$ and $p_6 A' = p_2$. Thus

$$\begin{aligned} h'_1 &= h_1 A' = \{[10\omega], [\omega 11], [1\bar{\omega}1], [0\omega 1], [100], [010]\}, \\ h'_2 &= h_2 A' = \{[01\omega], [1\omega 1], [\bar{\omega}11], [\omega 01], [100], [010]\}, \\ h'_3 &= h_3 A' = \{[11\bar{\omega}], [101], [011], [11\omega], [100], [010]\}, \end{aligned} \quad (9.33)$$

are all hexads, and so they are the three hexads intersecting h_0 in $\{p_1, p_2\}$. By inspection, we see $|h_i \cap h'_j| = 0$ or 2, for $1 \leq i, j \leq 3$.

Now consider the matrix

$$A'' = \begin{bmatrix} 1 & \bar{\omega} & \omega \\ 0 & \omega & 0 \\ 0 & 0 & \bar{\omega} \end{bmatrix}. \tag{9.34}$$

Since $A'' \in PGL_3(\mathbb{F}_4)$, A'' preserves hexads. We also have $p_5 A'' = p_1$ and $p_6 A' = p_5$. Thus

$$\begin{aligned} h''_1 &= h_1 A'' = \{[\omega 01], [1\omega 0], [0\omega 1], [011], [100], [\omega\bar{\omega}1]\}, \\ h''_2 &= h_2 A'' = \{[\omega 11], [1\omega 1], [\omega 10], [101], [100], [\omega\bar{\omega}1]\}, \\ h''_3 &= h_3 A'' = \{[\bar{\omega}11], [11\bar{\omega}], [10\omega], [110], [100], [\omega\bar{\omega}1]\}, \end{aligned} \tag{9.35}$$

are all hexads, and so they are the three hexads intersecting h_0 in $\{p_1, p_5\}$. By inspection, we see $|h_i \cap h''_j| = 2$, for $1 \leq i, j \leq 3$. □

Theorem 9.4.12. *If h_0, h_1, h'_1 are hexads such that $|h_0 \cap h_1| = 2$ and $|h_0 \cap h'_1| = 2$, and $|(h_0 \cap h_1) \cap (h_0 \cap h'_1)| = 0$, then $|h_1 \cap h'_1| = 0$ or 2.*

If h_0, h_1, h''_1 are hexads such that $|h_0 \cap h_1| = 2$ and $|h_0 \cap h''_1| = 2$, and $|(h_0 \cap h_1) \cap (h_0 \cap h''_1)| = 1$, then $|h_1 \cap h''_1| = 2$.

Proof. Since $PGL_3(\mathbb{F}_4)$ acts transitively on hexads in $P^2(\mathbb{F}_4)$, without loss of generality we can choose h_0 as in Theorem 9.4.5. Since the stabilizer of h_0 in $PGL_3(\mathbb{F}_4)$ is isomorphic to A_6, it is 4-transitive on the points of h_0. Thus, without loss of generality, we can choose $h_0 \cap h_1 = \{p_1, p_2\}$ and $h_0 \cap h'_1 = \{p_5, p_6\}$.

Similarly, without loss of generality, we can also choose $h_0 \cap h_1 = \{p_1, p_2\}$ and $h_0 \cap h''_1 = \{p_1, p_5\}$. By Lemma 9.4.11, we know that the intersection properties are as described for both statements. □

Theorem 9.4.13. *Hexads intersecting evenly is an equivalence relation whose equivalence classes are $PSL_3(\mathbb{F}_4)$-orbits. Thus hexads are in the same $PSL_3(\mathbb{F}_4)$-orbit if and only if they share an even number of points.*

Proof. Let h be a hexad. Theorems 9.4.4, 9.4.6, 9.4.10, and 9.4.12 establish that having an even number of intersections with h is a transitive relation on hexads. Transitivity, along with reflexivity and symmetry, makes the 56 hexads that intersect h evenly an equivalence class. Next, by Corollary 9.4.7, hexads h' such that $|h \cap h'| = 2$ are in the same $PSL_3(\mathbb{F}_4)$-orbit. Furthermore, by Theorem 9.4.10, we know that if $|h \cap h'| = 0$, there exists some hexad h'' such that $|h \cap h''| = 2$ and $|h' \cap h''| = 2$. Since h and h'' intersect in 2 points, they are in the same $PSL_3(\mathbb{F}_4)$-orbit, and the same holds true for h' and h''. Thus h and h' are in the same orbit, and all 56 hexads are part of the same $PSL_3(\mathbb{F}_4)$-orbit. By Theorem 9.2.7, each orbit has size 56, and thus equivalence classes are identical with $PSL_3(\mathbb{F}_4)$-orbits. □

Corollary 9.4.14. *Hexads in different $PSL_3(\mathbb{F}_4)$ orbits intersect in 1 or 3 points.*

Proof. By Theorem 9.4.13, hexads are in the same $PSL_3(\mathbb{F}_4)$-orbit if and only if they share an even number of points. Thus hexads in different $PSL_3(\mathbb{F}_4)$-orbits share an odd number of points. Since 4 points determine a hexad, this is either 1 or 3 points. □

Theorem 9.4.15. *Let $G = P\Gamma L_3(\mathbb{F}_4)$ and $N = PSL_3(\mathbb{F}_4)$. The action of G on $P^2(\mathbb{F}_4)$ preserves the N-orbits of hexads.*

Proof. Let h_1, h_2 be in the same N-orbit and let $T \in G$. Since h_1 and h_2 are in the same N-orbit, there is an $A \in SL_3(\mathbb{F}_4)$ such that $h_1 A = h_2$. By Theorem 5.2.15,

$SL_3(\mathbb{F}_4) \lhd \Gamma L_3(\mathbb{F}_4)$, and thus $A' = T^{-1}AT \in SL_3(\mathbb{F}_4)$. We have

$$h_2T = (h_1A)T = h_1(AT) = h_1(TA') = (h_1T)A'. \tag{9.36}$$

By definition, h_1T and $(h_1T)A' = h_2T$ are in the same N-orbit. Therefore, $T \in G$ preserves the N-orbits of hexads. $\square$

CHAPTER 10

BINARY LINEAR CODES

In this chapter, we introduce binary linear codes in preparation for Chapter 11. We cover basic properties of such codes, and then move on to special characteristics possessed by some binary linear codes. Lastly, we define the automorphism group of a code, giving an important criterion for determining whether a given permutation is in such a group. We follow the treatment of Pless [Ple89].

10.1 Binary Linear Codes

In this section, we define some basic properties of binary linear codes: Hamming distance, Hamming weight, the minimum weight of a code, and the weight distribution of a code. We also describe an important identification between codewords in a vector space of dimension n and subsets of $\{1, \ldots, n\}$.

Definition 10.1.1. A *binary linear code* is a k-dimensional subspace $\mathcal{C}$ of $\mathbb{F}_2^n$, and is referred to as an (n, k)-*code*. The vectors of $\mathcal{C}$ are called *codewords*. Using the standard basis $\beta = \{e_1, \ldots, e_n\}$, we often abbreviate the coordinate vector $(a_1, a_2 \ldots, a_n)$ as $(a_1 a_2 \ldots a_n)$, where $a_i \in \{0, 1\}$.

Definition 10.1.2. Let $x \in V$. If $X = \{i_1, i_2, \ldots, i_k\} \subseteq \{1, \ldots, n\}$ are the nonzero coordinates of x, then this set represents x. Thus $x = (1000011) = e_1 + e_6 + e_7$ is represented as $X = \{1, 6, 7\}$. The sum of vectors corresponds to the symmetric

difference of sets: if vectors x, y correspond to sets X, Y, then $x + y$ corresponds to

$$X \Delta Y = (X \cup Y) \setminus (X \cap Y) = (X \setminus Y) \cup (Y \setminus X). \tag{10.1}$$

We have $0 \in V$ represented by $\varnothing$.

Definition 10.1.3. Let $S = \{v_1, \ldots, v_m\} \subseteq \mathbb{F}_2^n$. The *code generated by* S is $\mathcal{C} = \operatorname{Span} S$. Thus if $\mathcal{C}$ is an (n, k)-code, then $m \geq k$, and we do not assume our generating set S is linearly independent.

Definition 10.1.4. The *Hamming distance* between vectors $u, v \in \mathcal{C}$ is the number of coordinates in which they differ, and is denoted $d(u, v)$. Thus $d(u, v) = \sum_{u_i \neq v_i} 1$.

Lemma 10.1.5. *The Hamming distance between vectors is a metric on* $\mathbb{F}_2^n$.

Proof. By definition, $d(u, v) = 0$ if and only if vectors $u, v \in \mathcal{C}$ have no coordinates in which they differ, that is, when $u = v$. Since differing in the ith coordinate is a symmetric relation, we have $d(u, v) = d(v, u)$. Let $u, v, w \in \mathcal{C}$, and let r be the number of coordinates that u differs from v, and s the number of coordinates that v differs from w. If these coordinates are all distinct, then $r + s$ is the number of coordinates that u differs from w. If they are not all distinct, then u differs from w in fewer than $r + s$ coordinates. Thus $d(u, w) \leq d(u, v) + d(v, w)$, and Hamming distance obeys the triangle inequality. □

Definition 10.1.6. The *Hamming weight* of a vector is the number of nonzero components it possesses, and is denoted $wt(v)$. Thus $wt(v) = \sum_{v_i \neq 0} 1$.

