Benchmark Papers
in Electrical Engineering and
Computer Science V. 27

DISTRIBUTED PARAMETER SYSTEMS THEORY, PART II
Estimation

Edited by
Peter Stavroulakis

Benchmark Papers in Electrical Engineering and Computer Science / 27

A BENCHMARK® Books Series

DISTRIBUTED PARAMETER SYSTEMS THEORY, PART II: Estimation

Edited by
PETER STAVROULAKIS
Mediterranean College, Athens

Hutchinson Ross Publishing Company
Stroudsburg, Pennsylvania

Dedicated to the memory of my father, Petros Ioannou Stavroulakis

Benchmark Papers in Electrical Engineering and Computer Science, Volume 27
Library of Congress Catalog Card Number: 82-21218
ISBN: 0-87933-072-4

85 84 83 1 2 3 4 5
Manufactured in the United States of America.

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA
Main entry under title:
Distributed parameter systems theory.
(Benchmark papers in electrical engineering and computer science ; v. 26-27)
Includes indexes.
Contents: pt. 1. Control—pt. 2. Estimation.
1. Distributed parameter systems. I. Stavroulakis, Peter. II. Series.
QA402.D58 1983 003 82-21218
ISBN 0-87933-071-6 (v. 1)
ISBN 0-87933-072-4 (v. 2)
ISBN 0-87933-103-8 (set)

Distributed worldwide by Van Nostrand Reinhold Company Inc.,
135 W. 50th Street, New York, NY 10020.

SERIES EDITOR'S FOREWORD

This Benchmark Series in Electrical Engineering and Computer Science is aimed at sifting, organizing, and making readily accessible to the reader the vast literature that has accumulated. Although the series is not intended as a complete substitute for a study of this literature, it will serve at least three major critical purposes. In the first place, it provides a practical point of entry into a given area of research. Each volume offers an expert's selection of the critical papers on a given topic as well as his views on its structure, development, and present status. In the second place, the series provides a convenient and time-saving means for study in areas related to but not contiguous with one's principal interests. Last, but by no means least, the series allows the collection, in a particularly compact and convenient form, of the major works on which present research activities and interests are based.

Each volume in the series has been collected, organized, and edited by an authority in the area to which it pertains. In order to present a unified view of the area, the volume editor has prepared an introduction to the subject, has included his comments on each article, and has provided a subject index to facilitate access to the papers.

We believe that this series will provide a manageable working library of the most important technical articles in electrical engineering and computer science. We hope that it will be equally valuable to students, teachers, and researchers.

This volume, *Distributed Parameter Systems Theory, Part II: Estimation,* is the second of a two-volume set edited by Peter Stavroulakis. It contains thirty papers on the estimation aspects of distributed parameter systems. All of the papers reproduced here have been published in the past decade, and, in fact, most of them date from 1975 or later. Although workers in this area should find both volumes of interest, they can be profitably read independently.

JOHN B. THOMAS

FOREWORD

It is my pleasure to write the prologue of this two-volume reprinted paper book on *Distributed Parameter Systems Theory,* edited by P. Stavroulakis.

Distributed parameter systems theory encompasses a variety of scientific and engineering fields. In countless physical situations one encounters systems involving parameters that are time varying and/or distributed over certain spatial domains. The dynamic behavior of these systems is governed by partial differential equations, integral, or integrodifferential equations, and sometimes by more general functional equations. Due to the fundamental nature of these problems and the importance of application areas, the study of distributed parameter systems has attracted the attention of great mathematicians and control theorists over the years. As an example, the pioneering work *Methods of Mathematical Physics* of Courant and Hilbert [1] is mentioned, which contains basic results in the study of partial differential equations.

From the applied engineering points of view, these types of systems have become the subject of serious study, since 1960, starting with the papers of Butkovsky, Wang, Lions, and others. The most important results that laid the foundations for the study of distributed parameter systems problems seem to be the Distributed Parameter Maximum Principle in optimal control, the Distributed Parameter Separation Principle in estimation and stochastic control, and the results related to the optimal location of pointwise control actuators and sensors.

