Chapter IX. Tensors and Multilinear Forms.

IX.1. Basic Definitions and Examples.

1.1. Definition. *A* **bilinear form** *is a map $B : V \times V \rightarrow \mathbb{C}$ that is linear in each entry when the other entry is held fixed, so that*

$$\begin{aligned} B(\alpha x, y) &= \alpha B(x,y) = B(x, \alpha y) \\ B(x_1 + x_2, y) &= B(x_1, y) + B(x_2, y) \qquad \text{for all } \alpha \in \mathbb{F},\ x_k \in V,\ y_k \in V \\ B(x, y_1 + y_2) &= B(x, y_1) + B(x, y_2) \end{aligned}$$

(*This of course forces $B(x,y) = 0$ if either input is zero.*) *We say B is* **symmetric** *if $B(x,y) = B(y,x)$, for all x, y and* **antisymmetric** *if $B(x,y) = -B(y,x)$.*

Similarly a **multilinear form** (*aka a* **k-linear form** *, or a* **tensor of rank** *k*) *is a map $B : V \times \cdots \times V \rightarrow \mathbb{F}$ that is linear in each entry when the other entries are held fixed. We write $V^{(0,k)} = V^* \otimes \ldots \otimes V^*$ for the set of k-linear forms. The reason we use V^* here rather than V, and the rationale for the "tensor product" notation, will gradually become clear.*

The set $V^* \otimes V^*$ of bilinear forms on V becomes a vector space over $\mathbb{F}$ if we define

1. Zero element: $B(x,y) = 0$ for all $x, y \in V$;
2. Scalar multiple: $(\alpha B)(x,y) = \alpha B(x,y)$, for $\alpha \in \mathbb{F}$ and $x, y \in V$;
3. Addition: $(B_1 + B_2)(x,y) = B_1(x,y) + B_2(x,y)$, for $x, y \in V$.

When $k > 2$, the space of k-linear forms $V^* \otimes \ldots \otimes V^*$ is also a vector space, using the same definitions. The space of 1-linear forms (= *tensors of rank* 1 on V) is the dual space $V^* = \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$ of all $\mathbb{F}$-linear maps $\ell : V \rightarrow \mathbb{F}$. By convention the space of **0-forms** is identified with the ground field: $V^{(0,0)} = \mathbb{F}$; its elements are not mappings on V. It is also possible (and useful) to define multilinear forms of mixed type, mappings $\theta : V_1 \times \ldots \times V_k \rightarrow \mathbb{F}$ in which the components V_j are not all the same. These forms also constitute a vector space. We postpone any discussion of forms of "mixed type."

If $\ell_1, \ell_2 \in V^*$ we can create a bilinear form $\ell_1 \otimes \ell_2$ by taking a "*tensor product*" of these forms

$$\ell_1 \otimes \ell_2(v_1, v_2) = \langle \ell_1, v_1 \rangle \cdot \langle \ell_2, v_2 \rangle \qquad \text{for } v_1, v_2 \in V$$

Bilinearity is easily checked. More generally, if $\ell_1, \cdots, \ell_k \in V^*$ we obtain a k-linear map from $V \times \ldots \times V \rightarrow \mathbb{F}$ if we let

$$\ell_1 \otimes \ldots \otimes \ell_k(v_1, \cdots, v_k) = \prod_{j=1}^{k} \langle \ell_j, v_j \rangle \quad .$$

We will show that "*monomials*" of the form $\ell_1 \otimes \ldots \otimes \ell_k$ span the space $V^{(0,k)}$ of rank-k tensors, but they do not by themselves form a vector space except when $k = 1$.

1.2. Exercise. If $A : V \rightarrow V$ is any linear operator on a real inner product space verify that

$$\phi(v_1, v_2) = (Av_1, v_2) \qquad \text{for } v_1, v_2 \in V$$

is a bilinear form.

Note: This would not be true if $\mathbb{F} = \mathbb{C}$. Inner products on a complex vector space are conjugate-linear in their second input, with $(x, z \cdot y) = \overline{z} \cdot (x, y)$ for $z \in \mathbb{C}$; for $\mathbb{C}$-linearity in the second entry we would need $(x, z \cdot y) = z \cdot (x, y)$. However, $\overline{c} = c$ for real scalars so an inner product on a *real* vector space is a linear function of each input when the other is held fixed. □

1.3. Example. Let $A \in \mathrm{M}(n, \mathbb{F})$ and $V = \mathbb{F}^n$. Regarding elements of $\mathbb{F}^n$ as $n \times 1$ column vectors, define

$$B(x, y) = x^{\mathrm{t}} A y = \sum_{ij=1}^{n} x_i A_{ij} y_j$$

where x^{t} is the $1 \times n$ transpose of the $n \times 1$ column vector x. If we interpret the 1×1 product as a scalar in $\mathbb{F}$, then B is a typical bilinear form on $V = \mathbb{F}^n$. □

The analogous construction for multilinear forms is more complicated. For instance, to describe a rank-3 linear form $B(x, y, z)$ on $V \times V \times V$ we would need a three-dimensional $n \times n \times n$ array of coefficients $\{B_{i_1,i_2,i_3} : 1 \le i_k \le n\}$, from which we recover the original multilinear form via

$$B(x, y, z) = \sum_{i_1,i_2,i_3=1}^{n} x_{i_1} y_{i_2} z_{i_3} B_{i_1,i_2,i_3} \qquad \text{for } (x, y, z) \in \mathbb{F}^3 \ .$$

The coefficient array is an *ntimesn* square matrix only for bilinear form ($k = 2$). For the time being we will focus on bilinear forms, which are quite important in their own right.

Many examples involve symmetric or antisymmetric bilinear forms, and in any case we have the following result.

1.4. Lemma. *Every bilinear form B is uniquely the sum $B = B_+ + B_-$ of a symmetric and antisymmetric form.*

Proof: $B_\pm$ are given by

$$B_+(v_1, v_2) = \frac{B(v_1, v_2) + B(v_2, v_1)}{2} \qquad \text{and} \qquad B_- = \frac{B(v_1, v_2) - B(v_2, v_1)}{2} \ .$$

As for uniqueness, you can't have $B = B'$ with B symmetric and B' antisymmetric without both being the zero form. □

Variants. If V is a vector space over $\mathbb{C}$, a map $B : V \times V \to \mathbb{C}$ is **sesquilinear** if it is a linear function of its first entry when the other is held fixed, but is conjugate-linear in its second entry, so that

$$\begin{aligned} B(x_1 + x_2, y) = B(x_1, y) + B(x_2, y) \quad &\text{and} \quad B(x, y_1 + y_2) \ = \ B(x, y_1) + B(x, y_2) \\ B(\alpha x, y) = \alpha B(x, y) \quad &\text{and} \quad B(x, \alpha y) = B(x, y)\,\overline{\alpha} \ \text{ for all } \alpha \in \mathbb{C} \ . \end{aligned}$$

This is the same as bilinearity when $\mathbb{F} = \mathbb{R}$. The map is **Hermitian symmetric** if

$$B(y, x) = \overline{B(x, y)}$$

On a vector space over $\mathbb{R}$, an **inner product** is a special type of bilinear form, one that is *strictly positive definite* in the sense that

$$B(x, x) \ge 0 \text{ for all } x \in V \qquad \text{and} \qquad B(x, x) = \|x\|^2 = 0 \Rightarrow x = 0 \tag{32}$$

Over $\mathbb{C}$, an inner product is a map $B : V \times V \to \mathbb{C}$ that is sesquilinear, Hermitian symmetric, and satisfies the nondegeneracy condition (32).

A bilinear form $B \in V^* \otimes V^*$ is completely determined by its action on a basis $\mathfrak{X} = \{e_i\}$ via the matrix $[B]_{\mathfrak{X}} = [B_{ij}]$ with entries

$$B_{ij} = B(e_i, e_j) \quad \text{for } 1 \leq i, j \leq n$$

This matrix is symmetric/antisymmetric if and only if B has these properties. Given $[B]_{\mathfrak{X}}$ we recover B by writing $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$; then

$$\begin{aligned} B(x,y) &= B\big(\sum_i x_i e_i\,,\, \sum_j y_j e_j\big) \;=\; \sum_i x_i\, B\big(e_i, \sum_j y_j e_j\big) \\ &= \sum_{i,j} x_i B_{ij} y_j \;=\; [x]^{\mathrm{t}}_{\mathfrak{X}} [B]_{\mathfrak{X}} [y]_{\mathfrak{X}}\ , \end{aligned}$$

a 1×1 matrix regarded as an element of $\mathbb{F}$. Conversely, given a basis and a matrix $A \in \mathrm{M}(n, \mathbb{F})$ the previous equality determines a bilinear form B (symmetric if and only if $B = B^{\mathrm{t}}$ etc) such that $[B]_{\mathfrak{X}} = A$. Thus we have isomorphisms between vector spaces over $\mathbb{F}$:

1. The space of rank-2 tensors $V^{(0,2)} = V^* \otimes V^*$ is $\cong \mathrm{M}(n, \mathbb{F})$ via $B \to [B]_{\mathfrak{X}}$;
2. The space of *symmetric* bilinear forms is isomorphic to the space of symmetric matrices , etc.

We next produce a basis for $V^* \otimes V^*$ and determine its dimension.

1.5. Proposition. *If $\mathfrak{X} = \{e_i\}$ is a basis in a finite-dimensional vector space V, and $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis in V^* such that $\langle e_i^*, e_j\rangle = \delta_{ij}$, then the monomials $e_i^* \otimes e_j^*$ given by*

$$e_i^* \otimes e_j^*(v_1, v_2) = \langle e_i^*, v_1\rangle \cdot \langle e_j^*, v_2\rangle$$

are a bases on $V^* \otimes V^*$. Hence, $\dim(V^* \otimes V^*) = n^2$.