Lemma 10.1.7. *For* $u, v \in \mathbb{F}_2{}^n$, *we have* $d(u, v) = wt(u - v)$.

Proof. We have:

$$wt(u - v) = \sum_{(u-v)_i \neq 0} 1 = \sum_{u_i - v_i \neq 0} 1 = \sum_{u_i \neq v_i} 1 = d(u, v). \tag{10.2}$$

□

Definition 10.1.8. A code $\mathcal{C}$ with *minimum weight* d has $d \in \mathbb{N}$ as the minimum weight of its nonzero codewords. An (n, k)-code of minimum weight d is referred to as an (n, k, d)-code.

Definition 10.1.9. Let $\mathcal{C}$ be a code, and let A_k denote the number of codewords of weight k. Let $k_1 < k_2 < \cdots < k_i$ be the Hamming weights that occur in $\mathcal{C}$. Then the *weight distribution* of $\mathcal{C}$ is denoted

$$k_1^{A_{k_1}} k_2^{A_{k_2}} \ldots k_i^{A_{k_i}}. \tag{10.3}$$

The *weight enumerator* of $\mathcal{C}$ is the polynomial

$$\sum_{k=0}^{n} A_k \, x^{n-k} y^k. \tag{10.4}$$

The weight distribution of $\mathcal{C}$ is *symmetric* if $A_k = A_{n-k}$, for $0 \leq k \leq n$.

10.2 Self-Orthogonal and Self-Dual Codes

In this section, we define dual codes. We also look at self-orthogonal and doubly even codes, developing a criterion for each of these characteristics. Finally, we consider self-dual codes, and a certain consequence of being self-dual.

Remark 10.2.1. In the remainder of this chapter, $\cdot$ denotes the *dot product*, modulo 2. By Lemma 6.4.13, the dot product is a nondegenerate symmetric bilinear form, and thus orthogonality is reflexive (Definition 6.4.2) for binary linear codes.

Definition 10.2.2. If $\mathcal{C}$ is a code in V, we refer to $\mathcal{C}^\perp$ as the *dual code* or *orthogonal code* of $\mathcal{C}$.

Theorem 10.2.3. *If $\mathcal{C}$ is an (n, k)-code, then $\mathcal{C}^\perp$ is an $(n, n-k)$-code.*

Proof. By Theorems 6.5.4 and 6.5.7, $\mathcal{C}^\perp$ is a $(n-k)$-dimensional subspace of V, making it an $(n, n-k)$-code. □

Corollary 10.2.4. *We have* $\mathcal{C}^{\perp\perp} = \mathcal{C}$.

Proof. Since $\cdot$ is a nondegenerate symmetric bilinear form on V, this follows from Corollary 6.5.8. □

Definition 10.2.5. If $\mathcal{C} \leq \mathcal{C}^{\perp}$, then we say that $\mathcal{C}$ is *self-orthogonal*. Thus $u \cdot v = 0$ for all $u, v \in \mathcal{C}$ if and only $\mathcal{C}$ is self-orthogonal.

Lemma 10.2.6. *If an* (n, k)*-code* $\mathcal{C}$ *is self-orthogonal, then* $k \leq n - k$*, that is,* $k \leq n/2$.

Proof. Since $\mathcal{C} \leq \mathcal{C}^{\perp}$,

$$k = \dim \mathcal{C} \leq \dim \mathcal{C}^{\perp} = n - k. \tag{10.5}$$

Therefore, $2k \leq n$, or $k \leq n/2$. □

Definition 10.2.7. A binary code is *even* if all its codewords have even weight.

Lemma 10.2.8. *If* $\mathcal{C}$ *is self-orthogonal, then* $\mathcal{C}$ *is even.*

Proof. Let $u \in \mathcal{C} \leq \mathcal{C}^{\perp}$. Thus $u \cdot u = \sum_i {u_i}^2 = 0$. This only happens if there are an even number of 1's in u, making u of even weight. □

Theorem 10.2.9. *The following are equivalent for a binary* (n, k)*-code* $\mathcal{C}$*:*

(1) $\mathcal{C}$ *is self-orthogonal.*

(2) The vectors in some generating set S *for* $\mathcal{C}$ *have even weight and are pairwise orthogonal.*

Proof. Let $\mathcal{C}$ be self-orthogonal. Now the vectors in a generating set S are codewords in $\mathcal{C}$. Thus by Lemma 10.2.8, they are of even weight. Let $v_i, v_j \in S \subseteq \mathcal{C}$. Since $\mathcal{C} \leq \mathcal{C}^{\perp}$, $v_i \cdot v_j = 0$, and the two are orthogonal.

Let the vectors in some generating set $S = \{v_1, \ldots, v_m\}$ have even weight and be pairwise orthogonal. Fix $x \in \mathcal{C}$ and let $y \in \mathcal{C}$. By the bilinearity of the dot product,

$$x \cdot y = \left(\sum_i x_i v_i\right) \cdot \left(\sum_j y_j v_j\right) = \sum_{i,j} x_i y_j (v_i \cdot v_j) = \sum_{i,j} x_i y_j (0) = 0. \tag{10.6}$$

Thus $x \in \mathcal{C}^\perp$, and so $\mathcal{C} \leq \mathcal{C}^\perp$. □

Definition 10.2.10. For $u, v \in \mathbb{F}_2^n$, define $u \cap v$ as the vector with a 1 in every coordinate in which u and v agree in having a 1, and 0 elsewhere. Thus $u \cap v$ corresponds to $U \cap V$, where U, V are the set representations of u, v (as in Remark 10.1.2).

Lemma 10.2.11. *For $u, v \in \mathbb{F}_2^n$, we have $wt(u + v) = wt(u) + wt(v) - 2wt(u \cap v)$.*

Proof. The number of 1's in a vector is its weight. If two vectors have a 1 in common, there is a 0 after addition. If we add the number of 1's in each of two vectors, and subtract the number of 1's in their sum, we almost have $wt(u \cap v)$; specifically, the 1's they had in common were double-counted, and so we must divide by 2. Thus

$$wt(u \cap v) = \left(wt(u) + wt(v) - wt(u + v)\right)/2. \tag{10.7}$$

Our result follows. □

Lemma 10.2.12. *For $u, v \in \mathbb{F}_2^n$, we have $wt(u \cap v) = u \cdot v \pmod 2$.*

Proof. We have

$$u \cdot v = \sum_i u_i v_i = \sum_{u_i = v_i = 1} 1 \pmod 2. \tag{10.8}$$

This value is 0 if u, v agree in an even number 1's, and 1 otherwise. Modulo 2, this is precisely when $wt(u \cap v)$ is 0 or 1. □

Definition 10.2.13. A binary code is *doubly even* if all its codewords have weight divisible by 4.

Lemma 10.2.14. *A doubly even code $\mathcal{C}$ is self-orthogonal.*

Proof. Let $\mathcal{C}$ be doubly even, and let $u, v \in \mathcal{C}$. It follows from Lemma 10.2.12 that

$$2wt(u \cap v) = 2(u \cdot v) \pmod 4. \tag{10.9}$$

By hypothesis, $u, v, u+v$ all have weight divisible by 4. Using Lemma 10.2.11, we have

$$\begin{aligned} 2(u \cdot v) = 2wt(u \cap v) &= 2wt(u \cap v) - wt(u) - wt(v) \\ &= -wt(u+v) = 0 \pmod 4. \end{aligned} \tag{10.10}$$

Thus $u \cdot v = 0 \pmod 2$, and $\mathcal{C}$ is self-orthogonal. □

Theorem 10.2.15. *Let $\mathcal{C}$ be a binary (n,k)-code. If the vectors of a generating set S have weight divisible by 4 and are pairwise orthogonal, then $\mathcal{C}$ is doubly even.*

Proof. By Theorem 10.2.9, $\mathcal{C}$ is self-orthogonal, and so it remains to verify that every codeword has weight divisible by 4. Let $S = \{v_1, \ldots, v_m\}$ generate $\mathcal{C}$, and let $v_i, v_j \in S$. By Lemma 10.2.11,

$$wt(v_i + v_j) = wt(v_i) + wt(v_j) - 2wt(v_i \cap v_j). \tag{10.11}$$

Now $v_i \cdot v_j = 0$ and so $wt(v_i \cap v_j)$ is even, by Lemma 10.2.12. Thus every term on the right hand side is divisible by 4, and thus $wt(v_i + v_j)$ is as well.

