Paul R. Halmos

Finite-Dimensional Vector Spaces

AMS Subject Classifications: 15-01, 15A03

Library of Congress Cataloging in Publication Data

Halmos, Paul Richard, 1916-
Finite-dimensional vector spaces.

(Undergraduate texts in mathematics)
Reprint of the 2d ed. published by Van Nostrand, Princeton, N.J., in series: The University series in undergraduate mathematics.
Bibliography: p.
1. Vector spaces. 2. Transformations (Mathematics) I. Title.
[QA186.H34 1974] 512'.523 74-10688

Printed in the United States of America.

9 8 7 6 5 4 3 2

ISBN 0-387-90093-4 Springer-Verlag New York
ISBN 3-540-90093-4 Springer-Verlag Berlin Heidelberg

PREFACE

My purpose in this book is to treat linear transformations on finite-dimensional vector spaces by the methods of more general theories. The idea is to emphasize the simple geometric notions common to many parts of mathematics and its applications, and to do so in a language that gives away the trade secrets and tells the student what is in the back of the minds of people proving theorems about integral equations and Hilbert spaces. The reader does not, however, have to share my prejudiced motivation. Except for an occasional reference to undergraduate mathematics the book is self-contained and may be read by anyone who is trying to get a feeling for the linear problems usually discussed in courses on matrix theory or "higher" algebra. The algebraic, coordinate-free methods do not lose power and elegance by specialization to a finite number of dimensions, and they are, in my belief, as elementary as the classical coordinatized treatment.

I originally intended this book to contain a theorem if and only if an infinite-dimensional generalization of it already exists. The tempting easiness of some essentially finite-dimensional notions and results was, however, irresistible, and in the final result my initial intentions are just barely visible. They are most clearly seen in the emphasis, throughout, on generalizable methods instead of sharpest possible results. The reader may sometimes see some obvious way of shortening the proofs I give. In such cases the chances are that the infinite-dimensional analogue of the shorter proof is either much longer or else non-existent.

A preliminary edition of the book (Annals of Mathematics Studies, Number 7, first published by the Princeton University Press in 1942) has been circulating for several years. In addition to some minor changes in style and in order, the difference between the preceding version and this one is that the latter contains the following new material: (1) A brief discussion of fields, and, in the treatment of vector spaces with inner products, special attention to the real case. (2) A definition of determinants in invariant terms, via the theory of multilinear forms. (3) Exercises.

The exercises (well over three hundred of them) constitute the most significant addition; I hope that they will be found useful by both student

and teacher. There are two things about them the reader should know. First, if an exercise is neither imperative ("prove that . . .") nor interrogative ("is it true that . . . ?") but merely declarative, then it is intended as a challenge. For such exercises the reader is asked to discover if the assertion is true or false, prove it if true and construct a counterexample if false, and, most important of all, discuss such alterations of hypothesis and conclusion as will make the true ones false and the false ones true. Second, the exercises, whatever their grammatical form, are not always placed so as to make their very position a hint to their solution. Frequently exercises are stated as soon as the statement makes sense, quite a bit before machinery for a quick solution has been developed. A reader who tries (even unsuccessfully) to solve such a "misplaced" exercise is likely to appreciate and to understand the subsequent developments much better for his attempt. Having in mind possible future editions of the book, I ask the reader to let me know about errors in the exercises, and to suggest improvements and additions. (Needless to say, the same goes for the text.)

None of the theorems and only very few of the exercises are my discovery; most of them are known to most working mathematicians, and have been known for a long time. Although I do not give a detailed list of my sources, I am nevertheless deeply aware of my indebtedness to the books and papers from which I learned and to the friends and strangers who, before and after the publication of the first version, gave me much valuable encouragement and criticism. I am particularly grateful to three men: J. L. Doob and Arlen Brown, who read the entire manuscript of the first and the second version, respectively, and made many useful suggestions, and John von Neumann, who was one of the originators of the modern spirit and methods that I have tried to present and whose teaching was the inspiration for this book.

