Chapter 7

Bilinear Forms

I presume that to the uninitiated
the formulae will appear cold and cheerless.

Benjamin Pierce

1. DEFINITION OF BILINEAR FORM

Our model for bilinear forms is the dot product

(1.1) $$(X \cdot Y) = X^t Y = x_1 y_1 + \cdots + x_n y_n$$

of vectors in $\mathbb{R}^n$, which was described in Section 5 of Chapter 4. The symbol $(X \cdot Y)$ has various properties, the most important for us being the following:

(1.2) *Bilinearity:*
$$(X_1 + X_2 \cdot Y) = (X_1 \cdot Y) + (X_2 \cdot Y)$$
$$(X \cdot Y_1 + Y_2) = (X \cdot Y_1) + (X \cdot Y_2)$$
$$(cX \cdot Y) = c(X \cdot Y) = (X \cdot cY)$$

Symmetry: $$(X \cdot Y) = (Y \cdot X)$$

Positivity: $$(X \cdot X) > 0, \quad \text{if } X \neq 0.$$

Notice that bilinearity says this: If one variable is fixed, the resulting function of the remaining variable is a linear transformation $\mathbb{R}^n \longrightarrow \mathbb{R}$.

We will study dot product and its analogues in this chapter. It is clear how to generalize bilinearity and symmetry to a vector space over any field, while positivity is, a priori, applicable only when the scalar field is $\mathbb{R}$. We will also extend the concept of positivity to complex vector spaces in Section 4.

Let V be a vector space over a field F. A *bilinear form* on V is a function of two variables on V, with values in the field: $V \times V \xrightarrow{f} F$, satisfying the bilinear axioms, which are

(1.3)
$$\begin{aligned} f(v_1 + v_2, w) &= f(v_1, w) + f(v_2, w) \\ f(cv, w) &= cf(v, w) \\ f(v, w_1 + w_2) &= f(v, w_1) + f(v, w_2) \\ f(v, cw) &= cf(v, w) \end{aligned}$$

for all $v, w, v_i, w_i \in V$ and all $c \in F$. Often a notation similar to dot product is used. We will frequently use the notation

(1.4)
$$\langle v, w \rangle$$

to designate the value $f(u, v)$ of the form. So $\langle v, w \rangle$ is a scalar, an element of F.

A form $\langle\, , \rangle$ is said to be *symmetric* if

(1.5)
$$\langle v, w \rangle = \langle w, v \rangle$$

and *skew-symmetric* if

(1.6)
$$\langle v, w \rangle = -\langle w, v \rangle,$$

for all $v, w \in V$. (This is actually not the right definition of skew-symmetry if the field F is of characteristic 2, that is, if $1 + 1 = 0$ in F. We will correct the definition in Section 8.)

If the form f is either symmetric or skew-symmetric, then linearity in the second variable follows from linearity in the first.

The main examples of bilinear forms are the forms on the space F^n of column vectors, obtained as follows: Let A be an $n \times n$ matrix in F, and define

(1.7)
$$\langle X, Y \rangle = X^t A Y.$$

Note that this product is a 1×1 matrix, that is, a scalar, and that it is bilinear. Ordinary dot product is included as the case $A = I$.

A matrix A is *symmetric* if

(1.8)
$$A^t = A, \quad \text{that is, } a_{ij} = a_{ji} \quad \text{for all } i, j.$$

(1.9) **Proposition.** The form (1.7) is symmetric if and only if the matrix A is symmetric.

Proof. Assume that A is symmetric. Since Y^tAX is a 1×1 matrix, it is equal to its transpose: $Y^tAX = (Y^tAX)^t = X^tA^tY = X^tAY$. Thus $\langle Y, X \rangle = \langle X, Y \rangle$. The other implication is obtained by setting $X = e_i$ and $Y = e_j$. We find $\langle e_i, e_j \rangle = e_i{}^t A e_j = a_{ij}$, while $\langle e_j, e_i \rangle = a_{ji}$. If the form is symmetric, then $a_{ij} = a_{ji}$, and so A is symmetric. □

Let $\langle\, , \rangle$ be a bilinear form on a vector space V, and let $\mathbf{B} = (v_1, \ldots, v_n)$ be a basis for V. We can relate the form to a product X^tAY by the *matrix of the form* with respect to the basis. By definition, this is the matrix $A = (a_{ij})$, where

$$a_{ij} = \langle v_i, v_j \rangle. \tag{1.10}$$

Note that A is a symmetric matrix if and only if $\langle\ ,\ \rangle$ is a symmetric form. Also, the symmetry of the bilinear form does not depend on the basis. So if the matrix of the form with respect to some basis is symmetric, its matrix with respect to any other basis will be symmetric too.

The matrix A allows us to compute the value of the form on two vectors $v, w \in V$. Let X, Y be their coordinate vectors, as in Section 4 of Chapter 3, so that $v = \mathbf{B}X$, $w = \mathbf{B}Y$. Then

$$\langle v, w \rangle = \langle \sum_i v_i x_i, \sum_j v_j y_j \rangle.$$

This expands using bilinearity to $\sum_{i,j} x_i y_j \langle v_i, v_j \rangle = \sum_{i,j} x_i a_{ij} y_j = X^t A Y$:

$$\langle v, w \rangle = X^t A Y. \tag{1.11}$$

Thus, if we identify F^n with V using the basis $\mathbf{B}$ as in Chapter 3 (4.14), the bilinear form $\langle\ ,\ \rangle$ corresponds to $X^t A Y$.

As in the study of linear operators, a central problem is to describe the effect of a change of basis on such a product. For example, we would like to know what happens to dot product when the basis of $\mathbb{R}^n$ is changed. This will be discussed presently. The effect of a change of basis $\mathbf{B} = \mathbf{B}'P$ [Chapter 3 (4.16)] on the matrix of the form can be determined easily from the rules $X' = PX$, $Y' = PY$: If A' is the matrix of the form with respect to a new basis $\mathbf{B}'$, then by definition of A', $\langle v, w \rangle = X'^t A' Y' = X^t P^t A' P Y$. But we also have $\langle v, w \rangle = X^t A Y$. So

$$P^t A' P = A. \tag{1.12}$$

Let $Q = (P^{-1})^t$. Since P can be any invertible matrix, Q is also arbitrary.

(1.13) **Corollary.** Let A be the matrix of a bilinear form with respect to a basis. The matrices A' which represent the same form with respect to different bases are the matrices $A' = QAQ^t$, where Q is an arbitrary matrix in $GL_n(F)$. □

Let us now apply formula (1.12) to our original example of dot product on $\mathbb{R}^n$. The matrix of the dot product with respect to the standard basis is the identity matrix: $(X \cdot Y) = X^t I Y$. So formula (1.12) tells us that if we change basis, the matrix of the form changes to

$$A' = (P^{-1})^t I (P^{-1}) = (P^{-1})^t (P^{-1}), \tag{1.14}$$

where P is the matrix of change of basis as before. If the matrix P happens to be *orthogonal,* meaning that $P^t P = I$, then $A' = I$, and dot product carries over to dot product: $(X \cdot Y) = (PX \cdot PY) = (X' \cdot Y')$, as we saw in Chapter 4 (5.13). But under a general change of basis, the formula for dot product changes to $X'^t A' Y'$, where A' is as in (1.14). For example, let $n = 2$, and let the basis $\mathbf{B}'$ be

$$v_1' = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad v_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then

$$(1.15) \qquad P^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad A' = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

The matrix A' represents dot product on $\mathbb{R}^2$, with respect to the basis $\mathbf{B}'$.

We can also turn the computation around. Suppose that we are given a bilinear form $\langle\,,\rangle$ on a real vector space V. Let us ask whether or not this form becomes dot product when we choose a suitable basis. We start with an arbitrary basis $\mathbf{B}$, so that we have a matrix A to work with. Then the problem is to change this basis in such a way that the new matrix is the identity, if that is possible. By formula (1.12), this amounts to solving the matrix equation $I = (P^{-1})^t A (P^{-1})$, or

$$(1.16) \qquad A = P^t P.$$

(1.17) **Corollary.** The matrices A which represent a form equivalent to dot product are the matrices $A = P^t P$, where P is invertible. □

This corollary gives a theoretical answer to our problem of determining the bilinear forms equivalent to dot product, but it is not very satisfactory because we don't yet have a practical method of deciding which matrices can be written as a product $P^t P$, let alone a practical method of finding P.

We can get some conditions on the matrix A from the properties of dot product listed in (1.2). Bilinearity imposes no condition on A, because the symbol $X^t A Y$ is always bilinear. However, symmetry and positivity restrict the possibilities. The easier property to check is symmetry: In order to represent dot product, the matrix A must be symmetric. Positivity is also a strong restriction. In order to represent dot product, the matrix A must have the property that

$$(1.18) \qquad X^t A X > 0, \quad \text{for all } X \neq 0.$$

A real symmetric matrix having this property is called *positive definite*.

(1.19) **Theorem.** The following properties of a real $n \times n$ matrix A are equivalent:

(i) A represents dot product, with respect to some basis of $\mathbb{R}^n$.
(ii) There is an invertible matrix $P \in GL_n(\mathbb{R})$ such that $A = P^t P$.
(iii) A is symmetric and positive definite.

We have seen that (i) and (ii) are equivalent [Corollary (1.17)] and that (i) implies (iii). So it remains to prove the remaining implication, that (iii) implies (i). It will be more convenient to restate this implication in vector space form.

A symmetric bilinear form $\langle\ ,\ \rangle$ on a finite-dimensional real vector space V is called *positive definite* if

$$(1.20)\qquad \langle v, v\rangle > 0$$

for every nonzero vector $v \in V$. Thus a real symmetric matrix A is positive definite if and only if the form $\langle X, Y\rangle = X^tAY$ it defines on $\mathbb{R}^n$ is a positive definite form. Also, the form $\langle\ ,\ \rangle$ is positive definite if and only if its matrix A with respect to any basis is a positive definite matrix. This is clear, because if X is the coordinate vector of a vector v, then $\langle v, v\rangle = X^tAX$ (1.11).

Two vectors v, w are called *orthogonal* with respect to a symmetric form if $\langle v, w\rangle = 0$. Orthogonality of two vectors is often denoted as

$$(1.21)\qquad v \perp w.$$

This definition extends the concept of orthogonality which we have already seen when the form is dot product on $\mathbb{R}^n$ [Chapter 4 (5.12)]. A basis $\mathbf{B} = (v_1,\dots, v_n)$ of V is called an *orthonormal basis* with respect to the form if

$$\langle v_i, v_j\rangle = 0 \quad \text{for all } i \neq j, \text{ and} \quad \langle v_i, v_i\rangle = 1 \quad \text{for all } i.$$

It follows directly from the definition that a basis $\mathbf{B}$ is orthonormal if and only if the matrix of the form with respect to $\mathbf{B}$ is the identity matrix.

