LINEAR TRANSFORMATIONS in n-dimensional vector space

LINEAR TRANSFORMATIONS

in n-dimensional vector space

AN INTRODUCTION TO THE THEORY OF HILBERT SPACE

BY

H. L. HAMBURGER

Professor in the University of Cologne

AND

M. E. GRIMSHAW

Fellow of Newnham College, Cambridge

CAMBRIDGE

AT THE UNIVERSITY PRESS

1956

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS

London Office: Bentley House, N.W.1
American Branch: New York

Agents for Canada, India, and Pakistan: Macmillan

First edition 1951
Reprinted 1956

First printed in Great Britain at the University Press, Cambridge
Reprinted by offset-litho by
Jarrold and Sons Ltd., Norwich

PREFACE

In this book we introduce the ideas and methods of the theory of linear transformations in Hilbert space by using them to present the elements of the theory in a finite dimensional vector space. We discuss problems fundamental in the study of general abstract linear spaces and use the powerful methods appropriate to the consideration of these spaces; both problems and methods, however, are simplified by our restriction to an n-dimensional space.

Some particular features of the presentation may be mentioned. The main results in the theory of linear transformations, namely, the spectral representations of Hermitian transformations, and the canonical representations and commutativity properties of general linear transformations, are developed here without the use of determinants because determinants can be used in Hilbert space only in very special cases. A concrete vector space has been chosen in preference to an abstract space so that the ideas may be more readily grasped, but the coordinates of vectors and the elements of matrices occur only exceptionally in the proofs; most of them could be applied unchanged for an abstract space. By defining the scalar product at the start instead of developing first the descriptive properties of linear manifolds and transformations (an alternative course with some obvious advantages, followed in the books by G. Julia [V]* and P. R. Halmos [III]) we can introduce the Schmidt orthogonalization process at an early stage and take advantage of its appeal to geometrical intuition; for instance, we obtain a natural geometrical illustration of the problem of the solution of a system of homogeneous linear equations. The resulting development on geometrical lines of the calculus of linear manifolds replaces the familiar and more formal calculus of matrices.

We use only orthogonal coordinate systems because the use of oblique coordinates would lead to difficulties in later generalizations to Hilbert space. This means that we consider general linear transformations only as mappings of the space on itself, while we interpret unitary transformations both as congruent mappings of the space on itself and also as transformations of coordinate systems.

* Numbers in square brackets refer to the list of references at the end of the book. Roman numbers are used for books and arabic numbers for papers.

Finally, although we deal with algebraical problems, we have not excluded analytical methods in view of their importance for certain problems in Hilbert space, and we use them to obtain the spectral representations of Hermitian transformations.

The book presents a self-contained account of the theory of linear transformations in finite dimensional vector space. Chapter I contains the elementary general properties of linear manifolds and linear transformations. Chapter II introduces the algebraical treatment of the normal transformations (with the Hermitian and unitary transformations as special examples) and of general and orthogonal projectors. In Chapter III the spectral representations of Hermitian and normal transformations are obtained by the analytical methods due to Hilbert and based on the property of the eigen-values of a Hermitian form as its maxima in certain closed sets of the vector space. The methods are developed further to give inequalities for the eigen-values that have proved of value in recent computational work. The chapter closes with a discussion of the functional calculus for Hermitian and normal transformations. In Chapter IV we return to algebraical methods in dealing with the more complicated problem of the reduction of a general linear transformation to the Jordan canonical form. The existence of eigen-values is established here, without the introduction of determinants, by the use of arguments contained in recent papers by N. Dunford and one of the present authors ([16] and [23]). The spectral representations of normal transformations are obtained again, this time as special examples of the representations of linear transformations with simple elementary divisors, and the chapter ends with a discussion of the commutativity properties of general linear transformations. These results include a series of theorems for linear transformations with simple elementary divisors, and a formulation, in terms of the complete reduction of the linear transformation A, of the commutativity properties of the polynomials $p(A)$, for which we have found no reference. The theme of Chapter III is taken up again in Chapter V, which is concerned with the spectral representation of the pencil $H-\lambda G$, where H and G are Hermitian transformations of which G is positive definite; a familiar problem with various applications in mathematical physics. The application to the dynamical theory of small oscillations is dealt with in detail in order to illustrate the significance of the inequalities of Chapter III. The introduction, in Chapter V, of a generalized n-dimensional vector space in which the scalar product is defined by a positive definite Hermitian form foreshadows some of our later work on abstract spaces and provides a link between this work and the elementary approach of Chapter I. Moreover, it leads to the theorem given towards the end of Chapter V, that every linear transformation with simple elementary divisors can be considered as a normal transformation in a vector space in which the scalar product is suitably defined.

