CHAPTER 10

British mathematician and pioneer computer scientist Ada Lovelace (1815–1852), as painted by Alfred Chalon in this 1840 portrait.

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 the trace and determinant of an operator and then connect these notions to the corresponding notions for matrices. The book concludes with an explanation of the important role played by determinants in the theory of volume and integration.

Our assumptions for this chapter are as follows:

10.1 **Notation** $\mathbf{F}$, V

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

LEARNING OBJECTIVES FOR THIS CHAPTER

- change of basis and its effect upon the matrix of an operator
- trace of an operator and of a matrix
- determinant of an operator and of a matrix
- determinants and volume

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-11080-6_10

10.A Trace

For our study of the trace and determinant, we will need to know how the matrix of an operator changes with a change of basis. Thus we begin this chapter by developing the necessary material about change of basis.

Change of Basis

With respect to every basis of V, the matrix of the identity operator $I \in \mathcal{L}(V)$ is the diagonal matrix with 1's on the diagonal and 0's elsewhere. We also use the symbol I for the name of this matrix, as shown in the next definition.

10.2 Definition ***identity matrix, I***

Suppose n is a positive integer. The n-by-n diagonal matrix

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

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.

10.3 Definition ***invertible, inverse,*** A^{-1}

A square matrix A is called ***invertible*** if there is a square matrix B of the same size such that $AB = BA = I$; we call B the ***inverse*** of A and denote it by A^{-1}.

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

The same proof as used in 3.54 shows that if A is an invertible square matrix, then there is a unique matrix B such that $AB = BA = I$ (and thus the notation $B = A^{-1}$ is justified).

In Section 3.C we defined the matrix of a linear map from one vector space to another with respect to two bases—one basis of the first vector space and another basis of 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 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 and display at most one basis because we are using one basis in two capacities.

The next result is one of the unusual cases in which we use two different bases even though we have operators from a vector space to itself. It is just a convenient restatement of 3.43 (with U and W both equal to V), but now we are being more careful to include the various bases explicitly in the notation. The result below holds because we defined matrix multiplication to make it true—see 3.43 and the material preceding it.

10.4 The matrix of the product of linear maps

Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ and $w_1, \dots, w_n$ are all bases of V. Suppose $S, T \in \mathcal{L}(V)$. Then

$$\mathcal{M}\big(ST, (u_1, \dots, u_n), (w_1, \dots, w_n)\big) = \\ \mathcal{M}\big(S, (v_1, \dots, v_n), (w_1, \dots, w_n)\big)\mathcal{M}\big(T, (u_1, \dots, u_n), (v_1, \dots, v_n)\big).$$

The next result deals with the matrix of the identity operator I with respect to two different bases. Note that the k^{th} column of the matrix $\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$ consists of the scalars needed to write u_k as a linear combination of $v_1, \dots, v_n$.

10.5 Matrix of the identity with respect to two bases

Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then the matrices $\mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$ and $\mathcal{M}\big(I, (v_1, \dots, v_n), (u_1, \dots, u_n)\big)$ are invertible, and each is the inverse of the other.

Proof In 10.4, replace w_j with u_j, and replace S and T with I, getting

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

Now interchange the roles of the u's and v's, getting

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

These two equations give the desired result. ■

10.6 Example Consider the bases $(4,2), (5,3)$ and $(1,0), (0,1)$ of $\mathbf{F}^2$. Obviously

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

because $I(4,2) = 4(1,0) + 2(0,1)$ and $I(5,3) = 5(1,0) + 3(0,1)$.

The inverse of the matrix above is

$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix},$$

as you should verify. Thus 10.5 implies that

$$\mathcal{M}\Big(I, \big((1,0),(0,1)\big), \big((4,2),(5,3)\big)\Big) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}.$$

Now we can see how the matrix of T changes when we change bases. In the result below, we have two different bases of V. Recall that the notation $\mathcal{M}\big(T, (u_1, \dots, u_n)\big)$ is shorthand for $\mathcal{M}\big(T, (u_1, \dots, u_n), (u_1, \dots, u_n)\big)$

10.7 Change of basis formula

Suppose $T \in \mathcal{L}(V)$. Let $u_1, \dots, u_n$ and $v_1, \dots, v_n$ be bases of V. Let $A = \mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$. Then

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

Proof In 10.4, replace w_j with u_j and replace S with I, getting

10.8 $$\mathcal{M}\big(T, (u_1, \dots, u_n)\big) = A^{-1}\mathcal{M}\big(T, (u_1, \dots, u_n), (v_1, \dots, v_n)\big),$$

where we have used 10.5.

Again use 10.4, this time replacing w_j with v_j. Also replace T with I and replace S with T, getting

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

Substituting the equation above into 10.8 gives the desired result. ∎

Trace: A Connection Between Operators and Matrices

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Let $n = \dim V$. Recall that we defined the multiplicity of λ to be the dimension of the generalized eigenspace $G(\lambda, T)$ (see 8.24) and that this multiplicity equals $\dim \operatorname{null}(T - \lambda I)^n$ (see 8.11). Recall also that if V is a complex vector space, then the sum of the multiplicities of all the eigenvalues of T equals n (see 8.26).

In the definition below, the sum of the eigenvalues "with each eigenvalue repeated according to its multiplicity" means that if $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues of T (or of $T_{\mathbf{C}}$ if V is a real vector space) with multiplicities $d_1, \dots, d_m$, then the sum is

$$d_1\lambda_1 + \cdots + d_m\lambda_m.$$

Or if you prefer to list the eigenvalues with each repeated according to its multiplicity, then the eigenvalues could be denoted $\lambda_1, \dots, \lambda_n$ (where the index n equals $\dim V$) and the sum is

$$\lambda_1 + \cdots + \lambda_n.$$

10.9 Definition ***trace of an operator***

Suppose $T \in \mathcal{L}(V)$.

- If $\mathbf{F} = \mathbf{C}$, then the ***trace*** of T is the sum of the eigenvalues of T, with each eigenvalue repeated according to its multiplicity.
- If $\mathbf{F} = \mathbf{R}$, then the ***trace*** of T is the sum of the eigenvalues of $T_{\mathbf{C}}$, with each eigenvalue repeated according to its multiplicity.

The trace of T is denoted by $\operatorname{trace} T$.

10.10 Example Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

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

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 find that $\operatorname{trace} T = 1 + (2 + 3i) + (2 - 3i)$; in other words, $\operatorname{trace} T = 5$.

The trace has a close connection with the characteristic polynomial. Suppose $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T (or of $T_{\mathbf{C}}$ if V is a real vector space) with each eigenvalue repeated according to its multiplicity. Then by definition (see 8.34 and 9.21), the characteristic polynomial of T equals

$$(z - \lambda_1) \cdots (z - \lambda_n).$$

Expanding the polynomial above, we can write the characteristic polynomial of T in the form

10.11 $$z^n - (\lambda_1 + \cdots + \lambda_n) z^{n-1} + \cdots + (-1)^n (\lambda_1 \cdots \lambda_n).$$

The expression above immediately leads to the following result.

10.12 Trace and characteristic polynomial

Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then trace T equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.

Most of the rest of this section is devoted to discovering how to compute 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 as in 8.29. With respect to that basis, T has an upper-triangular matrix with the diagonal of the matrix containing precisely the eigenvalues of T, each repeated according to its multiplicity. Thus trace T equals the sum of the diagonal entries of $\mathcal{M}(T)$ with respect to that basis.

