LINEAR ALGEBRA
with Linear Differential Equations

FRANKLIN LOWENTHAL

LINEAR ALGEBRA
with Linear Differential Equations

FRANKLIN LOWENTHAL

University of Wisconsin—Parkside

JOHN WILEY & SONS

New York London Sydney Toronto

Library of Congress Cataloging in Publication Data:

Lowenthal, Franklin, 1938–
Linear algebra with linear differential equations.

Bibliography: p.
1. Algebras, Linear. 2. Differential equations, Linear. I. Title.

QA184. L68 512'.15 74-8933
ISBN 0-471-55150-3

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

LINEAR ALGEBRA
with Linear Differential Equations

To my mother and father

PREFACE

This book is designed to serve as a text for a first course in linear algebra. Such a course will normally be taken immediately after the first year of calculus. However, it is quite feasible for an ambitious student to take it concurrent with the second semester (or third quarter) of the calculus sequence.

In a one-quarter course it should be possible to include most of the material in Chapters 1–4 as well as Sections 5.1 and 5.2 (and possibly 6.1 and 6.2). In a one-semester course there should be no difficulty in covering most of the material in Chapters 1–6. In a two-quarter sequence there should be no problem in finishing the whole book at a relaxed pace.

The content of this book is noteworthy in several respects. The treatment of inner product spaces and spectral theory is unusually complete for a book at this level. The finite dimensional versions of such classical topics as the projection theorem, the Fredholm alternative, and the spectral theorem for normal transformations are included. But perhaps the most distinctive feature of the book is the inclusion throughout of examples involving function spaces and linear differential operators. This serves a dual purpose: On the one hand, it presents the theory of linear differential equations as an integrated body of knowledge which illustrates the ideas and techniques of linear algebra; on the other hand, it provides a certain insight into linear algebra which cannot otherwise be achieved.

The material is organized so as to introduce as soon as possible the concepts of a vector space and of a linear transformation. Thus the reader will have had ample opportunity to become familiar with these new ideas by the time the formal development of the theory of finite dimensional vector spaces is begun in Chapter 3. This abstract approach is best suited to reveal the essential unity and beauty of linear algebra. The difficulties that are sometimes encountered in this kind of development are overcome by four principal devices: (a) suitably motivating each new development in the subject; (b) examining all aspects and implications of both definitions and theorems in a special category called Comment; (c) illustrating every new concept with copious examples, each designed to combine a minimum of tedious computation with a maximum of insight; (d) providing ample exercises of all types and at just the right level of difficulty.

This book is devoted primarily to the study of *real* vector spaces. The results of Sections 1.4 and 1.5 on complex numbers and complex vector spaces are required for only four sections: 5.6, 6.2, 6.3, and 7.4. In fact, all that is really needed for Sections 6.2 and 7.4 is the fundamental theorem of algebra, which is stated in Section 1.4.

The reader should have no difficulty with the numbering system that is used. Within every chapter, definitions are double-numbered consecutively beginning with the chapter number and one; the same is true for theorems, lemmas, examples, and equations. Corollaries and comments also have double numbers corresponding to the theorem, definition, etc., they pertain to, and they are further designated as a, b, etc., when more than one occurs. The word "equation" will usually be omitted from a reference; thus "by (5–6)" means "by equation (5–6)."

My chief acknowledgement is to Charles Wright of the University of Oregon, who presented me with about 30 pages of detailed suggestions. Others who were kind enough to offer various kinds of assistance were John Jacobs, Richard M. Koch, and Kenneth Ross of the University of Oregon, Nelo Allan, Ronald Gatterdam, Alan Wallace, and Kenneth Weston of the University of Wisconsin—Parkside, Harry Hochstadt of the Polytechnic Institute of Brooklyn, Paul Sally of the University of Chicago, Lawrence Kugler of the University of Michigan, Flint College, and Victor Katz of the Federal City College. Special thanks are due to Deborah Keller for an outstanding job in typing the manuscript. I also wish to thank my two editors, Frederick Corey and Gary Ostedt, as well as the whole crew at John Wiley & Sons, for their cooperation throughout. Finally, I owe much to my family for their interest and patience.