The notes at the end of the book give references for the main results. They are selected somewhat arbitrarily and make no claim to completeness. No references to the notes are made in the text; they are meant to be consulted after the reading of each chapter.

H. L. H.
M. E. G.

CAMBRIDGE
September 1947

ACKNOWLEDGEMENTS

The authors wish to record their indebtedness to Dr N. Aronszajn for permission to publish the proof of his inequalities which appears in Chapter III.

They also wish to thank Miss S. M. Edmonds, Miss N. Walls and Dr J. D. Weston for assistance in proof correction and for much valuable criticism.

H. L. H.
M. E. G.

CONTENTS

CHAPTER I. LINEAR MANIFOLDS AND LINEAR TRANSFORMATIONS IN $\mathfrak{B}_n$

§ 1. Vectors and operations on vectors *page* 1
§ 2. Linear manifolds 6
§ 3. Projections; orthonormal bases of a L.M.; complete orthonormal systems 12
§ 4. Products and sums of L.M.'s 17
§ 5. Orthogonal complements; systems of homogeneous linear equations 21
§ 6. Linear transformations; systems of non-homogeneous linear equations 24
§ 7. Sums and products of L.T.'s 34

CHAPTER II. SPECIAL LINEAR TRANSFORMATIONS IN $\mathfrak{B}_n$

§ 8. Hermitian transformations; eigen-values, eigen-manifolds and reduction 40
§ 9. Normal transformations and unitary transformations 45
§ 10. Projectors 49

CHAPTER III. SPECTRAL REPRESENTATION OF HERMITIAN TRANSFORMATIONS IN $\mathfrak{B}_n$

§ 11. Continuous functions in $\mathfrak{B}_n$ 58
§ 12. Orthonormal systems of eigen-solutions of a H.T. The spectral representations 60
§ 13. Unitary transformations of a Hermitian form to principal axes 67
§ 14. Further maximal properties of the eigen-values. Inequalities 73
§ 15. Functional calculus of H.T.'s. The resolvent and its properties 78
§ 16. Transformations commutative with H.T.'s. Normal transformations 86

Chapter IV. Spectral properties of linear transformations in $\mathfrak{B}_n$

§ 17. Eigen-values and principal manifolds *page* 97
§ 18. The minimal polynomial 103
§ 19. Functional calculus for L.T.'s. The resolvent 108
§ 20. Canonical bases of a principal manifold 117
§ 21. Jordan's canonical form for a nilpotent L.T. Commutativity 122
§ 22. Jordan's canonical form for any L.T.; elementary divisors and Segre characteristic; similar transformations 130
§ 23. Commutativity of L.T.'s 138

Chapter V. Vector spaces with positive Hermitian metric forms

§ 24. The spaces $\mathfrak{B}_n'$ and $\mathfrak{B}_{n0}$ 151
§ 25. Hermitian forms in $\mathfrak{B}_n'$ and $\mathfrak{B}_{n0}$ and the pencil $H - \lambda G$ in $\mathfrak{B}_n$ 159
§ 26. L.T.'s that are Hermitian or normal in some $\mathfrak{B}_n'$ 164
§ 27. Application to the dynamical theory of small oscillations 166

Notes 170

References 184

Index of Authors 190

General Index 191

CHAPTER I

LINEAR MANIFOLDS AND LINEAR TRANSFORMATIONS IN $\mathfrak{V}_n$

§ 1. Vectors and operations on vectors

1·0. In these first four chapters we deal for the most part with well-known algebraical problems which we describe in geometrical terms and interpret in a space of n-dimensional complex vectors (an n-dimensional vector space). Our object is to introduce the ideas and technique that are characteristic of the theory of Hilbert space, and so we give the proofs as far as possible without using the coordinates of vectors or the elements of matrices; most of them would have taken the same form if we had defined the vector space, as Hilbert space is defined, in terms of its abstract properties. We illustrate the applications of the general theory by considering for n-dimensional vector space such familiar problems as those of the solution of systems of homogeneous and non-homogeneous linear equations and of the transformation of a Hermitian form to principal axes, but we do not give the familiar arguments using determinants because they are not in keeping with our development of the theory and cannot be generalized for Hilbert space. Indeed, we have deliberately avoided the use of determinants throughout the book.

