Applications of Mathematics

3

A.V. Balakrishnan

Applied Functional Analysis

Second Edition

Springer-Verlag
New York Heidelberg Berlin

A. V. Balakrishnan

Applied Functional Analysis

Second Edition

Springer-Verlag

New York Heidelberg Berlin

A. V. Balakrishnan
University of California
Systems Science Department
Los Angeles, CA 90024
USA

AMS Subject Classifications
46Cxx, 46N05, 47B05, 47B10, 47B40, 47D05, 52A05, 90C25, 93Cxx, 93Exx

Library of Congress Cataloging in Publication Data

Balakrishnan, A V
Applied functional analysis.

(Applications of mathematics; 3)
Bibliography: p.
Includes index.
1. Hilbert space. 2. Mathematical optimization.
3. Functional analysis. I. Title.
QA322.4.B34 1981 515.7′33 80-21291

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90527-8 Springer-Verlag New York
ISBN 3-540-90527-8 Springer-Verlag Berlin Heidelberg

Preface to the Second Edition

In preparing the second edition, I have taken advantage of the opportunity to correct errors as well as revise the presentation in many places. New material has been included, in addition, reflecting relevant recent work.

The help of many colleagues (and especially Professor J. Stoer) in ferreting out errors is gratefully acknowledged. I also owe special thanks to Professor V. Sazonov for many discussions on the white noise theory in Chapter 6.

February, 1981 A. V. BALAKRISHNAN

Preface to the First Edition

The title "Applied Functional Analysis" is intended to be short for "Functional analysis in a Hilbert space and certain of its applications," the applications being drawn mostly from areas variously referred to as system optimization or control systems or systems analysis.

One of the signs of the times is a discernible tilt toward application in mathematics and conversely a greater level of mathematical sophistication in the application areas such as economics or system science, both spurred undoubtedly by the heightening pace of digital computer usage. This book is an entry into this twilight zone. The aspects of functional analysis treated here are rapidly becoming essential in the training at the advance graduate level of system scientists and/or mathematical economists. There are of course now available many excellent treatises on functional analysis. However, the very fact of the comprehensive coverage makes it difficult of access to the application-minded user. Also, the high degree of generality, the watermark of mathematical achievement, is often at the expense of the richer results obtainable in the more highly structured cases common in applications. It is with some of these thoughts in mind that I have dealt exclusively with analysis in a Hilbert space and emphasized such special topics as Volterra operators and Hilbert–Schmidt operators; dissipative compact semigroups; and factorization theorems for positive definite operators, to name a few. Many topics in functional analysis *per se* have had to be totally shelved or otherwise abridged considerably mostly based on considerations of significance in application, but also to keep the size of the volume within reasonable bounds.

Another point is that the abstract theory is sometimes easier than the applications. This is true for instance in the case of semigroup theory where the generation theorems, for example, are far easier than showing that a particular partial differential equation generates a semigroup. Indeed a novice is bewildered by a seemingly endless variety of approaches to boundary value problems, even to the notion of what is meant by boundary value. Here I have taken some pains to illustrate by examples how the abstract theory relates to problems in partial differential equations without of course any claim to completeness.

Of the six chapters in the book, three deal specifically with applications topics. These are Chapter 2 on convex sets and convex programming in a Hilbert space; Chapter 5 on deterministic control problems and Chapter 6 on stochastic optimization problems. Chapter 6 is unusual in that it exploits the theory of finitely additive probability measures on a Hilbert space (in contrast to the more standard Wiener measure on the space of continuous functions). This chapter also contains some original material.

