to computational complexity. Instead of investigating if there exists an algorithm to solve a given problem, more attention is given to study how economical (space or time intensive) an algorithm for a given problem potentially can be, thereby classifying decidable problems into (sub) hierarchies of difficulty. To the computer scientist, then, the book compiles material which is well understood for some time now. This is in line with the tutorial level at which the book is written.

References

1. A. Church, *An unsolvable problem of elementary number theory*, Amer. J. Math. **58** (1936), 345–363.

2. S. Kleene, *Introduction to metamathematics*, Van Nostrand, Princeton, N.J., 1952. MR **14**, 525.

3. L. Kalmár, *An argument against the plausibility of Church's thesis*, Constructivity in Mathematics: North-Holland, Amsterdam, 1959, pp. 72–80. MR **21** #5567.

CHRISTOPH M. HOFFMANN

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 1, January 1978

Applied functional analysis, By A. V. Balakrishnan, Springer-Verlag, New York, Heidelberg, Berlin, 1976, vii + 309 pp., $19.80.

For present purposes "functional analysis" will mean the study of Hilbert spaces and of (not necessarily continuous) linear operators between such spaces. Let us recall that Hilbert spaces are characterized among general Banach spaces by any one of numerous geometric properties, such as the parallelogram property of the norm (Jordan-von Neumann, 1935), or the existence of a norm-one projector onto every (closed linear) subspace (Kakutani, 1939; spaces of dimension $\leqslant 2$ are exceptional). These two conditions can be jointly utilized to show that any Banach space of dimension > 2 having sufficiently many finite dimensional linear metric projectors is a Hilbert space (Rudin-Smith, 1961). More recently, isomorphs of Hilbert spaces have been characterized in several spectacular ways; for example, as those Banach spaces in which every subspace is complemented [**15**], or those which obey a Central Limit Theorem property (due to several authors, see [**1**] for one presentation).

The operator theory for Hilbert spaces is dominated by the interplay between an operator and its adjoint which, thanks to the Riesz representation theorem, can be defined on the codomain of the given operator. Major achievements in this theory include the spectral theorem for normal operators (Hilbert, Riesz, von Neumann, Stone, Gelfand-Naimark, Segal, et al., 1906–1951), the polar decomposition of closed operators (von Neumann, 1932), the dilation theory of Halmos and Sz.-Nagy (1950–1955), which in turn has led to the characteristic function and canonical model approach to the study of contractive operators (Livšic, Sz.-Nagy-Foiaş, et al., 1946–1967), and the triangular representation of compact operators (Livšic, Brodskiĭ, et al., 1954–1969). Presentations of these theories and much more are given in [**7**], [**10**], [**18**], [**19**]. Note that we are for the most part leaving completely aside the vast subject of operator algebras.

There are of course innumerable applications of Hilbert spaces and operators thereon. Expansions in Fourier and other orthogonal series come immediately to mind. Marin's book **[16]** describes many other applications to mathematics (group representations, ergodic theory, the Dirichlet problem, etc.). The crucial role of hermitian operators in quantum mechanics (Dirac, von Neumann, 1930–1932) is also well known. In a different direction there are important applications to the statistical and engineering sciences which may be broadly classified as linear least squares estimation. The essential problem here is to make an "optimal estimate" of a random variable Z based on a collection $\{X\}$ of "observable" random variables. This problem may be cast into a Hilbert space framework by supposing all random variables under consideration to be defined on a common probability space and to have finite second moments. Then the dual requirement that our estimate be both linear and optimal can be satisfied by projecting Z orthogonally onto the closed linear span M of $\{X\}$. Since

$$\|Z - Y\|^2 = |E(Z - Y)|^2 + \mathrm{Var}(Z - Y)$$

for any second order Y, in particular for $Y \in M$, the optimal estimate is neatly characterized both intuitively and mathematically. The optimal nonlinear estimate of Z can be similarly obtained (in principle) by projecting Z onto the larger subspace of all second order random variables measurable with respect to the σ-algebra generated by $\{X\}$; the result is the usual conditional expectation $E(Z|\{X\})$. The case where $\{X\}$ is a stochastic process is particularly important (filtering and prediction problems), and in most such instances the computational difficulties are severe. Associated with the early history of this subject are the names of Wold, Kolmogorov, Wiener, and Krein (1938–1945), and with more recent developments are those of Kalman, Bucy, Parzen, Masani, Kailath, Rozanov and many others. For an excellent historical perspective see **[14]**; a general reference is **[17]**.

