IDENTIFICATION OF SPATIALLY-VARYING PARAMETERS IN DISTRIBUTED PARAMETER SYSTEMS

Thesis by

Costas Kravaris

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1984

(Submitted December 9, 1983)

Dedicated to my parents

Michalis and Evangelia Kravaris

ACKNOWLEDGEMENT

I am grateful to my advisor, Professor J. H. Seinfeld, for his advice, support and continuous encouragement throughout this work and also for giving me freedom with research and helping me to find a job. Also, I would like to thank Dr. P. T. Wu and his group (Drs. L. Padmanabhan, M. L. Wasserman, W. Bertiger and others) at Chevron Oil Field Research Company; my discussions with them have helped me to get exposed to major industrial problems in the area of reservoir engineering. Special thanks to Lenore Kerner and Fran Williams for the painstaking job of typing this thesis.

Last but not least, I would like to thank my wife Irene for her encouragement, patience and understanding and for making my life very happy.

ABSTRACT

Identification of spatially-varying parameters in distributed parameter systems given an observation of the state is as a rule an ill-posed problem in the sense of Hadamard. Even in case when the solution is unique, it does not depend continuously on the data. The identification problem that motivated this work arises in the description of petroleum reservoirs and subsurface aquifers; it consists of identifying the spatially-varying parameter $\alpha(x,y)$ in the diffusion equation $u_t = (\alpha u_x)_x + (\alpha u_y)_y + f$ given an observation of u at a discrete set of spatial locations.

The question of uniqueness of α (identifiability problem) is first investigated. The analysis is restricted to the one-dimensional version of the above equation i.e. to $u_t = (\alpha u_x)_x + f$ and an observation of u at a single point. The identifiability problem is formulated as an inverse Sturm-Liouville problem for $(\alpha y')' + \lambda y = 0$. It is proved that the eigenvalues and the normalizing constants determine the above Sturm-Liouville operator uniquely. Identifiability and non-identifiability results are obtained for three special cases.

The problem of constructing stable approximate solutions to identification problems in distributed parameter systems is next investigated. The concept of regularization, widely used in solving linear Fredholm integral equations, is developed for the solution of such problems. A general regularization identification theory is presented and applied to the identification of parabolic systems. Two alternative numerical approaches for the minimization of the smoothing functional are investigated: (i) classical Banach space gradient methods and

(ii) discretized minimization methods. The latter use finite-dimensional convergent approximations in Sobolev spaces and are based on an appropriate convergence theorem. The performance of the regularization identification method is evaluated by numerical experiments on the identification of spatially-varying diffusivity α in the diffusion equation.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

1. THE IDENTIFICATION PROBLEM OF PETROLEUM RESERVOIRS AND SUBSURFACE AQUIFERS AND ITS ILL-POSED NATURE

The classic identification problem arising in the modeling of petroleum reservoirs and subsurface aquifers is to estimate the parameter $\alpha(x,y)$ in the parabolic partial differential equation,

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x,y) \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\alpha(x,y) \frac{\partial u}{\partial y}\right) + q(x,y,t) \qquad (1)$$

based on noisy measurements of $u(x,y,t)$, $u^{obs}(x_i,y_i,t)$, at a set of discrete spatial locations, $i = 1,2,\ldots, m$. In such applications the dependent variable u represents pressure, q accounts for the withdrawal or injection of fluid in the region, and α is the transmissivity that determines the ease with which fluid flows in the reservoir. The initial condition to (1) is $u(x,y,0) = u_0(x,y)$, and a typical boundary condition is no fluid flow across the boundary of the region, i.e. $\frac{\partial u}{\partial \nu} = 0$, where ν represents the normal to the boundary. This identification problem has motivated the present thesis work.

Consider first the problem of solving (1) given $\alpha(x,y)$. It is well-known that under appropriate regularity assumptions for α, q and u_0 as well as strict positivity of α, there exists a unique solution $u(x,y,t)$ of the above boundary-value problem that depends continuously on α, q and u_0. Hence one can represent the solution u as

$$u = \mathscr{K}\alpha \qquad (2)$$

where $\mathscr{K}$ is some uniquely defined continuous mapping. Now the identification problem is the inverse of that just stated, namely given data on u, determine α. For this reason, identification problems of this sort are frequently referred to as *inverse problems*.

Assume that $u^{obs} = \mathscr{C}u$ is known, where $\mathscr{C}$ is an observation operator. For instance, $\mathscr{C}$ can be the operator that associates $u(x,y,t) \rightarrow u(x_i,y_i,t)$, $i = 1,2,\ldots, m$. The identification or inverse problem can be conceptually formulated as solving the (nonlinear) operator equation

$$(\mathscr{C} \circ \mathscr{K})(\alpha) = u^{obs} \tag{3}$$

It is natural to inquire:

(i) Does $\mathscr{C} \circ \mathscr{K}$ have a unique inverse?

(ii) Is the (unique or not) inverse operator $(\mathscr{C} \circ \mathscr{K})^{-1}$ continuous? Alternately stated, will small changes in u^{obs} result in arbitrarily small changes in α? If the answer to both questions is affirmative, then our problem is *well-posed* for all $u^{obs} \in \mathscr{C}(\mathscr{K}(\mathsf{F}))$, where F denotes the function space of α's. It is worth stressing that for the inverse problem to be well-posed it is not enough to ensure uniqueness of α. In fact, the crucial property will turn out to be the stability of the solution with respect to changes in the observation data. It will be shown later in this section that, even when the solution for α is guaranteed to be unique, the inverse problem is generally an unstable one.

The question of uniqueness of the identification problem is a very involved one. Many authors have speculated in the past that $\alpha(x,y)$ corresponding a given set of observations $u(x_i,y_i,t)$, $i = 1,\ldots, m$ is non-unique, the argument being that one cannot determine an infinite-dimensional parameter from a "finite" set of measurements. Such an argument is wrong, since the measurements are infinite-dimensional with respect to time and this permits (under certain assumptions) the unique determination of the spectrum of the differential operator $\nabla\cdot(\alpha(x,y)\nabla)$ and at the same time provides some information on its eigenfunctions. Thus, the question of uniqueness of the identification problem can be formulated within the framework of inverse Sturm-Liouville theory. This will be done in Chapter II. However, with the available mathematical tools, only special cases of the problem are solvable.

Although the complete discussion of these special cases will be postponed until Chapter II, we will consider one of these here. Consider the idealized situation of an one-dimensional one-phase reservoir and suppose that oil is extracted from just one well, located at $x = x_p$, and that the extraction started at $t = 0$. Then the governing equation is

$$\begin{aligned} &\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) = q(t)\,\delta(x-x_p) \\ &u(x,0) = u_0 = \text{constant} \\ &\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(\ell,t) = 0 \end{aligned} \qquad (4)$$

where $q(t)$ is the flow rate, ℓ is the "length" of the reservoir and $\delta(\cdot)$ is the Dirac delta "function".

If $u^{obs}(t) = u(x_p,t)$, $t \in [0,T]$ and $x_p = \frac{\ell}{2}$, then (see Chapter II) $\alpha(x)$ is *nonunique*. On the other hand, if $u^{obs}(t) = (u(0,t),\ u(x_p,t))$, then $\alpha(x)$ is *unique*.

We can now proceed to the question of continuous dependence of the (unique or not) α on the observation $u^{obs}(t) = (u(x_i,t),\ i = 1,\dots,\ m)$.

It is apparent from (1) that α depends on u only through its derivatives and it is well-known that given any function, a neighboring function can be selected that has its derivatives arbitrarily far from that of the given function. This ill-posedness manifests itself even in the numerical approximation of derivatives.

In order to rigorously establish non-continuous dependence of α on u one can use homogenization theory (Bensoussan, Lions and Papanicolaou, 1978). Homogenization theory shows that partial differential operators with highly oscillatory coefficients can be "replaced" by very different operators with flat coefficients and still yield practically the same response. In the above reference, one can find homogenization results for quite general partial differential operators. For the one-dimensional version of (1), the result is amazingly simple:

Proposition: Let a(y), $y \in \mathbb{R}$ be a function such that

$$\begin{cases} a(y) \text{ is bounded a.e. in } \mathbb{R} \\ a(y) \geq a_0 > 0 \quad \text{a.e. in } \mathbb{R} \\ a(y) \text{ is periodic with period } 1 \end{cases}$$

and denote
$$\begin{cases} \alpha^{\varepsilon}(x) = a\left(\frac{x}{\varepsilon}\right), \quad x \in [0,1], \quad \varepsilon > 0 \\ \overline{\alpha} = \dfrac{1}{\int_0^1 \frac{dy}{a(y)}} \end{cases}$$

Let $u^{\varepsilon}(x,t)$ be the solution of

$$\left.\begin{aligned}&\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha^{\varepsilon}(x)\,\frac{\partial u}{\partial x}\right) = f(x,t)\\&u(x,0) = u_o(x)\\&k_o\,\frac{\partial u}{\partial x}\,(0,t) + \ell_o\,u(0,t) = 0\\&k_1\,\frac{\partial u}{\partial x}\,(1,t) + \ell_1\,u(1,t) = 0\end{aligned}\right\}\tag{5}$$

and $\bar{u}(x,t)$ the solution of

$$\left.\begin{aligned}&\frac{\partial u}{\partial t} - \bar{\alpha}\,\frac{\partial^2 u}{\partial x^2} = f(x,t)\\&u(x,0) = u_o(x)\\&k_o\,\frac{\partial u}{\partial x}\,(0,t) + \ell_o\,u(0,t) = 0\\&k_1\,\frac{\partial u}{\partial x}\,(1,t) + \ell_1\,u(1,t) = 0\end{aligned}\right\}\tag{6}$$

Then $u_\varepsilon(x,t)$ converges *uniformly* to $\bar{u}(x,t)$ as $\varepsilon \to 0$.

The above lemma can be used to construct a broad class of counter-examples to disprove well-posedness of the identification problem. To see this, observe that as $\varepsilon \to 0$, α^{ε}'s are highly oscillatory functions whose amplitude can be selected to be arbitrarily large and whose maximum and minimum values can be arbitrarily far from $\bar{\alpha}$. For example, if

$$a(x) = \begin{cases}\dfrac{1}{1-\frac{1}{M}}\ , \text{ for } x \in [k,\ k + \frac{1}{2}[\ ,\ k \in \mathbb{Z}\\ M, \text{ for for } x \in [k + \frac{1}{2}\ ,\ k + 1[,\ k \in \mathbb{Z}\end{cases}$$

where M is a very big positive number, then $\alpha^{\varepsilon}(x)$ is a highly oscillatory "square-wave" function ranging from ~ 1 to M. On the other hand $\bar{\alpha} = 2$.

It is noteworthy that Lions (1978), based on homogenization results, has cited the main difficulty in identifying distributed coefficients in partial differential equations as preventing excess of oscillations in the coefficients. To overcome this difficulty, one has to somehow impose a certain degree of smoothness on the estimate based on physical considerations. In other words, one has to provide some extra qualitative information on the unknown solution.

As a concluding remark of this section, we want to emphasize that the essential difficulty associated with estimating α in (1) is to alleviate the inherent ill-posedness of the problem.

In the next two sections of this introductory chapter we will study this identification problem with special attention to its ill-posedness or ill-conditioning. §2 summarizes prior literature and §3 describes the basic approaches that are available. Because of our focus on the ill-posed nature of the current problem, we present a brief review of methods of solution of general ill-posed problems in §4. These methods have motivated the regularization identification approach developed in Chapters III and IV of the present thesis. Finally, in §5 we give an outline of the thesis.

2. SURVEY OF LITERATURE IN RESERVOIR AND AQUIFER IDENTIFICATION

The identification problem of interest has been defined in the previous section, namely to estimate $\alpha(x,y)$ in (1) based on spatially discrete, noisy measurements of $u(x,y,t)$, $u^{obs}(x_i,y_i,t)$, $i = 1,2,\ldots, m$. It appears that two principal lines of attack have been followed with regard to this problem. One consists of regarding (1) as a first-order partial differential equation in α, where u and its spatial and temporal derivatives are assumed to be known over the entire region based on the available data. Several authors have referred to this as the *direct approach*. The uniqueness of the equation for α can be insured by prescribing the value of α along a curve Γ that cuts all the streamlines of the flow or by prescribing the flow, rather than α, at every streamline. With either condition, the problem reduces simply to an initial-value problem for α. The second main line of attack, often referred to as the *indirect approach*, consists of a trial and error procedure that seeks to improve an existing estimate of α in an iterative manner until the predicted u is sufficiently close to the observed u. Specifically, the identification problem is formulated as the minimization of a suitable measure of the difference between observed and predicted $u(x_i,y_i,t)$. The so-called indirect approach seems to offer several advantages over the direct approach in that it is far less demanding in terms of the necessary observation data, specifically, the spatial and temporal derivatives of u over the region need not be known, only the temporal histories of u at the observation locations. The indirect approach is, in fact, the line of attack of most distributed parameter identification problems (Kubrusly, 1977), and is the approach on which we focus our attention later in this chapter.

Table 1 presents a list of many of the prior studies on the current identification problem. Upon a survey of the literature it becomes apparent that two more or less independent lines of research on this problem have existed, one associated with petroleum reservoirs and one directed toward subsurface hydrological systems, although from time to time one finds cross references between the two bodies of work. The papers listed in the table are divided into three major topics, direct methods, indirect methods, and those devoted to specific consideration of the ill-posed nature of the problem and the statistical aspects of the parameter estimation. These two aspects are treated together because a great deal of insight into the ill-conditioning of the problem is available from a statistical analysis of the estimates. In the papers adopting the indirect approach a large variety of minimization methods have been employed. Because such methods have received extensive coverage in several textbooks and review papers they need not be reviewed here (Eykhoff, 1974; Seinfeld and Lapidus, 1974; Beck and Arnold, 1977; Kubrusly, 1977). In the next section methods for solution of the identification problem are discussed, with emphasis on dealing with the ill-conditioning.

TABLE 1 Available Literature on the Estimation of Transmissivity $\alpha(x,y)$ in

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\alpha \frac{\partial u}{\partial y}\right) + q$$

Reference	Approach
Stallman (1963) Sammel (1964) Nelson (1968) Nelson and McCollum (1969) Deininger (1969) Kleinecke (1971) Frind and Pinder (1973) Sagar et al. (1975) Yakowitz and Noren (1976)	*Direct approach* in which (1) is regarded as a first-order p.d.e. in $\alpha(x,y)$ where u and its spatial and temporal derivatives are assumed known over the entire region. Later papers have employed finite-element approximations for u and α, e.g. Frind and Pinder (1973)
Jacquard and Jain (1965) Jahns (1966) Haimes et al. (1968) Vermuri and Karplus (1969) Vemuri et al. (1969) Coats et al. (1970) Slater and Durrer (1971) Yeh and Tauxe (1971ab) Chen and Seinfeld (1972) Marino and Yeh (1973) Carter et al. (1974) Chen et al. (1974) Chavent et al. (1975) Distefano and Rath (1975) Wasserman et al. (1975) Chang and Yeh (1976) Gavalas et al. (1976) Cooley (1977, 1979) Shah et al. (1978) Seinfeld and Chen (1978)	*Indirect approach* in which $\alpha(x,y)$ is determined to minimize a performance index such as $J = \int_0^T \sum_{i=1}^{m} [u^{obs}(x_i,y_i,t) - u(x_i,y_i,t)]^2 \, dt$ subject to (1). Many of the papers employ *zonation* in which the region is divided into N subregions (zones) having uniform α. Minimization methods used include direct gradient methods, steepest descent algorithms derived from optimal control theory, and linear programming.
Emsellem and de Marsily (1971) Lovell et al. (1972) Neuman (1973,1980ab) Gavalas et al. (1976) Yeh and Yoon (1976) Shah et al. (1978) Neuman and Yakowitz (1979) Yakowitz and Duckstein (1980)	Treatments of the ill-posed nature of the problem based on concepts of Bayesian estimation and incorporation of prior information about α into the problem. Statistical analysis used to determine optimum number of parameters, e.g. Yeh and Yoon (1976) and Shah et al. (1978).

3. METHODS FOR SOLUTION OF THE IDENTIFICATION PROBLEM

The solution of the identification problem can be divided into two steps: (i) formulation of the problem, including the performance index and the manner of representing $\alpha(x,y)$ and (ii) development of a method of actually carrying out the estimation. As noted above, the key difficulty is dealing with the ill-conditioning of the basic identification problem, and that aspect must generally be faced when formulating the problem. As noted, methods of actual parameter estimation have been discussed widely and need not be reviewed here. Therefore, in this section we focus on methods for solution of the identification problem that specifically address the matter of ill-conditioning.

The only way to improve the ill-conditioning of the identification problem is to incorporate additional a priori information into the problem beyond that available in the model (1) and the observational data. Such a priori information may take the form of assumptions concerning the nature of α, such as its spatial variation or an assumed functional form to which α is to adhere, or the a priori information may be introduced through terms added to the identification performance index. All of these approaches act to restrict the space of functions within which the estimate of α may lie. Thus, the essence of how one alleviates the ill-conditioning of the identification problem is how one incorporates a priori information into it.

Zonation

Since (1) is frequently solved numerically on a grid lying over the region, an obvious first approach is to assume that $\alpha(x,y)$ takes on different values in each grid cell. The number of such grid cells for even a modest sized reservoir can be quite large (the order of 10^3), leading to an equally large number of unknown parameters. It is well known that when one attempts to estimate such a large number of parameters, even though the optimization routine may lead to an "optimal" set of estimates, the accuracy associated with the estimates is poor (Gavalas et al., 1976; Shah et al., 1978).

The use of an array of grid cells to characterize the spatial heterogeneity of a reservoir is, of course, a computational convenience. That the spatial heterogeneity of the parameters should conform to the same grid is only a computational convenience. When this approach is used, however, not only do the estimated parameters lack uniqueness, but the estimated distributions do not preserve the smoothness features inherent in the geologists' concept of the reservoir. A common way to alleviate the problem of nonuniqueness is what is sometimes referred to as *zonation*, that is simply requiring that the parameters be uniform over regions of the reservoir that contain multiple numbers of grid cells (Jacquard and Jain, 1965; Jahns, 1966; Coats et al., 1970; Carter et al., 1974; Gavalas et al., 1976; Cooley, 1977, 1979).

When prior information about the transmissivities is not taken into account, the zonation approach makes intuitive sense, and, if the number N of zones is sufficiently small, should lead to a well-behaved solution. It is important to note, however, that when N is small, the modeling error becomes large since the true α distribution cannot be represented even if the solution is well-behaved. The question is how large N can become before ill-conditioning sets in, and what should be the optimum size and shape of each zone. The determination of an optimal zoning pattern was first addressed by Emsellem and de Marsily (1971), who proposed solving a sequence of inverse problems starting with a small number of zones and gradually increasing the number until there is no improvement in model fit. Distefano and Rath (1975) adopted a criterion for deciding on the optimal number of zones based on visual observation of the spatial behavior of the estimated α field. Yeh and Yoon (1976) were apparently the first to introduce statistical notions into the stepwise zoning process. They used the variance of the estimation error as a criterion for deciding whether a particular zone should be subdivided or kept intact at each step. Shah et al. (1978) developed a detailed theoretical basis for the selection of an optimum number of zones. The trace of the covariance matrix used as a measure of the over-all accuracy of estimation was studied as a function of N, and a procedure for selecting the optimum parametrization was developed. A brief summary of that analysis is presented later in this section.

Neuman (1973) suggested that incorporating prior information about the transmissivities directly into the statement of the identification problem could alleviate ill-conditioning. A similar idea was advanced by Gavalas et al. (1976) who formulated the identification problem in a Bayesian framework. By using a priori statistical information on the unknown parameters, the problem becomes statistically better determined. Shah et al. (1978) showed that when reliable a priori statistics are available, the Bayesian approach is superior in that it leads to a smaller variance of the estimation error. Neuman and Yakowitz (1979) employed statistical theory in estimating transmissivities on the basis of steady state data on u. They formulated the identification problem as one of minimizing a performance index consisting of two terms, one, the mismatch between predictions and observations, and second, a Bayesian-like term accounting for deviations from a priori estimate of α, similarly to Gavalas et al. (1976). The selection of the arbitrary scaling parameter that multiplies the second term of the performance index was studied in some detail. When reliable a priori statistics are available, the Bayesian approach has been shown to be effective in reducing the total estimation error. Such statistics may not be available however.

An alternative to zonation is to specify the functional form obeyed by α up to a set of unknown coefficients. For example, Distefano and Rath (1975) used zonation and then assumed that α is a linear function of position in each zone.

In either instance, the original problem of estimating $\alpha(x,y)$, a continuous function of spatial position, is converted into one of estimating a finite number of parameters, $\alpha_1, \alpha_2, \ldots, \alpha_N$. The ill-conditioned nature of the problem is, of course, not circumvented if N is sufficiently large. The underlying nature of the problem can be nicely illustrated by considering its statistical aspects.

Statistical Aspects of the Identification Process

We have expressed the solution of (1) as $u = \mathcal{K}\alpha$, and have discussed the problems arising because of the nonuniqueness and lack of continuity of the inverse operator. For the purpose of illustrating the ill-conditioning of the basic identification problem, let us assume that (1) has been discretized both spatially and temporally, so that $u = A\alpha$, where u and α are vectors and A is a matrix, all of appropriate dimension.

If A is invertible (this corresponds to dim u = dim α = rank A), then the solution is $\alpha = A^{-1}u$. When rank A $<$ dim α, we have an over-determined system of algebraic equations; the least-squares estimate is given by

$$\hat{\alpha} = A^{+}u \tag{7}$$

where A^{+} is the pseudo-inverse

$$A^{+} = (A^{*}A)^{-1}A^{*} \tag{8}$$

ane where A^{*} is the adjoint of A. When dim $\alpha >$ dim u, u = Aα has infinitely many solutions given by $\hat{\alpha} = A^{+}u + (I-A^{+}A)\kappa$ where κ is an arbitrary vector with dim κ = dim α. The first term is a particular solution and the second represents any element from the null space of A. Since κ is arbitrary, the estimate $\hat{\alpha}$ may also be arbitrarily different from the true value α. In other words, the estimation error has no upper bound. The dimension of the null space is the number of zero eigenvalues of $A^{*}A$. In this sense, A is singular. In a practical sense, overparametrization can occur even without A being strictly singular. It is enough if $A^{*}A$ has one or more of its eigenvalues much smaller than the rest. (A would then be ill-conditioned.) The parameters along the associated eigenspace are virtually indeterminable. As these eigenvalues approach zero, in the limit the eigenspace becomes the null space. In the eigenspace a slight variation in the data u would be interpreted as being induced by a large change in the parameters, since we have $\delta\hat{\alpha} = A^{+}(\delta u)$. Here A^{+} has large eigenvalues in the eigenspace so that $\| \delta\hat{\alpha} \|$ is large even though $\| \delta u \|$ is small. The estimation is ill-posed even though A is invertible in the strict sense.

We have noted that one approach to achieving better-conditioned estimates is to reduce the number of unknown parameters. This process can be represented as replacing the parameter vector α by its restriction to a smaller subspace of the parameter space F. If we characterize this subspace by its basis vectors M, the parameter space can be resolved into M and its orthogonal complement M' in F. Correspondingly, we can resolve α into α_M and $\alpha_{M'}$ by $\alpha = \alpha_M + \alpha_{M'}$, where $\alpha_M = MM^+\alpha$, the projection of α into M. The replacement of α by α_M leads to so-called *parameterization error.* Since α_M lies in a smaller space, we may characterize it as $\alpha_M = M\zeta$, in terms of a vector ζ with fewer components. Thus, the estimation of α is replaced as that of ζ. If the data are represented as $u^o = u + \eta$, where η is the error vector, then

$$
\begin{aligned}
u^o &= u + \eta = A\alpha + \eta \\
&= A(M\zeta + \alpha_{M'}) + \eta \\
&= [(AM)\zeta] \qquad + [A\alpha_{M'}] \qquad\qquad + \eta \qquad\qquad (9) \\
&= \underset{\text{Model}}{\text{Approximate}} \quad \underset{\text{Error}}{\text{Parameterization}} \quad \underset{\text{Error}}{\text{Measurement}}
\end{aligned}
$$

which shows the composition of the observed quantity u^o. The estimate is $\hat{\zeta} = (AM)^+u^o$, or, in terms of the original parameter vector, $\hat{\alpha} = M(AM)^+u^o$.

The error in matching observed and predicted performance can be defined as

$$J(\hat{\alpha}) = E\{\| u^o - A\hat{\alpha} \|^2\} \tag{10}$$

where $E\{\cdot\}$ is the expectation. Eq. (10) thus becomes

$$
\begin{aligned}
J(\hat{\alpha}) &= E\{\| [I - (AM)(AM)^+]u^o \|^2\} \\
&= \text{trace}\{\rho_{(AM)}\ E(u^o u^{o*})\rho_{(AM)}\} \qquad (11)
\end{aligned}
$$

where $\rho_{(AM)} = I - (AM)(AM)^+$ is the projection into the null space of $(AM)^+$. The expectation $E(u^o u^{o*})$ is $E\{(A\alpha + \eta)(A\alpha + \eta)^* = A\alpha\alpha^* A^* + E(\eta\eta^*)$. We see that the estimation error can be reduced if $\rho_{(AM)} \to 0$, i.e., as (AM) assumes full rank. In other words, by increasing the number of parameters, it is possible to produce a perfect match of observations and predictions, regardless of measurement errors.

The parameter estimation error is defined as

$$J_\alpha(\hat{\alpha}) = E\{ \| \hat{\alpha} - \alpha \|^2 \} \tag{12}$$

The estimate $\hat{\alpha}$ is given by

$$\begin{aligned} \hat{\alpha} &= M(AM)^+ u^o \\ &= M(AM)^+[A(\alpha_M + \alpha_{M'}) + \eta] \end{aligned} \tag{13}$$

Thus,

$$J_\alpha(\hat{\alpha}) = \| \alpha_{M'} \|^2 + \text{trace } \{M(AM)^+ \; E(\eta\eta^*)(M(AM)^+)^*\} \tag{14}$$

which consists of two terms, the parametrization error and the measurement error. The parametrization error, $\| \alpha_{M'} \|^2$, approaches zero as ν = rank M approaches dim α, i.e., as the number of parameters is increased. The second term contains $M(AM)^+$, which has the eigenvalues $1/\mu_1$, $1/\mu_2, \ldots,$ $1/\mu_\nu$, where $\{\mu_i^2\}$ are the largest ν eigenvalues of A*A. As ν increases, the degree of ill-conditioning of $M(AM)^+$ also worsens because $1/\mu_\nu$ increases with ν. Thus, the effect of measurement error is amplified as the number of parameters is increased. For every problem there is an optimum level of parametrization measured by ν at which $J_\alpha(\hat{\alpha})$ reaches a minimum. Padmanabhan (1980) has included modeling error in the above analysis. Although we have not considered modeling error here, i.e. the inexactness of A, it is frequently of more importance that measurement error in the identification problem.

4. REGULARIZATION METHODS FOR THE SOLUTION OF ILL-POSED PROBLEMS

Introduction

In the previous sections we have discussed the ill-posed nature of the identification problem arising from nonuniqueness and instability of solutions. No mathematical method can really alleviate nonuniqueness unless it selects some solution that has certain physically desirable properties (e.g. smoothness). However, as far decreasing instability is concerned, mathematical methods do have a great potential.

The regularization approach is a stability-oriented approach, with which one obtains stable approximate solutions to ill-posed problems that are in accordance with the level of error in the data. The regularization method has two crucial properties:

i) In problems with inherent nonuniqueness, it selects (for a given level of error) the smoothest solution, in the sense that it corresponds to the least value of an appropriately chosen functional (called the stabilizing functional).

ii) In problems where no nonuniqueness is present, the approximate solution it yields converges to the true solution in the limit as the level of error in the data tends to zero.

The Concepts of Well-Posed and Ill-Posed Problems

Consider the operator equation

$$Az = u \tag{15}$$

where u and z are elements of the metric spaces U and F with metrics

$\rho_U(u_1,u_2)$ for $u_1,u_2 \in U$ and $\rho_F(z_1,z_2)$ for $z_1,z_2 \in F$. We say that a solution $z = R(u)$ of (15) is *stable* on the spaces (F,U) if for every positive number ε, there exists a positive number $\delta(\varepsilon)$ such that the inequality $\rho_U(u_1,u_2) \leq \delta(\varepsilon)$ implies $\rho_F(z_1,z_2) \leq \varepsilon$, where $z_1 = R(u_1)$ and $z_2 = R(u_2)$ with u_1 and u_2 in U and z_1 and z_2 in F.

The problem of solving (15) is said to be *well-posed* on the pair of metric spaces (F,U) if the following three conditions are satisfied:

i) for every $u \in U$ there exists a solution $z \in F$

ii) the solution is unique

iii) the solution is stable on the spaces (F,U).

Problems that do not satisfy all the above conditions are said to be *ill-posed*. Note that existence of a solution is assured if and only if $U \subset AF$; uniqueness if and only if A is injective; stability if and only if A^{-1} is continuous. The (exact) solution of a well-posed problem of the form (15) is simply

$$z = A^{-1}u \tag{16}$$

For a long time, it was the accepted point of view that any mathematical problem corresponding to any physical problem has to be well-posed. This point of view, while natural as applied to certain phenomena that have been studied over the years, is not valid for all problems. It is well-established now that a great number of physically meaningful problems are mathematically ill-posed.

In most practical problems, the RHS of (15) is not known exactly. Rather, we have an approximation u_δ of the true, u, u_T, with accuracy δ, i.e. $\rho_U(u_\delta,u_T) \leq \delta$. It is natural then to seek an approximate solution to (15) in the class Q_δ of elements z for which $\rho_U(Az,u\ ) \leq \delta$. The question is whcih element of Q_δ to choose. If the problem is well-posed, we may choose

$$z = A^{-1}u_\delta \tag{17}$$

If the problem is ill-posed, on the other hand, the result (17) is of little value.

<u>The Regularization Method</u>

Consider again the solution of the operator equation (15) with approximately known right hand side $u_\delta(\rho_U(u_\delta,u_T) \leq \delta)$ and suppose that A^{-1} is unique but not continuous. Then, an approximate solution z_δ of (15) cannot be obtained by (17).

Since the numerical parameter δ characterizes the error in the right-hand side of (15), it is natural to define z_δ with the aid of an operator depending on a parameter having a value chosen in accordance with the error δ in the data u_δ. Specifically, as $\delta \to 0$, i.e. as u_δ approaches (in the metric of the space U) the exact value u_T, the approximate solution z_δ must approach (in the metric of the space F) the exact solution z_T.

Suppose that the elements $z_T \in F$ and $u_T \in U$ are related by $Az_T = u_T$. An operator $R(u,\beta)$ depending on a parameter β is said to be a *regularizing operator* for the equation $Az = u$ in a neighborhood of $u = u_T$ if:

(i) There exists $\delta_1 > 0$ such that $R(u,\beta)$ is defined for every $\beta > 0$ and every $u \in U$ for which $\rho_U(u,u_T) \leq \delta \leq \delta_1$, and

(ii) There exists a function $\beta = \beta(\delta)$ such that, for every $\varepsilon > 0$, there exists a number $\delta(\varepsilon) \leq \delta_1$ such that $u_\delta \in U$ and $\rho_U(u_T,u_\delta) \leq \delta(\varepsilon)$ imply that $\rho_F(z_T,z_\beta) \leq \varepsilon$, where $z_\beta = R(u_\delta,\beta(\delta))$.

If $\rho_U(u_T,u_\delta) \leq \delta$, we can take for an approximate solution of (15) with approximate RHS u_δ the element $z_\beta = R(u_\delta,\beta)$ obtained with the aid of the regularizing operator $R(u,\beta)$, where $\beta = \beta(\delta,u_\delta)$ is in accordance with the error in the initial data u_δ. This solution is called a *regularized solution* of (15), and the numerical parameter β is called the *regularization parameter*. This method of constructing approximate solutions is called the *regularization method*.

Tikhonov Regularization

Consider the operator equation

$$Az = u \tag{18}$$

where A is a Fredholm integral operator of the first kind, i.e. $Az = \int_a^b K(x,\zeta)z(\zeta)\, d\zeta$. The problem of solving (18) with K being any measurable kernel is an ill-posed one. This is a consequence of the Riemann-Lebesque lemma.

$$\int_a^b K(x,\zeta)\sin(\omega\zeta)\, d\zeta \to 0 \text{ as } \omega \to +\infty$$

In fact, if u_1 and z_1 satisfy (18) and $u_2(x) = u_1(x) + p \int_a^b K(x,\zeta)\sin(\omega\zeta)d\zeta$, then the difference $u = u_2-u_1$ can be arbitrarily small for sufficiently large ω. However, the corresponding change in the solution $\delta z = z_2 - z_1 = p \sin(\omega x)$ can be arbitrarily large with an appropriate choice of p.

Now suppose that u_δ is a measurement of u with mean square error $\leq \delta$, i.e.

$$\| u_\delta - u_T \| = \left(\int_a^b [u_\delta(x) - u_T(x)]^2 dx \right)^{\frac{1}{2}} \leq \delta$$

where u_T denotes the exact value of u, and let $\Omega(z)$ be a functional, called *p-th order stabilizing functional*, of the form

$$\Omega(z) = \int_a^b \sum_{r=0}^{p} q_r(\zeta) \left(\frac{d^r z}{d\zeta^r}\right)^2 d\zeta \tag{19}$$

where $q_r(\zeta)$ are continuous strictly positive functions.

Tikhonov (1963a, 1963b) proposed the minimization of the so-called *smoothing functional*

$$M^\beta(z,u_\delta) = \| Az - u_\delta \|^2 + \beta\Omega(z) \tag{20}$$

as a means of constructing regularized solutions to the operator equation (18).