The same formula works for the operator $T \in \mathcal{L}(\mathbf{C}^3)$ in Example 10.10 whose trace equals 5. In that example, the matrix is not in upper-triangular form. However, the sum of the diagonal entries of the matrix in that example equals 5, which is the trace of the operator T.

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 suspicion turns out to be true. To prove it, we start by making the following definition.

10.13 Definition *trace of a matrix*

The ***trace*** of a square matrix A, denoted trace A, is defined to be the sum of the diagonal entries of A.

Now we have defined the trace of an operator and the trace of a square matrix, using the same word "trace" in two different contexts. This would be bad terminology unless the two concepts turn out to be essentially the same. As we will see, it is indeed true that $\operatorname{trace} T = \operatorname{trace} \mathcal{M}\big(T, (v_1, \dots, v_n)\big)$, where $v_1, \dots, v_n$ is an arbitrary basis of V. We will need the following result for the proof.

10.14 Trace of AB equals trace of BA

If A and B are square matrices of the same size, then

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

Proof Suppose

$$A = \begin{pmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,n} \end{pmatrix}, \quad B = \begin{pmatrix} B_{1,1} & \dots & B_{1,n} \\ \vdots & & \vdots \\ B_{n,1} & \dots & B_{n,n} \end{pmatrix}.$$

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

$$\sum_{k=1}^{n} A_{j,k} B_{k,j}.$$

Thus

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

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.

10.15 Trace of matrix of operator does not depend on basis

Let $T \in \mathcal{L}(V)$. Suppose $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Then

$$\operatorname{trace} \mathcal{M}\big(T, (u_1, \ldots, u_n)\big) = \operatorname{trace} \mathcal{M}\big(T, (v_1, \ldots, v_n)\big).$$

Proof Let $A = \mathcal{M}\big(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n)\big)$. Then

$$\begin{aligned} \operatorname{trace} \mathcal{M}\big(T, (u_1, \ldots, u_n)\big) &= \operatorname{trace}\Big(A^{-1}\big(\mathcal{M}\big(T, (v_1, \ldots, v_n)\big)A\big)\Big) \\ &= \operatorname{trace}\Big(\big(\mathcal{M}\big(T, (v_1, \ldots, v_n)\big)A\big)A^{-1}\Big) \\ &= \operatorname{trace} \mathcal{M}\big(T, (v_1, \ldots, v_n)\big), \end{aligned}$$

where the first equality comes from 10.7 and the second equality follows from 10.14. The third equality completes the proof. ■

The result below, which is the most important result in this section, 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 result above, the sum of the diagonal entries of the matrix of an operator is the same for every choice of basis.

10.16 Trace of an operator equals trace of its matrix

Suppose $T \in \mathcal{L}(V)$. Then $\operatorname{trace} T = \operatorname{trace} \mathcal{M}(T)$.

Proof As noted above, $\operatorname{trace} \mathcal{M}(T)$ is independent of which basis of V we choose (by 10.15). Thus to show that

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

for every basis of V, we need only show that the equation above holds for some basis of V.

As we have already discussed, if V is a complex vector space, then choosing the basis as in 8.29 gives the desired result. If V is a real vector space, then applying the complex case to the complexification $T_{\mathbf{C}}$ (which is used to define $\operatorname{trace} T$) gives the desired result. ■

If we know the matrix of an operator on a complex vector space, the result above allows us to find the sum of all the eigenvalues without finding any of the eigenvalues, as shown by the next example.

10.17 Example Consider the operator on $\mathbf{C}^5$ whose matrix is

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

No one can find 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.

We can use 10.16 to give easy proofs of 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. The proof of the next result is an example of this technique. The eigenvalues of $S + T$ are not, in general, formed from adding together eigenvalues of S and eigenvalues of T. Thus the next result would be difficult to prove without using 10.16.

10.18 Trace is additive

Suppose $S, T \in \mathcal{L}(V)$. Then $\operatorname{trace}(S + T) = \operatorname{trace} S + \operatorname{trace} T$.

Proof Choose a basis of V. Then

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

where again the first and last equalities come from 10.16; the third equality is obvious from the definition of the trace of a matrix. ■

The techniques we have developed have the following curious consequence. A generalization of this result to infinite-dimensional vector spaces has important consequences in modern physics, particularly in quantum theory.

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

10.19 The identity is not the difference of ST and TS

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

Proof Suppose $S, T \in \mathcal{L}(V)$. Choose a basis of V. Then

$$\begin{aligned}\operatorname{trace}(ST - TS) &= \operatorname{trace}(ST) - \operatorname{trace}(TS)\\ &= \operatorname{trace} \mathcal{M}(ST) - \operatorname{trace} \mathcal{M}(TS)\\ &= \operatorname{trace}\bigl(\mathcal{M}(S)\mathcal{M}(T)\bigr) - \operatorname{trace}\bigl(\mathcal{M}(T)\mathcal{M}(S)\bigr)\\ &= 0,\end{aligned}$$

where the first equality comes from 10.18, the second equality comes from 10.16, the third equality comes from 3.43, and the fourth equality comes from 10.14. Clearly the trace of I equals $\dim V$, which is not 0. Because $ST - TS$ and I have different traces, they cannot be equal. ■

EXERCISES 10.A

1 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Prove that the matrix $\mathcal{M}\bigl(T, (v_1, \dots, v_n)\bigr)$ is invertible if and only if T is invertible.

2 Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.

3 Suppose $T \in \mathcal{L}(V)$ has the same matrix with respect to every basis of V. Prove that T is a scalar multiple of the identity operator.

4 Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let $T \in \mathcal{L}(V)$ be the operator such that $Tv_k = u_k$ for $k = 1, \dots, n$. Prove that

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

5 Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

6 Give an example of a real vector space V and $T \in \mathcal{L}(V)$ such that $\operatorname{trace}(T^2) < 0$.

7 Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and V has a basis consisting of eigenvectors of T. Prove that $\operatorname{trace}(T^2) \geq 0$.

8 Suppose V is an inner product space and $v, w \in V$. Define $T \in \mathcal{L}(V)$ by $Tu = \langle u, v \rangle w$. Find a formula for trace T.

9 Suppose $P \in \mathcal{L}(V)$ satisfies $P^2 = P$. Prove that

$$\operatorname{trace} P = \dim \operatorname{range} P.$$

10 Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{trace} T^* = \overline{\operatorname{trace} T}.$$

11 Suppose V is an inner product space. Suppose $T \in \mathcal{L}(V)$ is a positive operator and trace $T = 0$. Prove that $T = 0$.

12 Suppose V is an inner product space and $P, Q \in \mathcal{L}(V)$ are orthogonal projections. Prove that $\operatorname{trace}(PQ) \geq 0$.

13 Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

Someone tells you (accurately) that -48 and 24 are eigenvalues of T. Without using a computer or writing anything down, find the third eigenvalue of T.

14 Suppose $T \in \mathcal{L}(V)$ and $c \in \mathbf{F}$. Prove that $\operatorname{trace}(cT) = c \operatorname{trace} T$.

15 Suppose $S, T \in \mathcal{L}(V)$. Prove that $\operatorname{trace}(ST) = \operatorname{trace}(TS)$.

16 Prove or give a counterexample: if $S, T \in \mathcal{L}(V)$, then $\operatorname{trace}(ST) = (\operatorname{trace} S)(\operatorname{trace} T)$.

