Classical Groups and Self-Dual Binary Codes

by

Patrick H. King

A thesis submitted to the Department of Mathematics and Statistics in partial fulfillment of the requirements for the degree of Bachelor of Science (Joint Honours) in Computer Science and Pure Mathematics.

Memorial University of Newfoundland
St. John's, Newfoundland and Labrador, Canada
November 2024

Abstract

Suppose that V is a symplectic space, that is, a finite-dimensional vector space endowed with a nondegenerate alternating bilinear form. A subspace L of V is said to be Lagrangian if L coincides with its orthogonal complement. This thesis aims to construct a simple algorithm to compute the Lagrangians of $\mathbb{F}_2^{2n}$ as a vector space over the field $\mathbb{F}_2$ up to a permutation of coordinates. There will first, however, need to be a discussion of the classical linear groups to achieve such a goal. In particular, we will include a discussion of the symplectic groups.

Acknowledgements

First and foremost, I would like to sincerely thank my supervisor Dr. M. Kotchetov for agreeing to supervise me on this project. His advice and guidance have proven to be invaluable to the completion of this thesis. I would also like to thank all of my friends and family for their continuous support and encouragement whilst I was preparing the contents of this thesis. Last but not least, I would like to thank all of the staff and Faculty in the Department of Mathematics and Statistics at Memorial University for providing me with a nurturing environment to pursue my studies. I am truly grateful to those who have offered me aid of any kind on this journey.

Contents

Chapter 1

Introduction

A vector space V over a field $\mathbb{F}$ equipped with a nondegenerate alternating form B is called a symplectic space and, loosely speaking, symplectic geometry is devoted to the study of (V, B) [6, p. 21]. For a given symplectic space (V, B), define a Lagrangian subspace L of V to be a vector subspace of V such that $L^\perp = L$ with respect to B. The term "Lagrangian subspace" was supposedly first introduced by a mathematical physicist named Viktor Maslov in 1965 and the related notion of a Lagrangian submanifold has since proven to be invaluable to the study of virtually every physical system [1, p. 402–403]. In this thesis, we propose a simple algorithm for computing the Lagrangian subspaces of $V = \mathbb{F}_2^{2n}$ as a vector space over $\mathbb{F}_2$ up to a permutation of coordinates. We first, however, need to establish some theory concerning the classical linear groups. The most fundamental of which is the general linear group. We first study abstract bilinear forms and then specialize to alternating and symmetric forms, the theory of which will be needed in the discussion of their relevant isometry groups. Unless otherwise stated, we assume for this chapter that V is an n-dimensional vector space over a field $\mathbb{F}$.

1.1 Bilinear Forms

We now begin with our discussion of bilinear forms, following the relevant theory in chapters 2 and 4 in [6].

Definition 1.1. *A function $B : V \times V \to \mathbb{F}$ is a* bilinear form *on V if for any $w \in V$, the maps*

$$\begin{aligned} L_w &: V \to \mathbb{F},\ v \mapsto B(w, v), \\ R_w &: V \to \mathbb{F},\ v \mapsto B(v, w) \end{aligned}$$

are $\mathbb{F}$-linear.

If B is a bilinear form on V and $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis for V, then the matrix $\widehat{B}$ given by $\widehat{B}_{ij} = B(v_i, v_j)$ for all i, j is called the matrix of B relative to $\mathcal{B}$. Note that if $u, v \in V$ with $u = \sum\limits_i x_i v_i$ and $v = \sum\limits_i y_i v_i$, then

$$B(u, v) = B\left(\sum_i x_i v_i, \sum_j y_i v_i\right) = \sum_{i,j} x_i B(v_i, v_j) y_i = \hat{u}^T \widehat{B} \hat{v} \qquad (*)$$

where $\hat{u}$ and $\hat{v}$ are the column vectors of u and v relative to $\mathcal{B}$.

Conversely, suppose $\widehat{B} = [b_{ij}]$ is an $n \times n$ matrix over $\mathbb{F}$ and $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis for V. Then $\widehat{B}$ determines a bilinear form B on V by first defining B on the basis $\mathcal{B}$ via $B(v_i, v_j) = b_{ij}$ for all i, j and then extending B to a map $V \times V \to \mathbb{F}$ by linearity, as in equation $(*)$.

Suppose now that $\mathcal{B} = \{v_1, \ldots, v_n\}$ and $\mathcal{B}' = \{u_1, \ldots u_n\}$ are two bases for V. For each j, write $u_j = \sum\limits_i a_{ij} v_i$. We then have that for a bilinear form B on V with corresponding matrix $\widehat{B}_{\mathcal{B}}$ relative to $\mathcal{B}$,

$$B(u_i, u_j) = \sum_{k,l} a_{ki} B(v_k, v_l) a_{lj} = (A^T \widehat{B} A)_{ij}$$

where $A = [a_{ij}]$. Thus, $\widehat{B}_{\mathcal{B}'} = A^T \widehat{B}_{\mathcal{B}} A$. Note that A is invertible since its columns represent a basis for V. Thus, any two representing matrices of a bilinear form are equivalent under the equivalence relation $\sim$ on $n \times n$ matrices over $\mathbb{F}$ given by

$$X \sim Y \iff X = M^T Y M \text{ for some } M \in GL(n, \mathbb{F}).$$

For a bilinear form B on V, define the left and right radicals of V under B respectively as

$$\begin{aligned} \mathrm{rad}_L(V) &= \{v \in V \mid \text{for all } w \in V,\ B(v, w) = 0\}, \\ \mathrm{rad}_R(V) &= \{v \in V \mid \text{for all } w \in V,\ B(w, v) = 0\}. \end{aligned}$$

We say that B is *nondegenerate* if $\mathrm{rad}_L(V) = \{0\}$, or equivalently $\mathrm{rad}_R(V) = \{0\}$ thanks to the following proposition.

Denote by $\mathbb{F}^\times$ the *multiplicative group of* $\mathbb{F}$ consisting of all non-zero elements of $\mathbb{F}$.

Proposition 1.1. *Let B be a bilinear form on V with a representing matrix $\widehat{B}$. Then B is nondegenerate if and only if $\det \widehat{B} \in \mathbb{F}^\times$.*

Proof. Choose a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ for V and let $\widehat{B}$ be the matrix corresponding to B relative to $\mathcal{B}$. Let $v \in \mathrm{rad}_L(V)$ and write $v = \sum\limits_j \lambda_j v_j$. Then for all $1 \leq i \leq n$,

$$B(v, v_i) = B\left(\sum_j \lambda_j v_j, v_i\right) = \sum_j \lambda_j B(v_j, v_i) = 0.$$

Hence, the row vector $\hat{v}^T$ satisfies $\hat{v}^T \widehat{B} = 0$ and thus B is nondegenerate $\iff$ $\det \widehat{B} \neq 0$. □

In general, for any subset $S \subset V$ and any bilinear form B on V define

$$\begin{aligned} \perp_L (S) &= \{v \in V \mid \text{for all } w \in S, B(v, w) = 0\}, \\ \perp_R (S) &= \{v \in V \mid \text{for all } w \in S, B(w, v) = 0\}. \end{aligned}$$

It is evident that $\perp_L (S)$ and $\perp_R (S)$ are subspaces of V and in the case when $S = V$, we recover the definition of, respectively, the left and right radicals of V.

Proposition 1.2. *If B is a nondegenerate bilinear form on V and $W \subset V$ is a subspace, then*

$$\dim \perp_L (W) = \dim V - \dim W.$$

Proof. Denote by V^* the set of $\mathbb{F}$-linear maps $V \to \mathbb{F}$. Then V^* is an $\mathbb{F}$-vector space (called the *dual space of* V) under pointwise addition and scalar multiplication given by

$$(f+g)(v) = f(v) + g(v) \text{ and } (\lambda f)(v) = \lambda(f(v)), \text{ for all } v \in V \text{ and } \lambda \in \mathbb{F}.$$

Fix a basis $\{v_1, \ldots, v_m\}$ of W and extend it to a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V. For each $1 \leq i \leq n$, define $v_i^* \in V^*$ by

$$v_i^*(v_j) = \delta_{ij},$$

extending v_i^* by linearity. Then, $\{v_i^*\}_{i=1}^n$ is a basis for V^* (called the *dual basis* of $\mathcal{B}$) and in particular, dim $V =$ dim V^*.

Now, define $f_i \in V^*$ by $f_i(v) = B(v_i, v)$ for all $v \in W$. We claim that $\{f_1, \ldots, f_n\}$ is a basis for V^*. Since dim $V^* = n$, it suffices to show that $f_1, \ldots, f_n$ are linearly independent. Indeed, if there are $\lambda_1, \ldots, \lambda_n \in \mathbb{F}$ such that $\sum\limits_{i=1}^n \lambda_i f_i = 0$, then

$$\begin{aligned} B\left(\sum_{i=1}^n \lambda_i v_i,\, v\right) &= \sum_{i=1}^n \lambda_i f_i(v) = 0 \text{ for all } v \in V \\ &\implies \sum_{i=1}^n \lambda_i v_i = 0 \text{ by nondegeneracy.} \end{aligned}$$

Hence, $\lambda_i = 0$ for all $1 \leq i \leq n$ and so $\{f_i\}_{i=1}^n$ is a basis for V^*. Note then that for any $f \in V^*$, there exists $v \in V$ such that $f(u) = B(v, u)$ for all $u \in V$. Consider now the map

$$\psi : V \to W^*,\ v \mapsto L_v|_W,$$

where $L_v(u) = B(v, u)$ for all $u \in V$. For each $f \in W^*$, extend f to a map $f' : V \to \mathbb{F}$ by setting $f'(v_i) = 0$ for $i > m$, extending by linearity. Then, $f \in V^*$ and so there exists $v \in V$ such that for any $u \in V$, $f'(u) = B(v, u)$ and in particular, $f = L_v|_W$. Therefore ψ is surjective. It is evident that ψ is linear and by definition, ker $\psi =\perp_L (W)$. Hence, $W^* \cong V/ \perp_L (W)$ and in particular,

$$\text{dim } V = \text{dim } W + \text{dim } \perp_L (W).$$

□

Remark 1.1. *We may replace* $\perp_L (W)$ *with* $\perp_R (W)$ *in Proposition 1.2. The argument is essentially the same as in the case of* $\perp_L (W)$.

Corollary 1.3. *If* B *is a nondegenerate bilinear form on* V *and* $W \subset V$ *is a subspace, then*

$$\perp_R \big(\perp_L (W)\big) = W =\perp_L \big(\perp_R (W)\big).$$

Proof. By definition, $W \subset\perp_R \big(\perp_L (W)\big)$ and by Proposition 1.2,

$$\begin{aligned} \text{dim } \perp_R \big(\perp_L (W)\big) &= \text{dim } V - \text{dim } \perp_L (W) \\ &= \text{dim } V - (\text{dim } V - \text{dim } W) \\ &= \text{dim } W. \end{aligned}$$

Similarly, $W =\perp_L \big(\perp_R (W)\big)$. □

Definition 1.2. *Let V_1 and V_2 be $\mathbb{F}$-vector spaces each equipped with a bilinear form, say B_1 on V_1 and B_2 on V_2. An $\mathbb{F}$-isomorphism $\tau : V_1 \to V_2$ is an* isometry *if for all $u, v \in V_1$,*

$$B_1(u, v) = B_2(\tau u, \tau v).$$

If such an isometry exists, $(V_1,\ B_1)$ and $(V_2,\ B_2)$ are said to be isometric. In this case, we may also say that V_1 and V_2 are isometric *provided there is no confusion about the associated forms.*

Proposition 1.4. *With notation as in Definition 1.2, $(V_1,\ B_1)$ and $(V_2,\ B_2)$ are isometric if and only if there are bases for V_1 and V_2 relative to which $\widehat{B}_1 = \widehat{B}_2$.*

Proof. Let $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ be a basis for V_1 and suppose $\tau : V_1 \to V_2$ is an isometry. Since τ is an isomorphism, $\mathcal{B}_2 = \{\tau v_1, \ldots, \tau v_n\}$ is a basis for V_2. We then have that the matrix $\widehat{B}_1$ relative to $\mathcal{B}_1$ is equal to the matrix $\widehat{B}_2$ relative to $\mathcal{B}_2$. Conversely, suppose there are bases $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_n\}$ of V_1 and V_2 respectively, relative to which $\widehat{B}_1 = \widehat{B}_2$. Define

$$\tau : V_1 \to V_2,\ v_i \mapsto w_i$$

for each $1 \leq i \leq n$, extending by linearity. Then τ is an isomorphism. For $u, v \in V_1$, write $u = \sum\limits_i x_i v_i$ and $v = \sum\limits_j y_i w_i$. Then

$$B_1(u, v) = B_1\left(\sum_i x_i v_i, \sum_j y_i w_i\right) = \sum_{i,j} x_i y_j B_1(v_i, v_j).$$

But $B_1(v_i, v_j) = B_2(w_i, w_j)$ so we have

$$\begin{aligned} B_1(u, v) &= \sum_{i,j} x_i y_j B_2(w_i, w_j) \\ &= \sum_{i,j} x_i y_j B_2(\tau v_i, \tau v_j) \\ &= B_2(\tau u, \tau v). \end{aligned}$$

Thus, τ is an isometry. □

If $u, v \in V$ satisfy $B(u, v) = 0$ for a bilinear form B on V, we say that u is *orthogonal* to v and write $u \perp v$. If $u \perp v \implies v \perp u$, then B is said to be *reflexive*. Note that in this case, $\perp_L (W) =\perp_R (W)$ for any $W \subset V$ and so we instead write $W^\perp$. Furthermore, we say subspaces U and W are orthogonal if for all $u \in U$ and all $w \in W$, $B(u, w) = 0$. If U and W intersect trivially, denote their direct sum by $U \oplus W$. If W is a subspace, then $W^\perp$ is called the *orthogonal complement* of W. The restriction $B|_{W \times W}$ is a nondegenerate form on W if and only if $W \cap W^\perp = \{0\}$. We define the *radical* of W to be rad $W = W \cap W^\perp$, that is, the radical of $(W,\ B|_{W \times W})$, and say that W is a *nondegenerate subspace* of V relative to B if rad $W = \{0\}$.

Proposition 1.5. *Suppose B is a reflexive bilinear form on V and $W \subset V$ is a nondegenerate subspace relative to B. Then V is the (internal) direct sum*

$$V = W \oplus W^\perp.$$

Proof. Let $v \in V$ and define $L_v : V \to \mathbb{F}$ by $L_v(u) = B(v, u)$. Consider the restriction

$$f_v = L_v|_W.$$

Since $f_v \in W^*$ and W is nondegenerate, there exists $w' \in W$ such that $f_v(w) = f_{w'}(w)$ for all $w \in W$ (as in the proof of Proposition 1.2). Thus $f_{v-w'}(w) = 0$ for all $w \in W \implies v - w' \in W^\perp$ and since $v = w' + (v - w')$, we have $V = W \oplus W^\perp$, the sum being direct since $W \cap W^\perp = \{0\}$. □

1.1.1 Alternating Forms

We now discuss a specific type of bilinear form that is of particular interest. Namely, alternating bilinear forms or, simply, alternating forms. Such forms will play a vital role in the subsequent chapters.

Definition 1.3. *Let B be a bilinear form on V. We say that B is an* alternating *form on V if for every $v \in V$, $B(v, v) = 0$.*

Observe that if B is an alternating form on V and $u, v \in V$, then

$$B(u + v, u + v) = B(u, u) + B(v, v) + B(u, v) + B(v, u) = 0$$

and so $B(u, v) = -B(v, u)$. Therefore, B is *skew symmetric*. If the characteristic of $\mathbb{F}$ is not 2, then the converse is true.

It should also be noted that if B is an alternating form on V and $B(u, v) = 0$ for some $u, v \in V$ (i.e. u is orthogonal to v), then $B(v, u) = -B(u, v) = 0$. Therefore v is orthogonal to u and so any alternating form is reflexive.

Suppose B is an alternating form on V which is not identically zero. Then there are $u, v \in V$ such that $B(u, v) = \lambda \neq 0$. If $u = av$ for some $a \in \mathbb{F}$, then $B(u, v) = aB(v, v) = 0$, a contradiction. Hence, u and v are linearly independent. We may assume that $\lambda = 1$ by replacing u with $\frac{1}{\lambda}u$. We call (u, v) a *hyperbolic pair* and the subspace W spanned by u and v a *hyperbolic plane*. Note that the restriction $B|_{W\times W}$ has representing matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

relative to (u, v) and so W is a nondegenerate subspace.

