Tom M. Apostol

CALCULUS

VOLUME II

Multi Variable Calculus and Linear Algebra, with Applications to Differential Equations and Probability

SECOND EDITION

John Wiley & Sons

New York London Sydney Toronto

CONSULTING EDITOR

George Springer, *Indiana University*

Library of Congress Catalog Card Number: 67-14605
ISBN 0 471 00007 8 Printed in the United States of America.

1 0 9 8 7 6 5 4 3 2

To

Jane and Stephen

PREFACE

This book is a continuation of the author's ***Calculus, Volume I, Second Edition.*** The present volume has been written with the same underlying philosophy that prevailed in the first. Sound training in technique is combined with a strong theoretical development. Every effort has been made to convey the spirit of modern mathematics without undue emphasis on formalization. As in Volume I, historical remarks are included to give the student a sense of participation in the evolution of ideas.

The second volume is divided into three parts, entitled ***Linear Analysis, Nonlinear Analysis***, and ***Special Topics.*** The last two chapters of Volume I have been repeated as the first two chapters of Volume II so that all the material on linear algebra will be complete in one volume.

Part 1 contains an introduction to linear algebra, including linear transformations, matrices, determinants, eigenvalues, and quadratic forms. Applications are given to analysis, in particular to the study of linear differential equations. Systems of differential equations are treated with the help of matrix calculus. Existence and uniqueness theorems are proved by Picard's method of successive approximations, which is also cast in the language of contraction operators.

Part 2 discusses the calculus of functions of several variables. Differential calculus is unified and simplified with the aid of linear algebra. It includes chain rules for scalar and vector fields, and applications to partial differential equations and extremum problems. Integral calculus includes line integrals, multiple integrals, and surface integrals, with applications to vector analysis. Here the treatment is along more or less classical lines and does not include a formal development of differential forms.

The special topics treated in Part ***3*** are ***Probability*** and ***Numerical Analysis.*** The material on probability is divided into two chapters, one dealing with finite or countably infinite sample spaces; the other with uncountable sample spaces, random variables, and distribution functions. The use of the calculus is illustrated in the study of both one- and two-dimensional random variables.

The last chapter contains an introduction to numerical analysis, the chief emphasis being on different kinds of polynomial approximation. Here again the ideas are unified by the notation and terminology of linear algebra. The book concludes with a treatment of approximate integration formulas, such as Simpson's rule, and a discussion of Euler's summation formula.

There is ample material in this volume for a full year's course meeting three or four times per week. It presupposes a knowledge of one-variable calculus as covered in most first-year calculus courses. The author has taught this material in a course with two lectures and two recitation periods per week, allowing about ten weeks for each part and omitting the starred sections.

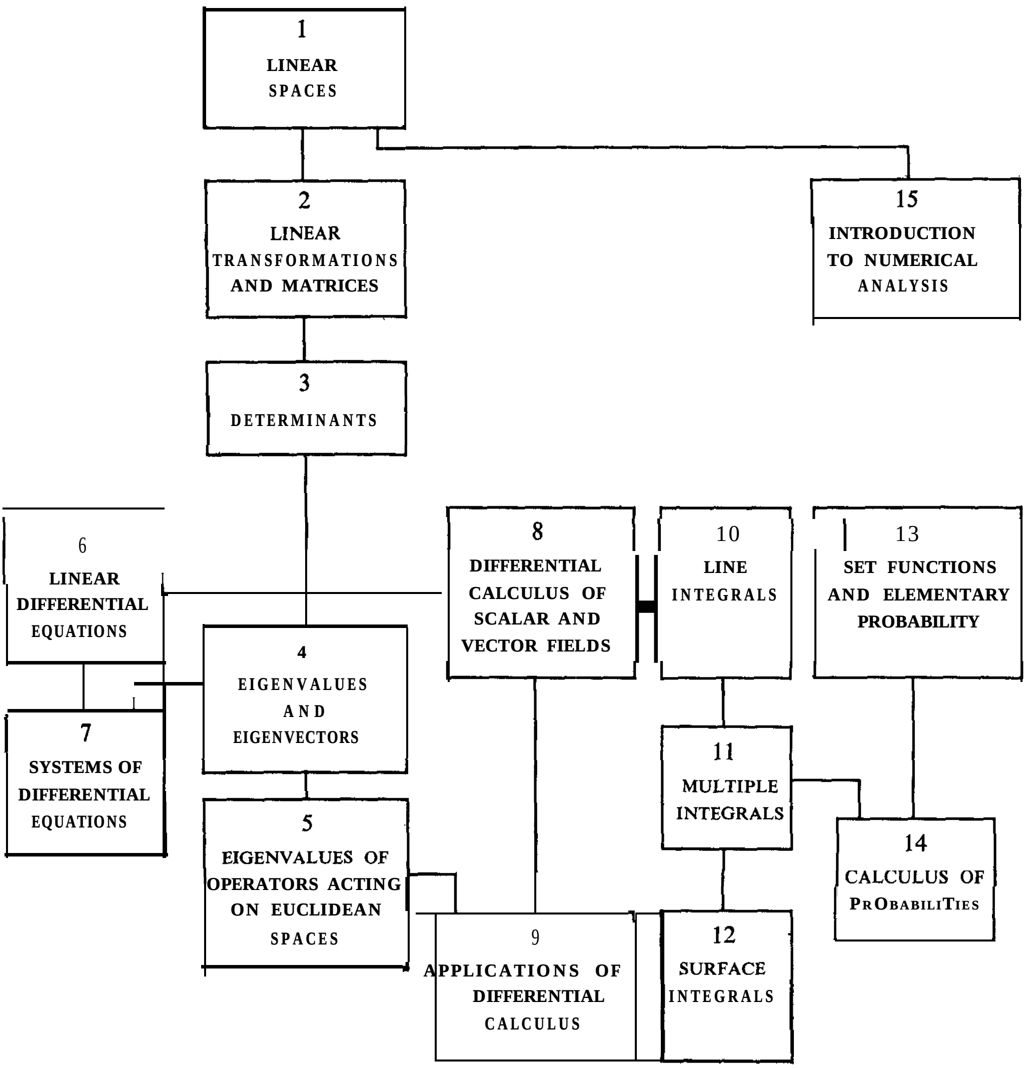
This second volume has been planned so that many chapters can be omitted for a variety of shorter courses. For example, the last chapter of each part can be skipped without disrupting the continuity of the presentation. Part 1 by itself provides material for a combined course in linear algebra and ordinary differential equations. The individual instructor can choose topics to suit his needs and preferences by consulting the diagram on the next page which shows the logical interdependence of the chapters.

