# 1

# INNER PRODUCT SPACES

## 1-1 Definitions and Examples

The fundamental concept in this chapter is that of a *symmetric bilinear form.*

Let $K$ be a field, $V$ a vector space over $K$. Until §1-15 our main concern will be with the finite-dimensional case, but we allow infinite-dimensionality, and state the theorems in maximum generality; (e.g., in Theorem 2, $V$ is allowed to be infinite-dimensional but the subspace $S$ must be finite-dimensional).

Normally our notation will be early letters of the alphabet for scalars (elements of $K$), and late letters for vectors. We consider a function of two variables, each of the variables ranging over $V$, with the values of the function lying in $K$. Our notation for the function will usually be $(\ ,\ )$; but when desirable we shall use a functional symbol such as $f(\ ,\ )$, $g(\ ,\ )$.

The assumptions are, in brief, symmetry and linearity in each variable. In detail:

$$(1) \qquad (x,y) = (y,x)$$

$$(2) \qquad (ax + by, z) = a(x,z) + b(y,z)$$

Thus we are explicitly assuming only linearity in the first variable. But (1) and (2) combine to imply linearity in the second variable as well:

$$(z, ax + by) = a(z,x) + b(z,y)$$

We shall usually call the form an *inner product*, and we speak of the structure consisting of $V$ together with the form as an *inner product space*.

We offer a suggestion to the reader: whenever he is faced with a new concept he should check to see whether there are any examples, however trivial. In that spirit we offer as a first example the inner product (defined on any $V$) that is 0 for all vectors: $(x,y) = 0$ for all $x,y$ in $V$.

Next we mention one of the main motivating examples. With $K$ the field of real numbers, we think of ordinary Euclidean space; $(x,x)$ is the square of the length of $x$, and $(x,y)$ is the scalar product of vector analysis (the product of the lengths of $x$ and $y$ and the cosine of the included angle). We shall return to this example in §2-1.

Our further discussion of examples will be in terms of a basis of $V$, and will be, in effect, a construction of all examples. Let then $u_1, \cdots, u_n$ be a basis of $V$. Write $(u_i,u_j) = a_{ij}$. Note that by the symmetry of $(\ ,\ )$ we have $a_{ij} = a_{ji}$. A typical element $x$ of $V$ has the form $x = x_1u_1 + \cdots + x_nu_n$. (We violate our convention concerning early and late letters of the alphabet when we take the coordinates of a vector relative to a basis.) If $y = y_1u_1 + \cdots + y_nu_n$ is a second element, we find from the bilinearity

$$(3) \qquad \begin{aligned}(x,y) &= a_{11}x_1y_1 + a_{12}x_1y_2 + \cdots + a_{ij}x_iy_j + \cdots + a_{nn}x_ny_n \\ &= \Sigma a_{ij}x_iy_j\end{aligned}$$

Conversely, we may take $a_{ij}$ to be any elements of $K$ subject to the condition $a_{ij} = a_{ji}$, define $(x,y)$ by (3), and obtain an inner product space.

In a way this is a description of all inner product spaces, but there is more to be said (in fact a good deal more), for we would like to know how the $a_{ij}$'s are affected by a change of basis.

For this purpose we first introduce the matrix $(a_{ij})$; let us write it $A$. We note that it is an $n \times n$ symmetric matrix with entries in the field $K$. We call it the *matrix of the form* relative to the basis $u_1, \cdots, u_n$.

Now let a second basis $v_1, \cdots, v_n$ of $V$ be given. We express the second basis in terms of the first:

$$\begin{aligned}v_1 &= p_{11}u_1 + p_{12}u_2 + \cdots + p_{1n}u_n \\ &\ \ \vdots \\ v_n &= p_{n1}u_1 + p_{n2}u_2 + \cdots + p_{nn}u_n\end{aligned}$$



Here the elements $(p_{ij})$ form an $n \times n$ non-singular matrix for which we write $P$. For the inner product $(v_i,v_j)$ our notation is $b_{ij}$, and this gives us a second $n \times n$ symmetric matrix $B$. Now comes a computation, which is brief when written under summation signs, but perhaps a little indigestible:

$$\begin{aligned}b_{ij} &= (v_i,v_j) \\ &= (\sum_k p_{ik}u_k, \sum_m p_{jm}u_m) \\ &= \sum_{k,m} p_{ik}a_{km}p_{jm}\end{aligned}$$

