

7. Prove that if  $T \in \mathcal{L}(V)$  and  $j$  is a positive integer such that  $j \leq \dim V$ , then  $T$  has an invariant subspace whose dimension equals  $j - 1$  or  $j$ .
8. Prove that there does not exist an operator  $T \in \mathcal{L}(\mathbf{R}^7)$  such that  $T^2 + T + I$  is nilpotent.
9. Give an example of an operator  $T \in \mathcal{L}(\mathbf{C}^7)$  such that  $T^2 + T + I$  is nilpotent.
10. Suppose  $V$  is a real vector space and  $T \in \mathcal{L}(V)$ . Suppose  $\alpha, \beta \in \mathbf{R}$  are such that  $\alpha^2 < 4\beta$ . Prove that

$$\text{null}(T^2 + \alpha T + \beta I)^k$$

has even dimension for every positive integer  $k$ .

11. Suppose  $V$  is a real vector space and  $T \in \mathcal{L}(V)$ . Suppose  $\alpha, \beta \in \mathbf{R}$  are such that  $\alpha^2 < 4\beta$  and  $T^2 + \alpha T + \beta I$  is nilpotent. Prove that  $\dim V$  is even and

$$(T^2 + \alpha T + \beta I)^{\dim V/2} = 0.$$

12. Prove that if  $T \in \mathcal{L}(\mathbf{R}^3)$  and 5, 7 are eigenvalues of  $T$ , then  $T$  has no eigenpairs.
13. Suppose  $V$  is a real vector space with  $\dim V = n$  and  $T \in \mathcal{L}(V)$  is such that

$$\text{null } T^{n-2} \neq \text{null } T^{n-1}.$$

Prove that  $T$  has at most two distinct eigenvalues and that  $T$  has no eigenpairs.

14. Suppose  $V$  is a vector space with dimension 2 and  $T \in \mathcal{L}(V)$ . Prove that if

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

is the matrix of  $T$  with respect to some basis of  $V$ , then the characteristic polynomial of  $T$  equals  $(z - a)(z - d) - bc$ .

15. Suppose  $V$  is a real inner-product space and  $S \in \mathcal{L}(V)$  is an isometry. Prove that if  $(\alpha, \beta)$  is an eigenpair of  $S$ , then  $\beta = 1$ .

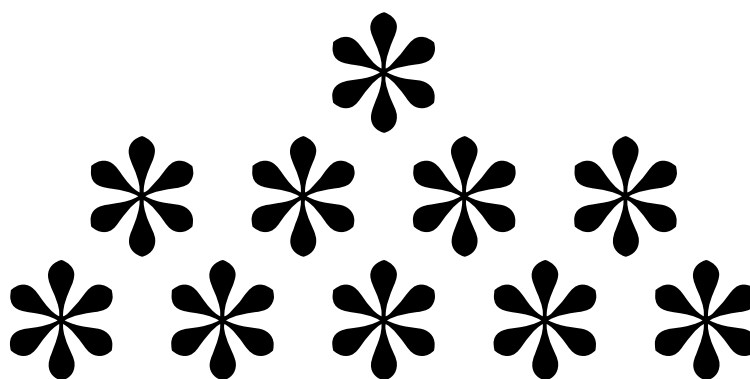
*You do not need to find the eigenvalues of  $T$  to do this exercise. As usual unless otherwise specified, here  $V$  may be a real or complex vector space.*

## CHAPTER 10

# *Trace and Determinant*

Throughout this book our emphasis has been on linear maps and operators rather than on matrices. In this chapter we pay more attention to matrices as we define and discuss traces and determinants. Determinants appear only at the end of this book because we replaced their usual applications in linear algebra (the definition of the characteristic polynomial and the proof that operators on complex vector spaces have eigenvalues) with more natural techniques. The book concludes with an explanation of the important role played by determinants in the theory of volume and integration.

Recall that  $F$  denotes  $\mathbf{R}$  or  $\mathbf{C}$ .  
Also,  $V$  is a finite-dimensional, nonzero vector space over  $F$ .



## Change of Basis

The matrix of an operator  $T \in \mathcal{L}(V)$  depends on a choice of basis of  $V$ . Two different bases of  $V$  may give different matrices of  $T$ . In this section we will learn how these matrices are related. This information will help us find formulas for the trace and determinant of  $T$  later in this chapter.

With respect to any basis of  $V$ , the identity operator  $I \in \mathcal{L}(V)$  has a diagonal matrix

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

This matrix is called the **identity matrix** and is denoted  $I$ . Note that we use the symbol  $I$  to denote the identity operator (on all vector spaces) and the identity matrix (of all possible sizes). You should always be able to tell from the context which particular meaning of  $I$  is intended. For example, consider the equation

$$\mathcal{M}(I) = I;$$

on the left side  $I$  denotes the identity operator and on the right side  $I$  denotes the identity matrix.

If  $A$  is a square matrix (with entries in  $\mathbf{F}$ , as usual) with the same size as  $I$ , then  $AI = IA = A$ , as you should verify. A square matrix  $A$  is called **invertible** if there is a square matrix  $B$  of the same size such that  $AB = BA = I$ , and we call  $B$  an **inverse** of  $A$ . To prove that  $A$  has at most one inverse, suppose  $B$  and  $B'$  are inverses of  $A$ . Then

$$B = BI = B(AB') = (BA)B' = IB' = B',$$

and hence  $B = B'$ , as desired. Because an inverse is unique, we can use the notation  $A^{-1}$  to denote the inverse of  $A$  (if  $A$  is invertible). In other words, if  $A$  is invertible, then  $A^{-1}$  is the unique matrix of the same size such that  $AA^{-1} = A^{-1}A = I$ .

Recall that when discussing linear maps from one vector space to another in Chapter 3, we defined the matrix of a linear map with respect to two bases—one basis for the first vector space and another basis for the second vector space. When we study operators, which are linear maps from a vector space to itself, we almost always use the same basis

Some mathematicians use the terms **nonsingular**, which means the same as invertible, and **singular**, which means the same as noninvertible.

for both vector spaces (after all, the two vector spaces in question are equal). Thus we usually refer to the matrix of an operator with respect to a basis, meaning that we are using one basis in two capacities. The next proposition is one of the rare cases where we need to use two different bases even though we have an operator from a vector space to itself.

Let's review how matrix multiplication interacts with multiplication of linear maps. Suppose that along with  $V$  we have two other finite-dimensional vector spaces, say  $U$  and  $W$ . Let  $(u_1, \dots, u_p)$  be a basis of  $U$ , let  $(v_1, \dots, v_n)$  be a basis of  $V$ , and let  $(w_1, \dots, w_m)$  be a basis of  $W$ . If  $T \in \mathcal{L}(U, V)$  and  $S \in \mathcal{L}(V, W)$ , then  $ST \in \mathcal{L}(U, W)$  and

$$\mathbf{10.1} \quad \mathcal{M}(ST, (u_1, \dots, u_p), (w_1, \dots, w_m)) = \mathcal{M}(S, (v_1, \dots, v_n), (w_1, \dots, w_m)) \mathcal{M}(T, (u_1, \dots, u_p), (v_1, \dots, v_n)).$$

The equation above holds because we defined matrix multiplication to make it true—see 3.11 and the material following it.

