Lecture Notes in Control and Information Sciences

Edited by A.V. Balakrishnan and M. Thoma

8

Ruth F. Curtain
Anthony J. Pritchard

Infinite Dimensional Linear Systems Theory

Springer-Verlag
Berlin · Heidelberg · New York

Lecture Notes in Control and Information Sciences

Edited by A.V. Balakrishnan and M. Thoma

8

Ruth F. Curtain

Anthony J. Pritchard

Infinite Dimensional Linear Systems Theory

Springer-Verlag

Berlin Heidelberg New York 1978

Authors
Dr. Ruth F. Curtain
Control Theory Centre,
University of Warwick
presently at:
Rijksuniversiteit te Groningen
Mathematisch Instituut
Groningen Netherlands

Dr. Anthony J. Pritchard
Director, Control Theory Centre
University of Warwick
Coventry CV4 7AL
United Kingdom

ISBN 3-540-08961-6 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-08961-6 Springer-Verlag New York Heidelberg Berlin

Printed in Germany

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2061/3020-543210

Lecture Notes in Control and Information Sciences

Edited by A. V. Balakrishnan and M. Thoma

Vol. 1: Distributed Parameter Systems: Modelling and Identification
Proceedings of the IFIP Working Conference, Rome, Italy, June 21–26, 1976
Edited by A. Ruberti
V, 458 pages. 1978

Vol. 2: New Trends in Systems Analysis
International Symposium, Versailles, December 13–17, 1976
Edited by A. Bensoussan and J. L. Lions
VII, 759 pages. 1977

Vol. 3: Differential Games and Applications
Proceedings of a Workshop, Enschede, Netherlands, March 16–25, 1977
Edited by P. Hagedorn, H. W. Knobloch, and G. J. Olsder
XII, 236 pages. 1977

Vol. 4: M. A. Crane, A. J. Lemoine
An Introduction to the Regenerative Method for Simulation Analysis
VII, 111 pages. 1977

Vol. 5: David J. Clements, Brian D. O. Anderson
Singular Optimal Control: The Linear Quadratic Problem
V, 93 pages. 1978

Vol. 6: Optimization Techniques
Proceedings of the 8th IFIP Conference on Optimization Techniques, Würzburg, September 5–9, 1977
Part 1
Edited by J. Stoer
XIII, 528 pages. 1978

Vol. 7: Optimization Techniques
Proceedings of the 8th IFIP Conference on Optimization Techniques, Würzburg, September 5–9, 1977
Part 2
Edited by J. Stoer
XIII, 512 pages. 1978

Vol. 8: R. F. Curtain, A. J. Pritchard
Infinite Dimensional Linear Systems Theory
VII, 298 pages. 1978

Vol. 9: Y. M. Ez-Fattah, C. Foulard
Learning Systems
Decision, Simulation, and Control
VII, 119 pages. 1978

Vol. 10: M. Maciejowski
The Modelling of Systems with Small Observation Sets
VII, 241 pages. 1978

Lecture Notes in Economics and Mathematical Systems

For information about Vols. 1–99 please contact your bookseller or Springer-Verlag

Vol. 100: B. S. Duran and P. L. Odell, Cluster Analysis. A Survey. VI, 137 pages. 1974.

Vol. 101: W. M. Wonham, Linear Multivariable Control. A Geometric Approach. X, 344 pages. 1974.

Vol. 102: Analyse Convexe et Ses Applications. Comptes Rendus, Janvier 1974. Edited by J.-P. Aubin. IV, 244 pages. 1974.

Vol. 103: D. E. Boyce, A. Farhi, R. Weischedel, Optimal Subset Selection. Multiple Regression, Interdependence and Optimal Network Algorithms. XIII, 187 pages. 1974.

Vol. 104: S. Fujino, A Neo-Keynesian Theory of Inflation and Economic Growth. V, 96 pages. 1974.

Vol. 105: Optimal Control Theory and its Applications. Part I. Proceedings 1973. Edited by B. J. Kirby. VI, 425 pages. 1974.

Vol. 106: Optimal Control Theory and its Applications. Part II. Proceedings 1973. Edited by B. J. Kirby. VI, 403 pages. 1974.