Now $\sum_k p_{ik}a_{km}$ is (by the definition of matrix multiplication) the $i,m$-term of the matrix $PA$. Let us write $c_{im}$ for $\sum_k p_{ik}a_{km}$. Then the remaining sum to be evaluated is $\sum_m c_{im}p_{jm}$. In order to recognize this as a matrix product we have to transpose the matrix $P$. To see this in detail, write $p_{jm} = q_{mj}$ and regard $q_{mj}$ as the $m,j$-entry of a matrix $Q$. Then $\sum_m c_{im}p_{jm} = \sum_m c_{im}q_{mj}$ is the $i,j$-entry of $PAQ$. Now $Q$ is the transpose of $P$, for which our notation will be $P'$. The upshot of all this is the equation

$$(4) \qquad B = PAP'$$

*Remark.* For a discussion in a more conceptual style, see §1-14.

The matrix relation (4) just encountered deserves recognition on its own merits, and so we define it for any (not necessarily symmetric) square matrices.

Definition. Two $n \times n$ matrices $A$ and $B$ are *congruent* if there exists a non-singular matrix $P$ such that $PAP' = B$.

To show that congruence is an equivalence relation will be left as an exercise (Ex. 1, page 4). However, we should also note that it follows from the above discussion that congruence is an equivalence relation, at least on symmetric matrices; if we had carried through our discussion for general bilinear forms (i.e., not necessarily symmetric ones) the remark would apply to congruence of arbitrary matrices.

We next introduce the concepts of radical and non-singularity.

Let $(\ ,\ )$ be a symmetric bilinear form on a vector space $V$. The set of vectors $x$ satisfying $(x,V) = 0$ ($(x,y) = 0$ for all $y$ in $V$) is called the *radical* of the form. It is routine to see that the radical is a subspace. When the radical is 0, we say that the form is *non-singular*. The theory of general inner product spaces can be reduced to non-singular ones. (See Ex. 2 or Ex. 3, page 4.)

We proceed to relate non-singularity of an inner product to non-singularity of the matrix attached to it.

Theorem 1. *Let $V$ be a finite-dimensional vector space over a field $K$, and let $(\ ,\ )$ be a symmetric bilinear form on $V$. Let $u_1, \cdots, u_n$ be a basis of $V$ and let $A = (a_{ij})$ be the matrix determined by the form relative to this basis: $a_{ij} = (u_i,u_j)$. Then $(\ ,\ )$ is non-singular if and only if $A$ is non-singular.*

Proof. We prove instead the equivalent statement that $(\ ,\ )$ is singular if and only if $A$ is singular.

Suppose the matrix $A$ is singular. Then there exist elements $c_1, \cdots, c_n$ in $K$, not all 0, such that the linear combination of the rows of $A$ with coefficients $c_1, \cdots, c_n$ is 0. Write $x = c_1u_1 + \cdots + c_nu_n$. Then for every $i$ we have $(x,u_i) = 0$. Thus $x$ is a non-zero element in the radical.

The proof of the converse merely requires retracing these steps. Suppose the form is singular, so that there is a non-zero element $x$ in the radical. Write $x = c_1u_1 + \cdots + c_nu_n$, and note that the $c$'s are not all 0. Then $(x,u_i) = 0$ for all $i$. This translates to the statement that the rows of $A$ are linearly dependent via the coefficients $c_1, \cdots, c_n$. Hence $A$ is singular.