The two decades (1960–1980) witnessed an unprecedented development of this field with applications to a wide range of scientific disciplines, such that the collection of basic results on distributed parameter systems theory has become long overdue. It is evident that new students and researchers in this field will find such a set of original results absolutely necessary for their work.

The Benchmark Books Series has been serving such a purpose in many scientific fields for a long time. In the present two volumes on distributed parameter systems, the reader will find many important results on control, estimation, and related topics, which are believed to make the searching process for more new results in this field much easier.

SPYROS G. TZAFESTAS

REFERENCE

1. R. Courant and D. Hilbert, *Methods of Mathematical Physics,* vol. 1, Interscience Publishers, Inc., New York, 1953, 561p.

PREFACE

In physical situations, one often encounters systems the parameters of which are distributed in both space and time. The dynamic behavior of these systems is governed by partial differential equations, integral equations, integrodifferential equations, and sometimes, more general functional equations. Optimal control problems for systems with distributed parameters frequently arise in mechanics, mathematical physics, and engineering. Sometimes situations exist in which systems that are described by high order ordinary differential equations can be considerably simplified, under certain assumptions, if we express them by partial differential equations. The fundamental equations of electromagnetism, fluid mechanics, thermodynamics, chemistry, and relativistic kinematics may be written directly as systems of partial differential equations.

This reprint book collects the major theoretical results that have been published in the general area of estimation of distributed parameter systems (DPS) in the last ten years in a cohesive volume. This book also covers the major developments on computational techniques that have been derived as well as the areas of the applications of the theoretical results. Important papers that have been published before 1970, even though they are excluded as reprint papers, are included as references in the discussion that precedes each group of papers. This book is divided into six sections: Observability, Filtering, Smoothing, and Prediction, Identification, Computational Methods, Sensitivity, Applications. Each section begins with editorial comments that are then followed, in the opinion of the editor, by original important publications that have appeared in the last ten years and that cover the subject of each section most adequately. A constraint that had to be met in selecting these papers among an impressively large set of papers was, besides the quality of each paper, the fact that the book should be kept at a manageable size of around 300 pages. We hope that this constraint did not affect our original goal to present in a balanced manner the general estimation area of distributed parameter systems to be useful to both students and researchers.

PETER STAVROULAKIS

CONTENTS

INTRODUCTION

Distributed parameter systems (DPS) theory has attained considerable maturity and sophistication in the last twenty years. The origin of this field of study usually is dated from the first papers of the Russian scientists Butkovsky and Lerner in 1960 [1]. Some work had appeared much earlier on the calculus of variations for systems described by partial differential equations in 1953 in the book *Methods of Mathematical Physics* by Courant and Hilbert [2]. This earlier book was, however, somewhat different in intent because it was purely theoretical.

The results of the theory of DPS recently have been applied to a large number of engineering fields. The key areas of the theory that have found an impressive number of applications are stability, control estimation, identification, and optimal design. Application areas in the last ten years include chemical engineering, petroleum and metallurgical industries, nuclear reactor control, plasma control, mechanical structure design problems (bridges, platforms), resource recovery (oil, water, coal) environmental problems (environmental quality modeling, control, water quality management, physiological systems (distribution and effect of drugs, for example), and sociological systems (dynamic modeling and control of behavior of groups of people), to name a few.

The fundamental aspects of observability as it relates to the ability to recover completely some prior state of a dynamic system based on partial observations of the state over some period of time for various models are covered in the first section. The question of observability is examined in conjunction with filter convergence of stochastic DPS.

The filtering technique as an estimation process is studied in the second section. Several statistical information-processing approaches are utilized. Among these stand out the Wiener-Hopf equation, orthogonal projection, maximum likelihood Bayesian approach, Fokker-Planck equations, adaptive estimation method via the partitioning estimates algorithm, and the Monte Carlo approach.

In the third section, fundamental aspects of identification as an estimation process are covered. The problems presented include parameter, initial state, and input identification.