17 Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{trace}(ST) = 0$ for all $S \in \mathcal{L}(V)$. Prove that $T = 0$.

18 Suppose V is an inner product space with orthonormal basis $e_1, \dots, e_n$ and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{trace}(T^*T) = \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

Conclude that the right side of the equation above is independent of which orthonormal basis $e_1, \dots, e_n$ is chosen for V.

19 Suppose V is an inner product space. Prove that

$$\langle S, T \rangle = \operatorname{trace}(ST^*)$$

defines an inner product on $\mathcal{L}(V)$.

20 Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T, repeated according to multiplicity. Suppose

$$\begin{pmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,n} \end{pmatrix}$$

is the matrix of T with respect to some orthonormal basis of V. Prove that

$$|\lambda_1|^2 + \dots + |\lambda_n|^2 \le \sum_{k=1}^{n} \sum_{j=1}^{n} |A_{j,k}|^2.$$

21 Suppose V is an inner product space. Suppose $T \in \mathcal{L}(V)$ and

$$\|T^* v\| \le \|T v\|$$

for every $v \in V$. Prove that T is normal.
[*The exercise above fails on infinite-dimensional inner product spaces, leading to what are called hyponormal operators, which have a well-developed theory.*]

10.B *Determinant*

Determinant of an Operator

Now we are ready to define the determinant of an operator. Notice that the definition below mimics the approach we took when defining the trace, with the product of the eigenvalues replacing the sum of the eigenvalues.

10.20 Definition ***determinant of an operator,*** $\det T$

Suppose $T \in \mathcal{L}(V)$.

- If $\mathbf{F} = \mathbf{C}$, then the ***determinant*** of T is the product of the eigenvalues of T, with each eigenvalue repeated according to its multiplicity.
- If $\mathbf{F} = \mathbf{R}$, then the ***determinant*** of T is the product of the eigenvalues of $T_{\mathbf{C}}$, with each eigenvalue repeated according to its multiplicity.

The determinant of T is denoted by $\det T$.

If $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues of T (or of $T_{\mathbf{C}}$ if V is a real vector space) with multiplicities $d_1, \dots, d_m$, then the definition above implies

$$\det T = \lambda_1^{d_1} \cdots \lambda_m^{d_m}.$$

Or if you prefer to list the eigenvalues with each repeated according to its multiplicity, then the eigenvalues could be denoted $\lambda_1, \dots, \lambda_n$ (where the index n equals $\dim V$) and the definition above implies

$$\det T = \lambda_1 \cdots \lambda_n.$$

10.21 Example Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

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

Then the eigenvalues of T are 1, $2 + 3i$, and $2 - 3i$, each with multiplicity 1, as you can verify. Computing the product of the eigenvalues, we find that $\det T = 1 \cdot (2 + 3i) \cdot (2 - 3i)$; in other words, $\det T = 13$.

The determinant has a close connection with the characteristic polynomial. Suppose $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of T (or of $T_{\mathbf{C}}$ if V is a real vector space) with each eigenvalue repeated according to its multiplicity. Then the expression for the characteristic polynomial of T given by 10.11 gives the following result.

10.22 Determinant and characteristic polynomial

Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\det T$ equals $(-1)^n$ times the constant term of the characteristic polynomial of T.

Combining the result above and 10.12, we have the following result.

10.23 Characteristic polynomial, trace, and determinant

Suppose $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T can be written as

$$z^n - (\operatorname{trace} T)z^{n-1} + \cdots + (-1)^n(\det T).$$

We turn now to some simple but important properties of determinants. Later we will discover how to calculate $\det T$ from the matrix of T (with respect to an arbitrary basis).

The crucial result below has an easy proof due to our definition.

10.24 Invertible is equivalent to nonzero determinant

An operator on V 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 consider the case where 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, which happens if and only if 0 is not an eigenvalue of $T_{\mathbf{C}}$ (because $T_{\mathbf{C}}$ and T have the same real eigenvalues by 9.11). Thus again we see that T is invertible if and only if $\det T \neq 0$. ■

Some textbooks take the result below as the definition of the characteristic polynomial and then have our definition of the characteristic polynomial as a consequence.

10.25 Characteristic polynomial of T equals $\det(zI - T)$

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. If $\lambda, z \in \mathbf{C}$, then λ is an eigenvalue of T if and only if $z - \lambda$ is an eigenvalue of $zI - T$, as can be seen from the equation

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

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

Let $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of T, repeated according to multiplicity. Thus for $z \in \mathbf{C}$, the paragraph above shows that 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,

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

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. Applying the complex case to $T_{\mathbf{C}}$ gives the desired result. ■

Determinant of a Matrix

Our next task is to discover how to compute $\det T$ 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 as in 8.29. With respect to that basis, T has an upper-triangular matrix with the diagonal of the matrix containing precisely the eigenvalues of T, each repeated according to its multiplicity. Thus $\det T$ equals the product of the diagonal entries of $\mathcal{M}(T)$ with respect to that basis.

When dealing with the trace in the previous section, we discovered that the formula (trace = sum of diagonal entries) that worked for the upper-triangular matrix given by 8.29 also worked with respect to an arbitrary basis. Could that also work for determinants? In other words, is the determinant of an operator equal to the product of the diagonal entries of the matrix of the operator with respect to an arbitrary basis?

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 in Example 10.21 has determinant 13 but the product of the diagonal entries of its matrix equals 0.

For each square matrix A, we want to define the determinant of A, denoted $\det A$, so 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.

10.26 **Example** Suppose $a_1, \dots, a_n \in \mathbf{F}$. Let

$$A = \begin{pmatrix} 0 & & & & a_n \\ a_1 & 0 & & & \\ & a_2 & 0 & & \\ & & \ddots & \ddots & \\ & & & a_{n-1} & 0 \end{pmatrix};$$

here all entries of the matrix are 0 except for the upper-right corner and along the line just below the diagonal. Suppose $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$ is such that $\mathcal{M}\big(T, (v_1, \dots, v_n)\big) = A$. Find the determinant of T.

Solution First assume $a_j \neq 0$ for each $j = 1, \dots, n-1$. Note that the list $v_1, Tv_1, T^2v_1, \dots, T^{n-1}v_1$ equals $v_1, a_1v_2, a_1a_2v_3, \dots, a_1 \cdots a_{n-1}v_n$.

Computing the minimal polynomial is often an efficient method of finding the characteristic polynomial, as is done in this example.

Thus $v_1, Tv_1, \dots, T^{n-1}v_1$ is linearly independent (because the a's are all nonzero). Hence if p is a monic polynomial with degree at most $n-1$, then $p(T)v_1 \neq 0$. Thus the minimal polynomial of T cannot have degree less than n.

As you should verify, $T^n v_j = a_1 \cdots a_n v_j$ for each j. Thus we have $T^n = a_1 \cdots a_n I$. Hence $z^n - a_1 \cdots a_n$ is the minimal polynomial of T. Because $n = \dim V$ and the characteristic polynomial is a polynomial multiple of the minimal polynomial (9.26), this implies that $z^n - a_1 \cdots a_n$ is also the characteristic polynomial of T.

Thus 10.22 implies that

$$\det T = (-1)^{n-1} a_1 \cdots a_n.$$

If some a_j equals 0, then $Tv_j = 0$ for some j, which implies that 0 is an eigenvalue of T and hence $\det T = 0$. In other words, the formula above also holds if some a_j equals 0.