P. R. H.

CONTENTS

CHAPTER I

SPACES

§ 1. Fields

In what follows we shall have occasion to use various classes of numbers (such as the class of all real numbers or the class of all complex numbers). Because we should not, at this early stage, commit ourselves to any specific class, we shall adopt the dodge of referring to numbers as *scalars*. The reader will not lose anything essential if he consistently interprets scalars as real numbers or as complex numbers; in the examples that we shall study both classes will occur. To be specific (and also in order to operate at the proper level of generality) we proceed to list all the general facts about scalars that we shall need to assume.

(A) To every pair, α and β, of scalars there corresponds a scalar $\alpha + \beta$, called the *sum* of α and β, in such a way that

(1) addition is commutative, $\alpha + \beta = \beta + \alpha$,

(2) addition is associative, $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$,

(3) there exists a unique scalar 0 (called *zero*) such that $\alpha + 0 = \alpha$ for every scalar α, and

(4) to every scalar α there corresponds a unique scalar $-\alpha$ such that $\alpha + (-\alpha) = 0$.

(B) To every pair, α and β, of scalars there corresponds a scalar $\alpha\beta$, called the *product* of α and β, in such a way that

(1) multiplication is commutative, $\alpha\beta = \beta\alpha$,

(2) multiplication is associative, $\alpha(\beta\gamma) = (\alpha\beta)\gamma$,

(3) there exists a unique non-zero scalar 1 (called *one*) such that $\alpha 1 = \alpha$ for every scalar α, and

(4) to every non-zero scalar α there corresponds a unique scalar α^{-1} $\left(\text{or } \frac{1}{\alpha}\right)$ such that $\alpha\alpha^{-1} = 1$.

(C) Multiplication is distributive with respect to addition, $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

If addition and multiplication are defined within some set of objects (scalars) so that the conditions **(A)**, **(B)**, and **(C)** are satisfied, then that set (together with the given operations) is called a *field*. Thus, for example, the set $\mathfrak{Q}$ of all rational numbers (with the ordinary definitions of sum and product) is a field, and the same is true of the set $\mathfrak{R}$ of all real numbers and the set $\mathfrak{C}$ of all complex numbers.

EXERCISES

1. Almost all the laws of elementary arithmetic are consequences of the axioms defining a field. Prove, in particular, that if $\mathfrak{F}$ is a field, and if α, β, and γ belong to $\mathfrak{F}$, then the following relations hold.

(a) $0 + \alpha = \alpha$.

(b) If $\alpha + \beta = \alpha + \gamma$, then $\beta = \gamma$.

(c) $\alpha + (\beta - \alpha) = \beta$. (Here $\beta - \alpha = \beta + (-\alpha)$.)

(d) $\alpha \cdot 0 = 0 \cdot \alpha = 0$. (For clarity or emphasis we sometimes use the dot to indicate multiplication.)

(e) $(-1)\alpha = -\alpha$.

(f) $(-\alpha)(-\beta) = \alpha\beta$.

(g) If $\alpha\beta = 0$, then either $\alpha = 0$ or $\beta = 0$ (or both).

2. (a) Is the set of all positive integers a field? (In familiar systems, such as the integers, we shall almost always use the ordinary operations of addition and multiplication. On the rare occasions when we depart from this convention, we shall give ample warning. As for "positive," by that word we mean, here and elsewhere in this book, "greater than or equal to zero." If 0 is to be excluded, we shall say "strictly positive.")

(b) What about the set of all integers?

(c) Can the answers to these questions be changed by re-defining addition or multiplication (or both)?

3. Let m be an integer, $m \geqq 2$, and let $\mathfrak{Z}_m$ be the set of all positive integers less than m, $\mathfrak{Z}_m = \{0, 1, \cdots, m - 1\}$. If α and β are in $\mathfrak{Z}_m$, let $\alpha + \beta$ be the least positive remainder obtained by dividing the (ordinary) sum of α and β by m, and, similarly, let $\alpha\beta$ be the least positive remainder obtained by dividing the (ordinary) product of α and β by m. (Example: if $m = 12$, then $3 + 11 = 2$ and $3 \cdot 11 = 9$.)

