THE STATE IDENTIFICATION OF A CLASS OF DISTRIBUTED SYSTEMS

GORDON ADOLPH PHILLIPSON

SRC-68-12

SYSTEMS RESEARCH CENTER

CASE WESTERN RESERVE UNIVERSITY

UNIVERSITY CIRCLE

CLEVELAND, OHIO 44106

JUNE, 1968

THE STATE IDENTIFICATION OF A CLASS

OF DISTRIBUTED SYSTEMS

by

Gordon Adolph Phillipson

ABSTRACT

The problem of state identification of a distributed system is phrased as a variational problem, that of minimizing a Quadratic functional on an appropriate Hilbert space. The systems considered are those described by linear parabolic and second order hyperbolic partial differential equations. A variety of state identification problems associated with these systems are posed.

The solutions of these variational problems are characterized by cannonical sets of equations, obtained by invoking a specialized "maximum principle" derived from the theory of variational inequalities. These canonical equations are of a special type and consideration is given to two methods of obtaining numerical solutions, one of which proceeds via a Ricatti-like decoupling, the other being an iterative technique involving conjugate directions of search on a quadratic error surface.

The proposed numerical schemes are carried out on an illustrative example.

ACKNOWLEDGEMENTS

I should like to express my sincere appreciation to Dr. Sanjoy K. Mitter, under whose guidance, insight and encouragement this thesis evolved. Thanks are also due to Dr. Jaques L. Lions for providing a flawless mid-course correction in the direction of the thesis and to Dr. John D. Pearson for his contribution of the initial idea.

The Systems Research Center provided a stimulating environment for this undertaking, made more so the the presence of staff and students ever willing to assist by providing meaningful criticism and discussion.

Words can only indicate the gratitude which I owe my wife Ellen for her understanding and support over those sometimes tortuous years. Appreciation is also due to Miss L. Cottrell for her competent typing of the manuscript.

Financial support was provided in part by the industrial sponsors of the Group C Project in the Systems Research Center.

TABLE OF CONTENTS

Page

LIST OF COMMON SYMBOLS

Matrices

$A\ B\ F\ G\ I\ WW^T$	Coefficients of Vector Differential Equations
P	Matrix of Ricatti Variables
$\Phi(\cdot,\cdot)$	System State Transition Matrix

Vectors

$\underline{f}$, $\underline{g}$, $\underline{z}$	Vector of Forcing Functions
$\underline{y}$, $\underline{p}$	System State and Adjoint State
$\underline{u}$, $\underline{v}$, $\underline{w}$	Vector of Initial and Boundary Conditions

Scalars

i, j, k, l, m, n	Integers
α, β, γ	Coefficients
λ_i	i-th Eigenvalue

Functions

$y(\cdot,\cdot)$, $p(\cdot,\cdot)$	Distributed System State, Adjoint State
$f(\cdot,\cdot)$	Distributed Forcing Functions
$z(\cdot,\cdot)$, $z_1(\cdot)$, $z_2(\cdot)$	Distributed, Boundary and Initial Condition Measurements, Respectively
$u_1(\cdot)$, $u_2(\cdot)$	Refined Estimate of True Boundary and Initial Conditions
$v_1(\cdot)$, $v_2(\cdot)$	Any Estimate of Boundary and Initial Condition

$w_i(\cdot)$	i-th Element of a Complete Basis in a Hilbert Space

Functionals

$a(\underline{u},\underline{v})$	Bilinear Functional of $\underline{u}$, $\underline{v}$
$l(\underline{v})$	Linear Functional of $\underline{v}$
$J(\underline{v})$	Quadratic Functional of $\underline{v}$

Operators

$A[\cdot]$	Elliptic Partial Differential Operator

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A.0 The Identification Problem in Context

The intelligent control of the behaviour of some abstract system experiencing constraining influences from its environment is an all-pervading theme in the engineering experience. Frequently, the exertion of this "intelligent control" must be performed in the face of inexact and incomplete data pertaining to the evolution of the system in its environment. This is the context in which the identification problem arises. Stated simply, the problem or goal of identification, is to provide the decision maker (intelligent controller) with a summary of the position (in a generalized coordinate space) of the system, which is optimal, (in some sense) with respect to the available data pertaining to environmental interaction and system evolution.

This study is concerned with systems whose evolution processes are described by a class of partial differential equations, so that in the context of the preceding discussion, the "position" of the system is "summarized" by a specification of the solution to the given partial differential equation. This specification is contingent upon a knowledge of the environmental interaction processes. That is, the generation of the solution to the partial differential equation requires a statement of the:

(i) Initial condition(s)

(ii) Environmental forcing terms, which include the boundary conditions.

It is assumed that the initial conditions (i) with which the system begins its evolution as well as the environmental interaction processes (ii) are not known with precision. However the "available data" includes inexact measurements of (i) and (ii) and in addition, provides inexact and possibly incomplete measurements of the state of evolution of the system. Thus the identification problems considered in this study are the following:

P Determine, on the basis of the available data, an estimate of the true initial and boundary conditions associated with a given partial differential equation which is, in some sense, optimal with respect to the given data.

Problem P is the so called State Identification problem, as distinguished from the system identification or parameter identification problems, wherin the objectives are to specify the system evolution process or to obtain estimates on parameters influencing the evolution of some known system, respectively.

The basis for selecting the estimates of the boundary and initial conditions associated with a given partial differential equation, that is, the criterion of optimality, is that of "least squares". To be more precise, we mean the following:

Given:

(i) The measurement data, which we denote here by Z, and

(ii) An (arbitrary) solution of the partial differential equation, denoted here by $Y(\underline{v})$, where $\underline{v}$ is an arbitrary estimate of the true initial state and boundary conditions, then

Obtain:

(i) $\underline{v}$ which extremizes the <u>error functional</u>

$$J(\underline{v}) = ||Z - Y(\underline{v})||^2,$$

where $||\cdot||^2$ is some appropriate squared metric.

Evidently then, the identification problem, as treated in this study, when stripped of all the descriptive embellishment, is a variational problem - that of characterizing extremals to a given functional, constrained by a partial differential equation. To the author's knowledge, there is no similar treatment of the identification problem extant in the literature. However, because of the variational framework of the problem, the results of other investigators in the area of optimal control of distributed systems is relevant. In that context, we should mention the pioneering work of Wang [25] in the USA and A.G. Butkovskii [3,4] in the USSR. Both investigators presented a "maximum principle" for a very general class of partial differential equations and "cost" functionals. These principles were (and are) weakened by certain formal assumptions which were necessitated (in part) by the generality of the class of variational problems.

A salient difference in the statement of the maximum principles of Wang and Butkovskii is that the latter used an integral equation approach to the "solution", whereas the former applied the dynamic programming formalism. Because of the algorithmic character of this formalism, it is used almost exclusively in contemporary studies of the optimal control problem for distributed systems. A feature of solutions obtained in this way is that they are formal, that is, the principle may characterize a non-existent solution.

It has been apparent to some investigators (principally Lions [15], Erzberger and Kim [6]) that a reduction in generality of the class of distributed systems treated would lead to a strengthened characterization of extremals to the associated variational problem. Erzberger and Kim treat a diffusion system via dynamic programming, while Lions generates a new maximum principle, giving necessary and sufficient conditions to the solution of certain specialized variational problems associated with general diffusion and "wave" systems.

We use this principle and obtain new results for the identification problem associated with diffusion and wave type systems, phrased as a variational problem. Lest there be any confusion, we state that the variational problems arising in optimal control and in the special phrasing of the identification problem are different.

Another approach to the distributed variational problem has been taken. There, the distributed systems are first "lumped" and then, to this equivalent lumped system, the established maximum principle of L.S. Pontriagin is applied. This approach, while intuitively appealing has mathematical pitfalls, which render the results obtained in this way unsound. See for example, [11].

We should say that numerical techniques for the solution of these variational problems are conspicuous by their absence, although

[6] does present a numerical solution. Two numerical methods for the solution of the variational problem are presented here, one of which is novel. In addition, a new numerical technique for the approximation of solutions to distributed systems using spline interpolation is also considered. Before presenting a review of the thesis content however, we first remark on the identification criterion adopted (namely least squares) and the connection with another approach.

A.1 Stochastic Respectability

The identification problem, as introduced, is customarily given a stochastic treatment. In that context, the error associated with the measurement data Z is endowed with statistical properties. That is, the error is considered to be a statistically random variable, whose values are "distributed" in a known way. The state identification or filtering problem, as it is called in this context, is to determine the a-posteriori probability density of the state, given the measurements Z. Equivalently, the definition of the sufficient statistics of this distribution solves the filtering problem. In a sense, these sufficient statistics are optimal estimates of the system state, since knowledge of them enables a statement of the "maximum likelihood estimate" or any other statistical estimate.

Under special statistical hypothesis on the error processes, namely that they be purely random with Gaussian probability density and in addition, are additive - that is

$$Z = Y(\underline{u}) + E$$

where $\underline{u}$ is the true "state of nature" and E is the error process, then if the system state evolution process is also linear, the a-posteriori density of the states is also Gaussian. We show, in Chapter III, that the filtered estimate (given in terms of the sufficient statistics of the Gaussian distribution of the states, the mean and variance) coincides with the "least squares" estimate. Thus, under these special hypothesis, the variational and stochastic approaches yield identical results, and the variational approach has "stochastic respectability".

The filtering problem introduced, except in the special case considered, is difficult. Indeed, even in the lumped case, where results have been obtained [2,12,26] solutions to the resulting equations characterizing the a-posteriori probability density of the states must be approximated. In the distributed case, Falb [7] has recently obtained equations for the sufficient statistics of the pertinent probability density under the specialized hypothesis mentioned.

It was the knowledge of these difficulties and also the "stochastic respectability" of the least squares estimate which persuaded the author to take the variational approach.

A.2 Distributed Systems

In section A.0, we indicated that the systems of immediate concern are those whose states are described by partial differential equations. Such systems are commonly called distributed parameter systems or simply, distributed systems. To be specific, the state evolution processes (that is, the differential equations) which are considered in this study are of two types:

(i) Linear Parabolic partial differential equations and

(ii) Linear second order hyperbolic partial differential equations.

We give, in Chapter III, a precise definition of these equations - along with an example, (in each case) of the sort of physical phenomena characterized by these equations. Suffice it to say here that one can construct examples (admittedly idealized) from across the industrial spectrum - from glass making to aerospace. For example, the temperature of molten glass flowing slowly in a forehearth is described by an equation of type (i), whereas the displacements occasioned by dynamic loading on a slender airframe can be described by equations of type (ii). In the former example, temperature measurements are available at selected points along the spatial domain (obtained by pyrometer or some other device), whereas in the latter case strain guage measurements at selected points on the airframe are reduced to yield the deflection data. In both cases, the measurements are incomplete in the sense that the entire spatial profile is not available. Moreover, the measurements are inexact in virtue of inherent errors of measurement associated with transducing elements and also because of the measurement environment.

A.3 Summary of Thesis Content

In Chapter II, we consider the notation and theoretical preliminaries necessary to the mathematical statement of the identification problem - which, together with a precise statement of the class of systems considered, is given in Chapter III. Chapter III also contains a demonstration of the stochastic respectability of the least squares estimate. In Chapter IV, a number of identification problems are posed, and their solutions are given in terms of a canonical set of partial differential equations. In Chapter V, we consider two methods for solving this canonical set of equations. These solution techniques are applied to an illustrative numerical example and the results are evaluated. Concluding remarks and suggestions for future work are given in Chapter VI.

We claim that the approach to the solution of the variational problems, arising from our phrasing of the distributed identification problem, is new, as are the results. The numerical methods posed in Chapter V are novel, as is a suggested scheme for the approximation of the solution to parabolic and second order hyperbolic partial differential equations based on cubic spline interpolation, which is omitted in this report, but which appears in [21].

CHAPTER II

NOTATION, DEFINITIONS, THEORETICAL PRELIMINARIES

A.0 Introduction

This chapter contains the theoretical foundations on which the solution to the identification problem is constructed. We begin, in section A.1, by consideration of the notation to be used in the sequel and the presentation of some pertinent definitions. In section A.2, we give the properties of the solutions to two classes of partial differential equations. In section A.3, a method for approximating the solution of partial differential equations, the Galerkin method, is presented along with some error estimates associated with the method. Section A.4 contains the fundamental theorems concerning the properties of extremals to quadratic functionals in a Hilbert space, which relate directly to the solution of the identification problem, appropriately phrased.(1) Those theorems are due to Lions and Stampaccia [16].

A.1 Notation and Definitions

We shall consider functions defined on the following sets:

Let Ω = A simply connected, open set in R^r. Points of Ω are denoted by $x=(x_1 x_2 \ldots x_r)$. Let t denote time, $t \varepsilon (0,T]$.

Let Γ denote the boundary of Ω.

Let $\Sigma = \Gamma \times (0,T]$.

$Q = \Omega \times (0,T]$.

---

1 See Chapter III, section A.2.

The set Q may be visualized, with the aid of figure II.1 as a cylindrical volume in (r+1)-dimensional Euclidean space enclosed within the sheath with "walls" Σ and "bottom" Ω.

We shall be concerned with equivalence classes of functions of a particular genre, namely, the (separable) Hilbert spaces. In particular, let $L^2(\Omega) = H$ denote the space (equivalence class) of real functions whose second powers are integrable for the measure $dx=(dx_1,dx_2 \dots dx_r)$. Define the inner product and norm of elements $f, g \varepsilon H$:

Inner Product: $(f,g)_{L^2(\Omega)} = \int_\Omega f(x)g(x)dx.$

Norm: $||f||_{L^2(\Omega)} = [\int_\Omega f(x)^2 dx]^{1/2}$

Let $L^2(\Sigma)$ denote the space of real functions whose second powers are square integrable for the measure $d\Sigma = (ds\ dt)$, where s is a point of Γ. Define the inner product and norm of elements $f, g\varepsilon L^2(\Sigma)$:

Inner Product: $(f,g)_{L^2(\Sigma)} = \int_\Sigma f(s,t)g(s,t)d\Sigma$

Norm: $||f||_{L^2(\Sigma)} = [\int_\Sigma f(s,t)^2 d\Sigma]^{1/2}$

Let $L^2(\Sigma)\times L^2(\Omega) = V$. The inner product and norm of elements $[f_\Sigma f_\Omega]^T$, $[g_\Sigma g_\Omega]^T \varepsilon V$ are given by:

CHAPTER II

NOTATION, DEFINITIONS, THEORETICAL PRELIMINARIES

A.0 Introduction

This chapter contains the theoretical foundations on which the solution to the identification problem is constructed. We begin, in section A.1, by consideration of the notation to be used in the sequel and the presentation of some pertinent definitions. In section A.2, we give the properties of the solutions to two classes of partial differential equations. In section A.3, a method for approximating the solution of partial differential equations, the Galerkin method, is presented along with some error estimates associated with the method. Section A.4 contains the fundamental theorems concerning the properties of extremals to quadratic functionals in a Hilbert space, which relate directly to the solution of the identification problem, appropriately phrased.(1) Those theorems are due to Lions and Stampaccia [16].

A.1 Notation and Definitions

We shall consider functions defined on the following sets:

Let Ω = A simply connected, open set in R^r. Points of Ω are denoted by $x=(x_1 x_2 \ldots x_r)$. Let t denote time, $t \varepsilon (0,T]$.

Let Γ denote the boundary of Ω.

Let $\Sigma = \Gamma \times (0,T]$.

$Q = \Omega \times (0,T]$.

---

1 See Chapter III, section A.2.

The set Q may be visualized, with the aid of figure II.1 as a cylindrical volume in (r+1)-dimensional Euclidean space enclosed within the sheath with "walls" Σ and "bottom" Ω.

We shall be concerned with equivalence classes of functions of a particular genre, namely, the (separable) Hilbert spaces. In particular, let $L^2(\Omega) = H$ denote the space (equivalence class) of real functions whose second powers are integrable for the measure $dx=(dx_1,dx_2...dx_r)$. Define the inner product and norm of elements $f,\ g\varepsilon H$:

Inner Product: $(f,g)_{L^2(\Omega)} = \int_\Omega f(x)g(x)dx.$

Norm: $||f||_{L^2(\Omega)} = [\int_\Omega f(x)^2 dx]^{1/2}$

Let $L^2(\Sigma)$ denote the space of real functions whose second powers are square integrable for the measure $d\Sigma = (ds\ dt)$, where s is a point of Γ. Define the inner product and norm of elements $f,\ g\varepsilon L^2(\Sigma)$:

Inner Product: $(f,g)_{L^2(\Sigma)} = \int_\Sigma f(s,t)g(s,t)d\Sigma$

Norm: $||f||_{L^2(\Sigma)} = [\int_\Sigma f(s,t)^2 d\Sigma]^{1/2}$

Let $L^2(\Sigma)\times L^2(\Omega) = V$. The inner product and norm of elements $[f_\Sigma f_\Omega]^T$, $[g_\Sigma g_\Omega]^T \varepsilon V$ are given by:

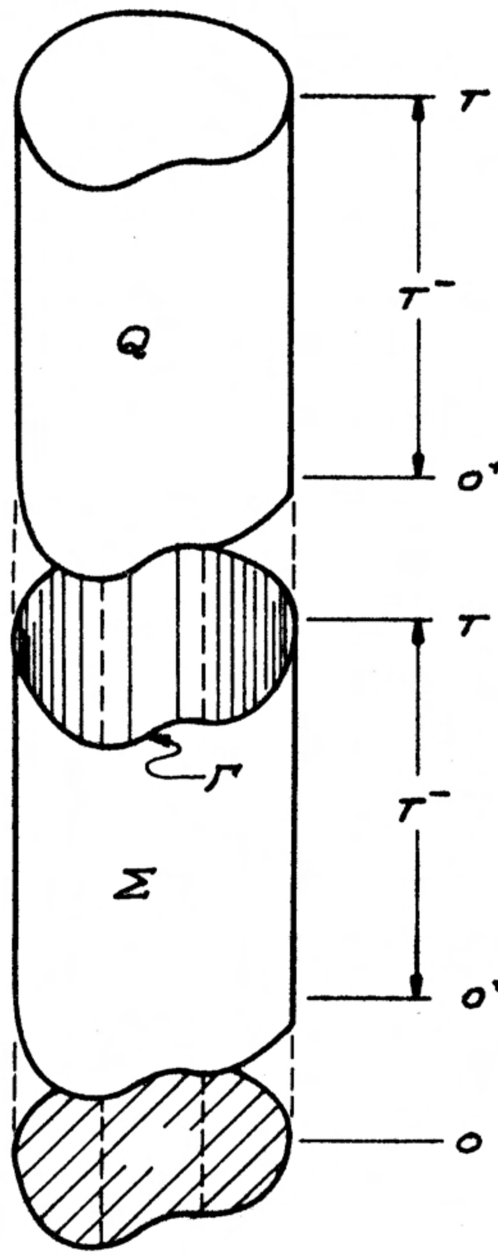

FIGURE II.1

"Exploded" Schematic of a Special Domain

Inner Product: $(\underline{f},\underline{g})_V=\int_\Sigma f_\Sigma(s,t)g_\Sigma(s,t)d\Sigma+\int_\Omega f_\Omega(x)g_\Omega(x)dx$

Norm: $||f||_V = [\int_\Sigma f_\Sigma(s,t)^2 d\Sigma+\int_\Omega f_\Omega^2(x)dx]^{1/2}$

Let $H^k(\Omega)$, k=1,2..., be the space of real functions whose squared powers and squared powers of the partial derivatives up to order k are integrable for the measure $dx=(dx_1dx_2...dx_r)$. Let $D^j_{x_i}f(x)$ denote the partial derivatives of order j, (j=1,2...k) with respect to all j-tuples of the r coordinate arguments, each evaluated at a point $x\varepsilon\Omega$. Define the inner product and norm of the two elements f, $g\varepsilon H^k$:

Inner Product: $(f,g)_{H^k(\Omega)}=\int_\Omega[f(x)g(x)+\sum_{j=1}^{k} D^j_{x_i}f(x)D^j_{x_i}g(x)]dx$

Norm: $||f||_{H^k(\Omega)}=\{\int_\Omega[f(x)^2+\sum_{j=1}^{k} D^j_{x_i}f(x)^2]dx\}^{1/2}$

Let $H^1_0(\Omega) = \{f : f\varepsilon H^1(\Omega)$ and $f = 0$ on $\Gamma\}$.

Finally, we shall be concerned with the space $L^2(0,T;H)$ of functions which for any time $t\varepsilon(0,T]$ are elements of H, and whose second powers are integrable with respect to the measure dx dt. Define the norm and inner product of two elements f, $g\varepsilon L^2(0,T;H)$:

Inner Product: $(f,g)_{L^2(0,T;H)}=\int_0^T(f,g)_H\,dt$

Norm: $||f||_{L^2(0,T;H)}=\{\int_0^T||f||_H^2\,dt\}^{\frac{1}{2}}$

For clarity, we adopt the following notational conventions regarding f:

$f(x,t)$	is a point in R^1, where x, $t \epsilon Q$
$f(\cdot,t)$	is an element of the Hilbert space $H(\Omega)$
$f(\cdot,\cdot)$	is an element of the Hilbert space $L^2(0,T;H)$
$f(\cdot,\cdot;\alpha)$	is an element of the Hilbert space $L^2(0,T;H)$ parameterized by α, taken from a collection of functions $\{f_\alpha\}$ each an element of $L^2(0,T;H)$.

As we indicated, the class of functions $f(\cdot)$ which we shall consider are elements of some Hilbert space H. In general, the functions need not be continuous. We therefore define the derivative of these functions in the following way:

<u>Definition II.1</u>

Given a test function $\phi(\cdot)\epsilon C^1(\Omega)$ with compact support in Ω, then for $f(\cdot)\epsilon L^2(\Omega)$, the mapping

$$\frac{\partial f}{\partial x_i} : \phi(\cdot) \rightarrow -\int_\Omega f \frac{\partial \phi}{\partial x_i} dx; \; i=1,2...r \qquad (2.2)$$

is called the distribution derivative or secant of the function f.

<u>Remark</u>

The definition of $\frac{\partial f}{\partial x_i}$ is made in terms of a "supporting" function ϕ, as $\frac{\partial f}{\partial x_i}$ is a "weak" function, called a distribution or generalized function. We arrive at (2.2) in the following way: Integrate, by parts, the product

$$\frac{\partial f(x)}{\partial x_i}\phi(x)$$

over Ω to obtain, via Green's Theorem:

$$\int_\Omega \frac{\partial f(x)}{\partial x_i}\phi(x)dx = - \int_\Omega f(x)\frac{\partial \phi(x)}{\partial x_i}dx$$

For an example of the determination of a distribution derivative, see [21].

It is convenient to express any function $f(\cdot)\varepsilon H$ in terms of a countable, everywhere dense set of (known) "elementary" functions belonging to H. Since we consider only separable Hilbert spaces, the existence of these functions is axiomatic. In particular, we consider as elementary functions the complete orthonormal system $\{w_i(x)\}_{i=1,2...}$ and state the following classical theorem:

<u>Theroem II.1</u>

In $L^2(\Omega)$ space, every <u>complete</u> orthonormal system is <u>closed</u>, and conversely.

Thus, for any $f\varepsilon L^2(\Omega)$,

$$\lim_{m\to\infty} ||f - \sum_{i=1}^{m} c_i w_i||^2_{L^2(\Omega)} \to 0$$

where

$$c_i = \int_\Omega f(x)w_i(x)dx.$$

A special system of complete orthonormal functions which is used in the sequel are those generated by solutions to the classical Sturm-Liouville equation:

$$A[w] - \lambda\rho w = 0 \quad \text{in } \Omega \tag{2.3}$$

with any one of the three boundary conditions

$$\begin{aligned} &(1) \quad w(\Gamma) = 0 \\ &(2) \quad \frac{\partial w(\Gamma)}{\partial \nu_A} = 0 \\ &(3) \quad \frac{\partial w(\Gamma)}{\partial \nu_A} + \gamma(\Gamma)w(\Gamma) = 0 \; ; \; \sigma(\Gamma) > 0 \end{aligned} \tag{2.4}$$

with the hypothesis

$$A[w] = -\sum_{i=j=1}^{r} \frac{\partial}{\partial x_i}[a_{ij}(x)\frac{\partial w_j}{\partial x_j}] + a_0(x)w(x)$$

$a_0(x)$ and $a_{ij}(x)$ are bounded, measureable

$a_{ij}(x) = a_{ji}(x)$ for all $x \varepsilon \Omega$

$a_0(x) \geq \alpha > 0$ almost everywhere in $L^2(\Omega)$

$\sum_{i,j=1}^{r} a_{ij}(x)\xi_i\xi_j \geq \alpha\ (\xi_1^2+\xi_2^2+\ldots+\xi_r^2)$ almost everywhere in Ω for all $\underline{\xi} = (\xi_1\xi_2\ldots\xi_r)$ in R^r

$f(x) \geq 0$ almost everywhere in Ω

In summary, we give the classical theorem:[1]

Theorem II.2

Solutions to (2.3) and any one of (2.4) constitute a complete orthonormal system in $L^2(\Omega)$. These solutions $w_i(x)$, $x \in \Omega$, are called the eigenfunctions of the Sturm-Liouville problem or more simply, eigenfunctions.[2]

Of special interest are the eigenfunctions of

$$\frac{\partial^2 w}{\partial x^2} + \lambda w = 0 \qquad \text{on } (0,1)$$

$$w(0) = 0$$

$$w(1) = 0$$

namely,

$$\{\sqrt{2}\,\sin\sqrt{\lambda_i}\,x\}_{i=1,2\ldots}\ ;\ \lambda_i = (i\pi)^2$$

A.2 Solution Properties of Linear Parabolic and Second Order Hyperbolic Partial Differential Equations

We give the following Lemmas which, under appropriate hypothesis (given), assert the existence and uniqueness of solutions to two classes of linear partial differential equations. These equations,

---

1 See for example, reference [5].

2 The term Eigenfunction is derived from the German,"Eigenfunktion" meaning characteristic function, the term "characteristic" expressing the complete character of the set $\{w_i\}$.

along with pertinent hypothesis are introduced formally here. A more natural introduction to these equations is given in Chapter III.

A.2.1 Linear Parabolic Equations

$$\frac{\partial y}{\partial t}(x,t;\underline{u})+A[y(x,t;\underline{u})] = f(x,t) \qquad x,t\varepsilon Q \tag{2.5}$$

Initial condition:

$$y(x,0) = u_2(x) \qquad x\varepsilon\Omega \tag{2.6}$$

Boundary condition - one of:

(I) $$y(s,t) = u_1(s,t) \qquad s,t\varepsilon\Sigma \tag{2.7}$$

(II) $$\frac{\partial y}{\partial \nu_A}(s,t) = u_1(s,t) \qquad s,t\varepsilon\Sigma \tag{2.8}$$

(III) $$\frac{\partial y}{\partial \nu_A}(s,t)+\beta(s,t)y(s,t) = u_1(s,t) \qquad s,t\varepsilon\Sigma \tag{2.9}$$

Remark

It is possible to consider boundary conditions which are combinations of (I), (II) and (III) in the following sense: Γ is partitioned into $\Gamma_1\Gamma_2\ldots\Gamma_k$ such that $\bigcup_k \Gamma_k = \Gamma$. Then, on each of the sets $\Sigma_i = \Gamma_i\times(0,T]$, one of the conditions (I), (II) or (III) holds. Such considerations introduce no new technical problems and are not considered.

Hypothesis

(i) $A[\phi(x,t)]=-\sum_{i=j=1}^{r} \frac{\partial}{\partial x_i}[a_{ij}(x,t)\frac{\partial}{\partial x_j}\phi(x,t)]+a_0(x,t)\phi(x,t)$

$a_0(x,t) \geq \alpha > 0$ almost everywhere in Ω for all $t\varepsilon(0,T]$.

$$\sum_{i=j=1}^{r} a_{ij}(x,t)\xi_i\xi_j \geq \alpha(\xi_1^2+\xi_2^2+\ldots+\xi_r^2)$$

almost everywhere in Ω for all $t\varepsilon(0,T]$.

$a_0(x,t)$, $a_{ij}(x,t)$ bounded, measureable $(i,j=1,2\ldots r)$

$\beta(s,t) \geq 0$ for all $s,t\varepsilon\Sigma$

(ii) $f(\cdot,\cdot)$ $L^2(Q)$

(iii) $u_2(\cdot)\varepsilon L^2(\Omega)$

(iv) $u_1(\cdot)\varepsilon L^2(\Sigma)$

Nomenclature

We shall refer to equations (2.5), (2.6) and (2.7) as system (I), (2.5), (2.6) and (2.8) as system (II) and (2.5), (2.6) and (2.9) as system (III).

Notation

$$\frac{\partial y}{\partial \nu_A} = \sum_{i=j=1}^{r} a_{ij}(s,t)\frac{\partial y}{\partial x_j} \cos(n,x_i)_{L^2(R^r)}$$

n = vector outward normal to the curve Γ.

Lemma II.1

If a solution exists for system (I), system (II) or system (III), such that $y(\cdot,\cdot)\varepsilon L^2(0,T;H^1)$, then in each case, that solution is unique.

A proof is given in [21].

Lemma II.2 (Lions-Magenes)

Under the given hypothesis, a solution to system (I) exists in $L^2(Q)$ and is unique. Note,

$$\frac{\partial \underline{y}(\cdot,t;\underline{u})}{\partial x_i} \notin L^2(\Omega), \quad (i=1,2\ldots r)$$

Lemma II.3

Under the given hypothesis, a solution to system (II) exists such that

$$\int_0^T ||y(t)||^2_{H^1(\Omega)} dt < \infty$$

Lemma II.4

Under the given hypothesis, a solution to system (III) exists such that

$$\int_0^T ||y(t)||^2_{H^1(\Omega)} dt < \infty$$

Remark

The proof of Lemmas II.2, II.3 and II.4 are to appear in a book authored by Lions and Magenes, Dunod (Paris) publishers, 1968.

Thus we have the existence and uniqueness of solutions to systems (I), (II) and (III) (of appropriate class).

A.2.2 Linear, Second Order Hyperbolic Equations

$$\frac{\partial^2 y(x,t;\underline{u})}{\partial t^2} + A[y(x,t;\underline{u})] = f(x,t) \qquad x,t\varepsilon Q \qquad (2.11)$$

Initial conditions:

$$y(x,0) = u_2(x) \qquad x\varepsilon\Omega \qquad (2.12)$$

$$\frac{dy(x,0)}{dt} = u_3(x) \qquad x\varepsilon\Omega \qquad (2.13)$$

Boundary conditions - one of:

(I) $y(s,t) = u_1(s,t)$ $\qquad s,t\varepsilon\Sigma \qquad (2.14)$

(II) $\dfrac{\partial y(s,t)}{\partial \nu_A} = u_1(s,t)$ $\qquad s,t\varepsilon\Sigma \qquad (2.14)$

(III) $\dfrac{\partial y(s,t)}{\partial \nu_A} + (s,t)y(s,t) = u_1(s,t)$ $\qquad s,t\varepsilon\Sigma \qquad (2.15)$

Hypothesis

We augment the hypothesis of section A.2.1 with the following:

$a_{ij}(x,t) = a_{ji}(x,t) \quad (i,j=1,2...r)$

for all $x,t\varepsilon Q$.

$a_{ij}(x,t)$ is $C^1(0,T)$

Nomenclature

Equations (2.11), (2.12) and (2.13) with the boundary conditions (I), (II) or (III) is called system (I), (II) and (III) respectively.

For systems (I), (II) and (III) we state the following three Lemmas, analogous to Lemmas II.2, II.3 and II.4 of section A.2.1:

Lemma II.5

Under the augmented hypothesis, a solution to system (I) exists in $L^2(0,T;H)$ and is unique.

Lemma II.6

Under the augmented hypothesis, a solution to system (II) exists in $L^2(0,T;H^1)$ and in unique.

Lemma II.7

Under the augmented hypothesis, a solution to system (III) exists in $L^2(0,T;H^1)$ and is unique.

Thus the solutions to each of systems (I), (II) and (III) of sections A.2.1 and A.2.2 exist (in the appropriate spaces) and moreover, are unique. Another property of these systems is that they are "well set" in the sense of Hadamard, a feature which is summarized in the following Lemma:

Lemma II.8

The solutions of each of the systems (I), (II) and (III) of section A.2.1 and A.2.2 vary continuously with the initial and boundary data. That is, for any $\varepsilon > 0$, there exists a δ such that

$$||\underline{u}-\underline{v}||_V < \delta \quad ; \quad u,v\varepsilon V \tag{2.16}$$

implies that

$$||y(x,t;u)-y(x,t;\underline{v})||_{L^2(Q)} < \varepsilon \tag{2.17}$$

An independent proof is given in [21].

Lemma II.8 establishes the stability of the solutions to systems (I), (II) and (III) of sections A.2.1 and A.2.2 (stability in the sense that "small" excursions from some nominal boundary and initial data give rise to "small" trajectory variations of the response).

Attention is now directed towards a technique which uses the result of Theorem II.1 to obtain approximate solutions to the systems (I), (II) and (III).

A.3 The Galerkin Approximation Scheme

As we indicated in section A.1, it is frequently convenient to describe a function from a given class in terms of a set of elementary functions which form a basis in that class. Here, in the context of solving a given type of partial differential equation, we are concerned with approximating the solution of that equation, which is known to belong to a certain class of functions.

We are given, for example, that $y(\cdot,t;\underline{u})$ is the solution of a certain differential equation. Moreover, $y(\cdot,t;\underline{u})$ H, some Hilbert space. If $\{w_i\}_{i=1,2...}$ are a basis in H, then we seek approximations to $y(\cdot,t;\underline{u})$ of the form

$$y_m(x,t;\underline{u}) = \sum_{i=1}^{m} y_i(t;\underline{u})w_i(x) \tag{2.18}$$

We shall consider the cases where $\{w_i\}_{i=1,2...}$ are

(A) The Normalized Eigenfunctions in H, and

(B) The cubic splines, for Ω a segment of R^1.

Remark

(i) The problem of obtaining the eigenfunctions to the Sturm-Liouville equation (section A.1) when $\Omega \cup \Gamma$ is "irregular", that is, when $\Omega \cup \Gamma$ is not a simple geometric figure, is complex. We accept this restriction, and proceed under its shadow.

(ii) No approximations arising from the common "differencing" methods are attempted, except in the sense that (B) is sophisticated differencing technique.

To illustrate the Galerkin method, consider its application to system (I) of section A.2.1.