Let the weight of the sum of any k elements of S be divisible by 4, and let $x \in \mathcal{C}$ be the sum of $k+1$ elements from S. Re-ordering elements of S if needed, we have

$$x = \sum_{i=1}^{k+1} v_i = \sum_{i=1}^{k} v_i + v_{k+1} = y + v_{k+1}. \tag{10.12}$$

Now y has weight divisible by 4, as does v_{k+1}. Since $\mathcal{C}$ is self-orthogonal, $y \cdot v_{k+1} = 0$, and so $wt(y \cap v_{k+1})$ is even, by the same reasoning as above. Therefore, every term on the right hand side of

$$wt(x) = wt(y + v_k) = wt(y) + wt(v_k) - 2wt(y \cap v_k) \tag{10.13}$$

is divisible by 4, and so $wt(x)$ is also. By induction, we conclude every codeword of $\mathcal{C}$ has weight divisible by 4. □

Definition 10.2.16. An (n, k)-code $\mathcal{C}$ is *self-dual* if $\mathcal{C} = \mathcal{C}^{\perp}$.

Theorem 10.2.17. *If $\mathcal{C}$ is a self-dual (n, k)-code, then $k = n/2$.*

Proof. If $\mathcal{C}$ is an (n, k)-code, then $\mathcal{C}^{\perp}$ is an $(n, n-k)$-code. If $\mathcal{C} = \mathcal{C}^{\perp}$, then $k = n - k$, implying $n = 2k$ and $k = n/2$. □

Definition 10.2.18. Let $\mathbf{1}$ denote the vector of $\mathbb{F}_2^n$ that has a 1 for every coordinate.

Theorem 10.2.19. *If $\mathcal{C}$ is a self-dual (n, k)-code, then n is even, $\mathbf{1} \in \mathcal{C}$, and the weight distribution of $\mathcal{C}$ is symmetric.*

Proof. If $\mathcal{C}$ is self-dual, then $\mathcal{C}$ is an $(n, n/2)$-code, by Theorem 10.2.17. Thus n is even, and so $\mathbf{1}$ is of even weight. Assume $\mathbf{1} \notin \mathcal{C}$. Then the code generated by $\mathcal{C}$ and $\mathbf{1}$, $\mathcal{C} + \mathbf{1}$, is of dimension $(n/2) + 1$.

Let $S + \mathbf{1}$ be a generating set of $\mathcal{C} + \mathbf{1}$, where S is a generating set of $\mathcal{C}$. Since $\mathcal{C}$ is self-orthogonal, the elements of S are pairwise orthogonal. Now $\mathbf{1}$ intersects each codeword of $\mathcal{C}$ in an even number of coordinates, and so each element of S is orthogonal to $\mathbf{1}$. Thus $\mathcal{C} + \mathbf{1}$ is self-orthogonal by Theorem 10.2.9. But Lemma 10.2.6 implies that

$$\dim(C + \mathbf{1}) = ((n/2) + 1) \leq n/2, \tag{10.14}$$

a contradiction. Thus $\mathbf{1} \in \mathcal{C}$.

If $x \in \mathcal{C}$ is of weight k, then $(x + \mathbf{1}) \in \mathcal{C}$ is of weight $n - k$. Thus adding $\mathbf{1}$ gives a bijection between codewords of weight k and those of weight $n - k$. The final result follows. □

10.3 Automorphism Group of a Binary Code

In this section, we define a code's automorphism group, and develop a criterion for determining whether a given coordinate permutation is in a code's automorphism group.

Definition 10.3.1. Let $\beta = \{e_1, \ldots, e_n\}$ be the standard basis for $\mathbb{F}_2^n$ and let $v = \sum_{i=1}^{n} a_i e_i \in \mathbb{F}_2^n$, for some $a_i \in \mathbb{F}_2$. For $\pi \in S_n$, using the natural action of S_n on $\{1, \ldots, n\}$, we define:

$$v\pi = \left(\sum_{i=1}^{n} a_i e_i \right) \pi = \sum_{i=1}^{n} (a_i e_i)\pi = \sum_{i=1}^{n} a_i (e_i \pi) = \sum_{i=1}^{n} a_i e_{i\pi}. \tag{10.15}$$

Lemma 10.3.2. *Definition 10.3.1 gives a group action.*

Proof. Let $v \in \mathbb{F}_2^n$ and $1 \in S_n$. We have

$$v1 = \left(\sum_{i=1}^{n} a_i e_i \right) 1 = \sum_{i=1}^{n} a_i (e_i 1) = \sum_{i=1}^{n} a_i e_i = v. \tag{10.16}$$

Now let $v \in \mathbb{F}_2^n$ and $\pi, \sigma \in S_n$. We have

$$(v\pi)\sigma = \left(\sum_{i=1}^{n} a_i e_{i\pi} \right) \sigma = \sum_{i=1}^{n} a_i e_{(i\pi)\sigma} = \sum_{i=1}^{n} a_i e_{i(\pi\sigma)} = v(\pi\sigma), \tag{10.17}$$

where the third equality follows from the action of S_n on $\{1, \ldots, n\}$. □

Remark 10.3.3. The group action defined in Definition 10.3.1 has an analogue if elements of $\mathbb{F}_2^n$ are viewed as subsets of $\{1, \ldots, n\}$, as in Definition 10.1.2. Represent

$x \in \mathbb{F}_2^n$ as $X = \{i_1, i_2, \ldots, i_k\}$, in which each element of the set is the coordinate index of a nonzero coordinate in x, and $k \leq n$. Thus

$$x = e_{i_1} + e_{i_2} + \ldots + e_{i_k}. \tag{10.18}$$

Then S_n act on $\mathbb{F}_2^n$, such that if $\pi \in S_n$, then

$$x\pi = X\pi = \{i_1, i_2, \ldots, i_k\}\pi = \{(i_1)\pi, (i_2)\pi, \ldots, (i_k)\pi\} = \sum_{i=1}^{n} a_i e_{i\pi}. \tag{10.19}$$

Example 10.3.4. Let $x = e_1 + e_4 = (1, 0, 0, 1) \in F_2^4$ and let $\pi = (1234) \in S_4$. We have

$$x\pi = (e_1 + e_4)\pi = e_{1\pi} + e_{4\pi} = e_2 + e_1 = (1, 1, 0, 0). \tag{10.20}$$

As a set, our vector is $X = \{1, 4\}$ and

$$\{1, 4\}(1234) = \{1(1234), 4(1234)\} = \{2, 1\} = (1, 1, 0, 0). \tag{10.21}$$

Definition 10.3.5. Let $\mathcal{C}$ be a binary (n, k)-code. An *automorphism of* $\mathcal{C}$ is an element of S_n that sends codewords to codewords, according to the action of Definition 10.3.1. Thus if π is an automorphism of $\mathcal{C}$ and $c \in \mathcal{C}$, then $c\pi \in \mathcal{C}$. The *automorphism group of* $\mathcal{C}$ is

$$\mathrm{Aut}(\mathcal{C}) = \{\pi \in S_n \mid c\pi \in \mathcal{C} \text{ for all } c \in \mathcal{C}\}. \tag{10.22}$$

Theorem 10.3.6. *Let $\sigma \in S_n$, and let $\mathcal{C}$ be a binary (n, k)-code generated by $S = \{v_1, \ldots, v_m\}$. Then the following are equivalent:*

(1) $\sigma \in \mathrm{Aut}(\mathcal{C})$.

(2) $v_i\sigma \in \mathcal{C}$ for all i.

Proof. Since $S \subseteq \mathcal{C}$, (1) $\Rightarrow$ (2) follows immediately from the definition of $\mathrm{Aut}(\mathcal{C})$. Conversely, let $v_i\sigma \in \mathcal{C}$ for all i. If $c \in \mathcal{C}$, then

$$c = \sum_{i=1}^{m} a_i v_i, \tag{10.23}$$

for some $a_i \in \mathbb{F}_2$. Also, for $v_i \in S$,

$$v_i = \sum_{j=1}^{n} b_j e_j, \tag{10.24}$$

for some $b_j \in \mathbb{F}_2$. Thus we have:

$$c\sigma = \left(\sum_{i=1}^{m} a_i v_i \right) \sigma = \left(\sum_{i=1}^{m} a_i \sum_{j=1}^{n} b_j e_j \right) \sigma = \sum_{i=1}^{m} a_i \sum_{j=1}^{n} b_j e_{j\sigma} = \sum_{i=1}^{m} a_i (v_i \sigma) \in \mathcal{C}. \tag{10.25}$$

Since $c \in \mathcal{C}$ was arbitrary, we have $\sigma \in \mathrm{Aut}(\mathcal{C})$. □

CHAPTER 11

LARGE MATHIEU GROUPS

In this chapter, we add 3 points to the 21 points of $P^2(\mathbb{F}_4)$, and embed the lines and hexads of $P^2(\mathbb{F}_4)$ in $\mathbb{F}_2^{24}$. The resulting vectors generate the Golay code $\mathcal{C}_{24}$. We define the Mathieu group M_{24} as the automorphism group of $\mathcal{C}_{24}$. The Mathieu groups M_{23}, M_{22} and M_{21} are defined as the pointwise stabilizers of 1, 2 and 3 points, respectively. We show that M_{24} is 5-transitive, and also show that M_{21} is isomorphic to $PSL_3(\mathbb{F}_4)$. Finally, we apply the multiple transitivity criterion for simplicity, developed in Chapter 3, to conclude that M_{22}, M_{23}, and M_{24} are simple. Our treatment is roughly based on that of Conway and Sloane [CS93].