The following proposition deals with the matrix of the identity operator when we use two different bases. Note that the  $k^{\text{th}}$  column of  $\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$  consists of the scalars needed to write  $u_k$  as a linear combination of the  $v$ 's. As an example of the proposition below, consider the bases  $((4, 2), (5, 3))$  and  $((1, 0), (0, 1))$  of  $\mathbb{F}^2$ . Obviously

$$\mathcal{M}(I, ((4, 2), (5, 3)), ((1, 0), (0, 1))) = \begin{bmatrix} 4 & 5 \\ 2 & 3 \end{bmatrix}.$$

The inverse of the matrix above is  $\begin{bmatrix} 3/2 & -5/2 \\ -1 & 2 \end{bmatrix}$ , as you should verify. Thus the proposition below implies that

$$\mathcal{M}(I, ((1, 0), (0, 1)), ((4, 2), (5, 3))) = \begin{bmatrix} 3/2 & -5/2 \\ -1 & 2 \end{bmatrix}.$$

**10.2 Proposition:** *If  $(u_1, \dots, u_n)$  and  $(v_1, \dots, v_n)$  are bases of  $V$ , then  $\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$  is invertible and*

$$\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))^{-1} = \mathcal{M}(I, (v_1, \dots, v_n), (u_1, \dots, u_n)).$$

**PROOF:** In 10.1, replace  $U$  and  $W$  with  $V$ , replace  $w_j$  with  $u_j$ , and replace  $S$  and  $T$  with  $I$ , getting

$$I = \mathcal{M}(I, (\nu_1, \dots, \nu_n), (u_1, \dots, u_n)) \mathcal{M}(I, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)).$$

Now interchange the roles of the  $u$ 's and  $\nu$ 's, getting

$$I = \mathcal{M}(I, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)) \mathcal{M}(I, (\nu_1, \dots, \nu_n), (u_1, \dots, u_n)).$$

These two equations give the desired result. ■

Now we can see how the matrix of  $T$  changes when we change bases.

**10.3 Theorem:** Suppose  $T \in \mathcal{L}(V)$ . Let  $(u_1, \dots, u_n)$  and  $(\nu_1, \dots, \nu_n)$  be bases of  $V$ . Let  $A = \mathcal{M}(I, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n))$ . Then

$$\mathbf{10.4} \quad \mathcal{M}(T, (u_1, \dots, u_n)) = A^{-1} \mathcal{M}(T, (\nu_1, \dots, \nu_n)) A.$$

PROOF: In 10.1, replace  $U$  and  $W$  with  $V$ , replace  $w_j$  with  $\nu_j$ , replace  $T$  with  $I$ , and replace  $S$  with  $T$ , getting

$$\mathbf{10.5} \quad \mathcal{M}(T, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)) = \mathcal{M}(T, (\nu_1, \dots, \nu_n)) A.$$

Again use 10.1, this time replacing  $U$  and  $W$  with  $V$ , replacing  $w_j$  with  $u_j$ , and replacing  $S$  with  $I$ , getting

$$\mathcal{M}(T, (u_1, \dots, u_n)) = A^{-1} \mathcal{M}(T, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)),$$

where we have used 10.2. Substituting 10.5 into the equation above gives 10.4, completing the proof. ■

## Trace

Let's examine the characteristic polynomial more closely than we did in the last two chapters. If  $V$  is an  $n$ -dimensional complex vector space and  $T \in \mathcal{L}(V)$ , then the characteristic polynomial of  $T$  equals

$$(z - \lambda_1) \dots (z - \lambda_n),$$

where  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $T$ , repeated according to multiplicity. Expanding the polynomial above, we can write the characteristic polynomial of  $T$  in the form

$$\mathbf{10.6} \quad z^n - (\lambda_1 + \dots + \lambda_n)z^{n-1} + \dots + (-1)^n(\lambda_1 \dots \lambda_n).$$

If  $V$  is an  $n$ -dimensional real vector space and  $T \in \mathcal{L}(V)$ , then the characteristic polynomial of  $T$  equals

$$(x - \lambda_1) \dots (x - \lambda_m)(x^2 + \alpha_1 x + \beta_1) \dots (x^2 + \alpha_M x + \beta_M),$$

where  $\lambda_1, \dots, \lambda_m$  are the eigenvalues of  $T$  and  $(\alpha_1, \beta_1), \dots, (\alpha_M, \beta_M)$  are the eigenpairs of  $T$ , each repeated according to multiplicity. Expanding the polynomial above, we can write the characteristic polynomial of  $T$  in the form

$$\begin{aligned} 10.7 \quad x^n - (\lambda_1 + \dots + \lambda_m - \alpha_1 - \dots - \alpha_M)x^{n-1} + \dots \\ + (-1)^m(\lambda_1 \dots \lambda_m \beta_1 \dots \beta_M). \end{aligned}$$

In this section we will study the coefficient of  $z^{n-1}$  (usually denoted  $x^{n-1}$  when we are dealing with a real vector space) in the characteristic polynomial. In the next section we will study the constant term in the characteristic polynomial.

For  $T \in \mathcal{L}(V)$ , the negative of the coefficient of  $z^{n-1}$  (or  $x^{n-1}$  for real vector spaces) in the characteristic polynomial of  $T$  is called the **trace** of  $T$ , denoted  $\text{trace } T$ . If  $V$  is a complex vector space, then 10.6 shows that  $\text{trace } T$  equals the sum of the eigenvalues of  $T$ , counting multiplicity. If  $V$  is a real vector space, then 10.7 shows that  $\text{trace } T$  equals the sum of the eigenvalues of  $T$  minus the sum of the first coordinates of the eigenpairs of  $T$ , each repeated according to multiplicity.

For example, suppose  $T \in \mathcal{L}(\mathbb{C}^3)$  is the operator whose matrix is

$$10.8 \quad \begin{bmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{bmatrix}.$$

Then the eigenvalues of  $T$  are  $1$ ,  $2 + 3i$ , and  $2 - 3i$ , each with multiplicity 1, as you can verify. Computing the sum of the eigenvalues, we have  $\text{trace } T = 1 + (2 + 3i) + (2 - 3i)$ ; in other words,  $\text{trace } T = 5$ .

As another example, suppose  $T \in \mathcal{L}(\mathbb{R}^3)$  is the operator whose matrix is also given by 10.8 (note that in the previous paragraph we were working on a complex vector space; now we are working on a real vector space). Then  $1$  is the only eigenvalue of  $T$  (it has multiplicity 1) and  $(-4, 13)$  is the only eigenpair of  $T$  (it has multiplicity 1), as you should have verified in the last chapter (see page 205). Computing the sum of the eigenvalues minus the sum of the first coordinates of the eigenpairs, we have  $\text{trace } T = 1 - (-4)$ ; in other words,  $\text{trace } T = 5$ .

Here  $m$  or  $M$  might equal 0.

Recall that a pair  $(\alpha, \beta)$  of real numbers is an eigenpair of  $T$  if  $\alpha^2 < 4\beta$  and  $T^2 + \alpha T + \beta I$  is not injective.

Note that  $\text{trace } T$  depends only on  $T$  and not on a basis of  $V$  because the characteristic polynomial of  $T$  does not depend on a choice of basis.

The reason that the operators in the two previous examples have the same trace will become clear after we find a formula (valid on both complex and real vector spaces) for computing the trace of an operator from its matrix.

Most of the rest of this section is devoted to discovering how to calculate trace  $T$  from the matrix of  $T$  (with respect to an arbitrary basis). Let's start with the easiest situation. Suppose  $V$  is a complex vector space,  $T \in \mathcal{L}(V)$ , and we choose a basis of  $V$  with respect to which  $T$  has an upper-triangular matrix  $A$ . Then the eigenvalues of  $T$  are precisely the diagonal entries of  $A$ , repeated according to multiplicity (see 8.10). Thus trace  $T$  equals the sum of the diagonal entries of  $A$ . The same formula works for the operator  $T \in \mathcal{L}(\mathbb{F}^3)$  whose matrix is given by 10.8 and whose trace equals 5. Could such a simple formula be true in general?

