Functional Analysis in Modern Applied Mathematics

RUTH F. CURTAIN

A. J. PRITCHARD

Functional Analysis in Modern Applied Mathematics

RUTH F. CURTAIN

A. J. PRITCHARD

Control Theory Centre
University of Warwick
Coventry, U.K.

This is Volume 132 in
MATHEMATICS IN SCIENCE AND ENGINEERING
A Series of Monographs and Textbooks
Edited by RICHARD BELLMAN, *University of Southern California*

The complete listing of books in this series is available from the Publisher upon request.

1977

Academic Press
London New York San Francisco
A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS INC. (LONDON) LTD.
24/28 Oval Road
London NW1

United States Edition published by
ACADEMIC PRESS INC.
111 Fifth Avenue
New York, New York 100003

Library of Congress Catalog Card Number: 77-71795
ISBN: 0-12-196250-4

Printed in Great Britain by Page Bros. (Norwich) Ltd., Norwich

Introduction

One of the most important problems for applied mathematicians, theoretical scientists or systems analysts is the investigation of a system by first obtaining a mathematical model, and then determining such properties as existence, uniqueness and regularity of solutions, stability of equilibrium points, controllability and so on. Although these modelling techniques are usually acquired from experience in a particular field, there is often a parallel technique in some other field. Whereas formulating and examining these models usually involves the use of *a priori* knowledge of the system and extensive manipulation the main ideas can be expressed very simply. In order to teach these skills it is necessary to use a framework which allows a large class of systems to be considered in the same formulation. Such an abstract approach is provided by functional analysis. In fact there is an ever increasing literature in engineering, theoretical physics, applied mathematics, economics and other applied fields written in this mathematical language which to the uninitiated seems very abstract and incomprehensible. So what is functional analysis, and why has it become so fashionable?

First you may ask if functional analysis is some powerful technique which leads to solutions unobtainable by traditional methods. Unfortunately this is rarely the case. The strength and appeal of functional analysis is that it is a convenient way of examining the behaviour of various mathematical models, and it clarifies, rigorizes, and unifies the underlying concepts.

It clarifies because functional analysis is a generalization and combination of linear algebra, analysis, and geometry expressed in a simple mathematical notation which allows these three aspects of the problem to be easily seen. It rigorizes, because it has the back up of a vast mathematical machinery which subsumes many of the classical results on differential equations, analysis, numerical methods, and applied mathematical techniques. It unifies, because the simple notation does away with many of the complicating details leaving the essentials standing out clearly, so that problems from many different fields have the same functional analytical symbolism.

Functional analysis is the mathematicians "black box" diagram where inputs and outputs belong to spaces and the black box is an operator.

There are several excellent books on functional analysis and we have referenced some of these at the end of each chapter. However in the last few years research workers in applied sciences have sought a working knowledge of functional analysis which would help them in their particular field, and this is not readily obtained by reading a standard text on functional analysis. This demand has been recognised by various institutes and we have lectured at vacation schools for the Institute of Mathematics and its Applications (IMA) at Warwick, U.K., Institute of Electrical Engineers (IEE) at Oxford, U.K., and United Nations Educations, Scientific and Cultural Organization (UNESCO) at Trieste, Italy. Although there are some excellent books which specialize in a particular area of application such as Blum [1] and Luenberger [2], we felt that there was a need for a book which illustrated the application of functional analysis in a variety of fields. Quite apart from the research considerations we have been concerned with establishing and teaching a three year undergraduate course in Modern Applied Mathematics at the University of Warwick. The students take courses in algebra, analysis, sets and groups, and differential equations in their first two years and we have taught most of the material in this book in third year courses on Applied Functional Analysis, Modern Control Theory, and Stability Theory. From this experience we have found that these courses provide a valuable link between pure mathematics and applied mathematics which is often missing from more traditional applied mathematics courses.

We expect the reader to have had a first course in functional analysis or to have read a book like Naylor and Sell [3]. So the first third of the book consists of a sequence of definitions and theorems interleavened with many examples which we hope will illustrate the main ideas. The purpose is to establish a uniform notation and cover the background material in topological spaces, linear operators and calculus which we require for the applications sections. The next sections on differential equations and spectral theory are also contained in many books on functional analysis, but because of the particularly important role of these areas in applied mathematics we have included most of the proofs of the fundamental results. The remaining applications sections are self-contained with complete proofs and we feel that the chapters on stability, linear systems theory, optimization and numerical methods could form the basis of lecture courses for final year undergraduates. Although the material we have chosen to illustrate the usefulness of functional analysis is entirely subjective, we have included references at the end of each chapter which give a much more complete picture of each application area.