Vol. 107: Control Theory, Numerical Methods and Computer Systems Modeling. International Symposium, Rocquencourt, June 17–21, 1974. Edited by A. Bensoussan and J. L. Lions. VIII, 757 pages. 1975.

Vol. 108: F. Bauer et al., Supercritical Wing Sections II. A Handbook. V, 296 pages. 1975.

Vol. 109: R. von Randow, Introduction to the Theory of Matroids. IX, 102 pages. 1975.

Vol. 110: C. Striebel, Optimal Control of Discrete Time Stochastic Systems. III. 208 pages. 1975.

Vol. 111: Variable Structure Systems with Application to Economics and Biology. Proceedings 1974. Edited by A. Ruberti and R. R. Mohler. VI, 321 pages. 1975.

Vol. 112: J. Wilhlem, Objectives and Multi-Objective Decision Making Under Uncertainty. IV, 111 pages. 1975.

Vol. 113: G. A. Aschinger, Stabilitätsaussagen über Klassen von Matrizen mit verschwindenden Zeilensummen. V, 102 Seiten. 1975.

Vol. 114: G. Uebe, Produktionstheorie. XVII, 301 Seiten. 1976.

Vol. 115: Anderson et al., Foundations of System Theory: Finitary and Infinitary Conditions. VII, 93 pages. 1976

Vol. 116: K. Miya[illegible], Input-Output Analysis and the Structure of Income Distributio[illegible] 135 pages. 1976.

Vol. 117: Optimization and Operations Research. Proceedings 1975. Edited by W. Oettli and K. Ritter. IV, 316 pages. 1976.

Vol. 118: Traffic Equilibrium Methods, Proceedings 1974. Edited by M. A. Florian. XXIII, 432 pages. 1976.

Vol. 119: Inflation in Small Countries. Proceedings 1974. Edited by H. Frisch. VI, 356 pages. 1976.

Vol. 120: G. Hasenkamp, Specification and Estimation of Multiple-Output Production Functions. VII, 151 pages. 1976.

Vol. 121: J. W. Cohen, On Regenerative Processes in Queueing Theory. IX, 93 pages. 1976.

Vol. 122: M. S. Bazaraa, and C. M. Shetty,Foundations of Optimization VI. 193 pages. 1976

Vol. 123: Multiple Criteria Decision Making. Kyoto 1975. Edited by M. Zeleny. XXVII, 345 pages. 1976.

Vol. 124: M. J. Todd. The Computation of Fixed Points and Applications. VII, 129 pages. 1976.

Vol. 125: Karl C. Mosler. Optimale Transportnetze. Zur Bestimmung ihres kostengünstigsten Standorts bei gegebener Nachfrage. VI, 142 Seiten. 1976.

Vol. 126: Energy, Regional Science and Public Policy. Energy and Environment I. Proceedings 1975. Edited by M. Chatterji and P. Van Rompuy. VIII, 316 pages. 1976.

Vol. 127: Environment, Regional Science and Interregional Modeling. Energy and Environment II. Proceedings 1975. Edited by M. Chatterji and P. Van Rompuy. IX, 211 pages. 1976.

Vol. 128: Integer Programming and Related Areas. A Classified Bibliography. Edited by C. Kastning. XII, 495 pages. 1976.

Vol. 129: H.-J. Lüthi, Komplementaritäts- und Fixpunktalgorithmen in der mathematischen Programmierung. Spieltheorie und Ökonomie. VII, 145 Seiten. 1976.

Vol. 130: Multiple Criteria Decision Making, Jouy-en-Josas, France. Proceedings 1975. Edited by H. Thiriez and S. Zionts. VI, 409 pages. 1976.

Vol. 131: Mathematical Systems Theory. Proccedings 1975. Edited by G. Marchesini and S. K. Mitter. X, 408 pages. 1976.

Vol. 132: U. H. Funke, Mathematical Models in Marketing. A Collection of Abstracts. XX, 514 pages. 1976.