We begin our investigation by considering  $T \in \mathcal{L}(V)$  where  $V$  is a real vector space. Choose a basis of  $V$  with respect to which  $T$  has a block upper-triangular matrix  $\mathcal{M}(T)$ , where each block on the diagonal is a 1-by-1 matrix containing an eigenvalue of  $T$  or a 2-by-2 block with no eigenvalues (see 9.4 and 9.9). Each entry in a 1-by-1 block on the diagonal of  $\mathcal{M}(T)$  is an eigenvalue of  $T$  and thus makes a contribution to trace  $T$ . If  $\mathcal{M}(T)$  has any 2-by-2 blocks on the diagonal, consider a typical one

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

The characteristic polynomial of this 2-by-2 matrix is  $(x-a)(x-d)-bc$ , which equals

$$x^2 - (a+d)x + (ad-bc).$$

*You should carefully review 9.9 to understand the relationship between eigenpairs and characteristic polynomials of 2-by-2 blocks.*

Thus  $(-a-d, ad-bc)$  is an eigenpair of  $T$ . The negative of the first coordinate of this eigenpair, namely,  $a+d$ , is the contribution of this block to trace  $T$ . Note that  $a+d$  is the sum of the entries on the diagonal of this block. Thus for any basis of  $V$  with respect to which the matrix of  $T$  has the block upper-triangular form required by 9.4 and 9.9, trace  $T$  equals the sum of the entries on the diagonal.

At this point you should suspect that trace  $T$  equals the sum of the diagonal entries of the matrix of  $T$  with respect to an arbitrary basis. Remarkably, this turns out to be true. To prove it, let's define the **trace** of a square matrix  $A$ , denoted  $\text{trace } A$ , to be the sum of the diagonal entries. With this notation, we want to prove that

$\text{trace } T = \text{trace } \mathcal{M}(T, (\nu_1, \dots, \nu_n))$ , where  $(\nu_1, \dots, \nu_n)$  is an arbitrary basis of  $V$ . We already know this is true if  $(\nu_1, \dots, \nu_n)$  is a basis with respect to which  $T$  has an upper-triangular matrix (if  $V$  is complex) or an appropriate block upper-triangular matrix (if  $V$  is real). We will need the following proposition to prove our trace formula for an arbitrary basis.

**10.9 Proposition:** *If  $A$  and  $B$  are square matrices of the same size, then*

$$\text{trace}(AB) = \text{trace}(BA).$$

PROOF: Suppose

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix}, \quad B = \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,n} \end{bmatrix}.$$

The  $j^{\text{th}}$  term on the diagonal of  $AB$  equals

$$\sum_{k=1}^n a_{j,k} b_{k,j}.$$

Thus

$$\begin{aligned} \text{trace}(AB) &= \sum_{j=1}^n \sum_{k=1}^n a_{j,k} b_{k,j} \\ &= \sum_{k=1}^n \sum_{j=1}^n b_{k,j} a_{j,k} \\ &= \sum_{k=1}^n k^{\text{th}} \text{ term on the diagonal of } BA \\ &= \text{trace}(BA), \end{aligned}$$

as desired. ■

Now we can prove that the sum of the diagonal entries of the matrix of an operator is independent of the basis with respect to which the matrix is computed.

**10.10 Corollary:** *Suppose  $T \in \mathcal{L}(V)$ . If  $(u_1, \dots, u_n)$  and  $(\nu_1, \dots, \nu_n)$  are bases of  $V$ , then*

$$\text{trace } \mathcal{M}(T, (u_1, \dots, u_n)) = \text{trace } \mathcal{M}(T, (\nu_1, \dots, \nu_n)).$$

PROOF: Suppose  $(u_1, \dots, u_n)$  and  $(v_1, \dots, v_n)$  are bases of  $V$ . Let  $A = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$ . Then

*The third equality here depends on the associative property of matrix multiplication.*

$$\begin{aligned} \text{trace } \mathcal{M}(T, (u_1, \dots, u_n)) &= \text{trace} \left( A^{-1} (\mathcal{M}(T, (v_1, \dots, v_n)) A) \right) \\ &= \text{trace} \left( (\mathcal{M}(T, (v_1, \dots, v_n)) A) A^{-1} \right) \\ &= \text{trace } \mathcal{M}(T, (v_1, \dots, v_n)), \end{aligned}$$

where the first equality follows from 10.3 and the second equality follows from 10.9. The third equality completes the proof. ■

The theorem below states that the trace of an operator equals the sum of the diagonal entries of the matrix of the operator. This theorem does not specify a basis because, by the corollary above, the sum of the diagonal entries of the matrix of an operator is the same for every choice of basis.

**10.11 Theorem:** *If  $T \in \mathcal{L}(V)$ , then  $\text{trace } T = \text{trace } \mathcal{M}(T)$ .*

PROOF: Let  $T \in \mathcal{L}(V)$ . As noted above,  $\text{trace } \mathcal{M}(T)$  is independent of which basis of  $V$  we choose (by 10.10). Thus to show that

$$\text{trace } T = \text{trace } \mathcal{M}(T)$$

for every basis of  $V$ , we need only show that the equation above holds for some basis of  $V$ . We already did this (on page 218), choosing a basis of  $V$  with respect to which  $\mathcal{M}(T)$  is an upper-triangular matrix (if  $V$  is a complex vector space) or an appropriate block upper-triangular matrix (if  $V$  is a real vector space). ■

If we know the matrix of an operator on a complex vector space, the theorem above allows us to find the sum of all the eigenvalues without finding any of the eigenvalues. For example, consider the operator on  $\mathbb{C}^5$  whose matrix is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

No one knows an exact formula for any of the eigenvalues of this operator. However, we do know that the sum of the eigenvalues equals 0 because the sum of the diagonal entries of the matrix above equals 0.

The theorem above also allows us easily to prove some useful properties about traces of operators by shifting to the language of traces of matrices, where certain properties have already been proved or are obvious. We carry out this procedure in the next corollary.

**10.12 Corollary:** *If  $S, T \in \mathcal{L}(V)$ , then*

$$\text{trace}(ST) = \text{trace}(TS) \quad \text{and} \quad \text{trace}(S + T) = \text{trace } S + \text{trace } T.$$

PROOF: Suppose  $S, T \in \mathcal{L}(V)$ . Choose any basis of  $V$ . Then

$$\begin{aligned} \text{trace}(ST) &= \text{trace } \mathcal{M}(ST) \\ &= \text{trace}(\mathcal{M}(S)\mathcal{M}(T)) \\ &= \text{trace}(\mathcal{M}(T)\mathcal{M}(S)) \\ &= \text{trace } \mathcal{M}(TS) \\ &= \text{trace}(TS), \end{aligned}$$

where the first and last equalities come from 10.11 and the middle equality comes from 10.9. This completes the proof of the first assertion in the corollary.

To prove the second assertion in the corollary, note that

$$\begin{aligned} \text{trace}(S + T) &= \text{trace } \mathcal{M}(S + T) \\ &= \text{trace}(\mathcal{M}(S) + \mathcal{M}(T)) \\ &= \text{trace } \mathcal{M}(S) + \text{trace } \mathcal{M}(T) \\ &= \text{trace } S + \text{trace } T, \end{aligned}$$

where again the first and last equalities come from 10.11; the third equality is obvious from the definition of the trace of a matrix. This completes the proof of the second assertion in the corollary. ■

The techniques we have developed have the following curious corollary. The generalization of this result to infinite-dimensional vector spaces has important consequences in quantum theory.