### EXERCISES

1. Prove that congruence of matrices is an equivalence relation.
2. Let $V$ be an inner product space, $N$ its radical. Write $\bar{V}$ for the quotient space $V/N$. For $\bar{x},\bar{y}$ in $\bar{V}$ define $(\bar{x},\bar{y})$ by taking representatives $x,y$ in $V$ and setting $(\bar{x},\bar{y}) = (x,y)$.
   (a) Prove that $(\bar{x},\bar{y})$ is well-defined.
   (b) Prove that in this way we get an inner product space on $\bar{V}$ and that it is non-singular.
   (c) Show that the inner product space $V$ is uniquely determined by $\bar{V}$ and the dimension of $N$.
3. (This is a less invariant but more "down to earth" version of Ex. 2.) Let $V$ be an inner product space with radical $N$. Let $W$ be a vector space complement of $N$. Show that the inner product is non-singular when restricted to $W$, and that the inner product on $V$ is uniquely determined by that on $W$ and the dimension of $N$.
4. Let $V$ be a finite-dimensional inner product space with matrix $(a_{ij})$ relative to a basis. Prove that the rank of $(a_{ij}) = \dim V - \dim(\text{radical of } V)$. (*Hint:* use the theorem that for a set of $n$ linear homogeneous equations in $n$ variables the dimension of the space of solutions is given by $n$ − (rank of the coefficient matrix). (See §1-14 for a more conceptual discussion.)



5. Let $V$ be a non-singular inner product space, $W$ a finite-dimensional subspace. Prove that $W$ can be embedded in a non-singular finite-dimensional subspace. (*Hint:* if $W$ is singular, pick $x \neq 0$ in its radical, and then $y$ in $V$ with $(x,y) \neq 0$. Show that in the subspace spanned by $W$ and $y$ the dimension of the radical has diminished by 1.)
6. Let $V$ have basis $u_1, \cdots, u_n$ over $K$. Define an inner product by $(u_i,u_j) = 1$ for $i \neq j$, $(u_i,u_i) = a$. Let $N$ be the radical. Prove that $N$ is $(n-1)$-dimensional for $a = 1$, one-dimensional if $a = 1 - n$ and the characteristic does not divide $n$, 0 if $a \neq 1$ or $1 - n$. (*Hint:* the determinant
$$\begin{vmatrix} a & 1 & 1 & \cdots & 1 \\ 1 & a & 1 & \cdots & 1 \\ 1 & 1 & a & \cdots & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & 1 & \cdots & a \end{vmatrix}$$
can be evaluated. Think of $a$ as a variable. The first $n - 1$ derivatives vanish at $a = 1$ so we get a factor $(a-1)^{n-1}$. The other factor can be shown to be $a + n - 1$.)
7. Let $V$ have basis $u_1, \cdots, u_n$ over $K$. Define an inner product by $(u_i,u_i) = a$, $(u_i,u_{i+1}) = 1$, $(u_i,u_j) = 0$ otherwise. Discuss the radical for $n = 3$ and any $a$, and for $a = 2$ and any $n$. (*Hint:* the determinant
$$\begin{vmatrix} 2 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2 \end{vmatrix}$$
can be evaluated. If we call it $D_n$, prove that $D_n = 2D_{n-1} - D_{n-2}$ by expanding along the first row, and deduce by induction $D_n = n + 1$. Conclude that the radical is one-dimensional if the characteristic divides $n + 1$, and is otherwise 0. For $n = 3$ and general $a$, the determinant is $a^3 - 2a$.)

## 1-2 The Direct Summand Theorem

The next idea we introduce is that of the orthogonal complement of a subspace. Let $V$ be an inner product space, $S$ any subspace of $V$. We define the *orthogonal complement* of $S$ to be the set of all $x$ in $V$ satisfying $(x,S) = 0$, i.e., satisfying $(x,y) = 0$ for all $y$ in $S$. Our notation for the orthogonal complement of $S$ is $S'$.

Note that $V'$, the orthogonal complement of all of $V$, is the radical of $V$.

As in the case of the radical, we leave to the reader the routine argument that $S'$ is a subspace.

The phrase "orthogonal complement" probably suggests that $S'$ may well behave as some sort of complement to $S$, that is, that the two subspaces $S$ and $S'$ are disjoint (or as disjoint as subspaces can be) and that $S$ and $S'$ together span $V$. We shall proceed to see that, under a mild and inevitable restriction, this is indeed the case.

First we need a basic remark. If $V$ is an inner product space, and $S$ is a subspace of $V$, then we may regard $S$ as an inner product space by using the very same form $(\ ,\ )$. Note that even if $V$ is non-singular, a subspace $S$ may perfectly well be singular.