Vol. 133: Warsaw Fall Seminars in Mathematical Economics 1975. Edited by M. W. Loś, J. Loś, and A. Wieczorek. V. 159 pages. 1976.

Vol. 134: Computing Methods in Applied Sciences and Engineering. Proceedings 1975. VIII, 390 pages. 1976.

Vol. 135: H. Haga, A Disequilibrium – Equilibrium Model with Money and Bonds. A Keynesian – Walrasian Synthesis. VI, 119 pages. 1976.

Vol. 136: E. Kofler und G. Menges, Entscheidungen bei unvollständiger Information. XII, 357 Seiten. 1976.

Vol. 137: R. Wets, Grundlagen Konvexer Optimierung. VI, 146 Seiten. 1976.

Vol. 138: K. Okuguchi, Expectations and Stability in Oligopoly Models. VI, 103 pages. 1976.

Vol. 139: Production Theory and Its Applications. Proceedings. Edited by H. Albach and G. Bergendahl. VIII, 193 pages. 1977.

Vol. 140: W. Eichhorn and J. Voeller, Theory of the Price Index. Fisher's Test Approach and Generalizations. VII, 95 pages. 1976.

Vol. 141: Mathematical Economics and Game Theory. Essays in Honor of Oskar Morgenstern. Edited by R. Henn and O. Moeschlin. XIV, 703 pages. 1977.

Vol. 142: J. S. Lane, On Optimal Population Paths. V, 123 pages. 1977.

Vol. 143: B. Näslund, An Analysis of Economic Size Distributions. XV, 100 pages. 1977.

Vol. 144: Convex Analysis and Its Applications. Proceedings 1976. Edited by A. Auslender. VI, 219 pages. 1977.

Vol. 145: J. Rosenmüller, Extreme Games and Their Solutions. IV, 126 pages. 1977

Vol. 146: In Search of Economic Indicators. Edited by W. H. Strigel. XVI, 198 pages. 1977.

Vol. 147: Resource Allocation and Division of Space. Proceedings. Edited by T. Fujii and R. Sato. VIII, 184 pages. 1977.

Vol. 148: C. E. Mandl, Simulationstechnik und Simulationsmodelle in den Sozial- und Wirtschaftswissenschaften. IX, 173 Seiten. 1977.

Vol. 149: Stationäre und schrumpfende Bevölkerungen: Demographisches Null- und Negativwachstum in Österreich. Herausgegeben von G. Feichtinger. VI, 262 Seiten. 1977.

Vol. 150: Bauer et al., Supercritical Wing Sections III. VI, 179 pages. 1977.

Vol. 151: C. A. Schneeweiß, Inventory-Production Theory. VI, 116 pages. 1977.

Vol. 152: Kirsch et al., Notwendige Optimalitätsbedingungen und ihre Anwendung. VI, 157 Seiten. 1978.

Vol. 153: Kombinatorische Entscheidungsprobleme: Methoden und Anwendungen. Herausgegeben von T. M. Liebling und M. Rössler. VIII, 206 Seiten. 1978.

Vol. 154: Problems and Instruments of Business Cycle Analysis. Proceedings 1977. Edited by W. H. Strigel. VI, 442 pages. 1978.

Vol. 155: Multiple Criteria Problem Solving. Proceedings 1977. Edited by S. Zionts. VIII, 567 pages. 1978.

Vol. 156: B. Näslund and B. Sellstedt, Neo-Ricardian Theory. With Applications to Some Current Economic Problems. VI, 165 pages. 1978.

PREFACE

Over the past five years the major research effort of the Control Theory Centre at the University of Warwick has been in the area of infinite dimensional system theory. The philosophy underlying the research has been to develop a mathematical framework which enables the generalisation of the finite dimensional results to infinite dimensions and which includes both distributed parameter systems and differential delay systems as special cases. So following the lead of Fattorini in [6], we describe the system dynamics in terms of a strongly continuous semigroup on an appropriate Banach space. Using this unifying mathematical approach it is possible to clarify the essential concepts of observability, controllability, the quadratic cost control problem, and the estimation and control problems for stochastic systems. At this stage we feel it is appropriate to present the culmination of this collective effort in a coordinated way in the form of lecture notes.