Thus in order to have $\det T = \det \mathcal{M}(T)$, we will have to make the determinant of the matrix in Example 10.26 equal to $(-1)^{n-1} a_1 \cdots a_n$. 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.

10.27 Definition ***permutation,*** perm n

- A ***permutation*** of $(1, \ldots, n)$ is a list $(m_1, \ldots, m_n)$ that contains each of the numbers $1, \ldots, n$ exactly once.
- The set of all permutations of $(1, \ldots, n)$ is denoted perm n.

For example, $(2, 3, 4, 5, 1) \in \text{perm } 5$. You should think of an element of perm n as a rearrangement of the first n integers.

10.28 Example Suppose $a_1, \ldots, a_n \in \mathbf{F}$ and $v_1, \ldots, v_n$ is a basis of V. Consider a permutation $(p_1, \ldots, p_n) \in \text{perm } n$ that can be obtained as follows: break $(1, \ldots, n)$ 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

$$(2, 3, 1, 5, 6, 7, 4, 9, 8)$$

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 the permutation displayed above.

Let $T \in \mathcal{L}(V)$ be the operator such that

$$T v_k = a_k v_{p_k}$$

for $k = 1, \ldots, n$. Find $\det T$.

Solution This generalizes Example 10.26, because if $(p_1, \ldots, p_n)$ is the permutation $(2, 3, \ldots, n, 1)$, then our operator T is the same as the operator T in Example 10.26.

With respect to the basis $v_1, \ldots, v_n$, the matrix of the operator T is a block diagonal matrix

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

where each block is a square matrix of the form of the matrix in 10.26.

Correspondingly, we can write $V = V_1 \oplus \cdots \oplus V_M$, where each V_j is invariant under T and each $T|_{V_j}$ is of the form of the operator in 10.26. Because $\det T = (\det T|_{V_1}) \cdots (\det T|_{V_M})$ (because the dimensions of the generalized eigenspaces in the V_j add up to $\dim V$), we have

$$\det T = (-1)^{n_1-1} \cdots (-1)^{n_M-1} a_1 \cdots a_n,$$

where V_j has dimension n_j (and correspondingly each A_j has size n_j-by-n_j) and we have used the result from 10.26.

The number $(-1)^{n_1-1} \cdots (-1)^{n_M-1}$ that appears above is called the sign of the corresponding permutation $(p_1, \dots, p_n)$, denoted $\operatorname{sign}(p_1, \dots, p_n)$ [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 the matrix A from Example 10.28. Thus

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

Example 10.28 shows that we want

10.29 $$\det A = \sum_{(m_1,\dots,m_n) \in \operatorname{perm} n} \bigl(\operatorname{sign}(m_1, \dots, m_n)\bigr) A_{m_1,1} \cdots A_{m_n,n};$$

note that each summand is 0 except the one corresponding to the permutation $(p_1, \dots, p_n)$ [which is why it does not matter that the sign of the other permutations is not yet defined].

We can now guess that $\det A$ should be defined by 10.29 for an arbitrary square matrix A. This will turn out to be correct. We will now dispense with the motivation and begin the more formal approach. First we will need to define the sign of an arbitrary permutation.

10.30 Definition ***sign of a permutation***

- 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.

10.31 Example ***sign of permutation***

- The only pair of integers (j, k) with $j < k$ such that j appears after k in the list $(2, 1, 3, 4)$ is $(1, 2)$. Thus the permutation $(2, 1, 3, 4)$ has sign -1.
- 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 Example 10.26).

The next result shows that interchanging two entries of a permutation changes the sign of the permutation.

10.32 Interchanging two entries in a permutation

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 interchanged entries 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.

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

Consider each entry between the two interchanged entries. If an intermediate entry was originally in the natural order with respect to both interchanged entries, then it is now in the natural order with respect to neither interchanged entry. Similarly, if an intermediate entry was originally in the natural order with respect to neither of the interchanged entries, then it is now in the natural order with respect to both interchanged entries. If an intermediate entry was originally in the natural order with respect to exactly one of the interchanged entries, then that is still true. Thus the net change for each intermediate entry in the number of pairs not in their natural order is 2, -2, or 0 (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. ■

Our motivation for the next definition comes from 10.29.

10.33 Definition ***determinant of a matrix,*** det A

Suppose A is an n-by-n matrix

$$A = \begin{pmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,n} \end{pmatrix}.$$

The ***determinant*** of A, denoted det A, is defined by

$$\det A = \sum_{(m_1,\dots,m_n)\in \operatorname{perm} n} \big(\operatorname{sign}(m_1,\dots,m_n)\big) A_{m_1,1}\cdots A_{m_n,n}.$$

10.34 Example ***determinants***

- If A is the 1-by-1 matrix $[A_{1,1}]$, then $\det A = A_{1,1}$, because perm 1 has only one element, namely (1), which has sign 1.
- Clearly perm 2 has only two elements, namely $(1,2)$, which has sign 1, and $(2,1)$, which has sign -1. Thus
$$\det \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} = A_{1,1}A_{2,2} - A_{2,1}A_{1,2}.$$

The set perm 3 *contains six elements. In general,* perm n *contains* $n!$ *elements. Note that* $n!$ *rapidly grows large as n increases.*

To make sure you understand this process, you should now find the formula for the determinant of an arbitrary 3-by-3 matrix using just the definition given above.

10.35 Example Compute the determinant of an upper-triangular matrix

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

Solution The permutation $(1,2,\dots,n)$ has sign 1 and thus contributes a term of $A_{1,1}\cdots A_{n,n}$ to the sum defining det A in 10.33. Any other permutation $(m_1,\dots,m_n) \in \operatorname{perm} n$ 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 in 10.33 make no contribution.

Hence $\det A = A_{1,1}\cdots A_{n,n}$. In other words, the determinant of an upper-triangular matrix equals the product of the diagonal entries.

Suppose V is a complex vector space, $T \in \mathcal{L}(V)$, and we choose a basis of V as in 8.29. With respect to that basis, T has an upper-triangular matrix with the diagonal of the matrix containing precisely the eigenvalues of T, each repeated according to its multiplicity. Thus Example 10.35 tells us that $\det T = \det \mathcal{M}(T)$, where the matrix is with respect to that basis.

Our goal is to prove that $\det T = \det \mathcal{M}(T)$ for every basis of V, not just the basis from 8.29. To do this, we will need to develop some properties of determinants of matrices. The result below is the first of the properties we will need.

10.36 Interchanging two columns in a matrix

Suppose A is a square matrix and B is the matrix obtained from A by interchanging two columns. Then

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

Proof Think of the sum defining $\det A$ in 10.33 and the corresponding sum defining $\det B$. The same products of $A_{j,k}$'s appear in both sums, although they correspond to different permutations. The permutation corresponding to a given product of $A_{j,k}$'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 10.32). Hence we see that $\det A = -\det B$. ■

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 $\det T = 0$. Although this comment makes the next result plausible, it cannot be used in the proof, because we do not yet know that $\det T = \det \mathcal{M}(T)$ for every choice of basis.