In preparation for the application of this approximation technique to a system with non-homogeneous boundary conditions, it is customary[1] to first employ an affine transformation of the dependent variables so as to convert the original system into two systems as follows:

Define

$$y(x,t) = u(x,t) + w(x,t) \tag{2.19}$$

---

1 See for example reference [23].

According to (2.5),

$$\frac{\partial u(x,t)}{\partial t}+A[u(x,t)] = f(x,t)-\{\frac{\partial w(x,t)}{\partial t}+A[w(x,t)]\} \qquad (2.20)$$

As usual, a function w(x,t) is constructed, satisfying

$$w(s,t) = u_1(s,t) \qquad s,t\varepsilon\Sigma$$

Now we observe that the hypothesis on $u(\Sigma)$ is that $u(\cdot)\varepsilon L^2(\Sigma)$. Evidently, difficulties may arise when the necessary derivatives are taken on the function w(x,t), $x,t\varepsilon Q$, since the RHS of (2.20) must be $L^2(Q)$. These difficulties are skirted by proceeding in a novel way, proposed by Lions [16].

Define

Ψ is a function with the following properties:

$$\frac{\partial\Psi(x,t)}{\partial t}+A[\Psi(x,t)]=\rho(x,t);\rho(\cdot,\cdot)\varepsilon L^2(Q) \qquad (2.21)$$

$$\Psi(s,t) = 0 \qquad (2.22)$$

$$\Psi(x,T) = 0 \qquad (2.23)$$

Multiply (2.5) by $\Psi(x,t)$ and integrate over Q to obtain:

$$\int_Q y(x,t)\rho(x,t)dx\ dt-\int_\Omega u_2(x)\Psi(x,0)dx$$

$$+\int_\Sigma u_1(\Sigma)\frac{\partial\Psi(\Sigma)}{\partial\nu_A}d\Sigma = \int_Q f(x,t)\Psi(x,t)dx\ dt \qquad (2.24)$$

Equation (2.24) is a linear integral equation of the first type and is equivalent to the system (I) of section A.2.1.

The application of the Galerkin technique to (2.24) is straightforward provided $\Psi(x,t)$ is defined.

Remark

The following developments are valid for type (A) approximations. Since type (B) approximations are essentially different, the corresponding approximation technique is considered in [21].

Define

$$\Psi(x,t) = g(t)w_j(x), \text{ j fixed but arbitrary} \tag{2.25}$$

$$g(\cdot)\varepsilon C^1\ (0,T]\ ;\ g(T) = 0 \tag{2.26}$$

$$u_{1i}(t) = \int_\Gamma \frac{\partial w_i(\Gamma)}{\partial \nu_{A*}} u_1(\Sigma)d\Gamma \tag{2.27}$$

$$y_m(x,t) = \sum_{i=1}^{m} y_i(t)w_i(x) \tag{2.28}$$

Evidently,

$$\rho(x,t) = [-\frac{dg(t)}{dt}+\lambda_j g(t)]w_j(x)\ ;\ \rho(\cdot,\cdot)\varepsilon L^2(Q)$$

$$\Psi(s,t) = 0 \qquad s,t\varepsilon\Sigma$$

$$\Psi(x,T) = 0 \qquad x\varepsilon\Omega$$

$$\lim_{m\to\infty} y_m(\cdot,t) \to y(\cdot,t)\varepsilon L^2(\Omega), \text{ for each } t\varepsilon(0,T].$$

Using definitions (2.25), (2.26), (2.27) and (2.28) and the properties of $\{w_i\}_{i=1,2...}$ in (2.24), there results:

$$\int_0^T y_j(t)\{-\frac{dg}{dt} + \lambda_j g(t)\}dt - g(0)u_{2j} + \int_0^T g(t)u_{1j}(t)dt$$

$$= \int_0^T f_j(t)g(t)dt \qquad (2.29)$$

where we used

$$f(x,t) = \lim_{m\to\infty} \sum_{i=1}^{m} f_i(t)w_i(x); \; f_i(t) = \int_\Omega f(x,t)w_i(x)dx$$

$$u_2(x) = \lim_{m\to\infty} \sum_{i=1}^{m} u_{21}w_i(x) \; ; \; u_{21} = \int_\Omega u_2(x)w_i(x)dx$$

Integrating (2.29) by parts and using the properties of $g(t)$, we obtain the following equation for $y_i(t)$, $(i=1,2...m)$:

$$\begin{aligned} \frac{dy_j(t)}{dt} + \lambda_j(t)y_j(t) &= f_j(t) - u_{1j}(t) \\ y_j(0) &= u_{2j} \end{aligned} \qquad (j=1,2...m) \qquad (2.30)$$

We call (2.28) the Galerkin Approximation to system (I), with $y_j(t)$ defined by (2.30), $(j=1,2,3...m)$.

It is possible to obtain an estimate of the error of approximation. Indeed, if we define:

$$M = \sup_{i} \sup_{t\varepsilon(0,T]} \{|f_i(\cdot)-u_{1i}(\cdot)|\} \tag{2.31}$$

Then it is shown in [21].

$$E_m(y_m)=||y(\cdot,t)-\sum_{i=1}^{m} y_i(t)w_i(\cdot)||_{L^2(\Omega)} \leq [\frac{M^2}{3\pi 4}]^{1/2}(\frac{1}{m})^{3/2}$$

that is, the error is of the order $(\frac{1}{m})^{3/2}$.

Remark

The error estimate obtained for the Galerkin approximation using the eigenfunctions as a basis indicates that the convergence may be slow. (It is of course possible that there is no error after m terms). Furthermore, the convergence of the derivatives of y_m to those of y is slow. A more strict hypothesis on the function $y(\cdot,t)$ being approximated, for example, $y(\cdot,t)\varepsilon C^4(\Omega)$ allows the adoption of splines as a "basis" and there result some striking error estimates. The spline approximation is considered in [ ]. There, an important drawback is indicated, hindering the application of splines in the context of this study.

We consider next some fundamental results pertaining to the characterization of extremals of quadratic functionals on H.

A.4 Linear, Bilinear, and Quadratic Functionals on a Hilbert Space

Let u, v and w be elements of the Hilbert space V defined in section A.1. Define the following:

(i) $l(\underline{v})$ is a continuous linear functional on V. That is,

$$|l(\underline{v})| \leq M||\underline{v}||_V \; ; \; M < \infty$$

for $\alpha, \beta \varepsilon R^1$,

$$l(\alpha u + \beta v) = \alpha l(\underline{u}) + \beta l(\underline{v})$$

(ii) $a(\underline{u},\underline{v})$ is a continuous bilinear functional on V. That is,

$$|a(\underline{u},\underline{v})| \leq N||\underline{u}||_V \; ||\underline{v}||_V \; ; \; N < \infty$$

for $\alpha, \beta \varepsilon R^1$,

$$a(\alpha \underline{u} + \beta \underline{v},w) = \alpha a(\underline{u},\underline{w}) + \beta a(\underline{v},\underline{w})$$

and

$$a(\underline{u}, \alpha \underline{v} + \beta \underline{w}) = \alpha a(\underline{u},\underline{v}) + \beta a(\underline{u},\underline{w})$$

(iii) K is a closed convex set in V.

(iv) $J(\underline{v})$ is a quadratic functional on V with the property

$$J(\underline{v}) = a(\underline{v},\underline{v}) - 2l(\underline{v})$$

Lemma II.9

$J(\underline{v})$ is a convex functional on V: for $\underline{u}$, $\underline{v} \varepsilon V$ and $0 \leq \lambda \leq 1$,

$$J(\lambda \underline{u} + (1-\lambda)\underline{v}) \leq \lambda J(u) + (1-\lambda) J(\underline{v})$$

The proof of Lemma II.9 is well known and is given in [21].

We give a theorem which asserts the existence and uniqueness of an element $\underline{u}\varepsilon K$ with the property:

$$J(\underline{u}) = \underset{\underline{v}\varepsilon V}{\mathrm{Inf}}\, J(\underline{v})$$

Theorem II.3

There exists a $\underline{u}\varepsilon K$, unique such that

$$J(\underline{u}) = \underset{\underline{v}\varepsilon K}{\mathrm{Inf}}\, J(\underline{v}) \ ; \ J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v}\varepsilon K.$$

Proof

See Appendix II.1.

The following two theorems which characterize the $\underline{u}\varepsilon K$ for which

$$J(\underline{u}) = \underset{\underline{v}\varepsilon K}{\mathrm{Inf}}\, J(\underline{v})$$

are due to Lions and Stampaccia [17]. The proofs are included here because of the importance of these theorems.

Theorem II.4

The unique element $\underline{u}\varepsilon K$ for which

$$J(\underline{u}) = \underset{\underline{v}\varepsilon K}{\mathrm{Inf}}\, J(\underline{v})$$

is characterized by:

$$a(\underline{u},\underline{v}-\underline{u}) \geq l(\underline{v}-\underline{u}) \text{ for all } \underline{v}\varepsilon K \tag{2.32}$$

Proof (i) Necessity

$$J(\underline{u}) \leq J(\underline{w}) \quad \text{for all } \underline{w}\varepsilon K \tag{2.33}$$

let $\underline{w} = (1-\lambda)\underline{u} + \lambda\underline{v} \qquad 0 \leq \lambda \leq 1$

Then

$$J(\underline{u}) \leq J((1-\lambda)\ \underline{u} + \lambda\underline{v}) \tag{2.34}$$

We have that

$$J((1-\lambda)\underline{u}+\lambda\underline{v})=J(\underline{u}+\lambda(\underline{v}-\underline{u}))=J(\underline{u})+2a(\underline{u},\underline{v}-\underline{u})+\lambda^2 a(\underline{v}-\underline{u},\underline{v}-\underline{u})-2l(\underline{v}-\underline{u})$$

Taking note of (2.34), then

$$2a(\underline{u},\underline{v}-\underline{u})-2l(\underline{v}-\underline{u})+\lambda a(\underline{v}-\underline{u},\underline{v}-\underline{u}) \geq 0 \qquad 0 \leq \lambda \leq 1 \tag{2.35}$$

Set $\lambda = 0$ in (2.34) and obtain the desired result given by (2.32).

(ii) Sufficiency

$$J(\underline{v})-J(\underline{u}) = a(\underline{v},\underline{v})-a(\underline{u},\underline{u})-2l(\underline{v}-\underline{u})$$

Use the result that

$$a(\underline{u},\underline{v}-\underline{u}) \geq l(\underline{v}-\underline{u})$$

and obtain

$$J(\underline{v})-J(\underline{u}) \geq a(\underline{v},\underline{v})-a(\underline{u},\underline{u})-2a(\underline{u},\underline{v}-\underline{u}) = a(\underline{v}-\underline{u},\underline{v}-\underline{u})$$

$$\geq \alpha||\underline{v}-\underline{u}||_V^2 \geq 0 \qquad \underline{v}\varepsilon K$$

Theorem II.5

For $K = V$, then theorem II.4 implies that the unique $\underline{u} \varepsilon V$ for which $J(\underline{u}) = \underset{\underline{v}\varepsilon V}{\mathrm{Inf}}\, J(\underline{v})$ is characterized by

$$a(\underline{u},\underline{v}) - l(\underline{v}) = 0 \qquad (2.36)$$

Proof

Let $\underline{v} = \underline{u} \pm \underline{e}$ $\qquad \underline{e},\ \underline{v}\varepsilon V$.

Put this choice of $\underline{v}$ into (2.32) and obtain

$$\pm a(\underline{u},\underline{e}) \geq \pm l(\underline{e})$$

$$\Longrightarrow$$

$$a(\underline{u},\underline{e}) = l(\underline{e}) \quad \text{for all } \underline{e}\varepsilon V.$$

Theorems II.4 and II.5 provide the theoretical basis for a variational approach to the identification problem to be introduced in Chapter III and developed in Chapters IV and V. We remark that theorems II.4 and II.5 are a maximum (or more properly, a minimum), principle which is at once more general and less general than the celebrated maximum principle of Pontryagin [21]. It is more general in the sense that, as we will show, it characterizes extremals to systems whose evolution processes are distributed. It is less general because of the hypothesis on $J(\underline{v})$ and also because it is applicable to linear evolution processes only.

It is possible to obtain other characterizations of the optimal $\underline{u}$ for other special types of closed and convex K, as for example when

K is a cone with vertex at the origin. However the identification problems posed in Chapter III do not need such hypothesis.

It should be remarked that theorems II.4 and II.5 have wider applicability than that which is given in this study. Indeed, Lions has used these results to obtain the first rigorous treatment of a class of distributed optimal control problems [15]. Applications are also immediate in the treatment of some classical problems in the mathematical physics, such as the "maximum principle" for elliptic partial differential equations.

CHAPTER III

FORMAL DEFINITION OF THE STATE IDENTIFICATION PROBLEM

A.0 Introduction

The state identification problem was informally introduced in Chapter I. Here, in section A.1, the systems whose states are to be identified are presented along with a precise statement of the criteria of identification (section A.2). Section A.3 deals with a special relation which exists between two identification criteria.

A.1 Definition of the Class of Systems Considered

The abstract system from which an input-output record has been obtained is assumed to have associated with it the state determined character. That is, the output at some time $t > t_0$ is uniquely determined by knowledge of the input record over $t_0 \leq \tau \leq t$ and the evolution of the <u>state</u> of the system, starting from a known value at $t = t_0$. The state evolution process is described mathematically by a partial differential equation. Conventionally, the system is classified by the type of equation satisfied by the associated state evolution process. We shall investigate the state identification of linear parabolic and second order hyperbolic systems.

The domain of dependence of these partial differential equations is customarily space and time. At any time $t > t_0$, $y(\cdot,t)$, the system state, is therefore a function of the space variable(s). Generally, an infinite number of elementary or basis functions would be necessary to describe the state at time $t > t_0$. Thus the state space for the systems under consideration are "infinite dimensional". We assume without much loss of generality that the state equations are scalar.

A.1.1 Systems Classification

The classification of parabolic or (second order) hyperbolic equations arises when the following general linear partial differential equation with real coefficients is transformed to a cannonical form at some point $\underline{x}$ in the domain of the process:

$$\sum_{j=1}^{r+1} \sum_{i=1}^{r+1} a_{ij}(\underline{x}) \frac{\partial^2 u(x)}{\partial x_i \partial x_j} + \sum_{i=1}^{r+1} b_i(\underline{x}) \frac{\partial u(\underline{x})}{\partial x_i} + c(\underline{x})u(\underline{x}) + f(\underline{x}) = 0 \quad (3.1)$$

Equation (3.1) is reducible to one of the following three cannonical forms[1]:

$$\frac{\partial^2 u(\underline{x})}{\partial x_1^2} + \frac{\partial^2 u(\underline{x})}{\partial x_2^2} + \ldots + \frac{\partial^2 u(\underline{x})}{\partial x_{r+1}^2} + g(\underline{x}) = 0 \qquad \text{Elliptical Type} \quad (3.2)$$

$$\frac{\partial^2 u(\underline{x})}{\partial x_1^2} = \sum_{i=2}^{r+1} \frac{\partial^2 u}{\partial x_i^2} + g(\underline{x}) \qquad \text{Hyperbolic Type} \quad (3.3)$$

$$\sum_{i=1}^{r+1-m} (\pm u_{x_i x_i}) + g(\underline{x}) = 0 \qquad 0 < m < r+1 \qquad \text{Parabolic Type} \quad (3.4)$$

---

1 We make no assumptions here on the coefficients $a_{ij}(\underline{x})$, $b_i(x)$ and $c(x)$ nor on the differentiability of $u(\underline{x})$.

For details of the method of classification, see (for example) [23]. The specific types of Linear Parabolic and second order Hyperbolic systems to be identified is now made more precise.

A.1.2 Linear Parabolic Systems

The system state $y(x,t)$, $x \varepsilon \Omega$; $t \varepsilon (0,T]$ evolves according to

$$\frac{\partial y}{\partial t}(x,t) + A[y(x,t)] = f(x,t) \text{ in } \Omega \tag{3.5}$$

with boundary data

$$u_1(\Sigma) \tag{3.6}$$

and initial data

$$u_2(x) \tag{3.7}$$

given in a way to be defined presently. At any time $t \varepsilon (0,T]$, the spatial operator $A[\cdot]$ has the representation:

$$A[\Psi(\underline{x})] = - \sum_{i=1}^{r} \sum_{j=1}^{r} \frac{\partial}{\partial x_i} [a(\underline{x},t) \frac{\partial}{\partial x_j} \Psi(\underline{x})] + a_0(\underline{x},t)\Psi(\underline{x}) \quad (3.9)$$

Moreover the coefficients $a_{ij}(x,t)$, $(i,j=1,2...r)$ and $a_0(x,t)$ satisfy an ellipticity condition: Given $\Psi(\cdot)\varepsilon H^1(\Omega)$, for $t\varepsilon(0,T]$

$$\sum_{i=1}^{r} \sum_{j=1}^{r} (a_{ij}(\cdot,t)\Psi_i(\cdot),\Psi_j(\cdot))_{H^1(\Omega)} \geq \alpha \sum_{i=1}^{r} (\Psi_i(\cdot),\Psi_i(\cdot))_{H^1(\Omega)} \quad (3.10)$$

$$\alpha \varepsilon R^1, \ \alpha > 0; \ a_0(\underline{x},t) > 0, \ \underline{x}\varepsilon\Omega.$$

$a_{ij}(\underline{x},t)$, $a_0(\underline{x},t)$ are bounded, measureable.

$$f(\cdot,\cdot) \ \varepsilon \ L^2(Q)$$
$$u_1(\cdot) \ \varepsilon \ L^2(\Sigma)$$
$$u_2(\cdot) \ \varepsilon \ L^2(\Omega)$$

The specification of the boundary data may be made in any one of the three following ways:

(I) $y(\Sigma) = u_1(\Sigma)$ (Dirichlet)

(II) $\frac{\partial \underline{y}(\Sigma)}{\partial \nu_A} = u_1(\Sigma)$ (Neumann)

$$\text{(III)} \quad \frac{\partial y}{\partial \nu_A} + \beta(\Sigma)y(\Sigma) = u_1(\Sigma) \qquad \text{(Mixed)}$$

where:

$$\frac{\partial y(\Sigma)}{\partial \nu_A} = \sum_{i,j=1}^{r} a_{ij}(\Sigma)\frac{\partial y(\Sigma)}{\partial x_j}[\cos(n,x_i)]$$

$$\frac{\partial y(\Sigma)}{\partial \nu_{A*}} = \sum_{i,j=1}^{r} a_{ji}(\Sigma)\frac{\partial y(\Sigma)}{\partial x_j}[\cos(n,x_i)]$$

n = outward normal to the boundary Γ.

(n,x_i) = angle between the i^{th} coordinate and the outward normal.

$\beta(\Sigma) > 0$

Theorem 3.1

The system (3.5), (3.6), (I), (II) or (III), (3.7) is well posed in the sense of Hadamard[1]: that is, the solution $y(x,t)$ exists (in some appropriate space which is defined, and the solution $y(x,t)$ varies continuously with $u_1(\Sigma)$ and $u_2(x)$).

Remark: (i) If the differential equation represents a physical system, the well posed property had better be present: evidently a solution "exists". Furthermore, the continuous variation of the state with respect to the initial and boundary data is a statement of the stability of the solution.

1 See Lemma II.8, Chapter II

A.1.3 Linear Hyperbolic Systems

The system state $y(x,t)$ evolves according to

$$\frac{\partial^2 y(x,t)}{\partial t^2} + A[y(x,t)] = f(x,t) \quad \text{in } \Omega \tag{3.11}$$

with boundary data:

$$u_1(\Sigma) \tag{3.12}$$

and initial data:

$$u_2(x) \tag{3.13}$$

In addition to properties (3.8) to (3.10), $A[y]$ satisfies:

$$a_{ij}(x,t) = a_{ji}(x,t) \tag{3.14}$$

As before,

$$f(\cdot,\cdot) \in L^2(Q)$$
$$u_1(\cdot) \in L^2(\Sigma)$$
$$u_2(\cdot),\ u_3(\cdot) \in L^2(\Omega)$$

The boundary conditions are any of (I), (II) or (III). The initial data is

$$y(x,0) = u_2(x);\ \frac{dy(x,0)}{dt} = u_3(x) \tag{3.15}$$

Theorem 3.2

The system (3.11), (3.12) (i), (II), (III), (3.13), (3.15) is well posed in the sense of Hadamard.[1]

---

1 A precise statement of this theorem is given in Chapter II, Lemma II.8

A.1.4 Examples of Parabolic and Hyperbolic Systems

The list of physical systems which may be categorized as either parabolic or hyperbolic is long.[1] Consider the following salient examples:

(a) Parabolic: Temperature $y(x,t)$ at a point $x\varepsilon\Omega$ and $t\varepsilon(0,T]$ of a medium which is exchanging heat in a predominantly diffusive manner with its environment, which at that point in space and that time, is at a temperature $f(x,t)$:

$$\frac{\partial y}{\partial t}(x,t) - k\,\frac{\partial^2 y}{\partial x^2}(x,t) = f(x,t) \qquad (3.16)^{(2)}$$

The boundary conditions of the system represented by (3.16) are exhausted by three considerations:

(i) No insulation

(ii) Full insulation

(iii) Partial insulation

These three physical orientations are mathematically represented by conditions (I), (II), (III) (respectively) of section A.1.2.

---

1 There are systems described by a set of vector equations which are hybrids: Parabolic in one dependent variable and hyperbolic in another. Since we consider the scalar differential equation exclusively, such systems are bypassed (though not because of any inherent difficulty).

2 The "transport" term $\frac{\partial y}{\partial x}$ occuring in the description of many heat exchange problems can be avoided by a linear transformation leading to the form given in (3.16).

(b) Hyperbolic: Vibrations in an elastic medium. Consider for example the case of longitudinal displacement y(x,t) at a point x in an elastic rod, at a certain time t, experiencing a dynamic axial load applied at the ends:

$$\frac{\partial^2 y(x,t)}{\partial t} - \frac{\rho}{EA}\,\frac{\partial^2 y(x,t)}{\partial x^2} = 0 \qquad (3.17)$$

The load is applied in one of three ways[1]

(i) Direct axial loading

(ii) Transverse bending

(iii) Loading through an elastic support.

Cases (i), (ii) and (iii) correspond to (I), (II), (III) of section A.1.2.

The examples given are classical. The question of identification of such systems may seem of academic importance until it is realized that more modern systems can be cast into the framework of the two illustrative classes. For example, a little reflection casts the problem of vibrations in a slender airframe into the perspective of example (b), appropriately modified. An exact definition of identification is now undertaken.

---

1 It is possible to consider combinations of (i), (ii) or (iii) on the boundary - for example, condition (i) at one end of the rod and (ii) at the other. In the more general setting of multidimensional spatial domains, such combinations are of rich variety.

All our results are valid for any such consistent combinations.

A.2 State Identification

The concept of state identification can be introduced by consideration of the following problem.

Suppose that a real physical system S has been evolving from some indeterminate time past and that the initial data with which the evolution began is known inexactly. Then for the distributed systems under study, the state at some point M in the space-time domain is parameterized by the initial data and the values of the forcing on the boundary and in the interior. Measurements I and O are allowed in the input and output spaces respectively of the system S. The question arises,[1] in what way can these measurements be utilized so that

(a) A "refinement" R of these inexact measurements is accomplished, consistent with

(b) An extension of the information contained in the results, accomplished via knowledge of the state evolution process.

The resolution of this question constitutes a solution to the state identification problem. The statement that a "refinement" of the measurement process is sought implies that a criteria of satisfaction is adopted. The criteria adopted throughout this report is that of "least squares", which is defined presently. The identification problem has the following,

---

1 There is considerable physical motivation to this question. Some examples giving rise to an identification problem were discussed in Chapter I.

A.2.1 Mathematical Description

System S The systems are as defined in section A.1.2 and A.1.3.

Input Measurements I The inexactly known inputs occur on the boundary of the system S:[1]

$$z_1(\cdot) = u_1(\cdot) + \varepsilon_1(\cdot) \qquad \text{on } \Sigma \qquad (3.18)$$

$$z_1(\cdot) \in L^2(\Sigma)$$

$$z_2(\cdot) = u_2(\cdot) + \varepsilon_2(\cdot) \qquad \text{in } \Omega^{(2)} \qquad (3.19)$$

$$z_2(\cdot) \in L^2(\Omega)$$

The input I can be considered as a vector

$$\underline{z} = \{z_1(\Sigma) \mid z_2(x)\}^T = \{u_1(\Sigma) \mid u_2(x)\}^T + \{\varepsilon_1(\Sigma) \mid \varepsilon_2(x)\}^T$$

$\underline{z} \in L^2(\Sigma) \times L^2(\Omega)$; $\underline{u} \in L^2(\Sigma) \times L^2(\Omega)$. That is,

$$\int_\Sigma z_1^2(\Sigma)d\Sigma + \int_\Omega z_2(x)^2 dx < \infty; \quad \int_\Sigma u_1(\Sigma)^2 d\Sigma + \int_\Omega u_2(x)^2 dx < \infty$$

Output Measurements O In general, the output measurements on the system state $y(x,t;\underline{u})$ are given by:

---

1 It is assumed that forcing of the states in the interior of the spatial domain if present, is known with precision.

2 The inexactly known initial condition is considered as an input to the system S.

$$z(x,t) = L(x,t)[y(\cdot,\cdot;\underline{u})] + \varepsilon(x,t) \quad x\varepsilon\Omega,\ t\varepsilon(0,T] \qquad (3.20)$$

where

$$L(x,t)[\cdot] : L^2(Q) \rightarrow L^2(Q) \qquad \text{Linear}$$

In particular, we consider the following three forms of (3.20) in detail:

$$\text{(I)} \quad L(x,t)[y(\cdot,\cdot;\underline{u})] = \int_\Omega y(\xi,t)\delta(\xi-x)d\xi = y(x,t) \qquad (3.21)$$

$$z(x,t) = y(x,t) + \varepsilon(x,t);\ z(\cdot,\cdot)\varepsilon L^2(Q)$$

$$\text{(II)} \quad L(x,t)[y(\cdot,\cdot;\underline{u})] = \int_\Omega \sum_{i=1}^{\nu} y(\xi,t)\delta(\xi-x^i)d\xi = \sum_{i=1}^{\nu} y(x^i,t) \qquad (3.22)$$

$$z(x^i,t) = y(x^i,t) + \varepsilon(x^i,t);\ z(x^i,\cdot)\varepsilon L^2((0,T]).$$

where x^i is a point in the spatial domain Ω.

$$\text{(III)} \quad L(x,t)[y(\cdot,\cdot;\underline{u})] = 0 \qquad (3.23)$$

$$z(x,t) = 0 \quad \text{everywhere in } Q.$$

In (3.18), (3.19) and (3.20), the error processes $\varepsilon_1(\Sigma)$, $\varepsilon_2(x)$ and $\varepsilon(x,t)$ are assumed to be unknown, (except with respect to the fact that they are square integrable). In section A.3, some knowledge of a statistical nature is assumed for these processes.

Data Refinement R

The dependence of the response $y(x,t)$ of the system S on the boundary and initial data $\underline{u}$ has already been mentioned. In order to display this dependence explicitly, write

$$y(x,t) = y(x,t;\underline{u})$$

Define

$\underline{v}$ = An arbitrary estimate of the "true state of nature", $\underline{u}^*$.

$\underline{u}$ = A refined estimate of the true state of nature $\underline{u}^*$ given measurements z, z_1 and z_2 and the equation of evolution of the system S.

The refinement R is chosen to be in the least squares sense: $\underline{u}$ minimizes the quadratic functional

$$J(v) = \int_Q [L[y(x,t;v)] - z(x,t)]^2 dx\, dt$$
$$+ \int_\Sigma [v_1(\Sigma) - u_1(\Sigma)]^2 d\Sigma$$
$$+ \int_\Omega [v_2(x) - u_2(x)]^2 dx \qquad (3.24)$$

In (3.24), y(x,t;v) evolves according to the definition of S. In order to stress some important properties of the refinement chosen, we generalize the statement of (3.24) in the following way

$$J(v) = a(v,v) - 2l(v) + k \qquad (3.25)$$

where a(v,v) and l(v) are endowed with the properties given in Chapter II, section A.4. These properties are restated here:

(i) $a(v,w)$ is a bilinear, continuous form on the Hilbert space $L^2(\Sigma)\times L^2(\Omega)$

(ii) There exists an $\alpha > 0$ such that

$$a(v,v) \geq ||\underline{v}||_{L^2(\Sigma)\times L^2(\Omega)}$$

(iii) $a(u,v) = a(v,u)$

(iv) $l(v)$ is a continuous linear form on $L^2(\Sigma)\times L^2(\Omega)$

It is essential that the functional $J(v)$ given by (3.24) have all the properties of the quadratic form (3.25).(1) Consequently, we now establish that for the system S and any of the measurement processes (I), (II) or (III), (3.24) can be phrased as in (3.25).

Theorem III.1

(3.24) and (3.25) are equivalent if $a(\underline{v},\underline{v})$ and $l(\underline{v})$ are defined by (3.26), (3.27) and (3.28). Moreover $a(\underline{v},\underline{v})$ and $l(\underline{v})$ obey the hypothesis of theorem II.5

$$a(\underline{v},\underline{v}) = \int_Q \{L[y(x,t;\underline{v})-y(x,t;\underline{0})]\}^2 dx\, dt$$

$$+ \int_\Sigma v_1(\Sigma)^2 d \;\; + \int_\Omega v_2(x)^2 dx \qquad (3.26)$$

---

1 The theorems given in Chapter II characterizing extremals of a quadratic functional will furnish a solution to the identification problem given. It is essential that the corresponding hypothesis be compatible.

$$l(\underline{v}) = -\int_Q L[y(x,t;\underline{v})-y(x,t;\underline{0})]\{L[y(x,t;\underline{0})]-z(x,t)]\}dx\,dt$$

$$-\int_\Sigma v_1(\Sigma)z_1(\Sigma)d \quad - \int_\Omega v_2(x)z_2(x)dx \tag{3.27}$$

$$k = \int_Q\{L[y(x,t;0)]-z(x,t)\}^2 dx\,dt$$

$$+ \int_\Sigma z_1(\Sigma)^2 d \quad + \int_\Omega z_2(x)^2 dx \tag{3.28}$$

Proof

Hypothesis on a(v,v)

(i) Evidently a(u,v) is a bilinear form, since S is a linear evolution process. In addition,

$$y(x,t;\underline{v})\varepsilon L^2(Q) \qquad \text{[Lemma II.2]}$$

Thus each term in (3.26) is bounded from above. Thus $a(\underline{v},\underline{v})$ is continuous in $\underline{v}$.

(ii) $a(\underline{v},\underline{v}) \geq \int_\Sigma v_1(\Sigma)^2 d\Sigma + \int_\Sigma v_2^2(x)dx = \alpha||\underline{v}||^2_{L^2(\Sigma)\times L^2(\Omega)}$
$\alpha = 1$.

(iii) The symmetry is self-evident.

(iv) Again, linearity of $l(\underline{v})$ is evident due to the linearity of the evolution process S. Boundedness (and hence continuity) is obtained by application of the triangle and Schwartz inequalities:

$$|l(\underline{v})| \leq |\int_Q L[y(x,t;\underline{v})-y(x,t;0)]\{L[y(x,t;0)]-z(x,t)\}dx\ dt|$$

$$+|\int_\Sigma v_1(\Sigma)z_1(\Sigma)d\Sigma|+|\int_\Omega v_2(x)z_2(x)dx|$$

$$\leq |M_0|\ ||L[y(x,t;\underline{v})-y(x,t;0)]||_{L^2(Q)}+|M|\ ||\underline{v}||_{L^2(\Sigma)\times L^2(\Omega)}$$

Since $y(x,t;\underline{v})\varepsilon L^2(Q)$, $\underline{v}\varepsilon L^2(\Sigma)\times L^2(\Omega)$ (M_0 and M are positive constants), then $|l(\underline{v})|$ is bounded and linear in $\underline{v}$, hence continuous in $\underline{v}$.

It is possible to consider many quadratic forms of the type (3.25). It is also possible to pose innocuous forms which resemble (3.25) superficially, but which violate the given hypothesis. Consider the following counter-example due to Lions [16].

System S_c

$$\frac{\partial y}{\partial t} + A[y] = f \text{ in } Q;\ Q = (0,1)\times(0,T]$$

$$y(0,t) = 0;\ y(1,t) = \frac{1}{(T-t)^\alpha};2\alpha < 1;u_1(\cdot)\varepsilon L^2(\Sigma)$$

$$y(x,0) = u_2(x) \qquad\qquad u_2(\cdot)\varepsilon L^2(\Omega)$$

A[y] defined by (3.8) and (3.9).

Assertion

For the system S_c,

$$y(\cdot,T;\underline{u}) \not\varepsilon L^2(\Omega)$$

Proof

It has been shown, (Lemma II.2, chapter II) that: $y(\cdot,\cdot)\varepsilon L^2(Q)$. Hence there exists a solution $y_m(x,t)$ defined by

$$y_m(x,t) = \sum_{i=1}^{m} y_i(t)w_i(x)$$

with the property that

$$\lim_{m\to\infty} y_m(x,t) \to y(x,t) \text{ for almost every } x,\ t\varepsilon Q.$$

where $\{w_i(\cdot)\}_{i=1,2,3...}$ are a basis in $L^2(\Omega)$. It can be shown that:

$$y_i(t) = -\frac{dw_i}{dx}(1) \int_0^t e^{-\lambda_i(t-\tau)}u(1,\tau)d\tau \tag{3.29}$$

$$|y_i(T)| = (\sqrt{2}\pi i) \int_0^T \frac{e^{-\lambda_i(T-\tau)}}{(T-\tau)^\alpha}\, d\tau \tag{3.30}$$

taking $\nu = \lambda_i(T-\tau)$, (3.30) becomes

$$|y_i(T)| = c \int_0^{\lambda_i T} [\frac{e^{-\nu}}{\lambda_i(1-\alpha)\nu^\alpha}]\, d\nu$$

Noting the classical result that $\lambda_i \geq \lambda_1;\ i \geq 1,$

$$|y_i(T)| \geq \frac{c_i}{\lambda_i(1-\alpha)} \int_0^{\lambda_1 T} \frac{e^{-\nu}}{\nu^{\alpha}} d\nu = c_1 \frac{i}{i^{2(1-\alpha)}}$$

$$\int_\Omega y_m(x,T)^2 d\Omega = \sum_{i=1}^{m} |y_i(T)|^2 \geq c_1^2 \frac{1}{i^{2(-4\alpha)}}$$

$$= \infty \quad \text{If } 2-4\alpha \leq 1.$$

If an identification problem is posed for the system S_c with the following identification criteria:

$$J(\underline{v}) = \int_\Omega |y(x,T;v)-z(x,T)|^2 dx + \int_\Sigma [v_1(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+\int_\Omega [v_2(x)-z_2(x)]^2 dx \tag{3.31}$$

Then it is possible that no solution exists which minimizes (3.31) as was illustrated by the counter-example. If we insist on constructing the error functional from measurements at the terminal time, $z(x,T)$, then $J(v)$ must be appropriately modified.