Theorem 1.6. *If B is an alternating form on V, then there are hyperbolic planes $W_1, \ldots, W_k$ such that*

$$V = W_1 \ominus \cdots \ominus W_k \ominus \text{rad } V.$$

Proof. Induction on $n = \dim V$. If $n = 1$, then V is spanned by some nonzero $v \in V$ and so for any $u, w \in V$ with $u = av$ and $w = bv$, $B(u, w) = abB(v, v) = 0$. Thus, $V = \text{rad } V$. Suppose that the result holds on spaces of dimension less than n. If V is of dimension n and B is not identically zero, then there exists $u, v \in V$ with $B(u, v) = 1$ as discussed above. Then if W_1 is the hyperbolic plane spanned by u and v, we have that $B|_{W_1\times W_1}$ is nondegenerate. Therefore, $V = W_1 \ominus W_1^\perp$ by Proposition 1.5 and thus, $\dim W_1^\perp = \dim V - \dim W = n - 2$. Therefore, by the induction hypothesis, there exists hyperbolic planes $W_2, \ldots, W_k$ such that

$$W_1^\perp = W_2 \ominus \cdots \ominus W_k \ominus \text{rad } W_1^\perp.$$

Note that $\text{rad } V = V^\perp = (W_1 \ominus W_1^\perp)^\perp = W_1^\perp \cap (W_1^\perp)^\perp = \text{rad } W_1^\perp$ and hence,

$$V = W_1 \ominus \cdots \ominus W_k \ominus \text{rad } V.$$

In the case when $B = 0$, $V = \text{rad } V$ and so the result is also true. □

Corollary 1.7. *If B is a nondegenerate alternating form on V, then* $\dim V$ *is even.*

Proof. By definition, rad $V = \{0\}$. Combining this with Theorem 1.6, we have

$$V = W_1 \oplus \cdots \oplus W_k$$

with each W_i a hyperbolic plane. Therefore, dim $V = 2k$. □

Definition 1.4. *If B is a nondegenerate alternating form on V, then we call $(V,\ B)$ (or simply V if the form is understood) a* symplectic space*.*

1.1.2 Symmetric Forms

We now turn our attention to another particular type of bilinear form. Namely, symmetric bilinear forms or, simply, symmetric forms. We will assume for the remainder of this chapter that char $\mathbb{F} \neq 2$.

Definition 1.5. *We say a bilinear form B on V is symmetric if for all $u, v \in V$,*

$$B(u, v) = B(v, u).$$

For our purposes, we define an *inner product* on a vector space V over a field $\mathbb{F}$ of any characteristic to be either a nondegenerate symmetric bilinear form on V, or a nondegenerate alternating form on V. In fact, such forms characterize all reflexive forms on V thanks to the following proposition.

Proposition 1.8. *Let V be a vector space over a field $\mathbb{F}$ of any characteristic. A bilinear form B on V is reflexive if and only if B is either symmetric or alternating.*

Proof. We have already seen that symmetric and alternating forms are reflexive. Suppose then that B is reflexive. Let $u, v, w \in V$ and set $x = B(u,\ v)w - B(u,\ w)v$. Then,

$$B(u,\ x) = B(u,\ v)B(u,\ w) - B(u,\ w)B(u,\ v) = 0$$

and so $u \perp x$. Therefore, $x \perp u$ and so

$$B(u,\ v)B(w, u) - B(u,\ w)B(v,\ u) = 0. \tag{$*$}$$

Letting $u = v$, we find that

$$B(v, v)[B(w, v) - B(v, w)] = 0. \tag{$**$}$$

Now, suppose to the contrary that B is neither alternating nor symmetric. Then, there exists $x, y, z \in V$ such that $B(x,\ x) \neq 0$ and $B(y,\ z) \neq B(z,\ y)$. Equation $(**)$ implies that

$$B(y,\ y) = 0 = B(z,\ z),\ B(x,\ y) = B(y,\ x),\ \text{and}\ B(x,\ z) = B(z,\ x).$$

Now, equation $(*)$ gives $B(y,\ x)B(z,\ y) - B(y,\ z)B(x,\ y) = 0 \implies B(x,\ y) = 0 = B(y,\ x)$. Moreover, $B(z,\ x)B(y,\ z) - B(z,\ y)B(x,\ z) = 0 \implies B(x,\ z) = 0\ = B(z,\ x)$. Hence,

$$B(x + z,\ y) = B(z,\ y) \neq B(y,\ z) = B(y, x + z).$$

Thus, by equation $(**)$, we have

$$\begin{aligned} &B(x + z,\ x + z)[B(y,\ x + z) - B(x + z,\ y)] = 0 \\ \iff &B(x + z,\ x + z) = 0. \end{aligned}$$

But,

$$\begin{aligned} B(x+z,\, x+z) &= B(x,\, x) + B(x,\, z) + B(z,\, x) + B(z,\, z) \\ &= B(x,\, x) \neq 0. \end{aligned}$$

a contradiction. □

If B is a symmetric form on V, define its associated quadratic form by

$$Q : V \to \mathbb{F},\ v \mapsto B(v, v).$$

We then have that for all $a \in \mathbb{F}$ and $v \in V$,

$$Q(av) = a^2 Q(v).$$

Furthermore, if $u \in V$,

$$Q(u+v) = B(u+v, u+v) = Q(u) + Q(v) + 2B(u, v).$$

Since char $\mathbb{F} \neq 2$, we have

$$B(u, v) = \frac{1}{2}[Q(u+v) - Q(u) - Q(v)].$$

Hence symmetric forms and quadratic forms are equivalent over fields of characteristic other than 2.

Proposition 1.9. *If B is a symmetric form on V which is not identically zero, then its associated quadratic form satisfies*

$$Q(v) \neq 0$$

for some $v \in V$.

Proof. Suppose to the contrary that Q is identically zero. Then if $u, v \in V$,

$$B(u, v) = \frac{1}{2}[Q(u+v) - Q(u) - Q(v)] = 0,$$

a contradiction. □

A set $\{v_1, \ldots v_k\} \subset V$ is *orthogonal* (relative to B) if $i \neq j \implies B(v_i, v_j) = 0$. If in addition $\{v_1, \ldots, v_k\}$ forms a basis, then it is called an *orthogonal basis.*

Theorem 1.10. *If B is a symmetric form on V, then V has an orthogonal basis.*

Proof. If B is identically zero, then every basis for V is orthogonal. Assume that B is nonzero. We argue by induction on $n = \dim V$. If $n = 1$, there is nothing to prove. Assume the result holds for spaces of dimension less than n. Suppose V is of dimension n. By proposition 1.9, there exists $v_1 \in V$ such that $Q(v_1) \neq 0$. Then the subspace W generated by v_1 is nondegenerate and so $V = W \ominus W^\perp$ by Proposition 1.5. Furthermore, $\dim W^\perp = n-1$ by Proposition 1.2 and so there exists an orthogonal basis $\{v_2, \ldots, v_n\}$ for $W^\perp$ and thus $\{v_1, \ldots, v_n\}$ is an orthogonal basis for V. □

Let B be a symmetric form on V. We say $v \in V \setminus \{0\}$ is *isotropic* if $Q(v) = 0$. Otherwise, v is said to be *anisotropic.* If V contains an isotropic vector, then V, B and Q are said to be *isotropic.* By convention, we take 0 to be anisotropic. Call V *totally isotropic* if every

$v \in V \setminus \{0\}$ is isotropic. If the image of V under Q is $\mathbb{F}$, say that B and Q are *universal.* If B is nondegenerate, call (V, B) a *quadratic space.*

If a subspace $W \subset V$ is totally isotropic and $u, w \in W$, then $u + w$ is isotropic and so

$$B(u, w) = \frac{1}{2}[Q(u + w) - Q(u) - Q(w)] = 0.$$

Thus, $W \subset W^{\perp}$. Conversely, if $W \subset W^{\perp}$ and $w \in W$, then $w \perp w \implies Q(w) = 0$ and so W is totally isotropic. Therefore, we have the characterization that a subspace is totally isotropic if and only if it is contained in its orthogonal complement.

Proposition 1.11. *If (V, B) is a quadratic space and B is isotropic, then B is universal.*

Proof. Let $v \in V$ be isotropic. Since B is nondegenerate, there exists $w \in V$ such that $B(v, w) \neq 0$. We may assume that $B(v, w) = \frac{1}{2}$, replacing w with $\frac{w}{2B(v,w)}$ if need be. For any $a \in \mathbb{F}$, set $u = (a - Q(w))v + w$. Then,

$$\begin{aligned} Q(u) &= \big(a - Q(w)\big)^2 Q(v) + Q(w) + 2\big(a - Q(w)\big)B(v, w) \\ &= Q(w) + a - Q(w) \\ &= a. \end{aligned}$$

□

Recall that if V is equipped with an alternate form B, a subspace of V is a hyperbolic plane if it has a basis $\{u, v\}$ such that $B(u, v) = 1$. Similarly if B is symmetric, a *hyperbolic plane* is a subspace W of V such that there exists a basis $\{u, v\}$ for W satisfying $Q(u) = Q(v) = 0$ and $B(u, v) = 1$. Again, we call (u, v) a *hyperbolic pair.* Note that in the symmetric case, the representing matrix of B relative to (u, v) is given by

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Proposition 1.12. *Suppose (V, B) is a quadratic space with $\dim V = 2$. Then V is a hyperbolic plane if and only if V is isotropic.*

Proof. Clearly, if V is a hyperbolic plane then it has a hyperbolic pair (u, v) as a basis and both u and v are isotropic vectors in V. Suppose now that V is isotropic and choose $v \in V \setminus \{0\}$ with $Q(v) = 0$. Since B is nondegenerate, $v \notin V^{\perp}$ and so there exists $w \in V$ such that $B(v, w) = a \neq 0$. If $w = \lambda v$ for some $\lambda \in \mathbb{F}$, then $B(v, w) = \lambda Q(v) = 0 \iff \lambda = 0$. This cannot happen since $w \neq 0$ and hence, $\{v, w\}$ is a basis for V. We may assume without loss of generality that $a = 1$ (again by replacing w with $\frac{1}{a}w$ if need be). Let $u = w - \frac{1}{2}Q(w)v$. Then,

$$Q(u) = Q(w) + \big(\frac{1}{2}Q(w)\big)^2 Q(v) - Q(w)B(w, v) = Q(w) - Q(w) = 0.$$

Note that $\{u, v\}$ is also basis for V and that

$$B(u, v) = B(w, v) - \frac{1}{2}Q(w)B(v, v) = 1.$$

Thus (u, v) is a hyperbolic pair. □

Proposition 1.13. *Suppose (V, B) is a quadratic space, $W \subset V$ is a subspace with* $\operatorname{rad} W \neq 0$, *and W' is a complement to* rad W *so that* $W =$ rad $W \obot W'$. *If $\{w_1, \dots, w_k\}$ is a basis for* rad W, *then there exists a subspace $U = \langle u_1, \dots, u_k \rangle$ such that*

(i) $W \cap U = \{0\}$.
(ii) $W \oplus U$ *is nondegenerate.*
(iii) *For $1 \leq i \leq k$, (w_i, u_i) is a hyperbolic pair for $H_i = \langle w_i, u_i \rangle$.*
(iv) $W \oplus U = W' \obot H_1 \obot H_2 \obot \cdots \obot H_k$.

Proof. Induction on $k =$ dim rad W. First note that rad $W =$ rad $W \oplus$ rad W' and thus rad $W' = \{0\}$ and so W' is a nondegenerate subspace. Define $f : W \to \mathbb{F}$ via

$$f(w_i) = \delta_{i1} \text{ for } 1 \leq i \leq n,\ f|_{W'} = 0,$$

and extending f by linearity so that $f \in W^*$. Then, as in the proof for Proposition 1.2, we may find $v_1 \in V$ so that $f = L_{v_1}$ where $L_{v_1}(w) = B(v_1, w)$. We then have that $v_1 \perp W'$ and for $1 \leq i \leq n$, $B(v_1, w_i) = \delta_{i1}$. If $a \in \mathbb{F}$, then

$$B(v_1 + aw_1, w) = B(v_1, w) + aB(w_1, w) = B(v_1, w)$$

so we may replace v_1 with $v_1 + aw_1$. Notice that

$$Q(v_1 + aw_1) = Q(v_1) + a^2Q(w_1) + 2aB(v_1, w_1) = Q(v_1) + 2a.$$

Taking $a = -\frac{1}{2}Q(v_1)$ and relabelling, we may assume that $Q(v_1) = 0$. Hence, (w_1, v_1) is a hyperbolic pair and so $V = H_1 \obot H_1^{\perp}$ where $H_1 = \langle w_1, v_1 \rangle$ (by Proposition 1.5). If $k = 1$, take $U = \langle v_1 \rangle$ so that $W \oplus U = \langle w_1 \rangle \oplus W' \oplus \langle v_1 \rangle = W' \obot H_1$ as desired. Suppose that $k > 1$. Let

$$S = W' \obot \langle w_2, \dots w_k \rangle \subset H_1^{\perp}.$$

Since W' is nondegenerate, rad $S = \langle w_2, \dots, w_k \rangle$ and in particular, dim rad $S = k - 1$. Thus, there exists $v_2, \dots, v_k \in V$ such that for $2 \leq i \leq k$, $v_i \perp W'$ and (w_i, v_i) is a hyperbolic pair for $H_i = \langle w_i, v_i \rangle$. Moreover, $S \cap \langle v_2, \dots, v_k \rangle = \{0\}$ and

$$S \oplus \langle v_2, \dots v_k \rangle = W' \obot H_2 \obot H_3 \obot \cdots \obot H_k.$$

Take $U = \langle v_1, \dots v_k \rangle$. Then $W \cap U = \{0\}$ and

$$W \oplus U = W' \obot H_1 \obot H_2 \obot \cdots \obot H_k.$$

□

Chapter 2

Classical Linear Groups

A group is a classical group if it is isomorphic to either the general linear group, the symplectic group, the orthogonal group, or the unitary group [11, p. 239]. This definition is hardly strict as some prefer to encompass other related groups such as $SL(V)$ and $PSL(V)$. This chapter first includes a brief discussion of the general linear and special linear groups following chapter 1 of [6]. Then, we introduce the symplectic group and the theory concerning such groups. Finally, we study the orthogonal group, leading to two famous theorems of Witt. The theory developed in this chapter will be vital for what is to come in the third chapter.

2.1 The General Linear and Special Linear Groups

Let V be an n-dimensional vector space over a field $\mathbb{F}$ and for $v_1, \ldots, v_k \in V$, denote their span by $\langle v_1, \ldots, v_k \rangle$.

Definition 2.1. *The general linear group of V is the group*

$$GL(V) = \{\tau : V \to V \mid \tau \text{ is linear and invertible}\}$$

with group operation given by composition, i.e., $GL(V)$ is the automorphism group of V.

If we fix a basis for V, then any linear map $V \to V$ is represented by a matrix over $\mathbb{F}$, and the composition of linear maps corresponds to matrix multiplication. Thus, we have that have that $GL(V)$ is isomorphic to the group $GL(n, \mathbb{F})$ of invertible $n \times n$ matrices over $\mathbb{F}$.

Consider now the determinant homomorphism

$$\det : GL(V) \to \mathbb{F}^\times$$

between the general linear group of V and the multiplicative group $\mathbb{F}^\times$ of $\mathbb{F}$. The kernel of the determinant homomorphism, ker(det), is a subgroup of $GL(V)$ which is of particular interest. We denote this subgroup by $SL(V)$ and call it the *special linear group.* We then have that

$$SL(V) = \{\tau \in GL(V) \mid \det \tau = 1\}.$$

Remark 2.1. *Note that $SL(V)$ is normal in $GL(V)$ as it is the kernel of a homomorphism.*

We also define the *projective special linear group* of V, denoted $PSL(V)$, to be the quotient of $SL(V)$ by its center. That is,

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

In the case when $\dim V \geq 2$, every $PSL(V)$ group is simple with the exceptions of when $\dim V = 2$ and $|\mathbb{F}|$ is either 2 or 3 (see [6, Theorem 1.13]).

If $|\mathbb{F}| = q$ is finite, then denote $GL(n, \mathbb{F})$ by $GL(n, q)$ and similarly for the matrix groups $SL(n, \mathbb{F})$ and $PSL(n, \mathbb{F})$. We now compute the order of these groups.

Proposition 2.1. *Suppose that $|\mathbb{F}| = q$ is finite. Then*

$$|GL(n, q)| = \prod_{i=0}^{n-1}(q^n - q^i) = q^{\frac{(n-1)n}{2}} \prod_{i=1}^{n}(q^i - 1).$$

Proof. Since an $n \times n$ matrix A over $\mathbb{F}$ lies in $GL(n, q)$ if and only if the rows of A form a basis for V, the problem reduces to counting the number of ordered bases of V. Let $\{v_1, \ldots, v_n\}$ be an ordered basis of V. There are q^n vectors in V so after excluding the zero vector, there are $q^n - 1$ choices for v_1. We then choose v_2 such that $v_2 \notin \langle v_1 \rangle$. This gives $q^n - q$ choices for v_2 as there are q vectors in $\langle v_1 \rangle$. In general, for $0 \leq i < n$, there are $q^n - q^i$ choices for v_{i+1}. The result then follows. □

Proposition 2.2. *Suppose that $|\mathbb{F}| = q$ is finite. Then*

$$|SL(n, q)| = \frac{|GL(n, q)|}{q - 1}.$$

Proof. Consider the subgroup D of $GL(n, q)$ consisting of diagonal block matrices of the form

$$\begin{bmatrix} a & 0 \\ 0 & I_{n-1} \end{bmatrix},$$

where $a \in \mathbb{F}^\times$ and I_{n-1} is the $(n-1) \times (n-1)$ identity matrix. Then $D \cong \mathbb{F}^\times$ and so $|D| = q - 1$. Let $A \in GL(n, q)$ and $\lambda = \det A$. Then the matrix

$$B = \begin{bmatrix} \lambda & 0 \\ 0 & I_{n-1} \end{bmatrix}$$

lies in D and $AB^{-1} \in SL(n, q)$ and so $GL(n, q) = SL(n, q)D$. Moreover, $SL(n, q) \cap D = \{I_n\}$ and since $SL(n, q)$ is normal in $GL(n, q)$ (as previously noted), we have that

$$GL(n, q) = SL(n, q) \rtimes D$$

is an inner semidirect product. Thus,

$$|GL(n, q)| = |SL(n, q)| \cdot |D|.$$

□

To compute the order of $PSL(n, q)$, we make use of the following result.