Once again I acknowledge with pleasure the assistance of many friends and colleagues. In preparing the second edition I received valuable help from Professors Herbert S. Zuckerman of the University of Washington, and Basil Gordon of the University of California, Los Angeles, each of whom suggested a number of improvements. Thanks are also due to the staff of Blaisdell Publishing Company for their assistance and cooperation.

As before, it gives me special pleasure to express my gratitude to my wife for the many ways in which she has contributed. In grateful acknowledgement I happily dedicate this book to her.

T. M. A.

Pasadena, California
September 16, 1968


1
LINEAR SPACES
2
LINEAR TRANSFORMATIONS AND MATRICES
3
DETERMINANTS
4
EIGENVALUES AND EIGENVECTORS
5
EIGENVALUES OF OPERATORS ACTING ON EUCLIDEAN SPACES
6
LINEAR DIFFERENTIAL EQUATIONS
7
SYSTEMS OF DIFFERENTIAL EQUATIONS
8
DIFFERENTIAL CALCULUS OF SCALAR AND VECTOR FIELDS
9
APPLICATIONS OF DIFFERENTIAL CALCULUS
10
LINE INTEGRALS
11
MULTIPLE INTEGRALS
12
SURFACE INTEGRALS
13
SET FUNCTIONS AND ELEMENTARY PROBABILITY
14
CALCULUS OF PROBABILITIES
15
INTRODUCTION TO NUMERICAL ANALYSIS

CONTENTS

PART 1. LINEAR ANALYSIS

1. LINEAR SPACES

2. LINEAR TRANSFORMATIONS AND MATRICES

3. DETERMINANTS

4. EIGENVALUES AND EIGENVECTORS

5. EIGENVALUES OF OPERATORS ACTING ON EUCLIDEAN SPACES

6. LINEAR DIFFERENTIAL EQUATIONS

7. SYSTEMS OF DIFFERENTIAL EQUATIONS

PART 2. NONLINEAR ANALYSIS

8. DIFFERENTIAL CALCULUS OF SCALAR AND VECTOR FIELDS

9. APPLICATIONS OF THE DIFFERENTIAL CALCULUS

10. LINE INTEGRALS

11. MULTIPLE INTEGRALS

12. SURFACE INTEGRALS

PART 3. SPECIAL TOPICS

13. SET FUNCTIONS AND ELEMENTARY PROBABILITY

14. CALCULUS OF PROBABILITIES

15. INTRODUCTION TO NUMERICAL ANALYSIS

Calculus

PART 1

LINEAR ANALYSIS

1

LINEAR SPACES

1.1 Introduction

Throughout mathematics we encounter many examples of mathematical objects that can be added to each other and multiplied by real numbers. First of all, the real numbers themselves are such objects. Other examples are real-valued functions, the complex numbers, infinite series, vectors in n-space, and vector-valued functions. In this chapter we discuss a general mathematical concept, called a *linear space,* which includes all these examples and many others as special cases.

Briefly, a linear space is a set of elements of any kind on which certain operations (called *addition* and *multiplication by numbers)* can be performed. In defining a linear space, we do not specify the nature of the elements nor do we tell how the operations are to be performed on them. Instead, we require that the operations have certain properties which we take as axioms for a linear space. We turn now to a detailed description of these axioms.

1.2 The definition of a linear space

Let V denote a nonempty set of objects, called *elements*. The set V is called a linear space if it satisfies the following ten axioms which we list in three groups.

Closure axioms

AXIOM 1. CLOSURE UNDER ADDITION. *For every pair of elements x and y in V there corresponds a unique element in V called the sum of x and y, denoted by $x + y$.*

AXIOM 2. CLOSURE UNDER MULTIPLICATION BY REAL NUMBERS. *For every x in V and every real number a there corresponds an element in V called the product of a and x, denoted by ax.*

Axioms for addition

AXIOM 3. COMMUTATIVE LAW. *For all x and y in V, we have $x + y = y + x$.*

AXIOM 4. ASSOCIATIVE LAW. *For all x, y, and z in V, we have $(x + y) + z = x + (y + z)$.*

AXIOM 5. EXISTENCE OF ZERO ELEMENT. *There is an element in V, denoted by 0, such that*

$$x + O = x \text{ for all } x \text{ in } V \quad .$$

AXIOM 6. EXISTENCE OF NEGATIVES. *For every x in V, the element $(-1)x$ has the property*

$$x + (-1)x = 0 \, .$$

Axioms for multiplication by numbers

AXIOM 7. ASSOCIATIVE LAW. *For every x in V and all real numbers a and b, we have*

$$a(bx) = (ab)x.$$

AXIOM 8. DISTRIBUTIVE LAW FOR ADDITION IN V. *For all x andy in V and all real a, we hare*

$$a(x + y) = ax + ay \, .$$

AXIOM 9. DISTRIBUTIVE LAW FOR ADDITION OF NUMBERS. *For all x in V and all real a and b, we have*

$$(a + b)x = ax + bx.$$

AXIOM 10. EXISTENCE OF IDENTITY. *For every x in V, we have* $1x = x$.

