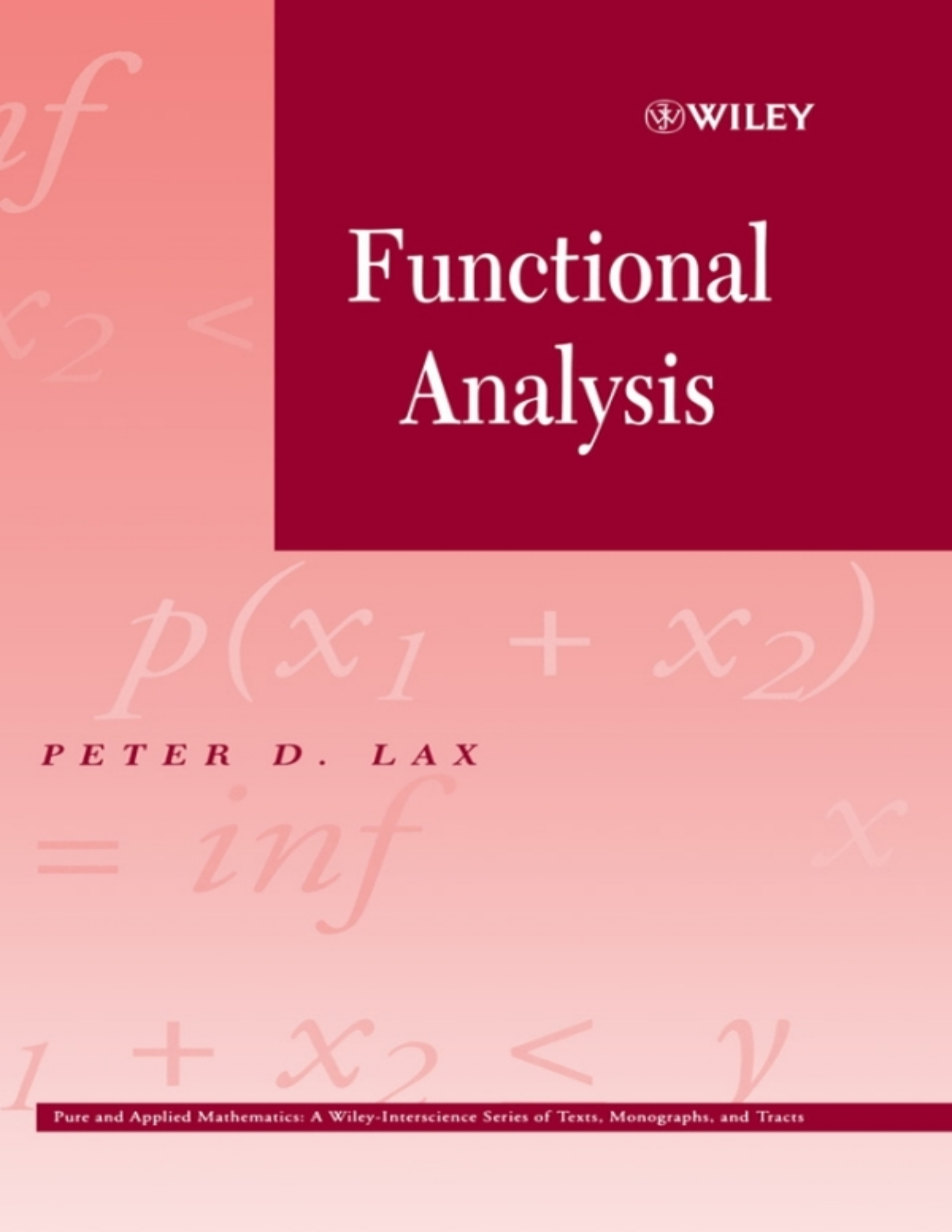

WILEY

Functional Analysis

PETER D. LAX

Pure and Applied Mathematics: A Wiley-Interscience Series of Texts, Monographs, and Tracts

PURE AND APPLIED MATHEMATICS

A Wiley-Interscience Series of Texts, Monographs, and Tracts

A complete list of the titles in this series appears at the end of this volume.

FUNCTIONAL ANALYSIS

PETER D. LAX
Courant Institute
New York University

A JOHN WILEY & SONS, INC., PUBLICATION

Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data

Lax, Peter D.