Lions has shown that for the system S_c,

$y(\cdot,T)\epsilon H^{-1}(\Omega)$ (Dual space of the Sobolev space H_0^1 defined in chapter II, section A.1).

Thus we choose

$$J(v)= ||y(\cdot,T;u)-z(\cdot,T)||^2_{H^{-1}(\Omega)}+||v||^2_{L^2(\Sigma)\times L^2(\Omega)} \qquad (3.32)$$

Where the norm in $H^{-1}(\Omega)$ is defined by the following operations:

$$f,\ g\ H^{-1}(\Omega),\quad \text{solve}$$

$$(-\frac{\partial^2}{\partial x^2}+1)\ \phi = f\ ;\ \phi = 0 \quad \text{on} \quad \Gamma$$

$$(-\frac{\partial^2}{\partial x^2}+1)\ \Psi = g\ ;\ \Psi = 0 \quad \text{on} \quad \Gamma$$

$$(f,g)_{H^{-1}(\Omega)} = \int_\Omega[\phi\Psi + \frac{\partial\Psi}{\partial x}\frac{\partial\phi}{\partial x}]dx$$

This problem under the refinement implied by the minimization of (3.32) is well posed and has a solution, considered in Chapter IV.

In summary, the state identification problem associated with the system S is that of minimizing the squared error arising from inaccurate measurements on S. This squared error is viewed as a quadratic functional of the boundary conditions $v_1(\Sigma)$ and initial data $v_2(x)$. The refinement of the incomplete and/or inexact measurements is accomplished by choosing v_1 and v_2 so that the quadratic functional is minimized, along a trajectory of the system S. The conditions characterizing these refined estimates u_1 and u_2 are given in chapter IV.

The term "refined estimate" as used here denotes a least squares fit of a trajectory of S to the given data I and O. Under certain hypothesis considered in [21], this least-squares refinement has a statistical significance. See [21].

CHAPTER IV

CHARACTERIZATION OF THE SOLUTION OF THE STATE IDENTIFICATION PROBLEM

A.0 Introduction

The state identification problem has been posed as a variational one, that of minimizing a quadratic functional along given system state trajectories, each parameterized by the initial and boundary data, denoted by $\underline{v}$. The indicated minimization is accomplished by the appropriate choice of an "admissible" $\underline{v}$. The purpose of this chapter is to indicate mathematically the "appropriate" choice of v, (denoted by $\underline{u}$), for a variety of systems, associated measurement processes and quadratic error functionals.[1] In particular, $\underline{u}$ will be characterized as the solution to a set of simultaneous equations.

Theorems establishing the existence and uniqueness of solutions to the various identification problems are given in each case.

The task of actually constructing a solution $\underline{u}$ from the characterization given by the simultaneous equations referred to in this introduction is the domain of Chapter V.

A.1 Parabolic Systems, $V = L^2(\Sigma)\times L^2(\Omega)$

This section contains the solution of the state identification problem for parabolic systems where the output measurements are:

---

1 The general result for extremals to special quadratic forms, given by Theorem II.5 must be examined in each of the example cases to ensure that the problems given obey the hypothesis of the general result.

(A)[1] $z(x,t)=y(x,t)+\varepsilon(x,t)$; $z(\cdot,\cdot)\varepsilon L^2(Q)$

(B)[2] $z(x^i,t)=y(x^i,t)+\varepsilon(x^i,t)$; $z(x^i,\cdot)\varepsilon L^2(0,T)$

(C)[3] $z(x,t) = 0$

Additionally, it is assumed that the admissible set of boundary and initial data, denoted by V is the entire space $L^2(\Sigma)\times L^2(\Omega)$.[4] In each problem induced by A, B and C, the measurement on the boundary and at the initial time are (respectively):

$$z(\Sigma) = u_1(\Sigma) + \varepsilon_1(\Sigma) \ ; \ z(\Sigma)\varepsilon L^2(\Sigma)$$

$$z(x) = u_2(x) + \varepsilon_2(x) \ ; \ z_2(x)\varepsilon L^2(\Omega)$$

We now consider the identification problem induced by (A), (B) and (C) for the parabolic system with Dirichlet boundary conditions. See [21] for results pertaining to the Neumann and mixed boundary condition cases.

---

1 It is assumed that $z(x,t)$ is measured everywhere in Q - A physically unrealistic situation.

2 Here, $z(x^i,t)$ correspond to pointwise measurements located at the spatial coordinates x^i, $i=1,2...\nu$.

3 No interior measurements taken.

4 Section A.3 considers the case where V is a closed convex subset of $L^2(\Sigma)\times L^2(\Omega)$.

PR.I: Dirichlet Boundary Conditions

System: S_D

$$\frac{\partial y}{\partial t}(x,t)+A[y(x,t)]=f(x,t) \qquad \text{in } Q \qquad (4.1)$$

$$y(\Sigma) \qquad =v_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.2)$$

$$y(x,0) \qquad =v_2(x) \qquad \text{in } \Omega \qquad (4.3)$$

Hypothesis on the system: H_D

$$f(x,t)\varepsilon L^2(Q):\int_0^T\int_\Omega|f(x,t)|^2 dx\ dt < \infty$$

$$\left.\begin{array}{l} v_1(\Sigma)\varepsilon L^2(\Sigma):\int_0^T \int_\Gamma|v_1(\Gamma,t)|^2 d\Gamma\ dt < \infty \\ v_2(x)\varepsilon L^2(\Omega):\int_\Omega|v_2(x)|^2 dx \qquad < \infty \end{array}\right\} = V$$

Problem PR.IA

Let $J(\underline{u}) = \operatorname*{Inf}_{\underline{v}\varepsilon V} J(v)$, where

$$J(\underline{v}) =\{\int_Q[y(x,t;\underline{v})-z(x,t)]^2 dx\ dt$$

$$+\int_\Sigma[v_1(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+\int_\Omega[v_2(x)-z_2(x)]^2 dx\} \qquad (4.4)$$

It is required to find a characterization of $\underline{u}$ which has the property that $J(\underline{u}) = \inf_{v \varepsilon V} J(\underline{v})$ for any $\underline{v}\varepsilon V$ along a trajectory $y(\cdot,\cdot,\underline{u})$ of the system S_D under the hypothesis H_D.

Problem PR.IB

Let $J(\underline{u}) = \inf_{\underline{v}\varepsilon V} J(\underline{v})$

$$\inf_{\underline{v}\varepsilon V} J(\underline{v}) = \sum_{i=1}^{\nu} \{\int_0^T [|y(x^i,t;\underline{v}) - z(x^i,t)|^2]dt$$

$$+\int_\Sigma [v_1(\Sigma) - z_1(\Sigma)]^2 d\Sigma$$

$$+\int_\Omega [v_2(x) - z_2(x)]^2 dx \tag{4.5}$$

It is required to find a characterization of $\underline{u}$ which has the property $J(\underline{u}) = \inf_{\underline{v}\varepsilon V} J(\underline{v})$ for any $\underline{v}\varepsilon V$ along trajectories $y(x^i,\cdot,\underline{u})$ of the system S_D under the hypothesis H_D.

Problem PR.IC

$$\inf_{\underline{v}\varepsilon V} J(v) = \int_\Sigma [v_1(\Sigma) - z_1(\Sigma)]^2 d\Sigma$$

$$+ \int_\Omega [v_2(x) - z_2(x)]^2 dx \tag{4.6}$$

The solution to each of the three problems PR.IA, PR.IB and PR.IC stated in conjunction with the state evolution process S_D is given by the three theorems, Theorem IV.1, IA, Theorem IV.1, IB and Theorem IV.1, IC which follows:

Theorem IV.1,IA

Given the system S_D, hypothesis H_D and the functional $J(\underline{v})$ of PR.IA. Then:

(i) There exists one and only one $u \epsilon V$ with the property that

$$J(u) \leq J(v) \quad \text{for all } v \epsilon V$$

(ii) The boundary and initial data $\underline{u}$ is uniquely characterized by the simultaneous solution of the following system of equations:

$$\frac{\partial \underline{y}(x,t;\underline{u})}{\partial t} + A[y(x,t;u)] = f(x,t) \quad \text{in } Q \qquad (4.7)$$

$$y(\Sigma) = u_1(\Sigma) \quad \text{on } \Sigma \qquad (4.8)$$

$$y(x,0) = u_2(x) \quad \text{in } \Omega \qquad (4.9)$$

$$-\frac{\partial \underline{p}(x,t;\underline{u})}{\partial t} + A[p(x,t;\underline{u})] = y(x,t;\underline{u}) - z(x,t) \quad \text{in } Q \qquad (4.10)$$

$$p(\Sigma) = 0 \quad \text{on } \Sigma \qquad (4.11)$$

$$p(x,T) = 0 \quad \text{in } \Omega \qquad (4.12)$$

$$-\frac{\partial \underline{p}(\Sigma)}{\partial \nu_{A^*}} + u_1(\Sigma) - z_1(\Sigma) = 0 \quad \text{on } \Sigma \qquad (4.13)$$

$$p(x,0) + u_2(x) - z_2(x) = 0 \quad \text{in } \Omega \qquad (4.14)$$

Theorem IV.1,IB

Given the system S_D, hypothesis H_D and the functional $J(\underline{v})$ of PR.IB. Then:

(i) There exists one and only one $\underline{u} \varepsilon V$ with the property that

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } v \varepsilon V$$

(ii) The boundary and initial data $\underline{u}$ is uniquely characterized by the simultaneous solution of the following system of equations:

$$\frac{\partial y}{\partial t}(x,t;\underline{u}) + A[y(x,t;\underline{u})] = f(x,t) \qquad \text{in } Q \qquad (4.15)$$

$$y(\Sigma) = u_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.16)$$

$$y(x,0) = u_2(x) \qquad \text{in } \Omega \qquad (4.17)$$

$$-\frac{\partial p}{\partial t}(x,t;u) + A[p(x,t;\underline{u}) = \sum_{i=1}^{\nu} [y(x^i,t;\underline{u}) - z(x^i,t)]\delta(x-x^i) \qquad (4.18)$$

$$p(\Sigma) = 0 \qquad (4.19)$$

$$p(x,T) = 0 \qquad (4.20)$$

($\delta(x-x^i)$ is the dirac delta function)

$$-\frac{\partial p(\Sigma)}{\partial \nu_{A*}} + u_1(\Sigma) - z_1(\Sigma) = 0 \qquad (4.21)$$

$$p(x,0) + u_2(x) - z_2(x) = 0 \qquad (4.22)$$

Theorem IV.1,IC

Given the system S_D, hypothesis H_D and the functional $J(\underline{v})$ of PR.IC, then:

(i) There exists one and only one $\underline{u} \varepsilon V$ with the property that

$$J(u) \leq J(\underline{v}) \text{ for all } v \varepsilon V$$

(ii) The unique solution for $\underline{u}$ is trivial, and is given by:

$$u_1(\Sigma) = z_1(\Sigma) \qquad (4.23)$$

$$u_2(x) = z_2(x) \qquad (4.24)$$

These three theorems are now proven in turn. Before proceeding however, the following Lemmas which assist in the proofs are presented:

Lemma IV.1

There exists a unique solution $y(x,t;\underline{v})$ to the system S_D and hypothesis H_D with the property that

$$y(\cdot,\cdot;\underline{v}) \varepsilon L^2(Q)$$

This Lemma was stated earlier in Chapter II, namely, Lemma II.2.

Lemma IV.2

For $y(x,t;\underline{v})$, $z(x,t;\underline{v}) \varepsilon L^2(Q)$ then there exists one and only one solution to the system of equations (4.10), (4.11) and (4.12) with the properties[1]

(i) $p(\cdot,\cdot,;\underline{v})\varepsilon L^2(Q)$

(ii) $\frac{\partial p(\cdot)}{\partial \nu_{A^*}}\varepsilon L^2(\Sigma)$

Proof of Lemma IV.2

Part (i) of the Lemma is the same as Lemma IV.1, simply reversing time from t = T to t = 0.

Part (ii) is due to Lions [18].

Lemma IV.3

For $y(x^i,\cdot;\underline{v})$ and $z(x^i,\cdot;\underline{v})\varepsilon L^2(0,T)$, then there exists one and only one solution to the system of equations (4.18), (4.19) and (4.20) with the properties

(i) $p(\cdot,\cdot;\underline{v})\varepsilon L^2(Q)$

(ii) $\frac{\partial p(\cdot)}{\partial \nu_{A^*}}\varepsilon L^2(\Sigma)$

Proof of Lemma IV.3

Same as for Lemma IV.2.

---

1 In fact, a stronger result is possible (i.e., it is possible to restrict $p(x,t;\underline{v})$ to the Sobolev space H_0^1, H_0^2 etc. See reference [18]).

Lemma IV.4

The functional $J(\underline{v})$ of PR.IA has the representation

$$J(\underline{v}) = a(\underline{v},\underline{v})-2l(\underline{v})+c \qquad (4.25)$$

where

(i) $a(w,v)$ is a continuous bilinear form on $L^2(\Sigma)\times L^2(\Omega)$

(ii) $l(\underline{v})$ is a continuous linear form on $L^2(\Sigma)\times L^2(\Omega)$

(iii) c is a nonnegative constant

Proof of Lemma IV.4

If we define

$$a(\underline{v},\underline{v}) = \int_Q [y(x,t;\underline{v})-y(x,t;\underline{0})]^2 dx\, dt$$

$$+\int_\Sigma [v_1(\Sigma)]^2 d\Sigma + \int_\Omega [v_2(x)]^2 dx \qquad (4.26)$$

$$l(\underline{v}) = -\{\int_Q [y(x,t;\underline{v})-y(x,t;\underline{0})][y(x,t;0)-z(x,t)]dx\, dt$$

$$- \int_\Omega v_1(\Sigma)z_1(\Sigma)d\Sigma - \int_\Omega v_2(x)z_2(x)dx\} \qquad (4.27)$$

$$c = \int_Q [y(x,t;\underline{0})-z(x,t)]^2 dx\, dt$$

$$+ \int_\Sigma [z_1(\Sigma)]^2 d\Sigma + \int_\Omega [z_2(x)]^2 dx \qquad (4.28)$$

Then is is clear that:

$$J(\underline{v}) = a(\underline{v},\underline{v})-2l(\underline{v})+c$$

It remains to be shown that $a(\underline{v},\underline{v})$ given by (4.26) and $l(\underline{v})$ given by (4.27) are continuous bilinear and linear forms (respectively) on $L^2(\Sigma)\times L^2(\Omega)$.

Owing to the linearity of the process S_D, bilinearity of $a(\underline{v},\underline{v})$ and linearity of $l(\underline{v})$ are evident. Continuity of $a(\underline{v},\underline{v})$ and $l(\underline{v})$ is immediate from the boundedness of (4.26) and (4.27) arising from the hypothesis H_D and Lemma IV.1. Finally, it is clear from (4.26) that $a(\underline{v},\underline{v})$ is coercive, that is,

$$a(\underline{v},\underline{v}) \geq \alpha||\underline{v}||^2_{L^2(\Sigma)\times L^2(\Omega)} \text{ (for } \alpha = 1)$$

Lemma IV.5

The functional $J(\underline{v})$ of PR.IB has the representation

$$J(\underline{v}) = a(v,v)-2l(v)+c \qquad (4.29)$$

where

(i) $a(v,w)$ is a continuous bilinear form on $L^2(\Sigma)\times L^2(\Omega)$

(ii) $l(v)$ is a continuous linear form on $L^2(\Sigma)\times L^2(\Omega)$

(iii) c is a real nonnegative constant.

Proof of Lemma IV.5

Same as for Lemma IV.4.

Lemma IV.6

The functional $J(\underline{v})$ of PR.IC has the representation

$$J(\underline{v}) = a(\underline{v},\underline{v})-2l(\underline{v})+c$$

where

(i) $a(v,w)$ is a continuous bilinear form on $L^2(\Sigma)\times L^2(\Omega)$

(ii) $l(\underline{v})$ is a continuous linear form on $L^2(\Omega)\times L^2(\Omega)$

(iii) c is a real nonnegative constant.

Proof of Lemma IV.6

The proof is the same as for Lemma IV.4.

The groundwork for the proof of Theorems IV.1.IA,.IB,.IC has been laid. We now proceed directly to the proofs.

Proof of Theorem IV.1,IA

Invoke Lemma IV.4 and write

$$J(\underline{v}) = a(\underline{v},\underline{v})-2l(\underline{v})+c$$

with $V = L^2(\Sigma)\times L^2(\Omega)$, we have that

(a) V is convex

(b) V is closed

Then, all the hypothesis of Theorem II.5 are satisfied and thus:

(i) There exists one and only one $u \varepsilon V$ such that

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v}\varepsilon V$$

Furthermore, $\underline{u}$ is characterized by

$$a(\underline{u},\underline{v})-l(\underline{v}) = 0 \quad \text{for all } \underline{v}\varepsilon V \tag{4.30}$$

Using the definitions (4.26) and (4.27), (4.30) is

$$\int_Q [y(x,t;\underline{u})-z(x,t)][y(x,t;\underline{v})-y(x,t;\underline{0})]dx\ dt$$

$$+\int_\Sigma [u_1(\Sigma)-z_1(\Sigma)]v_1(\Sigma)d\Sigma+\int_\Omega [u_2(x)-z_2(x)]v_2(x)dx=0 \qquad (4.31)$$

The system adjoint to $S_D^{(1)}$ evolves according to

$$-\frac{\partial p(x,t;\underline{u})}{\partial t}+A[p(x,t;\underline{u})]=y(x,t;\underline{u})-z(x,t) \qquad (4.32)$$

$$p(\Sigma) = 0 \qquad (4.33)$$

$$p(x,T) = 0 \qquad (4.34)$$

If (4.32) is multiplied through by $[y(x,t;\underline{v})-y(x,t;\_)]$, $v \varepsilon V$ (arbitrary) and the result integrated over Q, we obtain:

$$\int_Q \{[-\frac{\partial p(x,t;\underline{u})}{\partial t}+A[p(x,t;\underline{u})][y(x,t;\underline{v})-y(x,t;\underline{0})]\}dx\ dt$$

$$=\int_Q [y(x,t;\underline{u})-z(x,t)][y(x,t;\underline{v})-y(x,t;\underline{0})]dx\ dt \qquad (4.35)$$

The reason for introducing the adjoint system (4.32) to (4.34) and then generating (4.35) is now apparent: The R.H.S. of (4.35) is the same as the first term in (4.31). Thus the characterization of

---

1 Equations (4.32), (4.33) and (4.34) define the "adjoint" variable $p(x,t;\underline{u})$.

$\underline{u}$ given by (4.30) is equivalent, via (4.31) and (4.35) to:

$$\int_Q\{[-\frac{\partial p(x,t;\underline{u})}{\partial t}+A[p(x,t;\underline{u})][y(x,t;\underline{v})-y(x,t;\underline{0})]\}dx\ dt$$

$$+\int_\Sigma[u_1(\Sigma)-z_1(\Sigma)]v_1(\Sigma)d\Sigma+\int_\Omega[u_2(x)-z_2(x)]v_2(x)dx=0$$

$$\text{for all } \underline{v}\varepsilon V \qquad (4.36)$$

The result which is desired (equations (4.13) and (4.14)) is at hand. Using Green's theorem[1] for "integration by parts", observe that

$$\int_Q A[p(x,t;\underline{u})][y(x,t;\underline{v})-y(x,t;\underline{0})]dx\ dt$$

$$=\int_Q p(x,t;\underline{u})[\frac{\partial y(x,t;\underline{0})}{\partial t}-\frac{\partial y(x,t;\underline{v})}{\partial t}]dx\ dt$$

$$+\int_\Sigma\{-\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A*}}\underline{v}_1(\Sigma)+p(\Sigma;\underline{u})[\frac{\partial y(\Sigma;\underline{v})}{\partial \nu_A}-\frac{\partial y(\Sigma;0)}{\partial \nu}]\}d\Sigma \qquad (4.37)$$

In addition,

$$\int_Q -\frac{\partial p(x,t;\underline{u})}{\partial t}[y(x,t;\underline{v})-y(x,t;\underline{0})]dx\ dt$$

$$=\int_Q -p(x,t;\underline{u})[\frac{\partial y(x,t;\underline{0})}{\partial t}-\frac{\partial y(x,t;\underline{v})}{\partial t}]dx\ dt$$

$$+\int_\Omega\{p(x,0;\underline{u})[v_2(x)] - p(x,T;\underline{u})[y(x,T;\underline{v})]\}dx \qquad (4.38)$$

---

1 See Remark IV.1

Putting (4.37) and (4.38) into (4.36) where appropriate and noting that $p(x,T;\underline{u})=p(\Sigma;\underline{u})=0$, we obtain:

$$\int_\Sigma\{[-\frac{\partial p(\Sigma;\underline{u})}{\partial\nu_{A*}}+u_1(\Sigma)-z_1(\Sigma)]v_1(\Sigma)d\Sigma$$

$$+\int_\Omega\{[p(x,0;\underline{u})+u_2(x)-z_2(x)]v_2(x)dx = 0$$

$$\text{for all } \underline{v}\varepsilon V \tag{4.39}$$

Finally, the arbitrariness of $v_1(\Sigma)$ and $v_2(x)$ give the required result:

$$-\frac{\partial p(\Sigma;\underline{u})}{\partial\nu_{A*}}+u_1(\Sigma)-z_1(\Sigma)=0 \tag{4.40}$$

$$p(x,0;\underline{u})+u_2(x)-z_2(x)=0 \tag{4.41}$$

Assertion (ii) of Theorem IV.1,IA has been established, since (by Lemmas IV.1 and IV.2) the solution of the system of equations (4.7) through (4.14) is unique.

<u>Remark IV.1</u>

Consider the application of Green's theorem in the following situation:

$$\int_Q A[p(x,t;\underline{u})][y(x,t;\underline{v})-y(x,t;\underline{0})]dx\ dt$$

$$=\int_Q A[y(x,t;\underline{v})-y(x,t;0)]p(x,t;\underline{u})dx\ dt$$

$$+\int_\Sigma\{-\frac{\partial \underline{p}(\Sigma;\underline{u})}{\partial \nu_{A*}}v_1(\Sigma)-p(\Sigma;\underline{u})[\frac{\Sigma \underline{y}(\Sigma;\underline{v})}{\partial\nu}-\frac{\underline{y}(\Sigma;\underline{0})}{\partial\nu}]\}d\Sigma \qquad (4.42)$$

It is assumed that $p(x,t;\underline{u})$ is sufficiently differentiable ($\frac{\partial p}{\partial x^i}\varepsilon L^2(Q)$; $\frac{\partial^2 p}{\partial x^{i2}}\varepsilon L^2(Q)$) so that the differentiation of $[y(x,t;v)-y(x,t;\underline{0})]$ can be taken in the distribution sense.(1) This remark is essential, since $A[y(x,t;\underline{v})]\not\varepsilon L^2(Q)$.

The proof of theorem IV.1,IA has been given in complete detail. The proof of theorem IV.1,IB and IV.1,IC proceed in precisely the same way. Thus we give a formal outline of the salient features of the proof in each case.

Proof of Theorem IV.1,IB

The proof of part (i) is the saem as given in the proof of Theorem IV.1,IA. The structure of the proof of part (ii) is (again) the saem as given in the proof of Theorem IV.1,IA. As before, the

---

1 Given a "test function" $\phi(x)$, in C^1 with compact support in , with $\phi(\cdot)\varepsilon L^2(\Omega)$, $\frac{\partial\phi}{\partial X_i}\varepsilon L^2(\Omega)$, i=1,2...r. Then the distribution derivtive, or "secant" of the function $f(\cdot)\varepsilon L^2(\Omega)$ is defined as the linear map:

$$\frac{\partial f}{\partial x^i} : \phi \rightarrow -\int_\Omega \frac{\partial\phi(\ )}{\partial x^i} f(x)dx$$

existence of an appropriately smooth $p(x,t;\underline{u})$ satisfying (4.17), (4.18) and (4.19) is again (as in Remark IV.1) assumed. The assumption is vindicated by the results of Lions [18].

Proof of Theorem IV.1,IC

Part (i) carries over exactly as before. Part (ii) is trivial. For we have

$$-\frac{\partial p}{\partial t}(x,t;\underline{u})+A[p(x,t;\underline{u})] = 0$$

$$p(\Sigma;\underline{u}) = 0$$

$$p(x,T) = 0$$

which implies that

$$p(x,t;\underline{u}) \equiv 0 \quad \text{everywhere.}$$

which, in turn implies

$$u_1(\Sigma) = z_1(\Sigma) \quad \text{almost everywhere in } L^2(\Sigma)$$

$$u_2(x) = z_2(x) \quad \text{almost everywhere in } L^2(\Omega)$$

The solution of PR.I is complete. PR.II and PR.III differ from PR.I in that the boundary conditions of the parabolic system S are Neumann and "mixed", respectively. The criteria of identification given for PR.II and PR.III are the same as for PR.I. We now define PRII, the identification problem for the parabolic system with

A.1.1 Remarks on the Results

Having solved a variety of identification problems for a linear parabolic system with spatial operator A[·], it is tempting, in the light of these results, to argue that these results could have been embodied in a general theorem or "maximum principle". In fact, each of the theorems given have the essence, if not the substance, of such a principle. To see this, recall that each of the problems considered had the structure

$$\inf_{\substack{\underline{v}\varepsilon V \\ y\varepsilon S}} J(\underline{v}) = [a(\underline{v},\underline{v})-2l(\underline{v})+c] \qquad (4.43)$$

The solution to (4.85) was characterized by the system of equations

S	:	System evolution process	(4.44)
S*	:	System adjoint evolution process	(4.45)
P	:	"Maximum Principle", $a(\underline{u},\underline{v})-l(\underline{v})=0$	(4.46)

Tempting as it may be, the formal procedure outlined (in abstract) by (4.85) to (4.86) cannot be immediately applied as an algorithmic method for the solution of variational problems arising in the context of this study. The reason for this has been hinted at in Chapter III, section A.2.1. Let it be said that the procedure (4.44) to (4.46) would provide the necessary, (but not necessarily sufficient) conditions for optimality. However, considerable insight would be lost by such a blind application. The stumbling block is the nature of the solutions to S and S*, and whether these solutions play a role in the definition of $J(\underline{v})$.

It is possible that the system S has such a response that $J(\underline{v})$ has no minimum, for $\underline{v} \varepsilon V$. In such a situation, the procedure (4.44) to (4.46) may characterize a nonexistent minimum. It is for this reason that the response $y(x,t;\underline{v})$ was categorized for each of the three parabolic systems considered. Further, the effect of $y(x,t;\underline{v})$ on the functional $J(\underline{v})$ was tested. In this way, a solution could be guaranteed. The characterization of this solution in terms of the adjoint variable $p(x,t;\underline{u})$ required the establishment of existence of solutions of a particular class for $p(x,t;\underline{u})$. Clearly, if a solution to the adjoint system S*, denoted by $p(x,t;\underline{u})$ does not exist, then the characterization in terms of $p(x,t;\underline{u})$ is not viable.[1]

The solutions to the adjoint equations (4.10)-(4.12), (4.18)-(4.20),

have been established [18] to be of the appropriate class.

---

1 It is conceivable that a solution to the identification problem exists in spite of this. It cannot be represented in terms of $p(x,t;\underline{u})$ however. On the other hand, the pathological behaviour of $p(x,t;\underline{u})$ might be an indicator of non-existence of extremals to the variational problem.

We remarked that it is possible that S has such a response that $J(\underline{v})$ has no minimum. Such a case has already been considered, Chapter III, section A.2.1. In terms of the notation of this chapter the problem is:

$$\operatorname*{Inf}_{\substack{\underline{v}\varepsilon V \\ \underline{y}\varepsilon S_D}} J(\underline{v}) = \{\int_\Omega [y(x,T;\underline{v})-z(x,T)]^2 dx\ dt$$

$$+ \int_\Sigma [v_1(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+ \int_\Omega [v_2(x)-z_2(x)]^2 d\Omega$$

If we would proceed directly to the solution as implied via (4.44) to (4.46), then we would have as a characterization of $\underline{u}\varepsilon V$:

$$\frac{\partial y}{\partial t}(x,t;\underline{u})+A[y(x,t;\underline{u})] = f(x,t) \qquad \text{in } Q \qquad (4.47)$$

$$y(\Sigma) = u_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.48)$$

$$y(x,0) = u_2(x) \qquad \text{in } \Omega \qquad (4.49)$$

$$-\frac{\partial p}{\partial t}(x,t;\underline{u})+A[p(x,t;\underline{u})] = 0 \qquad \text{in } Q \qquad (4.50)$$

$$p(\Sigma) = 0 \qquad \text{on } \Sigma \qquad (4.51)$$

$$p(x,T) = - y(x.T;\underline{v})-z(x,T) \qquad \text{in } \Omega \qquad (4.52)$$

$$\frac{\partial p}{\partial \nu_{A*}}(\Sigma;\underline{u})+u_1(\Sigma)-z_1(\Sigma) = 0 \qquad \text{on } \Sigma \qquad (4.53)$$

$$p(x,0;\underline{u})+u_2(x)-z_2(x) = 0 \qquad \text{in } \Omega \qquad (4.54)$$

It is possible that the system of equations (4.47) to (4.54) has a solution - if so, it is unique and the choice of $\underline{u}$ from (4.53) and (4.54) is optimal. However, under the hypothesis H_D, there are responses $y(\cdot,T;\underline{u})$ which are not in $L^2(\Omega)$. Thus (4.47)-(4.54) is not a maximum principle.

From an engineering point of view, we might be persuaded to regard those inputs causing unboundedness in $y(x,T;\underline{u})$ as occurring with probability zero and therefore accept (4.47)-(4.54) as the solution. However, we have demonstrated only one counterexample. Arguing from the same engineering point of view, we would be inclined to admit to the existence of others, and also to the possibility that these others contain "realistic" hypothesis. Thus, there would be no certainty that a formal solution to the system (4.47)-(4.54) would yield the optimal $\underline{u}$.

There is another approach. We could insist upon the realistic constraint that the class of inputs $\underline{u}$ be bounded. This added restriction to the problem complicates the mathematics, as it is necessary to employ more sophisticated methods in the proofs of some of the theorems. For completeness, such an example will be considered in section A.3. In the context of the identification problem, the assumption on boundedness of $\underline{u}$ corresponds to a sort of "confidence" that the "worst" possible estimates on $\underline{u}$ are known.

With these remarks in mind, some identification problems for hyperbolic systems are now considered.

A.2 Hyperbolic Systems, $V = L^2(\Sigma)\times L^2(\Omega)$

As for the parabolic system considered in section A.1, there are three hyperbolic systems associated with the spatial operator A[·]:

(I) Dirichlet boundary conditions

(II) Neumann boundary conditions

(III) "Mixed" boundary conditions

For each of these systems, we can pose an identification problem induced by an error functional $J(\underline{v})$. Provided that this error functional can be phrased in cannonical form

$$J(\underline{v}) = a(\underline{v},\underline{v}) - 2l(\underline{v}) + c$$

where $a(\underline{v},\underline{v})$ is a continuous bilinear form on $L^2(\Sigma)\times L^2(\Omega)$ and $l(\underline{v})$ is a continuous linear form on $L^2(\Sigma)\times L^2(\Omega)$, then the solution of the problem is given by the solution to

S	:	System evolution process	(4.55)
S*	:	System adjoint evolution process	(4.56)
P,		"Maximum Principle": $a(\underline{u},\underline{v}) - l(\underline{v}) = 0$	(4.57)

As in section A.1, we shall pose the identification problem then present its solution as a theorem. The burden of proof of the theorem will lie in the area of establishing the appropriate properties of solutions to S, S* and the cannonical nature of $J(\underline{v})$.

Again, as in section A.1, we shall consider for system (I), and (III) identification problems induced by error functionals of the following "output" measurement processes:

$$\text{(A)}\quad z(x,t) = y(x.t)+\varepsilon(x,t) \qquad \text{in } Q \qquad (4.58)$$

$$z(x,t)\varepsilon L^2(Q)$$

$$\text{(B)}\quad z(x^i,t) = y(x^i,t)+ \ (x^i,t) \qquad \text{in } Q \qquad (4.59)$$

$$z(x^i,t)\ L^2(0,T)$$

$$\text{(C)}\quad z(x,t) = 0 \qquad (4.60)$$

The "input" measurement processes are (in each of (A), (B) and (C))

$$z_1(\Sigma) = u_1(\Sigma)+\varepsilon_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.61)$$

$$z(\Sigma)\varepsilon L^2(\Sigma)$$

$$z_2(x,0) = u_2(x)+\varepsilon_2(x) \qquad \text{in } \Omega \qquad (4.62)$$

$$z_3(x,0) = u_3(x)+\varepsilon_3(x) \qquad \text{in } \Omega \qquad (4.63)$$

Again, the treatment of systems (II) and (III) is given in [21].