The remaining chapters (about two thirds of the book) are devoted to functional analysis and semigroups within a Hilbert space framework. The basic properties of Hilbert spaces and some of the fundamental theorems central to what follows are in the beginning chapter. The background so built up is sufficient to consider applications to convex programming problems in the second chapter. It is possible to proceed directly from Chapter 1 to Chapter 3 featuring the theory of linear operators in a Hilbert space. L_2-distributional derivatives are studied as examples of unbounded operators and associated notion of Sobolev spaces. Operators over separable Hilbert spaces receive special attention as well as L_2 spaces over Hilbert spaces. The final section in Chapter 3 is devoted to nonlinear operators, or more accurately, polynomials and analytic functions. We go on to semigroup theory in Chapter 5, again emphasizing the more specialized cases such as compact semigroups and Hilbert–Schmidt semigroups. Semigroup theory in a Hilbert space strikes the right balance for our purposes between the too general and the too particular; for example, it provides a general enough framework for optimization problems involving partial differential equations without getting lost in the details of the particular equations. The concepts of controllability and observability important in system theory are examined in the semigroup theoretic setting. An example illustrates the application to nonhomogeneous boundary value problems. A final section deals with a special class of evolution equations that arise as perturbations of the semigroup equation.

The book is a revised and enlarged version of the author's *Introduction to Optimization Theory in a Hilbert Space*, No. 42 in the Springer-Verlag Lecture Series on Economics and Mathematical Systems Theory, and has been used in graduate courses given in the Department of Mathematics and the Department of System Science. The prerequisites are the standard graduate courses in real and complex variables and concommitant material such as Fourier transforms; material on function spaces usually included in real analysis texts would be helpful background since the bare definitions given here in introductory sections may be inadequate for a firm grasp.

Similarly in the applications chapters, some familiarity with control problems in finite dimensions would be helpful.

Many students, past and present, have helped in improving the presentation: D. Washburn, Claude Benchimol and Frank Tung, in particular. Dr. J. Mersky helped with proofreading. Dr. J. Ruzicka rendered much needed assistance throughout the various stages of the manuscript.

I am indebted to Regina Safdie for her endurance typing and to W. Kaufmann-Bühler for interest and encouragement. Grateful acknowledgement is made of the financial support in part under a research grant from the Applied Mathematics Division, AFOSR-USAF, monitored by Colonel W. Rabe.

Contents

Basic Properties of Hilbert Spaces 1

1.0 Introduction

This is an introductory chapter in which we study the basic properties of Hilbert spaces, indispensable for an understanding of the sequel. Although it is fairly complete in itself, this chapter is necessarily brief in many areas and the reader would find it helpful to have had an elementary introduction to linear spaces, and Hilbert spaces in particular, such as one finds in the standard texts on real analysis.

We begin with the basic definitions in Section 1.1. Some of the standard examples of Hilbert spaces are given in Section 1.2, while the more sophisticated ways (of importance in applications) in which Hilbert spaces are made up of Hilbert spaces are indicated in Section 1.3, including in particular tensor products of Hilbert spaces. We discuss next the simplest optimization problem in a Hilbert space—namely, projections on convex sets—in Section 1.4, and go on to the concepts of orthogonality and orthonormal bases in Section 1.5. After a brief discussion of continuous linear functionals in Section 1.6, we prove the basic Riesz representation theorem in Section 1.7. Section 1.8 contains some of the main theorems. We study weak convergence and prove the weak compactness property of bounded sets characteristic of Hilbert spaces, the Mazur theorem on convex sets, as well as the more general uniform boundedness principle. Section 1.9 treats a rather specialized topic: the *generalized curves* of L. C. Young in the context of nonlinear functionals on a Hilbert space, illustrating its importance in control theory (*chattering controls*). We close in Section 1.10 with a statement of the Hahn–Banach theorem as needed in Chapter 2.

Much of the material is standard and can be found in many places: notably [1], [5], [25], [30], [35], [39]. A useful reference for real analysis is [33]. For the specialized material of Section 1.10, the basic reference is [40].