10.37 Matrices with two equal columns

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 10.36 (with $B = A$), we have

$$\det A = -\det A,$$

which implies that $\det A = 0$. ■

Recall from 3.44 that if A is an n-by-n matrix

$$A = \begin{pmatrix} A_{1,1} & \dots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \dots & A_{n,n} \end{pmatrix},$$

then we can think of the k^{th} column of A as an n-by-1 matrix denoted $A_{\cdot,k}$:

$$A_{\cdot,k} = \begin{pmatrix} A_{1,k} \\ \vdots \\ A_{n,k} \end{pmatrix}.$$

> *Some books 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 10.38 and* det $I = 1$. *To prove that such a function exists and that it is unique takes a nontrivial amount of work.*

Note that $A_{j,k}$, with two subscripts, denotes an entry of A, whereas $A_{\cdot,k}$, with a dot as a placeholder and one subscript, denotes a column of A. This notation allows us to write A in the form

$$\begin{pmatrix} A_{\cdot,1} & \dots & A_{\cdot,n} \end{pmatrix},$$

which will be useful.

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

10.38 Permuting the columns of a matrix

Suppose $A = \begin{pmatrix} A_{\cdot,1} & \dots & A_{\cdot,n} \end{pmatrix}$ is an n-by-n matrix and $(m_1, \dots, m_n)$ is a permutation. Then

$$\det\begin{pmatrix} A_{\cdot,m_1} & \dots & A_{\cdot,m_n} \end{pmatrix} = \big(\operatorname{sign}(m_1, \dots, m_n)\big) \det A.$$

Proof We can transform the matrix $\begin{pmatrix} A_{\cdot,m_1} & \dots & A_{\cdot,m_n} \end{pmatrix}$ into A through a series of steps. In each step, we interchange two columns and hence multiply the determinant by -1 (see 10.36). 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 10.32, along with the observation that the permutation $(1, \dots, n)$ has sign 1). ■

The next result about determinants will also be useful.

10.39 Determinant is a linear function of each column

Suppose k, n are positive integers with $1 \le k \le n$. Fix n-by-1 matrices $A_{\cdot,1}, \ldots, A_{\cdot,n}$ except $A_{\cdot,k}$. Then the function that takes an n-by-1 column vector $A_{\cdot,k}$ to

$$\det\begin{pmatrix} A_{\cdot,1} & \ldots & A_{\cdot,k} & \ldots & A_{\cdot,n} \end{pmatrix}$$

is a linear map from the vector space of n-by-1 matrices with entries in $\mathbf{F}$ to $\mathbf{F}$.

Proof The linearity follows easily from 10.33, where each term in the sum contains precisely one entry from the k^{th} column of A. ■

The result below was first proved in 1812 by French mathematicians Jacques Binet and Augustin-Louis Cauchy.

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 10.14).

10.40 Determinant is multiplicative

Suppose A and B are square matrices of the same size. Then

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

Proof Write $A = \begin{pmatrix} A_{\cdot,1} & \ldots & A_{\cdot,n} \end{pmatrix}$, where each $A_{\cdot,k}$ is an n-by-1 column of A. Also write

$$B = \begin{pmatrix} B_{1,1} & \ldots & B_{1,n} \\ \vdots & & \vdots \\ B_{n,1} & \ldots & B_{n,n} \end{pmatrix} = \begin{pmatrix} B_{\cdot,1} & \ldots & B_{\cdot,n} \end{pmatrix},$$

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

First we will prove $\det(AB) = (\det A)(\det B)$. As we observed earlier (see 3.49), the definition of matrix multiplication easily implies that $AB = \begin{pmatrix} AB_{\cdot,1} & \ldots & AB_{\cdot,n} \end{pmatrix}$. Thus

$$\begin{aligned}
\det(AB) &= \det\begin{pmatrix} AB_{\cdot,1} & \dots & AB_{\cdot,n} \end{pmatrix} \\
&= \det\begin{pmatrix} A(\sum_{m_1=1}^{n} B_{m_1,1}e_{m_1}) & \dots & A(\sum_{m_n=1}^{n} B_{m_n,n}e_{m_n}) \end{pmatrix} \\
&= \det\begin{pmatrix} \sum_{m_1=1}^{n} B_{m_1,1}Ae_{m_1} & \dots & \sum_{m_n=1}^{n} B_{m_n,n}Ae_{m_n} \end{pmatrix} \\
&= \sum_{m_1=1}^{n} \cdots \sum_{m_n=1}^{n} B_{m_1,1}\cdots B_{m_n,n} \det\begin{pmatrix} Ae_{m_1} & \dots & Ae_{m_n} \end{pmatrix},
\end{aligned}$$

where the last equality comes from repeated applications of the linearity of det as a function of one column at a time (10.39). 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 10.37). 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,

$$\begin{aligned}
\det(AB) &= \sum_{(m_1,\dots,m_n)\in \operatorname{perm} n} B_{m_1,1}\cdots B_{m_n,n} \det\begin{pmatrix} Ae_{m_1} & \dots & Ae_{m_n} \end{pmatrix} \\
&= \sum_{(m_1,\dots,m_n)\in \operatorname{perm} n} B_{m_1,1}\cdots B_{m_n,n} \big(\operatorname{sign}(m_1,\dots,m_n)\big) \det A \\
&= (\det A) \sum_{(m_1,\dots,m_n)\in \operatorname{perm} n} \big(\operatorname{sign}(m_1,\dots,m_n)\big) B_{m_1,1}\cdots B_{m_n,n} \\
&= (\det A)(\det B),
\end{aligned}$$

where the second equality comes from 10.38.

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. ■

Note the similarity of the proof of the next result to the proof of the analogous result about the trace (see 10.15).

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.

10.41 Determinant of matrix of operator does not depend on basis

Let $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then

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

Proof Let $A = \mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big)$. Then

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

where the first equality follows from 10.7 and the second equality follows from 10.40. The third equality completes the proof. ■

The result 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 result above, the determinant of the matrix of an operator is the same for every choice of basis.

10.42 Determinant of an operator equals determinant of its matrix

Suppose $T \in \mathcal{L}(V)$. Then $\det T = \det \mathcal{M}(T)$.

Proof As noted above, 10.41 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 result holds for some basis of V.

As we have already discussed, if V is a complex vector space, then choosing a basis of V as in 8.29 gives the desired result. If V is a real vector space, then applying the complex case to the complexification $T_{\mathbf{C}}$ (which is used to define $\det T$) gives the desired result. ■

If we know the matrix of an operator on a complex vector space, the result above allows us to find the product of all the eigenvalues without finding any of the eigenvalues.

10.43 **Example** Suppose T is the operator on $\mathbf{C}^5$ whose matrix is

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

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.

We can use 10.42 to give easy proofs of 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 result.

10.44 Determinant is multiplicative

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

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

Proof Choose a basis of V. Then

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

where the first and last equalities come from 10.42 and the third equality comes from 10.40.

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. ■

The Sign of the Determinant

We proved the basic results of linear algebra before introducing determinants in this final chapter. Although 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).

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

Determinants do have one important application in undergraduate mathematics, namely, in computing certain volumes and integrals. In this subsection we interpret the meaning of the sign of the determinant on a real vector space. Then in the final subsection 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 will begin with some purely linear algebra results that will also be useful when investigating volumes. Our setting will be inner product spaces. 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.

10.45 Isometries have determinant with absolute value 1

Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is an isometry. Then $|\det S| = 1$.

Proof First consider the case where V is a complex inner product space. Then all the eigenvalues of S have absolute value 1 (see the proof of 7.43). 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. We present two different proofs in this case.

