Werner Greub

Linear Algebra

Fourth Edition

Werner Greub

Linear Algebra

Fourth Edition

Springer-Verlag New York Heidelberg Berlin

Werner Greub

University of Toronto
Department of Mathematics
Toronto M5S 1A1
Canada

AMS Subject Classifications
15-01, 15A03, 15A06, 15A18, 15A21, 16-01

Library of Congress Cataloging in Publication Data

Greub, Werner Hildbert, 1925-
Linear algebra
(Graduate texts in mathematics v. 23)
Bibliography p. 445
1. Algebras, Linear I. Title II. Series
QA184.G7313 1974 512'.5 74-13868

Printed in the United States of America.

ISBN 0-387-90110-8 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90110-8 Springer-Verlag Berlin Heidelberg New York

Preface to the fourth edition

This textbook gives a detailed and comprehensive presentation of linear algebra based on an axiomatic treatment of linear spaces. For this fourth edition some new material has been added to the text, for instance, the intrinsic treatment of the classical adjoint of a linear transformation in Chapter IV, as well as the discussion of quaternions and the classification of associative division algebras in Chapter VII. Chapters XII and XIII have been substantially rewritten for the sake of clarity, but the contents remain basically the same as before. Finally, a number of problems covering new topics – e.g. complex structures, Caylay numbers and symplectic spaces – have been added.

I should like to thank Mr. M.L. Johnson who made many useful suggestions for the problems in the third edition. I am also grateful to my colleague S. Halperin who assisted in the revision of Chapters XII and XIII and to Mr. F. Gomez who helped to prepare the subject index.

Finally, I have to express my deep gratitude to my colleague J.R. Vanstone who worked closely with me in the preparation of all the revisions and additions and who generously helped with the proof reading.

Toronto, February 1975 WERNER H. GREUB

Preface to the third edition

The major change between the second and third edition is the separation of linear and multilinear algebra into two different volumes as well as the incorporation of a great deal of new material. However, the essential character of the book remains the same; in other words, the entire presentation continues to be based on an axiomatic treatment of vector spaces.

In this first volume the restriction to finite dimensional vector spaces has been eliminated except for those results which do not hold in the infinite dimensional case. The restriction of the coefficient field to the real and complex numbers has also been removed and except for chapters VII to XI, § 5 of chapter I and § 8, chapter IV we allow any coefficient field of characteristic zero. In fact, many of the theorems are valid for modules over a commutative ring. Finally, a large number of problems of different degree of difficulty has been added.

Chapter I deals with the general properties of a vector space. The topology of a real vector space of finite dimension is axiomatically characterized in an additional paragraph.

In chapter II the sections on exact sequences, direct decompositions and duality have been greatly expanded. Oriented vector spaces have been incorporated into chapter IV and so chapter V of the second edition has disappeared. Chapter V (algebras) and VI (gradations and homology) are completely new and introduce the reader to the basic concepts associated with these fields. The second volume will depend heavily on some of the material developed in these two chapters.

Chapters X (Inner product spaces) XI (Linear mappings of inner product spaces) XII (Symmetric bilinear functions) XIII (Quadrics) and XIV (Unitary spaces) of the second edition have been renumbered but remain otherwise essentially unchanged.

Chapter XII (Polynomial algebra) is again completely new and developes all the standard material about polynomials in one indeterminate. Most of this is applied in chapter XIII (Theory of a linear transformation). This last chapter is a very much expanded version of chapter XV of the second edition. Of particular importance is the generalization of the results in the second edition to vector spaces over an arbitrary coefficient field of characteristic zero. This has been accomplished without reversion to the cumbersome calculations of the first edition. Furthermore the concept of a semisimple transformation is introduced and treated in some depth.

One additional change has been made: some of the paragraphs or sections have been starred. The rest of the book can be read without reference to this material.

Last but certainly not least, I have to express my sincerest thanks to everyone who has helped in the preparation of this edition. First of all I am particularly indebted to Mr. S. HALPERIN who made a great number of valuable suggestions for improvements. Large parts of the book, in particular chapters XII and XIII are his own work. My warm thanks also go to Mr. L. YONKER, Mr. G. PEDERZOLI and Mr. J. SCHERK who did the proof reading. Furthermore I am grateful to Mrs. V. PEDERZOLI and to Miss M. PETTINGER for their assistance in the preparation of the manuscript. Finally I would like to express my thanks to professor K. BLEULER for providing an agreeable milieu in which to work and to the publishers for their patience and cooperation.