Lemma 2.3. *The center of $SL(n, \mathbb{F})$ is given by*

$$Z(SL(n, \mathbb{F})) = \{\lambda I_n \mid \lambda \in \mathbb{F}^\times,\ \lambda^n = 1\}.$$

Proof. Let $k, l \in \{1, \ldots, n\}$ be distinct and consider the elementary matrix

$$X_{kl} = I_n + E_{kl},$$

where E_{kl} is a matrix unit. Since $k \neq l$, we have that $X_{kl} \in SL(n, \mathbb{F})$. Therefore if $A = [a_{ij}] \in Z(SL(n, \mathbb{F}))$, then

$$AX_{kl} = X_{kl}A.$$

Note that right multiplication by X_{kl} adds the k^{th} column of A to the l^{th} column of A and left multiplication by X_{kl} adds the l^{th} row of A to the k^{th} row of A. Thus if $i \neq k$ and $j \neq l$, we have that

$$\begin{aligned} a_{il} + a_{ik} &= a_{il} \\ a_{kj} + a_{lj} &= a_{kj}. \end{aligned}$$

Therefore, $a_{ik} = a_{lj} = 0$. Moreover,

$$a_{kl} + a_{kk} = a_{kl} + a_{ll}$$

and so $a_{kk} = a_{ll}$. Since k and l were arbitrary, we have that A is of the form

$$A = \lambda I_n$$

with the property that $\lambda^n = 1$ as det $A = 1$. □

We are now equipped to compute the order of $PSL(n, \mathbb{F})$.

Proposition 2.4. *Suppose that $|\mathbb{F}| = q$ is finite. Then*

$$|PSL(n, q)| = \frac{|SL(n, q)|}{gcd(n, q-1)}$$

Proof. We make use of the fact that the multiplicative group of a finite field is cyclic. Since $PSL(n, q)$ is the quotient of $SL(n, q)$ by its center, it suffices to compute $|Z(SL(n, q))|$. Note that $Z(SL(n, q)) \cong \{\lambda \in \mathbb{F}^\times \mid \lambda^n = 1\}$, the subgroup of $\mathbb{F}^\times$ consisting of all n^{th} roots of unity. Let $d = gcd(n, q-1)$. We claim that if $\lambda \in \mathbb{F}^\times$, then $\lambda^n = 1$ if and only if $\lambda^d = 1$. Given that d divides n, there exists $k \in \mathbb{Z}$ so that $n = dk$. Then if $\lambda^d = 1$, we have that $\lambda^n = (\lambda^d)^k = 1$. Conversely, suppose $\lambda^n = 1$. Note that there exists $a, b \in \mathbb{Z}$ such that

$$d = an + b(q-1).$$

Then, $\lambda^d = (\lambda^n)^a(\lambda^{(q-1)})^b = 1$ and so in fact, $Z(SL(V)) \cong \{\lambda \in \mathbb{F}^\times \mid \lambda^d = 1\}$, the cyclic subgroup of order d. □

2.1.1 Transvections

We now introduce the notion of a *transvection* which will help us both in the study of $SL(V)$. The contents explored here will also prove to be useful in the following section.

Definition 2.2. *A subspace W of an n-dimensional vector space V is called a hyperplane if* $\dim W = n-1$. *If $\tau \in GL(V)$, $\tau \neq \mathrm{id}_V$, then τ is called a transvection if there exists a hyperplane W such that $\tau|_W = \mathrm{id}_W$ and $\tau v - v \in W$ for all $v \in V$. We say that W is the fixed hyperplane of τ.*

Suppose that τ is a transvection with fixed hyperplane W. Take $v_1 \in V \setminus W$ and choose a basis $\{v_2, \ldots, v_n\}$ for W. For $2 \leq i \leq n$, $\tau v_i = v_i$. Write $\tau v_1 = \sum_{i=1}^{n} x_i v_i$. Given that $\tau v_1 - v_1 \in W$, we have that

$$\tau v_1 - v_1 = \sum_{i=2}^{n} y_i v_i = \sum_{i=1}^{n} x_i v_i - v_1$$

and thus,

$$(x_1 - 1)v_1 = \sum_{i=2}^{n} (y_i - x_i) v_i.$$

The right-hand side of this equation lies in W yet $v_1 \notin W$ so we must have that $x_1 = 1$. Therefore, the matrix representing τ relative to $\{v_1, \ldots, v_n\}$ is upper triangular with ones along the main diagonal and so $\det \tau = 1 \implies \tau \in SL(V)$.

Proposition 2.5. *If $u, v \in V$ are linearly independent, then there exists a transvection τ such that $\tau u = v$.*

Proof. Choose a hyperplane W in V so that $u \notin W$ but $u - v \in W$. There is a linear map $\tau : V \to V$ satisfying $\tau|_W = \mathrm{id}_W$ and $\tau u = v$. If $x \in V$, then there exists $\lambda \in \mathbf{F}$ and $w \in W$ such that $x = \lambda u + w$. Then,

$$\tau x - x = \lambda v + w - (\lambda u + w) = -\lambda(u - v) \in W.$$

Hence, τ is a transvection. □

Proposition 2.6. *Let W_1 and W_2 be two distinct hyperplanes in V and $v \in V \setminus (W_1 \cup W_2)$. Then there exists a transvection τ such that $\tau v = v$ and W_2 is the image of W_1 under τ.*

Proof. Since W_1 and W_2 are distinct hyperplanes in V, we have that $V = W_1 + W_2$. Therefore, $\dim W_1 \cap W_2 = \dim W_1 + \dim W_2 - \dim V = n - 2$ and hence $W = W_1 \cap W_2 + \langle v \rangle$ is another hyperplane. Write $v = x + y$ where $x \in W_1$ and $y \in W_2$. Then $x \notin W_2 \implies W_1 = W_1 \cap W_2 + \langle x \rangle$ and $y \notin W_1 \implies W_2 = W_1 \cap W_2 + \langle y \rangle$ and thus $V = W_1 \cap W_2 + \langle x, y \rangle$. We then have that $x \notin W$. Otherwise, $y = v - x \in W$ and we would have that $V \subset W$. Now, define $\tau : V \to V$ via $\tau|_W = \mathrm{id}_W$ and $\tau x = -y$, extending τ by linearity. We then have that τ is a transvection as in the proof of Proposition 2.5. Furthermore, $\tau v = v$ since $v \in W$ and also,

$$\tau(W_1) = \tau(W_1 \cap W_2 + \langle x \rangle) = W_1 \cap W_2 + \langle y \rangle = W_2.$$

□

We now show that $SL(V)$ is generated by transvections. To do this, we require the following two lemmas.

Lemma 2.7. *For a transvection τ on V with fixed hyperplane W, we have that τ^{-1} is also a transvection on V with fixed hyperplane W.*

Proof. Note that $\tau \neq id_V \implies \tau^{-1} \neq \mathrm{id}_V$. For $w \in W$, we have that $\tau w = w \implies w = \tau^{-1}w$ and so $\tau^{-1}|_W = \mathrm{id}_W$. If $v \in V$, then $\tau v - v = u \in W$ and so $\tau^{-1}v - v = -u \in W$. □

Lemma 2.8. *Suppose that U is a subspace of V, $v \in V \setminus U$, and that T is a transvection on U with fixed hyperplane W'. Then there is a transvection τ on V with fixed hyperplane W such that $T = \tau|_U$ and $v \in W$.*

Proof. Choose a basis $\mathcal{B}_1 = \{v_1, \dots, v_m\}$ for U such that $\mathcal{B}_1 \setminus \{v_1\}$ is a basis for W'. Set $v_{m+1} = v$ and extend $\mathcal{B}_1$ to a basis $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ for V where $\mathcal{B}_2 = \{v_{m+1}, \dots, v_n\}$. Let W be the subspace of V spanned by $\mathcal{B} \setminus \{v_1\}$ and define $\tau : V \to V$ via

$$\tau|_U = T,\ \tau v_i = v_i \text{ for } m < i \leq n,$$

and extending by linearity. Then $\tau|_W = id_W$ and if $x = \sum_{i=1}^{n} \lambda_i v_i \in V$, then

$$\begin{aligned} \tau x - x &= \sum_{i=1}^{n} \lambda_i(\tau v_i) - \sum_{i=1}^{n} \lambda_i v_i \\ &= \sum_{i=1}^{n} \lambda_i(T v_i - v_i) \in W' \subset W. \end{aligned}$$

Hence, τ is a transvection on V with fixed hyperplane W. □

Theorem 2.9. *$SL(V)$ is generated by the set of transvections on V.*

Proof. We saw earlier that every transvection lies in $SL(V)$. It remains to show that every $\sigma \in SL(V)$ is a product of transvections. Let $\sigma \in SL(V)$. Choose a hyperplane W in V and let $v \in V \setminus W$. If v and σv are linearly independent, then by Proposition 2.5 there exists a transvection τ_1 such that $\tau_1\sigma v = v$. Otherwise, there exists a transvection τ_0 such that v and $\tau_0\sigma v$ are linearly independent and then find a transvection τ_1' such that $\tau_1\sigma v = v$ where $\tau_1 = \tau_1'\tau_0$. In either case, we have that $\tau_1\sigma v = v$ with τ_1 a product of transvections. Note that $v \notin \tau_1\sigma(W)$. If $\tau_1\sigma(W) = W$, then set $\tau_2 = id_V$. Otherwise, there exists a transvection τ_2 such that $\tau_2\tau_1\sigma(W) = W$ and $\tau_2 v = v$ by Proposition 2.6. Set $\rho = \tau_2\tau_1\sigma \implies \rho v = v$. If we set $v_1 = v$ and choose a basis $\{v_2, \dots, v_n\}$ for W so that $\mathcal{B} = \{v_1, \dots, v_n\}$ is a basis for V, then, relative to $\mathcal{B}$, ρ has a representing matrix of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & A \end{bmatrix}$$

where A is a matrix representing $\rho|_W$ relative to $\mathcal{B} \setminus \{v_1\}$. Since $\rho \in SL(V)$, we have that

$$\det\, \rho = 1 \cdot \det\, A = 1$$

and so $\rho|_W \in SL(W)$. We now argue by induction on $n = \dim V$. If $n = 2$, then we can see from the matrix representing ρ that $\rho = id_V$ and so $\sigma = \tau_1^{-1}\tau_2^{-1}$ is a product of transvections by Lemma 2.7. Suppose that $n > 2$. By the induction hypothesis, $\rho|_W$ is a product of transvections on W and by Lemma 2.8, each transvection on W extends to a transvection on V whose fixed hyperplane contains v. Therefore, ρ is a product of transvections on V and hence $\sigma = \tau_1^{-1}\tau_2^{-2}\rho$ is a product of transvections. □

2.2 The Symplectic Group

About two centuries ago, symplectic geometry provided a lens for physicists to study classical mechanics through [12, p. xiii]. However, the study of symplectic geometry has since proven to be rich and highly appreciated as a field in its own right. The symplectic group was formerly called the "abelian linear group", attributed to Abel who first studied this group. In the modern day, this terminology is highly ambiguous since abelian groups generally refer to groups whose elements commute under the given operation. The term "symplectic", as we know it today, was coined by Hermann Weyl and was a supplement for the term "complex" for which Weyl was formerly using [14, p. 165].

In this section, we assume that V is a vector space over $\mathbb{F}$ with dimension $n = 2k$ for some $k \in \mathbb{N}$ and that B is a nondegenerate alternating form on V, i.e., (V, B) is a *symplectic space.* We call the isometries of V (relative to B) *symplectic* and a *symplectic basis* for V is a basis $\{u_1, v_1, u_2, v_2, \dots, u_k, v_k\}$ for V where each (u_i, v_i) is a hyperbolic pair such that V decomposes into an orthogonal direct sum of the subspaces spanned by such pairs. Note that Theorem 1.6 guarantees the existence of a symplectic basis. We now embark on our study of the symplectic group following chapter 3 of [6].

Definition 2.3. *The symplectic group on V is the subgroup of $GL(V)$ given by*

$$Sp(V) = \{\tau \in GL(V) \mid \tau \text{ is symplectic}\}.$$

If $|\mathbb{F}| = q$ is finite and a basis has been fixed, we denote the matrix group $Sp(n, \mathbb{F})$ by $Sp(n, q)$.

We also define the *projective symplectic group*, denoted $PSp(V)$, to be the quotient of $Sp(V)$ by its center. That is,

$$PSp(V) = Sp(V)/Z(Sp(V)).$$

Much like the projective special linear group, $PSp(V)$ is simple with few exceptions. It is known that $PSp(V)$ is always simple with the exceptions of $PSp(2, 2)$, $PSp(2, 3)$, and $PSp(4, 2)$ (see [6, Theorem 3.11]).

Proposition 2.10. *Let $\mathcal{B}$ be a symplectic basis for V and suppose $\tau \in GL(V)$ is represented, relative to $\mathcal{B}$, by the matrix A. Then $\tau \in Sp(V)$ if and only if $\widehat{B} = A^T \widehat{B} A$.*

Proof. Let $u, v \in V$ with corresponding coordinate vectors $\hat{u}$ and $\hat{v}$ relative to $\mathcal{B}$. Then

$$\begin{aligned} B(u, v) &= B(\tau u, \tau v) \\ \hat{u}^T \widehat{B} \hat{v} &= (A\hat{u})^T \widehat{B} (A\hat{v}) \\ \hat{u}^T \widehat{B} \hat{v} &= \hat{u}^T (A^T \widehat{B} A) \hat{v}. \end{aligned}$$

□

Corollary 2.11. *If V is a hyperbolic plane, then $Sp(V) = SL(V)$.*

Proof. From our discussion of alternating forms, we may assume that B is represented by the matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Let $A \in GL(2, \mathbb{F})$ and write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

By Proposition 2.10, $A \in Sp(2, \mathbb{F}) \iff \widehat{B} = A^T \widehat{B} A$ where

$$A^T \widehat{B} A = \begin{bmatrix} 0 & ad - bc \\ -(ad - bc) & 0 \end{bmatrix}.$$

Thus, $A \in Sp(2, \mathbb{F}) \iff \det A = 1 \iff A \in SL(2, \mathbb{F})$. □

2.2.1 The Structure of $Sp(V)$

We are now prepared to examine the relation between transvections and the general structure of the group $Sp(V)$.

Proposition 2.12. *Let $\tau \in Sp(V)$ be a transvection with fixed hyperplane W. Then there exists $\alpha \in \mathbb{F}^\times$ and $u \in W$ such that for all $v \in V$,*

$$\tau v = v + \alpha B(v, u)u.$$

Proof. By Proposition 1.2, $\dim W^\perp = 1$ and consequently $\dim (W^\perp)^\perp = n - 1$. Furthermore, $W \subset (W^\perp)^\perp$ and so in fact, $W = (W^\perp)^\perp$. Suppose now that $W^\perp = \langle u \rangle$. Note that $u \in (W^\perp)^\perp \implies u \in W$. Choose $x \in V \setminus W$ so that $V = \langle x \rangle \oplus W$. Let $z = \tau x - x$ and note that $z \in W$ and $z \neq 0$ (as $\tau x \neq x$). Define

$$f : V \to \mathbb{F},\ \lambda x + w \mapsto \lambda.$$

Note that f is linear $\implies f \in V^*$ and so, as in the proof of Proposition 1.2, we may find $y \in V$ such that $f(v) = B(v, y)$ for all $v \in V$. Then for every $v = \lambda x + w \in V$,

$$\begin{aligned} \tau v &= \lambda \tau x + w \\ &= \lambda(z + x) + w \\ &= f(v)z + v \\ &= B(v, y)z + v. \end{aligned}$$

Note that if $w \in W$, we then have that $\tau w = w = B(w, y)z + w$ and hence $B(w, y) = 0$ and so $y \in W^\perp$. Write $y = au$ for some $a \in \mathbb{F}^\times$. Since $\tau \in Sp(V)$, we also have that

$$\begin{aligned} B(w, x) &= B(\tau w, \tau x) \\ &= B(w, x + z) \\ &= B(w, x) + B(w, z) \end{aligned}$$

and hence $B(w, z) = 0$ for all $w \in W$ and so $z \in W^\perp$. Write $z = bu$ for some $b \in \mathbb{F}^\times$ and let $\alpha = ab$. Then,

$$\begin{aligned} \tau v &= v + B(v, y)z \\ &= v + abB(v, u)u \\ &= v + \alpha B(v, u)u. \end{aligned}$$

□

Remark 2.2. *Proposition 2.12 tells us that any transvection $\tau \in Sp(V)$ is completely determined by the scalar α and the vector u. We call τ a* symplectic transvection *and write $\tau = \tau_{u,\alpha}$.*

Suppose now that $\alpha \in \mathbb{F}^\times$ and $u \in V \setminus \{0\}$. Define

$$\tau : V \to V,\ v \mapsto v + \alpha B(v,u)u$$

and write $\tau = \tau_{u,\alpha}$. If we set $W = u^\perp = \{v \in V \mid B(v,u) = 0\}$, then $\tau v - v = \alpha B(v,u)u \in W$ and for any $w \in W$, $\tau w = w + \alpha B(w,u)u = w$. Hence τ is a transvection with fixed hyperplane W. Moreover,

$$\begin{aligned} B(\tau x, \tau y) &= B(x + \alpha B(x,u)u, y + \alpha B(y,u)u) \\ &= B(x,y) + \alpha B(y,u)B(x,u) + \alpha B(x,u)B(u,y) + \alpha^2 B(x,u)B(y,u)B(u,u) \\ &= B(x,y) + \alpha B(y,u)B(x,u) - \alpha B(x,u)B(y,u) \\ &= B(x,y). \end{aligned}$$

Therefore, $\tau_{u,\alpha} \in Sp(V)$ and so the converse of Proposition 2.12 is true. We denote by $\mathcal{T}$ the subgroup of $Sp(V)$ generated by such transvections, i.e.,

$$\mathcal{T} = \langle \tau_{u,\alpha} \mid u \in V \setminus \{0\},\ \alpha \in \mathbb{F}^\times \rangle.$$

Proposition 2.13. *Let the group $Sp(V)$ act naturally on V, i.e., for $\tau \in Sp(V)$, $v \in V$,*

$$\tau \cdot v = \tau v.$$

Then $\mathcal{T}$ acts transitively on $V \setminus \{0\}$.