Kenosha, Wisconsin *Franklin Lowenthal*

CONTENTS

1

VECTOR SPACES

THE ENGLISH LANGUAGE

In this book, many new ideas and concepts will be introduced and studied; many theorems will be formulated and proved. To understand what is said here it is essential to understand the meaning of certain key words. These are words that the reader has known and used all his life. However, they are used in a very precise sense in mathematics, a sense that differs somewhat from their colloquial usage.

The words *each*, *every*, and *all* are used synonymously by the mathematician. Thus the three statements "Each man is mortal," "Every man is mortal," and "All men are mortal" have exactly the same meaning. In contrast, the word *some* has an entirely different meaning. The statement "Some women are sexy" asserts that there exists at least one woman who is sexy; the statement gives no information as to whether all, many, or even two women are sexy. The distinction between the two types of statements above is clarified by examining what would be involved in disproving them. To disprove "All men are mortal" it suffices to find just one immortal man while to disprove "Some women are sexy" it is necessary to show that every woman is not sexy.

The statement "A and B are true" means that both are true; to establish it we would have to show that A is true and that B is true. In contrast, the statement "A or B is true" means that at least one of the two is true—possibly both are true, but no information is given on this point. An often convenient way of establishing the assertion "A or B is true" is to show that if A were false, then B is true (or that if B were false, then A is true) and consequently at least one of the two is definitely true. This crucial distinction between the words *and* and *or* can also be seen by considering the negations of the two kinds of statements above. Remember that to disprove a proposition or a theorem, exhibiting just one instance where it is false is sufficient—whether it is always false may be interesting but is irrelevant. Thus, to disprove "A and B are true," it suffices to show that either A or B is false, while to disprove "A or B is true," it is necessary to show that both A and B are false.

In this book there will be an abundance of both theorems and definitions; we should like to briefly examine their salient features. A theorem has the form "A implies B" or "if A, then B"; its converse is "B implies A" or "if B, then A." The reader, hopefully, realizes that these two statements are quite different; in fact, one could be true and the other false (see Exercise 1). If both are true, we say that A is true just in case B is true or A and B are equivalent or "A if and only if B." This single statement consists, in fact, of two assertions; "A implies B" and "B implies A." Now a definition by its very nature asserts that a new concept is equivalent to certain known old concepts. Hence, in a *definition*, the word *if* means "if and only if." For example, the definition "A triangle is isosceles if two sides are equal" really means "A triangle is isosceles if and only if two sides are equal." We could reformulate this definition to eliminate the word if (and hence this ambiguity) as follows: "An isosceles triangle is a triangle with two equal sides."

To assert that "if A, then B" is the same as asserting that "A holds *only if* B holds," i.e., "if not B, then not A." In other words, the two assertions "A implies B" and "not B implies not A" are equivalent. The latter is called the *contrapositive*. We shall sometimes find it simpler to prove the contrapositive. For example, the reader may recall that in plane geometry, instead of proving "If the alternate interior angles are equal, then the lines are parallel," one proves the contrapositive "If the lines are not parallel, then the alternate interior angles are not equal."

The word *set* will be undefined; synonyms, equally undefined, are collection, class, and system. A subset of a set is a collection of objects each of which belongs to the original set. Every set is a subset of itself; the term *proper subset* is used to designate a subset of a set that is not equal to the whole set. The fact that the empty set is a subset of every set is of importance elsewhere in mathematics; the reader of this book need have no nightmares about empty sets.