(a) Prove that $\mathfrak{Z}_m$ is a field if and only if m is a prime.

(b) What is -1 in $\mathfrak{Z}_5$?

(c) What is $\frac{1}{3}$ in $\mathfrak{Z}_7$?

4. The example of $\mathfrak{Z}_p$ (where p is a prime) shows that not quite all the laws of elementary arithmetic hold in fields; in $\mathfrak{Z}_2$, for instance, $1 + 1 = 0$. Prove that if $\mathfrak{F}$ is a field, then either the result of repeatedly adding 1 to itself is always different from 0, or else the first time that it is equal to 0 occurs when the number of summands is a prime. (The *characteristic* of the field $\mathfrak{F}$ is defined to be 0 in the first case and the crucial prime in the second.)

5. Let $\mathbb{Q}(\sqrt{2})$ be the set of all real numbers of the form $\alpha + \beta\sqrt{2}$, where α and β are rational.
(a) Is $\mathbb{Q}(\sqrt{2})$ a field?
(b) What if α and β are required to be integers?

6. (a) Does the set of all polynomials with integer coefficients form a field?
(b) What if the coefficients are allowed to be real numbers?

7. Let $\mathfrak{F}$ be the set of all (ordered) pairs (α, β) of real numbers.
(a) If addition and multiplication are defined by

$$(\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta)$$

and

$$(\alpha, \beta)(\gamma, \delta) = (\alpha\gamma, \beta\delta),$$

does $\mathfrak{F}$ become a field?
(b) If addition and multiplication are defined by

$$(\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta)$$

and

$$(\alpha, \beta)(\gamma, \delta) = (\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma),$$

is $\mathfrak{F}$ a field then?
(c) What happens (in both the preceding cases) if we consider ordered pairs of complex numbers instead?

§ 2. Vector spaces

We come now to the basic concept of this book. For the definition that follows we assume that we are given a particular field $\mathfrak{F}$; the scalars to be used are to be elements of $\mathfrak{F}$.

DEFINITION. A *vector space* is a set $\mathcal{V}$ of elements called *vectors* satisfying the following axioms.

(A) To every pair, x and y, of vectors in $\mathcal{V}$ there corresponds a vector $x + y$, called the *sum* of x and y, in such a way that
(1) addition is commutative, $x + y = y + x$,
(2) addition is associative, $x + (y + z) = (x + y) + z$,
(3) there exists in $\mathcal{V}$ a unique vector 0 (called the *origin*) such that $x + 0 = x$ for every vector x, and
(4) to every vector x in $\mathcal{V}$ there corresponds a unique vector $-x$ such that $x + (-x) = 0$.

(B) To every pair, α and x, where α is a scalar and x is a vector in $\mathcal{V}$, there corresponds a vector αx in $\mathcal{V}$, called the *product* of α and x, in such a way that
(1) multiplication by scalars is associative, $\alpha(\beta x) = (\alpha\beta)x$, and
(2) $1x = x$ for every vector x.

(C) (1) Multiplication by scalars is distributive with respect to vector addition, $\alpha(x + y) = \alpha x + \alpha y$, and

(2) multiplication by vectors is distributive with respect to scalar addition, $(\alpha + \beta)x = \alpha x + \beta x$.

These axioms are not claimed to be logically independent; they are merely a convenient characterization of the objects we wish to study. The relation between a vector space $\mathcal{V}$ and the underlying field $\mathcal{F}$ is usually described by saying that $\mathcal{V}$ is a vector space *over* $\mathcal{F}$. If $\mathcal{F}$ is the field $\mathcal{R}$ of real numbers, $\mathcal{V}$ is called a *real vector space;* similarly if $\mathcal{F}$ is $\mathcal{Q}$ or if $\mathcal{F}$ is $\mathcal{C}$, we speak of *rational vector spaces* or *complex vector spaces.*

§ 3. Examples

Before discussing the implications of the axioms, we give some examples. We shall refer to these examples over and over again, and we shall use the notation established here throughout the rest of our work.