It will be seen that we define a metric form for the space at a very early stage in the discussion, and that this leads us at once to the idea of orthogonality and to the introduction of Schmidt's orthogonalization process which plays an essential part in the theory of bounded linear transformations in Hilbert space.

1·1. Definition of a vector. *We define a vector x as an ordered system of n complex numbers $x_1, x_2, \dots, x_n$,* called the *coordinates* of the vector, and we write $x = (x_1, x_2, \dots, x_n)$. We define the *zero vector* as $0 = (0, 0, \dots, 0)$.

A system of n complex numbers may also be called a *point*, and the point $(0, 0, \dots, 0)$ may be called the *origin*.

When $x' = (x_1', x_2', \dots, x_n')$ we write $x = x'$ if, and only if, $x_\nu = x_\nu'$ for $\nu = 1, 2, \dots, n$. We write $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)$, where $\bar{x}_\nu$ denotes the number conjugate to x_ν, and we say that the vector $\bar{x}$ is conjugate to x. We say that x is real if $x = \bar{x}$; that is, if the n coordinates x_ν are all real.

Throughout the book vectors are represented by letters of the same types as those representing complex numbers, but in the case of vectors suffixes are raised. Thus, $x^1, x^2, \ldots, x^n$ denotes a set of n vectors, while n numbers (for example, the coordinates of a vector) are denoted by $x_1, x_2, \ldots, x_n$.

The set of all vectors x defined by n coordinates forms the n-dimensional vector space which we denote by $\mathfrak{V}_n$. We write $x \in \mathfrak{V}_n$ to mean that x is an element of $\mathfrak{V}_n$ and, more generally, we write $x \in \mathfrak{X}$ if x is an element of any set $\mathfrak{X}$ of vectors.

Sets of vectors are denoted by German capital letters; sets in the plane of complex numbers, such as curves and domains, are denoted by small German letters.

1·2. Multiplication of a vector by a number. If x is any vector, $x = (x_1, x_2, \ldots, x_n)$, and if α is any complex number, then x multiplied by α is defined as the vector $(\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$ and is written as αx. In particular, we write $-x$ for αx when $\alpha = -1$.

The vector αx coincides with x if $\alpha = 1$; it is the zero vector if, and only if, either $\alpha = 0$ or $x = 0$. We see at once that $\overline{\alpha x} = \bar{\alpha}\bar{x}$, and that, if β is any complex number, then $\alpha(\beta x) = \beta(\alpha x) = \alpha\beta x$.

1·3. Addition of vectors. If

$$x = (x_1, x_2, \ldots, x_n) \quad \text{and} \quad y = (y_1, y_2, \ldots, y_n),$$

then the sum of x and y is defined as the vector

$$x + y = (x_1 + y_1,\ x_2 + y_2,\ \ldots,\ x_n + y_n).$$

We write $x - y$ for $x + (-y)$. We see that $x - x = 0$.

1·31. The commutative, associative and distributive laws of addition all hold, since we obtain at once, for vectors x, y, z and any numbers α, β,

$$x + y = y + x, \quad (x + y) + z = x + (y + z)$$

and

$$\alpha(x + y) = \alpha x + \alpha y, \quad (\alpha + \beta)x = \alpha x + \beta x.$$

If $x + y = x + z$, then $y = z$.

1·32. For any non-real vector z there are two real vectors x and y such that $z = x + iy$; they are

$$x = \tfrac{1}{2}(z + \bar{z}), \quad y = \frac{1}{2i}(z - \bar{z}).$$

1·33. The vectors

$$u^1 = (1, 0, \ldots, 0), \quad u^2 = (0, 1, 0, \ldots, 0), \quad \ldots, \quad u^n = (0, 0, \ldots, 0, 1)$$

are called the *coordinate vectors*. We have, for any vector x,

$$x = \sum_{\nu=1}^{n} x_\nu u^\nu. \tag{1·33·1}$$

1·4. Definition of the scalar product of two vectors. The *scalar product* of two vectors x and y is a complex number associated with the vectors. We denote it by (x, y) and define it so that it shall satisfy the following three conditions for any vectors x, y, z and any number α:

$$\text{(i)}\ (\alpha x, y) = \alpha(x, y), \quad \text{(ii)}\ (x+y, z) = (x, z) + (y, z),$$
$$\text{(iii)}\ (y, x) = \overline{(x, y)} = (\bar{x}, \bar{y}).$$