<u>Define the following error functionals</u>

<u>PR.A</u>

$$J(\underline{v}) = \int_Q [y(x,t;\underline{v})-z(x,t)]^2 dx\ dt$$

$$+ \int_\Sigma [v_1(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+ \int_\Omega \{[v_2(x)-z_2(x)]^2+[v_3(x)-z_3(x)]^2\} dx \qquad (4.64)$$

PR.B $$J(\underline{v}) = \sum_{i=1}^{\nu} \int_0^T [y(x^i,t;\underline{v})-z(x^i,t)]^2 dx\, dt$$

$$+ \int_\Sigma [v_1(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+ \int_\Omega \{[u_2(x)-z_2(x)]^2+[u_3(x)-z_3(x)]^2 dx \qquad (4.65)$$

PR.C $$J(\underline{v}) = \int_\Sigma [v_1(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+\int_\Omega \{[v_2(x)-z_2(x)]^2+[v_3(x)-z_3(x)]^2\} dx \qquad (4.66)$$

In each of PR.A, PR.B and PR,C, $y(\cdot,\cdot;\underline{v})$ is a trajectory of a hyperbolic system (I), (II) or (III). Consider the following identification problems and their solution:

PRI : Dirichlet Boundary Conditions

System S_D

$$\frac{\partial^2 y}{\partial t^2}(x,t)+A[y(x,t)] = f(x,t) \qquad \text{in } Q \qquad (4.67)$$

$$y(\Sigma) = v_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.68)$$

$$y(x,0) = v_2(x) \qquad \text{in } \Omega \qquad (4.69)$$

$$\frac{\partial y}{\partial t}(x,0) = v_3(x) \qquad \text{in } \Omega \qquad (4.70)$$

Hypothesis H_D

$$f(\cdot,\cdot)\varepsilon L^2(Q) \quad ; \int_Q |f(x,t)|^2 dx\, dt < \infty$$

$$\left.\begin{array}{ll} v_1(\cdot)\varepsilon L^2(\Sigma) & ;\ \int_\Sigma |v_1(\Sigma)|^2 d\Sigma < \infty \\ v_2(\cdot)\varepsilon L^2(\Omega) & ;\ \int_\Omega |v_2(x)|^2 dx < \infty \\ v_3(\cdot)\varepsilon L^2(\Omega) & ;\ \int_\Omega |v_3(x)|^2 dx < \infty \end{array}\right\} = V$$

Additional hypothesis on A[·]:[2]

$$a_{ij}(x,t) = a_{ji}(x,t) \quad i,j=1,2,3...r$$

Remark

The symmetry on the coefficients of the differential operator A[·] was not necessary in the parabolic case but is essential here. For the system S_D and hypothesis H_D, we give the following results as theorems:

Theorem IV.2,IA

Given the system S_D, hypothesis H_D and the functional $J(\underline{v})$ of PRA, then:

(i) There exists one and only one $\underline{u}\varepsilon V$[1] with the property that

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v}\varepsilon V$$

---

1 The vector $\underline{u}$ now represents $[u_1(\Sigma)\ u_2(x)\ u_3(x)]^T$

2 See the definition of A[·], Chapter III section A.1.2.

(ii) The boundary and initial data $\underline{u}$ is uniquely characterized by the simultaneous solution of the following system of equations:

$$\frac{\partial^2 y(x,t;\underline{u})}{\partial t^2}+A[y(x,t;\underline{u})] = f(x,t) \qquad \text{in } Q \qquad (4.71)$$

$$y(\Sigma) = u_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.72)$$

$$y(x,0) = u_2(x) \qquad \text{in } \Omega \qquad (4.73)$$

$$\frac{\partial y(x,0)}{\partial t} = u_3(x) \qquad \text{in } \Omega \qquad (4.74)$$

$$\frac{\partial^2 p(x,t;\underline{u})}{\partial t^2}+A[p(x,t;\underline{u})] = y(x,t;\underline{u})-z(x,t) \qquad \text{in } Q \qquad (4.75)$$

$$p(\Sigma) = 0 \qquad \text{on } \Sigma \qquad (4.76)$$

$$p(x,T) = 0 \qquad \text{in } \Omega \qquad (4.77)$$

$$\frac{\partial p(x,T)}{\partial t} = 0 \qquad \text{in } \Omega \qquad (4.78)$$

$$-\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A*}}+u_1(\Sigma)-z_1(\Sigma) = 0 \qquad \text{on } \Sigma \qquad (4.79)$$

$$p(x,0;\underline{u})+u_2(x)-z_2(x) = 0 \qquad \text{in } \Omega \qquad (4.80)$$

$$-\frac{\partial p(x,0;\underline{u})}{\partial t}+u_3(x)-z_3(x) = 0 \qquad \text{in } \Omega \qquad (4.81)$$

Theorem IV.2,IB

Given the system S_D, Hypothesis H_D and functional $J(\underline{v})$ of PRB, then:

(i) There exists one and only one $\underline{u}\varepsilon V$ with the property that

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v}\varepsilon V$$

(ii) The boundary and initial data $\underline{u}$ is uniquely characterized by the simultaneous solution of the following system of equations:

$$\frac{\partial^2 y(x,t;\underline{u})}{\partial t^2} + A[y(x,t;\underline{u})] = f(x,t) \qquad \text{in } Q \qquad (4.82)$$

$$y(\Sigma) = u_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.83)$$

$$y(x,0) = u_2(x) \qquad \text{in } \Omega \qquad (4.84)$$

$$\frac{\partial y(x,0)}{\partial t} = u_3(x) \qquad \text{in } \Omega \qquad (4.85)$$

$$\frac{\partial^2 p(x,t;\underline{u})}{\partial t^2} + A[p(x,t;\underline{u})] = \sum_{i=1}^{\nu} [y(x^i,t;\underline{u}) - z(x^i,t)]\delta(x-x^i) \qquad \text{in } Q \qquad (4.86)$$

$$p(\Sigma;\underline{u}) = 0 \qquad \text{on } \Sigma \qquad (4.87)$$

$$p(x,T) = 0 \qquad \text{in } \Omega \qquad (4.88)$$

$$\frac{\partial p(x,T)}{\partial t} = 0 \qquad \text{in } \Omega \qquad (4.89)$$

$$-\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A^*}} + u_1(\Sigma) - z_1(\Sigma) = 0 \qquad \text{on } \Sigma \qquad (4.90)$$

$$p(x,0;\underline{u}) + u_2(x) - z_2(x) = 0 \qquad \text{in } \Omega \qquad (4.91)$$

$$-\frac{\partial p(x,0;\underline{u})}{\partial t} + u_3(x) - z_3(x) = 0 \qquad \text{in } \Omega \qquad (4.92)$$

Remark

$\frac{\partial p}{\partial \nu_{A*}} = \frac{\partial p}{\partial \nu_A}$ for hyperbolic systems under the hypothesis on A[·] = A*[·].

Theorem IV.2,IC

Given the system S_D, hypothesis H_D and the functional $J(\underline{v})$ of PRC, then:

(i) There exists one and only one $\underline{u} \varepsilon V$ with the property that:

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v} \varepsilon V$$

(ii) The unique solution for $\underline{u}$ is trivial and is given by

$$u_1(\Sigma) = z_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.93\ )$$

$$u_2(x) = z_2(x) \qquad \text{in } \Omega \qquad (4.94\ )$$

$$u_3(x) = z_3(x) \qquad \text{in } \Omega \qquad (4.95\ )$$

Remark on the Proof of Theorems IV.2,IA, IV.2,IB and IV.2,IC

The proofs of these theorems carry through exactly as for Theorems IV.1,IA, IV.1,IB and IV.1,IC. The additional hypothesis on A[·] given in H_D over and above the hypothesis on A[·] for parabolic systems (namely that A[·] = A*[·]) allows this exact equivalence. The ramification of this additional hypothesis is that the responses for $y(x,t;\underline{u})$ and $p(x,t;\underline{u})$ are of the same classes as in the corresponding parabolic systems.

A.3 An Example of Constrained State Identification

As we mentioned in the closing paragraph of section A.1, it is reasonable to pose the following identification problem for a given system S:

P Given a system S, Input and Output measurements I and O respectively, where the input measurements are known with a specified confidence, denoted by C. Then, obtain a refinement of these measurements, R, along a trajectory of S.

Problem P corresponds to a physically appealing notion, that the error processes ε associated with the boundary and initial data measurements is bounded by known quantities λ_1 and λ_2:

$$\underline{\lambda}_1 \leq \varepsilon \leq \underline{\lambda}_2$$

P: Identification Problem for Bounded Input $v \varepsilon V \subset L^2(\Sigma) \varepsilon L^2(\Omega)$

System S_D

$$\frac{\partial y}{\partial t}(x,t) + A[y(x,t)] = f(x,t) \quad \text{in } Q \qquad (4.96)$$

$$y(\Sigma) = u_1(\Sigma) \quad \text{on } \Sigma \qquad (4.97)$$

$$y(x,0) = u_2(\Sigma) \quad \text{in } \Omega \qquad (4.98)$$

Hypothesis H_D

$$f(\cdot,\cdot)\ L^2(Q)$$

$$\left.\begin{array}{l} v_1(\cdot) \varepsilon L^2(\Sigma) \text{ and } \lambda_1^L \leq v_1(\Sigma) \leq \lambda_1^U \\ v_2(\cdot) \varepsilon L^2(\Omega) \text{ and } \lambda_2^L \leq v_2(x) \leq \ _2^U \end{array}\right\} = V$$

Then we have the following theorem:

Theorem IV.3

Given the system S_D, hypothesis H_D and the functional $J(\underline{v})$ of PRA, then:

(i) There is one and only one $\underline{u} \varepsilon V$ with the property that:

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v} \varepsilon V$$

(ii) The boundary and initial data $\underline{u}$ is uniquely characterized by the simultaneous solution of the following system of equations:

$$\frac{\partial y(x,t;\underline{u})}{\partial t}+A[y(x,t;\underline{u})] = f(x,t) \qquad \text{in } Q \qquad (4.99)$$

$$y(\Sigma) = u_1(\Sigma) \qquad \text{on } \Sigma \qquad (4.100)$$

$$y(x,0) = u_2(x) \qquad \text{in } \Omega \qquad (4.101)$$

$$\frac{\partial p(x,t\underline{u})}{\partial t}+A[p(x,t;\underline{u})] = y(x,t;\underline{u})-z(x,t) \qquad \text{in } Q \qquad (4.102)$$

$$p(\Sigma) = 0 \qquad \text{on } \Sigma \qquad (2.103)$$

$$p(x,T) = 0 \qquad \text{in } \Omega \qquad (2.104)$$

$$[-\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A*}}+u_1(\Sigma)-z_1(\Sigma)][\lambda_1 - u_1(\Sigma)] \geq 0$$

$$\lambda_1^L \leq \lambda_1 \leq \lambda_1^U \; ; \qquad \text{on } \Sigma \qquad (4.105)$$

$$[p(x,0;\underline{u}) + u_2(x) - z_2(x)][\lambda_2 - u_2(x)] \geq 0$$

$$\lambda_2^L \leq \lambda_2 \leq \lambda_2^U \; ; \qquad \text{in } \Omega \qquad (4.106)$$

In the proof of Theorem IV.3, it is necessary to invoke the following Lemma due to Lebesque:

Lemma IV.7

Given a measureable function $\Psi(x,t)$ defined on Q. If O_j is an elemental volume of Q at M_0 with

$$\lim_{j\to\infty} 0_j \to 0,$$

then

$$\frac{1}{\mu(0_j)} \int_{0_j} \Psi(x,t)dx\, dt \to 0$$

$$\lim_{j\to\infty} \frac{1}{\mu(0_j)} \int_{0_j} \Psi(x,t)dx\, dt = \Psi(M_0) \geq 0$$

almost everywhere in Q.

Remark on Lemma IV.7

Note the weak hypothesis on $\Psi(x,t)$. In particular, $\Psi(x,t)$ is not necessarily continuous in Q.

Proof of Theorem IV.3

Invoke Lemma IV.4 and write $J(\underline{v})$ of PRA as:

$$J(\underline{v}) = a(\underline{v},\underline{v})-2l(\underline{v})+c \tag{4.107}$$

in addition,

$$V = \{\underline{v} : \underline{v}\varepsilon L^2(\Sigma)\times L^2(\Omega);\ \underline{\lambda}^L \leq \underline{v} \leq \underline{\lambda}^U\}$$

Evidently,

(a) V is closed

(b) V is convex[1]

---

1 for k_1, $k_2 \varepsilon V$, $0 \leq \mu \leq 1$
$\underline{v}_\mu = \mu k_1+(1-\mu)k_2\varepsilon V$. $(v_\mu \varepsilon L^2(\Sigma)\times L^2(\Omega);\ \lambda^L \leq v_\mu \leq \lambda^U)$.

Thus the hypothesis of Theorem II.4 is satisfied and so we have immediately that

(i) There exists one and only one $\underline{u} \varepsilon V$ such that

$$J(\underline{u}) \leq J(\underline{v}) \text{ for all } \underline{v} \varepsilon V$$

Moreover, $\underline{u}$ is characterized by

$$a(\underline{u},\underline{v}-\underline{u}) \geq l(\underline{v}-\underline{u}) \text{ for all } \underline{v} \varepsilon V \tag{4.108}$$

Using the definitions (4.26) and (4.27), (4.216) is equivalent to:

$$\int_Q [y(x,t;\underline{u})-z(x,t)][y(x,t;\underline{v})-y(x,t;\underline{u})]dx\,dt$$
$$+\int_\Sigma [u_1(\Sigma)-z_1(\Sigma)][v_1(\Sigma)-u_1(\Sigma)]d\Sigma$$
$$+\int_\Omega [u_2(x)-z_2(x)][v_2(x)-u_2(x)]dx = 0 \tag{4.109}$$

Introduce the system adjoint to S_D:

$$\frac{\partial p(x,t;\underline{u})}{\partial t}+A[p(x,t;\underline{u})]=y(x,t;\underline{u})-z(x,t) \quad \text{in } Q \tag{4.110}$$

$$p(\Sigma;\underline{u}) = 0 \quad \text{on } \Sigma \tag{4.111}$$

$$p(x,T;\underline{u}) = 0 \quad \text{in } \Omega \tag{4.112}$$

As we saw in the proof of Theorem IV.1,IA, the adjoint system can be arranged to give an alternative representation to (4.109). By multiplying (4.110) by $[y(x,t;\underline{v})-y(x,t;\underline{u})]$ and then integrating the result over Q, it is seen that substitution of (4.217) into the resulting expression yields upon application of Green's Theorem, the following result:

$$\int_{\Sigma}[-\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A*}}+u_1(\Sigma)-z_1(\Sigma)][v_1(\Sigma)-u_1(\Sigma)]d\Sigma \geq 0 \qquad (4.113)$$

$$\int_{\Omega}[p(x,0;\underline{u})+u_2(x)-z_2(x)][v_2(x)-z_2(x)]dx \geq 0 \qquad (4.114)$$

Assertion

(4.221) and (4.222) are equivalent to

$$[-\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A*}}+u_1(\Sigma)-z_1(\Sigma)][\lambda_1-u_1(\Sigma)] \geq 0;\ \lambda_1^L \leq \lambda_1 \leq \lambda_1^U \qquad (4.115)$$

$$[p(x,0;\underline{u})+u_2(x)-z_2(x)][\lambda_2-u_2(x)] \geq 0;\ \lambda_1^L \leq \lambda_2 \leq \lambda_2^U \qquad (4.116)$$

Proof

Evidently, (4. 115) and (4. 116) imply (4. 113) and (4. 114) respectively. Simply replace λ_1 by $v_1(\Sigma)$ and λ_2 by $v_2(x)$ and integrate appropriately.

However, it is not obvious that (4.113) and (4.114) imply (4.115) and (4.116), respectively. If the integrands of (4.113) and (4.114) were continuous, then the implication would be obvious. However the hypothesis H_D does not permit such a restrictive assumption.

We shall obtain the desired result by the application of Lemma IV.7.

Consider Figure IV.1. Σ is the surface of the cyclindrical sheath to the volume Q. Ω is the "bottom" surface of this volume Q. The boundary or perimeter of Ω is Γ. Let M_0 be a point on the surface Σ with coordinates (Γ_0, t_0). Let M_0 be contained in an elemental area A_j with sides h_j and 0_j. Let

$$\lim_{j\to\infty} 0_j \to 0 \ ; \ \lim_{j\to\infty} h_j \to 0$$

Consider as a "test function" $v_j(\Sigma)$, defined as follows:

$$v_j(\Sigma) = \begin{cases} \lambda_1 \text{ inside } A_j \ ; \ \lambda_1^L \leq \lambda_1 \leq \lambda_1^U \\ u_1(\Sigma) \text{ outside } A_j ; \ u_1 \varepsilon V \end{cases} \tag{4.117}$$

Evidently, $v_j \varepsilon V$. Thus, using (4.113), we obtain:

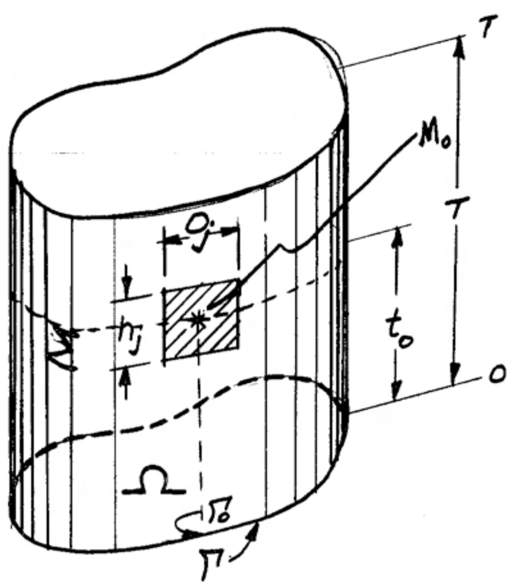

FIGURE IV.1

$$\int_{A_j} [- \frac{\partial p}{\partial \nu_{A^*}} + u_1 - z_1][\lambda - u_1] d\mu$$

$$+ \int_{\Sigma - A_j} [- \frac{\partial p}{\partial \nu_{A^*}} + u_1 - z_1][u_1 - u_1] d\mu \geq 0 \qquad (4.118)$$

where $d\mu$ is a differential area for the surface Σ. Since the second term in (4.226) is zero, we have, upon multiplication by $\frac{1}{\mu(A_j)}$,

$$\frac{1}{\mu(A_j)} \int_{A_j} [- \frac{\partial p}{\partial \nu_{A^*}} + u_1 - z_1][\lambda - u_1] d\mu \geq 0 \qquad (4.119)$$

since $\mu(A_j) \geq 0$. $\mu(A_j)$ is the "measure" of A_j - (for engineering purposes, its area).

We now invoke Lemma IV.7 and obtain that

$$[- \frac{\partial p}{\partial \nu_{A^*}}(\Gamma_0, t_0) + u_1(\Gamma_0, t_0) - z_1(\Gamma_0, t_0)][\lambda_1 - u_1(\Gamma_0, t_0)] \geq 0 \qquad (4.120)$$

But (4.120) is true for almost every Γ_0, t_0 in Σ. Arguing similarily in Ω, the assertion follows.

Thus part (ii) of the theorem follows, in virtue of the uniqueness of the solution to (4.99)-(4.104).

A.4 Concluding Remarks

The state identification problem for parabolic and hyperbolic systems associated with the operator A[·] has been solved (i.e., a characterization of the solution has been obtained) for some rather simple quadratic error functionals. The technique of solution was demonstrated via several complete proofs concerning existence and uniqueness of optimal $\underline{v} \varepsilon V$, for V bounded and unbounded. The technique is valid for any choice of quadratic error functional, as it filters out those functionals which are non-well set.

The proof of Theorem IV.3 was of special importance, since the characterization of the extremal to the identification problem was in terms of a pair of inequations. These inequations were shown to be a valid and equivalent representation of the condition of optimality,

$$a(\underline{u},\underline{v}-\underline{u}) \geq l(\underline{v}-\underline{u})$$

Previous representations of the type obtained have been formally assumed (in another context by another author) without the rigorous justification given here, due to Lions [16].

CHAPTER V

SOLUTION OF THE STATE IDENTIFICATION PROBLEM

A.0 Introduction

In Chapter IV, the solution to the state identification problem associated with the distributed system S was characterized by a set of simultaneous equations, denoted here by E.

Here, we consider two methods which construct a solution to E. We remark that the construction of solutions to E is not straightforward, owing to the split end conditions in time contained therein.

Specifically, a Ricatti-like decoupling of the two point "time boundary" value problem is considered in section A.2. In section A.3, we solve E by solving directly the variational problem leading to E. This solution is accomplished by hill climbing (descending) on the error functional $J(\underline{v})$. Directions of search on the hill are conjugate.

Both methods for solving E lead to the consideration of a set of partial differential equations, the solution of which must be approximated.[1] Section A.4 deals with the application of the method of Galerkin to the results of sections A.2 and A.3. Chapter V concludes with the consideration of an illustrative numerical example.

A.1 Schemes for the Solution of the Identification Equations

In order to fully appreciate the salient features of E and the

---

1 The solution of the differential equations involved defy (in general) an analytic representation which is of immediate numerical utility.

need for special methods of solution, consider the following typical set of equations E, arising out of Theorem IV.1.IA:

E: Obtain $\underline{u}$ which satisfies:

$$\frac{\partial y}{\partial t}(x,t;\underline{u}) + A[y(x,t;\underline{u})] = f(x,t) \quad \text{in } Q \quad (5.1)$$

$$y(\Sigma) = u_1(\Sigma) \quad \text{on } \Sigma \quad (5.2)$$

$$y(x,0) = u_2(x) \quad \text{in } \Omega \quad (5.3)$$

$$-\frac{\partial p}{\partial t}(x,t;\underline{u}) + A[p(x,t;\underline{u})] = y(x,t;\underline{u}) - z(x,t) \quad \text{in } Q \quad (5.4)$$

$$p(\Sigma) = 0 \quad \text{on } \Sigma \quad (5.5)$$

$$p(x,T) = 0 \quad \text{in } \Omega \quad (5.6)$$

$$-\frac{\partial p(\Sigma;u)}{\partial \nu_{A^*}} + u_1(\Sigma) - z_1(\Sigma) = 0 \quad \text{on } \Sigma \quad (5.7)$$

$$p(x,0;u) + u_2(x) - z_2(x) = 0 \quad \text{in } \Omega \quad (5.8)$$

Observe that the system of equations (5.1) through (5.6) cannot simultaneously be solved forwards in time, since the initial condition on $p(x,0;\underline{u})$ is missing. Nor can the system be solved backwards from $t = T$, since $y(x,T;\underline{u})$ is missing. However, if $y(x,T;\underline{u})$ could be defined compatibly with (5.1) through (5.8), then the system (5.1) through (5.7) could be solved backwards from $t = T$ to recover $y(x,0) = u_2(x)$. On the way, we would obtain (from (5.7)) $u_1(\Sigma)$. We shall present, in section A.2, a scheme for the compatible determination of $y(x,T;\underline{u})$ (with respect to (5.1) through (5.8)). This scheme is based on the fact that there exists a continuous linear transformation between $p(x,T;\underline{u})$ and $y(x,T;\underline{u})$. We shall see that this transformation derives its representation from the solution to a Ricatti-like partial differential equation.

An alternative to the simultaneous solution of (5.1) through (5.8) is available. Recall that the functional to be minimized (as a criterion of identification) is given by

$$J(\underline{v}) = a(\underline{v},\underline{v}) - 2l(\underline{v}) + c \qquad (5.9)$$

The $\underline{u} \epsilon V$ which minimizes $J(\underline{v})$ is characterized by the solution of

$$a(\underline{u},\underline{v}) - l(\underline{v}) = 0 \qquad (5.10)$$

along a trajectory $y(x,t;\underline{u})$.

By introducing the adjoint system $p(x,t;\underline{u})$, we saw that (5.10) was equivalent to (5.7) and (5.8). We now assert that (5.10) is the formal variational derivative of (5.9).[1]

It is this recognition which gives the alternative approach to the simultaneous solution of (5.1) through (5.8).

---

1 See Appendix V.3

Define

$$\delta J = a(\underline{u},\underline{v}) - l(\underline{v}) \tag{5.11}$$

$$G = \begin{bmatrix} -\frac{\partial p(\Sigma;\underline{u})}{\partial \nu_{A*}} + u_1(\Sigma) - z_1(\Sigma) \\ \\ p(x,0;\underline{u}) + u_2(x) - z_2(x) \end{bmatrix} \tag{5.12}$$

Again, we note that

$$\underline{\delta J} = 0 \Longleftrightarrow \underline{G} = 0$$

The technique proceeds by relaxing the condition

$$\delta J \sim \underline{G} = 0$$

This relaxation (or violation) of (5.7) and (5.8) is successively modified by the iterative scheme (at the k+1st iteration)

$$\underline{u}^{k+1} = \underline{u}^k + \alpha^k \underline{s}^k(\underline{G}) \tag{5.13}$$

In (5.12) and (5.13), $\underline{G}$ is the gradient of the functional $J(\underline{v})$. $s^k(G)$ is a conjugate gradient and α^k is the modification of $\underline{u}^k$ in the direction s^k. In section A.3, we shall deal fully with the properties of this iterative scheme. Consider next the Ricatti-like decoupling.

A.1.1 Ricatti-like Decoupling

As was indicated in section A.1, the task at hand is to define, in a way compatible with (5.1) through (5.8), the terminal time condition on the system trajectory, $y(x,T;\underline{u})$. We pointed out that if such a condition were specified, then $\underline{u}$ could be recovered (in principle) from a suitable solution to (5.1) through (5.7).

While $\underline{u}$ can be recovered from (5.1) through (5.7), frequently the definition of $y(x,T;\underline{u})$ is sufficient. That is, the identification problem would be considered solved upon obtaining $y(x,T;\underline{u})$. Thus we shall consider our task complete once having obtained an equation yielding as a result $y(x,T;\underline{u})$. In fact, we shall obtain an equation for the evolution of $y(x,T;\underline{u})$ with T considered variable. Specifically, consider the characterization of $\underline{u}$ for the Parabolic system with Dirichlet boundary conditions and error functional PRA given by equations (5.1) through (5.8). We give the result as a

Theorem V.1

Given the system of equations (5.1) through (5.8) and the hypothesis of Theorem IV.1, IA, and:

(a) If $P(x,\xi,t)$ satisfies

$$\frac{\partial P}{\partial t}(x,\xi,t)-A_{\xi}[P(x,\xi,t)]-A_{x}[P(x,\xi,t)]-\delta(\xi-x)$$

$$+\int_{\Gamma_s}\frac{\partial P}{\partial \nu_{A_s^*}}(x,\Sigma_s)\frac{\partial P}{\partial \nu_{A_s^*}}(\Sigma_s,\xi)d\Gamma_s = 0 \quad \text{in } \Omega\times\Omega\times(0,T] \tag{5.13}$$

$$P(x,\Sigma_s) = P(\Sigma_s,\xi) = 0 \tag{5.14}$$

$$P(x,\xi,0) = \delta(x-\xi) \tag{5.15}$$

(b) If $P(\cdot,\cdot,t)\varepsilon H^2(\Omega\times\Omega)$ and $\frac{\partial P}{\partial t}(\cdot,\cdot,t)\varepsilon L^2(\Omega\times\Omega)$ $t\varepsilon(0,T]$, then:

(i) There exists one and only one $\hat{y}(\cdot,\cdot)\varepsilon L^2(Q)$ such that

$$y(x,T;\underline{u}) = \hat{y}(x,T)$$

where $\hat{y}(x,t)$ is the unique solution of the following Linear Integral Equation of the Second Kind[1]:

$$\int_\Omega P(x,\xi,t)\{\frac{\partial\hat{y}}{\partial t}(\xi,t)+A[\hat{y}(\xi,t)]-f(\xi,t)\}d\xi=z(x,t)-\hat{y}(x,t) \quad (5.16)$$

where (5.16) is valid in the partially open set Q. The conditions satisfied by $\hat{y}(x,t)$ on the closure of Q are:

$$\hat{y}(\Sigma) = z_1(\Sigma) \quad (5.17)$$

$$\hat{y}(x,0) = z_2(x) \quad (5.18)$$

<u>Remarks on Theorem V.1</u>

(i) We shall see, in the proof of this theorem, that there exists an affine transformation between $p(\cdot,t)$ and $y(\cdot,t)$, denoted by π.

$$\pi : L^2(\Omega) \rightarrow H_0^1$$
$$y(\cdot,t) \mapsto p(\cdot,t)$$

---

1 This representation is not standard, but by Green's theorem can be reduced to standard form:

$$\hat{y}(x,t) = g(x,t) + \int_\Omega \Phi(x,\xi,t)y(\xi,t)d\xi$$

This transformation π has, by the Schwartz kernel theorem the following representation:

$$p(\cdot,t)=\pi[y(\cdot,t)]= -\int_{\Omega}P(\cdot,\xi,t)[y(\xi,t;\underline{u})-y(x,t)]dx \qquad (5.19)$$

Using (5.1) through (5.6), we obtain the equation which $P(x,\xi,t)$ satisfies, namely (5.13). Now we know that there exists a unique $P(x,\xi,t)$ in virtue of (5.19) and the existence and uniqueness of $p(x,t)$ and $y(x,t;\underline{u})$. However it is non-trivial to establish the class of functions to which $P(\cdot,\cdot,\cdot)$ belongs, since equation (5.13) is nonlinear. In any event, $P(x,\xi,t)$ must satisfy (5.13). The point of this remark is that $\hat{y}(\cdot,\cdot)$ exists in $L^2(Q)$ and is unique. Furthermore, $\hat{y}(x,t)$ must satisfy (5.16), (5.17) and (5.18). Sufficient conditions for the system (5.16), (5.17) and (5.18) to have a unique solution of the appropriate class impose restrictions on the class of functions $P(\cdot,\cdot,\cdot)$. These restrictions are given as hypothesis (b) and for which (5.16), (5.17) and (5.18) have a unique solution in $L^2(Q)$.

(ii) The theorem asserts that $y(x,T;\underline{u}) = \hat{y}(x,t)$ for all $x\varepsilon\Omega$. Equation (5.16) gives the evolution of $\hat{y}(x,T)$ for T considered variable. In the context of lumped parameter systems and in the statistical framework given in Chapter III section A.3, $\hat{y}(x,T)$ is called the Filtered Estimate. Note that for T considered variable, the open sets Q and Σ must be given an extended definition. Thus for some $T_E > T$,

$$Q_E = Q \times [T, T_E]$$

$$\Sigma_E = \Sigma \times [T, T_E]$$

The hypothesis on $f(x,t)$, $z(x,t)$, $z_1(\Sigma)$, $z_2(x)$ must be given on the appropriate extended sets:

$$f(\cdot,\cdot)\varepsilon L^2(Q_E) \; ; \; z(\cdot,\cdot)\varepsilon L^2(Q_E) \; ; \; z_1(\cdot)\varepsilon L^2(\Sigma_E)$$

In this way, we avoid a contradiction. For, as we have seen, (Chapter III section A.2.1)

$$y(\cdot,T;u) \not\in L^2(\Omega)$$

additionally, it is easy to show that

$$\hat{y}(\cdot,T) \not\in L^2(\Omega)$$

<u>for T fixed</u>. However if T varies, $T < T_E$

$$\hat{y}(\cdot,T)\varepsilon L^2(\Omega), \; T\varepsilon(0,T_E].$$

<u>Proof of Theorem V.1</u>

The result of the theorem is that it is possible to find an expression for the missing end condition (in time) for $y(x,t;\underline{u})$, $y(x,T;\underline{u})$. This expression is $\hat{y}(x,T)$, where $\hat{y}(x,t)$ is the solution of a given integral equation, (5.16). It is possible to find $y(x,T;\underline{u})$

because of the existence of a continuous affine mapping from $y(\cdot,t;\underline{u}) \to p(\cdot,t;\underline{u})$ for $t\varepsilon(0,T]$. Thus if this mapping can be constructed explicitly, given $p(\cdot,T)$, $y(x,T;\underline{u})$ can be recovered. Thus the task of the proof is twofold: first, to show the existence of the map and second to give an explicit formula for it. This formulation includes the result that:

$$y(x,T;\underline{u}) = \hat{y}(x,T).$$

We assert here that there exists a unique affine continuous map from $y(\cdot,t;\underline{u}) \to p(\cdot,t;\underline{u})$. That is,

$$\pi[y(\cdot,t;\underline{u})]+g(\cdot,t)=p(\cdot,t;\underline{u});\ t\varepsilon(0,T] \qquad (5.20)$$

We demonstrate (5.20) in Appendix V.1 and in addition, show, by construction, that the following Lemma holds:

Lemma V.1 (L. Schwartz):

Consider the mapping

$$\pi : \begin{array}{l} L^2(\Omega) \to L^2(\Omega) \\ y(\cdot,t;\underline{u}) \to \pi[y(\cdot,t;\underline{u})] \end{array} \qquad \text{Linear, continuous}$$

Then there exists a "Distribution Kernel" $P(x,\xi,t)$ on $\Omega\times\Omega$ such that:

$$\pi[y(\cdot,t;\underline{u})]= -\int_\Omega P(\cdot,\xi,t)y(\xi,t)d\xi \qquad (5.21)^{(1)}$$

1 The minus sign is introduced for manipulative convenience in the sequel.

It will be convenient to define $\hat{y}(x,t)$ as the unique solution of the following integral equation (of the first kind):

$$g(x,t) = \int_{\Omega} P(x,\xi,t)\hat{y}(\xi,t)d\xi \tag{5.22}$$

Thus, by (5.20)

$$p(x,t;\underline{u}) = -\int_{\Omega} P(x,\xi,t)[y(\xi,t;\underline{u}) - \hat{y}(\xi,t)]d\xi \tag{5.23}$$

We remark that $P(x,\xi,t)$ is unique, since $p(x,t;\underline{u})$, $y(x,t;\underline{u})$ and $\hat{y}(x,t)$ are. Thus we have established the existence of a unique $P(x,\xi,t)$.[1] In order to obtain a formula for $P(x,\xi,t)$, we shall use (5.23) in (5.4) through (5.8).

As a first result, we observe that (5.23) into (5.5) gives

$$\int_{\Omega} P(\Sigma_x,\xi)[y(\xi,t;\underline{u}) - \hat{y}(\xi,t)]d\xi = 0$$

which implies that

$$P(\Sigma_x,\xi) = 0 \tag{5.24}$$

and similarily

$$P(x,\Sigma_\xi) = 0 \tag{5.25}$$

Next, (5.23) into (5.8) gives:

---

1 Refer to the remark of Theorem V.1.

$$-\int_{\Omega} P(x,\xi,0)[y(\xi,0;\underline{u})-\hat{y}(\xi,0)]d\xi+u_1(x,0)-z_1(x,0)=0$$

Thus

$$P(x,\xi,0) = \delta(\xi-x) \tag{5.26}$$

$$\hat{y}(x,0) = z_2(x). \tag{5.27}$$

To obtain evolution equations for $P(x,\xi,t)$ and $y(\xi,t)$, put (5.23) into (5.4) and obtain:

$$\int_{\Omega}\{\frac{\partial P}{\partial t}(x,\xi,t)[y(\xi,t;\underline{u})-\hat{y}(\xi,t)]+P(x,\xi,t)[\frac{\partial y}{\partial t}(\xi,t;\underline{u})-\frac{\partial y}{\partial t}(\xi,t)]$$

$$-A_x[P(x,\xi,t)][y(\xi,t;\underline{u})-\hat{y}(\xi,t)]\}d\xi=y(x,t;\underline{u})-z(x,t) \tag{5.28}$$

Equation (5.28) implies ([21]) that:

$$\frac{\partial P}{\partial t}(x,\xi,t)-A_\xi[P(x,\xi,t)]-A_x[P(x,\xi,t)]-\delta(\xi-x)$$

$$+\int_{\Gamma_s}\frac{\partial P}{\partial \nu_{A_s^*}}(x,\Sigma_s)\frac{\partial P}{\partial \nu_{A^*_s}}(\Sigma_s,\xi)d\Gamma_s = 0 \quad \text{in } \Omega\times\Omega\times(0,T] \tag{5.29}$$

$$\int_{\Omega}P(x,\xi,t)\{\frac{\partial \hat{y}}{\partial t}(\xi,t)+A_\xi[\hat{y}(\xi,t)]-f(\xi,t)]\}d\xi = z(x,t)-\hat{y}(x,t)$$
$$\text{in } \Omega\times(0,T] \tag{5.30}$$

$$\hat{y}(\Sigma) = z_1(\Sigma). \tag{5.31}$$

Now for $P(x,\xi,t)$ defined by (5.29), (5.30) has a unique solution $\hat{y}(\cdot,\cdot)\varepsilon L^2(Q)$, provided hypothesis (b) of the theorem holds. Finally, we see that (5.23) substituted into (5.6) implies:

$$-\int_{\Omega} P(x,\xi,T)[y(\xi,T;\underline{u}) - \hat{y}(\xi,T)]d\xi = 0 \qquad (5.32)$$

Since $P(x,\xi,t)$ is defined by (5.24), (5.25), (5.26) and (5.29), we must have

$$y(x,T;\underline{u}) = \hat{y}(x,T) \quad \text{for all} \quad x\varepsilon\Omega.$$

We have shown how it is possible to obtain an expression for $y(x,T;\underline{u})$ compatible with equations (5.1) through (5.8) and thus it is possible (in principle) to obtain $u_1(\Sigma)$ and $u_2(x)$ by the simultaneous solution of (5.1) through (5.7) "backwards" in time, starting with $t = T$. As we remarked, the evolution of $y(x,T;\underline{u})$ from T forwards (as more measurement information contained in $\varepsilon(x,T)$ and $\varepsilon_1(\Gamma,T)$ becomes available) is often of more interest, especially in the case of the "on-line" control problem. In that situation, a refinement $y^*(x,t)$ of the true system state of nature is used to define a control strategy $m(x,t;y^*)$. The important feature of this control is that it is a function of the refinement $y^*(x,t)$ at the "present time", t. If we were to choose $y^*(\cdot,\cdot) = y(\cdot,\cdot;\underline{u})$, an inherent time delay, composed of two parts - one due to the data gathering process and the other due to computation, in incurred. Thus, at the present time, t, we would have available $y(\cdot,t-\Delta t;\underline{u})$, where Δt is the time delay. However, using $y^*(\cdot,t) = \hat{y}(\cdot,t)$; $t \geq T$ there is no data gathering time delay. We expect that $m(x,t;y)$ would be a better control action than $m(x,t;y(\cdot,t-\Delta t;\underline{u}))$, providing T is large enough.