He showed:

Result 1: For any square-integrable function u_δ and for every $\beta > 0$ there exists a unique continuous p-times differentiable function z_β that minimizes the smoothing functional $M^\beta(z,u_\delta)$.

Result 2:

$$\text{If} \begin{cases} Az_T = u_T, \text{ with } \Omega(z_T) < \infty \\ u_\delta \text{ is such that } \| u_\delta - u_T \| \leq \delta \\ z_\delta \text{ minimizes } M^\beta(z,u_\delta), \text{ where } c_1\delta^2 \leq \beta \leq c_2\delta^2 \\ \qquad \text{with } 0 < c_1 < c_2 \end{cases}$$

then z_δ converges uniformly to z as $\delta \to 0$. Also $\frac{d^r z}{dx^r}$ converges uniformly to $\frac{d^r z}{dx^r}$, for $1 \leq r \leq p-1$.

Tikhonov's results are readily extendable to the case where A is an arbitrary bounded linear operator from a Banach space F into a Banach space U with unbounded inverse (Ivanov, 1966; Morozov, 1966). In this case one can construct regularized solutions of (18) by minimizing the smoothing functional

$$M^{\beta}(z,u_{\delta}) = \| Az - u_{\delta} \|_{U}^{2} + \beta\Omega(z)$$

where

$$\Omega(z) = \| z \|_{W}^{2}$$

with W being a Hilbert space compactly imbedded in F. One can show:

Result 1: For every $u_{\delta} \in U$ and $\beta > 0$ there exists a unique $z_{\beta} \in W$ that minimizes the smoothing functional $M^{\beta}(z,u_{\delta})$.

Result 2:

$$\text{If} \quad \begin{cases} Az_T = u_T, & \text{with } z_T \in W \\ u_{\delta} \text{ is such that } \| u_{\delta} - u_T \|_{U} \leq \delta \\ z_{\delta} \text{ minimizes } M^{\beta}(z,u_{\delta}), & \text{where } c_1\delta^2 \leq \beta \leq c_2\delta^2 \\ & \text{with } 0 < c_1 < c_2 \end{cases}$$

then $\| z_{\delta} - z \|_{F} \to 0$ as $\delta \to 0$.

The regularization approach can be further extended to a class of nonlinear problems (Tikhonov and Arsenin, 1977). We will not discuss this extension at this point. The reader is referred to Chapter III of this work, where we present a generalization of the Tikhonov-Arsenin theory which is suitable for identification problems in partial differential equations.

The Approach of Miller

Consider again the problem of solving an operator equation of the form (15). Miller (1970) reformulated the problem as follows:

Given the Hilbert spaces F, U and W, the bounded linear operators $A: F \to U$ and $B: F \to W$, and an element $u \in U$, and assuming that B^{-1} exists and is bounded, suppose that $z_T \in F$ satisfies

$$\| Az_T - u \| \leq \delta \tag{21}$$

$$\| Bz_T \| \leq \Delta \tag{22}$$

The problem is to find an element $\tilde{z} \in F$ that approximates z_T in the sense that the seminorm $<\tilde{z}-z_T>$ is small for δ small.

Consider first the case where both δ and Δ are known. Miller defines

$$\mu(\delta,\Delta) = \sup\{<z>:z \in F, \| Az \| \leq \delta, \| Bz \| \leq \Delta\}$$

$$\mu_1(\delta\ \Delta) = \sup\{<z>:z \in F, \| Az \|^2 + \left(\frac{\delta}{\Delta}\right)^2 \| Bz \|^2 \leq 2\delta^2\}$$

and shows that if z_T satisfies (21) and (22) then it also satisfies

$$\| Az_T - u \|^2 + \left(\frac{\delta}{\Delta}\right)^2 \| Bz_T \|^2 \leq 2\delta^2 \tag{23}$$

Conversely, any z_T satisfying (23) also satisfies (21) and (22), except for a factor of at most $\sqrt{2}$. Furthermore,

$$\mu(\delta,\Delta) \leq \mu_1(\delta,\Delta) \leq \sqrt{2}\, \mu(\delta,\Delta)$$

This result suggests that we may combine the two constraints (21) and (22) and sacrifice at most a factor of $\sqrt{2}$.

<u>Method</u> 1. Miller proposed to take as an approximate solution the element $\tilde{z} \in F$ that minimizes

$$\| Az-u \|^2 + \left(\frac{\delta}{\Delta}\right)^2 \| Bz \|^2$$

This element is given by

$$\tilde{z} = \left(A^*A + \left(\frac{\delta}{\Delta}\right)^2 B^*B \right)^{-1} A^*u$$

where superscript * denotes the adjoint of an operator. Note that the operator $(A^*A + \left(\frac{\delta}{\Delta}\right)^2 B^*B)$ has a bounded inverse under the assumptions made for the operators A and B. This guarantees stability of the approximate solution $\tilde{z}$.

Miller also gives an error estimate

$$<\tilde{z} - z_T> \leq \mu_1(\delta,\Delta)$$

provided that z_T satisfies (23).

Next Miller presents a more general approach for which it is not necessary that both δ and Δ are specified. A pair (δ,Δ) is said to be *permissible* if there exists a $z \in F$ satisfying (21) and (22).

Now let z_β be a minimizer of

$$\| Az-u \|^2 + \beta \| Bz \|^2$$

and let $\delta_\beta = \| Az_\beta-u \|$ and $\Delta_\beta = \| Bz_\beta \|$. Clearly z_β minimizes $\| Az-u \|$ with respect to the constraint $\| Bz \| \leq \Delta_\beta$. Likewise z_β minimizes $\| Bz \|$ with respect to the constraint $\| Az-u \| \leq \delta_\beta$. It is also easily seen that δ_β and Δ_β are continuously increasing and decreasing functions of β, respectively. Thus, the set E of permissible pairs is exactly the set of points that are above and to the right of the curve $(\delta_\beta,\Delta_\beta)$, $0 \leq \beta \leq \infty$

(Figure 1). Here the case $\beta = 0$ corresponds to the minimization of $\| A\bar{z}-u \|$ alone, in which case Δ_β may be ∞, and the case $\beta = \infty$ corresponds to minimization of $\| Bz \|$ alone, in which case $\delta_\beta = \| u \|$ and $\Delta_\beta = 0$. Moreover, it is easily shown that E is a convex set, hence computation of only a finite number of points on its boundary curve is needed to ascertain the curve.

Method 2. If δ and Δ are known, the unknown solution must lie in the shaded area of Figure 1. Furthermore, any z_β whose corresponding $(\delta_\beta, \Delta_\beta)$ touches the shaded area can be taken as an approximate solution. Note that $\langle z_\beta - z_T \rangle \leq 2\mu(\delta,\Delta)$.

Method 3. Assume that only δ is known. Miller proposes to take as an approximate solution $\tilde{z}$ that element of F that minimizes $\| Bz \|$ subject to $\| Az-u \| \leq \delta$. Then $\tilde{z} = z_{\beta'}$, where β' is the value of β such that $\delta_{\beta'} = \delta$, as shown in Figure 1.

Method 4. Assume that only Δ is known. Miller proposes to take as an approximate solution $\tilde{z}$ that element of F that minimizes $\| Az-u \|$ subject to $\| Bz \| \leq \Delta$. Then $\tilde{z} = z_{\beta''}$, where β'' is the value of β such that $\Delta_{\beta''} = \Delta$, as shown in Figure 1.

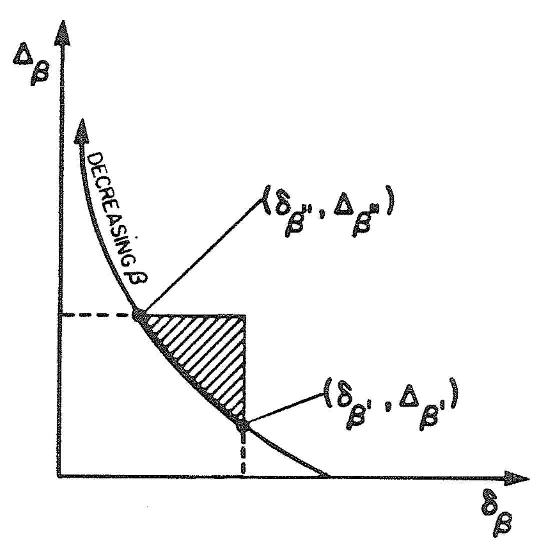


Figure 1

Discussion. Upon examining Method 1, we note that Miller's theory relies heavily on the fact that A is a linear operator, whereas the regularization method of the previous section is applicable to nonlinear operators as well. $\| Bz \|^2$ can be interpreted as a stabilizing functional, but the assumption that B is bounded is too restrictive for the problems of interest. The operator B corresponding to stabilizing functionals of the form (19) is unbounded. The regularization parameter $(\frac{\delta}{\Delta})^2$ is a quadratic function of the error in the data, which is consistent with Result 2 $(c_1 = c_2 = \frac{1}{\Delta^2})$.

In methods 2-4, $\| Az-u \|^2 + \beta \| Bz \|^2$ stands for a smoothing functional with β the regularization parameter. Methods 2-4 suggest an interesting method of selection of the regularization parameter. Note however that if the operator A is nonlinear, the set E will no longer in general be convex and the functions $\delta_\beta(\beta)$ and $\Delta_\beta(\beta)$ no longer single-valued. Hence, for a nonlinear operator A it is possible that β' and β'' do not exist for given δ and Δ. This fact creates a serious complication when dealing with nonlinear problems, in which case the local convexity of E must be investigated either theoretically or numerically.

Wahba's Version of the Regularization Method - Weighted Cross Validation

Wahba's approach to regularization consists of constructing regularized approximate solutions by minimizing a smoothing functional. The difference of this approach from the previously described ones is the method of selection of the regularization parameter. The method, termed *weighted cross-validation* does not require any a priori statistical information concerning measurement error. The weighted cross-validation technique seems to have originated from problems of interpolation, and has been applied to solution of Fredholm integral equations of the first kind. The idea is quite general, however, so that extension to general nonlinear ill-posed problems seems possible but far from trivial (Neuman and Yakowitz, 1979; Wahba, 1980).

Consider for the moment the problem of solving the operator equation

$$Az(x) = u(x) \tag{24}$$

from discrete noisy measurements u_1^{obs}, u_2^{obs}, ..., u_n^{obs}, of $u(x)$ at the points x_1, x_2, ..., x_n. If A is the identity operator, then this is the standard interpolation problem. (Note that although the identity operator is obviously continuous, the interpolation problem is ill-posed since solutions are not unique). The interpolation problem can be approached by the regularization method. Assuming that the true $z(x)$ has continous second derivatives and choosing

$$\int_{x_1}^{x_n} [z''(x)]^2 \, dx$$

as a stabilizing functional, one can obtain regularized solutions to the interpolation problem by minimizing the smoothing functional

$$M^{\beta}(z,u^{obs}) = \frac{1}{n} \sum_{i=1}^{n} \left[u_i^{obs} - z(x_i)\right]^2 + \beta \int_{x_1}^{x_n} \left[z''(x)\right]^2 dx \tag{25}$$

It can be shown that the solution of the minimization problem leads to a natural cubic spline,

$$s(x) = a_i + b_i(x-x_i) + c_i(x-x_i)^2 + d_i(x-x_i)^3$$
$$x \in [x_i, x_{i+1}] \tag{26}$$

Expressions for the coefficients are given in the literature (Reinsch, 1967; Merz, 1978). For this discussion we are interested only in the expression for $a = (a_1, a_2, \ldots, a_n)$.

Let P be a symmetric, tridiagonal, (n-2)x(n-2) dimensional matrix with elements,

$$p_{ii} = \frac{2}{3}(x_{i+2}-x_i) \qquad p_{i,i+1} = p_{i+1,i} = \frac{1}{3}(x_{i+2}-x_{i+1})$$

and let Q be the tridiagonal nx(n-2) dimensional matrix with elements

$$q_{ii} = \frac{1}{x_{i+1}-x_i}, \quad q_{i+1,i} = -\left[\frac{1}{x_{i+1}-x_i} + \frac{1}{x_{i+2}-x_{i+1}}\right], \quad q_{i+2,i} = \frac{1}{x_{i+2}-x_{i+1}}$$

then (Merz, 1978)

$$a = Lu^{obs} \tag{27}$$

where $L = I - Q(Q^TQ + n\beta P)Q^T$

Reinsch (1967) suggested that if the standard deviation σ of the errors in the data is approximately known, β should be chosen such that

$$\frac{1}{n}\sum_{i=1}^{n}(u_i^{obs} - s(x_i))^2 = \sigma^2$$

or

$$\frac{1}{n}\,\|(I-L(\beta))u^{obs}\|^2 = \sigma^2 \tag{28}$$

where $\|\cdot\|$ denotes Euclidian norm. This method for determining β gives excellent results for most practical purposes, but breaks down whenever error statistics are unavailable.

The determination of β by cross-validation proceeds as follows (Wahba and Wold, 1975; Craven and Wahba, 1979). Let $s_\beta^{[k]}$ be the spline using all data points except the kth. We take the ability of $s_\beta^{[k]}$ to predict the missing data point u_k^{obs} as a measure of the effectiveness of β. An overall measure of this effectiveness is the value of

$$V_0(\beta) = \frac{1}{n} \sum_{k=1}^{n} \left[s_\beta^{[k]}(x_k) - u_k^{obs} \right]^2 \qquad (29)$$

It can be shown from (27) and some algebra that

$$V_0(\beta) = \frac{1}{n} \sum_{k=1}^{n} \frac{[u_k^{obs} - a_k(\beta)]^2}{[1-L_{kk}(\beta)]^2} \qquad (30)$$

a weighted sum of squared residuals of the fit to the data.

One could use the minimizer of $V_0(\beta)$ as the optimum β, a procedure known as ordinary cross-validation. Using statistical arguments, Wahba suggested that ordinary cross-validation be used only for periodic splines with equally-spaced data points. Note that in this case the weights $(1-L_{kk}(\beta))^{-2}$ of squared residuals are all equal. In the general case one should minimize

$$V(\beta) = \frac{1}{n} \sum_{k=1}^{n} \left[s^{[k]}(x_k) - u_k^{obs} \right]^2 w_k(\beta) \qquad (31)$$

where w_k account for non-equally spaced data and possible non-periodicity of the spline. Equation (31) can be written as

$$V(\beta) = \frac{1}{n} \sum_{k=1}^{n} \frac{w_k(\beta)}{[1-L_{kk}(\beta)]^2} [u_k^{obs} - a_k(\beta)]^2 \qquad (32)$$

According to Wahba, $w_k(\beta)$ have to be chosen in such a way that the weights $w_k(\beta)/[1-L_{kk}(\beta)]^2$ are all equal and $V(\beta) = V_0(\beta)$ in the special case of equally-spaced data and periodic spline. Then

$$w_k(\beta) = \left[\frac{1 - L_{kk}(\beta)}{\frac{1}{n}\,\mathrm{tr}(I-L(\beta))}\right]^2 \tag{33}$$

and

$$V(\beta) = \frac{\frac{1}{n}\|(I-L(\beta))u^{obs}\|^2}{\left[\frac{1}{n}\,\mathrm{tr}(I-L(\beta))\right]^2} \tag{34}$$

where $\|\cdot\|$ denotes Euclidian norm. The minimization of $V(\beta)$ given by (34) yields the weighted cross-validation estimate of β.

Consider now the problem of approximately solving the integral equation

$$\int_0^1 K(x,\zeta)z(\zeta)\,d\zeta = u(x) \tag{35}$$

given discrete noisy measurements $u_1^{obs}, u_2^{obs}, \ldots, u_n^{obs}$ at the points $x_1, x_2, \ldots, x_n$.

Assuming that $z \in L_2[0,1]$ and choosing a stabilizing functional of zeroth order, one can obtain regularized solutions $z_\beta(x)$ by minimizing the smoothing functional,

$$M^\beta(z,u^{obs}) = \frac{1}{n}\sum_{i=1}^{n}\left[\int_0^1 K(x_i,\zeta)z(\zeta)\,d\zeta - u_i^{obs}\right]^2 + \beta\int_0^1 [z(\zeta)]^2 d\zeta \tag{36}$$

Wahba (1977) has shown that

$$z_\beta(x) = \Big(K(x_1,x), K(x_2,x), \ldots, K(x_n,x)\Big)\ (\tilde{Q} + n\beta I)^{-1}(u_1^{obs}, u_2^{obs}, \ldots, u_n^{obs})^T \tag{37}$$

where

$$\tilde{Q}_{jk} = \int_0^1 K(x_j,\zeta)K(x_k,\zeta)\,d\zeta$$

Now the weighted cross-validation function is given by

$$V(\beta) = \frac{1}{n} \sum_{n=1}^{n} \left[\int_0^1 K(x_k,\zeta) z_\beta^{[k]}(\zeta)\, d\zeta - u_k^{obs} \right]^2 w_k(\beta) \tag{38}$$

where $z_\beta^{[k]}(x)$ is the regularized solution obtained by omitting the k-th data point, u_k^{obs}.

Wahba (1977, 1980), using a similar argument as in the interpolation problem, has shown that the appropriate choice of $w_k(\beta)$ is

$$w_k(\beta) = \left[\frac{1-L_{kk}(\beta)}{\frac{1}{n}\, tr(I-L(\beta))} \right]^2 \tag{39}$$

where

$$L(\beta) = \tilde{Q}(\tilde{Q} + n\beta I)^{-1}$$

Then

$$V(\beta) = \frac{\frac{1}{n} \| (I-L(\beta)) u^{obs} \|^2}{\left[\frac{1}{n}\, tr(I-L(\beta))\right]^2} \tag{40}$$

The idea of cross-validation is quite general and is applicable nonlinear problems. The key problem is how to select the weights $w_k(\beta)$ in the cross-validation function $V(\beta)$. As pointed out by Wahba (1980), the problem of choosing the weights $w_k(\beta)$ for nonlinear problems has not yet been solved. However, one may use ordinary cross-validation, or even (40) when the operator is "mildly" nonlinear.

5. THESIS OUTLINE

The identification problem of petroleum reservoirs and subsurface aquifers, due to its extreme economic importance, has motivated the present thesis. The work was oriented in two directions. First, to study the fundamental question of uniqueness of the identification problem; second, and most important, to develop well-conditioned algorithms for its numerical solution. The work in Chapter II is along the first direction; Chapters III and IV are along the second.

It is noteworthy that the ill-posed nature of the petroleum reservoir/subsurface aquifer problem is typical in all identification problems in distributed parameter systems with spatially-varying parameters (Lions, 1978). This motivates the study of more general identification problems, the objective being to develop well-conditioned numerical algorithms. For this reason, in the regularization identification approaches presented in Chapters III and IV, we have not restricted ourselves to (1); we have considered much more general distributed parameter systems. It is also noteworthy that the inverse Sturm-Liouville framework, developed in Chapter II for the study of the uniqueness question, is potentially applicable to more general identification problems. However, since the theory is still at early stages and far from answering the uniqueness problem for (1) in its full generality, such a consideration would not be meaningful.

In Chapter II, we consider the one-dimensional analog of (1) and address the question of uniqueness of $\alpha(x)$ from a measurement of u at a single point. More specifically, we consider

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) = f(x,t), \qquad \text{in }]0,\ell[x]0,T]$$

$$u(x,0) = u_o \text{ (constant)} \qquad , \qquad \text{in }]0,\ell[$$

$$\alpha(0)\frac{\partial u}{\partial x}(0,t) = g(t) \qquad , \qquad \text{in }]0,T]$$

$$\alpha(\ell)\frac{\partial u}{\partial x}(\ell,t) = G(t) \qquad , \qquad \text{in }]0,T]$$

and the point measurement $u^{obs}(t) = u(x_p,t)$, $t \in]0,T]$ at some $x_p \in [0,\ell]$. The uniqueness question is answered in the following special cases

(i) $f = 0$, either g or $G = 0$, $x_p = 0$

(ii) $f = 0$, either g or $G = 0$, $x_p = \ell$

(iii) $f = q(t)\delta(x-x_p)$, $g = G = 0$, $x_p = \frac{\ell}{2}$

by proving their equivalence to an inverse Sturm-Liouville problem and extending the available Gel'fand-Levitan theory.

In Chapter III we develop a very general regularization identification approach by extending the Tikhonov regularization. The theory is presented in three levels: (i) abstract identification problems (ii) Identification of general 2m-order linear parabolic systems (iii) identification of spatially-varying diffusivity in the diffusion equation (1).

In level (iii), the n-dimensional version of (1) is considered with two different types of observation: distributed observation and point observation. In level (ii), the dynamic system

$$\frac{\partial u}{\partial t} + A(t)u = f \ , \qquad \text{in } \Omega \times]0,T[$$

$$u(x,0) = u_0 \qquad , \qquad \text{in } \Omega$$

$$B_j u = g_j \qquad , \qquad j = 0,\ldots, m=1, \text{ on } \Gamma \times]0,T[$$

where $\Omega \in \mathbb{R}^n$ with boundary Γ and $0 < T < \infty$ and where

$$A(t)\ u = \sum_{|p|,|q| \leq m} (-1)^{|p|} D_x^p (a_{pq}(x,t)\ D_x^q u)$$

$$B_j u = \sum_{|h| \leq m_j} b_{jh}(x,t) D_x^h u \quad , \ j = 0,\ldots, m-1$$

with $0 \leq m_j$ = order of $B_j \leq 2m-1$

is considered with distributed or point observation. Level (i) refers to abstract identification problems of the form

$$\Psi(A(\lambda),u) = f$$

where Ψ is an abstract mapping, A is a partial differential operator depending on a set of spatially-varying parameters λ (which are to be identified) and u is the state of the dynamic system (for which an observation $u^{obs} = \mathscr{C}u$ is given).

The last section of Chapter III refers to a numerical implementation of the regularization approach. A gradient method for the minimization of the smoothing functional is presented, as a natural consequence of the theory. The gradient approach is tested by numerical experiments in the one-dimensional version of (1).

Chapter IV focuses on the development of practical, computationally efficient numerical schemes for the minimization of the smoothing functional. The proposed approach is to perform the minimization of the smoothing functional over an appropriate finite-dimensional space and thus reduce the original infinite-dimensional identification problem to a finite-parameter estimation problem. The approach is based on a convergence theorem for the discretized minimization of the smoothing functional as well as available techniques for constructing piecewise-polynomial approximations in Sobolev spaces. The performance of this discrete regularization approach is tested by numerical experiments in the identification problem associated with (1).

REFERENCES

Beck, J. V. and K. J. Arnold (1977). Parameter Estimation in Engineering and Science. John Wiley, New York.

Bensoussan, A., J. L. Lions, and G. Papanicolaou (1978). Asymptotic Analysis for Periodic Structures. North Holland, Amsterdam.

Carter, R. D., L. F. Kemp, Jr., A. C. Pierce, and D. L. Williams (1974). Performance matching with constraints. Soc. Pet. Eng. J., 14, 187-196.

Chang, S. and W. Yeh (1976). A proposed algorithm for the solution of the large-scale inverse problem in groundwater. Water Resources Research, 12, 365-374.

Chavent, C., M. Dupuy, and P. Lemonnier (1975). History matching by use of optimal control theory. Soc. Pet. Eng. J., 15, 74-86.

Coats, K. H., J. R. Dempsey, and J. H. Henderson (1970). A new technique for determining reservoir description from field performance data. Soc. Pet. Eng. J., 10, 66-74.

Chen, W. H. and J. H. Seinfeld (1972). Estimation of spatially varying parameters in partial differential equations. Int. J. Control, 15, 487-495.

Chen, W. H., G. R. Gavalas, J. H. Seinfeld, and M. L. Wasserman (1974). A new algorithm for automatic history matching. Soc. Pet. Eng. J., 14, 593-608.

Cooley, R. L. (1977). A method of estimating parameters and assessing reliability for models of steady state groundwater flow, 1. Theory and numerical properties. Water Resour. Res., 13, 318-324.

Cooley, R. L. (1979). A method of estimating parameters and assessing reliability for models of steady state groundwater flow. 2. Application and statistical analysis. Water Resour. Res., 15, 603-617.

Craven, P. and G. Wahba (1979). Smoothing noisy data with spline functions. Numerische Mathematik, 31, 377-403.

Deininger, R. A. (1969). Linear programming for hydrologic analyses. Water Resour. Res., 5, 1105-1109.

Distefano, N. and A. Rath (1975). An identification approach to subsurface hydrological systems. Water Resour. Res., 11, 1005-1012.

Emsellem, Y. and G. de Marsily (1971). An automatic solution for the inverse problem. Water Resour. Res., 7, 1264-1283.

Eykhoff, P. (1974). System Identification, John Wiley, New York.

Frind, E. O. and G. F. Pinder (1973). Gelerkin solution of the inverse problem for aquifer transmissivity. Water Resour. Res., 9, 1397-1410.

Gavalas, G. R., P. C. Shah, and J. H. Seinfeld (1976). Reservoir history matching by Bayesian estimation. Soc. Pet. Eng. J., 16, 337-350.

Haimes, Y. Y., R. L. Perrine, and D. A. Wismer (1968). Identification of aquifer parameters by decomposition in multi-level optimization. Israel J. of Tech., 6, 322.

Ivanov, V. K. (1966). The approximate solutions of operator equations of the first kind. Zhurnal Vychislistel'noi Matematiki i Matematicheskoi Fiziki, 6, 1089-1094.

Jacquard, P. and C. Jain (1965). Permeability distribution from field pressure data. Soc. Pet. Eng. J., 5, 281-294.

Jahns, H. O. (1966). A rapid method for obtaining a two-dimensional reservoir description from well pressure response data. Soc. Pet. Eng. J., 6, 315-327.

Kleinecke, D. (1971). Use of linear programming for estimating geohydrologic parameters of groundwater basins. Water Resour. Res., 7, 367-375.

Kubrusly, C. (1977). Distributed parameter system identification: A survey. Int. J. Control, 26, 509-535.

Lions, J. L. (1978). Some aspects of modeling problems in distributed parameter systems in Lecture Notes in Control and Information Sciences Vol. I, edited by A. Ruberti, Springer-Verlag, Berlin, 11-41.

Lovell, R. E., L. Duckstein, and C. C. Kisiel (1972). Use of subjective information in estimation of aquifer parameters. Water Resour. Res., 8, 680-690.

Marino, M.A. and W-G. Yeh (1973). Identification of parameters in finite leaky aquifer system. J. Hydraul. Div. Amer. Soc. Civil Eng., 99, 319-336.

Merz, P. (1978). Spline smoothing by generalized cross-validation. Chevron Research Company, Richmond, California.

Miller, K. (1970). Least-squares methods for ill-posed problems with a prescribed bound. SIAM J. on Math. Anal., 1, 52-74.

Morozov, V. A. (1966). On the solution of functional equations by the method of regularization. Doklady Akademii Nauk SSSR, 167, 510-512.

Nelson, R. N. (1968). In-place determination of permeability distribution for heterogeneous porous media through analysis of energy dissipation. Soc. Pet. Eng. J., 8, 33-42.

Nelson, R. N. and W. L. McCollum (1969). Transient energy dissipation methods of measuring permeability distributions in heterogeneous porous materials. Rep. CSC 691229, Water Resour. Div., U.S. Geol. Surv., Washington, D.C.

Neuman, S. P. (1973). Calibration of distributed parameter groundwater flow models viewed as a multiple-objective decision process under uncertainty. Water Resour. Res., 9, 1006-1021.

Neuman, S. P. (1980b). A statistical approach to the inverse problem of aquifer hydrology. 3. Improved solution method and added perspective. Water Resour. Res., 16, 33-58.

Neuman, S. P. and S. Yakowitz (1979). A statistical approach to the inverse problem of aquifer hydrology. 1. Theory. Water Resour. Res., 15, 845-860.

Neuman, S. P., G. Fogg, and E. Jacobson (1980a). A statistical approach to the inverse problem of aquifer hydrology. 2. Case study. Water Resour. Res., 16, 33-58.

Padmanabhan, L. (1980). Private communication.

Reinsch, C. H. (1967). Smoothing by spline functions. Numerische Mathematik, 10, 177-183.

Sagar, B., S. Yakowitz, and L. Duckstein (1975). A direct method for the identification of the parameters of dynamic nonhomogeneous aquifers, Water Resour. Res., 11, 563-570.

Sammel, E. A. (1964). Evaluation of numerical analysis methods for determining variations in transmissibilities. Int. Ass. Sci. Hydrol. Publ., 64, 239-251.

Seinfeld, J. H. and W. H. Chen (1978). Identification of petroleum reservoir properties in Distributed Parameter Systems. Identification, Estimation and Control, edited by W. H. Ray and D. G. Lainiotis, Marcel Dekker, 497-554.

Seinfeld, J. H. and L. Lapidus (1974). Mathematical Methods in Chemical Engineering, Vol. 3. Process modeling estimation and identification. Prentice-Hall, Englewood Cliffs.

Shah, P. C., G. R. Gavalas, and J. H. Seinfeld (1978). Error analysis in history matching: The optimum level of parametrization. Soc. Pet. Eng. J., 18, 219-228.

Slater, G. E. and E. J. Durrer (1971). Adjustment of reservoir simulation models to match field performance. Soc. Pet. Eng. J., 11, 295-305.

Stallman, R. W. (1963). Calculation of resistance and error in an electric analog of steady flow through nonhomogeneous aquifers. U. S. Geol. Surv. Water Supply Pap. 20, 1544-1546.

Tikhonov, A. N. (1963a). Solution of incorrectly formulated problems and the regularization method. Doklady Akademii Nauk SSSR, 151, 501-504.

Tikhonov, A. N. (1963b). Regularization of incorrectly posed problems. Doklady Akademii Nauk SSSR, 153, 49-52.

Tikhonov, A. N. and V. Ya. Arsenin (1977). Solutions of Ill-Posed Problems. Winston-Wiley.

Vermuri, V. and W. J. Karplus (1969). Identification of nonlinear parameters of groundwater basins by hybrid computation. Water Resour. Res., 5, 172-185.

Vermuri, V., J. A. Dracup, and R. C. Erdmann (1969). Sensitivity analysis method of system identification and its potential in hydrologic research. Water Resour. Res., 5, 341-349.

Wahba, G. (1975). Smoothing noisy data with spline functions. Numerische Mathematik, 24, 383-393.

Wahba, G. and S. Wold (1975). A completely automatic french curve: Fitting spline functions by cross-validation. Communications in Statistics, 4, 1-17.

Wahba, G. (1977). Practical approximate solutions to linear operator equations when data are noisy. SIAM J. on Numer. Anal., 14, 651-667.

Wahba, G. (1980). Ill-posed problems: Numerical and statistical methods for mildly, moderately, and severely ill-posed problems with noisy data. Technical Report No. 595, University of Wisconsin.

Wasserman, M. L., A. S. Emanuel, and J. H. Seinfeld (1975). Practical application of optimal control theory for history-matching multiphase simulator models. Soc. Pet. Eng. J., 15, 347-355.

Yakowitz, S. and L. Duckstein (1980). Instability in aquifer identification: Theory and case studies. Water Resour. Res., 16, 1054-1064.

Yakowitz, S. and P. Noren (1976). On the identification of space variable parameters in dynamic linear partial differential equations. J. Math. Anal. Appl., 53, 521-538.

Yeh, W. W. and Y. S. Yoon (1976). A systematic optimization procedure for the identification of inhomogeneous aquifer parameters, in Advances in Groundwater Hydrology. American Water Resources Association, Minneapolis, Minnesota, 72-82.

Yeh, W-G. and G. W. Tauxe (1971a). Optimal identification of aquifer diffusivity using quasi-linearization. Water Resour. Res., 7, 955-962.

Yeh, W-G. and G. W. Tauxe (1971b). A proposed technique for identification of unconfined aquifer parameters. J. Hydrol., 12, 117-128.

CHAPTER II: IDENTIFIABILITY OF SPATIALLY-VARYING CONDUCTIVITY FROM POINT OBSERVATION AS AN INVERSE STURM-LIOUVILLE PROBLEM

1. INTRODUCTION. The partial differential equation

$$(1.1) \qquad \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x,y)\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\alpha(x,y)\frac{\partial u}{\partial y}\right) = f(x,y,t)$$

governs the temperature distribution in a nonhomogeneous isotropic solid or the pressure distribution in a fluid-containing porous medium. The conductivity $\alpha(x,y)$ is inaccessible to direct measurement and, consequently, its value must be inferred from measurements of u at a finite number of points. A fundamental question arising in such problems is that of *identifiability*, namely, do the measurements provide sufficient information to determine α uniquely.

Relatively little work has been carried out on the identifiability of α in (1.1). Early work by Cannon and coworkers [1-3] is concerned with the steady-state version of (1.1) and identifiability given the temperature u and the heat flux along the boundary. Kitamura and Nakagiri [4] have studied the identifiability of $\alpha(x)$ in the one-dimensional version of (1.1) given measurements of $u(x,t)$ at all x and t. The most relevant measurement configuration is that of one or more point measurements of u and we concentrate on that situation here. Specifically, we consider the problem of identifying $\alpha(x)$ in the one-dimensional version of (1.1),

$$(1.2) \qquad \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) = f(x,t)$$

given measurement of u at a single point x_p, $u(x_p,t)$. The appropriate method of attack to obtain uniqueness and non-uniqueness results is to formulate the problem as an inverse Sturm-Liouville problem.

2. INVERSE STURM-LIOUVILLE PROBLEMS. THEIR RELATION TO IDENTIFIABILITY PROBLEMS. The inverse Sturm-Liouville problem was first posed in 1946 by Borg [5] as follows: *Given the eigenvalues* λ_n *of the Sturm-Liouville problem*

(2.1) $$y'' + [\lambda - q(x)]y = 0$$

(2.2) $$y'(0) - hy(0) = 0$$

(2.3) $$y'(\ell) + Hy(\ell) = 0$$

determine q(x). Borg showed that knowledge of the spectrum alone is not sufficient to determine q(x) uniquely. Since that early work, two not altogether equivalent inverse Sturm-Liouville problems have been considered.