The conditions (i) and (iii) imply that

$$(x, \alpha y) = \overline{(\alpha y, x)} = \bar{\alpha}\overline{(y, x)} = \bar{\alpha}(x, y), \tag{1·4·1}$$

and (ii) and (iii) imply that

$$(x, y+z) = (\overline{y+z, \bar{x}}) = \overline{(y, x)} + \overline{(z, x)} = (x, y) + (x, z). \tag{1·4·2}$$

We now add the definition of the special scalar products of the coordinate vectors u^ν by writing

$$\text{(iv)}\ (u^\mu, u^\nu) = \delta_{\mu\nu} \quad (\mu, \nu = 1, 2, \ldots, n),$$

where here, and throughout the book, $\delta_{\mu\nu}$ is the Kronecker symbol with the interpretation

$$\delta_{\mu\nu} = 0 \quad (\mu \neq \nu), \qquad \delta_{\mu\nu} = 1 \quad (\mu = \nu).$$

Writing $x = \sum_{\nu=1}^{n} x_\nu u^\nu$, $y = \sum_{\nu=1}^{n} y_\nu u^\nu$ we see by (i), (ii) and (iv) that $(x, u^\nu) = x_\nu$ and hence, by (1·33·1), that $x = \sum_{\nu=1}^{n} (x, u^\nu) u^\nu$. We also obtain the analytical expression for the scalar product

$$(x, y) = \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} x_\nu \overline{y_\mu} (u^\nu, u^\mu) = \sum_{\nu=1}^{n} \sum_{\mu=1}^{n} x_\nu \overline{y_\mu} \delta_{\mu\nu}$$
$$= \sum_{\nu=1}^{n} x_\nu \overline{y_\nu}, \tag{1·4·3}$$

which could have been used for the definition. Had we defined the scalar product by (1·4·3) we should have deduced at once the conditions (i), (ii), (iii) and (iv).

We deduce at once from the condition (iii) that (x, y) is real if x and y are both real and that $(y, x) = 0$ if $(x, y) = 0$.

1·41. Definition of the absolute value of a vector. *The absolute value, or norm, of a vector x is defined as the non-negative real number*

$$\| x \| = (x, x)^{\frac{1}{2}} = \left(\sum_{\nu=1}^{n} | x_\nu |^2 \right)^{\frac{1}{2}}.$$

We also call $\| x \|$ the *length* of the vector x. We see that $\| x \| = 0$ if and only if $x = 0$. If $\| x \| = 1$ we say that x is a *unit vector*; the coordinate vectors are unit vectors. The absolute value $\| x - y \|$ is called the *distance* between the points, or vectors, x and y. Clearly, $\| x - y \| = \| y - x \|$.

By 1·4 (i) and (1·4·1), we have

$$\| \alpha x \|^2 = (\alpha x, \alpha x) = \alpha \bar{\alpha} (x, x) = | \alpha |^2 \| x \|^2,$$

and hence

$$\| \alpha x \| = | \alpha | \, \| x \|. \tag{1·41·1}$$

We notice that this proof of (1·41·1) is independent of the special condition 1·4 (iv) for the scalar product.

1·42. For any vectors x and y and any number λ we have

$$\begin{aligned} \| x + \lambda y \|^2 &= (x + \lambda y, x + \lambda y) \\ &= \| x \|^2 + \lambda (y, x) + \bar{\lambda} (x, y) + \lambda \bar{\lambda} \| y \|^2 \\ &\geqslant 0, \end{aligned}$$

with equality if, and only if, $x + \lambda y = 0$. Taking $\| y \| \neq 0$ and $\lambda \| y \|^2 = -(x, y)$, we obtain

$$\| x \|^2 \| y \|^2 - | (x, y) |^2 \geqslant 0,$$

and we deduce that

$$| (x, y) | \leqslant \| x \| \, \| y \|, \tag{1·42·1}$$

with equality if, and only if, one of the vectors x and y is zero or is a numerical multiple of the other.

The inequality (1·42·1) follows from the conditions (i), (ii) and (iii) of 1·4 for the scalar product. If condition (iv) is also used, (1·42·1) follows from (1·4·3) by the Cauchy inequality.

If neither x nor y is zero, and if we write

$$\frac{(x, y)}{\| x \| \, \| y \|} = \cos \theta, \tag{1·42·2}$$

θ is real so long as (x, y) is real, and we can then interpret it as the angle between the vectors x and y. Thus the scalar product defines both length and angle; it is the metric form in $\mathfrak{V}_n$.

If $(x, y) = 0$ we say that the vectors x and y are *orthogonal*. Thus, condition (iv) of 1·4 means that any two of the coordinate vectors u^ν are orthogonal. If $x = \alpha y \neq 0$ we say that x and y are *parallel*.