Linear spaces, as defined above, are sometimes called ***real*** linear spaces to emphasize the fact that we are multiplying the elements of V by real numbers. If ***real number*** is replaced by ***complex number*** in Axioms 2, 7, 8, and 9, the resulting structure is called a ***complex linear space.*** Sometimes a linear space is referred to as a ***linear vector space*** or simply a ***vector space;*** the numbers used as multipliers are also called ***scalars.*** A real linear space has real numbers as scalars; a complex linear space has complex numbers as scalars. Although we shall deal primarily with examples of real linear spaces, all the theorems are valid for complex linear spaces as well. When we use the term linear space without further designation, it is to be understood that the space can be real or complex.

1.3 Examples of linear spaces

If we specify the set V and tell how to add its elements and how to multiply them by numbers, we get a concrete example of a linear space. The reader can easily verify that each of the following examples satisfies all the axioms for a real linear space.

EXAMPLE 1. Let $V = R$, the set of all real numbers, and let $x + y$ and ax be ordinary addition and multiplication of real numbers.

EXAMPLE 2. Let $V = C$, the set of all complex numbers, define $x + y$ to be ordinary addition of complex numbers, and define ax to be multiplication of the complex number x

by the real number a. Even though the elements of V are complex numbers, this is a real linear space because the scalars are real.

EXAMPLE 3. Let $V = V_n$, the vector space of all n-tuples of real numbers, with addition and multiplication by scalars defined in the usual way in terms of components.

EXAMPLE 4. Let V be the set of all vectors in V_n orthogonal to a given nonzero vector N. If $n = 2$, this linear space is a line through 0 with N as a normal vector. If $n = 3$, it is a plane through 0 with N as normal vector.

The following examples are called *function spaces*. The elements of V are real-valued functions, with addition of two functions f and g defined in the usual way:

$$(f + g)(x) = f(x) + g(x)$$

for every real x in the intersection of the domains off and g. Multiplication of a function f by a real scalar a is defined as follows: af is that function whose value at each x in the domain off is af (x). The zero element is the function whose values are everywhere zero. The reader can easily verify that each of the following sets is a function space.

EXAMPLE 5. The set of all functions defined on a given interval.

EXAMPLE 6. The set of all polynomials.

EXAMPLE 7. The set of all polynomials of degree $\leq n$, where n is fixed. (Whenever we consider this set it is understood that the zero polynomial is also included.) The set of all polynomials of degree *equal* to n is not a linear space because the closure axioms are not satisfied. For example, the sum of two polynomials of degree n need not have degree n.

EXAMPLE 8. The set of all functions continuous on a given interval. If the interval is $[a, b]$, we denote this space by $C(a, b)$.

EXAMPLE 9. The set of all functions differentiable at a given point.

EXAMPLE 10. The set of all functions integrable on a given interval.

EXAMPLE 11. The set of all functions f defined at 1 with $f(1) = 0$. The number 0 is essential in this example. If we replace 0 by a nonzero number c, we violate the closure axioms.

EXAMPLE 12. The set of all solutions of a homogeneous linear differential equation $y'' + ay' + by = 0$, where a and b are given constants. Here again 0 is essential. The set of solutions of a nonhomogeneous differential equation does not satisfy the closure axioms.

These examples and many others illustrate how the linear space concept permeates algebra, geometry, and analysis. When a theorem is deduced from the axioms of a linear space, we obtain, in one stroke, a result valid for each concrete example. By unifying diverse examples in this way we gain a deeper insight into each. Sometimes special knowledge of one particular example helps to anticipate or interpret results valid for other examples and reveals relationships which might otherwise escape notice.

1.4 Elementary consequences of the axioms

The following theorems are easily deduced from the axioms for a linear space.

THEOREM 1.1. UNIQUENESS OF THE ZERO ELEMENT. ***In any linear space there is one and only one zero element.***

Proof. Axiom 5 tells us that there is at least one zero element. Suppose there were two, say 0, and 0,. Taking x = O_1 and 0 = 0, in Axiom 5, we obtain $O_1 + O_2 = O_1$. Similarly, taking x = 02 and 0 = O_1, we find 02 + 0, = 02. But O_1 + 02 = 02 + 0, because of the commutative law, so 0, = 02.

THEOREM **1.2.** UNIQUENESS OF NEGATIVE ELEMENTS. ***In any linear space every element has exactly one negative. That is, for every x there is one and only one y such that x + y = 0.***

Proof. Axiom 6 tells us that each x has at least one negative, namely (- 1)x. Suppose x has two negatives, say y_1 and y_2. Then x + y_1 = 0 and x + y_2 = 0. Adding y_2 to both members of the first equation and using Axioms 5, 4, and 3, we find that

$$y_2 + (x + y_1) = y_2 + O = y_2,$$

and

$$y_2 + (x + y_1) = (y_2 + x) + y_1 = 0 + y_1 = y_1 + O = y_1.$$

Therefore $y_1 = y_2$, so x has exactly one negative, the element (- 1)x.