One approach, which has become associated with Gel'fand and Levitan [6], uses the spectral function $\sigma(\lambda)$ as a starting point. If $\phi(x;\lambda)$ denotes the solution of (2.1) satisfying $y(0) = 1$ and $y'(0) = h$ and if we define

$$E_f(\lambda) = \int_0^\ell f(x)\phi(x;\lambda)\, dx$$

where f is an arbitrary element of $L^2(0,\ell)$, then by Parseval's theorem,

$$\int_0^\ell f^2(x)\, dx = \int_{-\infty}^{\infty} E_f^2(\lambda)\, d\sigma(\lambda)$$

where

$$\sigma(\lambda) = \sum_{\lambda_n<\lambda} \frac{1}{\xi_n}$$

and

$$\xi_n = \int_0^\ell \phi^2(x;\lambda_n)\, dx$$

Gel'fand and Levitan have shown that knowledge of $\sigma(\lambda)$, or equivalently, of the spectrum $\{\lambda_n\}$ and the normalizing constants $\{\xi_n\}$, determine the potential $q(x)$ uniquely. Furthermore, they provided a method of constructing $q(x)$ from $\sigma(\lambda)$, as well as necessary and sufficient conditions for existence. Note that since $\xi_n = [y_n(0)]^{-2}$, where $y_n(x)$ are the normalized eigenfunctions of the Sturm-Liouville problem (2.1)-(2.3), the results of Gel'fand and Levitan can be interpreted as applicable to the problem of constructing a Sturm-Liouville operator of the form (2.1) given $\{\lambda_n\}$ and $\{|y_n(0)|\}$.

The other approach to the inverse Sturm-Liouville problem consists in using two spectra, such as $\{\lambda_n\}$ associated with (2.1)-(2.3) and $\{\mu_n\}$ associated with (2.1) and a different set of boundary conditions, to determine $q(x)$ [7]. Krein [8,9] provided a method of constructing $q(x)$ from two spectra as well as necessary and sufficient conditions for existence. The issue of existence was investigated further by Levitan [10] who showed how the normalizing constants ξ_n can be evaluated from $\{\lambda_n\}$ and $\{\mu_n\}$.

It is noteworthy that the prior work on inverse Sturm-Liouville theory is based on the Liouville normal form (2.1)-(2.3). As we shall see shortly, the problem of interest in the present work requires us to consider inverse Sturm-Liouville problems that are not in normal form.

Consider for a moment the parabolic system,

$$\begin{aligned}
&\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) + q(x)u = f(x,t), && \text{in }]0,\ell[\times]0,T] \\
&u(x,0) = u_o(x) && \text{, in }]0,\ell[\\
&\frac{\partial u}{\partial x}(0,t) - h\,u(0,t) = g(t) && \text{, in }]0,T] \\
&\frac{\partial u}{\partial x}(\ell,t) + H\,u(\ell,t) = G(t) && \text{, in }]0,T]
\end{aligned} \tag{2.4}$$

where $f(x,t), u_o(x)$, $g(t)$, $G(t)$, h, and H are known. Given the point measurement $z_d(t) = u(x_p,t)$, $t \in]0,T]$, at some $x_p \in [0,\ell]$, the question is can $\alpha(x)$ and $q(x)$ be uniquely determined.

Note that the solution [1] of (2.4) can be expressed in terms of the eigenvalues λ_n and the eigenfunctions $y_n(x)$ of

$$\frac{d}{dx}\left(\alpha(x)\frac{dy}{dx}\right) + [\lambda - q(x)]y = 0$$

(2.5)
$$y'(0) - h\,y(0) = 0$$
$$y'(\ell) + H\,y(\ell) = 0$$

as follows

(2.6)
$$u(x,t) = \sum_{n=1}^{\infty}\left[\int_0^{\ell} u_o(x)y_n(x)dx\right] y_n(x)e^{-\lambda_n t}$$
$$+ \int_0^t\int_0^{\ell}\left\{\sum_{n=1}^{\infty} y_n(x)y_n(x')e^{-\lambda_n(t-\tau)}\right\}f(x',\tau)\,dx'd\tau$$
$$- \int_0^t \left\{\sum_{n=1}^{\infty} \alpha(0)y_n(0)y_n(x)e^{-\lambda_n(t-\tau)}\right\}g(\tau)\,d\tau$$
$$+ \int_0^t \left\{\sum_{n=1}^{\infty} \alpha(\ell)y_n(\ell)y_n(x)e^{-\lambda_n(t-\tau)}\right\}G(\tau)\,d\tau$$

and thus the measurement

$$z_d(t) = u(x_p,t) = \sum_{n=1}^{\infty} y_n(x_p)\left[\int_0^{\ell} u_o(x)y_n(x)\,dx\right]e^{-\lambda_n t}$$
$$+ \int_0^t\int_0^{\ell}\left\{\sum_{n=1}^{\infty} y_n(x_p)y_n(x')e^{-\lambda_n(t-\tau)}\right\}f(x',\tau)\,dx'd\tau$$

[1] With $\alpha(x)$ strictly positive and α, q, f, u_o, g and G sufficiently regular, there exists a unique strong solution of (2.4). See [11, pp. 320-1] for appropriate Hölder continuity and compatibility conditions.

$$- \int_0^t \left\{ \sum_{n=1}^{\infty} \alpha(0) y_n(0) y_n(x_p) e^{-\lambda_n(t-\tau)} \right\} g(\tau)\, d\tau$$

$$(2.7) \qquad + \int_0^t \left\{ \sum_{n=1}^{\infty} \alpha(\ell) y_n(\ell) y_n(x_p) e^{-\lambda_n(t-\tau)} \right\} G(\tau) d\tau$$

Kitamura and Nakagiri [4] (see also [12]) considered (2.4) with α and q both being constant. Using (2.7), they have shown that (under certain assumptions) in the following special cases

(i) $f(x,t) = 0,\ g(t) = G(t) = 0$

(ii) $f(x,t) = 0,\ u_o(x) = 0$, one of $g(t)$ or $G(t)$ vanishes

(iii) $u_o(x) = 0,\ g(t) = G(t) = 0,\ f(x,t) = f_1(x) f_2(t)$

the eigenvalues λ_n can be uniquely determined. Thus the constants α and q can be easily obtained.

When α and q are spatially varying, one can still (under certain assumptions) determine the eigenvalues λ_n as well as some information on the eigenfunctions $y_n(x)$ for the cases (i)-(iii) ([13]). Thus the identifiability problem reduces to an inverse Sturm-Liouville problem related to (2.5). Notice that (2.5) is in normal form only if $\alpha(x) \equiv 1$, in which case the results of Gel'fand and Levitan [6] and Levinson [7] provide identifiability in very special cases.

In the present work we are going to consider the case $q(x) = 0$, i.e.

$$(2.8) \qquad \begin{aligned} &\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x) \frac{\partial u}{\partial x}\right) = f(x,t) \ , && \text{in }]0,\ell[\times]0,T] \\ &u(x,0) = u_o(x) \ , && \text{in }]0,\ell[\\ &\frac{\partial u}{\partial x}(0,t) - h\, u(0,t) = g(t) \ , && \text{in }]0,T] \\ &\frac{\partial u}{\partial x}(\ell,t) + H\, u(\ell,t) = G(t) \ , && \text{in }]0,T] \end{aligned}$$

the general problem being to determine $\alpha(x)$ knowing $f(x,t)$, $u_0(x)$, $g(t)$, $G(t)$, h and H and given the point measurement

$$z_d(t) = u(x_p,t) , \qquad t \in \,]0,T] \tag{2.9}$$

at some $x_p \in [0,\ell]$. The Sturm-Liouville problem associated with (2.8) is

$$\begin{aligned} &\frac{d}{dx}\left(\alpha(x) \frac{dy}{dx}\right) + \lambda y = 0 \\ &y'(0) - h\, y(0) = 0 \\ &y'(\ell) + H\, h(\ell) = 0 \end{aligned} \tag{2.10}$$

and the eigenfunction expansion of the solution of (2.8) is still given by (2.6).

In Section 3 we define three special cases of (2.8) corresponding to models of physical systems and formulate identifiability problems as inverse Sturm-Liouville problems. In Section 4 we state and prove the analog of Gel'fand and Levitan's result for the Sturm-Liouville problem (2.10). In Section 5 we obtain uniqueness and nonuniqueness results for the identification problems of Section 3.

3. PROBLEM STATEMENT. In the previous section we have stated a general identifiability problem associated to the system (2.8) and the measurement (2.9). With the available tools it does not appear to be possible to attack the problem in its full generality; rather, it is necessary to consider special cases. It will, however, be very important to select cases that are physically relevant and of practical significance. Our selection is based on the following considerations:

(i) In practice one generally has point actuators and thus boundary control and/or point control at some interior point(s) exist as opposed to distributed control

(ii) Before performing a heat conduction experiment, it is natural to assume that the system is at ambient temperature, i.e. u = constant. Thus, the most important special case for $u_o(x)$ is $u_o(x)$ = constant.

In the present work the analysis will be restricted to the SISO case. In other words, we are going to assume either that only one of $g(t)$, $G(t)$ is nonzero and $f(x,t) = 0$ or $g(t) = G(t) = 0$ and $f(x,t) = Q(t)\delta(x-x_p)$. Also, we will restrict ourselves to the special case of $u_o(x) = u_o$ (constant.)

We consider

(3.1)
$$\begin{aligned}
&\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) &&, \text{ in }]0,\ell[\, \times\,]0,T] \\
&u(x,0) = u_o &&, \text{ in }]0,\ell[\\
&\alpha(0)\,\frac{\partial u}{\partial x}\,(0,t) = Q(t) &&, \text{ in }]0,T] \\
&\frac{\partial u}{\partial x}\,(\ell,t) = 0 &&, \text{ in }]0,T]
\end{aligned}$$

where $\alpha \in C^1([0,\ell])$ and $\exists\, \alpha_o > 0\colon\ \alpha(x) \geq \alpha_o\ \forall x \in [0,\ell]$

$Q \in H^1(0,T)$ and $\nexists\, \varepsilon > 0\colon\ Q(t) = 0\ \ \forall t \in\,]0,\varepsilon[$

$u_o \in \mathbb{R}$

PROBLEM 1. *To a known input* $Q(t)$, *a known initial state* u_o *and a given measurement* $z_d(t) = u(0,t)$, $t \in]0,T]$, *does there correspond a unique* $\alpha(x)$?

PROBLEM 2. *To a known input* $Q(t)$, *a known initial state* u_o *and a given measurement* $z_d(t) = u(\ell,t)$, $t \in]0,T]$, *does there correspond a unique* $\alpha(x)$?

Now consider

$$\text{(3.2)}\qquad \begin{aligned} &\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) + Q(t)\delta(x-x_p) && \text{, in }]0,\ell[\times]0,T] \\ &u(x,0) = u_o && \text{, in }]0,\ell[\\ &\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(\ell,t) = 0 && \text{, in }]0,T] \end{aligned}$$

where

$$\alpha \in C^1([0,\ell]) \text{ and } \exists\, \alpha_o > 0: \quad \alpha(x) \geq \alpha_o \quad \forall x \in [0,\ell]$$

$$Q \in H^1(0,T) \text{ and } \quad \nexists\, \varepsilon > 0: \quad Q(t) = 0 \quad \forall t \in]0,\varepsilon[$$

$$u_o \in \mathbb{R}$$

$$x_p \in]0,\ell[$$

PROBLEM 3. *To a known input* $Q(t)$, *a known initial state* u_o *and a given measurement* $z_d(t) = u(x_p,t)$, $t \in]0,T]$, *does there correspond a unique* $\alpha(x)$?

To be able to formulate Problems 1-3 as inverse Sturm-Liouville problems, we will need the following lemmata:

LEMMA 1: *Let* $\{\lambda_n\}$ *be a strictly increasing sequence tending to infinity. If* $\{c_n\}$ *is a bounded sequence and*

$$\sum_{n=1}^{\infty} c_n e^{-\lambda_n t} = 0 \qquad \forall\, t \in]0,T]$$

then

$$c_n = 0 \qquad \forall n \in \mathbb{N}$$

Proof: See [14, p. 170].

LEMMA 2: *Let* $\{\lambda_n\}$ *and* $\{\hat{\lambda}_n\}$ *be strictly increasing sequences tending to infinity and* $\{c_n\}$ *and* $\{\hat{c}_n\}$ *be bounded sequences. Suppose*

$$(3.3) \qquad \sum_{n=1}^{\infty} c_n e^{-\lambda_n t} = \sum_{n=1}^{\infty} \hat{c}_n e^{-\hat{\lambda}_n t} \qquad \forall\, t \in]0,T]$$

(i) *If* $c_n \neq 0$ *and* $\hat{c}_n \neq 0$, $n = 1,\dots, N$ *then* $\lambda_n = \hat{\lambda}_n$ *and* $c_n = \hat{c}_n$, $n = 1,\dots, N$

(ii) *If* $c_n \neq 0$ *and* $\hat{c}_n \neq 0 \quad \forall\, n \in \mathbb{N}$ *then* $\lambda_n = \hat{\lambda}_n$ *and* $c_n = \hat{c}_n \quad \forall\, n \in \mathbb{N}$

Proof: The proof is similar to that of a similar proposition in [15, p. 476].

We will first prove (i) for $N = 1$ i.e. we will prove that if $c_1 \neq 0$ and $\hat{c}_1 \neq 0$ then $\lambda_1 = \hat{\lambda}_1$ and $c_1 = \hat{c}_1$. Suppose that $\lambda_1 \neq \hat{\lambda}_1$ and assume $\lambda_1 < \hat{\lambda}_1$ without loss of generality. Multiplying both sides of (3.3) by $e^{\lambda_1 t}$, we obtain

$$c_1 - \hat{c}_1 e^{-(\hat{\lambda}_1-\lambda_1)t} + \sum_{n=2}^{\infty} c_n e^{-(\lambda_n-\lambda_1)t} - \sum_{n=2}^{\infty} \hat{c}_n e^{-(\hat{\lambda}_n-\lambda_1)t} = 0 \qquad \forall\, t \in]0,T]$$

Since all the exponents in the above equation are $\neq 0$ it follows from Lemma 1 that $c_1 = 0$. This contradicts our hypothesis. Hence $\lambda_1 = \hat{\lambda}_1$. Thus we can write (3.3) as follows

$$(c_1 - \hat{c}_1)e^{-\lambda_1 t} + \sum_{n=2}^{\infty} c_n e^{-\lambda_n t} - \sum_{n=2}^{\infty} \hat{c}_n e^{-\hat{\lambda}_n t} = 0 \quad \forall\, t \in]0,T]$$

or

$$(c_1-\hat{c}_1) + \sum_{n=2}^{\infty} c_n e^{-(\lambda_n-\lambda_1)t} - \sum_{n=2}^{\infty} \hat{c}_n e^{-(\hat{\lambda}_n-\lambda_1)t} = 0 \qquad \forall\, t \in]0,T]$$

With the same argument as before we conclude that $c_1 = \hat{c}_1$. So (i) is true for N = 1.

Assume that (i) is true for N = k. Let $c_n \neq 0$ and $\hat{c}_n \neq 0$, n = 1,..., k + 1. Then (3.3) becomes

$$\sum_{n=k+1}^{\infty} c_n e^{-\lambda_n t} = \sum_{n=k+1}^{\infty} \hat{c}_n e^{-\hat{\lambda}_n t}$$

and by repeating the same argument we deduce that $\lambda_{k+1} = \hat{\lambda}_{k+1}$ and $c_{k+1} = \hat{c}_{k+1}$, which shows that (i) is true for N=k+1.

This completes the proof of (i). Condition (ii) is an immediate consequence of (i).

LEMMA 3 [16, p. 325]: *Let* Ψ, $Q \in L^1(0,T)$ *and assume*

$$\nexists \varepsilon > 0: \quad Q(t) = 0 \qquad \textit{a.e. in } [0,\varepsilon[.$$

If

$$\int_0^t \Psi(t-\tau)Q(\tau)\, d\tau = 0 \qquad \textit{a.e. in }]0.T[$$

then

$$\Psi(t) = 0 \qquad \textit{a.e. in }]0,T[$$

COROLLARY: *Let* $Q \in L^1(0,T)$ *and assume*

$$\nexists \varepsilon > 0: \quad Q(t) = 0 \qquad \textit{a.e. in }]0,\varepsilon[$$

If the integral equation

$$\int_0^t \Psi(t-\tau)Q(\tau)\, d\tau = R(t)$$

admits a solution $\Psi \in L^1(0,T)$ *then* Ψ *is unique*

Next consider

$$\frac{d}{dx}\left(\alpha(x)\,\frac{dy}{dx}\right) + \lambda y = 0$$

$$(3.4) \qquad y'(0) = 0$$

$$y'(\ell) = 0$$

where $\alpha \in C^1([0,\ell])$ and $\exists\, \alpha_0$: $\alpha(x) \geq \alpha_0 \quad \forall\, x \in [0,\ell]$ and denote by λ_n the eigenvalues of the above Sturm-Liouville problem and by $y_n(x)$ its normalized eigenfunctions.

PROBLEM 1'. *Referring to* (3.4), *is knowledge of* $\{\lambda_n\}_{n=1}^{\infty}$ *and* $\{|y_n(0)|\}_{n=1}^{\infty}$ *sufficient to determine* $\alpha(x)$ *uniquely?*

Proof of Equivalence of Problems 1 and 1'.

We will show that knowing $Q(t)$ and u_0

(i) $z_d(t)$ is sufficient to determine $\{\lambda_n\}_{n=1}^{\infty}$ and $\{|y_n(0)|\}_{n=1}^{\infty}$ uniquely.

(ii) $\{\lambda_n\}_{n=1}^{\infty}$ and $\{|y_n(0)|\}_{n=1}^{\infty}$ are sufficient to determine $z_d(t)$ uniquely.

The eigenfunction expansion of the solution of (3.1) is given by

$$u(x,t) = u_0 - \int_0^t \left\{ \sum_{n=1}^{\infty} y_n(0) y_n(x) e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

Hence

$$(3.5) \quad z_d(t) = u(0,t) = u_0 - \int_0^t \left\{ \sum_{n=1}^{\infty} |y_n(0)|^2\, e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

Now given $Q(t)$, u_0 and $z_d(t)$, it follows from the Corollary of Lemma 3 that the function

$$\Psi(t) = \sum_{n=1}^{\infty} |y_n(0)|^2\, e^{-\lambda_n t}$$

is uniquely determined. Now since $\{\lambda_n\}$ and $\{y_n(x)\}$ are eigenvalues and eigenfunctions, respectively, of the regular Sturm-Liouville problem (3.4), they clearly satisfy the assumptions of Lemma 2. Hence $\{\lambda_n\}$ and $\{|y_n(0)|\}$ are uniquely determined by $\Psi(t)$. This proves (i). Condition (ii) is an obvious consequence of (3.5).

PROBLEM 2'. *Referring to (3.4), is knowledge of* $\{\lambda_n\}_{n=1}^{\infty}$ *and* $\{y_n(0)y_n(\ell)\}_{n=1}^{\infty}$ *sufficient to determine* $\alpha(x)$ *uniquely?*

Proof of Equivalence of Problems 2 and 2'

The eigenfunction expansion of the solution of (3.1) is given by

$$u(x,t) = u_o - \int_o^t \left\{ \sum_{n=1}^{\infty} y_n(0)y_n(x)e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

Hence

$$z_d(t) = u(\ell,t) = u_o - \int_o^t \left\{ \sum_{n=1}^{\infty} y_n(0)y_n(\ell)\, e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau \tag{3.6}$$

Using (3.6) and repeating the same argument as with Problems 1 and 1', we conclude that Problems 2 and 2' are equivalent.

PROBLEM 3'. *Referring to (3.4), is knowledge of* $\{\lambda_n\}_{n=1}^{\infty}$ and $\{|y_n(x_p)|\}_{n=1}^{\infty}$ *sufficient to determine* $\alpha(x)$ *uniquely?*

Problem 3 is equivalent to Problem 3' provided that $y_n(x_p) \neq 0 \quad \forall\, n \in \mathbb{N}$. Indeed, using the eigenfunction expansion of the solution of (3.2),

$$u(x,t) = u_o + \int_o^t \left\{ \sum_{n=1}^{\infty} y_n(x_p)y_n(x)e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

we have

$$z_d(t) = u(x_p,t) = u_o + \int_o^t \left\{ \sum_{n=1}^{\infty} |y_n(x_p)|^2\, e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau \tag{3.7}$$

Thus, repeating the same argument as with Problems 1 and 1', we conclude that Problems 3 and 3' are equivalent, provided that $y_n(x_p) \neq 0 \quad \forall n \in \mathbb{N}$.

4. AN ASSOCIATED INVERSE STURM-LIOUVILLE PROBLEM.

As mentioned in Section 2, Gel'fand and Levitan ([6]) have solved the inverse Sturm-Liouville problem for a Sturm-Liouville operator in *normal* form. Their result (as applied to a finite interval) can be stated as follows:

THEOREM 1 (GEL'FAND AND LEVITAN): *Let* $\{\lambda_n\}$ *and* $\{\xi_n\}$ *be two sequences of positive real numbers obeying the asymptotic formulas*

$$\sqrt{\lambda_n} = \frac{\pi}{\ell} n + \frac{b_1}{n} + \frac{b_3}{n^3} + 0\left(\frac{1}{n^4}\right)$$

$$\xi_n = \frac{\ell}{2} + \frac{a_1}{n^2} + 0\left(\frac{1}{n^4}\right)$$

where a_1, b_1, b_3 *are constants. Then there exists a unique differential operator, defined by a differential expression of the form*

$$L(y) = y'' - q(x)y \qquad 0 \leq x \leq \ell$$

with $q(x) \in C([0,\ell])$ *and by boundary conditions of the form*

$$y'(0) - h\, y(0) = 0$$

$$y'(\ell) + H\, y(\ell) = 0$$

which has $\{\lambda_n\}$ *as eigenvalues and* $\{\xi_n\}$ *as normalizing constants. The function* $q(x)$ *and the number* h *can be computed via the formulae*

$$q(x) = \frac{1}{2} \frac{\partial K(x,x)}{\partial x} \quad , \qquad h = K(0,0)$$

where $K(x,t)$ *is the solution of the linear integral equation*

$$F(x,t) + K(x,t) + \int_0^x K(x,s)F(s,t)\, ds = 0$$

and where

$$F(x,t) = \frac{1}{\xi_o}\cos(\sqrt{\lambda_o}x).\ \cos(\sqrt{\lambda_o}t) - \frac{1}{\ell} +$$

$$+ \sum_{n=1}^{\infty}\left[\frac{\cos(\sqrt{\lambda_n}x)\cos(\sqrt{\lambda_n}t)}{\xi_n} - \frac{2}{\ell}\cos\left(\frac{n\pi}{\ell}x\right)\cos\left(\frac{n\pi}{\ell}t\right)\right]$$

Note that the above theorem gives at the same time existence, uniqueness and method of construction of the differential operator from its eigenvalues and its normalizing constants. For the purpose of studying identifiability problems, one needs only uniqueness. With this in mind, and the fact that $\xi_n = [y_n(0)]^{-2}$, where $y_n(x)$ are the normalized eigenfunctions of the Sturm-Liouville operator, what we wish to retain is

THEOREM 1' (GEL'FAND AND LEVITAN)

Consider

(4.1)
$$\begin{aligned} y'' + [\lambda - q(x)]y &= 0 \\ y'(0) - h\,y(0) &= 0 \\ y'(\ell) + H\,y(\ell) &= 0 \end{aligned}$$

where $q \in C([0,\ell])$ *and denote by* λ_n *its eigenvalues and* $y_n(x)$ *its normalized eigenfunctions. Also, consider*

(4.2)
$$\begin{aligned} y'' + [\lambda - r(x)]y &= 0 \\ y'(0) - \hat{h}\,y(0) &= 0 \\ y'(\ell) + \hat{H}\,y(\ell) &= 0 \end{aligned}$$

where $r \in C([0,\ell])$ *and denote by* μ_n *its eigenvalues and* $z_n(x)$ *its normalized eigenfunctions. If*

(4.3)
$$\begin{aligned} \lambda_n &= \mu_n \\ |y_n(0)| &= |z_n(0)| \end{aligned} \qquad \forall n \in \mathbb{N}$$

then

$$q(x) = r(x)$$
$$h = \hat{h}$$
$$H = \hat{H}$$

The purpose of this section will be to obtain a similar result for the Sturm-Liouville problem (3.4). Note that (3.4) can be reduced to (4.1) via the so-called Liouville transform ([17]). Therefore, it is natural to try to "back Liouville-transform" the result of Gel'fand and Levitan. In fact, this is possible and it leads to the following result:

Given $\alpha(0)$, λ_n *and* $|y_n(0)|$ *for a differential operator of the form* (3.4), *there corresponds a unique* $\alpha(x)$.

Using an entirely different approach than that of Gel'fand and Levitan we will show that λ_n and $|y_n(0)|$ are sufficient to determine $\alpha(x)$ uniquely. In fact, we will prove the following theorem:

THEOREM 2: *Consider*

$$\text{(4.4)} \qquad \begin{aligned} &\frac{d}{dx}\left(\alpha(x)\frac{dy}{dx}\right) + \lambda y = 0 \\ &y'(0) - h\,y(0) = 0 \\ &y'(\ell) + H\,y(\ell) = 0 \end{aligned}$$

with $\alpha \in C^1([0,\ell])$, *and* $\exists\, \alpha_0 > 0$: $\alpha(x) \geq \alpha_0 \quad \forall x \in [0,\ell]$ *and denote by* λ_n *its eigenvalues and* $y_n(x)$ *its normalized eigenfunctions. Also, consider*

$$\text{(4.5)} \qquad \begin{aligned} &\frac{d}{dx}\left(\beta(x)\frac{dy}{dx}\right) + \lambda y = 0 \\ &y'(0) - \hat{h}\,y(0) = 0 \\ &y'(\ell) + \hat{H}\,y(\ell) = 0 \end{aligned}$$

with $\beta \in C^1([0,\ell])$, *and* $\exists\, \beta_0 > 0$: $\beta(x) \geq \beta_0 \quad \forall x \in [0,\ell]$ *and denote by* μ_n *its eigenvalues and* $z_n(x)$ *its normalized eigenfunctions. If*

$$\lambda_n = \mu_n \qquad \forall\, n \in \mathbb{N} \tag{4.6}$$
$$|y_n(0)| = |z_n(0)|$$

then

$$\alpha(x) = \beta(x)$$
$$h = \hat{h}$$
$$H = \hat{H}$$

Before proving Theorem 2, we will first prove a number of lemmata.

LEMMA 4: *Consider the class $\mathcal{M}$ of differential operators of the form*

$$A = \frac{d}{dx}\left(\alpha(x)\frac{d}{dx}\right)$$

densely defined in $L^2(0,\ell)$, *where* $\alpha \in C^1([0,\ell])$ *bounded below by a positive constant. If* $A_1 \in \mathcal{M}$ *and* $A_2 \in \mathcal{M}$ *have the same spectrum, then*

$$\int_0^\ell \frac{dx}{\sqrt{\alpha_1(x)}} = \int_0^\ell \frac{dx}{\sqrt{\alpha_2(x)}}$$

Proof: For every $A \in \mathcal{M}$, the eigenvalues λ_n satisfy the following asymptotic formula (see e.g. [18])

$$\sqrt{\lambda_n} = \frac{n\pi}{\int_0^\ell \frac{dx}{\sqrt{\alpha(x)}}} + O(1)$$

Now if the operators $A_1 = \frac{d}{dx}\left(\alpha_1(x)\frac{d}{dx}\right)$ and $A_2 = \frac{d}{dx}\left(\alpha_2(x)\frac{d}{dx}\right)$ have the same eigenvalues λ_n, then

$$\sqrt{\lambda_n}\int_0^\ell \frac{dx}{\sqrt{\alpha_1(x)}} = n\pi + O(1) = \sqrt{\lambda_n}\int_0^\ell \frac{dx}{\sqrt{\alpha_2(x)}}$$

$$\Rightarrow \qquad \sqrt{\lambda_n}\left(\int_0^\ell \frac{dx}{\sqrt{\alpha_1(x)}} - \int_0^\ell \frac{dx}{\sqrt{\alpha_2(x)}}\right) = O(1)$$

$$\Rightarrow \qquad \int_0^\ell \frac{dx}{\sqrt{\alpha_1(x)}} - \int_0^\ell \frac{dx}{\sqrt{\alpha_2(x)}} = O\left(\frac{1}{\sqrt{\lambda_n}}\right) = O\left(\frac{1}{n}\right) \qquad \forall\, n \in \mathbb{N}$$

Hence

$$\int_0^\ell \frac{dx}{\sqrt{\alpha_1(x)}} = \int_0^\ell \frac{dx}{\sqrt{\alpha_2(x)}}$$

Q.E.D.

LEMMA 5: *Let* α, $\beta \in C^1([0,\ell])$ *bounded below by positive constants and satisfying* $\int_0^\ell \frac{dx}{\sqrt{\alpha(x)}} = \int_0^\ell \frac{dx}{\sqrt{\beta(x)}}$. *Denote* $\rho \in C^2([0,\ell])$ *the solution of*

$$\frac{d\rho}{dx} = [\beta(x)]^{-1/2}[\alpha(\rho)]^{1/2}$$

$$\rho(0) = 0$$

Then ρ *is a bijection of* $[0,\ell]$ *onto itself.*

Proof: Since $\frac{d\rho}{dx} > 0$, ρ is strictly increasing. Furthermore, from the definition of ρ, we have

$$\frac{d\rho}{[\alpha(\rho)]^{1/2}} = \frac{dx}{[\beta(x)]^{1/2}}$$

which upon integration gives

$$\int_0^{\rho(\ell)} \frac{dx}{\sqrt{\alpha(x)}} = \int_0^\ell \frac{dx}{\sqrt{\beta(x)}}$$

Hence

$$\int_0^{\rho(\ell)} \frac{dx}{\sqrt{\alpha(x)}} = \int_0^\ell \frac{dx}{\sqrt{\alpha(x)}} \Leftrightarrow \int_\ell^{\rho(\ell)} \frac{dx}{\sqrt{\alpha(x)}} = 0$$

But since $\alpha(x)$ is strictly positive, $\int_0^x \frac{dx}{\sqrt{\alpha(x)}}$ is a strictly increasing function. Hence $\rho(\ell) = \ell$. So ρ is a strictly increasing continuous mapping of $[0,\ell]$ onto $[0,\ell]$. Hence ρ is a bijection.

LEMMA 6: *Let* α, β, ρ *as in Lemma 5 and let* $U; L^2(0,\ell) \to L^2(0,\ell)$ *be defined by*

$$U(f) = \kappa\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} f(\rho(x)) + \int_0^x K(x,t)f(\rho(t))\,dt$$

where $\kappa \in \mathbb{R}^+$ *and* $K \in C([0,\ell] \times [0,\ell])$. U *will be unitary iff* $\kappa = 1$ *and* $K = 0$.

Proof: Let

$$T(f) = \left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} f(\rho(x))$$

$$V(f) = \int_0^x K(x,t) f(\rho(t))\, dt$$

so that

$$U = \kappa T + V$$

To prove the "if" part of the lemma, we need to show that T is unitary. A straightforward calculation gives the adjoint of T

$$T^*(f) = \left[\frac{\beta(\rho^{-1}(x))}{\alpha(x)}\right]^{1/4} f(\rho^{-1}(x))$$

and thus $TT^* = T^*T = I$.

To prove the "only if" part of the lemma, we will first show that the spectral radius of V is zero. To see this, observe that

$$|(V^n f)(x)|^2 \le \frac{\gamma\Gamma^{(n-1)^2} M^n}{2^{n-1}(n-1)!} \|f\|^2_{L^2}\, x^{2n-1} \qquad \forall x \in [0,\ell]$$

where

$$\gamma = \sup_{0\le x\le \ell} \left[\frac{\beta(x)}{\alpha(\rho(x))}\right]^{1/2}$$

$$\Gamma = \sup_{0\le x\le \ell} \left[\frac{\alpha(\rho(x)}{\beta(x)}\right]^{1/2}$$

$$M = \sup_{0\le t\le x\le \ell} [K(x,t)]^2$$

(This can be easily shown by induction.) Hence

$$\|V^n f\|_{L^2} \le \frac{\gamma}{\sqrt{n!}} \, \Gamma^{\frac{(n-1)^2}{2}} \left(\frac{M\ell^2}{2}\right)^{\frac{n}{2}} \|f\|_{L^2}$$

$$\|V^n\|_{L^2} \le \frac{\gamma}{\sqrt{n!}} \, \Gamma^{\frac{(n-1)^2}{2}} \left(\frac{M\ell^2}{2}\right)^{\frac{n}{2}}$$

$$r(V) = \lim_{n\to\infty} \|V^n\|^{\frac{1}{n}} = 0$$

Now

$$V = U - \kappa T$$

Since U and T are unitary, it is easy to show that $(U-\kappa T)$ and $(U-\kappa T)^*$ commute, i.e. $V = U-\kappa T$ is normal. But for every normal operator, the spectral radius equals the norm of the operator. Hence $\|V\|_{L^2} = 0 \Rightarrow K(x,t) = 0$. So

$$U = \kappa T$$

and since they are both unitary and $\kappa > 0$, it follows that $\kappa = 1$. This completes the proof of the lemma.