Toronto, December 1966 WERNER H. GREUB

Preface to the second edition

Besides the very obvious change from German to English, the second edition of this book contains many additions as well as a great many other changes. It might even be called a new book altogether were it not for the fact that the essential character of the book has remained the same; in other words, the entire presentation continues to be based on an axiomatic treatment of linear spaces.

In this second edition, the thorough-going restriction to linear spaces of finite dimension has been removed. Another complete change is the restriction to linear spaces with real or complex coefficients, thereby removing a number of relatively involved discussions which did not really contribute substantially to the subject. On p. 6 there is a list of those chapters in which the presentation can be transferred directly to spaces over an arbitrary coefficient field.

Chapter I deals with the general properties of a linear space. Those concepts which are only valid for finitely many dimensions are discussed in a special paragraph.

Chapter II now covers only linear transformations while the treatment of matrices has been delegated to a new chapter, chapter III. The discussion of dual spaces has been changed; dual spaces are now introduced abstractly and the connection with the space of linear functions is not established until later.

Chapters IV and V, dealing with determinants and orientation respectively, do not contain substantial changes. Brief reference should be made here to the new paragraph in chapter IV on the trace of an endomorphism – a concept which is used quite consistently throughout the book from that time on.

Special emphasis is given to tensors. The original chapter on Multilinear Algebra is now spread over four chapters: Multilinear Mappings (Ch. VI), Tensor Algebra (Ch. VII), Exterior Algebra (Ch. VIII) and Duality in Exterior Algebra (Ch. IX). The chapter on multilinear mappings consists now primarily of an introduction to the theory of the tensor-product. In chapter VII the notion of vector-valued tensors has been introduced and used to define the contraction. Furthermore, a treatment of the transformation of tensors under linear mappings has been added. In Chapter VIII the antisymmetry-operator is studied in greater detail and the concept of the skew-symmetric power is introduced. The dual product (Ch. IX) is generalized to mixed tensors. A special paragraph in this chapter covers the skew-symmetric powers of the unit tensor and shows their significance in the characteristic polynomial. The paragraph "Adjoint Tensors" provides a number of applications of the duality theory to certain tensors arising from an endomorphism of the underlying space.

There are no essential changes in Chapter X (Inner product spaces) except for the addition of a short new paragraph on normed linear spaces. In the next chapter, on linear mappings of inner product spaces, the orthogonal projections (§ 3) and the skew mappings (§ 4) are discussed in greater detail. Furthermore, a paragraph on differentiable families of automorphisms has been added here.

Chapter XII (Symmetric Bilinear Functions) contains a new paragraph dealing with Lorentz-transformations.

Whereas the discussion of quadrics in the first edition was limited to quadrics with centers, the second edition covers this topic in full.

The chapter on unitary spaces has been changed to include a more thorough-going presentation of unitary transformations of the complex plane and their relation to the algebra of quaternions.

The restriction to linear spaces with complex or real coefficients has of course greatly simplified the construction of irreducible subspaces in chapter XV. Another essential simplification of this construction was achieved by the simultaneous consideration of the dual mapping. A final paragraph with applications to Lorentz-transformation has been added to this concluding chapter.

Many other minor changes have been incorporated – not least of which are the many additional problems now accompanying each paragraph.