*The statement of this corollary does not involve traces, though the short proof uses traces. Whenever something like this happens in mathematics, we can be sure that a good definition lurks in the background.*

**10.13 Corollary:** *There do not exist operators  $S, T \in \mathcal{L}(V)$  such that  $ST - TS = I$ .*

PROOF: Suppose  $S, T \in \mathcal{L}(V)$ . Then

$$\begin{aligned}\text{trace}(ST - TS) &= \text{trace}(ST) - \text{trace}(TS) \\ &= 0,\end{aligned}$$

where the second equality comes from 10.12. Clearly the trace of  $I$  equals  $\dim V$ , which is not 0. Because  $ST - TS$  and  $I$  have different traces, they cannot be equal. ■

## Determinant of an Operator

*Note that  $\det T$  depends only on  $T$  and not on a basis of  $V$  because the characteristic polynomial of  $T$  does not depend on a choice of basis.*

For  $T \in \mathcal{L}(V)$ , we define the **determinant** of  $T$ , denoted  $\det T$ , to be  $(-1)^{\dim V}$  times the constant term in the characteristic polynomial of  $T$ . The motivation for the factor  $(-1)^{\dim V}$  in this definition comes from 10.6.

If  $V$  is a complex vector space, then  $\det T$  equals the product of the eigenvalues of  $T$ , counting multiplicity; this follows immediately from 10.6. Recall that if  $V$  is a complex vector space, then there is a basis of  $V$  with respect to which  $T$  has an upper-triangular matrix (see 5.13); thus  $\det T$  equals the product of the diagonal entries of this matrix (see 8.10).

If  $V$  is a real vector space, then  $\det T$  equals the product of the eigenvalues of  $T$  times the product of the second coordinates of the eigenpairs of  $T$ , each repeated according to multiplicity—this follows from 10.7 and the observation that  $m = \dim V - 2M$  (in the notation of 10.7), and hence  $(-1)^m = (-1)^{\dim V}$ .

For example, suppose  $T \in \mathcal{L}(\mathbb{C}^3)$  is the operator whose matrix is given by 10.8. As we noted in the last section, the eigenvalues of  $T$  are 1,  $2 + 3i$ , and  $2 - 3i$ , each with multiplicity 1. Computing the product of the eigenvalues, we have  $\det T = (1)(2 + 3i)(2 - 3i)$ ; in other words,  $\det T = 13$ .

As another example, suppose  $T \in \mathcal{L}(\mathbb{R}^3)$  is the operator whose matrix is also given by 10.8 (note that in the previous paragraph we were working on a complex vector space; now we are working on a real vector space). Then, as we noted earlier, 1 is the only eigenvalue of  $T$  (it

has multiplicity 1) and  $(-4, 13)$  is the only eigenpair of  $T$  (it has multiplicity 1). Computing the product of the eigenvalues times the product of the second coordinates of the eigenpairs, we have  $\det T = (1)(13)$ ; in other words,  $\det T = 13$ .

The reason that the operators in the two previous examples have the same determinant will become clear after we find a formula (valid on both complex and real vector spaces) for computing the determinant of an operator from its matrix.

In this section, we will prove some simple but important properties of determinants. In the next section, we will discover how to calculate  $\det T$  from the matrix of  $T$  (with respect to an arbitrary basis). We begin with a crucial result that has an easy proof with our approach.

**10.14 Proposition:** *An operator is invertible if and only if its determinant is nonzero.*

PROOF: First suppose  $V$  is a complex vector space and  $T \in \mathcal{L}(V)$ . The operator  $T$  is invertible if and only if 0 is not an eigenvalue of  $T$ . Clearly this happens if and only if the product of the eigenvalues of  $T$  is not 0. Thus  $T$  is invertible if and only if  $\det T \neq 0$ , as desired.

Now suppose  $V$  is a real vector space and  $T \in \mathcal{L}(V)$ . Again,  $T$  is invertible if and only if 0 is not an eigenvalue of  $T$ . Using the notation of 10.7, we have

$$\det T = \lambda_1 \dots \lambda_m \beta_1 \dots \beta_M,$$

where the  $\lambda$ 's are the eigenvalues of  $T$  and the  $\beta$ 's are the second coordinates of the eigenpairs of  $T$ , each repeated according to multiplicity. For each eigenpair  $(\alpha_j, \beta_j)$ , we have  $\alpha_j^2 < 4\beta_j$ . In particular, each  $\beta_j$  is positive. This implies (see 10.15) that  $\lambda_1 \dots \lambda_m \neq 0$  if and only if  $\det T \neq 0$ . Thus  $T$  is invertible if and only if  $\det T \neq 0$ , as desired. ■

If  $T \in \mathcal{L}(V)$  and  $\lambda, z \in \mathbb{F}$ , then  $\lambda$  is an eigenvalue of  $T$  if and only if  $z - \lambda$  is an eigenvalue of  $zI - T$ . This follows from

$$-(T - \lambda I) = (zI - T) - (z - \lambda)I.$$

Raising both sides of this equation to the  $\dim V$  power and then taking null spaces of both sides shows that the multiplicity of  $\lambda$  as an eigenvalue of  $T$  equals the multiplicity of  $z - \lambda$  as an eigenvalue of  $zI - T$ .

The next lemma gives the analogous result for eigenpairs. We will use this lemma to show that the characteristic polynomial can be expressed as a certain determinant.

*Real vector spaces are harder to deal with than complex vector spaces. The first time you read this chapter, you may want to concentrate on the basic ideas by considering only complex vector spaces and ignoring the special procedures needed to deal with real vector spaces.*

**10.16 Lemma:** *Suppose  $V$  is a real vector space,  $T \in \mathcal{L}(V)$ , and  $\alpha, \beta, x \in \mathbf{R}$  with  $\alpha^2 < 4\beta$ . Then  $(\alpha, \beta)$  is an eigenpair of  $T$  if and only if  $(-2x - \alpha, x^2 + \alpha x + \beta)$  is an eigenpair of  $xI - T$ . Furthermore, these eigenpairs have the same multiplicities.*

**PROOF:** First we need to check that  $(-2x - \alpha, x^2 + \alpha x + \beta)$  satisfies the inequality required of an eigenpair. We have

$$\begin{aligned} (-2x - \alpha)^2 &= 4x^2 + 4\alpha x + \alpha^2 \\ &< 4x^2 + 4\alpha x + 4\beta \\ &= 4(x^2 + \alpha x + \beta). \end{aligned}$$

Thus  $(-2x - \alpha, x^2 + \alpha x + \beta)$  satisfies the required inequality.

Now

$$T^2 + \alpha T + \beta I = (xI - T)^2 - (2x + \alpha)(xI - T) + (x^2 + \alpha x + \beta)I,$$

as you should verify. Thus  $(\alpha, \beta)$  is an eigenpair of  $T$  if and only if  $(-2x - \alpha, x^2 + \alpha x + \beta)$  is an eigenpair of  $xI - T$ . Furthermore, raising both sides of the equation above to the  $\dim V$  power and then taking null spaces of both sides shows that the multiplicities are equal. ■

Most textbooks take the theorem below as the definition of the characteristic polynomial. Texts using that approach must spend considerably more time developing the theory of determinants before they get to interesting linear algebra.

**10.17 Theorem:** *Suppose  $T \in \mathcal{L}(V)$ . Then the characteristic polynomial of  $T$  equals  $\det(zI - T)$ .*

**PROOF:** First suppose  $V$  is a complex vector space. Let  $\lambda_1, \dots, \lambda_n$  denote the eigenvalues of  $T$ , repeated according to multiplicity. Thus for  $z \in \mathbf{C}$ , the eigenvalues of  $zI - T$  are  $z - \lambda_1, \dots, z - \lambda_n$ , repeated according to multiplicity. The determinant of  $zI - T$  is the product of these eigenvalues. In other words,

$$\det(zI - T) = (z - \lambda_1) \dots (z - \lambda_n).$$

The right side of the equation above is, by definition, the characteristic polynomial of  $T$ , completing the proof when  $V$  is a complex vector space.