Proof: The monomials $e_i^* \otimes e_j^*$ span $V^* \otimes V^*$, for if B is any bilinear form and $B_{ij} = B(e_i, e_j)$, then $\tilde{B} = \sum_{i,j} B_{ij} e_i^* \otimes e_j^*$ has the same action on pairs $e_k, e_\ell \in V$ as the original tensor B.

$$\begin{aligned} \tilde{B}(e_k, e_l) &= \Big(\sum_{i,j} B_{ij} \cdot e_i^* \otimes e_j^*\Big)\langle e_k, e_\ell\rangle \;=\; \sum_{i,j} B_{ij} \langle e_i^*, e_k\rangle \cdot \langle e_j^*, e\ell\rangle \\ &= \sum_{i,j} B_{ij} \delta_{ik} \delta_{j\ell} = B_{k\ell} = B(e_k, e_\ell)\ , \end{aligned}$$

so $\tilde{B} = B \in \mathbb{F}\text{-span}\{e_i^* \otimes e_j^*\}$. As for linear independence, if $\tilde{B} = \sum_{i,j} b_{ij} e_i^* \otimes e_j^* = 0$ in $V^{(0,2)}$, then $\tilde{B}(x,y) = 0$ for all x, y, so $b_{k\ell} = \tilde{B}(e_k, e_\ell) = 0$ for $1 \leq k, \ell \leq n$. $\square$

A similar discussion shows that the space $V^{(0,r)}$ of rank-k tensors has dimension

$$\dim(V^{(0,r)}) = \dim(V^* \otimes \ldots \otimes V^*) = \dim(V)^r = n^r\ .$$

If $\mathfrak{X} = \{e_1, \ldots, e_n\}$ is a basis for V and $\{e_i^*\}$ is the dual basis in V^*, the monomials

$$e_{i_1}^* \otimes \ldots \otimes e_{i_r}^* \qquad 1 \leq i_1, \ldots, i_r \leq n$$

are a basis for $V^{(0,r)}$.

1.6. Theorem (Change of Basis) *Given $B \in V^* \otimes V^*$ and a basis $\mathfrak{X}$ in V, we describe B by its matrix via* (32). *If $\mathfrak{Y} = \{f_j\}$ is another basis, and if*

$$\mathrm{id}(f_j) = f_j = \sum_k s_{kj} e_k \quad \textit{for} \quad 1 \leq j \leq n\ , \tag{33}$$

then $S = [s_{ij}] = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}$ is the transition matrix for basis vectors and we have

$$\begin{aligned}
\big([B]_{\mathfrak{Y}}\big)_{ij} &= B(f_i, f_j) \;=\; B\Big(\sum_{k,\ell} s_{ki}e_k\,,\, \sum_{\ell} s_{\ell j}e_\ell\Big) \\
&= \sum_{k,\ell} S_{ki}B_{k\ell}S_{\ell j} \;=\; \sum_{k,\ell}(S^{\mathrm{t}})_{ik}B_{k\ell}S_{\ell j} \\
&= (S^{\mathrm{t}}[B]_{\mathfrak{X}}S)_{ij}
\end{aligned}$$

Note: We can also write this as $[B]_{\mathfrak{Y}} = P[B]_{\mathfrak{X}}P^{\mathrm{t}}$, taking $P = S^{\mathrm{t}} = [\mathrm{id}]^{\mathrm{t}}_{\mathfrak{X}\mathfrak{Y}}$. $\square$

Thus change of basis is effected by "*congruence*" of matrices $A \mapsto SAS^{\mathrm{t}}$, with $\det(S) \neq 0$. This differs considerably from the "*similarity transforms*" $A \mapsto SAS^{-1}$ that describe the effect of change of basis on the matrix of a linear operator $T : V \to V$. Notethet S^{t} is generally not equal to S^{-1}, so congruence and similarity are not the same thing. The difference between these concepts will emerge when we seek "normal forms" for various kinds of bilinear (or sesquilinear) forms.

1.7. Definition. *A bilinear form B is* **nondegenerate** *if*

$$B(v, V) = 0 \Rightarrow v = 0 \quad \textit{and} \quad B(V, v) = 0 \Rightarrow v = 0$$

If B is either symmetric or antisymmetric we only need the one-sided version. The **radical** *of B is the subspace*

$$\mathrm{rad}(B) = \{v \in V : B(v, v') = 0 \textit{ for all } v' \in V\}\ ,$$

*which measures the degree of degeneracy of the form B The B-***orthocomplement** *of a subspace $W \subseteq V$ is defined to be*

$$W^{\perp,B} = \{v \in V : B(v, W) = (0)\}\ .$$

Obviously, $W^{\perp,B}$ is a subspace. When B is symmetric or antisymmetric the conditions $B(v, W) = 0$ and $B(W, v) = 0$ yield the same subspace $B^{\perp,B}$. Then nondegeneracy means that $V^{\perp,B} = \{0\}$, and in general $V^{\perp,B}$ is equal to the radical of B.

1.8. Exercise (Dimension Formula). If B is nondegenerate and either symmetric or antisymmetric, and if $W \subseteq V$ is a subspace, prove that

$$\dim(W) + \dim\big(W^{\perp,B}\big) = \dim(V) \qquad \square.$$

The notion of "nondegeneracy" is a little ambiguous when the bilinear form B is neither symmetric nor antisymmetric: Is there a difference between "right nondegenerate," in which $B(V, y) = 0 \Rightarrow y = 0$, and nondegeneracy from the left: $B(x, V) = 0 \Rightarrow x = 0$? The answer is no. In fact if we view vectors $x, y \in V$ as $n \times 1$ columns, we may write $B(x, y) = [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[y]_{\mathfrak{X}}$, and if $[B]_{\mathfrak{X}}$ is singular there would be some $y \neq 0$ such that $[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} = 0$, hence $B(V, y) = 0$. That can't happen if B is right nondegenerate so B right-nondegenerate implies $[B]_{\mathfrak{X}}$ is nonsingular. The same argument shows B left-nondegenerate also implies $[B]_{\mathfrak{X}}$ nonsingular.

But in fact, this works in both directions, so

1.9. Lemma. *B is right nondegenerate if and only if $[B]_{\mathfrak{X}}$ is non singular.*

Proof: We have already proved ($\Leftarrow$) for both left- and right nondegeneracy. Conversely, if $B(V, y) = 0$ for some $y \neq 0$, then $[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} \neq 0$ if $\det([B]_{\mathfrak{X}}) \neq 0$, and we would have

$$B(e_i, y) = e_i^{\mathrm{t}}\,[B]_{\mathfrak{X}}\,[y]_{\mathfrak{X}} \neq 0$$

for some i. This conflicts with the fact that $[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} \neq 0$. Contradiction. $\square$

Thus for any basis $\mathfrak{X}$, B is right-nondegenerate $\Leftrightarrow$ $[B]_{\mathfrak{X}}$ is nonsingular $\Leftrightarrow$ left-nondegenerate, and it is legitimate to drop the "left/right" conditions on nondegeneracy.

Hereafter we will often abbreviate $\dim(V) = |V|$, which is convenient in this and other situations.

1.10. Lemma. *If B is a nondegenerate bilinear form on a finite dimensional space V, and M is a vector subspace, we let $M^{\perp,B} = \{w : B(V,w) = 0\}$. Then*

$$|M| + |M^{\perp,B}| = |V| \ ,$$

even though we need not have $M \cap M^{\perp,B} = (0)$.

Proof: If $|V| < \infty$ any nondegenerate bilinear form B mediates a natural bijection $J : V \rightarrow V^*$ that identifies each vector $v \in V$ with a functional $J(v)$ in V^* such that

$$\langle J(v), \ell\rangle = \langle \ell, v\rangle \qquad \text{for all } v \in V, \ell \in V^* \ .$$

This map is clearly $\mathbb{F}$-linear and $J(w) = 0 \Rightarrow B(V,w) = 0 \Rightarrow w = 0$ by non degeneracy of B, so J is one-to-one and also a bijection because $|V| = |V^*|$.

In Section III.3 of the *Linear Algebra I Course Notes*, we defined the "**annihilator**" of a subspace $M \subseteq V$ to be

$$M^\circ = \{\ell \in V^* : \langle \ell, M\rangle = 0\}$$

and discussed its properties, indicating that

$$(M^\circ)^\circ = M \qquad \text{and} \qquad |V| = |M| + |M^\circ|$$

when $|V| < \infty$. The annihilator M° is analogous to the orthogonal complement $M^\perp$ in an inner product space, but it lives in the dual space V^* instead of V; it has the advantage that M° makes sense in *any* vector space V, whether or not it is equipped with an inner product or a nondegenerate bilinear form. (Also, orthogonal complements $M^\perp$ depend on the particular inner product on V, while the annihilator M° has an absolute meaning.)

1.11. Exercise. When V is equipped with a nondegenerate bilinear form B we may invoke the natural isomorphism $V \cong V^*$ it induces to identify an annihilator M° in V^* with a uniquely defined subspace $J^{-1}(M^\circ)$ in V. From the definitions, verify that $M^\circ \subseteq V^*$ becomes the B-orthocomplement $M^{\perp,B} \subseteq V$ under these identifications. $\square$

1.12. Exercise. *If B is a nondegenerate bilinear form on a finite dimensional vector space, and if M* is any subspace, prove that

$$|M| + |M^{\perp,B}| = |V| \qquad \text{and} \qquad \left(M^{\perp,B}\right)^{\perp,B} = M. \tag{34}$$

Hint: Identifiying B-orthocomplements with annihilators, apply the basic properties of annihilators mentioned in Exercise 1.12. $\square$

If B is degenerate, so the radical $\mathrm{rad}(B)$ is nonzero, the role of the radical can be eliminated for most practical purposes, allowing us to focus on nondegenerate forms.

1.13. Exercise. Let $M = \mathrm{rad}(B)$ and form the quotient space $\tilde{V} = V/M$. Show that

1. B induces a well-defined bilinear form $\tilde{B} : \tilde{V} \times \tilde{V} \rightarrow \mathbb{F}$ if we let

$$\tilde{B}(x+M, y+M) = B(x,y) \qquad \text{for all } x, y \in V$$

2. $\tilde{B}$ is symmetric (or antisymmetric) $\Leftrightarrow$ B is.

3. Prove that $\tilde{B}$ is now nondegenerate on V/M. $\square$

1.14. Exercise. Given $n \times n$ matrices A, B show that

$$x^{\mathrm{t}} B y = x^{\mathrm{t}} A y \text{ for all } x, y \in \mathbb{F}^n \text{ if and only if } A = B. \quad \square$$

IX.2. Canonical Models for Bilinear Forms.

Bilinear forms arise often in physics and many areas of mathematics are concerned with these objects, so it is of some importance to find natural "*canonical forms*" for B that reveal its properties. This is analogous to the diagonalization problem for linear operators, and we will even speak of "diagonalizing" bilinear forms, although these problems are quite different and have markedly different outcomes.