11.1 Action of $P\Gamma L_3(\mathbb{F}_4)$ on $PSL_3(\mathbb{F}_4)$-Orbits of Hexads

In this section, we add 3 points to the 21 points of $P^2(\mathbb{F}_4)$, such that these 3 points represent the 3 orbits of hexads in $P^2(\mathbb{F}_4)$ under the action of $PSL_3(\mathbb{F}_4)$. We then define an action on these 24 points by $P\Gamma L_3(\mathbb{F}_4)$.

We first repeat Theorems 9.2.7 and 9.4.13, the main results of Chapter 9.

Theorem 11.1.1. *There are* 168 *hexads in* $P^2(\mathbb{F}_4)$ *that split into* 3 *orbits of* 56 *under the action of* $PSL_3(\mathbb{F}_4)$ *such that hexads are in the same orbit if and only if they intersect in an even number of points.* $\square$

Remark 11.1.2. We let $G = P\Gamma L_3(\mathbb{F}_4)$ and $N = PSL_3(\mathbb{F}_4)$ in the remainder of this section.

Lemma 11.1.3. *Let* $h_0 = \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}$ *be a representative hexad for an* N*-orbit, which we designate orbit* I*, and let* $A = \begin{bmatrix} \omega & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$*. Then*

$$h_0' = h_0 A = \{[100], [010], [001], [\omega 11], [1\omega 1], [\bar{\omega}\bar{\omega}1]\},$$

$$h_0'' = h_0 A^2 = \{[100], [010], [001], [\bar{\omega}11], [1\bar{\omega}1], [\omega\omega 1]\}$$

are representative hexads of the other two orbits, which we designate as orbits II *and* III*, respectively.*

Proof. A routine calculation shows $h_0' = h_0 A$ and $h_0'' = h_0' A$. Since $A \in GL_3(\mathbb{F}_4)$, h_0' and h_0'' are also hexads. We observe that each pair of h_0, h_0', h_0'' intersect in the 3 points $\{[100], [010], [001]\}$. By Theorem 11.1.1, they are thus each in one of the three N-orbits. □

Theorem 11.1.4. G *acts on hexad orbits* I*,* II *and* III*.*

Proof. By Theorem 9.4.15, the action of G on $P^2(\mathbb{F}_4)$ preserves the N-orbits. Thus $T \in G$ sends hexad orbits to hexad orbits, and $I \in G$ sends every orbit to itself. If $x \in \{I, II, III\}$ and $T_1, T_2 \in G$, then $(xT_1)T_2 = x(T_1T_2)$ because $x \subseteq P^2(\mathbb{F}_4)$, and G's action on $P^2(\mathbb{F}_4)$ has already been defined. □

Lemma 11.1.5. *We have* N *as a subgroup of the kernel of the action of* G *on hexad orbits* I*,* II *and* III*, inducing an action of* G/N *on the points* $\{I, II, III\}$*.*

Proof. Since I, II and III are N-orbits, $T \in N$ sends each of I, II and III to itself. Now the kernel of the action is the subgroup fixing I, II and III pointwise, and so N is contained in the kernel. Since $N \lhd G$, there is an induced action of G/N on $\{I, II, III\}$. □

Theorem 11.1.6. *$G/N \cong S_3$ in its induced action on $\{I, II, III\}$. Thus N is the kernel of the action of G on points $\{I, II, III\}$.*

Proof. Now

$$|G/N| = |G|/|N| = 120960/20160 = 6. \tag{11.1}$$

Let A be the transformation in Lemma 11.1.3 which sends h_0 in orbit I to h_0' in orbit II, and h_0' in orbit II to h_0'' in orbit III. Now

$$h_0''A = (h_0A^2)A = h_0A^3 = h_0I = h_0, \tag{11.2}$$

and so h_0'' is sent to h_0 by A. Thus $AN \in G/N$ cyclically permutes orbits I, II and III as the permutation $(I\ II\ III)$.

Let $f(\sigma) \in \Gamma L_3(\mathbb{F}_4)$ be the mapping that applies the nontrivial automorphism of $\mathbb{F}_4$ to all three vector coordinates, as in Lemma 5.2.5. Thus $f(\sigma)N \in G/N$, and $f(\sigma)$ is a coset representative. Now

$$[\omega\bar{\omega}1]f(\sigma) = [\omega^\sigma\bar{\omega}^\sigma 1^\sigma] = [\bar{\omega}\omega 1] \tag{11.3}$$

and similarly, $[\bar{\omega}\omega 1]f(\sigma) = [\omega\bar{\omega}1]$. Thus h_0 in orbit I is sent to itself. We have

$$\{[100], [010], [001], [\omega 11], [1\omega 1], [\bar{\omega}\bar{\omega}1]\}^\sigma = \{[100], [010], [001], [\bar{\omega}11], [1\bar{\omega}1], [\omega\omega 1]\}, \tag{11.4}$$

and thus $h_0'f(\sigma) = h_0''$. Similarly, $h_0''f(\sigma) = h_0'$. Thus $f(\sigma)$ fixes orbit I and exchanges orbits II and III, giving the permutation $(II\ III)$.

These two permutations generate a subgroup isomorphic to S_3 on $\{I, II, III\}$. Since G/N has 6 elements, with a homomorphism onto S_3, we conclude $G/N \cong S_3$. It follows that G/N acts faithfully on I, II and III. By Definition 3.2.15, this implies N is the kernel of the action. □

11.2 Golay Code $\mathcal{C}_{24}$

In this section, we consider a certain set X of size 24 as coordinate values in $\mathbb{F}_2^{24}$, allowing an imbedding of the 21 lines and 168 hexads of $P^2(\mathbb{F}_4)$ into $\mathbb{F}_2^{24}$. We then define the Golay code $\mathcal{C}_{24}$ as the subspace generated by these vectors, and show this code is self-dual and doubly even. Thus $\mathcal{C}_{24}$ is of dimension 12, with codewords all of weight divisible by 4.

Remark 11.2.1. The hexad orbits I, II and III will be referred to as *Romans*.

Definition 11.2.2. Let $X = P^2(\mathbb{F}_4) \cup \{I, II, III\}$. Following Conway and Sloane (Ch. 11, Sec. 11 in [CS93]), we choose to have these 24 elements correspond to coordinates in $\mathbb{F}_2^{24}$ in the following way:

$$\begin{bmatrix} [010] & [100] & [001] & [101] & [\omega 01] & [10\omega] \\ I & [110] & [011] & [111] & [\omega 11] & [\bar{\omega} 11] \\ II & [1\omega 0] & [0\omega 1] & [1\omega 1] & [11\bar{\omega}] & [\bar{\omega}\omega 1] \\ III & [\omega 10] & [01\omega] & [1\bar{\omega} 1] & [\omega\bar{\omega} 1] & [11\omega] \end{bmatrix} = \begin{bmatrix} 1 & 5 & 9 & 13 & 17 & 21 \\ 2 & 6 & 10 & 14 & 18 & 22 \\ 3 & 7 & 11 & 15 & 19 & 23 \\ 4 & 8 & 12 & 16 & 20 & 24 \end{bmatrix}. \quad (11.5)$$

We identify subsets of X and vectors of $\mathbb{F}_2^{24}$ as in Definition 10.1.2. Thus $S \subseteq X$ corresponds to vector $v \in \mathbb{F}_2^{24}$ such that if $x_i \in S$ corresponds to the ith coordinate, $1 \leq i \leq 24$, then v has a 1 in the ith coordinate. If $x_i \notin S$, then the ith coordinate of v is 0. Thus X corresponds to $\mathbf{1}$, $\varnothing$ corresponds to 0, and $\{[010], II\}$ corresponds to $e_1 + e_3$.

Definition 11.2.3. We define an action of $P\Gamma L_3(\mathbb{F}_4)$ on X. Let $T \in P\Gamma L_3(\mathbb{F}_4)$ act on the points of $P^2(\mathbb{F}_4)$ in the usual way, defined in Lemma 5.3.1, and let T permute points I, II and III according to the coset of $P\Gamma L_3(\mathbb{F}_4)/PSL_3(\mathbb{F}_4)$ containing T, as in Theorem 11.1.6.

Definition 11.2.4. If ℓ is a line in $P^2(\mathbb{F}_4)$, then $\ell \cup \{I, II, III\} \subseteq X$ is referred to as a *line octad*. Thus the corresponding vector is of weight 8.

Definition 11.2.5. If h is a hexad in $P^2(\mathbb{F}_4)$, we associate h and the Romans in the following way:

- If h is in orbit I, then $h \cup \{II, III\} \subseteq X$.
- If h is in orbit II, then $h \cup \{I, III\} \subseteq X$.
- If h is in orbit III, then $h \cup \{I, II\} \subseteq X$.