Of course it has not been possible to cover all of the systems theory concepts, and significant omissions are realization theory and identification. Readers interested in these areas should see for instance [11] and watch out for the forthcoming monograph by Fuhrmann [7]. We hope, however, that these notes will provide an introduction to infinite dimensional system theory, accessible to readers with a knowledge of finite dimensional theory and some functional analysis. The treatment of the material reflects our own personal approach and is by no means the only way or even the most commonly accepted way. While it is aesthetically pleasing that the abstract formulation yields results which mirror those for finite dimensions; because the state space is infinite dimensional this superficial resemblance can be misleading. Consequently we have examined the implications of the abstract theory to specific examples of distributed and delay systems.

Many other researchers have contributed to infinite dimensional systems theory, some using a semigroup approach and others using methods appropriate for special classes of systems. For example Butkovskii [2] and Lions [9] have examined optimal control problems for deterministic distributed parameter systems and Bensoussan [1] estimation and control of stochastic distributed systems. The pioneering work on controllability and observability of hyperbolic and parabolic partial differential equations was developed by Fattorini [5] and Russell[12],whilst Kirillova [8], Manitius [10], Banks [4] and Lee [3] have contributed to the study of differential delay systems. It has not been possible to include all of these special results for particular systems in these notes. What we have tried to do, however, is to illustrate through the examples how many of these special results may be obtained using an abstract semigroup approach and at the end of each chapter included references of other contributions.

The first chapter reviews the types of finite dimensional systems theory results which will be generalized to infinite dimensions in chapters 3-7 and gives several examples of how such problems may arise in distributed and delay systems. To make these notes self-contained, chapter 2 presents the known results on semigroups, which we shall need in later chapters. Chapters 3-7 form the core of the book, namely the extension of the finite dimensional systems theory results outlined in Chapter 1 for time-invariant linear systems. All these results are proved in detail and are illustrated by several examples and would be appropriate for an introductory graduate course in linear systems theory. Chapters 8 & 9 are concerned with extensions of the results of Chapters 3-7 to more complicated systems, namely time dependent systems and distributed systems allowing for boundary control and point observations. The difficulties which arise in trying to extend the results to more general systems are technical mathematical ones rather than conceptual ones. As the technical details are already available in the literature, we have chosen to motivate the approach using simple examples and have omitted proofs which are heavy technical extensions of those in earlier chapters. So although a complete presentation of the results is available in Chapters 8 and 9, many proofs are given in outline only. Again, considerable attention is given to analyzing the implications for distributed and delay systems by means of examples.

These lecture notes have been influenced by the many visitors to the Control Theory Centre and especially by the SRC [1] funded research fellows, S.P. Banks, A. Ichikawa, E.P. Ryan, R. Triggiani, A. Wirth and J. Zabczyk. It is also a pleasure to pay tribute to the former directors of the Centre, Professor L. Markus and Professor P.C. Parks, for their guidance

---

[1] Science Research Council (United Kingdom) which has supported the Control Theory Centre from its inception in 1970.

and encouragement in our research, and to Professor J.L. Douce and Professor J.A. Shercliff of the Engineering Department for their continuing support of the Centre. Finally special thanks to Frances Ryan for transforming a largely illegible manuscript into a respectable typescript.

A.J. Pritchard, Director of Control Theory Centre, University of Warwick.

Ruth F. Curtain, Control Theory Centre, University of Warwick.
(presently at: Mathematics Institute, Rijksuniversiteit Groningen.)

References

[1] Bensoussan, A. Filtrage optimal des systemes lineaires, Dunod 1971.

[2] Butkovskii, A.G. Theory of optimal control of distributed parameter systems, American Elsevier, 1969.

[3] Chyung, D.H. and Lee, E.B. Delayed action control problems, Automatica, 6 (1970), pp.395-400.

[4] Eller, D.H., Aggarwal, J.K. and Banks, H.T. Optimal control of linear time-delay systems, IEEE Trans., AC 14 (1969), pp.678-687.

[5] Fattorini, H.O. On complete controllability of linear systems, J. Diff. Eqns., 3 (1967), pp.391-402.