Functional analysis / Peter D. Lax.
p. cm.
Includes bibliographical references and index.
ISBN 0-471-55604-1 (cloth : alk. paper)
1. Functional analysis. I. Title.

QA320.L345 2002

515—dc21 2001046547

20 19 18

CONTENTS

FOREWORD

This book grew out of a course of lectures on functional analysis taught over many years to second-year graduate students at the Courant Institute of New York University. It is a graduate text, not a treatise or a monograph. Most of the chapters are short, for it is easier to digest material in small chunks. Not all topics can be presented briefly, so some of the chapters are longer. Theorems and lemmas, as well as equations, are numbered consecutively in each chapter.

The first 23 chapters make only a modest technical demand on the reader; this material would serve very well as text for an introductory graduate course on functional analysis. The rest of the material is well suited as text for a more advanced graduate course on functional analysis in general, or on Hilbert space in particular.

When I was a student, the only text on functional analysis was Banach's original classic, written in 1932; Hille's book appeared in time to serve as my graduation present. For Hilbert space there was Stone's Colloquium publication, also from 1932, and Sz.-Nagy's Ergebnisse volume. Since then, our cup hath run over; first came Riesz and Sz.-Nagy, then Dunford and Schwartz, Yosida, later Reed and Simon, and Rudin. For Hilbert space, there was Halmos's elegant slender volume, and Achiezer and Glazman, all of which I read with pleasure and profit. Many, many more good texts have appeared since. Yet I believe that my book offers something new: the order in which the material is arranged, the interspersing of chapters on theory with chapters on applications, so that cold abstractions are made flesh and blood, and the inclusion of a very rich fare of mathematical problems that can be clarified and solved from the functional analytic point of view.

In choosing topics I heeded the warning of my teacher Friedrichs: “It is easy to write a book if you are willing to put into it everything you know about the subject.” I present the basic structure of the subject, and those more advanced topics that loom large in the body of mathematics. Among these are the spectral resolution and spectral representation of self-adjoint operators, the theory of compact operators, the Krein-Milman theorem, Gelfand’s theory of commutative Banach algebras, invariant subspaces, strongly continuous one-parameter semigroups. I discuss the index of operators, so important in calculating topological invariants; the celebrated trace formula of Lidskii, a powerful tool in analysis; the Fredholm determinant and its generalizations, rising again after almost a hundred years of hibernation; and scattering theory, another gift from physics to mathematics. I have also included some (but not all) special topics close to my heart.

What has been omitted? All of nonlinear functional analysis, for which I recommend the four-volume treatise by Zeidler. Operator algebras, except for Gelfand’s theory of commutative Banach algebras. I slight the geometric theory of Banach spaces; happily a handbook on this subject, edited by Bill Johnson and Joram Lindenstrauss is about to be published by North Holland.

What are the prerequisites? What every second-year graduate student-and many undergraduates—knows:

- *Naive set theory.* Denumerable sets, the continuum, Zorn’s lemma.
- *Linear algebra.* The alternative for linear maps, trace and determinant of a matrix, the spectral theory of general and symmetric matrices, functions of a matrix.
- *Point set topology.* Complete metric spaces, the Baire category principle, Hausdorff spaces, compact sets, Tychonov’s theorem.
- *Basic theory of functions of a complex variable.*

- *Real variables.* The Arzela-Ascoli theorem, the Lebesgue decomposition of measures on $\mathbb{R}$, Borel measure on compact spaces.

It is an accident of history that measure theory was invented before functional analysis. The usual presentations of measure theory fail to take advantage of the concepts and constructions of functional analysis. In an appendix on the Riesz-Kakutani representation theorem I show how to use the tools of functional analysis in measure theory. Another appendix summarizes the basic facts of Laurent Schwartz's theory of distributions.

Many of the applications are to problems of partial differential equations. Here a nodding aquaintance with the Laplace and the wave equation would help, although an alert uninformed reader could pick up some of the basic results from these pages.

Like most mathematicians, I am no historian. Yet I have included historical remarks in some of the chapters, mainly where I had some firsthand knowledge, or where conventional history has been blatantly silent concerning the tragic fate of many of the founding fathers of functional analysis during the European horrors of the 1930s and 1940s.