Each such 8-element subset is referred to as an *oval octad*.

Definition 11.2.6. We define $\mathcal{C}_{24} \leq \mathbb{F}_2^{24}$ to be the code generated by the 21 line octads and 168 oval octads, and refer to it as the *Golay 24-code*.

Lemma 11.2.7. *The line octads obtained from the following lines of* $P^2(\mathbb{F}_4)$ *(as in Definition 11.2.4):*

$$[001]^\perp, [100]^\perp, [101]^\perp, [10\omega]^\perp, [\omega 01]^\perp, [010]^\perp, [011]\ \perp, [01\omega]^\perp, [110]^\perp, [\omega 10]^\perp \tag{11.6}$$

form a linearly independent set of 10 *vectors. If oval octads obtained from the following hexads of* $P^2(\mathbb{F}_4)$ *(as in Definition 11.2.5):*

$$\begin{aligned} h_0 &= \{[100], [010], [001], [111], [\omega\bar{\omega}1], [\bar{\omega}\omega 1]\}, \\ h_0' &= \{[100], [010], [001], [\omega 11], [1\omega 1], [\bar{\omega}\bar{\omega}1]\}, \end{aligned} \tag{11.7}$$

are added, they form a linearly independent set of 12 *vectors. Thus* $\dim \mathcal{C}_{24} \geq 12$.

Proof. In the coordinates of Definition 11.2.2, we have

$$\begin{aligned}
v_1 &= (1111\ 1111\ 0000\ 0000\ 0000\ 0000) && [001]^\perp \\
v_2 &= (1111\ 0000\ 1111\ 0000\ 0000\ 0000) && [100]^\perp \\
v_3 &= (1111\ 0000\ 0000\ 1111\ 0000\ 0000) && [101]^\perp \\
v_4 &= (1111\ 0000\ 0000\ 0000\ 1111\ 0000) && [10\omega]^\perp \\
v_5 &= (1111\ 0000\ 0000\ 0000\ 0000\ 1111) && [\omega 01]^\perp \\
v_6 &= (0111\ 1000\ 1000\ 1000\ 1000\ 1000) && [010]^\perp \\
v_7 &= (0111\ 1000\ 0100\ 0100\ 0100\ 0100) && [011]^\perp \\
v_8 &= (0111\ 1000\ 0010\ 0010\ 0010\ 0010) && [01\omega]^\perp \\
v_9 &= (0111\ 0100\ 1000\ 0100\ 0010\ 0001) && [110]^\perp \\
v_{10} &= (0111\ 0010\ 1000\ 0010\ 0001\ 0100) && [\omega 10]^\perp \\
v_{11} &= (1011\ 1000\ 1000\ 0100\ 0001\ 0010) && h_0 \\
v_{12} &= (1101\ 1000\ 1000\ 0010\ 0100\ 0001) && h_0'. \qquad (11.8)
\end{aligned}$$

A direct calculation shows linear independence. □

Theorem 11.2.8. *$\mathcal{C}_{24}$ is doubly even.*

Proof. Every generating element of $\mathcal{C}_{24}$ is an octad, and thus the weight of each is divisible by 4. Let v_i and v_j be two such octads. Now $v_i \cdot v_j = 0$ if and only if v_i and v_j intersect in an even number of 1's.

Assume v_i and v_j are line octads. Since lines in $P^2(\mathbb{F}_4)$ intersect in 1 point, v_i and v_j share it, plus the 3 Romans, for 4 points of intersection.

Assume v_i is a line octad and v_j is an oval octad. A line and oval in $P^2(\mathbb{F}_4)$ intersect in either 0 or 2 points, by Theorem 8.5.11, and 2 of the Romans will be shared between these two octads, for 2 or 4 points of intersection.

Assume v_i and v_j are oval octads with their hexads in the same $PSL_3(\mathbb{F}_4)$-orbit. By Theorem 9.4.13 and Corollary 9.2.3, they have 0 or 2 points in common, plus 2 Romans, for 2 or 4 points of intersection.

Assume v_i and v_j are oval octads with their hexads in different $PSL_3(\mathbb{F}_4)$-orbits. By Corollary 9.4.14, the hexads intersect in 1 or 3 points and the octads share only 1 Roman, for 2 or 4 points of intersection.

In each case, $v_i \cdot v_j = 0$, and the generating elements are mutually orthogonal. By Theorem 10.2.15, $\mathcal{C}_{24}$ is doubly even. □

Theorem 11.2.9. *$\mathcal{C}_{24}$ is of dimension 12, and is thus self-dual. Thus the vectors of Lemma 11.2.7 form a basis for $\mathcal{C}_{24}$.*

Proof. Since $\mathcal{C}_{24}$ is self-orthogonal, we have by Lemma 10.2.6 that:

$$\dim \mathcal{C}_{24} \leq n/2 = 24/2 = 12. \tag{11.9}$$

By Lemma 11.2.7, $\dim \mathcal{C}_{24} \geq 12$. Thus $\dim \mathcal{C}_{24} = 12$ and $\dim \mathcal{C}_{24}^{\perp} = 12$, by Theorem 10.2.3. Since $\mathcal{C}_{24} \leq \mathcal{C}_{24}^{\perp}$, we conclude $\mathcal{C}_{24} = \mathcal{C}_{24}^{\perp}$. □

Corollary 11.2.10. *The weight distribution of $\mathcal{C}_{24}$ is symmetric.*

Proof. This follows from Theorems 11.2.9 and 10.2.19. □

11.3 Large Mathieu Groups

In this section, we define the Mathieu group M_{24} to be the automorphism group of $\mathcal{C}_{24}$. We show $P\Gamma L_3(\mathbb{F}_4) \leq M_{24}$ and exhibit an element $\pi \in M_{24}$, where $\pi \notin P\Gamma L_3(\mathbb{F}_4)$. We use π and $P\Gamma L_3(\mathbb{F}_4)$ to establish the 5-transitivity of M_{24}, and then use 5-transitivity to determine the minimum weight of $\mathcal{C}_{24}$, which is 8. The codewords of weight 8, the octads, have the property that any 5 of their 24 points lie in exactly one octad, and thus M_{24} acts transitively on the octads.

Definition 11.3.1. We define the *Mathieu group*, M_{24}, to be the automorphism group of $\mathcal{C}_{24}$.

Theorem 11.3.2. *We have* $P\Gamma L_3(\mathbb{F}_4) \leq M_{24}$.

Proof. By definition, each line octad in $\mathcal{C}$ is obtained from a line in $P^2(\mathbb{F}_4)$, and each oval octad in $\mathcal{C}$ is obtained from a hexad in $P^2(\mathbb{F}_4)$. By Corollary 8.2.16, $P\Gamma L_3(\mathbb{F}_4)$ sends lines to lines. Since ovals are defined in terms of lines, $P\Gamma L_3(\mathbb{F}_4)$ also sends ovals to ovals.

If $T \in P\Gamma L_3(\mathbb{F}_4)$, and $\ell \cup \{I, II, III\}$ is a line octad, then T sends ℓ to ℓ', another line in $P^2(\mathbb{F}_4)$. Now T permutes the 3 Romans by the induced action of Theorem 11.1.6, and thus $\{I, II, III\}T = \{I, II, III\}$. Thus T sends $\ell \cup \{I, II, III\}$ to $\ell' \cup \{I, II, III\}$, and line octads are sent to line octads by the action of $P\Gamma L_3(\mathbb{F}_4)$ on X (or equivalently, by its action on $\mathbb{F}_2^{24}$).

Now Definition 11.2.3 guarantees that the action of $T \in P\Gamma L_3(\mathbb{F}_4)$ on a hexad h matches the action of T on the 2 Romans associated with h. Thus if $hT = h'$, another hexad, then T sends the 2 hexad orbits not containing h to the 2 hexad orbits not containing h'. Thus oval octads are sent to oval octads by the action of $P\Gamma L_3(\mathbb{F}_4)$ on X (or equivalently, by its action on $\mathbb{F}_2^{24}$).

Since generators of the Golay code are sent to generators, this implies that $P\Gamma L_3(\mathbb{F}_4) \leq \mathrm{Aut}(\mathcal{C}_{24})$, by Theorem 10.3.6. □

Lemma 11.3.3. *The coordinate permutation*

$$\pi = (1\ 2)(3\ 4)(9\ 10)(11\ 12)(17\ 19)(18\ 20)(21\ 24)(22\ 23) \tag{11.10}$$

is an element of M_{24}, *and* $\pi \notin P\Gamma L_3(\mathbb{F}_4)$.