1.1 Basic Definitions

Def. 1.1.1. *A* linear space *is a nonvoid set* $\mathscr{E}$ *for which two operations called* addition *(denoted* $+$*) and* scalar multiplication *(*$\cdot$*) are defined. Addition is commutative and associative, making* $\mathscr{E}$ *into a commutative group under addition. Multiplication is by scalars (either from the complex field, in which case we have a* complex linear space, *or the real number field, in which case we have a* real linear space). *Multiplication is associative, and distributive with respect to* $(+)$ *as well as addition of scalars.*

In this book we shall deal almost exclusively with *function spaces*; that is, linear spaces of functions—in which case the operations will be natural. We shall, as a rule, use the letters $\mathscr{E}$, $\mathscr{F}$, $\mathscr{H}$, $\mathscr{X}$, and $\mathscr{Y}$ to denote linear spaces; x, y, z to denote elements of $\mathscr{H}$; f, g, h to denote elements in a function space where $f(\cdot)$, $g(\cdot)$, $h(\cdot)$ are the functions, and α, β, γ to denote elements in the scalar field.

Def. 1.1.2. *A set of elements in* $\mathscr{E}$ *is* linearly dependent *if the zero element can be expressed as a finite linear combination of elements in the set; i.e.,*

$$0 = \sum_{1}^{n} a_k x_k,$$

where x_k *are elements in the set and not all the scalar coefficients* a_k *are zero. Otherwise the set is* linearly independent.

Def. 1.1.3. *A* linear subspace, *or simply a* subspace, *of the linear space* $\mathscr{E}$ *is a subset which is itself a linear space under the same operations.*

Def. 1.1.4. *A* linear functional *on* $\mathscr{E}$ *is a function defined on* $\mathscr{E}$ *with range in the scalar field such that if* $f(\cdot)$ *denotes the functional*

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y); \qquad x, y \in \mathscr{E};\ \alpha, \beta \text{ scalars}.$$

Def. 1.1.5. *The* Cartesian product $\mathscr{E}_1 \times \mathscr{E}_2$ *of two linear spaces* $\mathscr{E}_1$, $\mathscr{E}_2$ *(over the same scalar field) is the set of all pairs* (x, y), $x \in \mathscr{E}_1$, $y \in \mathscr{E}_2$.

Def. 1.1.6. *A* bilinear functional *on* $\mathscr{E}$ *is a functional defined on the Cartesian product space* $\mathscr{E} \times \mathscr{E}$, *with range in the scalar field such that, denoting the functional by* $f(x, y)$, *we have*

(i) $f(x, y)$ *is a linear functional on* $\mathscr{E}$ *for fixed* y.
(ii) $f(x, y) = \overline{f(y, x)}$, *the bar denoting conjugate complex.*

We now come to definitions closer to our needs.

Def. 1.1.7. *An* inner product *on a linear space is a bilinear functional* $f(x, y)$ *which satisfies the additional condition*

(iii) $f(x, x) \geq 0$

and equality holds if and only if x is zero.

An inner product will usually be denoted $[x, y]$, to be distinguished from the same notation to denote a closed interval, but the confusion should be minimal.

EXAMPLE 1.1.1. Let $C[a, b]$ denote the space of continuous functions defined on the closed finite interval $[a, b]$ of the real line. Let $C^{(k)}[a, b]$ denote the space of k-times continuously differentiable functions on $[a, b]$. An inner product of some importance on $C^{(k)}[a, b]$ is

$$[f, g] = \sum_{j=0}^{k} \int_a^b f^{(j)}(t)\overline{g^{(j)}(t)}\, dt, \tag{1.1.1}$$

where $f^{j}(\cdot)$ denotes the jth derivative of $f(\cdot)$.

An example of a different kind of inner product on $C[a, b]$ is provided by

$$[f, g] = \int_a^b \int_a^b \frac{\sin \pi(t - s)}{\pi(t - s)} f(s)\overline{g(t)}\, ds\, dt. \tag{1.1.2}$$

The verification that (1.1.1) is an inner product is immediate, and for (1.1.2), only condition (iii) requires a proof. But this follows from the fact that we can express (1.1.2) as

$$[f, g] = \int_{-1}^{1} \left(\int_a^b e^{i\pi\lambda t} f(t)\, dt \right)\left(\int_a^b e^{-i\pi\lambda t}\overline{g(t)}\, dt \right) d\lambda$$

so that $[f, f] \geq 0$, and if equality holds we must have

$$\int_a^b e^{i\pi\lambda t} f(t)\, dt = 0, \qquad -1 \leq \lambda \leq +1.$$

However, the integral defines an analytic function of λ, and hence must be zero for all λ, implying in turn that (the continuous function) $f(\cdot)$ must be identically zero.

