Senior Honors Thesis

in

Mathematics

The Structure of Orthogonal and Symplectic Geometries

Author:

Samir Chowdhury

Advisor:

George McNinch

Thesis Committee:

Richard Weiss

George McNinch

Tufts University

Medford, Massachusetts

May 3, 2013

Contents

Chapter 1

Nondegenerate Forms on Vector Spaces

Throughout this chapter (and the next), we will assume that our vector spaces are finite-dimensional. Once the theory of bilinear forms has been developed, we will further assume (unless mentioned otherwise) that our forms are nondegenerate. For technical reasons, we will also assume that our base field K has characteristic $\neq 2$. After proving Theorem 1.18, we will always assume that our forms are either symmetric or alternating.

1 Linear Algebra Review

We motivate our study of forms on vector spaces with a preliminary discussion of matrix groups. Let $\mathrm{M}_n(K)$ denote the set of $n \times n$ matrices with coefficients in K. The *identity matrix* is denoted I_n and consists of 1's on the diagonal and zeros everywhere else. Multiplication of matrices $A = (a_{ij})$, $B = (b_{ij})$, gives us another matrix in $\mathrm{M}_n(K)$, defined by:

$$AB = (\Sigma_{k=1}^{n} a_{ik} b_{kj})$$

Our first observation is that restricting to the set of *invertible* matrices gives us a group, called the *general linear group* and denoted $\mathrm{GL}_n(K)$. From linear algebra, we note that these are the matrices with *non-zero determinant.* We denote the determinant of a matrix $A \in \mathrm{M}_n(K)$ by det A and its transpose by A^{T}.

Properties of the determinant and transpose:

- $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$
- $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$
- $(cA)^{\mathrm{T}} = cA^{\mathrm{T}}$
- $\det\, A^{\mathrm{T}} = \det\, A$
- $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$
- $\vec{a} \cdot \vec{b} = (\vec{b})^{\mathrm{T}}(\vec{a})$
- $\det\,(AB) = (\det\, A)\ (\det\, B)$
- $\det\,(cA) = c^n\ \det(A)$
- $\det\,(A^{-1}) = \dfrac{1}{\det\, A}$

An important subset of $\mathrm{GL}_n(K)$ is the set of all matrices with $\det = 1$. This set contains the identity and is closed under products and taking inverses. This subgroup is called the *special linear group* and is denoted by $\mathrm{SL}_n(K)$.

2 Linear Maps and Dual Spaces

$\mathrm{GL}_n(K)$ has a number of other interesting subgroups, and we will eventually develop two of them. For now, we will take a detour into linear maps and dual spaces.

Given a field K, let V and W be vector spaces over K of dimensions n and m, respectively.

Definition 1.1. A *linear transformation* between V and W is a map $T : V \to W$ such that the following hold:

(i) $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all vectors $v_1, v_2 \in V$

(ii) $T(\alpha v) = \alpha T(v)$ for any $\alpha \in K$

The set of all linear transformations between V and W is denoted $\mathcal{L}(V, W)$.

Once we fix a basis for V, we can find a canonical correspondence between $\mathrm{M}_n(K)$ and $\mathcal{L}(V, V)$. More generally, if we fix bases for V and W, there is a vector space isomorphism between $\mathcal{L}(V, W)$ and $M_{m\times n}(K)$, ***with respect to the bases we fixed.*** ([1], Theorem 11.10)

Given two elements $\phi, \psi \in \mathcal{L}(V, V)$, we take the product $\phi\psi$ by function composition. Under this operation, the invertible elements of $\mathcal{L}(V, V)$ form a group that we denote by $GL(V)$ and call the *general linear group of V.* In particular, the subset consisting of maps with determinant 1 form a subgroup called the *special linear group of V*, and we denote this by $SL(V)$.

Note that a linear map $\phi \in SL(V)$ will have determinant 1 independent of the choice of basis. Suppose A is the matrix associated to ϕ with determinant 1. If we now fix a different basis and get a new matrix B, a fact from linear algebra tells us that we have $B = P^{-1}AP$ for some invertible *change of basis* matrix P ([1] p. 419). Well, that gives us:

$$\begin{aligned}\det\, P^{-1}AP &= (\det\, P^{-1})(\det\, A)(\det\, P)\\ &= (\det\, A)(\det\, P^{-1})(\det\, P)\\ &= (\det\, A)(\det\, P)^{-1}(\det\, P)\\ &= \det\, A = 1\end{aligned}$$

Now let ϕ and ψ be two elements of $\mathcal{L}(V, W)$. Notice that $\mathcal{L}(V, W)$ becomes a vector space over K when equipped with the following operations:

$$(\phi + \psi)(x) = \phi(x) + \psi(x)$$

$$(\alpha\phi)(x) = \alpha\phi(x)$$

In particular, $\mathcal{L}(V, K)$ is known as the *dual space* of V.

Definition 1.2. The *dual space of V*, denoted V^*, is the set of all linear maps $f : V \to K$.

3 Dual Bases

Constructing a basis for the dual space is not too difficult, and we will go through it carefully. First note that any linear transformation $f : V \to K$ is completely determined by where f sends the basis elements $(v_1, v_2, \dots v_n)$. Suppose we have $f(v_1) = a_1,\ f(v_2) = a_2,\ \cdots,\ f(v_n) = a_n$. For $x \in V$, let $x = c_1v_1 + c_2v_2 + \cdots + c_nv_n$. Then we have:

$$\begin{aligned} f(x) &= f(c_1v_1) + f(c_2v_2) + \cdots + f(c_nv_n) \\ &= c_1f(v_1) + c_2f(v_2) + \cdots + c_nf(v_n) \\ &= c_1a_1 + c_2a_2 + \cdots + c_na_n \end{aligned}$$

We define the dual basis $(v_1^*, v_2^*, \dots v_n^*)$ in the following way:

$$v_i^*(v_j) = \begin{cases} 1, & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Now we claim that $f = a_1v_1^* + a_2v_2^* + \cdots + a_nv_n^*$. Indeed, we have:

$$\begin{aligned} (a_1v_1^* + a_2v_2^* + \cdots + a_nv_n^*)(x) &= a_1v_1^*(x) + a_2v_2^*(x) + \cdots + a_nv_n^*(x) \\ &= a_1v_1^*(c_1v_1 + \cdots + c_nv_n) + \cdots + a_nv_n^*(c_1v_1 + \cdots + c_nv_n) \\ &= a_1c_1 + a_2c_2 + \cdots + a_nc_n \\ &= f(x) \end{aligned}$$

$$\therefore f = a_1v_1^* + a_2v_2^* + \cdots + a_nv_n^*$$

We claim that the $\{v_i^*\}_{i\in I}$ are linearly independent. Suppose $c_1v_1^* + \cdots + c_nv_n^* = 0$, the zero function. Then $(c_1v_1^* + \cdots + c_nv_n^*)(v_i) = 0 \implies c_i = 0$ for all i. And because f was arbitrary, the $\{v_i^*\}_{i\in I}$ span V^*. Therefore the list $(v_1^*, v_2^*, \dots v_n^*)$ forms a basis for V^*.

The Double Dual Space

Since V^* is a vector space, we can form the *double dual space* V^{**}, which consists of all linear maps $V^* \to K$. In particular, there is a canonical map between V and V^{**}, which gives us an isomorphism in the finite dimensional case. Given $v \in V$, consider the element $\overline{v} : V^* \to K$ defined in the following way:

$$\overline{v}(f) = f(v)$$

Note that $\overline{v}$ is a linear map, so it is indeed an element of V^{**}:

$$\overline{v}(\alpha f + g) = (\alpha f + g)(v) = \alpha f(v) + g(v)$$

Now we define the *canonical map* $\phi : V \to V^{**}$ by $\phi(v) = \overline{v}$. Well, this map is also linear:

$$\begin{aligned}
\phi(\alpha v + w) &= \overline{(\alpha v + w)} \\
\overline{(\alpha v + w)}(f) &= f(\alpha v + w) \\
&= \alpha f(v) + f(w) \\
&= \alpha\overline{v}(f) + \overline{w}(f) \\
&= (\alpha\overline{v} + \overline{w})(f) \\
\therefore \phi(\alpha v + w) &= \alpha\phi(v) + \phi(w)
\end{aligned}$$

We will prove that ker ϕ is trivial, but first we need a lemma:

Lemma 1.3. *A vector $v \in V$ is the zero vector if and only if $f(v) = 0$ for all $f \in V^*$.*

Proof. Suppose $v = 0$. Then, because linear maps take 0 to 0, we have $f(v) = 0$ for all $f \in V^*$.

For the converse, let $(v_1, v_2, \ldots v_n)$ be a basis for V. Define the family of functions $(f_1, f_2, \ldots f_n)$ by setting:

$$f_i(v_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Let $v = c_1 v_1 + \cdots + c_n v_n$. Well, $f_1(v) = c_1 = 0$, and similarly we find $c_i = 0$ for all i. Therefore, $v = 0$. □

Now we prove that ker ϕ is trivial.

$$\begin{aligned} & \phi(v) = \overline{v} = 0 \\ \iff & \overline{v}(f) = 0 \text{ for all } f \in V^* \\ \implies & f(v) = 0 \text{ for all } f \in V^* \\ \implies & v = 0 \text{ by Lemma 1.3} \end{aligned}$$

In the finite dimensional case, ϕ is also surjective, and hence it gives an isomorphism between V and V^{**}.

4 Dual Operators

Given $\phi \in \mathcal{L}(V, W)$, it is reasonable to ask if there is a corresponding map between elements in the dual spaces V^* and W^*.

Definition 1.4. For $\phi \in \mathcal{L}(V, W)$, we can define the *dual* or *adjoint operator* $\phi^* : W^* \to V^*$ by

$$\phi^*(f) = f \circ \phi$$

There is a nice relation between the matrices of ϕ and ϕ^*, which we describe in the following theorem.

Theorem 1.5. *Let V and W be finite dimensional. Fix a set of bases for V, W and the corresponding dual bases for V^*, W^*. Then, for $\phi \in \mathcal{L}(V, W)$, we have*

$$M(\phi)^T = M(\phi^*),$$

where $M(\phi)$ is the matrix associated to the linear transformation ϕ.