Proof 1: With respect to the inner product on the complexification $V_{\mathbf{C}}$ given by Exercise 3 in Section 9.B, it is easy to see that $S_{\mathbf{C}}$ is an isometry on $V_{\mathbf{C}}$. Thus by the complex case that we have already done, we have $|\det S_{\mathbf{C}}| = 1$. By definition of the determinant on real vector spaces, we have $\det S = \det S_{\mathbf{C}}$ and thus $|\det S| = 1$, completing the proof.

Proof 2: By 9.36, there is an orthonormal basis of V with respect to which $\mathcal{M}(S)$ is 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

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

with $\theta \in (0, \pi)$. Note that the determinant of each 2-by-2 matrix of the form above equals 1 (because $\cos^2\theta + \sin^2\theta = 1$). Thus the determinant of S, which is the product of the determinants of the blocks (see Exercise 6), is the product of 1's and -1's. Hence, $|\det S| = 1$, as desired. ■

The Real Spectral Theorem 7.29 states that a self-adjoint operator T on a real inner product space has an orthonormal basis consisting of eigenvectors. With respect to such a basis, the number of times each eigenvalue appears on the diagonal of $\mathcal{M}(T)$ is its multiplicity. Thus $\det T$ equals the product of its eigenvalues, counting multiplicity (of course, this holds for every 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 7.35 and 7.36). Because $\sqrt{T^*T}$ is positive, all its eigenvalues are nonnegative (again, see 7.35), and hence $\det \sqrt{T^*T} \geq 0$. These considerations play a role in next example.

10.46 **Example** Suppose V is a real inner product space and $T \in \mathcal{L}(V)$ is invertible (and thus $\det T$ is either positive or negative). Attach a geometric meaning to the sign of $\det T$.

Solution First we consider an isometry $S \in \mathcal{L}(V)$. By 10.45, the determinant of S equals 1 or -1. Note that

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

is the eigenspace $E(-1, S)$. Thinking geometrically, we could say that this is the subspace on which S reverses direction. An examination of proof 2 of 10.45 shows that $\det S = 1$ if this subspace has even dimension and $\det S = -1$ if this subspace has odd dimension.

We are not formally defining the phrase "reverses direction" because these comments are meant only as an intuitive aid to our understanding.

Returning to our arbitrary invertible operator $T \in \mathcal{L}(V)$, by the Polar Decomposition (7.45) there is an isometry $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T}.$$

Now 10.44 tells us that

$$\det T = (\det S)(\det \sqrt{T^*T}).$$

The remarks just before this example pointed out that $\det \sqrt{T^*T} \geq 0$. Thus whether $\det T$ is positive or negative depends on whether $\det S$ is positive or negative. As we saw in the paragraph above, this depends on whether the subspace on which S reverses direction has even or odd dimension.

Because T is the product of S and an operator that never reverses direction (namely, $\sqrt{T^*T}$), we can reasonably say that whether $\det T$ is positive or negative depends on whether T reverses vectors an even or an odd number of times.

Volume

The next result will be a key tool in our investigation of volume. Recall that our remarks before Example 10.46 pointed out that $\det \sqrt{T^*T} \geq 0$.

10.47 $|\det T| = \det \sqrt{T^*T}$

Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then

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

Proof

By the Polar Decomposition (7.45), there is an isometry $S \in \mathcal{L}(V)$ such that

Another proof of this result is suggested in Exercise 8.

$$T = S\sqrt{T^*T}.$$

Thus

$$\begin{aligned}|\det T| &= |\det S| \det \sqrt{T^*T}\\ &= \det \sqrt{T^*T},\end{aligned}$$

where the first equality follows from 10.44 and the second equality follows from 10.45. ■

Now we turn to the question of volume in $\mathbf{R}^n$. Fix a positive integer n for the rest of this subsection. We will consider only the real inner product space $\mathbf{R}^n$, with its standard inner product.

We would like to assign to each subset Ω of $\mathbf{R}^n$ its n-dimensional volume (when $n = 2$, this is usually called area instead of volume). We begin with boxes, where we have a good intuitive notion of volume.

10.48 Definition ***box***

A ***box*** in $\mathbf{R}^n$ is a set of the form

$$\{(y_1, \ldots, y_n) \in \mathbf{R}^n : x_j < y_j < x_j + r_j \text{ for } j = 1, \ldots, n\},$$

where $r_1, \ldots, r_n$ are positive numbers and $(x_1, \ldots, x_n) \in \mathbf{R}^n$. The numbers $r_1, \ldots, r_n$ are called the ***side lengths*** of the box.

You should verify that when $n = 2$, a box is a rectangle with sides parallel to the coordinate axes, and that when $n = 3$, a box is a familiar 3-dimensional box with sides parallel to the coordinate axes.

The next definition fits with our intuitive notion of volume, because we define the volume of a box to be the product of the side lengths of the box.

10.49 Definition ***volume of a box***

The ***volume*** of a box B in $\mathbf{R}^n$ with side lengths $r_1, \dots, r_n$ is defined to be $r_1 \cdots r_n$ and is denoted by volume B.

Readers familiar with outer measure will recognize that concept here.

To define the volume of an arbitrary set $\Omega \subset \mathbf{R}^n$, the idea is to write Ω as a subset of a union of many small boxes, then add up the volumes of these small boxes. As we approximate Ω more accurately by unions of small boxes, we get a better estimate of volume Ω.

10.50 Definition ***volume***

Suppose $\Omega \subset \mathbf{R}^n$. Then the ***volume*** of Ω, denoted volume Ω, is defined to be the infimum of

$$\text{volume } B_1 + \text{volume } B_2 + \cdots,$$

where the infimum is taken over all sequences $B_1, B_2, \dots$ of boxes in $\mathbf{R}^n$ whose union contains Ω.

We will work only with an intuitive notion of volume. Our purpose in this book is to understand linear algebra, whereas notions of volume belong to analysis (although volume is intimately connected with determinants, as we will soon see). Thus for the rest of this section we will rely on intuitive notions of volume rather than on a rigorous development, although we shall maintain our usual rigor in the linear algebra parts of what follows. Everything said here about volume will be correct if appropriately interpreted—the intuitive approach used here can be converted into appropriate correct definitions, correct statements, and correct proofs using the machinery of analysis.

10.51 Notation $T(\Omega)$

For T a function defined on a set Ω, define $T(\Omega)$ by

$$T(\Omega) = \{Tx : x \in \Omega\}.$$

For $T \in \mathcal{L}(\mathbf{R}^n)$ and $\Omega \subset \mathbf{R}^n$, we seek a formula for volume $T(\Omega)$ in terms of T and volume Ω. We begin by looking at positive operators.

10.52 Positive operators change volume by factor of determinant

Suppose $T \in \mathcal{L}(\mathbf{R}^n)$ is a positive operator and $\Omega \subset \mathbf{R}^n$. Then

$$\text{volume}\, T(\Omega) = (\det T)(\text{volume}\, \Omega).$$

Proof To get a feeling for why this result is true, first consider the special case where $\lambda_1, \dots, \lambda_n$ are positive numbers and $T \in \mathcal{L}(\mathbf{R}^n)$ is defined by

$$T(x_1, \dots, x_n) = (\lambda_1 x_1, \dots, \lambda_n x_n).$$

This operator stretches the j^{th} standard basis vector by a factor of λ_j. If B is a box in $\mathbf{R}^n$ with side lengths $r_1, \dots, r_n$, then $T(B)$ is a box in $\mathbf{R}^n$ with side lengths $\lambda_1 r, \dots, \lambda_n r$. The box $T(B)$ thus has volume $\lambda_1 \cdots \lambda_n r_1 \cdots r_n$, whereas the box Ω has volume $r_1 \cdots r_n$. Note that $\det T = \lambda_1 \cdots \lambda_n$. Thus

$$\text{volume}\, T(B) = (\det T)(\text{volume}\, B)$$

for every box B in $\mathbf{R}^n$. Because the volume of Ω is approximated by sums of volumes of boxes, this implies that volume $T(\Omega) = (\det T)(\text{volume}\, \Omega)$.