There appear to be trade-offs in this context. The relative accuracy of the refinement $y(x,t;\underline{u})$ makes its incorporation in the control strategy attractive. While the computation time necessary to recover $y(x,t;\underline{u})$ via the Ricatti-like transformation just discussed is prohibitive, another technique (to be presented in section A.3) generates $y(x,t;\underline{u})$ with relative rapidity. Thus if $M[y(\cdot,t;\underline{u}) - y(\cdot,t-\Delta t)]$ is small ($M[\cdot]$ is some appropriate metric), where Δt is the data gathering and computation time combined, then it may be advantageous to use the refinement $y(x,t;\underline{u})$ in the control scheme rather than $\hat{y}(x,t)$.

Of course, we are interested in $y(x,t;\underline{u})$ in its own right, and presented a justification of the usefullness of the refinement $\hat{y}(x,t)$ in the context of on line control as a special case. Consequently we return our attention to the problem of solving for $\underline{u}$ (and thus $y(x,t;\underline{u})$) from the equations characterizing $\underline{u}$. In particular, a direct method" is considered in detail.

A.1.2 Successive Approximation Technique

As we have repeatedly stated, equations (5.1) through (5.8) are the characterization of $\underline{u}$ which solves the identification problem.[1]

---

1 The specific identification problem is stated in section A.1

It is worth recapitulating on the generation of these equations for the purpose of introducing a direct method for their solution.

Recall that the identification problem was phrased as the selection of $\underline{v} \varepsilon V$ which extremized a quadratic functional $J(\underline{v})$,

$$J(\underline{v}) : V \rightarrow R^1$$

By convention, we defined $\underline{u}$ implicitly as follows:

$$J(\underline{u}) = \operatorname*{Inf}_{\underline{v} \varepsilon V} J(\underline{v}) = \operatorname*{Inf}_{\underline{v} \varepsilon V} \{a(\underline{v},\underline{v}) - 2l(\underline{v}) + c\} \qquad (5.26)$$

Under appropriate hypothesis, the unique $\underline{u}$ is given by

$$a(\underline{u},\underline{v}) - l(\underline{v}) = 0 \quad \text{for all } \underline{v} \varepsilon V \qquad (5.27)$$

The determination of $\underline{u}$ given in (5.26) is nothing more than a mathematical programming problem in the Hilbert space V. There are several iterative schemes for the direct solution of (5.26), each of the following genre:

<u>PA: Programming Algorithm</u>

(i) Initial selection of $\underline{u}^0 \varepsilon V$

(ii) Compute "gradient" of $J(\underline{v})$ with respect to $\underline{v}$, evaluated at $\underline{u}^0$. Denote gradient by $\underline{G}$. (The term gradient will be defined presently)

(iii) If G is non zero, then update the initial choice of $\underline{u}$ according to:

$$\underline{u}^{i+1} = \underline{u}^i + \alpha^i \underline{s}^i \qquad i = 0,1,2...$$

where $\underline{s}^i$ is a meaningful modification of $\underline{G}$ and α^i is the magnitude of the step in the direction s^i on the surface $J(\underline{v})$. By meaningful modification of $\underline{G}$ we mean that after n iterations,

$$J(\underline{u}^n) \leq J(\underline{\bar{u}}^n), \quad \underline{\bar{u}}^0 = \underline{u}^0$$

where $\{\underline{\bar{u}}^n\}$ are generated by stepping along $\{G^n\}$ and $\{\underline{u}^n\}$ are generated by stepping along $\{s^n\}$.

The properties of this computational algorithm are considered in some detail presently. Suffice it to say here that $J(\underline{u}^i)$ ($i = 0,1,2...$) is a monotone decreasing sequence. Also

$$\lim_{i\to\infty} \underline{u}^i \to \underline{u} \varepsilon V$$

As a meaningful modification to $\underline{G}^i$ we chose $\underline{s}^i$ as the conjugate gradient [8,13,24]. Consider then, the following mathematical programming problem.

A.1.3 The Mathematical Programming Problem Hypothesis

Given a Quadratic Functional $J(\underline{v})$ on the Hilbert space V, $J(\underline{v}) : V \to R^1$

$$J(\underline{v}) = a(\underline{v},\underline{v}) - 2l(\underline{v}) + c$$

where

$a(\underline{v},\underline{v})$ is a continuous bilinear functional of $\underline{v}\varepsilon V$

$l(\underline{v})$ is a continuous linear functional of $\underline{v}\varepsilon V$

In addition, $a(\underline{v},\underline{v})$ is coercive (bounded from below):

$$a(\underline{v},\underline{v}) \geq \gamma||\underline{v}||_V^2 \quad ; \qquad \gamma > 0$$

<u>Problem</u>: select $\underline{u}\varepsilon V$ such that

$$J(\underline{u}) = \operatorname*{Inf}_{v\varepsilon V} J(\underline{v}) \tag{5.28}$$

As we have seen, (Theorem II.5), there is a unique $\underline{u}\varepsilon V$ with the property (5.28) and it is characterized by

$$a(\underline{u},\underline{v}) - l(\underline{v}) = 0 \tag{5.29}$$

For a given $\underline{u}\varepsilon V$, (5.29) is a continuous linear functional of $\underline{v}$ and thus has the representation

$$a(\underline{u},\underline{v}) - l(\underline{v}) = (\underline{G}(\underline{u}),\underline{v})_V \tag{5.30}$$

We show, in Appendix V.3, that G($\underline{u}$) is the variational derivative of the functional J($\underline{v}$) evaluated at $\underline{u}$. This is what we have previously called the gradient of the functional J(·).

We now specialize the programming algorithm for generating $\underline{u}$ presented earlier as follows:

Algorithm

(i) Select $\underline{u}^0 \varepsilon V$ (Initial guess)

(ii) Evaluate $G(\underline{u}^0)$. If $G(\underline{u}^0) = 0$, then by (5.29) and (5.30), $\underline{u}^0$ is the solution. If $G(\underline{u}^0) \neq 0$ then for the (i+1)st iteration, (i = 0,1,2...), proceed as follows:

(iii) $\underline{u}^{i+1} = \underline{u}^i + \alpha^i s^i$

$$\underline{s}^0 = -G(\underline{u}^0)$$

$$\underline{s}^{i+1} = -G(\underline{u}^{i+1}) + \beta^i \underline{s}^i \tag{5.31}$$

$$\beta^i = \frac{(G(\underline{u}^{i+1}), G(\underline{u}^{i+1}))_V}{(G(\underline{u}^i), G(\underline{u}^i))_V} \tag{5.32}$$

In addition, α^i is chosen so that

$$J(\underline{u}^i + \alpha^i \underline{s}^i) = \operatorname*{Inf}_{\beta^i \varepsilon R^1} J(\underline{u}^i + \beta^i \underline{s}^i) \tag{5.33}$$

It is possible to obtain an explicit expression for α^i.

$$\alpha^i = -\frac{a(\underline{s}^i, \underline{u}^i) - l(\underline{s}^i)}{a(\underline{s}^i, \underline{s}^i)} = \frac{(G(\underline{u}^i), G(\underline{u}^i))_V}{a(\underline{s}^i, \underline{s}^i)} \tag{5.34}$$

Before proceeding to show the connection of this technique with equations (5.1) through (5.8), we review some properties of the algorithm in the following theorems.

Theorem V.2

Given the hypothesis of section A.3.1, then if $\underline{G}(\underline{u}^i) \neq 0$, $J(\underline{u}^{i+1}) < J(\underline{u}^i)$. We prove theorem V.2 in Appendix V.2.

Corollary V.2

The sequence of real numbers $J(\underline{u}^i)$ is monotone decreasing and hence has a limit in the extended real numbers:

$$\lim_{i \to \infty} J(\underline{u}^i) = J_\infty = \operatorname*{Inf}_{\underline{v} \varepsilon V} J(\underline{v})$$

Theorem V.3 (convergence)

The sequence $\{\underline{u}^i\}$ converges weakly to a unique $\underline{u} \varepsilon V$ and the limit $\underline{u}$ has the property that

$$J(\underline{u}) = \operatorname*{Inf}_{\underline{v} \varepsilon V} J(\underline{v})$$

That is,

$$\lim_{i \to \infty} \underline{u}_i \overset{\text{weakly}}{\rightarrow} \underline{u} \varepsilon V \quad \text{unique}$$

$$J(\underline{u}) = J_\infty$$

Proof

$\underline{u}^i$ is a minimizing sequence and so theorem II.3 gives the result.

We have shown that the sequence $\{\underline{u}^i\}_{i=1,2...}$ generated by the algorithm converges (weakly) to a unique $\underline{u} \varepsilon V$ which realizes the infimum of $J(\underline{v})$.

For a discussion of the relative merits of the conjugate gradient direction of search compared with a steepest descent (or gradient technique), see [13].

It is now possible to display the connection between equations (5.1) through (5.8) with the programming technique presented. To be explicit, the functional $J(\underline{v})$ associated with (5.1) through (5.8) was given as

$$J(\underline{v}) = a(\underline{v},\underline{v}) - 2l(\underline{v}),$$

with

$$a(\underline{v},\underline{v}) = \int_Q (y(x,t;\underline{v})-y(x,t;\underline{0}))^2 dx\, dt$$

$$+\int_\Sigma (v_1(\Sigma)-z_1(\Sigma))^2 d\Sigma + \int_\Omega (v_2(x)-z_2(x))^2 dx$$

$$l(\underline{v}) = -\{\int_Q (y(x,t;\underline{v})-y(x,t;\underline{0}))(y(x,t;\underline{0})-z_0)dx\, dt$$

$$-\int_\Sigma u_1(\Sigma)z_1(\Sigma)d\Sigma - \int_\Omega u_2(x)z_2(x)dx\}$$

We saw that the characterization of $\underline{u}$ given by

$$a(\underline{u},\underline{v}) - l(\underline{v}) = 0 \qquad \text{for all } \underline{v}\varepsilon V \tag{5.35}$$

could be expressed in terms of an adjoint variable $p(x,t;\underline{u})$ (which was introduced for the sole purpose of obtaining such an expression). The expression equivalent to (5.35) was shown to be (equations (4.13), (4.14)).

$$(G(\underline{u}),\underline{v})_V = 0 = \int_\Sigma [-\frac{\partial p(\Sigma)}{\partial \nu_{A*}} + u_1(\Sigma) - z_1(\Sigma)] v_1(\Sigma) d\Sigma$$

$$+ \int_\Omega [p(x,0) + u_2(x) - z_2(x)] v_2(x) dx$$

From which we recover

$$\underline{G}(\underline{u}^i) = \begin{bmatrix} -\frac{\partial p(\Sigma;\underline{u}^i)}{\partial \nu_{A*}} + u_1^i - z_1^i \\ p(x,0;\underline{u}^i) + u_2^i - z_2^i \end{bmatrix} \quad \begin{matrix} (5.36) \\ (5.37) \end{matrix}$$

$p(x,t;\underline{u}^i)$ evolves according to

$$-\frac{\partial p(x,t;\underline{u}^i)}{\partial t} + A[p(x,t;\underline{u}^i)] = y(x,t;\underline{u}^i) - z(x,t) \qquad (5.38)$$

$$p(\Sigma) = 0 \qquad (5.39)$$

$$p(x,T) = 0 \qquad (5.40)$$

and $y(x,t;\underline{u}^i)$ evolves according to

$$\frac{\partial y(x,t;\underline{u}^i)}{\partial t} + A[y(x,t;\underline{u}^i)] = f(x,t) \qquad (5.41)$$

$$y(\Sigma) = u_1^i \qquad (5.42)$$

$$y(x,0) = u_2^i \qquad (5.43)$$

(5.43) through (5.36) are equivalent to (5.1) through (5.8) with $\underline{u}$ in the latter replaced by $\underline{u}^i$ in the former.

Thus the algorithm would proceed as follows:

(i) Select an initial $\underline{u}^0$

(ii) Compute $\underline{G}(\underline{u}^0)$. To do so, we first compute $y(\cdot,\cdot;\underline{u}^0)$, starting from $u_2^0(x,0)$, evolving according to (5.44), under the influence (on the boundary Σ) of $u_1^0(\Sigma)$. We remark that this is the numerically stable direction of solution. Having obtained $y(\cdot,\cdot;\underline{u}^0)$, it is possible to generate $p(\cdot,\cdot;\underline{u}^0)$, starting from $p(x,T) = 0$, evolving backwards in time according to (5.41). Again, we remark that this is the numerically stable direction of solution. Thus $p(\cdot,\cdot;\underline{u}^i)$ has been recovered. It is now possible to evaluate $G(\underline{u}^0)$ given by (5.36) and (5.37).

(iii) Using $G(\underline{u}^0)$, compute

$$\underline{u}^1 = \underline{u}^0 + \alpha^0 G(\underline{u}^0)$$

where

$$\alpha^0 = \frac{(\underline{G}(\underline{u}^0),\underline{G}(\underline{u}^0))_V}{a(G(\underline{u}^0),G(\underline{u}^0))}$$

Explicitly,

$$(\underline{G}(\underline{u}^0),G(\underline{u}^0))_V = \int_\Sigma [-\frac{\partial p}{\partial \nu_A}(\Sigma;\underline{u}^0)+u_1^0(\Sigma)-z_1(\Sigma)]^2 d\Sigma$$

$$+\int_\Omega [p(x,0;\underline{u}^0)+u_2^0(x)-z_2(x)]^2 dx$$

$$a(\underline{G}(\underline{u}^0),G(\underline{u}^0))=\int_Q [y(x,t;\underline{G}(\underline{u}^0))-y(x,t;\underline{0})]^2 dx\, dt$$

$$+ (\underline{G}(\underline{u}^0),\underline{G}(\underline{u}^0))_V$$

The (i+1)st iteration proceeds as previously outlined (i=1,2,...):

(ii)' Evaluate $\underline{G}(\underline{u}^{i+1})$ in a way analogous to that outlined in (ii)

(iii)' Using $\underline{G}(\underline{u}^{i+1})$, compute:

$$\underline{u}^{i+1} = \underline{u}^{i} + \alpha^{i}\underline{s}^{i}$$

$$\alpha^{i} = \frac{(\underline{G}(\underline{u}^{i}),\underline{G}(\underline{u}^{i}))_V}{a(\underline{s}^{i},\underline{s}^{i})}$$

$$\underline{s}^{i+1} = -\underline{G}(\underline{u}^{i+1}) + \beta^{i}\underline{s}^{i}$$

$$\beta^{i} = \frac{(\underline{G}(\underline{u}^{i+1}),\underline{G}(\underline{u}^{i+1}))_V}{(G(\underline{u}^{i}),G(\underline{u}^{i}))_V}$$

The explicit representation formulas for the quantities α^i β^i, etc. are now obvious.

The iteration continues until $(G(\underline{u}^i),G(\underline{u}^i))_V$ is acceptably close to zero.

In sections A.2 and A.3, we have outlined two distinct methods for obtaining the solution to the simultaneous equations characterizing $\underline{u}\varepsilon V$ for a specific Identification problem. That problem was, as we saw, concerned with obtaining $\underline{u}\varepsilon V$ which minimized the quadratic error functional $J(\underline{v})$ along a system trajectory $y(\cdot,\cdot\ \underline{v})$, where $y(x,t;\underline{v})$ satisfied a parabolic evolution equation with Dirichlet boundary conditions. The other (non-trivial) identification problems treated in Chapter IV can be solved by the methods of sections A.2

and A.3 of this chapter, the results being different in detail only. However, the following remarks concerning the direct method of section A.3 are in order.

(i) The response of the parabolic system s, denoted by $y(\cdot,\cdot;\underline{v})$, is stable - in the sense that the homogeneous solution starting from a non zero initial condition $u_2(x,0)$ at $t = 0$ approaches zero as time increases.

(ii) The response of the adjoint parabolic system s^*; denoted by $p(\cdot,\cdot;\underline{v})$, <u>is unstable</u> in the sense considered in (i). However, the response $p(\cdot,\cdot;\underline{v})$ is stable in reverse time - the homogeneous solution starting from a non-zero terminal condition $p(x,T)$ at $t = T$ approaches zero <u>as time decreases</u>.

(iii) For the second order hyperbolic systems, the response of the system s or its adjoint s* is marginally stable in either time direction. (For a particular $x\varepsilon\Omega$, $y(x,\cdot;\underline{v})$ and $p(x,\cdot;\underline{v})$ are a sum of harmonic functions in time).

With these remarks, we observe that the numerical computation of the algorithm posed in section A.3 is always stable, since s is solved forward in time and s* is solved backwards in time.

The numerical solution of the Integro-partial and partial differential equations of sections A.2 and A.3 respectively is accomplished by the approximation technique of Galerkin, the application of which is now considered.

A.2 <u>The Galerkin Technique</u>

We address ourselves at first to the solution of equations (5.13) through (5.18), repeated here for convenience:

$$\frac{\partial P}{\partial t}(x,\xi,t) - A_\xi[P(x,\xi,t)] - A_x[P(x,\xi,t)] - \delta(\xi - x)$$

$$+\int_{\Gamma_s} \frac{\partial P(x,\Sigma_s)}{\partial \nu_{A_s^*}} \frac{P(\Sigma_s,\xi)}{\partial \nu_{A_s^*}} d\Gamma_s = 0 \quad \text{in } \Omega\times\Omega\times(0,T] \tag{5.13}$$

$$P(x,\Sigma_s) = P(\Sigma_s,\xi) = 0 \qquad \text{on } \Sigma_s\times\Omega \tag{5.14}$$

$$P(x,\xi,0) = \delta(x-\xi) \qquad \text{in } \Omega\times\Omega \tag{5.15}$$

$$\int_\Omega P(x,\xi,t)\{\frac{\partial \hat{y}}{\partial t}(\xi,t) + A_\xi[\hat{y}(\xi,t)] - f(\xi,t)\}dt = z(x,t) - \hat{y}(x,t)$$

$$\text{in } \Omega\times(0,T] \tag{5.16}$$

$$\hat{y}(\Sigma) = z_1(\Sigma) \qquad \text{on } \Sigma \tag{5.17}$$

$$\hat{y}(x,0) = z_2(x) \qquad \text{in } \Omega \tag{5.18}$$

By Hypothesis (b) of theorem V.1, $P(\cdot,\cdot\ t)\varepsilon H^2(\Omega\times\Omega)$ $P(\cdot,\cdot\ t)\varepsilon L^2(\Omega\times\Omega)$. Thus there exists an approximation to $P(x,\xi,t)$, denoted by $P_m(x,\xi,t)$ given by

$$P_m(x,\xi,t) = \sum_{i=j=1}^{m} P_{ij}(t) w_i(x) w_j(\xi) \tag{5.44}$$

where $\{w_i(\cdot)w_j(\cdot)\}_{i,j=1,2,\ldots}$ is an orthonormalized basis in $L^2(\Omega\times\Omega)$ $P_m(\cdot,\cdot,t)$ has the property that for each $t\varepsilon(0,T]$

$$\lim_{m\to\infty} P_m(\cdot,\cdot,t) \to P(\cdot,\cdot,t) \quad \text{in} \quad L^2(\Omega\times\Omega).$$

In particular, we choose the orthonormal basis $\{w_i(\cdot)w_j(\cdot)\}_{i,j=1,2,\ldots}$ afforded by the normalized eigenfunctions which are generated as solutions to

$$A_x[w_i(x)]+A_\xi[w_j(\xi)] - (\lambda_i+\lambda_j)w_i(x)w_j(\xi) = 0$$

$$w_i(\Gamma_x) = 0$$

$$w_j(\Gamma_\xi) = 0 \qquad i,j = 1,2,\ldots$$

Define

$$\Psi(x,\xi,t) = w_k(x)w_l(\xi)g(t) \quad k,l \text{ fixed but arbitrary}$$

$$g(t)\varepsilon c^1, \quad g(T) = 0.$$

An equation defining the coefficients $P_{ij}(t)$ of (5.44) can be obtained via the Galerkin approximation scheme. We proceed as follows:

Multiply equation (5.13) by $\Psi(x,\xi,t)$ and integrate over $(\Omega\times\Omega)\times(0,T]$. There results:

$$\int_0^T\{-\frac{dg(t)}{dt}P_{kl}(t)-(\lambda_k+\lambda_l)P_{kl}(t)g(t)-\Delta_{kl}g(t)$$

$$+\int_{\Gamma_s}[\sum_{j=1}^{m}P_{kj}(t)\frac{\partial w_j(\Gamma_s)}{\partial\nu_{A_s^*}}][\sum_{i=1}^{m}P_{il}(t)\frac{\partial w_i(\Gamma_s)}{\partial\nu_{A_s^*}}]g(t)d\Gamma_s\}dt$$

$$-\ g(0)\Delta_{kl} = 0 \qquad (5.45)$$

where

$$\Delta_{kl} = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{otherwise} \end{cases}$$

Integrate the first term of (5.4) by parts and collect terms in $g(t)$. Owing to the arbitrariness of $g(t)$ and the fact that $g(t)\varepsilon c^1$, we obtain an equation for $P_{kl}(t)$ $(k,l = 1,2,\ldots m)$:

$$\frac{dP_{kl}}{dt}-(\lambda_k+\lambda_l)P_{kl}(t)-\Delta_{kl}+\int_{\Gamma_s}\Big[\sum_{j=1}^{m}P_{kj}(t)\frac{\partial w_j(\Gamma_s)}{\partial \nu_{A_s^*}}\Big]\Big[\sum_{i=1}^{m}P_{il}(t)\frac{\partial w_i(\Gamma_s)}{\partial \nu_{A_s^*}}\Big]d\Gamma_s=0 \tag{5.46}$$

$$P_{kl}(0) = \Delta_{kl} \tag{5.47}$$

Equation (5.46) can be solved by any of the many numerical integration schemes available, starting with the initial condition (5.47). Thus, using (5.44), we can obtain $P_m(\cdot,\cdot,\cdot)$ $x,\xi,t \in \Omega\times\Omega\times[0,T]$.

Equation (5.16) is a linear integral equation of the second kind. Having obtained the solution $P_m(x,\xi,t)$, (5.16) can be solved by the successive substitution or successive approximation methods [19]. Rather than proceed in that manner however, (because we seek an equation for the evolution of $\hat{y}_i(t)$, not $\hat{y}_i(t)$), consider the following approach: Multiply (5.16) by $g(t)$ and integrate over $(0,T]$ to obtain:

$$\int_Q\{-\frac{\partial}{\partial t}[P(x,\xi,t)g(t)]+g(t)A_\xi[P(x,\xi,t)]\}\hat{y}(\xi,t)d\xi\ dt$$

$$-\int_Q f(\xi,t)P(x,\xi,t)g(t)d\xi\ dt-\int_\Omega P(x,\xi,0)\hat{y}(\xi,0)g(0)d\xi$$

$$+\int_{\Sigma_s}[y(\Sigma_s)\frac{\partial P(x,\Sigma_s)}{\partial\nu_{A_s^*}}g(t)d\Sigma_s=\int_0^T[z(x,t)-\hat{y}(x,t)]dt \tag{5.48}$$

We seek an approximate solution to (5.48) denoted by $\hat{y}_m(x,t)$, where as before

$$y_m(x,t)=\sum_{i=1}^{m}\hat{y}_i(t)w_i(x);\ \lim_{m\to\infty}\hat{y}_m(\cdot,\cdot)\to\hat{y}(\cdot,\cdot)\text{ in }L^2(Q)$$

Substituting the expression for $y_m(x,t)$ into (5.48) along with the definition of $P_m(x,\xi,t)$ given by (5.44) yields:

$$\int_Q\{-\sum_{k,l=1}^{m}\frac{\partial}{\partial t}[P_{kl}(t)g(t)]w_k(x)w_l(\xi)\sum_{i=1}^{m}\hat{y}_i(t)w_i(\xi)$$

$$+g(t)\sum_{i=1}^{m}y_i(t)w_i(\xi)\sum_{k,l=1}^{m}P_{kl}(t)w_k(x)A[w_l(\xi)]$$

$$-f(\xi,t)g(t)\sum_{k,l=1}^{m}P_{kl}(t)w_k(x)w_l(\xi)\}d\xi\ dt\ -\ z_2(x)g(0)$$

$$+\int_{\Sigma_s}[z_1(\Sigma_s)\sum_{k,l=1}^{m}P_{kl}(t)w_k(x)\frac{\partial w_l}{\partial\nu_{A_s^*}}(\Gamma_s)g(t)d\Sigma_s$$

$$=\int_0^T[z(x,t)-\hat{y}(x,t)]dt \tag{5.49}$$

Using the orthonormal properties of the set $\{w_i(\cdot)\}_{i=1,2,...}$ and the intentioned arbitrariness of $g(\cdot)$ in $c^1[0,T)$, then (5.49) is equivalent to:

$$\sum_{k,l=1}^{m} P_{kl}(t)w_k(x)[\frac{d\hat{y}_l}{dt} + \lambda_l\hat{y}_l(t) - f_l(t) + z_{1i}(t)] = \hat{y}(x,t) - z(x,t) \tag{5.50}$$

$$-z_2(x) + \sum_{k,l=1}^{m} P_{kl}(0)\hat{y}_l(0)w_k(x) = 0 \tag{5.51}$$

where

$$z_{1i}(t) = \int_{\Gamma_s} z_1(\Sigma_s)\frac{\partial w_i}{\partial \nu_{A_s^*}}\, d\Gamma_s$$

$$f_i(t) = \int_{\Omega} f(x,t)w_i(x)dx$$

Before proceeding further with (5.50) and (5.51), define the following matrix and vector guantities:

$$P = \{P_{ij}(t)\}_{i=j=1,2,...m}$$

$$A = \begin{bmatrix} \lambda_1 & 0 & ---- & 0 & 0 \\ 0 & \lambda_2 & ---- & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & ---- & \lambda_{m-1} & 0 \\ 0 & 0 & ---- & 0 & \lambda_m \end{bmatrix}$$

$$\underline{f} = [f_1(t)\ f_2(t)\ \ldots\ f_m(t)]^T$$

$$z_1 = [z_{11}(t)\ z_{12}(t)\ \dots\ z_{1m}(t)]^T$$

$$z_2 = [z_{21}\ z_{22}\ \dots\ z_{2m}]^T$$

$$z = [z_1(t)\ z_2(t)\ \dots\ z_m(t)];\ z_i = \int_\Omega z(x,t)w_i(x)dx$$

$$WW^T = \int_{\Gamma_s}\left\{\begin{bmatrix} \frac{\partial w_1(\Gamma_s)}{\partial \nu_{A^*}} & \frac{\partial w_1(\Gamma_s)}{\partial \nu_{A^*}} & \cdots & \frac{\partial w_1(\Gamma_s)}{\partial \nu_{A^*}} & \frac{\partial w_m(\Gamma_s)}{\partial \nu_{A^*}} \\ \vdots & \vdots & & \vdots & \vdots \\ \frac{\partial w_m(\Gamma_s)}{\partial \nu_{A^*}} & \frac{\partial w_1(\Gamma_s)}{\partial \nu_{A^*}} & \cdots & \frac{\partial w_m(\Gamma_s)}{\partial \nu_{A^*}} & \frac{\partial w_m(\Gamma_s)}{\partial \nu_{A^*}} \end{bmatrix}\right\} d\Gamma_s$$

Then if equations (5.50) and (5.51) are each multiplied by $w_i(x)$ and the results are integrated over Ω, the following equations, given in matrix-vector form, result:

$$P(t)[\frac{d\hat{y}(t)}{dt}+A\hat{y}(t)-f(t)+z_1(t)] = z(t)-\hat{y}(t) \tag{5.52}$$

$$\hat{y}(0) = z_2 \tag{5.53}$$

For completeness, we give the matrix representation for $\{P_{ij}(t)\}$ satisfying (5.46) and (5.47):

$$\frac{dP(t)}{dt}-AP(t)-P(t)A-I+P(t)WW^TP(t) = 0 \tag{5.54}$$

$$P(0) = I \tag{5.55}$$

In order to solve (5.52) for y(t), we need an equation for $P^{-1}(t)$. Consider the following identity:

$$\frac{d}{dt}(P^{-1}(t)) = -P^{-1}(t)\frac{dP(t)}{dt}P^{-1}(t)$$

Formally then, pre and post multiplication of (5.54) gives an equation for the evolution of $P^{-1}(t)$:

$$\frac{d}{dt}(P^{-1}(t))+P^{-1}(t)A+AP^{-1}(t)+P^{-1}(t)IP^{-1}(t)-WW^T=0 \tag{5.56}$$

$$P^{-1}(0) = I \tag{5.57}$$

The existence of an unique solution to (5.56) and (5.57) is given by the following theorem due to Kalman [10].

<u>Theorem V.4 (Kalman)</u>

If $P^{-1}(0)$ is positive definite, then $P^{-1}(t)$ exists for all $t > 0$ and is the unique solution of (5.56) having the initial value $P^{-1}(0)$ given by (5.57).

Consequently, since $P^{-1}(0) = I > 0$, a solution to (5.56) exists and is unique. Thus we can write equation (5.55) as follows:

$$\frac{d\hat{y}(t)}{dt}+A\hat{y}(t)-f(t)+z_1(t) = P^{-1}(t)[z(t)-\hat{y}(t)] \tag{5.63}$$

$$\hat{y}(0) = z_2 \tag{5.64}$$

rewriting (5.56) and (5.57) here for convenience, $P^{-1}(t)$ is given by:

$$\frac{d}{dt}(P^{-1}(t))+P^{-1}(t)A+AP^{-1}(t)+P^{-1}(t)IP^{-1}(t)-WW^T=0 \tag{5.65}$$

$$P(0)^{-1} = 1 \tag{5.66}$$

$\hat{y}_m(x,t)$ is obtained by solving (5.63) and (5.64) by a numerical integration scheme, thereby obtaining $\hat{y}_i(t)$, and

$$\hat{y}_m(x,t) = \sum_{i=1}^{m} \hat{y}_i(t)w_i(x).$$

The numerical technique for solving the equations arising in section A.2 has been demonstrated. In summary,the partial differential and integral equations were solved by approximation of the solutions in terms of eigenfunction expansions. These expansions were carried out according to the technique of Galerkin (Chapter II section A.3). The expansion led to the consideration of a system of ordinary differential equations, given by (5.54), (5.58), (5,63) and (5.66). We established as a subsidiary result that the system (5.56) and (5.57) had a unique solution hence was equivalent to the system (5.54) and (5.55).

The equations arising out of the "direct method" of section A.3 can be solved using the Galerkin approximation technique. The method is straightforward and the results can be outlined.

(a) As before, we have the existence of $y_m(x,t;\underline{u}^k)$ and $p_m(x,t;\underline{u}^k)$, where

$$y_m(x,t;\underline{u}^k) = \sum_{i=1}^{m} y_i(t;\underline{u}^k)w_i(x)$$

$$p_m(x,t;\underline{u}^k) = \sum_{i=1}^{m} p_i(t;\underline{u}^k)w_i(x)$$

$$\lim_{m\to\infty} y_m(x,t;\underline{u}^k) \to y(x,t;\underline{u}^k)$$

$$\lim_{m\to\infty} p_m(x,t;\underline{u}^k) \to p(x,t;\underline{u}^k).$$

(b) $y_i(t;\underline{u}^k)$ and $p_i(t;\underline{u}^k)$ are defined, according to Galerkin by:

$$\frac{dy_i}{dt}(t;\underline{u}^k) + \lambda_i y_i(t;\underline{u}^k) = f_i(t) - u_{1i}^k(t)$$

$$y_i(0) = u_{2i}^k$$

$$-\frac{dp_i}{dt}(t;\underline{u}^k) + \lambda_i p_i(t;\underline{u}^k) = y_i(t;\underline{u}^k) - z_i(t)$$

$$p_i(T) = u_{2i}^k$$

with

$$f_i(t) = \int_\Omega f(x,t) w_i(x) dx$$

$$z_i(t) = \int_\Omega z(x,t) w_i(x) dx$$

$$u_{1i}(t) = \int_\Omega z_1(\Sigma) \frac{\partial w_i(\Gamma)}{\partial \nu_{A*}} d\Gamma$$

$$u_{2i} = \int_\Omega z_2(x) w_i(x) dx \ .$$

(c) The gradient vector $\underline{G}(\underline{u}^k)$ is given by:

$$G(\underline{u}^k) = \begin{bmatrix} -\sum_{i=1}^{m} p_i(t)\frac{\partial w_i(\Gamma)}{\partial \nu_{A^*}}+u_2(\Sigma)-z_1(\Sigma) \\ \sum_{i=1}^{m} p_i(0)w_i(x)+u_2(x)-z_2(x) \end{bmatrix}$$

The algorithm outlined in section A.3 is carried out with y(x,t) and p(x,t) replaced by $y_m(x,t)$ and $p_m(x,t)$. The specific technique is best illustrated by the example considered in section A.5. In that section, two illustrative examples are considered, each of which are solved by the methods described in section A.2 and A.3.