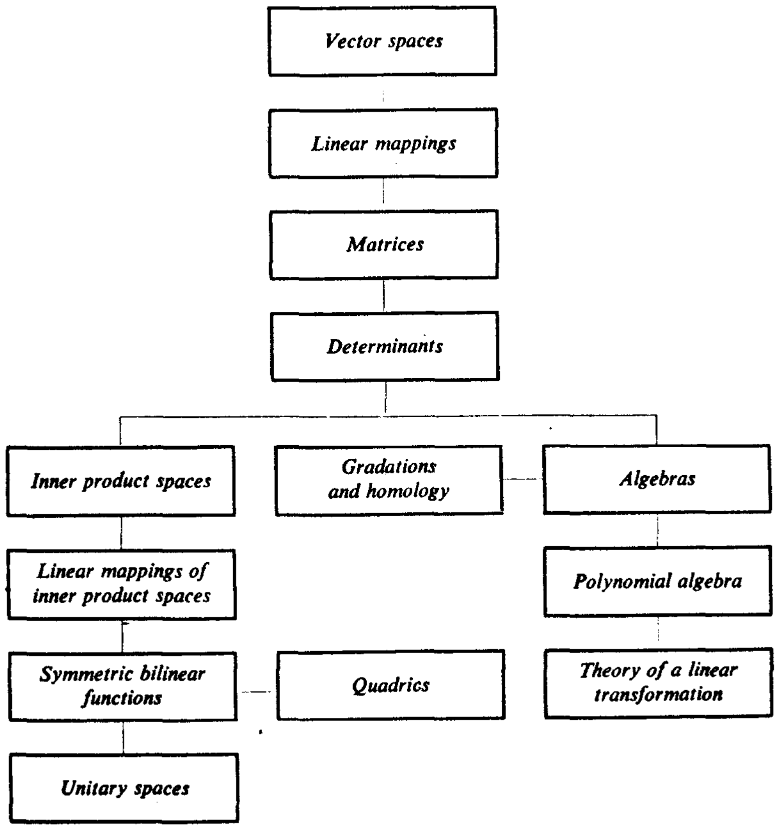
Last, but certainly not least, I have to express my sincerest thanks to everyone who has helped me in the preparation of this second edition. First of all, I am particularly indebted to CORNELIE J. RHEINBOLDT who assisted in the entire translating and editing work and to Dr. WERNER C. RHEINBOLDT who cooperated in this task and who also made a number of valuable suggestions for improvements, especially in the chapters on linear transformations and matrices. My warm thanks also go to Dr. H. BOLDER of the Royal Dutch/Shell Laboratory at Amsterdam for his criticism on the chapter on tensor-products and to Dr. H. H. KELLER who read the entire manuscript and offered many

important suggestions. Furthermore, I am grateful to Mr. GIORGIO PEDERZOLI who helped to read the proofs of the entire work and who collected a number of new problems and to Mr. KHADJA NESAMUDDIN KHAN for his assistance in preparing the manuscript.

Finally I would like to express my thanks to the publishers for their patience and cooperation during the preparation of this edition.

Toronto, April 1963 WERNER H. GREUB

Interdependence of Chapters



Contents

Chapter 0

Prerequisites

0.1. Sets. The reader is expected to be familiar with naive set theory up to the level of the first half of [11]. In general we shall adopt the notations and definitions of that book; however, we make two exceptions. First, the word *function* will in this book have a very restricted meaning, and what Halmos calls a function, we shall call a *mapping* or a *set mapping*. Second, we follow Bourbaki and call mappings that are one-to-one (onto, one-to-one and onto) injective (surjective, bijective).

0.2. Topology. Except for § 5 chap. I, § 8, Chap. IV and parts of chapters VII to IX we make no use at all of topology. For these parts of the book the reader should be familiar with elementary point set topology as found in the first part of [16].

0.3. Groups. A *group* is a set G, together with a binary law of composition

$$\mu: G \times G \to G$$

which satisfies the following axioms ($\mu(x, y)$ will be denoted by xy):

1. *Associativity:* $(xy)z = x(yz)$
2. *Identity:* There exists an element e, called the *identity* such that

$$xe = ex = x.$$

3. To each element $x \in G$ corresponds a second element x^{-1} such that

$$xx^{-1} = x^{-1}x = e.$$

The identity element of a group is uniquely determined and each element has a unique inverse. We also have the relation

$$(xy)^{-1} = y^{-1}x^{-1}.$$

As an example consider the set S_n of all permutations of the set $\{1 \dots n\}$ and define the product of two permutations σ, τ by

$$(\sigma\tau)i = \sigma(\tau i) \qquad i = 1 \dots n.$$

In this way S_n becomes a group, called the *group of permutations of n objects*. The identity element of S_n is the identity permutation.

Let G and H be two groups. Then a mapping

$$\varphi : G \to H$$

is called a *homomorphism* if

$$\varphi(x\,y) = \varphi x\, \varphi y \quad x, y \in G .$$

A homomorphism which is injective (resp. surjective, bijective) is called a monomorphism (resp. epimorphism, isomorphism). The inverse mapping of an isomorphism is clearly again an isomorphism.