Now consider an arbitrary positive operator $T \in \mathcal{L}(\mathbf{R}^n)$. By the Real Spectral Theorem (7.29), there exist an orthonormal basis $e_1, \dots, e_n$ of $\mathbf{R}^n$ and nonnegative numbers $\lambda_1, \dots, \lambda_n$ such that $Te_j = \lambda_j e_j$ for $j = 1, \dots, n$. In the special case where $e_1, \dots, e_n$ is the standard basis of $\mathbf{R}^n$, this operator is the same one as defined in the paragraph above. For an arbitrary orthonormal basis $e_1, \dots, e_n$, this operator has the same behavior as the one in the paragraph above—it stretches the j^{th} basis vector in an orthonormal basis by a factor of λ_j. Your intuition about volume should convince you that volume behaves the same with respect to each orthonormal basis. That intuition, and the special case of the paragraph above, should convince you that T multiplies volume by a factor of $\lambda_1 \cdots \lambda_n$, which again equals $\det T$. ■

Our next tool is the following result, which states that isometries do not change volume.

10.53 An isometry does not change volume

Suppose $S \in \mathcal{L}(\mathbf{R}^n)$ is an isometry and $\Omega \subset \mathbf{R}^n$. Then

$$\text{volume}\, S(\Omega) = \text{volume}\, \Omega.$$

Proof For $x, y \in \mathbf{R}^n$, we have

$$\begin{aligned}\|Sx - Sy\| &= \|S(x-y)\| \\ &= \|x-y\|.\end{aligned}$$

In other words, S does not change the distance between points. That property alone may be enough to convince you that S does not change volume.

However, if you need stronger persuasion, consider the complete description of isometries on real inner product spaces provided by 9.36. According to 9.36, S can be decomposed into pieces, each of which is the identity on some subspace (which clearly does not change volume) or multiplication by -1 on some subspace (which again clearly does not change volume) or a rotation on a 2-dimensional subspace (which again does not change volume). Or use 9.36 in conjunction with Exercise 7 in Section 9.B to write S as a product of operators, each of which does not change volume. Either way, you should be convinced that S does not change volume. ■

Now we can prove that an operator $T \in \mathcal{L}(\mathbf{R}^n)$ changes volume by a factor of $|\det T|$. Note the huge importance of the Polar Decomposition in the proof.

10.54 T changes volume by factor of $|\det T|$

Suppose $T \in \mathcal{L}(\mathbf{R}^n)$ and $\Omega \subset \mathbf{R}^n$. Then

$$\text{volume}\, T(\Omega) = |\det T|(\text{volume}\, \Omega).$$

Proof By the Polar Decomposition (7.45), there is an isometry $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T}.$$

If $\Omega \subset \mathbf{R}^n$, then $T(\Omega) = S\big(\sqrt{T^*T}(\Omega)\big)$. Thus

$$\begin{aligned}\text{volume}\, T(\Omega) &= \text{volume}\, S\big(\sqrt{T^*T}(\Omega)\big) \\ &= \text{volume}\, \sqrt{T^*T}(\Omega) \\ &= (\det \sqrt{T^*T})(\text{volume}\, \Omega) \\ &= |\det T|(\text{volume}\, \Omega),\end{aligned}$$

where the second equality holds because volume is not changed by the isometry S (by 10.53), the third equality holds by 10.52 (applied to the positive operator $\sqrt{T^*T}$), and the fourth equality holds by 10.47. ■

The result that we just proved leads to the appearance of determinants in the formula for change of variables in multivariable integration. To describe this, we will again be vague and intuitive.

Throughout this book, almost all the functions we have encountered have been linear. Thus please be aware that the functions f and σ in the material below are not assumed to be linear.

The next definition aims at conveying the idea of the integral; it is not intended as a rigorous definition.

10.55 Definition ***integral***, $\int_\Omega f$

If $\Omega \subset \mathbf{R}^n$ and f is a real-valued function on Ω, then the ***integral*** of f over Ω, denoted $\int_\Omega f$ or $\int_\Omega f(x)\,dx$, is defined by breaking Ω into pieces small enough that f is almost constant on each piece. On each piece, multiply the (almost constant) value of f by the volume of the piece, then add up these numbers for all the pieces, getting an approximation to the integral that becomes more accurate as Ω is divided into finer pieces.

Actually, Ω in the definition above needs to be a reasonable set (for example, open or measurable) and f needs to be a reasonable function (for example, continuous or measurable), but we will not worry about those technicalities. Also, notice that the x in $\int_\Omega f(x)\,dx$ is a dummy variable and could be replaced with any other symbol.

Now we define the notions of differentiable and derivative. Notice that in this context, the derivative is an operator, not a number as in one-variable calculus. The uniqueness of T in the definition below is left as Exercise 9.

10.56 Definition ***differentiable, derivative***, $\sigma'(x)$

Suppose Ω is an open subset of $\mathbf{R}^n$ and σ is a function from Ω to $\mathbf{R}^n$. For $x \in \Omega$, the function σ is called ***differentiable*** at x if there exists an operator $T \in \mathcal{L}(\mathbf{R}^n)$ such that

$$\lim_{y \to 0} \frac{\|\sigma(x+y) - \sigma(x) - Ty\|}{\|y\|} = 0.$$

If σ is differentiable at x, then the unique operator $T \in \mathcal{L}(\mathbf{R}^n)$ satisfying the equation above is called the ***derivative*** of σ at x and is denoted by $\sigma'(x)$.

If $n = 1$, then the derivative in the sense of the definition above is the operator on $\mathbf{R}$ of multiplication by the derivative in the usual sense of one-variable calculus.

The idea of the derivative is that for x fixed and $\|y\|$ small,

$$\sigma(x+y) \approx \sigma(x) + \big(\sigma'(x)\big)(y);$$

because $\sigma'(x) \in \mathcal{L}(\mathbf{R}^n)$, this makes sense.

Suppose Ω is an open subset of $\mathbf{R}^n$ and σ is a function from Ω to $\mathbf{R}^n$. We can write

$$\sigma(x) = \big(\sigma_1(x), \dots, \sigma_n(x)\big),$$

where each σ_j is a function from Ω to $\mathbf{R}$. The partial derivative of σ_j with respect to the k^{th} coordinate is denoted $D_k\sigma_j$. Evaluating this partial derivative at a point $x \in \Omega$ gives $D_k\sigma_j(x)$. If σ is differentiable at x, then the matrix of $\sigma'(x)$ with respect to the standard basis of $\mathbf{R}^n$ contains $D_k\sigma_j(x)$ in row j, column k (this is left as an exercise). In other words,

10.57
$$\mathcal{M}\big(\sigma'(x)\big) = \begin{pmatrix} D_1\sigma_1(x) & \dots & D_n\sigma_1(x) \\ \vdots & & \vdots \\ D_1\sigma_n(x) & \dots & D_n\sigma_n(x) \end{pmatrix}.$$

Now we can state the change of variables integration formula. Some additional mild hypotheses are needed for f and σ' (such as continuity or measurability), but we will not worry about them because the proof below is really a pseudoproof that is intended to convey the reason the result is true.