In doing calculations it is natural to work with the matrices $[B]_{\mathfrak{X}}$ that represent B with respect to various bases, and seek bases yielding the simplest possible form. If a bilinear form B is represented by $A = [B]_{\mathfrak{X}}$ we must examine the effect of a change of basis $\mathfrak{X} \rightarrow \mathfrak{Y}$, and describe the new matrix $[B]_{\mathfrak{Y}}$ in terms of the transition matrix $S = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}$ that tells us how to write vectors in the $\mathfrak{Y}$-basis in terms of vectors in $\mathfrak{X}$, as in (32). Thus if $\mathfrak{X} = \{e_i\}$ and $\mathfrak{Y} = \{f_j\}$, $S = [s_{ij}]$ is the matrix such that

$$f_j = \sum_k s_{kj} e_k \qquad \text{for } 1 \leq j \leq n \tag{35}$$

Obviously $\det(S) \neq 0$ because this system of vector equations must be invertible.

In Theorem 1.6 we worked out the effect of such a basis change: $[B]_{\mathfrak{Y}} = S^{\mathrm{t}}[B]_{\mathfrak{X}}S$, which takes the form

$$[B]_{\mathfrak{Y}} = P[B]_{\mathfrak{X}}P^{\mathrm{t}} \qquad \text{if we set } P = S^{\mathrm{t}}\ . \tag{36}$$

We now show that the matrix of a nondegenerate B has a very simple standard form, at least when B is either symmetric or antisymmetric, the forms of greatest interest in applications. We might also ask whether these canonical forms are unique. (Answer: not very.)

The Automorphism Group of a Form B. If a vector space is equipped with a nondegenerate bilinear form B, a natural (and important) *automorphism group* $\mathrm{Aut}(B) \subseteq \mathrm{GL}_{\mathbb{F}}(V)$ comes along with it. It consists of the invertible linear maps $T : V \rightarrow V$ that "*leave the form invariant*," in the sense that $B\big(T(x), T(y)\big) = B(x, y)$ for all vectors. We have encountered such automorphism groups before, by various names. For example,

1. The **real orthogonal group** $\mathrm{O}(n)$ consists of the invertible linear maps T on $\mathbb{R}^n$ that preserve the usual inner product,

$$B(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i y_i \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n\ .$$

As explained in Section VI.5 of the *Linear Algebra I Notes*, the automorphisms that preserve this symmetric bilinear form are precisely the linear *rigid motions* on Euclidean space, those that leave invariant lengths of vectors and distances between them, so that

$$\|T(\mathbf{x})\| = \|\mathbf{x}\| \quad \text{and} \quad \|T(\mathbf{x}) - T(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

where $\|\mathbf{x}\| = \big(\sum_{i=1}^{n} |x_i|^2\big)^{1/2}$ (Pythagoras' formula).

2. The **unitary group** $\mathrm{U}(n)$ is the group of invertible linear operators on $V = \mathbb{C}^n$ that preserve the (Hermitian, sesquilinear) standard inner product

$$B(\mathbf{z}, \mathbf{w}) = \sum_{k=1}^{n} z_k \overline{w_k}$$

on complex n-space. For these operators the following conditions are equivalent (see *Linear Algebra I Notes*, Section VI.4).

$$\begin{aligned} T \in \mathrm{U}(n) \quad &\Leftrightarrow \quad B\big(T(\mathbf{z})\,,\, T(\mathbf{w})\big) = B(\mathbf{z}, \mathbf{w}) \\ &\Leftrightarrow \quad \|T(\mathbf{z})\| = \|\mathbf{z}\| \\ &\Leftrightarrow \quad \|T(\mathbf{z}) - T(\mathbf{w})\| = \|\mathbf{z} - \mathbf{w}\| \end{aligned}$$

for $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$, where

$$\|\mathbf{z}\| = B(\mathbf{z}, \mathbf{z})^{1/2} = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2} \quad \text{(Pythagoras' formula for complex } n\text{-space).}$$

2.1. Exercise. Explain why $\mathrm{U}(n)$ is a closed and bounded subset in matrix space $\mathrm{M}(n, \mathbb{C}) \cong \mathbb{C}^{n^2}$ $\square$

3. The **complex orthogonal group** $\mathrm{O}(n, \mathbb{C})$ is the automorphism group of the bilinear form on complex n-space $\mathbb{C}^n$

$$B(\mathbf{z}, \mathbf{w}) = \sum_{k=1}^{n} z_k w_k \qquad (\mathbf{z}, \mathbf{w} \in \mathbb{C}^n)$$

This is bilinear over $\mathbb{F} = \mathbb{C}$, but is *not* an inner product because it is not conjugate-linear in the entry $\mathbf{w}$ because w_k appears in B instead of $\overline{w_k}$; furthermore, not all vectors have $B(\mathbf{z}, \mathbf{z}) \geq 0$ (try $\mathbf{z} = (1, i)$ in $\mathbb{C}^2$).

In the present section we will systematically examine the canonical forms and associated automorphism groups for nondegenerate symmetric or antisymmetric forms over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. The number of possibilities is surprisingly small.

2.1A. Definition. *The* **automorphism group** *of a nondegenerate symmetric or antisymmetric form* $B : V \times V \to \mathbb{F}$ *is*

$$\mathrm{Aut}(B) = \{T \in \mathrm{GL}_{\mathbb{F}}(V) : B\big(T(v), T(w)\big) = B(v, w) \ \ \textit{for all } v, w \in V\}\ , \tag{37}$$

where $\mathrm{GL}_{\mathbb{F}}(V) = \{T : \det(T) \neq 0\}$ *is the* **general linear group** *consisting of all invertible* $\mathbb{F}$*-linear operators* $T : V \to V$*.*

$\mathrm{Aut}(B)$ is a *group* because it contains: the identity $I = \mathrm{id}_V$; the composition product $S \circ T$ of any two elements; and the inverse T^{-1} of any element.

Given a basis $\mathfrak{X}$ for V, each element $T \in \mathrm{Aut}(B)$ corresponds to an invertible matrix $[B]_{\mathfrak{X}} = [\,B(e_i, e_j)]$, and these matrices form a group

$$G_{B,\mathfrak{X}} = \{[T]_{\mathfrak{X}} : T \in \mathrm{Aut}(B)\}$$

under matrix multiplication $(\cdot)$. The group $(\mathrm{Aut}(B), \circ)$ and the matrix group $(G_{B,\mathfrak{X}}, \cdot\,)$ are isomorphic and are often identified.

Matrices in $G_{B,\mathfrak{X}}$ are characterized by their special algebraic properties,

$$G_{B,\mathfrak{X}} = \{E \in \mathrm{GL}(n, \mathbb{F}) : E^{\mathrm{t}}[B]_{\mathfrak{X}} E = [B]_{\mathfrak{X}}\}\ , \tag{38}$$

This identification follows because

$$\begin{aligned} T \in \mathrm{Aut}(B) \quad &\Leftrightarrow \quad B\big(T(x),T(y)\big) = B(x,y) \qquad \text{for all } x,y \in V \\ &\Leftrightarrow \quad [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} \ = \ [T(x)]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}\,[T(y)]_{\mathfrak{X}} \\ &\qquad\qquad\qquad\quad = \ [x]^{\mathrm{t}}_{\mathfrak{X}}\big([T]^{\mathrm{t}}_{\mathfrak{X}}\,[B]_{\mathfrak{X}}[T]_{\mathfrak{X}}\big)[y]_{\mathfrak{X}} \\ &\Leftrightarrow \quad [B]_{\mathfrak{X}} \ = \ [T]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}\,[T]_{\mathfrak{X}} \qquad \text{for all } x,y \in V\ . \end{aligned}$$

Given basis $\mathfrak{X}$, T is an automorphism of the bilinear form B if and only if the matrix $[T]_{\mathfrak{X}}$ satisfies the identity $[B]_{\mathfrak{X}} = [T]^{\mathrm{t}}_{\mathfrak{X}}[\,B]_{\mathfrak{X}}[T]_{\mathfrak{X}}$, and this must be true for any basis $\mathfrak{X}$. Matrices in $G_{B,\mathfrak{X}}$ are precisely the matrix realizations (with respect to basis $\mathfrak{X}$) of all the automorphisms in $\mathrm{Aut}(B)$.

2.2. Exercise. If B is a non degenerate bilinear form, show that $G_B = \mathrm{Aut}(B)$ is a subgroup in the general linear group $\mathrm{GL}_{\mathbb{F}}(V)$ – i.e. that (i) $I \in G_B$, (ii) $T_1, T_2 \in G_B \Rightarrow T_1, T_2 \in G_B$, and (iii) $T \in G_B \Rightarrow T^{-1} \in G_B$. $\square$

We can also assess the effect of change of basis $\mathfrak{X} \rightarrow \mathfrak{Y}$: $G_{B,\mathfrak{Y}}$ is a conjugate of $G_{B,\mathfrak{X}}$ under the action of $\mathrm{GL}(n,\mathbb{F})$.

2.3. Exercise. If $\mathfrak{X}$, $\mathfrak{Y}$ are bases in V, define $G_{B,\mathfrak{X}}$ and $G_{B,\mathfrak{Y}}$ as in (38) and prove that

$$G_{B,\mathfrak{Y}} = S^{-1} G_{B,\mathfrak{X}} S \qquad \text{where } S = [\mathrm{id}]_{\mathfrak{Y},\mathfrak{X}}$$

(or equivalently $G_{B,\mathfrak{Y}} = \tilde{S} G_{B,\mathfrak{X}} \tilde{S}^{-1}$ where $\tilde{S} = [\mathrm{id}]_{\mathfrak{Y},\mathfrak{X}}$ since $[\mathrm{id}]_{\mathfrak{Y},\mathfrak{X}} \cdot [\mathrm{id}]_{\mathfrak{X},\mathfrak{Y}} = I$). $\square$

Recall that S is the matrix such that $f_i = \sum_{k=1}^{n} s_{ji} e_j$ if $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_j\}$.

The general linear group $\mathrm{GL}_{\mathbb{F}}(V)$ in which all these automorphism groups live is defined by the condition $\det(T) \neq 0$, which makes no reference to a bilinear form. The **special linear group** $\mathrm{SL}_{\mathbb{F}}(V) = \{T \in \mathrm{GL}_{\mathbb{F}}(V) : \det(T) = 1\}$ is another "classical group" that does not arise as the automorphism group of a bilinear form B. All the other classical groups of physics and geometry are automorphism groups, or their intersections with $\mathrm{SL}_{\mathbb{F}}(V)$

Canonical Forms for Symmetric and Antisymmetric B. We classify the congruence classes of nondegenerate bilinear forms according to whether B is symmetric or antisymmetric, and whether the ground field is $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, always assuming B is nondegenerate. The analysis is the same for antisymmetric forms over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, so there are really only three cases to deal with.