A *subgroup* H of a group G is a subset H such that with any two elements $y \in H$ and $z \in H$ the product yz is contained in H and that the inverse of every element of H is again in H. Then the restriction of μ to the subset $H \times H$ makes H into a group.

A group G is called *commutative* or *abelian* if for each $x, y \in G$ $xy = yx$. In an abelian group one often writes $x + y$ instead of xy and calls $x + y$ the *sum* of x and y. Then the unit element is denoted by 0. As an example consider the set $\mathbb{Z}$ of integers and define addition in the usual way.

0.4. Factor groups of commutative groups.* Let G be a commutative group and consider a subgroup H. Then H determines an equivalence relation in G given by

$$x \sim x' \quad \text{if and only if} \quad x - x' \in H .$$

The corresponding equivalence classes are the sets $\{H + x\}$ and are called the *cosets* of H in G. Every element $x \in G$ is contained in precisely one coset $\bar{x}$. The set G/H of these cosets is called the *factor set* of G by H and the surjective mapping

$$\pi : G \to G/H$$

defined by

$$\pi x = \bar{x}, \qquad x \in \bar{x}$$

is called the *canonical projection* of G onto G/H. The set G/H can be made into a group in precisely one way such that the canonical projection becomes a homomorphism; i.e.,

$$\pi(x + y) = \pi x + \pi y . \tag{0.1}$$

To define the addition in G/H let $\bar{x} \in G/H$, $\bar{y} \in G/H$ be arbitrary and choose $x \in G$ and $y \in G$ such that

$$\pi x = \bar{x} \quad \text{and} \quad \pi y = \bar{y} .$$

*) This concept can be generalized to non-commutative groups.

Then the element $\pi(x+y)$ depends only on $\bar{x}$ and $\bar{y}$. In fact, if x', y' are two other elements satisfying $\pi x' = \bar{x}$ and $\pi y' = \bar{y}$ we have

$$x' - x \in H \quad \text{and} \quad y' - y \in H$$

whence

$$(x' + y') - (x + y) \in H$$

and so $\pi(x'+y') = \pi(x+y)$. Hence, it makes sense to define the sum $\bar{x}+\bar{y}$ by

$$\bar{x} + \bar{y} = \pi(x + y) \qquad \pi x = \bar{x}, \pi y = \bar{y}.$$

It is easy to verify that the above sum satisfies the group axioms. Relation (0.1) is an immediate consequence of the definition of the sum in G/H. Finally, since π is a surjective map, the addition in G/H is uniquely determined by (0.1).

The group G/H is called the *factor group of G with respect to the subgroup H*. Its unit element is the set H.

0.5. Fields. A *field* is a set Γ on which two binary laws of composition, called respectively addition and multiplication, are defined such that

1. Γ is a commutative group with respect to the addition.
2. The set $\Gamma - \{0\}$ is a commutative group with respect to the multiplication.
3. Addition and multiplication are connected by the *distributive law*,

$$(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma, \qquad \alpha, \beta, \gamma \in \Gamma.$$

The rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$ are fields with respect to the usual operations, as will be assumed without proof.

A homomorphism $\varphi: \Gamma \to \Gamma'$ between two fields is a mapping that preserves addition and multiplication.

A subset $\Delta \subset \Gamma$ of a field which is closed under addition, multiplication and the taking of inverses is called a *subfield*. If Δ is a subfield of Γ, Γ is called an *extension field* of Δ.

Given a field Γ we define for every positive integer k the element $k\varepsilon$ (ε unit element of Γ) by

$$k\varepsilon = \underbrace{\varepsilon + \cdots + \varepsilon}_{k}$$

The field Γ is said to have *characteristic zero* if $k\varepsilon \neq 0$ for every positive integer k. If Γ has characteristic zero it follows that $k\varepsilon \neq k'\varepsilon$ whenever $k \neq k'$. Hence, a field of characteristic zero is an infinite set. Throughout this book it will be assumed without explicit mention that all fields are of characteristic zero.

For more details on groups and fields the reader is referred to [29].

0.6. Partial order. Let $\mathscr{A}$ be a set and assume that for some pairs X, Y ($X \in \mathscr{A}$, $Y \in \mathscr{A}$) a relation, denoted by $X \leqq Y$, is defined which satisfies the following conditions:

(i) $X \leqq X$ for every $X \in \mathscr{A}$ (Reflexivity)

(ii) if $X \leqq Y$ and $Y \leqq X$ then $X = Y$ (Antisymmetry)

(iii) If $X \leqq Y$ and $Y \leqq Z$, then $X \leqq Z$ (Transitivity).