[6] Fattorini, H.O. Time optimal control of solutions of operational differential equations, SIAM J. Control, 21 (1964) pp.54-59.

[7] Fuhrmann, P.A. Notes on operators and systems, to appear.

[8] Kirillova, F.M. and Chirakova, S.V. Relative controllability of systems with time lag, Dok. Akad. Nauk. USSR, 176 (1967) pp.1260-1263.

[9] Lions, J.L. Optimal control of systems described by partial differential equations, Springer 1971.

[10] Manitius, A. Optimal control of time lag systems with quadratic performance indices, lV IFAC Congress, Warsaw 1969.

[11] Ray, W.H. and Lainiotis, D.G. (Eds), Identification, estimation and control of distributed parameter systems, Marcel-Dekker, 1977.

[12] Russell, D.L. Nonharmonic Fourier series in the control theory of distributed parameter systems, J. Math. Anal. Appl., 18 (1967), pp.542-559.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

Most dynamical systems which describe processes in engineering, physics, and economics are extremely complex and the identification of mathematical models is difficult. Consequently, early investigations of systems were confined mainly to analysing stability behaviour of very simple models using frequency domain methods. In the last fifteen or so years the state space approach has also become very popular and several new systems theory concepts such as controllability, observability, linear quadratic control, estimation and filtering, and realisation have been introduced and analysed [1],[4],[5],[9],[10]. However, these concepts are only well understood for simple systems, namely linear difference equations and linear ordinary differential equations. In the engineering jargon these simple systems are called lumped systems and they can be described by linear maps on finite dimensional linear vector spaces. For systems described by partial differential equations (distributed systems) or by delay equations the appropriate state space is an infinite dimensional function space and there has been some work on generalizing the systems theory concepts to special classes of these systems [2],[6],[7],[8].

Using a semigroup representation, we develop a self contained abstract theory for a wide class of linear systems, both finite and infinite dimensional which includes lumped, delay and distributed systems. Results are obtained which, when interpreted for a particular class of system, yields the known results. Moreover, the abstract approach clarifies the main ideas and mathematical problems so that new results are more easily obtained.

We do not consider all the systems theory concepts here, but first concentrate on controllability, observability and stabilizability which turn out to be more complicated in infinite dimensions. We then consider the quadratic cost control problem and its dual, the filtering problem, and obtain the separation principle for infinite dimensional stochastic systems. To motivate our approach we present a brief survey of the finite dimensional theory concepts and results which we will generalize in Chapters 2 - 7.

Finite dimensional linear systems theory

Here we restrict ourselves to systems which are described by linear ordinary differential equations with a given initial state. Without loss of generality, we suppose that by a suitable choice of state vector the system has been expressed in the canonical form

$$\dot{z} = Az + Bu \ ; \quad z(0) = z_o \tag{1.1}$$

where z_o, $z(t) \in R^n$, $u(t) \in R^m$, A and B are $n \times n$ and $n \times m$ real matrices respectively. u is the control term.
(1.1) is a differential equation on the state space R^n and has the unique solution

$$z(t) = e^{At} z_o + \int_o^t e^{A(t-s)} Bu(s)ds \tag{1.2}$$

We also suppose that we have an associated observation of (1.1)

$$y = Cz \tag{1.3}$$

where C is a real $k \times n$ matrix, so $y \in R^k$.

The following concepts of stability, controllability and observability for (1.1) and (1.3) are now standard.

Definition 1.1 Exponential stability

The matrix A in (1.1) is exponentially stable if there exists positive constants M, ω such that

$$\| e^{At} \| \leq Me^{-\omega t} \quad \text{for all } t \geq 0.$$

This implies that for the uncontrolled system (1.1) (with $u = 0$), $\|z(t)\| \to 0$ as $t \to \infty$. A necessary and sufficient condition for the exponential stability of A is that the real parts of its eigenvalues are strictly negative.

If one uses a feedback control $u = -Fz$, then (1.1) becomes $\dot{z} = (A-BF)z$ and this type of control can be used to stabilize an unstable system $\dot{z} = Az$.