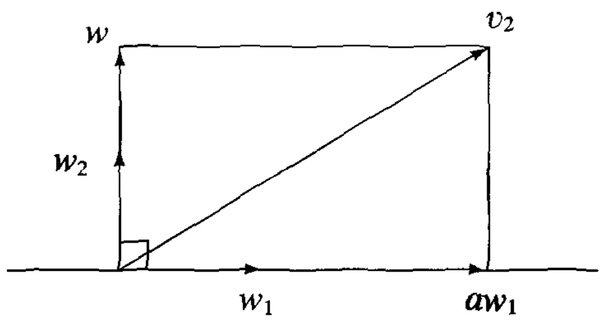
(1.22) **Theorem.** Let $\langle\ ,\ \rangle$ be a positive definite symmetric form on a finite-dimensional vector real space V. There exists an orthonormal basis for V.

Proof. We will describe a method called the *Gram–Schmidt procedure* for constructing an orthonormal basis, starting with an arbitrary basis $\mathbf{B} = (v_1,\dots, v_n)$. Our first step is to normalize v_1, so that $\langle v_1, v_1\rangle = 1$. To do this we note that

$$(1.23)\qquad \langle cv, cv\rangle = c^2v.$$

Since the form is positive definite, $\langle v_1, v_1\rangle > 0$. We set $c = \langle v_1, v_1\rangle^{-\frac{1}{2}}$, and replace v_1 by $w_1 = cv_1$.

Next we look for a linear combination of w_1 and v_2 which is orthogonal to w_1. The required linear combination is $w = v_2 - aw_1$, where $a = \langle v_2, w_1\rangle : \langle w, w_1\rangle = \langle v_2, w_1\rangle - a\langle w_1, w_1\rangle = \langle v_2, w_1\rangle - a = 0$. We normalize this vector w to length 1. obtaining a vector w_2 which we substitute for v_2. The geometric interpretation of this operation is illustrated below for the case that the form is dot product. The vector aw_1 is the orthogonal projection of v_2 onto the subspace (the line) spanned by w_1.



This is the general procedure. Suppose that the $k-1$ vectors $w_1, \ldots, w_{k-1}$ are orthonormal and that $(w_1, \ldots, w_{k-1}, v_k, \ldots, v_n)$ is a basis. We adjust v_k as follows: We let $a_i = \langle v_k, w_i \rangle$ and

$$w = v_k - a_1 w_1 - a_2 w_2 - \cdots - a_{k-1} w_{k-1}. \tag{1.24}$$

Then w is orthogonal to w_i for $i = 1, \ldots, k-1$, because

$$\langle w, w_i \rangle = \langle v_k, w_i \rangle - a_1 \langle w_1, w_i \rangle - a_2 \langle w_2, w_i \rangle - \cdots - a_{k-1} \langle w_{k-1}, w_i \rangle.$$

Since $w_1, \ldots, w_{k-1}$ are orthonormal, all the terms $\langle w_j, w_i \rangle$, $1 \leq j \leq k-1$, are zero except for the term $\langle w_i, w_i \rangle$, which is 1. So the sum reduces to

$$\langle w, w_i \rangle = \langle v_k, w_i \rangle - a_i \langle w_i, w_i \rangle = \langle v_k, w_i \rangle - a_i = 0.$$

We normalize the length of w to 1, obtaining a vector w_k which we substitute for v_k as before. Then $(w_1, \ldots, w_k)$ is orthonormal. Since v_k is in the span of $(w_1, \ldots, w_k; v_{k+1}, \ldots, v_n)$, this set is a basis. The existence of an orthonormal basis follows by induction on k. □

End of the proof of Theorem (1.19). The fact that part (iii) of Theorem (1.19) implies (i) follows from Theorem (1.22). For if A is symmetric and positive definite, then the form $\langle X, Y \rangle = X^t A Y$ it defines on $\mathbb{R}^n$ is also symmetric and positive definite. In that case, Theorem (1.22) tells us that there is a basis $\mathbf{B}'$ of $\mathbb{R}^n$ which is orthonormal with respect to the form $\langle X, Y \rangle = X^t A Y$. (But the basis will probably not be orthonormal with respect to the usual dot product on $\mathbb{R}^n$.) Now on the one hand, the matrix A' of the form $\langle X, Y \rangle$ with respect to the new basis $\mathbf{B}'$ satisfies the relation $P^t A' P = A$ (1.12), and on the other hand, since $\mathbf{B}'$ is orthonormal, $A' = I$. Thus $A = P^t P$. This proves (ii), and since (i) and (ii) are already known to be equivalent, it also proves (i). □

Unfortunately, there is no really simple way to show that a matrix is positive definite. One of the most convenient criteria is the following: Denote the upper left $i \times i$ submatrix of A by A_i. Thus

$$A_1 = [a_{11}], \quad A_2 = \begin{bmatrix} a_{11} a_{12} \\ a_{21} a_{22} \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_{11} a_{12} a_{13} \\ a_{21} a_{22} a_{23} \\ a_{31} a_{32} a_{33} \end{bmatrix}, \ldots, A_n = A.$$

(1.25) **Theorem.** A real symmetric $n \times n$ matrix A is positive definite if and only if the determinant $\det A_i$ is positive for each $i = 1, \ldots, n$.

For example, the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \tag{1.26}$$

is positive definite if and only if $a > 0$ and $ad - bc > 0$. Using this criterion, we can check immediately that the matrix A' of (1.15) is positive definite, which agrees with the fact that it represents dot product.

The proof of Theorem (1.25) is at the end of the next section.

2. SYMMETRIC FORMS: ORTHOGONALITY

In this section, we consider a finite-dimensional real vector space V on which a symmetric bilinear form $\langle\, ,\rangle$ is given, but we drop the assumption made in the last section that the form is positive definite. A form such that $\langle v, v\rangle$ takes on both positive and negative values is called *indefinite*. The *Lorentz form*

$$X^tAY = x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$$

of physics is a typical example of an indefinite form on "space-time" $\mathbb{R}^4$. The coefficient c representing the speed of light can be normalized to 1, and then the matrix of the form with respect to the given basis becomes

$$(2.1)\qquad \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix}.$$

We now pose the problem of describing all symmetric forms on a finite-dimensional real vector space. The basic notion used in the study of such a form is still that of orthogonality. But if a form is not positive definite, it may happen that a nonzero vector v is self-orthogonal: $\langle v, v\rangle = 0$. For example, this is true for the vector $(1,0,0,1)^t \in \mathbb{R}^4$, when the form is defined by (2.1). So we must revise our geometric intuition. It turns out that there is no need to worry about this point. There are enough vectors which are not self-orthogonal to serve our purposes.

(2.2) **Proposition.** Suppose the symmetric form $\langle\, ,\rangle$ is not identically zero. Then there is a vector $v \in V$ which is not self-orthogonal: $\langle v, v\rangle \neq 0$.

Proof. To say that $\langle\, ,\rangle$ is not identically zero means that there is a pair of vectors $v, w \in V$ such that $\langle v, w\rangle \neq 0$. Take these vectors. If $\langle v, v\rangle \neq 0$, or if $\langle w, w\rangle \neq 0$, then the proposition is verified. Suppose $\langle v, v\rangle = \langle w, w\rangle = 0$. Let $u = v + w$, and expand $\langle u, u\rangle$ using bilinearity:

$$\langle u, u\rangle = \langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle = 0 + 2\langle v, w\rangle + 0.$$

Since $\langle v, w\rangle \neq 0$, it follows that $\langle u, u\rangle \neq 0$. □

If W is a subspace of V, then we will denote by $W^\perp$ the set of all vectors v which are orthogonal to every $w \in W$:

$$(2.3)\qquad W^\perp = \{v \in V \mid \langle v, W\rangle = 0\}.$$

This is a subspace of V, called the *orthogonal complement* to W.

(2.4) **Proposition.** Let $w \in V$ be a vector such that $\langle w, w\rangle \neq 0$. Let $W = \{cw\}$ be the span of w. Then V is the direct sum of W and its orthogonal complement:

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

Proof. According to Chapter 3 (6.4, 6.5), we have to show two things:

(a) $W \cap W^{\perp} = 0$. This is clear. The vector cw is not orthogonal to w unless $c = 0$, because $\langle cw, w\rangle = c\langle w, w\rangle$ and $\langle w, w\rangle \neq 0$.

(b) W and $W^{\perp}$ span V: Every vector $v \in V$ can be written in the form $v = aw + v'$, where $v' \in W^{\perp}$. To show this, we solve the equation $\langle v - aw, w\rangle = 0$ for a: $\langle v - aw, w\rangle = \langle v, w\rangle - a\langle w, w\rangle = 0$. The solution is $a = \dfrac{\langle v, w\rangle}{\langle w, w\rangle}$. We set $v' = v - aw$. □

Two more concepts which we will need are the null space of a symmetric form and nondegenerate form. A vector $v \in V$ is called a *null vector* for the given form if $\langle v, w\rangle = 0$ for all $w \in V$, that is, if v is orthogonal to the whole space V. The *null space* of the form is the set of all null vectors

(2.5) $$N = \{v \mid \langle v, V\rangle = 0\} = V^{\perp}.$$

A symmetric form is said to be *nondegenerate* if the null space is $\{0\}$.

(2.6) **Proposition.** Let A be the matrix of a symmetric form with respect to a basis.

(a) The null space of the form is the set of vectors v such that the coordinate vector X of v is a solution of the homogeneous equation $AX = 0$.
(b) The form is nondegenerate if and only if the matrix A is nonsingular.

Proof. Via the basis, the form corresponds to the product X^tAY [see (1.11)]. We might as well work with this product. If Y is a vector such that $AY = 0$, then $X^tAY = 0$ for all X; hence Y is in the null space. Conversely, suppose that $AY \neq 0$. Then AY has at least one nonzero coordinate. The ith coordinate of AY is $e_i{}^tAY$. So one of the products $e_i{}^tAY$ is not zero. This shows that Y is not a null vector, which proves (a). Part (b) of the proposition follows from (a). □

Here is a generalized version of (2.4):

(2.7) **Proposition.** Let W be a subspace of V, and consider the restriction of a symmetric form $\langle\,,\rangle$ to W. Suppose that this form is nondegenerate on W. Then $V = W \oplus W^{\perp}$.

We omit the proof, which closely follows that of (2.4). □

(2.8) **Definition.** An *orthogonal basis* $\mathbf{B} = (v_1,\ldots, v_n)$ for V, with respect to a symmetric form $\langle\,,\rangle$, is a basis such that $v_i \perp v_j$ for all $i \neq j$.

Since the matrix A of a form is defined by $a_{ij} = \langle v_i, v_j\rangle$, the basis $\mathbf{B}$ is orthogonal if and only if A is a *diagonal* matrix. Note that if the symmetric form $\langle\,,\rangle$ is nondegenerate and the basis $\mathbf{B} = (v_1, \ldots, v_n)$ is orthogonal, then $\langle v_i, v_i \rangle \neq 0$ for all i: the diagonal entries of A are nonzero.