Now suppose  $V$  is a real vector space. Let  $\lambda_1, \dots, \lambda_m$  denote the eigenvalues of  $T$  and let  $(\alpha_1, \beta_1), \dots, (\alpha_M, \beta_M)$  denote the eigenpairs of  $T$ , each repeated according to multiplicity. Thus for  $x \in \mathbf{R}$ , the eigenvalues of  $xI - T$  are  $x - \lambda_1, \dots, x - \lambda_m$  and, by 10.16, the eigenpairs of  $xI - T$  are

$$(-2x - \alpha_1, x^2 + \alpha_1 x + \beta_1), \dots, (-2x - \alpha_M, x^2 + \alpha_M x + \beta_M),$$

each repeated according to multiplicity. Hence

$$\det(xI - T) = (x - \lambda_1) \dots (x - \lambda_m) (x^2 + \alpha_1 x + \beta_1) \dots (x^2 + \alpha_M x + \beta_M).$$

The right side of the equation above is, by definition, the characteristic polynomial of  $T$ , completing the proof when  $V$  is a real vector space. ■

## *Determinant of a Matrix*

Most of this section is devoted to discovering how to calculate  $\det T$  from the matrix of  $T$  (with respect to an arbitrary basis). Let's start with the easiest situation. Suppose  $V$  is a complex vector space,  $T \in \mathcal{L}(V)$ , and we choose a basis of  $V$  with respect to which  $T$  has an upper-triangular matrix. Then, as we noted in the last section,  $\det T$  equals the product of the diagonal entries of this matrix. Could such a simple formula be true in general?

Unfortunately the determinant is more complicated than the trace. In particular,  $\det T$  need not equal the product of the diagonal entries of  $\mathcal{M}(T)$  with respect to an arbitrary basis. For example, the operator on  $\mathbf{F}^3$  whose matrix equals 10.8 has determinant 13, as we saw in the last section. However, the product of the diagonal entries of that matrix equals 0.

For each square matrix  $A$ , we want to define the determinant of  $A$ , denoted  $\det A$ , in such a way that  $\det T = \det \mathcal{M}(T)$  regardless of which basis is used to compute  $\mathcal{M}(T)$ . We begin our search for the correct definition of the determinant of a matrix by calculating the determinants of some special operators.

Let  $c_1, \dots, c_n \in \mathbf{F}$  be nonzero scalars and let  $(v_1, \dots, v_n)$  be a basis of  $V$ . Consider the operator  $T \in \mathcal{L}(V)$  such that  $\mathcal{M}(T, (v_1, \dots, v_n))$  equals

$$10.18 \quad \begin{bmatrix} 0 & & & & c_n \\ c_1 & 0 & & & \\ & c_2 & 0 & & \\ & & \ddots & \ddots & \\ & & & c_{n-1} & 0 \end{bmatrix};$$

here all entries of the matrix are 0 except for the upper-right corner and along the line just below the diagonal. Let's find the determinant of  $T$ . Note that

$$(\nu_1, T\nu_1, T^2\nu_1, \dots, T^{n-1}\nu_1) = (\nu_1, c_1\nu_2, c_1c_2\nu_3, \dots, c_1 \dots c_{n-1}\nu_n).$$

Thus  $(\nu_1, T\nu_1, \dots, T^{n-1}\nu_1)$  is linearly independent (the  $c$ 's are all non-zero). Hence if  $p$  is a nonzero polynomial with degree at most  $n-1$ , then  $p(T)\nu_1 \neq 0$ . In other words, the minimal polynomial of  $T$  cannot have degree less than  $n$ . As you should verify,  $T^n\nu_j = c_1 \dots c_n\nu_j$  for each  $j$ , and hence  $T^n = c_1 \dots c_n I$ . Thus  $z^n - c_1 \dots c_n$  is the minimal polynomial of  $T$ . Because  $n = \dim V$ , we see that  $z^n - c_1 \dots c_n$  is also the characteristic polynomial of  $T$ . Multiplying the constant term of this polynomial by  $(-1)^n$ , we get

$$10.19 \quad \det T = (-1)^{n-1} c_1 \dots c_n.$$

If some  $c_j$  equals 0, then clearly  $T$  is not invertible, so  $\det T = 0$  and the same formula holds. Thus in order to have  $\det T = \det \mathcal{M}(T)$ , we will have to make the determinant of 10.18 equal to  $(-1)^{n-1} c_1 \dots c_n$ . However, we do not yet have enough evidence to make a reasonable guess about the proper definition of the determinant of an arbitrary square matrix.

To compute the determinants of a more complicated class of operators, we introduce the notion of permutation. A **permutation** of  $(1, \dots, n)$  is a list  $(m_1, \dots, m_n)$  that contains each of the numbers  $1, \dots, n$  exactly once. The set of all permutations of  $(1, \dots, n)$  is denoted  $\text{perm } n$ . For example,  $(2, 3, \dots, n, 1) \in \text{perm } n$ . You should think of an element of  $\text{perm } n$  as a rearrangement of the first  $n$  integers.

For simplicity we will work with matrices with complex entries (at this stage we are providing only motivation—formal proofs will come later). Let  $c_1, \dots, c_n \in \mathbb{C}$  and let  $(\nu_1, \dots, \nu_n)$  be a basis of  $V$ , which we are assuming is a complex vector space. Consider a permutation  $(p_1, \dots, p_n) \in \text{perm } n$  that can be obtained as follows: break  $(1, \dots, n)$

*Recall that if the minimal polynomial of an operator  $T \in \mathcal{L}(V)$  has degree  $\dim V$ , then the characteristic polynomial of  $T$  equals the minimal polynomial of  $T$ . Computing the minimal polynomial is often an efficient method of finding the characteristic polynomial.*

into lists of consecutive integers and in each list move the first term to the end of that list. For example, taking  $n = 9$ , the permutation

$$\mathbf{10.20} \quad (2, 3, 1, 5, 6, 7, 4, 9, 8)$$

is obtained from  $(1, 2, 3), (4, 5, 6, 7), (8, 9)$  by moving the first term of each of these lists to the end, producing  $(2, 3, 1), (5, 6, 7, 4), (9, 8)$ , and then putting these together to form 10.20. Let  $T \in \mathcal{L}(V)$  be the operator such that

$$\mathbf{10.21} \quad T\mathbf{v}_k = c_k \mathbf{v}_{p_k}$$

for  $k = 1, \dots, n$ . We want to find a formula for  $\det T$ . This generalizes our earlier example because if  $(p_1, \dots, p_n)$  happens to be the permutation  $(2, 3, \dots, n, 1)$ , then the operator  $T$  whose matrix equals 10.18 is the same as the operator  $T$  defined by 10.21.

With respect to the basis  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ , the matrix of the operator  $T$  defined by 10.21 is a block diagonal matrix

$$A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_M \end{bmatrix},$$

where each block is a square matrix of the form 10.18. The eigenvalues of  $T$  equal the union of the eigenvalues of  $A_1, \dots, A_M$  (see Exercise 3 in Chapter 9). Recalling that the determinant of an operator on a complex vector space is the product of the eigenvalues, we see that our definition of the determinant of a square matrix should force

$$\det A = (\det A_1) \dots (\det A_M).$$

However, we already know how to compute the determinant of each  $A_j$ , which has the same form as 10.18 (of course with a different value of  $n$ ). Putting all this together, we see that we should have

$$\det A = (-1)^{n_1-1} \dots (-1)^{n_M-1} c_1 \dots c_n,$$

where  $A_j$  has size  $n_j$ -by- $n_j$ . The number  $(-1)^{n_1-1} \dots (-1)^{n_M-1}$  is called the sign of the permutation  $(p_1, \dots, p_n)$ , denoted  $\text{sign}(p_1, \dots, p_n)$  (this is a temporary definition that we will change to an equivalent definition later, when we define the sign of an arbitrary permutation).