Definition 1.2 Stabilizability

(1.1) or (A,B) is stabilizable if there exists an $m \times n$ matrix F, such that A-BF is exponentially stable.

Another important systems theory concept is whether or not a preassigned final state can be reached.

Definition 1.3 Controllability

(1.1) or (A,B) is controllable if any initial point z_o can be steered to an arbitrary final point z in some finite time t_1 by some control $u \in L_\infty(0,t_1;R^m)$.

(A,B) is controllable if and only if the following n × nm matrix has rank n : $[B : AB : A^2B : \dots : A^{n-1}B]$.

It happens that if (A,B) is controllable, then (A,B) is stabilizable. Controllability is also related to observability.

Definition 1.4 Observability

We say that (1.1),(1.3) or (A,C) is observable if $Ce^{At}z_o = 0$ for all $t \geq 0$ implies $z_o = 0$. That is for the controlled system (1.1),(1.3), a knowledge of y(t) and u(t) on a finite time interval $[0,t_1]$ uniquely determines the initial state z_o.

(A,C) is observable if and only if the rank of the following n × kn matrix is n : $[C' : A'C' : \dots : (A')^{n-1}C']$.

So controllability and observability are dual concepts in the sense that (A,C) is observable if and only if (A',C') is controllable.

If the controller and/or observer is designed so that the system is stable and controllable or observable, then the question of optimality can be considered.

The quadratic control problem (regulator problem)

Consider (1.1) and the cost functional

$$(1.4) \qquad J(u) = z(t_1)'Gz(t_1) + \int_0^{t_1} \{z(t)'Mz(t) + u(t)'Ru(t)\}dt$$

when G, M and R are real symmetric n × n, n × n, and m × m matrices respectively, with $G \geq 0$, $M \geq 0$ and $R > 0$. The regulator problem is to find an optimal control $u^* \in L_2(0,t_1;R^m)$ such that (1.4) is minimized. Under the above assumptions, there exists a unique optimal control

$$(1.5) \qquad u^*(t) = -R^{-1}B'Q(t)z(t)$$

where Q(t) is an n × n real symmetric matrix which is the unique solution of the Riccati equation

$$(1.6) \qquad \begin{cases} \dot{Q}(t) + Q(t)A + A'Q(t) + M = Q(t)BR^{-1}B'Q(t) \\ Q(t_1) = G \end{cases}$$

The optimal cost is $J(u^*) = z_o'Q(0)z_o$.

For $G = 0$ and $t_1 = \infty$ we have the infinite time regulator problem and if (A,B) and $(A',M^{\frac{1}{2}})$ are stabilizable, there exists a unique optimal feedback control of the form (1.5), where Q is time invariant and is the unique solution of the algebraic Riccati equation

$$QA + A'Q + M = QBR^{-1}B'Q \tag{1.7}$$

The filtering problem

We consider a noisy signal process and a noisy observation process described by the following stochastic differential equations

$$dz(t) = Az(t)dt + Ddw(t) \quad ; \quad z(0) = z_o \tag{1.8}$$

$$dy(t) = Cz(t)dt + Fdv(t) \tag{1.9}$$

where A, D, C and F are real $n \times n$, $n \times m$, $k \times n$, $k \times k$ matrices, z_o is a Gaussian zero mean vector random variable with covariance matrix P_o, and $w(t)$ and $v(t)$ are independent vector-valued Wiener processes of dimensions n and k and incremental covariance matrices W and V respectively. The solution of (1.8) is a zero mean Gaussian stochastic process with continuous sample paths and is given by

$$z(t) = e^{At}z_o + \int_o^t e^{A(t-s)}Ddw(s) \tag{1.10}$$

The filtering problem is to find the best estimate $\hat{z}(t)$ of the signal process $z(t)$ based on the observations $y(s)$, $0 \leq s \leq t$.