Finally the chapter on infinite dimensional control theory reflects our research interests and in fact comprises part of a Masters Course in Control Theory we teach at Warwick. Consequently it is not appropriate in an undergraduate course. Our reason for including this chapter is to illustrate how, by using functional analysis, it is possible to study very large classes of systems with very different physical behaviour using the same mathematical formulation.

Warwick R.F.C.
January 1977 A.J.P.

REFERENCES

1. E. K. Blum. "Numerical Analysis and Computation Theory and Practice". Addison-Wesley. 1972.
2. D. G. Luenberger. "Optimisation by Vector Space Methods", Wiley, New York, 1969.
3. A. W. Naylor and G. R. Sell. "Linear Operators in Engineering and Science", Holt Rinehart and Winson, Inc. 1971.

Contents

Part III
Applications

PART I

Basic Functional Analysis

CHAPTER 1

Normed Linear Spaces

The concept of a normed linear space is fundamental to functional analysis and is most easily thought of as a generalization of n-dimensional Euclidean vector space $\mathscr{R}^n = \{x : x = (x_1, \dots, x_n),\ x_i \in \mathscr{R}$, the real numbers$\}$ with the euclidean length function $\|.\| : \mathscr{R}^n \to \mathscr{R}^+$ (the positive real numbers) given by

$$\|x\| = \sum_{i=1}^{n} |x_i|^2$$

In fact it is just a linear vector space with a length function or norm defined on it. First we review the basic properties of linear vector spaces.

Definition 1.1. Linear vector space

A linear vector space W over a scalar field $\mathscr{F}$ is a nonempty set W with a mapping: $(x_1, x_2) \to x_1 \oplus x_2$ from $W \times W$ into W, which we call addition, and a mapping: $(\alpha, x) \to \alpha x$ from $\mathscr{F} \times W$ into W which we call scalar multiplication. These mappings satisfy the conditions:

(1) $x \oplus y = y \oplus x$, for all $x, y \in W$
   (the commutative property).
(2) $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ for all $x, y, z \in W$.
   (the associative property).
(3) For each $x \in W$, there exists a unique element 0 in W such that

$$x \oplus 0 = 0 \oplus x = x$$

   (the existence of the zero element 0).
(4) For each $x \in W$, there is a unique element $-x \in W$ such that

$$x \oplus -x = 0.$$

   (the existence of an inverse).
(5) $\alpha(\beta x) = (\alpha\beta)x$ for all $x \in W$ and all $\alpha, \beta \in \mathscr{F}$.

(6) $(\alpha + \beta)x = \alpha x \oplus \beta y$ for all $x \in W$ and all $\alpha, \beta \in \mathscr{F}$.
(7) $\alpha(x \oplus y) = \alpha x \oplus \alpha y$ for all $x, y \in W$ and all $\alpha \in \mathscr{F}$.
(8) $1x = x$ for all $x \in W$, where 1 is the unit element of the scalar field $\mathscr{F}$.

In this book, $\mathscr{F}$ will be either the real number field, $\mathscr{R}$, or the complex number field, $\mathscr{C}$; W over $\mathscr{R}$ is called a real vector space, and W over $\mathscr{C}$ is called a complex linear vector space. (Where we do not explicitly mention $\mathscr{F}$ in examples we shall be taking $\mathscr{F} = \mathscr{R}$). We illustrate this concept by considering a number of very common linear vector spaces.

Example 1.1. Take $W = \mathscr{R}$ and $\mathscr{F} = \mathscr{R}$ with $\oplus$ ordinary addition, and scalar multiplication ordinary multiplication.

Example 1.2. Take W to be the set of all real polynomials of degree n and $\mathscr{F} = \mathscr{R}$. This is clearly a real linear vector space, and if we consider complex polynomials and $\mathscr{F} = \mathscr{C}$, we obtain a complex linear vector space.