Proof. ([1], [2])

Let $(v_1, v_2, \dots v_n)$ be a basis for V, $(w_1, w_2, \dots w_m)$ a basis for W. Let $(v_1^*, v_2^*, \dots v_n^*)$ and $(w_1^*, w_2^*, \dots w_m^*)$ be corresponding dual bases for V^* and W^*.

The map ϕ is determined by where it takes basis elements. We start by writing out the matrix $M(\phi)$:

$$M(\phi) = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

From $M(\phi)$, we can write:

$$\phi(v_j) = \sum_{i=1}^{m} \alpha_{ij} w_i, \qquad 1 \leq j \leq n$$

Next we need to compute $M(\phi^*)$. Applying ϕ^* to the basis elements of W^* gives:

$$\phi^*(w_k^*)(v_j) = (w_k^*\phi)(v_j) = (w_k^*)(\phi(v_j)) = (w_k^*)\Big(\sum_{i=1}^{m} \alpha_{ij} w_i\Big) = \alpha_{kj}, \text{ for } 1 \leq k \leq m$$

We also have:

$$\Big(\sum_{i=1}^{n} \alpha_{ki} v_i^*\Big)(v_j) = \alpha_{kj}, \qquad 1 \leq j \leq n$$

Since these two functions agree on all the basis elements of V, they must be the same. So we conclude that $\phi^*(w_k^*) = \sum_{i=1}^{n}(\alpha_{ki} v_i^*)$. Writing out the matrix, we find:

$$M(\phi^*) = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{mn} \end{bmatrix}$$

This proves that $M(\phi)^{\mathrm{T}} = M(\phi^*)$.

□

Here are some more useful properties of the adjoint operator:

Theorem 1.6. *Let $\phi, \psi \in \mathcal{L}(V, W)$ and let $\phi^* \in \mathcal{L}(W^*, V^*)$.*

(i) ϕ^* is linear.

(ii) $(\phi\psi)^* = \psi^*\phi^*$.

(iii) If ϕ *is invertible, then so is* ϕ^*, *and in particular,* $(\phi^{-1})^* = (\phi^*)^{-1}$.

Proof. Let $f, g \in W^*$, $\alpha \in K$. Then:

$$\phi^*(\alpha f + g) = (\alpha f + g)\phi = \alpha(f\phi) + (g\phi) = \alpha\phi^*(f) + \phi^*(g).$$

This proves assertion (i). Next, we have:

$$(\phi\psi)^*(f) = f(\phi\psi) = (f\phi)\psi = \psi^*(f\phi) = \psi^*(\phi^* f) = \psi^*\phi^* f.$$

Finally, we compute:

$$(\phi^{-1})^*(\phi^*)(f) = (\phi^*(f))(\phi^{-1}) = (f\phi)(\phi^{-1}) = f$$

$$\therefore (\phi^*)^{-1} = (\phi^{-1})^*$$

□

With these tools, we can now provide a neat proof of the following useful property:

Proposition 1.7. *Let* A, B *be matrices. Then* $(AB)^T = B^T A^T$.

Proof. Let A, B correspond to $M(\phi)$ and $M(\psi)$ for some pair of linear transformations ϕ, ψ. By Theorem 1.5, we have $A^{\mathrm{T}} = M(\phi^*)$ and $B^{\mathrm{T}} = M(\psi^*)$.

Well, then we can write:

$$\begin{aligned}(AB)^{\mathrm{T}} &= M((\phi\psi)^*)\\ &= M(\psi^*\phi^*)\\ &= M(\psi^*)M(\phi^*), \text{ by definition of matrix multiplication}\\ &= B^{\mathrm{T}}A^{\mathrm{T}}\end{aligned}$$

□

5 Bilinear Forms

Now that we have studied dual spaces, we can move closer to studying the subgroups of $\mathrm{GL}_n(K)$ mentioned before. These will appear naturally as we study *bilinear forms.*

Definition 1.8. [3] Let V_1, V_2 and W be vector spaces over a field K. A map $B : V_1 \times V_2 \to W$ is *bilinear* if the following properties hold for each $c \in K$ and for each pair of vectors $x_i, y_i \in V_i$:

(i) $\phi(x_1 + y_1, x_2) = \phi(x_1, x_2) + \phi(y_1, x_2)$

(ii) $\phi(x_1, x_2 + y_2) = \phi(x_1, x_2) + \phi(x_1, y_2)$

(iii) $\phi(cx_1, x_2) = c \cdot \phi(x_1, x_2) = \phi(x_1, cx_2)$

In particular, a bilinear map $\langle\ ,\ \rangle : V \times V \to K$ is a *bilinear form.*

Given a bilinear form as defined above, we can talk about an associated linear map to the dual space V^*. Let $\langle x, \cdot\ \rangle$ be the element in V^* defined in the following way:

$$\langle x, \cdot\ \rangle : V \to K$$

$$\langle x, \cdot\ \rangle(y) = \langle x, y\rangle$$

Then the bilinear form $\langle\ ,\ \rangle$ is associated to the map:

$$\phi : V \to V^*$$

$$\phi(x) = \langle x, \cdot\ \rangle$$

Remark. The kernel of ϕ is the set $\{x \in V \mid \phi(x) = 0, \text{ the zero function}\}$. So ker $\phi = \{x \in V \mid \langle x, v\rangle = 0 \text{ for all } v \in V\}$.

This kernel has its own name: it is the *orthogonal complement of* V, denoted by $V^{\perp}$. A vector x is *orthogonal* to y if $\langle x, y\rangle = 0$. This is denoted by writing $x \perp y$.

Let S be a subset of V. A vector $x \in V$ is orthogonal to S if $\langle x, s \rangle = 0$ for all $s \in S$. We can express this by writing $x \perp S$. The set of vectors orthogonal to S, known as the *orthogonal complement of* S, is denoted by $S^{\perp}$. We can formally write

$$S^{\perp} = \{x \in V \mid \langle x, s \rangle = 0 \text{ for all } s \in S\}$$

Definition 1.9. A bilinear form $\langle \ , \ \rangle : V \times V \to K$ is *nondegenerate* if $\langle x, v \rangle = 0$ for all $v \in V$ implies $x = 0$, i.e., if $V^{\perp} = \{0\}$.

Remark. If V is finite dimensional and ϕ is injective, then we actually have an isomorphism between V and V^*. In particular, Ψ is invertible, and thus det $(M(\Psi)) \neq 0$. This gives one way to check if Ψ is injective.

Remark. We will limit ourselves to studying only vector spaces with a nondegenerate form. In Theorem 1.23, we will prove that a vector space can always be decomposed into a nondegenerate subspace and an additional subspace, so this restriction is quite reasonable.

Proposition 1.10. *For a nondegenerate bilinear form on* V*,* $\langle v, y \rangle = 0$ *for all* $v \in V$ *implies* $y = 0$*.*

Proof. ([3]) Suppose y is fixed and $\langle v, y \rangle = 0$ for all $v \in V$. Then $\langle v, \cdot \ \rangle(y) = 0$ for all $v \in V$. Recall that V and V^* have the same number of basis elements, and hence have the same dimension. If $\phi : V \to V^*$ is injective, then it gives an isomorphism between V and V^*. By surjectivity, we would then know that for all $f \in V^*, \exists\, v \in V$ such that $\phi(v) = f = \langle v, \cdot \ \rangle$.

Assume, towards a contradiction, that the y we fixed was non-zero. So if we write $y = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$, there is at least one $c_i \neq 0$. Fix $f = v_i^*$. Well, then we have:

$$\begin{aligned} f(y) &= (v_i^*)(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n) \\ &= c_i \\ &= 0 \text{ by the assumption that } f(y) = \langle v, y \rangle = 0. \end{aligned}$$

The contradiction proves that $y = 0$. Note that this last argument is similar to the one used in Lemma 1.3. □

This proposition shows that the nondegeneracy condition is symmetric in both arguments. In particular, let the *left perp of* x be the set ${}^{\perp}x = \{z \mid \langle x, z\rangle = 0\}$, and let the *right perp of* x be the set $x^{\perp} = \{z \mid \langle z, x\rangle = 0\}$. Then if ${}^{\perp}x = V \implies x = 0$, or $x^{\perp} = V \implies x = 0$, we conclude that the form is nondegenerate.

Remark. In Theorem 1.18, we will prove that assuming ${}^{\perp}x = x^{\perp}$ guarantees that the form $\langle\ ,\ \rangle$ is symmetric or alternating.

We have already mentioned that the map $\phi : V \to V^*$ is an isomorphism when V has a nondegenerate form. A related result holds for subspaces of V.

Theorem 1.11. *Let V be a vector space, S a subspace. Suppose the form $\langle\ ,\ \rangle$ is nondegenerate on either V or S, and consider the linear map $\phi : V \to S^*$ defined by $\phi(x) = \langle x, \cdot\ \rangle|_S$. Then for any $f \in S^*$, there exists a vector $x \in V$ such that $f = \langle x, \cdot\ \rangle$. This x is unique (and belongs to S) if $\langle\ ,\ \rangle$ is nondegenerate on S.*

Proof. ([2]) In both cases, $\ker \phi = \{x \in V \mid \langle s, x\rangle = 0 \text{ for all } s \in S\} = S^{\perp}$. Suppose first that $\langle\ ,\ \rangle$ is nondegenerate on S. Then $\phi|_S : S \to S^*$ is injective, hence an isomorphism. In particular, for any $f \in S^*$, there exists a unique $x \in S$ such that $f(s) = \langle s, x\rangle$ for all $s \in S$.

Now suppose that $\langle\ ,\ \rangle$ is nondegenerate on V. Extend $\phi : V \to S^*$ to $\hat{\phi} : V \to V^*$ by setting $\hat{\phi}(x) = \langle x, \cdot\ \rangle$. Then $\hat{\phi} : V \to V^*$ is an isomorphism. In particular, we can pick $f \in S^*$ and find $x \in V$ such that $f = \langle x, \cdot\ \rangle$. Furthermore, we can extend f to a map $\hat{f} \in V^*$ such that $\hat{f}|_S = f$.