I am indebted to many. I learned the rudiments of functional analysis, and how to apply them, from my teacher Friedrichs. Subsequently my views were shaped by the work of Tosio Kato, who has brought the power of functional analysis to bear on an astonishing range of problems. My happy and long collaboration with Ralph Phillips has led to some unusual uses of functional analysis. r learned much from Israel Gohberg, especially about the index of Toeplitz operators, from Bill Johnson about the fine points of the geometry of Banach spaces, and from Bob Phelps about Choquet's theorem. I thank Reuben Hersh and Louise Raphael for their critique of the appendix on

distributions, and Jerry Goldstein for his expert comments on the material on semigroups and scattering theory. To all of them, as well as to Gabor Francsics, my thanks.

Jerry Berkowitz and I alternated teaching functional analysis at the Courant Institute. This would be a better book had he lived and looked the manuscript over critically.

I thank Jeff Rosenbluth and Paul Chernoff for a careful reading of the early chapters, and Keisha Grady for TEXing the manuscript, and cheerfully making subsequent changes and corrections.

The lecture course was popular and successful with graduate students of the Courant Institute. I hope this printed version retains the spirit of the lectures.

PETER D. LAX

New York. NY
November 2001

1

LINEAR SPACES

A *linear space X* over a *field* **F** is a mathematical object in which two operations are defined: addition and multiplication by scalars.

Addition, denoted by +,as in

(1) $x+y$

is assumed to be *commutative*,

(2) $x + y = y + x,$

associative,

(3) $x + (y + z) = (x + y) + z,$

and to form a *group*, with the neutral element denoted as 0:

(4) $x + 0 = x.$

The inverse of addition is denoted by-:

(5) $x + (-x) \equiv x - x = 0.$

The second operation is the *multiplication* of elements of *X by elements k of the field* **F**:

$kx.$

The result of this multiplication is again an element of *X*. Multiplication by elements of **F** is assumed to be *associative*,

(6) $k(ax) = (ka)x,$

and *distributive*,

(7) $k(x + y) = kx + ky$

as well as

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

We assume that multiplication by the *unit* of **F**, denoted as 1, acts as the identity:

(9) $1x = x.$

These are the *axioms* of linear algebra. From them proceed to draw some deductions.

Set $b = 0$ in (8). It follows that for all x,

(10) $0x = 0.$

Set $a = 1$, $b = -1$ in (8). Using (9) and (10), we deduce that for all x,

(11) $(-1)x = -x.$

The finite-dimensional linear spaces are dealt with in courses on linear algebra. In this book the emphasis is on the infinite-dimensional ones—those that are not finitedimensional. The field **F** will be either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$ Here are some examples.

Example 1. X is the space of all polynomials in a single variable s, with real coefficients, here **F** = $\mathbb{R}$

Example 2. X is the space of all polynomials in N variables $s_1, \ldots, s_N$, with real coefficients, here **F** = $\mathbb{R}$

Example 3. G is a domain in the complex plane, and X the space of all functions complex analytic in G, here **F** = $\mathbb{C}$.

Example 4. X = space of all vectors

$$x = (a_1, a_2, \ldots)$$

with infinitely many real components, here **F** = $\mathbb{R}$

Example 5. Q is a Hausdorff space, X the space of all continuous real-valued functions on Q, here **F** = $\mathbb{R}$

Example 6. M is a C^∞ differentiable manifold, $X = C^\infty(M)$, the space of all differentiable functions on M.

Example 7. Q is a measure space with measure m, $X = L^1(Q, m)$.

Example 8. $X = L^P(Q,m)$.

Example 9. X = harmonic functions in the upper half-plane.

Example 10. X = all solutions of a linear partial differential equation in a given domain.

Example 11. All meromorphic functions on a given Riemann surface; **F** = $\mathbb{C}$.

We start the development of the theory by giving the basic *constructions* and *concepts*. Given two subsets S and T of a linear space X, we define their *sum*, denoted as $S + T$ to be the set of all points x of the form $x = y + z$, y in S, z in T. The *negative* of a set S, denoted as $-S$, consists of all points x of the form $x = -y$, y in S.

Given two linear spaces Z and U over the same field, their *direct sum* is a linear space denoted as $Z \oplus U$, consisting of *ordered pairs* (z,u), z in Z, u in U. Addition and multiplication by scalars is componentwise.

Definition. A subset Y of a linear space X is called a *linear subspace* of X if sums and scalar multiples of Y belong to Y.

Theorem 1.

(i) The sets $\{0\}$ *and* X *are linear subspaces of* X.