Example 1.3. Take $W = \mathscr{R}^n$

$$\mathscr{R}^n = \{x : x = (x_1, \ldots, x_n), x_i \in \mathscr{R}; i = 1, \ldots, n\}.$$

with

$$x \oplus y = (x_1 + y_1, \ldots, x_n + y_n)$$

and

$$\alpha x = (\alpha x_1, \ldots, \alpha x_n) \text{ for } \alpha \in \mathscr{R}.$$

It is easily verified that W is a linear vector space over $\mathscr{R}$, but W is not a linear vector space over $\mathscr{C}$.

Example 1.4. Take W to be the set of all complex valued $m \times n$ matrices and $\mathscr{F} = \mathscr{C}$.

Example 1.5. Let W be the set of all scalar-valued functions $u : S \to \mathscr{F}$, where S is any nonempty set. Then for any $s \in S$, $u(s)$ is in $\mathscr{F}$, and we may define $u \oplus v$ and $\alpha u : (u \oplus v)(s) = u(s) + v(s)$ for all $s \in S$; $u, v \in W$. $(\alpha u)(s) = \alpha u(s)$ for all $\alpha \in \mathscr{F}$, $u \in W$.

Example 1.6. Take W to be the set of Riemann integrable real-valued functions on (0, 1), such that $\int_0^1 |u(s)|^2 \, ds < \infty$. If we let $\mathscr{F} = \mathscr{R}$ and define addition and scalar multiplication as in Example 1.5 above, then $u, v \in W$ implies that $u \oplus v$ and $\alpha u \in W$.

$$\text{i.e.} \quad \int_0^1 |u(s) + v(s)|^2 \, ds < \infty \quad \text{and} \quad \int_0^1 |\alpha u(s)|^2 \, ds < \infty.$$

The last inequality is trivial and for the first we have

$$\int_0^1 |u(s) + v(s)|^2 \,\mathrm{d}s \leqslant \int_0^1 (|u(s)|^2 + 2|u(s)\|v(s)| + |v(s)|^2)\,\mathrm{d}s$$

$$\leqslant \int_0^1 |u(s)|^2 \,\mathrm{d}s + 2\left(\int_0^1 |u(s)|^2 \,\mathrm{d}s\right)^{1/2} \left(\int_0^1 |v(s)| \,\mathrm{d}s\right)^{1/2}$$

$$+ \int_0^1 |v(s)|^2 \,\mathrm{d}s$$

where we have used the very useful inequality:

Schwarz's inequality

$$\int_0^1 |u(s)\, v(s)| \,\mathrm{d}s \leqslant \left(\int_0^1 |u(s)|^2 \,\mathrm{d}s\right)^{1/2} \left(\int_0^1 |v(s)|^2 \,\mathrm{d}s\right)^{1/2} \tag{1.1}$$

provided u and v are square Riemann integrable (see Chapter 2, 2.13 for a generalization of this inequality).

This example is a special type of subset of the vector space of Example 1.5 with $S = (0, 1]$, which we call a linear subspace.

Definition 1.2. Linear subspace

If W is a linear vector space over $\mathscr{F}$, then a subset S of W is a linear subspace if $x, y \in S \Rightarrow \alpha x \oplus \beta y \in S$, for all scalars $\alpha, \beta \in \mathscr{F}$ (i.e. S is closed under addition and scalar multiplication and so is itself a linear vector space over $\mathscr{F}$). Other examples of linear subspaces are:

Example 1.7. In Example 1.3 let S be the set of n-tuples of the form

$$x = (x_1, x_2, 0, \ldots, 0).$$

Example 1.8. In Example 1.4, let S be the set of matrices with certain blocks zero.

Example 1.9. In Example 1.2, let W be the set of all rth order polynomials where $r \leqslant n$.

A feature of linear subspaces is that they all contain the zero element. If we translate the origin of a linear subspace, we obtain an affine subset (sometimes called a linear manifold or a linear variety).

Definition 1.3. Affine subset

If W is a linear vector space over $\mathscr{F}$, then an affine subset M has the form

$$M = \left\{ \begin{array}{l} x : x = c + x_0, \text{ where } c \text{ is a fixed element of } W \\ \text{and } x_0 \in S, \text{ a linear subspace of } W \end{array} \right\}.$$

Example 1.10. In Example 1.3, let M be the set of n-tuples of the form

$$x = (x_1, x_2, 1, \ldots, 1).$$

Example 1.11. In Example 1.4, let M be the set of matrices with certain blocks of 1's.