Notice that there exists $x \in V$ such that $\hat{f} = \langle x, \cdot\ \rangle$. Restricting to S, we observe that for all $s \in S$, we have $f(s) = \hat{f}(s) = \langle x, s\rangle$. In this case, we do not necessarily have $x \in S$. □

6 Symmetric and Alternating Forms

We now introduce some more terminology:

Definition 1.12 ([2])**.** A bilinear form is *symmetric* if $\langle x, y\rangle = \langle y, x\rangle$ for all $x, y \in V$. Similarly, a bilinear form is *alternating* if $\langle x, x\rangle = 0$ for all $x \in V$. Finally, it is *skew-symmetric* if $\langle x, y\rangle = -\langle y, x\rangle$ for all $x, y \in V$.

Lemma 1.13. *An alternating form is always skew-symmetric.*

Proof. Let $v, w \in V$. Then we compute:

$$\begin{aligned} \langle v+w, v+w\rangle &= \langle v, v+w\rangle + \langle w, v+w\rangle \\ &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ \implies 0 &= \langle v, w\rangle + \langle w, v\rangle \\ \therefore \langle v, w\rangle &= -\langle w, v\rangle \text{ for all } v, w \in V. \end{aligned}$$

□

It is possible to formulate these definitions in terms of matrices to get a more visual understanding of what they mean.

Definition 1.14 ([3])**.** Suppose we have a bilinear form $B : V \times V \to K$, and a fixed basis $(v_1, v_2, \dots v_n)$ for V. The *matrix of* B is given by $M(B) = (a_{ij})$, where $a_{ij} = \langle v_i, v_j\rangle$.

In matrix notation:

$$M(B) = \begin{bmatrix} \langle v_1, v_1\rangle & \langle v_1, v_2\rangle & \cdots & \langle v_1, v_n\rangle \\ \langle v_2, v_1\rangle & \langle v_2, v_2\rangle & \cdots & \langle v_2, v_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_n, v_1\rangle & \langle v_n, v_2\rangle & \cdots & \langle v_n, v_n\rangle \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Note. We denoted the bilinear form $\langle\ ,\ \rangle$ by B here for clarity.

This matrix completely determines a bilinear form in the following way: suppose $x, y \in V$, let $x = c_1v_1 + c_2v_2 + \cdots + c_nv_n$, and let $y = d_1v_1 + d_2v_2 + \cdots d_nv_n$. Then we have $\langle x, y\rangle = x^{\mathrm{T}}M(B)y$.

In matrix form, we write this as:

$$\langle x, y\rangle = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = \sum_{i,j=1}^{n} a_{ij} c_i d_j$$

Going back to the terminology we introduced at the beginning of this section, we observe the following:

(i) For a symmetric form, the terms on opposite sides of the diagonal are the same. So $M(B) = M(B)^{\mathrm{T}}$.

(ii) For a skew-symmetric form, the terms on one side of the diagonal are negatives of the corresponding terms on the opposite side, i.e. $a_{ij} = -a_{ji}$.

(iii) For an alternating form, the diagonal terms are all 0. By Lemma 1.13, the matrix is also skew-symmetric.

Example 1.15. Given vectors v, w where $v = (v_1, v_2, \dots v_n)^{\mathrm{T}}$ and $w = (w_1, w_2, \dots w_n)^{\mathrm{T}}$, the dot product is defined as:

$$v \cdot w = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$$

The dot product is a bilinear form where the matrix $M(B)$ is the identity matrix. In the 2-dimensional case, observe the following:

$$v \cdot w = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \sum_{i=1}^{n} v_i w_i$$

7 Orthogonal and Symplectic Geometries

We will continue our discussion of orthogonal vectors from the section on Bilinear Forms. Recall that a bilinear form is non-degenerate if $V^{\perp} = \{0\}$, i.e. if there are no non-zero vectors in V that are orthogonal to all other vectors in V. This is the kind of behavior we are used to in Euclidean space. However, other vector spaces may have vectors that are orthogonal to every vector in the space, including itself. Furthermore, orthogonality does not need to be symmetric; i.e. we might not always have ${}^{\perp}x = x^{\perp}$.

Now we introduce some terms that are useful for studying orthogonality in vector spaces.

Definition 1.16. Let V be a vector space with a bilinear form $\langle\ ,\ \rangle$.

(i) A nonzero vector $v \in V$ is *isotropic* if $\langle x, x\rangle = 0$, and *anisotropic* or *nonisotropic* otherwise. Equivalently, a vector is isotropic if it is orthogonal to itself.

(ii) V is *isotropic* if it contains at least one isotropic vector. Otherwise, we say that V is *nonisotropic* or *anisotropic.*

(iii) V is *totally isotropic* if every $v \in V$ is isotropic.

Remark. If $v \in V$ is isotropic, then so is av for any scalar $a \in K$. We can see this from the linearity of $\langle av, av\rangle$.

Definition 1.17. Let V be a vector space with a bilinear form $\langle\ ,\ \rangle$.

(i) A vector $v \in V$ is *degenerate* if $v \perp V$, i.e. if v is orthogonal to every vector in V.

(ii) Given a subset S of V, we define the *radical of* S to be the set of vectors that are degenerate in S. We write this formally as:

$$\text{rad }(S) = S \cap S^{\perp}.$$

Note that rad (S) is different from $S^{\perp}$, which consists of *all* vectors in V that are orthogonal to S. So $S^{\perp}$ could include elements $\notin S$, whereas rad $(S) \subset S$.

(iii) V is *singular* or *degenerate* if rad $(V) \neq \{0\}$. Equivalently, $V^{\perp} \neq \{0\}$. In words, V contains a vector that is orthogonal to every other vector in V.

(iv) V is *nonsingular* or *nondegenerate* if rad $(V) = \{0\}$.

(v) V is *totally singular* or *totally degenerate* if rad $(V) = V$, i.e. every vector is orthogonal to every other vector in V.

The interesting result here is that if we have a vector space where orthogonality is a symmetric relation, i.e. ${}^{\perp}x = x^{\perp}$ for all $x \in V$, then the bilinear form is either symmetric or alternating. This gives rise to *orthogonal* or *symplectic geometries.*

Theorem 1.18. *Suppose* ${}^{\perp}x = x^{\perp}$ *for all* $x \in V$*, i.e.* $\langle x, y\rangle = 0 \iff \langle y, x\rangle = 0$ *for all* $x, y \in V$*. Then* $\langle\ ,\ \rangle$ *is either symmetric or alternating.*

Proof. ([4]) Let $x, y, z \in V$.

$$\begin{aligned}
&\langle x, z\rangle\langle x, y\rangle - \langle x, y\rangle\langle x, z\rangle = 0\\
\iff\ &\langle x\langle x, z\rangle, y\rangle - \langle x, y\rangle\langle x, z\rangle = 0 \text{ by linearity, since } \langle x, z\rangle \in K\\
\iff\ &\langle x\langle x, z\rangle, y\rangle - \langle x, \langle x, y\rangle z\rangle = 0 \text{ also by linearity, since } \langle x, y\rangle \in K\\
\iff\ &\langle x, \langle x, z\rangle y - \langle x, y\rangle z\rangle = 0\\
\implies\ &\langle\langle x, z\rangle y - \langle x, y\rangle z, x\rangle = 0 \text{ by the symmetry assumption}\\
\iff\ &\langle x, z\rangle\langle y, x\rangle - \langle x, y\rangle\langle z, x\rangle = 0\\
\iff\ &\langle x, z\rangle\langle y, x\rangle = \langle x, y\rangle\langle z, x\rangle \qquad (1.1)
\end{aligned}$$

Now fix $z = x$. Then we get

$$\langle x, x\rangle\langle y, x\rangle = \langle x, x\rangle\langle x, y\rangle.$$

If $\langle x, x\rangle \neq 0$, we get $\langle x, y\rangle = \langle y, x\rangle$ for all $x, y \in V$, and we conclude that $\langle\ ,\ \rangle$ is a symmetric form.

Note that if $\langle x, y\rangle \neq \langle y, x\rangle$, then we necessarily have $\langle x, x\rangle = 0$. By a similar argument, in which we replace x with y, we also have $\langle y, y\rangle = 0$. We will record this property and use it again later:

$$\langle x, y\rangle \neq \langle y, x\rangle \implies \langle x, x\rangle = 0 \text{ and } \langle y, y\rangle = 0 \qquad (1.2)$$

Now we look at the case where $\langle\ ,\ \rangle$ is not symmetric. Then there exist $x, y \in V$ such that $\langle x, y\rangle \neq \langle y, x\rangle$. We want to show that $\langle w, w\rangle = 0$ for all $w \in V$. This is trivially true by eq. (1.2) if $\langle x, w\rangle \neq \langle w, x\rangle$ or if $\langle y, w\rangle \neq \langle w, y\rangle$. So assume $\langle x, w\rangle = \langle w, x\rangle$ and $\langle y, w\rangle = \langle w, y\rangle$.

Comparing with eq. (1.1), we see the following:

$$\langle x,w\rangle\langle y,x\rangle = \langle x,w\rangle\langle x,y\rangle$$

$$\implies \langle x,w\rangle\langle y,x\rangle - \langle x,w\rangle\langle x,y\rangle = 0$$

$$\implies \langle x,w\rangle\Big(\langle y,x\rangle - \langle x,y\rangle\Big) = 0$$

$$\implies \langle x,w\rangle = \langle w,x\rangle = 0$$

Similarly, we have:

$$\langle y,w\rangle = \langle w,y\rangle = 0$$

Now we observe:

$$\langle x,y\rangle \neq \langle y,x\rangle$$

$$\implies \langle x,y\rangle + \langle w,y\rangle \neq \langle y,x\rangle + \langle y,w\rangle \text{ (adding 0)}$$

$$\implies \langle x+w,y\rangle \neq \langle y,x+w\rangle$$

Then from eq. (1.2):

$$\langle x+w,x+w\rangle = 0$$

$$\implies \langle x,x\rangle + \langle x,w\rangle + \langle w,x\rangle + \langle w,w\rangle = 0$$

$$\implies \langle w,w\rangle = 0$$

We have just shown that *every vector in V is isotropic.* So we conclude that $\langle\ ,\ \rangle$ is an alternating form. □

Remark. We can use Lemma 1.13 to prove that an alternating form $\langle\ ,\ \rangle$ is also skew-symmetric. Or we can see this directly. Let $x, y \in V$. Then:

$$\langle x+y,x+y\rangle = 0$$

$$\implies \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle = 0$$

$$\implies \langle x,y\rangle = -\langle y,x\rangle$$

Definition 1.19. Let V be a vector space with a nondegenerate bilinear form $\langle\ ,\ \rangle$. If $\langle\ ,\ \rangle$ is symmetric, we say that V is an *orthogonal geometry.* If $\langle\ ,\ \rangle$ is alternating, we say that V is a *symplectic geometry.*

Remark. From now on, we will always assume that our nondegenerate forms $\langle\ ,\ \rangle$ are either symmetric or alternating.