(2.9) **Theorem.** Let $\langle\, , \rangle$ be a symmetric form on a real vector space V.

(a) There is an orthogonal basis for V. More precisely, there exists a basis $\mathbf{B} = (v_1, \ldots, v_n)$ such that $\langle v_i, v_j \rangle = 0$ for $i \neq j$ and such that for each i, $\langle v_i, v_i \rangle$ is either 1, -1, or 0.

(b) *Matrix form:* Let A be a real symmetric $n \times n$ matrix. There is a matrix $Q \in GL_n(\mathbb{R})$ such that QAQ^t is a diagonal matrix each of whose diagonal entries is 1, -1, or 0.

Part (b) of the theorem follows from (a), and (1.13), taking into account the fact that any symmetric matrix A is the matrix of a symmetric form. □

We can permute an orthogonal basis $\mathbf{B}$ so that the indices with $\langle v_i, v_i \rangle = 1$ are the first ones, and so on. Then the matrix A of the form will be

$$A = \begin{bmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{bmatrix}, \tag{2.10}$$

where p is the number of $+1$'s, m is the number of -1's, and z is the number of 0's, so that $p + m + z = n$. These numbers are uniquely determined by the form or by the matrix A:

(2.11) **Theorem.** *Sylvester's Law:* The numbers p, m, z appearing in (2.10) are uniquely determined by the form. In other words, they do not depend on the choice of orthogonal basis $\mathbf{B}$ such that $\langle v_i, v_i \rangle = \pm 1$ or 0.

The pair of integers (p, m) is called the *signature* of the form.

Proof of Theorem (2.9). If the form is identically zero, then the matrix A, computed with respect to any basis, will be the zero matrix, which is diagonal. Suppose the form is not identically zero. Then by Proposition (2.2), there is a vector $v = v_1$ with $\langle v_1, v_1 \rangle \neq 0$. Let W be the span of v_1. By Proposition (2.4), $V = W \oplus W^\perp$, and so a basis for V is obtained by combining the basis (v_1) of W with any basis $(v_2, \ldots, v_n)$ of $W^\perp$ [Chapter 3 (6.6)]. The form on V can be restricted to the subspace $W^\perp$, and it defines a form there. We use induction on the dimension to conclude that $W^\perp$ has an orthogonal basis $(v_2, \ldots, v_n)$. Then $(v_1, v_2, \ldots, v_n)$ is an orthogonal basis for V. For, $\langle v_1, v_i \rangle = 0$ if $i > 1$ because $v_i \in W^\perp$, and $\langle v_i, v_j \rangle = 0$ if $i, j > 1$ and $i \neq j$, because $(v_2, \ldots, v_n)$ is an orthogonal basis.

It remains to normalize the orthogonal basis just constructed. If $\langle v_i, v_i \rangle \neq 0$, we solve $c^{-2} = \pm\langle v_i, v_i \rangle$ and change the basis vector v_i to cv_i. Then $\langle v_i, v_i \rangle$ is changed to ± 1. This completes the proof of (2.9.) □

Proof of Theorem (2.11). Let $r = p + m$. (This is the rank of the matrix A.) Let $(v_1, \ldots, v_n)$ be an orthogonal basis of V of the type under consideration, that is, so that the matrix is (2.10). We will first show that the number z is determined by proving that the vectors $v_{r+1}, \ldots, v_n$ form a basis for the null space $N = V^{\perp}$. This will show that $z = \dim N$, hence that z does not depend on the choice of a basis.

A vector $w \in V$ is a null vector if and only if it is orthogonal to every element v_i of our basis. We write our vector as a linear combination of the basis: $w = c_1 v_1 + \cdots + c_n v_n$. Then since $\langle v_i, v_j \rangle = 0$ if $i \neq j$, we find $\langle w, v_i \rangle = c_i \langle v_i, v_i \rangle$. Now $\langle v_i, v_i \rangle = 0$ if and only if $i > r$. So in order for w to be orthogonal to every v_i, we must have $c_i = 0$ for all $i \leq r$. This shows that $(v_{r+1}, \ldots, v_n)$ spans N, and, being a linearly independent set, it is a basis for N.

The equation $p + m + z = n$ proves that $p + m$ is also determined. We still have to show that one of the two remaining integers p,m is determined. This is not quite so simple. It is not true that the span of $(v_1, \ldots, v_p)$, for instance, is uniquely determined by the form.

Suppose a second such basis $(v_1', \ldots, v_n')$ is given and leads to integers p', m' (with $z' = z$). We will show that the $p + (n - p')$ vectors

$$v_1, \ldots, v_p;\ v_{p'+1}', \ldots, v_n' \tag{2.12}$$

are linearly independent. Then since V has dimension n, it will follow that $p + (n - p') \leq n$, hence that $p \leq p'$, and, interchanging the roles of p and p', that $p = p'$.

Let a linear relation between the vectors (2.12) be given. We may write it in the form

$$b_1 v_1 + \cdots + b_p v_p = c_{p'+1} v_{p'+1}' + \cdots + c_n v_n'. \tag{2.13}$$

Let v denote the vector defined by either of these two expressions. We compute $\langle v, v \rangle$ in two ways. The left-hand side gives

$$\langle v, v \rangle = b_1^2 \langle v_1, v_1 \rangle + \cdots + b_p^2 \langle v_p, v_p \rangle = b_1^2 + \cdots + b_p^2 \geq 0,$$

while the right-hand side gives

$$\langle v, v \rangle = c_{p'+1}^2 \langle v_{p'+1}, v_{p'+1} \rangle + \cdots + c_n^2 \langle v_n', v_n' \rangle = -c_{p'+1}^2 - \cdots - c_{p'+m'}^2 \leq 0.$$

It follows that $b_1^2 + \cdots + b_p^2 = 0$, hence that $b_1 = \ldots = b_p = 0$. Once this is known, the fact that $(v_1', \ldots, v_n')$ is a basis combines with (2.13) to imply $c_{p'+1} = \ldots = c_n = 0$. Therefore the relation was trivial, as required. □

For dealing with indefinite forms, the notation $I_{p,m}$ is often used to denote the diagonal matrix

$$I_{p,m} = \begin{bmatrix} I_p & \\ & -I_m \end{bmatrix}. \tag{2.14}$$

With this notation, the matrix representing the Lorentz form (2.1) is $I_{3,1}$.

We will now prove Theorem (1.25)—that a matrix A is positive definite if and only if $\det A_i > 0$ for all i.

Proof of Theorem (1.25). Suppose that the form X^tAY is positive definite. A change of basis in $\mathbb{R}^n$ changes the matrix to $A' = QAQ^t$, and

$$\det A' = (\det Q)(\det A)(\det Q^t) = (\det Q)^2(\det A).$$

Since they differ by a square factor, $\det A'$ is positive if and only if $\det A$ is positive. By (1.19), we can choose a matrix Q so that $A' = I$, and since I has determinant 1, $\det A > 0$.

The matrix A_i represents the restriction of the form to the subspace V_i spanned by $(v_1, \ldots, v_i)$, and of course the form is positive definite on V_i. Therefore $\det A_i > 0$ for the same reason that $\det A > 0$.

Conversely, suppose that $\det A_i$ is positive for all i. By induction on n, we may assume the form to be positive definite on V_{n-1}. Therefore there is a matrix $Q' \in GL_{n-1}$ such that $Q'A_{n-1}Q'^t = I_{n-1}$. Let Q be the matrix

$$Q = \begin{bmatrix} Q' & \\ & 1 \end{bmatrix}.$$

Then

$$QAQ^t = \begin{bmatrix} & & * \\ & I & \vdots \\ * & \cdots & * \end{bmatrix}.$$

We now clear out the bottom row of this matrix, except for the (n,n) entry, by elementary row operations $E_1, \ldots, E_{n-1}$. Let $P = E_{n-1} \cdots E_1 Q$. Then

$$A' = PAP^t = \begin{bmatrix} \begin{bmatrix} & & \\ & I & \\ & & \end{bmatrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ 0 \quad \cdots \quad 0 & c \end{bmatrix},$$

for some c. The last column has also been cleared out because $A' = PAP^t$ is symmetric. Since $\det A > 0$, we have $\det A' = (\det A)(\det P)^2 > 0$ too, and this implies that $c > 0$. Therefore the matrix A' represents a positive definite form. It also represents the same form as A does. So A is positive definite. $\square$

3. THE GEOMETRY ASSOCIATED TO A POSITIVE FORM

In this section we return to look once more at a positive definite bilinear form $\langle\, , \rangle$ on an n-dimensional real vector space V. A real vector space together with such a form is often called a *Euclidean space*.

It is natural to define the *length* of a vector v by the rule

(3.1) $$|v| = \sqrt{\langle v, v\rangle},$$

in analogy with the length of vectors in $\mathbb{R}^n$ [Chapter 4 (5.10)]. One important consequence of the fact that the form is positive definite is that we can decide whether a vector v is zero by computing its length:

(3.2) $$v = 0 \quad \textit{if and only if} \quad \langle v, v\rangle = 0.$$

As was shown in Section 1, there is an orthonormal basis $\mathbf{B} = (v_1,\ldots,v_n)$ for V, and thereby the form corresponds to dot product on $\mathbb{R}^n$:

$$\langle v, w\rangle = X^tY,$$

if $v = \mathbf{B}X$ and $w = \mathbf{B}Y$. Using this correspondence, we can transfer the geometry of $\mathbb{R}^n$ over to V. Whenever a problem is presented to us on a Euclidean space V, a natural procedure will be to choose a convenient orthonormal basis, thereby reducing the problem to the familiar case of dot product on $\mathbb{R}^n$.

When a subspace W of V is given to us, there are two operations we can make. The first is to *restrict* the form $\langle\,,\rangle$ to the subspace, simply by defining the value of the form on a pair w_1, w_2 of vectors in W to be $\langle w_1, w_2\rangle$. The restriction of a bilinear form to a subspace W is a bilinear form on W, and if the form is symmetric or if it is symmetric and positive definite, then so is the restriction.

Restriction of the form can be used to define the unoriented *angle* between two vectors v, w. If the vectors are linearly dependent, the angle is zero. Otherwise, (v, w) is a basis of a two-dimensional subspace W of V. The restriction of the form to W is still positive definite, and therefore there is an orthonormal basis (w_1, w_2) for W. By means of this basis, v, w correspond to their coordinate vectors X, Y in $\mathbb{R}^2$. This allows us to interpret geometric properties of the vectors v, w in terms of properties of X, Y.

Since the basis (w_1, w_2) is orthonormal, the form corresponds to dot product on $\mathbb{R}^2$: $\langle v, w\rangle = X^tY$. Therefore

$$|v| = |X|, \quad |w| = |Y|, \quad \text{and} \quad \langle v, w\rangle = (X \cdot Y).$$

We define the angle θ between v and w to be the angle between X and Y, and thereby obtain the formula

(3.3) $$\langle v, w\rangle = |v||w| \cos\theta,$$

as a consequence of the analogous formula [Chapter 4 (5.11)] for dot product in $\mathbb{R}^2$. This formula determines $\cos\theta$ in terms of the other symbols, and $\cos\theta$ determines θ up to a factor of ± 1. Therefore the angle between v and w is determined up to sign by the form alone. This is the best that can be done, even in $\mathbb{R}^3$.