Canonical Forms. Case 1: B symmetric, $\mathbb{F} = \mathbb{R}$.

If B is a nondegenerate symmetric bilinear form on a vector space over $\mathbb{R}$ with $\dim(V) = n$, there are $n+1$ possible canonical forms.

2.4. Theorem (B symmetric; $\mathbb{F} = \mathbb{R}$). *is an $\mathbb{R}$-basis $\mathfrak{X} \subseteq V$ such that the matrix describing B has the form*

$$(39) \qquad [B]_{\mathfrak{X}} = \begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} \qquad \textit{with } p + q = n = \dim(V)\ .$$

In this case, we say B has **signature** (p,q).

Proof: First observe that we have a polarization identity for symmetric B that determines $B(v,w)$ from homogheneous expressions of the form $B(u,u)$, just as with inner products over $\mathbb{R}$.

(40) Polarization Identity: $\quad B(v,w) = \frac{1}{2}\left[\, B(v+w, v+w) - B(v,v) - B(w,w) \right]$

for all $v, w \in V$.

2.5. Definition. *The map* $Q(v) = B(v,v)$ *from* V *to* $\mathbb{R}$ *is the* **quadratic form** *associated with a symmetric is bilinear form. Note that* $B(\lambda v, \lambda v) = \lambda^2 B(v,v)$*, and the quadratic form* $Q : V \to \mathbb{R}$ *determines the full bilinear form* $B : V \times V \to \mathbb{F}$ *via the polarization identity* (40).

Therefore, since $B \not\equiv 0$ there is some $v_1 \neq 0$, such that $B(v_1, v_1) \neq 0$, and after scaling v_1 by some $a \neq 0$ we can insure that $B(v_1, v_1) = \pm 1$. But because $\mathbb{F} = \mathbb{R}$ we can't control whether the outcome will be $+1$ or -1.

Let $M_1 = \mathbb{R} \cdot v_1$ and

$$M_1^{\perp, B} = \{v \in V : B(V, v_1) = 0\} \ .$$

We have $M_1 \cap M_1^{\perp, B} = \{0\}$ because any w in the intersection must have the form $w = c_1 v_1$, $c_1 \in \mathbb{R}$. But $w \in M_1^{\perp, B}$ too, so $0 = B(w,w) = c_1^2 B(v_1, v_1) = \pm c_1^2$, hence, $c_1 = 0$ and $w = 0$. Therefore $M_1 \oplus M_1^{\perp, B} = V$ because $|W| + |W^{\perp, B}| = |V|$ for any $W \subseteq V$ (Exercise 1.12). $\big[$For an alternative proof: recall the general result about the dimensions of subspaces W_1, W_2 in a vector space V: $|W_1 + W_2| = |W_1| + |W_2| - |W_1 \cap W_2|$.$\big]$

If B_1 is the restriction of B to $M_1^{\perp}$ we claim that $B_1 : M_1^{\perp, B} \times M_1^{\perp, B} \to \mathbb{R}$ is nondegenerate on the lower-dimensional subspace $M_1^{\perp, B}$. Otherwise, there would be an $x \in M_1^{\perp, B}$ such that $B\big(x, M_1^{\perp, B}\big) = 0$. But since $x \in M_1^{\perp, B}$ too, we also have $B(x, M_1) = 0$, and therefore by additivity of B in each entry,

$$B(x, V) = B(x\,,\, M_1^{\perp, B} + M_1) = 0 \ .$$

Nondegenerancy of B on V then forces $x = 0$.

We may therefore continue by induction on $\dim(V)$. Choosing a suitable basis $\mathfrak{X}' = \{v_2, \cdots, v_n\}$ in $M_1^{\perp, B}$ and $\mathfrak{X} = \{v_1, v_2, \ldots, v_n\}$ in V we get

$$[B]_{\mathfrak{X}} = \begin{pmatrix} \boxed{\pm 1} & \cdots & 0 \\ 0 & \boxed{I_{p \times p}} & \\ 0 & & \boxed{-I_{q \times q}} \end{pmatrix} \qquad \text{with } p + q = n - 1.$$

If the top left entry is -1, we may switch vectors $e_1 \leftrightarrow e_p$, which replaces $[B]_{\mathfrak{X}}$ with $[B]_{\mathfrak{Y}} = E^{\mathrm{t}} [B]_{\mathfrak{X}} E$, where E is the following permutation matrix (the zero on the diagonal is at the position p)

$$E = \begin{pmatrix} 0 & 0 & \cdot & \cdot & 1 & \cdot & \cdot & 0 \\ 0 & +1 & & & & & & 0 \\ \cdot & & \ddots & & & & & \\ \cdot & & & +1 & & & & \\ 1 & & & & 0 & & & \\ \cdot & & & & & -1 & & \\ \cdot & & & & & & \ddots & \\ 0 & 0 & & & & & & -1 \end{pmatrix}$$

(Note that $E^{\mathrm{t}} = E$ for this particular permutation matrix). Then $[B]_{\mathfrak{Y}}$ has the block-diagonal form (39), completing the proof. $\square$

Later on, we will describe an algorithmic procedure for putting B into canonical form $\mathrm{diag}(+1, \cdots, +1, -1, \cdots, -1)$; these algorithms work the same way over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We will also see that an antisymmetric B cannot be diagonalized by any congruence, but they do have a different (and equally useful) canonical form.

The Real Orthogonal Groups $\mathrm{O}(\mathbf{p},\mathbf{q})$, $\mathbf{p}+\mathbf{q}=\mathbf{n}$. The outcome in Theorem 2.4 breaks into $n+1$ possibilities. If $\mathfrak{X}$ is a basis such that $[B]_{\mathfrak{X}}$ has the standard form (39), then $A \in G_{B,\mathfrak{X}}$ if and only if

$$A^{\mathrm{t}} \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right) A = \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right) \tag{41}$$

This condition can be written concisely as $A^{\mathrm{t}}JA = J$ where $J = \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right)$.

The members of this family of classical matrix groups over $\mathbb{R}$ are denoted by $\mathrm{O}(p,q)$, and each one contains as a subgroup the **special orthogonal group of signature** (p,q),

$$\mathrm{SO}(p,q) = \mathrm{O}(p,q) \cap \mathrm{SL}(n,\mathbb{R}) \ .$$

Several of the groups $\mathrm{O}(p,q)$ and $\mathrm{SO}(p,q)$, are of particular interest.

The real Orthogonal Groups $\mathrm{O}(n,0) = \mathrm{O}(n)$ and $\mathrm{SO}(n)$. With respect to the standard basis in $\mathbb{R}^n$ we have $B_{\mathfrak{X}} = I_{n\times n}$, so $J = I_{n\times n}$ in (41) and

$$\mathrm{O}(n,0) = G_{B,\mathfrak{X}} = \{A : A^{\mathrm{t}}A = A^{\mathrm{t}}IA = I\} \ .$$

Thus $\mathrm{O}(n,0)$ is the familar group of orthogonal transformations on $\mathbb{R}^n$, traditionally denoted $\mathrm{O}(n)$. This group is a closed and bounded set in matrix space $\mathrm{M}(n,\mathbb{R}) \cong \mathbb{R}^{n^2}$. $\square$

The Lorentz Group $\mathrm{O}(n-1,1)$. This is the group of space-time symmetries at the center of Einstein's theory of special relativity for $n-1$ space dimensions $x_1,\ldots,x_{n-1}$ and one time dimension x_n which is generally labeled "t" by physicists. For a suitably chosen basis $\mathfrak{X}$ in $\mathbb{R}^n$ the matrix describing an arbitrary nondegenerate symmetric bilinear form B of signature $(n-1,1)$ becomes

$$[B]_{\mathfrak{X}} = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix} , \tag{42}$$

and the associated quadratic form is

$$B(x,x) = [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[x]_{\mathfrak{X}} = x_1^2 + \ldots + x_{n-1}^2 - x_n^2$$

Note: The physicists' version of this is a little different:

$$B(x,x) = x_1^2 + \ldots + x_{n-1}^2 - c^2t^2 \ ,$$

where c is the speed of light. But the numerical value of c depends on the physical units used to describe it – *feet per second*, etc – and one can always choose the units of (*length*) and (*time*) to make the experimentally measured speed of light have numerical value $c=1$. For instance we could take $t =$ (*seconds*) and measure lengths in (*light seconds*) = the distance a light ray travels in one second; or, we could measure t in (*years*) and lengths in (*light years*). Either way, the numerical value of the speed of light is $c=1$. $\square$

From (41) it is clear that A is in $\mathrm{O}(n-1,1)$ if and only if

$$A^{\mathrm{t}} \left(\begin{array}{c|c} I_{n-1} & 0 \\ \hline 0 & -1 \end{array}\right) A = \left(\begin{array}{c|c} I_{n-1} & 0 \\ \hline 0 & -1 \end{array}\right) \tag{43}$$

$O(n-1,1)$ contains the subgroup $SO(n-1,1) = O(n-1,1) \cap SL(n,\mathbb{R})$ of "*proper*" Lorentz transformations, those having determinant $+1$. Within $SO(n-1,1)$ we find a copy $\widetilde{SO}(n-1)$ of the standard orthogonal group $SO(n-1) \subseteq M(n-1,\mathbb{R})$, embedded in $M(n,\mathbb{R})$ via the one-to-one homomorphism

$$A \in \widetilde{SO}(n-1) \subseteq M(n-1,\mathbb{R}) \ \mapsto \ \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right) \in SO(n-1,1) \subseteq M(n,\mathbb{R}) \ .$$

The subgroup $\widetilde{SO}(n-1)$ acts only on the "space coordinates" $x_1, \cdots, x_{n-1}$ in $\mathbb{R}^n$, leaving the time coordinate $t = x_n$ fixed.

The following family of matrices in $O(n-1,1)$ is of particular interest in understanding the meaning of special relativity.

$$(44) \qquad A = \begin{pmatrix} 1/\sqrt{1-v^2} & 0 & & 0 & -v/\sqrt{1-v^2} \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 1 & 0 \\ -v/\sqrt{1-v^2} & 0 & \cdots & 0 & 1/\sqrt{1-v^2} \end{pmatrix}$$

When we employ units that make the speed of light $c = 1$, the parameter v must have values $|v| < 1$ to prevent the corner entries in this array from having physically meaningless imaginary values; as $v \rightarrow 1$ these entries blow up, so $SO(n-1,1)$ is indeed an unbounded set in matrix space $M(n,\mathbb{R})$.