8 Orthogonal Direct Sums

We now discuss subspaces of V that are orthogonal to each other.

Definition 1.20. Let V be a vector space, and let P, Q be subspaces. Then V is the *orthogonal direct sum* of P and Q if the following hold:

(i) $V = P \oplus Q$

(ii) $P \perp Q$, i.e. every vector of P is orthogonal to every vector of Q.

We write this as $V = P \bigoplus\!\!\!\!\!\perp Q$.

Now we will prove that every vector space with a bilinear form can be decomposed into a radical and a nondegenerate subspace, but first we will do a quick check to make sure that the radical is a subspace.

Lemma 1.21. *Let V be a vector space, and suppose S is a subspace. Then rad (S) is a subspace of V.*

Proof. Note that $0 \in$ rad (S) because the zero vector is orthogonal to every other vector. Suppose $u, v \in$ rad (S). Then $\langle u + v, x\rangle = \langle u, x\rangle + \langle v, x\rangle = 0$ for all $x \in S$, and likewise for the other argument. So rad (S) is closed under addition. Finally, let $\alpha \in K$. Then $\langle \alpha u, x\rangle = \alpha\langle u, x\rangle = 0$ for all $x \in S$, so rad (S) is closed under scalar multiplication. Thus rad (S) is a subspace. □

Lemma 1.22. *Suppose $V = U_1 + U_2 + \cdots + U_m$, with $U_i \perp U_j$ for $i \neq j$. Then rad $(V) =$ rad $(U_1) +$ rad $(U_2) + \cdots +$ rad (U_m).*

Proof. ([4]) Let $x, y \in V$, where $x = u_1 + u_2 + \cdots + u_m$ and $y = w_1 + w_2 + \cdots + w_m$. Suppose $u_i \in \text{rad }(U_i)$ for each i. We compute the following:

$$\begin{aligned} \langle x, y\rangle &= \langle u_1, y\rangle + \langle u_2, y\rangle + \cdots + \langle u_m, y\rangle \\ &= \langle u_1, w_1\rangle + \langle u_2, w_2\rangle + \cdots + \langle u_m, w_m\rangle, \text{ since } u_i \perp U_j \text{ if } i \neq j \\ &= 0 \\ \therefore x &\in \text{rad }(V) \end{aligned}$$

Conversely, let $x \in \text{rad }(V)$ and let y be arbitrary. Then observe:

$$\langle x, y\rangle = \langle x, w_1\rangle + \langle x, w_2\rangle + \cdots + \langle x, w_m\rangle$$

But each $\langle x, w_i\rangle = 0$ and $\langle x, w_i\rangle = \langle u_i, w_i\rangle$, so $\langle u_i, w_i\rangle = 0$ for all i

$$\therefore u_i \in \text{rad }(U_i) \text{ for all } i$$

We have shown both containments, so this concludes our proof. □

Theorem 1.23. *Let V be a vector space with a bilinear form $\langle\ ,\ \rangle$. Then $V = rad\ (V) \operp S$ for a maximal subspace S on which $\langle\ ,\ \rangle$ is nondegenerate.*

Proof. Since rad (V) is a subspace by Lemma 1.21, it has a basis $(w_1, w_2, \ldots w_k)$. We can construct S by taking vectors that are linearly independent relative to the rad (V) basis, and then building them up to form a basis $(s_1, s_2, \ldots s_j)$. Note that if dim $V = n$, then $k + j = n$.

We now have $V = \text{rad }(V) \oplus S$. Since rad $(V) \perp V$, we also know rad $(V) \perp S$, and so $V = \text{rad }(V) \operp S$.

Finally, note that rad $(V) = \text{rad }(\text{rad }(V)) \operp \text{rad }(S)$ by Lemma 1.22. Since rad $(\text{rad }(V)) = \text{rad }(V)$, we have $\text{rad }(V) = \text{rad }(V) \operp \text{rad }(S)$. But then $\text{rad }(V) = \text{rad }(V) \oplus \text{rad }(S)$, so we conclude that rad $(S) = \{0\}$. Thus $\langle\ ,\ \rangle$ is nondegenerate on S. □

The next lemma tells us something useful about the dimensions of subspaces and their orthogonal complements, and will be used in the theorem that follows.

Lemma 1.24. *Let V be a (finite dimensional) vector space with a bilinear form, and let S be a subspace. If $\langle\ ,\ \rangle$ is nondegenerate on either V or S, then*

$$dim\ S + dim\ S^{\perp} = dim\ V$$

Proof. Consider the map $\phi : V \to S^*$ given by $\phi(x) = \langle x, \cdot\ \rangle|_S$. By Theorem 1.11,ker $\phi = S^{\perp}$, and im $\phi = S^*$. By the rank-nullity theorem, dim $V =$ dim $S^* +$ dim $S^{\perp}$. But in the finite dimensional case, ϕ restricts to an isomorphism on S, so dim $S =$ dim S^*. Therefore, dim $V =$ dim $S +$ dim $S^{\perp}$. □

Theorem 1.25 ([2]). *Let V be a vector space with a bilinear form, and let S be a subspace. If $\langle\ ,\ \rangle$ is nondegenerate on either V or S, then the following are equivalent:*

(i) $V = S \bigoplus S^{\perp}$

(ii) *rad* $(S) = \{0\}$

(iii) $V = S + S^{\perp}$

In particular, if $\langle\ ,\ \rangle$ is nondegenerate on V, then we can also say:

(a) $S = S^{\perp\perp}$

(b) *rad* $(S) =$ *rad* $(S^{\perp})$

(c) $\langle\ ,\ \rangle$ *is nondegenerate on S if and only if it is nondegenerate on $S^{\perp}$*

Proof. First, suppose that $\langle\ ,\ \rangle$ is nondegenerate on either V or S. Assume $V = S \bigoplus S^{\perp}$. Then by Lemma 1.22, we have rad $(V) =$ rad $(S) +$ rad $(S^{\perp})$. If $\langle\ ,\ \rangle$ is nondegenerate on S, we are done. Suppose we only know that $\langle\ ,\ \rangle$ is nondegenerate on V. Then rad $(V) = \{0\}$, so we must have rad $(S) = \{0\}$. So $(i) \implies (ii)$.

Now assume rad $(S) = \{0\}$. Then $S \cap S^{\perp} = \{0\}$. But by Lemma 1.24, dim $S +$ dim $S^{\perp} =$ dim V, so $V = S \oplus S^{\perp}$. So $(ii) \implies (iii)$.

Finally, assume $V = S + S^{\perp}$. If $\langle\ ,\ \rangle$ is nondegenerate on S, then we have a direct sum by the argument in the previous paragraph. Since $S \perp S^{\perp}$, we conclude $V = S \textcircled{\perp} S^{\perp}$. Otherwise, suppose $\langle\ ,\ \rangle$ is nondegenerate on V, i.e. rad $(V) = \{0\}$. We still have $V = S + S^{\perp}$, with $S \perp S^{\perp}$. By Lemma 1.22, rad $(S) = \{0\}$, and we repeat the argument for the case where $\langle\ ,\ \rangle$ is nondegenerate on S.

Because we have shown $(iii) \implies (i)$, we may conclude the first part of our proof.

Now suppose $\langle\ ,\ \rangle$ is nondegenerate on V. We will start by proving assertion (a). Let $s \in S$. We want to show that $s \in S^{\perp\perp}$.

$$S^{\perp} = \{x \in V \mid \langle x, s\rangle = 0 \text{ for all } s \in S\}$$

$$S^{\perp\perp} = \{y \in V \mid \langle y, x\rangle = 0 \text{ for all } x \in S^{\perp}\}$$

To show $s \in S^{\perp\perp}$, we need $\langle s, x\rangle = 0$ for all $x \in S^{\perp}$. Well, $\langle s, x\rangle = k\langle x, s\rangle$ for some scalar k, and $\langle x, s\rangle = 0$ from the definition of $S^{\perp}$. So $\langle s, x\rangle = 0$, and therefore $S \subset S^{\perp\perp}$.

By applying Lemma 1.24 to the case where our subspace is $S^{\perp}$, we get the following relation:

$$\dim\, S^{\perp\perp} + \dim\, S^{\perp} = \dim\, V = \dim\, S + \dim\, S^{\perp}$$

$$\implies \dim\, S^{\perp\perp} = \dim\, S$$

$$\implies S = S^{\perp\perp}, \text{ since } S \subset S^{\perp\perp}$$

Assertion (b) follows immediately. Finally, if rad $(S) =$ rad $(S^{\perp\perp})$, then rad $(S) = 0 \iff$ rad $(S^{\perp\perp}) = 0$. This proves (c). □

Before we end this section, we will prove a useful theorem about the decomposition of orthogonal geometries into orthogonal sums of one-dimensional subspaces. But first, we will need a lemma:

Lemma 1.26. *Let V be an orthogonal geometry. Then it contains an anisotropic vector—some $v \in V$ such that $\langle v, v\rangle \neq 0$.*

Proof. ([4]) Assume (towards a contradiction) that every vector is isotropic. Then the form $\langle\ ,\ \rangle$ is *alternating* on V, so our space is actually symplectic. So for $x, y \in V$:

$$\langle x, y\rangle = \langle y, x\rangle = -\langle y, x\rangle \implies 2\langle y, x\rangle = 0 \implies \langle x, y\rangle = 0$$

But x, y were arbitrary, so this contradicts the nondegeneracy condition. Therefore, there exists at least one anisotropic vector in V. □

Theorem 1.27 (Orthogonal sum decomposition.). *Let V be an n-dimensional orthogonal geometry. Then we can write:*

$$V = V_1 \bigoplus V_2 \bigoplus \cdots \bigoplus V_n$$

where each V_i = span (v_i) for some $v_i \in V$. Furthermore, the form $\langle\ ,\ \rangle$ is nondegenerate on each V_i.