Proof. Let $v, w \in V \setminus \{0\}$ be distinct. If $B(v,w) \neq 0$, take $\alpha = \frac{1}{B(v,w)}$ and $u = v - w$. Then

$$\begin{aligned} \tau_{u,\alpha}(v) &= v + \frac{B(v, v-w)}{B(v,w)}(v-w) \\ &= v + \frac{B(v,v) - B(v,w)}{B(v,w)}(v-w) = w. \end{aligned}$$

If $B(v,w) = 0$, choose $f \in V^*$ so that $f(v), f(w) \neq 0$. As in the proof of Proposition 1.2, there exists $u \in V$ such that $f(v) = B(v,u)$ and $f(w) = B(w,u)$. Since $B(v,u), B(w,u) \neq 0$, there exists $\tau, \tau' \in \mathcal{T}$ such that $\tau v = u$ and $\tau' u = w$. We then have that $\tau' \tau v = w$. □

Proposition 2.14. *Let the group $Sp(V)$ act naturally on $V \times V$, i.e., for $\tau \in Sp(V)$, $(u,v) \in V \times V$,*

$$\tau \cdot (u,v) = (\tau u, \tau v).$$

Then $\mathcal{T}$ acts transitively on the set

$$S = \{(u,v) \in V \times V \mid B(u,v) = 1\},$$

i.e., the set of all hyperbolic pairs.

Proof. Suppose that $(u_1, v_1), (u_2, v_2) \in S$. By Proposition 2.13, there exists $\tau \in \mathcal{T}$ such that $\tau u_1 = u_2$. Let $v_3 = \tau v_1$. We then have that $\tau \cdot (u_1, v_1) = (u_2, v_3) \in S$. If we can find $\sigma \in \mathcal{T}$

such that $\sigma u_2 = u_2$ and $\sigma v_3 = v_2$, then $\sigma\tau \cdot (u_1, v_1) = \sigma \cdot (\tau u_1, \tau v_1) = (u_2, v_2)$ and we are done. If $B(v_3, v_2) \neq 0$, let $\alpha = \frac{1}{B(v_3,v_2)}$, $u = v_3 - v_2$, and $\sigma = \tau_{u,\alpha}$. Then,

$$\begin{aligned}\sigma v_3 &= v_3 + \alpha B(v_3, u)u \\ &= v_3 + \frac{B(v_3, v_3) - B(v_3, v_2)}{B(v_3, v_2)}(v_3 - v_2) \\ &= v_2.\end{aligned}$$

Moreover, $B(u_2, u) = B(u_2, v_3) - B(u_2, v_2) = 0$ and so $\sigma u_2 = u_2$. Suppose that $B(v_3, v_2) = 0$. Note that $B(u_2, u_2 + v_3) = B(u_2, v_3) = 1 \implies (u_2, u_2 + v_3) \in S$. Also, $B(v_3, u_2 + v_3) = -1$ and so, as in the case of $B(v_3, v_2) \neq 0$, we may find $\sigma_1 \in \mathcal{T}$ such that $\sigma_1 v_3 = u_2 + v_3$ and $\sigma_1 u_2 = u_2$. Note also that $B(u_2 + v_3, v_2) = 1$ and so, as in the case of $B(v_3, v_2)$, we may find $\sigma_2 \in \mathcal{T}$ such that $\sigma_2(u_2 + v_3) = v_2$ and $\sigma_2 u_2 = u_2$. Write $\sigma = \sigma_2\sigma_1$. Then,

$$\begin{aligned}\sigma \cdot (u_2, v_3) &= \sigma_2 \cdot (\sigma_1 u_2, \sigma_1 v_3) \\ &= \sigma_2 \cdot (u_2, u_2 + v_3) \\ &= (u_2, v_2).\end{aligned}$$

□

We are now ready to show that in fact $Sp(V) = \mathcal{T}$, i.e., $Sp(V)$ is generated by symplectic transvections.

Theorem 2.15. *The group $Sp(V)$ is generated by symplectic transvections.*

Proof. Induction on k where $\dim V = 2k$. If $k = 1$, then $Sp(V) = SL(V)$ by Corollary 2.11 and $SL(V)$ is generated by transvections by Theorem 2.9. Suppose $k > 1$. Choose a hyperbolic pair $(u, v) \in V$ and let $W = \langle u, v\rangle$. By Proposition 1.5, we have that

$$V = W \oplus W^{\perp}.$$

Let $\sigma \in Sp(V) \implies \sigma \cdot (u, v)$ is a hyperbolic pair and so there exists $\tau \in \mathcal{T}$ such that $\tau\sigma \cdot (u, v) = (u, v)$ by Proposition 2.14. Thus, $\tau\sigma|_W = id_W$ and so $\tau\sigma|_{W^\perp} \in Sp(W^\perp)$. Since $\dim W^\perp = 2(k - 1)$, we have that $\tau\sigma|_{W^\perp}$ is a product of symplectic transvections on $W^\perp$ by the induction hypothesis. If $\tau_{w,\alpha}$ is a symplectic transvection appearing in said product, then $\tau_{w,\alpha}$ extends to a symplectic transvection $V \to V$ via the mapping

$$v \mapsto v + \alpha B(v, w)w.$$

The fixed hyperplane of this transvection contains W and since $\tau\sigma|_W = id_W$, we have that $\tau\sigma \in \mathcal{T}$ and hence $\sigma \in \mathcal{T}$. □

Remark 2.3. *Due to Theorem 2.15 and Theorem 2.9, we have that $Sp(V)$ is a subgroup of $SL(V)$.*

In general, if a group G acts on a set X and $x \in X$, we define the *stabilizer* of x to be

$$Stab_G(x) = \{g \in G \mid g \cdot x = x\}.$$

We also define the *orbit* of x to be

$$Orb_G(x) = \{g \cdot x \mid g \in G\}.$$

$Stab_G(x)$ is subgroup of G for all $x \in X$ and the *Orbit-Stabilizer Theorem* states

$$|Orb_G(x)| = [G : Stab_G(x)].$$

We use these notions to assist in computing the order of $Sp(V)$ when V is over a finite field. We then compute its center, concluding our discussion of the symplectic group. We first, however, require the following lemma.

Lemma 2.16. *Let (u, v) be a hyperbolic pair and $W = \langle u, v\rangle$. Then*

$$Stab_{Sp(V)}((u, v)) \cong Sp(W^{\perp}).$$

Proof. Note that $\sigma \in Stab_{Sp(V)}((u, v))$ if and only if $\sigma|_W = id_W$. By Proposition 1.5,

$$V = W \oplus W^{\perp}$$

and so every $\tau' \in Sp(W^{\perp})$ extends to $\tau \in Sp(V)$ such that $\tau|_W = id_W$. The result follows. □

Theorem 2.17. *Suppose that $|\mathbb{F}| = q$ is finite. Then if* $\dim V = n = 2k$,

$$|Sp(n, q)| = q^{k^2} \prod_{i=1}^{k} (q^{2i} - 1).$$

Proof. We may choose any one of the $q^n - 1$ nonzero vectors in V to serve as the first vector u in a hyperbolic pair. For $v \in V$, define $R_u(v) = B(v, u)$. The number of vectors orthogonal to u is then the cardinality of $\ker R_u$. We may find $w \in V$ such that $B(u, w) = 1$ and so $B(u, \lambda w) = \lambda$ for any $\lambda \in \mathbb{F}$ and hence R_u is surjective. Thus,

$$\dim \ker R_u = \dim V - \dim \operatorname{im} R_u = n - 1.$$

Thus, there are q^{n-1} vectors that are orthogonal to u and, consequently, the number of choices for the second vector in a hyperbolic pair is

$$\frac{q^n - q^{n-1}}{q - 1} = q^{n-1}.$$

Hence, there are $q^{n-1}(q^n - 1)$ distinct hyperbolic pairs in $V \times V$. Let (u, v) be a hyperbolic pair and set $W = \langle u, v\rangle$. By Lemma 2.16, $Stab_{Sp(V)}((u, v)) \cong Sp(W^{\perp}) = Sp(n-2, q)$ and by Proposition 2.14, $Sp(V)$ acts transitively on the set S of hyperbolic pairs and so $Orb_{Sp(V)}((u, v)) = S$. Thus, by the Orbit-Stabilizer Theorem,

$$|Sp(n, q)| = q^{n-1}(q^n - 1) \cdot |Sp(n - 2, q)|. \qquad (*)$$

We now argue by induction on k. If $k = 1$, then by Corollary 2.11, $Sp(2, q) = SL(2, q)$ and hence $|Sp(2, q)| = q(q^2 - 1)$ by Proposition 2.2. Assume that

$$|Sp(n - 2, q)| = q^{(k-1)^2} \prod_{i=1}^{k-1} (q^{2i} - 1).$$

The result then follows from equation $(*)$. □

Proposition 2.18. *The center of $Sp(V)$ is*

$$Z(Sp(V)) = \{-id_V, id_V\}.$$

Proof. For $\tau_{u,\alpha} \in Sp(V)$ and $v \in V$, $\tau_{u,\alpha}(v) = v \iff B(v,u) = 0$, i.e., v is orthogonal to u. Let $\sigma \in Z(Sp(V))$. If $v \perp u$, then

$$\tau_{u,\alpha}(\sigma v) = \sigma\tau_{u,\alpha}(v) = \sigma v$$

and thus $\sigma v \perp u$, i.e., $\sigma(u^\perp) = u^\perp$ where $u^\perp = \{v \in V \mid B(v,u) = 0\}$. Note that $\mathrm{rad}\, u^\perp = \langle u \rangle$ as $\dim \mathrm{rad}\, u^\perp = 1$. Now, let $w \in \mathrm{rad}\, u^\perp \implies w \in (u^\perp)^\perp$ and $\sigma w \in u^\perp$. If $x \in u^\perp$, then $x = \sigma y$ for some $y \in u^\perp$ and so

$$B(\sigma w, x) = B(\sigma w, \sigma y) = B(w, y) = 0.$$

Hence, $\sigma(\langle u \rangle) = \langle u \rangle$ and consequently, $\sigma u = \lambda_u u$ for some $\lambda_u \in \mathbb{F}$. Choose a basis $v_1, \dots, v_n$ for V and set $v = v_1 + \cdots + v_n \implies \sigma(v) = \lambda_v v$. On the other hand,

$$\sigma(v) = \sum_{i=1}^{n} \sigma v_i = \sum_{i=1}^{n} \lambda_{v_i} v_i.$$

Thus, $\lambda = \lambda_v = \lambda_{v_i}$ for all $1 \leq i \leq n$ and so $\sigma = \lambda\, id_V$. We then have that for all $u, w \in V$,

$$B(u,w) = B(\sigma u, \sigma w) = \lambda^2 B(u,w) \implies \lambda \in \{-1, 1\}.$$

□

Corollary 2.19. *Suppose that $|\mathbb{F}| = q$ is finite. Then*

$$|PSp(n,q)| = \frac{|Sp(n,q)|}{gcd(2, q-1)}.$$

Proof. If $\mathrm{char}\, \mathbb{F} = 2$, i.e., $q = 2^m$ for some $m \in \mathbb{N}$, then $-id_V = id_V \implies |Z(Sp(n,q))| = 1$. Otherwise, $|Z(Sp(n,q))| = 2$. Since $PSp(V) = Sp(V)/Z(Sp(V))$, the result is immediate. □

2.3 The Orthogonal Group

We now turn our attention to when B is a nondegenerate symmetric bilinear form on V. Recall that, in this case, the pair (V, B) is called a quadratic space. We further assume for this section that $\mathrm{char}\, \mathbb{F} \neq 2$ and that $\dim V = n \geq 2$. We call the group of isometries of V (relative to B) the *orthogonal group* and denote it by $O(V)$. Such isometries are called *orthogonal transformations.*

Note that $\tau \in GL(V)$ lies in $O(V)$ if and only if $Q(\tau v) = Q(v)$ for all $v \in V$, where Q is the quadratic form associated to B. Also, if $\tau \in GL(V)$ is represented by the matrix A relative to some basis $\mathcal{B}$ of V, then, as in Proposition 2.10, $\tau \in O(V)$ if and only if

$$A^T \widehat{B} A = \widehat{B},$$

where $\widehat{B}$ is the representing matrix for B relative to $\mathcal{B}$. Consequently, we have that

$$(\det A)^2 \cdot \det \widehat{B} = \det \widehat{B}$$

and thus $\det A \in \{-1,1\}$. If $\tau \in O(V)$ satisfies $\det \tau = 1$ call τ a *rotation.* Otherwise, call τ a *reversion.*

Suppose that $u \in V \setminus \{0\}$ is such that $Q(u) \neq 0$, i.e., u is anisotropic (note that Proposition 1.9 ensures the existence of such a $u \in V$). Define a linear map $\sigma_u : V \to V$ by

$$\sigma_u(v) = v - 2\frac{B(v,u)}{Q(u)}u.$$

Then for all $v, w \in V$,

$$\begin{aligned} B(\sigma_u(v), \sigma_u(w)) &= B\left(v - 2\frac{B(v,u)}{Q(u)}u,\ w - 2\frac{B(w,u)}{Q(u)}u\right) \\ &= B(v,w) - 4\frac{B(w,u)}{Q(u)}B(v,u) + 4\frac{B(v,u)B(w,u)}{Q(u)^2}B(u,u) \\ &= B(v,w). \end{aligned}$$

Therefore, $\sigma_u \in O(V)$. We call the map σ_u a *reflection along* u or, simply, a *reflection.*

Note that since u is anisotropic, $\langle u \rangle$ is a nondegenerate subspace and so $V = \langle u \rangle \ominus \langle u \rangle^\perp$ by Proposition 1.5. If we set $u_1 = u$ and choose a basis $\{u_2, \ldots, u_n\}$ for $\langle u \rangle^\perp$, then, relative to the basis $\{u_1, \ldots, u_n\}$, σ_u is represented by the matrix

$$\begin{bmatrix} -1 & 0 \\ 0 & I_{n-1} \end{bmatrix}$$

since $\sigma_u(u) = -u$ and $v \perp u \implies \sigma_u(v) = v$. Therefore σ_u is a reversion of order 2.

The existence of reversions in $O(V)$ implies that the image of the determinant homomorphism is, in fact, equal to $\{-1,1\}$. Denote the kernel of $\det : O(V) \to \mathbb{F}^\times$ by $SO(V)$. Then,

$$O(V)/SO(V) \cong \{-1,1\}$$

and so $SO(V)$ has index 2 inside $O(V)$. We call $SO(V)$ the *special orthogonal group.*

2.3.1 The Structure of $O(V)$

We now wish to study the structure of $O(V)$ which will give us a sense of what orthogonal transformations “look like”. This will lead us to a famous theorem attributed to Cartan and Dieudonné which says that every orthogonal transformation is a product of at most n reflections. Elie Cartan provided a proof for a special case of this theorem in 1937 which was later generalized by Dieudonné [3, p. 202]. It should be emphasized however that this theorem does not hold over fields of characteristic 2. Now, following chapter 6 of [6], we begin with some preparation that will be needed in proving this theorem.

Proposition 2.20. *If* $\tau \in O(V)$, *then* $\ker(\tau - id_V) = \operatorname{im}(\tau - id_V)^\perp$.

Proof. Let $v \in \ker(\tau - id_V)$ and $w \in V$. Then,

$$\begin{aligned} B(v, (\tau - id_V)w) &= B(v, \tau w) - B(v,w) \\ &= B(v, \tau w) - B(\tau v, \tau w) \\ &= -B(\tau v - v, \tau w) \\ &= 0. \end{aligned}$$

On the other hand, if $v \in \operatorname{im}(\tau - id_V)^\perp$ and $w \in V$, then,

$$\begin{aligned} B((\tau - id_V)v, \tau w) &= B(\tau v, \tau w) - B(v, \tau w) \\ &= B(v, w) - B(v, \tau w) \\ &= -B(v, \tau w - w) \\ &= 0. \end{aligned}$$

Thus, $(\tau - id_V)v \in \operatorname{rad} V = \{0\}$. □

For $\tau \in O(V)$, $\operatorname{im}(\tau - id_V)$ is totally isotropic if and only if $\operatorname{im}(\tau - id_V) \subset \operatorname{im}(\tau - id_V)^\perp$ and so, by Proposition 2.20, $\operatorname{im}(\tau - id_V) \subset \ker(\tau - id_V)$ which happens if and only if $(\tau - id_V)^2 = 0_V$. Furthermore, if $v \in V \setminus \{0\}$, then $(\tau - id_V)v$ is isotropic if and only if

$$\begin{aligned} 0 &= B(\tau v - v, \tau v - v) \\ &= 2[B(v, v) - B(\tau v, v)] \\ &= -2B(\tau v - v, v), \end{aligned}$$

i.e., $\tau v - v \perp v$.