In special relativity, an *event* is described by a point $(\mathbf{x}, t)$ in space-time $\mathbb{R}^{n-1} \times \mathbb{R}$ that specifies the location $\mathbf{x}$ and the time t at which the event occurred. Now suppose two observers are moving through space at constant velocity with respect to one another (*no acceleration* as time passes). Each will use his or her own frame of reference in observing an event to assign space-time coordinates to it. The matrix A in (44) tells us how to make the (relativistic) transition from the values $(\mathbf{x}, t)$ seen by Observer #1 to those recorded by Observer #2:[1]

$$\begin{pmatrix} \mathbf{x}' \\ t' \end{pmatrix} = A \cdot \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}$$

2.6. Exercise. Verify that the matrices in (44) all lie in $SO(n-1,1)$. Be sure to check that $\det(A) = +1$.

Note: Show that (41) $\Rightarrow \det(A)^2 = 1$, so $\det(A) = \pm 1$, and then argue that $\det(I) = 1$ and $\det(A)$ is a continuous function of the real-valued parameter $-1 < v < +1$. $\square$.

2.7. Exercise. Show that

$$B = \begin{pmatrix} \cosh(y) & 0 & 0 & \sinh(y) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(y) & 0 & 0 & \cosh(y) \end{pmatrix}$$

is in $SO(3,1)$ for all $y \in \mathbb{R}$. $\square$

A Final Remark about (44). If we work with physical units that do not make $c = 1$, as assumed in (44), we must replace "$\sqrt{1-v^2}$" everywhere it appears with

$$\sqrt{1 - \left(\frac{v}{c}\right)^2}$$

[1]To keep things simple, the transition matrix (44) describes what happens when Observer #2 is moving with velocity v in the positive x_1-direction, as seen by Observer #1, so that $x'_1 = x_1 - vt, x'_2 = x_2, \ldots, x'_{n-1} = x_{n-1}$. The general formula is more complicated.

in which the speed of light c appears explicitly $\square$.

Invariance of the Signature for $A \in \mathrm{O}(p,q)$. One way to compute the signature would be to find a basis that puts $[B]_{\mathfrak{X}}$ into the block-diagonal form (39), but how do we know the signature does not depend on the basis used to compute it? That it does not is the subject of the next theorem. Proving this amounts to showing that the signature is a **congruence invariant**: you cannot transform

$$\begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} \quad \text{to} \quad S^{\mathrm{t}} \begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} S = \begin{pmatrix} \boxed{I_{p'\times p'}} & 0 \\ 0 & \boxed{-I_{q'\times q'}} \end{pmatrix}$$

unless $p' = p$ and $q' = q$. This fact is often referred to as "*Sylvesters's Law of Intertia.*"

2.8. Theorem (Sylvester). *If A is a nondegenerate real symmetric $n \times n$ matrix, then there is some $P \in \mathrm{GL}(n,\mathbb{R})$ such that $P^{\mathrm{t}}AP = \mathrm{diag}(1,\cdots,1,-1,\cdots,-1)$. The number p of $+1$ entries and the canonical form* (39) *are uniquely determined.*

Proof: The existence of a diagonalization has already been proved. If $B(\mathbf{x},\mathbf{y}) = \sum_{i,j} x_i A_{ij} y_j = \mathbf{x}^{\mathrm{t}}A\mathbf{y}$ is a nondegenerate symmetric bilinear form on $\mathbb{R}^n$, so $[B] = [A_{ij}]$ with respect to the standard basis, then there is a basis $\mathfrak{X}$ such that $[B]_{\mathfrak{X}} = \mathrm{diag}(1,\cdots,1,-1,\cdots,-1)$. Suppose $p = \#(\text{entries} = +1)$ for $\mathfrak{X}$, and that there is another diagonalizing basis $\mathfrak{Y}$ such that $p' = \#(\text{entries} = +1)$ is $\neq p$. We may assume $p < p'$. Writing $\mathfrak{X} = \{v_1,\cdots,v_p,v_{p+1},\cdots,v_n\}$ and $\mathfrak{Y} = \{w_1,\cdots,w_{p'},w_{p'+1},\cdots,w_n\}$, define $L : V \rightarrow \mathbb{R}^{p-p'+n}$ via

$$L(\mathbf{x}) = \big(B(\mathbf{x},v_1),\cdots,B(\mathbf{x},v_p),B(\mathbf{x},w_{p'+1}),\cdots,B(\mathbf{x},w_n)\big)$$

The rank $\mathrm{rk}(L)$ of this linear operator is at most $\dim(\mathbb{R}^{p-p'+n}) = p - p' + n < n$, hence $\dim(\ker(L)) = \dim(V) - \mathrm{rk}(L) > 0$ and there is some $v_0 \neq 0$ in V such that $L(v_0) = 0$. That means

$$B(v_0,v_i) = 0 \quad \text{for } 1 \leq i \leq p \qquad \text{and} \qquad B(v_0,w_i) = 0 \quad \text{for } p'+1 \leq i \leq n\ .$$

Writing v_0 in terms of the two bases we have $v_0 = \sum_{j=1}^n a_j v_j = \sum_{k=1}^n b_k w_k$.

For $i \leq p$ we get

$$\begin{aligned} 0 = B(v_0,v_i) &= B\big(\sum_j a_j v_j\ ,\ v_i\big) = \sum_j a_j B(v_j,v_i) \\ &= \sum_j a_j \delta_{ij} = a_i = a_i B(v_i,v_i)\ , \end{aligned}$$

since $[B]_{\mathfrak{X}} = \mathrm{diag}(1,\cdots,1,-1,\cdots,-1)$. But $B(v_i,v_i) > 0$ for $i \leq p$ while $B(v_0,v_i) = 0$, so we conclude that $a_i = 0$ for $0 \leq i \leq p$. Similarly, $b_j = 0$ for $p'+1 \leq j \leq n$.

It follows that $a_i \neq 0$ for some $p' < i \leq n$, and hence

$$\begin{aligned} B(v_0,v_0) &= B\big(\sum_{j=1}^n a_j v_j\ ,\ \sum_{\ell=1}^n a_\ell v_\ell\big) = \sum_{j=1}^n a_j^2 B(v_j,v_j) \\ &= \sum_{j=p+1}^n a_j^2 B(v_j,v_j) < 0\ . \end{aligned}$$

Furthermore,

$$\begin{aligned} B(v_0,v_0) &= B\big(\sum_{j=1}^n b_j w_j\ ,\ \sum_{\ell=1}^n b_\ell w_\ell\big) = \sum_{j=1}^n b_j^2 B(w_j,w_j) \\ &= \sum_{j=1}^{p'} b_j^2 B(w_j,w_j) > 0\ . \end{aligned}$$

Thus $B(v_0, v_0) < 0$ and $B(v_0, v_0) \geq 0$, which is a contradiction. $\square$

2.9. Corollary. *Two non singular symmetric matrices in $M(n, \mathbb{R})$ are congruent via $A \to P^{\mathrm{t}}AP$ for some $P \in \mathrm{GL}(n, \mathbb{R})$ if and only if they have the same signature (p, q).*

Let A be a symmetric $n \times n$ matrix with entries from a field $\mathbb{F}$ not of characteristic two. We know that there are matrices $Q, D \in \mathrm{M}(n, \mathbb{F})$ such that Q is invertible and $Q^{\mathrm{t}}AQ = D$ is diagonal. We now give a method for computing suitable Q and diagonal form D via elementary row and column operations; a short additional step then yields the signature (p, q) when $\mathbb{F} = \mathbb{R}$.

The Diagonalization Algorithm. Recall that the effect of an elementary row operation on A is obtained by right multiplication $A \mapsto AE$ by a suitable "elementary matrix" E, as explained in *Linear Algebra I Notes*, Sections I-1 and IV-2. Furthermore, the same elementary operation on columns is effected by a left multiplication $A \mapsto E^{\mathrm{t}}A$ using the same E. If we perform an elementary operation on rows followed by the same elementary operation on columns, this is effected by

$$A \mapsto E^{\mathrm{t}}AE$$

(The order of the operations can be reversed because matrix multiplication is associative.)

Now suppose that Q is an invertible matrix such that $Q^{\mathrm{t}}AQ = D$ is diagonal. Any invertible Q is a product of elementary matrices, say $Q = E_1 E_2 \cdots E_k$, hence

$$D = Q^{\mathrm{t}}AQ = E_k^{\mathrm{t}} E_{k-1}^{\mathrm{t}} \cdot \ldots \cdot E_1^{\mathrm{t}} A E_1 E_2 \cdot \ldots \cdot E_k$$

Putting these observations together we get

2.10. Lemma. *A sequence of paired elementary row and column operations can transform any real symmetric matrix A into a diagonal matrix D. Furthermore, if $E_1, \cdots, E_k$ are the appropriate elementary matrices that yield the necessary row operations (indexed in the order performed), then $Q^{\mathrm{t}}AQ = D$ if we take $Q = E_1 E_2 \cdots E_k$.*

2.11. Example. *Let A be the symmetric matrix in* $\mathrm{M}(3, \mathbb{R})$

$$A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

We apply the procedure just described to find an invertible matrix Q such that $Q^{\mathrm{t}}AQ = D$ is diagonal.