Proof. ([4]) We prove this by induction on the dimension of V. By Lemma 1.26, there exists $v \in V$ such that $\langle v, v\rangle \neq 0$. Set $v_1 = v$ and V_1 = span (v_1). Let $W = {V_1}^{\perp}$. Note that rad (V_1) = $\{0\}$, so by Theorem 1.25, we can write $V = V_1 \bigoplus W$. In particular, dim $W = n - 1 <$ dim V. We conclude our proof by induction on the dimension of W. Note that rad $(V_i) = \{0\}$ for all i, so $\langle\ ,\ \rangle$ is nondegenerate on each subspace V_i. □

9 Hyperbolic Space

In this section, we introduce a special vector space with some "nice" properties.

Definition 1.28. Let V be a vector space with a bilinear form $\langle\ ,\ \rangle$. A *hyperbolic pair* is an ordered pair of vectors $u, v \in V$ such that

$$\langle u, u\rangle = \langle v, v\rangle = 0 \text{ and } \langle u, v\rangle = 1$$

The subspace H = span (u, v) is called a *hyperbolic plane.* An orthogonal sum of the form

$$\mathcal{H} = H_1 \bigoplus H_2 \bigoplus \cdots \bigoplus H_n$$

is called a *hyperbolic space.*

Let (u_i, v_i) be a hyperbolic pair for H_i. Then $(u_1, v_1, u_2, v_2, \ldots u_n, v_n)$ is a *hyperbolic basis* for $\mathcal{H}$.

Remark. The form $\langle\ ,\ \rangle$ is nondegenerate on each H_i, so rad $(H_i) = \{0\}$. By Lemma 1.22, rad $(\mathcal{H}) = \{0\}$, so $\langle\ ,\ \rangle$ is also nondegenerate on $\mathcal{H}$.

There is an analog of Theorem 1.27 which uses hyperbolic spaces. It states that any vector space with a nondegenerate, symplectic form can be written as an orthogonal sum of hyperbolic planes. We state this in the following theorem:

Theorem 1.29. *Let V be a symplectic geometry. Then V is a hyperbolic space, and we can write:*

$$V = H_1 \bigoplus H_2 \bigoplus \cdots \bigoplus H_k$$

In particular, notice that dim V is even.

Proof. The proof is by induction on the dimension of V. Let $v_1 \in V$. Because $\langle\ ,\ \rangle$ is alternating, we have $\langle v_1, v_1\rangle = 0$. The nondegenerate condition implies that rad $(V) = \{0\}$, so there exists $u \in V$ such that $\langle u, v_1\rangle = \alpha \neq 0$ for some scalar α. In particular, note that $\dfrac{\langle u, v_1\rangle}{\alpha} = 1$. Set $z_1 = \dfrac{1}{\alpha}u$; then we have $\langle z_1, v_1\rangle = 1$. Then span (z_1, v_1) forms a hyperbolic plane; call this H_1.

Now let $W = {H_1}^{\perp}$. Since rad $(H_1) = \{0\}$, Theorem 1.25 implies $V = H_1 \bigoplus W$. Furthermore, it asserts that since H_1 is nondegenerate, W is also nondegenerate. Since dim $W <$ dim V, we conclude our proof by induction on the dimension of V.

Since each hyperbolic plane has dimension 2, we also note that dim V is even.

□

10 Isometries

Orthogonal and symplectic geometries motivate the study of two subgroups of $\mathrm{GL}_n(K)$, known as the orthogonal and symplectic groups. To discuss these groups, we first need the notion of an isometry.

Definition 1.30. Let V and W be vector spaces, each with an associated bilinear form. Let $\phi : V \to W$ be an isomorphism (i.e. a bijective linear map). We say that ϕ is an *isometry* if the following holds for all $u, v \in V$:

$$\langle \phi u, \phi v \rangle = \langle u, v \rangle$$

If ϕ is an isometry, we say that V and W are *isometric* and write $V \approx W$.

Proposition 1.31. *The set of all isometries from V to V forms a group under composition.*

Proof. The identity map is an isometry, so it is in the set. For closure, suppose ϕ and σ are both isometries. Composition then gives the following result:

$$\langle \phi(\sigma u), \phi(\sigma v) \rangle = \langle \sigma u, \sigma v \rangle = \langle u, v \rangle$$

Hence the set is closed under composition. Finally, suppose ϕ is an isometry. We check that the inverse of ϕ is also in the set:

$$\begin{aligned} \langle \phi^{-1}u, \phi^{-1}v \rangle &= \langle \phi(\phi^{-1}u), \phi(\phi^{-1}v) \rangle \text{ (because } \phi \text{ is an isometry)} \\ &= \langle (\phi\phi^{-1})u, (\phi\phi^{-1})v \rangle \\ &= \langle u, v \rangle \end{aligned}$$

□

In the following theorem, we discuss some properties of isometries:

Theorem 1.32 ([2])**.** *Let V, W be vector spaces, each with a bilinear form $\langle\ ,\ \rangle$. Let $(v_1, v_2, \dots v_n)$ be a basis for V. Suppose ϕ is a linear transformation between V and W. Then the following hold:*

(i) ϕ is an isometry if and only if $\langle \phi v_i, \phi v_j \rangle = \langle v_i, v_j \rangle$ for all i, j

(ii) Suppose V is an orthogonal geometry. Then ϕ is an isometry if and only if $\langle \phi v, \phi v\rangle = \langle v, v\rangle$ for all $v \in V$

(iii) Suppose ϕ is an isometry and we know the following:

$$V = S \bigoplus S^{\perp} \qquad\qquad W = T \bigoplus T^{\perp}$$

Then if $\phi(S) = T$, we also have $\phi(S^{\perp}) = T^{\perp}$

Proof. Part (i):

Suppose ϕ is an isometry. Then $\langle \phi u, \phi v\rangle$ for all $u, v \in V$, and this includes the basis vectors $(v_1, v_2, \dots v_n)$.

Now suppose $\langle \phi v_i, \phi v_j\rangle = \langle v_i, v_j\rangle$ for all i, j. Let $x = c_1v_1 + \cdots + c_nv_n$ and let $y = d_1v_1 + \cdots + d_nv_n$. Then by linearity,

$$\begin{aligned}\langle \phi x, \phi y\rangle &= \sum_{i=1,\, j=1}^{n} k_{ij}\langle \phi v_i, \phi v_j\rangle \\ &= \sum_{i=1,\, j=1}^{n} k_{ij}\langle v_i, v_j\rangle \\ &= \langle x, y\rangle\end{aligned}$$

Part (ii): Now suppose V is an orthogonal geometry. For one direction, suppose that ϕ is an isometry. Then we are trivially done, because $\langle \phi v, \phi v\rangle = \langle v, v\rangle$ for all $v \in V$.

For the other direction, suppose $\langle \phi v, \phi v\rangle = \langle v, v\rangle$ for all $v \in V$. Then we have:

$$\langle \phi(u + v), \phi(u + v)\rangle = \langle u + v, u + v\rangle$$

$$\implies \langle \phi u + \phi v, \phi u + \phi v\rangle = \langle u + v, u + v\rangle$$

$$\implies \langle \phi u, \phi u\rangle + \langle \phi u, \phi v\rangle + \langle \phi v, \phi u\rangle + \langle \phi v, \phi v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle$$

$$\implies \langle u, u\rangle + \langle v, v\rangle + 2\langle \phi u, \phi v\rangle = \langle u, u\rangle + \langle v, v\rangle + 2\langle u, v\rangle$$

$$\therefore \langle \phi u, \phi v\rangle = \langle u, v\rangle$$

Part (iii): Let $x \in S^{\perp}$. We want to show that $\phi x \in T^{\perp}$, i.e. $\phi x \perp t$ for all $t \in T$. Let $t \in T$. Since $\phi(S) = T$, there is some $s \in S$ such that $\phi s = t$.

Well, $\langle \phi x, t\rangle = \langle \phi x, \phi s\rangle = \langle x, s\rangle = 0 \implies \phi x \in T^\perp$ for all $x \in S^\perp \implies \phi(S^\perp) \subset T^\perp$.

Since ϕ is an isomorphism, we compare dimensions to conclude that $\phi(S^\perp) = T^\perp$. □

Definition 1.33. An isometry of an orthogonal geometry is called an *orthogonal transformation.* Similarly, an isometry of a symplectic geometry is called a *symplectic transformation.*

Definition 1.34. The group of all orthogonal transformations of an n-dimensional orthogonal geometry V is denoted by $\mathcal{O}_n(V)$ and is called the *orthogonal group of V* . Likewise, the group of all symplectic transformations of a $2n$-dimensional symplectic geometry is denoted $Sp_n(V)$ and is called the *symplectic group of V.*

11 Witt's Theorem

In this section, we discuss Witt's theorem, which states that any isometry between two subspaces of isometric orthogonal or symplectic geometries can be extended to an isometry of the whole geometry. We start by discussing a preliminary lemma.

Lemma 1.35. *Suppose V is an orthogonal or symplectic geometry, and let S be a subspace. If $P =$ span $(v) \bigoplus S$ for some isotropic vector v, then there exists some z and a hyperbolic plane $H =$ span (v, z) for which $Q = H \bigoplus S$ is an extension of P.*

Proof. ([2] p.274) Let $P =$ span $(v) \bigoplus S$ for some isotropic vector v. Well $v \notin S$, and because $S = S^{\perp\perp}$ by Theorem 1.25, we also know $v \notin S^{\perp\perp}$. Then there exists some $x \in S^\perp$ such that $\langle v, x\rangle \neq 0$. By assumption, we have $v \in S^\perp$.

Now we divide our argument into two cases. If V is symplectic, then all vectors are isotropic. Let $z = \dfrac{x}{\langle v, x\rangle}$. Then $\langle v, z\rangle = \langle v, \dfrac{x}{\langle v, x\rangle}\rangle = \dfrac{\langle v, x\rangle}{\langle v, x\rangle} = 1$. Note that $H =$ span (v, z) is thus a hyperbolic plane.