(1) Let $\mathcal{C}^1 (= \mathcal{C})$ be the set of all complex numbers; if we interpret $x + y$ and αx as ordinary complex numerical addition and multiplication, $\mathcal{C}^1$ becomes a complex vector space.

(2) Let $\mathcal{P}$ be the set of all polynomials, with complex coefficients, in a variable t. To make $\mathcal{P}$ into a complex vector space, we interpret vector addition and scalar multiplication as the ordinary addition of two polynomials and the multiplication of a polynomial by a complex number; the origin in $\mathcal{P}$ is the polynomial identically zero.

Example (1) is too simple and example (2) is too complicated to be typical of the main contents of this book. We give now another example of complex vector spaces which (as we shall see later) is general enough for all our purposes.

(3) Let $\mathcal{C}^n$, $n = 1, 2, \cdots$, be the set of all n-tuples of complex numbers. If $x = (\xi_1, \cdots, \xi_n)$ and $y = (\eta_1, \cdots, \eta_n)$ are elements of $\mathcal{C}^n$, we write, by definition,

$$x + y = (\xi_1 + \eta_1, \cdots, \xi_n + \eta_n),$$

$$\alpha x = (\alpha\xi_1, \cdots, \alpha\xi_n),$$

$$0 = (0, \cdots, 0),$$

$$-x = (-\xi_1, \cdots, -\xi_n).$$

It is easy to verify that all parts of our axioms (A), (B), and (C), § 2, are satisfied, so that $\mathcal{C}^n$ is a complex vector space; it will be called *n-dimensional complex coordinate space.*

(4) For each positive integer n, let $\mathcal{P}_n$ be the set of all polynomials (with complex coefficients, as in example (2)) of degree $\leqq n-1$, together with the polynomial identically zero. (In the usual discussion of degree, the degree of this polynomial is not defined, so that we cannot say that it has degree $\leqq n-1$.) With the same interpretation of the linear operations (addition and scalar multiplication) as in (2), $\mathcal{P}_n$ is a complex vector space.

(5) A close relative of $\mathcal{C}^n$ is the set $\mathcal{R}^n$ of all n-tuples of real numbers. With the same formal definitions of addition and scalar multiplication as for $\mathcal{C}^n$, except that now we consider only real scalars α, the space $\mathcal{R}^n$ is a real vector space; it will be called *n-dimensional real coordinate space*.

(6) All the preceding examples can be generalized. Thus, for instance, an obvious generalization of (1) can be described by saying that every field may be regarded as a vector space over itself. A common generalization of (3) and (5) starts with an arbitrary field $\mathcal{F}$ and forms the set $\mathcal{F}^n$ of n-tuples of elements of $\mathcal{F}$; the formal definitions of the linear operations are the same as for the case $\mathcal{F} = \mathcal{C}$.

(7) A field, by definition, has at least two elements; a vector space, however, may have only one. Since every vector space contains an origin, there is essentially (i.e., except for notation) only one vector space having only one vector. This most trivial vector space will be denoted by $\mathcal{O}$.

(8) If, in the set $\mathcal{R}$ of all real numbers, addition is defined as usual and multiplication of a real number by a rational number is defined as usual, then $\mathcal{R}$ becomes a rational vector space.

(9) If, in the set $\mathcal{C}$ of all complex numbers, addition is defined as usual and multiplication of a complex number by a real number is defined as usual, then $\mathcal{C}$ becomes a real vector space. (Compare this example with (1); they are quite different.)