Discussion: We begin by eliminating all of the nonzero entries in the first row and first column except for the entry a_{11}. To this end we start by performing the column operation $Col(2) \to Col(2) + Col(1)$; this yields a new matrix to which we apply the same operation on rows, $Row(2) \to Row(2) + Row(1)$. These first steps yield

$$A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 \\ -1 & 1 & 1 \\ 3 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 3 & 4 & 1 \end{pmatrix} = E_1^{\mathrm{t}} A E_1$$

where

$$E_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The second round of moves is: $Col(3) \to Col(3) - 3 \cdot Col(1)$ followed by $Row(3) \to Row(3) - 3 \cdot Row(1)$, which yields

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 3 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 3 & 4 & -8 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 4 & -8 \end{pmatrix} = E_2^{\mathrm{t}} E_1^{\mathrm{t}} A E_1 E_2$$

where

$$E_2 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Finally we achieve a diagonal form by applying $Col(3) \to Col(3) - 4 \cdot Col(2)$ and then the corresponding operation on rows to get

$$E_3^{\mathrm{t}} E_2^{\mathrm{t}} E_1^{\mathrm{t}} A E_1 E_2 E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -24 \end{pmatrix} \qquad \text{where} \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} .$$

Since the outcome is a diagonal matrix, the process is complete. To summarize: taking

$$Q = E_1 E_2 E_3 = \begin{pmatrix} 1 & 1 & -7 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{we get a diagonal form} \quad D = Q^{\mathrm{t}} A Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -24 \end{pmatrix}$$

To obtain the canonical form (39) we need one more pair of operations

$$Row(3) \to \frac{1}{\sqrt{24}} \cdot Row(3) \qquad \text{and} \qquad Col(3) \to \frac{1}{\sqrt{24}} \cdot Col(3) ,$$

both of which correspond to the (diagonal) elementary matrix

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{24}} \end{pmatrix} .$$

The canonical form is

$$\mathrm{diag}(1, 1, 1, -1) = \tilde{Q}^{\mathrm{t}} A \tilde{Q} \qquad \text{where} \qquad \tilde{Q} = E_4 \cdot \ldots \cdot E_1 \quad \square$$

This example also shows that the diagonal form of a real symmetric matrix achieved through congruence transformations $A \to Q^{\mathrm{t}} A Q$ is not unique; both $\mathrm{diag}(1, 1, 1, -24)$ and $\mathrm{diag}(1, 1, 1, -1)$ are congruent to A. Only the signature $(3, 1)$ is a true congruence invariant.

In Section IV-2 of the *Linear Algebra I Notes* we showed that the inverse A^{-1} of an invertible matrix can be obtained multiplying on the left by a sequence of elementary matrices (or equivalently, by executing the corresponding sequence of elementary row operations). We also developed the *Gauss-Seidel Algorithm* does this efficiently.

GAUSS-SEIDEL ALGORITHM. *Starting with the $n \times 2n$ augmented matrix $[A : I_{n \times n}]$, apply row operations to bring the left-hand block into reduced echelon form, which must equal $I_{n \times n}$ since A is invertible. Applying the same moves to the entire $n \times 2n$ augmented matrix we arrive at a matrix $[I_{n \times n} : A^{-1}]$ whose right-hand block is the desired inverse.*

An algorithm similar to Gauss-Seidel yields a matrix Q such that $Q^{\rm t}AQ = D$ is diagnonal; the signature (r,s) can then be determined by inspection as in the last steps of Example 2.11. The reader should justify the method, illustrated below, for computing an appropriate Q without recording each elementary matrix separately. Starting with an augmented $n\times 2n$ matrix $\left[A : I_{n\times n}\right]$, we apply paired row and column operations to drive the left-hand block into diagonal form; but we apply them to the *entire augmented matrix*. When the left-hand block achieves diagonal form D the right-hand block in $\left[\,D : Q^{\rm t}\right]$ is a matrix such that $Q^{\rm t}AQ = D$. The steps are worked out below; we leave the reader to verify that $Q^{\rm t}AQ = D$.

Details: Starting with $Col(2)\rightarrow Col(2)+Col(1)$ and then the corresponding operation on rows, we get

$$[A:I\,] \;=\; \left(\begin{array}{ccc|ccc} 1 & -1 & 3 & 1 & 0 & 0\\ -1 & 2 & 1 & 0 & 1 & 0\\ 3 & 1 & 1 & 0 & 0 & 1\end{array}\right) \xrightarrow{\text{paired R/C opns.}} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0\\ 0 & 1 & 4 & 1 & 1 & 0\\ 3 & 4 & 1 & 0 & 0 & 1\end{array}\right)$$

$$\xrightarrow{\text{paired R/C opns.}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 4 & 1 & 1 & 0\\ 0 & 4 & -8 & -3 & 0 & 1\end{array}\right)$$

$$\xrightarrow{\text{paired R/C opns.}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 1 & 1 & 0\\ 0 & 0 & -24 & -7 & -4 & 1\end{array}\right) \;\rightarrow\; [\,D:Q^{\rm t}]$$

Therefore,

$$Q^{\rm t} \;=\; \begin{pmatrix} 1 & 0 & 0\\ 1 & 1 & 0\\ -7 & -4 & 1\end{pmatrix} \qquad\qquad Q \;=\; \begin{pmatrix} 1 & 1 & -7\\ 0 & 1 & -4\\ 0 & 0 & 1\end{pmatrix}$$

and a diagonalized form $Q^{\rm t}AQ$ is

$$D \;=\; \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -24\end{pmatrix}$$

We now turn to the the next type of bilinear form to be analyzed.

Canonical Forms. Case 2: B symmetric, $\mathbb{F}=\mathbb{C}$.

In this case there is just one canonical form.

2.12. Theorem (B symmetric; $\mathbb{F}=\mathbb{C}$). *If B is a nondegenerate, symmetric bilinear form over $\mathbb{F}=\mathbb{C}$ there is a basis $\mathfrak{X}$ such that $[B]_{\mathfrak{X}} = I_{n\times n}$. In coordinates, for this basis we have*

$$B(x,y) = \sum_{j=1}^{n} x_j y_j \qquad (\textit{no conjugate, even though } \mathbb{F}=\mathbb{C})\ .$$

Proof: We know (by our discussion of $\mathbb{F}=\mathbb{R}$), we can put B in diagonal form $[B]_{\mathfrak{X}} = \mathrm{diag}(\lambda_1,\cdots,\lambda_n)$, with each $\lambda_i\neq 0$ since B is nondegenerate. Now take square roots in $\mathbb{C}$ and let $P=\mathrm{diag}(1/\sqrt{\lambda_1},\cdots,1/\sqrt{\lambda_n})$ to get $P^{\rm t}[B]_{\mathfrak{X}}P = I_{n\times n}$. $\square$

There is just one matrix automorphism group, modulo conjugations in $\mathrm{GL}(n,\mathbb{C})$. Taking a basis such that $[B]_{\mathfrak{X}} = I$, we get the **complex orthogonal group** in $\mathrm{M}(n,\mathbb{C})$,

$$\mathrm{O}(n,\mathbb{C}) = G_{B,\mathfrak{X}} = \{A\in \mathrm{M}(n,\mathbb{C}) : \det(A)\neq 0 \text{ and } A^{\rm t}A = I\}$$

(Note our use of the transpose $A^{\rm t}$ here, not the adjoint $A^* = \overline{A^{\rm t}}$, even though $\mathbb{F}=\mathbb{C}$. As a subgroup we have the **special orthogonal group** over $\mathbb{C}$,

$$\mathrm{SO}(n,\mathbb{C}) = \mathrm{O}(n,\mathbb{C})\cap \mathrm{SL}(n,\mathbb{C})$$

These are closed unbounded subsets in and $\mathrm{M}(n,\mathbb{C})$.

2.13. Exercise.

1. Show that $\mathrm{SO}(2,\mathbb{C})$ is abelian and isomorphic to the direct product group $S^1 \times \mathbb{R}$ where $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and the product operation is
$$(z,x)\cdot(z'x') = (zz', x+x')$$
2. Show that $A \in \mathrm{SO}(2,\mathbb{C})$ if and only if
$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$
with $a, b \in \mathbb{C}$ and $a^2 + b^2 = 1$.
3. Show that $\mathrm{SO}(2,\mathbb{C})$ is an unbounded subset in $\mathrm{M}(2,\mathbb{C})$, and hence that $\mathrm{SO}(n,\mathbb{C})$ is unbounded in $\mathrm{M}(n,\mathbb{C})$ because we may embed $\mathrm{SO}(2,\mathbb{C})$ in $\mathrm{SO}(n,\mathbb{C})$ via
$$A \in \mathrm{SO}(2,\mathbb{C}) \mapsto \left(\begin{array}{c|ccc} A & 0 & \cdot & 0 \\ \hline 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{array}\right)$$
if $n \geq 2$.

Hints: For (1.) you must produce an explicit bijection $\Phi : S^1 \times \mathbb{R} \to \mathrm{SO}(2,\mathbb{C})$ such that $\Phi(q_1,q_2) = \Phi(q_1)\cdot\Phi(q_2)$ (matrix product of elements in $\mathrm{M}(2,\mathbb{C})$). In (2.), if we write $A = [a,b;c,d]$ the identities $A^{\mathrm{t}}A = I = AA^{\mathrm{t}}$ plus $\det(A) = 1$ yield 9 equations in the complex unknowns a,b,c,d, which reduce to 7 when duplicate are deleted. But there is a lot of redundancy in the remaining system, and it can actually be solved by algebraic elimination despite its nonlinearity . In (3.) use the sup-norm $\|A\| = \max_{i,j}\{|A_{ij}|\}$ to discuss bounded sets in matrix space. □ □

Note: A similar problem was posed in the *Linear Algebra I Notes* regarding the group of real matrices $\mathrm{SO}(3) \subseteq \mathrm{M}(3,\mathbb{R})$ – see *Notes* Section VI-5, especially Euler's Theorem VI-5.6. The analog for $\mathrm{SO}(3)$ of the problem posed above for $\mathrm{SO}(2,\mathbb{C})$ i:wq is crucial in understanding the geometric meaning of the corresponding linear operators $L_A : \mathbb{R}^3 \to \mathbb{R}^3$. By Euler's Theorem $\mathrm{SO}(3)$ gets identified as the group of all *rotations* $R_{\ell,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$, by any angle θ about any oriented axis ℓ through the origin. □

2.14. Exercise. Is $\mathrm{SO}(n,\mathbb{C})$ a *closed subset* in $\mathrm{M}(n,\mathbb{C}) \simeq \mathbb{C}^{n^2}$? Prove or disprove. Which scalar matrices λI lie in $\mathrm{SO}(n,\mathbb{C})$ or $\mathrm{O}(n,\mathbb{C})$?

Canonical Forms. Case 3: B Antisymmetric; $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

In the antisymmetric case, the same argument applies whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Note that $B(v,v) = 0$ for all v, and if $W \subseteq V$ the B-annihilator $W^{\perp,B} = \{v : B(v,W) = 0\}$ need not be complementary to W. We might even have $W^{\perp,B} \supseteq W$, although the identity $\dim(W) + \dim(W^{\perp,B}) = \dim(V)$ remains valid.