If V is orthogonal, let $z = \alpha v + \beta x$ for some scalars α, β. Then we have two equations:

$$1 = \langle v, z\rangle = \langle v, \alpha v + \beta x\rangle = \alpha \cdot 0 + \beta\langle v, x\rangle \tag{1.3}$$

$$0 = \langle z, z\rangle = \langle \alpha v + \beta x, \alpha v + \beta x\rangle = \alpha^2\langle v, v\rangle + \beta^2\langle x, x\rangle + 2\alpha\beta\langle v, x\rangle = \beta^2\langle x, x\rangle + 2\alpha \tag{1.4}$$

We can solve eq. (1.3) for β and then plug the result into eq. (1.4) to solve for α.

So for both cases, we can construct a vector z such that $H = \text{span }(v, z)$ is a hyperbolic plane. Recall that $x, v \in S^{\perp}$. So in both cases, we also know $z \in S^{\perp}$ and therefore $H \subset S^{\perp}$.

Since H is hyperbolic, rad $(H) = \{0\}$. Therefore, $H \cap H^{\perp} = \{0\}$. So if $H \subset S^{\perp}$, then $S = S^{\perp\perp} \subset H^{\perp} \implies H \cap S = \{0\}$. Thus we have the orthogonal sum $Q = H \bigoplus S$. □

We should step back and understand what this means. Given V, a subspace S and an isotropic vector v that is orthogonal to S, the lemma asserts that we can find another vector z that is orthogonal to S and forms a hyperbolic pair with v. The next theorem uses this statement to extend any subspace of V to one on which $\langle\ ,\ \rangle$ is nondegenerate.

Theorem 1.36. *Suppose V is an orthogonal or symplectic geometry and U is a subspace with a degenerate form. By Theorem 1.23, we can write $U = rad\ (U) \bigoplus W$ for a subspace W on which $\langle\ ,\ \rangle$ is nondegenerate. Let $(u_1, u_2, \ldots u_n)$ be a basis for rad (U). Then V contains a hyperbolic space $\mathcal{H}_n = H_1 \bigoplus \cdots \bigoplus H_n$ with hyperbolic basis $(u_1, v_1, u_2, v_2, \ldots u_n, v_n)$. Furthermore, the subspace $\overline{U} = \mathcal{H}_n \bigoplus W$ is an extension of U on which $\langle\ ,\ \rangle$ is nondegenerate.*

Finally, let ϕ be an isometry of U into a subspace of a nondegenerate geometry $\overline{V}$. Then we can extend ϕ to an isometry $\overline{\phi} : \overline{U} \to \overline{\phi}(\overline{U}) \subset \overline{V}$.

Proof. We prove this by induction on the dimension of rad (U). Observe that the basis vectors for rad (U) are all isotropic. For $n = 1$, we can use Lemma 1.35 to show that $\overline{U} = \mathcal{H}_1 \bigoplus W$ is in V. Because $\langle\ ,\ \rangle$ is nondegenerate on W and $\mathcal{H}_1$, we conclude that it is nondegenerate on $\overline{U}$.

Assume that the result holds for all dimensions up to $n-1$. For the case where rad (U) has dimension n, we have the following:

$$\begin{aligned} U &= \text{rad }(U) \bigoplus W \\ &= \text{span }(u_1, u_2, \ldots u_n) \bigoplus W \\ &= \text{span }(u_n) \bigoplus \Big(\text{span }(u_1, u_2, \ldots u_{n-1}) \bigoplus W\Big) \end{aligned}$$

By Lemma 1.35, there exists a vector v_n and a hyperbolic plane $H_n = \text{span}\ (u_n, v_n)$ such that $H_n \perp \text{span}\ (u_1, u_2, \dots u_{n-1}) \perp W$ is a subspace of V. By the induction hypothesis, there also exist vectors $(v_1, v_2, \dots v_{n-1})$ such that $(u_1, v_1, u_2, v_2, \dots u_{n-1}, v_{n-1})$ form a hyperbolic basis for $\mathcal{H}_{n-1} = H_1 \perp \cdots \perp H_{n-1}$.

Putting everything together, we conclude that $\overline{U} = \mathcal{H}_n \perp W$ is a subspace of V. Since $\langle\ ,\ \rangle$ is nondegenerate on each H_i and also on W, it is nondegenerate on $\overline{U}$.

For the next part, suppose $\phi : U \hookrightarrow \overline{V}$ is an isometry. Suppose $(w_1, w_2, \dots w_m)$ is a basis for W. Define $\overline{u_i} = \phi(u_i)$ and $\overline{w_j} = \phi(w_j)$. Also let $\overline{W} = \phi(W)$. Because ϕ is an isometry, the orthogonal sum $U = \text{rad}\ (U) \perp W$ maps to $Q := \text{span}\ (\overline{u}_1, \overline{u}_2, \dots \overline{u}_n) \perp \overline{W}$.

But now we can apply the theorem to Q and find isotropic vectors $(\overline{v}_1, \overline{v}_2, \dots \overline{v}_n)$ such that $(\overline{u}_1, \overline{v}_1, \overline{u}_2, \overline{v}_2, \dots \overline{u}_n, \overline{v}_n)$ form a hyperbolic basis. We extend ϕ to $\overline{\phi}$ by assigning $\overline{\phi}(v_i) = \overline{v_i}$. By construction, the map $\overline{\phi} : \overline{U} \hookrightarrow \overline{V}$ is an isometry. □

Before moving on, we should understand the implications of this theorem. It says the following: (i) a degenerate subspace of an orthogonal or symplectic geometry can be extended to one on which $\langle\ ,\ \rangle$ is nondegenerate, (ii) for each degenerate vector in U, we can find a vector in V with which it forms a (nondegenerate) hyperbolic pair, and (iii) we can always find a suitable hyperbolic space in V that contains the degenerate part of U.

Another key part of this theorem is this: given an isometry between two degenerate subspaces of an orthogonal or symplectic geometry, we can always extend the isometry to an isometry between nondegenerate subspaces. This involves finding extensions of the domain and image subspaces where $\langle\ ,\ \rangle$ is nondegenerate.

Witt's theorem is a special case of the previous statement. It asserts that given two isometric orthogonal or symplectic geometries, an isometry between two subspaces (which could be degenerate) can always be extended to an isometry between the two geometries. Here is the formal statement of the theorem:

Theorem 1.37 (Witt's theorem). *Let V and V' be orthogonal or symplectic geometries which are isometric under an isometry Φ. Suppose ϕ is an isometry on a subspace S of V, i.e. the map $\phi : S \to \phi(S)$ is an isometry. Then ϕ can be extended to an isometry between V and V'.*

Proof. ([4] p. 121) Let S' denote the image of S under ϕ. First note that if S is degenerate, we can always extend it to a nondegenerate subspace $\overline{S}$ by Theorem 1.36. By the same theorem, we can extend ϕ to an isometry between nondegenerate subspaces, i.e. we can extend S' to be nondegenerate. So we can assume that S and S' are both nondegenerate.

Case 1 (symplectic): If V is symplectic, we use Theorem 1.25 to write $V = S \bigoplus S^{\perp}$ and $V' = S' \bigoplus S'^{\perp}$. We also use the theorem to observe that $S^{\perp}$ and $S'^{\perp}$ are nondegenerate. By Theorem 1.29, $S^{\perp}$ and $S'^{\perp}$ are both hyperbolic spaces, and we can find for them hyperbolic bases $(v_1, z_1, v_2, z_2, \dots v_k, z_k)$ and $(v'_1, z'_1, v'_2, z'_2, \dots v'_k, z'_k)$, respectively. Now we extend the isometry $\phi : S \to S'$ by setting $\phi(v_i) = v'_i$ and $\phi(z_i) = z'_i$. This gives us the isometric map $\phi : V \to V'$.

Case 2 (orthogonal, characteristic $\neq 2$): We proceed by induction on the dimension of S, starting with the base case dim $S = 1$. Then $S =$ span (u) for some $u \in V$, and the nondegeneracy condition guarantees that $\langle u, u\rangle \neq 0$. Let $\phi(u)$ be denoted by v. Also note that there exists some vector x such that $\Phi(x) = v$. Next, we observe that $\langle u, u\rangle = \langle v, v\rangle = \langle x, x\rangle \neq 0$. Our objective is to find an isometry $\sigma : V \to V$ such that $\sigma(u) = x$; then the composition $\Phi \circ \sigma : V \to V'$ will give us $\Phi\sigma(u) = \Phi(x) = v$. In particular, $\Phi\sigma|_S = \phi$, which is what we need.

Consider the vectors $(u + x)$ and $(u - x)$; since $\langle u + x, u - x\rangle = \langle u, u\rangle - \langle x, x\rangle = 0$, they are orthogonal. If both these vectors were isotropic, then $(u + x) + (u - x) = 2u$ would be isotropic. Since that is not the case, one of these vectors must be anisotropic. Denote this vector by $Y = u + \delta x$, where $\delta = \pm 1$ (the base field has characteristic $\neq 2$). Let $Y' = u - \delta x$. Then $Y' \in$ span $(Y)^{\perp}$.

Since span (Y) is nondegenerate, we can write (by Theorem 1.25):

$$V = \text{span } (Y) \bigoplus \text{span } (Y)^{\perp}$$

Let $\sigma : V \to V$ be the map that fixes span $(Y)^{\perp}$ and maps $Y \mapsto -Y$. This is an isometry because $\langle -Y, -Y \rangle = \langle Y, Y \rangle$. Now we look at the following equations:

$$\sigma(u + \delta x) = -u - \delta x$$

$$\sigma(u - \delta x) = u - \delta x$$

$$\implies \sigma(2u) = -2\delta x$$

$$\therefore \sigma(u) = -\delta x$$

Suppose $\delta = -1$. Then $\sigma(u) = x$, and we have the isometry we need. Otherwise, suppose $\delta = 1$. Define a new map $\rho : V \to V$ by $\rho(z) = -z$. Then $\rho\sigma(u) = x$, and we are done.

Now for the inductive step, we assume that the theorem holds for dim $S < n$. Suppose dim $S = n$. By Theorem 1.27, we can write $S = S_1 \textcircled{\perp} S_{n-1}$, where S_1 is a one-dimensional subspace. By the nondegeneracy condition on S, rad $(S_1) = \{0\}$. So by Theorem 1.25, we can write $V = S_1 \textcircled{\perp} {S_1}^{\perp}$.