Notation. The negative of x is denoted by $-x$. The difference y - x is defined to be the sum y + (-x) .

The next theorem describes a number of properties which govern elementary algebraic manipulations in a linear space.

THEOREM 1.3. ***In a given linear space, let x and y denote arbitrary elements and let a and b denote arbitrary scalars. Then we'have the following properties:***

(a) Ox = 0.

(b) ***a0 = 0.***

(c) (-a)x = -(ax) = a(-x).

(d) ***Ifax=O,theneithera=Oorx=O.***

(e) *If $ax = ay$ and $a \neq 0$,* ***thenx=y.***

(f) *If $ax = bx$ and $x \neq O$, then $a = b$.*

(g) $-(x + y) = (-x) + (-y) = -x - y.$

(h) x + x = $2x$, ***x + x + x = 3x, andingeneral,*** $\sum_{i=1}^{n} x = nx$.

We shall prove (a), (b), and (c) and leave the proofs of the other properties as exercises.

Proof of (a). Let $z = 0x$. We wish to prove that $z = 0$. Adding z to itself and using Axiom 9, we find that

$$z + z = 0x + 0x = (0 + 0)x = 0x = z.$$

Now add $-z$ to both members to get $z = 0$.

Proof of (b). Let $z = aO$, add z to itself, and use Axiom 8.

Proof of (c). Let $z = (-a)x$. Adding z to ax and using Axiom 9, we find that

$$z + ax = (-a)x + ax = (-a + a)x = 0x = O,$$

so z is the negative of ax, $z = -(ax)$. Similarly, if we add $a(-x)$ to ax and use Axiom 8 and property (b), we find that $a(-x) = -(ax)$.

1.5 Exercises

In Exercises 1 through 28, determine whether each of the given sets is a real linear space, if addition and multiplication by real scalars are defined in the usual way. For those that are not, tell which axioms fail to hold. The functions in Exercises 1 through 17 are real-valued. In Exercises 3, 4, and 5, each function has domain containing 0 and 1. In Exercises 7 through 12, each domain contains all real numbers.

1. All rational functions.
2. All rational functions f/g, with the degree of $f \leq$ the degree of g (including $f = 0$).
3. All f with $f(0) = f(1)$.
4. All f with $2f(0) = f(1)$.
5. All f with $f(1) = 1 + f(0)$.
6. All step functions defined on [0, 1].
7. All f with $f(x) \to 0$ as $x \to +\infty$.
8. All even functions.
9. All odd functions.
10. All bounded functions.
11. All increasing functions.
12. All functions with period 2π.
13. All f integrable on [0, 1] with $\int_0^1 f(x)\,dx = 0$.
14. All f integrable on [0, 1] with $\int_0^1 f(x)\,dx \geq 0$.
15. All f satisfying $f(x) = f(1 - x)$ for all x.
16. All Taylor polynomials of degree $\leq n$ for a fixed n (including the zero polynomial).
17. All solutions of a linear second-order homogeneous differential equation' $y'' + P(x)y' + Q(x)y = 0$, where P and Q are given functions, continuous everywhere.
18. All bounded real sequences.
19. All convergent real sequences.
20. All convergent real series.
21. All absolutely convergent real series.
22. All vectors (x, y, z) in V_3 with $z = 0$.
23. All vectors (x, y, z) in V_3 with $x = 0$ or $y = 0$.
24. All vectors (x, y, z) in V_3 with $y = 5x$.
25. All vectors (x, y, z) in V_3 with $3x + 4y = 1$, $z = 0$.
26. All vectors (x, y, z) in V_3 which are scalar multiples of $(1, 2, 3)$.
27. All vectors (x, y, z) in V_3 whose components satisfy a system of three linear equations of the form :

$$a_{11}x + a_{12}y + a_{13}z = 0, \qquad a_{21}x + a_{22}y + a_{23}z = 0, \qquad a_{31}x + a_{32}y + a_{33}z = 0.$$

28. All vectors in V_n that are linear combinations of two given vectors A and B.
29. Let $V = \mathbf{R}^+$, the set of positive real numbers. Define the "sum" of two elements x and y in V to be their product $x\,y$ (in the usual sense), and define "multiplication" of an element x in V by a scalar c to be x^c. Prove that V is a real linear space with 1 as the zero element.
30. (a) Prove that Axiom 10 can be deduced from the other axioms.
    (b) Prove that Axiom 10 cannot be deduced from the other axioms if Axiom 6 is replaced by Axiom 6': For every x in V there is an element y in V such that $x + y = 0$.
31. Let S be the set of all ordered pairs (x_1, x_2) of real numbers. In each case determine whether or not S is a linear space with the operations of addition and multiplication by scalars defined as indicated. If the set is not a linear space, indicate which axioms are violated.
    (a) $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$, $\quad a(x_1, x_2) = (ax_1, 0)$.
    (b) $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, 0)$, $\quad a(x_1, x_2) = (ax_1, ax_2)$.
    (c) $(x_1, x_2) + (y_1, y_2) = (x_1, x_2 + y_2)$, $\quad a(x_1, x_2) = (ax_1, ax_2)$.
    (d) $(x_1, x_2) + (y_1, y_2) = (|x_1 + x_2|, |y_1 + y_2|)$, $\quad a(x_1, x_2) = (|ax_1|, |ax_2|)$.
32. Prove parts (d) through (h) of Theorem 1.3.