Proof. We observe that $v_i\pi = v_i$ for $1 \leq i \leq 5$ and

$$\begin{aligned}
v_6\pi &= (1011\ 1000\ 0100\ 1000\ 0010\ 0001) = v_4 + v_5 + v_6 + v_7 + v_{11} \\
v_7\pi &= (1011\ 1000\ 1000\ 0100\ 0001\ 0010) = v_{11} \\
v_8\pi &= (1011\ 1000\ 0001\ 0010\ 1000\ 0100) = v_2 + v_4 + v_7 + v_8 + v_{11} \\
v_9\pi &= (1011\ 0100\ 0100\ 0100\ 1000\ 1000) = v_4 + v_5 + v_7 + v_9 + v_{11} \\
v_{10}\pi &= (1011\ 0010\ 0100\ 0010\ 0100\ 0010) = v_7 + v_{10} + v_{11} \\
v_{11}\pi &= (0111\ 1000\ 0100\ 0100\ 0100\ 0100) = v_7 \\
v_{12}\pi &= (1110\ 1000\ 0100\ 0010\ 0001\ 1000) = v_5 + v_7 + v_{11} + v_{12}. \qquad (11.11)
\end{aligned}$$

By Theorems 10.3.6 and 11.2.9, $\pi \in \mathrm{Aut}(\mathcal{C}_{24})$.

By Definition 11.2.3, $P\Gamma L_3(\mathbb{F}_4)$ stabilizes $\{I, II, III\}$. However, π exchanges points I and [010]. □

Definition 11.3.4. We define $G_{24} = \langle P\Gamma L_3(\mathbb{F}_4), \pi \rangle$.

Lemma 11.3.5. *We have $G_{24} \leq M_{24}$, and G_{24} acts on X.*

Proof. By Theorem 11.3.2, $P\Gamma L_3(\mathbb{F}_4) \leq M_{24}$, and by Lemma 11.3.3, $\pi \in M_{24}$, and $\langle P\Gamma L_3(\mathbb{F}_4), \pi \rangle$ is the subgroup generated by this subgroup and this element. Now M_{24} acts on X, and so its subgroups do as well. □

Definition 11.3.6. Following Conway and Sloane [CS93], we use the following convention to denote the coordinates of $\mathbb{F}_2^{24}$:

$$\begin{bmatrix}
\infty & 0 & * & * & * & * \\
I & * & * & * & * & * \\
II & * & * & * & * & * \\
III & * & * & * & * & *
\end{bmatrix}. \qquad (11.12)$$

Thus $[010] = \infty$ and $[100] = 0$, and the remainder have the same names.

Theorem 11.3.7. *G_{24} acts 5-transitively on X.*

Proof. Now $PSL_3(\mathbb{F}_4) \lhd P\Gamma L_3(\mathbb{F}_4)$ acts 2-transitively on $P^2(\mathbb{F}_4)$, and fixes the Romans pointwise, by Theorem 4.2.8, Lemma 11.1.5, and Definition 11.2.3.

Also, given $\sigma \in S_3$, there is an $A \in P\Gamma L_3(\mathbb{F}_4) \leq G_{24}$, such that A gives the same permutation of $\{I, II, III\}$ as σ does of $\{1, 2, 3\}$, by Theorem 11.1.6.

Lastly, $\pi \in G_{24}$ transposes ∞ and Roman I, transposes II and III, and permutes the rest of X as in Lemma 11.3.3.

Let a, b, c, x, y be 5 distinct points of X. We claim there exists a $T \in G_{24}$ such that $aT = I$, $bT = II$, $cT = III$, $xT = \infty$, and $yT = 0$. By Lemma 3.3.2, this will give a 5-transitive action of G_{24} on X.

First, if a is a point of $P^2(\mathbb{F}_4)$, use an element of $PSL_3(\mathbb{F}_4)$ to move it to ∞, followed by π to move it to I, followed by an element of $P\Gamma L_3(\mathbb{F}_4)$ to move it from I to II. If a is a Roman, use an element of $P\Gamma L_3(\mathbb{F}_4)$ to move it to II.

Second, if b is a point of $P^2(\mathbb{F}_4)$, use an element of $PSL_3(\mathbb{F}_4)$ to move it to ∞, which fixes a at II, followed by π to move it to I, which also moves a from II to III. Lastly, use an element of $P\Gamma L_3(\mathbb{F}_4)$ that moves b to II, and fixes a at III. If b is a Roman, use an element of $P\Gamma L_3(\mathbb{F}_4)$ to move b to II, and a to III.

Third, if c is a point of $P^2(\mathbb{F}_4)$, use an element of $PSL_3(\mathbb{F}_4)$ to move it to ∞, which fixes a and b at II and III, followed by π to move it to I, which swaps a and b at II and III, and finally an element of $P\Gamma L_3(\mathbb{F}_4)$ to move a, b, c to I, II, III respectively. If c is Roman I, use an element of $P\Gamma L_3(\mathbb{F}_4)$ to move a, b, c to I, II, III, respectively.

The 2-transitivity of $PSL_3(\mathbb{F}_4)$ guarantees an element that fixes the Romans, but moves points x and y to points ∞ and 0, respectively. The composition of the above transformations gives $T \in G_{24}$ which moves a, b, c, x, y to $I, II, III, \infty, 0$. $\square$

Corollary 11.3.8. *M_{24} acts 5-transitively on X.*

Proof. This follows because $G_{24} \leq M_{24}$ acts 5-transitively on X. □

Corollary 11.3.9. *$\mathcal{C}_{24}$ has minimum weight $d = 8$, with no codewords of weight 4 or 20.*

Proof. Now $\mathcal{C}_{24}$ is doubly even, by Theorem 11.2.8, which means every codeword has weight divisible by 4. Thus possible codeword weights are $0, 4, 8, 12, 16, 20, 24$. Assume $x \in \mathcal{C}_{24}$ is of weight 4. Since M_{24} acts 5-transitively on X, by Theorem 11.3.7, there is an element in $\mathrm{Aut}(\mathcal{C}_{24})$ fixing three of the 1's in x and sending the fourth to another coordinate, to yield $x' \in \mathcal{C}_{24}$, also of weight 4. Since x and x' intersect in three 1's, $x + x'$ has weight 2. This contradiction shows there are no codewords of weight 4. By $\mathcal{C}$'s symmetric weight distribution, from Corollary 11.2.10, there are no codewords of weight 20. Finally, our generators are of weight 8, and thus $d = 8$. □

Definition 11.3.10. We refer to any weight-8 codeword of $\mathcal{C}_{24}$ as an *octad*, which can also be viewed as an 8-point subset of X.

Corollary 11.3.11. *Any 5 points of X lie in exactly 1 octad of $\mathcal{C}_{24}$.*

Proof. The points I, II, III, $\infty, 0$ of Theorem 11.3.7 are in the line octad obtained from $L_\infty \subseteq P^2(\mathbb{F}_4)$. By the 5-transitivity of M_{24}, any 5 points of X lie in at least one octad of $\mathcal{C}_{24}$. Assume the same 5 points are in 2 distinct octads, O_1 and O_2. Now $|O_1 \cap O_2| \geq 5$ and so $wt(O_1 + O_2) \leq 6$. But $O_1 + O_2 \in \mathcal{C}_{24}$, which has minimum weight 8. This contradiction shows any 5-point subset is in at most one octad. □

Corollary 11.3.12. *G_{24} and M_{24} act transitively on the octads of $\mathcal{C}_{24}$.*

Proof. Let O_1, O_2 be octads of $\mathcal{C}_{24}$, and choose 5-point subsets P_1, P_2 of each. Now P_1 lies only in O_1 and similarly, P_2 lies only in O_2. Now there is an element of $G_{24} \leq M_{24}$ that sends P_1 to P_2, and thus O_1 is sent to O_2. □

11.4 Steiner System of Octads

Given a 3-point subset of X, it turns out that this subset is in exactly 21 of the 759 octads of $\mathcal{C}_{24}$, an important fact for establishing $PSL_3(\mathbb{F}_4)$ as the pointwise stabilizer (in M_{24}) of 3 points of X. In this section, we define Steiner systems, and show that the octads of $\mathcal{C}_{24}$ form a Steiner system $S(5, 8, 24)$. Finally, we formulate an alternate definition of M_{24} as the automorphism group of this $S(5, 8, 24)$.

Definition 11.4.1. Let X be a set such that $|X| = v$, and let subsets of X containing k elements be denoted as *k-subsets*. An $S(t, k, v)$ *Steiner system* is a collection of distinct k-subsets of X (called *blocks*) with the property that any t-subset of X is contained in exactly 1 block.