Once the following lemma is established, we will be equipped with the tools to prove the theorem of Cartan and Dieudonné which will conclude our discussion of the structure of $O(V)$.

Lemma 2.21. *Let $\tau \in O(V)$ with $(\tau - id_V)^2 \neq 0_V$. Then there exists a nonzero anisotropic vector $v \in V \setminus \{0\}$ such that $z = \tau v - v$ is anisotropic. Moreover, if $z \neq 0$, then $\sigma_z \tau v = v$.*

Proof. Suppose to the contrary that for every nonzero anisotropic vector $v \in V \setminus \{0\}$, $\tau v - v$ is isotropic. Then, $\tau v - v \perp v$ and hence $\tau v - v \in \operatorname{rad}\langle \tau v - v,\, v\rangle \implies \langle \tau v - v,\, v\rangle$ is a degenerate subspace. Since V is nondegenerate, it follows that $\dim V \geq 3$. Now, suppose $w \in V$ is isotropic. By Proposition 1.9, there exists a nonzero anisotropic vector $x \in \langle w\rangle^\perp$ and, for a fixed $a \in \mathbb{F}^\times$, set $u = w + ax$. Then,

$$\begin{aligned} Q(u) &= Q(w) + a^2 Q(x) + 2aB(w, x) \\ &= a^2 Q(x) \\ &\neq 0 \end{aligned}$$

and so u is a nonzero anisotropic vector. Therefore, $\tau u - u \perp u$ and so

$$\begin{aligned} 0 &= B(\tau w + a\tau x - w - ax, w + ax) \\ &= B(\tau w - w, w) + a[B(\tau x - x, w) + B(\tau w - w, x)] + a^2 B(\tau x - x, x) \\ &= B(\tau w - w, w) + a[B(\tau x - x, w) + B(\tau w - w, x)]. \end{aligned}$$

Since $a \in \mathbb{F}$ is arbitrary, we must have that

$$B(\tau x - x, w) + B(\tau w - w, x) = 0 \implies B(\tau w - w, w) = 0,$$

i.e., $\tau w - w \perp w$. Therefore, by the preceding remarks as well as Proposition 1.9, $\operatorname{im}(\tau - id_V)$ is totally isotropic and so $(\tau - id_V)^2 = 0_V$, a contradiction. Hence, there exists an anisotropic vector $v \in V \setminus \{0\}$ such that $z = \tau v - v$ is anisotropic. If $z \neq 0$, then

$$\begin{aligned}\sigma_z \tau v &= \sigma_z(z+v)\\ &= -z + v - 2\frac{B(v,z)}{Q(z)}z\\ &= v - \frac{Q(z)+B(2v,z)}{Q(z)}z\\ &= v - \frac{B(\tau v + v, \tau v - v)}{Q(z)}z\\ &= v\end{aligned}$$

since $B(\tau v + v, \tau v - v) = B(\tau v, \tau v) - B(\tau v, v) + B(v, \tau v) - B(v,v) = 0$. □

We now prove the anticipated theorem following [6, p.48].

Theorem 2.22 (Cartan-Dieudonné). *Let V be a quadratic space of dimension n. Then every $\tau \in O(V)$ is a product of at most n reflections.*

Proof. Induction on n. If $n = 1$, then $O(V) = \{-id_V, id_V\}$. We view id_V as being a product of zero reflections and $-id_V$ is itself a reflection since any nonzero $v \in V \setminus \{0\}$ is a basis for V and $-id_V(v) = -v = \sigma_v(v) \implies -id_V = \sigma_v$. Suppose the result holds for spaces of dimension less than n and assume to the contrary that $\tau \in O(V)$ is not a product of at most n reflections. We first show that $(\tau - id_V)^2 = 0_V$. If not, we may, by Lemma 2.21, choose a nonzero anisotropic vector $v \in V \setminus \{0\}$ such that $z = \tau v - v$ is anisotropic.

If $z = 0$, then $\tau v = v$ so τ maps $\langle v\rangle^\perp$ onto itself. Thus by the induction hypothesis, $\tau|_{\langle v\rangle^\perp}$ is a product of at most $n-1$ reflections in $O(\langle v\rangle^\perp)$, each of which extends to a reflection in $O(V)$ fixing v. Hence, τ is a product of at most $n-1$ reflections.

Suppose then that $z \neq 0$. By Lemma 2.21, v is fixed by $\sigma_z\tau$ and so, as above, $\sigma_z\tau$ is a product of at most $n-1$ reflections in $O(\langle v\rangle^\perp)$, each of which extends to a reflection in $O(V)$. Thus, τ is a product of at most n reflections. Thus, in either case, we obtain a contradiction and therefore, $(\tau - id_V)^2 = 0_V \iff \operatorname{im}(\tau - id_V) \subset \ker(\tau - id_V)$ is totally isotropic. Observe that $\ker(\tau - id_V)$ is itself totally isotropic. Otherwise, there exists a nonzero anisotropic $v \in \ker(\tau - id_V) \implies \tau v = v$ and so, as in the case of $z = 0$, we see that τ is a product of at most $n-1$ reflections. Therefore,

$$\ker(\tau - id_V) \subset \ker(\tau - id_V)^\perp = \operatorname{im}(\tau - id_V)$$

by Proposition 2.20. Hence,

$$\operatorname{im}(\tau - id_V) = \ker(\tau - id_V) = \ker(\tau - id_V)^\perp.$$

Denote the space of fixed points of τ by $\operatorname{Fix}(\tau)$, i.e.,

$$v \in \operatorname{Fix}(\tau) \iff \tau v = v \iff v \in \ker(\tau - id_V).$$

We then have that

$$n = \dim \ker(\tau - id_V) + \dim \ker(\tau - id_V)^\perp = 2\dim \operatorname{Fix}(\tau).$$

Note that $\tau|_{\mathrm{Fix}(\tau)} = id_{\mathrm{Fix}(\tau)}$. Furhtermore, the natural action of τ on the quotient space $V/\operatorname{Fix}(\tau)$ coincides with the identity map on $V/\operatorname{Fix}(\tau)$. To see this, let $v \in V$. Then,

$$\begin{aligned}\tau \cdot (v + \operatorname{Fix}(\tau)) &= \tau v + \operatorname{Fix}(\tau)\\ &= [v + \operatorname{Fix}(\tau)] + [\tau v - v + \operatorname{Fix}(\tau)]\\ &= v + \operatorname{Fix}(\tau)\end{aligned}$$

since $\tau v - v \in \operatorname{im}(\tau - id_V) = \operatorname{Fix}(\tau)$. Now, Let $k = \dim \operatorname{Fix}(\tau)$. Choose a basis $\{v_1, \dots, v_k\}$ for $\operatorname{Fix}(\tau)$ and extend it to a basis $\{v_1, \dots v_n\}$ for V. For $1 \le j \le k$, we have that $\tau v_j = v_j$. For $k < j \le n$, we have that $\tau v_j - v_j \in \operatorname{Fix}(\tau)$ and so there exists $\lambda_{(1,j)}, \dots, \lambda_{(k,j)} \in \mathbb{F}$ such that

$$\tau v_j = v_j + \sum_{i=1}^{k} \lambda_{(i,j)} v_i.$$

Therefore, $\det \tau = 1$ and so $\tau \in SO(V)$. Let $\sigma \in O(V)$ be a reflection $\implies \sigma\tau \notin SO(V)$. If $\sigma\tau$ is not a product of at most n reflections, then repeating the above argument would yield that $\sigma\tau \in SO(V)$, a contradiction. Thus, $\sigma\tau$ is a product of at most n reflections and hence τ is a product of at most $n+1$ reflections. If τ is a product of $n+1$ reflections, then we would have that $\det \tau = (-1)^{2k+1}$ contradicting the fact that $\tau \in SO(V)$. Thus, τ is a product of at most n reflections, contradicting our hypothesis on τ. □

2.3.2 Theorems of Witt

We conclude this chapter with two theorems due to Witt, namely, Witt's Cancellation Theorem and Witt's Extension Theorem, keeping in mind that we are still under the assumption that (V, B) is a quadratic space of dimension $n \ge 2$ over a field of characteristic different from two. Witt's Cancellation Theorem was proved in his famous 1937 paper "Theorie der quadratischen Formen in beliebigen Körpern" [15] which was, oddly enough, the same year that Cartan presented his proof of a special case to the above theorem concerning the structure of $O(V)$. Witt's Cancellation Theorem has since proven to be vital to the entire study of quadratic forms over fields [8, p. 12]. After proving Witt's Cancellation Theorem (following [6, p.40]), we will prove Witt's Extension Theorem (following [6, p.41]) which will play an important role in our case.

Theorem 2.23 (Witt's Cancellation Theorem). *Suppose that U_1 and U_2 are nondegenerate isometric subspaces of a quadratic space (V, B). Then $U_1^{\perp}$ and $U_2^{\perp}$ are also isometric.*

Proof. Induction on $\dim U_1$. Suppose that $U_1 = \langle u_1 \rangle$ and $U_2 = \langle u_2 \rangle$. Then both $Q(u_1)$ and $Q(u_2)$ are nonzero by nondegeneracy. We may assume that there is an isometry $\sigma : U_1 \to U_2$ such that $\sigma u_1 = u_2$ and so $Q(u_1) = Q(u_2)$. We then have that

$$\begin{aligned}Q(u_1 + u_2) = 2Q(u_1) + 2B(u_1, u_2)\\ Q(u_1 - u_2) = 2Q(u_1) - 2B(u_1, u_2).\end{aligned}$$

We then have that one of $Q(u_1 + u_2)$ and $Q(u_1 - u_2)$ are nonzero. Otherwise,

$$B(u_1, u_2) = Q(u_1) = -B(u_1, u_2)$$

contradicting the fact that $Q(u_1) \ne 0$. Suppose then that $Q(u_1 + u_2) \ne 0$. We have that

$$B(u_1 + u_2, u_1 - u_2) = Q(u_1) - Q(u_2) = 0,$$

i.e., $u_1 + u_2 \perp u_1 - u_2$ and so the reflection $\sigma_{u_1+u_2} \in O(V)$ satisfies

$$\sigma_{u_1+u_2}(u_1 - u_2) = u_1 - u_2.$$

Hence,

$$\begin{aligned}\sigma_{u_1+u_2}(u_1) &= \frac{1}{2}\,\sigma_{u_1+u_2}((u_1+u_2)+(u_1-u_2))\\ &= \frac{1}{2}\,[-(u_1+u_2)+(u_1-u_2)]\\ &= -u_2\end{aligned}$$

and, consequently, $\sigma_{u_1+u_2}$ maps $\langle u_1\rangle$ onto $\langle u_2\rangle$ and $\langle u_1\rangle^\perp$ onto $\langle u_2\rangle^\perp$. Therefore, $U_1^\perp$ and $U_2^\perp$ are isometric via the restriction of $\sigma_{u_1+u_2}$ to $U_1^\perp$.

If $Q(u_1 - u_2) \neq 0$, then the reflection $\sigma_{u_1-u_2} \in O(V)$ satisfies

$$\sigma_{u_1-u_2}(u_1 + u_2) = u_1 + u_2.$$

Hence,

$$\begin{aligned}\sigma_{u_1-u_2}(u_1) &= \frac{1}{2}\,\sigma_{u_1-u_2}((u_1+u_2)+(u_1-u_2))\\ &= \frac{1}{2}\,[(u_1+u_2)-(u_1-u_2)]\\ &= u_2\end{aligned}$$

and, consequently, $\sigma_{u_1-u_2}$ maps $\langle u_1\rangle$ onto $\langle u_2\rangle$ and $\langle u_1\rangle^\perp$ onto $\langle u_2\rangle^\perp$. Therefore, $U_1^\perp$ and $U_2^\perp$ are isometric via the restriction of $\sigma_{u_1-u_2}$ to $U_1^\perp$.

Suppose now that dim $U_1 > 1$ and that the result holds for spaces of lower dimension. Choose a nonzero anisotropic vector $u_1 \in U_1 \setminus \{0\}$ and let W_1 be the orthogonal complement of u_1 in U_1, i.e.,

$$W_1 = \{v \in U_1 \mid B(v, u_1) = 0\}.$$

Then, $U_1 = \langle u_1\rangle \oplus W_1$. Let $\sigma : U_1 \to U_2$ be an isometry and set $u_2 = \sigma u_1$, $W_2 = \sigma(W_1)$. Then, $U_2 = \langle u_2\rangle \oplus W_2$ and

$$\begin{aligned} V &= U_1 \oplus U_1^\perp = U_2 \oplus U_2^\perp\\ \Longrightarrow\ V &= \langle u_1\rangle \oplus W_1 \oplus U_1^\perp = \langle u_2\rangle \oplus W_2 \oplus U_2^\perp.\end{aligned}$$

As in the case of dim $U_1 = 1$, one of $Q(u_1 + u_2)$ and $Q(u_1 - u_2)$ are nonzero and so there is a reflection $\Phi \in O(V)$ that maps $\langle u_1\rangle$ onto $\langle u_2\rangle$. Consequently, $W_1 \oplus U_1^\perp$ and $W_2 \oplus U_2^\perp$ are isometric via the restriction $\phi = \Phi|_{W_1 \oplus U_1^\perp}$. Therefore, $\phi\sigma^{-1}$ is as isometry $W_2 \to \phi(W_1)$. Furthermore, $U_2^\perp$ and $\phi(U_1^\perp)$ are the orthogonal complements of W_2 and $\phi(W_1)$, respectively, inside $W_2 \oplus U_2^\perp$. Thus, $U_2^\perp$ and $\phi(U_1^\perp)$ are isometric by induction, say through $\tau : \phi(U_1^\perp) \to U_2^\perp$. Then, $\tau\phi|_{U_1^\perp}$ is an isometry $U_1^\perp \to U_2^\perp$. □

Theorem 2.24 (Witt's Extension Theorem)**.** *If U_1 and U_2 are subspaces of a quadratic space (V, B) and $\sigma : U_1 \to U_2$ is an isometry, then there exists an isometry $\tau \in O(V)$ such that $\tau|_{U_1} = \sigma$.*

Proof. Suppose first that U_1 and U_2 are nondegenerate. By Witt's Cancellation Theorem 2.23, there exists an isometry $\phi : U_1^\perp \to U_2^\perp$. Since $V = U_1 \oplus U_1^\perp = U_2 \oplus U_2^\perp$, we may define an isometry $\tau \in O(V)$ by $\tau(u + u') = \sigma u + \phi u'$ for all $u \in U_1$, $u' \in U_1^\perp$ and hence, $\tau|_{U_1} = \sigma$.

Suppose then that U_1 is degenerate. Let U_1' be a complement to rad U_1 so that

$$U_1 = \text{rad } U_1 \oplus U_1'.$$

By Proposition 1.13, there is a subspace $W_1 \subset V$ such that $U_1 \oplus W_1$ is nondegenerate and

$$U_1 \oplus W_1 = U_1' \oplus H_1 \oplus H_2 \oplus \cdots \oplus H_k,$$

where each H_i is a hyperbolic plane with, say, hyperbolic pair $(u_i,\, v_i)$ and $k = \dim \text{rad } U_1$. We claim that σ maps rad U_1 onto rad U_2. Let $x \in \sigma(\text{rad } U_1)$ and $u' \in U_2$. Then, $x = \sigma w$ for some $w \in \text{rad } U_1$ and $u' = \sigma u$ for some $u \in U_1$. Then, $B(x, u') = B(w, u) = 0 \implies x \in \text{rad } U_2$. Conversely, if $y \in \text{rad } U_2$, then $y = \sigma w$ for some $w \in U_1$ and for any $u \in U_1$, we have $B(y, \sigma u) = 0 \iff B(w, u) = 0$. Hence, $y \in \sigma(\text{rad } U_1)$ and thus $U_2' = \sigma(U_1')$ is a complement to rad U_2 and so there exists a subspace $W_2 \subset V$ such that $U_2 \oplus W_2$ is nondegenerate and

$$U_2 \oplus W_2 = U_2' \oplus H_1' \oplus H_2' \oplus \cdots \oplus H_k',$$

where each H_i' is a hyperbolic plane with, say, $(\sigma u_i, v_i')$ as a hyperbolic pair. Extend σ to $\sigma' : U_1 \oplus W_1 \to U_2 \oplus W_2$ by setting $\sigma'(v_i) = v_i'$, extending by linearity. Since σ' is an isometry and $U_1 \oplus W_1$ is nondegenerate, there exists $\tau \in O(V)$ such that $\tau|_{U_1 \oplus W_1} = \sigma'$ and hence, $\tau|_{U_1} = \sigma$. □

Remark 2.4. *It should be noted that Witt's Cancellation Theorem 2.23 and Witt's Extension Theorem 2.24 both hold over fields of any characteristic in the case when B is a nondegenerate alternating form. See [6, Theorem 12.10] for Witt's Extension Theorem. Witt's Cancellation Theorem follows from 1.6.*

Corollary 2.25. *Any two maximal (in the sense of inclusion) totally isotropic subspaces of V have the same dimension and every totally isotropic subspace is contained in one of maximal dimension.*

Proof. Let $U \subset V$ be a totally isotropic subspace of maximal dimension m. If $W \subset V$ is totally isotropic, then there exists an isometry σ from W to a subspace of U. By Witt's Extension Theorem 2.24, there exists $\tau \in O(V)$ such that $\tau|_W = \sigma$. Then, $W \subset \tau^{-1}(U)$, a totally isotropic subspace of dimension m. □

Chapter 3

Binary Linear Codes

Coding theory is a vast field of practical importance that makes use of elegant mathematical theory [2, p. vii]. A frequent question in coding theory is how to construct a code or structure a code such that it satisfies imposed mathematical or practical constraints [10, p. vii]. This chapter analyzes a small subset of the classes of codes and considers their construction. In particular, this chapter introduces linear codes and aims to tie the structure of these codes together with the theory studied thus far.