LEMMA 7: *Let* $\alpha(x)$ *and* $\beta(x)$ *be* C^1 *-functions bounded below by positive constants. Furthermore, let* $\phi(x;\lambda)$ *be the solution of*

$$(4.7) \qquad \begin{aligned} &\frac{d}{dx}\left(\alpha(x)\frac{dw}{dx}\right) + \lambda w = 0, \qquad x \geq 0 \\ &w(0) = 1 \\ &w'(0) = h \end{aligned}$$

$\psi(x;\lambda)$ *be the solution of*

$$(4.8) \qquad \begin{aligned} &\frac{d}{dx}\left(\beta(x)\frac{dw}{dx}\right) + \lambda w = 0, \qquad x \geq 0 \\ &w(0) = 1 \\ &w'(0) = \hat{h} \end{aligned}$$

Then there exists a continuous kernel $K(x,t)$ *such that*

$$\psi(x;\lambda) = \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4} \left[\frac{\alpha(\rho(x)}{\beta(x)}\right]^{1/4} \phi(\rho(x);\lambda) + \int_0^x K(x,t)\phi(\rho(t);\lambda)\, dt$$

$$\forall x \geq 0$$

where $\rho(x)$ *is the solution of*

(4.9)
$$\frac{d\rho}{dx} = [\beta(x)]^{-1/2} [\alpha(\rho)]^{1/2}$$
$$\rho(0) = 0$$

Proof: Let $u(x,y;\lambda) = \phi(x;\lambda)\psi(y;\lambda)$.

It can be easily seen that u satisfies the hyperbolic P.D.E.

(4.10)
$$\frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\beta(y)\frac{\partial u}{\partial y}\right) = 0$$

and the initial conditions

(4.11)
$$u)_{y=0} = \phi(x;\lambda)$$

(4.12)
$$\frac{\partial u}{\partial y})_{y=0} = \hat{h}\,\phi(x;\lambda)$$

The Cauchy problem (4.10)-(4.12) admits a unique solution which can be computed by Riemann's method (See Appendix for details). First, we make the change of variable

(4.13)
$$X = \int_0^x \frac{dx'}{\sqrt{\alpha(x')}} = \mathscr{A}(x)$$

(4.14)
$$Y = \int_0^y \frac{dy'}{\sqrt{\beta(y')}} = \mathscr{B}(y)$$

Then, applying (A.9) we find

$$\tilde{u}(X,Y;\lambda) = \frac{1}{2}\left[\frac{\beta(0)}{\tilde{\beta}(Y)\tilde{\alpha}(X)}\right]^{1/4} \left\{[\tilde{\alpha}(X-Y)]^{1/4}\tilde{\phi}(X-Y;\lambda) + [\tilde{\alpha}(X+Y)]^{1/4}\tilde{\phi}(X+Y\ ;\lambda)\right\}$$

$$(4.15) \qquad + \frac{1}{2}\int_{X-Y}^{X+Y} \tilde{W}(X,Y,t)\tilde{\phi}(t;\lambda)\, dt$$

where $\tilde{W}(X,Y,t) = W_1(X,Y,t) + \hat{h}W_2(X,Y,t)$. Applying (4.15) at $X = 0$ we obtain

$$\tilde{\psi}(Y;\lambda) = \tilde{u}(0,Y;\lambda) = \frac{1}{2}\left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\frac{[\tilde{\alpha}(-Y)]^{1/4}\tilde{\phi}(-Y;\lambda)+[\tilde{\alpha}(Y)]^{1/4}\tilde{\phi}(Y;\lambda)}{[\tilde{\beta}(Y)]^{1/4}}$$

$$(4.16) \qquad + \frac{1}{2}\int_{-Y}^{Y} \tilde{W}(0,Y,t)\tilde{\phi}(t;\lambda)dt$$

Now, if the function $\alpha(x)$ is continued so as to be even, it is easy to see that $\phi(x;\lambda) = \phi(-x;\lambda)$. Thus we easily deduce that

$$(4.17) \qquad \begin{aligned} \tilde{\alpha}(Y) &= \tilde{\alpha}(-Y) \\ \tilde{\phi}(Y;\lambda) &= \tilde{\phi}(-Y\ ;\lambda) \end{aligned}$$

Furthermore

$$(4.18) \qquad \begin{aligned} \int_{-Y}^{Y} \tilde{W}(0,Y,t)\tilde{\phi}(t;\lambda)dt &= \int_{0}^{Y} \tilde{W}(0,Y,t)\tilde{\phi}(t;\lambda)\, dt + \int_{0}^{Y} \tilde{W}(0,Y,-t)\tilde{\phi}(t;\lambda)\, dt \\ &= \int_{0}^{Y} \left\{\tilde{W}(0,Y,t) + \tilde{W}(0,Y,-t)\right\}\tilde{\phi}(t;\lambda)\, dt \end{aligned}$$

Thus setting

$$\tilde{K}(Y,t) = \frac{1}{2}\left\{\tilde{W}(0,Y,t) + \tilde{W}(0,Y,-t)\right\}$$

it follows from (4.16), (4.17) and (4.18) that

$$(4.19) \qquad \tilde{\psi}(Y;\lambda) = \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\tilde{\alpha}(Y)}{\tilde{\beta}(Y)}\right]^{1/4}\tilde{\phi}(Y;\lambda) + \int_0^Y \tilde{K}(Y,t)\tilde{\phi}(t,\lambda)dt$$

It remains to back-transform to the original variables. It is easy to see that

$$\rho = \mathscr{A}^{-1} \circ \mathscr{B}$$

Also recall (from Appendix) the notation $\tilde{\alpha} = \alpha \circ \mathscr{A}^{-1}$, $\tilde{\beta} = \beta \circ \mathscr{B}^{-1}$, $\tilde{\phi} = \phi \circ \mathscr{A}^{-1}$ etc. Thus (4.19) becomes

$$\psi(y,\lambda) = \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(y))}{\beta(y)}\right]^{1/4}\phi(\rho(y);\lambda) + \int_0^y K(y,t)\phi(\rho(t);\lambda)\, dt$$

Q E D

Proof of Theorem 2. Let $\phi(x;\lambda)$ be the solution of

$$(4.20) \qquad \begin{aligned} &\frac{d}{dx}\left(\alpha(x)\frac{dy}{dx}\right) + \lambda y = 0 \\ &y(0) = 1 \\ &y'(0) = h \end{aligned}$$

and $\psi(x;\lambda)$ the solution of

$$(4.21) \qquad \begin{aligned} &\frac{d}{dx}\left(\beta(x)\frac{dy}{dx}\right) + \lambda y = 0 \\ &y(0) = 1 \\ &y'(0) = \hat{h} \end{aligned}$$

Clearly,

$$\left.\begin{aligned} y_n(x) &= y_n(0)\phi(x;\lambda_n) \\ z_n(x) &= z_n(0)\psi(x;\lambda_n) \end{aligned}\right\} \quad \forall n \in \mathbb{N}$$

From Lemma 7 we have

(4.22) $$\psi(x;\lambda) = \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4}\phi(\rho(x);\lambda) + \int_0^x K(x,t)\phi(\rho(t);\lambda)dt$$

Hence,

(4.23) $$\frac{z_n(x)}{z_n(0)} = \frac{\left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} y_n(\rho(x)) + \int_0^x K(x,t)y_n(\rho(t))dt}{y_n(0)}$$

Now define the operator $U; L^2(0,\ell) \to L^2(0,\ell)$ by

$$U(f) = \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} f(\rho(x)) + \int_0^x K(x,t)f(\rho(t))dt$$

Since every $f \in L^2(0,\ell)$ can be expanded as

$$f(x) = \sum_{n=1}^{\infty} f_n y_n(x)$$

where

$$f_n = \int_0^{\ell} f(x)y_n(x)\,dx$$

we have

$$(Uf)(x) = \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} f(\rho(x)) + \int_0^x K(x,t)f(\rho(t))dt$$

$$= \left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(x)}{\beta(x)}\right]^{1/4} \sum_{n=1}^{\infty} f_n y_n(\rho(x)) + \int_0^x K(x,t) \sum_{n=1}^{\infty} f_n y_n(\rho(t))\,dt$$

$$= \sum_{n=1}^{\infty} f_n \left\{\left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4}\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} y_n(\rho(x)) + \int_0^y K(x,t)y_n(\rho(t))\,dt\right\}$$

and taking into account (4.23),

$$(Uf)(x) = \sum_{n=1}^{\infty} f_n \frac{y_n(0)}{z_n(0)} z_n(x)$$

Thus, from Parseval's theorem we obtain

$$\|Uf\|_{L^2}^2 = \sum_{n=1}^{\infty} f_n^2 \left|\frac{y_n(0)}{z_n(0)}\right|^2$$

and since $|y_n(0)| = |z_n(0)| \quad \forall\, n \in \mathbb{N}$ (by (4.6)),

$$\|Uf\|_{L^2}^2 = \sum_{n=1}^{\infty} f_n^2 = \|f\|_{L^2}^2$$

which means that U is unitary. But from Lemma 6, this implies $\left[\frac{\beta(0)}{\alpha(0)}\right]^{1/4} = 1$ and $K(x,t) = 0$. So

(4.24) $\qquad \alpha(0) = \beta(0)$

(4.25) $\qquad \psi(x;\lambda) = \left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} \phi(\rho(x);\lambda)$

Now, due to (4.25), we have

$$\frac{d}{dx}\left(\beta(x)\frac{d\psi}{dx}\right) + \lambda\psi = [\alpha(\rho(x))]^{1/4}[\beta(x)]^{3/4} \frac{d^2\phi(\rho(x);\lambda)}{dx^2} +$$

$$+ \frac{1}{2}\left\{\left[\frac{\beta(x)}{\alpha(\rho(x))}\right]^{1/4} \alpha'(\rho(x)) + \left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} \beta'(x)\right\} \frac{d\phi(\rho(x);\lambda)}{dx} +$$

$$+ \left\{\frac{d}{dx}\left(\beta(x)\frac{d}{dx}\left(\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4}\right)\right)\right\} \phi(\rho(x);\lambda) + \lambda\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} \phi(\rho(x);\lambda) =$$

$$= \left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} \left\{ \frac{d}{d\rho(x)} \left(\alpha(\rho(x)) \frac{d\phi(\rho(x);\lambda)}{d\rho(x)}\right) + \lambda\phi(\rho(x);\lambda)\right\}$$

$$+ \left\{ \frac{d}{dx} \left(\beta(x) \frac{d}{dx}\left(\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4}\right)\right)\right\}\phi(\rho(x);\lambda)$$

Since $\phi(x;\lambda)$ and $\psi(x;\lambda)$ are solutions of (4.20) and (4.21) respectively, it follows that

$$\frac{d}{dx}\left(\beta(x) \frac{d}{dx}\left(\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4}\right)\right) = 0$$

Integrating and taking into account (4.24) we find

$$\left[\frac{\alpha(\rho(x))}{\beta(x)}\right]^{1/4} = 1 + c \int_0^x \frac{d\xi}{\beta(\xi)}$$

hence

$$\rho(x) = \int_0^x \left(1 + c \int_0^\xi \frac{d\xi'}{\beta(\xi')}\right)^2 d\xi$$

Finally, since $\rho(\ell) = \ell$ and $\int_0^x \frac{d\xi}{\beta(\xi)}$ is a strictly increasing positive function, it easily follows that $c = 0$. So $\rho(x) = x$, hence $\alpha(x) = \beta(x)$, hence $\psi(x;\lambda) = \phi(x;\lambda)$, hence $h = \hat{h}$ and $H = \hat{H}$. This completes the proof.

5. IDENTIFIABILITY AND NON-IDENTIFIABILITY RESULTS. Using Theorem 2 we can now solve completely Problems 1 and 2, which were posed in Section 3. Also we can solve a special case of Problem 3, namely that for which $x_p = \frac{\ell}{2}$.

RESULT 1: *Consider Problem 1. To a known input* $Q(t)$, *a known initial state* u_0 *and a given measurement,* $z_d(t) = u(0,t)$, $t \in]0,T]$ *there corresponds a unique* $\alpha(x)$.

Proof: Immediate consequence of Theorem 2.

It is noteworthy that Kitamura and Nakagiri [4, p. 794] claimed that it is impossible to identify a spatially varying parameter uniquely from a single point measurement $u(x_p,t)$. The above result is a counterexample to their claim. The next result will establish the fact that Problem 2 has in general a non-unique solution. We first establish the following lemma:

LEMMA 8: *Consider* (3.4) *with eigenvalues* λ_n *and normalized eigenfunctions* $y_n(x)$. *Also consider*

$$\frac{d}{dx}\left(\alpha(\ell-x)\frac{dy}{dx}\right) + \lambda y = 0$$

$$(5.1) \qquad y'(0) = 0$$

$$y'(\ell) = 0$$

with eigenvalues $\bar{\lambda}_n$ *and normalized eigenfunctions* $\bar{y}_n(x)$. *Then*

$$(5.2) \qquad \begin{aligned} \bar{\lambda}_n &= \lambda_n \\ \bar{y}_n(x) &= y_n(\ell-x) \end{aligned} \qquad \forall n \in \mathbb{N}$$

Proof: Under the affine transformation

$$\bar{x} = \ell - x$$

(5.1) reduces to (3.4). (5.2) follows immediately.

RESULT 2. *If* $u(x,t)$ *is the solution of* (3.1) *and* $\bar{u}(x,t)$ *is the solution of*

$$(5.3)\qquad \begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial}{\partial x}\left(\alpha(\ell-x)\,\dfrac{\partial u}{\partial x}\right) & , \textit{ in } \;]0,\ell[\times]0,T] \\ u(x,0) = u_o & , \textit{ in } \;]0,\ell[\\ \alpha(0)\,\dfrac{\partial u}{\partial x}(0,t) = Q(t) & , \textit{ in } \;]0,T] \\ \dfrac{\partial u}{\partial x}(\ell,t) = 0 & , \textit{ in } \;]0,T] \end{cases}$$

then

$$(5.4)\qquad \bar{u}(\ell,t) = u(\ell,t)$$

Proof: We have

$$u(\ell,t) = u_o - \int_0^t \left\{ \sum_{n=1}^{\infty} y_n(0)y_n(\ell)e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

$$\bar{u}(\ell,t) = u_o - \int_0^t \left\{ \sum_{n=1}^{\infty} \bar{y}_n(0)\bar{y}_n(\ell)e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

Using Lemma 8, we immediately conclude that

$$y_n(0)y_n(\ell) = \bar{y}_n(0)\bar{y}_n(\ell) \qquad \forall\, n \in \mathbb{N}$$

Hence the result.

LEMMA 9: *Consider again* (3.4). *If* $\alpha(x)$ *is symmetric, i.e.* $\alpha(x) = \alpha(\ell-x)$, *then*

$$(5.5)\qquad y_n(x) = y_n(\ell-x).$$

Proof: Immediate consequence of Lemma 8.

RESULT 3: *Consider again* (3.1) *and assume that* $\alpha(x)$ *is symmetric. Then to a known* $Q(t)$, *a known* u_o *and a given measurement*

$$z_d(t) = u(\ell,t)$$

there corresponds a unique $\alpha(x)$.

Proof: We have

$$u(\ell,t) = u_o - \int_0^t \left\{ \sum_{n=1}^{\infty} y_n(0)y_n(\ell)e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

From (5.5) we have $y_n(0) = y_n(\ell)$, hence

$$z_d(t) = u(\ell,t) = u_o - \int_0^t \left\{ \sum_{n=1}^{\infty} |y_n(0)|^2 e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

Thus the problem of identifying $\alpha(x)$ from $Q(t)$, u_o and $z_d(t)$ reduces to the one of identifying $\alpha(x)$ from $\{\lambda_n\}$ and $\{|y_n(0)|\}$. Hence by Theorem 2 $\alpha(x)$ is unique.

The next two results solve the special case $x_p = \frac{\ell}{2}$ for Problem 3.

RESULT 4: *If* $u(x,t)$ *is the solution of* (3.2) *and* $\bar{u}(x,t)$ *is the solution of*

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(\ell-x)\frac{\partial u}{\partial x} \right) + Q(t)\delta(x - \frac{\ell}{2}) \quad , \quad \textit{in }]0,\ell[\times]0,T]$$

$$u(x,0) = u_o \quad , \quad \textit{in }]0,\ell[$$

$$\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(\ell,t) = 0 \quad , \quad \textit{in }]0,T]$$

then $\bar{u}(\frac{\ell}{2}, t) = u(\frac{\ell}{2}, t)$

Proof: We have

$$u\left(\frac{\ell}{2},t\right) = u_o + \int_0^t \left\{ \sum_{n=1}^{\infty} [y_n(\tfrac{\ell}{2})]^2 e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

$$\bar{u}\left(\frac{\ell}{2},t\right) = u_o + \int_0^t \left\{ \sum_{n=1}^{\infty} [\bar{y}_n(\tfrac{\ell}{2})]^2 e^{-\lambda_n(t-\tau)} \right\} Q(\tau)\, d\tau$$

Using Lemma 8, we immediately conclude that

$$y_n(\tfrac{\ell}{2}) = \bar{y}_n(\tfrac{\ell}{2}) \qquad \forall\, n \in \mathbb{N}$$

Hence the result.

RESULT 5: *Consider again* (3.2) *and assume that* $\alpha(x)$ *is symmetric. Then to a known* $Q(t)$, *a known* u_o *and a given measurement*

$$z_d(t) = u(\tfrac{\ell}{2},t)$$

there corresponds a unique $\alpha(x)$.

Proof: Since $\alpha(x)$ is symmetric, $y_n(x)$ are symmetric $\forall\, n \in \mathbb{N}$ by Lemma 9. Hence,

$$y_n'(\tfrac{\ell}{2}) = 0 \qquad \forall\, n \in \mathbb{N}$$

Now consider the Sturm-Liouville problem

$$\frac{d}{dx}\left(\alpha(x)\frac{dy}{dx}\right) + \lambda y = 0$$

$$(5.7) \qquad y'(0) = 0$$

$$y'(\tfrac{\ell}{2}) = 0$$

and observe that (3.4) and (5.7) have the same eigenvalues λ_n and the same normalized eigenfunctions $y_n(x)$, $x \in [0, \frac{\ell}{2}]$. Thus the problem of identifying

$\alpha(x)$ from $Q(t)$, u_o and $z_d(t)$ reduces to the one of identifying $\alpha(x)$ in (5.7) from $\{\lambda_n\}$ and $\{|y_n(\frac{\ell}{2})\ |\}$. Hence by Theorem 2 $\alpha(x)$ is unique.

6. CONCLUSIONS AND SIGNIFICANCE

In this paper the problem of identifiability of spatially varying conductivity from point measurement of temperature in the linear, one-dimensional heat equation is addressed. Uniqueness and non-uniqueness results are derived referring to special cases of the above general problem. More specifically, in Problem 1 (system (3.1)) we have shown that in a rod, which is insulated at one end and heated at the other end (with known heat flux), measurement of the temperature as a function of time at the heated end determines uniquely the conductivity as a function of position. Uniqueness is not obtained, however, when in the above physical system the measurement is placed at the insulated end (Problem 2). In Problem 3 (system (3.2)) we have addressed the case of a rod, which is insulated at both ends, with a known heat source at the point $x = x_p$ and a temperature measurement as a function of time at x_p. We have shown that in the special case where both the heat source and the sensor are placed in the middle of the rod ($x_p = \frac{\ell}{2}$), there corresponds in general a non-unique conductivity. Only in the highly exceptional case where the conductivity is a symmetric function (with respect to the middle of the rod) Problems 2 and 3 can have a unique solution.

There are still important questions that remain unanswered, such as

(i) Is the system described by (3.1) identifiable if $0 < x_p < \ell$? In other words, is uniqueness the "rule" or the "exception"?

(ii) In the system described by (3.2), is the point $x_p = \frac{\ell}{2}$ an "exceptional" or a "typical" point? What happens for other x_p's?

The above questions reduce to inverse Sturm-Liouville problems for which, at the moment, results are not available. However, we intuitively expect that uniqueness will be the case for all x_p's, except for a set of measure zero.

The motivation of undertaking this work is to establish identifiability conditions for (1.1). Since, for example, (1.1) governs the pressure distribution in petroleum reservoirs and subsurface aquifers, and since the identification of $\alpha(x,y)$ is a key problem in describing these systems, elucidation of the fundamental question of identifiability will have a significant impact on the estimation of such parameters.

APPENDIX: SOLUTION OF THE CAUCHY PROBLEM FOR $(\alpha u_x)_x - (\beta u_y)_y = 0$ *BY RIEMANN'S METHOD.*

Consider

$$\frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\beta(y)\,\frac{\partial u}{\partial y}\right) = 0$$

(A.1) $$u)_{y=o} = f(x)$$

$$\frac{\partial u}{\partial y})_{y=o} = g(x)$$

By making the change of variable

$$X = \int_0^x \frac{dx'}{\sqrt{\alpha(x')}} = \mathscr{A}(x)$$

(A.2)

$$Y = \int_0^y \frac{dy'}{\sqrt{\beta(y')}} = \mathscr{B}(y)$$

(A.1) becomes

$$\frac{\partial^2\tilde{u}}{\partial X^2} - \frac{\partial^2\tilde{u}}{\partial Y^2} + 2\tilde{a}(X)\,\frac{\partial\tilde{u}}{\partial X} + 2\tilde{b}(Y)\,\frac{\partial\tilde{u}}{\partial Y} = 0$$

(A.3) $$\tilde{u})_{Y=o} = \tilde{f}(X)$$

$$\frac{\partial\tilde{u}}{\partial Y})_{Y=o} = \sqrt{\beta(0)}\;\tilde{g}(X)$$

where

$$a = \frac{\frac{d\alpha}{dx}}{4\sqrt{\alpha}}$$

(A.4)

$$b = -\frac{\frac{d\beta}{dy}}{4\sqrt{\beta}}$$

and where we have used the symbols $\tilde{u}(X,Y)$, $\tilde{a}(X)$, $\tilde{b}(Y)$, $\tilde{f}(X)$, $\tilde{g}(X)$ in place of $u(\mathscr{A}^{-1}(X), \mathscr{B}^{-1}(Y))$, $a(\mathscr{A}^{-1}(X))$, $b(\mathscr{B}^{-1}(Y))$, $f(\mathscr{A}^{-1}(X))$, $g(\mathscr{A}^{-1}(X))$ respectively.

For the solution of (A.3) we make use of Riemann's formula [19, p. 81] (See Figure 1),

$$\text{(A.5)}\quad \tilde{u}(P) = \frac{1}{2}\,[\tilde{u}(Q)\tilde{\nu}(Q) + \tilde{u}(R)\tilde{\nu}(R)] + \frac{1}{2}\int_{QR}\left(\tilde{\nu}\,\frac{\partial\tilde{u}}{\partial Y} - \tilde{u}\,\frac{\partial\tilde{\nu}}{\partial Y} - 2\tilde{b}\tilde{u}\tilde{\nu}\right)dX + \left(\tilde{\nu}\,\frac{\partial\tilde{u}}{\partial X} - \tilde{u}\,\frac{\partial\tilde{\nu}}{\partial X} + 2\tilde{a}\tilde{u}\tilde{\nu}\right)dY$$

where $\tilde{\nu}(X,Y;X_o,Y_o)$ is the solution of

$$\frac{\partial^2\tilde{\nu}}{\partial X^2} - \frac{\partial^2\tilde{\nu}}{\partial Y^2} - 2\tilde{a}(X)\,\frac{\partial\tilde{u}}{\partial X} - 2\tilde{b}(Y)\,\frac{\partial\tilde{\nu}}{\partial Y} - \left(\frac{d\tilde{a}}{dX} + \frac{d\tilde{b}}{dY}\right)\tilde{\nu} = 0$$

$$\text{(A.6)}\quad \begin{aligned} &\frac{\partial\tilde{\nu}}{\partial X} - \frac{\partial\tilde{\nu}}{\partial Y} = (\tilde{a}+\tilde{b})\tilde{\nu} && \text{on } Y + X = Y_o + X_o \\ -&\frac{\partial\tilde{\nu}}{\partial X} + \frac{\partial\tilde{\nu}}{\partial Y} = (\tilde{a}-\tilde{b})\tilde{\nu} && \text{on } Y - X = X_o - X_o \end{aligned}$$

$$\tilde{\nu}(X_o,Y_o) = 1$$

It can be easily seen that the boundary conditions in (A.6) are equivalent to

$$\text{(A.7)}\quad \tilde{\nu} = \exp\left[\int_{X_o}^{X}\tilde{a}(X')dX' - \int_{Y_o}^{Y}\tilde{b}(Y')dY'\right] \text{ on the lines } \begin{cases} Y + X = Y_o + X_o \\ Y - X = Y_o - X_o \end{cases}$$

Thus (A.5) gives

(A.8) $$\tilde{u}(X_o,Y_o) = \frac{1}{2} \exp\left[\int_0^{Y_o} \tilde{b}(Y')dY'\right]\left\{ \tilde{f}(X_o-Y_o) \exp\left[\int_{X_o}^{X_o-Y_o} \tilde{a}(X')dX'\right]\right.$$

$$\left. + \tilde{f}(X_o + Y_o) \exp\left[\int_{X_o}^{X_o+Y_o} \tilde{a}(X')\, dX'\right]\right\}$$

$$+ \frac{1}{2} \int_{X_o-Y_o}^{X_o+Y_o} W_1(X_o,Y_o,t)\tilde{f}(t)\, dt + \frac{1}{2} \int_{X_o-Y_o}^{X_o+Y_o} W_2(X_o,Y_o,t)\tilde{g}(t)\, dt$$

where

$$W_1(X_o,Y_o,t) = -\frac{\partial\tilde{\nu}}{\partial Y}(t,0;X_o,Y_o) - 2b(0)\tilde{\nu}(t,0;X_o,Y_o)$$

$$W_2(X_o,Y_o,t) = \sqrt{\beta(0)}\tilde{\nu}(t,0;X_o,Y_o)$$

Finally, taking into account (A.4), the integrals of the first term of (A.8) can be easily evaluated. Thus (A.8) (dropping the subscript o) gives

(A.9) $$\tilde{u}(X,Y) = \frac{1}{2}\left[\frac{\beta(0)}{\tilde{\beta}(Y)\tilde{\alpha}(X)}\right]^{1/4}\left\{[\tilde{\alpha}(X-Y)]^{1/4}\tilde{f}(X-Y) + [\tilde{\alpha}(X+Y)]^{1/4}\tilde{f}(X+Y)\right\}$$

$$+ \frac{1}{2} \int_{X-Y}^{X+Y} W_1(X,Y,t)\tilde{f}(t)\, dt + \frac{1}{2} \int_{X-Y}^{X+Y} W_2(X,Y,t)\tilde{g}(t)\, dt$$

where $\tilde{\alpha}(\cdot) = \alpha(\mathscr{A}^{-1}(\cdot))$ and $\tilde{\beta}(\cdot) \equiv \beta(\mathscr{B}^{-1}(\cdot))$.

REFERENCES

[1] J. R. CANNON, J. DOUGLAS AND B. F. JONES, *Determination of the diffusivity of an isotropic medium*, Int. J. Engng. Sci., 1 (1963), pp. 453-455.

[2] J. R. CANNON AND B. F. JONES, *Determination of the diffusivity of an anisotropic medium*, Int. J. Engng. Sci., 1, (1963), pp. 457-460.

[3] J. R. CANNON AND PAUL DUCHATEAU, *Determination of the conductivity of an isotropic medium*, J. Math. Anal. Appl., 48 (1974), pp. 699-704.

[4] S. KITAMURA AND S. NAKAGIRI, *Identifiability of spatially-varying and constant parameters in distributed systems of parabolic type*, SIAM J. Control Optim., 15 (1977), pp. 785-802.

[5] G. BORG, *Eine umkehrung der Sturm-Liouvilleschen eigenwertaufgabe*, Acta Math., 78 (1946), pp. 1-96.

[6] I. M. GEL'FAND AND B. M. LEVITAN, *On the determination of a differential equation from its spectral function*, Izv. Akad. Nauk. SSSR Ser. Mat., 15 (1951), pp. 309-360. Amer. Math. Soc. Transl. (Ser. 2), 1 (1955), pp. 253-304.

[7] N. LEVINSON, *The inverse Sturm-Liouville problem*, Mat. Tidsskr. B, (1949), pp. 25-30.

[8] M. G. KREIN, *Solution of the inverse Sturm-Liouville problem*, Dokl. Akad. Nauk. SSSR (N.S.), 76 (1951), pp. 21-24.

[9] M. G. KREIN, *Determination of the density of a non-homogeneous symmetric cord by its frequency spectrum*, Dokl. Akad. Nauk. SSSR (N.S.) 76 (1951), pp. 345-348.

[10] B. M. LEVITAN, *On the determination of a differential equation by two of its spectra*, Izv. Akad. Nauk. SSSR, 28 (1964), pp. 63-78. Amer. Math. Soc. Transl. (Ser. 2), 68 (1968), pp. 1-20.

[11] O. A. LADYZENSKAJA, V. A. SOLONNIKOV AND N. N. URAL'CEVA, *Linear and Quasi-linear Equations of Parabolic Type*, American Mathematical Society, Providence, RI, 1968.

[12] M. COURDESSES, *Comments on identifiability of spatially varying and constant parameters in distributed systems of parabolic type*, SIAM J. Control Optim., 21 (1983), pp. 410-412.

[13] A. PIERCE, *Unique identification of eigenvalues and coefficients in a parabolic problem*, SIAM J. Control Optim., 17 (1979), pp. 494-499.

[14] R. E. GOODSON AND R. E. KLEIN, *A definition and some results for distributed system observability,* IEEE Trans. Autom. Control, AC-15 (1970), pp. 165-174.

[15] M. COURDESSES, M. P. POLIS AND M. AMOUROUX, *On identifiability of parameters in a class of parabolic distributed systems,* IEEE Trans. Autom. Control, AC-26 (1981), pp. 474-477.

[16] E. C. TITCHMARSH, *Introduction to the Theory of Fourier Integrals,* 2nd Ed., Clarendon Press, Oxford, 1948.

[17] G. BIRKHOFF AND G. C. ROTA, *Ordinary Differential Equations,* 3rd. Ed., Wiley, New York, 1968.

[18] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics,* Vol. 1, Interscience, New York, 1953.

[19] E. T. COPSON, *Partial Differential Equations,* Cambridge University Press, Cambridge, 1975.

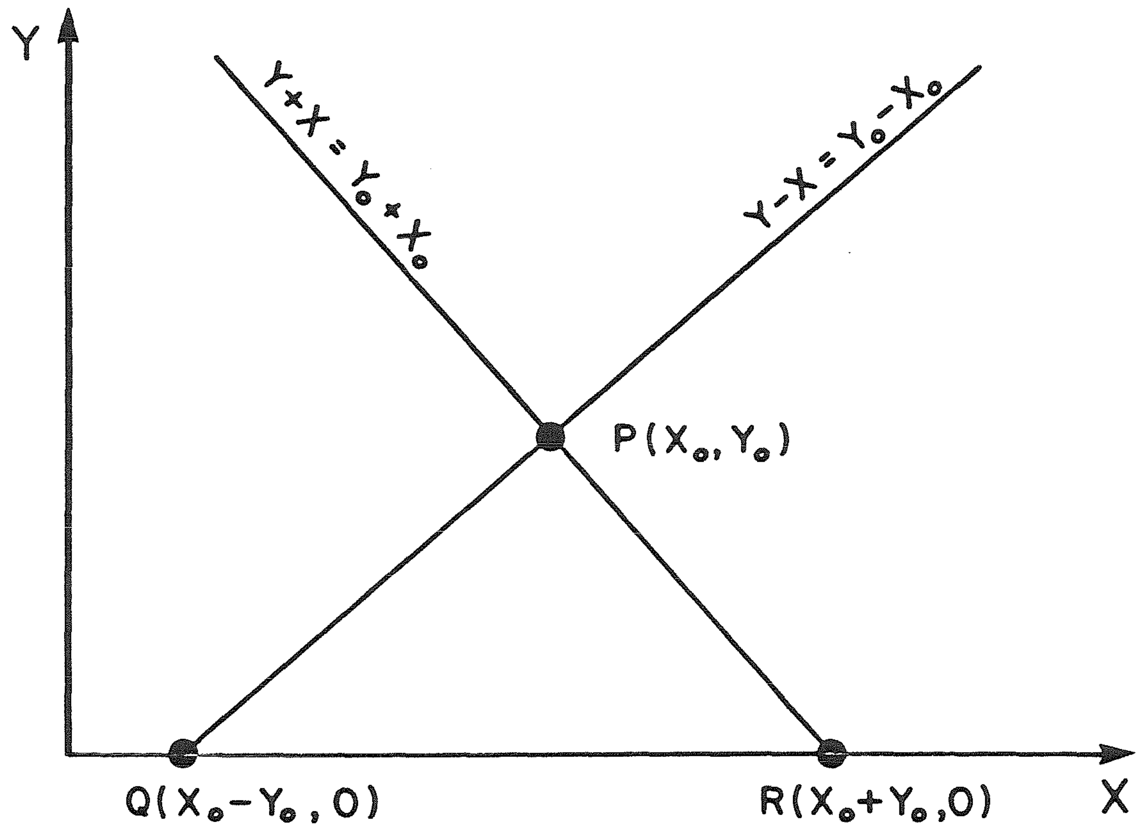


Figure 1

CHAPTER III: IDENTIFICATION OF PARAMETERS IN DISTRIBUTED PARAMETER SYSTEMS BY REGULARIZATION

1. INTRODUCTION

Consider the following distributed parameter dynamic system:

$$\frac{\partial u}{\partial t} + A(t)u = f \quad , \qquad \text{in } \Omega \times]0,T[\tag{1.1}$$

$$u(x,0) = u_o \quad , \qquad \text{in } \Omega \tag{1.2}$$

$$B_j u = g_j \quad , \; j = 0,\ldots, m-1, \text{ on } \Gamma\times]0,T[\tag{1.3}$$

where $\Omega \subset \mathbb{R}^n$ with boundary Γ and $0 < T < \infty$ and where

$$A(t)\, u = \sum_{|p|,|q|\leq m} (-1)^{|p|} D_x^p(a_{pq}(x,t)\; D_x^q u) \tag{1.4}$$

$$B_j u = \sum_{|h|\leq m_j} b_{jh}(x,t) D_x^h u \quad , \quad j = 0,\ldots, m-1 \tag{1.5}$$

with $0 \leq m_j$ = order of $B_j \leq 2m-1$.