Standard facts such as the *Schwarz Inequality*

(3.4) $$|\langle v, w\rangle| \le |v||w|$$

and the *Triangle Inequality*

(3.5) $$|v + w| \le |v| + |w|$$

can also be proved for arbitrary Euclidean spaces by restriction to a two-dimensional subspace.

The second operation we can make when a subspace W is given is to project V onto W. Since the restriction of the form to W is positive definite, it is nondegenerate. Therefore $V = W \oplus W^{\perp}$ by (2.17), and so every $v \in V$ has a unique expression

$$(3.6) \qquad v = w + w', \quad \text{with} \quad w \in W \quad \text{and} \quad \langle w, w' \rangle = 0.$$

The *orthogonal projection* $\pi: V \longrightarrow W$ is defined to be the linear transformation

$$(3.7) \qquad v \rightsquigarrow \pi(v) = w$$

where w is as in (3.6).

The projected vector $\pi(v)$ can be computed easily in terms of an orthonormal basis $(w_1, \ldots, w_r)$ of W. What follows is important:

(3.8) **Proposition.** Let $(w_1, \ldots, w_r)$ be an orthonormal basis of a subspace W, and let $v \in V$. The orthogonal projection $\pi(v)$ of v onto w is the vector

$$\pi(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_r \rangle w_r.$$

Thus if π is defined by the above formula, then $v - \pi(v)$ is orthogonal to W. This formula explains the geometric meaning of the Gram–Schmidt procedure described in Section 1.

Proof. Let us denote the right side of the above equation by $\tilde{w}$. Then $\langle \tilde{w}, w_i \rangle = \langle v, w_i \rangle \langle w_i, w_i \rangle = \langle v, w_i \rangle$ for $i = 1, \ldots, r$, hence $v - \tilde{w} \in W^{\perp}$. Since the expression (3.6) for v is unique. $w = \tilde{w}$ and $w' = v - \tilde{w}$. □

The case $W = V$ is also important. In this case, π is the identity map.

(3.9) **Corollary.** Let $\mathbf{B} = (v_1, \ldots, v_n)$ be an orthonormal basis for a Euclidean space V. Then

$$v = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_n \rangle v_n.$$

In other words, the coordinate vector of v with respect to the orthonormal basis $\mathbf{B}$ is

$$X = (\langle v, v_1 \rangle, \ldots, \langle v, v_n \rangle)^t. \; □$$

4. HERMITIAN FORMS

In this section we assume that our scalar field is the field $\mathbb{C}$ of complex numbers. When working with complex vector spaces, it is desirable to have an analogue of the concept of the length of a vector, and of course one can define length on $\mathbb{C}^n$ by identifying it with $\mathbb{R}^{2n}$. If $X = (x_1, \ldots, x_n)^t$ is a complex vector and if $x_r = a_r + b_r i$, then the *length* of X is

$$(4.1) \qquad |X| = \sqrt{a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2} = \sqrt{\overline{x}_1 x_1 + \cdots + \overline{x}_n x_n},$$

where the bar denotes complex conjugation. This formula suggests that dot product is "wrong" for complex vectors and that we should define a product by the formula

(4.2) $$\langle X, Y\rangle = \overline{X}^{t}Y = \overline{x}_1 y_1 + \cdots + \overline{x}_n y_n.$$

This product has the *positivity* property:

(4.3) $\langle X, X\rangle$ *is a positive real number if* $X \neq 0$.

Moreover, (4.2) agrees with dot product for real vectors.

The product (4.2) is called the *standard hermitian product*, or the *hermitian dot product*. It has these properties:

(4.4)

Linearity in the second variable:

$$\langle X, cY\rangle = c\langle X, Y\rangle \quad \text{and} \quad \langle X, Y_1 + Y_2\rangle = \langle X, Y_1\rangle + \langle X, Y_2\rangle;$$

Conjugate linearity in the first variable:

$$\langle cX, Y\rangle = \overline{c}\langle X, Y\rangle \quad \text{and} \quad \langle X_1 + X_2, Y\rangle = \langle X_1, Y\rangle + \langle X_2, Y\rangle;$$

Hermitian symmetry:

$$\langle Y, X\rangle = \overline{\langle X, Y\rangle}.$$

So we can have a positive definite product at a small cost in linearity and symmetry.

When one wants to work with notions involving length, the hermitian product is the right one, though symmetric bilinear forms on complex vector spaces also come up in applications.

If V is a complex vector space, a *hermitian form* on V is any function of two variables

(4.5) $$\begin{aligned} V \times V &\longrightarrow \mathbb{C} \\ v, w &\rightsquigarrow \langle v, w\rangle \end{aligned}$$

satisfying the relations (4.4). Let $\mathbf{B} = (v_1, \ldots, v_n)$ be a basis for V. Then the *matrix* of the form is defined in the analogous way as the matrix of a bilinear form:

$$A = (a_{ij}), \quad \text{where } a_{ij} = \langle v_i, v_j\rangle.$$

The formula for the form now becomes

(4.6) $$\langle v, w\rangle = \overline{X}^{t}AY,$$

if $v = \mathbf{B}X$ and $w = \mathbf{B}Y$.

The matrix A is not arbitrary, because hermitian symmetry implies that

$$a_{ij} = \langle v_i, v_j\rangle = \overline{\langle v_j, v_i\rangle} = \overline{a}_{ji},$$

that is, that $A = \overline{A}^{t}$. Let us introduce the *adjoint* of a matrix A [different from the one defined in Chapter 1 (5.4)] as

(4.7) $$A^* = \bar{A}^t.$$

It satisfies the following rules:

$$(A + B)^* = A^* + B^*$$
$$(AB)^* = B^*A^*$$
$$(A^*)^{-1} = (A^{-1})^*$$
$$A^{**} = A.$$

These rules are easy to check. Formula (4.6) can now be rewritten as

(4.8) $$\langle v, w \rangle = X^*AY,$$

and the standard hermitian product on $\mathbb{C}^n$ becomes $\langle X, Y \rangle = X^*Y$.

A matrix A is called *hermitian* or *self-adjoint* if

(4.9) $$A = A^*,$$

and it is the hermitian matrices which are matrices of hermitian forms. Their entries satisfy $a_{ji} = \bar{a}_{ij}$. This implies that the diagonal entries are real and that the entries below the diagonal are complex conjugates of those above it:

$$A = \begin{bmatrix} r_1 & & a_{ij} \\ & \ddots & \\ \bar{a}_{ij} & & r_n \end{bmatrix}, \quad r_i \in \mathbb{R}, \quad a_{ij} \in \mathbb{C}.$$

For example, $\begin{bmatrix} 2 & i \\ -i & 1 \end{bmatrix}$ is a hermitian matrix.

Note that the condition for a real matrix to be hermitian is $a_{ji} = a_{ij}$:

(4.10) *The real hermitian matrices are the real symmetric matrices.*

The discussion of change of basis in Sections 1 and 2 has analogues for hermitian forms. Given a hermitian form, a change of basis by a matrix P leads as in (1.12) to

$$X'^*A'Y' = (PX)^*A'PY = X^*(P^*A'P)Y.$$

Hence the new matrix A' satisfies

(4.11) $$A = P^*A'P \quad \text{or} \quad A' = (P^*)^{-1}AP^{-1}.$$

Since P is arbitrary, we can replace it by $Q = (P^*)^{-1}$ to obtain the description analogous to (1.13):

(4.12) **Corollary.** Let A be the matrix of a hermitian form with respect to a basis. The matrices which represent the same hermitian form with respect to different bases are those of the form $A' = QAQ^*$, for some invertible matrix $Q \in GL_n(\mathbb{C})$. □

For hermitian forms, the analogues of orthogonal matrices are the unitary matrices. A matrix P is called *unitary* if it satisfies the condition

$$(4.13) \qquad P^*P = I \quad \text{or} \quad P^* = P^{-1}.$$

For example, $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}$ is a unitary matrix.

Note that for a real matrix P, this condition becomes $P^tP = I$:

(4.14) *The real unitary matrices are the real orthogonal matrices.*

The unitary matrices form a group, the *unitary group* U_n:

$$(4.15) \qquad U_n = \{P \mid P^*P = I\}.$$

Formula (4.11) tells us that unitary matrices represent changes of basis which leave the standard hermitian product X^*Y invariant:

(4.16) **Corollary.** A change of basis preserves the standard hermitian product, that is, $X^*Y = X'^*Y'$, if and only if its matrix P is unitary. □

But Corollary (4.12) tells us that a general change of basis changes the standard hermitian product X^*Y to $X'^*A'Y'$, where $A' = QQ^*$, and $Q \in GL_n(\mathbb{C})$.

The notion of orthogonality for hermitian forms is defined exactly as for symmetric bilinear forms: v is called *orthogonal* to w if $\langle v, w\rangle = 0$. Since $\overline{\langle v, w\rangle} = \langle w, v\rangle$, orthogonality is still a symmetric relation. We can now copy the discussion of Sections 1 and 2 for hermitian forms without essential change, and Sylvester's Law (2.11) for real symmetric forms carries over to the hermitian case. In particular, we can speak of *positive definite* forms, those having the property that

(4.17) $\langle v, v\rangle$ *is a positive real number if* $v \neq 0$,

and of *orthonormal bases* $\mathbf{B} = (v_1, \dots, v_n)$, those such that

$$(4.18) \qquad \langle v_i, v_i\rangle = 1 \quad \text{and} \quad \langle v_i, v_j\rangle = 0 \quad \text{if } i \neq j.$$

(4.19) **Theorem.** Let $\langle\, ,\rangle$ be a hermitian form on a complex vector space V. There is an orthonormal basis for V if and only if the form is positive definite.

(4.20) **Proposition.** Let W be a subspace of a hermitian space V. If the restriction of the form to W is nondegenerate, then $V = W \oplus W^\perp$

The proofs of these facts are left as exercises. □

5. THE SPECTRAL THEOREM

In this section we will study an n-dimensional complex vector space V and a positive definite hermitian form $\langle\, ,\rangle$ on V. A complex vector space on which a positive definite hermitian form is given is often called a *hermitian space*. You can imagine that V is $\mathbb{C}^n$, with its standard hermitian product X^*Y, if you want to. The choice of an orthonormal basis in V will allow such an identification.