1.6 Subspaces of a linear space

Given a linear space V, let S be a nonempty subset of V. If S is also a linear space, with the same operations of addition and multiplication by scalars, then S is called a *subspace* of V. The next theorem gives a simple criterion for determining whether or not a subset of a linear space is a subspace.

THEOREM 1.4. ***Let S be a nonempty subset of a linear space V. Then S is a subspace if and only if S satisfies the closure axioms.***

Proof. If S is a subspace, it satisfies all the axioms for a linear space, and hence, in particular, it satisfies the closure axioms.

Now we show that if S satisfies the closure axioms it satisfies the others as well. The commutative and associative laws for addition (Axioms 3 and 4) and the axioms for multiplication by scalars (Axioms 7 through 10) are automatically satisfied in S because they hold for all elements of V. It remains to verify Axioms 5 and 6, the existence of a zero element in S, and the existence of a negative for each element in S.

Let x be any element of S. (S has at least one element since S is not empty.) By Axiom 2, ax is in S for every scalar a. Taking $a = 0$, it follows that $0x$ is in S. But $0x = O$, by Theorem 1.3(a), so $O \in S$, and Axiom 5 is satisfied. Taking $a = -1$, we see that $(-1)x$ is in S. But $x + (-1)x = O$ since both x and $(-1)x$ are in V, so Axiom 6 is satisfied in S. Therefore S is a subspace of V.

DEFINITION. ***Let S be a nonempty subset of a linear space V. An element x in V of the form***

$$x = \sum_{i=1}^{k} c_i x_i ,$$

where $x_1, \ldots, x_k$ are all in S and $c_1, \ldots, c_k$ are scalars, is called a finite linear combination of elements of S. The set of all finite linear combinations of elements of S satisfies the closure axioms and hence is a subspace of V. We call this the subspace spanned by S, or the linear span of S, and denote it by L(S). If S is empty, we define L(S) to be {O}, the set consisting of the zero element alone.

Different sets may span the same subspace. For example, the space V_2 is spanned by each of the following sets of vectors: $\{i, j\}$, $\{i, j, i + j\}$, $\{O, i, -i, j, -j, i + j\}$. The space of all polynomials $p(t)$ of degree $\leq n$ is spanned by the set of $n + 1$ polynomials

$$\{1, t, t^2, \ldots, t^n\}.$$

It is also spanned by the set $\{1, t/2, t^2/3, \ldots, t^n/(n + 1)\}$, *and by* $\{1, (1 + t), (1 + t)^2, \ldots, (1 + t)^n\}$. The space of all polynomials is spanned by the infinite set of polynomials $\{1, t, t^2, \ldots\}$.

A number of questions arise naturally at this point. For example, which spaces can be spanned by a finite set of elements? If a space can be spanned by a finite set of elements, what is the smallest number of elements required? To discuss these and related questions, we introduce the concepts of ***dependence, independence, bases,*** and ***dimension.*** These ideas were encountered in Volume I in our study of the vector space V_n. Now we extend them to general linear spaces.

1.7 Dependent and independent sets in a linear space

DEFINITION. *A set S of elements in a linear space V is called dependent if there is a finite set of distinct elements in S, say $x_1, \ldots, x_k$, and a corresponding set of scalars $c_1, \ldots, c_k$, not all zero, such that*

$$\sum_{i=1}^{k} c_i x_i = O.$$

An equation $\sum c_i x_i = 0$ with not all $c_i = 0$ is said to be a nontrivial representation of 0. The set S is called independent if it is not dependent. In this case, for all choices of distinct elements $x_1, \ldots, x_k$ in S and scalars $c_1, \ldots, c_k$,

$$\sum_{i=1}^{k} c_i x_i = O \qquad \textit{implies} \qquad c_1 = c_2 = \cdots = c_k = 0.$$

Although dependence and independence are properties of sets of elements, we also apply these terms to the elements themselves. For example, the elements in an independent set are called independent elements.

If S is a finite set, the foregoing definition agrees with that given in Volume I for the space V_n. However, the present definition is not restricted to finite sets.

EXAMPLE 1. If a subset T of a set S is dependent, then S itself is dependent. This is logically equivalent to the statement that every subset of an independent set is independent.

EXPMPLE 2. If one element in S is a scalar multiple of another, then S is dependent.

EXAMPLE 3. If $0 \in S$, then S is dependent.

EXAMPLE 4. The empty set is independent, Many examples of dependent and independent sets of vectors in V_n were discussed in Volume I. The following examples illustrate these concepts in function spaces. In each case the underlying linear space V is the set of all real-valued functions defined on the real line.

EXAMPLE 5. Let $u_1(t) = \cos^2 t$, $u_2(t) = \sin^2 t$, $u_3(t) = 1$ for all real t. The Pythagorean identity shows that $u_1 + u_2 - u_3 = 0$, so the three functions u_1, u_2, u_3 are dependent.

EXAMPLE 6. Let $u_k(t) = t^k$ for $k = 0, 1, 2, \ldots$, and t real. The set $S = \{u_0, u_1, u_2, \ldots\}$ is independent. To prove this, it suffices to show that for each n the $n + 1$ polynomials $u_0, u_1, \ldots, u_n$ are independent. A relation of the form $\sum c_k u_k = 0$ means that

$$\sum_{k=O}^{n} c_k t^k = 0 \tag{1.1}$$

for all real t. When $t = 0$, this gives $c_0 = 0$. Differentiating (1.1) and setting $t = 0$, we find that $c_1 = 0$. Repeating the process, we find that each coefficient c_k is zero.