Corollary 11.4.2. *The octads of $\mathcal{C}_{24}$ form a Steiner system $S(5, 8, 24)$.*

Proof. This follows from Corollary 11.3.11. □

Corollary 11.4.3. *There are 759 octads in $\mathcal{C}_{24}$.*

Proof. Now X has $\binom{24}{5}$ 5-subsets. Each octad contains $\binom{8}{5}$ 5-subsets, each lying in only that octad (by Corollary 11.3.11). Thus if λ is the number of octads in $\mathcal{C}_{24}$, the number of 5-subsets of X is equal to the product of λ with the number of 5-subsets per octad. Thus $\binom{24}{5} = \lambda\binom{8}{5}$, and there are

$$\lambda = \binom{24}{5} \Big/ \binom{8}{5} = 42504/56 = 759 \tag{11.13}$$

octads in $\mathcal{C}_{24}$. □

Corollary 11.4.4. *The weight distribution of $\mathcal{C}_{24}$ is*

$$0^1 \;\; 8^{759} \;\; 12^{2576} \;\; 16^{759} \;\; 24^1. \tag{11.14}$$

Proof. By Theorem 11.2.8, $\mathcal{C}_{24}$ is doubly even, and thus by definition, all codewords have weight divisible by 4. By Corollary 11.2.10, $\mathcal{C}_{24}$ has a symmetric weight distribution. We have one 0 vector and one $\mathbf{1}$. There are 759 octads by Corollary 11.4.3, and thus 759 codewords of weight 16. By Corollary 11.3.9, $\mathcal{C}_{24}$ has no codewords of weight 4 or of weight 16. There are a total of $2^{12} = 4096$ codewords, and thus

$$4096 - 2 - 2(759) = 2576 \tag{11.15}$$

codewords remain, which must be of weight 12. □

Lemma 11.4.5. *Let $\{x_1, \ldots, x_i\}$ be an i-subset of X contained in some octad (for $i < 5$ there is more than one such octad). Let λ_i be the number of octads containing this i-subset. Then for $0 \leq i \leq 5$,*

$$\lambda_i = \binom{24-i}{5-i} \Big/ \binom{8-i}{5-i}, \tag{11.16}$$

and $\lambda_i = 1$ for $6 \leq i \leq 8$.

Proof. Every octad has $\varnothing$ as a subset, and if $i = 0$, our result agrees with Corollary 11.4.3. If $1 \leq i \leq 5$, we consider the 5-subsets containing $\{x_1, \ldots, x_i\}$. Each are completed by choosing $5 - i$ of $24 - i$ points, for $\binom{24-i}{5-i}$ such. Each also determines an octad containing $\{x_1, \ldots, x_i\}$, with $\binom{8-i}{5-i}$ such 5-subsets for each octad containing $\{x_1, \ldots, x_i\}$. Thus

$$\binom{24-i}{5-i} = \lambda_i \binom{8-i}{5-i}, \tag{11.17}$$

and our result follows. Each 5-subset determines a unique octad, and thus $\lambda_i = 1$ for $6 \leq i \leq 8$. □

Corollary 11.4.6. *Let* $\{x_1, x_2, x_3\} \subseteq X$. *There are exactly* 21 *octads containing* $\{x_1, x_2, x_3\}$. *There are also exactly* 77 *octads which contain* $\{x_1, x_2\}$, *and thus* 56 *octads which contain* $\{x_1, x_2\}$, *but not* x_3.

Proof. From Lemma 11.4.5, we have

$$\lambda_2 = \binom{24-2}{5-2} \Big/ \binom{8-2}{5-2} = \binom{22}{3} \Big/ \binom{6}{3} = 1540/20 = 77, \tag{11.18}$$

and

$$\lambda_3 = \binom{24-3}{5-3} \Big/ \binom{8-3}{5-3} = \binom{21}{2} \Big/ \binom{5}{2} = 210/10 = 21. \tag{11.19}$$

Thus there are $77 - 21 = 56$ octads that contain $\{x_1, x_2\}$, but not x_3. □

Definition 11.4.7. If $\sigma \in S_{24}$ is a permutation sending octads of $S(5, 8, 24)$ to octads, then σ is said to be an *automorphism* of $S(5, 8, 24)$. The collection of such automorphisms is the *automorphism group* of the Steiner system, $\mathrm{Aut}(S(5, 8, 24))$.

Theorem 11.4.8. *We have* $M_{24} = \mathrm{Aut}(S(5, 8, 24))$.

Proof. From Definition 11.3.1, $M_{24} = \mathrm{Aut}(\mathcal{C}_{24})$. Let $T \in \mathrm{Aut}(\mathcal{C}_{24})$. Thus T sends codewords to codewords, preserving their weight. Thus octads are sent to octads, and $T \in \mathrm{Aut}(S(5, 8, 24))$. Thus $M_{24} \leq \mathrm{Aut}(S(5, 8, 24))$.

Conversely, let $T \in \mathrm{Aut}(S(5, 8, 24))$. Thus T sends octads of $\mathcal{C}_{24}$ to octads. Now $\mathcal{C}_{24}$ is generated by the 21 line octads and the 168 oval octads, by Definition 11.2.6. Thus T sends the generators of $\mathcal{C}_{24}$ to other codewords of $\mathcal{C}_{24}$. By Theorem 10.3.6, this implies $T \in \mathrm{Aut}(\mathcal{C}_{24})$. This gives $\mathrm{Aut}(S(5, 8, 24)) \leq M_{24}$, and equality follows. □

11.5 Simplicity of the Large Mathieu Groups

In this section, we first prove $P\Gamma L_3(\mathbb{F}_4)$ is the collineation group of $P^2(\mathbb{F}_4)$. We then establish that $P\Gamma L_3(\mathbb{F}_4)$ is the setwise stabilizer of any 3 points of X, using

the main result from Section 11.4. Then we show that $PSL_3(\mathbb{F}_4)$ is the pointwise stabilizer of any 3 points, using a result from Section 11.1, and thus identify $PSL_3(\mathbb{F}_4)$ with M_{21}. Finally, we calculate the orders of M_{22}, M_{23} and M_{24}, and use Theorem 3.5.11 to prove their simplicity.

Definition 11.5.1. Let P be a projective space. A *collineation* is a bijection $P \to P$ sending collinear points to collinear points. The set of all such bijections is the *collineation group* of P.

In general, for $n \geq 3$, $P\Gamma L_n(E)$ is the collineation group of $P^{n-1}(E)$ (Gruenberg and Weir [GW77]). We prove this for the special case of $n = 3$ and $E = \mathbb{F}_4$.

Theorem 11.5.2. *$P\Gamma L_3(\mathbb{F}_4)$ is the collineation group of $P^2(\mathbb{F}_4)$.*

Proof. Let κ be a collineation of $P^2(\mathbb{F}_4)$ and let q_0 be the ordered tetrad:

$$q_o = \{p_1 = [100], p_2 = [010], p_3 = [001], p_4 = [111]\}. \tag{11.20}$$

Since κ preserves lines, $q_0\kappa = q_1$ is another ordered tetrad. By Theorem 9.1.7, $PGL_3(\mathbb{F}_4)$ acts sharply transitively on ordered tetrads, and there is an $A \in PGL_3(\mathbb{F}_4)$ such that

$$q_0 = q_1 A = (q_0\kappa)A = q_0(\kappa A), \tag{11.21}$$

where the ordering of q_0 is preserved. Thus κA fixes q_0 pointwise, implying every join of q_0 is fixed setwise. By Theorem 9.3.6, there are exactly 2 points in $P^2(\mathbb{F}_4)$ not on any of the 6 joins of q_0, namely $p_5 = [\omega\bar{\omega}1]$ and $p_6 = [\bar{\omega}\omega 1]$, conjugates under $\mathbb{F}_4$'s nontrivial automorphism σ. Since p_5 and p_6 are the two points not on any join of q_0, they remain such under κA. Thus κA either fixes or switches them. If κA

switches them, then $\kappa Af(\sigma)$ fixes them, as the two points are conjugate. Thus one of κA, $\kappa Af(\sigma)$ fixes h_0 pointwise. Call this transformation T.

By Corollary 9.3.11, each of the 15 points in $P^2(\mathbb{F}_4) \setminus h_0$ is the point of intersection for exactly 3 opposite joins of h_0. Now 2 lines determine a unique point, and thus each point not in h_0 is uniquely determined by 2 of the joins of h_0. Since T fixes all the joins of h_0, points of intersection are sent to themselves. Thus the 15 points of $P^2(\mathbb{F}_4) \setminus h_0$ are fixed by T, and the entire plane is fixed by T.

By Definition 11.5.1, T is the identity transformation of the collineation group. But then κ and one of A or $Af(\sigma)$ are inverses, implying $\kappa \in P\Gamma L_3(\mathbb{F}_4)$. Since $P\Gamma L_3(\mathbb{F}_4)$ preserves lines by Corollary 8.2.16, the reverse inclusion follows, and $P\Gamma L_3(\mathbb{F}_4)$ is the collineation group of $P^2(\mathbb{F}_4)$. □

Theorem 11.5.3. *In $G_{24} = \langle P\Gamma L_3(\mathbb{F}_4), \pi \rangle$ and M_{24}, the setwise stabilizer of 3 points is isomorphic to $P\Gamma L_3(\mathbb{F}_4)$.*

Proof. By Corollary 11.4.6, there are 21 octads having $\{I, II, III\}$ as a subset, which by Definition 11.2.4, are precisely the 21 line octads. Now M_{24} preserves codeword weight in $\mathcal{C}_{24}$, and thus G_{24} does as well.