2.15. Theorem (*B* antisymmetric; $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). *If B is a nondegenerate antisymmetric form over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, there is a basis $\mathfrak{X}$ such that*
$$[B]_{\mathfrak{X}} = J = \begin{pmatrix} 0 & I_{m\times m} \\ -I_{m\times m} & 0 \end{pmatrix}$$
In particular $\dim_{\mathbb{F}}(V)$ *must be even if* V *carries a nondegenerate skew-symmetric bilinear form.*

Proof: Recall that $\dim(W) + \dim\left(W^{\perp,B}\right) = \dim(V)$ for any nondegenerate bilinear form B on V. Fix $v_1 \neq 0$. Then $M_1 = (\mathbb{F}v_1)^{\perp,B}$ has dimension $n-1$ if $\dim(V) = n$, but it includes $\mathbb{F}v_1 \subseteq (\mathbb{F}v_1)^{\perp,B}$. Now take any $v_2 \notin M_1$ (so $v_2 \neq 0$) and scale it to get $B(v_1, v_2) = -1$. Let $M_2 = (\mathbb{F}v_2)^{\perp,B}$; again we have $\dim(M_2) = n-1 = \dim(M_1)$. But $M_2 \neq M_1$ since $v_2 \in M_2$ and $v_2 \notin M_1$, so $\dim(M_1 \cap M_2) = n-2$. The space $M = M_1 \cap M_2$ is B-orthogonal to $\mathbb{F}$-span$\{v_1, v_2\}$ by definition of these vectors. Furthermore, $B|_M$ is antisymmetric and nondegenerate. [In fact, we already know that $B(w, w_1) = B(w, v_2) = 0$ and $V = \mathbb{F}v_1 \oplus \mathbb{F}v_2 \oplus M$, so if $B(w, M) = 0$ for some $w \in M$, then $B(w, V) = B(w, \mathbb{R}v_1 + \mathbb{R}v_2 + M) = 0$ and $w = 0$ by nondegeneracy.] Furthermore, if $N = \mathbb{F}$-span$\{v_1, v_2\}$ we have $V = N \oplus M$. (Why?)

We can now argue by induction on $n = \dim(V)$: $\dim(M)$ must be even and there is a basis $\mathfrak{X}_0 = \{v_3, \cdots, v_n\}$ in M such that

$$[\,B|_M\,]_{\mathfrak{X}_0} = \begin{pmatrix} \boxed{R} & & 0 \\ & \ddots & \\ 0 & & \boxed{R} \end{pmatrix}$$

with

$$R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Hence, $\mathfrak{X} = \{v_1, v_2\} \cup \mathfrak{X}_0$ is a basis for V such that

$$[B]_{\mathfrak{X}} = \left(\begin{array}{c|c} R & 0 \\ \hline 0 & \left[\,B|_M\right]_{\mathfrak{X}_0} \end{array}\right) = \begin{pmatrix} \boxed{R} & & & 0 \\ & \boxed{R} & & \\ & & \ddots & \\ 0 & & & \boxed{R} \end{pmatrix}$$

A single permutation of basis vectors (corresponding to some permutation matrix E such that $E^{\mathrm{t}} = E^{-1}$) gives the standard form

$$E^{\mathrm{t}}[B]_{\mathfrak{X}}E = [B]_{\mathfrak{Y}} = \left(\begin{array}{c|c} 0 & I_{m\times m} \\ \hline -I_{m\times m} & 0 \end{array}\right)$$

where $m = \frac{1}{2}\dim(V)$. $\square$

A skew-symmetric nondegenerate form B is called a **symplectic structure** on V. The dimension $\dim_{\mathbb{F}}(V)$ must be even, and as we saw earlier there is just one such nondegenerate structure up to congruence of the representative matrix.

2.16. Definition. *The automorphism group* $\mathrm{Aut}(B)$ *of a nondegenerate skew-symmetric form on* V *is called a* **symplectic group**. *If* $\mathfrak{X}$ *is a basis that puts* B *into standard form, we have*

$$B(x, y) = [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} = [x]^{\mathrm{t}}_{\mathfrak{X}}J\,[y]_{\mathfrak{X}} \qquad \text{where} \qquad J = \begin{pmatrix} 0 & I_{m\times m} \\ -I_{m\times m} & 0 \end{pmatrix}\ .$$

By (38)*, elements of* $\mathrm{Aut}(B)$ *are determined by the condition*

$$A \text{ is in } G_{\mathfrak{X},B} \Leftrightarrow A^{\mathrm{t}}JA = J.$$

on $V \simeq \mathbb{R}^{2m}$. *The corresponding matrix group*

$$\mathrm{Sp}(n, \mathbb{F}) = G_{B,\mathfrak{X}} = \{A \in \mathrm{M}(n, \mathbb{F}) : A^{\mathrm{t}}JA = J\}$$

is the classical **symplectic group of degree** $m = \frac{1}{2}\dim(V)$.

The related matrix

$$J' = \begin{pmatrix} \boxed{R} & & 0 \\ & \ddots & \\ 0 & & \boxed{R} \end{pmatrix} \qquad \textit{with} \qquad R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

is a GL*-conjugate of* J*, with* $J' = CJC^{-1}$ *for some* $C \in \mathrm{GL}(2m, \mathbb{R})$*, and the algebraic condition*

$$A^{\mathrm{t}} J' A = J'$$

determines a subgroup $G' \subseteq \mathrm{GL}(n, \mathbb{F})$ *that is conjugate* (*hence isomorphic to*) *the matrix group* $G_{B,\mathfrak{X}} = \mathrm{Sp}(n, \mathbb{F})$*.*

Both versions of the commutation relations determining matrix versions of $\mathrm{Aut}(B)$ are used in the literature.

Note: $\det(A) \neq 0$ automatically because $\det(J) = (-1)^m \neq 0$. In fact, $A \in \mathrm{Sp}(n, \mathbb{F})$ implies $\det(J) = \det(A^{\mathrm{t}} J A) \Rightarrow (\det(A))^2 = 1$, so $\det(A) = \pm 1$ whether the underlying field $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. $\square$

The only scalar matrices λI in $\mathrm{Sp}(n, \mathbb{F})$ are those such that $\lambda^2 = 1$. The fact that $\det(J) = (-1)^m$ follows because m row transpositions send $J \to I_{2m \times 2m}$.

IX-3. Sesquilinear Forms ($\mathbb{F} = \mathbb{C}$).

Finally we take up sesquilinear forms $B : V \times V \to \mathbb{C}$ (over complex vector spaces), which are linear functions of the first entry in $B(v, w)$, but conjugate-linear in the second, so that $B(x, \lambda y) = \overline{\lambda} B(x, y)$, $B(\lambda x, y) = \lambda B(x, y)$. There are only a limited number of possibilities.

3.1. Lemma. *A sesquilinear form on* V *cannot be symmetric or antisymmetric unless it is zero.*

Proof: We know that $\lambda B(x, y) = B(\lambda x, y)$, and if B is (anti-)symmetric this would be equal to $\pm B(x, \lambda y) = \pm \overline{\lambda} B(x, y)$ for all $\lambda \in \mathbb{C}$, $x, y \in V$. This is impossible if $B(x, y) \neq 0$. $\square$

Thus the only natural symmetry properties for sesquilinear forms over $\mathbb{C}$ are

1. Hermitian symmetry: $B(x, y) = \overline{B(y, x)}$
2. Skew-hermitian symmetry: $B(x, y) = -\overline{B(y, x)}$.

However, if B is Hermitian then iB (where $i = \sqrt{-1}$) is skew-Hermitian and vice-versa, so once we analyze Hermitian sesquilinear forms there is nothing new to say about skew-Hermitian forms.

The sesquilinear forms on V are a vector space over $\mathbb{C}$. Every such form is uniquely a sum $B = B_H + B_S$ of a Hermitian and skew-Hermitian form

$$B(v, w) = \frac{B(v, w) + \overline{B(w, v)}}{2} + \frac{B(v, w) - \overline{B(w, v)}}{2} \qquad \text{for all } v, w \in V$$

As usual, a sesquilinear form B is determined by its matrix representation relative to a basis $\mathfrak{X} = \{e_1, \ldots, e_n\}$ in V, given by

$$[B]_{\mathfrak{X}} = [B_{ij}] \quad \text{where } B_{ij} = B(e_i, e_j) \ .$$

Given any basis $\mathfrak{X}$, the form B is

1. Nondegenerate if and only if $[B]_{\mathfrak{X}}$ is nonsingular (nonzero determinant).

2. Hermitian symmetric if and only if $[B]_{\mathfrak{X}}$ is self-adjoint ($= [B]^*_{\mathfrak{X}}$).
3. The correspondence $B \mapsto [B]_{\mathfrak{X}}$ is a $\mathbb{C}$-linear isomorphism between the vector space of sesquilinear forms on V and matrix space $\mathrm{M}(n, \mathbb{C})$.

The change of basis formula is a bit different from that for bilinear forms. If $\mathfrak{Y} = \{f_j\}$ is another basis, related to $\mathfrak{X} = \{e_i\}$ via

$$f_i = \sum_{j=1}^{n} s_{ji} e_j \qquad \text{where} \qquad S = [\mathrm{id}]_{\mathfrak{X},\mathfrak{Y}} \ .$$

we then have

$$\begin{aligned} \big([B]_{\mathfrak{Y}}\big)_{ij} &= B(f_i, f_j) \ = \ B\Big(\sum_k s_{ki} e_k \, , \, \sum_{\ell} s_{\ell j} e_{\ell}\Big) \\ &= \sum_{k,l} s_{ki} \, \overline{s_{\ell j}} \, ([B]_{\mathfrak{X}})_{k\ell} \\ &= (S^{\mathrm{t}}[B]_{\mathfrak{X}} \overline{S})_{ij} \quad \text{where } \overline{S} \text{ is the complex conjugate matrix: } \overline{S}_{ij} = \overline{s_{ij}} \end{aligned}$$

Letting $P = \overline{S}$, we may rewrite the result of this calculation as

$$[B]_{\mathfrak{Y}} = P^*[B]_{\mathfrak{X}} P \tag{45}$$

where $\det(P) \neq 0$, $P^* = (\overline{P})^{\mathrm{t}}$. In terms of the transition matrix S between bases, we have $P = \overline{S} = \overline{[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}}$.

Note that P^* need not to be equal to P^{-1}, so P need not be a unitary matrix in $\mathrm{M}(n, \mathbb{C})$. Formula (45) differs from that for orthogonal matrices in that P^{t} has been replaced by P^*.