To put this into a form that does not depend on the particular permutation  $(p_1, \dots, p_n)$ , let  $a_{j,k}$  denote the entry in row  $j$ , column  $k$ , of  $A$ ; thus

$$a_{j,k} = \begin{cases} 0 & \text{if } j \neq p_k; \\ c_k & \text{if } j = p_k. \end{cases}$$

Then

$$\mathbf{10.22} \quad \det A = \sum_{(m_1, \dots, m_n) \in \text{perm } n} (\text{sign}(m_1, \dots, m_n)) a_{m_1,1} \dots a_{m_n,n},$$

because each summand is 0 except the one corresponding to the permutation  $(p_1, \dots, p_n)$ .

Consider now an arbitrary matrix  $A$  with entry  $a_{j,k}$  in row  $j$ , column  $k$ . Using the paragraph above as motivation, we guess that  $\det A$  should be defined by 10.22. This will turn out to be correct. We can now dispense with the motivation and begin the more formal approach. First we will need to define the sign of an arbitrary permutation.

*Some texts use the unnecessarily fancy term **signum**, which means the same as **sign**.*

The **sign** of a permutation  $(m_1, \dots, m_n)$  is defined to be 1 if the number of pairs of integers  $(j, k)$  with  $1 \leq j < k \leq n$  such that  $j$  appears after  $k$  in the list  $(m_1, \dots, m_n)$  is even and  $-1$  if the number of such pairs is odd. In other words, the sign of a permutation equals 1 if the natural order has been changed an even number of times and equals  $-1$  if the natural order has been changed an odd number of times. For example, in the permutation  $(2, 3, \dots, n, 1)$  the only pairs  $(j, k)$  with  $j < k$  that appear with changed order are  $(1, 2), (1, 3), \dots, (1, n)$ ; because we have  $n - 1$  such pairs, the sign of this permutation equals  $(-1)^{n-1}$  (note that the same quantity appeared in 10.19).

The permutation  $(2, 1, 3, 4)$ , which is obtained from the permutation  $(1, 2, 3, 4)$  by interchanging the first two entries, has sign  $-1$ . The next lemma shows that interchanging any two entries of any permutation changes the sign of the permutation.

**10.23 Lemma:** *Interchanging two entries in a permutation multiplies the sign of the permutation by  $-1$ .*

**PROOF:** Suppose we have two permutations, where the second permutation is obtained from the first by interchanging two entries. If the two entries that we interchanged were in their natural order in the first permutation, then they no longer are in the second permutation, and

vice versa, for a net change (so far) of 1 or  $-1$  (both odd numbers) in the number of pairs not in their natural order.

Consider each entry between the two interchanged entries. If an intermediate entry was originally in the natural order with respect to the first interchanged entry, then it no longer is, and vice versa. Similarly, if an intermediate entry was originally in the natural order with respect to the second interchanged entry, then it no longer is, and vice versa. Thus the net change for each intermediate entry in the number of pairs not in their natural order is 2, 0, or  $-2$  (all even numbers).

For all the other entries, there is no change in the number of pairs not in their natural order. Thus the total net change in the number of pairs not in their natural order is an odd number. Thus the sign of the second permutation equals  $-1$  times the sign of the first permutation. ■

If  $A$  is an  $n$ -by- $n$  matrix

$$10.24 \quad A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix},$$

then the **determinant** of  $A$ , denoted  $\det A$ , is defined by

$$10.25 \quad \det A = \sum_{(m_1, \dots, m_n) \in \text{perm } n} (\text{sign}(m_1, \dots, m_n)) a_{m_1,1} \dots a_{m_n,n}.$$

For example, if  $A$  is the 1-by-1 matrix  $[a_{1,1}]$ , then  $\det A = a_{1,1}$  because  $\text{perm } 1$  has only one element, namely,  $(1)$ , which has sign 1. For a more interesting example, consider a typical 2-by-2 matrix. Clearly  $\text{perm } 2$  has only two elements, namely,  $(1, 2)$ , which has sign 1, and  $(2, 1)$ , which has sign  $-1$ . Thus

$$10.26 \quad \det \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} = a_{1,1}a_{2,2} - a_{2,1}a_{1,2}.$$

To make sure you understand this process, you should now find the formula for the determinant of the 3-by-3 matrix

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

using just the definition given above (do this even if you already know the answer).

*Our motivation for this definition comes from 10.22.*

*The set  $\text{perm } 3$  contains 6 elements. In general,  $\text{perm } n$  contains  $n!$  elements. Note that  $n!$  rapidly grows large as  $n$  increases.*

Let's compute the determinant of an upper-triangular matrix

$$A = \begin{bmatrix} a_{1,1} & & * \\ & \ddots & \\ 0 & & a_{n,n} \end{bmatrix}.$$

The permutation  $(1, 2, \dots, n)$  has sign 1 and thus contributes a term of  $a_{1,1} \dots a_{n,n}$  to the sum 10.25 defining  $\det A$ . Any other permutation  $(m_1, \dots, m_n) \in \text{perm } n$  contains at least one entry  $m_j$  with  $m_j > j$ , which means that  $a_{m_j, j} = 0$  (because  $A$  is upper triangular). Thus all the other terms in the sum 10.25 defining  $\det A$  make no contribution. Hence  $\det A = a_{1,1} \dots a_{n,n}$ . In other words, the determinant of an upper-triangular matrix equals the product of the diagonal entries. In particular, this means that if  $V$  is a complex vector space,  $T \in \mathcal{L}(V)$ , and we choose a basis of  $V$  with respect to which  $\mathcal{M}(T)$  is upper triangular, then  $\det T = \det \mathcal{M}(T)$ . Our goal is to prove that this holds for every basis of  $V$ , not just bases that give upper-triangular matrices.

Generalizing the computation from the paragraph above, next we will show that if  $A$  is a block upper-triangular matrix

$$A = \begin{bmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{bmatrix},$$

where each  $A_j$  is a 1-by-1 or 2-by-2 matrix, then

$$\mathbf{10.27} \quad \det A = (\det A_1) \dots (\det A_m).$$

To prove this, consider an element of  $\text{perm } n$ . If this permutation moves an index corresponding to a 1-by-1 block on the diagonal anywhere else, then the permutation makes no contribution to the sum 10.25 defining  $\det A$  (because  $A$  is block upper triangular). For a pair of indices corresponding to a 2-by-2 block on the diagonal, the permutation must either leave these indices fixed or interchange them; otherwise again the permutation makes no contribution to the sum 10.25 defining  $\det A$  (because  $A$  is block upper triangular). These observations, along with the formula 10.26 for the determinant of a 2-by-2 matrix, lead to 10.27. In particular, if  $V$  is a real vector space,  $T \in \mathcal{L}(V)$ , and we choose a basis of  $V$  with respect to which  $\mathcal{M}(T)$  is a block upper-triangular matrix with 1-by-1 and 2-by-2 blocks on the diagonal as in 9.9, then  $\det T = \det \mathcal{M}(T)$ .

Our goal is to prove that  $\det T = \det \mathcal{M}(T)$  for every  $T \in \mathcal{L}(V)$  and every basis of  $V$ . To do this, we will need to develop some properties of determinants of matrices. The lemma below is the first of the properties we will need.