EXERCISE

1. Give an example from (a) geometry, (b) algebra, and (c) calculus of a theorem whose converse is false.

1.1 THE VECTOR SPACE R^n

The reader may already have some idea of what a vector is; probably, he pictures a vector as an arrow. In this book we will call many other things vectors. Therefore, the reader is urged at the very outset to abandon the idea that a vector is an arrow; at the same time, he should never forget that an arrow is an *example* of a vector.

In this section we shall study vectors in n-dimensional space. The reader is cautioned that these, too, are just *examples* of vectors; the definition of the concept of a vector is postponed until the next section.

DEFINITION 1.1. A *vector in real n space* is an n-tuple of real numbers

$$v = (x_1, x_2, \ldots, x_n) \tag{1-1}$$

Two vectors $v = (x_1, x_2, \ldots, x_n)$ and $w = (y_1, y_2, \ldots, y_n)$ in n space are said to be *equal* if they look identical, i.e., if $x_i = y_i$ for all i, $i = 1, 2, \ldots, n$.

Notation. Throughout this book we shall use the letters u, v, and w to denote vectors. Lowercase letters at the beginning of the Latin $(a, b, c, \ldots)$ and Greek $(\alpha, \beta, \gamma, \ldots)$ alphabets as well as the letters s, t, x, and y will denote real numbers (also called *scalars*). The letters i, j, k, l, m, n, p, q, and r will denote integers. There will be, unfortunately, one flagrant violation of these rules: The notation $i_1, i_2, \ldots, i_n$ will be used for certain special vectors in n space.

Example 1.1. The vectors (0, 1), (1, 0), and (1, 1) are three different vectors in 2 space.

Comment 1.1a. The reader should carefully distinguish between the vector, which is the whole n-tuple, and the real numbers that appear as entries in the n-tuple. These entries are called the *components* of the vector.

Example 1.2. The vector $(5, \pi, -\frac{3}{2}, +\sqrt{2})$ in 4 space has the real number 5 as its first component, the real number π as its second component, the real number $-\frac{3}{2}$ as its third component, and the real number $+\sqrt{2}$ as its fourth component.

The special n-tuple all of whose components are zero is called the *zero vector* and is denoted by 0. Thus the symbol 0 is ambiguous since it represents both the real number zero and the zero vector; however, the author guarantees that the context will always clarify the sense in which the symbol 0 is used.

Comment 1.1b. The vector $v = (x_1, x_2, x_3)$ in 3 space may be identified with the arrow starting at the origin and terminating at the point whose first, second, and third coordinates are just x_1, x_2, and x_3, respectively. The arrow that represents the zero vector degenerates to a point. Similar identifications can, of course, be made for vectors in 1 and 2 space.

Now the reader must suspect (if for no reason other than that the book has just begun) that there is more to the study of vectors in n space than just writing down n-tuples. In fact, mathematicians have no interest in objects per se; rather, they wish to study operations that can be performed on these objects.

DEFINITION 1.2. Let $v = (x_1, x_2, \ldots, x_n)$ and $w = (y_1, y_2, \ldots, y_n)$ be vectors in real n space. The *sum* of the vectors v and w, denoted by $v + w$, is the n-tuple

$$v + w = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \tag{1-2}$$

Comment 1.2a. This operation is called *vector addition*; note that the result of this operation, i.e., the sum $v + w$, is again a vector in n space. Observe also that vector addition is defined only for vectors in (the same) n space, e.g., $(3, 1) + (5, 0, -1)$ makes no sense.

Example 1.3

$$(1, -2, -\tfrac{3}{2}, 0) + (3, 7, \tfrac{5}{2}, \pi) = (4, 5, 1, \pi)$$

Comment 1.2b. To find the ith component of $v + w$ we merely add the ith component of v and the ith component of w. Hence in 2 or 3 space, if a vector is identified with an arrow, the vector $v + w$ may be found geometrically as follows: Rigidly translate the arrow that represents w until it starts at the terminal point of v; then the arrow that starts at the origin and terminates at the terminal point of this translate of w represents the vector $v + w$ (see Fig. 1.1). The rule just described is often referred to as the parallelogram law since $v + w$ is represented by a diagonal of the parallelogram determined by v and w.