Note next that by definition, $S \cap S'$ is the radical of $S$. If we are to have $S \cap S' = 0$, as hinted above, it is indispensable to assume $S$ non-singular. It turns out that for finite-dimensional $S$ this assumption is sufficient.

Theorem 2. *Let $V$ be an inner product space and let $S$ be a finite-dimensional subspace that is non-singular relative to the induced inner product. Then $V$ is the direct sum of $S$ and its orthogonal complement $S'$.*

Proof. We have already remarked that $S \cap S' = 0$. It remains for us to prove that $S$ and $S'$ span $V$. Let $x$ be an arbitrary element of $V$; we must express $x$ as a sum of a vector in $S$ and a vector orthogonal to $S$.

Pick a basis of $S$, say $u_1, \cdots, u_n$. To solve our problem we must find scalars $a_1, \cdots, a_n \in K$ and an element $y$ in $S'$ such that

$$x = a_1u_1 + a_2u_2 + \cdots + a_nu_n + y \tag{5}$$

Suppose $(u_i,u_j) = c_{ij}$. Apply to (5) the process of taking an inner product with $u_i$, $i = 1, \cdots, n$. Since $(y,u_i) = 0$, we find

$$\begin{aligned} c_{11}a_1 + c_{12}a_2 + \cdots + c_{1n}a_n &= (x,u_1) \\ c_{21}a_1 + c_{22}a_2 + \cdots + c_{2n}a_n &= (x,u_2) \\ &\ \ \vdots \\ c_{n1}a_1 + c_{n2}a_2 + \cdots + c_{nn}a_n &= (x,u_n) \end{aligned} \tag{6}$$

We look at (6) as a system of $n$ equations for the $n$ unknowns $a_1, \ldots, a_n$. The matrix of coefficients, $(c_{ij})$, is non-singular by Theorem 1 and by our hypothesis that $S$ is non-singular. Hence the equations (6) have a solution (indeed a unique solution), and Theorem 2 is proved.



### EXERCISES

1. Let $\{S_i\}$ be subspaces of an inner product space $V$. Prove: $(\cup S_i)' = \cap S_i'$.
2. Let $V$ be an $n$-dimensional non-singular inner product space, $S$ an $r$-dimensional subspace. Prove that $S'$ is $(n - r)$-dimensional.
3. Let $x$ be a vector with $(x,x) = 0$ in a non-singular inner product space, and let $S$ be the subspace spanned by $x$. Prove that $S'$ has radical $S$.

## 1-3 Diagonalization

The process of diagonalization is normally thought of as referring to matrices. We shall indeed state the main theorem of this section in matrix form, but we prefer first to derive it in the inner product notation.

Given an inner product space $V$, we shall be seeking for it a basis $u_1$, $u_2$, $\cdots$ with the property that any two different $u$'s are orthogonal: $(u_i,u_j) = 0$ for $i \neq j$. We call such a basis an *orthogonal basis*. If in addition we have $(u_i,u_i) = 1$ for all $i$, we speak of the $u$'s as *orthonormal* and say they form an orthonormal basis. Orthonormality can usually be achieved only when the underlying field contains suitable square roots.

In the process of getting an orthogonal basis, characteristic 2 begins to create difficulties, as exemplified in the next theorem.

We make a definition at this point (*cf.* Ex. 3 in §1-2): a vector $x$ is *null* (or *isotropic*) if $(x,x) = 0$; otherwise it is *non-null* (*non-isotropic*). We say that an inner product space is non-isotropic if it contains no null vectors.

Theorem 3. *Let $V$ be an inner product space over a field of characteristic $\neq 2$. Assume that the inner product is not identically 0 on $V$. Then $V$ contains a non-null vector.*

Proof. Assume the contrary, that $(x,x) = 0$ for all $x$. In view of the identity

$$(x + y, x + y) = (x,x) + (y,y) + 2(x,y)$$

it follows that $2(x,y) = 0$ for all $x$ and $y$. For characteristic $\neq 2$ we get the contradiction $(x,y) = 0$ for all $x$ and $y$.