**10.28 Lemma:** *Suppose  $A$  is a square matrix. If  $B$  is the matrix obtained from  $A$  by interchanging two columns, then*

$$\det A = -\det B.$$

PROOF: Suppose  $A$  is given by 10.24 and  $B$  is obtained from  $A$  by interchanging two columns. Think of the sum 10.25 defining  $\det A$  and the corresponding sum defining  $\det B$ . The same products of  $a$ 's appear in both sums, though they correspond to different permutations. The permutation corresponding to a given product of  $a$ 's when computing  $\det B$  is obtained by interchanging two entries in the corresponding permutation when computing  $\det A$ , thus multiplying the sign of the permutation by  $-1$  (see 10.23). Hence  $\det A = -\det B$ . ■

If  $T \in \mathcal{L}(V)$  and the matrix of  $T$  (with respect to some basis) has two equal columns, then  $T$  is not injective and hence  $\det T = 0$ . Though this comment makes the next lemma plausible, it cannot be used in the proof because we do not yet know that  $\det T = \det \mathcal{M}(T)$ .

**10.29 Lemma:** *If  $A$  is a square matrix that has two equal columns, then  $\det A = 0$ .*

PROOF: Suppose  $A$  is a square matrix that has two equal columns. Interchanging the two equal columns of  $A$  gives the original matrix  $A$ . Thus from 10.28 (with  $B = A$ ), we have  $\det A = -\det A$ , which implies that  $\det A = 0$ . ■

This section is long, so let's pause for a paragraph. The symbols \* that appear on the first page of each chapter are decorations intended to take up space so that the first section of the chapter can start on the next page. Chapter 1 has one of these symbols, Chapter 2 has two of them, and so on. The symbols get smaller with each chapter. What you may not have noticed is that the sum of the areas of the symbols at the beginning of each chapter is the same for all chapters. For example, the diameter of each symbol at the beginning of Chapter 10 equals  $1/\sqrt{10}$  times the diameter of the symbol in Chapter 1.

*An entire book could be devoted just to deriving properties of determinants.*

*Fortunately we need only a few of the basic properties.*

We need to introduce notation that will allow us to represent a matrix in terms of its columns. If  $A$  is an  $n$ -by- $n$  matrix

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix},$$

then we can think of the  $k^{\text{th}}$  column of  $A$  as an  $n$ -by-1 matrix

$$a_k = \begin{bmatrix} a_{1,k} \\ \vdots \\ a_{n,k} \end{bmatrix}.$$

We will write  $A$  in the form

$$[ a_1 \quad \dots \quad a_n ],$$

with the understanding that  $a_k$  denotes the  $k^{\text{th}}$  column of  $A$ . With this notation, note that  $a_{j,k}$ , with two subscripts, denotes an entry of  $A$ , whereas  $a_k$ , with one subscript, denotes a column of  $A$ .

The next lemma shows that a permutation of the columns of a matrix changes the determinant by a factor of the sign of the permutation.

*Some texts define the determinant to be the function defined on the square matrices that is linear as a function of each column separately and that satisfies 10.30 and  $\det I = 1$ . To prove that such a function exists and that it is unique takes a nontrivial amount of work.*

**10.30 Lemma:** Suppose  $A = [ a_1 \quad \dots \quad a_n ]$  is an  $n$ -by- $n$  matrix. If  $(m_1, \dots, m_n)$  is a permutation, then

$$\det[ a_{m_1} \quad \dots \quad a_{m_n} ] = (\text{sign}(m_1, \dots, m_n)) \det A.$$

**PROOF:** Suppose  $(m_1, \dots, m_n) \in \text{perm } n$ . We can transform the matrix  $[ a_{m_1} \quad \dots \quad a_{m_n} ]$  into  $A$  through a series of steps. In each step, we interchange two columns and hence multiply the determinant by  $-1$  (see 10.28). The number of steps needed equals the number of steps needed to transform the permutation  $(m_1, \dots, m_n)$  into the permutation  $(1, \dots, n)$  by interchanging two entries in each step. The proof is completed by noting that the number of such steps is even if  $(m_1, \dots, m_n)$  has sign 1, odd if  $(m_1, \dots, m_n)$  has sign  $-1$  (this follows from 10.23, along with the observation that the permutation  $(1, \dots, n)$  has sign 1). ■

Let  $A = [ a_1 \quad \dots \quad a_n ]$ . For  $1 \leq k \leq n$ , think of all columns of  $A$  except the  $k^{\text{th}}$  column as fixed. We have

$$\det A = \det[ a_1 \quad \dots \quad a_k \quad \dots \quad a_n ],$$

and we can think of  $\det A$  as a function of the  $k^{\text{th}}$  column  $a_k$ . This function, which takes  $a_k$  to the determinant above, is a linear map from the vector space of  $n$ -by-1 matrices with entries in  $\mathbf{F}$  to  $\mathbf{F}$ . The linearity follows easily from 10.25, where each term in the sum contains precisely one entry from the  $k^{\text{th}}$  column of  $A$ .

Now we are ready to prove one of the key properties about determinants of square matrices. This property will enable us to connect the determinant of an operator with the determinant of its matrix. Note that this proof is considerably more complicated than the proof of the corresponding result about the trace (see 10.9).

**10.31 Theorem:** *If  $A$  and  $B$  are square matrices of the same size, then*

$$\det(AB) = \det(BA) = (\det A)(\det B).$$

*This theorem was first proved in 1812 by the French mathematicians Jacques Binet and Augustin-Louis Cauchy.*

PROOF: Let  $A = [ a_1 \quad \dots \quad a_n ]$ , where each  $a_k$  is an  $n$ -by-1 column of  $A$ . Let

$$B = \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,n} \end{bmatrix} = [ b_1 \quad \dots \quad b_n ],$$

where each  $b_k$  is an  $n$ -by-1 column of  $B$ . Let  $e_k$  denote the  $n$ -by-1 matrix that equals 1 in the  $k^{\text{th}}$  row and 0 elsewhere. Note that  $Ae_k = a_k$  and  $Be_k = b_k$ . Furthermore,  $b_k = \sum_{m=1}^n b_{m,k} e_m$ .

First we will prove that  $\det(AB) = (\det A)(\det B)$ . A moment's thought about the definition of matrix multiplication shows that  $AB = [ Ab_1 \quad \dots \quad Ab_n ]$ . Thus

$$\begin{aligned} \det(AB) &= \det[ Ab_1 \quad \dots \quad Ab_n ] \\ &= \det[ A(\sum_{m=1}^n b_{m,1} e_m) \quad \dots \quad A(\sum_{m=1}^n b_{m,n} e_m) ] \\ &= \det[ \sum_{m=1}^n b_{m,1} A e_m \quad \dots \quad \sum_{m=1}^n b_{m,n} A e_m ] \\ &= \sum_{m_1=1}^n \dots \sum_{m_n=1}^n b_{m_1,1} \dots b_{m_n,n} \det[ A e_{m_1} \quad \dots \quad A e_{m_n} ], \end{aligned}$$

where the last equality comes from repeated applications of the linearity of  $\det$  as a function of one column at a time. In the last sum above,

all terms in which  $m_j = m_k$  for some  $j \neq k$  can be ignored because the determinant of a matrix with two equal columns is 0 (by 10.29). Thus instead of summing over all  $m_1, \dots, m_n$  with each  $m_j$  taking on values  $1, \dots, n$ , we can sum just over the permutations, where the  $m_j$ 's have distinct values. In other words,

$$\begin{aligned}
 \det(AB) &= \sum_{(m_1, \dots, m_n) \in \text{perm } n} b_{m_1,1} \dots b_{m_n,n} \det[ Ae_{m_1} \dots Ae_{m_n} ] \\
 &= \sum_{(m_1, \dots, m_n) \in \text{perm } n} b_{m_1,1} \dots b_{m_n,n} (\text{sign}(m_1, \dots, m_n)) \det A \\
 &= (\det A) \sum_{(m_1, \dots, m_n) \in \text{perm } n} (\text{sign}(m_1, \dots, m_n)) b_{m_1,1} \dots b_{m_n,n} \\
 &= (\det A)(\det B),
 \end{aligned}$$

where the second equality comes from 10.30.