The next operation differs from vector addition in that it combines not two vectors but rather a real number and a vector. Note, however, that as in vector addition the result of the operation is again a vector.

DEFINITION 1.3. Let $v = (x_1, x_2, \ldots, x_n)$ be a vector in real n space and let α be a real number. The *product* of the scalar α and the vector v, denoted by αv, is the n-tuple.

$$\alpha v = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) \tag{1-3}$$

Comment 1.3a. This operation is called *multiplication of a vector by a scalar*; note that the result of this operation, i.e., αv, is again a vector in n space.

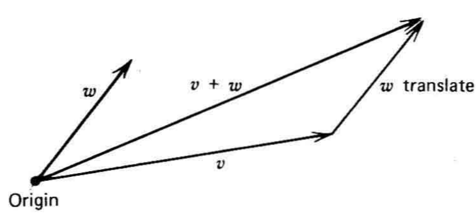

Figure 1.1

Figure 1.2

Example 1.4

$$\pi(3, -\pi, 0, 2/\pi, 0) = (3\pi, -\pi^2, 0, 2, 0)$$

Comment 1.3b. To find the ith component of αv we merely take the product of α and the ith component of v. Hence in 2 or 3 space, if a vector is identified with an arrow, the vectors v and αv must lie on the same straight line through the origin. These arrows point in the same direction if $\alpha > 0$ and in opposite directions if $\alpha < 0$ (see Fig. 1.2). Note that if $\alpha = 0$, then αv is the zero vector.

Notation. The symbol R^n is used to denote the set of all vectors in real n space *together with* the two operations of vector addition and multiplication of a vector by a scalar. (Strictly speaking, R^n is not the set of all n-tuples but that set together with the two operations defined above; this distinction can be safely ignored by the reader.)

A brief comment on R^1 is in order. It consists of 1-tuples; these can in a natural way be identified with the real numbers themselves: Identify the 1-tuple (x) with the real number x. Moreover, addition of 1-tuple looks just like addition of real numbers and multiplication of 1-tuples by scalars looks just like multiplication of real numbers. Thus there is no actual distinction between real numbers and vectors in R^1; nevertheless, such a distinction, artificial as it may seem, will be adhered to in this book. The 1-tuple (x) will be called a vector; the real number x will be called a scalar.

The operations of vector addition and multiplication of a vector by a scalar have many properties. Some of these are listed below. Note that $-v$ denotes the vector all of whose components are the negative of the corresponding components of the vector v, i.e., $-v = (-1)v$.

Properties of Vector Addition

(a) $v + w = w + v$ commutative law
(b) $v + (w + u) = (v + w) + u$ associative law

(c) $v + 0 = v$ zero vector is the identity
(d) $v + -v = 0$ $-v$ is the inverse of v

Properties of Multiplication of a Vector by a Scalar

(e) $1v = v$
(f) $\alpha(\beta v) = \beta(\alpha v) = (\alpha\beta)v$
(g) $\alpha(v + w) = \alpha v + \alpha w$ distributive law
(h) $(\alpha + \beta)v = \alpha v + \beta v$ distributive law

To prove (a), note that if $v = (x_1, x_2, \ldots, x_n)$ and $w = (y_1, y_2, \ldots, y_n)$, then

$$v + w = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

while

$$w + v = (y_1 + x_1, y_2 + x_2, \ldots, y_n + x_n).$$

Since addition of real numbers is commutative, we have that $x_i + y_i = y_i + x_i$ for all i, $i = 1, 2, \ldots, n$. Hence $v + w = w + v$. The proofs of (b–h) are equally simple and are left to the reader (see Exercise 2).