The parameter identification problem associated with (1.1)-(1.5) can be stated as follows: Assuming the input function f, the initial condition and the boundary condition(s) to be known, and given an observation of u, determine the system operator A(t), i.e. the parameters $a_{pq}(x,t)$.[1]

A number of important physical identification problems fall within the above framework. For example, the partial differential equation

[1]The boundary condition parameters $b_{jh}(x,t)$ may also be unknown, although we do not consider that case here.

$$\frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x,y)\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\alpha(x,y)\frac{\partial u}{\partial y}\right) = f(x,y,t) \qquad (1.6)$$

governs the temperature distribution in an inhomogeneous solid or the pressure distribution in a fluid-containing porous medium. The local flux of energy or fluid is dependent on the value of the parameter α. For example, in the case of fluid flow in a porous medium, α is termed the transmissivity. For models of petroleum reservoirs and subsurface aquifers the transmissivity is generally inaccessible to direct measurement, and its value must be inferred from measurements of the pressure u at wells. Because of the economic importance of knowing accurately the properties of subsurface aquifers and petroleum reservoirs, a great deal of effort has been expended in developing techniques for determining transmissivity from measurements of pressure ([5], [9], [10], [12], [14], [24], [25], [32]). The determination of α from data on u is a special case of the general linear parabolic system identification problem introduced at the outset. Specifically, given f, the initial condition and appropriate boundary conditions, and given measurements z_{d_i} of $u(x_i,y_i,t)$ at a set of discrete spatial locations, $i = 1,2,\ldots,\mu$, it is desired to determine, or identify, $\alpha(x,y)$.

The key difficulty in developing successful numerical techniques for identifying spatially-dependent parameters is the fact that such problems are ill-posed. To see this, consider (1.6) as a first order hyperbolic equation in α. One can easily show that the characteristics $\omega(x,y) = c$ are orthogonal to the lines of constant u. Thus, one can define a new curvilinear coordinate system $(\vec{e}_u,\vec{e}_\omega)$ so that $\vec{e}_u$ is unitary in $\mathbb{R}^2$ and the metric factor in the ω-coordinate is 1. Equation (1.6) can be written as

$$|\nabla u| \frac{\partial}{\partial u}\left(\alpha|\nabla u|\right) = \frac{\partial u}{\partial t} - f \tag{1.7}$$

where $|\nabla u| = \left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right)^{1/2}$. Upon integration of (1.7), we obtain

$$\alpha(x,y) = \frac{\int(\frac{\partial u}{\partial t} - f)d\ell_\omega}{|\nabla u|} \tag{1.8}$$

where integration is performed along the characteristics and ℓ_ω denotes Lebesgue measure along the characteristics. Ill-posedness follows from the fact that the differentiation operator is not continuous with respect to any physically meaningful observation topology. The fact that the identification problem associated with (1.6) is not well-posed can also be illustrated by counterexample ([32]). In summary, the problem of identifying spatially-dependent coefficients appearing in the differential operator of a partial differential equation is, in general, both nonlinear and ill-posed ([16],[17]).

The customary way to approach the identification of α in (1.6) has been by least-squares, i.e. by minimizing the functional

$$J_{LS} = \int_0^T \sum_{i=1}^{\mu} [u(x_i,y_i,t) - z_{d_i}]^2 \, dt \tag{1.9}$$

subject to (1.6), initial and boundary conditions. There have been two ways of treating the unknown parameter α. In the first, α is considered as an element of an infinite-dimensional function space ([9],[5]), whereas in the second, the minimization is performed over a finite-dimensional subspace, reducing the problem to one of determining a finite number of constant parameters ([10],[12]). When the number of parameters is kept small in this approach, a well-behaved solution results. However, the modeling error introduced is significant, since the corresponding subspace of α's is too restricted to provide a good approximation of an arbitrary α. As the number of parameters is increased, on the other hand, numerical instabilities appear, manifested by spatial oscillations in the estimated α, the frequency and amplitude of which are inconsistent with the expected smoothness of the true α. The symptoms usually also include a flat global minimum in J_{LS} ([24], [25], [32]). The same instability phenomena characterize the minima of J_{LS} over an infinite-dimensional function space. One approach that has been somewhat successful in alleviating numerical instabilities involves the incorporation of a priori statistics concerning α into the minimization by adding a Bayesian term in the performance index (1.9) ([14], [24]). The major drawback to this approach is that reliable a priori statistics for α are seldom available. Thus, there is a need to develop a rigorously based approach for identifying parameters in partial differential equations from noisy data that is numerically stable and physically consistent with the expected character of the unknown parameters.

The numerical instabilities and ill-posed nature of the problem of interest strongly suggest a regularization approach. "Regularization" of a problem refers in general to solving a related problem, called the regularized problem, the solution of which is more regular, in a sense, than that of the original problem and approximates the solution of the original problem. When referring to ill-posed problems, regularization is an approach to circumvent lack of continuous dependence on the data. The regularized problem is a well-posed problem whose solution yields a physically meaningful answer to the given ill-posed problem.

The idea of regularization of ill-posed problems was first proposed by Tikhonov ([27], [28]) as a method of solving linear Fredholm integral equations of the first kind. Further development of the theory for ill-posed linear operator equations followed ([23]). Modern practical numerical methods for the solution of linear Fredholm integral equations involve regularization ([31]).

The object of the present work is to develop a regularization theory for the identification of parameters in distributed parameter systems. In §2 we define the parameter identification problem in an abstract manner that facilitates proof of the major theorems. The concept of identifiability is discussed in §3. In §4 and §5 a general regularization identification theory is presented. In §6, §7 and §8 the theory is applied to the identification of parabolic systems from distributed and point observations. Finally, in §9 numerical results are given for the identification of a spatially-varying diffusivity in the one-dimensional diffusion equation.

2. PROBLEM STATEMENT

To develop a general identification theory, we introduce the following abstract problem.

Let $\mathcal{A}$, U and F be Banach spaces. Consider a system described by

$$\Psi(A,u) = f \tag{2.1}$$

where Ψ is a mapping, not necessarily linear, from $\mathcal{A} \times U$ into F. We assume:

(A1) Ψ is of C^k-class $(k \geq 1)$

(A2) There is an open subset $\mathcal{A}_c$ of $\mathcal{A}$ and an open subset U_c of U such that $\forall A \in \mathcal{A}_c$ (2.1) admits a unique solution $u \in U_c$.

(A3) $\forall A \in \mathcal{A}_c \quad \forall u \in U_c \quad \frac{\partial \Psi}{\partial u}(A,u)$ is a linear homeomorphism of U onto F.

Furthermore, consider that A depends on a set of parameters λ belonging to the Banach space Λ. The set of physically admissible λ is $\Lambda_{ad} \subseteq \Lambda$. We assume:

(A4) $A;\ \Lambda \to \mathcal{A}$ is of C^k-class $(k \geq 1)$

(A5) Λ_{ad} is a norm-closed convex subset of Λ

(A6) $A(\Lambda_{ad}) \subseteq \mathcal{A}_c$

Now from the implicit function theorem ([26, pp. 277-304]) we have:

PROPOSITION 2.1: *Assume* (A1)-(A3) *are valid. Then the implicit function,* $u = \Phi(A)$, *defined as the solution of* (2.1) *is of* C^k*-class from* $\mathcal{A}_c$ *into* U_c. *Its first derivative is given by*

$$\Phi'(A) = -\left[\frac{\partial\Psi}{\partial u}(A,u)\right]^{-1} \circ \left[\frac{\partial\Psi}{\partial A}(A,u)\right] \qquad \forall A \in \mathcal{A}_c \tag{2.2}$$

Equivalently, $\Phi'(A)$ *associates* $\delta A \in \mathcal{A} \to \delta u \equiv \Phi'(A) \cdot \delta A \in U$*, where* δu *is the solution of*

$$\frac{\partial\Psi}{\partial u}(A,u) \cdot \delta u + \frac{\partial\Psi}{\partial A}(A,u) \cdot \delta A = 0 \tag{2.3}$$

As an immediate consequence, we have:

PROPOSITION 2.2. *Assume that* (A1)-(A4) *and* (A6) *are valid. Then* $\Phi \circ A;\ \Lambda_{ad} \to U$ *is of* C^k*-class. Its first derivative* $(\Phi \circ A)'(\lambda)$ *associates* $\delta\lambda \in \Lambda_{ad} \to \delta u \in U$*, where* δu *is the solution of*

$$\frac{\partial\Psi}{\partial u}\Big(A(\lambda),u\Big) \cdot \delta u + \frac{\partial\Psi}{\partial A}\Big(A(\lambda),u\Big) \circ A'(\lambda) \cdot \delta\lambda$$

Now the identification problem can be posed as follows:

Knowing the mappings $\Psi;\mathcal{A} \times U \to F$ and $A;\ \Lambda \to \mathcal{A}$ and the element $f \in F$ and given an observation of u, find $\lambda \in \Lambda_{ad}$ to satisfy (2.1).

We need to be precise about the nature of the observation of u. Thus, consider a Hilbert Space $\mathcal{H}$ (Observation Space). Denote by $\Lambda_{\mathcal{H}}$ the canonical isomorphism of $\mathcal{H}$ onto $\mathcal{H}'$. Also, consider an observation operator, not necessarily linear, $\mathcal{C};\ U \to \mathcal{H}$ and assume

(A7) $\mathcal{C}$ is of C^k-class $(k \geq 1)$

The situation is depicted in Figure 1.

3. IDENTIFIABILITY

The identification problem, as defined in §2 can be viewed as solving in Λ_{ad} the (nonlinear) operator equation

$$(\mathscr{C} \circ \Phi \circ A)(\lambda) = z_d \tag{3.1}$$

Before one develops an identification method, a key issue to be examined is the well-posedness (in the sense of Hadamard) of the problem of solving (3.1).

Existence of a solution of (3.1) for arbitrary $z_d \in \mathscr{H}$ is of no interest, since in any physical identification problem there is always a true λ in Λ_{ad} corresponding to an error-free observation of u. In other words, well-posedness needs to be examined for $z_d \in \mathscr{C}(\Phi(A(\Lambda_{ad})))$.

Uniqueness of the solution of (3.1) is commonly termed identifiability.

DEFINITION 3.1 ([4], [15], [7]): A parameter λ is said to be *identifiable* in Λ_{ad} for the observation operator $\mathscr{C}$, if the mapping (parameter → observation) is injective, i.e. if $\mathscr{C} \circ \Phi \circ A ; \Lambda_{ad} \to \mathscr{H}$ has a unique inverse.

DEFINITION 3.2: A parameter λ is said to be *stable* in Λ_{ad} for the observation operator $\mathscr{C}$, if $(\mathscr{C} \circ \Phi \circ A)^{-1}$ is continuous.

Remark: In case of non-identifiable λ, stability is understood in the sense of continuity of multiple-valued mappings.

The identifiability of the parameter $\alpha(x)$ in

$$\frac{\partial u}{\partial t} - \sum_{j=1}^{n} \frac{\partial}{\partial x_j} \left(\alpha(x) \frac{\partial u}{\partial x_j} \right) = f \tag{3.2}$$

from a distributed observation of u, i.e. an observation of $u(x,t)$ in $\Omega \times]0,T[$, has been studied by [15] (one spatial dimension) and [4] (several spatial dimensions). It has been shown that in general $\alpha(x)$ need not be unique [15]. However, if the set $E(t) = \left\{ x \in \bar{\Omega} \,\middle|\, \frac{\partial u}{\partial x}(x,t) = 0 \right\}$ is nonempty for every $t \in]0,T[$ and $\bigcap_{t \in]0,T[} E(t)$ is of measure zero, then $\alpha(x)$ is unique. The result in [4] is similar to that in [15], but involves quite restrictive assumptions concerning ∇z_d and $\nabla^2 z_d$. Thus, the available results establish that $\alpha(x)$ in (3.2) is not identifiable in the sense of Definition 3.1; however, under certain additional conditions on z_d, there corresponds a unique $\alpha(x)$.

No results appear to be available concerning the identifiability of $\alpha(x)$ in (3.2) from point observation of u. It is, however, possible to establish uniqueness of $\alpha(x)$ in special cases, by reducing the problem to an inverse Sturm-Liouville problem.

For a review of the (limited) available identifiability results in various distributed parameter identification problems, the reader is referred to [7].

Due to the conditional nature of most identifiability results, we find it important to introduce a weaker concept of identifiability.

DEFINITION 3.3: A parameter λ is said to be *conditionally identifiable* in Λ_{ad} with respect to $\mathscr{H}_c \subset \mathscr{H}$, if the restriction of the mapping $\mathscr{C} \circ \Phi \circ A$; $\Lambda_{ad} \to \mathscr{H}$ on the set $(\mathscr{C} \circ \Phi \circ A)^{-1}\, \mathscr{H}_c$ has a unique inverse.

A degenerate case of conditional identifiability is obtained when $\mathscr{H}_c$ is a point set, i.e. $\mathscr{H}_c = \{\tilde{z}_d\} \subset \mathscr{C}(\Phi(A(\Lambda_{ad})))$.

DEFINITION 3.4: A parameter λ is said to be *pointwise identifiable* in Λ_{ad} for the observation $\tilde{z}_d \in \mathscr{C}(\Phi(A(\Lambda_{ad})))$, if $\tilde{z}_d$ has a unique preimage with respect to the mapping $\mathscr{C} \circ \Phi \circ A$; $\Lambda_{ad} \to \mathscr{H}$.

The concept of pointwise identifiability is the weakest possible concept of identifiability. It will be used in §4 and §5 (Theorems 4.3 and 5.3).

As we have noted, the identification of distributed coefficients appearing in the differential operator of a partial differential equation is, as a rule, an unstable problem ([17]). The homogenization theory ([2]) shows that operators with highly oscillatory coefficients can be "replaced" by very different operators and still yield practically the same response. Lions [17] has, in fact, cited the main difficulty in identifying distributed coefficients in partial differential equations as preventing excess of oscillations in the coefficients.

To illustrate the power of homogenization theory in proving instability of identification problems, let us consider the problem of identifying $\alpha(x)$ in (3.2).

Let $Y =]0,y_1^o[\times]0,y_2^o[\times\cdots\times]0,y_n^o[\subset \mathbb{R}^n$ and $\alpha;\ \mathbb{R}^n \to \mathbb{R}$ a function with properties

(i) $\alpha \in L^\infty(\mathbb{R}^n)$

(ii) $\alpha(y) \geq \alpha_o > 0$ a.e. in y

(iii) $\alpha(y)$ is Y-periodic i.e. it admits a period y_j^o in the direction y_j, $j = 1,\ldots, n$

Denote $\alpha^\varepsilon(x) = \alpha\left(\frac{x}{\varepsilon}\right)$, $\varepsilon > 0$. Now given Ω a bounded open subset of $\mathbb{R}^n$ and $T > 0$, consider

$$\left.\begin{array}{l} \dfrac{\partial u_\varepsilon}{\partial t} - \displaystyle\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\alpha^\varepsilon(x)\,\frac{\partial u_\varepsilon}{\partial x_j}\right) = f, \text{ in } \Omega\times]0,T[\\ u_\varepsilon(x,0) = u_o(x) \quad , \quad \text{in } \Omega \\ \text{Boundary Condition} \end{array}\right\} \tag{3.3}$$

Observe that as $\varepsilon \to 0$, the α^ε's are *highly oscillating functions*. They converge in a weak sense:

$$\alpha^{\varepsilon} \rightarrow M(\alpha) \text{ in } L^{\infty}(\Omega) \text{ weak - } * \tag{3.4}$$

$$\left(\text{i.e. } \int_{\Omega} \alpha^{\varepsilon}\phi \, dx \rightarrow \int_{\Omega} M(\alpha)\phi \, dx \qquad \forall \phi \in L^{1}(\Omega)\right)$$

where $M(\alpha) = \frac{1}{\text{meas } Y} \int_{Y} \alpha(y) \, dy$

The question now concerns the behavior of the solution u_{ε} of (3.3) as $\varepsilon \rightarrow 0$. It is tempting to believe that u converges to the solution of

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} - M(\alpha) \, \Delta u = f \quad , \text{ in } \Omega \times]0,T[\\ &u(x,0) = u_0(x) \qquad , \text{ in } \Omega \\ &\text{Boundary Condition} \end{aligned}\right\} \tag{3.5}$$

But this is untrue ([2 , p.242]). The correct result is given by the following proposition, which is an immediate consequence of a general result for second-order parabolic systems ([2, pp. 241-243]).

PROPOSITION 3.1: *The solution* u_{ε} *of* (3.3) *converges in* $L^2(Q)$ *to the solution of the following homogenized problem*

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} - \frac{1}{M\left(\frac{1}{\alpha}\right)} \, \Delta u = f \quad , \quad \text{in } \Omega \times]0,T[\\ &u(x,0) = u_0(x) \qquad , \quad \text{in } \Omega \\ &\text{Boundary Condition} \end{aligned}\right\} \tag{3.6}$$

Thus, for sufficiently small ε, u_{ε} is approximately equal to the solution of (3.6); however α^{ε} and $\frac{1}{M\left(\frac{1}{\alpha}\right)}$ can be very different.

The least-squares approach to distributed parameter system identification ([4], [8]) can be stated as follows:

Given $z_d \in \mathscr{H}$, find $\bar{\lambda} \in \Lambda_{ad}$ to minimize the functional

$$J_{LS}(\lambda) = \left\| \mathscr{C}\left(\Phi(A(\lambda))\right) - z_d \right\|^2_{\mathscr{H}} \tag{3.7}$$

Conceptually, the least-squares approach consists of two steps:

(a) Project z_d in $\bar{z}_d$ on the set $\mathscr{C}\left(\Phi(A(\Lambda_{ad}))\right)$

(b) Find in Λ_{ad} a preimage $\bar{\lambda}$ of $\bar{z}_d$ for the mapping $\mathscr{C} \circ \Phi \circ A$

It is therefore natural to inquire if a projection of an arbitrary $z_d \in \mathscr{H}$ on the set $\mathscr{C}\left(\Phi(A(\Lambda_{ad}))\right)$ exists and is unique. Also, when z_d is perturbed slightly, does the perturbation correspond to a small change in $\bar{\lambda}$?

DEFINITION 3.3 ([6]): A parameter λ is said to be *output least-square identifiable* (OLSI) in Λ_{ad} for the observation operator $\mathscr{C}$, if there exists a neighborhood $\mathscr{N} \supset \mathscr{C}\left(\Phi(A(\Lambda_{ad}))\right)$ such that for every $z_d \in \mathscr{N}$ the least-squares estimate is unique and depends continuously on z_d.

It is easy to see that the following are necessary conditions for OLSI:

(i) Existence of a neighborhood $\mathscr{N} \supset \mathscr{C}\left(\Phi(A(\Lambda_{ad}))\right)$ such that every $z_d \in \mathscr{N}$ has a unique projection on $\mathscr{C}\left(\Phi(A(\Lambda_{ad}))\right)$.

(ii) Well-posedness of the identification problem for every $z_d \in \mathscr{C}\left(\Phi(A(\Lambda_{ad}))\right)$, i.e. both identifiability and stability of λ in Λ_{ad} w.r.t $\mathscr{C}$.

It has been shown in [6] that with Λ_{ad} convex and $\mathscr{C} \circ \Phi \circ A$ sufficiently regular, satisfaction of (i) can be guaranteed. Condition (ii) is the key one; unless a parameter is both identifiable and stable, the least squares approach will not produce a reliable estimate.

4. IDENTIFICATION BY REGULARIZATION

Let us return to the general identification problem of §2. In order to regularize the parameter λ, we introduce a more regular space $\mathscr{R}$, for which we assume:

(A8) $\mathscr{R}$ is a Hilbert space.

(A9) $\mathscr{R}$ is imbedded in Λ.

(A10) The imbedding operator from $\mathscr{R}$ into Λ is compact.

Define $\mathscr{R}_{ad} = \mathscr{R} \cap \Lambda_{ad}$. With (A5) and (A9) it readily follows that $\mathscr{R}_{ad}$ is a norm-closed convex subset of $\mathscr{R}$.

We now introduce the *stabilizing functional*

$$J_S(\lambda) = \|\lambda\|^2_{\mathscr{R}} , \qquad \lambda \in \mathscr{R}_{ad} \tag{4.1}$$

and the *smoothing functional*

$$J_\beta(\lambda) = J_{LS}(\lambda) + \beta J_S(\lambda), \quad \lambda \in \mathscr{R}_{ad} \tag{4.2}$$

where $\beta > 0$ is the *regularization parameter*. Identification by regularization proceeds as follows. Given $z_d \in \mathscr{H}$ and $\beta > 0$, find $\lambda_\beta \in \mathscr{R}_{ad}$ so as to minimize $J_\beta(\lambda)$.

In this section we establish the basic results concerning the regularization method. Our first result concerns differentiability of the functional $J_\beta(\lambda)$.

THEOREM 4.1: *Assume that* (A1)-(A4) *and* (A6)-(A9) *are valid. Then the functional*

$$J_\beta(\lambda) = \|\mathscr{C}[\Phi(A(\lambda))] - z_d\|^2_{\mathscr{H}} + \beta\|\lambda\|^2_{\mathscr{R}} \tag{4.3}$$

is of C^k*-class. Its first derivative* $J'_\beta(\lambda); \mathscr{R} \to \mathbb{R}$ *is given by*[2]

$$J'_\beta(\lambda)\cdot\delta\lambda = \left(\frac{\partial\Psi}{\partial A}(A(\lambda),u)\circ A'(\lambda)\cdot\delta\lambda,\rho\right)_{FF'} + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}} \tag{4.4}$$

where u *is the solution of* $\Psi(A(\lambda),u) = f$ *and* ρ *is the solution of*

$$\left[\frac{\partial\Psi}{\partial u}\left(A(\lambda),u\right)\right]^*\rho = -2[\mathscr{C}'(u)]^*\Lambda_{\mathscr{H}}\left(\mathscr{C}(u) - z_d\right) \tag{4.5}$$

Proof: (A1)-(A4) and (A7) imply that $J_{LS}(\lambda) = \|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathscr{H}}$ is of C^k-class with respect to the Λ-topology. Due to (A9), $J_{LS}(\lambda)$ will also be of C^k-class with respect to the $\mathscr{R}$-topology. Hence $J_\beta(\lambda)$ is also of C^k-class.

Existence and uniqueness of the solution of $\psi(A(\lambda),u) = f$ is guaranteed by (A2) and (A6).

Existence and uniqueness of the solution of (4.5) follows from the following facts:

(a) $-2[\mathscr{C}'(u)]^*\Lambda_{\mathscr{H}}(\mathscr{C}(u) - z_d) \in U'$,

since $\begin{cases} (\mathscr{C}(u) - z_d) \in \mathscr{H} \\ \Lambda_{\mathscr{H}}; \mathscr{H} \to \mathscr{H}' \\ [\mathscr{C}'(u)]^*; \mathscr{H}' \to U' \end{cases}$

(b) $\left[\frac{\partial\Psi}{\partial u}\ (A(\lambda),u)\right]^*$ is a linear homeomorphism of F' onto U', as a result of (A3).

---

[2] Given a Banach space X and its dual X', we denote by $(\cdot\,,\cdot)_{XX'}$ the duality between X and X'. Given a Hilbert space *H*, we denote by $(\cdot\,,\cdot)_H$ the inner product in *H*.

Let us now calculate the first derivative of $J_\beta(\lambda)$. For every $\delta\lambda \in \Lambda$ we have

$$J'_\beta(\lambda)\cdot\delta\lambda = 2\left(\mathscr{C}'(u)\cdot\delta u,\ \mathscr{C}(u) - z_d\right)_{\mathscr{H}} + 2\beta\ (\delta\lambda,\lambda)_{\mathscr{R}}$$

$$= 2\left(\mathscr{C}'(u)\cdot\delta u, \Lambda_{\mathscr{H}}(\mathscr{C}(u) - z_d)\right)_{\mathscr{H}\mathscr{H}'} + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}}$$

$$= 2\left(\delta u, [\mathscr{C}'(u)]^{*}\Lambda_{\mathscr{H}}(\mathscr{C}(u) - z_d)\right)_{UU'} + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}}$$

Taking into account (4.5),

$$J'_\beta(\lambda)\cdot\delta\lambda = -\left(\delta u, \left[\frac{\partial\Psi}{\partial u}\ (A,u)\right]^{*}\rho\right)_{UU'} + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}}$$

$$= \left(-\frac{\partial\Psi}{\partial u}\ (A,u)\cdot\delta u, \rho\right)_{FF'} + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}}$$

Finally, from Proposition 2.2,

$$J'_\beta(\lambda)\cdot\delta\lambda = \left(\frac{\partial\Psi}{\partial A}\ (A,u)\circ A'(\lambda)\cdot\delta\lambda, \rho\right)_{FF'} + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}}$$

This completes the proof.

The next theorem establishes the existence of a global minimum of $J_\beta(\lambda)$ on $\mathscr{R}_{ad}$. We first prove the following lemma:

LEMMA 4.1: *Let $\mathcal{R}$ and Λ be Banach spaces and assume that $\mathcal{R}$ is compactly imbedded in Λ. If* $x_n \xrightarrow{\text{weak-top of } \mathcal{R}} x$, *then* $x_n \xrightarrow{\text{norm-top of } \Lambda} x$.

Proof: Take $x_n \xrightarrow{\text{weak-top of } \mathcal{R}} x$.

Then there exists $M > 0$ so that $\|x_n\|_{\mathcal{R}} \leq M \ \forall n \in \mathbb{N}$. In other words, the family of functions $\{x_n\}_{n=1}^{\infty}$ is a subset of the ball $B_M = \{x \in \mathcal{R} / \|x\|_{\mathcal{R}} \leq M\}$. Since the imbedding $\mathcal{R} \to \Lambda$ is compact, it follows that B_M is precompact in the norm-topology of Λ. Hence, there is a subsequence $\{x_{n_k}\}$ that (strongly) converges in Λ. Due to the uniqueness of the limit, it follows that

$$\begin{cases} x_{n_k} \xrightarrow{\text{norm-top. of } \Lambda} x \\ \{x_n\} \text{ cannot have cluster point other than } x \end{cases}$$

So x is the *unique* cluster point of the sequence $\{x_n\}$ with respect to the norm-topology of Λ. And since $\{x_n\}$ is contained in the precompact set B_M, it follows ([1,p.68]) that $x_n \xrightarrow{\text{norm-top. of } \Lambda} x$. This completes the proof.

THEOREM 4.2: *Under assumptions* (A1)-(A10), *the functional*

$$J_\beta(\lambda) = \|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathcal{H}} + \beta\|\lambda\|^2_{\mathcal{R}}$$

admits a global minimum on $\mathcal{R}_{ad}$.

Proof: Let $m = \inf_{\lambda \in \mathcal{R}_{ad}} J_\beta(\lambda)$. Clearly, $m \geq 0$. There is a minimizing sequence $\{\lambda_n\}$ in $\mathcal{R}_{ad}$ such that $\lim_{n\to\infty} J_\beta(\lambda_n) = m$. Clearly, we may assume that

$$\cdots \leq J_\beta(\lambda_{n+1}) \leq J_\beta(\lambda_n) \leq \cdots \leq J_\beta(\lambda_1)$$

Then, for every $n \in \mathbb{N}$

$$\|\lambda_n\|^2_{\mathscr{R}} \leq \frac{1}{\beta} J_\beta(\lambda_n) \leq \frac{1}{\beta} J_\beta(\lambda_1)$$

i.e. $\{\lambda_n\}$ is norm-bounded in $\mathscr{R}$. Hence, there is a subsequence $\{\lambda_{n_k}\}$ that converges in the weak topology of $\mathscr{R}$ to some $\lambda \in \mathscr{R}$. Since $\mathscr{R}_{ad}$ is norm-closed and convex, it is also weakly closed and hence $\lambda \in \mathscr{R}_{ad}$ so

$$\lambda_{n_k} \xrightarrow{\text{weak top. of } \mathscr{R}} \lambda \in \mathscr{R}_{ad}$$

It follows from Lemma 4.1 that

$$\lambda_{n_k} \xrightarrow{\text{norm top. of } \Lambda} \lambda \in \mathscr{R}_{ad}$$

Finally, using the continuity of the functional $J_{LS}(\lambda) = \|\mathscr{C}[\Phi(A(\lambda))] - z_d\|^2_{\mathscr{H}}$ in the norm - topology of Λ, as well as the weak lower semicontinuity of $J_S(\lambda) = \|\lambda\|^2_{\mathscr{R}}$ in $\mathscr{R}$, we conclude

$$m = \lim_{k\to\infty} J_\beta(\lambda_{n_k}) = \lim_{k\to\infty} \|\mathscr{C}[\Phi(A(\lambda_{n_k}))] - z_d\|^2_{\mathscr{H}} + \beta \lim_{k\to\infty} \|\lambda_{n_k}\|^2_{\mathscr{R}}$$

$$\geq \|\mathscr{C}[\Phi(A(\lim_{k\to\infty} \lambda_{n_k}))] - z_d\|^2_{\mathscr{H}} + \beta \|\lim_{k\to\infty} \lambda_{n_k}\|^2_{\mathscr{R}}$$

$$= J_\beta(\lambda)$$

Hence $J_\beta(\lambda) = m$. This completes the proof of Theorem 4.2.

Remark: We can say nothing about uniqueness of the minimum, since the functional $J_\beta(\lambda)$ will in general be non-convex.

Now we can give a necessary condition for optimality (following [20, p. 9]).

PROPOSITION 4.1: *A necessary condition for* $\lambda \in \mathscr{R}_{ad}$ *to be global minimum of* $J_\beta(\lambda)$ *on the set* $\mathscr{R}_{ad}$ *is*

$$J'_\beta(\lambda)\cdot(\nu-\lambda) \geq 0 \qquad \forall\, \nu \in \mathscr{R}_{ad}$$

Proof: Let λ be a global minimum of $J_\beta(\lambda)$ on $\mathscr{R}_{ad}$, i.e. $J_\beta(\lambda) \leq J_\beta(\mu)$ $\forall \mu \in \mathscr{R}_{ad}$. Since $\mathscr{R}_{ad}$ is convex, for any $\nu \in \mathscr{R}_{ad}$ and $\theta \in [0,1]$ we have $((1-\theta)\lambda + \theta\nu) \in \mathscr{R}_{ad}$. Thus $J_\beta(\lambda) \leq J_\beta((1-\theta)\lambda + \theta\nu)$ $\forall \nu \in \mathscr{R}_{ad}$. Hence,

$$\frac{1}{\theta}\,[J_\beta(\lambda+\theta(\nu-\lambda)) - J_\beta(\lambda)] \geq 0$$

Since $J_\beta(\lambda)$ is differentiable, we can pass to the limit and obtain

$$\lim_{\theta\to 0} \frac{1}{\theta}\,[J_\beta(\lambda+\theta(\nu-\lambda)) - J_\beta(\lambda)] \geq 0$$

i.e.
$$J'_\beta(\lambda)\cdot(\nu-\lambda) \geq 0$$

This completes the proof.

So far we have established existence of a minimum of the smoothing functional on $\mathscr{R}_{ad}$ and have given a necessary condition for optimality. Now we will show that minima of J_β depend continuously on the observation. This will be the key result of the regularization approach. Roughly speaking, what the next theorem says is the following:

Let $\tilde{\lambda}$ be the "true" value of the parameter and $\tilde{z}_d := \mathscr{C}[\Phi(A(\tilde{\lambda}))]$, what we would have observed with a zero-error observation. Provided that

(i) $\tilde{\lambda}$ is the unique preimage of $\tilde{z}_d$

(ii) β is an appropriately chosen function of the observation error

any minimum of $J_\beta(\lambda)$ converges (in the norm of Λ) to $\tilde{\lambda}$, as the observation error tends (in the norm of $\mathscr{H}$) to zero.

Note that our theorem is a local version of the Tikhonov-Arsenin convergence theorem [29, p.65] in the sense that:

(a) We relax their global identifiability assumption (i.e. in the sense of Definition 3.1).

(b) We refer to a specific pair $(\tilde{\lambda}, \tilde{z}_d)$ for which it is assumed that $\tilde{\lambda}$ is the unique preimage of $\tilde{z}_d$ in Λ_{ad} (pointwise identifiability assumption).

The need of such a generalization has been motivated by the fact that identifiability results are as a rule conditional identifiability results (see §3) Note that the pointwise identifiability assumption (b) is the weakest possible assumption to ensure that the estimated parameter is "close enough" to the true parameter. (If $\tilde{z}_d$ has e.g. two preimages $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$, no mathematical method can "predict" which one is the true λ).

We first prove the following lemma:

LEMMA 4.2: *Let* $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$ *be metric spaces,* $f: X \to Y$ *a continuous mapping,* K *a precompact subset of* X. *Furthermore, we are given* $y^\circ \in f(K)$ *to which there corresponds a unique* $x^\circ \in X$ *with* $y^\circ = f(x^\circ)$. Then $\forall \varepsilon > 0$ $\exists \gamma(\varepsilon) > 0$ *such that* $\forall x \in K \quad d_Y(f(x), y^\circ) \leq \gamma \Rightarrow d_X(x, x^\circ) \leq \varepsilon$.

Proof: It suffices to prove that for every sequence $\{x_n\}$ in K such that $f(x_n) \to y°$ we have $x_n \to x°$. Since K is precompact, $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ that converges to some $\tilde{x} \in X$. Since f is continuous, $f(x_{n_k}) \to f(\tilde{x})$. But $f(x_n) \to y°$. Hence, $f(\tilde{x}) = y°$. And since x° is the unique preimage of y°, $\tilde{x} = x°$. So, $x_{n_k} \to x°$.

The same argument shows that $\{x_n\}$ cannot have any cluster point other than x°. Thus x° is the unique cluster point of $\{x_n\}$, which is contained in the precompact set K. Hence ([1, p.68]) $x_n \to x°$. This completes the proof.