1·43. The absolute value satisfies the 'triangle inequality'

$$\| x+y \| \leqslant \| x \| + \| y \|,$$

since, by (1·42·1),

$$\begin{aligned} \| x+y \|^2 &= (x+y, x+y) = (x,x)+(x,y)+\overline{(x,y)}+(y,y) \\ &\leqslant \| x \|^2 + 2\,|(x,y)| + \| y \|^2 \leqslant \| x \|^2 + 2 \| x \| \, \| y \| + \| y \|^2 \\ &= (\| x \| + \| y \|)^2. \end{aligned}$$

When x and y are orthogonal vectors we obtain the Theorem of Pythagoras

$$\| x+y \|^2 = \| x \|^2 + \| y \|^2.$$

1·5. We now prove that the space $\mathfrak{V}_n$ is complete and separable. We shall not again refer explicitly to these properties of n-dimensional vector space. We introduce them here in preparation for the discussion of the same important properties of Hilbert vector space.

1·51. Definition of a limit vector of a sequence. We say that the vector $\mathring{x}$ of $\mathfrak{V}_n$ is the *limit vector of the sequence* $\{x^m\}$ of vectors of $\mathfrak{V}_n$, and we write $\lim_{m\to\infty} x^m = \mathring{x}$, if $\lim_{m\to\infty} \| x^m - \mathring{x} \| = 0$.

It follows from **1·43** that a sequence cannot have two distinct limit vectors. For suppose that $\mathring{x}$ and $\mathring{y}$ are both limit vectors of the sequence $\{x^m\}$. Then, for any integer m,

$$\| \mathring{x} - \mathring{y} \| \leqslant \| \mathring{x} - x^m \| + \| x^m - \mathring{y} \| \to 0$$

as $m \to \infty$, so that $\| \mathring{x} - \mathring{y} \| = 0$ and $\mathring{y}$ coincides with $\mathring{x}$.

1·52. THEOREM. *Let* $\lim_{m\to\infty} x^m = \mathring{x}$ *and let* y *be any vector of* $\mathfrak{V}_n$. *Then* $\lim_{m\to\infty} (x^m, y) = (\mathring{x}, y)$.

PROOF. We have, by **1·4** (ii) and (1·42·1),

$$|(x^m, y) - (\mathring{x}, y)| = |(x^m - \mathring{x}, y)| \leqslant \| x^m - \mathring{x} \| \, \| y \| \to 0$$

as $m \to \infty$.

1·53. THEOREM. *The space* $\mathfrak{V}_n$ *is complete; that is, if a sequence* $\{x^m\}$ *of vectors of* $\mathfrak{V}_n$ *satisfies the condition that there corresponds to any positive number* ϵ *an integer* m_0 *such that* $\| x^p - x^q \| < \epsilon$ *for* $p, q \geqslant m_0(\epsilon)$, *then there exists a vector* $\mathring{x}$ *of* $\mathfrak{V}_n$ *which is the limit vector of the sequence.*

PROOF. Let $x^m = \sum_{\nu=1}^{n} x_{m\nu} u^\nu$. Then

$$\begin{aligned} |x_{p\nu} - x_{q\nu}| &\leqslant \left(\sum_{\nu=1}^{n} |x_{p\nu} - x_{q\nu}|^2 \right)^{\frac{1}{2}} \\ &= \| x^p - x^q \| < \epsilon \quad (p, q \geqslant m_0), \end{aligned}$$

and so $\lim_{m\to\infty} x_{m\nu}$ exists; denote it by $\hat{x}_\nu$ and write $\hat{x} = \sum_{\nu=1}^{n} \hat{x}_\nu u^\nu$. We obtain

$$\|\hat{x}-x^m\| = \left(\sum_{\nu=1}^{n} |\hat{x}_\nu - x_{m\nu}|^2\right)^{\frac{1}{2}} \to 0$$

as $m \to \infty$, and this gives the desired result.

A sequence that satisfies the condition of Theorem 1·53 is said to be *convergent*. We readily see that a sequence that has a limit vector $\hat{x}$ is convergent and we say that it *converges to* $\hat{x}$.

1·6. Definition of the closure of a set of vectors, of a closed set and of an everywhere dense subset. Let $\mathfrak{S}$ be any set of vectors of $\mathfrak{B}_n$. The set obtained by *adding all those limit vectors of sequences of elements of* $\mathfrak{S}$ *that do not belong to* $\mathfrak{S}$ is called *the closure of* $\mathfrak{S}$; we denote it by $\bar{\mathfrak{S}}$. If $\bar{\mathfrak{S}}$ coincides with $\mathfrak{S}$ we say that $\mathfrak{S}$ is *closed*.