EXAMPLE 1.1.2. As an example of a bilinear functional which is *not* an inner product, define on $C[a, b]$

$$[f, g] = \int_a^b \int_a^b f(s)\overline{g(t)}\, ds\, dt. \tag{1.1.3}$$

Obviously there are many nonzero functions in $C[a, b]$ such that

$$[f, f] = \left| \int_a^b f(t)\, dt \right|^2 = 0.$$

A fundamental property of inner products is the Cauchy–Buniakowski–Schwarz inequality (referred to hereafter as the Schwarz inequality):

$$|[x, y]|^2 \leq [x, x] \cdot [y, y]. \tag{1.1.4}$$

This can be deduced as follows: For any λ we note that $0 \leq [x + \lambda y, x + \lambda y] = [x, x] + |\lambda|^2[y, y] + \bar{\lambda}[x, y] + \lambda\overline{[x, y]}$. The inequality is obviously satisfied if y is zero. Hence we need only consider the case when y is nonzero. Hence we may choose $\lambda = -[x, y]/[y, y]$, and this choice of λ yields

$$0 \leq [x, x] - |[x, y]|^2/[y, y],$$

which is the inequality sought. Further, we see from this that equality holds in (1.1.4) if and only if one element is a scalar multiple of the other. (This simple observation provides the basis for the theory of matched filters in communication (detection) theory; see [3].)

Def. 1.1.8. *A* norm *on a linear space is a nonnegative functional $f(\cdot)$ such that*

$$\begin{aligned} f(x) &= 0 \quad \textit{if and only if} \quad x = 0 \\ f(\alpha x) &= |\alpha| f(x). \\ f(x + y) &\leq f(x) + f(y) \qquad \textit{(triangle inequality)}. \end{aligned}$$

A norm is usually denoted $\|\cdot\|$.

Def. 1.1.9. *A* normed linear space *is a linear space with the topology induced by the norm defined on it: neighborhoods of any point x_0 are the spheres $\|x - x_0\| < r, r > 0$.*

Def. 1.1.10. *An* inner product space *(also called a pre-Hilbert space) is a normed linear space with the norm defined by $\|x\| = \sqrt{[x, x]}$. Implicit in this definition is of course the verification that $\sqrt{[x, x]}$ yields a norm. The only nontrivial part in this verification is the triangle inequality, which is an easy consequence of the Schwarz inequality.*

Def. 1.1.11. *In a normed linear space, a sequence x_n is said to be a* convergent sequence *if there is an element x in the space such that $\|x_n - x\| \to 0$, and we say that x_n converges to x.*

In a normed linear space, if x is a limit point of a set, then we can find a sequence $\{x_n\}$ in the set such that x_n converges to x, and of course conversely.

Def. 1.1.12. *A sequence x_n in a normed linear space is said to be a* Cauchy sequence *if, given $\varepsilon > 0$, we can find an integer $N(\varepsilon)$, such that $\|x_n - x_m\| < \varepsilon$ for all $n, m > N(\varepsilon)$.*

Whereas every convergent sequence is a Cauchy sequence, the reverse is not necessarily true in a normed linear space. For example it is well known (see [33], for instance) that in $C[0, 1]$ under the norm defined by (1.1) with $k = 0$, it is possible to find a Cauchy sequence of functions which does not converge to a continuous function in $C[0, 1]$.

Def. 1.1.13. *A normed linear space in which every Cauchy sequence is a convergent sequence is called a* Banach space.

Def. 1.1.14. *An inner product (normed linear) space in which every Cauchy sequence is a convergent sequence is said to be* complete.

A complete inner product space is called a Hilbert space.