THEOREM 4.3: *For any* $\beta > 0$ *and* $z_d \in \mathscr{H}$, *denote by* $\lambda_\beta \in \mathscr{R}_{ad}$ *any minimum of* $J_\beta(\lambda)$ *on* $\mathscr{R}_{ad}$. *Also, denote by* T_{δ_1} *the class of functions that are nonnegative, nondecreasing and continuous on the interval* $[0,\delta_1]$. *Suppose*

$$\begin{cases} \tilde{z}_d \in \mathscr{H} \\ \exists \text{ a unique } \tilde{\lambda} \in \mathscr{R}_{ad} \text{ with } \tilde{z}_d = \mathscr{C}[\Phi(A(\tilde{\lambda}))] \end{cases}$$

Then $\forall \varepsilon > 0 \quad \forall B_1, B_2 \in T_{\delta_1}$ *with*

$$\begin{Bmatrix} B_2(0) = 0 \\ \dfrac{\delta^2}{B_1(\delta)} \leq B_2(\delta) \end{Bmatrix}$$

$\exists \delta_0(\varepsilon, B_1, B_2) \leq \delta_1$ *such that* $\forall z_d \in \mathscr{H} \quad \forall \delta \leq \delta_0$

$$\| z_d - \tilde{z}_d \|_{\mathscr{H}} \leq \delta \implies \| \lambda_\beta - \tilde{\lambda} \|_\Lambda \leq \varepsilon$$

for all β *satisfying* $\dfrac{\delta^2}{B_1(\delta)} \leq \beta \leq B_2(\delta)$.

Proof: We have

$$\beta\|\lambda_\beta\|^2_{\mathscr{R}} \le \|\mathscr{C}\big(\Phi(A(\lambda_\beta))\big) - z_d\|^2_{\mathscr{H}} + \beta\|\lambda_\beta\|^2_{\mathscr{R}}$$

$$\le \|\mathscr{C}\big(\Phi(A(\tilde{\lambda}))\big) - z_d\|^2_{\mathscr{H}} + \beta\|\tilde{\lambda}\|^2_{\mathscr{R}}$$

$$= \|\tilde{z}_d - z_d\|^2_{\mathscr{H}} + \beta\|\tilde{\lambda}\|^2_{\mathscr{R}}$$

$$\le \delta^2 + \beta\|\tilde{\lambda}\|^2_{\mathscr{R}}$$

$$= \beta\left[\frac{\delta^2}{\beta} + \|\tilde{\lambda}\|^2_{\mathscr{R}}\right]$$

$$\le \beta[B_1(\delta) + \|\tilde{\lambda}\|^2_{\mathscr{R}}]$$

$$\le \beta[B_1(\delta_1) + \|\tilde{\lambda}\|^2_{\mathscr{R}}]$$

Denote $H_o = [B_1(\delta_1) + \|\tilde{\lambda}\|^2_{\mathscr{R}}]^{\frac{1}{2}}$. Clearly, $\|\lambda_\beta\|_{\mathscr{R}} \le H_o$ and $\|\tilde{\lambda}\|_{\mathscr{R}} \le H_o$. Thus we have shown that the elements $\tilde{\lambda}$ and λ_β belong to the set

$$\Lambda_{H_o} = \{\lambda \in \mathscr{R}_{ad} / \|\lambda\|_{\mathscr{R}} \le H_o\}$$

which is precompact in the norm-topology of Λ. It follows from Lemma 4.2. that

$\forall \varepsilon > 0 \;\; \exists \gamma(\varepsilon) > 0$ such that $\forall \hat{\lambda} \in \Lambda_{H_o}$

$$\|\mathscr{C}\big(\Phi(A(\hat{\lambda}))\big) - \tilde{z}_d\|_{\mathscr{H}} \le \gamma \Rightarrow \|\hat{\lambda} - \tilde{\lambda}\|_\Lambda \le \varepsilon$$

Now observe that

$$\|\mathscr{C}(\Phi(A(\lambda_\beta))) - z_d\|_{\mathscr{H}}^2 \le \|\mathscr{C}(\Phi(A(\lambda_\beta))) - z_d\|_{\mathscr{H}}^2 + \beta\|\lambda_\beta\|_{\mathscr{R}}^2$$

$$\le \|\mathscr{C}(\Phi(A(\tilde{\lambda}))) - z_d\|_{\mathscr{H}}^2 + \beta\|\tilde{\lambda}\|_{\mathscr{R}}^2$$

$$= \|\tilde{z}_d - z_d\|_{\mathscr{H}}^2 + \beta\|\tilde{\lambda}\|_{\mathscr{R}}^2$$

$$\le \delta^2 + B_2(\delta)\|\tilde{\lambda}\|_{\mathscr{R}}^2$$

It follows that

$$\|\mathscr{C}(\Phi(A(\lambda_\beta))) - \tilde{z}_d\|_{\mathscr{H}} \le \|\mathscr{C}(\Phi(A(\lambda_\beta))) - z_d\|_{\mathscr{H}} + \|z_d - \tilde{z}_d\|_{\mathscr{H}}$$

$$\le \left(\delta^2 + B_2(\delta)\|\tilde{\lambda}\|_{\mathscr{R}}^2\right)^{\frac{1}{2}} + \delta$$

The function $\psi(\delta) = (\delta^2 + B_2(\delta)\|\tilde{\lambda}\|_{\mathscr{R}}^2)^{\frac{1}{2}} + \delta$ is a continuous monotonically increasing function satisfying $\psi(0) = 0$. Hence, one can choose $\delta_o = \psi^{-1}(\gamma(\varepsilon))$ and have $\|\mathscr{C}(\Phi(A(\lambda_\beta))) - \tilde{z}_d\|_{\mathscr{H}} \le \gamma(\varepsilon) \quad \forall \delta \le \delta_o$. Thus we see that for all β satisfying $\frac{\delta^2}{B_1(\delta)} \le \beta \le B_2(\delta)$, the inequality $\|z_d - \tilde{z}_d\|_{\mathscr{H}} \le \delta$ implies the inequality $\|\lambda_\beta - \tilde{\lambda}\|_\Lambda \le \varepsilon$. This completes the proof.

5. SELECTION OF THE REGULARIZATION PARAMETER

In §4 we established that the regularization approach provides a stable method for distributed system identification. One question was not addressed, the selection of the regularization parameter β. In this section we will discuss two methods for selection of β.

Let $\tilde{\lambda} \in \mathscr{R}_{ad}$ be the "true" value of the parameter λ and $\tilde{z}_d$ be the error-free observation, $\tilde{z}_d = \mathscr{C}\big(\Phi(A(\tilde{\lambda}))\big)$. We assume that (i) $\tilde{\lambda}$ is the unique preimage of $\tilde{z}_d$; (ii) An upper bound in the observation error is known, i.e. $\|z_d - \tilde{z}_d\|_{\mathscr{H}} \leq \delta$.

Method 1. When an a priori upper bound on $\|\tilde{\lambda}\|_{\mathscr{R}}$ is known, i.e. $\|\tilde{\lambda}\|_{\mathscr{R}} \leq \Delta$, Miller [22] suggests $\beta(\delta) = (\delta/\Delta)^2$. (When $\mathscr{R}$ is a Sobolev space, $\|\ \|_{\mathscr{R}}$ is a measure of smoothness.) We note first that this choice of β satisfies the assumptions of Theorem 4.3. Furthermore, if $\lambda_{\beta(\delta)}$ is a minimizer of

$$J_\beta(\lambda) = \|\mathscr{C}\big(\Phi(A(\lambda))\big) - z_d\|^2_{\mathscr{H}} + \left(\frac{\delta}{\Delta}\right)^2 \|\lambda\|^2_{\mathscr{R}} \tag{5.1}$$

on $\mathscr{R}_{ad}$, then

$$\begin{aligned}
\|\mathscr{C}\big(\Phi(A(\lambda_{\beta(\delta)}))\big) - z_d\|^2_{\mathscr{H}} + \left(\frac{\delta}{\Delta}\right)^2 \|\lambda_{\beta(\delta)}\|^2_{\mathscr{R}} &= \\
&= J_\beta(\lambda_{\beta(\delta)}) \\
&\leq J_\beta(\tilde{\lambda}) \\
&= \|\mathscr{C}\big(\Phi(A(\tilde{\lambda}))\big) - z_d\|^2_{\mathscr{H}} + \left(\frac{\delta}{\Delta}\right)^2 \|\tilde{\lambda}\|^2_{\mathscr{R}} \\
&\leq 2\delta^2
\end{aligned}$$

Hence,

$$\left.\begin{aligned} &\|\mathscr{C}\Big(\Phi(A(\lambda_{\beta(\delta)}))\Big) - z_d\|_{\mathscr{H}} \leq \sqrt{2}\,\delta \\ &\|\lambda_{\beta(\delta)}\|_{\mathscr{R}} \leq \sqrt{2}\,\Delta \end{aligned}\right\} \tag{5.2}$$

i.e. regularized solutions satisfy the constraints up to a factor of $\sqrt{2}$.

Method 2: This method has been suggested by Tikhonov and Arsenin [29]. Their suggestion is to choose $\beta(\delta)$ so that

$$\|\mathscr{C}\Big(\Phi(A(\lambda_{\beta(\delta)}))\Big) - z_d\|_{\mathscr{H}} = \delta$$

where $\lambda_{\beta(\delta)}$ minimizes

$$J_\beta(\lambda) = \|\mathscr{C}\big(\Phi(A(\lambda))\big) - z_d\|^2_{\mathscr{H}} + \beta(\delta)\|\lambda\|^2_{\mathscr{R}}$$

Before one discusses the stability of the method, one has to examine the existence of such a β.

To this end, we follow a different approach than Tikhonov and Arsenin, who give a simple sufficient condition for existence. We give here a much weaker condition which is both necessary and sufficient. In the development of this condition we have used concepts and results from the theory of minimization of vector-valued functionals.

DEFINITION 5.1 (*Ordering Relations in* $\mathbb{R}^n$): Let $x = (x_1,\ldots, x_n)$ and $y = (y_1,\ldots, y_n)$ be two arbitrary elements of $\mathbb{R}^n$. We will write

(i) $x \leq y$ iff $x_i \leq y_i$ for all i

(ii) $x \mathrel{\overline{<}} y$ iff $x \leq y$ and $x_i < y_i$ for at least one i

(iii) $x < y$ iff $x_i < y_i$ for all i.

DEFINITION 5.2:

Let $T; \Lambda \to \mathbb{R}^n$. We will say that $\hat{\lambda}$ is a *Pareto-minimal point* of the vector-valued functional T if $\nexists\, \lambda \in \Lambda$ with $T(\lambda) \mathrel{\bar{<}} T(\hat{\lambda})$. The set $\{T(\hat{\lambda})/\hat{\lambda}$ is a Pareto-minimal point of $T\}$ is called the *Pareto-minimal set* of T.

PROPOSITION 5.1 ([30, p.94]):

Let $T(\lambda) \equiv (T_1(\lambda),\ldots, T_n(\lambda))$ *be a vector-valued functional on* Λ. *An element* $\hat{\lambda} \in \Lambda$ *is a Pareto-minimal point iff for every* $j \in \{1,\ldots, n\}$ $\hat{\lambda}$ *minimizes* $T_j(\lambda)$ *on the set*

$$\Lambda_j = \{\lambda \in \Lambda/T_i(\lambda) \leq T_i(\hat{\lambda}) \quad \forall i \in \{1,\ldots, n\} \text{ with } i \neq j\}$$

LEMMA 5.1: *Let*

$$\lambda_{min} = \text{the minimum-norm element of } \mathscr{R}_{ad}$$
$$\delta_{max} = J_{LS}(\lambda_{min}) = \|\mathscr{C}\left(\Phi(A(\lambda_{min}))\right) - z_d\|^2_{\mathscr{H}}$$
$$\delta_{min} = \inf_{\lambda \in \mathscr{R}_{ad}} J_{LS}(\lambda) = \inf_{\lambda \in \mathscr{R}_{ad}} \|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathscr{H}}$$

Given $\delta > \delta_{min}$, *there exists an element* λ_δ *minimizing the functional* $J_S(\lambda)$ *on the set* $\{\lambda \in \mathscr{R}_{ad}/\|\mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathscr{H}} \leq \delta\}$. *Furthermore, if* $\delta \leq \delta_{max}$, *then* $\|\mathscr{C}\left(\Phi(A(\lambda_\delta))\right) - z_d\|^2_{\mathscr{H}} = \delta$.

Remark: Lemma 5.1 holds for $\delta = \delta_{min}$ if $J_{LS}(\lambda)$ admits a minimum on $\mathscr{R}_{ad}$.

THEOREM 5.1: *Let* λ_δ, δ_{max} *as in Lemma 5.1. The function*

$$\Theta(\delta) = J_S(\lambda_\delta) \quad , \qquad \delta \leq \delta_{max}$$

is monotonically decreasing. Its graph coincides with the Pareto-minimal set of the vector-valued functional $T; \mathscr{R}_{ad} \to \mathbb{R}^2$ *defined by* $T(\lambda) = (J_{LS}(\lambda), J_S(\lambda))$.

Proof of Lemma 5.1: The proof of existence of a minimum is almost the same as that of Theorem 4.2.

Consider a minimizing sequence $\{\lambda_n\}$. This will have a subsequence $\{\lambda_{n_k}\}$ that converges in the weak topology of $\mathscr{R}$ to some $\lambda \in \mathscr{R}$. We conclude that

$$\begin{cases} \lambda \in \mathscr{R}_{ad} \\ \lambda_{n_k} \xrightarrow[\text{of } \Lambda]{\text{strong topology}} \lambda \end{cases}$$

Also, due to the continuity of $\mathscr{C} \circ \Phi \circ A$, it is easy to see that the limit has to satisfy

$$\|\mathscr{C}\{\Phi(A(\lambda))\} - z_d\|^2_{\mathscr{H}} \leq \delta$$

Finally, using the weak lower semicontinuity of $J_S(\lambda)$ in $\mathscr{R}$, we conclude that λ minimizes $J_S(\lambda)$ on the set $\{\lambda \in \mathscr{R}_{ad} / \|\mathscr{C}\{\Phi(A(\lambda))\} - z_d\|^2_{\mathscr{H}} \leq \delta\}$.

To prove the second part of the lemma, suppose

$$\|\mathscr{C}\{\Phi(A(\lambda_\delta))\} - z_d\|^2_{\mathscr{H}} = \delta^* < \delta$$

Since $\mathscr{C} \circ \Phi \circ A$ is continuous, there is a ball $B(\lambda_\delta)$, centered at λ_δ, such that

$$\|\mathscr{C}\{\Phi(A(\lambda))\} - \mathscr{C}\{\Phi(A(\lambda_\delta))\}\|^2_{\mathscr{H}} < \frac{\delta - \delta^*}{2} \qquad \forall \lambda \in B(\lambda_\delta) \cap \mathscr{R}_{ad}$$

Now observe that

(i) We can always have $\lambda_{min} \notin B(\lambda_\delta)$, since $\delta^* < \delta_{max}$ implies $\lambda_\delta \neq \lambda_{min}$

(ii) $B(\lambda_\delta) \cap \mathscr{R}_{ad} \subset \{\lambda \in \mathscr{R}_{ad} / \|\mathscr{C}\{\Phi(A(\lambda))\} - z_d\|^2_{\mathscr{H}} \leq \delta\}$

since

$$\|\mathscr{C}[\Phi(A(\lambda))] - z_d\|_{\mathscr{H}} \leq \|\mathscr{C}[\Phi(A(\lambda))] - \mathscr{C}[\Phi(A(\lambda_\delta))]\|_{\mathscr{H}} + \|\mathscr{C}[\Phi(A(\lambda_\delta))] - z_d\|_{\mathscr{H}}$$

But from (i) and convexity of $\mathscr{R}_{ad}$ it follows that $\exists\, \lambda^* \cap B(\lambda_\delta) \cap \mathscr{R}_{ad}$ so that $\|\lambda^*\|_{\mathscr{R}} < \|\lambda_\delta\|_{\mathscr{R}}$.

This contradicts with the definition of λ_δ and (ii).

Proof of Theorem 5.1: If λ_δ minimizes $J_S(\lambda)$ subject to the constraint $J_{LS}(\lambda) \leq \delta < \delta_{max}$, then from Lemma 5.1,

$$T(\lambda_\delta) = (\delta, \Theta(\delta))$$

It is clear that $\delta_1 \leq \delta_2 \leq \delta_{max}$ implies $\Theta(\delta_2) \leq \Theta(\delta_1)$ i.e. Θ is monotonically decreasing. Furthermore, Proposition 5.1 implies that the Pareto-minimal set of T is a subset of the graph of Θ.

Finally, if $T(\hat{\lambda}) \lneqq T(\lambda_\delta)$ for some $\hat{\lambda} \in \mathscr{R}_{ad}$ and $\delta \leq \delta_{max}$, this would mean

$$\text{either} \quad \begin{cases} J_{LS}(\hat{\lambda}) < \delta \\ J_S(\hat{\lambda}) \leq J_S(\lambda_\delta) \end{cases}$$

$$\text{or} \quad \begin{cases} J_{LS}(\hat{\lambda}) \leq \delta \\ J_S(\hat{\lambda}) < J_S(\lambda_\delta) \end{cases}$$

Both cases are impossible since they contradict the definition of λ_δ and/or Lemma 5.1. Hence, $T(\lambda_\delta) \equiv (\delta, \Theta(\delta))$ is Pareto-minimal.

This completes the proof of the theorem.

Before we proceed to the main result of this section, we will state an important proposition by Yu [33], which will be needed in the proof. We first give the definition of cone convexity, introduced in the same paper.

DEFINITION 5.3: Let $S \subset \mathbb{R}^n$ and C a convex cone in $\mathbb{R}^n$. S will be called *C-convex* if S + C is convex.

PROPOSITION 5.2. ([33, p.28])

Let $T: \Lambda \to \mathbb{R}^n$ *and suppose that* Ran T *is* $\mathbb{R}^n_+$*-convex, where* $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n / x \geq 0\}$. *If* $\hat{\lambda}$ *is a Pareto-minimal point, there exists* $\eta \gneq 0$ *such that*

$$\eta^T T(\hat{\lambda}) \leq \eta^T T(\lambda) \qquad \forall \lambda \in \Lambda$$

THEOREM 5.2.

Let Θ *and* T *be as in Theorem 5.1. Then the following assertions are equivalent:*

(i) *For every* $\delta \in]\delta_{min}, \delta_{max}[$ *there is* $\beta > 0$ *and a minimizer* λ_β *of the functional*

$$J_\beta(\lambda) = J_{LS}(\lambda) + \beta J_S(\lambda)$$

on $\mathcal{R}_{ad}$, *such that*

$$J_{LS}(\lambda_\beta) = \delta$$

(ii) Ran T *is* $\mathbb{R}^n_+$*-convex.*

(iii) Θ *is convex.*

We first prove the following lemma:

LEMMA 5.2:

$$\text{Let} \quad \tilde{\Theta}(\delta) = \begin{cases} \Theta(\delta), \ \text{for } \delta \leq \delta_{max} \\ J_S(\lambda_{min}), \ \text{for } \delta \geq \delta_{max} \end{cases}$$

Then[4] $\text{Ran } T + \mathbb{R}^n_+ = \text{Epi } \tilde{\Theta}$.

Proof of Lemma 5.2: Take $(x,y) \in \text{Epi } \tilde{\Theta}$. If $x \geq \delta_{max}$, then $(x,y) \in \{T(\lambda_{min})\} + \mathbb{R}^n_+$. If $x \leq \delta_{max}$, then $(x,y) \in \{T(\lambda_x)\} + \mathbb{R}^n_+$, where λ_x minimizes $J_S(\lambda)$ on $\mathscr{R}_{ad}$ subject to $J_{LS}(\lambda) \leq x$. So: $\text{Ran } T + \mathbb{R}^n_+ \supseteq \text{Epi } \tilde{\Theta}$.

Now take $z \in \text{Ran } T + \mathbb{R}^n_+$. This means $\exists \lambda \in \mathscr{R}_{ad}$ $\exists \nu \in \mathbb{R}^n_+$ with $z = T(\lambda) + \nu$. Assume that $z \notin \text{Epi } \tilde{\Theta}$, hence $\exists \delta \geq 0$ with $z \lneqq (\delta, \Theta(\delta))$.

Case 1: $\delta \leq \delta_{max}$

Then $T(\lambda) + \nu \lneqq T(\lambda_\delta)$, where λ_δ minimizes $J_S(\lambda)$ on $\mathscr{R}_{ad}$ subject to $J_{LS}(\lambda) \leq \delta$. This implies that λ_δ is not Pareto-minimal.

Case 2: $\delta \geq \delta_{max}$

Then $T(\lambda) + \nu \lneqq T(\lambda_{min}) \Rightarrow \lambda_{min}$ is not Pareto-minimal.

Thus, we see that in all cases the assumption $z \notin \text{Epi } \tilde{\Theta}$ leads to contradiction. Hence, $\text{Ran } T + \mathbb{R}^n_+ \subseteq \text{Epi } \tilde{\Theta}$. This completes the proof.

---

[4] By the symbol Epi F we mean the epigraph of a function F, i.e. the set $\{(x,y) \in \mathbb{R}^2 / y \geq F(x)\}$.

Proof of Theorem 5.2.

(i) ⇒ (ii)

Given $\beta > 0$ denote by λ_β a minimizer of $J_\beta(\lambda) = J_{LS}(\lambda) + \beta J_S(\lambda)$ on $\mathscr{R}_{ad}$ and define

$$P_\beta = \{(x,y) \in \mathbb{R}^2 / x + \beta y \geq J_{LS}(\lambda_\beta) + \beta J_S(\lambda_\beta)\}$$

$$P_\infty = \{(x,y) \in \mathbb{R}^2 / y \geq J_S(\lambda_{min})\}$$

Furthermore, define

$$P_o = \begin{cases} \{(x,y) \in \mathbb{R}^2 / x > 0\} \cup \{(x,y) \in \mathbb{R}^2 / x = 0,\ y \geq J_S(\bar\lambda)\}, \text{ if } \bar\lambda \text{ minimizes } J_{LS}(\lambda) \text{ on } \mathscr{R}_{ad} \\ \{(x,y) \in \mathbb{R}^2 / x > 0\}, \text{ if } J_{LS}(\lambda) \text{ does not admit a minimum of } \mathscr{R}_{ad} \end{cases}$$

We will show that $\bigcap_{0 \leq \beta \leq \infty} P_\beta = \operatorname{Ran} T + \mathbb{R}^n_+$. Clearly $\forall \lambda \in \mathscr{R}_{ad}$ $\forall \beta > 0$

$T(\lambda) \in P_\beta$. Hence $\forall \lambda \in \mathscr{R}_{ad}$ $\forall z \in \mathbb{R}^n_+$ $\forall \beta > 0$ $T(\lambda) + z \in P_\beta$.

Thus $\operatorname{Ran} T + \mathbb{R}^n_+ \subseteq P_\beta$ $\forall \beta > 0$. It is also trivial to see that the above relation holds for $\beta = 0$ and $\beta = \infty$. Hence $\operatorname{Ran} T + \mathbb{R}^n_+ \subseteq \bigcap_{0 \leq \beta \leq \infty} P_\beta$.

To show that $\bigcap_{0<\beta<\infty} P_\beta \subseteq \operatorname{Ran} T + \mathbb{R}^n_+$ we will take $(\hat{x},\hat{y}) \notin \operatorname{Ran} T + \mathbb{R}^n_+$ and show that $\exists \beta$ such that $(\hat{x},\hat{y}) \notin P_\beta$. We only need to consider the case $\hat{x} \in\ ]\delta_{min}, \delta_{max}[$, since $\hat{x} = \delta_{min}$ clearly implies $(\hat{x},\hat{y}) \notin P_o$ and $\hat{x} \geq \delta_{max}$ implies $(\hat{x},\hat{y}) \notin P_\infty$. By (i), $\exists \hat\beta > 0$ $\exists$ minimizer $\lambda_{\hat\beta}$ of the functional $J_{\hat\beta}(\lambda) = J_{LS}(\lambda) + \hat\beta J_s(\lambda)$ on $\mathscr{R}_{ad}$ satisfying $J_{LS}(\lambda_{\hat\beta}) = \hat{x}$. Observe that $J_S(\lambda_{\hat\beta}) > \hat{y}$, since otherwise $T(\lambda_{\hat\beta}) \leq (\hat{x},\hat{y})$ which would imply that $(\hat{x},\hat{y}) \in \operatorname{Ran} T + \mathbb{R}^n_+$. But

$$\left.\begin{array}{l} J_{LS}(\lambda_{\hat\beta}) = \hat{x} \\ \\ J_S(\lambda_{\hat\beta}) > \hat{y} \end{array}\right\} \Rightarrow J_{LS}(\lambda_{\hat\beta}) + \hat\beta J_S(\lambda_{\hat\beta}) > \hat{x} + \hat\beta\hat{y} \Rightarrow (\hat{x},\hat{y}) \notin P_{\hat\beta}$$

So:

$$\bigcap_{0 \leq \beta \leq \infty} P_\beta = \operatorname{Ran} T + \mathbb{R}^n_+$$

Taking into account the convexity of the sets P_β we conclude that $\operatorname{Ran} T + \mathbb{R}^n_+$ is convex.

(ii) $\Rightarrow$ (i)

Consider an arbitrary $\delta \in]\delta_{min}, \delta_{max}[$ and denote by λ_δ a minimizer of $J_S(\lambda)$ on $\mathscr{R}_{ad}$ subject to the constraint $J_{LS}(\lambda) \leq \delta$.

Since Ran T is $\mathbb{R}^n_+$ - convex, by Proposition 5.2, there is $\eta \equiv (\eta_1, \eta_2) \overset{>}{=} 0$ such that

$$\eta_1 J_{LS}(\lambda_\delta) + \eta_2 J_S(\lambda_\delta) \leq \eta_1 J_{LS}(\lambda) + \eta_2 J_S(\lambda) \qquad \forall \lambda \in \mathscr{R}_{ad}$$

Since $\delta > \delta_{min}$, λ_δ does not minimize J_{LS} on $\mathscr{R}_{ad}$. Hence $\eta_2 \neq 0$.

Since $\delta < \delta_{max}$, we have $\lambda_\delta \neq \lambda_{min}$. Hence $\eta_1 \neq 0$.

So we may choose $\beta = \frac{\eta_2}{\eta_1} > 0$ and have

$$J_{LS}(\lambda_\delta) + \beta J_S(\lambda_\delta) \leq J_{LS}(\lambda) + \beta J_S(\lambda) \qquad \forall \lambda \in \mathscr{R}_{ad}$$

But by construction of λ_δ and Lemma 5.1 it follows that $J_{LS}(\lambda_\delta) = \delta$.

(ii) $\Leftrightarrow$ (iii)

$$\text{Ran } T \text{ is } \mathbb{R}^n_+ \text{ - convex} \quad \overset{\text{Lemma 5.2}}{\Longleftrightarrow} \quad \text{Epi } \tilde{\Theta} \text{ is convex} \Leftrightarrow \tilde{\Theta} \text{ is convex}$$

And since Θ is decreasing, the latter is equivalent to Θ convex.

This completes the proof of the theorem.

What remains to show is that regularized solutions obtained by this method converge (in the norm of Λ) to $\tilde{\lambda}$ as the observation error tends (in the norm of $\mathscr{H}$) to zero. This will be done independently of the theory of Section 4. Note, however, that the argument is almost the same as in Theorem 4.3.

THEOREM 5.3

Suppose $\begin{cases} \tilde{z}_d \in \mathscr{H} \\ \exists \text{ a unique } \tilde{\lambda} \in \mathscr{R}_{ad} \text{ with } \tilde{z}_d = \mathscr{C}(\Phi(A(\tilde{\lambda}))) \\ \text{The function } \Theta \text{ defined in Theorem 5.1 is convex} \end{cases}$

Then $\forall \varepsilon > 0 \quad \exists \delta_0(\varepsilon) > 0$ *such that* $\forall z_d \in \mathscr{H} \quad \forall \delta \leq \delta_0$

$$\| z_d - \tilde{z}_d \|_{\mathscr{H}} \leq \delta \Rightarrow \| \lambda_{\beta(\delta)} - \tilde{\lambda} \|_{\Lambda} \leq \varepsilon$$

where $\begin{cases} \beta(\delta) \text{ denotes a regularization parameter} \\ \lambda_\beta(\delta) \text{ denotes a minimizer of } J_{\beta(\delta)}(\lambda) \text{ on } \mathscr{R}_{ad} \end{cases}$

satisfying $\| \mathscr{C}(\Phi(A(\lambda_{\beta(\delta)}))) - z_d \|_{\mathscr{H}} = \delta$

Proof: Denote $\hat{\Lambda} = \{\lambda \in \mathscr{R}_{ad} / \| \lambda \|_{\mathscr{R}} \leq \| \tilde{\lambda} \|_{\mathscr{R}} \}$ which is precompact in the norm-topology of Λ. It follows from Lemma 4.2 that

$$\forall \varepsilon > 0 \quad \exists \gamma(\varepsilon) > 0 \text{ such that } \forall \hat{\lambda} \in \hat{\Lambda}$$

$$\| \mathscr{C}(\Phi(A(\hat{\lambda}))) - \tilde{z}_d \|_{\mathscr{H}} \leq \gamma \Rightarrow \| \hat{\lambda} - \tilde{\lambda} \|_{\Lambda} \leq \varepsilon$$

Clearly, $\lambda_{\beta(\delta)}$ minimizes $J_S(\lambda)$ on $\mathscr{R}_{ad}$ subject to the constraint $\| \mathscr{C}(\Phi(A(\lambda))) - z_d \|_{\mathscr{H}} \leq \delta$. Since

$$\| \mathscr{C}(\Phi(A(\tilde{\lambda}))) - z_d \|_{\mathscr{H}} = \| \tilde{z}_d - z_d \|_{\mathscr{H}} \leq \delta$$

it is obvious that $J_S(\lambda_{\beta(\delta)}) \leq J_S(\tilde{\lambda})$ i.e. $\lambda_{\beta(\delta)} \in \hat{\Lambda}$. Also, observe that

$$\| \mathscr{C}(\Phi(A(\lambda_{\beta(\delta)}))) - \tilde{z}_d \|_{\mathscr{H}} \leq \| \mathscr{C}(\Phi(A(\lambda_{\beta(\delta)}))) - z_d \|_{\mathscr{H}} + \| z_d - \tilde{z}_d \|_{\mathscr{H}} \leq 2\delta$$

Thus we can choose $\delta_0 = \frac{\gamma(\varepsilon)}{2}$ and have

$$\| \mathscr{C}(\Phi(A(\lambda_{\beta(\delta)}))) - \tilde{z}_d \|_{\mathscr{H}} \leq \gamma(\varepsilon) \qquad \forall \delta \leq \delta_0$$

Thus we see that for all $\delta \leq \delta_o = \frac{\gamma(\varepsilon)}{2}$, the inequality $\|z_d - \tilde{z}_d\|_{\mathcal{H}} \leq \delta$ implies the inequality $\|\lambda_{\beta(\delta)} - \tilde{\lambda}\|_{\Lambda} \leq \varepsilon$.

This completes the proof.

6. IDENTIFICATION OF LINEAR PARABOLIC SYSTEMS FROM DISTRIBUTED OBSERVATION

Let V and H be Hilbert spaces such that

$$\left.\begin{array}{l} V \subset H \\ \text{continuous injection} \\ \text{dense inclusion} \end{array}\right\} \tag{6.1}$$

Identifying H with its dual and denoting by V' the dual of V, we have

$$\left.\begin{array}{l} V \subset H \subset V' \\ \text{continuous injections} \\ \text{dense inclusions} \end{array}\right\} \tag{6.2}$$

Consider the following operator differential equation

$$\left.\begin{array}{l} \frac{du}{dt} + A(t)u = f \\ u(0) = u_0 \end{array}\right\} \tag{6.3}$$

and take

Operator space:

$$\mathscr{A} = L^{\infty}(0,T;\mathscr{L}(V,V')) \tag{6.4}$$

$= \text{space of essentially bounded measurable mappings } [0,T] \to \mathscr{L}(V,V')$

Subspace of coercive operators:

$$\mathscr{A}_c = \Big\{ A \in \mathscr{A} \;/\; \exists\ \zeta \in \mathbf{R}\ \ \exists \eta > 0:\ (A(t)u,u)_{V'V} + \zeta\|u\|_H^2 \geq \eta\|u\|_V^2 \quad \text{a.e. in }]0,T[\Big\} \tag{6.5}$$

Space of right-hand sides:

$$F = L^2(0,T;V') \times H \tag{6.6}$$

Space of solutions:

$$U = W(0,T) = \left\{u/u \in L^2(0,T;V),\ \frac{du}{dt} \in L^2(0,T;V')\right\} \qquad (6.7)$$

PROPOSITION 6.1 ([20, p.102]): *Given* $A(t) \in \mathscr{A}_c$, $f \in L^2(0,T;V')$ *and* $u_o \in H$, *the problem* (6.3) *admits a unique solution* $u \in W(0,T)$. *The solution depends continuously on the data* f *and* u_o.

Now define the mapping

$$\Psi;(A,u) \in \mathscr{A} \times U \to \left(\frac{du}{dt} + Au,\ u(0)\right) \in F \qquad (6.8)$$

Clearly, Ψ is of C^∞-class, hence assumption (A1) is satisfied. Also, we have

$$\frac{\partial\Psi}{\partial u}(A,u)\cdot\delta u = \left(\frac{d(\delta u)}{dt} + A\delta u,\ \delta u(0)\right) \qquad (6.9)$$

$$\frac{\partial\Psi}{\partial A}(A,u)\cdot\delta A = (\delta A u,\ 0) \qquad (6.10)$$

Proposition 6.1 establishes satisfaction of (A2) and (A3).