§ 4. Comments

A few comments are in order on our axioms and notation. There are striking similarities (and equally striking differences) between the axioms for a field and the axioms for a vector space over a field. In both cases, the axioms (A) describe the additive structure of the system, the axioms (B) describe its multiplicative structure, and the axioms (C) describe the connection between the two structures. Those familiar with algebraic terminology will have recognized the axioms (A) (in both § 1 and § 2) as the defining conditions of an abelian (commutative) group; the axioms (B) and (C) (in § 2) express the fact that the group admits scalars as operators. We mention in passing that if the scalars are elements of a ring (instead of a field), the generalized concept corresponding to a vector space is called a *module*.

Special real vector spaces (such as $\mathfrak{R}^2$ and $\mathfrak{R}^3$) are familiar in geometry. There seems at this stage to be no excuse for our apparently uninteresting insistence on fields other than $\mathfrak{R}$, and, in particular, on the field $\mathfrak{C}$ of complex numbers. We hope that the reader is willing to take it on faith that we shall have to make use of deep properties of complex numbers later (conjugation, algebraic closure), and that in both the applications of vector spaces to modern (quantum mechanical) physics and the mathematical generalization of our results to Hilbert space, complex numbers play an important role. Their one great disadvantage is the difficulty of drawing pictures; the ordinary picture (Argand diagram) of $\mathfrak{C}^1$ is indistinguishable from that of $\mathfrak{R}^2$, and a graphic representation of $\mathfrak{C}^2$ seems to be out of human reach. On the occasions when we have to use pictorial language we shall therefore use the terminology of $\mathfrak{R}^n$ in $\mathfrak{C}^n$, and speak of $\mathfrak{C}^2$, for example, as a plane.

Finally we comment on notation. We observe that the symbol 0 has been used in two meanings: once as a scalar and once as a vector. To make the situation worse, we shall later, when we introduce linear functionals and linear transformations, give it still other meanings. Fortunately the relations among the various interpretations of 0 are such that, after this word of warning, no confusion should arise from this practice.

EXERCISES

1. Prove that if x and y are vectors and if α is a scalar, then the following relations hold.
(a) $0 + x = x$.
(b) $-0 = 0$.
(c) $\alpha \cdot 0 = 0$.
(d) $0 \cdot x = 0$. (Observe that the same symbol is used on both sides of this equation; on the left it denotes a scalar, on the right it denotes a vector.)
(e) If $\alpha x = 0$, then either $\alpha = 0$ or $x = 0$ (or both).
(f) $-x = (-1)x$.
(g) $y + (x - y) = x$. (Here $x - y = x + (-y)$.)

2. If p is a prime, then $\mathfrak{Z}_p^n$ is a vector space over $\mathfrak{Z}_p$ (cf. § 1, Ex. 3); how many vectors are there in this vector space?

3. Let $\mathfrak{U}$ be the set of all (ordered) pairs of real numbers. If $x = (\xi_1, \xi_2)$ and $y = (\eta_1, \eta_2)$ are elements of $\mathfrak{U}$, write

$$x + y = (\xi_1 + \eta_1, \xi_2 + \eta_2)$$
$$\alpha x = (\alpha \xi_1, 0)$$
$$0 = (0, 0)$$
$$-x = (-\xi_1, -\xi_2).$$

Is $\mathfrak{U}$ a vector space with respect to these definitions of the linear operations? Why?

4. Sometimes a subset of a vector space is itself a vector space (with respect to the linear operations already given). Consider, for example, the vector space $\mathcal{C}^3$ and the subsets $\mathcal{U}$ of $\mathcal{C}^3$ consisting of those vectors (ξ_1, ξ_2, ξ_3) for which

(a) ξ_1 is real,
(b) $\xi_1 = 0$,
(c) either $\xi_1 = 0$ or $\xi_2 = 0$,
(d) $\xi_1 + \xi_2 = 0$,
(e) $\xi_1 + \xi_2 = 1$.

In which of these cases is $\mathcal{U}$ a vector space?