We note that every inner product space can be *completed*. That is to say, denoting the inner product space by $\mathscr{E}$, we can find a Hilbert space $\mathscr{H}$ such that:

(i) There is a function $L(\cdot)$ defined on $\mathscr{E}$ with range in $\mathscr{H}$ such that $L(\cdot)$ is linear and *one-to-one*:

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$$
$$L(x) = 0 \quad \text{implies} \quad x = 0$$

(ii) $L(\cdot)$ is an *inner product–preserving* map:

$$[L(x), L(y)]_{\mathscr{H}} = [x, y]_{\mathscr{E}}$$

where the subscripts denote the space in which the inner product is taken, and

(iii) The closure of the set $L(\mathscr{E})$ is equal to $\mathscr{H}$.

We shall indicate briefly how the existence of such a Hilbert space can be established [see, e.g., [39] for details]. Let us say that two Cauchy sequences $\{x_n\}$, $\{y_n\}$ in $\mathscr{E}$ are *equivalent* if the difference sequence $\{x_n - y_n\}$ converges to zero. We consider the linear space of (equivalence classes) of Cauchy sequences in $\mathscr{E}$ made into an inner product space under the inner-product $[\{x_n\}, \{y_n\}] = \lim_n [x_n, y_n]$. The important point is that it may be verified that this space is actually complete, yielding the Hilbert space sought. The map $L(\cdot)$ is defined by $L(x) =$ the (equivalence class containing the) Cauchy sequence, defined by $x_n = x$ for every n. We call the Hilbert space so obtained the completion of $\mathscr{E}$, "identifying" each x in $\mathscr{E}$ with the corresponding Cauchy sequence $(L(x))$ in $\mathscr{H}$. Note that the procedure is analogous to the way in which we define real numbers as Cauchy sequences of rational numbers.

1.2 Examples of Hilbert Spaces

Perhaps the simplest example of a function space which is a Hilbert space is the space of real-(complex-) valued functions $f(\cdot)$, Lebesgue measurable and square integrable on the interval $[a, b]$, $-\infty \leq a < b \leq +\infty$, with inner product defined by $[f, g] = \int_a^b f(t)\overline{g(t)}\, dt$. That the inner product space so defined is complete is a standard result in analysis; see, for example [33]. We usually denote the space $L_2(a, b)$. For many purposes, particularly in dealing with partial differential equations, we need a slightly more general form of this space. Thus let $\mathscr{D}$ denote an open set of the Euclidean space E_n of dimension n. By $L_2(\mathscr{D})^{qp}$ we shall mean the space of all q-by-p martix functions Lebesgue measurable on $\mathscr{D}$, such that $\int_{\mathscr{D}} \text{Tr.} f(s) f(s)^* \, dm < \infty$, (where m denotes Lebesgue measure, * denotes conjugate transpose, and Tr. indicates the "trace") under the inner product $[f, g] = \int_{\mathscr{D}} \text{Tr.}\ f(s) g(s)^* \, dm$. This is a Hilbert space.

As may be surmised immediately, the restriction to Lebesgue measure is not necessary. For example, in the theory of stochastic processes we shall often need to work with the following generic Hilbert space, which we shall denote $L_2(\Omega, \boldsymbol{\beta}, \mu)$. This is the space of q-by-p matrix functions $f(\cdot)$ defined on the abstract set Ω, measurable with respect to a sigma-algebra $\boldsymbol{\beta}$ of sets in Ω, such that $\int_\Omega \text{Tr.}\ f(s) f(s)^* \, d\mu < \infty$, where μ is a countably additive (σ-finite, generally-finite for the needs of probability theory) measure defined on $\boldsymbol{\beta}$, with inner product defined by $[f, g] = \int_\Omega \text{Tr.}\ f(s) g(s)^* \, d\mu$.

Problem 1.2.1. Let R be a self adjoint nonnegative definite $(n \times n)$ matrix. Consider the class of $n \times 1$ functions $u(\cdot)$, Lebesgue measurable on (0, 1), and such that

$$\|u\|^2 = \int_0^1 [R\, u(t), u(t)] \, dt < \infty.$$

Can this be made into a Hilbert space with norm $\|u\|$?