The fourth section covers several aspects of computational methods in DPS estimation. Emphasis is placed on filter convergence by the appropriate choice of measurements and sensor location so that observability of the system is preserved.

System sensitivity to the measurement conditions is studied in the fifth section.

Finally, the last section presents a comprehensive survey of recent applications of DPS theory as it relates to a wide variety of estimation problems.

REFERENCES

1. A. G. Butkovsky and A. Y. Lerner, The Optimal Control of Systems with Distributed Parameters, *Autom. Remote Control* **21**:472–477 (1960).
2. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, Interscience Publishers, Inc., New York, 1953, 561p.

OBSERVABILITY

Editor's Comments on Papers 1 Through 4

1 **GOODSON and KLEIN**
A Definition and Some Results for Distributed System Observability

2 **YU and SEINFELD**
Observability of a Class of Hyperbolic Distributed Parameter Systems

3 **YU and SEINFELD**
Observability and Optimal Measurement Location in Linear Distributed Parameter Systems

4 **KOBAYASHI**
Discrete-time Observability for Distributed Parameter Systems

Observability refers to the ability to recover completely some prior state of a dynamic system based on partial observations of the state over some period of time. Observability is also a fundamental property in state estimation. It is true that estimates of the states converge to the best possible estimates when the measurements are taken so that the system is observable. Observability for DPS has been defined by Wang [1]. Sakawa [2] studies the determination of initial conditions of DPS on the basis of observed measurement as an observability problem. In Paper 1, Goodson and Klein define observability as the ability to establish the uniqueness of a solution of the system under study. Necessary and sufficient conditions for observability of a class of hyperbolic systems are derived by Yu and Seinfeld in Paper 2. In Paper 3, Yu and Seinfeld study the concepts of observability in conjunction with filter convergence of a class of stochastic DPS. The questions they examine are the effect of measurement locations on observability and the optimal location of measurement for state estimation. It is shown that for systems whose solutions can be expressed as eigenfunction expansions, only a few measurements need suffice for observability. Finally, in Paper 4, Kobayashi examines observability on the basis of the observed measurement data from a finite number of sensors over a finite period of time. It is shown that for finite sensor locations, a general class of DPS cannot be observable by utilizing finite step observations (finite time). Certain cases for which this is true are also investigated.

REFERENCES

1. P. K. C. Wang, Control of Distributed Parameter Systems, *Advances in Control Systems,* vol. 1, Academic Press, New York, pp. 75–172 (1964).
2. Y. Sakawa, Observability and Related Problems for Partial Differential Equations of Parabolic Type, *SIAM J. Control* **13:**14–27 (1975).

1

A Definition and Some Results for Distributed System Observability

RAYMOND E. GOODSON, MEMBER, IEEE, AND RICHARD E. KLEIN, MEMBER, IEEE

Introduction

A FUNDAMENTAL problem in distributed system control is measurement. The spatial location of measurement transducers and the information content of the resulting signals with respect to the distributed state and partial differential equation model are the primary questions. The location of sensors is not necessarily dictated by physical consideration and should be included as one question in the optimal control formulation. The information content of the measurement signals relative to the model is an indication of the capability of a feedback system to accurately control a process.

Both of these problems concerning distributed control measurement may be posed as observability questions. The observability of systems with ordinary differential equation models is concerned primarily with the information content of the measurement signals. For distributed systems, sensor location is equally important. Recognizing the measurement problems in feedback control of distributed systems, a new definition of observability is introduced. The definition states observability as a uniqueness question and does not consider the subsequent estimation problem. However, the definition is independent of a particular analytical solution form as is the present theory for ordinary differential equations. Therefore, the full range of classical and modern mathematical techniques may be employed to answer observability questions.

For systems with modal solutions, a weakened definition of observability is offered. The application of this definition to several examples provides considerable insight into the sensor location problem. While a complete theory for distributed system theory is not developed, the definition and results should provide an impetus to further research.

An observability theory for distributed systems should consider the following questions.

1) Where should the measurement transducers be located in order to provide feedback information with respect to a specified control domain?