Since the form $\langle\, ,\rangle$ is given, we will not want to choose an arbitrary basis for V in order to make computations. It is natural to work exclusively with orthonormal bases. This changes all previous calculations in the following way: It will no longer be true that the matrix P of a change of basis is an arbitrary invertible matrix. Rather, if $\mathbf{B} = (v_1,\dots,v_n)$, $\mathbf{B}' = (v_1',\dots,v_n')$ are two *orthonormal* bases, then the matrix P relating them will be unitary. The fact that the bases are orthonormal means that the matrix of the form $\langle\, ,\rangle$ with respect to each basis is the identity I, and so (4.11) reads $I = P^*IP$, or $P^*P = I$.

We are going to study a linear operator

$$T\colon V \longrightarrow V \tag{5.1}$$

on our space. Let $\mathbf{B}$ be an orthonormal basis, and let M be the associated matrix of T. A change of orthonormal basis changes M to $M' = PMP^{-1}$ [Chapter 4 (3.4)] where P is unitary; hence

$$M' = PMP^*. \tag{5.2}$$

(5.3) **Proposition.** Let T be a linear operator on a hermitian space V, and let M be the matrix of T with respect to an orthonormal basis $\mathbf{B}$.

(a) The matrix M is hermitian if and only if $\langle v, Tw\rangle = \langle Tv, w\rangle$ for all $v, w \in V$. If so, T is called a *hermitian operator*.

(b) The matrix M is unitary if and only if $\langle v, w\rangle = \langle Tv, Tw\rangle$ for all $v, w \in V$. If so, T is called a *unitary operator*.

Proof. Let X, Y be the coordinate vectors of v, w: $v = \mathbf{B}X$, $w = \mathbf{B}Y$, so that $\langle v, w\rangle = X^*Y$ and $Tv = \mathbf{B}MX$. Then $\langle v, Tw\rangle = X^*MY$, and $\langle Tv, w\rangle = X^*M^*Y$. So if $M = M^*$, then $\langle v, Tw\rangle = \langle Tv, w\rangle$ for all v,w; that is, T is hermitian. Conversely, if T is hermitian, we set $v = e_i$, $w = e_j$ as in the proof of (1.9) to obtain $b_{ij} = e_i^*(Me_j) = (e_i^*M^*)e_j = \bar{b}_{ji}$. Thus $M = M^*$. Similarly, $\langle v, w\rangle = X^*Y$ and $\langle Tv, Tw\rangle = X^*M^*MY$, so $\langle v, w\rangle = \langle Tv, Tw\rangle$ for all v, w if and only if $M^*M = I$. □

(5.4) **Theorem.** *Spectral Theorem:*

(a) Let T be a hermitian operator on a hermitian vector space V. There is an orthonormal basis of V consisting of eigenvectors of T.

(b) *Matrix form:* Let M be a hermitian matrix. There is a unitary matrix P such that PMP^* is a real diagonal matrix.

Proof. Choose an eigenvector $v = v_1$, and normalize so that its length is 1: $\langle v, v\rangle = 1$. Extend to an orthonormal basis. Then the matrix of T becomes

$$M = \begin{bmatrix} a & * & \cdots & * \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{bmatrix}.$$

Since T is hermitian, so is the matrix M (5.3). This implies that $* \cdots * = 0 \cdots 0$ and that N is hermitian. Proceed by induction. □

To diagonalize a hermitian matrix M by a unitary P, one can proceed by determining the eigenvectors. If the eigenvalues are distinct, the corresponding eigenvectors will be orthogonal. This follows from the Spectral Theorem. Let $\mathbf{B}'$ be the orthonormal basis obtained by normalizing the lengths of the eigenvectors to 1. Then $P = [\mathbf{B}']^{-1}$ [Chapter 3 (4.20)].

For example, let

$$M = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}.$$

The eigenvalues of this matrix are 3, 1, and the vectors

$$v_1' = \begin{bmatrix} 1 \\ -i \end{bmatrix}, \quad v_2' = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

are eigenvectors with these eigenvalues. We normalize their lengths to 1 by the factor $\frac{1}{\sqrt{2}}$. Then

$$P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}^* = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \quad \text{and} \quad PMP^* = \begin{bmatrix} 3 & \\ & 1 \end{bmatrix}.$$

But the Spectral Theorem asserts that a hermitian matrix can be diagonalized *even if its eigenvalues aren't distinct*. This statement becomes particularly simple for 2×2 matrices: If the characteristic polynomial of a 2×2 hermitian matrix M has a double root, then there is a unitary matrix P such that $PMP^* = aI$. Bringing the P's over to the other side of the equation, we obtain $M = P^*aIP = aP^*P = aI$. So it follows from the Spectral Theorem that $M = aI$. The only 2×2 hermitian matrices whose characteristic polynomials have a double root are the matrices aI, where a is a real number. We can verify this fact directly from the definition. We write $M = \begin{bmatrix} a & \beta \\ \overline{\beta} & d \end{bmatrix}$, where a, d are real and β is complex. Then the characteristic polynomial is $t^2 - (a + d)t + (ad - \beta\overline{\beta})$. This polynomial has a double root if and only if its discriminant vanishes, that is, if

$$(a + d)^2 - 4(ad - \beta\overline{\beta}) = (a - d)^2 + 4\beta\overline{\beta} = 0.$$

Both of the terms $(a - d)^2$ and $\beta\overline{\beta}$ are nonnegative real numbers. So if the discriminant vanishes, then $a = d$ and $\beta = 0$. In this case, $M = aI$, as predicted.

Here is an interesting consequence of the Spectral Theorem for which we can give a direct proof:

(5.5) **Proposition.** The eigenvalues of a hermitian operator T are real numbers.

Proof. Let a be an eigenvalue, and let v be an eigenvector for T such that $T(v) = av$. Then by (5.3) $\langle Tv, v\rangle = \langle v, Tv\rangle$; hence $\langle av, v\rangle = \langle v, av\rangle$. By conjugate linearity (4.4),

$$\overline{a}\langle v, v\rangle = \langle av, v\rangle = \langle v, av\rangle = a\langle v, v\rangle,$$

and $\langle v, v\rangle \neq 0$ because the form $\langle\, ,\rangle$ is positive definite. Hence $a = \overline{a}$. This shows that a is real. □

The results we have proved for hermitian matrices have analogues for real symmetric matrices. Let V be a real vector space with a positive definite bilinear form $\langle\, ,\rangle$. Let T be a linear operator on V.

(5.6) **Proposition.** Let M be the matrix of T with respect to an orthonormal basis.

(a) The matrix M is symmetric if and only if $\langle v, Tw\rangle = \langle Tv, w\rangle$ for all $v, w \in V$. If so, T is called a *symmetric operator*.

(b) The matrix M is orthogonal if and only if $\langle v, w\rangle = \langle Tv, Tw\rangle$ for all $v, w \in V$. If so, T is called an *orthogonal operator*. □

(5.7) **Proposition.** The eigenvalues of a real symmetric matrix are real.

Proof. A real symmetric matrix is hermitian. So this is a special case of (5.5). □

(5.8) **Theorem.** *Spectral Theorem (real case):*

(a) Let T be a symmetric operator on a real vector space V with a positive definite bilinear form. There is an orthonormal basis of eigenvectors of T.

(b) *Matrix form:* Let M be a real symmetric $n \times n$ matrix. There is an orthogonal matrix $P \in O_n(\mathbb{R})$ such that PMP^t is diagonal.

Proof. Now that we know that the eigenvalues of such an operator are real, we can copy the proof of (5.4). □

6. CONICS AND QUADRICS

A *conic* is the locus in the plane $\mathbb{R}^2$ defined by a quadratic equation in two variables, of the form

(6.1) $$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0.$$

More precisely, the locus (6.1) is a conic, meaning an ellipse, a hyperbola, or a parabola, or else it is called *degenerate*. A degenerate conic can be a pair of lines, a single line, a point, or empty, depending on the particular equation. The term *quadric* is used to designate the analogous loci in three or more dimensions.

The quadratic part of $f(x_1, x_2)$ is called a quadratic form:

$$q(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2. \tag{6.2}$$

In general, a *quadratic form* in n variables $x_1, \dots, x_n$ is a polynomial each of whose terms has degree 2 in the variables.

It is convenient to express the form $q(x_1, x_2)$ in matrix notation. To do this, we introduce the symmetric matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}. \tag{6.3}$$

Then $q(x_1, x_2) = X^tAX$, where X denotes the column vector $(x_1, x_2)^t$. We also introduce the row vector $B = (b_1, b_2)$. Then equation (6.1) can be written in matrix notation as

$$X^tAX + BX + c = 0. \tag{6.4}$$

We put the coefficient 2 into formulas (6.1) and (6.2) in order to avoid some coefficients $\frac{1}{2}$ in the matrix (6.3). An alternative way to write the quadratic form would be

$$q(x_1, x_2) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{12}x_2x_1 + a_{22}x_2^2.$$

We propose to describe the congruence classes of conics as geometric figures or, what is the same, their orbits under the action of the group M of rigid motions of the plane. A rigid motion will produce a change of variable in equation (6.1).

(6.5) Theorem. Every nondegenerate conic is congruent to one of the following:

(i) *Ellipse:* $\quad a_{11}x_1^2 + a_{22}x_2^2 - 1 = 0,$

(ii) *Hyperbola:* $\quad a_{11}x_1^2 - a_{22}x_2^2 - 1 = 0,$

(iii) *Parabola:* $\quad a_{11}x_1^2 - x_2 = 0, \quad$ where $a_{11}, a_{22} > 0$.

Proof. We simplify equation (6.1) in two steps, first applying an orthogonal transformation (a rotation or reflection) to diagonalize A and then applying a translation to eliminate, as much as possible, the linear and constant terms $BX + c$.

By the Spectral Theorem (5.8), there is an orthogonal matrix P such that PAP^t is diagonal. We make the change of variable $X' = PX$, or $X = P^tX'$. Substitution into equation (6.4) yields

$$X'^t(PAP^t)X' + (BP^t)X' + c = 0. \tag{6.6}$$

Hence there is an orthogonal change of variable such that the quadratic form becomes diagonal, that is, the coefficient a_{12} of x_1x_2 is zero.

Suppose that A is diagonal. Then f has the form

$$f(x_1, x_2) = a_{11}x_1^2 + a_{22}x_2^2 + b_1x_1 + b_2x_2 + c = 0.$$

We eliminate b_i by completing the squares, making the substitution

$$x_i = \left(x_i' - \frac{b_i}{2a_{ii}}\right). \tag{6.7}$$

This substitution results in

$$f(x_1, x_2) = a_{11}x_1'^2 + a_{22}x_2'^2 + c', \tag{6.8}$$

where c' is a number which can be determined if desired. This substitution corresponds to translation by the vector $(b_1/2a_{11}, b_2/2a_{22})^t$, and we can make it provided a_{11}, a_{22} are not zero.