3.1 Preliminaries

Following the relevant theory from chapters 1, 4 and 5 of [2], we introduce some basic notions and results.

Definition 3.1. *Let A and B be non-empty finite sets. A word in B is an element*

$$(b_1, \ldots, b_n) \in B \times \cdots \times B$$

for some $n \in \mathbb{N}$. We write $b_1 \cdots b_n$ instead of $(b_1, \ldots, b_n)$ and say that $b_1 \cdots b_n$ has length n. We denote the set of all words in B by $\omega(B)$. A coding is a map $K : A \to \omega(B)$. We call A the source alphabet, B the code alphabet, and for each $a \in A$, the element $K(a)$ a code word. Elements of A are called source symbols. If $|B| = 2$, K is called a binary code and elements of $\omega(B)$ are called binary words.

A coding $K : A \to \omega(B)$ induces a map $K^* : \omega(A) \to \omega(B)$ by setting

$$K^*(a_1 \cdots a_n) = K(a_1) \cdots K(a_n).$$

We call K^* the *coding of source messages* and say that K is *uniquely decodable* provided that K^* is injective.

We now introduce a fundamental class of codes and consider their construction. Namely, the so-called *instantaneous* codes.

Definition 3.2. *A coding $K : A \to \omega(B)$ is called* instantaneous *if for all $a, \alpha \in A$, $a \neq \alpha$, and all $b \in \omega(B)$, $K(a) \neq K(\alpha)\, b$.*

Suppose now that $A = \{a_1, \ldots, a_n\}$ is a source alphabet and B is a code alphabet of cardinality $k \geq 2$. We wish to construct an instantaneous code $K : A \to \omega(B)$. Assign an arbitrary word of length $d_1 \geq 1$ to a_1. Now, fix $d_2 \geq d_1$ and note that the number of words of length d_2 is k^{d_2} and therefore, the number of words b of length d_2 such that $b = K(a_1)\,\beta$ is $k^{d_2-d_1}$. Thus, we may choose an arbitrary word for a_2 of length d_2 such that $K(a_2) \neq K(a_1)\,b$ for any word b in B since $k^{d_2} - k^{d_2-d_1} \geq 1$.

If we fix $d_3 \geq d_2$, then there are $k^{d_3-d_1} + k^{d_3-d_2}$ words b of length d_3 such that either $b = K(a_1)\,\beta_1$, or $b = K(a_2)\,\beta_2$. Therefore if we want to assign a word to a_3 such that this is not the case, we require that

$$\begin{aligned} k^{d_3} - k^{d_3-d_1} - k^{d_3-d_2} &\geq 1 \\ \iff k^{-d_1} + k^{-d_2} + k^{-d_3} &\leq 1. \end{aligned}$$

Analogously, we require that

$$\sum_{i=1}^{n} k^{-d_i} \leq 1$$

in order to construct our instantaneous code K. Indeed, this inequality is both a sufficient and necessary condition to construct an instantaneous code.

The following theorem connects uniquely decodable codes with instantaneous codes in the sense that for every uniquely decodable code K_1 with source alphabet $A = \{a_1, \ldots, a_n\}$ and code alphabet B, there exists an instantaneous code $K_2 : A \to \omega(B)$ such that for all $1 \leq i \leq n$, $K_1(a_i)$ and $K(\sigma \cdot a_i)$ have the same lengths for some permutation $\sigma \in S_n$.

Theorem 3.1 (McMillan's Theorem). *Let K be a uniquely decodable coding with source alphabet $\{a_1, \ldots, a_n\}$ and code alphabet B of cardinality $k \geq 2$. Then,*

$$\sum_{i=1}^{n} k^{-d_i} \leq 1,$$

where d_i is the length of $K(a_i)$.

Proof. For each $1 \leq i \leq n$, there are k^i words of length i in B. Since K^* is injective, the number of source messages $a_{j_1} \cdots a_{j_r}$ such that $d_{j_1} + \cdots + d_{j_r} = i$ is bounded by k^i. Now, let

$$c = \sum_{i=1}^{n} k^{-d_i}.$$

We claim that for each $r \in \mathbb{N}$,

$$c^r = \sum_{i_1,\ldots,i_r=1}^{n} k^{-(d_{i_1}+\cdots+d_{i_r})}.$$

The case $r = 1$ is satisfied by the definition of c. Assume the result holds for c^{r-1}. Then,

$$\begin{aligned} c^r &= \Bigg(\sum_{i_1,\ldots,i_{r-1}=1}^{n} k^{-(d_{i_1}+\cdots+d_{i_{r-1}})}\Bigg)\Bigg(\sum_{i_r=1}^{n} k^{-d_{i_r}}\Bigg) \\ &= \sum_{i_1,\ldots,i_r=1}^{n} k^{-(d_{i_1}+\cdots+d_{i_{r-1}})}\, k^{-d_{i_r}} \end{aligned}$$

as desired. Collect the summands k^{-i} in c^r for which $i = d_{i_1} + \cdots + d_{i_r}$. As mentioned, the number of such summands is bounded by k^i. Furthermore, $i \leq rd$, where $d = \max(d_1, \ldots, d_n)$. We then have that

$$c^r \leq \sum_{i=1}^{rd} k^{-i} k^i = rd.$$

Hence, $\frac{c^r}{r} \leq d$ for all $r \in \mathbb{N} \iff c \leq 1$. □

3.1.1 Hamming Distance

We now introduce a crucial notion in coding theory, namely, that of Hamming distance. We will consider the role of this notion in an important area of coding theory known as error detection and correction. We begin with some rudimentary definitions.

Definition 3.3. *A coding K with source alphabet A is called a* block coding of length n *if K is injective and for all $a \in A$, $K(a)$ has length n.*

Remark 3.1. *Let K be a block coding of length n with source alphabet A and let $\alpha \in \omega(A)$ have length m. Then, the length of $K^*(\alpha)$ is $n \cdot m$. In particular, n divides the length of $K^*(\alpha)$. Moreover, if $x, y \in \omega(A)$, then $K^*(x) = K^*(y) \implies x$ and y have the same length, say k. Write $x = x_1 \cdots x_k$ and $y = y_1 \cdots y_k$. If $x \neq y$, then there exists $1 \leq i \leq k$ such that $x_i \neq y_i \implies K(x_i) \neq K(y_i) \implies K^*(x) \neq K^*(y)$ and hence K is uniquely decodable.*

Remark 3.2. *The condition that a block coding has length n for each of its code words is necessary to conclude that it is uniquely decodable. For example, define $K : \{a, b, c\} \to \omega(\{0, 1\})$ by*

$$K(a) = 0, \;\; K(b) = 1, \;\; K(c) = 01.$$

Then K is injective yet $K^*(ab) = 01 = K^*(c)$.

Definition 3.4. *Let $a = a_1 \cdots a_n$ and $b = b_1 \cdots b_n$ be two words in a set. We define the* Hamming distance *between a and b as*

$$d(a, b) = |\{i \in \{1, \ldots, n\} \mid a_i \neq b_i\}|.$$

Proposition 3.2. *For any set A and any $n \in \mathbb{N}$, the Hamming distance is a metric on the set $\omega_n(A)$ of words of length n in A. That is to say, $(\omega_n(A), d)$ is a metric space.*

Proof. Clearly, $d(a, a) = 0$, $d(a, b) > 0$ for $a \neq b$, and $d(a, b) = d(b, a)$. Suppose now that $d(a, b) = k$ and $d(b, c) = l$. Write $a = a_1 \cdots a_n$ and similarly for b and c. Then, there are indices $I = \{i_1, \ldots, i_k\}$ and $J = \{j_1, \ldots, j_l\}$ such that $a_i \neq b_i$ for all $i \in I$ and $b_j \neq c_j$ for all $j \in J$. We then have that $a_i = b_i$ whenever $i \notin I$ and $b_j = c_j$ whenever $j \notin J$. Consequently, $a_i = c_i$ whenever $i \notin I \cup J$ and thus, $d(a, c) \leq k + l$. □

Definition 3.5. *The* minimum distance *of a nontrivial (nonconstant) block code K is*

$$d(K) = \min\{d(a, b) \mid a \text{ and } b \text{ are code words and } a \neq b\}.$$

We now introduce the notions of error detection and correction as well as some results to illustrate how minimum distance is used to measure a code's ability to detect or correct errors.

Definition 3.6. *A block code K is said to detect t errors if for all distinct code words a and b,*

$$d(a, b) > t.$$

It then follows that a nontrivial block code K detects t errors if and only if $d(K) > t$.

Definition 3.7. *A block code K is said to correct t errors if for all distinct code words a and b and any word c in the code alphabet satisfying $1 \leq d(a, c) \leq t$, we have that $d(a, c) < d(b, c)$.*

Proposition 3.3. *A nontrivial block code K corrects t errors if and only if $d(K) > 2t$.*

Proof. Suppose that K is of length n and that K corrects t errors. Assume to the contrary that $d(K) \leq 2t$. Let a and b be code words with $d(a, b) = d(K) = k$. Then, there are indices $I = \{i_1, \ldots, i_k\}$ such that $a_i \neq b_i$ for all $i \in I$. Define a word $c = c_1 \cdots c_n$ in the code alphabet by setting

$$c_i = \begin{cases} a_i & \text{if } i = i_{2j} \text{ for some } 1 \leq j < k \\ b_i & \text{if } i \neq i_{2j} \text{ for all } 1 \leq j < k \end{cases}$$

for each $1 \leq i \leq n$. We then have that $d(a, c) = \lfloor \frac{k+1}{2} \rfloor \leq \frac{k+1}{2} \leq t + \frac{1}{2} \implies d(a, c) \leq t$. However, $d(b, c) = \lfloor \frac{k}{2} \rfloor \leq d(a, c)$, a contradiction. Suppose now that $d(K) > 2t$. Let a and b be distinct code words and suppose c is a word in the code alphabet such that $1 \leq d(a,c) \leq t$. We then have that

$$2t < d(a, b) \leq d(a, c) + d(c, b)$$

and hence,

$$d(b, c) > 2t - d(a, c) \geq d(a, c).$$

□

3.1.2 Parity Check Matrices of Binary Linear Codes

Many important binary codes can be described by a system of linear equations over $\mathbb{F}_2$. We now turn our attention to this class of codes.

Formally, a block code K of length n is called a *binary linear code* if im $K \subset \mathbb{F}_2^n$ is a subspace. However, this definition is quite tedious. Since K maps bijectively onto its image, elements of the source alphabet can be identified with the vectors lying in im K. It is therefore convenient to instead adopt the common definition of a binary linear code as being a subspace $K \subset \mathbb{F}_2^n$. It should then be understood that in this setting, code words correspond to elements of K.

Definition 3.8. *The* Hamming weight *of a word $a = a_1 \cdots a_n$ in $\mathbb{F}_2$ is*

$$w(a) = |\{i \in \{1, \ldots, n\} \mid a_i \neq 0\}|.$$

The minimum weight of a binary linear code K is

$$w(K) = \min\{w(x) \mid x \in K, x \neq 0\}.$$

Proposition 3.4. *If K is a binary linear code, then $d(K) = w(K)$.*

Proof. Suppose that a is a code word of Hamming weight $w(K)$. Then, $w(K) = d(a,0) \geq d(K)$. Now, choose code words a and b with $d(a, b) = d(K)$. Then, $d(a, b) = w(a + b) \leq w(K)$. □

Corollary 3.5. *A binary linear code K detects t errors if and only if $w(K) > t$. K corrects t errors if and only if $w(K) > 2t$.*

We now introduce the notion of a parity check matrix, which will then lead us to arguably the most important class of binary linear codes, namely, Hamming codes.

Definition 3.9. *A matrix H over $\mathbb{F}_2$ is called a* parity check matrix *for a binary linear code K if K coincides with the set of solutions to the homogeneous system $Hx = 0$.*

Proposition 3.6. *A binary linear code K corrects single errors if and only if every parity check matrix of K has nonzero, pairwise distinct columns.*

Proof. Let $e_i \in \mathbb{F}_2^n$ be the i-th standard basis vector. Suppose that K corrects single errors, i.e., $w(K) > 2$. If H is a parity check matrix for K with a zero column, say the i-th column, then $He_i = 0$ and so e_i is a code word which contradicts $w(K) > 2$. Moreover, if $i \neq j$ are such that the i-th and j-th columns of H are the same, then $H(e_i + e_j) = 0$. Therefore, $e_i + e_j$ is a code word which again contradicts $w(K) > 2$. Suppose then that every parity check matrix H of K has nonzero, pairwise distinct columns. Then for all i, $j \in \{1, \dots, n\}$, $i \neq j$, both e_i and $e_i + e_j$ are not code words. Thus, every code word a satisfies $w(a) > 2 \implies w(K) > 2$. □

Definition 3.10. *A binary linear code is called a Hamming code if for some $m \in \mathbb{N}$, it has an $m \times n$ parity check matrix H such that $2^m - 1 = n$ and each $v \in \mathbb{F}_2^m \setminus \{0\}$ is a column of H.*

Consider the case when $m = 3$. To construct a Hamming code of length 7, we need a 3×7 parity check matrix H whose columns consist of all nonzero $v \in \mathbb{F}_2^3$. Note that the choice of such a matrix is not unique since applying any elementary row operation to H or permuting any of the columns of H would yield an appropriate matrix. Here, we use the binary expansion of the integers $1, \dots, 7$ for columns $1, \dots, 7$ of H. This gives

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

Thus, our Hamming code is precisely the solution set to $Hx = 0$. This Hamming code is called the *Hamming* (7, 4)*-code.*

For each code word $x_1 \cdots x_n$ in a binary linear code K, choose $x_{n+1} \in \mathbb{F}_2$ such that

$$\sum_{i=1}^{n+1} x_i = 0.$$

The set K' of all words of the form $x_1 \cdots x_{n+1}$ is again a binary linear code (of length $n + 1$) called the *extension of K*. Note then that the extension of the Hamming (7, 4)-code has a parity check matrix given by

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

The set of solutions to $Hx = 0$ is called the *extended Hamming* (8, 4)*-code.*

Definition 3.11. *Let K be a binary linear code and suppose $v_1, \dots, v_k$ form a basis for K. The matrix*

$$G = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$$

whose rows are comprised of the vectors $v_1, \dots, v_k$ is called a generator matrix *for K.*

For example, the Hamming (7, 4)-code has

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

as a parity check matrix. H has rank 3 and so the nullity of H is $7-3=4$. A generator matrix for this code is given by

$$G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

3.2 Dual Codes of Binary Linear Codes

We focus now on the case when K is a binary linear code of an inner product space, i.e., K is a subspace of an inner product space $(\mathbb{F}_2^n,\, B)$.

3.2.1 MacWilliams' Identity

We assume for the rest of this section that K is a binary linear code of an inner product space $(\mathbb{F}_2^n,\, B)$. This allows us to define the notion of a dual code. Note that thanks to Proposition 1.8 and the fact that we are working in characteristic 2, we may assume that B is symmetric.

Definition 3.12. *The dual code $K^\perp$ of K is the orthogonal complement of K in $\mathbb{F}_2^n$. That is,*

$$K^\perp = \{v \in \mathbb{F}_2^n \mid \text{for any } w \in\, K,\ B(v,\, w) = 0\}.$$

Recalling that the orthogonal complement in V of any subspace of V is again a subspace, we see that the dual code is again a binary linear code.

For $0 \le i \le n$, define $A_i = |\{x \in K \mid w(x) = i\}|$. We call the sequence $(A_0, \dots, A_n)$ the *weight distribution* of K and the polynomial

$$W_K(x,y) = \sum_{i=0}^{n} A_i x^i y^{n-i}$$

the *weight enumerator* of K. Note that if $A_K(x) = \sum_{i=0}^{n} A_i x^i$, then $W_K(x,y) = y^n A_K(\frac{x}{y})$ and $A_K(x) = W_K(x,1)$ so either polynomial can be recovered from the other.