The problem which was posed in section A.1 was the identification problem associated with the parabolic system with Dirichlet boundary conditions and quadratic functional PRA defined in Chapter IV. One of the illustrative examples of section A.5 is concerned with the same system and boundary conditions, but with the quadratic functional of PRB, Chapter IV. That is, the case of pointwise output measurements in the spatial domain. We give in [21], a theorem similar in content to Theorem V.1, which deals with the "discrete measurement" case. The resulting equations, analagous to (5.63) through (5.66) are:

$$\frac{d\hat{y}}{dt}(t)+A\hat{y}(t)-f(t)+z_1(t) = P^{-1}(t)Q_\nu(z(t)-\hat{y}(t)) \tag{5.67}$$

$$\hat{y}(0) = z_2 \tag{5.68}$$

$$\frac{d}{dt}(P^{-1}(t))+P^{-1}(t)A+AP^{-1}(t)+P^{-1}(t)Q_\nu P^{-1}(t)-WW^T=0 \tag{5.69}$$

$$P^{-1}(0) = I \tag{5.70}$$

where the matrices and vectors appearing in (5.67) through (5.70) are as previously defined with the exception of Q_ν, defined below:

$$Q_\nu = \sum_{i=1}^{\nu} Q^i \ ; \quad Q^i = \{q^i_{kl}\}_{k,l=1,2,\ldots m}$$

$$q^i_{kl} = w_k(x^i)w_l(x^i)$$

<u>Remark</u>

(i) The difference in the equations for $\hat{y}(x,t)$ (the case of continuous interior measurements and discrete interior measurements), is encapsulated by a comparison of (5.63) and (5.64) with (5.67) and (5.68) respectively. In the latter case, the feedback error term on the R.H.S. of (5.81) reflects a weighted sum of errors arising at the discrete measurement locations x^i, $(i=1,2,...\nu)$.

The numerical solution of two illustrative examples is considered next. These examples serve to consolidate the theoretical results.

A.3 Numerical Examples

Two numerical examples serve as vehicles for the display of some salient features of the solution technique presented in sections A.2 and A.3. In each case, the system is of parabolic type with Dirichlet boundary conditions. The point of departure between the examples is in the definition of the output measurement process, and in consequence a direct comparison between the resulting refined estimates is afforded. In the first case, output measurements are taken for each $t\varepsilon(0,T]$, over the entire spatial profile Ω. In the second case, output measurements are taken for each $t\varepsilon(0,T]$ at selected[1] points along the spatial profile, denoted by x^i, $x^i\varepsilon\Omega$; $(i=1,2,...\nu)$.

Consider then the following two examples in turn.

Example 1

System S_D

$$\frac{\partial \underline{y}(x,t;\underline{u})}{\partial t} - \frac{\partial^2 \underline{y}(x,t)}{\partial x^2} = 212.0 \qquad x,t\varepsilon(0,1)\times(0,1] \qquad (5.71)$$

---

1 In the simulation, the selection was arbitrary.

$$y(0,t) = u_1(0,t) \qquad t\varepsilon(0,T] \qquad (5.72)$$

$$y(1,t) = u_1(1,t) \qquad t\varepsilon(0,T] \qquad (5.73)$$

$$y(x,0) = u_2(x) \qquad x\varepsilon(0,1) \qquad (5.74)$$

Input Measurements I

$$z_1(0,t) = u_1^*(0,t)+k_1N_1(t) \qquad t\varepsilon(0,1] \qquad (5.75)$$

$$z_1(1,t) = u_1^*(1,t)+k_2N_2(t) \qquad t\varepsilon(0,1] \qquad (5.76)$$

$$z_2(x) = u_2^*(x) + k_3 \qquad x\varepsilon(0,1) \qquad (5.77)$$

where $\underline{u}^*$ is the true state of nature, and is defined by:

$$u_1^*(0,t) = 70+10\sin 2\Pi t \qquad (5.78)$$

$$u_1^*(1,t) = 54.5 \qquad (5.79)$$

$$u_2^*(x) = 70e^{-0.25x} \qquad (5.80)$$

and

$N_1(t)$, $N_2(t)$ are purely random functions of time, each independent of the other with amplitude ±1.0. k_1 and k_2 are variables chosen to affect the signal to noise ratio. k_3 is an arbitrary bias on the initial condition, also chosen to affect the signal to noise ratio.

Evidently, from (5.75) through (5.80) and the fact that $k_1N_1(t)$ and $k_2N_2(t)$ are amplitude limited,

$$z_1(\cdot)\varepsilon L^2(\Sigma)$$

$$z_2(\cdot)\varepsilon L^2(\Omega)$$

Output Measurements 0

$$z(x,t) = y(x,t;\underline{u}^*)+k_0N_0(t) \qquad x,t\varepsilon(0,1)\times(0,1] \qquad (5.81)$$

where

$y(x,t;\underline{u}^*)$ is the response of the system S_D at the point $x,t\varepsilon(0,1)\times(0,1]$ which evolves according to (5.73) with initial condition $u_2^*(x)$ and boundary conditions $u_1^*(0,t)$ and $u_1^*(1,t)$. $N_0(t)$ is a purely random function of time with amplitude ±1.0. The set $\{N_0(t), N_1(t), N_2(t)\}$ are pairwise independent for all $t\varepsilon(0,1]$. k_0 is a variable which alters the signal to noise ratio.

The problem is to choose $\underline{u}\varepsilon V$ which extremizes an error functional $J(\underline{v})$, $\underline{v}\varepsilon V$.

Define

$$V = L^2(\Sigma)\varepsilon L^2(\Omega)$$

$$J(\underline{v}) = \int_0^1\int_0^1[y(x,t;\underline{v})-z(x,t)]^2dx\,dt$$

$$+ \int_0^1[v_1(0,t)-z_1(0,t)]^2dt + \int_0^1[v_1(1,t)-z_1(1,t)]^2dt$$

$$+ \int_0^1[v_2(x)-z_2(x)]^2dx. \qquad (5.82)$$

As we have seen (theorem IV.1 IA), the unique $\underline{u}\varepsilon V$ which extremizes $J(\underline{v})$ is characterized by:

$$\frac{\partial y}{\partial t}(x,t;\underline{u}) - \frac{\partial^2 y}{\partial x^2} = 212.0 \qquad x,t\varepsilon(0,1)\times(0,1] \qquad (5.83)$$

$$y(0,t) = u_1(0,t) \qquad t\varepsilon(0,1] \tag{5.84}$$

$$y(1,t) = u_1(1,t) \qquad t\varepsilon(0,1] \tag{5.85}$$

$$y(x,0) = u_2(x) \qquad x\varepsilon(0,1) \tag{5.86}$$

$$\frac{\partial p}{\partial t}(x,t;\underline{u}) - \frac{\partial^2 p}{\partial x^2}(x,t;\underline{u}) = y(x,t;\underline{u}) - z(x,t) \tag{5.87}$$

$$p(0,t) = 0 \tag{5.88}$$

$$p(1,t) = 0 \tag{5.89}$$

$$p(x,1) = 0 \tag{5.90}$$

$$\frac{\partial p}{\partial x}(0,t;\underline{u}) + u_1(0,t) - z_1(0,t) = 0 \qquad t\varepsilon(0,1] \tag{5.91}$$

$$-\frac{\partial p}{\partial x}(1,t;\underline{u}) + u_1(1,t) - z_1(1,t) = 0 \qquad t\varepsilon(0,1] \tag{5.92}$$

$$p(x,0;\underline{u}) + u_2(x) - z_2(x) = 0 \qquad x\varepsilon(0,1) \tag{5.93}$$

The system state $y(x,t;\underline{u})$ and its adjoint $p(x,t;\underline{u})$ were approximated, in accordance with the developments of section A.4, by:

$$y_m(x,t;\underline{u}) = \sum_{i=1}^{4} y_i(t;\underline{u}) w_i(x) \tag{5.94}$$

$$p_m(x,t;\underline{u}) = \sum_{i=1}^{4} p_i(t;\underline{u}) w_i(x) \tag{5.95}$$

$$\hat{y}_m(x,t) = \sum_{i=1}^{8} \hat{y}_i(t)w_i(x) \tag{5.96}$$

$$w_i(x) = \sqrt{2} \quad \sin(\sqrt{\lambda_i})x \; ; \; \lambda_i = (i\pi)^2 \tag{5.97}$$

Remark

$\{\sqrt{2} \quad \sin(i\pi)(\cdot)\}_{i=1,2...}$ are complete in $L^2(\Omega)$

$$\int_0^1 \sqrt{2} \quad \sin(i\pi)x \ \sqrt{2} \quad \sin(j\pi)x \ dx = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

Thus, in the notation of section A.4, (See [21] for details):

$$f_i(t) = \frac{212\sqrt{2}}{(i\pi)} [1.0-(-1.0)^i] \tag{5.98}$$

$$z_{1i}(t) = \sqrt{2}(i\pi)\{-[70+10\sin(2\pi)t+k_1N_1(t)] +(-1.0)^i[54.5+k_2N_2(t)]\} \tag{5.99}$$

$$z_{2i} \doteq \frac{70\sqrt{2}}{(i\pi)}[1.0-e^{-0.25}(-1.0)^i]+ \frac{k_3\sqrt{2}}{(i\pi)}[1.0-(-1.0)^i] \tag{5.100}$$

$$z_i(t) = y_i(t;\underline{u}^*)+ \frac{k_0N_0(t)\sqrt{2}}{(i\pi)} [1.0-(-1.0)^i] \tag{5.101}$$

$$y_i(t;\underline{u}^*) = \frac{70\sqrt{2}}{(i\pi)} [1.0-e^{-0.25}(-1.0)^i]e^{-(i\pi)^2t} \tag{5.102}$$

$$+ \{\frac{212\sqrt{2}}{(i\pi)^3} [1.0-(-1.0)^i]+ \frac{70\sqrt{2}}{(i\pi)} - \frac{54.5(-1.0)^i\sqrt{2}}{(i\pi)}\}\{1.0-e^{-(i\pi)^2t}\}$$

$$+ \frac{10\sqrt{2}\,(i\pi)}{(i\pi)^4+4\pi^2}\left[(i\pi)^2\sin 2\pi t-2\pi\cos 2\pi t+2\pi e^{-(i\pi)^2 t}\right] \qquad (5.103)$$

$$\{WW^T\}_{k,l} = 2\pi^2 kl[1.0+(-1)^k(-1)^l] \qquad (5.104)$$

Using the identities (5.83) through (5.104), we obtained results for the following, via the developments of section A.2:

(a) $\hat{y}_i(t)$, $\{P_{ij}(t)\}^{-1}$ $(i,j=1,2,3,4)$ evaluated for ($t = 0+\Delta t$, $0+2\Delta t,\ldots.1.0$). $\Delta t = 0.001$ from equations (5.63), (5.64), (5.65) and (5.66).

(b) $\hat{y}_m(x,t)$ from equation (5. 95), with ($x=0+\Delta x$, $0+2\Delta x,\ldots 1$), $\Delta x = 0.01$.

(c) Using the algorithm of section A.3, we recovered $u_1(0,t)$, $u_1(1,t)$ for ($t = 0+\Delta t$, $0+2\Delta t,\ldots 1$), $\Delta t = 0.01$.

(d) $u_2(x)$ for ($x = 0+\Delta x$, $0+2\Delta x,\ldots 1$);$\Delta x = 0.01$.

(e) $y_i(t;\underline{u})$ for ($t = 0+\Delta t$, $0+2\Delta t,\ldots 1$); $\Delta t = 0.01$.

(f) $y_m(x,t;\underline{u})$ for $x = (0+\Delta x$, $0+2\Delta x,\ldots 1)$; $\Delta x = 0.01$.

A discussion on these results is delayed until after consideration of the second numerical example, in order that comparisons may be made.

Example 2

The system S_D is the same as that of example 1, evolving according to (5.71) through (5.74). The input measurements are also the same as those in example 1, given by equations (5.75) through (5.80). As we indicated earlier, the point of departure was in the definition of the output measurements:

Output Measurements 0

Measurements were taken at four points in the spatial domain, for each t. The four points were chosen at (x = 0.2, x = 0.4, x = 0.6, x = 0.8). That is,

$$z(x^i,t) = y(x^i,t;\underline{u}^*)+k_0^iN_0^i(t), \quad (i=1,2,3,4) \tag{5.105}$$
$$x^1 = 0.2,\ x^2 = 0.4,\ x^3 = 0.6,\ x^4 = 0.8$$

$y(x^i,t;\underline{u}^*)$ is the response of S_D to initial and boundary conditions $\underline{u}^*$ (defined by (5.78) through (5.80)) at the selected points x^i. For each i, k_0^i and $N_0^i(t)$ are as defined in example 1.

As before, the problem is to choose $\underline{u}\varepsilon V$ which extremizes an error functional $J(\underline{v})$, $\underline{v}\varepsilon V$. Here, we define the following:

$$V = L^2(\Sigma)\times L^2(\Omega)$$

$$J(\underline{v}) = \int_0^1 \sum_{i=1}^4 [y(x^i,t;\underline{v})-z(x^i,t)]^2dt$$
$$+ \int_0^1 [v_1(0,t)-z_1(0,t)]^2dt + \int_0^1 [v_1(1,t)-z_1(1,t)]^2dt$$
$$+ \int_0^1 [v_2(x)-z_2(x)]^2dx \tag{5.106}$$

Remark

The output measurements O for example 2 are taken at selected points $x^i \varepsilon \Omega$. Supposedly, we have access to the system only at the points x^i. Note however that one of the "input" measurements, namely $z_2(x)$ is assumed to be available over the entire spatial profile, an apparent inconsistency. First, let it be said that the assumption that $z_2(x)$ is available for all $x \varepsilon \Omega$ is essential to the theoretical framework. That framework is, in abstract, the minimization of quadratic functionals on a Hilbert space V. As we saw, V was chosen to be $L^2(\Sigma)\times L^2(\Omega)$. Discrete spatial measurements on the boundary would lead to spaces V such as:

$$V = L^2(\Sigma)\times l_1^2 \times l_2^2 \times \ldots \times l_\nu^2$$

where l_i^2 is the space of squared real numbers. No consideration of this type of space is made. In fact, such a consideration would be extraordinarily difficult. However, we can rationalize this apparent inconsistency. We do so by asserting that $z_2(x)$, the initial "measurement" of the true state of nature $u_2^*(x)$ is obtained by <u>calculation of a steady state profile</u>. This calculation is assumed to be in error by an amount k_3 (See equation (5.98)). Thus we assume, for the purposes of this example, that the system S_D is initially at some steady state. This steady state is calculated and used as an initial condition $z_2(x)$, considered as an "input measurement".

The solution of example 2 is accomplished via theorem IV.1 IB, which asserts that the unique $\underline{u}\varepsilon V$ which extremizes $J(\underline{v})$ (equation (5.106)) is characterized by the following system of equations:

$$\frac{\partial y(x,t;\underline{u})}{\partial t} - \frac{\partial^2 y(x,t;\underline{u})}{\partial x^2} = 212.0 \qquad x,t\varepsilon(0,1)\times(0,1] \quad (5.107)$$

$$y(0,t) = u_1(0,t) \qquad t\varepsilon(0,1] \quad (5.108)$$

$$y(1,t) = u_1(1,t) \qquad t\varepsilon(0,1] \quad (5.109)$$

$$y(x,0) = u_2(x) \qquad x\varepsilon(0,1) \quad (5.110)$$

$$-\frac{\partial p(x,t;\underline{u})}{\partial t} - \frac{\partial^2 p}{\partial x^2} = \sum_{i=1}^{4} [y(x^i,t;\underline{u}) - z(x^i,t)]\delta(x-x^i) \quad (5.111)$$

$$x,t\varepsilon(0,1)\times(0,1] \quad (5.112)$$

$$p(0,t) = 0 \qquad t\varepsilon(0,1] \quad (5.113)$$

$$p(1,t) = 0 \qquad t\varepsilon(0,1] \quad (5.114)$$

$$p(x,1) = 0 \qquad x\varepsilon(0,1) \quad (5.115)$$

$$\frac{\partial p(0,t;\underline{u})}{\partial x} + u_1(0,t) - z_1(0,t) = 0 \qquad t\varepsilon(0,1] \quad (5.116)$$

$$-\frac{\partial p(1,t;\underline{u})}{\partial x} + u_1(1,t) - z_1(1,t) = 0 \qquad t\varepsilon(0,1] \quad (5.117)$$

$$p(x,0;\underline{u}) + u_2(x) - z_2(x) = 0 \qquad x\varepsilon(0,1) \quad (5.118)$$

As in example 1, the system state $y(x,t;\underline{u})$ and its adjoint $p(x,t;\underline{u})$ were approximated by:

$$y_m(x,t;\underline{u}) = \sum_{i=1}^{4} y_i(t;\underline{u})w_i(x) \quad (5.119)$$

$$p_m(x,t;\underline{u}) = \sum_{i=1}^{4} p_i(t;\underline{u}) w_i(x) \qquad (5.120)$$

$$\hat{y}_m(x,t) = \sum_{i=1}^{8} \hat{y}_i(t) w_i(x) \qquad (5.121)$$

$$w_i(x) = \sqrt{2}\,\sin(\sqrt{\lambda_i})x;\; \lambda_i=(i\pi)^2 \qquad (5.122)$$

In the computation of $\hat{y}_i(t)$, the system of identities (5. 98) through (5.104) is augmented by the addition of:

$$Q_\nu = \sum_{i=1}^{\nu} Q^i \;;\quad \nu=1,2,3,4. \quad \text{(The four measurement locations)}$$

$$\{Q^i\}_{k,l} = \sqrt{2}\,\sin(k\pi)x^i \sqrt{2}\,\sin(l\pi)x^i;\; (k,l=1,2\ldots8) \qquad (5.123)$$

Using the appropriate identities contained in this section, we obtained results via the developments of section A.2, for the following:

(a) $\hat{y}_i(t)$; $\{P_{ij}(t)\}^{-1}$ $\quad i,j=1,2,\ldots 8$

(b) $\hat{y}_m(x,t)$ using (5.121)

(c) Using the algorithm of section A.3, we recovered $u_1(0,t), u_1(1,t)$ and

(d) $u_2(x)$

(e) $y_i(t;\underline{u})$

(f) $y_m(x,t;\underline{u})$

The intervals of definition of these quantities are the same as in example 1.

We are now in a position to give an evaluation of the numerical results for examples 1 and 2.

A.4 Evaluation of the Numerical Results

Numerical examples 1 and 2 are considered in tandem because of the pervading structural similarity of the equations defining the solutions. We consider first solutions for $\hat{y}(x,t)$, the "filtered estimate" for both examples.

The equations of interest are those arising out of an approximation to $\hat{y}(x,t)$ (See sections A.4, A.5):

Example 1

$$\frac{d\hat{y}}{dt}(t)+A\hat{y}(t)-f(t)+z_1(t) = P^{-1}(t)[z(t)-\hat{y}(t)] \qquad (5.63)$$

$$\hat{y}(0) = z_2 \qquad (5.64)$$

$$\frac{dP^{-1}}{dt}(t)+AP^{-1}(t)+P^{-1}(t)A+P^{-1}(t)IP^{-1}(t)-WW^T=0 \qquad (5.65)$$

$$P^{-1}(0) = I \qquad (5.66)$$

$$\hat{y}_m(x,t) = \sum_{i=1}^{m} \hat{y}_i(t)w_i(x).$$

Example 2

$$\frac{d\hat{y}}{dt} + A\hat{y}(t) - f(t) + z_1(t) = P^{-1}(t)Q_\nu[z(t) - \hat{y}(t)] \quad (5.67)$$

$$\hat{y}(0) = z_2 \quad (5.68)$$

$$\frac{dP^{-1}(t)}{dt} + AP^{-1}(t) + P^{-1}(t)A + P^{-1}(t)Q_\nu P^{-1}(t) - WW^T = 0 \quad (5.69)$$

$$P^{-1}(0) = I \quad (5.70)$$

$$\hat{y}_m(x,t) = \sum_{i=1}^{m} \hat{y}_i(t)w_i(x).$$

In both examples, a fourth order Runge-Kutta integration routine was used to generate solutions. It was found that a stepsize of 0.001 had to be taken to ensure the numerical stability of the (8×8)-dimensional P equations. This was a consequence of the small time constant τ, $\tau = \frac{1}{2(i\pi)^2}$, for higher order modes. This small stepsize resulted in the algorithm being slow, although the computation time required for each integration step (manipulating and integrating 52 equations), was about 3 seconds.

Figure V.1 shows a typical set of responses of elements of $P^{-1}(t)$. Note the very fast "rise time" of the higher order mode $\{P^{-1}\}_{5,3}$. Note also the offset at steady state between the results for examples 1 and 2. A positive definite character of $P^{-1}(t)$ for all t was observed for t_{ss}=0.008, for example 1. The elements of P^{-1} were at steady state for t_{ss}=0.008. The results are displayed in

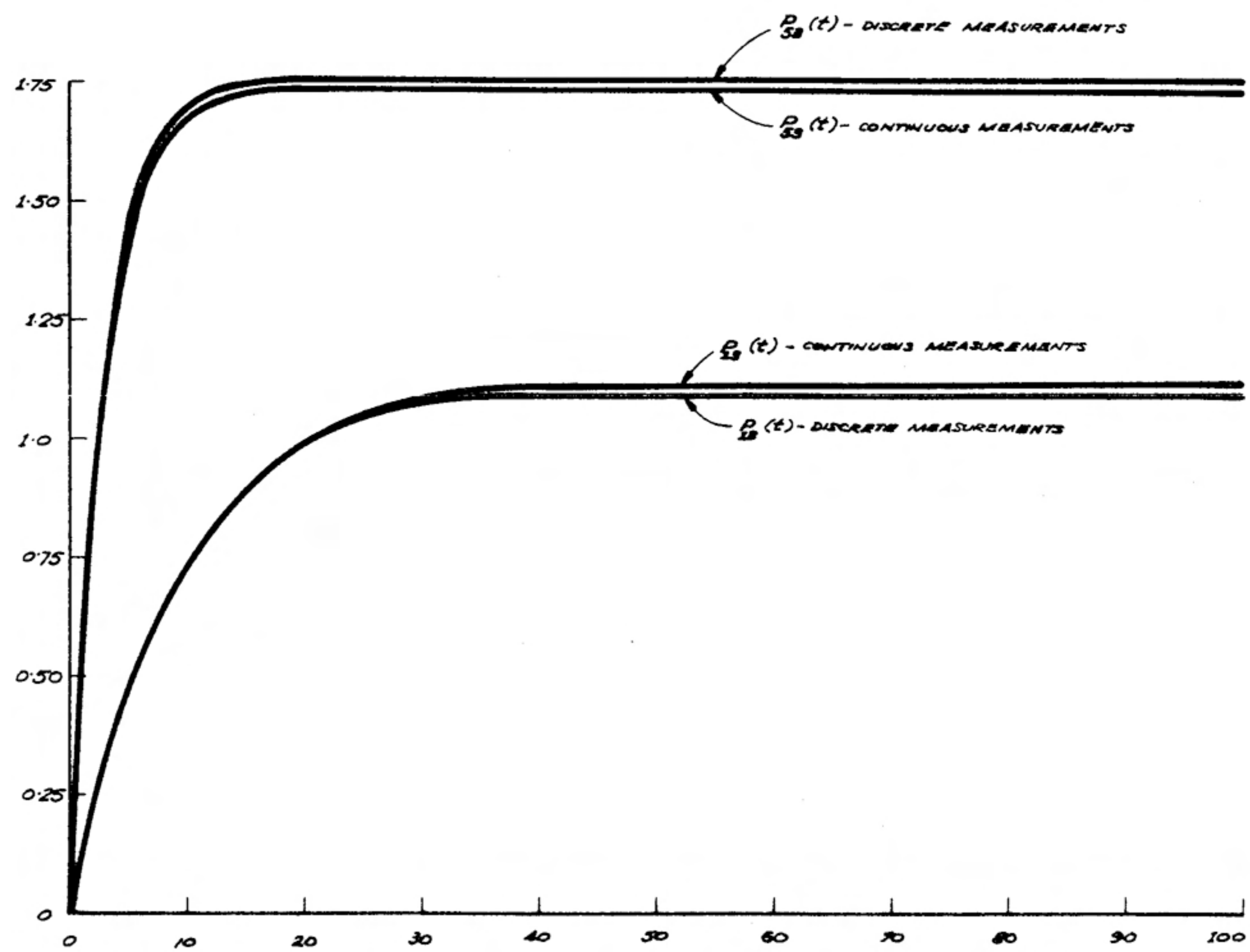

FIGURE V.1 Response of Selected Elements $\{P_{ij}\}^{-1}$

Table V.1:

TABLE V.1

Principal Minors of $P^{-1}(t_{ss})$

Principal Minor No.	Value
1	1.7349 > 0
2	3.2860 > 0
3	3.9406 > 0
4	2.7196 > 0
5	0.9916 > 0
6	0.1749 > 0
7	0.0158 > 0
8	0.0050 > 0

An important feature of the matrix $P^{-1}(t)$ in example 1, for all $t\varepsilon(0,T]$ is the fact that:

$$\{P^{-1}(t)\}_{i,j} = 0 \qquad \text{if } (i+j) \text{ is odd.}$$

We see this property immediately from the equations defining $P^{-1}(t)$, wherein it is observed that the odd terms satisfy a homogeneous linear ordinary differential equation with zero initial conditions. This property plays an important role in the definition of the responses $\hat{y}(t)$, to which we now turn.

In example 1, the equation for the i-th component of $\hat{y}(t)$ is (equation 5.67):

$$\frac{d\hat{y}_i}{dt}(t)+(i\pi)^2\hat{y}_i(t)+f_i(t)-z_{1i}(t)=\sum_{j=1}^{8}\{P^{-1}(t)\}_{ij}[z_j(t)-\hat{y}_j(t)] \tag{5.124}$$

Now observe that i even, z_{2i}, $z_{1i}(t)$ and $z_i(t)$ contain no noise - a consequence of the eigenfunction expansion chosen. Thus the LHS of (5.124) describes exactly the evolution of the i-th mode of the true state of nature. The property that $\{P^{-1}(t)\}_{ij} = 0$ if (i+j) odd implies that the RHS of (5.124) is zero for <u>i even</u>. Thus for i even, (5.124) describes the evolution of the i-th mode of the true state of nature.

A similar result holds for the equations of example 2, but there the result is not as readily established.

The foregoing discussion is summarized by the response shown for $\hat{y}_2(t)$, examples 1 and 2, in figure V.2. A typical response of an odd-numbered element of $\hat{y}(t)$, examples 1 and 2 is also given in figure V.2. It is noted that the responses $\hat{y}_5(t)$ shown in figure V.2 for both examples are virtually indistinguishable. However it is possible to distinguish between the responses $\hat{y}_1(t)$ for the two examples, as shown in figure V.3. We observed that:

$$[\hat{y}_i(\cdot)]_{\text{Example 2}} \to [\hat{y}(\cdot)]_{\text{Example 1}} \text{ as } i \to \infty$$

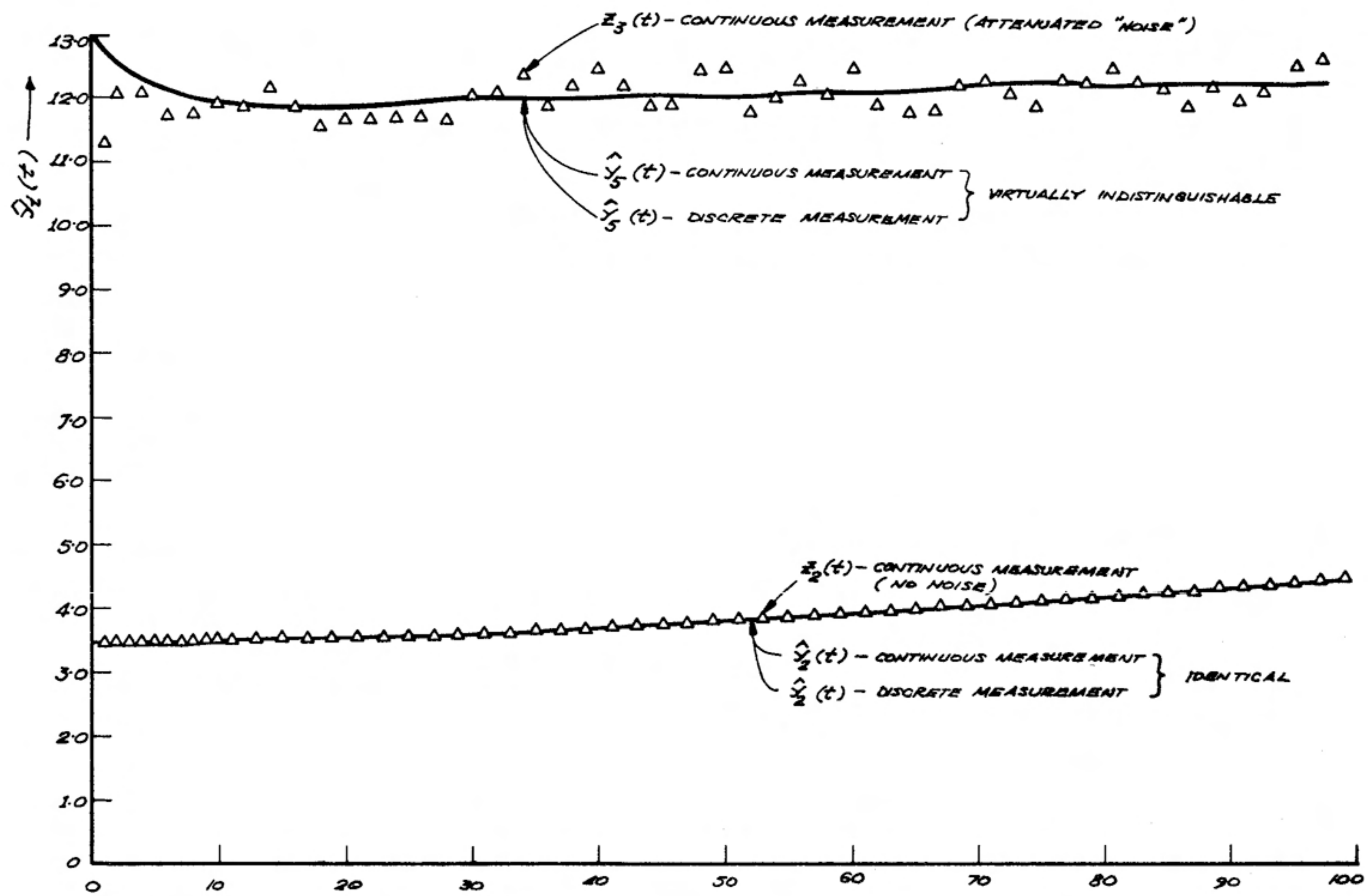

FIGURE V.2 Response of Selected Modes of the Filtered Estimates $\hat{y}_i(t)$ Continuous and Discrete Interior Measurements

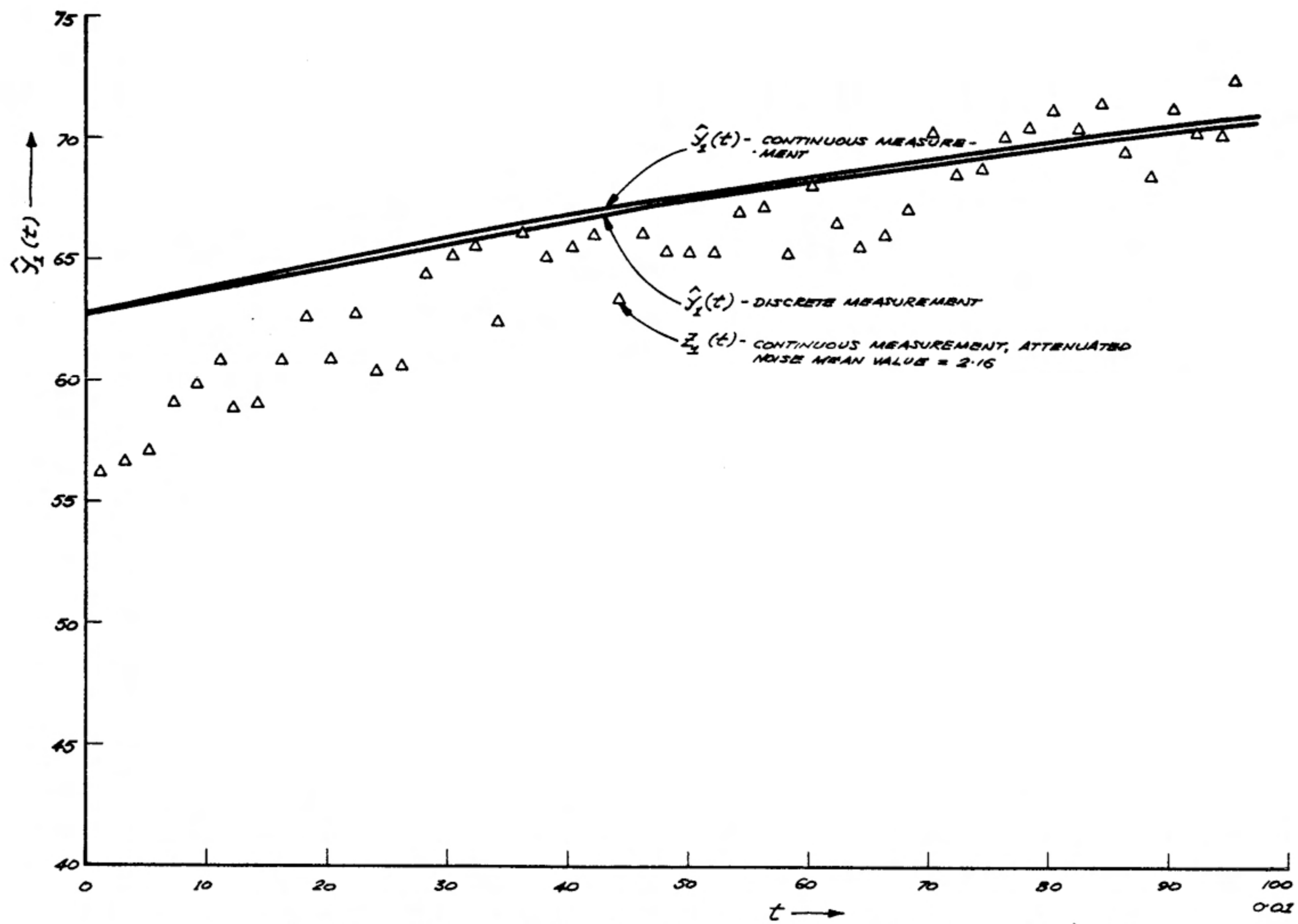

FIGURE V.3 Response of Selected Modes of the Filtered Estimates $\hat{y}_i(t)$ Continuous and Discrete Interior Measurements

A graphical display of $\hat{y}_m(x,t)$ (and $P_m(x,\xi,t)$) is omitted, as the salient features of this kind of approximation are treated in the consideration of the approximation to the "smoothed estimate", $y_m(x,t;\underline{u})$, which follows. To illustrate the method of solution, we shall go through an iteration of the algorithm for obtaining the "smoothed estimate" $y(x,t;\underline{u})$ given in section A.3.1, using therein the approximations outlined in section A.4. The specific examples are those of section A.5.