5. Consider the vector space $\mathcal{P}$ and the subsets $\mathcal{U}$ of $\mathcal{P}$ consisting of those vectors (polynomials) x for which

(a) x has degree 3,
(b) $2x(0) = x(1)$,
(c) $x(t) \geqq 0$ whenever $0 \leqq t \leqq 1$,
(d) $x(t) = x(1 - t)$ for all t.

In which of these cases is $\mathcal{U}$ a vector space?

§ 5. Linear dependence

Now that we have described the spaces we shall work with, we must specify the relations among the elements of those spaces that will be of interest to us.

We begin with a few words about the summation notation. If corresponding to each of a set of indices i there is given a vector x_i, and if it is not necessary or not convenient to specify the set of indices exactly, we shall simply speak of a set $\{x_i\}$ of vectors. (We admit the possibility that the same vector corresponds to two distinct indices. In all honesty, therefore, it should be stated that what is important is not which vectors appear in $\{x_i\}$, but how they appear.) If the index-set under consideration is finite, we shall denote the sum of the corresponding vectors by $\sum_i x_i$ (or, when desirable, by a more explicit symbol such as $\sum_{i=1}^{n} x_i$). In order to avoid frequent and fussy case distinctions, it is a good idea to admit into the general theory sums such as $\sum_i x_i$ even when there are no indices i to be summed over, or, more precisely, even when the index-set under consideration is empty. (In that case, of course, there are no vectors to sum, or, more precisely, the set $\{x_i\}$ is also empty.) The value of such an "empty sum" is defined, naturally enough, to be the vector 0.

DEFINITION. A finite set $\{x_i\}$ of vectors is *linearly dependent* if there exists a corresponding set $\{\alpha_i\}$ of scalars, not all zero, such that

$$\sum_i \alpha_i x_i = 0.$$

If, on the other hand, $\sum_i \alpha_i x_i = 0$ implies that $\alpha_i = 0$ for each i, the set $\{x_i\}$ is *linearly independent.*

The wording of this definition is intended to cover the case of the empty set; the result in that case, though possibly paradoxical, dovetails very satisfactorily with the rest of the theory. The result is that the empty set of vectors is linearly independent. Indeed, if there are no indices i, then it is not possible to pick out some of them and to assign to the selected ones a non-zero scalar so as to make a certain sum vanish. The trouble is not in avoiding the assignment of zero; it is in finding an index to which something can be assigned. Note that this argument shows that the empty set is not linearly dependent; for the reader not acquainted with arguing by "vacuous implication," the equivalence of the definition of linear independence with the straightforward negation of the definition of linear dependence needs a little additional intuitive justification. The easiest way to feel comfortable about the assertion "$\sum_i \alpha_i x_i = 0$ implies that $\alpha_i = 0$ for each i," in case there are no indices i, is to rephrase it this way: "if $\sum_i \alpha_i x_i = 0$, then there is no index i for which $\alpha_i \neq 0$." This version is obviously true if there is no index i at all.

Linear dependence and independence are properties of sets of vectors; it is customary, however, to apply the adjectives to vectors themselves, and thus we shall sometimes say "a set of linearly independent vectors" instead of "a linearly independent set of vectors." It will be convenient also to speak of the linear dependence and independence of a not necessarily finite set, $\mathfrak{X}$, of vectors. We shall say that $\mathfrak{X}$ is linearly independent if every finite subset of $\mathfrak{X}$ is such; otherwise $\mathfrak{X}$ is linearly dependent.

To gain insight into the meaning of linear dependence, let us study the examples of vector spaces that we already have.

(1) If x and y are any two vectors in $\mathcal{C}^1$, then x and y form a linearly dependent set. If $x = y = 0$, this is trivial; if not, then we have, for example, the relation $yx + (-x)y = 0$. Since it is clear that every set containing a linearly dependent subset is itself linearly dependent, this shows that in $\mathcal{C}^1$ every set containing more than one element is a linearly dependent set.