EXAMPLE 7. If a,,..., a, are distinct real numbers, the n exponential functions

$$u_1(x) = e^{a_1 x}, \ldots, u,(x) = e^{a_n x}$$

are independent. We can prove this by induction on n. The result holds trivially when $n = 1$. Therefore, assume it is true for $n - 1$ exponential functions and consider scalars $c_1, \ldots, c_n$ such that

$$\sum_{k=1}^{n} c_k e^{a_k x} = 0. \tag{1.2}$$

Let a_M be the largest of the n numbers a,, . . . , a,. Multiplying both members of (1.2) by $e^{-a_M x}$, we obtain

$$\sum_{k=1}^{n} c_k e^{(a_k - a_M)x} = 0. \tag{1.3}$$

If $k \neq M$, the number $a_k - a_M$ is negative. Therefore, when x → + co in Equation (1.3), each term with $k \neq M$ tends to zero and we find that $c_M = 0$. Deleting the Mth term from (1.2) and applying the induction hypothesis, we find that each of the remaining $n - 1$ coefficients c_k is zero.

THEOREM 1.5. *Let $S = \{x_1, \ldots, x_k\}$ be an independent set consisting of k elements in a linear space V and let L(S) be the subspace spanned by S. Then every set of $k + 1$ elements in L(S) is dependent.*

Proof. The proof is by induction on k, the number of elements in S. First suppose $k = 1$. Then, by hypothesis, S consists of one element x_1, where $x_1 \neq 0$ since S is independent. Now take any two distinct elements y_1 and y_2 in $L(S)$. Then each is a scalar multiple of x_1, say $y_1 = c_1x_1$ and $y_2 = c_2x_1$, where c_1 and c_2 are not both 0. Multiplying y_1 by c_2 and y_2 by c_1 and subtracting, we find that

$$c_2y_1 - c_1y_2 = 0.$$

This is a nontrivial representation of 0 soy, and y_2 are dependent. This proves the theorem when $k = 1$.

Now we assume the theorem is true for $k - 1$ and prove that it is also true for k. Take any set of $k + 1$ elements in L(S), say $T = \{y_1, y_2, \ldots, y_{k+1}\}$. We wish to prove that Tis dependent. Since each y_i is in $L(S)$ we may write

$$y_i = \sum_{j=1}^{k} a_{ij}x_j \tag{1.4}$$

foreachi= $1, 2, \ldots, k + 1$. We examine all the scalars a_{i1} that multiply x_1 and split the proof into two cases according to whether all these scalars are 0 or not.

CASE 1. $a_{i1} = 0$ *for every* $i = 1, 2, \ldots, k + 1$. In this case the sum in (1.4) does not involve x_1, so each y_i in T is in the linear span of the set $S' = \{x_2, \ldots, x_k\}$. But S' is independent and consists of $k - 1$ elements. By the induction hypothesis, the theorem is true for $k - 1$ so the set T is dependent. This proves the theorem in Case 1.

CASE 2. Not all the scalars a_{i1} *are zero.* Let us assume that $a_{11} \neq 0$. (If necessary, we can renumber the y's to achieve this.) Taking $i = 1$ in Equation (1.4) and multiplying both members by c_i, where $c_i = a_{i1}/a_{11}$, *we* get

$$c_iy_1 = a_{i1}x_1 + \sum_{j=2}^{k} c_ia_{1j}x_j .$$

From this we subtract Equation (1.4) to get

$$c_iy_1 - y_i = \sum_{j=2}^{k}(c_ia_{1j} - a_{ij})x_j ,$$

for $i = 2, \ldots, k + 1$. This equation expresses each of the k elements $c_iy_1 - y_i$ as a linear combination of the $k - 1$ independent elements $x_2, \ldots, x_k$. By the induction hypothesis, the k elements $c_iy_1 - y_i$ must be dependent. Hence, for some choice of scalars $t_2, \ldots, t_{k+1}$, not all zero, we have

$$\sum_{i=2}^{k+1} t_i(c_iy_1 - y_i) = O ,$$

from which we find

$$\left(\sum_{i=2}^{k+1} t_ic_i\right) y_1 - \sum_{i=2}^{k+1} t_iy_i = O .$$

But this is a nontrivial linear combination of $y_1, \ldots, y_{k+1}$ which represents the zero element, so the elements $y_1, \ldots, y_{k+1}$ must be dependent. This completes the proof.

1.8 Bases and dimension

DEFINITION. *A jinite set S of elements in a linear space V is called a finite basis for V if S is independent and spans V. The space V is called finite-dimensional if it has a jinite basis, or if V consists of 0 alone. Otherwise, V is called injinite-dimensional.*

THEOREM 1.6. *Let V be a finite-dimensional linear space. Then every finite basis for V has the same number of elements.*

Proof. Let S and T be two finite bases for V. Suppose S consists of k elements and T consists of m elements. Since S is independent and spans V, Theorem 1.5 tells us that every set of $k + 1$ elements in $V$is dependent. Therefore, every set of more thank elements in V is dependent. Since T is an independent set, we must have $m \leq k$. The same argument with S and T interchanged shows that $k \leq m$. Therefore $k = m$.