By the induction hypothesis, we can extend the map $\phi|_{S_1}$ to an isometry $V \to V'$; call this isometry σ_1. Denote $\sigma_1(S_1)$ by S_1'. Note that σ maps ${S_1}^{\perp} \mapsto {S_1'}^{\perp}$, so these two vector spaces are isometric. Furthermore, they are nondegenerate, so we can apply the theorem to ${S_1}^{\perp}$ and ${S_1'}^{\perp}$.

Next, we observe that S_{n-1} is a subspace of ${S_1}^{\perp}$, so by the induction hypothesis, the map $\phi|_{S_{n-1}}$ can be extended to an isometry $\sigma_2 : {S_1}^{\perp} \to {S_1'}^{\perp}$.

Finally, we define $\Sigma : V \to V'$ by $\Sigma(e + f) = \sigma_1(e) + \sigma_2(f)$, where $e \in S_1$ and $f \in {S_1}^{\perp}$. The theorem follows by induction. □

Chapter 2

Orthogonal and Symplectic Geometries

1 Rotations and reflections

We start by introducing some standard linear algebra notation. Let V denote a vector space over a field K, and let $\beta = (v_1, v_2, \dots v_n)$ be a fixed, ordered basis for V. Then for each $v \in V$, there exists a unique ordered n-tuple $(c_1, c_2, \dots c_n)$ of scalars such that $v = c_1 v_1 + \cdots + c_n v_n$. We can define the *coordinate map* $V \to K^n$ by setting:

$$[v]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Next, we quote (without proving) a standard result in linear algebra: given a linear operator $\tau : V \to V$, the matrix of τ can be written as the following partitioned matrix $[\tau]_\beta$:

$$[\tau]_\beta = ([\tau v_1]_\beta \mid \cdots \mid [\tau v_n]_\beta)$$

It follows that we can write $[\tau v]_\beta = [\tau]_\beta [v]_\beta$.

Recall from the discussion following Definition 1.14 that a bilinear form has an associated matrix, with respect to a fixed basis. If we let B denote the map $\langle\ ,\ \rangle : V \times V \to K$, then we can write $\langle x, y\rangle = x^{\mathrm{T}} M(B) y$. We can be more precise with our new terminology: given an ordered basis β, we can write $\langle x, y\rangle = [x]_\beta^{\mathrm{T}} M(B)[y]_\beta$

Suppose we have an isometry σ such that $\langle \sigma x, \sigma y\rangle = \langle x, y\rangle$. Then we can write:

$$\begin{aligned}
\langle \sigma x, \sigma y\rangle &= [\sigma x]_\beta^{\mathrm{T}} M(B)[\sigma y]_\beta \\
&= \Big([\sigma]_\beta [x]_\beta\Big)^{\mathrm{T}} M(B)[\sigma]_\beta [y]_\beta \\
&= [x]_\beta^{\mathrm{T}} \Big([\sigma]_\beta^{\mathrm{T}} M(B)[\sigma]_\beta\Big)[y]_\beta \\
&= \langle x, y\rangle \\
\implies [x]_\beta^{\mathrm{T}} M(B)[y]_\beta &= [x]_\beta^{\mathrm{T}} \Big([\sigma]_\beta^{\mathrm{T}} M(B)[\sigma]_\beta\Big)[y]_\beta \\
\implies M(B) &= [\sigma]_\beta^{\mathrm{T}} M(B)[\sigma]_\beta
\end{aligned}$$

Taking determinants, we find:

$$\begin{aligned}
\det M(B) &= \det [\sigma]_\beta^{\mathrm{T}} \det M(B) \det [\sigma]_\beta \\
&= (\det [\sigma]_\beta)^2 \det M(B) \qquad (2.1)
\end{aligned}$$

We now claim that if B is nondegenerate, $\det M(B) \neq 0$. Suppose B is a nondegenerate form. Then the kernel of the map $\phi : V \to V^*$ is trivial. Note that a vector x is in $\ker \phi$ if and only if $\langle x, y\rangle = 0$ for all $y \in V$.

Let $(v_1, v_2, \ldots v_n)$ denote the ordered basis β. Also let $x = c_1 v_1 + \cdots + c_n v_n$ and let $y = d_1 v_1 + \cdots + d_n v_n$. By the discussion following Definition 1.14, we can write $\langle x, y\rangle = \sum_{i,j=1}^n a_{ij} c_i d_j$. So a vector x is in $\ker \phi \iff \langle x, y\rangle = 0$ for all $y \iff \sum_{i,j=1}^n a_{ij} c_i = 0$ for all j.

This last system of equations can be expressed in the following matrix form:

$$\begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = 0$$

Well, this is precisely the equation $\Big(M(B)\Big)^{\mathrm{T}} [x]_\beta = 0$. Since $\ker \phi = \{0\}$, we know that only $x = 0$ solves this system of equations, and hence $\det M(B)^{\mathrm{T}} \neq 0 \implies \det M(B) \neq 0$.

Returning to eq. (2.1), we cancel out $\det M(B)$ from both sides and find $\det [\sigma]_\beta = \pm 1$

Definition 2.1. Let σ be an isometry on an orthogonal or symplectic geometry. If $\det \sigma = 1$, then σ is called a *rotation.* If $\det \sigma = -1$, then σ is called a *reflection.*

2 Symplectic Transvections

Recall from Definition 1.34 that the group of all isometries of a $2n$-dimensional symplectic geometry is the symplectic group $Sp_n(V)$. We would like to say something about the generators of this group, so we start by asking how an isometry σ affects each vector in V. By looking at the net effect, we can say something useful about the behavior of σ. Fix a vector $v \in V$; what are the vectors that are translated in the direction of v? This is equivalent to finding the set $\{x \mid \sigma x = x + kv \text{ for some } k \in K\}$. Well, we have the following setup:

$$\sigma x = x + k_1 v \qquad \sigma y = y + k_2 v \qquad \langle \sigma x, \sigma y \rangle = \langle x, y \rangle$$

$$\implies \langle x + k_1 v, y + k_2 v \rangle = \langle x, y \rangle$$

$$\langle x, y \rangle + k_2 \langle x, v \rangle + k_1 \langle v, y \rangle + k_1 k_2 \langle v, v \rangle = \langle x, y \rangle \qquad \langle v, v \rangle = 0, \langle x, y \rangle \text{ terms cancel}$$

$$k_1 \langle v, y \rangle = k_2 \langle v, x \rangle$$

$$\implies k_1 = \frac{k_2}{\langle v, y \rangle} \langle v, x \rangle$$

For this to make sense, choose $y \in V$ such that y is not orthogonal to v. We know from the nondegeneracy condition that V contains elements not orthogonal to v, so this is fine. Finally, we can write:

$$\sigma x = x + k \langle v, x \rangle v \tag{2.2}$$

This gives us some useful information about x: if x is in $\langle v \rangle^{\perp}$, then $\sigma x = x$. So our isometries σ *leave* $\langle v \rangle^{\perp}$ *as invariant subspaces.*

We could have also taken a different perspective here, by first looking at invariant subspaces of the isometries σ of V. For some vector v, let $\langle v \rangle$ denote the 1-dimensional subspace generated by v. Next, let H denote the subspace $\langle v \rangle^{\perp}$ and suppose H is an invariant subspace of σ.

Fix $y \in H$. Then we have $\sigma y = y \implies \sigma y - y = 0$. So for any $x \in V$, we know the following:

$$\begin{aligned} \langle \sigma x, \sigma y - y\rangle &= 0 \\ \implies \langle \sigma x, \sigma y\rangle - \langle \sigma x, y\rangle &= 0 \\ \implies \langle x, y\rangle - \langle \sigma x, y\rangle &= 0 \\ \implies \langle x - \sigma x, y\rangle &= 0 \\ \therefore x - \sigma x &\in H^{\perp} \end{aligned}$$

Well, $H^{\perp} = \langle v\rangle^{\perp\perp}$, so by Theorem 1.25, $H^{\perp} = \langle v\rangle$. Therefore, $\sigma x = x + k_1 v$. Now we can apply the same argument as before to obtain eq. (2.2).

Definition 2.2. An isometry of the type $\sigma_{k,v}x = x + k\langle v, x\rangle v$ is called a *symplectic transvection in the direction of v.*

Next, we claim that every symplectic transformation is the product of symplectic transformations. We will break up the proof into parts.

Proposition 2.3. *Suppose x, y are two vectors in a symplectic geometry V. Then there exists a transvection that sends $x \mapsto y$.*

Proof. ([4])

Case 1: $\langle x, y\rangle \neq 0$. The direction for the transvection would be $y - x$. Observe:

$$\sigma_{k,y-x}x = x + k\langle y - x, x\rangle(y - x) = x + k\langle y, x\rangle(y - x)$$

Set $k = \dfrac{1}{\langle y, x\rangle}$. Then we get $\sigma_{k,y-x} = x + (y - x) = y$ as needed.

Case 2: $\langle x, y\rangle = 0$. We need to find a vector z such that $\langle x, z\rangle \neq 0$ and $\langle z, y\rangle \neq 0$. Suppose $\langle x\rangle^{\perp} = \langle y\rangle^{\perp}$. Then let z be any vector not in $\langle x\rangle^{\perp}$; such a vector exists because of the nondegeneracy condition.

Now suppose $\langle x\rangle^\perp \neq \langle y\rangle^\perp$. Pick vectors m, n such that $m \in \langle x\rangle^\perp$, $m \notin \langle y\rangle^\perp$ and $n \in \langle y\rangle^\perp$, $n \notin \langle x\rangle^\perp$ (if $y^\perp \subset x^\perp$, pick $n \notin \langle x\rangle^\perp$). Let $z = m + n$. Then $\langle x, z\rangle = \langle x, n\rangle \neq 0$ and $\langle y, z\rangle = \langle y, m\rangle \neq 0$.