Iteration 0

(i) Guess the functions $u_1(0,\cdot)$, $u_1(1,\cdot)$ and $u_2(\cdot)$ over the appropriate grid (section A.5).

Remark A reasonable first guess is to set

$$u_1^0(0,\cdot) = z_1(0,\cdot)$$

$$u_1^0(1,\cdot) = z_1(1,\cdot)$$

$$u_2^0(\cdot) = z_2(\cdot)$$

For the purposes of illustrating convergence properties, the first guess of $\underline{u}^0$ was made independently of the measurements $\underline{z}$. See for example the first guess of $u_2^0(\cdot)$ illustrated in figure V.5.

(ii) Evaluate $\underline{G}(\underline{u}^0)$, given by:

$$\underline{G}(\underline{u}^0)= \begin{bmatrix} \sum_{i=1}^{4} p_i(\cdot)[\sqrt{2}(i\Pi)\cos(i\Pi)x]_{x=0}+u_1^0(0,\cdot)-z_1(0,\cdot) \\ \\ -\sum_{i=1}^{4} p_i(\cdot)[\sqrt{2}(i\Pi)\cos(i\Pi)x]_{x=1}+u_1^0(1,\cdot)-z_1(1,\cdot) \\ \\ \sum_{i=1}^{4} p_i(0)\sqrt{2}\ \sin(i\Pi)(\cdot)+u_2^0(\cdot)-z_2(\cdot) \end{bmatrix} \qquad (5.125)$$

In order to carry out the evaluation, we must first obtain $p_i(\cdot)$ and $p_i(0)$, (i=1,2,3,4). $p_i(\cdot)$ and $p_i(0)$ are obtained by integrating "backwards" in time (from t = 1) the following equations:

$$-\frac{dp_i(t)}{dt}+(i\Pi)^2 p_i(t) = y_i(t;\underline{u}^0)-z_i(t) \qquad (5.126)$$

$$p_i(1) = 0 \qquad (5.127)$$

$y_i(\cdot;\underline{u}^0)$ is obtained by solving (from t = 0):

$$\frac{dy_i(t)}{dt}+(i\Pi)^2 y_i(t) = f_i(t)-u_{1i}^0(t) \qquad (5.128)$$

$$y_i(0) = u_{2i}^0 \qquad (5.129)$$

where

$$u_{1i}^0(t) = \sqrt{2}\ (i\Pi)\{-[u_1^0(0,t)]+(-1)^i[u_1^0(1,t)]\}$$

$$u_{2i} = \int_0^1 \sqrt{2}\ u_2^0(x)\sin(i\Pi)x\ dx$$

Remark

It was found that an integration step size of 0.01 was sufficient to guarantee the numerical stability of a fourth order Runge-Kutta integration scheme.

Having solved for $p_i(t)$, $t\varepsilon(0,1]$ from (5.126) through (5.129), $\underline{G}(\underline{u}^0)$ is evaluated via (5.125).

(iii) If $\underline{G}(\underline{u}^0)$ is non-zero (as it would be except by most propitious circumstance), then the choice of $\underline{u}^0$ is updated according to:

$$u_1^1(0,\cdot) = u_1^0(0,\cdot) + \alpha^0 s_1^0(0,\cdot)$$

$$u_1^1(1,\cdot) = u_1^0(1,\cdot) + \alpha^0 s_1^0(1,\cdot)$$

$$u_2^1(\cdot) = u_2^0(\cdot) + \alpha^0 s_2^0(\cdot)$$

As we indicated (section A.3.1), $\underline{s}^0$ is chosen to be the negative gradient which has already been computed:

$$s_1^0(0,\cdot)=-G_1^0(0,\cdot)=-\sum_{i=1}^{4} p_i(\cdot)[\sqrt{2}\ (i\Pi)\cos(i\Pi)x]_{x=0}-u_1^0(0,\cdot)+z_1(0,\cdot)$$

$$s_1^0(1,\cdot)=-G_1^0(1,\cdot)=\ \sum_{i=1}^{4} p_i(\cdot)[\sqrt{2}\ (i\Pi)\cos(i\Pi)x]_{x=1}-u_1^0(1,\cdot)+z_1(1,\cdot)$$

$$s_2^0(\cdot) = -G_2^0(\cdot) = -\sum_{i=1}^{4} p_i(0)\sqrt{2}\ \sin(i\Pi)(\cdot)-u_2^0(\cdot)+z_2(\cdot)$$

α^0 is given explicitly by:

$$\alpha^0 = \frac{\int_0^1 \{G_1^0(0,t)^2 + G_1^0(1,t)^2\}dt + \int_0^1 G_2^0(x)^2 dx}{\sum_{i=1}^{4} \int_0^1 [y_i(t;\underline{s}^0) - y_i(t;\underline{0})]^2 dt + \int_0^1 \{s_1^0(0,t)^2 + s_1^0(1,t)^2\}dt + \int_0^1 s_2^1(x)^2 dx}$$

Thus we obtain the $\underline{u}^1$ vector. The computational effort is less forbidding than the algebra implies.

Successive stages in the algorithm are as outlined in section A.3.1. The appropriate modifications to the explicit formulas given for the "zero-th" iteration are straightforward.

With this understanding of the iterative scheme, we can proceed to a discussion of its application to the two examples of section A.5. We obtained the following results, common to both examples:

(i) Convergence of the algorithm is rapid: 3 iterations sufficed for example 1 and 4 iterations for example 2. (that is, $\underline{G}(\underline{u}^3) \doteq 0$ and $\underline{G}(\underline{u}^4) \doteq 0$ respectively). Computation time per iteration was 25 seconds.

(ii) Using the refined estimate $\underline{u}$ obtained after the appropriate number of iterations, the approximation to $p(x,t;\underline{u})$ afforded by the four mode expansion

$$p_m(x,t;\underline{u}) = \sum_{i=1}^{4} p_i(t;\underline{u})w_i(x)$$

was precise - $p_1(t;\underline{u})$ being the only non-zero mode in the expansion. That is,

$$\lim_{m\to\infty} p_m(x,t;\underline{u}) = p_1(t;u)w_1(x) = p(x,t;\underline{u})$$

for almost every $x,t\varepsilon(0,1)\times(0,1]$.

A consequence of this result was that $\underline{u}$, defined in terms of $p(x,t;u)$, was accurate.

(iii) While a 4 mode expansion was adequate to define $p(x,t;\underline{u})$, the approximation of $y_m(x,t;\underline{u})$ to $y(x,t;\underline{u})$ was, as will be seen, poor. This fact is of little consequence, since the announced goal was to find $\underline{u}$, not $y(x,t;\underline{u})$.

Figures V.4 through V.15 illustrate the numerical results for examples 1 and 2. In figure V.4 we show the refined estimate of one of the boundary conditions. The convergence of the first guess of the other boundary condition $u_1^0(1,t)$ is displayed in figure V.5. In figure V.6, we show the refined estimate of the initial condition. Figures V.7 and V.8 show the responses of typical modes of the refined estimate $y(t;\underline{u})$ and the effect of successive improvement in the choice of $\underline{u}$.

Using $y(t;\underline{u})$, we recover the approximation to $y(x,t;u)$ as follows:

$$y_m(x,t;\underline{u}) = \sum_{i=1}^{4} y_i(t;\underline{u})w_i(x)$$

This approximation is shown in figure V.9 for t = 0.4. We note that the approximation is singularily bad. Of course, the addition of more terms in the approximation would improve the fit, but the

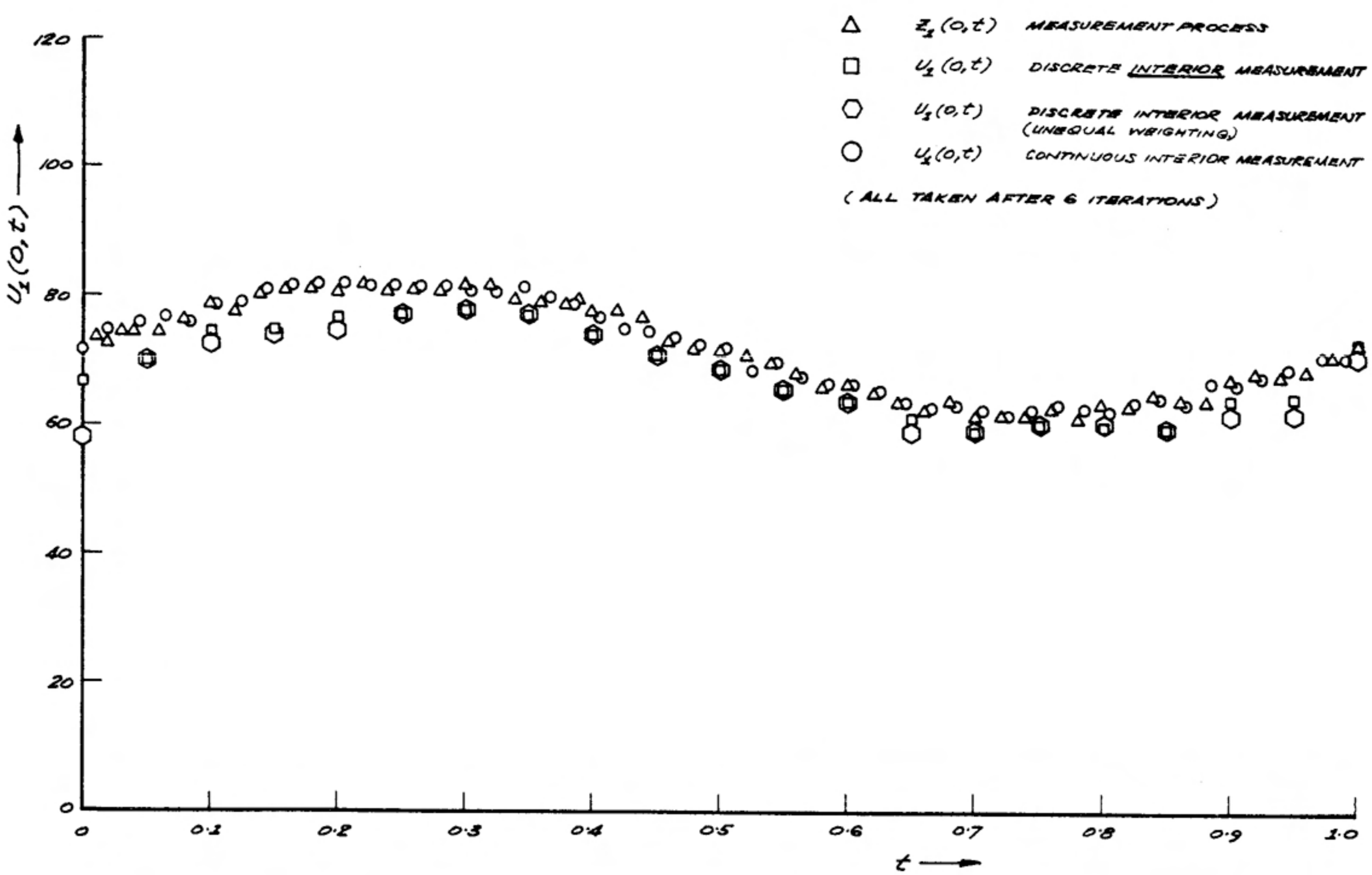

FIGURE V.4 Boundary Condition $u_1(0,t)$

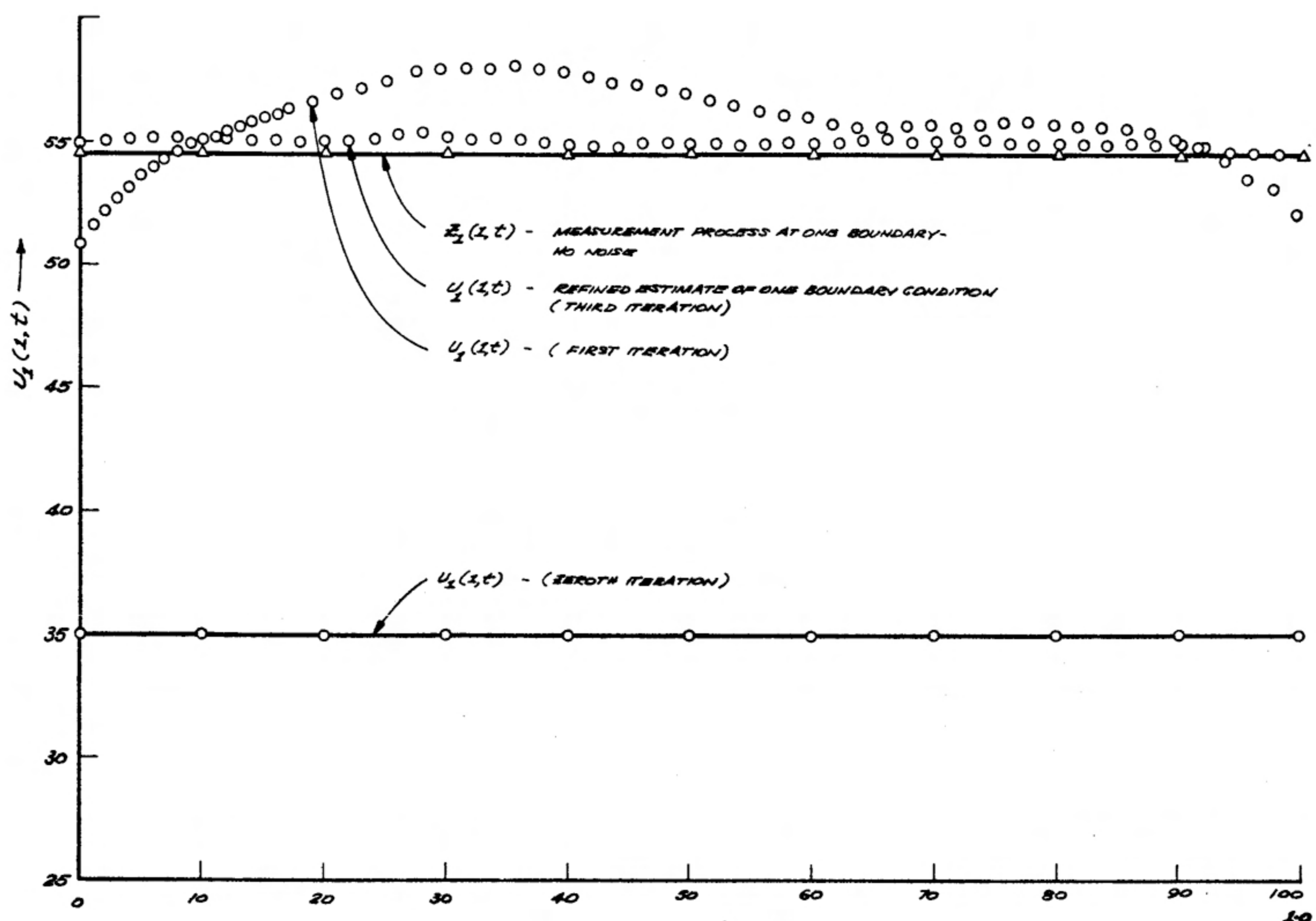

FIGURE V.5 Boundary Condition $u_1(1,t)$

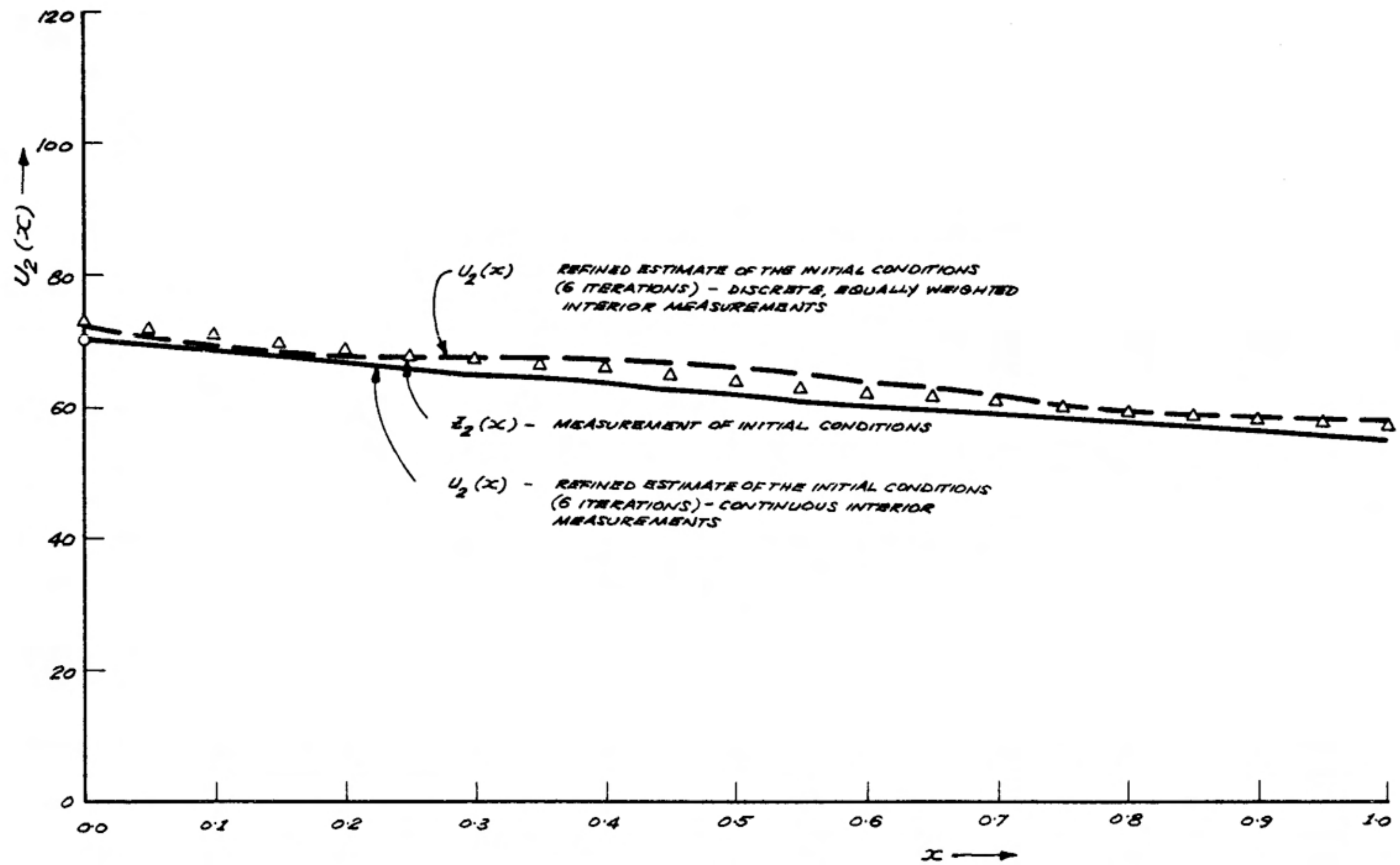

FIGURE V.6 Initial Condition $u_2(x)$

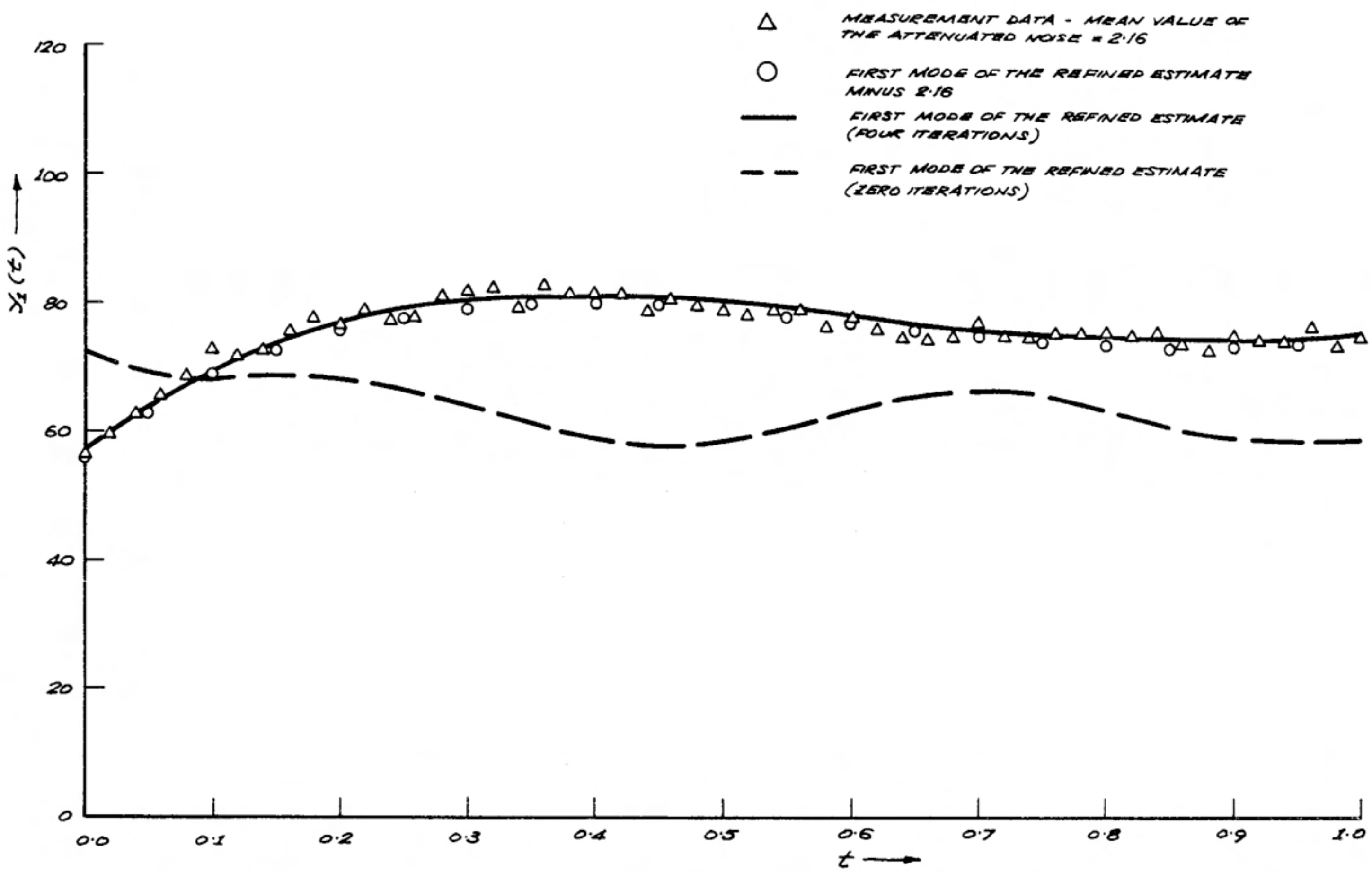

FIGURE V.7 Response of the First Mode of the Refined Estimate, $y_1(t)$

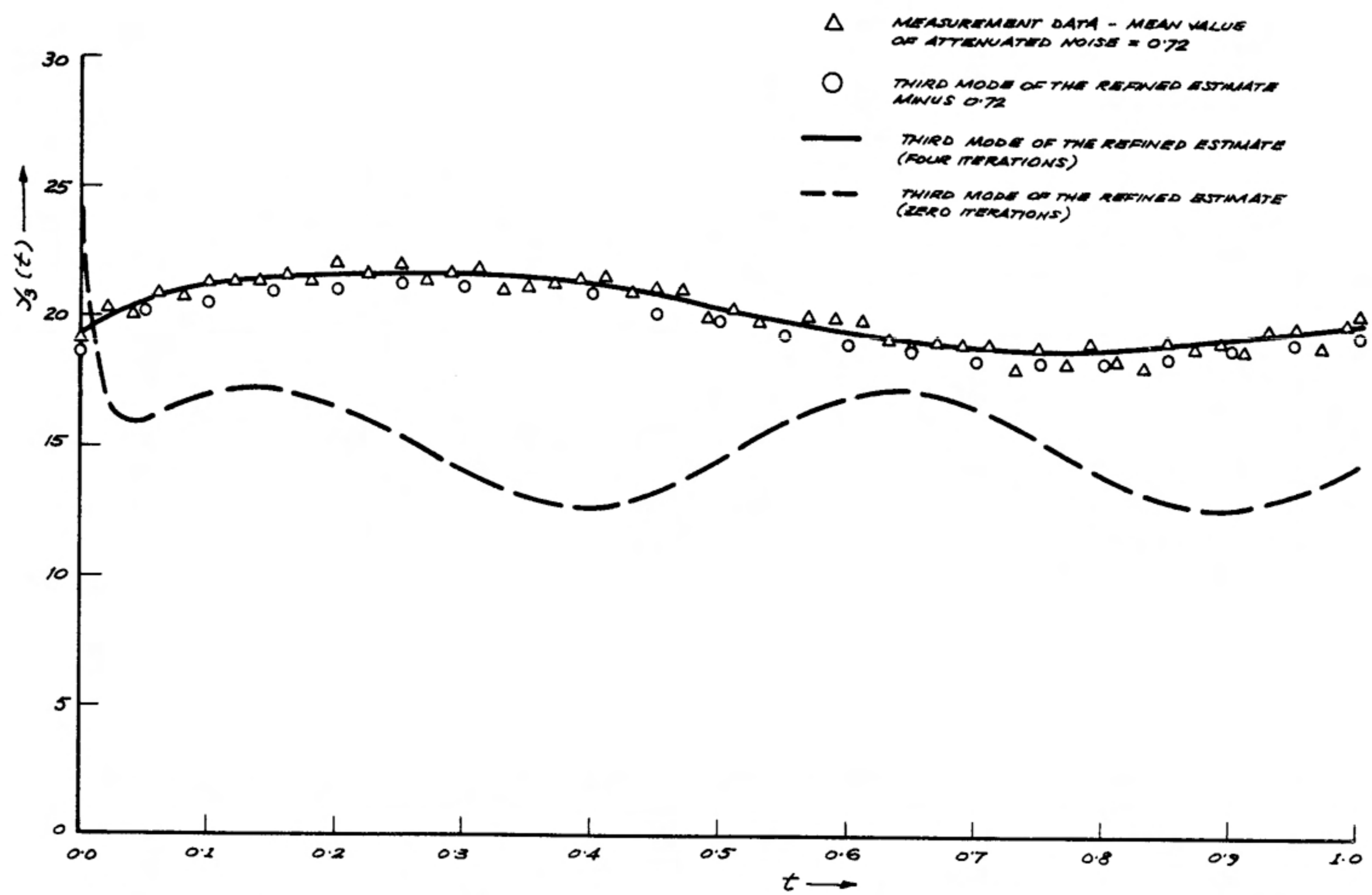

FIGURE V.8 Response of the Third Mode of the Refined Estimate, $y_3(t)$

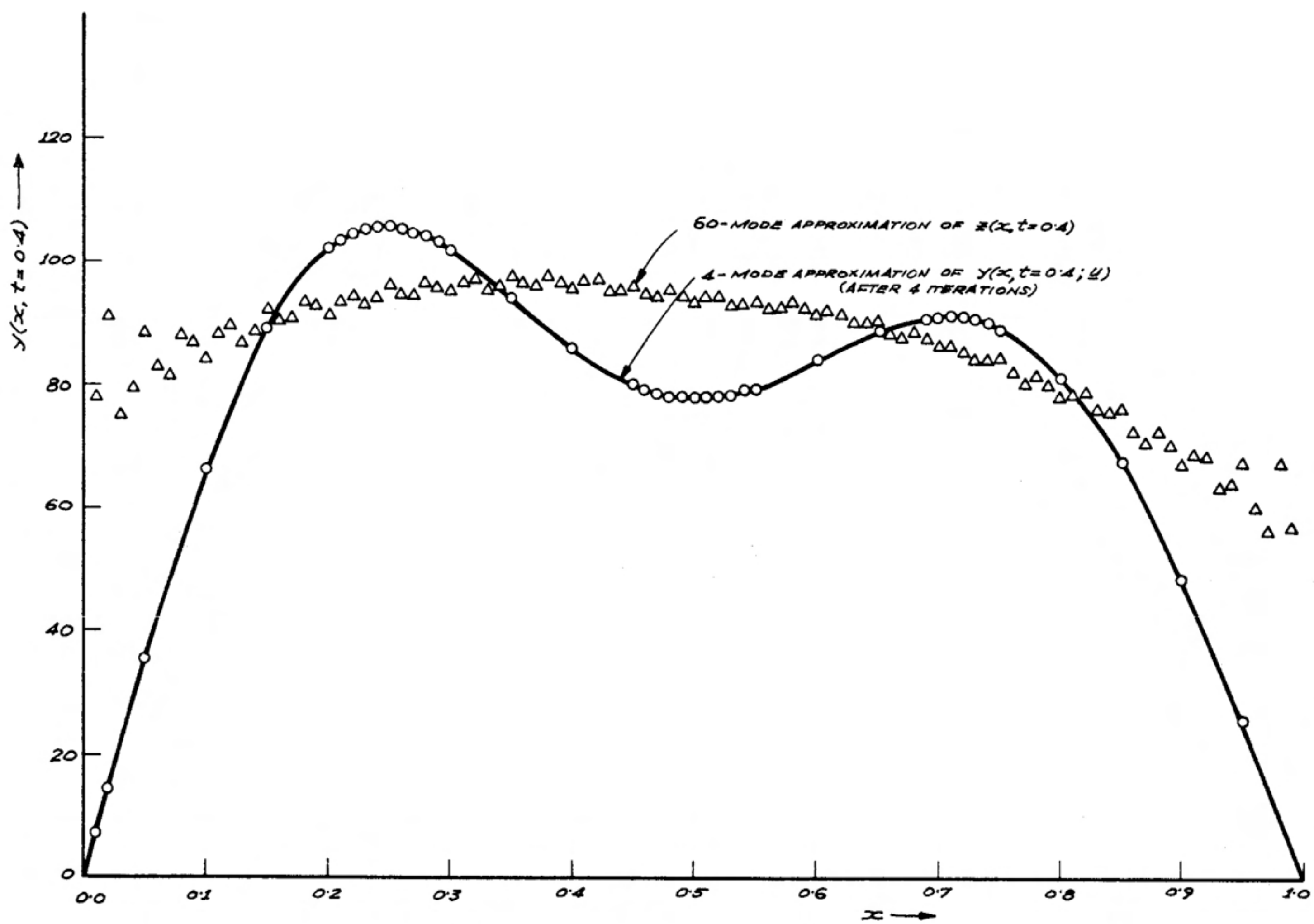

FIGURE V.9 Spatial Profile of Four Mode Refined Estimate at t = 0.4

accuracy of $\underline{u}$ would not be improved, in the light of preceding arguments. Note that $y_m(x,t;\underline{u})$ could be obtained to any desired accuracy, following the recovery of $\underline{u}$ with a four mode model.

A comparison of the four mode refined estimates for examples 1 and 2 is afforded by figure V.10, wherein we show y(0.2,t;u) for examples 1 and 2. Note from figure V.9, that at x = 0.6, the four mode approximations fall below the measured data. We show in figure V.11 the response at x = 0.6 of the four mode refined estimate for example 2. By "equal weighting" and "unequal weighting" we mean that the measurement errors at the discrete spatial locations were equally or unequally weighted. Unequal weighting caused a deterioration in the accuracy of the refined estimate. The apparent improvement shown in figure V.11 is illusory, since the fit was inordinately worsened at the other spatial locations.

The responses of two of the elements of p(t), $p_1(t)$ and $p_2(t)$, are shown in figure V.12. Note the effect of successive iterations on the responses. We observed that after the third iteration, $p_2(\cdot)$, $p_3(\cdot)$, and $p_4(\cdot)$ were all identically zero. $p_1(\cdot)$ was non-zero as shown. An indication of the rate of convergence of the algorithm is shown in figures V.13 and V.14. In the former, the convergence of one of the gradients of the functional $J(\underline{u})$ to zero is shown. In the latter, we show the reduction of $J(\underline{u})$ with successive iterations. Note that the value of the minimum J for example 1 is less than that for example 2.