These prefatory remarks having been made, it must next be stated that, although Balakrishnan's book is subtitled *Functional analysis in a Hilbert space and certain of its applications*, its content is essentially disjoint from all the aforementioned topics, excepting the final one. Indeed, the author views Hilbert spaces as the proper setting for a variety of deterministic and stochastic optimization problems. His choice of theoretical material is thereby guided largely by the intended applications. Having indicated (some of) the topics that are omitted under this program, let us now see in more detail what is covered. Incidentally, a preliminary version of this book appeared in 1971 **[4]**.

Chapters 1 and 2 cover the basic geometry of Hilbert space and its application to standard convex optimization problems. These include the Kuhn-Tucker theorem of convex programming, the minimax theorem for continuous concave-convex functions, and the classical (finite dimensional) version of Farkas' lemma. Some of this geometric theory is later put to good use to analyse the linear quadratic regulator problem with pointwise constraints on the control.

One of the highlights of this book is its wealth of interesting examples. Already in these first two chapters we see for instance a variety of properties of convex subsets of $L^2[a, b]^q$ of the form $\{f: f(\cdot) \in C \text{ a.e.}\}$, where C is a closed convex subset of R^q. We learn about the generalized (relaxed) functions of L. C. Young and we encounter a closed convex subset of l^2 which cannot be supported at every boundary point. With "^" denoting Fourier transform we also see that a unit vector sequence $\{f_n\}$ in $L^2[a, b]$ converges weakly to 0 if and only if $\int_{-B}^{B}|\hat{f}_n| \to 0$, $\forall B > 0$ (intuitively, the energy in any band of frequencies goes to zero).

The study of operators and semigroups is the subject of the next two chapters which together comprise half the book. Here the motivation comes ultimately from the intended applications to "systems analysis". Roughly, a (dynamical) system is an organized structure which evolves in time, which can be influenced externally, and whose behavior can be monitored. At any time t the output $y(t)$ of such a system depends on the current input $u(t)$ plus some (internal) attribute $x(t)$ of the system, called its state. Given sufficiently many axioms pertaining to causality, smoothness, linearity, time invariance, etc., the system can be modelled by a pair of equations

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{1}$$

$$y(t) = Cx(t), \tag{2}$$

where A, B, C are linear operators with appropriate domains. In the classical finite dimensional case the functions $x(\cdot)$, $y(\cdot)$, $u(\cdot)$ assume values in various Euclidean spaces, the model is fully deterministic, and the machinery of linear algebra and ordinary differential equations is adequate for a complete study **[6]**. Further, the resulting theory has been successfully applied to the control and synthesis of real physical systems. But, when the state and input spaces are allowed to have infinite dimension, and/or the forcing term $Bu(t)$ is a stochastic process then the very meaning of (1) becomes less clear and some sort of more sophisticated analysis is necessary.

Elementary spectral theory culminating in the spectral theorem for compact hermitian operators is given in Chapter 3, followed by introductions to special compact operators of Volterra, Hilbert-Schmidt, and nuclear type. Also briefly considered are nonlinear operators of polynomial type, and Fréchet analyticity of general nonlinear mappings. Among the many examples in this chapter are a variety of (compact) integral operators and (unbounded) differential operators, noncompact integral operators of convolution type on $L^2(R^n)$ (they being unitarily equivalent under Fourier transformation to multiplication operators), Volterra operators on $L^2[a, b]^q$ (these being compact, quasinilpotent integral operators of the form

$$y(t) = \int_a^t W(s, t)u(s)\, ds, \tag{3}$$

and viewed as explicitly describing the response y of a linear system to input u, in contrast to the implicit description provided by (1), (2); in systems theory W is known as the impulse response function or weighting pattern), Sobolev spaces, the Krein factorization theorem, and the Bochner spaces $L^2([a, b]; H)$. We also learn that integral operators on $L^2[a, b]^q$ with continuous, symmetric, positive semidefinite kernels (covariance kernels) are nuclear.