Let $\mathfrak{E}$ be any set of vectors of $\mathfrak{B}_n$ and let $\mathfrak{S}$ be a subset of $\mathfrak{E}$. If $\mathfrak{E}$ is contained in $\bar{\mathfrak{S}}$ we say that $\mathfrak{S}$ *is everywhere dense in* $\mathfrak{E}$.

A necessary and sufficient condition that $\mathfrak{S}$ is everywhere dense in $\mathfrak{E}$ is that there corresponds to any element x of $\mathfrak{E}$ and any positive number ϵ an element s of $\mathfrak{S}$ such that $\|x-s\| < \epsilon$, for this condition is equivalent to the condition that x is either an element of $\mathfrak{S}$ or else the limit vector of a sequence of elements of $\mathfrak{S}$.

1·61. Theorem. *The space $\mathfrak{B}_n$ is separable; that is, there exists a denumerably infinite set $\mathfrak{S}$ of vectors of $\mathfrak{B}_n$ that is everywhere dense in $\mathfrak{B}_n$.*

Proof. Consider the set $\mathfrak{S}$ of all vectors s of $\mathfrak{B}_n$ of which the coordinates $s_\nu = \sigma_\nu + i\tau_\nu$ are formed from rational numbers σ_ν and τ_ν. Now the set of rational numbers is denumerable and so is the set of ordered pairs of rationals. Thus the set of the numbers s_ν is denumerable. The elements of $\mathfrak{S}$ may be regarded as ordered sets of n numbers s_ν, so that $\mathfrak{S}$ is also denumerable. Further, since we can express any complex number as the limit of a sequence of numbers s_ν, it follows that we can express any vector of $\mathfrak{B}_n$ that is not a vector of $\mathfrak{S}$ as the limit vector of a sequence of vectors of $\mathfrak{S}$. Thus the set $\mathfrak{S}$ fulfils the requirements of the theorem.

§2. Linear Manifolds

2·0. Definition of a linear manifold. *A set $\mathfrak{M}$ of vectors of $\mathfrak{B}_n$ is called a linear manifold* (denoted by the abbreviation L.M.) *if, whenever $a \in \mathfrak{M}$ and $b \in \mathfrak{M}$, then also $\alpha a \in \mathfrak{M}$ and $a+b \in \mathfrak{M}$,* where α is any complex number.

If $a^1, a^2, \ldots, a^r$ are any vectors of $\mathfrak{B}_n$ and $\alpha_1, \alpha_2, \ldots, \alpha_r$ any complex numbers, then the vector $\alpha_1 a^1 + \alpha_2 a^2 + \ldots + \alpha_r a^r$ is said to be a *linear combination* of the vectors $a^1, a^2, \ldots, a^r$. It follows immediately from Definition 2·0 that, if $a^1, a^2, \ldots, a^r$ are contained in the L.M. $\mathfrak{M}$, then so is any linear combination of these vectors. In particular, any L.M. contains the zero vector.

If the L.M. $\mathfrak{M}$ consists of all linear combinations of the vectors $a^1, a^2, \ldots, a^r$ we write it as $\mathfrak{M} = [a^1, a^2, \ldots, a^r]$ and we say that $\mathfrak{M}$ is *spanned* by the vectors $a^1, a^2, \ldots, a^r$. The space $\mathfrak{B}_n$ is the L.M. spanned by the coordinate vectors, so that $\mathfrak{B}_n = [u^1, u^2, \ldots, u^n]$. The zero vector constitutes a L.M. which we denote by $\mathfrak{M} = \mathfrak{D}$.

If $\mathfrak{M}$ and $\mathfrak{M}'$ are L.M.'s such that every vector of each is contained in the other we write $\mathfrak{M}' = \mathfrak{M}$; if every vector of $\mathfrak{M}'$ is contained in $\mathfrak{M}$ we write $\mathfrak{M}' \subseteq \mathfrak{M}$ and $\mathfrak{M} \supseteq \mathfrak{M}'$. If there is a vector of $\mathfrak{M}$ that is not contained in $\mathfrak{M}'$ while every vector of $\mathfrak{M}'$ is contained in $\mathfrak{M}$ we write $\mathfrak{M}' \subset \mathfrak{M}$ and $\mathfrak{M} \supset \mathfrak{M}'$. Plainly, the L.M. $\mathfrak{D}$ is contained in every L.M. $\mathfrak{M}$ so that we always have $\mathfrak{M} \supseteq \mathfrak{D}$.