If $a_{ii} = 0$ but $b_i \neq 0$, then we can use the substitution

$$x_i = x_i' - c/b_i \tag{6.9}$$

to eliminate the constant term instead. We may normalize one coefficient to -1. Doing so and eliminating degenerate conics leaves us with the three cases listed in the theorem. It is not difficult to show that a change of the coefficients a_{11}, a_{22} results in a different congruence class, except for the interchange of a_{11}, a_{22} in the equation of an ellipse. □

The method used above can be applied in any number of variables to classify quadrics in n dimensions. The general quadratic equation has the form

$$f(x_1,\ldots,x_n) = \sum_i a_{ii}x_i^2 + \sum_{i<j} 2a_{ij}x_ix_j + \sum_i b_ix_i + c = 0. \tag{6.10}$$

We could also write this equation more compactly as

$$f(x_1,\ldots,x_n) = \sum_{i,j} a_{ij}x_ix_j + \sum_i b_ix_i + c = 0, \tag{6.11}$$

where the first sum is over all pairs of indices, and where we set $a_{ji} = a_{ij}$.

We define the matrices A,B to be

$$A = \begin{bmatrix} a_{11}a_{12} & \cdots & a_{1m} \\ a_{12} & & \vdots \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{mm} \end{bmatrix}, \quad B = (b_1,\ldots,b_n).$$

Then the quadratic form is

$$q(x_1,\ldots,x_n) = X^tAX, \tag{6.12}$$

and

$$f(x_1,\ldots,x_n) = X^tAX + BX + c. \tag{6.13}$$

By a suitable orthogonal transformation P, the quadric is carried to (6.6), where PAP^t is diagonal. When A is diagonal, linear terms are eliminated by the translation (6.7), or else (6.9) is used.

Here is the classification in three variables:

(6.14) **Theorem.** The congruence classes of nondegenerate quadrics in $\mathbb{R}^3$ are represented by

(i) *Ellipsoids:* $a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2-1 = 0,$

(ii) *1-sheeted hyperboloids:* $a_{11}x_1^2+a_{22}x_2^2-a_{33}x_3^2-1 = 0,$

(iii) *2-sheeted hyperboloids:* $a_{11}x_1^2-a_{22}x_2^2-a_{33}x_3^2-1 = 0,$

(iv) *Elliptic paraboloids:* $a_{11}x_1^2+a_{22}x_2^2-x_3 = 0,$

(v) *Hyperbolic paraboloids:* $a_{11}x_1^2-a_{22}x_2^2-x_3 = 0,$

where $a_{11}, a_{22}, a_{33} > 0$. □

If a quadratic equation $f(x_1, x_2) = 0$ is given, we can determine the type of conic it represents most easily by allowing nonorthogonal changes of coordinates. For example, if the associated quadratic form q is positive definite, then the conic is either an ellipse, or else it is degenerate (a single point or empty). To distinguish these cases, arbitrary changes of coordinates are permissible. A nonorthogonal coordinate change will distort the conic, but it will not change an ellipse into a hyperbola or a degenerate conic.

As an example, consider the locus

$$x_1^2+x_1x_2+x_2^2 + 4x_1+3x_2 + 4 = 0. \tag{6.15}$$

The associated matrix is

$$A = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix},$$

which is positive definite by (1.25). We diagonalize A by the nonorthogonal substitution $X' = PX$, where

$$P = \begin{bmatrix} 1 & \frac{1}{2} \\ & 1 \end{bmatrix}, \quad PAP^t = \begin{bmatrix} 1 & \\ & \frac{3}{4} \end{bmatrix}, \quad BP^t = (4, 1),$$

to obtain

$$x_1'^2+\tfrac{3}{4}x_2'^2 + 4x_1'+x_2' + 4 = 0.$$

Completing the square yields

$$x_1''^2 + \tfrac{3}{4}x_2''^2 - \tfrac{1}{3} = 0,$$

an ellipse. Thus (6.15) represents an ellipse too. On the other hand, if we change the constant term of (6.15) to 5, the locus becomes empty.

7. THE SPECTRAL THEOREM FOR NORMAL OPERATORS

The Spectral Theorem (5.4) tells us that any hermitian matrix M can be transformed into a real diagonal matrix D by a unitary matrix P: $D = PMP^*$. We now ask for the matrices M which can be transformed in the same way to a diagonal matrix D, but where we no longer require D to be real. It turns out that there is an elegant formal characterization of such matrices.

(7.1) **Definition.** A matrix M is called *normal* if it commutes with its adjoint, that is, if $MM^* = M^*M$.

(7.2) **Lemma.** If M is normal and P is unitary, then $M' = PMP^*$ is also normal, and conversely.

Proof. Assume that M is normal. Then $M'M'^* = PMP^*(PMP^*)^* = PMM^*P^* = PM^*MP^* = (PMP^*)^*(PMP^*) = M'^*M'$. So PMP^* is normal. The converse follows by replacing P by P^*. □

This lemma allows us to define a *normal operator* $T: V \longrightarrow V$ on a hermitian space V to be a linear operator whose matrix M with respect to any orthonormal basis is a normal matrix.

(7.3) **Theorem.** A complex matrix M is normal if and only if there is a unitary matrix P such that PMP^* is diagonal. □

The most important normal matrices, aside from hermitian ones, are unitary matrices: Since $M^* = M^{-1}$ if M is unitary, $MM^* = M^*M = I$, which shows that M is normal.

(7.4) **Corollary.** Every conjugacy class in the unitary group contains a diagonal matrix. □

Proof of Theorem (7.3). First, any two diagonal matrices commute, so a diagonal matrix is normal: $DD^* = D^*D$. The lemma tells us that M is normal if $PMP^* = D$. Conversely, suppose that M is normal. Choose an eigenvector $v = v_1$ of M, and normalize so that $\langle v, v\rangle = 1$, as in the proof of (5.4). Extend $\{v_1\}$ to an orthonormal basis. Then M will be changed to a matrix

$$M_1 = PMP^* = \left[\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1n} \\ \hline 0 & & & \\ \vdots & & N & \\ 0 & & & \end{array}\right], \quad \text{and} \quad M_1{}^* = PM^*P^* = \left[\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline \overline{a}_{12} & & & \\ \vdots & & N^* & \\ \overline{a}_{1n} & & & \end{array}\right].$$

The upper left entry of $M_1{}^*M_1$ is $a_{11}\overline{a}_{11}$, while the same entry of $M_1M_1{}^*$ is $a_{11}\overline{a}_{11}+a_{12}\overline{a}_{12}+\cdots+a_{1n}\overline{a}_{1n}$. Since M is normal, so is M_1, that is, $M_1{}^*M_1{}^* = M_1M_1{}^*$. It follows that $a_{12}\overline{a}_{12}+\cdots+a_{1n}\overline{a}_{1n} = 0$. Since $a_{1j}\overline{a}_{1j} \geq 0$, this shows that the entries a_{1j} with $j > 1$ are zero and that

$$M_1 = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{bmatrix}.$$

We continue, working on N. $\square$

8. SKEW-SYMMETRIC FORMS

The theory of skew-symmetric forms is independent of the field of scalars. One might expect trouble with fields of characteristic 2, in which $1 + 1 = 0$. They look peculiar because $a = -a$ for all a, so the conditions for symmetry (1.5) and for skew symmetry (1.6) are the same. It turns out that fields of characteristic 2 don't cause trouble with skew-symmetric forms, if the definition of skew symmetry is changed to handle them. The definition which works for all fields is this:

(8.1) **Definition.** A bilinear form $\langle\, ,\rangle$ on a vector space V is *skew-symmetric* if

$$\langle v, v\rangle = 0$$

for all $v \in V$.

The rule

$$\langle v, w\rangle = -\langle w, v\rangle \tag{8.2}$$

for all $v, w \in V$ continues to hold with this definition. It is proved by expanding

$$\langle v + w, v + w\rangle = \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle,$$

and by using the fact that $\langle v, v\rangle = \langle w, w\rangle = \langle v + w, v + w\rangle = 0$. If the characteristic of the field of scalars is not 2, then (8.1) and (8.2) are equivalent. For if (8.2) holds for all v, w, then setting $w = v$ we find $\langle v, v\rangle = -\langle v, v\rangle$. This implies that $2\langle v, v\rangle = 0$, hence that $\langle v, v\rangle = 0$ unless $2 = 0$ in the field.

Note that if F has characteristic 2, then $1 = -1$ in F, so (8.2) shows that the form is actually symmetric. But most symmetric forms don't satisfy (8.1).

The matrix A of a skew-symmetric form with respect to an arbitrary basis is characterized by the properties

$$a_{ii} = 0 \quad \text{and} \quad a_{ij} = -a_{ji}, \quad \text{if } i \neq j. \tag{8.3}$$

We take these properties as the definition of a *skew-symmetric matrix*. If the characteristic is not 2, then this is equivalent with the condition

$$A^t = -A. \tag{8.4}$$

(8.5) **Theorem.**

(a) Let V be a vector space of dimension m over a field F, and let $\langle\,,\rangle$ be a nondegenerate skew-symmetric form on V. Then m is an even integer, and there is a basis $\mathbf{B}$ of V such that the matrix A of the form with respect to that basis is

$$J_{2n} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

where $0, I$ denote the $n \times n$ matrices and $n = \frac{1}{2}m$.

(b) *Matrix form:* Let A be a nonsingular skew-symmetric $m \times m$ matrix. Then m is even, and there is a matrix $Q \in GL_m(F)$ such that QAQ^t is the matrix J_{2n}.

A basis $\mathbf{B}$ as in (8.6a) is called a *standard symplectic basis*. Note that rearranging the standard symplectic basis in the order $(v_1, v_{n+1}, v_2, v_{n+2}, \ldots, v_n, v_{2n})$ changes the matrix J_{2n} into a matrix made up of 2×2 blocks

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

along the diagonal. This is the form which is most convenient for proving the theorem. We leave the proof as an exercise. □

9. SUMMARY OF RESULTS, IN MATRIX NOTATION

Real numbers: A square matrix A is *symmetric* if $A^t = A$ and *orthogonal* if $A^t = A^{-1}$.

(1) *Spectral Theorem:* If A is a real symmetric matrix, there is an orthogonal matrix P such that $PAP^t(= PAP^{-1})$ is diagonal.

(2) If A is a real symmetric matrix, there is a real invertible matrix P such that

$$PAP^t = \begin{bmatrix} I_p & & \\ & -I_m & \\ & & 0_z \end{bmatrix},$$

for some integers p, m, z.

(3) *Sylvester's Law:* The numbers p, m, z are determined by the matrix A.

Complex numbers: A complex square matrix A is *hermitian* if $A^* = A$, *unitary* if $A^* = A^{-1}$, and *normal* if $AA^* = A^*A$.

(1) *Spectral Theorem*: If A is a hermitian matrix, there is a unitary matrix P such that PAP^* is a real diagonal matrix.

(2) If A is a normal matrix, there is a unitary matrix P such that PAP^* is diagonal.