3.2. Exercise. If B is sesquilinear, $\mathfrak{X}$ is a basis in V, and $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$ in V, show that

$$B(x, y) = [x]^{\mathrm{t}}_{\mathfrak{X}} [B]_{\mathfrak{X}} [y]^{-}_{\mathfrak{X}}, \quad \text{so that } B(x, y) = \sum_{ij} x_i B_{ij} \overline{y_j} \ . \quad \square$$

3.3. Definition. *A non degenerate sesquilinear form is an* **inner product** *if*

1. Hermitian: *$B(x, y) = \overline{B(y, x)}$;*
2. Positive Definite: *$B(x, x) \geq 0$, $\forall x$*
3. Nondegenerate:*: $B(x, V) = (0) \Leftrightarrow x = 0$.*

Conditions 2. + 3. amount to saying $B(x, x) \geq 0$ and $B(x, x) = 0 \Rightarrow x = 0$ – i.e. the form *strictly* positive definite. This equivalence follows from the polarization identity for Hermitian sesquilinear forms.

3.4. Lemma (Polartization Identity). *If B is a Hermitian sesquilinear form then*

$$B(v, w) = \frac{1}{4}\Big[\sum_{k=0}^{3} i^k B(v + i^k w, v + i^k w)\Big], \qquad \textit{where } i = \sqrt{-1}$$

Proof: Trivial expansion of the sum. $\square$

If B is a nondegenerate Hermitian sesquilinear form and $v \neq 0$ there must be some $w \in V$ such that $B(v, w) \neq 0$, but by the polarization identity nondegeneracy of B

implies that there is some $v \neq 0$ such that $B(v,v) \neq 0$ (and if B is positive definite it must be strictly positive definite). If v_1 is such a vector and $M_1 = \mathbb{R}v_1$, we obviously have $M_1 \cap M_1^{\perp} = (0)$ because $w \in M_1 \cap M_1^{\perp} \Rightarrow w = cv_1$ and also $0 = (w, v_1) = c(v_1, v_1)$, which implies $c = 0$. The restricted form $B|_{M_1^{\perp}}$ is again Hermitian symmetric; it is also nondegenerate because if $B(w, M_1^{\perp}) = 0$ for some nonzero $w \in M_1^{\perp}$, then $B(w, V) = (0)$ too, contrary to nondegeneracy of B on V. So, by an induction argument there is a basis $\mathfrak{X} = \{e_1 = v_1, e_2, \cdots, e_n\}$ in V such that

$$[B]_{\mathfrak{X}} = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$$

where $\mu_k \in \mathbb{C}$ and $\mu_k \neq 0$ (B being non degenerate).

Since $B(e_i, e_j) = \overline{B(e_j, e_i)}$ we get $\mu_k = \overline{\mu_k}$, so all entries are real and nonzero. Taking $P = \mathrm{diag}\big(1/\sqrt{|\mu_1|}, \cdots, 1/\sqrt{|\mu_n|}\,\big)$, we see that

$$P^*[B]_{\mathfrak{X}}P = \begin{pmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \end{pmatrix}$$

$= [B]_{\mathfrak{Y}}$ for some new basis $\mathfrak{Y}$; recall the change of basis formula.) Finally apply a permutation matrix (relabel basis vectors) to get

$$(46) \qquad [B]_{\mathfrak{Y}} = E^*P^*[B]_{\mathfrak{X}}PE = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}$$

where $P = I_{p\times p}$, $Q = -I_{q \times q}$, and $p + q = n = \dim_{\mathbb{C}}(V)$. We have proved

3.5. Proposition. *Every nondegenerate Hermitian sesquilinear form B can be put into the canonical form* (46) *by a suitable choice of basis in V. If $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$ with respect to a basis such that $[B]_{\mathfrak{X}}$ has canonical form, we get*

$$B(x,y) = \sum_{i=1}^{p} x_i\overline{y_i} - \sum_{i=p+1}^{n} x_i\overline{y_i}$$

In particular, if $p = n$ and $q = 0$ we obtain the standard inner product $(x,y) = \sum_{j=1}^n x_j\overline{y_j}$ in $\mathbb{C}^n$ when we identify V with $\mathbb{C}^n$ using the basis $\mathfrak{X}$ such that $[B]_{\mathfrak{X}}$ has the form (46).

There are just $n+1$ $\mathfrak{X}$-congruence classes of nondegenerate Hermitian sesquilinear forms on a complex vector space of dimension n; they are distinguished by their signatures (p,q). The possible automorphism groups

$$\mathrm{Aut}(B) = \{T \in \mathrm{Hom}_{\mathbb{C}}(V,V) : \det(T) \neq 0 \text{ and } B(T(v), T(w)) = B(v,w) \text{ for all } v, w\}$$

are best described as matrix groups $G_{B,\mathfrak{X}}$ with respect to a basis that puts B into canonical form. This yields the **unitary groups of type (p,q)**. $\mathrm{Aut}(B)$ is isomorphic to the matrix group

$$(47) \qquad \mathrm{U}(p,q) \;=\; \{A \in \mathrm{GL}(n,\mathbb{C}) : A^*JA = J\} \quad \text{where} \quad J = \begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{I_{q\times q}} \end{pmatrix}$$

There is a slight twist in the correspondence between operators $T \in \mathrm{Aut}(B)$ and matrices $A \in \mathrm{U}(p,q)$.

3.6. Exercise. Let B be nondegenerate Hermitian sesquilinear and let $\mathfrak{X} = \{e_i\}$ be a basis such that $[B]_{\mathfrak{X}}$ is in canonical form. If $[T]_{\mathfrak{X}}$ is the matrix associated with $T \in \mathrm{Aut}(B)$, verify that the complex conjugate $A = \big([T]_{\mathfrak{X}}\big)^{-}$ satisfies the identity (47), and conversely if $A \in \mathrm{U}(p,q)$ then $A = \big([T]_{\mathfrak{X}}\big)^{-}$ for some $T \in \mathrm{Aut}(B)$. $\square$

Thus the correspondence $\Phi : T \mapsto A = \big([T]_{\mathfrak{X}}\big)^{-}$ (rather than $T \mapsto A = [A]_{\mathfrak{X}}$) is a bijection between $\mathrm{Aut}(B)$ and the matrix group $\mathrm{U}(p,q) \subseteq \mathrm{GL}(n,\mathbb{C})$ such that $\Phi(T_1 \circ T_2) = \Phi(T_1) \cdot \Phi(T_2)$ (matrix product), and Φ is an group isomorphism between $\mathrm{Aut}(B)$ and $\mathrm{U}(p,q)$.

When $p = n$, we get the classical group $\mathrm{U}(n)$ of unitary operators on an inner product space, and when we identify $V \simeq \mathbb{C}^n$ via a basis such that $[B]_{\mathfrak{X}} = I_{n\times n}$, we get the group of **unitary matrices** in $\mathrm{M}(n,\mathbb{C})$,

$$\mathrm{U}(n) = \mathrm{U}(n,0) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^*A = I\} \quad \text{(because } A^*IA = A^*A)$$

As a closed subgroup of $\mathrm{U}(n)$ we have the **special unitary group**

$$\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C}) \subseteq \mathrm{U}(n) \ .$$

There are also special unitary group of type (p,q), the matrix groups

$$\mathrm{SU}(p,q) = \mathrm{U}(p,q) \cap \mathrm{SL}(n,\mathbb{C}) \ .$$

For $A \in \mathrm{U}(p,q)$ the identity (46) implies

$$\det(A^*) \cdot \det \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array} \right) \cdot \det(A) \ = \ (-1)^q$$

so $|\det(A)|^2 = (-1)^q$ (remember: $\mathbb{F} = \mathbb{C}$ so this could be negative). In particular, $|\det(A)|^2 = 1$ if $A \in \mathrm{U}(n)$. We already know that unitary matrices are orthogonally diagonalizable since they are normal operators ($A^*A = AA^*$, so $A^*A = I \Leftrightarrow AA^* = I$). Since $\|Ax\|^2 = \|x\|^2$ for all x, all eigenvalues λ_i have absolute value 1, so the specturm $\mathrm{sp}(A)$ is a subset of the unit circle $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ for unitary matrices (or operators). $\mathrm{U}(n)$ contains a copy of the unit circle (which is a group under the usual multiplication of complex number because $|zw| = |z| \cdot |w|$ and $|z| = 1 \Rightarrow |1/z| = 1$); namely $(S^1, \cdot) \cong \{\lambda I_{n\times n} : |\lambda| = 1\}$. In $\mathrm{SU}(n)$, however, the only scalar matrices are of the form λI where λ is an n^{th} root of unity, $\lambda = e^{2\pi i k/n}$ with $0 \le k \le n$.

Notice the parallel between certain groups over $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$.

1. $\mathrm{SO}(p,q)$ and $\mathrm{O}(p,q)$ over $\mathbb{R}$ are the "real parts" of $\mathrm{SU}(p,q)$ and $\mathrm{U}(p,q)$. In fact we have
$$\mathrm{O}(p,q) = \mathrm{U}(p,q) \cap \big(\mathrm{M}(n,\mathbb{R}) + i0\big) \ .$$
when we identity $\mathrm{M}(n,\mathbb{C}) = \mathrm{M}(n,\mathbb{R}) + \sqrt{-1}\,\mathrm{M}(n,\mathbb{R})$ by splitting a complex matrix $A = [z_{ij}]$ as $[x_{ij}] + \sqrt{-1}\,[y_{ij}]$ if $z_{ij} = x_{ij} + \sqrt{-1}\,y_{ij}$.

2. We also recognize $\mathrm{SO}(n)$ and $\mathrm{O}(n)$ as the real parts of the complex natrix groups $\mathrm{SO}(n,\mathbb{C})$ and $\mathrm{O}(n,\mathbb{C})$, as well as being the real parts of $\mathrm{SU}(n)$ and $\mathrm{U}(n)$.

3.7. Exercise. Prove that $\mathrm{U}(n)$ is a closed bounded subset when we identify $\mathrm{M}(n,\mathbb{C}) \approx \mathbb{C}^{n^2}$; hence it is a compact matrix group. $\square$

3.8. Exercise. If $p \neq n$, prove that $\mathrm{U}(p,q)$ and $\mathrm{SU}(p,q)$ are closed but unbounded subsets in $\mathrm{M}(n,\mathbb{C})$ when $q \neq 0$. $\square$