(ii) The sum of any collection of subspaces is a subspace.

(iii) The intersection of any collection of subspaces is a subspace.

(iv) The union of a collection of subspaces totally ordered by inclusion is a sub-space.

Exercise 1. Prove theorem 1.

Let S be some subset of the linear space X. Consider the collection $\{Y_\sigma\}$ of all linear subspaces that contain the setS. This collection is not empty, since it certainly contains X.

Definition. The intersection $\cap Y_\sigma$ of all linear subspaces Y_σ containing the set S is called the *linear span* of the set S.

Theorem2.

(i) The linear span of a set S is the smallest linear subspace containing S.

(ii) The linear span of S consists of all elements x of the form

$$x = \sum_1^n a_i x_i, \qquad x_i \in S, a_i \in \mathbf{F}, n \text{ any natural number.} \tag{12}$$

Proof Part (i) is merely a rephrasing of the definition of linear span. To prove part (ii), we remark that on the one hand, the elements of the form (12) form a linear subspace of X; on the other hand, every x of form (12) is contained in any subspace Y containing S.

REMARK 1. An element x of form (12) is called a *linear combination* of the points $x_1, \ldots, x_n$. So theorem 1 can be restated as follows:

The linear span of a subset S of a linear space consists of all linear combinations of elements of S.

Definition. X a linear space, Y a linear subspace of X. Two points x_1 and x_2 of X are called *equivalent* modulo Y, denoted as $x_1 = x_2$ (mod Y), if $x_1 - x_2$ belongs to Y.

It follows from the properties of addition that equivalence mod Y is an equivalence relation, meaning that it is symmetric, reflexive, and transitive. That being the case, we can divide X into distinct equivalence classes mod Y. We denote the set of equivalence classes as X / Y. The set X / Y has a natural linear structure; the sum of two equivalence classes is defined by choosing arbitrary points in each equivalence class, adding them and forming the equivalence class of the sum. It is easy to check that the last equivalence class is independent of the representatives we picked; put differently, if $x_1 \equiv z_1$, $x_2 \equiv z_2$. then $x_1 + x_2 \equiv z_1 + z_2$ mod Y. Similarly we define multiplication by a scalar by picking arbitrary elements in the equivalence class. The resulting operation does not depend on the choice, since, if $x_1 \equiv z_1$, then $kx_1 \equiv kz_1$ mod Y.The quotient set X / Y endowed with this natural linear structure is called

the *quotient space* of X mod Y. We define codim Y = dim X / Y.

Exercise 2. Verify the assertions made above.

As with all algebraic structures, so with linear structures we have the concept of *isomorphism*.

Definition. Two linear spaces X and Z over the same field are *isomorphic* if there is a one-to-one correspondence **T** carrying one into the other that maps sums into sums, scalar multiples into scalar multiples; that is,

$$\mathbf{T}(x_1 + x_2) = \mathbf{T}(x_1) + \mathbf{T}(x_2),$$
$$\mathbf{T}(kx) = k\mathbf{T}(x). \tag{13}$$

We define similarly *homomorphism*, called in this context a *linear map*.

Definition. X and U are linear spaces over the same field. A *mapping* $\mathbf{M} : X \to U$ is called *linear* if it carries sums into sums, and scalar multiples into scalar multiples; that is, if for all x, y in X and all k in **F**

$$\mathbf{M}(x + y) = \mathbf{M}(x) + \mathbf{M}(y),$$
$$\mathbf{M}(kx) = k\mathbf{M}(x). \tag{14}$$

X is called the domain of **M**, U its target.

REMARK 2. An isomorphism of linear spaces is a linear map that is one-to-one and onto.

Theorem 3.

(i) The image of a linear subspace Y of X under a linear map $\mathbf{M} : X \to U$ *is a linear subspace of U.*

(ii) The inverse imaRe under **M** *of a linear subspace V of U is a linear subspace of X.*

Exercise 3. Prove theorem 3.

A very important concept in a linear space over the *reals* is *convexity:*

Definition. X is a linear space over the reals; a subset K of X is called *convex* if, whenever x and y belong to K, the whole segment with endpoints x, y, meaning all points of the form

(15) $ax + (1-a)y, \qquad 0 \le a \le 1,$

also belong to K.