2) Does the output measurement contain sufficient information to uniquely prescribe a solution with respect to the dynamical model and the control domain?

3) Can a stable estimator be built which operates in the presence of noise?

4) Given an estimator, can something be said about the accuracy of the estimates in the presence of noise?

For finite-state, linear ordinary differential equations, all of the applicable questions have been considered and answered for the case of Gaussian noise [1]–[3]. For nonlinear ordinary differential equation systems, estimator equations have been developed for certain cases even though questions 2) and 4) have remained unanswered [4], [5]. In addition, for certain partial differential equations [6] and ordinary differential equations with pure time delays [7], estimator equations have been derived with particular answers to questions 2) and 4). For nonlinear systems, the sparsity of results related to observability and estimation theory is partially attributed to the absence of a general theory for nonlinear differential equations. In contrast, for certain classes of partial differential equations, theory is available which provides a basis for investigating observability.

Previously, the ordinary differential equation observability results of Kalman [2] and also of Gilbert [1] have required, in total, four basic system properties, which are:

1) linearity of the differential operator,
2) group property [8] with respect to time of the system's transition operator,
3) that the class of differential operators be with respect to one independent variable only, and
4) structural decomposition properties (in the sense employed by Gilbert).

For distributed parameter systems, Wang [9] has discussed certain aspects of observability and has given a definition with respect to initial state recovery of systems expressible in state function space notation. The recovery of initial state is a useful notion in ordinary differential

Manuscript received November 21, 1968; revised June 9, 1969.
R. E. Goodson is with the School of Mechnical Engineering, Purdue University, Lafayette, Ind. 47907.
R. E. Klein is with the Department of Mechanical Engineering, University of Illinois, Urbana, Ill. 61801.

equations; however, it is of limited utility in partial differential equations and, hence, the definition by Wang [9] will not be adhered to.

When the observability of distributed systems is considered, the differential operator is defined with respect to two or more independent variables since the solution is defined over a continuum in space. In addition, only semigroup properties hold for certain partial differential operators of importance. Consequently, the definition for observability of distributed systems must, therefore, allow for both measurement data and solutions defined over a continuum. Further, backward time extrapolation of measurement data is extraneous to observability and must not be required by the definition.

This paper considers questions 1) and 2); specifically, where should measurement transducers be located and when does noise-free measurement data provide sufficient information to guarantee a unique solution to a partial differential equation in the absence of initial conditions and possibly boundary conditions?

The Partial Differential Equation Model

The general form for the partial differential equations considered in the definition and results is given by

$$B_0(u,x,t)\frac{\partial u}{\partial t} + \sum_{i=1}^{k} B_i(u,x,t)\frac{\partial u}{\partial x_i} = E(u,x,t) \quad (1)$$

where

u	is an n-column vector representing the dependent variables,
k	is the dimensionality of the spatial domain,
x	is a k vector of spatial coordinates defined over a simply connected domain $\Omega(t)$,
t	is time,
$B_i, i = 0,\cdots,k$	is an $n \times n$ matrix with elements of class C^1,
E	is an n-column vector with elements of class C^1.

The boundary conditions are assumed to satisfy

$$f(u,x,t)\,|_{x=\partial\Omega} = 0, \quad t \geq 0 \quad (2)$$

where $\partial\Omega(t)$ denotes the boundary of the spatial domain $\Omega(t)$. The elements of the column vector f are assumed to be piecewise continuous. The unknown initial condition $u(x,0)$ is assumed to belong to some class of functions to be prescribed.

The form of (1) and (2) includes most equations from mathematical physics where the continuum hypothesis [10] has been invoked to define point functions. No general solutions to (1) and (2) are available. In fact, the existence and uniqueness of solutions to these equations is an open question. However, a canonical form is useful and that of (1) and (2) is appealing both from a physics and a mathematical viewpoint since most equations of interest can be expressed by it.