There are other properties that could have been included in our list, e.g., $0v = 0$ and $\alpha 0 = \alpha$. The reader may justifiably inquire why the eight properties above were singled out for special attention; in fact, he might even question the usefulness of describing properties as obvious as those above. The answer to the first question is that it turns out that from the eight properties listed above all other properties of vector addition and multiplication of a vector by a scalar can be derived *without* recourse to the explicit definitions of these operations. The second question is harder to answer. The point is that mathematicians would like to develop for sets bearing no apparent likeness to R^n a structure that resembles as much as possible that of R^n. In turns out that the key to accomplishing this is the ability to find a pair of operations that satisfy the eight properties listed above. Such a system will be called a vector space and an element in it will be called a vector; R^n will be only one—albeit a very important one—example of a vector space. The detailed development of this subject is postponed until the next section.

DEFINITION 1.4. Let $v_1, v_2, \ldots, v_k$ be k vectors in R^n. A *linear combination* of these vectors is any vector of the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \sum_{i=1}^{k} \alpha_i v_i \tag{1-4}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are any k scalars.

Comment 1.4a. A linear combination of one vector v_1 is just any scalar multiple of v_1; a linear combination of two vectors v_1, v_2 is any vector of the form $\alpha_1 v_1 + \alpha_2 v_2$.

Comment 1.4b. Observe that the zero vector is always a linear combination of any k vectors: Choose all the scalars equal to 0.

Example 1.5. Every vector in R^3 can be written as a linear combination of the three vectors (1, 0, 0), (0, 1, 0), and (0, 0, 1). For if $v = (x_1, x_2, x_3)$ is any vector in R^3, we have

$$v = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1) \tag{1-5}$$

This result is perhaps known to the reader, except that he is probably familiar with the vectors (1, 0, 0), (0, 1, 0), and (0, 0, 1) under the aliases of $\vec{i}, \vec{j}$, and $\vec{k}$, respectively.

To generalize the result of Example 1.5 to R^n, consider the n vectors $i_1 = (1, 0, \ldots, 0)$, $i_2 = (0, 1, \ldots, 0), \ldots, i_n = (0, 0, \ldots, 1)$; the kth component of i_k is 1 and all other components are zero. Then any vector $v = (x_1, x_2, \ldots, x_n)$ in R^n is a linear combination of the n vectors $i_1, i_2, \ldots, i_n$ since

$$v = x_1 i_1 + x_2 i_2 + \cdots + x_n i_n = \sum_{k=1}^{n} x_k i_k \tag{1-6}$$

It is left to the reader to show that the scalars in the expansion (1-6) are uniquely determined by v and hence every vector in R^n can be written in exactly one way as a linear combination of the n vectors $i_1, i_2, \ldots, i_n$ (see Exercise 5). For the present, we shall call the set of n vectors $\{i_1, i_2, \ldots, i_n\}$ a *basic set* for R^n.

The reader may have detected a notational abuse above. Note that i_1 in R^1 is the 1-tuple (1), i_1 in R^2 is the 2-tuple (1, 0), i_1 in R^3 is the 3-tuple (1, 0, 0), etc.; all of these different vectors are denoted by the same symbol i_1. Fortunately, the context will always make clear which i_1 is meant.

EXERCISES

2. Verify properties (b–h).

3. Verify that (a) $0v = 0$; (b) $\alpha 0 = 0$; (c) $(-1)v = -v$; (d) $\alpha v = 0$ implies that $\alpha = 0$ or $v = 0$.

4. Let v and w be vectors in R^n and assume that both u_1 and u_2 are linear combinations of v and w, i.e., $u_1 = \alpha v + \beta w$, $u_2 = \gamma v + \delta w$. Show that

(a) $u_1 + u_2$ is a linear combination of v and w

(b) $3u_1$ is a linear combination of v and w

(c) any linear combination of u_1 and u_2 is itself a linear combination of v and w

(d) Is v necessarily a linear combination of u_1 and u_2? If your answer is yes, prove it; if no, give a counterexample.