(2) More interesting is the situation in the space $\mathcal{P}$. The vectors x, y, and z, defined by

$$x(t) = 1 - t,$$

$$y(t) = t(1 - t),$$

$$z(t) = 1 - t^2,$$

are, for example, linearly dependent, since $x + y - z = 0$. However, the infinite set of vectors $x_0, x_1, x_2, \cdots$, defined by

$$x_0(t) = 1, \quad x_1(t) = t, \quad x_2(t) = t^2, \cdots,$$

is a linearly independent set, for if we had any relation of the form

$$\alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_n x_n = 0,$$

then we should have a polynomial identity

$$\alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n = 0,$$

whence

$$\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0.$$

(3) As we mentioned before, the spaces $\mathfrak{C}^n$ are the prototype of what we want to study; let us examine, for example, the case $n = 3$. To those familiar with higher-dimensional geometry, the notion of linear dependence in this space (or, more properly speaking, in its real analogue $\mathfrak{R}^3$) has a concrete geometric meaning, which we shall only mention. In geometrical language, two vectors are linearly dependent if and only if they are collinear with the origin, and three vectors are linearly dependent if and only if they are coplanar with the origin. (If one thinks of a vector not as a point in a space but as an arrow pointing from the origin to some given point, the preceding sentence should be modified by crossing out the phrase "with the origin" both times that it occurs.) We shall presently introduce the notion of linear manifolds (or vector subspaces) in a vector space, and, in that connection, we shall occasionally use the language suggested by such geometrical considerations.

§ 6. Linear combinations

We shall say, whenever $x = \sum_i \alpha_i x_i$, that x is a *linear combination* of $\{x_i\}$; we shall use without any further explanation all the simple grammatical implications of this terminology. Thus we shall say, in case x is a linear combination of $\{x_i\}$, that x is linearly dependent on $\{x_i\}$; we shall leave to the reader the proof that if $\{x_i\}$ is linearly independent, then a necessary and sufficient condition that x be a linear combination of $\{x_i\}$ is that the enlarged set, obtained by adjoining x to $\{x_i\}$, be linearly dependent. Note that, in accordance with the definition of an empty sum, the origin is a linear combination of the empty set of vectors; it is, moreover, the only vector with this property.

The following theorem is the fundamental result concerning linear dependence.

THEOREM. *The set of non-zero vectors $x_1, \cdots, x_n$ is linearly dependent if and only if some x_k, $2 \leq k \leq n$, is a linear combination of the preceding ones.*

PROOF. Let us suppose that the vectors $x_1, \cdots, x_n$ are linearly dependent, and let k be the first integer between 2 and n for which $x_1, \cdots, x_k$ are linearly

dependent. (If worse comes to worst, our assumption assures us that $k = n$ will do.) Then

$$\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$$

for a suitable set of α's (not all zero); moreover, whatever the α's, we cannot have $\alpha_k = 0$, for then we should have a linear dependence relation among $x_1, \cdots, x_{k-1}$, contrary to the definition of k. Hence

$$x_k = \frac{-\alpha_1}{\alpha_k} x_1 + \cdots + \frac{-\alpha_{k-1}}{\alpha_k} x_{k-1},$$

as was to be proved. This proves the necessity of our condition; sufficiency is clear since, as we remarked before, every set containing a linearly dependent set is itself such.

§ 7. Bases

DEFINITION. A (linear) *basis* (or a *coordinate system*) in a vector space $\mathcal{V}$ is a set $\mathfrak{X}$ of linearly independent vectors such that every vector in $\mathcal{V}$ is a linear combination of elements of $\mathfrak{X}$. A vector space $\mathcal{V}$ is *finite-dimensional* if it has a finite basis.

Except for the occasional consideration of examples we shall restrict our attention, throughout this book, to finite-dimensional vector spaces.