Some general remarks concerning the simulation are in order:

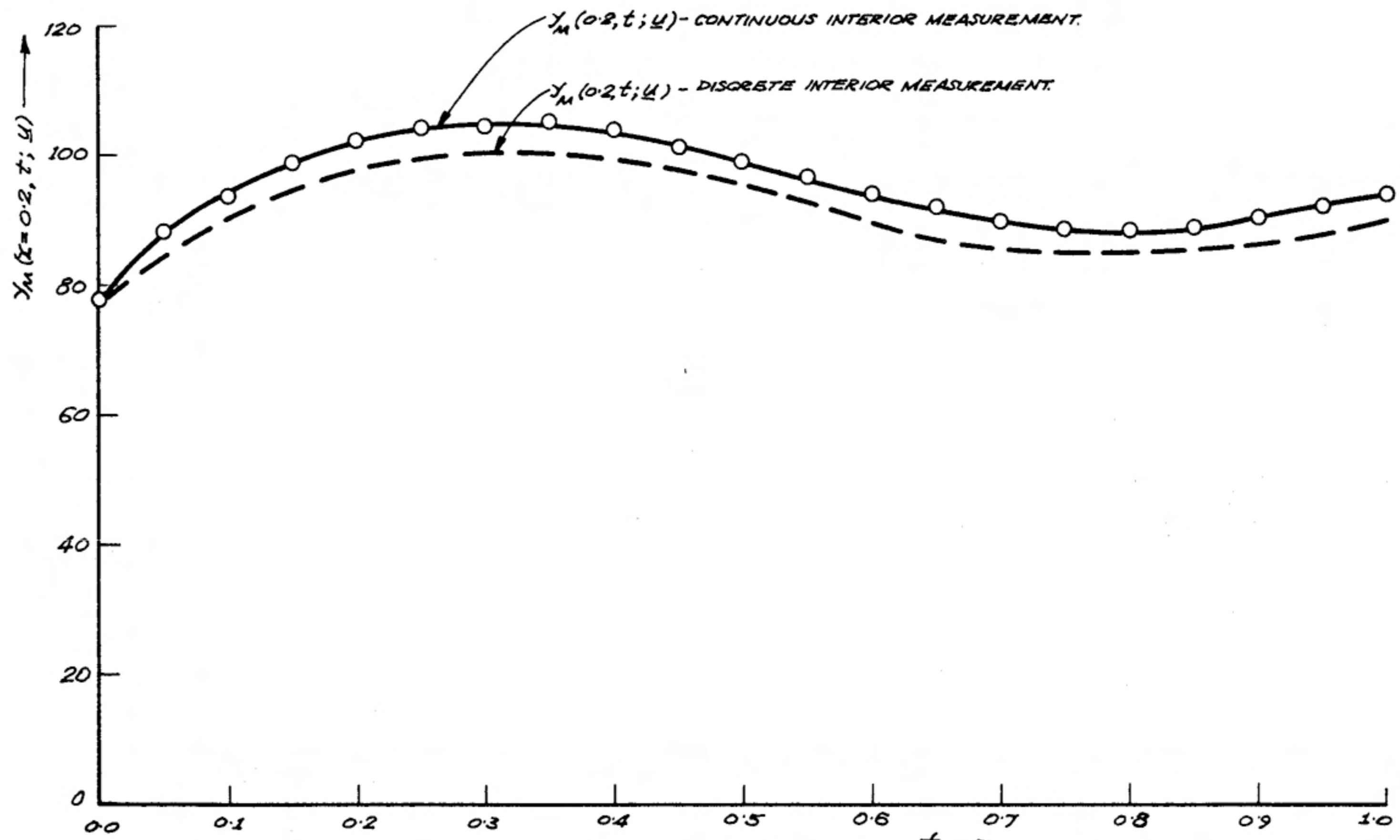

FIGURE V.10 Comparison of Four Mode Refined Estimates

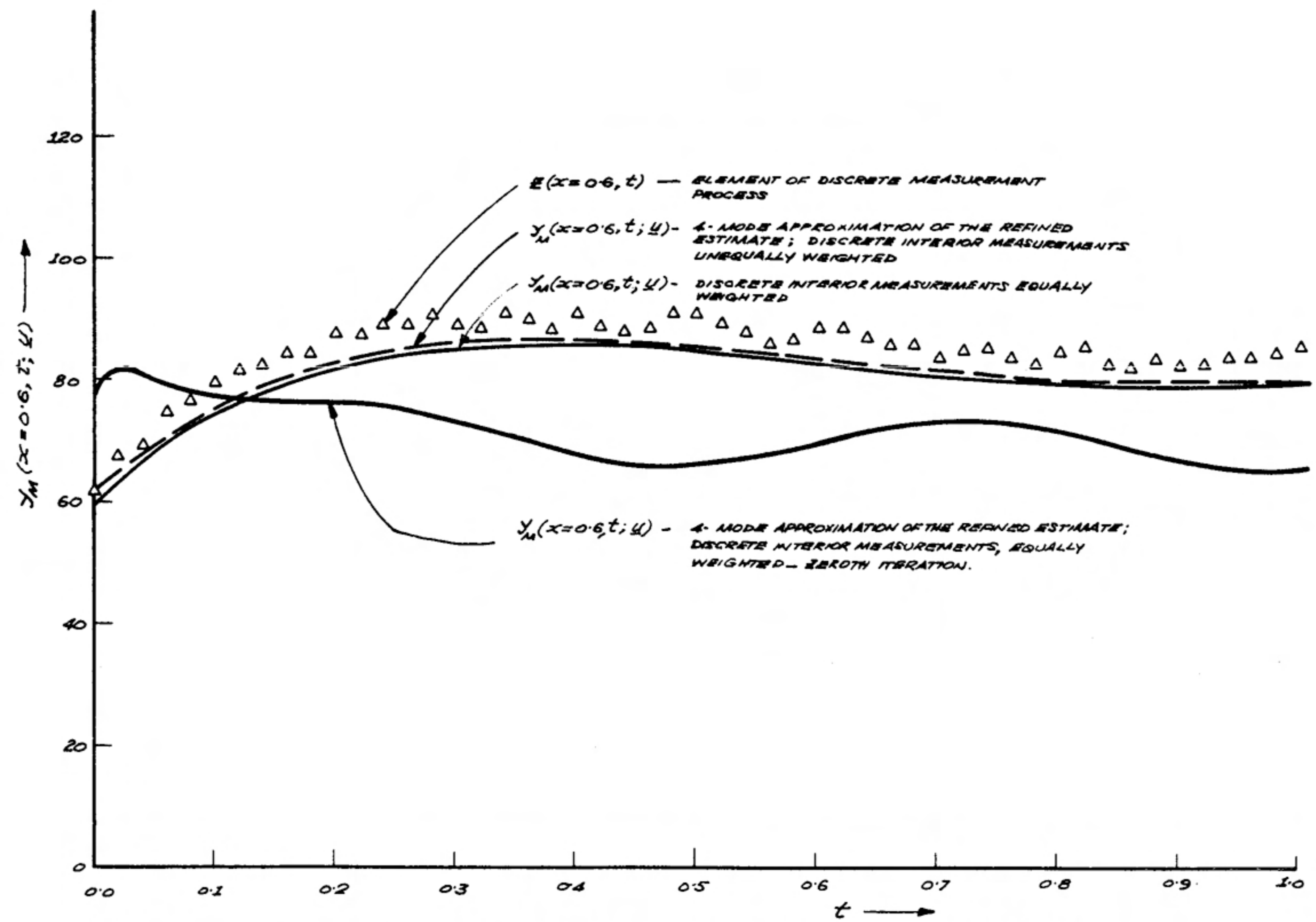

FIGURE V.11 Comparison of Four Mode Refined Estimates

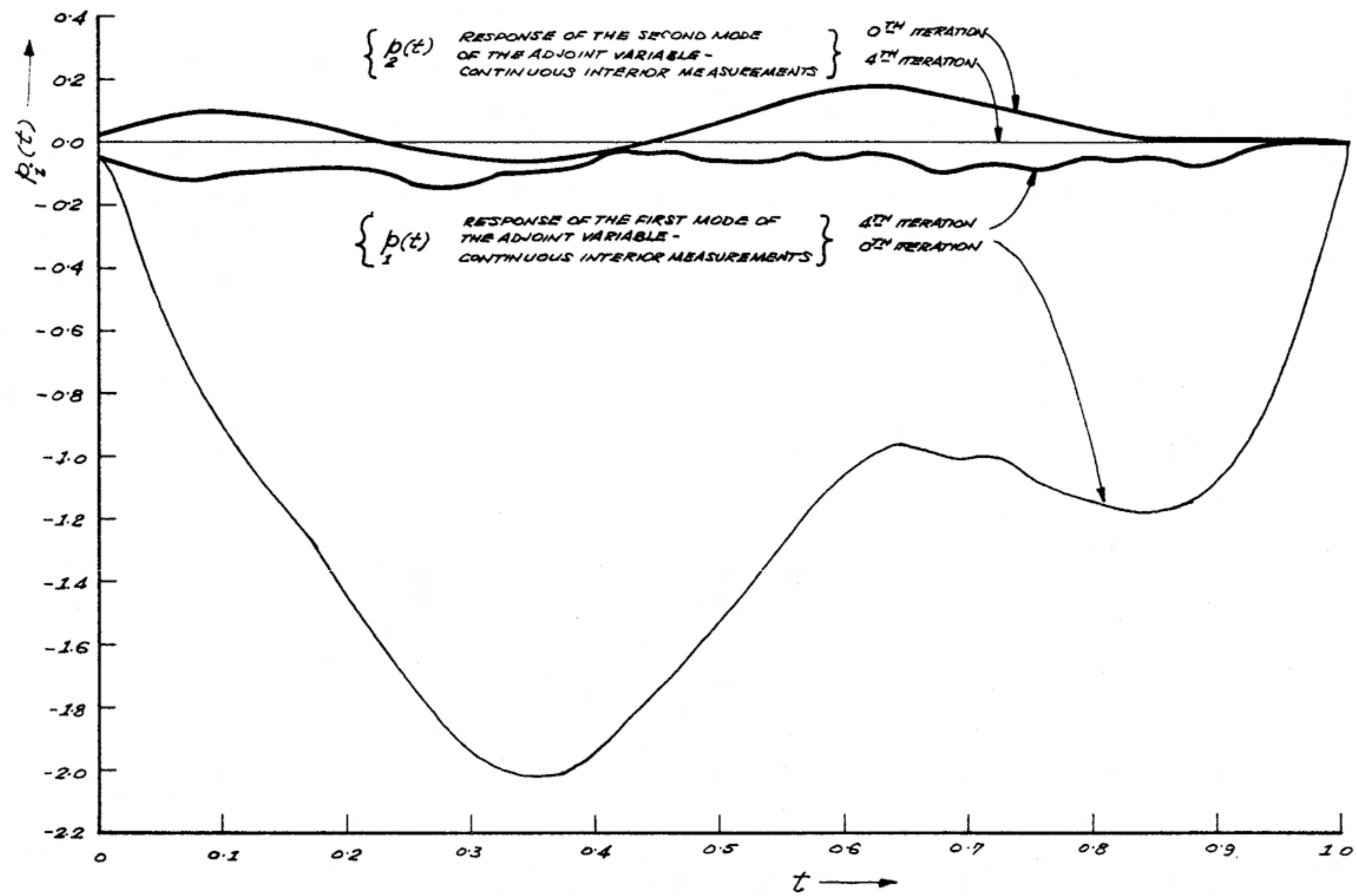

FIGURE V.12 Response of Selected Modes of the Adjoint Variable, $p_i(t)$

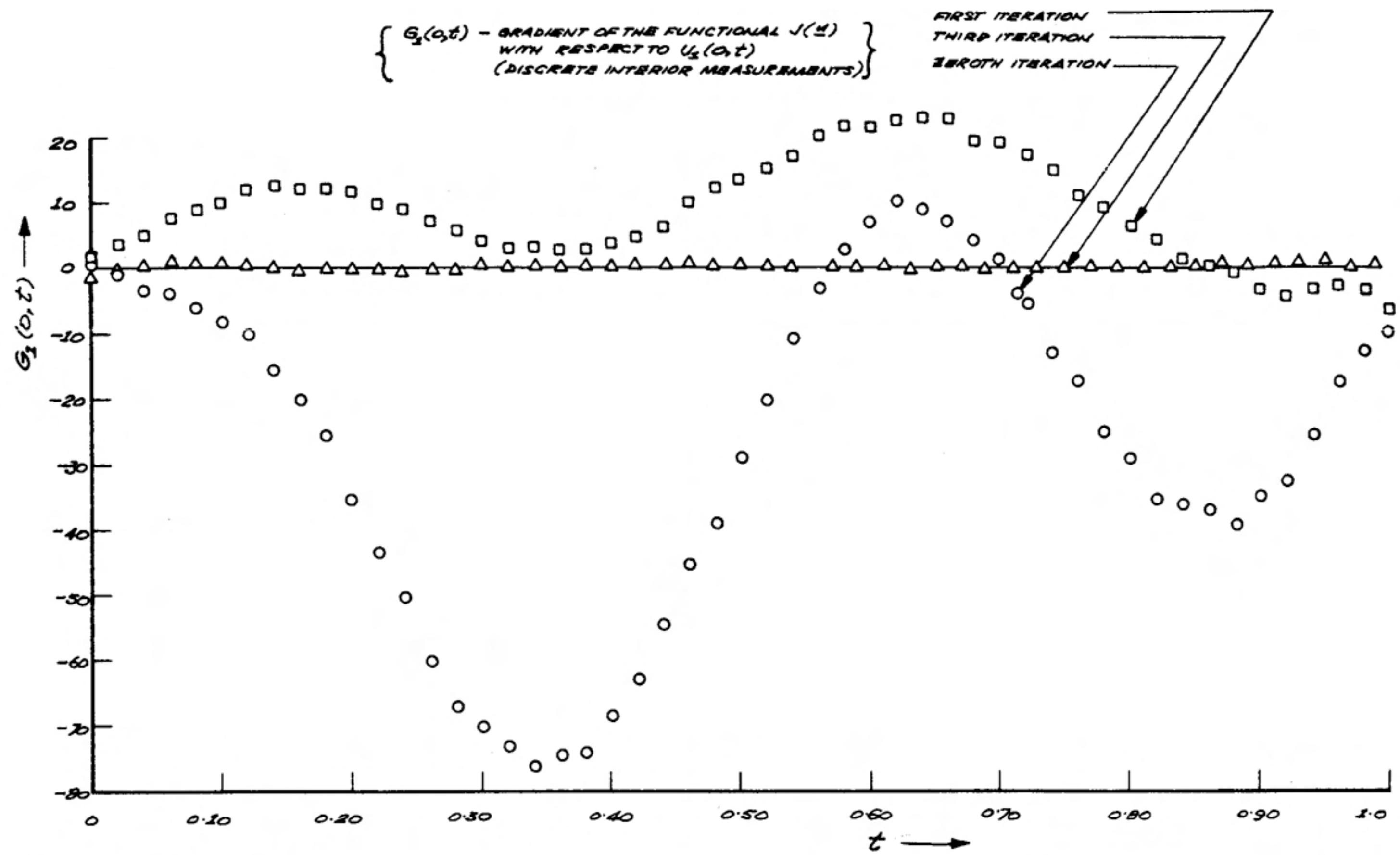

FIGURE V.13 Response of a Gradient of the Error Functional

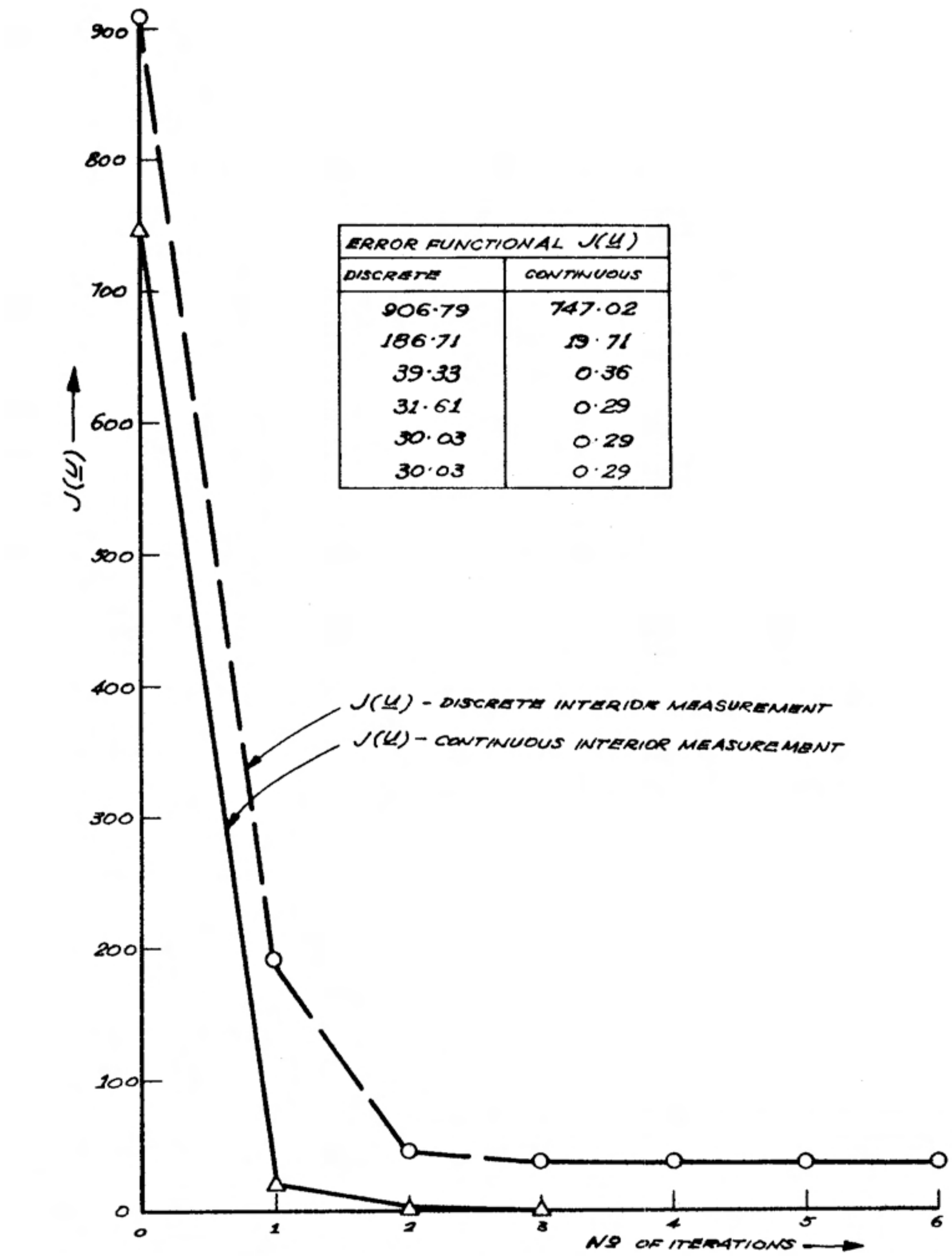

FIGURE V.14 Evaluation of the Error Functional

(i) No significant loss in the accuracy of the refined boundary and initial estimates was incurred by taking measurements at discrete interior spatial locations. This is not an obvious result and it cannot be generalized, though it may indicate that some of the information contained in the continuous spatial measurements is redundant.

(ii) The computational efficiency of the successive approximation algorithm presented should be stressed. Rapid convergence and relative computational speed per iteration of the scheme are noteworthy. A solution for the optimal $\underline{u}$ was obtained in 75 seconds on a Univac 1107. The computing time could undoubtedly be reduced by the use of more sophisticated programming. By comparison, obtaining an 8-mode approximation of the filtered estimate, $\hat{y}_m(x,t)$ would take 3000 seconds. Of course, the filtered estimate is a sequential or on-line estimator, whereas the smoothed estimate is obtained off-line, so direct comparisons are not meaningful.

A.5 Summary

In summary, this chapter has dealt with two distributed identification problems from formulation to numerical evaluation of two special examples. While the numerical evaluation was relatively exhaustive, there remain several alternatives for the numerical approximation of the solutions to the partial differential equations. This consideration, along with others, are proposed as meaningful extensions in the next chapter.

CHAPTER VI

SUMMARY AND EXTENSIONS

A.1 Summary

The identification problem associated with two classes of linear distributed systems has been posed as a variational one. This special interpretation allowed the application of some recent results in the theory of variational inequalities. This application resulted in a characterization of extremals to an error functional in terms of a set of cannonical partial differential equations, the solution of which was accomplished in two ways. First, we exploited the linearity of the equations, which enabled a decoupling of the two point (time) boundary value problem, which the cannonical equations constitute. This decoupling led to the consideration of a distributed Ricatti-like equation.

Secondly, the cannonical equations were solved by solving directly the associated variational problem in an iterative manner. The iterative technique, using conjugate directions of search on the error functional surface, was shown to converge, yielding a solution of the cannonical equations.

Both methods of solution required an approximation of a set of partial differential equations, which was made, in each case, via the technique of Galerkin. An illustrative example was chosen, and numerical evaluation of both solution schemes was made. It was found, for the example chosen (a diffusion system) that the iterative method of solution was rapid, yielding accurate results. The numerical solution of the distributed Ricatti-like equation was plagued by the necessity to take a small step size in numerical integration, in order to avoid numerical instability.

A.2 Extensions

The state evolution equations considered are general equations of their class, and hopefully there are immediate industrial applications. However, we acknowledge the idealized nature of the systems considered, especially in the assumption of linearity. A useful though admittedly formal expedient of linearizing nonlinear state evolution processes would lead to the consideration of the parameter identification problem. It is worthwhile to dwell momentarily on such a consideration. Consider the following parameter identification problem:

State evolution process:

$$\frac{\partial y(x,t)}{\partial t} - a(x)\,\frac{\partial^2 y(x,t)}{\partial x^2} = f(x,t) \qquad \text{in } Q \qquad (6.1)$$

$$y(\Sigma) = u_1(\Sigma) \qquad \text{on } \Sigma \qquad (6.2)$$

$$y(x,0) = u_2(x) \qquad \text{in } \Omega \qquad (6.3)$$

As before, $u_1(\Sigma)$ and $u_2(x)$ are unknown, and are to be determined by minimization of an appropriate quadratic error functional. Suppose that in addition, $a(x)$ is unknown, except for an initial estimate, given by

$$[a(x)]_{\text{INITIAL ESTIMATE}} = u_3(x) \tag{6.4}$$

We reformulate the system as follows:

<u>Define</u>

$$y_1(x,t) = y(x,t) \qquad x,t \in Q$$

$$y_2(x,t) = a(x) \qquad x,t \in Q$$

Then the system (6.1), (6.2), (6.3) and (6.4) may be written:

$$\begin{aligned} \frac{\partial y_1}{\partial t}(x,t) - y_2(x,t)\,\frac{\partial^2 y_1}{\partial x^2}(x,t) &= f(x,t) \\ \frac{\partial y_2}{\partial t}(x,t) &= 0 \end{aligned} \tag{6.5}$$

with the initial and boundary conditions:

$$y_1(x,0) = u_2(x) \tag{6.6}$$

$$y_2(x,0) = u_3(x) \tag{6.7}$$

$$y_1(\Sigma) = u_1(x) \tag{6.8}$$

We note that equations given by (6.5) are not of the type treated in the thesis. We can, however, reduce (6.5) to the type considered in the thesis by carrying out two approximations.

(I) Linearize (6.5) about an initial guess $\bar{y}_1(x,t)$ and $\bar{y}_2(x,t)$. Obtain equations for the principal-linear part of the perturbations, $p_1(x,t)$ and $p_2(x,t)$, induced by perturbing $u_1(\Sigma)$, $u_2(x)$ and $u_3(x)$:

$$y_1(x,t) = \bar{y}_1(x,t) + p_1(x,t)$$

$$y_2(x,t) = \bar{y}_2(x,t) + p_2(x,t)$$

$$u_1(\Sigma)=\bar{u}_1(\Sigma)+\delta u_1(\Sigma): \; u_2(x)=\bar{u}_2(x)+\delta u_2(x); \; u_3(x)=\bar{u}_3(x)+\delta u_3(x)$$

$$\frac{\partial p_1}{\partial t}(x,t) - c_1(x,t)p_2(x,t) - c_2(x,t)\frac{\partial^2 p_1}{\partial x^2} = 0 \qquad (6.9)$$

$$\frac{\partial p_2}{\partial t} = 0 \qquad (6.10)$$

$$p_1(\Sigma) = \delta u_1(\Sigma); \; p_2(\Sigma) = 0 \qquad (6.11)$$

$$p_1(x,0) = \delta u_2(x); \; p_2(x) = \delta u_3(x) \qquad (6.12)$$

where

$$c_1(x,t) = \frac{\partial^2 \bar{y}_1(x,t)}{\partial x^2}$$

$$c_2(x,t) = \bar{y}_2(x,t)$$

The second approximation necessary would be to introduce, in the manner suggested by Lions and Lattes [14] diffusion terms into (6.9) and (6.10):

(II) Modify (6.9) and (6.10) as follows:

$$\frac{\partial p_1}{\partial t}(x,t) - c_1(x,t)p_2(x,t) - \varepsilon_1 \frac{\partial^2 p_1}{\partial x^2}(x,t) - c_2(x,t)\frac{\partial^2 p_1}{\partial x^2}(x,t) = 0 \quad (6.13)$$

$$\frac{\partial p_2}{\partial t}(x,t) - \varepsilon_2 \frac{\partial^2 p_2}{\partial x^2}(x,t) = 0 \quad (6.14)$$

Where ε_1 and ε_2 are allowed to go to zero (in the limit).

Evidently (6.13) and (6.14) are in the form considered in the thesis. The object of the parameter identification problem would be to choose $\delta u_1(\Sigma)$, $\delta u_2(x)$, $\delta u_3(x)$ which minimize an error functional constrained by (6.13) and (6.14). The optimal choice of $\delta u_3(x)$ yields a solution to the parameter identification problem.

We feel, intuitively, that given an initial nominal guess of the vector $\underline{u}$ close to the optimal solution, the linearization would be effective. Unfortunately, proceeding formally as suggested, the problem loses its mathematical rigor, a cost often incurred but willingly accepted in many practical applications.

Another aspect of the problem rich with possibilities but fraught with difficulty is that of numerical technique. Given that a Galerkin-like technique is adopted, the choice of "basis functions" $w_i(x)$ is virtually limitless, and it is not obvious which is most acceptable from the standpoint of accuracy and/or computational effort. We suspect that the choice should be based on:

(i) The type of partial differential equation

(ii) A compromise between accuracy and computational effort.

Finally, the solution of the integral equation characterizing the filtered estimate might be attempted by the two conventional techniques of successive substitution or successive approximation. We suggest this expedient because of the large amount of time required to solve the distributed Ricatti-like equation.

APPENDIX II. 1

EXISTENCE AND UNIQUENESS OF EXTREMALS TO $J(\underline{v})$

Proof of Theorem II.3

(i) Existence: To show existence of $\underline{u} \epsilon K$, we proceed as follows:

(a) Show that a minimizing sequence of elements $v_n \epsilon K$ is such that

$$||v_n||_V \quad \text{bounded.}$$

(a) implies that there exists a subsequence $v_{\underline{n}_\mu} \epsilon K$ which converges weakly to $\underline{w}$.

(b) Show that $\underline{w} \epsilon K$

(c) Use the lower semi-continuity property of the norm $||\cdot||_V$ to deduce that

$$J(\underline{w}) = \operatorname*{Inf}_{\underline{v} \epsilon K} J(\underline{v})$$

Thus we start by showing (a). First we assume a non-trivial solution, that is, $K \neq \phi$

$$J(\underline{v}) = a(\underline{v},\underline{v}) - 2l(\underline{v}) \geq \alpha||\underline{v}||_V^2 - 2l(\underline{v}) \geq \alpha||\underline{v}||_V^2 - 2|l(\underline{v})|$$

$$J(\underline{v}) \geq \alpha||\underline{v}||_V^2 - 2||L|| \; ||\underline{v}||_V$$

$$\frac{J(\underline{v})}{||\underline{v}||_V} \to \infty \text{ as } ||\underline{v}|| \to \infty \quad ||v_n||_V \Longrightarrow \text{is bounded.}$$

Since $||v_n||_V$ is bounded, there exists a subsequence v_{n_μ} such that

$$v_{n_\mu} \overset{\text{weakly}}{\rightarrow} v_n$$

that is,

$$(v_{n_\mu}, v)_V \rightarrow (w,v)_V \qquad \text{for all} \quad v \varepsilon V$$

We show (b) by stating the following Lemma:

<u>Lemma II.10</u>

Given K closed and convex, then K is weakly closed.

Thus, for

$$v_{n_\mu} \varepsilon K, \; v_{n_\mu} \overset{\text{weakly}}{\rightarrow} w \Longrightarrow w \varepsilon K.$$

The characterization of the lower semi-continuity of the norm $||\cdot||_V$ is given in

<u>Lemma II.11</u>

$$\lim_{\mu\rightarrow\infty} \inf \; ||v_{n_\mu}||_V \geq ||w||_V \quad \text{if} \quad v_{n_\mu} \overset{\text{weakly}}{\rightarrow} w$$

We now deduce that <u>w</u> has the minimum property. First. we make the following assertion:

Assertion II.1

Lemma II.11 implies that

$$\lim_{\mu\to\infty} \text{Inf} \; J(\underline{v}_{n_\mu}) \geq J(\underline{w})$$

Proof

$$J(\underline{v}_{n}) - J(\underline{w}) = a(v_{n_\mu}, v_{n_\mu}) - a(\underline{w},\underline{w}) + (f, v_{n_\mu} - \underline{w})$$

$$\geq \alpha(||v_{n_\mu}||_V^2 - ||w||_V^2) + (f, v_{n_\mu} - \underline{w})$$

Using Lemma II.11 and the weak convergence of v_{n_μ} to $\underline{w}$,

$$\lim_{\mu\to\infty} \text{Inf} \; \{J(v_{n_\mu}) - J(\underline{w})\} \geq 0$$

$$\lim_{\mu\to\infty} \text{Inf} \; J(v_{n_\mu}) \geq J(\underline{w})$$

By hypothesis, v_n is a minimizing sequence, that is,

$$\lim_{\mu\to\infty} \text{Inf} \; J(v_{n_\mu}) \to \text{Inf}_{v \in K} \; J(\underline{v}) = j$$

By assertion II.1,

$$J(\underline{w}) \leq j$$

which is a contradiction, unless

$$J(\underline{w}) = j = \operatorname*{Inf}_{v\varepsilon K} J(\underline{v}) \ ; \ \underline{w}\varepsilon K$$

(ii) The uniqueness follows from the convexity of $J(\underline{v})$. We assume the contrary. Namely, take $\underline{u}_1$ and $\underline{u}_2 \varepsilon K$ such that:

$$J(\underline{u}_1) = J(\underline{u}_2) = j \ ; \ u_1 \neq u_2$$

By Lemma II.9, we have

$$J((1-\lambda)u_1+\lambda u_2) \leq (1-\lambda)J(u_1)+\lambda J(u_2) = j$$

Let

$$(1-\lambda)u_1+\lambda u_2 = u_3 \varepsilon K. \qquad \text{Then}$$

$$J(\underline{u}_3) \leq j, \qquad \text{a contradiction.}$$

Thus $\underline{u}_1 = \underline{u}_2$ and uniqueness follows.

APPENDIX V.1

CONVERGENCE OF CONJUGATE GRADIENT ALGORITHM

Theorem V.2

Given the hypothesis of section A.3.1, then if $\underline{G}(\underline{u}^i) \neq 0$, $J(u^{i+1}) < J(u^i)$.

Proof

Assume there is no $\alpha > 0$ such that

$$J(\underline{u}^{i+1}) = J(\underline{u}^i+\alpha\underline{s}^i) < J(\underline{u}^i).$$

Hence for all $\alpha > 0$,

$$\frac{J(\underline{u}^i+\alpha\underline{s}^i) - J(u^i)}{\alpha} \geq 0 \tag{5.35}$$

In the limit as α approaches zero, (5.35) gives

$$(G(u^i),s^i)_V \geq 0$$

$$\therefore \quad -(G(u^i),G(u^i))_V + \beta^i(G(u^i),s^{i-1}) \geq 0$$

But (5.33) implies

$$\frac{d}{d\alpha} J(u^i+\alpha s^i)=0 \Longrightarrow (G(u^i),s^i)_V = 0$$

Thus (5.36) gives a contradiction, namely

$$-(G(u^i),G(u^i))_V \geq 0 \ ; \ G(u^i) \neq 0$$

Hence there is an α which gives $J(u^i+1) < J(u^i)$. Since α is chosen according to (5.33), the theorem is proven.

REFERENCES

[1] Ahlberg, J.H., Nilson, E.N., Walsh, J.L., "The Theory of Splines and their Applications", New York: Academic Press, 1967.

[2] Bucy, R.S., "Nonlinear Filtering Theory", IEEE Series G., Correspondence item, 1965, pp. 198.

[3] Butkovskii, A.G., Lerner, A.Ya., "The Optimal Control of Systems with Distributed Parameters", Automation and Remote Control, Vol. 21, pp. 13-21, 1961.

[4] Butkovskii, A.G., "The Maximum Principle for Systems with Distributed Parameters", Automation and Remote Control, Vol. 22, pp. 1156-1169, 1962.

[5] Dettman, J.W., "Mathematical Methods in Physics and Engineering", New York: McGraw Hill, 1962.

[6] Erzberger, H., Kim, M., "Optimum Boundary Control of Distributed Parameter Systems", Information and Control, Vol. 9, pp. 265-278, 1966.

[7] Falb, P., "Infinite Dimensional Filtering - The Kalman-Bucy Filter in Hilbert Space", International Journal of Control, Vol. 11, pp. 102-137, 1967.

[8] Fletcher, R., Reeves, C.M., "Function Minimization by Conjugate Gradients", British Computer J., pp. 149-154, July 1964.

[9] Gelfand, I.M., Fomin, S.V., "Calculus of Variations", New Jersey: Prentice Hall, 1963.

[10] Kalman, R.E., "New Methods and Results in Linear Prediction and Filtering Theory", R.I.A.S., Technical Report No. 61-1, 1960.

[11] Kitamore, T., "Transformation of Distributed Parameter Systems into Lumped Parameter Systems for the Studies of Optimum Control", I.F.A.C., London, 1966, pp. 24A.3-24A.9.

[12] Kushner, H.J., "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes, with Applications", J. S.I.A.M. Control, Series A, Vol. 2, No. 1, 1964.

[13] Lasdon, L.S., Mitter, S.K. Warren, A.D., "The Conjugate Gradient Method for Optimal Control Problems", IEEE Series G, Vol. AC-12, No. 2, pp. 132-138, 1968.

[14] Lattes, R., Lions, J.L., "Methode de Quasi-Reversibilite et Applications", Paris: Dunod, 1967.

[15] Lions, J.L., "Control Problems in Systems Described by Partial Differential Equations", in "Mathematical Theory of Control", New York: Academic Press, 1967.

[16] ——— "Functional Analysis and Optimization", A two week short course, July 31 - August 11, 1967 at U.C.L.A.

[17] Lions, J.L., Stampacchia, G., "Variational Inequalities", - to appear in the Comm. Pure and Applied Math.

[18] Lions, J.L., Magenes, E., "Remarques sur les Problemes aux Limites pour Operateurs Paraboliques", C.R. Acad. Sc., Paris, Vol. 25, (1960) pp. 2118-2120.

[19] Lovitt, W.V., "Linear Integral Equations", New York: Dover, 1950.

[20] Pearson, J.D., "On the Duality Between Estimation and Control", J. S.I.A.M. Control, Series A, 1966.

[21] Phillipson, G.A., "The State Identification of a Class of Distributed Systems", Ph.D. Thesis, Case Institute of Technology, June 1968.

[22] Pontryagin, L.S., et. al., "The Mathematical Theory of Optimal Processes", New York: Wiley, 1962.

[23] Schultz, M.H., Varga, R.S., "L-Splines", Numerische Mathematik, Vol. 10, pp. 345-369, 1967.

[24] Tikhonov, A.N., Samarskii, A.A., "Equations of Mathematical Physics", New York: Macmillan, 1963.

[25] Todd, J. Ed.. "Survey of Numerical Analysis", New York: McGraw Hill, 1962.

[26] Wang, P.K.C., "Control of Distributed Parameter Systems", in "Advances in Control Systems: Theory and Applications", New York: Academic Press, 1964, pp. 75-172.

[27] Wonham, W.M., "Some Applications of Stochastic Differential Equations to Nonlinear Filtering", J. S.I.A.M. Control, Series A, Vol. 2, No. 3, 1965.