Our discussion continues with the restriction, characteristic $\neq 2$. Later (§1-10) we shall devote a special section to the case of characteristic 2.

Theorem 4. *Let $V$ be a finite-dimensional inner product space over a field of characteristic $\neq 2$. Then $V$ possesses an orthogonal basis.*

*Proof.* If $(\ ,\ )$ is identically 0 any basis will do. Otherwise, by Theorem 3 there exists an element, say $u_1$, with $(u_1,u_1) \neq 0$. Let $S$ denote the one-dimensional subspace of $V$ spanned by $u_1$. Obviously the inner product restricted to $S$ is non-singular. By Theorem 2, $V$ is the direct sum of $S$ and its orthogonal complement $S'$. If $V$ is $n$-dimensional, $S'$ is $(n-1)$-dimensional. By induction we may assume the theorem known for $S'$. Let $u_2, \cdots, u_n$ be a basis of the desired kind for $S'$. Together with $u_1$ these elements form an orthogonal basis for $V$.

Let $u_1, \cdots, u_n$ be an orthogonal basis for an inner product space $V$. Write $(u_i,u_i) = b_i$. Then the matrix of inner products is the diagonal matrix

$$\begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & 0 & \cdots & b_n \end{pmatrix}$$

Let $A = (a_{ij})$ be a given symmetric $n \times n$ matrix with entries in a field $K$ of characteristic $\neq 2$. We can invent an inner product space $V$ to go with it: we equip $V$ with a basis $v_1, \cdots, v_n$, define $(v_i,v_j)$ to be $a_{ij}$ and extend the form to all of $V$ by linearity. Now Theorem 4 asserts that there exists a second basis $u_1, \cdots, u_n$ with $(u_i,u_j) = 0$ for $i \neq j$. Say $(u_i,u_i) = b_i$, and let $B$ denote the diagonal matrix with diagonal entries $b_1, \cdots, b_n$. Let $P$ be the matrix effecting the change of basis from the $v$'s to the $u$'s, as in the discussion preceding Theorem 1. (Note that the roles of the $u$'s and $v$'s got interchanged.) Then we have $PAP' = B$. We summarize:

Theorem 5. *Let $A$ be a symmetric matrix with entries in a field $K$ of characteristic $\neq 2$. Then there exists a non-singular matrix $P$ with entries in $K$ such that $PAP'$ is diagonal.*

The inevitable question that arises is: to what extent are the diagonal entries $b_1, \cdots, b_n$ unique? There is one easy, unqualified assertion that can be made (Ex. 1): the number of zeros is invariant.

On the other hand there are two superficial changes that can be made: the $b$'s can be permuted (just perform the analogous permutation of the $u$'s), and each can be multiplied by a non-zero square. (If we replace $u_1$ by $cu_1$, then $b_1 = (u_1,u_1)$ gets replaced by $(cu_1,cu_1) = c^2b_1$.) These remarks are merely the starting point of a profound theory, of which we shall indicate only the beginnings. The important point to be borne in mind is that the



nature of the field $K$ plays a controlling role. For instance, if every element in $K$ is a square, then every non-zero $b_i$ can be converted to 1. (In the notation above, replace $u_i$ by $cu_i$ where $c^2 = 1/b_i$.) We state this as a formal theorem, and note at once the matrix version.

Theorem 6. *Let $K$ be a field of characteristic $\neq 2$ in which every element is a square. Let $V$ be a finite-dimensional inner product space over $K$. Then there exists a basis $u_1, \cdots, u_n$ of $V$ such that $(u_i,u_j) = 0$ for $i \neq j$ and each $(u_i,u_i) = 1$ or $0$. If $V$ is non-singular, each $(u_i,u_i) = 1$, i.e., $V$ has an orthonormal basis.*

Theorem 7. *Let $K$ be a field of characteristic $\neq 2$ in which every element is a square. Let $A$ be a symmetric matrix over $K$. Then there exists a non-singular matrix $P$ over $K$ such that $PAP'$ is diagonal with 1's and 0's on the diagonal. If $A$ is non-singular, $PAP' = I$, the identity matrix.*

We introduce at this point the idea of equivalence of forms. If $V$ and $W$ are inner product spaces over the same field, and there exists a one-to-one linear transformation of $V$ onto $W$ preserving the inner product, we say that $V$ and $W$ are *equivalent*, and write $V \sim W$. In symbols: if $f$ and $g$ are the inner products on $V$ and $W$, and $T$ is the linear transformation, the requirement is $g(Tx,Ty) = f(x,y)$. We also speak of $T$ as an *isometry* of $V$ onto $W$.