F arbitrary: A square $n \times n$ matrix is *skew-symmetric* if $a_{ii} = 0$ and $a_{ij} = -a_{ji}$ for all i, j. If A is an invertible skew-symmetric matrix, then n is even, and there is an invertible matrix P so that PAP^t has the form

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

(9.1) *Note*. The rule $A' = (P^t)^{-1}A(P^{-1})$ for change of basis in a bilinear form (see (1.12)) is rather ugly because of the way the matrix P of change of coordinates is defined. It is possible to rearrange equations (4.17) of Chapter 3, by writing

$$(9.2) \qquad v_i' = \sum_j q_{ij}v_j \quad \text{or} \quad \mathbf{B}'^{t} = Q\mathbf{B}^t.$$

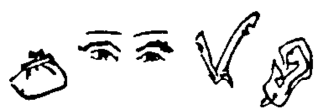
This results in $Q = (P^{-1})^t$, and with this rule we obtain the nicer formula

$$A' = QAQ^t,$$

to replace (1.12). We can use it if we want to.

The problem with formula (9.2) is that change of basis on a linear transformation gets messed up; namely the formula $A' = PAP^{-1}$ [Chapter 4 (3.4)] is replaced by $A' = (Q^{-1})^t AQ^t$. Trying to keep the formulas neat is like trying to smooth a bump in a rug.

This brings up an important point. Linear operators on V and bilinear forms on V are each given by an $n \times n$ matrix A, once a basis has been chosen. One is tempted to think that the theories of linear operators and of bilinear forms are somehow equivalent, but they are not, unless a basis is fixed. For under a *change* of basis the matrix of a bilinear form changes to $(P^t)^{-1}AP^{-1}$ (1.12), while the matrix of a linear operator changes to PAP^{-1} [Chapter 4 (3.4)]. So the new matrices are no longer equal. To be precise, this shows that the theories diverge when the basis is changed, unless the matrix P of change of basis happens to be orthogonal. If P is orthogonal, then $P = (P^t)^{-1}$, and we are all right. The matrices remain equal. This is one benefit of working with orthonormal bases.



Yvonne Verdier

EXERCISES

1. Definition of Bilinear Form

1. Let A and B be real $n \times n$ matrices. Prove that if $X^tAY = X^tBY$ for all vectors X, Y in $\mathbb{R}^n$, then $A = B$.

2. Prove directly that the bilinear form represented by the matrix $\begin{bmatrix} a & b \\ b & d \end{bmatrix}$ is positive definite if and only if $a > 0$ and $ad - b^2 > 0$.
3. Apply the Gram–Schmidt procedure to the basis $(1, 1, 0)^t$, $(1, 0, 1)^t$, $(0, 1, 1)^t$, when the form is dot product.
4. Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$. Find an orthonormal basis for $\mathbb{R}^2$ with respect to the form X^tAY.
5. (a) Prove that every real square matrix is the sum of a symmetric matrix and a skew-symmetric matrix ($A^t = -A$) in exactly one way.
   (b) Let $\langle\, ,\rangle$ be a bilinear form on a real vector space V. Show that there is a symmetric form $(\, ,)$ and a skew-symmetric form $[\, ,]$ so that $\langle\, ,\rangle = (\, ,) + [\, ,]$.
6. Let $\langle\, ,\rangle$ be a symmetric bilinear form on a vector space V over a field F. The function q: $V \longrightarrow F$ defined by $q(v) = \langle v, v\rangle$ is called the *quadratic form* associated to the bilinear form. Show how to recover the bilinear form from q, if the characteristic of the field F is not 2, by expanding $q(v + w)$.

*7. Let X, Y be vectors in $\mathbb{C}^n$, and assume that $X \neq 0$. Prove that there is a symmetric matrix B such that $BX = Y$.

2. Symmetric Forms: Orthogonality

1. Prove that a positive definite form is nondegenerate.
2. A matrix A is called *positive semidefinite* if $X^tAX \geq 0$ for all $X \in \mathbb{R}^n$. Prove that A^tA is positive semidefinite for any $m \times n$ real matrix A.
3. Find an orthogonal basis for the form on $\mathbb{R}^n$ whose matrix is as follows.

   (a) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$
4. Extend the vector $X_1 = (1, 1, 1)^t/\sqrt{3}$ to an orthonormal basis for $\mathbb{R}^3$.

*5. Prove that if the columns of an $n \times n$ matrix A form an orthonormal basis, then the rows do too.

6. Let A, A' be symmetric matrices related by $A = P^tA'P$, where $P \in GL_n(F)$. Is it true that the ranks of A and of A' are equal?
7. Let A be the matrix of a symmetric bilinear form $\langle\, ,\rangle$ with respect to some basis. Prove or disprove: The eigenvalues of A are independent of the basis.
8. Prove that the only real matrix which is orthogonal, symmetric, and positive definite is the identity.
9. The vector space P of all real polynomials of degree $\leq n$ has a bilinear form, defined by
$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx.$$
   Find an orthonormal basis for P when n has the following values. (a) 1 (b) 2 (c) 3
10. Let V denote the vector space of real $n \times n$ matrices. Prove that $\langle A, B\rangle = \text{trace}(A^tB)$ is a positive definite bilinear form on V. Find an orthonormal basis for this form.

11. A symmetric matrix A is called *negative definite* if $X^t AX < 0$ for all $X \neq 0$. Give a criterion analogous to (1.26) for a symmetric matrix A to be negative definite.

12. Prove that every symmetric nonsingular complex matrix A has the form $A = P^t P$.

13. In the notation of (2.12), show by example that the span of $(v_1, \dots, v_p)$ is not determined by the form.

14. **(a)** Let W be a subspace of a vector space V on which a symmetric bilinear form is given. Prove that $W^\perp$ is a subspace.

(b) Prove that the null space N is a subspace.

15. Let W_1, W_2 be subspaces of a vector space V with a symmetric bilinear form. Prove each of the following.

(a) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$ **(b)** $W \subset W^{\perp\perp}$ **(c)** If $W_1 \subset W_2$, then $W_1^\perp \supset W_2^\perp$.

16. Prove Proposition (2.7), that $V = W \oplus W^\perp$ if the form is nondegenerate on W.

17. Let $V = \mathbb{R}^{2\times 2}$ be the vector space of real 2×2 matrices.

(a) Determine the matrix of the bilinear form $\langle A, B \rangle = \operatorname{trace}(AB)$ on V with respect to the standard basis $\{e_{ij}\}$.

(b) Determine the signature of this form.

(c) Find an orthogonal basis for this form.

(d) Determine the signature of the form on the subspace of V of matrices with trace zero.

***18.** Determine the signature of the form $\langle A, B \rangle = \operatorname{trace} AB$ on the space $\mathbb{R}^{n\times n}$ of real $n \times n$ matrices.

19. Let $V = \mathbb{R}^{2\times 2}$ be the space of 2×2 matrices.

(a) Show that the form $\langle A, B \rangle$ defined by $\langle A, B \rangle = \det(A + B) - \det A - \det B$ is symmetric and bilinear.

(b) Compute the matrix of this form with respect to the standard basis $\{e_{ij}\}$, and determine the signature of the form.

(c) Do the same for the subspace of matrices of trace zero.

20. Do exercise 19 for $\mathbb{R}^{3\times 3}$, replacing the quadratic form $\det A$ by the coefficient of t in the characteristic polynomial of A.

21. Decide what the analogue of Sylvester's Law for symmetric forms over complex vector spaces is, and prove it.

22. Using the method of proof of Theorem (2.9), find necessary and sufficient conditions on a field F so that every finite-dimensional vector space V over F with a symmetric bilinear form $\langle\, ,\rangle$ has an orthogonal basis.

23. Let $F = \mathbb{F}_2$, and let $A = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}$.

(a) Prove that the bilinear form $X^t AY$ on F^2 can not be diagonalized.

(b) Determine the orbits for the action $P, A \rightsquigarrow PAP^t$ of $GL_2(F)$ on the space of 2×2 matrices with coefficients in F.

3. The Geometry Associated to a Positive Form

1. Let V be a Euclidean space. Prove the Schwarz Inequality and the Triangle Inequality.

2. Let W be a subspace of a Euclidean space V. Prove that $W = W^{\perp\perp}$.

3. Let V be a Euclidean space. Show that if $|v| = |w|$, then $(v + w) \perp (v - w)$. Interpret this formula geometrically.

4. Prove the parallelogram law $|v + w|^2 + |v - w|^2 = 2|v|^2 + 2|w|^2$ in a Euclidean space.
5. Prove that the orthogonal projection (3.7) is a linear transformation.
6. Find the matrix of the projection $\pi: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ such that the image of the standard bases of $\mathbb{R}^3$ forms an equilateral triangle and $\pi(e_1)$ points in the direction of the x-axis.

*7. Let W be a two-dimensional subspace of $\mathbb{R}^3$, and consider the orthogonal projection π of $\mathbb{R}^3$ onto W. Let $(a_i, b_i)^t$ be the coordinate vector of $\pi(e_i)$, with respect to a chosen orthonormal basis of W. Prove that (a_1, a_2, a_3) and (b_1, b_2, b_3) are orthogonal unit vectors.

*8. Let $w \in \mathbb{R}^n$ be a vector of length 1.
   (a) Prove that the matrix $P = I - 2ww^t$ is orthogonal.
   (b) Prove that multiplication by P is a reflection through the space W orthogonal to w, that is, prove that if we write an arbitrary vector v in the form $v = cw + w'$, where $w' \in W^\perp$, then $Pv = -cw + w'$.
   (c) Let X, Y be arbitrary vectors in $\mathbb{R}^n$ with the same length. Determine a vector w such that $PX = Y$.

*9. Use exercise 8 to prove that every orthogonal $n \times n$ matrix is a product of at most n reflections.

10. Let A be a real symmetric matrix, and let T be the linear operator on $\mathbb{R}^n$ whose matrix is A.
   (a) Prove that $(\ker T) \perp (\operatorname{im} T)$ and that $V = (\ker T) \oplus (\operatorname{im} T)$.
   (b) Prove that T is an orthogonal projection onto im T if and only if, in addition to being symmetric, $A^2 = A$.
11. Let A be symmetric and positive definite. Prove that the maximal matrix entries are on the diagonal.

4. Hermitian Forms

1. Verify rules (4.4).
2. Show that the dot product form $(X \cdot Y) = X^tY$ is not positive definite on $\mathbb{C}^n$.
3. Prove that a matrix A is hermitian if and only if the associated form X^*AX is a hermitian form.
4. Prove that if X^*AX is real for all complex vectors X, then A is hermitian.
5. Prove that the $n \times n$ hermitian matrices form a real vector space, and find a basis for that space.
6. Let V be a two-dimensional hermitian space. Let (v_1, v_2) be an orthonormal basis for V. Describe all orthonormal bases (v_1', v_2') with $v_1 = v_1'$.
7. Let $X, Y \in \mathbb{C}^n$ be orthogonal vectors. Prove that $|X + Y|^2 = |X|^2 + |Y|^2$.
8. Is $\langle X, Y \rangle = x_1y_1 + ix_1y_2 - ix_2y_1 + ix_2y_2$ on $\mathbb{C}^2$ a hermitian form?
9. Let A, B be positive definite hermitian matrices. Determine which of the following matrices are positive definite hermitian: A^2, A^{-1}, AB, $A + B$.
10. Prove that the determinant of a hermitian matrix is a real number.
11. Prove that A is positive definite hermitian if and only if $A = P^*P$ for some invertible matrix P.
12. Prove Theorem (4.19), that a hermitian form on a complex vector space V has an orthonormal basis if and only if it is positive definite.