Finally, we can apply Case 1 twice to move $x \mapsto z \mapsto y$. □

Proposition 2.4. *For any two hyperbolic pairs (u_1, u_2) and (v_1, v_2) in a symplectic geometry V, there exists a product of symplectic transvections that carries one pair to another. We need at most 4 transvections to carry out this operation.*

Proof. ([2], [4]) From Proposition 2.3, we know that we can map $(u_1, u_2) \mapsto (v_1, y)$ by a product π of at most 2 transvections. Note that isometries map hyperbolic pairs to hyperbolic pairs, so (v_1, y) is a hyperbolic pair. In particular, observe that $\langle v_1, y\rangle = \langle \pi u_1, \pi u_2\rangle = \langle u_1, u_2\rangle = 1$.

Now we want to show that the hyperbolic pair (v_1, y) can be mapped to (v_1, v_2) by at most 2 transvections. This holds in the general case for hyperbolic pairs with the first coordinate fixed. There are two cases to consider:

Case 1: $\langle y, v_2\rangle \neq 0$. The transvection $\sigma_{k, v_2 - y}$ with $k = \dfrac{1}{\langle v_2, y\rangle}$ maps $y \mapsto v_2$. The same transvection, when applied to v_1, gives the following:

$$\sigma_{k,v_2-y} v_1 = v_1 + k\Big(\langle v_2, v_1\rangle - \langle y, v_1\rangle\Big)(v_1 - y)$$

But $\langle v_2, v_1\rangle = -1 = \langle y, v_1\rangle$, so $\sigma_{k,v_2-y} v_1 = v_1$.

Case 2: $\langle y, v_2\rangle = 0$. We use an argument similar to the one used in the second part of Proposition 2.3; find a vector z such that $\langle y, z\rangle \neq 0$, $\langle z, v_2\rangle \neq 0$, and then use Case 1 twice to map $(v_1, y) \mapsto (v_1, z) \mapsto (v_1, v_2)$.

Putting everything together, we get the map $(u_1, u_2) \mapsto (v_1, v_2)$ using at most 4 transvections. □

Theorem 2.5. *Every transformation on a symplectic geometry V is a product of symplectic transvections. In other words, the symplectic group $Sp_n(V)$ is generated by symplectic transvections.*

Proof. ([2]) The proof is by induction on the dimension of V. Since dim V is always even, we start with the case dim $V = 2$. Then $V =$ span (x, y) for some hyperbolic pair (x, y). Suppose we have a symplectic transformation $\tau : V \to V$ such that $(x, y) \mapsto (\tau x, \tau y)$. These are both hyperbolic pairs, so by Proposition 2.4, we have a product of symplectic transvections σ that maps $(x, y) \mapsto (\tau x, \tau y)$.

Now suppose dim $V = 2k$ and the result holds for all subspaces of lower dimension. From the orthogonal direct sum decomposition, we can write $V = H \textcircled{\perp} W$ by letting $W = H^{\perp}$ for some hyperbolic plane $H =$ span (x, y). Suppose we are given a symplectic transformation $\tau : V \to V$ such that $(x, y) \mapsto (\tau x, \tau y)$. Once again, we can use Proposition 2.4 to find a product of symplectic transvections σ such that $\sigma|_H = \tau|_H$.

Note that $\sigma^{-1}\tau$ is the identity transformation on H; i.e. $\sigma^{-1}\tau(H) = H$. By part (*iii*) of Theorem 1.32, we then know that $\sigma^{-1}\tau(H^{\perp}) = H^{\perp} \iff \sigma^{-1}\tau(W) = W$. Since dim $W <$ dim V, we can apply the induction hypothesis to W and find a product of symplectic transvections, ϕ, such that $\phi(W) = \sigma^{-1}\tau(W)$.

Finally, we take the product of transvections $\sigma\phi$, and observe that $\sigma\phi(W) = \tau(W)$ and $\sigma\phi(H) = \tau(H)$. Therefore, $\sigma\phi = \tau$ on V. □

Theorem 2.6. *Every element of $Sp_n(V)$ is a rotation, i.e., has determinant = 1.*

Proof. First observe that we can write $\sigma_{k,v} = \sigma_{k/2,v} \cdot \sigma_{k/2,v}$:

$$\begin{aligned}
\sigma_{k/2,v}\Big(\sigma_{k/2,v}(x)\Big) &= \sigma_{k/2,v}(x + k/2\langle v, x\rangle v) \\
&= (x + k/2\langle v, x\rangle v) + k/2\langle v, x + k/2\langle v, x\rangle v\rangle v \\
&= x + k/2\langle v, x\rangle v + k/2\langle v, x\rangle v + k^2/4\langle v, x\rangle\langle v, v\rangle v \\
&= x + k\langle v, x\rangle v \\
&= \sigma_{k,v}(x)
\end{aligned}$$

So we have:

$$\sigma_{k,v} = (\sigma_{k/2,v})^2$$
$$\implies \det\, \sigma_{k,v} = (\det\, \sigma_{k/2,v})^2$$

And if we assume det $\sigma_{k/2,v} = \pm 1$, we must have det $\sigma_{k,v} = 1$. □

3 Maximal Hyperbolic Subpsaces of an Orthogonal Geometry

We now come to some applications of Witt's Theorem. Recall the statement: given isometric orthogonal or symplectic geometries V, V' and an isometry $\phi : S \to \phi(S)$ (where $S \subset V$), the map ϕ can be extended to an isometry between V and V'.

Definition 2.7. A *maximal totally degenerate subspace* of V is a totally degenerate subspace that is not properly contained in any other totally degenerate subspace of V. Similarly, A *maximal hyperbolic subspace* of V is a hyperbolic subspace that is not properly contained in any other hyperbolic subspace of V.

Theorem 2.8. *All maximal totally degenerate subspaces of* V *have the same dimension. This dimension is called the Witt index of* V *and is denoted* $w(V)$.

Proof. Let S and S' be two maximal totally degenerate subspaces, and assume (towards a contradiction) that dim $S <$ dim S'. We can find a vector space isomophism from S into S' (e.g. by sending basis vectors of S to basis vectors of S'), given by $\sigma : S \to \sigma S \subset S'$. Since both S and S' are totally degenerate, σ is an isometry. But now we can apply Witt's theorem and extend this to an isometry $\sigma : V \to V$ (where σ has been renamed by abuse of notation). Then $S \subset \sigma^{-1}S'$, where $\sigma^{-1}S'$ is another totally degenerate subspace. But since S is maximal, we must have dim $S =$ dim $\sigma^{-1}S' =$ dim S', which is a contradiction. □

A similar result holds for *maximal hyperbolic subspaces* of V.

Theorem 2.9. *All maximal hyperbolic subspaces of* V *have dimension* $2w(V)$.

Proof. First we show that any two maximal hyperbolic subspaces have the same dimension. Let $\mathcal{H}_{2m} = H_1 \bigoplus H_2 \bigoplus \cdots \bigoplus H_m$ and $\mathcal{J}_{2n} = J_1 \bigoplus J_2 \bigoplus \cdots \bigoplus J_n$ be two maximal hyperbolic subspaces, where $H_i = \text{span } (u_i, v_i)$ and $J_i = \text{span } (x_i, y_i)$. Without loss of generality, assume (towards a contradiction) $m < n$. We can define an isomorphism (and moreover, an isometry) $\sigma : \mathcal{H}_{2m} \to \sigma(\mathcal{H}_{2m}) \subset \mathcal{J}_{2n}$ by setting $\sigma(u_i) = x_i$ and $\sigma(v_i) = y_i$. Using Witt's theorem, we can extend this to an isometry $\sigma : V \to V$ (σ has been renamed by abuse of notation). Then $\mathcal{H}_{2m} \subset \sigma^{-1}\mathcal{J}_{2n}$, where $\sigma^{-1}\mathcal{J}_{2n}$ is a hyperbolic space. But $\mathcal{H}_{2m}$ was a *maximal* hyperbolic space, so we must have dim $\mathcal{H}_{2m} =$ dim $\sigma^{-1}\mathcal{J}_{2n} =$ dim $\mathcal{J}_{2n}$, which is a contradiction because we assumed $m < n$.

Now we want to show that a maximal hyperbolic subspace of V has dimension $2w(V)$. Let $\mathcal{H}_{2m}$ be such a space. Note that $\mathcal{H}_{2m}$ contains a totally degenerate subspace S_m, and by Theorem 2.8, dim $S_m \leq w(V) \implies$ dim $\mathcal{H}_{2k} \leq 2w(V)$. On the other hand, any maximal totally degenerate subspace S_m can be extended to a hyperbolic space of dimension 2(dim S_m) $=$ $2w(V)$, so a *maximal* hyperbolic space must have dimension $\geq 2w(V)$. We conclude that dim $\mathcal{H}_{2m} = 2w(V)$. □

We will end with a stronger version of Theorem 1.23, which stated that a degenerate geometry V could be decomposed into its radical and a subspace S on which $\langle \; , \; \rangle$ is nondegenerate. Note that S can still contain vectors that are isotropic; in our next theorem, we decompose S even further by extracting its isotropic vectors.

Theorem 2.10. (Anisotropic decomposition) *Let V be an orthogonal geometry. Then we can write:*

$$V = Q \bigoplus \mathcal{H},$$

where Q is anisotropic and $\mathcal{H}$ is hyperbolic.

Proof. Let $\mathcal{H}$ be a maximal hyperbolic subspace of V. We can write $V = \mathcal{H} \bigoplus \mathcal{H}^{\perp}$. We claim that $\mathcal{H}^{\perp}$ is anisotropic; if there were an isotropic vector v in $\mathcal{H}^{\perp}$, then we could use Lemma 1.35 to extend the sum $\mathcal{H} \textcircled{\perp} \operatorname{span}(v)$ to a hyperbolic space that is strictly larger than $\mathcal{H}$. This would contradict the maximality of $\mathcal{H}$, so we conclude that $\mathcal{H}^{\perp}$ is anisotropic. Setting $Q = \mathcal{H}^{\perp}$ gives us our result. □

Bibliography

[1] D. S. Dummit and R. M. Foote. *Abstract Algebra.* John Wiley & Sons, Inc., 3rd edition, 2004.

[2] S. Roman. *Advanced Linear Algebra.* Springer, 3rd edition, 2007.

[3] M. Boij and D. Laksov. An introduction to algebra and geometry via matrix groups. Lecture Notes, Spring 1995. URL `http://www.math.kth.se/~laksov/courses/alggeom02/Main.pdf`.

[4] E. Artin. *Geometric Algebra.* Interscience Publishers, Inc., 1st edition, 1957.