If $V$ is finite-dimensional and we diagonalize $V$ with diagonal entries $(a_1, \cdots, a_n)$, we write $V \sim (a_1, \cdots, a_n)$. If a second diagonalization yields $(b_1, \cdots, b_n)$, we write $(a_1, \cdots, a_n) \sim (b_1, \cdots, b_n)$.

## EXERCISES

1. Prove that the number of zero diagonal entries occurring in Theorem 4 or Theorem 5 is the dimension of the radical.
2. Return to the $V$ of Ex. 6 in §1-1.
   (a) For $a = 1$, prove $V \sim (1, 0, \cdots, 0)$.
   (b) For $n = 2$ and $a = 0$, prove $V \sim (1, -1)$ for characteristic $\neq 2$.
   (c) For $n = 2$ and $a \neq 0$, prove $V \sim (a, a(a^2 - 1))$.
3. Let $x_1, \cdots, x_n$ be elements of an inner product space $V$.
   (a) If the determinant of inner products $(x_i,x_j)$ is non-zero, prove that $x_1, \cdots, x_n$ are linearly independent.
   (b) Is the converse true?
   (c) Prove the converse true if $V$ is non-isotropic.

4. Let the inner product space $V$ have an orthonormal basis $u_1, \cdots, u_n$. Let $W$ be the subspace consisting of all $a_1u_1 + \cdots + a_nu_n$ with $\Sigma\, a_i = 0$. Prove that $W$ is non-singular if and only if the characteristic of the field does not divide $n$.
5. Let the inner product space $V$ have an infinite orthonormal basis $\{u_i\}$. Let $W$ be the subspace consisting of all elements having the sum of their coefficients 0. Prove that the orthogonal complement $W' = 0$, and that $W$ is non-singular.
6. Return to the $V$ of Ex. 7 in §1-1, and take $a = 2$. Let $W$ be as in Ex. 4, but with $n$ replaced by $n + 1$. Prove $V \sim W$. (*Hint:* use the basis $u_1 - u_2$, $-u_2 + u_3$, $u_3 - u_4$, $-u_4 + u_5, \cdots$ of $W$.)
7. Let $K$ be any field of characteristic $\neq 2$ and $e \in K$. Prove $(1, 1) \sim (e, e)$ if and only if $e$ is a sum of two squares.
8. Let $K$ be any field of characteristic $\neq 2$ and $e \in K$. Prove $(1, 1, 1, 1) \sim (e, e, e, e)$ if and only if $e$ is a sum of 4 squares. (*Hint:* the necessity is obvious. The sufficiency can be shown explicitly by a change of basis inspired by quaternions. Let $u_1, \cdots, u_4$ be an orthonormal basis. Say $e^2 = a^2 + b^2 + c^2 + d^2$. Take $v_1 = au_1 + bu_2 + cu_3 + du_4$, $v_2 = bu_1 - au_2 + du_3 - cu_4$, $v_3 = -cu_1 + du_2 + au_3 - bu_4$, $v_4 = du_1 + cu_2 - bu_3 - au_4$. See page 135.)
9. In Ex. 5 note that $V$ and $W$ are both non-singular but that $W + W' \neq V$. Thus observe that Theorem 2 requires the finite-dimensionality of $S$.
10. In the $V$ of Ex. 6 in §1-1, assume $n = a^2 - a + 1$. Prove: $V \sim (1, a-1, \cdots, a-1)$. (*Hint:* we have already done $a = 1$ (Ex. 2(a)). Assume $a \neq 1$. Take the diagonal inner product space with $(v_1,v_1) = 1$, $(v_i,v_i) = a - 1$ for $i = 2, \cdots, n$. Set $u_i = v_1 + v_i$ $(i = 2, \cdots, n)$, $u_1 = (a-1)^{-1}(v_2 + \cdots + v_n)$. Check $(u_i, u_j) = 1$ for $i \neq j$, $(u_i,u_i) = a$.)
11. (a) If the characteristic is 2, prove that the null vectors in an inner product space form a subspace.
    (b) Suppose given an inner product space with the property that the null vectors form a subspace. Prove that one of the following must be true:
    (i) any null vector is in the radical,
    (ii) the characteristic is 2. (*Hint:* if (i) is false take $x$ null, $(x,y) \neq 0$, say $(x,y) = 1$. If $a = -(y,y)/2$, then $y + ax$ is null but $y + bx$ is not null if $b \neq -(y,y)/2$.)