2·1. Definition of linear dependence and linear independence of vectors. The vector x is said to be *linearly dependent* on the vectors $a^1, a^2, \ldots, a^r$ if it is a linear combination of these vectors; that is, if there exist numerical coefficients α_ν such that $x = \sum_{\nu=1}^{r} \alpha_\nu a^\nu$. The vectors $a^1, a^2, \ldots, a^r$ are said to be linearly dependent, or to form a *linearly dependent set*, if there exist numerical coefficients α_ν, not all zero, such that

$$\sum_{\nu=1}^{r} \alpha_\nu a^\nu = 0. \qquad (2\cdot1\cdot1)$$

The vectors $a^1, a^2, \ldots, a^r$ are said to be *linearly independent*, or to form a *linearly independent set*, if a relation (2·1·1) implies that all the coefficients α_ν are zero.

It follows at once that any subset of a linearly independent set of vectors is a linearly independent set. We also see that a linearly independent set cannot contain the zero vector, since the relation $\alpha 0 = 0$ does not imply that $\alpha = 0$. An example of a linearly independent set of n vectors is the set of coordinate vectors $u^1, u^2, \ldots, u^n$.

2·11. The vectors $\overline{a^1}, \overline{a^2}, \ldots, \overline{a^r}$ are linearly dependent or linearly independent according as $a^1, a^2, \ldots, a^r$ are linearly dependent or linearly independent.

2·12. Let $\mathfrak{M}$ be a L.M. in $\mathfrak{B}_n$, and let $\overline{\mathfrak{M}}$ denote the set of all vectors conjugate to vectors of $\mathfrak{M}$. Then $\overline{\mathfrak{M}}$ is also a L.M. for, if $\overline{a} \in \overline{\mathfrak{M}}$ and $\overline{b} \in \overline{\mathfrak{M}}$, then $\alpha a \in \mathfrak{M}$ and $a + b \in \mathfrak{M}$ for every number α, and so $\overline{\alpha a} \in \overline{\mathfrak{M}}$ and $\overline{a} + \overline{b} \in \overline{\mathfrak{M}}$.

If $\mathfrak{M} = [a^1, a^2, \ldots, a^r]$ then, clearly, $\overline{\mathfrak{M}} = [\overline{a^1}, \overline{a^2}, \ldots, \overline{a^r}]$. We say that the L.M.'s $\mathfrak{M}$ and $\overline{\mathfrak{M}}$ are conjugate to one another.

2·2. Definition of a basis of a L.M. The set of vectors $a^1, a^2, \ldots, a^r$ is called a *basis* of the L.M. $\mathfrak{M}$ if the following two conditions are satisfied:

(i) *the vectors* $a^1, a^2, \ldots, a^r$ *are linearly independent,*

(ii) $\mathfrak{M} = [a^1, a^2, \ldots, a^r]$.

The set of coordinate vectors u^ν is a basis of the L.M. $\mathfrak{V}_n$.

2·21. We prove by the next two theorems that every L.M. in $\mathfrak{V}_n$ has a basis of not more than n elements, but we need first the following lemma.

Lemma. *If the set of vectors* $a^1, a^2, \ldots, a^r$ *is a basis of* $\mathfrak{M}$, *and if*

$$b = \beta_1 a^1 + \beta_2 a^2 + \ldots + \beta_r a^r,$$

where β_ν *is not zero, then the set of vectors* $a^1, a^2, \ldots, a^{\nu-1}, b, a^{\nu+1}, \ldots, a^r$ *is also a basis of* $\mathfrak{M}$.

Proof. Without loss of generality we may take $\nu = 1$ and $\beta_1 \neq 0$. We first show that if there exists a relation

$$\alpha_1 b + \alpha_2 a^2 + \ldots + \alpha_r a^r = 0$$

then all the coefficients are zero. We substitute for b in the relation and obtain

$$\alpha_1 \beta_1 a^1 + (\alpha_2 + \alpha_1 \beta_2) a^2 + \ldots + (\alpha_r + \alpha_1 \beta_r) a^r = 0.$$

But, since $a^1, a^2, \ldots, a^r$ are linearly independent, and since $\beta_1 \neq 0$, this gives successively $\alpha_1 = 0$, $\alpha_2 = 0, \ldots, \alpha_r = 0$. Thus the vectors $b, a^2, \ldots, a^r$ are linearly independent.