DEFINITION. *If a linear space V has a basis of n elements, the integer n is called the dimension of V. We write $n = \dim V$. If $V = \{O\}$, we say V has dimension 0.*

EXAMPLE 1. The space V_n has dimension n. One basis is the set of n unit coordinate vectors.

EXAMPLE 2. The space of all polynomials $p(t)$ of degree $\leq n$ has dimension $n + 1$. One basis is the set of $n + 1$ polynomials $\{1, t, t^2, \ldots, t^n\}$. Every polynomial of degree $\leq n$ is a linear combination of these $n + 1$ polynomials.

EXAMPLE 3. The space of solutions of the differential equation $y'' - 2y' - 3y = 0$ has dimension 2. One basis consists of the two functions $u_1(x) = e^{-x}$, $u_2(x) = e^{3x}$. Every solution is a linear combination of these two.

EXAMPLE 4. The space of all polynomials $p(t)$ is infinite-dimensional. Although the infinite set $\{1, t, t^2, \ldots\}$ spans this space, no *finite* set of polynomials spans the space.

THEOREM 1.7. *Let V be a jinite-dimensional linear space with $\dim V = n$. Then we have the following:*

(a) *Any set of independent elements in V is a subset of some basis for V.*
(b) *Any set of n independent elements is a basis for V.*

Proof. To prove (a), let $S = \{x_1, \ldots, x_k\}$ be any independent set of elements in V. If $L(S) = V$, then S is a basis. If not, then there is some element y in V which is not in $L(S)$. Adjoin this element to S and consider the new set $S' = \{x_1, \ldots, x_k, y\}$. If this set were dependent there would be scalars $c_1, \ldots, c_{k+1}$, not all zero, such that

$$\sum_{i=1}^{k} c_i x_i + c_{k+1} y = O .$$

But $c_{k+1} \neq 0$ since $x_1, \ldots, x_k$ are independent. Hence, we could solve this equation for y and find that $y \in L(S)$, contradicting the fact that y is not in $L(S)$. Therefore, the set S' is independent but contains $k+1$ elements. If $L(S') = V$, then S' is a basis and, since S is a subset of S', part (a) is proved. If S' is not a basis, we can argue with S' as we did with S, getting a new set S'' which contains $k+2$ elements and is independent. If S'' is a basis, then part (a) is proved. If not, we repeat the process. We must arrive at a basis in a finite number of steps, otherwise we would eventually obtain an independent set with $n+1$ elements, contradicting Theorem 1.5. Therefore part (a) is proved.

To prove (b), let S be any independent set consisting of n elements. By part (a), S is a subset of some basis, say B. But by Theorem 1.6, the basis B has exactly n elements, so $S = B$.

1.9 Components

Let V be a linear space of dimension n and consider a basis whose elements $e_1, \ldots, e_n$ are taken in a given order. We denote such an ordered basis as an n-tuple $(e_1, \ldots, e_n)$. If $x \in V$, we can express x as a linear combination of these basis elements:

$$x = \sum_{i=1}^{n} c_i e_i . \tag{1.5}$$

The coefficients in this equation determine an n-tuple of numbers $(c_1, \ldots, c_n)$ that is uniquely determined by x. In fact, if we have another representation of x as a linear combination of $e_1, \ldots, e_n$, say $x = \sum_{i=1}^{n} d_i e_i$, then by subtraction from (1.5), we find that $\sum_{i=1}^{n} (c_i - d_i)e_i = 0$. But since the basis elements are independent, this implies $c_i = d_i$ for each i, so we have $(c_1, \ldots, c_n) = (d_1, \ldots, d_n)$.

The components of the ordered n-tuple $(c_1, \ldots, c_n)$ determined by Equation (1.5) are called *the components of x relative to the ordered basis $(e_1, \ldots, e_n)$.*

1.10 Exercises

In each of Exercises 1 through 10, let S denote the set of all vectors (x, y, z) in V_3 whose components satisfy the condition given. Determine whether S is a subspace of V_3. If S is a subspace, compute $\dim S$.

1. $x = 0$.
2. $x + y = 0$.
3. $x + y + z = 0$.
4. $x = y$.
5. $x = y = z$.
6. $x = y$ or $x = z$.
7. $x^2 - y^2 = 0$.
8. $x + y = 1$.
9. $y = 2x$ and $z = 3x$.
10. $x + y + z = 0$ and $x - y - z = 0$.

Let P_n denote the linear space of all real polynomials of degree $\le n$, where n is fixed. In each of Exercises 11 through 20, let S denote the set of all polynomials f in P_n satisfying the condition given. Determine whether or not S is a subspace of P_n. If S is a subspace, compute $\dim S$.

11. $f(0) = 0$.
12. $f'(0) = 0$.
13. $f''(0) = 0$.
14. $f(0) + f'(0) = 0$.
15. $f(0) = f(1)$.
16. $f(0) = f(2)$.
17. f is even.
18. f is odd.
19. f has degree $\le k$, where $k < n$, or $f = 0$.
20. f has degree k, where $k < n$, or $f = 0$.

21. In the linear space of all real polynomials $p(t)$, describe the subspace spanned by each of the following subsets of polynomials and determine the dimension of this subspace.
(a) $\{1, t^2, t^4\}$; (b) $\{t, t^3, t^5\}$; (c) $\{t, t^2\}$; (d) $\{1 + t, (1 + t)^2\}$.