Next consider that A depends on a parameter λ that belongs to the Banach space Λ. If

$$(\lambda \in \Lambda \to A(\lambda;t) \in \mathscr{A}) \in \mathscr{L}(\Lambda,\mathscr{A}) \qquad (6.11)$$

$$A(\Lambda_{ad}) \subseteq \mathscr{A}_c \qquad (6.12)$$

$$\mathscr{C} \in \mathscr{L}(W(0,T),\mathscr{H}) \qquad (6.13)$$

$$\mathscr{R} \text{ is a Hilbert space imbedded in } \Lambda \qquad (6.14)$$

(A1)-(A4) and (A6)-(A9) are satisfied.

THEOREM 6.1: *The functional*

$$J_\beta(\lambda) = \|\mathscr{C}(\phi(A(\lambda))) - z_d\|^2_{\mathscr{H}} + \beta\|\lambda\|^2_{\mathscr{R}}$$

is of C^∞*-class. Its first derivative is given by*

$$J'_\beta(\lambda)\cdot\delta\lambda = \int_0^T (A(\delta\lambda)u, p)_{V'V} dt + 2\beta(\delta\lambda,\lambda)_{\mathcal{R}} \tag{6.15}$$

where $u \in W(0,T)$ *is the solution of* (6.3) *and* $p \in W(0,T)$ *is the solution of the adjoint equation*

$$\begin{aligned} &-\frac{dp}{dt} + A^*(t)p = -2\mathscr{C}^*\Lambda_{\mathscr{H}}(\mathscr{C}u - z_d) \\ &p(T) = 0 \end{aligned} \tag{6.16}$$

Proof: Apply Theorem 4.1. The adjoint state $\rho \in F'$ is of the form $\rho = (p, \tilde{p})$ with $p \in L^2(0,T;V)$, $\tilde{p} \in L^2(\Omega)$. Taking into account (6.10), we immediately get (6.15) from (4.4). Taking into account (6.9), (4.5) becomes

$$\int_0^T \left(p, \frac{dv}{dt} + Av\right)_{VV'} dt + (\tilde{p}, v(0))_H = -2[\mathscr{C}^*\Lambda_{\mathscr{H}}(\mathscr{C}(u) - z_d)](v)$$

or

$$\int_0^T \left(-\frac{dp}{dt} + A^*p, v\right)_{V'V} dt + (p(T), v(T))_H + (\tilde{p} - p(0), v(0))_H =$$

$$= -2[\mathscr{C}^*\Lambda_{\mathscr{H}}(\mathscr{C}(u) - z_d)](v)$$

from which (6.16) follows.

This completes the proof of the theorem.

As an application of the above theory, let

$$\begin{cases} \Omega & \text{a bounded open subset of } \mathbb{R}^n \\ \Gamma & \text{the boundary of } \Omega\text{, a } C^1\text{-manifold with } \Omega \text{ locally on one side of } \Gamma \\ T & \text{a real number with } 0 < T < \infty \\ Q & = \Omega \times]0,T[\\ \Sigma & = \Gamma \times]0,T[\end{cases}$$

and consider the following linear parabolic system

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} + \sum_{|p|,|q|\le m} (-1)^{|p|} D_x^p(a_{pq}(x,t)D_x^q u) = f_o(x,t), \text{ in } Q \\ &u(x,0) = u_o(x) \qquad , \text{ in } \Omega \\ &B_j u = \sum_{|h|\le m_j} b_{jh}(x,t)D_x^h u = 0 \quad , j = 0,\ldots, m-1 \text{ , on } \Sigma \end{aligned}\right\} \tag{6.17}$$

with $a_{pq} \in L^\infty(Q)$, $f_o \in L^2(Q)$, $u_o \in L^2(\Omega)$, $b_{jh} \in L^\infty(\Sigma)$ and $0 \le m_j \le 2m-1$. The boundary operators are numbered in increasing order and let s be the number of stable boundary conditions, i.e.

$$\left.\begin{aligned} &m_j < m \quad , \text{ for } j = 0,\ldots, s-1 \\ &m_j \ge m \quad , \text{ for } j = s,\ldots, m-1 \end{aligned}\right\} \tag{6.18}$$

If there exist[5] boundary operators $\{B'_j\}_{j=s}^{m-1}$ such that

(i) $\{B_o,\ldots, B_{s-1}, B'_s,\ldots, B'_{m-1}\}$ is a Dirichlet system.

(ii) $\Phi_j = B_j$, $j = s,\ldots, m-1$ where $\{\Phi_j\}_{j=0}^{m-1}$ is the adjoint to $\{B_o,\ldots, B_{s-1}, B'_s,\ldots, B'_{m-1}\}$ with respect to Green's formula

$$\begin{aligned} \int_\Omega \sum_{|p|,|q|\le m} a_{pq}D_x^q u D_x^p v\,dx = &\int_\Omega \sum_{|p|,|q|\le m} (-1)^{|p|}[D_x^p(a_{pq}D_x^q u)]v\,dx \\ &- \int_\Gamma \sum_{j=o}^{s-1} \Phi_j u B_j v\,d\sigma - \int_\Gamma \sum_{j=s}^{m-1} \Phi_j u B'_j v\,d\sigma \end{aligned} \tag{6.19}$$

[5]In general, such boundary operators need not exist ([21] Vol. I, p. 205). However, in all practical applications they do.

then problem (6.17) can be put in a variational formulation as follows:

$$\left.\begin{aligned} &\int_\Omega \frac{\partial u}{\partial t} v\,dx + \int_\Omega \sum_{|p|,|q|\le m} a_{pq} D_x^q u D_x^p v\,dx = \int_\Omega f_o v\,dx \\ &\qquad \forall\, v \text{ such that } B_j v = 0, \quad j = 0,\dots, s-1, \text{ on } \Sigma \\ &u(x,0) = u_o(x), \quad \text{in } \Omega \end{aligned}\right\} \tag{6.20}$$

To see that (6.20) is a special case of (6.3), make the following choices of function spaces

$$V = \{v/v \in H^m(\Omega),\ \ B_j v = 0,\ j = 0,\dots, s-1\} \tag{6.21}$$

which is a Hilbert space with the norm induced by $H^m(\Omega)$

$$H = L^2(\Omega) \tag{6.22}$$

$$\Lambda = \bigtimes_{|p|,|q|\le m} L^\infty(Q) \tag{6.23}$$

which is a Banach space with norm $\|\lambda\|_\Lambda = \max\limits_{|p|,|q|\le m} \|a_{pq}\|_{L^\infty(Q)}$ and define f and $A(\lambda;t)$ by

$$f(v) = \int_\Omega f_o v\,dx \tag{6.24}$$

$$\Big(A(\lambda;t)u,v\Big)_{V'V} = \int_\Omega \sum_{|p|,|p|\le m} a_{pq}(x,t) D_x^q u D_x^p v\,dx \quad \forall\, u,v \in V \tag{6.25}$$

Observe that the mapping $(\lambda \in \Lambda \to A(\lambda;t) \in \mathcal{A}) \in \mathcal{L}(\Lambda,\mathcal{A})$, hence it is of C^∞-class, i.e. (A4) is satisfied. Also, if, given $\zeta \in \mathbb{R}$ and $\eta > 0$, one defines

$$\Lambda_{ad} = \{\lambda \in \Lambda / (A(\lambda;t)v,v)_{V',V} + \zeta \|v\|_H^2 \geq \eta \|v\|_V^2 \quad \forall v \in V \quad \text{a.e. in }]0,T[\} \tag{6.26}$$

then Λ_{ad} is closed and convex and $A(\Lambda_{ad}) \supseteq \mathscr{A}_c$, i.e. (A5) and (A6) are satisfied.

Finally, suppose that we want to identify $\lambda = (a_{pq})$ by observing $u(x,t)$ in Q. We take

$$\mathscr{H} = L^2(Q) \tag{6.27}$$

$$\Lambda_{\mathscr{H}} = \text{identity} \tag{6.28}$$

$$\mathscr{C} = \text{injection of } W(0,T) \text{ into } L^2(Q) \tag{6.29}$$

$$\mathscr{R} = \bigtimes_{|p|,|q| \leq m} H^{\ell_1,\ell_2}(Q) \tag{6.30}$$

with

$$\begin{cases} \ell_1, \ell_2 \text{ such that } H^{\ell_1,\ell_2}(Q) \text{ is compactly imbedded in } L^\infty(Q) \\ \|\lambda\|_{\mathscr{R}} = \left(\sum_{|p|,|q| \leq m} w_{pq} \|a_{pq}\|^2_{H^{\ell_1,\ell_2}(Q)} \right)^{1/2}, \quad w_{pq} > 0 \end{cases}$$

and (A7)-(A10) are satisfied.

The smoothing functional $J_\beta(\lambda)$ now becomes

$$J_\beta(\lambda) = \int_Q (u(x,t;\lambda) - z_d(x,t))^2 \, dxdt + \beta \sum_{|p|,|q| \leq m} w_{pq} \|a_{pq}\|^2_{H^{\ell_1,\ell_2}(Q)} \tag{6.31}$$

where $u(x,t;\lambda)$ denotes the weak solution of (6.17). Theorem 6.1 gives the following result:

THEOREM 6.2: *The first derivative of the functional* $J_\beta(\lambda)$ *given by* (6.31) *is*

$$J'_\beta(\lambda) \cdot \delta\lambda = \sum_{|p|,|q| \leq m} \left[\int_Q \delta a_{pq} D_x^q u D_x^p p \, dxdt + 2\beta w_{pq} \left(\delta a_{pq}, a_{pq} \right)_{H^{\ell_1,\ell_2}(Q)} \right] \tag{6.32}$$

where u $\in$ W(0,T) *is the weak solution of* (6.17) *and* $p \in$ W(0,T) *is the weak solution of*

$$\left.\begin{aligned} &-\frac{\partial p}{\partial t} + \sum_{|p|,|q|\leq m} (-1)^{|p|} D_x^p(a_{qp} D_x^q p) = -2(u(x,t) - z_d(x,t)) \\ &B_j p = 0 \qquad , j = 0,\ldots, s-1 \\ &\Phi^* p = 0 \qquad , j = s,\ldots, m-1 \\ &p(x,T) = 0 \end{aligned}\right\} \qquad (6.33)$$

In the above theorem, $\{\Phi_j^*\}_{j=o}^{m-1}$ is the adjoint of $\{B_o,\ldots, B_{s-1}, B'_s,\ldots, B'_{m-1}\}$ with respect to Green's formula

$$\int_\Omega \sum_{|p|,|p|\leq m} a_{qp} D_x^q v D_x^p u dx = \int_\Omega \sum_{|p|,|q|\leq m} (-1)^{|p|} [D_x^p(a_{qp} D_x^q v)]\, udx$$

$$- \int_\Gamma \sum_{j=o}^{s-1} \Phi_j^* v B_j u d\sigma - \int_\Gamma \sum_{j=s}^{m-1} \Phi^*_j v B'_j u d\sigma \qquad (6.34)$$

As a further application of the theory, consider the Neumann problem in the isotropic diffusion equation

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} - \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\alpha(x)\,\frac{\partial u}{\partial x_j}\right) = f_o(x,t) \quad , \text{ in } Q \\ &u(x,0) = u_o(x) \qquad , \text{ in } \Omega \\ &\frac{\partial u}{\partial \nu} = 0 \qquad , \text{ on } \Sigma \end{aligned}\right\} \qquad (6.35)$$

As a result of Proposition 6.1 we have that for every

$$\begin{cases} \alpha(x) \in L^{\infty}(\Omega) \text{ with } \alpha(x) \geq \alpha_o > 0 & \text{a.e. in } \Omega \\ f_o(x,t) \in L^2(Q) \\ u_o(x) \in L^2(\Omega) \end{cases}$$

(6.35) admits a unique weak solution $u \in L^2(0,T;H^1(\Omega))$.

Now take

$$\Lambda = L^{\infty}(\Omega)$$

$$\Lambda_{ad} = \{\alpha \in L^{\infty}(\Omega) \,|\, \alpha(x) \geq \alpha_o > 0 \quad \text{a.e. in } \Omega\}$$

Λ_{ad} is a closed and convex subset of $L^{\infty}(\Omega)$.

$$\mathcal{R} = H^{\ell}(\Omega) \quad , \text{ with } \ell > \frac{n}{2}$$

$$\mathcal{R}_{ad} = \mathcal{R} \cap \Lambda_{ad} = \{\alpha \in H^{\ell}(\Omega) \,|\, \alpha(x) \geq \alpha_o > 0 \quad \text{a.e. in } \Omega\}$$

When we observe $u(x,t)$ in Q, we can estimate $\alpha(x)$ by minimizing the functional

$$J_{\beta}(\alpha) = \int_Q [u(x,t;\alpha) - z_d(x,t)]^2 \, dxdt + \beta \|\alpha\|^2_{H^{\ell}(\Omega)} \tag{6.36}$$

Theorem 6.2 shows that $J_{\beta}(\alpha)$ is of C^{∞}-class and

$$J'_{\beta}(\alpha)\cdot\delta\alpha = \int_{\Omega} \delta\alpha \left[\int_o^T \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} \frac{\partial p}{\partial x_j} \, dt \right] dx + 2\beta(\delta\alpha,\alpha)_{H^{\ell}(\Omega)} \tag{6.37}$$

where u is the solution of (6.35) and p is the solution of

$$\left.\begin{aligned} \frac{\partial p}{\partial t} + \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\alpha(x) \frac{\partial p}{\partial x_j}\right) &= 2(u(x,t) - z_d(x,t)), \text{ in } Q \\ \frac{\partial p}{\partial \nu} &= 0 \quad , \text{ on } \Sigma \\ p(x,T) &= 0 \quad , \text{ in } \Omega \end{aligned}\right\} \tag{6.38}$$

7. IDENTIFICATION OF LINEAR PARABOLIC SYSTEMS FROM POINT OBSERVATION

Given a set of discrete points $x_1, \ldots, x_\mu \in \Omega$, we now consider the identification of a linear parabolic system by observing $u(x_i,t)$, $i = 1,\ldots, \mu$.

We have seen in §6 that weak solutions of linear parabolic systems lie in $L^2(0,T;H^m(\Omega))$. Thus, for a weak solution u, the point value $u(x_i,t)$ has meaning if $H^m(\Omega) \subset C^0(\Omega) \Leftrightarrow n < 2m$. Since such an assumption is overly restrictive, we will consider here strong solutions, which are more regular.

Let Ω, Γ, T, Q and Σ as in §6. Consider

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} + A(t)u = f && \text{, in } Q \\ &u(x,0) = u_o && \text{, in } \Omega \\ &B_j u = g_j && \text{, } j = 0,\ldots, m-1, \text{on } \Sigma \end{aligned}\right\} \tag{7.1}$$

where

$$A(t)\, u = \sum_{|p|,|q|\leq m} (-1)^{|p|} D_x^p \left(a_{pq}(x,t) D_x^q u\right) \tag{7.2}$$

$$\left.\begin{aligned} &B_j u = \sum_{|h|\leq m_j} b_{jh}(x,t) D_x^h u \qquad , j = 0,\ldots, m-1 \\ &\text{with } 0 \leq m_j = \text{order of } B_j \leq 2m-1 \end{aligned}\right\} \tag{7.3}$$

Following [21], suppose[6]

Γ is an (n-1)-dimensional C^{2m}-manifold, with Ω locally on one side of Γ (7.4)

$$a_{pq} \in C^{|p,q|,1}(\bar{Q}) \text{ , where } |p,q| = \max(|p|,|q|) \tag{7.5}$$

$$b_{jh} \in C^{2m-m_j,\,2m-m_j+1}(\bar{\Sigma}) \tag{7.6}$$

$$\forall t_o \in [0,T], \text{ the system } \{B_j(x,t_o,D_x)\}_{j=o}^{m-1} \text{ is normal on } \Gamma \tag{7.7}$$

[6] Lions and Magenes use sharper regularity conditions for Γ, a_{pq} and b_{jh} than (7.4)-(7.6). However, the result remains unaltered. See Remark 6.1 in [21, p. 35] and Theorem 3.3 in [4, p. 32].

$$\left.\begin{array}{l}\forall\theta\in[-\frac{\pi}{2},\frac{\pi}{2}]\quad\forall t_o\in[0,T]\\ \{A(x,t_o,D_x)+(-1)^m e^{i\theta}D_y^{2m},\ B_j(x,t_o,D_x)\}\\ \text{is a regular elliptic system in } \bar{\Omega}\times\mathbb{R}_y\end{array}\right\}\tag{7.8}$$

PROPOSITION 7.1 ([21] Vol. II.p.33): *Assume that* (7.2)-(7.8) *hold. Given* $f\in L^2(Q)$, $u_o\in H^m(\Omega)$ *and* $g_j\in H^{2m-m_j-\frac{1}{2},\ (2m-m_j-\frac{1}{2})/2m}(\Sigma)$ *satisfying the compatibility relations*

$$B_j(x,0,D_x)u_o(x)\Big|_\Gamma = g_j(x,0)\quad,\ \text{for all } j \text{ such that } m_j\le m-1\tag{7.9}$$

problem (7.1) *admits a unique solution* $u\in H^{2m,1}(Q)$. *The solution depends continuously on the data* f, u_o *and* g_j.

Now we take

Parameter space:

$$\Lambda=\bigtimes_{|p|,|q|\le m} C^{|p,q|,1}(\bar{Q})\tag{7.10}$$

which is a Banach space with norm $\|\lambda\|_\Lambda=\max\limits_{|p|,|q|\le m}\|a_{pq}\|_{C^{|p,q|,1}(\bar{Q})}$

Operator space:

$$\mathscr{A}=\left\{A\in\mathscr{L}(H^{2m,1}(Q),L^2(Q))/A=\sum_{|p|,|q|\le m}(-1)^{|p|}D_x^p(a_{pq}D_x^q)\right.$$
$$\left.\text{with } a_{pq}\in C^{|p,q|,1}(\bar{Q})\right\}\tag{7.11}$$

which is a Banach space with norm $\|A\|_{\mathscr{A}}=\max\limits_{|p|,|q|\le m}\|a_{pq}\|_{C^{|p,q|,1}(\bar{Q})}$

and we denote by $\mathscr{A}_c$ its open subset

$$\mathscr{A}_c = \{A \in \mathscr{A} / (7.8) \text{ is satisfied}\} \tag{7.12}$$

Space of right-hand sides:

$$F = L^2(Q) \times \tilde{F} \tag{7.13}$$

where $\tilde{F} = \Big\{(u_o, g_j, j = 0, \ldots, m-1)/u_o \in H^m(\Omega), g_j \in H^{2m-m_j-\frac{1}{2},(2m-m_j-\frac{1}{2})/2m}(\Sigma)$

and (7.9) is satisfied$\Big\}$

Space of Solutions:

$$U = H^{2m,1}(Q) \tag{7.14}$$

Now define the mapping

$$\Psi;\ (A,u) \in \mathscr{A} \times U \to \left(\frac{\partial u}{\partial t} + Au,\ u(x,0),\ B_o u, \ldots,\ B_{m-1}u\right) \in F \tag{7.15}$$

Clearly, Ψ is of C^∞-class, hence assumption (A1) is satisfied. Also, we have

$$\frac{\partial \Psi}{\partial u}(A,u)\cdot\delta u = \left(\frac{\partial \delta u}{\partial t} + A\delta u,\ \delta u(x,0),\ B_o\delta u, \ldots,\ B_{m-1}\delta u\right) \tag{7.16}$$

$$\frac{\partial \Psi}{\partial A}(A,u)\cdot\delta A = (\delta A u,\ 0,\ 0,\ \ldots,\ 0) \tag{7.17}$$

Proposition 7.1 establishes satisfaction of (A2) and (A3). Observe that the mapping $\lambda = (a_{pq}) \in \Lambda \to A(\lambda;t) \in \mathscr{A}$ is clearly of C^∞-class, hence (A4) is also satisfied.

Before we proceed to the statement of the main result of this section, we will briefly state a number of results from [21] and [4] that will permit us to interpret the adjoint state as a distributional solution of an irregular parabolic problem.

Let $\{B_j\}_{j=o}^{m-1}$ be a system of boundary operators normal on Γ. If $\{B_j\}_{j=o}^{m-1}$ is "completed" by a system $\{S_j\}_{j=o}^{m-1}$ with $0 \le \mu_j$ = order of $S_j \le 2m-1$ so that $\{B_o, \dots, B_{m-1}, S_o, \dots, S_{m-1}\}$ is normal and of Dirichlet on Γ, then there exist 2m boundary operators C_j, T_j, $j = 0, \dots, m-1$, uniquely defined, with properties

a) the order of C_j is $2m - \mu_j - 1$

b) the order of T_j is $2m - m_j - 1$

c) $\{C_o, \dots, C_{m-1}, T_o, \dots, T_{m-1}\}$ is a Dirichlet system on Γ

such that the following *(formal)* Green's formula holds:

$$\int_Q u\left(Av + \frac{\partial v}{\partial t}\right) dxdt - \int_Q \left(A^*u - \frac{\partial u}{\partial t}\right) v\, dxdt =$$

$$= \sum_{j=o}^{m-1} \int_\Sigma S_j v C_j u d\Sigma - \sum_{j=o}^{m-1} \int_\Sigma B_j v T_j u d\Sigma +$$

$$+ \int_\Omega u(x,T)v(x,T)dx - \int_\Omega u(x,0)v(x,0)dx \qquad \forall\, u,\ v \in \mathcal{D}(\bar{Q}) \tag{7.18}$$

where A^* denotes the formal adjoint of A

$$A^*u = \sum_{|p|,|q| \le m} (-1)^{|p|} D_x^p (a_{qp} D_x^q u) \tag{7.19}$$

Now denote

$$X(Q) = \{v/v \in H^{2m,1}(Q),\ B_j v = 0,\ j = 0, \dots, m-1,\ v(x,0) = 0\} \tag{7.20}$$

which is a Hilbert space with norm

$$\|v\|_{X(Q)} = \left(\|v\|^2_{H^{2m,1}(Q)} + \left\|\frac{\partial v}{\partial t} + Av\right\|^2 \right)^{\frac{1}{2}}$$

and

$$D_{p*}(Q) = \left\{u/u \in L^2(Q),\ -\frac{\partial u}{\partial t} + A^*u \in \Xi^{-2m,-1}(Q)\right\} \tag{7.21}$$

which is a Hilbert space with norm

$$\|u\|_{D_{p*}(Q)} = \left(\|u\|^2_{L^2(Q)} + \|-\frac{\partial u}{\partial t} + A^*u\|^2_{\Xi^{-2m,-1}(Q)}\right)^{\frac{1}{2}}$$

PROPOSITION 7.2 (*Trace Theorem. [21] Vol. II pp. 49,55 and [4] p. 83):*

a) The mapping $u \to C_j u$ *of* $\mathcal{D}(\bar{Q})$ *in* $\mathcal{D}(\bar{\Sigma})$ *extends by continuity to a continuous linear mapping, still denoted by* C_j*, of*

$$D_{p*}(Q) \to H^{-2m+\mu_j+\frac{1}{2},(-2m+\mu_j+\frac{1}{2})/2m}(\Sigma)$$

b) The mapping $u \to u(x,T)$ *of* $\mathcal{D}(\bar{Q})$ *in* $\mathcal{D}(\bar{\Omega})$ *extends by continuity to a continuous linear mapping of* $D_{p*}(Q)$ *in* $H^{-m}(\Omega)$.

PROPOSITION 7.3 (*Partial Green's Formulas. [21] Vol. II pp.53,57 and [4] p.84)*

a) $\langle u, Av + \frac{\partial v}{\partial t}\rangle = \langle A^*u - \frac{\partial u}{\partial t}, v\rangle + \sum_{j=0}^{m-1} \langle S_j v, C_j u\rangle$

for $u \in D_{p*}(Q)$, $v \in H^{2m,1}_{,0}(Q)$, $B_j v = 0$, $j = 0,\ldots, m-1$

b) $\langle u, Av + \frac{\partial v}{\partial t}\rangle = \langle A^*u - \frac{\partial u}{\partial t}, v\rangle + \langle u(x,T), v(x,T)\rangle$

for $u \in D_{p*}(Q)$, $v \in H^{2m,1}(Q)$, $v(x,0) = 0$, $B_j v = S_j v = 0$, $j = 0,\ldots, m-1$

Note that Propositions 7.2 and 7.3 will not be used in the sequel. They only serve for the precise interpretation of Proposition 7.4, which follows.

PROPOSITION 7.4 *([21] Vol. II p. 60 and [4] p.86): Assume that* (7.2)-(7.8) *hold. Given* $\phi \in \Xi^{-2m,-1}(Q)$, $h_j \in H^{-2m+\mu_j+\frac{1}{2}}\,\Xi^{(-2m+\mu_j+\frac{1}{2})/2m}(\Sigma)$ *and* $u_T \in \Xi^{-m}(\Omega)$ *there exists a unique distribution* $u \in D_{p*}(Q)$ *satisfying*

$$\langle u, \frac{\partial v}{\partial t} + Av\rangle = \langle \phi, v\rangle + \sum_{j=0}^{m-1} \langle h_j, S_j v\rangle + \langle u_T, v(x,T)\rangle$$

for every $v \in X(Q)$.

In other words, u *is a distributional solution of*

$$\left.\begin{aligned} -\frac{\partial u}{\partial t} + A^*u &= \phi \quad &, \text{ in } Q \\ C_ju &= h_j,\ j = 0,\dots, m-1, &\text{ on } \Sigma \\ u(x,T) &= u_T \quad &, \text{ in } \Omega \end{aligned}\right\} \tag{7.22}$$

where C_ju *and* $u(x,T)$ *are understood in the sense of Proposition* 7.2. *Furthermore,* u *depends continuously on the data* ϕ, h_j *and* u_T.

Now to identify $\lambda = (a_{pq})$ in (7.1) from an observation of u at the points x_i, $i = 1,\dots,\mu$ take

$$\mathscr{H} = (L^2(0,T))^\mu \tag{7.23}$$

$$\Lambda_{\mathscr{H}} = \text{identity} \tag{7.24}$$

$$\mathscr{C};u(x,t) \in H^{2m,1}(Q) \to (u(x_i,t),\ i = 1,\dots, \mu) \in (L^2(0,T))^\mu \tag{7.25}$$

$$z_d = (z_{d_1}(t),\dots, z_{d_\mu}(t)) \in (L^2(0,T))^\mu \tag{7.26}$$

$$\mathscr{R} = \text{a Hilbert space compactly imbedded in } \Lambda \tag{7.27}$$

and minimize the smoothing functional

$$J_\beta(\lambda) = \sum_{i=1}^{\mu} \int_0^T [u(x_i,t;\lambda) - z_{d_i}(t)]^2 dt + \beta\|\lambda\|^2_{\mathscr{R}} \tag{7.28}$$

where $u(x,t;\lambda)$ denotes the strong solution of (7.1).

Notice that since $u \in H^{2m,1}(Q)$, it follows that $u \in L^2(0,T;H^{2m}(\Omega))$. Hence $u(x_i,t)$ has meaning and $(t \to u(x_i,t)) \in L^2(0,T)$ if $H^{2m}(\Omega) \subset C^0(\Omega) \Leftrightarrow n < 4m$. So we assume

$$n \leq 4m-1 \tag{7.29}$$

Also notice that clearly $\mathscr{C}$ is of C^∞-class, so that (A7) is satisfied.
To each observation $(z_{d_1}(t),\ldots, z_{d_\mu}(t)) \in (L^2(0,T))^\mu$ we associate the distribution

$$\phi = -2 \sum_{i=1}^{\mu} (u(x_i,t) - z_{d_i}(t)) \otimes \delta(x-x_i) \in \mathscr{D}'(\Omega) \tag{7.30}$$

Due to (7.29), it is easy to verify that

$$\phi \in \Xi^{-2m,-1}(Q) \tag{7.31}$$

Thus we have the following

THEOREM 7.1: *The functional* $J_\beta(\lambda)$ *given by* (7.28) *is of* C^∞-*class. Its first derivative is given by*

$$J'_\beta(\lambda)\cdot\delta\lambda = \int_Q A(\delta\lambda)u\,p\,dxdt + 2\beta(\delta\lambda,\lambda)_{\mathscr{R}} \tag{7.32}$$

where $u \in H^{2m,1}(Q)$ *is the unique solution of* (7.1) *and* $p \in L^2(Q)$ *is the unique solution of*

$$\left.\begin{aligned}&\int_Q p\left(\frac{\partial v}{\partial t} + Av\right) dxdt = -2\sum_{i=1}^{\mu} \int_0^T (u(x_i,t) - z_{d_i}(t))v(x_i,t)\,dt \\ &\forall v \in H^{2m,1}(Q) \text{ with } B_j v = 0,\ j = 0,\ldots, m-1,\ v(x,0) = 0\end{aligned}\right\} \tag{7.33}$$

In other words, p *is a distributional solution of*

$$\left.\begin{aligned}&-\frac{\partial p}{\partial t} + A^*p = -2\sum_{i=1}^{\mu} (u(x_i,t) - z_{d_i}(t)) \otimes \delta(x-x_i), && \text{in } Q\\ &C_j p = 0, \quad j = 0,\ldots, m-1, && \text{on } \Sigma\\ &p(x,T) = 0, && \text{in } \Omega\end{aligned}\right\} \tag{7.34}$$

Proof: Apply Theorem 4.1. The adjoint state $\rho \in F'$ is of the form $\rho = (p, \tilde{p}, q_j, j = 0,\ldots, m-1)$ with $p \in L^2(Q)$, $\tilde{p} \in H^{-m}(\Omega)$, $q_j \in H^{-2m+m_j+\frac{1}{2},(-2m+m_j+\frac{1}{2})/2m}(\Sigma)$. Taking into account (7.17), we immediately get (7.32) from (4.4). Taking into account (7.16), (4.5) becomes

$$\int_Q p\left(\frac{\partial v}{\partial t} + Av\right) dxdt + \langle \tilde{p}, v(x,0)\rangle + \sum_{j=o}^{m-1} \langle q_j, B_j v\rangle =$$

$$= -2\left[\mathscr{C}^*(\mathscr{C}u - z_d)\right](v)$$

from which (7.33) follows.

Uniqueness of the solution of (7.33) follows from Proposition 7.4. (7.34) also follows from Proposition 7.4. This completes the proof of the theorem.

Remark 7.1: For $A(\lambda) = \sum_{|p|,|q|\leq m} (-1)^{|p|} D_x^p(a_{pq}(x,t)D_x^q)$ with $a_{pq} \in C^{|p,q|,1}(\bar{Q})$, the first term in (7.32) can be *formally* rewritten as

$$\sum_{|p|,|q|\leq m} \int_Q \delta a_{pq} D^q u D^p p \; dxdt$$

by using Green's formula.

Remark 7.2: As in the case of distributed observations, the usual choice of $\mathscr{R}$ will be $\displaystyle\bigtimes_{|p|,|q|\leq m} H^{\ell_1,\ell_2}(Q)$ with norm $\left(\sum_{|p|,|q|\leq m} w_{pq} \|a_{pq}\|^2_{H^{\ell_1,\ell_2}(Q)}\right)^{\frac{1}{2}}$, $w_{pq} > 0$. $\ell_1(p,q)$ and $\ell_2(p,q)$ will be chosen so that the compact imbedding $H^{\ell_1,\ell_2} \to C^{|p,q|,1}$ is valid.

As an application of the theory, consider the Neumann problem in the n-dimensional isotropic diffusion equation ($n \leq 3$):

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} - \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\alpha(x)\,\frac{\partial u}{\partial x_j}\right) = f(x,t) && \text{, in } Q \\ &u(x,0) = u_o(x) && \text{, in } \Omega \\ &\frac{\partial u}{\partial \nu} = 0 && \text{, on } \Sigma \end{aligned}\right\} \qquad (7.35)$$

Applying Proposition 7.1 we have that for every

$\alpha(x) \in C^1(\bar{\Omega})$ satisfying $\alpha(x) \geq \alpha_o \quad \forall\, x \in \Omega$ for some $\alpha_o > 0$

$f(x,t) \in L^2(Q)$

$u_o(x) \in H^1(\Omega)$

the problem (7.35) admits a unique solution $u \in H^{2,1}(Q)$.

Now take

$$\Lambda = C^1(\bar{\Omega})$$

$$\Lambda_{ad} = \{\alpha \in C^1(\bar{\Omega}) / \alpha(x) \geq \alpha_o > 0 \qquad \forall\, x \in \Omega\}$$

Λ_{ad} is a closed and convex subset of $C^1(\bar{\Omega})$.

$$\mathscr{R} = H^{\ell}(\Omega) \qquad \text{, with } \ell > \frac{n}{2} + 1$$

$$\mathscr{R}_{ad} = \mathscr{R} \cap \Lambda_{ad} = \{\alpha \in H^{\ell}(\Omega)/\alpha(x) \geq \alpha_o > 0 \qquad \forall\, x \in \Omega\}$$

When we observe u at a number of points x_i, $i = 1,\ldots,\mu$ in Ω, we can estimate $\alpha(x)$ by minimizing the functional

$$J_\beta(\alpha) = \sum_{i=1}^{\mu} \int_o^T [u(x_i,t;\alpha) - z_{d_i}(t)]^2\, dt + \beta\|\alpha\|^2_{H^{\ell}(\Omega)} \qquad (7.36)$$

Theorem 7.1 shows that $J_\beta(\alpha)$ is of C^∞-class and

$$J'_\beta(\alpha)\cdot\delta\alpha = \int_\Omega \delta\alpha\left[\int_0^T \sum_{j=1}^n \frac{\partial u}{\partial x_j}\frac{\partial p}{\partial x_j}\,dt\right]dx + 2\beta(\delta\alpha,\alpha)_{H^\ell(\Omega)} \qquad (7.37)$$

where u is the solution of (7.35) and p is the solution of

$$\left.\begin{aligned} &\frac{\partial p}{\partial t} + \sum_{j=1}^n \frac{\partial}{\partial x_j}\left(\alpha(x)\frac{\partial p}{\partial x_j}\right) = 2\sum_{i=1}^{\mu}(u(x_i,t) - z_{d_i}(t))\otimes\delta(x-x_i), \text{ in } Q \\ &\frac{\partial p}{\partial \nu} = 0 \qquad , \text{ on } \Sigma \\ &p(x,T) = 0 \quad , \text{ in } \Omega \end{aligned}\right\} \qquad (7.38)$$

8. IDENTIFICATION OF A NONLINEAR PARABOLIC SYSTEM

In §6 and §7 we have studied the identification of linear parabolic systems. The purpose of this section is to show that the regularization approach is applicable to systems described by nonlinear evolution equations and, at the same time, point out the difficulties involved in regularizing nonlinear identification problems.