Reasonably conventional C_0-semigroup theory over Hilbert spaces is the subject of Chapter 4, with a view towards solving equation (1). Semigroups of the form $t \mapsto \exp(tA)$, $0 \leqslant t < \infty$, are uniformly continuous for bounded operators A, and these are adequate for handling ordinary differential equations. Since the author wishes to deal with partial differential equations it can only be assumed that A, typically a partial differential operator or matrix of such in spatial variables, is a closed d.d. operator, in which case the semigroup generated by A can only be expected to be strongly continuous. The Hille-Yosida theorem characterizing the generators of strongly continuous contraction semigroups is proved, and there are discussions of compact and holomorphic semigroups, and evolution equations. Of course, Stone's theorem on unitary groups is missing. With this background the author then establishes

$$x(t) = T(t)x(0) + \int_0^t T(t-s)Bu(s)\,ds \tag{4}$$

as the unique solution of (1), assuming, inter alia, that A generates a C_0-semigroup $t \mapsto T(t)$. The function (4) can be a solution in either a weak or strong sense depending on what is assumed about the smoothness of u and B. In the examples we see how the one-dimensional heat, wave, and Schrödinger equations can be cast into this framework, and the specific semigroups which provide their solution.

Next the author applies this machinery to introduce the notions of controllability and observability of the system (1)–(2), and to the resolution of a few selected problems of optimal control of the system (1). With assumptions sufficient to guarantee a unique solution of (1) on the time interval $[0, T]$ (for prescribed $x(0)$), these latter problems pertain to the minimization of quadratic cost functionals of the form

$$J(u) = \int_0^T (Rx, x)\,dt + \lambda \int_0^T (u, u)\,dt + (Sx(T), x(T)), \tag{5}$$

defined on the (Bochner) space of controls u. Optimal closed-loop (feedback) controls of the form

$$u(t) = -B^*P(t)x(t) \tag{6}$$

are derived for various combinations of R, S, λ being nonzero, where P weakly solves various operator equations of Riccati type. Some of these results represent generalizations of the finite dimensional work of Kalman (1960). The case of an infinite time interval ($T = \infty$) is also discussed under the strong assumption of exponential stability of the semigroup generated by A. Finally, problems with pointwise constraints on the controls are considered: first with $\lambda = 0$, $S = 0$ in (5), and then the famous time-optimal problem. Here the maximum principle and the concept of a "bang-bang" control make their appearance. Whether these optimality conditions actually obtain, however, depends on an assumption about the existence of support points to certain convex sets which, in infinite dimensional spaces, need not be valid, as the author points out.

Finally we come to the more difficult topic of stochastic optimization, especially state estimation ("filtering") and control (but not prediction). The conventional finite dimensional versions of this topic are of considerable practical importance and have been thoroughly studied [2], [8], [13], [17], etc. The intuitive idea is that both the system dynamics (1) and the observation process (2) have been corrupted by an additive "noise" term, so that we now have the stochastic system

$$\dot{x}(t, \omega) = Ax(t, \omega) + Bu(t, \omega) + F\omega(t),$$
$$y(t, \omega) = Cx(t, \omega) + Du(t, \omega) + G\omega(t)$$