In the paragraph above, we proved that  $\det(AB) = (\det A)(\det B)$ . Interchanging the roles of  $A$  and  $B$ , we have  $\det(BA) = (\det B)(\det A)$ . The last equation can be rewritten as  $\det(BA) = (\det A)(\det B)$ , completing the proof. ■

Now we can prove that the determinant of the matrix of an operator is independent of the basis with respect to which the matrix is computed.

**10.32 Corollary:** Suppose  $T \in \mathcal{L}(V)$ . If  $(u_1, \dots, u_n)$  and  $(v_1, \dots, v_n)$  are bases of  $V$ , then

$$\det \mathcal{M}(T, (u_1, \dots, u_n)) = \det \mathcal{M}(T, (v_1, \dots, v_n)).$$

*Note the similarity of this proof to the proof of the analogous result about the trace (see 10.10).*

**PROOF:** Suppose  $(u_1, \dots, u_n)$  and  $(v_1, \dots, v_n)$  are bases of  $V$ . Let  $A = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$ . Then

$$\begin{aligned}
 \det \mathcal{M}(T, (u_1, \dots, u_n)) &= \det(A^{-1}(\mathcal{M}(T, (v_1, \dots, v_n))A)) \\
 &= \det((\mathcal{M}(T, (v_1, \dots, v_n))A)A^{-1}) \\
 &= \det \mathcal{M}(T, (v_1, \dots, v_n)),
 \end{aligned}$$

where the first equality follows from 10.3 and the second equality follows from 10.31. The third equality completes the proof. ■

The theorem below states that the determinant of an operator equals the determinant of the matrix of the operator. This theorem does not specify a basis because, by the corollary above, the determinant of the matrix of an operator is the same for every choice of basis.

**10.33 Theorem:** *If  $T \in \mathcal{L}(V)$ , then  $\det T = \det \mathcal{M}(T)$ .*

PROOF: Let  $T \in \mathcal{L}(V)$ . As noted above, 10.32 implies that  $\det \mathcal{M}(T)$  is independent of which basis of  $V$  we choose. Thus to show that

$$\det T = \det \mathcal{M}(T)$$

for every basis of  $V$ , we need only show that the equation above holds for some basis of  $V$ . We already did this (on page 230), choosing a basis of  $V$  with respect to which  $\mathcal{M}(T)$  is an upper-triangular matrix (if  $V$  is a complex vector space) or an appropriate block upper-triangular matrix (if  $V$  is a real vector space). ■

If we know the matrix of an operator on a complex vector space, the theorem above allows us to find the product of all the eigenvalues without finding any of the eigenvalues. For example, consider the operator on  $\mathbb{C}^5$  whose matrix is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

No one knows an exact formula for any of the eigenvalues of this operator. However, we do know that the product of the eigenvalues equals  $-3$  because the determinant of the matrix above equals  $-3$ .

The theorem above also allows us easily to prove some useful properties about determinants of operators by shifting to the language of determinants of matrices, where certain properties have already been proved or are obvious. We carry out this procedure in the next corollary.

**10.34 Corollary:** *If  $S, T \in \mathcal{L}(V)$ , then*

$$\det(ST) = \det(TS) = (\det S)(\det T).$$

PROOF: Suppose  $S, T \in \mathcal{L}(V)$ . Choose any basis of  $V$ . Then

$$\begin{aligned}\det(ST) &= \det \mathcal{M}(ST) \\ &= \det(\mathcal{M}(S)\mathcal{M}(T)) \\ &= (\det \mathcal{M}(S))(\det \mathcal{M}(T)) \\ &= (\det S)(\det T),\end{aligned}$$

where the first and last equalities come from 10.33 and the third equality comes from 10.31.

In the paragraph above, we proved that  $\det(ST) = (\det S)(\det T)$ . Interchanging the roles of  $S$  and  $T$ , we have  $\det(TS) = (\det T)(\det S)$ . Because multiplication of elements of  $\mathbf{F}$  is commutative, the last equation can be rewritten as  $\det(TS) = (\det S)(\det T)$ , completing the proof. ■

## Volume

*Most applied mathematicians agree that determinants should rarely be used in serious numeric calculations.*

We proved the basic results of linear algebra before introducing determinants in this final chapter. Though determinants have value as a research tool in more advanced subjects, they play little role in basic linear algebra (when the subject is done right). Determinants do have one important application in undergraduate mathematics, namely, in computing certain volumes and integrals. In this final section we will use the linear algebra we have learned to make clear the connection between determinants and these applications. Thus we will be dealing with a part of analysis that uses linear algebra.

We begin with some purely linear algebra results that will be useful when investigating volumes. Recall that an isometry on an inner-product space is an operator that preserves norms. The next result shows that every isometry has determinant with absolute value 1.

**10.35 Proposition:** *Suppose that  $V$  is an inner-product space. If  $S \in \mathcal{L}(V)$  is an isometry, then  $|\det S| = 1$ .*

PROOF: Suppose  $S \in \mathcal{L}(V)$  is an isometry. First consider the case where  $V$  is a complex inner-product space. Then all the eigenvalues of  $S$  have absolute value 1 (by 7.37). Thus the product of the eigenvalues of  $S$ , counting multiplicity, has absolute value one. In other words,  $|\det S| = 1$ , as desired.

Now suppose  $V$  is a real inner-product space. Then there is an orthonormal basis of  $V$  with respect to which  $S$  has a block diagonal matrix, where each block on the diagonal is a 1-by-1 matrix containing 1 or  $-1$  or a 2-by-2 matrix of the form

$$10.36 \quad \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix},$$

with  $\theta \in (0, \pi)$  (see 7.38). Note that the constant term of the characteristic polynomial of each matrix of the form 10.36 equals 1 (because  $\cos^2 \theta + \sin^2 \theta = 1$ ). Hence the second coordinate of every eigenpair of  $S$  equals 1. Thus the determinant of  $S$  is the product of 1's and  $-1$ 's. In particular,  $|\det S| = 1$ , as desired. ■

Suppose  $V$  is a real inner-product space and  $S \in \mathcal{L}(V)$  is an isometry. By the proposition above, the determinant of  $S$  equals 1 or  $-1$ . Note that

$$\{v \in V : Sv = -v\}$$

is the subspace of  $V$  consisting of all eigenvectors of  $S$  corresponding to the eigenvalue  $-1$  (or is the subspace  $\{0\}$  if  $-1$  is not an eigenvalue of  $S$ ). Thinking geometrically, we could say that this is the subspace on which  $S$  reverses direction. A careful examination of the proof of the last proposition shows that  $\det S = 1$  if this subspace has even dimension and  $\det S = -1$  if this subspace has odd dimension.

A self-adjoint operator on a real inner-product space has no eigenpairs (by 7.11). Thus the determinant of a self-adjoint operator on a real inner-product space equals the product of its eigenvalues, counting multiplicity (of course, this holds for any operator, self-adjoint or not, on a complex vector space).

Recall that if  $V$  is an inner-product space and  $T \in \mathcal{L}(V)$ , then  $T^*T$  is a positive operator and hence has a unique positive square root, denoted  $\sqrt{T^*T}$  (see 7.27 and 7.28). Because  $\sqrt{T^*T}$  is positive, all its eigenvalues are nonnegative (again, see 7.27), and hence its determinant is nonnegative. Thus in the corollary below, taking the absolute value of  $\det \sqrt{T^*T}$  would be superfluous.

**10.37 Corollary:** *Suppose  $V$  is an inner-product space. If  $T \in \mathcal{L}(V)$ , then*

$$|\det T| = \det \sqrt{T^*T}.$$