1.3 Hilbert Spaces from Hilbert Spaces

Very often we need to study Hilbert spaces derived from a given Hilbert space or, more generally, a collection of Hilbert spaces. The simplest such example is the Cartesian product.

Def. 1.3.1. *The* Cartesian product *of two Hilbert spaces $\mathscr{H}_1$, $\mathscr{H}_2$ is the linear space of all pairs (x_1, x_2), $x_1 \in \mathscr{H}_1$, $x_2 \in \mathscr{H}_2$, under the operations*

$$(x_1, x_2) + (y_1, y_2) = (x_1 + x_2, y_1 + y_2)$$
$$\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2),$$

and endowed with the inner product $[(x_1, x_2), (y_1, y_2)] = [x_1, y_1] + [x_2, y_2]$.

The Cartesian product space will be denoted $\mathscr{H}_1 \times \mathscr{H}_2$, and is readily verified to be a Hilbert space. The definition can clearly be extended to a finite number of Hilbert spaces, $\mathscr{H}_i$, $i = 1, \ldots, n$. When the Hilbert spaces are the same we sometimes use the notation $\mathscr{H}^n$ for $\mathscr{H} \times \mathscr{H} \times \mathscr{H} \cdots \mathscr{H}$ n-times. Note that $L_2(\mathscr{D})^{qp}$ can be identified as the qp Cartesian product of $L_2(\mathscr{D})$. The Cartesian product of a sequence $\{\mathscr{H}_n\}$ of Hilbert spaces may also be defined in analogous manner. Thus we first consider the linear spaces of sequences $\{x_i\}$, $x_i \in \mathscr{H}_i$ with the vector space operations defined in obvious manner.

If x denotes the sequence $\{x_i\}$, and y similarly the sequence $\{y_i\}$, then we define $x + y$ as the sequence $\{x_i + y_i\}$ and $\alpha x =$ the sequence $\{\alpha x_i\}$. The Cartesian product $\prod_{i=1}^{\infty} \mathscr{H}_i$ is the subspace of the linear space of sequences $\{x_i\}$ such that

$$\sum_1^{\infty} \|x_i\|^2 < \infty$$

endowed with the inner product

$$[x, y] = \sum_1^{\infty} [x_i, y_i].$$

It may readily be verified that the space is complete. We note that the space l_2 of square summable sequences may be defined in this way, taking the space of real (complex) numbers as the basic Hilbert space.

Somewhat more involved is the notion of the *tensor product* of Hilbert spaces. To explain this, let $\mathscr{H}_1$, $\mathscr{H}_2$ be two Hilbert spaces. We first consider the *algebraic* tensor product considering the spaces merely as linear spaces. This is the linear space of all formal finite sums

$$z = \sum_{i=1}^{n} (x_i \otimes y_i), \qquad x_i \in \mathscr{H}_1,\ y_i \in \mathscr{H}_2,$$

with the following expressions identified:

$$\alpha(x \otimes y) = (x \otimes \alpha y) = (\alpha x \otimes y)$$
$$((x_1 + x_2) \otimes y) = (x_1 \otimes y) + (x_2 \otimes y)$$
$$(x \otimes (y_1 + y_2)) = (x \otimes y_1) + (x \otimes y_2).$$

We endow this space with the inner product

$$\left[\sum_{i=1}^{n_1} (x_i \otimes y_i), \sum_{j=1}^{n_2} (s_j \otimes t_j)\right] = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} [x_i, s_j][y_i, t_j].$$

(The indices n_1, n_2 may be taken to be same without loss of generality by adding zero entries.) It needs to be verified that this is indeed an inner product. The bilinearity being obvious, let us proceed to the consequence of

$$\left[\sum_i^n (x_i \otimes y_i), \sum_j^n (x_j \otimes y_j)\right] = 0.$$

Letting $r_{ij} = [x_i, x_j]$, $m_{ij} = [y_i, y_j]$, we see that this is the same as

$$\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} m_{ij} = 0.$$