Definition of Observability with Reference to the Uniqueness Question

Definition 1: The measurement vector $y(t)$ is defined by

$$y(t) = \int_{\Omega_m(t)} dx C(x,t)u(x,t), \quad 0 \leq t_1 \leq t \leq t_2,\ t_2 > t_1 \quad (3)$$

where $C(x,t)$ is an $r \times n$ matrix, and $\Omega_m(t)$, where $\Omega_m(t) \subset \Omega(t)$, is the measurement domain.

Definition 2: Let $\Omega_c(t)$, where $\Omega_c(t) \subset \Omega(t)$, be that spatial domain (or space-time domain $\Omega_c(x,t)$ if x assumes prescribed values over some time interval) with respect to which the observability question is posed.

Definition 3—Definition of Observability: Given the partial differential equation model in (1), the system is said to be observable in the domain $\Omega_c(t)$ [or $\Omega_c(x,t)$] if and only if a unique solution $u(x,t)$ in $\Omega_c(t)$ [or $\Omega_c(x,t)$] is established by the boundary conditions in (2) and the measurements $y(t)$.

Remark 1: The introduction of an Ω_c domain is a major departure from linear ordinary differential equation observability theory. The need for this is particularly evident for partial differential equations of hyperbolic type with characteristic lines. In fact, an Ω_c domain over which the solution is observable is an important consequence of the observability investigation.

Remark 2: The solution vector $u(x,t)$ within Ω_c is not the system state in the sense that $u(x,\tau)$ is required to generate the solution for $t > \tau$. Furthermore, the class of admissible solutions must be prescribed for each particular case under study and this often may be accomplished by prescribing the class of initial functions.

Remark 3: Although the boundary conditions must be in the form of (2), they may be partially or entirely unknown; i.e., observability may be established by the measurements $y(t)$ alone.

Remark 4: A partial differential equation may be observable in only one or several of the elements of the vector u. Thus, a distributed system might be decomposed into observable and nonobservable solution parts.

Remark 5: The spatial location of the measurement transducers is determined by the matrix $C(x,t)$. The measurement domain $\Omega_m(t)$ is usually specified. The weighting $C(x,t)$ is part of the control design problem.

Remark 6: In the case of point measurements, $C(x,t)$ would have Dirac delta functions as elements providing the appropriate measurements.

Remark 7: For certain linear partial differential equations, particularly those governed by the Sturm–Liouville theory [11], [12], solutions may be expressed as infinite summations of the weighted modes or eigenfunctions. For such systems, a particular definition related to the eigenfunctions is useful.

Definition 4—N-Mode Observability: Let (1) be linear with stationary boundaries and linear boundary conditions. For all admissible solutions which satisfy

$$u_i(x,t) = \sum_{m=0}^{\infty} a_m{}^i g_m{}^i(t)\phi_m{}^i(x,t), \quad i = 1,\cdots,n \quad (4)$$

where $g_m{}^i(t)$ and $\phi_m{}^i(x,t)$ are known, then the system is N-mode observable in Ω if and only if the uniqueness of the coefficients $a_m{}^i$, for $m \leq N$, is established by $y(t)$ for each i.

Results

In order to establish observability with respect to the preceding definitions, uniqueness theories for partial differential equations are required. There are three main techniques used in this paper to establish uniqueness. They are analytic continuation [13], characteristic theory [14], [15], and eigenfunction expansions [11], [12], all of which are classical. The formal application of these techniques to relevant problems in the form of (1)–(3) is the subject of the remainder of this paper. The problems considered are of sufficient variety to encourage the application of the definition to other problems of interest.

The type of theory to employ in the investigation of observability for classical and nonclassical boundary value problems depends strongly on the class of partial differential equations. Since the physical phenomena actually occurring in the process and the class of the appropriate partial differential equation describing the process are strongly linked, a proper technique can often be determined from the physics. For specific information on classification, see Courant and Hilbert [15] and Petrovsky [16].

Result 1—First-Order Quasilinear Equation: In (1), let B_i and E be scalars with $B_0 = 1$. Also, assume that the class of unknown initial functions $u(x,0)$ satisfies a Lipschitz condition. This form is of importance in fluid mechanics, reactors, heat exchangers, and distillation columns, for example.