for almost all $t \in [0, T]$, and with some additional information on the initial state $x(0)$. Rather surprisingly the author assumes that $x(0)$ is perfectly known–a most uncommon situation in practice. Here B, C, D, F, G are various bounded linear operators (with $B = 0$, $D = 0$ in the pure filtering problem), A has its usual significance, and $\omega(\cdot)$ is a "standard white noise process". The first problem is to make rigorous sense of this last term, a task which is not trivial even in the finite dimensional setting. Intuitively white noise is a stochastic process whose values at distinct times are independent, or at least uncorrelated. This causes no difficulties in discrete time but such processes simply do not exist in continuous time, as the associated spectral density would perforce be constant over all frequencies, leading to infinite variance of the process. However, use of either the Ito stochastic calculus or distribution theory leads to a viable theory. Now, as already noted, the optimal estimate of the state at time t given the observations $y(s)$, $0 \leqslant s \leqslant t$, is the conditional expectation. This expectation $\hat{x}(t)$ is then shown, assuming Gaussian noise and initial condition $x(0)$, to evolve according to the famous Kalman-Bucy (1961) equation

$$\hat{x}(t, \omega) = A\hat{x}(t, \omega) + Bu(t, \omega) + K(t)(y(t, \omega) - C\hat{x}(t, \omega) - Du(t, \omega)),$$
$$\hat{x}(0, \omega) = E(x(0)),$$

and where the gain function $K(t) = P(t)C^*$ is independent of the observations and can hence be precomputed; doing so involves solving a Riccati equation for $P(\cdot)$, which is the covariance of the error $x - \hat{x}$. Finally, the "separation principle" (Wonham, 1968) states that under various hypotheses, including the linear-quadratic case, the process of optimal control can be carried out in two stages: first, make an optimal state estimate $\hat{x}$ as above, then determine a feedback control as in (6) with x replaced by $\hat{x}$.

The author gives infinite dimensional generalizations of these fundamental results, involving in particular a novel definition of white noise as a process defined over a "probability space" whose measure is only finitely additive. This measure is in fact the standard Gauss measure on a Bochner L^2-space. This definition is preceded by a discussion of cylinder measures and Gaussian measures on Hilbert spaces, but without the use of abstract Wiener spaces to obtain countably additive extensions.

It is by now apparent that the prerequisites for an intelligent reading of this book are substantial. They are real and complex analysis, Fourier transforms and PDE, some functional analysis, probability theory (characteristic functions, Ito integrals, stochastic processes and their associated function space measures), and enough classical systems, control, and filtering theory to grasp the motivation for the author's selection of topics and results. Not all of these, especially the probability theory, are acknowledged in the Preface. But since this is alleged to be suitable as a text (having been so used at UCLA), they should be fully recognized. Do *your* students have this background?

Oftentimes one wishes that the author would slow down a bit and explain more carefully what motivates the introduction of certain ideas. Why for example are nuclear operators or holomorphic semigroups or Gaussian measures of interest? Sometimes we eventually get an answer (nonnegative nuclear operators serve as the covariance operators for countably additive Gaussian measures), but sometimes we don't (parabolic PDE's are governed by holomorphic semigroups; Gaussian measures can be induced on L^2 spaces by stationary Gaussian processes, for example). One also wishes for some general perspective on the subject of infinite dimensional systems analysis. What is its history, its purpose, etc.? Why is the author's definition of white noise to be preferred over some other possibility, such as one based on the more intuitive concept of an infinite dimensional Wiener process [5], [11]? Fortunately, much of this missing perspective is available elsewhere in the literature [3], [5], [9], [12].

There is also the related question of utility–is all this heavy theory suitable for application? Are the basic mathematical models realistic descriptions of interesting physical situations, or are they chosen more for reasons of mathematical convenience? There is little discussion of this type of question in Balakrishnan's book, and a corresponding lack of examples. It's certainly well understood that (mathematical) systems involving time delays or distributed parameters require infinite dimensional state spaces, and that there are important physical systems leading naturally to such models (transmission lines, semiconductors, vibrating beams, moving fluids, etc.) However, the proper modelling of such systems is a somewhat controversial topic [12], as is the computational problem of making finite dimensional approximations and assessing their accuracy [3]. In general it seems fair to say that present infinite dimensional systems theory is far ahead of potential applications to interesting engineering problems.

In summary, this book is a tour-de-force by the author who has contributed extensively to optimization and systems theory. It deals rigorously and at times idiosyncratically with sophisticated mathematical models of mostly uncertain applicability to the real world. From this viewpoint it is a significant contribution to a currently active area of research.