The result below is called a change of variables formula because you can think of $y = \sigma(x)$ as a change of variables, as illustrated by the two examples that follow the proof.

10.58 Change of variables in an integral

Suppose Ω is an open subset of $\mathbf{R}^n$ and $\sigma : \Omega \to \mathbf{R}^n$ is differentiable at every point of Ω. If f is a real-valued function defined on $\sigma(\Omega)$, then

$$\int_{\sigma(\Omega)} f(y)\,dy = \int_{\Omega} f\big(\sigma(x)\big)|\det \sigma'(x)|\,dx.$$

Proof Let $x \in \Omega$ and let Γ be a small subset of Ω containing x such that f is approximately equal to the constant $f\big(\sigma(x)\big)$ on the set $\sigma(\Gamma)$.

Adding a fixed vector [such as $\sigma(x)$] to each vector in a set produces another set with the same volume. Thus our approximation for σ near x using the derivative shows that

$$\text{volume}\,\sigma(\Gamma) \approx \text{volume}\big[\big(\sigma'(x)\big)(\Gamma)\big].$$

Using 10.54 applied to the operator $\sigma'(x)$, this becomes

$$\text{volume}\,\sigma(\Gamma) \approx |\det \sigma'(x)|(\text{volume}\,\Gamma).$$

Let $y = \sigma(x)$. Multiply the left side of the equation above by $f(y)$ and the right side by $f\big(\sigma(x)\big)$ [because $y = \sigma(x)$, these two quantities are equal], getting

$$f(y)\,\text{volume}\,\sigma(\Gamma) \approx f\big(\sigma(x)\big)|\det \sigma'(x)|(\text{volume}\,\Gamma).$$

Now break Ω into many small pieces and add the corresponding versions of the equation above, getting the desired result. ■

The key point when making a change of variables is that the factor of $|\det \sigma'(x)|$ must be included when making a substitution $y = f(x)$, as in the right side of 10.58. We finish up by illustrating this point with two important examples.

10.59 **Example** ***polar coordinates***

Define $\sigma : \mathbf{R}^2 \to \mathbf{R}^2$ by

$$\sigma(r, \theta) = (r\cos\theta, r\sin\theta),$$

where we have used r, θ as the coordinates instead of x_1, x_2 for reasons that will be obvious to everyone familiar with polar coordinates (and will be a mystery to everyone else). For this choice of σ, the matrix of partial derivatives corresponding to 10.57 is

$$\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix},$$

as you should verify. The determinant of the matrix above equals r, thus explaining why a factor of r is needed when computing an integral in polar coordinates.

For example, note the extra factor of r in the following familiar formula involving integrating a function f over a disk in $\mathbf{R}^2$:

$$\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} f(x, y)\,dy\,dx = \int_{0}^{2\pi}\int_{0}^{1} f(r\cos\theta, r\sin\theta)r\,dr\,d\theta.$$

10.60 Example ***spherical coordinates***

Define $\sigma\colon \mathbf{R}^3 \to \mathbf{R}^3$ by

$$\sigma(\rho, \varphi, \theta) = (\rho \sin\varphi \cos\theta, \rho \sin\varphi \sin\theta, \rho \cos\varphi),$$

where we have used ρ, θ, φ as the coordinates instead of x_1, x_2, x_3 for reasons that will be obvious to everyone familiar with spherical coordinates (and will be a mystery to everyone else). For this choice of σ, the matrix of partial derivatives corresponding to 10.57 is

$$\begin{pmatrix} \sin\varphi \cos\theta & \rho\cos\varphi\cos\theta & -\rho\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & \rho\cos\varphi\sin\theta & \rho\sin\varphi\cos\theta \\ \cos\varphi & -\rho\sin\varphi & 0 \end{pmatrix},$$

as you should verify. The determinant of the matrix above equals $\rho^2 \sin\varphi$, thus explaining why a factor of $\rho^2 \sin\varphi$ is needed when computing an integral in spherical coordinates.

For example, note the extra factor of $\rho^2 \sin\varphi$ in the following familiar formula involving integrating a function f over a ball in $\mathbf{R}^3$:

$$\int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} f(x, y, z)\, dz\, dy\, dx$$

$$= \int_{0}^{2\pi} \int_{0}^{\pi} \int_{0}^{1} f(\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi)\rho^2 \sin\varphi\, d\rho\, d\varphi\, d\theta.$$

EXERCISES 10.B

1 Suppose V is a real vector space. Suppose $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that $\det T > 0$.

2 Suppose V is a real vector space with even dimension and $T \in \mathcal{L}(V)$. Suppose $\det T < 0$. Prove that T has at least two distinct eigenvalues.

3 Suppose $T \in \mathcal{L}(V)$ and $n = \dim V > 2$. Let $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of T (or of $T_{\mathbf{C}}$ if V is a real vector space), repeated according to multiplicity.

(a) Find a formula for the coefficient of z^{n-2} in the characteristic polynomial of T in terms of $\lambda_1, \dots, \lambda_n$.

(b) Find a formula for the coefficient of z in the characteristic polynomial of T in terms of $\lambda_1, \dots, \lambda_n$.

4 Suppose $T \in \mathcal{L}(V)$ and $c \in \mathbf{F}$. Prove that $\det(cT) = c^{\dim V} \det T$.

5 Prove or give a counterexample: if $S, T \in \mathcal{L}(V)$, then $\det(S + T) = \det S + \det T$.

6 Suppose A is a block upper-triangular matrix

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

where each A_j along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \cdots (\det A_m).$$

7 Suppose A is an n-by-n matrix with real entries. Let $S \in \mathcal{L}(\mathbf{C}^n)$ denote the operator on $\mathbf{C}^n$ whose matrix equals A, and let $T \in \mathcal{L}(\mathbf{R}^n)$ denote the operator on $\mathbf{R}^n$ whose matrix equals A. Prove that $\operatorname{trace} S = \operatorname{trace} T$ and $\det S = \det T$.

8 Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\det T^* = \overline{\det T}.$$

Use this to prove that $|\det T| = \det \sqrt{T^*T}$, giving a different proof than was given in 10.47.

9 Suppose Ω is an open subset of $\mathbf{R}^n$ and σ is a function from Ω to $\mathbf{R}^n$. Suppose $x \in \Omega$ and σ is differentiable at x. Prove that the operator $T \in \mathcal{L}(\mathbf{R}^n)$ satisfying the equation in 10.56 is unique.
[*This exercise shows that the notation $\sigma'(x)$ is justified.*]

10 Suppose $T \in \mathcal{L}(\mathbf{R}^n)$ and $x \in \mathbf{R}^n$. Prove that T is differentiable at x and $T'(x) = T$.

11 Find a suitable hypothesis on σ and then prove 10.57.

12 Let a, b, c be positive numbers. Find the volume of the ellipsoid

$$\left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}$$

by finding a set $\Omega \subset \mathbf{R}^3$ whose volume you know and an operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that $T(\Omega)$ equals the ellipsoid above.