Example 1.12. In $\mathscr{R}^3$ all lines and planes through the origin are subspaces, whereas lines and planes not passing through the origin are affine subsets.

Definition 1.4. Hyperplane

A hyperplane M of a vector space W is a maximal proper affine subset of W.

$$\text{i.e. } M = \left\{ \begin{array}{l} x \in W : x = c + x_0, \text{ where } c \in W \text{ is fixed and} \\ \qquad\qquad\qquad x_0 \in S, \text{ a subspace of } W \end{array} \right\}$$

and S is maximal in the sense that any other subspace is either contained in S or is W itself.

Example 1.13. Hyperplanes in $\mathscr{R}^2$ are lines; hyperplanes in $\mathscr{R}^3$ are planes.

A fundamental concept for linear vector spaces is that of dimension, but first we need a few more definitions.

Definition 1.5. Linear dependence and independence

If $x_1, x_2, \ldots, x_n$ are elements of W, a linear vector space over $\mathscr{F}$, and there are scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ not all zero such that

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0,$$

then we say that $x_1, x_2, \ldots, x_n$ is a linearly dependent set. If no such set of scalars exist, then we say that $x_1, \ldots, x_n$ are linearly independent

Example 1.14. $\{1, t, t^2, \ldots, t^n\}$ is a linearly independent set in W of Example 1.2. However $\{1 + t, \frac{1}{2} + 3t, 2t\}$ is a linearly dependent set, since

$$\tfrac{1}{2}(1 + t) + \tfrac{5}{4}(2t) - (\tfrac{1}{2} + 3t) = 0.$$

Example 1.15. In $\mathscr{R}^3$ any 3 noncoplanar vectors form a linearly independent set.

Linearly independent vectors $\{x_1, x_2, \ldots, x_n\}$ may be used to generate a linear subspace. This is accomplished by considering all possible linear combinations of $x_1, x_2, \ldots, x_n$.

Definition 1.6. Span of vectors

If $x_1, x_2, \ldots, x_n$ are elements of a linear vector space W, then

$$\text{Span}\{x_1, \ldots, x_n\} = \left\{x \in W : x = \sum_{i=1}^{n} \alpha_i x_i;\ \alpha_i \in \mathscr{F}\right\}$$

Then Span $\{x_1, \ldots, x_n\}$ is a subspace of W and we say that it is spanned by $\{x_1, \ldots, x_n\}$. This leads us to the concept of dimension.

Definition 1.7. Dimension of linear vector spaces

If the linear vector space W is spanned by a finite set of linearly independent vectors $x_1, \ldots, x_n$ (i.e. W = Span $\{x_1, \ldots, x_n\}$). Then we say W has dimension n. If there exists no such set of vectors, W is said to be infinite dimensional.

Definition 1.8. Hamel basis

If the linear vector space W = Span $\{x_1, \ldots, x_n\}$ for some set of linearly independent vectors $x_1, \ldots, x_n$, then this set is called a (Hamel) basis for W.

Note that the basis for W is not unique, but the dimension of W is.

Example 1.16. The dimension of $\mathscr{R}$ of Example 1.1 is of course 1 and a basis is $\{\alpha\}$, $\alpha \neq 0$.

Example 1.17. The dimension of Example 1.2 is $(n + 1)$ and a basis is $\{1, t, t^2, \ldots, t^n\}$. Another possible basis is the set of Legendre polynomials $\{1, t, 3t^2 - 1/2, \ldots, d^n(t^2 - 1)^n/dt^n\}$, (see Example 4.8).

Example 1.18. In Example 1.3, the dimension of $\mathscr{R}^n$ is, of course, n and a basis is $(1, 0, \ldots)$, $(0, 1, 0, \ldots)$, $\ldots$ $(0, 0, \ldots, 1)$.

Example 1.19. Example 1.6 is an infinite dimensional linear vector space and so is Example 1.5 if S is an infinite set.

Finite dimensional linear vector spaces are very special and are "just like" $\mathscr{R}^n$ or $\mathscr{C}^n$. To make this comparison more precise we define.