Examples of convex sets in the plane are the circular disk, triangle, and semicircular disk. The following property of convex sets is an immediate consequence of the definition:

Theorem 4. *Let K be a convex subset of a linear space X over the reals. Suppose that $x_1, \ldots, x_n$ belong to K; then so does every x of the form*

$$x = \sum_1^n a_j x_j, \qquad a_j \ge 0,$$

(16) $$\sum_1^n a_j = 1.$$

Exercise 4. Prove theorem 4.

An x of form (16) is called a *convex combination* of $x_1, x_2, \ldots, x_n$.

Theorem 5. *Let X be a linear space over the reals.*

(i) The empty set is convex.

(ii) A subset consisting of a single point is convex.

(iii) Every linear subspace of X is convex.

(iv) The sum of two convex subsets is convex.

(v) If K is convex, so is $-K$.

(vi) The intersection of an arbitrary collection of convex sets is convex.

(vii) Let $\{K_j\}$ be a collection of convex subsets that is totally ordered by inclusion. Then their union $\cup K_j$ is convex.

(viii) The image of a convex set under a linear map is convex.

(ix) The inverse image of a convex set under a linear map is convex.

Exercise 5. Prove theorem 5.

Definition. Let S be any subset of a linear space X over the reals. The *convex hull* of S is defined as the *intersection* of all convex sets containing S. The hull is denoted as $\hat{S}$.

Theorem 6.

(i) The convex hull of S is the smallest convex set containing S.

(ii) The convex hull of S consists of all convex combinations (*16*) *of points of S.*

Exercise 6. Prove theorem 6.

Definition. A subset E of a convex set K is called an *extreme subset* of K if:

(i) E is convex and nonempty.

(ii) whenever a point x of E is expressed as

$$x = \frac{y+z}{2}, \qquad y, z \text{ in } K,$$

then both y and z belong to E.

An extreme subset consisting of a single point is called an *extreme point* of K.

Example 1. K is the interval $0 \leq x \leq 1$; the two endpoints are extreme points.

Example 2. K is the closed disk

$$x^2 + y^2 \leq 1.$$

Every point on the circle $x^2 + y_2 = 1$ is an extreme point.

Example 3. The open disk

$$x^2 + y^2 < 1$$

has no extreme points.

Example 4. K a polyhedron, including faces. Its extreme subsets are its faces, edges, vertices, and of course K itself.

Theorem 7. *Let K be a convex set, E an extreme subset of K, and Fan extreme subset of E. Then F is an extreme subset of K.*

Exercise 7. Prove theorem 7.

Theorem 8. *Let* **M** *be a linear map of the linear space X into the linear space U. Let K be a convex subset of U, E an extreme subset of K. Then the inverse image of E is either empty or an extreme subset of the inverse image of K*.

Exercise 8. Prove theorem 8.

Exercise 9. Give an example to show that the image of an extreme subset under a linear map need not be an extreme subset of the image.

Taking *U* to be one dimensional, we get

Corollary 8′. *Denote by H a convex subset of a linear space X, ℓ a linear map of X into* $\mathbb{R}$, $H_{\min}$ *and* $H_{\max}$ *the subsets of H, where ℓ achieves its minimum and maximum, respectively*.

Assertion. *When nonempty,* $H_{\min}$ *and* $H_{\max}$ *are extreme subsets of H*.

2

LINEAR MAPS

2.1 ALGEBRA OF LINEAR MAPS

We recall from chapter 1 that a *linear map* from one linear space X into another, U, both over 1he same field of scalars, is a mapping of X into U,

$$\mathbf{M} : X \longrightarrow U,$$

that is an algebraic homomorphism:

$$\mathbf{M}(x + y) = \mathbf{M}(x) + \mathbf{M}(y),$$

(1) $$\mathbf{M}(kx) = k\mathbf{M}(x).$$

In this section we explore those properties of linear maps that depend on the purely algebraic properties (1), without any topological restrictions imposed on the spaces X,U.

The *sum* of two linear maps **M** and **N** of X into U, and the *scalar multiple* is defined as

(2) $$(\mathbf{M} + \mathbf{N})(x) = \mathbf{M}(x) + \mathbf{N}(x),$$

(3) $$(k\mathbf{M})(x) = k\mathbf{M}(x).$$

This makes a *linear space* out of the set of linear maps of X into U. The space is denoted as $\mathcal{L}(X, U)$. Given two linear maps, one, **M** from $X \to U$, the other, **N** from $U \to W$, we can define their *product* as the *composite map*

(4) $$(\mathbf{NM})(x) = \mathbf{N}(\mathbf{M}(x)).$$

Since compositon of maps in general is *associative*, so is in particular the composition of linear maps. As we will see, composition is far from being commutative.