For examples of bases we turn again to the spaces $\mathcal{P}$ and $\mathcal{C}^n$. In $\mathcal{P}$, the set $\{x_n\}$, where $x_n(t) = t^n$, $n = 0, 1, 2, \cdots$, is a basis; every polynomial is, by definition, a linear combination of a finite number of x_n. Moreover $\mathcal{P}$ has no finite basis, for, given any finite set of polynomials, we can find a polynomial of higher degree than any of them; this latter polynomial is obviously not a linear combination of the former ones.

An example of a basis in $\mathcal{C}^n$ is the set of vectors x_i, $i = 1, \cdots, n$, defined by the condition that the j-th coordinate of x_i is δ_{ij}. (Here we use for the first time the popular Kronecker δ; it is defined by $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.) Thus we assert that in $\mathcal{C}^3$ the vectors $x_1 = (1, 0, 0)$, $x_2 = (0, 1, 0)$, and $x_3 = (0, 0, 1)$ form a basis. It is easy to see that they are linearly independent; the formula

$$x = (\xi_1, \xi_2, \xi_3) = \xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3$$

proves that every x in $\mathcal{C}^3$ is a linear combination of them.

In a general finite-dimensional vector space $\mathcal{V}$, with basis $\{x_1, \cdots, x_n\}$, we know that every x can be written in the form

$$x = \sum_i \xi_i x_i;$$

we assert that the ξ's are uniquely determined by x. The proof of this

assertion is an argument often used in the theory of linear dependence. If we had $x = \sum_i \eta_i x_i$, then we should have, by subtraction,

$$\sum_i (\xi_i - \eta_i) x_i = 0.$$

Since the x_i are linearly independent, this implies that $\xi_i - \eta_i = 0$ for $i = 1, \cdots, n$; in other words, the ξ's are the same as the η's. (Observe that writing $\{x_1, \cdots, x_n\}$ for a basis with n elements is not the proper thing to do in case $n = 0$. We shall, nevertheless, frequently use this notation. Whenever that is done, it is, in principle, necessary to adjoin a separate discussion designed to cover the vector space 0. In fact, however, everything about that space is so trivial that the details are not worth writing down, and we shall omit them.)

THEOREM. *If $\mathcal{V}$ is a finite-dimensional vector space and if $\{y_1, \cdots, y_m\}$ is any set of linearly independent vectors in $\mathcal{V}$, then, unless the y's already form a basis, we can find vectors $y_{m+1}, \cdots, y_{m+p}$ so that the totality of the y's, that is, $\{y_1, \cdots, y_m, y_{m+1}, \cdots, y_{m+p}\}$, is a basis. In other words, every linearly independent set can be extended to a basis.*

PROOF. Since $\mathcal{V}$ is finite-dimensional, it has a finite basis, say $\{x_1, \cdots, x_n\}$. We consider the set $\mathcal{S}$ of vectors

$$y_1, \cdots, y_m, x_1, \cdots, x_n,$$

in this order, and we apply to this set the theorem of § 6 several times in succession. In the first place, the set $\mathcal{S}$ is linearly dependent, since the y's are (as are all vectors) linear combinations of the x's. Hence some vector of $\mathcal{S}$ is a linear combination of the preceding ones; let z be the first such vector. Then z is different from any y_i, $i = 1, \cdots, m$ (since the y's are linearly independent), so that z is equal to some x, say $z = x_i$. We consider the new set $\mathcal{S}'$ of vectors

$$y_1, \cdots, y_m, x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n.$$

We observe that every vector in $\mathcal{V}$ is a linear combination of vectors in $\mathcal{S}'$, since by means of $y_1, \cdots, y_m, x_1, \cdots, x_{i-1}$ we may express x_i, and then by means of $x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n$ we may express any vector. (The x's form a basis.) If $\mathcal{S}'$ is linearly independent, we are done. If it is not, we apply the theorem of § 6 again and again the same way till we reach a linearly independent set containing $y_1, \cdots, y_m$, in terms of which we may express every vector in $\mathcal{V}$. This last set is a basis containing the y's.