Characteristic lines in the x,t space exist for this scalar equation, which represents the simplest example of hyperbolic-like behavior. Define the domains

$$\Omega(t) = \{x \mid 0 \leq x_i \leq a_i;\quad i = 1,\cdots,k\} \tag{5}$$

for $0 \leq t \leq T$, $T > 0$. The Cauchy initial value theory is directly applicable to this problem [15]. Equation (1) for B_i, $i = 0,\cdots,k$, reduces to the ordinary differential equation set

$$du/dt = E \tag{6}$$

$$dx_i/dt = B_i,\quad i = 1,2,\cdots,k. \tag{7}$$

Equation (7) defines the characteristic directions in x,t along which the rate of change of the solution is defined by (6).

For B_i not functions of $u(x,t)$, the characteristic directions are known in advance and the observability requirements may be determined by a construction procedure as follows. For a given domain of interest $\Omega_c(x,t) \subset \Omega(t) \times t$, construct, by integrating (7), the family of characteristic manifolds passing through $\Omega_c(x,t)$ and contained in $\Omega(t) \times t$. A necessary and sufficient condition for uniqueness of the solution $u(x,t)$ in $\Omega_c(x,t)$ is that the solution value u be given over a measurement manifold $\Omega_m(t)$ which intersects once the entire family of characteristic manifolds determined by $\Omega_c(x,t)$ and (7). This construction procedure is indicated in Fig. 1, for $k = 2$.

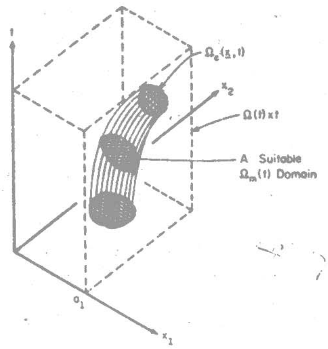

Fig. 1. Envelope of characteristic manifold and suitable measurement domain.

Remark 1a): An immediate conclusion from this result is that for more than one space dimension ($k > 1$) no finite set of point transducers recording data continuously in time is sufficient to observe the solution over any domain $\Omega_c(x,t)$ except for certain characteristic subsets of $\Omega \times t$ (which have zero measure).

Remark 1b): The general quasilinear problem occurs when the B_i coefficients are also functions of the dependent variable u. Two modifications to the discussion become necessary. First, the characteristic directions given by (7) become functions of u; thus, the space-time locations of $\Omega_m(t)$ for observability in $\Omega_c(x,t)$ may vary. Second, the solution u may be extended only locally, in general, since the nonlinear Cauchy theory is valid only in the small. However, for particular nonlinear problems, solutions in the large may be obtained using the characteristic theory.

Remark 1c): Results similar to those of Remarks 1a) and 1b) hold for the vector case of (1) when the equations are hyperbolic.

Result 2—Linear One-Dimensional Parabolic Equation: Let (1) take the form

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \frac{\partial u}{\partial t} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{\partial u}{\partial x_1} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} u \tag{8}$$

where $\Omega = \{x_1 \mid 0 \leq x_1 \leq 1\}$ and $|u_1(x_1,0)| < U$, where U is a prescribed bound. Also, let

$$C(x_1,t) = \begin{bmatrix} \delta(x_1 - x^*) & 0 \\ 0 & \delta(x_1 - x^*) \end{bmatrix} \tag{9}$$

where

$$\Omega_m(t) = \{x_1 \mid 0 \leq x_1 \leq 1\}$$

for $0 \leq t_1 \leq t \leq t_2$, $t_2 > t_1$ and where $\delta(x)$ is the Dirac delta function. Hence, x^* corresponds to the point of measurement. Now, for f_i at $x_1 = 0$ and $x_1 = 1$ analytic in t for $t \geq 0$, the solution u is observable over the domain

$$\Omega_c(x,t) = \{(x_1,t) \mid 0 \leq x_1 \leq 1,\ t \geq t_2\}.$$