From now on we omit the bracket and denote the action of a linear map on x as

$$\mathbf{M}(x) = \mathbf{M}x.$$

This notation suggests that the action of **M** on x is a kind of multiplication; indeed (1) and (2) give the distributive property of this kind of multiplication.

Exercise 1. Verify that the composite of two linear maps is linear, and that the distributive law holds:

$$\mathbf{M}(\mathbf{N} + \mathbf{K}) = \mathbf{MN} + \mathbf{MK},$$
$$(\mathbf{M} + \mathbf{K})\mathbf{N} = \mathbf{MN} + \mathbf{KN}.$$

Definition. A mapping is *invertible* if it maps X one-to-one and onto U.

If **M** is invertible, it has an inverse, denoted as $\mathbf{M}^{-1}$, that satisfies

$$\mathbf{M}^{-1}\mathbf{M} = \mathbf{I}, \quad \mathbf{M}\mathbf{M}^{-1} = \mathbf{I},$$

where **I** on the left is the *identity mapping* in X, on the right on U. If **M** is linear, so is $\mathbf{M}^{-1}$.

Definition. The *nullspace* of **M**, denoted by $N_{\mathbf{M}}$, is the set of points mapped into zero.

The *range* of **M**, denoted by $R_{\mathbf{M}}$, is the image of X under **M** in U.

Theorem 1. *Let* **M** *be a linear map of* $X \to U$.

(i) The nullspace N_{M} is a linear subspace of X, the range $R_{\mathbf{M}}$ a linear subspace of U.

(ii) **M** is invertible iff $N_{\mathbf{M}} = \{0\}$ and $R_{\mathbf{M}} = U$.

(iii) **M** maps the quotient space $X/N_{\mathbf{M}}$ one-to-one onto $R_{\mathbf{M}}$.

(iv) If $\mathbf{M} : X \to U$ and $\mathbf{K} : U \to W$ are both invertible, so is their product, and

$$(\mathbf{KM})^{-1} = \mathbf{M}^{-1}\mathbf{K}^{-1}.$$

(v) If **KM** is invertible, then

$$N_{\mathbf{M}} = \{0\}, \quad R_{\mathbf{K}} = W.$$

Exercise 2. Prove theorem 1.

We remark that when $x = U = W$ are finite dimensional, then the invertibility of the product **NM** implies that **N** and **M** separately are invertible. This is not so in the infinite-dimensional case; take, for instance, X to be the space of infinite sequences

$$x = (a_1, a_2, \ldots)$$

and define **R** and **L** to be right and left shift: $\mathbf{R}x = (0, a_1, a_2, \ldots)$, $\mathbf{L}x = (a_2, a_3, \ldots)$. Clearly, **LR** is the identity map, but neither **R** nor **L** are invertible; nor is **RL** the identity.

We formulate now a number of useful notions and results concerning mappings of a linear space *into itself:*

$$\mathbf{M} : X \longrightarrow X.$$

We denote by N_j the *nullspace* of the jth power of **M**:

(5) $$N_j = N_{\mathbf{M}^j}.$$

Theorem 2. *The subspaces N_j defined in (5) have these properties:*

(6) $$N_j \subset N_{j+1} \qquad \text{for all } j$$

and

(7) $$\dim\left(\frac{N_j}{N_{j-1}}\right) \geq \dim\left(\frac{N_{j+1}}{N_j}\right) \qquad \text{for all } j.$$

Proof. Equation (6) is an immediate consequence of (5). To show (7), we claim that **M** maps N_{j+1}/N_j into N_j/N_{j-1} in a one-to-one fashion. To see this, note that a nonzero element of N_{j+1}/N_j is represented by a point z in N_{j+1} that does not lie in N_j. Clearly, $\mathbf{M}z$ lies in N_j but not in N_{j-1}; this shows the one-to-oneness. It follows that N_{j+1}/N_j is isomorphic to a subspace of N_j/N_{j-1}, from which the statement (7) about dimension follows. When N_{j+1}/N_j is infinite-dimensional, so is N_j/N_{j-1}.