## 1-4 The Inertia Theorem

A new circle of ideas arises when the base field $K$ is ordered.

Definitions. Let $V$ be an inner product space over an ordered field $K$. We say that $V$ is *positive semi-definite* if $(x,x) \geqq 0$ for all $x$ in $V$. We say that $V$ is *positive definite* if $(x,x) > 0$ for every non-zero $x$ in $V$. The definitions of negative semi-definite and negative definite are analogous.



Let $K$ be ordered and let $V$ be any finite-dimensional inner product space over $K$. Pick an orthogonal basis $u_1, \cdots, u_n$ (Theorem 4); let $(u_i,u_i) = b_i$. We can arrange the notation so that $b_1, \cdots, b_r$ are positive, $b_{r+1}, \cdots, b_s$ are negative, and $b_{s+1}, \cdots, b_n$ are 0. Let $W_1$, $W_2$ and $W_3$ be the subspaces of $V$ spanned by these three batches of $b$'s respectively. It is routine to see that the inner product restricted to $W_1$ is positive definite, to $W_2$ is negative definite, and to $W_3$ is identically 0. This proves the first part of Theorem 8.

Theorem 8. *Let $V$ be a finite-dimensional inner product space over an ordered field $K$. Then $V$ can be written as an orthogonal direct sum $V = W_1 \oplus W_2 \oplus W_3$, where the inner product is positive definite on $W_1$, negative definite on $W_2$, and identically 0 on $W_3$. In such a decomposition the dimensions of $W_1$, $W_2$, and $W_3$ are uniquely determined.*

Proof. We have already demonstrated the existence of the decomposition. The uniqueness of the dimension of $W_3$ is immediate; in fact $W_3$ itself is unique, since it is obviously the radical of the form (compare Ex. 1 in §1-3).

The uniqueness of the dimensions of $W_1$ and $W_2$ can be handled expeditiously by inventing a suitable linear transformation. Let $X_1 \oplus X_2 \oplus X_3$ be a second orthogonal decomposition of $V$ into portions that are (like the $W$'s) positive definite, negative definite, and with inner product 0. Let $T$ denote the projection of $V$ onto $W_1$; the kernel of $T$ is $W_2 \oplus W_3$. Consider $T$ restricted to $X_1$; this is a linear transformation from $X_1$ into $W_1$ with kernel $(W_2 \oplus W_3) \cap X_1$. But the latter is 0; for if $x$ is a non-zero vector lying in both $X_1$ and $W_2 \oplus W_3$ then $(x,x) > 0$ since $X_1$ is positive definite, and $(x,x) \leqq 0$ since $W_2 \oplus W_3$ is negative semi-definite. Now the existence of a one-to-one linear transformation from $X_1$ into $W_1$ shows $\dim (X_1) \leqq \dim (W_1)$. We get the reverse inequality symmetrically, and hence equality. The equality of the dimensions of $W_2$ and $X_2$ is entirely analogous, and Theorem 8 is proved.

*Remark.* This proof of the inertia theorem (the uniqueness of the dimensions) is valid verbatim in the infinite-dimensional case. But we cannot, in the uncountable case, assert the *existence* of such a decomposition. (See [32] p. 523, for an example that Savage attributes to Mackey.) For the countable case, see §1-15.