$\mathrm{Stab}_{M_{24}}(\{I, II, III\})$ permutes $P^2(\mathbb{F}_4)$, sending the set of 21 line octads to itself, a collineation of $P^2(\mathbb{F}_4)$. By Theorem 11.5.2, $P\Gamma L_3(\mathbb{F}_4)$ is the collineation group of $P^2(\mathbb{F}_4)$, and so $\mathrm{Stab}_{M_{24}}(\{I, II, III\}) \leq P\Gamma L_3(\mathbb{F}_4)$. Conversely, by Theorem 11.1.6 and Definition 11.2.3, $P\Gamma L_3(\mathbb{F}_4) \leq \mathrm{Stab}_{M_{24}}(\{I, II, III\})$. Thus

$$\mathrm{Stab}_{M_{24}}(\{I, II, III\}) = P\Gamma L_3(\mathbb{F}_4) \leq G_{24}. \tag{11.22}$$

This gives

$$\mathrm{Stab}_{M_{24}}(\{I, II, III\}) = P\Gamma L_3(\mathbb{F}_4) = \mathrm{Stab}_{G_{24}}(\{I, II, III\}). \tag{11.23}$$

Since G_{24} and M_{24} are 3-transitive, the setwise stabilizer of any 3 points is isomorphic to $P\Gamma L_3(\mathbb{F}_4)$. □

Corollary 11.5.4. *In G_{24} and M_{24}, the pointwise stabilizer of 3 points is isomorphic to $PSL_3(\mathbb{F}_4)$.*

Proof. Now the the pointwise stabilizer is a subgroup of the setwise stabilizer, and by Theorem 11.1.6, $PSL_3(\mathbb{F}_4) \lhd P\Gamma L_3(\mathbb{F}_4)$ is the pointwise stabilizer of I, II, and III. Since G_{24} and M_{24} are 3-transitive, the pointwise stabilizer of any 3 points is isomorphic to $PSL_3(\mathbb{F}_4)$. □

Theorem 11.5.5. *G_{24} and M_{24} are both groups of order* 244823040*. Therefore, $M_{24} = G_{24} = \langle P\Gamma L_3(\mathbb{F}_4), \pi \rangle$.*

Proof. Since G_{24} acts 3-transitively on 24 points, and the pointwise stabilizer of 3 points is $PSL_3(\mathbb{F}_4)$, which is of order 20160, then by Theorem 3.3.4, we have

$$\begin{aligned} |G_{24}| &= (24)(24-1)(24-2)|\operatorname{Stab}(x_1, x_2, x_3)| \\ &= (24)(23)(22)(20160) \\ &= 244823040. \end{aligned} \tag{11.24}$$

The order of M_{24} is identical, by similar reasoning. Since $G_{24} \leq M_{24}$, we have $G_{24} = M_{24}$. □

Definition 11.5.6. Let $M_{24-i} = \operatorname{Stab}(x_1, \ldots, x_i)$, the pointwise stabilizer in M_{24} of i points, for $0 \leq i \leq 4$. Since M_{24} is 5-transitive, we may choose any i points for these stabilizers.

Theorem 11.5.7. *M_{24-i} is a $(5-i)$-transitive group on $24-i$ points, for*

$0 \leq i \leq 4$, *and*

$$\begin{aligned}
|M_{24}| &= 244823040 = (24)(23)(22)(21)(20)(16)(3), \\
|M_{23}| &= 10200960 = (23)(22)(21)(20)(16)(3), \\
|M_{22}| &= 443520 = (22)(21)(20)(16)(3), \\
|M_{21}| &= 20160 = (21)(20)(16)(3), \\
|M_{20}| &= 960 = (20)(16)(3). \qquad (11.25)
\end{aligned}$$

Proof. Since M_{24} acts 5-transitively on X, Lemma 3.3.3 implies that, for $x_1 \in X$, $\mathrm{Stab}(x_1) = M_{23}$ acts 4-transitively on $X - \{x_1\}$, which has 23 points. Now for $x_2 \in X - \{x_1\}$,

$$\mathrm{Stab}_{\mathrm{Stab}(x_1)}(x_2) = \mathrm{Stab}(x_1) \cap \mathrm{Stab}(x_2) = \mathrm{Stab}(x_1, x_2) = M_{22}, \qquad (11.26)$$

and so M_{22} acts 3-transitively on $X - \{x_1, x_2\}$, again by Lemma 3.3.3. We continue in this way through M_{20}. Now Theorem 11.5.5 gives $|M_{24}|$, and Theorem 3.3.4 gives the orders of the pointwise stabilizers. □

Theorem 11.5.8. *M_{22}, M_{23}, and M_{24} are simple groups.*

Proof. For $0 \leq i \leq 3$, M_{24-i} is defined to be a group of permutations that fixes i elements of X pointwise. Thus M_{24-i} acts faithfully on the remaining $24 - i$ points.

By Corollary 11.5.4, $M_{21} \cong PSL_3(\mathbb{F}_4)$, and thus by Theorem 4.3.2, M_{21} is simple.

Now $M_{21} \leq M_{22}$, and M_{21} is the stabilizer of 1 of the 22 points acted on by M_{22}. Since M_{22} acts 3-transitively on 22 points, and 22 is not a prime power, the conditions of Theorem 3.5.11 are satisfied, and M_{22} is simple.

Similarly, $M_{22} \leq M_{23}$, and M_{22} is the stabilizer of 1 of the 23 points acted on by M_{23}. Because M_{22} is simple, and since M_{23} acts 4-transitively on these 23 points, the conditions of Theorem 3.5.11 are satisfied, and M_{23} is simple.

Similarly, $M_{23} \leq M_{24}$, and M_{23} is the stabilizer of 1 of the 24 points acted on by M_{23}. Because M_{23} is simple, and since M_{24} acts 5-transitively on these 24 points, the conditions of Theorem 3.5.11 are satisfied, and M_{24} is simple. $\square$

BIBLIOGRAPHY

[Asc04] M. Aschbacher, *The status of the classification of the finite simple groups*, Notices of the AMS **51** (2004), no. 7, 736–740.

[CS93] J. Conway and N. Sloane, *Sphere packings, lattices and groups*, 2nd ed., Springer-Verlag, 1993.

[Edg65] W. Edge, *Some implications of the geometry of the 21-point plane*, Math. Zeitschr. **87** (1965).

[FIS03] S. Friedberg, A. Insel, and L. Spence, *Linear algebra*, 4th ed., Prentice-Hall, 2003.

[Fra03] J. Fraleigh, *A first course in abstract algebra*, 7th ed., Addison-Wesley, 2003.

[FT63] W. Feit and J. Thompson, *Solvability of groups of odd order*, Pacific Journal of Mathematics **13** (1963).

[Ful08] W. Fulton, *Algebraic curves: an introduction to algebraic geometry*, http://www.math.lsa.umich.edu/ wfulton/CurveBook.pdf, 2008.

[Gal46] E. Galois, *Lettre de Galois a M. Auguste Chevalier*, J. Math. Pures Appl. **11** (1846).

[Gal02] J. Gallian, *Contemporary abstract algebra*, 5th ed., Houghton-Mifflin, 2002.

[Gri82] R. Griess, *The friendly giant*, Invent. Math. **69** (1982).

[Gro02] L. Grove, *Classical groups and geometric algebra*, American Mathematical Society, 2002.

[GW77] K. Gruenberg and A. Weir, *Linear geometry*, 2nd ed., Springer-Verlag, 1977.

[Hir79] J. Hirschfeld, *Projective geometries over finite fields*, Clarendon, 1979.

[HP85] D. Hughes and F. Piper, *Design theory*, Cambridge, 1985.

[Jac85] N. Jacobson, *Basic algebra 1*, 2nd ed., Dover, 1985.

[Jan66] Z. Janko, *A new finite simple group with abelian 2-Sylow subgroups and its characterization*, J. Alg. **3** (1966).

[Jor70] C. Jordan, *Traite des substitutions et des equations algebriques*, Gauthier-Villars, 1870.

[Mat61] E. Mathieu, *Memoire sur l'etude des fonctions de plusieurs quantites*, J. Math. Pures Appl. **6** (1861).

[Mat73] ———, *Sur la fonction cinq fois transitive de 24 quantites*, J. Math. Pures Appl. **18** (1873).

[Maz] V. Mazurov, *Simple finite group*, Encylopaedia of mathematics, http://eom.springer.de/s/s085210.htm. Viewed March 30, 2011.

[Ple89] V. Pless, *Introduction to the theory of error-correcting codes*, 2nd ed., Wiley-Interscience, 1989.

[Ron06] M. Ronan, *Symmetry and the monster*, 4th ed., Oxford, 2006.

[Rot95] J. Rotman, *An introduction to the theory of groups*, Springer, 1995.

[Tit64] J. Tits, *Sur les systemes de Steiner associes aux trois "grands" groupes de Mathieu*, Rend. Math. e Appl. **23** (1964).

[Wit38a] E. Witt, *Die 5-fach transitiven gruppen von Mathieu*, Abhand. Math. Sem. Univ. Hamb. **12** (1938).

[Wit38b] ———, *Uber Steinersche systeme*, Abhand. Math. Sem. Univ. Hamb. **12** (1938).