We now provide an important result in coding theory which shows that the weight enumerator of $K^\perp$ is completely determined by the weight enumerator of K in the case when the inner product is given by the ordinary dot product. We must first establish the following lemma and introduce some notation. In what follows, $u \cdot v$ denotes the dot product of vectors $u = u_1 \cdots u_n$ and $v = v_1 \cdots v_n$ in $\mathbb{F}_2^n$ defined by

$$u \cdot v = \sum_{i=1}^{n} u_i v_i.$$

Furthermore, $\|a\|$ denotes the image of $a \in \mathbb{F}_2$ under the mapping

$$\mathbb{F}_2 \to \mathbb{Z},\ 0 \mapsto 0,\ 1 \mapsto 1.$$

Lemma 3.7. *For any $v \in K$ and any $w \in \mathbb{F}_2^n$,*

$$\frac{1}{|K|}\sum_{v\in K}(-1)^{\|v\cdot w\|} = \begin{cases} 1 & \text{if } w \in K^\perp \\ 0 & \text{if } w \notin K^\perp \end{cases}.$$

Proof. If $w \in K^\perp$, then $\|v \cdot w\| = 0$ for all $v \in K$ and so

$$\sum_{v\in K}(-1)^{\|v\cdot w\|} = |K|.$$

Suppose then that $w \notin K^\perp$, say $v \in K$ satisfies $v \cdot w = 1$. Let $v_1, \dots, v_k$ be all code words of K orthogonal to w. Then, $(v_i + v) \cdot w = 1$ and if $u \in K$ is any other code word such that $u \cdot w = 1$, then $(u + v) \cdot w = 0 \implies u + v = v_i$ for exactly one $i \in \{1, \dots, k\}$. Thus,

$$\sum_{v\in K}(-1)^{\|v\cdot w\|} = \sum_{i=1}^{k}(-1)^{\|v_i\cdot w\|} + \sum_{i=1}^{k}(-1)^{\|(v_i+v)\cdot w\|} = 0.$$

□

Theorem 3.8 (MacWilliams' Identity)**.** *The weight enumerator of the dual code $K^\perp$ satisfies the following identity*

$$W_{K^\perp}(x, y) = \frac{1}{|K|}W_K(y - x, y + x).$$

Proof. Note that the weight enumerator $W_K(x, y)$ can be written in terms of Hamming weights as

$$W_K(x, y) = \sum_{i=0}^{n} A_i x^i y^{n-i} = \sum_{v\in K} x^{w(v)}y^{n-w(v)}.$$

Consequently, we have

$$\begin{aligned} W_K(y - x, y + x) &= \sum_{v\in K}(y - x)^{w(v)}(y + x)^{n-w(v)} \\ &= \sum_{v\in K}\prod_{i=1}^{n}[\, y + (-1)^{\|v_i\|}x\,], \end{aligned}$$

where each $v \in K$ is written as $v_1 \cdots v_n$. Now, by Lemma 3.7, the weight enumerator $W_{K^\perp}(x, y)$ can be written as

$$\begin{aligned} W_{K^\perp}(x, y) &= \sum_{v\in K^\perp} x^{w(v)}y^{n-w(v)} \\ &= \sum_{v\in\mathbb{F}_2^n}\left(\frac{1}{|K|}\sum_{u\in K}(-1)^{\|u\cdot v\|}\right)x^{w(v)}y^{n-w(v)} \\ &= \frac{1}{|K|}\sum_{u\in K}\sum_{v\in\mathbb{F}_2^n}(-1)^{\|u\cdot v\|}x^{w(v)}y^{n-w(v)}. \end{aligned}$$

We now use induction on n to show that, for a given $u = u_1 \cdots u_n \in \mathbb{F}_2^n$,

$$\sum_{v\in\mathbb{F}_2^n}(-1)^{\|u\cdot v\|}x^{w(v)}y^{n-w(v)} = \prod_{i=1}^{n}[\, y + (-1)^{\|u_i\|}x\,],$$

completing our proof of the Theorem. If $n = 1$, then

$$\sum_{v \in \{0,1\}} (-1)^{\|u \cdot v\|} x^{w(v)} y^{n-w(v)} = y + (-1)^{\|u\|} x.$$

Suppose that the equation holds for $\mathbb{F}_2^{n-1}$. Let $A = \{v \in \mathbb{F}_2^n \mid v_n = 0\}$, A^c be the complement of A in $\mathbb{F}_2^n$, and $u' = u_1 \cdots u_{n-1}$. Then,

$$\begin{aligned}
\sum_{v \in \mathbb{F}_2^n} (-1)^{\|u \cdot v\|} x^{w(v)} y^{n-w(v)} &= \sum_{v \in A} (-1)^{\|u \cdot v\|} x^{w(v)} y^{n-w(v)} + \sum_{v \in A^c} (-1)^{\|u \cdot v\|} x^{w(v)} y^{n-w(v)} \\
&= \sum_{v \in \mathbb{F}_2^{n-1}} (-1)^{\|u' \cdot v\|} x^{w(v)} y^{n-w(v)} + \sum_{v \in \mathbb{F}_2^{n-1}} (-1)^{\|u' \cdot v\|} (-1)^{\|u_n\|} x^{w(v)+1} y^{n-(w(v)+1)} \\
&= y \sum_{v \in \mathbb{F}_2^{n-1}} (-1)^{\|u' \cdot v\|} x^{w(v)} y^{(n-1)-w(v)} + (-1)^{\|u_n\|}\, x \sum_{v \in \mathbb{F}_2^{n-1}} (-1)^{\|u' \cdot v\|} x^{w(v)} y^{(n-1)-w(v)} \\
&= [\, y + (-1)^{\|u_n\|} x\,] \prod_{i=1}^{n-1} [\, y + (-1)^{\|u_i\|} x\,] = \prod_{i=1}^{n} [\, y + (-1)^{\|u_i\|} x\,].
\end{aligned}$$

□

Note that our proof of Theorem 3.8 relies heavily on the definition of the dot product. In particular, we make use of the fact that for vectors $u = u_1 \cdots u_n$ and $v = v_1 \cdots v_n$ in $\mathbb{F}_2^n$,

$$(-1)^{\|u \cdot v\|} = \prod_{i=1}^{n} (-1)^{\|u_i v_i\|},$$

noting that the above equation holds despite the fact that $\| \cdot \|$ is not additive. A natural question is whether MacWilliams' Identity holds for more general inner products.

If we define the dual of a code K with respect to an arbitrary nondegenerate bilinear form (not necessarily symmetric) to be $\perp_R (K)$, then it is claimed in [13, Theorem 11] that MacWilliams' Identity holds for any nondegenerate bilinear form. However, this is, unfortunately, not true as pointed out in [5, Example 28]. The counterexample constructed there is as follows. Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in GL(\mathbb{F}_2^2)$$

and define a bilinear form $B : \mathbb{F}_2^2 \times \mathbb{F}_2^2 \to \mathbb{F}_2$ by $B(u,\, v) = u^T A v$ where $u, v \in \mathbb{F}_2^2$ are written as column vectors. Consider the code K generated by the first standard basis vector e_1, i.e., $K = \langle e_1 \rangle$. Then, $W_K(x, y) = y^2 + xy$. The dual code $K^\perp$ of K with respect to B is then the subspace generated by $e_1 + e_2$. Thus, $W_{K^\perp}(x, y) = x^2 + y^2$. We compute

$$\frac{1}{|K|} W_K(y - x, y + x) = \frac{1}{2}[(y + x)^2 + (y - x)(y + x)] = y^2 + xy.$$

Therefore, MacWilliams' Identity fails in this case and hence we cannot say that it holds in general for nondegenerate bilinear forms.

Notice that in the above example, B is not symmetric. The next question to ask is whether MacWilliams' Identity holds for inner products, i.e., nondegenerate symmetric forms or nondegenerate alternating forms. The answer to this question is again, unfortunately, no. We now construct a counterexample. We will make use of the fact that if A is an invertible $n \times n$ matrix over a field $\mathbb{F}$, $u, v \in \mathbb{F}^n$ are column vectors, and $\operatorname{adj} A$ denotes the *adjugate matrix of* A, then

$$\det(A + uv^T) = \det A + v^T(\operatorname{adj} A)u$$

which may be called *Cauchy's formula for the determinant of a rank-one pertubation* according to [7, p.26].

Given $k \in \mathbb{Z}$, let $[k]$ denote the image of k under the quotient map $\mathbb{Z} \to \mathbb{F}_2$. Define $\omega : \mathbb{F}_2^n \times \mathbb{F}_2^n \to \mathbb{F}_2$ by

$$\omega(u, v) = \sum_{i=1}^{n} u_i v_i + \left(\sum_{i=1}^{n} u_i\right)\left(\sum_{i=1}^{n} v_i\right) = u \cdot v + [w(u)][w(v)],$$

for all $u = u_1 \cdots u_n$ and $v = v_1 \cdots v_n$ in $\mathbb{F}_2^n$. We first show that ω is alternating and nondegenerate whenever n is even. To show that ω is bilinear, it suffices to show that for all u, $v \in \mathbb{F}_2^n$, $[w(u + v)] = [w(u)] + [w(v)]$. Write $u = u_1 \cdots u_n$, $v = v_1 \cdots v_n$, and let

$$\begin{aligned} k &= |\{i \in \{1, \dots, n\} \mid u_i = 1 = v_i\}|, \\ l &= |\{i \in \{1, \dots, n\} \mid u_i = 0 = v_i\}|. \end{aligned}$$

Observe that $n = w(u) + w(v) - k + l$. Then, $w(u + v) = n - k - l = w(u) + w(v) - 2k$ and the result follows.

We now show that ω is alternating and nondegenerate for even n. For any $v \in \mathbb{F}_2^n$, $\omega(v, v) = [w(v)] + [w(v)] = 0$ which shows that ω is alternating. For $1 \leq i \leq n$, let e_i be the i-th standard basis vector. Then, $\omega(e_i, e_j) = \delta_{ij} + 1$. Therefore, the matrix representing ω relative to $\{e_i\}_{i=1}^n$ has the form $\widehat{\omega} = I_n + ee^T$, where I_n is the $n \times n$ identity matrix and $e = e_1 + \cdots + e_n$ is the word of all 1's. Thus, by Cauchy's formula for the determinant of a rank-one perturbation,

$$\begin{aligned} \det \widehat{\omega} &= \det\left(I_n + ee^T\right) \\ &= \det\, I_n + e^T I_n e \\ &= 1 + [w(e)]. \end{aligned}$$

Hence, ω is nondegenerate whenever n is even.

Before constructing our counterexample, note that if $1 \leq i \leq n$ is fixed, then any $v \in \mathbb{F}_2^n$ can be written in the form

$$v = \lambda e_i + \sum_{j=1}^{k} e_{i_j},$$

where $\lambda \in \mathbb{F}_2$ and $\{e_{i_j}\}_{j=1}^k$ are pairwise distinct with $i \neq i_j$ for all $1 \leq j \leq k$. We then have that

$$\omega(e_i, v) = \lambda\omega(e_i, e_i) + \sum_{j=1}^{k} \omega(e_i, e_{i_j}) = \begin{cases} 1 & \text{if } k \text{ is odd} \\ 0 & \text{if } k \text{ is even} \end{cases}.$$

We now begin the construction. Consider the code $K \subset \mathbb{F}_2^4$ generated by e_1. By the above remark, we may easily compute the dual code

$$K^\perp = \{0,\ e_1,\ e_1 + e_2 + e_3,\ e_1 + e_2 + e_4,\ e_1 + e_3 + e_4,\ e_2 + e_3,\ e_2 + e_4,\ e_3 + e_4\}.$$

Now, the weight enumerator for K is $W_K(x,y) = y^4 + xy^3$ and the weight enumerator for $K^\perp$ is $W_{K^\perp}(x,y) = y^4 + xy^3 + 3x^3y + 3x^2y^2$. However, we find that

$$\begin{aligned}\frac{1}{|K|}W_K(y-x, y+x) &= \frac{1}{2}[(y+x)^4 + (y-x)(y+x)^3] \\ &= \frac{1}{2}[2y^4 + 6xy^3 + 2x^3y + 6x^2y^2] \\ &= y^4 + 3xy^3 + x^3y + 3x^2y^2 \neq W_{K^\perp}(x,y).\end{aligned}$$

Therefore, MacWilliams' identity fails in this case and hence we cannot say that it holds in general for inner products.

3.2.2 Self-Dual Binary Linear Codes

We have introduced the dual code of a binary linear code. We now introduce a class of binary linear codes induced by this notion, namely, self-dual codes.

Definition 3.13. *A binary linear code K is called self-dual if $K = K^\perp$.*

It should be noted that if K is self-dual, then by Proposition 1.2, we have

$$\dim\, V = \dim\, K + \dim\, K^\perp$$

and so V has even dimension and $\dim\, K = \frac{1}{2}n$.

We may also wish to note that if K is self-dual with respect to the dot product, then Theorem 3.8 gives

$$W_K(x,y) = \frac{1}{|K|}W_K(y-x, y+x).$$

Many important codes are self-dual. For example, if we equip $\mathbb{F}_2^8$ with the dot product, then the extended Hamming (8, 4)-code is self-dual. To see this, let K denote the extended Hamming (8, 4)-code and recall that a parity check matrix for K is given by

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

and hence a generator matrix for K is given by

$$G = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Note that in the case of the dot product, a generator matrix for a code is a parity check matrix for its dual. Thus, G is a parity check matrix for $K^\perp$ and an easy computation shows that the nullity of G is 4 and that $GG^T = 0 \implies K^\perp = K$.

If one wishes to classify self-dual codes, then there should be a notion of equivalence of codes. Although there are many ways to formulate what it means for two codes to be equivalent, we only introduce a standard one.

Let $K \subset \mathbb{F}_2^n$ be a binary linear code. The symmetric group S_n acts on $\mathbb{F}_2^n$ by permuting coordinates, i.e., for $\sigma \in S_n$ and $v = v_1 \cdots v_n \in \mathbb{F}_2^n$,

$$\sigma \cdot v = v_{\sigma^{-1}(1)} \cdots v_{\sigma^{-1}(n)}.$$

If K' is another binary linear code, then we say K and K' are *equivalent* if there exists a $\sigma \in S_n$ such that $\sigma \cdot K = K'$. In fact, if B is an S_n-invariant inner product on $\mathbb{F}_2^n$, i.e., $B(\sigma \cdot u, \sigma \cdot v) = B(u, v)$ for any $\sigma \in S_n$, then S_n acts on the set C of all self-dual codes. This is because $v \in K^\perp \iff \sigma \cdot v \in \sigma \cdot K^\perp$ and hence, $\sigma \cdot K^\perp = (\sigma \cdot K)^\perp$.

Given a binary linear code K, define the *automorphism group of* K to be the subgroup of S_n given by

$$\operatorname{Aut}(K) = \{\sigma \in S_n \mid \sigma \cdot K = K\}.$$

We then see that the number of codes equivalent to K is

$$\frac{|S_n|}{|\operatorname{Aut}(K)|} = \frac{n!}{|\operatorname{Aut}(K)|}.$$

Hence, if C is the collection of self-dual codes that are pairwise inequivalent, then the number of distinct self-dual codes is

$$T_n = \sum_{K \in C} \frac{n!}{|\operatorname{Aut}(K)|}.$$

We then obtain the following formula which is known as a *mass formula*

$$\frac{T_n}{n!} = \sum_{K \in C} \frac{1}{|\operatorname{Aut}(K)|}. \tag{$*$}$$

If we can determine T_n, then the left-hand side of equation $(*)$ gives us a stopping condition for an algorithm to find inequivalent self-dual codes. We will compute T_n in the following section for symplectic forms. It should first, however, be noted that in the case of the dot product,

$$T_n = \prod_{i=1}^{\frac{1}{2}n-1} (2^i + 1)$$

as shown in [9, Corollary 19 p. 630]. We provide an alternative proof of this in the next section.

3.3 Lagrangian Binary Self-Dual Codes

In this final section, we will consider binary self-dual codes of a symplectic space, i.e., codes $K \subset \mathbb{F}_2^n$ satisfying $K = K^\perp$ with respect to a nondegenerate alternating form on $\mathbb{F}_2^n$. Such codes are examples of *Lagrangians*, i.e., subspaces of a symplectic space that coincide with their orthogonal complement. We first give a gentle introduction to Lagrangian subspaces of a general finite-dimensional symplectic space (V, ω).

The following proposition concerns the existence of Lagrangians and gives a characterization of such subspaces. In what follows, V is a finite-dimensional symplectic space. In the spirit of symmetric bilinear forms, we call a subspace $W \subset V$ satisfying $W \subset W^\perp$ a *totally isotropic subspace*.

Remark 3.3. *Although we have already seen that a symplectic space has even dimension, the following proposition provides an alternative way to see this. Given that V has a Lagrangian L, we have*

$$\dim V = \dim L + \dim L^{\perp} = 2 \dim L.$$

Proposition 3.9. *A subspace of V is Lagrangian if and only if it is totally isotropic and maximal (in the sense of inclusion). In particular, there exists a Lagrangian subspace of V.*

Proof. Choose a maximal totally isotropic subspace L of V. Suppose to the contrary that $L \neq L^{\perp}$. Then, there exists $v \in L^{\perp} \setminus L \implies L \oplus \langle v \rangle$ is a totally isotropic subspace containing L, contradicting the maximality of L. Conversely, suppose L is a Lagrangian subspace of V and that $M \subset V$ is a totally isotropic subspace containing L. Then,

$$L \subset M \implies M^{\perp} \subset L^{\perp} = L \subset M.$$

Thus, $M = M^{\perp}$, i.e., M is Lagrangian. But then $\dim M = \dim L$ by the remark above and hence, $M = L$. □

We denote the set of all Lagrangian subspaces of V by $\mathrm{Lag}(V)$. After establishing the following lemma, we will compute $|\mathrm{Lag}(\mathbb{F}_q^{2n})|$ for any prime power q. In particular, taking $q = 2$ tells us the number of binary self-dual codes of $(\mathbb{F}_2^{2n},\, \omega)$.