22. In this exercise, $L(S)$ denotes the subspace spanned by a subset S of a linear space V. Prove each of the statements (a) through (f).
    (a) $S \subseteq L(S)$.
    (b) If $S \subseteq T \subseteq V$ and if T is a subspace of V, then $L(S) \subseteq T$. This property is described by saying that $L(S)$ is the *smallest* subspace of V which contains S.
    (c) A subset S of V is a subspace of V if and only if $L(S) = S$.
    (d) If $S \subseteq T \subseteq V$, then $L(S) \subseteq L(T)$.
    (e) If S and T are subspaces of V, then so is $S \cap T$.
    (f) If S and T are subsets of V, then $L(S \cap T) \subseteq L(S) \cap L(T)$.
    (g) Give an example in which $L(S \cap T) \neq L(S) \cap L(T)$.
23. Let V be the linear space consisting of all real-valued functions defined on the real line. Determine whether each of the following subsets of V is dependent or independent. Compute the dimension of the subspace spanned by each set.
    (a) $\{1, e^{ax}, e^{bx}\}$, $a \neq b$.
    (b) $\{e^{ax}, xe^{ax}\}$.
    (c) $\{1, e^{ax}, xe^{ax}\}$.
    (d) $\{e^{ax}, xe^{ax}, x^2e^{ax}\}$.
    (e) $\{e^x, e^{-x}, \cosh x\}$.
    (f) $\{\cos x, \sin x\}$.
    (g) $\{\cos^2 x, \sin^2 x\}$.
    (h) $\{1, \cos 2x, \sin^2 x\}$.
    (i) $\{\sin x, \sin 2x\}$.
    (j) $\{e^x \cos x, e^{-x} \sin x\}$.
24. Let V be a finite-dimensional linear space, and let S be a subspace of V. Prove each of the following statements.
    (a) S is finite dimensional and $\dim S \leq \dim V$.
    (b) $\dim S = \dim V$ if and only if $S = V$.
    (c) Every basis for S is part of a basis for V.
    (d) A basis for V need not contain a basis for S.

1.11 Inner products, Euclidean spaces. Norms

In ordinary Euclidean geometry, those properties that rely on the possibility of measuring lengths of line segments and angles between lines are called *metric* properties. In our study of V_n, we defined lengths and angles in terms of the dot product. Now we wish to extend these ideas to more general linear spaces. We shall introduce first a generalization of the dot product, which we call an *inner product,* and then define length and angle in terms of the inner product.

The dot product $x \cdot y$ of two vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in V_n was defined in Volume I by the formula

$$x \cdot y = \sum_{i=1}^{n} x_i y_i . \tag{1.6}$$

In a general linear space, we write (x, y) instead of $x \cdot y$ for inner products, and we define the product axiomatically rather than by a specific formula. That is, we state a number of properties we wish inner products to satisfy and we regard these properties as *axioms.*

DEFINITION. *A real linear space V is said to have an inner product if for each pair of elements x and y in V there corresponds a unique real number (x, y) satisfying the following axioms for all choices of x, y, z in V and all real scalars c.*

(1) $(x, y) = (y, x)$ *(commutativity, or symmetry).*
(2) $(x, y + z) = (x, y) + (x, z)$ *(distributivity, or linearity).*
(3) $c(x, y) = (cx, y)$ *(associativity, or homogeneity).*
(4) $(x, x) > 0$ *if* $x \neq O$ *(positivity).*

A real linear space with an inner product is called a ***real Euclidean space.***

Note: Taking c = 0 in (3), we find that $(O, y) = 0$ for all y.

In a complex linear space, an inner product (x, y) is a complex number satisfying the same axioms as those for a real inner product, except that the symmetry axiom is replaced by the relation

$$(1') \qquad (x, y) = (y, x), \qquad (\textit{Hermitian}\dagger\ \textbf{\textit{symmetry}})$$

where $(y$, x) denotes the complex conjugate of (y, x). In the homogeneity axiom, the scalar multiplier c can be any complex number. From the homogeneity axiom and $(1')$, we get the companion relation

$$(3') \qquad (x, cy) = \overline{(cy, x)} = \bar{c}\,\overline{(y, x)} = \bar{c}(x, y).$$

A complex linear space with an inner product is called a ***complex Euclidean 'space.*** (Sometimes the term ***unitary space*** is also used.) One example is complex vector space $V_n(C)$ discussed briefly in Section 12.16 of Volume I.

Although we are interested primarily in examples of real Euclidean spaces, the theorems of this chapter are valid for complex Euclidean spaces as well. When we use the term Euclidean space without further designation, it is to be understood that the space can be real or complex.

The reader should verify that each of the following satisfies all the axioms for an inner product.

EXAMPLE 1. In V_n let $(x, y) = x \cdot y$, the usual dot product of x and y.

EXAMPLE 2. If x = (x_1, x_2) and y = (y_1, y_2) are any two vectors in V_2, define (x, y) by the formula

$$(x, y) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 .$$

This example shows that there may be more than one inner product in a given linear space.

EXAMPLE 3. Let $C(a, b)$ denote the linear space of all real-valued functions continuous on an interval *[a,* b]. Define an inner product of two functions f and g by the formula

$$(f, g) = \int_a^b f(t)g(t)\,dt .$$

This formula is analogous to Equation (1.6) which defines the dot product of two vectors in V_n. The function values $f(t)$ and $g(t)$ play the role of the components x_i and y_i , and integration takes the place of summation.

† In honor of Charles Hermite (1822–1901), a French mathematician who made many contributions to algebra and analysis.