13. Extend the criterion (1.26) for positive definiteness to hermitian matrices.

14. State and prove an analogue of Sylvester's Law for hermitian matrices.

15. Let $\langle\, ,\rangle$ be a hermitian form on a complex vector space V, and let $\{v, w\}$ denote the real part of the complex number $\langle v, w\rangle$. Prove that if V is regarded as a real vector space, then $\{\, ,\}$ is a symmetric bilinear form on V, and if $\langle\, ,\rangle$ is positive definite, then $\{\, ,\}$ is too. What can you say about the imaginary part?

16. Let P be the vector space of polynomials of degree $\leq n$.

(a) Show that

$$\langle f, g\rangle = \int_0^{2\pi} \overline{f(e^{i\theta})}g(e^{i\theta})d\theta$$

is a positive definite hermitian form on P.

(b) Find an orthonormal basis for this form.

17. Determine whether or not the following rules define hermitian forms on the space $\mathbb{C}^{n\times n}$ of complex matrices, and if so, determine their signature.

(a) $A, B \rightsquigarrow \text{trace}(A^*B)$ **(b)** $A, B \rightsquigarrow \text{trace}(\overline{A}B)$

18. Let A be a unitary matrix. Prove that $|\det A| = 1$.

19. Let P be a unitary matrix, and let X_1, X_2 be eigenvectors for P, with distinct eigenvalues λ_1, λ_2. Prove that X_1 and X_2 are orthogonal with respect to the standard hermitian product on $\mathbb{C}^n$.

***20.** Let A be any complex matrix. Prove that $I + A^*A$ is nonsingular.

21. Prove Proposition (4.20).

5. The Spectral Theorem

1. Prove that if T is a hermitian operator then the rule $\{v, w\} = \langle v, Tw\rangle = X^*MY$ defines a second hermitian form on V.

2. Prove that the eigenvalues of a real symmetric matrix are real numbers.

3. Prove that eigenvectors associated to distinct eigenvalues of a hermitian matrix A are orthogonal.

4. Find a unitary matrix P so that PAP^* is diagonal, when

$$A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

5. Find a real orthogonal matrix P so that PAP^t is diagonal, when

$$\textbf{(a)}\ A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad \textbf{(b)}\ A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \textbf{(c)}\ A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

6. Prove the equivalence of conditions (a) and (b) of the Spectral Theorem.

7. Prove that a real symmetric matrix A is positive definite if and only if its eigenvalues are positive.

8. Show that the only matrix which is both positive definite hermitian and unitary is the identity I.

9. Let A be a real symmetric matrix. Prove that e^A is symmetric and positive definite.

10. Prove that for any square matrix A, $\ker A = (\operatorname{im} A^*)^\perp$.

*11. Let $\zeta = e^{2\pi i/n}$, and let A be the $n \times n$ matrix $a_{jk} = \zeta^{jk}/\sqrt{n}$. Prove that A is unitary.

12. Show that for every complex matrix A there is a unitary matrix P such that PAP^* is upper triangular.

13. Let A be a hermitian matrix. Prove that there is a unitary matrix P with determinant 1 such that PAP^* is diagonal.

*14. Let A, B be hermitian matrices which commute. Prove that there is a unitary matrix P such that PAP^* and PBP^* are both diagonal.

15. Use the Spectral Theorem to give a new proof of the fact that a positive definite real symmetric $n \times n$ matrix P has the form $P = AA^t$ for some $n \times n$ matrix A.

16. Let λ, μ be distinct eigenvalues of a complex symmetric matrix A, and let X, Y be eigenvectors associated to these eigenvalues. Prove that X is orthogonal to Y with respect to dot product.

6. Conics and Quadrics

1. Determine the type of the quadric $x^2+4xy+2xz+z^2 + 3x+z - 6 = 0$.
2. Suppose that (6.1) represents an ellipse. Instead of diagonalizing the form and then making a translation to reduce to the standard type, we could make the translation first. Show how to compute the required translation by calculus.
3. Discuss all degenerate loci for conics.
4. Give a necessary and sufficient condition, in terms of the coefficients of its equation, for a conic to be a circle.
5. (a) Describe the types of conic in terms of the signature of the quadratic form.
   (b) Do the same for quadrics in $\mathbb{R}^3$.
6. Describe the degenerate quadrics, that is, those which are not listed in (6.14).

7. The Spectral Theorem for Normal Operators

1. Show that for any normal matrix A, $\ker A = (\operatorname{im} A)^\perp$.
2. Prove or disprove: If A is a normal matrix and W is an A-invariant subspace of $V = \mathbb{C}^n$, then $W^\perp$ is also A-invariant.
3. A matrix is skew-hermitian if $A^* = -A$. What can you say about the eigenvalues and the possibility of diagonalizing such a matrix?
4. Prove that the cyclic shift operator

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

   is normal, and determine its diagonalization.
5. Let P be a real matrix which is normal and has real eigenvalues. Prove that P is symmetric.
6. Let P be a real skew-symmetric matrix. Prove that P is normal.

*7. Prove that the *circulant*

$$\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_n \\ c_n & c_0 & c_1 & \cdots & c_{n-1} \\ \vdots & & & & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}$$

is a normal matrix.

8. (a) Let A be a complex symmetric matrix. Prove that eigenvectors of A with distinct eigenvalues are orthogonal with respect to the bilinear form X^tX.
*(b) Give an example of a complex symmetric matrix A such that there is no $P \in O_n(\mathbb{C})$ with PAP^t diagonal.

9. Let A be a normal matrix. Prove that A is hermitian if and only if all eigenvalues of A are real, and that A is unitary if and only if every eigenvalue has absolute value 1.

10. Let V be a finite-dimensional complex vector space with a positive definite hermitian form $\langle\, ,\rangle$, and let T: $V \longrightarrow V$ be a linear operator on V. Let A be the matrix of T with respect to an orthonormal basis **B**. The adjoint operator T^*: $V \longrightarrow V$ is defined as the operator whose matrix with respect to the same basis is A^*.
(a) Prove that T and T^* are related by the equations $\langle Tv, w\rangle = \langle v, T^*w\rangle$ and $\langle v, Tw\rangle = \langle T^*v, w\rangle$ for all $v, w \in W$. Prove that the first of these equations characterizes T^*.
(b) Prove that T^* does not depend on the choice of orthonormal basis.
(c) Let v be an eigenvector for T with eigenvalue λ, and let $W = v^\perp$ be the space of vectors orthogonal to v. Prove that W is T^*-invariant.

11. Prove that for any linear operator T, TT^* is hermitian.

12. Let V be a finite-dimensional complex vector space with a positive definite hermitian form $\langle\, ,\rangle$. A linear operator T: $V \longrightarrow V$ is called *normal* if $TT^* = T^*T$.
(a) Prove that T is normal if and only if $\langle Tv, Tw\rangle = \langle T^*v, T^*w\rangle$ for all $v, w \in V$, and verify that hermitian operators and unitary operators are normal.
(b) Assume that T is a normal operator, and let v be an eigenvector for T, with eigenvalue λ. Prove that v is also an eigenvector for T^*, and determine its eigenvalue.
(c) Prove that if v is an eigenvector, then $W = v^\perp$ is T-invariant, and use this to prove the Spectral Theorem for normal operators.

8. Skew-Symmetric Forms

1. Prove or disprove: A matrix A is skew-symmetric if and only if $X^tAX = 0$ for all X.
2. Prove that a form is skew-symmetric if and only if its matrix has the properties (8.4).
3. Prove or disprove: A skew-symmetric $n \times n$ matrix is singular if n is odd.
4. Prove or disprove: The eigenvalues of a real skew-symmetric matrix are purely imaginary.

*5. Let S be a real skew-symmetric matrix. Prove that $I + S$ is invertible, and that $(I - S)(I + S)^{-1}$ is orthogonal.

*6. Let A be a real skew-symmetric matrix.
(a) Prove that $\det A \geq 0$.
(b) Prove that if A has integer entries, then $\det A$ is the square of an integer.

7. Let $\langle\ ,\ \rangle$ be a skew-symmetric form on a vector space V. Define orthogonality, null space, and nondegenerate forms as in Section 2.
   (a) Prove that the form is nondegenerate if and only if its matrix with respect to any basis is nonsingular.
   (b) Prove that if W is a subspace such that the restriction of the form to W is nondegenerate, then $V = W \oplus W^{\perp}$.
   (c) Prove that if the form is not identically zero, then there is a subspace W and a basis of W such that the restriction of the form to W has matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.
   (d) Prove Theorem (8.6).

9. Summary of Results, in Matrix Notation

1. Determine the symmetry of the matrices $AB + BA$ and $AB - BA$ in the following cases.
   (a) A, B symmetric (b) A, B hermitian (c) A, B skew-symmetric (d) A symmetric, B skew-symmetric
2. State which of the following rules define operations of $GL_n(\mathbb{C})$ on the space $\mathbb{C}^{n\times n}$ of all complex matrices:
$$P, A \rightsquigarrow PAP^t,\ (P^{-1})^t A (P^{-1}),\ (P^{-1})^t A P^t,\ P^{-1}AP^t,\ AP^t,\ P^tA.$$
3. (a) With each of the following types of matrices, describe the possible determinants:
   (i) real orthogonal (ii) complex orthogonal (iii) unitary (iv) hermitian
   (v) symplectic (vi) real symmetric, positive definite (vii) real symmetric, negative definite
   (b) Which of these types of matrices have only real eigenvalues?
4. (a) Let E be an arbitrary complex matrix. Prove that the matrix $\begin{bmatrix} I & E^* \\ -E & I \end{bmatrix}$ is invertible.
   (b) Find the inverse in block form $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$.

*5. (a) What is wrong with the following argument? Let P be a real orthogonal matrix. Let X be a (possibly complex) eigenvector of P, with eigenvalue λ. Then $X^tP^tX = (PX)^tX = \lambda X^tX$. On the other hand, $X^tP^tX = X^t(P^{-1}X) = \lambda^{-1}X^tX$. Therefore $\lambda = \lambda^{-1}$, and so $\lambda = \pm 1$.
   (b) State and prove a correct theorem based on this argument.

*6. Show how to describe any element of SO_4 in terms of rotations of two orthogonal planes in $\mathbb{R}^4$.

*7. Let A be a real $m \times n$ matrix. Prove that there are orthogonal matrices $P \in O_m$ and $Q \in O_n$ such that $PAQ = D$ is diagonal, with nonnegative diagonal entries.