The solution is the well-known Kalman Bucy filter; $\hat{z}(t) = E\{z(t) |\ y(s);\ 0 \leq s \leq t\}$, which is given by

$$\begin{cases} d\hat{z}(t) = \{A-P(t)C'(FVF')^{-1}C\}\hat{z}(t)dt + P(t)C'(FVF')^{-1}dy(t) \\ \hat{z}(0) = 0 \end{cases} \tag{1.11}$$

where $P(t)$ is the unique solution of the Riccati equation

$$\begin{cases} \dot{P}(t) = AP(t) + P(t)A' + D'WD - P(t)C'(FVF')^{-1}CP(t) \\ P(0) = P_o \end{cases} \tag{1.12}$$

$P(t)$ is the covariance of the error process, i.e.

$$P(t) = E\{[z(t) - \hat{z}(t)]\ [z(t) - \hat{z}(t)]'\} \tag{1.13}$$

If (A',C') and $(A,DW^{\frac{1}{2}})$ are stabilizable, then the filter is stable in the sense that measures induced by $P(t)$ converge as $t_1 \to \infty$ to P, the unique solution of the algebraic Riccati equation

$$PA' + AP + DWD' = PC'(FVF')^{-1}CP \tag{1.14}$$

The Riccati equations (1.12) for filtering and (1.6) for the regulator problem are equivalent if in (1.12) we replace A by A', P(t) by $Q(t_1-t)$, D'WD by M, C by B', FVF' by R and P_o by G. A similar substitution shows that (1.14) and (1.7) are also of the same form. This relationship is known as the duality between the control and the filtering problem.

The smoothing problem and prediction problem are also of interest and are concerned with estimating z(t) based on the observations y(s), $0 \leq s \leq t_1$, where $t_1 > t$ for the smoothing problem and $t_1 < t$ for the prediction problem. The best optimal predictor $\hat{z}(t|t_1)$ satisfies the stochastic differential equation

$$(1.15) \qquad \begin{aligned} \dot{\hat{z}}(t|t_1) &= A\hat{z}(t|t_1) \\ \hat{z}(t_1|t_1) &= \hat{z}(t_1) \end{aligned}$$

where $\hat{z}(t_1)$ is the best optimal filter up to time t_1. The best linear smoother is also given in terms of the optimal filter $\hat{z}(t)$

$$(1.16) \quad \begin{cases} d\hat{z}(t|t_1) = A\hat{z}(t|t_1)dt + D'WD\lambda(t)dt + P(t)C'(FVF')^{-1}C\hat{z}(t)dt \\ \hat{z}(t_1|t_1) = \hat{z}(t_1) \end{cases}$$

where

$$(1.17) \quad \begin{cases} d\lambda(t) = -(A-P(t)C(FVF')^{-1}C)'\lambda(t)-C'(FVF')^{-1}(dy(t)-C\hat{z}(t)dt) \\ \lambda(t_1) = 0 \end{cases}$$

and P(t) is the covariance operator of (1.12).

The stochastic quadratic cost control problem

We consider the stochastic controlled differential equation

$$(1.18) \qquad dz(t) = Az(t)dt + Bu(t)dt + Ddw(t) \; ; \qquad z(0) = z_o$$

with the noisy observation process

$$(1.19) \qquad dy(t) = Cz(t)dt + Fdv(t)$$

where we make the same assumptions on A, D, W, z_o, C, F and V as for the filtering problem. B is a real $n \times m$ matrix and u(t) is an admissible stochastic control. By admissible we mean that $u \in L_2(\Omega,p;L_2(0,t_1;R^m))$ and depends only on the past observations, y(s); $0 \leq s \leq t$. The stochastic control problem is to find the admissible control which minimizes the expected value of the cost J(u) given by (1.4). The solution is usually termed the separation principle, because the optimal control strategy is to use the deterministic feedback law of (1.5), replacing z(t) by its conditional expectation $\hat{z}(t)$, which is obtained using the

Kalman Bucy filter results. More precisely, the optimal control strategy is given by

$$u^*(t) = -R^{-1}B'Q(t)\hat{z}(t) \tag{1.20}$$

$$\left\{\begin{aligned} d\hat{z}(t) &= \{A-P(t)C'(FVF')^{-1} - BR^{-1}B'Q(t)\}\hat{z}(t)dt + \\ &\qquad + P(t)C'(FVF')^{-1}dy(t) \\ \hat{z}(0) &= 0 \end{aligned}\right. \tag{1.21}$$

where $Q(t)$ and $P(t)$ are the solutions of (1.6) and (1.12) respectively. The optimal cost is then

$$\text{trace}\{GP(t_1)\} + \int_0^{t_1} \text{trace}\{MP(s)\}ds + \int_0^{t_1} \text{trace}\{Q(s)P(s)C'(FVF')^{-1}CP(s)\}ds$$

To motivate the generalisation of the above systems theory to infinite dimensions, we present several simple examples of delay and distributed systems where questions of stability, controllability, and optimality might arise.