References

1. D. Aldous, *A characterization of Hilbert space using the central limit theorem*, J. London Math. Soc. **14** (1976), 376–380.

2. K. Aström, *Introduction to stochastic control theory*, Academic Press, New York, 1970.

3. M. Athans, *Towards a practical theory for distributed parameter systems*, IEEE Trans. Automatic Control, **AC-15** (1970), 245–247.

4. A. Balakrishnan, *Introduction to optimization theory in a Hilbert space*, Lecture Notes in Econ. and Math. Systems no. 42, Springer-Verlag, Berlin and New York, 1971.

5. A. Bensoussan, *Filtrage optimal des systèmes linéaires*, Dunod, Paris, 1971.

6. R. Brockett, *Finite dimensional linear systems*, Wiley, New York, 1970.

7. M. Brodskiĭ, *Triangular and Jordan representations of linear operators*, Transl. Math. Monographs, no. 32, Amer. Math. Soc.,Providence, R. I., 1972.

8. R. Bucy and P. Joseph, *Filtering for stochastic processes with applications to guidance*, Wiley, New York, 1968.

9. R. Curtain, *A survey of infinite dimensional filtering*, SIAM Review **17** (1975), 395–411.

10. N. Dunford and J. Schwartz, *Linear operators* (Part II: *Spectral theory*), Interscience, New York, 1963.

11. P. Falb, *Infinite dimensional filtering*, Information and Control **11** (1967), 102–137.

12. W. Helton, *Systems with infinite dimensional state space: The Hilbert space approach*, Proc. IEEE **64** (1976), 145–159.

13. A. Jazwinski, *Stochastic processes and filtering theory*, Academic Press, New York, 1970.

14. T. Kailath, *A view of three decades of linear filtering theory*, IEEE Trans. Information Theory, **IT-20** (1974), 145–181.

15. J. Lindenstrauss and L. Tzafriri, *On the complemented subspaces problem*, Israel J. Math. **9** (1971), 263–269.

16. K. Maurin, *Methods of Hilbert spaces*, Polish Scientific Publishers, Warsaw, 1967.

17. T. McGarty, *Stochastic systems and state estimation*, Interscience, New York, 1974.

18. C. Pearcy (Editor), *Topics in operator theory*, Math. Surveys, no. 13, Amer. Math. Soc., Providence, R. I., 1974.

19. B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space*, North-Holland, Amsterdam, 1970.

RICHARD B. HOLMES

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 1, January 1978

Geometry of spheres in normed spaces, by Juan Jorge Schäffer, Lecture Notes in Pure and Appl. Math., vol. 20, Dekker, New York, 1976, vi + 228 pp., \$24.50.

Geometric properties of the unit sphere of a Banach space have proved to give much information about the general nature of the space. For example, it has long been known that a Banach space is reflexive if its unit sphere is uniformly convex; this has been strengthened, so that it is now known that X is isomorphic to a space for which no two-dimensional sections of the unit sphere are nearly squares if and only if X is super-reflexive (no nonreflexive space has all its finite-dimensional subspaces "nearly isometric" to subspaces of X). Another spectacular example is the fact that all infinite-dimensional Banach spaces have arbitrarily large finite-dimensional subspaces that are nearly Euclidean, which has been widely useful and revealing. This book contains much new information about certain aspects of the geometry of unit spheres. It might be described as a detailed and comprehensive study of the girth, perimeter, radius, and diameter of unit spheres of Banach spaces. This field is new and interesting, perhaps even weird. It is not yet clear how important it will be for the study of Banach spaces, but it has connections with several concepts of current research interest, e.g., super-reflexivity, the Radon-Nikodým property, infinite trees, and preduals of $L^1(\mu)$-spaces. Although accessible to beginning students, the book seems primarily of value to research mathematicians interested in some of the concepts mentioned in this review. A nonspecialist might be confused by the frequent mixing of important and not-so-important facts.

With the aim of minimizing details and giving a feeling of the type of results involved, it seems best to describe some interesting facts about the