Then $\leqq$ is called a *partial order* in $\mathscr{A}$.

A *homomorphism of partially ordered sets* is a map $\varphi: \mathscr{A} \to \mathscr{B}$ such that $\varphi X \leqq \varphi Y$ whenever $X \leqq Y$.

Clearly a subset of a partially ordered set is again partially ordered.

Let $\mathscr{A}$ be a partially ordered set and suppose $A \in \mathscr{A}$ is an element such that the relation $A \leqq X$ implies that $A = X$. Then A is called a *maximal element* of $\mathscr{A}$. A partial ordered set need not have a maximal element.

A partially ordered set is called *linearly ordered* or a *chain* if for every pair X, Y either $X \leqq Y$ or $Y \leqq X$.

Let $\mathscr{A}_1$ be a subset of the partially ordered set $\mathscr{A}$. Then an element $A \in \mathscr{A}$ is called an *upper bound* for $\mathscr{A}_1$ if $X \leqq A$ for every $X \in \mathscr{A}_1$.

In this book we shall assume the following axiom:

A partially ordered set in which every chain has an upper bound, contains a maximal element.

This axiom is known as Zorn's lemma, and is equivalent to the axiom of choice (cf. [11]).

0.7. Lattices. Let $\mathscr{A}$ be a partially ordered set and let $\mathscr{A}_1 \subset \mathscr{A}$ be a subset. An element $A \in \mathscr{A}$ is called a *least upper bound* (l.u.b.) for $\mathscr{A}_1$ if

1) A is an upper bound for $\mathscr{A}_1$.

2) If X is any upper bound, then $A \leqq X$. It follows from (ii) that if a l.u.b. for $\mathscr{A}_1$ exists, then it is unique.

In a similar way, lower bounds and the greatest lower bound (g.l.b.) for a subset of $\mathscr{A}$ are defined.

A partially ordered set $\mathscr{A}$ is called a *lattice*, if for any two elements X, Y the subset $\{X, Y\}$ has a l.u.b. and a g.l.b. They are denoted by $X \vee Y$ and $X \wedge Y$. It is easily checked that any finite subset $(X_1, \ldots, X_r)$ of a lattice has a l.u.b. and a g.l.b. They are denoted by $\bigvee_{i=1}^{r} X_i$ and $\bigwedge_{i=1}^{r} X_i$.

As an example of a lattice, consider the collection of subsets of a given set, X, ordered by inclusion. If U, V are any two subsets, then

$$U \wedge V = U \cap V \quad \text{and} \quad U \vee V = U \cup V.$$

Chapter I

Vector Spaces

§ 1. Vector spaces

1.1. Definition. A *vector (linear) space*, E, over the field Γ is a set of elements $x, y, \ldots$ called *vectors* with the following algebraic structure:

I. E is an additive group; that is, there is a fixed mapping $E \times E \to E$ denoted by

$$(x, y) \to x + y \tag{1.1}$$

and satisfying the following axioms:

I.1. $(x+y)+z=x+(y+z)$ (associative law)

I.2. $x+y=y+x$ (commutative law)

I.3. there exists a zero-vector 0; i.e., a vector such that $x+0=0+x=x$ for every $x \in E$.

I.4. To every vector x there is a vector $-x$ such that $x+(-x)=0$.

II. There is a fixed mapping $\Gamma \times E \to E$ denoted by

$$(\lambda, x) \to \lambda x \tag{1.2}$$

and satisfying the axioms:

II.1. $(\lambda\mu)x=\lambda(\mu x)$ (associative law)

II.2. $(\lambda+\mu)x=\lambda x+\mu x$
$\lambda(x+y)=\lambda x+\lambda y$ (distributive laws)

II.3. $1 \cdot x = x$ (1 unit element of Γ)

(The reader should note that in the left hand side of the first distributive law, + denotes the addition in Γ while in the right hand side, + denotes the addition in E. In the sequel, the name addition and the symbol + will continue to be used for both operations, but it will always be clear from the context which one is meant). Γ is called the *coefficient field* of the vector space E, and the elements of Γ are called *scalars*. Thus the mapping