Examples of linear infinite dimensional systems

Example 1.1

Suppose we have a thin, narrow, homogeneous, continuous material strip which is fed into a furnace by means of a variable-speed transport mechanism. Then its temperature distribution can be modelled by the diffusion equation

$$z_t(x,t) = \mu z_{xx}(x,t) + v(t)z_x(x,t) + \sigma\{z(x,t)-u(x,t)\}$$

$$z_x(0,t) = 0 = z_x(1,t)$$

where z is the temperature distribution, μ is the coefficient of diffusivity, σ is a constant proportional to the surface conductivity of the material, v is the material-strip velocity, and $u(x,t)$ is the external temperature distribution of the strip. We shall suppose that we can control u and that it is desirable to keep the outlet temperature $z(1,t)$ at some preassigned temperature $\theta(t)$ say. Thus we are led to the controllability problem of whether the desired outlet temperature can be achieved and maintained. In general this will be impossible, and so instead we may seek to minimize the functional

$$J = \int_0^{t_1} \{z(1,t)-\theta(t)\}^2 dt$$

Usually the controls are constrained and we can express this by assuming $|u(t)| \leq 1$ say, or by including a penalty for using too much control in the functional J.

Example 1.2

In steel making plants it is necessary to estimate the temperature distribution of metal slabs based on measurements at certain points on the surface. A possible model for the temperature distribution is

$$\rho C_1 z_t(x,t) = kz_{xx}(x,t) - \alpha[z(x,t)-z_o(x,t)] + \xi(x,t) \; ; \quad 0 < x < 1$$

$$z_x(0,t) = 0 = z_x(1,t)$$

where ρ, C_1, k are the density, heat capacity, and effective thermal conductivity of the metal slab, α is a heat transfer parameter, z_o is the average coolant temperature, and $\xi(x,t)$ is some distributed white noise disturbance. The problem is to estimate the temperature profile $z(x,t)$; $0 \le x \le 1$, $t > 0$, based on the noisy measurements

$$y_i(t) = z(x_i,t) + \eta_i(t) \; ; \quad i=1,2,\ldots,k$$

where x_i, $i=1,\ldots,k$ are points on the surface of the slab and $\eta_i(t)$ represents the measurement error.

Example 1.3

The evolution of the population of a country can be described by the following linear hyperbolic partial differential equation

$$\frac{\partial p(t,r)}{\partial t} + \frac{\partial p(t,r)}{\partial r} = -\mu(t,r)p(t,r)$$

$$p(0,r) = p_o(r) \; ; \quad 0 \le r \le 1$$

$$p(t,0) = u(t) \; ; \quad 0 \le t \le t_1$$

where $p(t,r)$ represents the population density of individuals of age r at time t, $\mu(t,r)$ is the mortality function, $p_o(r)$ is the given initial age distribution and $u(t)$ is the birth rate which we assume is the control variable. The problem is to choose u so as to achieve a desired age profile $q(r)$ at the final time t_1, and mathematically we could interpret this as minimizing

$$J(u) = \int_0^1 \{p(t_1,r)-q(r)\}^2 dr + \int_0^{t_1} \lambda u^2(s)ds$$

where the second term measures the social cost of controlling birthrate.

Example 1.4

Suppose we have a stretched nonuniform string whose motion is described by

$$\rho(x)z_{tt}(x,t) - \Big(\alpha(x)z_x(x,t)\Big)_x = v(x,t)$$

$$z(0,t) = 0 \; ; \quad z(1,t) = u(t)$$