Lemma 3.10. *The group $Sp(V)$ acts on $\mathrm{Lag}(V)$ naturally, i.e., for $\tau \in Sp(V)$ and $L \in \mathrm{Lag}(V)$, $\tau \cdot L = \tau(L) \in \mathrm{Lag}(V)$. Moreover, this group action is transitive.*

Proof. Clearly, the mapping satisfies the axioms of a group action. We show that if $\tau \in Sp(V)$ and $L \in \mathrm{Lag}(V)$, then $\tau(L) \in \mathrm{Lag}(V)$. If $\tau v,\, \tau u \in \tau(L)$, then $\omega(\tau v,\, \tau u) = \omega(v,\, u) = 0$ since $u, v \in L = L^{\perp}$. Conversely, if $v \in \tau(L)^{\perp}$, then $v = \tau u$ for some $u \in V$. Let $\tau w \in \tau(L)$. Then, $\omega(v,\, \tau w) = 0 = \omega(u, w) \implies u \in L^{\perp} = L \implies v \in \tau(L)$. We now proceed to show that $Sp(V)$ acts transitively. As previously noted, dim V is even, say $2n$. Then, any $L_1, L_2 \in \mathrm{Lag}(V)$ are of dimension n and are hence isomorphic, say via $\sigma : L_1 \to L_2$. Since $L_i^{\perp} = L_i$, $i = 1, 2$, we have $u, v \in L_1 \implies \omega(u,\, v) = 0 = \omega(\sigma u,\, \sigma v)$ and so in fact σ is an isometry. Thus, by Witt's Extension Theorem 2.24 (see also Remark 2.4), there exists $\tau \in Sp(V)$ such that $\tau|_{L_1} = \sigma \implies \tau \cdot L_1 = L_2$. □

Theorem 3.11. *Let $V = \mathbb{F}_q^{2n}$ for any prime power q. Then, $|\mathrm{Lag}(V)| = \prod_{i=1}^{n}(q^i + 1)$.*

Proof. By Theorem 1.6, there is a basis $\{u_1, \ldots, u_n, v_1, \ldots, v_n\}$ of V such that

$$\omega(u_i,\, v_j) = \delta_{ij} = -\omega(v_j,\, u_i),$$
$$\omega(u_i,\, u_j) = 0 = \omega(v_i,\, v_j),$$

and thus, the matrix representing ω relative to this basis is

$$\widehat{\omega} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Let e_i be the i-th standard basis vector and set $L = \langle e_1, \ldots e_n \rangle$. Then, L is Lagrangian as $e_i^T \,\widehat{\omega}\, e_j = 0$ for all $i, j \in \{1, \ldots, n\}$. Let S be the stabilizer of L in $Sp(V)$. Since $Sp(V)$ acts transitively on $\mathrm{Lag}(V)$, we have

$$|\mathrm{Lag}(V)| = [\, Sp(V) : S\,] = \frac{|Sp(V)|}{|S|}.$$

Recall from Proposition 2.10 that a matrix $M \in GL(V)$ lies in $Sp(V) \iff M^T \,\widehat{\omega}\, M = \widehat{\omega}$. We may express M as a 2×2 block matrix whose entries are $n \times n$ matrices, say,

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Then, $M \in Sp(V)$ if and only if

$$\begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} \begin{bmatrix} C & D \\ -A & -B \end{bmatrix} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$$

which gives the following system of matrix equations:

$$\begin{aligned} A^T\, C - C^T\, A &= 0 \\ A^T\, D - C^T\, B &= I_n \\ B^T\, D - D^T\, B &= 0. \end{aligned}$$

If we impose further that $M \in S$, then $Me_i \in L$ for all $1 \leq i \leq n \implies C = 0$. Therefore, $D = (A^T)^{-1}$ and $B = AE$, where $E = B^T\, D$. Hence, we see that S is of the form

$$S = \left\{ \begin{bmatrix} A & AE \\ 0 & (A^T)^{-1} \end{bmatrix} \,\middle|\, A \in GL(n,\, q),\, E^T = E \right\}.$$

Notice that if E is any $n \times n$ symmetric matrix and $A \in GL(n,\, q)$, then

$$\begin{bmatrix} A & 0 \\ 0 & (A^T)^{-1} \end{bmatrix} \begin{bmatrix} I_n & E \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} A & AE \\ 0 & (A^T)^{-1} \end{bmatrix}.$$

Thus, if we let

$$G = \left\{ \begin{bmatrix} A & 0 \\ 0 & (A^T)^{-1} \end{bmatrix} \,\middle|\, A \in GL(n,\, q) \right\},\; H = \left\{ \begin{bmatrix} I_n & E \\ 0 & I_n \end{bmatrix} \,\middle|\, E^T = E \right\},$$

we have that S decomposes into an inner semidirect product $S = G \rtimes H$. Now, the order of H is the number of $n \times n$ symmetric matrices over $\mathbb{F}_q$. There are q^i choices for the i-th row of such a matrix and so,

$$|H| = \prod_{i=1}^{n} q^i = q^{\frac{n(n+1)}{2}}.$$

Furthermore, $G \cong GL(n,\, q)$ and so, by Proposition 2.1,

$$|G| = q^{\frac{(n-1)n}{2}} \prod_{i=1}^{n} (q^i - 1).$$

Thus, $|S| = |G| \cdot |H|$ and so the result follows from Theorem 2.17. □

Corollary 3.12. *If $V = \mathbb{F}_2^n$ is equipped with the dot product, then the number of binary self-dual codes of V is*

$$T_n = \prod_{i=1}^{\frac{1}{2}n-1} (2^i + 1).$$

Proof. Denote the dot product of two vectors $u, v \in V$ by $u \cdot v$ and for $k \in \mathbb{Z}$, write $[k]$ for the image of k under the quotient map $\mathbb{Z} \to \mathbb{F}_2$. Any subspace of V that is self-dual with respect to the dot product must consist entirely of vectors $v \in V$ with $[w(v)] = 0$, i.e., of even Hamming weight. Otherwise, $v \cdot v = 1$ and the space would not be self-dual. Moreover, we may assume that n is even, otherwise no subspace of V would be self-dual. Hence, the alternating form ω on V defined by $\omega(u,\, v) = u \cdot v + [w(u)][w(v)]$ is nondegenerate. Let e_i be the i-th standard basis vector and set $e = e_1 + \cdots + e_n$, i.e., the word consisting of all 1's. Then, the space of even Hamming weight vectors is given by $\langle e \rangle^\perp$ and, by the nondegeneracy of ω, we have $(\langle e \rangle^\perp)^\perp = \langle e \rangle = \operatorname{rad} \langle e \rangle^\perp$. Then, $\langle e \rangle^\perp / \langle e \rangle$ is a symplectic space with respect to the induced symplectic form $\bar{\omega}$ defined by $\bar{\omega}(\bar{u},\, \bar{v}) = \omega(u,\, v)$, where $\bar{u}, \bar{v} \in \langle e \rangle^\perp / \langle e \rangle$ are the cosets with representatives u and v respectively. Thus, the number of binary self-dual codes of V is the number of Lagrangians of $\langle e \rangle^\perp / \langle e \rangle$. Since $\dim(\langle e \rangle^\perp / \langle e \rangle) = n - 2$, the result follows from Theorem 3.11. □

In terms of practicality, we would like a way to construct the basis appearing in the proof of Theorem 3.11. This can be achieved recursively in a process which may be called the symplectic Gram-Schmidt process. It goes as follows. Choose a nonzero $u_1 \in V$ and find $v_1 \in V$ with $\omega(u_1,\, v_1) = 1$. Set $W_1 = \langle u_1,\, v_1 \rangle$. Then, $V = W_1 \oplus W_1^\perp$. Choose a nonzero $u_2 \in W_1^\perp$ and find $v_2 \in W_1^\perp$ with $\omega(u_2,\, v_2) = 1$. Set $W_2 = \langle u_2,\, v_2 \rangle$. Then, $V = W_1 \oplus W_2 \oplus W_2^\perp$. Iterating this process, we find that $V = W_1 \oplus \cdots \oplus W_n$, where each $W_k = \langle u_k,\, v_k \rangle$ is a hyperbolic pair and hence, $\{u_1, v_1, \ldots, u_n, v_n\}$ is a symplectic basis for V.

We observed in the proof of Theorem 3.11 that the subspace $L = \langle e_1, \ldots, e_n \rangle$ generated by the standard basis vectors $\{e_i\}_{i=1}^n$ is Lagrangian. Given that $Sp(V)$ acts transitively on $\operatorname{Lag}(V)$, we may determine all Lagrangian subspaces of V by computing the orbit of L. If $V = \mathbb{F}_2^{2n}$, then once $\operatorname{Lag}(V)$ has been computed, we may determine the classes of inequivalent Lagrangians by computing the S_{2n}-orbit of $\operatorname{Lag}(V)$. This can be accomplished through GAP [4] using a naive algorithm. Let us demonstrate this for the case when ω is the inner product on $\mathbb{F}_2^{2n}$ defined by $\omega(u, v) = u \cdot v + [w(u)][w(v)]$.

First, we construct a symplectic basis $\{f_i\}_{i=1}^{2n}$ for V as follows. Let $f_1 = e_1$ and $f_2 = e_2$ be the standard basis vectors. Then, $\omega(f_1, f_2) = 1$. For $k \geq 3$ define

$$f_k = \begin{cases} e_k + \sum_{i=1}^{k-1} f_i & \text{if } k \text{ is odd} \\ e_k + \sum_{i=1}^{k-1} (f_i + e_i) & \text{if } k \text{ is even} \end{cases}.$$

Note that if $k > 2$ is odd, then

$$\begin{aligned} f_k &= e_k + \sum_{i=1}^{k-1} f_i \\ &= e_k + f_{k-1} + \sum_{i=1}^{k-2} f_i \\ &= e_k + e_{k-1} + \sum_{i=1}^{k-2} (f_i + e_i) + \sum_{i=1}^{k-2} f_i \\ &= \sum_{i=1}^{k} e_i. \end{aligned}$$

Furthermore, if $k > 2$ is even, then

$$\begin{aligned}
f_k &= e_k + \sum_{i=1}^{k-1}(f_i + e_i) \\
&= e_k + f_{k-1} + \sum_{i=1}^{k-2} f_i + \sum_{i=1}^{k-1} e_i \\
&= e_k + e_{k-1} + 2\sum_{i=1}^{k-2} f_i + \sum_{i=1}^{k-1} e_i \\
&= e_{k-1} + \sum_{i=1}^{k} e_i.
\end{aligned}$$

In fact, the above formulas hold when $k = 1, 2$ as well by the definition of f_1 and f_2. Note also that $\langle f_1, \dots, f_{2n}\rangle = \langle e_1, \dots, e_{2n}\rangle$ so $\{f_i\}_{i=1}^{2n}$ is in fact a basis. We now show that if $k > 2$ is odd and $m < k$ is odd, then $f_k, f_{k+1} \in \langle f_m, f_{m+1}\rangle^\perp$ and that $\omega(f_k, f_{k+1}) = 1$. Indeed,

$$\omega(f_k, f_m) = \omega\left(\sum_{i=1}^{k} e_i,\ \sum_{i=1}^{m} e_i\right) = 1 + 1 = 0$$

and

$$\begin{aligned}
\omega(f_k, f_{m+1}) &= \omega\left(\sum_{i=1}^{k} e_i,\ e_m + \sum_{i=1}^{m+1} e_i\right) \\
&= \omega\left(\sum_{i=1}^{k} e_i,\ e_m\right) + \omega\left(\sum_{i=1}^{k} e_i,\ \sum_{i=1}^{m+1} e_i\right) \\
&= (1 + 1) + (0 + 0) = 0
\end{aligned}$$

which shows that $f_k \in \langle f_m, f_{m+1}\rangle^\perp$. Now,

$$\begin{aligned}
\omega(f_{k+1}, f_m) &= \omega\left(e_k + \sum_{i=1}^{k+1} e_i,\ \sum_{i=1}^{m} e_i\right) \\
&= \omega\left(e_k,\ \sum_{i=1}^{m} e_i\right) + \omega\left(\sum_{i=1}^{k+1} e_i,\ \sum_{i=1}^{m} e_i\right) \\
&= (0 + 1) + (1 + 0) = 0
\end{aligned}$$

and

$$\begin{aligned}
\omega(f_{k+1}, f_{m+1}) &= \omega\left(e_k + \sum_{i=1}^{k+1} e_i,\ e_m + \sum_{i=1}^{m+1} e_i\right) \\
&= \omega(e_k, e_m) + \omega\left(e_k,\ \sum_{i=1}^{m+1} e_i\right) + \omega\left(\sum_{i=1}^{k+1} e_i,\ e_m\right) + \omega\left(\sum_{i=1}^{k+1} e_i,\ \sum_{i=1}^{m+1} e_i\right) \\
&= (0 + 1) + (0 + 0) + (1 + 0) + (0 + 0) = 0
\end{aligned}$$

which shows that $f_{k+1} \in \langle f_m, f_{m+1}\rangle^{\perp}$. Finally,

$$\begin{aligned}\omega(f_k, f_{k+1}) &= \omega\Big(\sum_{i=1}^{k} e_i,\, e_k + \sum_{i=1}^{k+1} e_i\Big)\\ &= \omega\Big(\sum_{i=1}^{k} e_i,\, e_k\Big) + \omega\Big(\sum_{i=1}^{k} e_i,\, \sum_{i=1}^{k+1} e_i\Big)\\ &= (1+1) + (1+0) = 1.\end{aligned}$$

Hence, $\{f_i\}_{i=1}^{2n}$ is a symplectic basis for V with respect to ω.

Now, as we saw in the proof of Theorem 3.11, ω is represented by the matrix

$$\widehat{\omega} = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$$

relative to the basis $\mathcal{F} = \{f_1, f_3, \ldots f_{2n-1}, f_2, f_4, \ldots, f_{2n}\}$. Therefore, if each f_i and e_i are represented by row vectors, then the change of basis matrix from $\mathcal{F}$ to $\mathcal{E} = \{e_i\}_{i=1}^{2n}$ given by

$$P = \begin{bmatrix} f_1 \\ f_3 \\ \vdots \\ f_{2n-1} \\ f_2 \\ f_4 \\ \vdots \\ f_{2n} \end{bmatrix}$$

respects the identity $\widehat{\omega}_{\mathcal{F}} = P\,\widehat{\omega}_{\mathcal{E}}\,P^T$. These are all the tools needed for our naive algorithm. Code written in GAP [4] (see A) produced the following matrices for $n = 1, 2, 3, 4$. The row spaces of these matrices are representatives of each of the inequivalent classes of Lagrangians.

$n = 1:$

$$\begin{bmatrix} 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$n = 2:$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

$n = 3:$

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

$n = 4:$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Bibliography

[1] Abraham, R. and Marsden, J. E., *Foundations of Mechanics*, 2nd edition. Addison-Wesley, 1978.

[2] Adámek, J., *Foundations of Coding.* Wiley-Interscience, 1991.

[3] Gallier, J., *Geometric Methods and Applications.* Springer, 2001.

[4] *GAP – Groups, Algorithms, and Programming, Version 4.13.0*, The GAP Group, 2024. [Online]. Available: `%5Curl%7Bhttps://www.gap-system.org%7D`.

[5] Gómez-Torrecillas, J., Hieta-aho, E., Lobillo, F. J., López-Permouth, S., and Navarro, G., *Some remarks on non projective frobenius algebras and linear codes*, 2019. arXiv: `1903.08410 [math.RA]`.

[6] Grove, L. C., *Classical Groups and Geometric Algebra.* American Mathematical Society, 2002.

[7] Horn, R. A. and Johnson, C. R., *Matrix Analysis*, 2nd edition. Cambridge University Press, 2013.

[8] Lam, T. Y., *Introduction to Quadratic Forms over Fields.* American Mathematical Society, 2005.

[9] MacWilliams, F. J. and Sloan, N. J. A., *The Theory of Error-Correcting Codes.* North-Holland, 1977.

[10] Pless, V. and Huffman, W., *Handbook of Coding Theory.* Elsevier, 1998.

[11] Rotman, J. J., *An Introduction to the Theory of Groups*, 4th edition. Springer, 1999.

[12] Silva, A. C. da, *Lectures on Symplectic Geometry.* Springer, 2001.

[13] Szabo, S. and Wood, J. A., *Properties of dual codes defined by nondegenerate forms*, Journal of Algebra Combinatorics Discrete Structures and Applications, 2017.

[14] Weyl, H., *The Classical Groups. Their Invariants and Representations*, 2nd edition. Princeton University Press, 1997.

[15] Witt, E., *Theorie der quadratischen formen in beliebigen körpern*, J. reine angew. Math. 176, 1937.

Appendix A

Code

The following code, written in GAP [4], is used to compute representatives of each of the inequivalent classes of Lagrangians of $\mathbb{F}_2^{2n}$ for small values of n.

```
V := GF(2)^(2*n);

J := NullMat(2*n, 2*n, GF(2));
for i in [1..n] do
    J[i][n + i] := Z(2)^0;
    J[n + i][i] := Z(2)^0;
od;
G := Sp(2 * n, 2, J);

B := Basis(V);
LB := [];
for k in [1..n] do
    Add(LB, B[k]);
od;
L := VectorSpace(GF(2), LB);
L_Orbit := Orbit(G,L);

f := [B[1], B[2]];
sumf := f[1] + f[2];
sume := B[1] + B[2];
for k in [3..2*n] do
    Add(f, B[k] + sumf + sume * ((k + 1) mod 2));
    sumf := sumf + f[k];
    sume := sume + B[k];
od;

P := NullMat(2*n, 2*n, GF(2));
for k in [1..n] do
    P[k] := f[2*k - 1];
    P[k + n] := f[2*k];
od;
```

```
LG := [];
for l in L_Orbit do
    mat := BasisVectors(Basis(l)) * P;
    Add(LG, VectorSpace(GF(2), mat));
od;

action := function(l, sigma);
    return VectorSpace(GF(2),
    BasisVectors(Basis(l)) * PermutationMat(sigma, 2*n, GF(2)));
end;

S_2n := SymmetricGroup(2*n);
gens := GeneratorsOfGroup(S_2n);
S_Orbit := OrbitsDomain(S_2n, LG, gens, gens, action);
rep := [];

for i in [1..Size(S_Orbit)] do
    Add(rep, BasisVectors(Basis(S_Orbit[i][1])));
od;
```