If we strengthen the hypothesis on $K$ by assuming that every positive element is a square, in the above discussion we can convert all positive diagonal elements to 1's and all negative diagonal elements to $-1$'s. Hence:

Theorem 9. *Let $K$ be an ordered field in which every positive element is a* *square and let $V$ be a finite-dimensional inner product space over $K$. Then we can find an orthogonal basis $u_1, \cdots, u_n$ with each $(u_i,u_i) = 1, -1$, or $0$. The number of terms that are $1$, $-1$, or $0$ does not depend on the choice of basis.*

We leave it to the reader to state the matrix versions of Theorems 8 and 9.

### EXERCISE

Let $K$ be an ordered field in which every positive element is a square. Let $V$ be a non-isotropic inner product space over $K$. Prove: $V$ is either positive definite or negative definite.

## 1-5 The Discriminant

Return to Equation (4) in §1-1, and take determinants. The result (after we note $|P'| = |P|$):

$$|B| = |P|^2|A|$$

Thus the determinant of the matrix of inner products changes when the basis is changed. But if we grant that multiplication by a non-zero square in any case has to be absorbed (as it does in the diagonal entries after diagonalization), then we may speak of $|A|$ as an invariant of the form, and we call it the *discriminant.*

If one prefers, the ambiguity caused by multiplication by non-zero squares may be treated formally. Let $K^*$ denote the multiplicative group of non-zero elements in the underlying field $K$, and write $(K^*)^2$ for its subgroup of squares. Then (for non-singular forms) the discriminant, and also the diagonal entries obtained after diagonalization, may be regarded as lying in the group $(K^*)/(K^*)^2$. Note that the discriminant is the product of the diagonal entries (in $(K^*)/(K^*)^2$, or up to a square).

The discriminant offers a way of computing the result of diagonalization, which is reasonably effective for hand work (at least up to $3 \times 3$ matrices!). We illustrate with two examples.

For the matrix

$$\begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$



the first basis vector could start the diagonalization, giving us the entry 2. Since the discriminant is 5, the remaining entry is $5/2$, or 10 if we prefer to avoid fractions.

For the matrix

$$\begin{pmatrix} 2 & 1 & -1 \\ 1 & 3 & 0 \\ -1 & 0 & -2 \end{pmatrix},$$

we use the already effected diagonalization of the subspace spanned by the first two basis vectors. To get the third entry, we compute the discriminant, $-13$, and then know that $-13/5$, or $-65$ if we prefer, is the required third number. (We are assuming characteristic $\neq 5$.) To summarize: the $3 \times 3$ matrix above is congruent to the diagonal matrix with entries $2, 10, -65$.

A systematic statement of the method is as follows: given an $n \times n$ matrix $A$, call the determinant of the $i \times i$ upper-left corner the $i$th *leading minor* and denote it by $d_i$. Assume that $A$ is symmetric and the $d_i$'s are non-zero. (What happens if we encounter zeros is illustrated in Exs. 3, 4, and 5 of §1-8.) Then $d_1, d_2/d_1, d_3/d_2, \cdots, d_n/d_{n-1}$ are diagonal entries in a diagonalization of $A$.

Now let $A$ be a symmetric matrix over an ordered field $K$. We say that $A$ is positive definite if the form it determines is positive definite. This certainly implies that the discriminant is positive, i.e., the determinant of $A$ is positive. The form restricted to any subspace is likewise positive definite. This is in particular true for the subspace spanned by the first $i$ basis vectors. Thus we see that positive definiteness of $A$ implies that all the leading minors are positive. The converse is also true: if in the notation above we have $d_1, \cdots, d_n$ all positive, then the diagonal entries $d_1, d_2/d_1, \cdots, d_n/d_{n-1}$ are positive. We state this as a formal theorem.

Theorem 10. *A symmetric matrix over an ordered field is positive definite if and only if all the leading minors are positive.*

### EXERCISES

1. A symmetric matrix over an ordered field is called positive semi-definite if the form it determines is positive semi-definite. Prove: the leading minors in a positive semi-definite matrix are $\geqq 0$. Show that the converse is false. (*Hint:* try $\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$.)