Now any vector linearly dependent on $b, a^2, \ldots, a^r$ belongs to $\mathfrak{M}$ and further, since every element of $\mathfrak{M}$ is linearly dependent on $a^1, a^2, \ldots, a^r$, and since

$$a^1 = \frac{1}{\beta_1}(b - \beta_2 a^2 - \ldots - \beta_r a^r),$$

it follows that every element of $\mathfrak{M}$ is linearly dependent on $b, a^2, \ldots, a^r$. Thus these vectors form a basis of $\mathfrak{M}$.

2·3 Theorem. *Let* $\mathfrak{M}$ *be a L.M. in* $\mathfrak{V}_n$ *with a basis of* r *elements. Then any set of* r *linearly independent elements of* $\mathfrak{M}$ *is a basis of* $\mathfrak{M}$ *and any set of* $r+1$ *elements of* $\mathfrak{M}$ *is linearly dependent.*

Proof. Let $\mathfrak{M} = [a^1, a^2, \ldots, a^r]$ and let $b^1, b^2, \ldots, b^r$ be any set of r linearly independent elements of $\mathfrak{M}$. Then

$$b^1 = \alpha_{11} a^1 + \alpha_{12} a^2 + \ldots + \alpha_{1r} a^r,$$

where at least one of the coefficients, say α_{11}, is not zero, since no element b^μ can be zero because the b's are linearly independent. Then, by Lemma 2·21,

$$\mathfrak{M} = [a^1, a^2, \ldots, a^r] = [b^1, a^2, \ldots, a^r].$$

Suppose we have proved that

$$\mathfrak{M} = [b^1, b^2, \ldots, b^{s-1}, a^s, \ldots, a^r],$$

where $2 \leqslant s \leqslant r$. Then we have

$$b^s = \beta_{s1} b^1 + \ldots + \beta_{s,s-1} b^{s-1} + \alpha_{ss} a^s + \ldots + \alpha_{sr} a^r,$$

where the coefficients $\alpha_{s\mu}$ are not all zero since $b^1, b^2, \ldots, b^s$ are linearly independent, and we may assume $\alpha_{ss} \neq 0$. Then, by Lemma 2·21,

$$\mathfrak{M} = [b^1, b^2, \ldots, b^s, a^{s+1}, \ldots, a^r].$$

We obtain in this way, by an induction process, $\mathfrak{M} = [b^1, b^2, \ldots, b^r]$, and we see that the elements $b^1, b^2, \ldots, b^r$ form a basis of $\mathfrak{M}$ since, by hypothesis, they are linearly independent.

It follows at once that no set of $r+1$ elements of $\mathfrak{M}$ is linearly independent.

2·31. Corollary of Theorem 2·3. *Every set of $n+1$ vectors of $\mathfrak{V}_n$ is linearly dependent.*

Proof. The corollary follows at once from Theorem 2·3, since the coordinate vectors $u^1, u^2, \ldots, u^n$ form a basis of $\mathfrak{V}_n$.

2·4. Theorem. *There corresponds to every L.M. $\mathfrak{M}$ in $\mathfrak{V}_n$ (with the trivial exception $\mathfrak{M} = \mathfrak{O}$) a positive integer r, where $r \leqslant n$, which is the maximum number of linearly independent elements contained in $\mathfrak{M}$. Every set of r linearly independent elements of $\mathfrak{M}$ is a basis of $\mathfrak{M}$, and every basis of $\mathfrak{M}$ has exactly r elements.*

Note. This theorem establishes the existence of a basis of at most n elements for every L.M. in $\mathfrak{V}_n$. We describe a method of constructing such a basis in 2·5.

Proof. By Corollary 2·31, every set of $n+1$ vectors of $\mathfrak{V}_n$ is linearly dependent so that $\mathfrak{M}$ cannot contain a linearly independent set of more than n vectors. On the other hand, any non-zero vector of $\mathfrak{M}$ constitutes a linearly independent set. Hence, if $\mathfrak{M} \neq \mathfrak{O}$, there are linearly independent subsets of $\mathfrak{M}$, and the number of elements they contain has a greatest value r, where $r \leqslant n$.

Now let $a^1, a^2, \ldots, a^r$ be any set of r linearly independent elements of $\mathfrak{M}$ and let x be any other element of $\mathfrak{M}$. The $r+1$ elements $x, a^1, a^2, \ldots, a^r$ cannot be linearly independent, and there exists therefore a relation of the form $\alpha x + \beta_1 a^1 + \ldots + \beta_r a^r = 0$ in which