Definition 1.9. Isomorphic linear vector spaces

Linear vector spaces W_1 and W_2 are isomorphic if there is a bijective map T: $W_1 \to W_2$, such that

$$T(\alpha x \oplus \beta y) = \alpha(Tx) \oplus \beta(Ty) \quad \text{for all} \quad \alpha, \beta \in \mathscr{F};\ x, y \in W.$$

(This property of T is known as linearity—see Definition 3.1).

A major result of linear algebra is that all finite dimensional real linear vector spaces are isomorphic to $\mathscr{R}^n$ and all finite dimensional complex linear vector spaces are isomorphic to $\mathscr{C}^n$.

So far we have only been discussing algebraic properties of sets, but now we introduce a generalization of the concept of "length" for abstract sets.

Definition 1.10. Norm

A norm is a non-negative set function on a linear vector space, $\|.\|: W \to \mathscr{R}^+$ such that

(1) $\|x\| = 0$, if and only if $x = 0$.
(2) $\|x \oplus y\| \leqslant \|x\| + \|y\|$ for all $x, y \in W$
(the triangle inequality).
(3) $\|\alpha x\| = |\alpha|\ \|x\|$ for all $x \in W$ and $\alpha \in \mathscr{F}$.
(If (1) is not true, we call $\|.\|$ a seminorm—see Definition 7.12).

Definition 1.11. Normed linear space

A normed linear space is a linear vector space X with a norm $\|.\|_X$ on it, and is denoted by $(X, \|.\|_X)$. Usually where there will be no confusion we just write X, and use $\|.\|_X$ for the norm, and $+$ for addition in X.

As our first notion of norm came from the euclidean length function, it is appropriately our first example.

Example 1.20. Let $X = \mathscr{R}^n$ and $\|.\|$ be defined by

$$\|x\| = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}.$$

This may be proved using the *Schwarz inequality*

$$\left(\sum_{i=1}^{n} |x_i y_i|\right)^2 \leqslant \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2} \tag{1.2}$$

which is valid for n finite or infinite.

This is only one of several norms we can define for $\mathscr{R}^n$, for example

$$\|x\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|$$

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}; 1 \leqslant p < \infty.$$

all define norms on $\mathscr{R}^n$.

The proof that the p-norm is in fact a norm may be proved using *Minkowski's inequality*

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p} \tag{1.3}$$

which is valid for n finite or infinite.

$\mathscr{C}^n$ also becomes a normed linear space under these norms. We have a special notation for $\mathscr{R}^n$ (or $\mathscr{C}^n$) under the p-norm, namely l_p^n, and under the ∞-norm, l_∞^n.

Example 1.21. Take X to be the linear vector space of infinite sequences $x = (x_1, x_2, \ldots, x_n \ldots)$. Then using Minkowski's inequality we may also define a p-norm.

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} \quad \text{for} \quad 1 \leqslant p < \infty.$$

although it may not always be finite. However, restricting ourselves to those elements with finite p-norm, we have a linear vector space $l_p = \{x = (x_1, x_2, \ldots)$ with $\|x\|_p < \infty\}$, for $1 \leqslant p < \infty$. Notice that for $p \neq q$, l_p and l_q will not contain the same set of elements, because there exists constants c_i; $i = 1, 2, \ldots$ such that $c_1 \|x\|_1 \geqslant c_2 \|x\|_2 \geqslant \ldots$ and so $l_1 \subset l_2 \subset \ldots$ Another norm on the space of infinite sequences is $\|.\|_\infty$, where $\|x\|_\infty = \sup_{1 \leqslant i \leqslant \infty} |x_i|$, and this defines the normed linear space

$$l_\infty = \{x = (x_1, x_2, \ldots) \text{ with } \|x\|_\infty < \infty\}$$

and $l_\infty \supset l_p$ for $p \geqslant 1$.

Example 1.22. Consider the function space $X = \mathrm{C}[a, b]$, the space of continuous complex-valued functions defined on $[a, b]$ with norm $\|u\| = \sup_{a \leqslant t \leqslant b} |u(t)|$. This is called the uniform or sup norm.

Example 1.23. Consider $X = \mathrm{C}[a, b]$ under the p-norm

$$\|u\|_p = \left(\int_a^b |u(t)|^p \, \mathrm{d}t\right)^{1/p}$$