It has been apparent from §6 and §7 that in order to apply regularization methods to the identification of distributed parameter systems, what is needed is an appropriate P.D.E. framework. And it is a well-known fact that nonlinear P.D.E. theories are not sufficiently general. Furthermore, unlike the linear case, an existence-uniqueness-continuous dependence theorem for the P.D.E. is insufficient to satisfy assumption (A3) and give (By Proposition 2.2) differentiability of the solution with respect to the coefficients. To satisfy (A3), additional assumptions will usually be required.

For the above reasons, it does not seem to be possible to develop a fairly general theory for nonlinear parabolic identification problems; they must be treated on a case-by-case basis. In this section we study the identification of the spatially varying coefficients $a_{ij}(x)$ in

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} - \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(x)\, \frac{\partial u}{\partial x_j}\right) + b(x)u + |u|^{\gamma}u = f \quad , \text{ in } Q \\ &u(x,0) = u_0(x) \qquad , \text{ in } \Omega \\ &u = 0 \qquad , \text{ on } \Sigma \end{aligned}\right\} \tag{8.1}$$

from distributed observation, where $\gamma > 0$ and $b(x)$ is bounded from below by a positive number. Note that the initial/boundary-value problem (8.1) is a classical example of application of the theory of monotone operators. The resulting existence-uniqueness-continuous dependence theorem (Proposition 8.1) is not enough to verify assumption (A3). The result of Proposition 8.2, that uses additional assumptions, will, however, establish satisfaction of (A3).

Following [13] and [19], we take

$$V = H_o^1(\Omega) \tag{8.2}$$

$$H = L^2(\Omega) \tag{8.3}$$

$$\mathscr{A} = \mathscr{L}(V,V') \qquad \textit{(Operator space)} \tag{8.4}$$

$$\mathscr{A}_c = \{A \in \mathscr{A} / \exists\ \eta > 0 \text{ with } (Av,v)_{V'V} \geq \eta \|v\|_V^2 \qquad \forall\ v \in V\} \tag{8.5}$$

(subspace of coercive operators)

$$\Lambda = \bigtimes_{i,j=1}^{n} L^\infty(\Omega) \qquad \textit{(Parameter space)} \tag{8.6}$$

$$\left.\begin{array}{l} \forall \lambda = (a_{ij}) \in \Lambda \text{ we correspond } A \in \mathscr{A} \text{ given by} \\ (Au,v)_{V'V} = \displaystyle\int_\Omega \sum_{i,j=1}^{n} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i}\, dx \qquad \forall u,\ v \in V \end{array}\right\} \tag{8.7}$$

Using methods of monotone operators ([13] Ch. VI §1, [19] Ch. 2 §1) it is straightforward to prove

PROPOSITION 8.1: *Given*

$$\begin{cases} \gamma > 0 \\ A \in \mathscr{A}_c \\ b \in L^\infty(\Omega) \textit{ satisfying } b(x) \geq b_o > 0 \ \textit{a.e in } \Omega \\ f \in L^2(0,T;V') \oplus L^{\frac{\gamma+2}{\gamma+1}}(Q) \\ u_o \in L^2(\Omega) \end{cases}$$

the problem

$$\left.\begin{aligned} &\frac{du}{dt} + Au + bu + |u|^{\gamma}u = f \\ &u(0) = u_o \end{aligned}\right\} \tag{8.8}$$

admits a unique solution $u \in L^2(0,T;V) \cap L^{\gamma+2}(Q)$ *such that* $\frac{du}{dt} \in L^2(0,T;V') \oplus L^{\frac{\gamma+2}{\gamma+1}}(Q)$.[7]

It is easily seen that the above result is not enough to verify (A3) unless additional assumptions are made. Using the maximum principle one obtains (see [4, p. 37])

PROPOSITION 8.2: *Under the additional assumption*

$$\left.\begin{aligned} &\exists\ c_1, c_2 \textit{ with } 0 < c_1 < c_2 \textit{ such that} \\ &0 < c_1 \le \frac{f(x,t)}{b(x)} \le c_2 \qquad \textit{a.e. in } Q \\ &0 < c_1 \le u_o(x) \le c_2 \qquad \textit{a.e in } \Omega \end{aligned}\right\} \tag{8.9}$$

problem (8.8) *admits a unique solution* u *in*

$$U = \left\{ u/u \in L^2(0,T;V) \cap L^\infty(Q),\ \frac{du}{dt} + Au \in L^\infty(Q),\ u(0) \in L^\infty(\Omega) \right\} \tag{8.10}$$

satisfying

$$u(x,t) \ge c_1 \qquad \textit{a.e. in } Q \tag{8.11}$$

Now denote

$$U_c = \left\{ u \in U/\exists\, c_1 > 0 \text{ such that } u(x,t) \ge c_1 \quad \text{a.e. in } Q \right\} \tag{8.12}$$

[7]Footnote on following page

Footnote for page 137

[7]The space $L^2(0,T;V) \cap L^{\gamma+2}(Q)$ is a Banach space with norm

$$\|u\|_{L^2(0,T;V)} + \|u\|_{L^{\gamma+2}(Q)}$$

The space $L^2(0,T;V') \oplus L^{\frac{\gamma+2}{\gamma+1}}(Q) = \left\{u + v \,/ u \in L^2(0,T;V'),\ v \in L^{\frac{\gamma+2}{\gamma+1}}(Q)\right\}$ is the dual of $L^2(0,T;V) \cap L^{\gamma+2}(Q)$. It is a Banach space with norm

$$\|w\| = \inf_{\substack{u \in L^2(0,T;V') \\ v \in L^{\frac{\gamma+2}{\gamma+1}}(Q) \\ w = u + v}} \max\left(\|u\|_{L^2(0,T;V')},\ \|v\|_{L^{\frac{\gamma+2}{\gamma+1}}(Q)} \right)$$

$$F = L^{\infty}(Q) \times L^{\infty}(\Omega) \tag{8.13}$$

Observe that $\mathscr{A}_c$ and U_c are open in $\mathscr{A}$ and U, respectively, and define

$$\Psi; (A,u) \in \mathscr{A}_c \times U_c \to \left(\frac{du}{dt} + Au + bu + |u|^{\gamma}u,\ u(0) \right) \in F \tag{8.14}$$

As an immediate consequence of Proposition 8.2, (A2) is satisfied. Clearly, Ψ is of C^1-class, hence (A1) is satisfied. Furthermore,

$$\frac{\partial \Psi}{\partial u}(A,u) \cdot \delta u = \left(\frac{d(\delta u)}{dt} + A\delta u + b\delta u + (\gamma+1)|u|^{\gamma}\delta u,\ \delta u(0) \right) \tag{8.15}$$

$$\frac{\partial \Psi}{\partial A}(A,u) \cdot \delta A = (\delta A u, 0) \tag{8.16}$$

To show (A3), observe that the equation

$$\begin{aligned} &\frac{dz}{dt} + Az + bz + (\gamma+1)|u|^{\gamma}z = \tilde{f} \\ &z(0) = \tilde{z}_o \end{aligned} \tag{8.17}$$

admits for every $\tilde{f} \in L^{\infty}(Q)$ and every $\tilde{z}_o \in L^{\infty}(\Omega)$ a unique solution $z \in L^2(0,T;V)$. Furthermore, $z \in L^{\infty}(Q)$ by using the maximum principle. Hence, $z \in U$.

Observe that the mapping $\lambda = (a_{ij}) \in \Lambda \to A(\lambda) \in \mathscr{A}$ is clearly of C^{∞}-class, hence (A4) is satisfied. Also, given $a_o > 0$, define

$$\Lambda_{ad} = \left\{ \lambda = (a_{ij}) \in \Lambda \Big/ \sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq a_o |\xi|^2 \quad \forall \xi \in \mathbb{R}^n \text{ a.e. in } \Omega \right\} \tag{8.18}$$

Λ_{ad} is closed and convex and $A(\Lambda_{ad}) \subseteq \mathscr{A}_c$, i.e. (A5) and (A6) are satisfied. Finally, suppose that we want to identify $\lambda = (a_{ij})$ by observing $u(x,t)$. We take

$$\mathscr{H} = L^2(Q) \tag{8.19}$$

$$\Lambda_{\mathscr{H}} = \text{identity} \tag{8.20}$$

$$\mathscr{C} = \text{injection of } L^\infty(Q) \text{ into } L^2(Q) \tag{8.21}$$

$$\mathscr{R} = \bigtimes_{i,j=1}^{n} H^\ell(\Omega) \tag{8.22}$$

with
$$\begin{cases} \ell > \frac{n}{2} \\ \|\lambda\|_{\mathscr{R}} = \left(\sum_{i,j=1}^{n} w_{ij} \|a_{ij}\|^2_{H^\ell(\Omega)} \right)^{\frac{1}{2}} , \quad w_{ij} > 0 \end{cases}$$

and have (A7)-(A10) being satisfied.

THEOREM 8.1: *The smoothing functional*

$$J_\beta(\lambda) = \int_Q (u(x,t;\lambda) - z_d(x,t))^2 \, dxdt + \sum_{i,j=1}^{n} w_{ij} \|a_{ij}\|^2_{H^\ell(\Omega)} \tag{8.23}$$

is of C^1*-class. Its derivative is given by*

$$J'_\beta(\lambda)\cdot\delta\lambda = \sum_{i,j=1}^{n} \left[\int_Q \delta a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial p}{\partial x_i} \, dxdt + 2\beta w_{ij} (\delta a_{ij}, a_{ij})_{H^\ell(\Omega)} \right] \tag{8.24}$$

where $u \in U$ *is the weak solution of* (8.1) *and* $p \in L^2(0,T;H^1_0(\Omega))$ *is the weak solution of*

$$\left.\begin{aligned} &\frac{\partial p}{\partial t} + \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ji}(x) \frac{\partial p}{\partial x_j} \right) - [b(x) + (\gamma+1)|u|^\gamma]\, p = 2(u-z_d), \text{ in } Q \\ &p = 0, \text{ on } \Sigma \\ &p(x,T) = 0 , \text{ in } \Omega \end{aligned}\right\} \tag{8.25}$$

Proof: Apply Theorem 4.1. The adjoint state ρ is of the form $\rho = (p, \tilde{p})$ with $p \in [L^\infty(Q)]'$ and $\tilde{p} \in [L^\infty(\Omega)]'$. Taking into account (8.15), (4.5) becomes:

$$\left.\begin{array}{l} \langle p, \frac{dv}{dt} + Av + bv + (\gamma+1)|u|^\gamma v \rangle + \langle \tilde{p}, v(0) \rangle = -2 \int_Q (u-z_d)v \, dxdt \\ \forall\, v \in V \end{array}\right\} \quad (8.26)$$

(8.26) is satisfied if we define

$$\begin{cases} p \in L^2(0,T;H_0^1(\Omega)) \subset L^2(Q) \subset L^1(Q) \subset [L^\infty(Q)]' \text{ as solution of (8.25)} \\ \tilde{p} \in L^2(\Omega) \subset L^1(\Omega) \subset [L^\infty(\Omega)]' \text{ by } \tilde{p} = p(0) \end{cases}$$

Finally, (8.24) results from equation (4.4).

This completes the proof of the theorem.

Remark: If the data are more regular, e.g. $a_{ij} \in C^1(\bar{\Omega})$, $u_0 \in H_0^1(\Omega)$, the solution u of (8.1) will be in $H^{2,1}(Q)$. Hence, one will be able to consider point observation as well.

9. NUMERICAL IMPLEMENTATION OF THE REGULARIZATION METHOD

The minimization of $J_\beta(\lambda)$ can be conveniently carried out by a gradient method ([3],[11]), in which J_β is iteratively minimized along the gradient direction, $\partial J_\beta/\partial\lambda$, which is defined as the unique element $\phi \in \mathscr{R}$ satisfying $J'_\beta(\lambda)\cdot h = (\phi,h)_{\mathscr{R}}$ $\forall h \in \mathscr{R}$. To illustrate the theory, we will consider the identification of $\alpha(x)$ in the one-dimensional diffusion equation from point observations $z_{d_i}(t)$ of $u(x_i,t)$, $i = 1,\ldots,\mu$.

$$\left.\begin{aligned} &\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) + f && ,\ \text{in } \Omega \times\,]0,T[\\ &u(x,0) = u_o(x) && ,\ \text{in } \Omega \\ &\frac{\partial u}{\partial x} = 0 && ,\ \text{on } \Gamma \times\,]0,T[\end{aligned}\right\} \tag{9.1}$$

The smoothing functional is

$$J_\beta(\alpha) = \sum_{i=1}^{\mu} \int_o^T [u(x_i,t) - z_{d_i}(t)^2]dt + \beta\|\alpha\|^2_{H^2(\Omega)} \tag{9.2}$$

If $H^2(\Omega)$ is equipped with the norm

$$\|f\|_{H^2(\Omega)} = \left(\int_\Omega (f^2 + f''^2)dx\right)^{\frac{1}{2}} \tag{9.3}$$

the gradient $\partial J_\beta/\partial\alpha$ is given by

$$\frac{\partial J_\beta}{\partial\alpha}(\alpha) = \psi + 2\beta\alpha \tag{9.4}$$

where ψ is the weak solution of

$$\left.\begin{aligned} &\frac{d^4\psi}{dx^4} + \psi = \int_o^T \frac{\partial u}{\partial x}\frac{\partial p}{\partial x}\,dt \quad , \text{ in } \Omega \\ &\psi'' = 0 \quad , \text{ on } \Gamma \\ &\psi''' = 0 \ , \text{ on } \Gamma \end{aligned}\right\} \tag{9.5}$$

with u and p being the solutions of the state and adjoint equations respectively. The gradient algorithm in this case proceeds as follows:

(1) Initialize $\alpha \in H^2(\Omega)$

(2) Solve the state and adjoint equations

(3) Calculate $J_\beta(\alpha)$ and $\int_o^T \frac{\partial u}{\partial x}\frac{\partial p}{\partial x}\,dt$

(4) If $|J_\beta(\alpha^{old}) - J_\beta(\alpha^{new})| <$ Tolerance, stop

(5) Solve (9.5) for ψ and calculate $\frac{\partial J_\beta}{\partial \alpha}(\alpha)$

(6) Set $\alpha^{new} = \alpha^{old} + \varepsilon \frac{\partial J_\beta}{\partial \alpha}(\alpha^{old})$ where ε is a step length parameter to be determined by one-dimensional line-search.

(7) Go to (2).

We have considered the three cases given in Table 1. Data were generated by first numerically solving (9.1) using the Crank-Nicholson scheme with 50 grid points and then adding to $u(x_i,t)$ random numbers with zero mean and standard deviation $\sigma = 0.2$.

The smoothing functional

$$J_\beta(\alpha) = \frac{1}{5} \int_0^{0.5} \sum_{i=1}^{10} (u(x_i,t) - z_{d_i}(t))^2 \, dt + \\ + \beta \int_0^1 [(\alpha(x))^2 + (\alpha''(x))^2] \, dx \qquad (9.6)$$

was minimized by applying the gradient algorithm described above. The state and adjoint equations were solved by the Crank-Nicholson method. The fourth order O.D.E. giving ψ was solved by a finite-difference scheme. The one-dimensional line search for the step length was performed by the golden section search method. Finally, the test for stopping the iterations was

$$|J_\beta(\alpha^{new}) - J_\beta(\alpha^{old})| < 10^{-3}$$

The initial guess for $\alpha(x)$, the true $\alpha(x)$ and the result after six iterations of the gradient method are shown for Case 1 in Figure 2. Similarly, the estimated $\alpha(x)$ after six iterations is shown for Case 2 in Figure 3. In each of Cases 1 and 2 the value of the regularization parameter β was selected based on the suggestion of Miller (see §5). In Case 1, with an assumed upper bound of 0.05 for the squared error and an assumed upper bound for smoothness of $\|\alpha^{true}\|^2_{H^2} \leq 1$, we obtain $\beta = 5 \times 10^{-2}$. In Case 2, with the same assumed upper bound of 0.05 for the squared error and that for smoothness of $\|\alpha^{true}\|^2_{H^2} \leq 10$, we have $\beta = 5 \times 10^{-3}$.

The effect of the choice of β is examined in Case 3. Figure 4 shows three estimated $\alpha(x)$ profiles corresponding to $\beta = 10^{-4}$, 10^{-2}, and 1. The value $\beta = 10^{-2}$ is consistent with the suggestion of Miller. We note that when $\beta = 10^{-4}$ the oscillations in $\alpha(x)$, characteristic of numerical instability, are setting in. In the absence of a good estimate for the errors and/or smoothness, it is a good idea to examine the solution as a function of β.

REFERENCES

[1] J. P. AUBIN, *Applied Abstract Analysis*, Wiley, New York, 1977.

[2] A. BENSOUSSAN, J. L. LIONS AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structures*, North Holland, Amsterdam, 1978.

[3] J. CÉA, *Optimization: Théorie et Algorithmes*, Dunod, Paris, 1971.

[4] G. CHAVENT, *Analyse Fonctionelle et Identification de Coefficients Répartis dans les Équations aux Dérivées Partielles*, Thèse d' État, Faculté des Sciences de Paris, 1971.

[5] G. CHAVENT, M. DUPUY AND P. LEMONNIER, *History Matching by Use of Optimal Control Theory*, Soc. Pet. Eng. J., 15 (1975), pp. 74-86.

[6] G. CHAVENT, *About the Stability of the Optimal Control Solution of Inverse Problems*, Inverse and Improperly Posed Problems in Differential Equations, G. Anger, ed., Akademie-Verlag, Berlin, 1979, pp. 45-78.

[7] G. CHAVENT, *Identification of Distributed Parameter Systems: About the Output Least Square Method, its Implementation and Identifiability*, Proceedings 5th IFAC Symposium on Identification and System Parameter Estimation, Vol. I, R. Isermann, ed., Pergamon Press, 1980, pp. 85-97.

[8] W. H. CHEN AND J. H. SEINFELD, *Estimation of Spatially Varying Parameters in Partial Differential Equations*, Int. J. Control, 15 (1972), pp. 487-495.

[9] W. H. CHEN, G. R. GAVALAS, J. H. SEINFELD AND M. L. WASSERMAN, *A New Algorithm for Automatic History Matching*, Soc. Pet. Eng. J., 14 (1974), pp. 593-608.

[10] K. H. COATS, J. R. DEMPSEY AND J. H. HENDERSON, *A New Technique for Determining Reservoir Descriptions from Field Performance Data*, Soc. Pet. Eng. J., 10 (1970), pp. 66-74.

[11] J. W. DANIEL, *The Approximate Minimization of Functionals*, Prentice Hall, Englewood Cliffs, 1971

[12] N. DISTEFANO AND A. RATH, *An Identification Approach to Subsurface Hydrological Systems*, Water Resour. Res., 11 (1975), pp. 1005-1012

[13] H. GAJEWSKI, K. GRÖGER AND K. ZACHARIAS, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Akademie-Verlag, Berlin, 1974.

[14] G. R. GAVALAS, P. C. SHAH AND J. H. SEINFELD, *Reservoir History Matching by Bayesian Estimation*, Soc. Pet. Eng. J., 16 (1976), pp. 337-350.

[15] S. KITAMURA AND S. NAKAGIRI, *Identifiability of Spatially-Varying and Constant Parameters in Distributed Systems of Parabolic Type,* SIAM J. Control, 15 (1977), pp. 785-802.

[16] M. M. LAVRENTIEV, V. G. ROMANOV AND V. G. VASILIEV, *Multidimensional Inverse Problems for Differential Equations,* Springer-Verlag, Berlin, 1970.

[17] J. L. LIONS, *Some Aspects of Modeling Problems in Distributed Parameter Systems,* Proceedings of the IFIP Working Conference held in Rome 1976, A. Ruberti, ed., Lecture Notes in Control and Information Sciences Vol. 1, Springer-Verlag, Berlin, 1978, pp. 11-41.

[18] J. L. LIONS, *Équations Différentielles Opérationelles et Problèmes aux Limites,* Springer-Verlag, Berlin, 1961

[19] J. L. LIONS, *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires,* Dunod, Paris, 1969.

[20] J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations,* Springer-Verlag, Berlin, 1971.

[21] J. L. LIONS AND E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications,* Vols. I and II, Springer-Verlag, Berlin, 1972.

[22] K. MILLER, *Least-Squares Methods for Ill-Posed Problems with a Prescribed Bound,* SIAM J. Math. Anal, 1 (1970), pp. 52-74.

[23] M. Z. NASHED, *Approximate Regularized Solutions to Improperly Posed Linear Integral and Operator Equations,* Constructive and Computational Methods for Differential and Integral Equations, D. L. Colton and R. P. Gilbert, eds., Springer-Verlag, Berlin, 1974, pp. 289-332.

[24] S. P. NEUMAN AND S. YAKOWITZ, *A Statistical Approach to the Inverse Problem of Acquifer Hydrology. 1. Theory,* Water Resour. Res., 15 (1979), pp. 845-860.

[25] P. C. SHAH, G. R. GAVALAS AND J. H. SEINFELD, *Error Analysis in History Matching: The Optimum Level of Parameterization,* Soc. Pet. Eng. J., 18 (1978), pp. 219-228.

[26] L. SCHWARTZ, *Cours d' Analyse,* Hermann, Paris, 1967

[27] A. N. TIKHONOV, *Solution of Ill-Posed Problems and the Regularization Method,* Dokl. Akad. Nauk. SSSR, 151 (1963), pp. 501-504. Soviet Math. Dokl. 4(1963), pp. 1035-1038.

[28] A. N. TIKHONOV, *Regularization of Ill-Posed Problems,* Dokl. Akad. Nauk. SSSR, 153 (1963), pp. 49-52, Soviet Math. Dokl. 4(1963), pp. 1624-1627.

[29] A. N. TIKHONOV AND V. YA. ARSENIN, *Solutions of Ill-Posed Problems*, Winston-Wiley, New York, 1977.

[30] T. L. VINCENT AND W. J. GRANTHAM, *Optimality in Parametric Systems*, Wiley, New York, 1981

[31] G. WAHBA, *Practical Approximate Solutions to Linear Operator Equations When Data are Noisy*, Siam. N. Numer. Anal., 14 (1977), pp. 651-667

[32] S. YAKOWITZ AND L. DUCKSTEIN, *Instability in Acquifer Identification: Theory and Case Studies*, Water Resour. Res., 16 (1980), pp. 1054-1064.

[33] P. L. YU, *Cone Convexity, Cone Extreme Points, and Nondominated Solutions in Decision Problems with Multiobjectives*, Multicriteria Decision Making and Differential Games, G. Leitmann, ed., Plenum, New York, 1976, pp. 1-60.

Table 1. Numerical Values for Identification of $\alpha(x)$ in

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x)\,\frac{\partial u}{\partial x}\right) \quad , \; 0 < x < 1 \; , \quad 0 < t < 0.5$$

$$u(x,0) = 10 + 270x^2 - 180x^3$$

$$\frac{\partial u}{\partial x}(0,t) = \frac{\partial u}{\partial x}(1,t) = 0$$

based on noisy observations of u at the points $x_i = \frac{5i-3}{49}$, $i = 1,\ldots, 10$

Case	True $\alpha(x)$	$\lVert\alpha\rVert^2_{H^2}$	Standard deviation of errors in the data	Regularization Parameter
1	$0.5 + 0.5x$	0.583	0.2	5×10^{-2}
2	$0.5 + x - 5x^4 + 6x^5 - 2x^6$	6.236	0.2	5×10^{-3}
3	$0.5 + 0.5x - 0.5x^2$	1.342	0.2	10^{-4}
				10^{-2}
				1

FIGURE CAPTIONS

Figure 1. Function spaces for the abstract identification problem

Figure 2. True and estimated profiles of α for Case 1

Figure 3. True and estimated profiles of α for Case 2

Figure 4. True and estimated profiles of α for Case 3

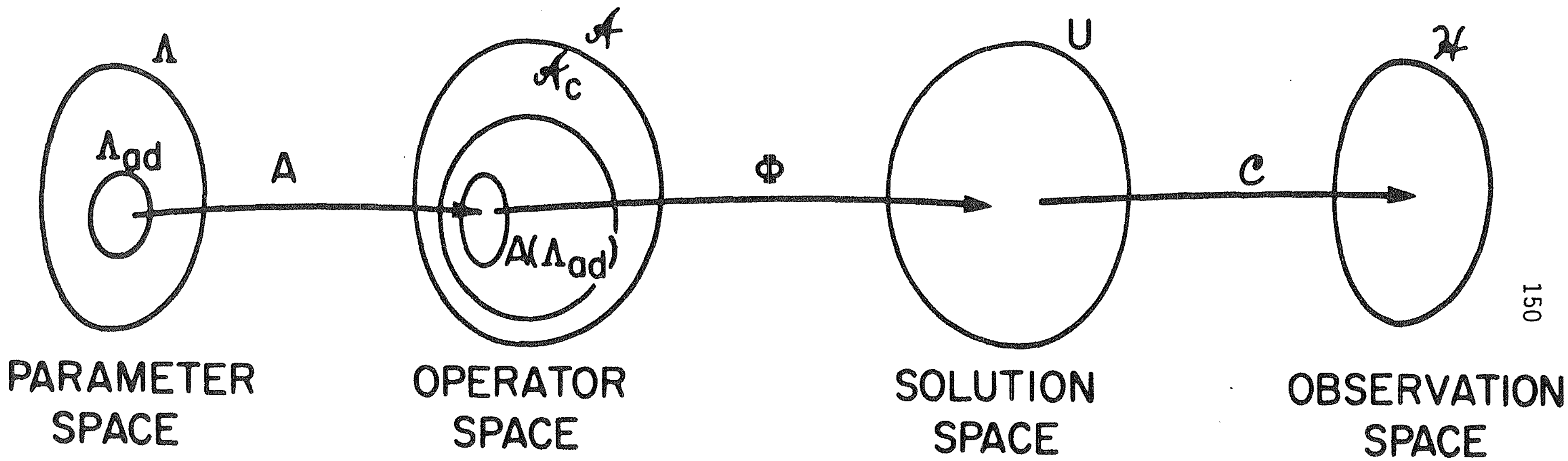

Λ
Λ_{ad}
A
𝒜
$\mathcal{A}_c$
A(Λ_{ad})
Φ
U
𝒞
ℋ
PARAMETER SPACE
OPERATOR SPACE
SOLUTION SPACE
OBSERVATION SPACE

Figure 1

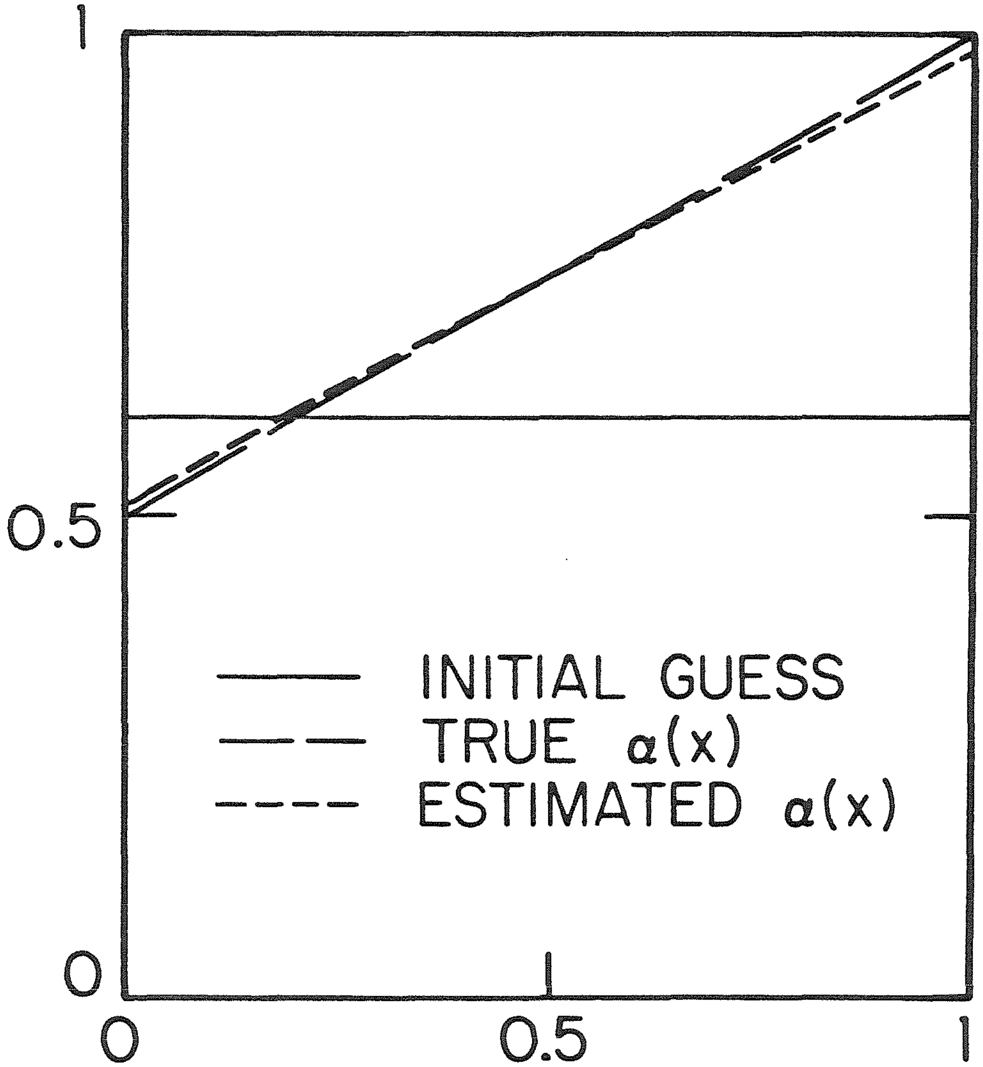


Figure 2

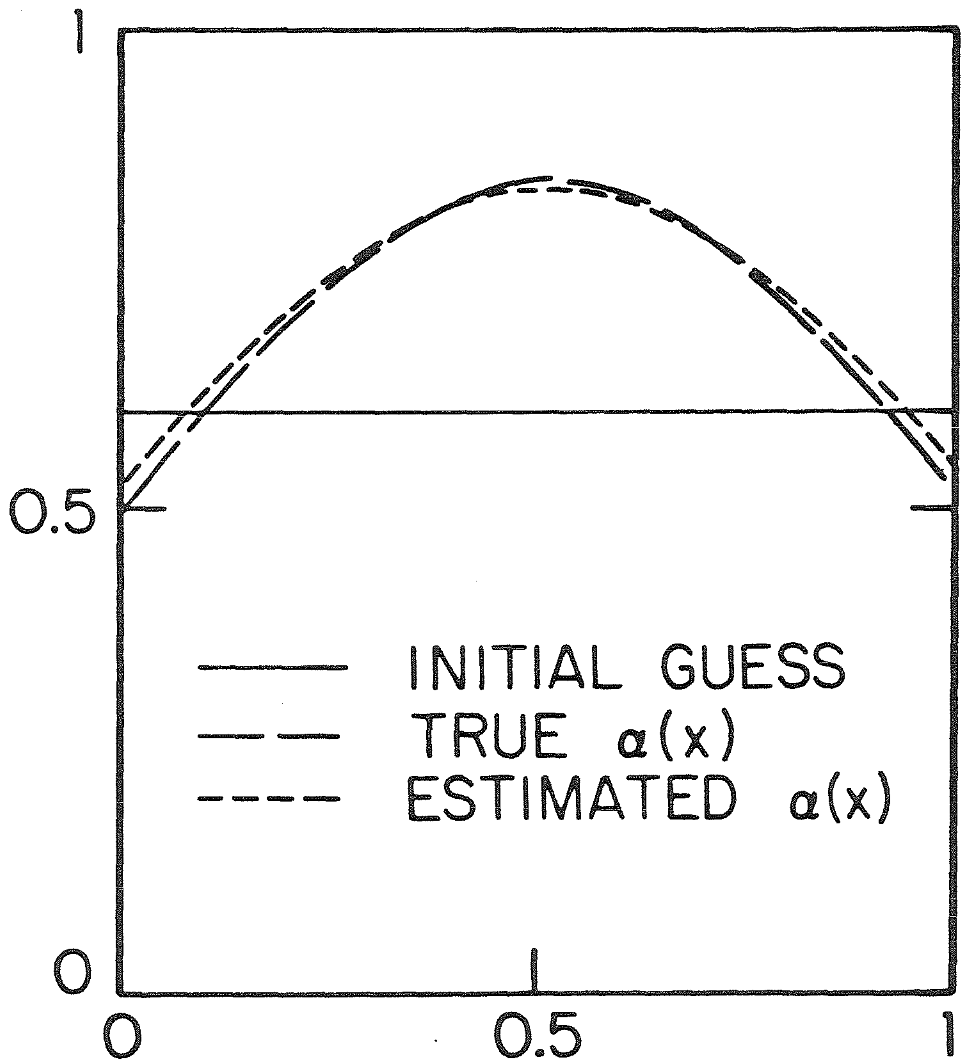


Figure 3

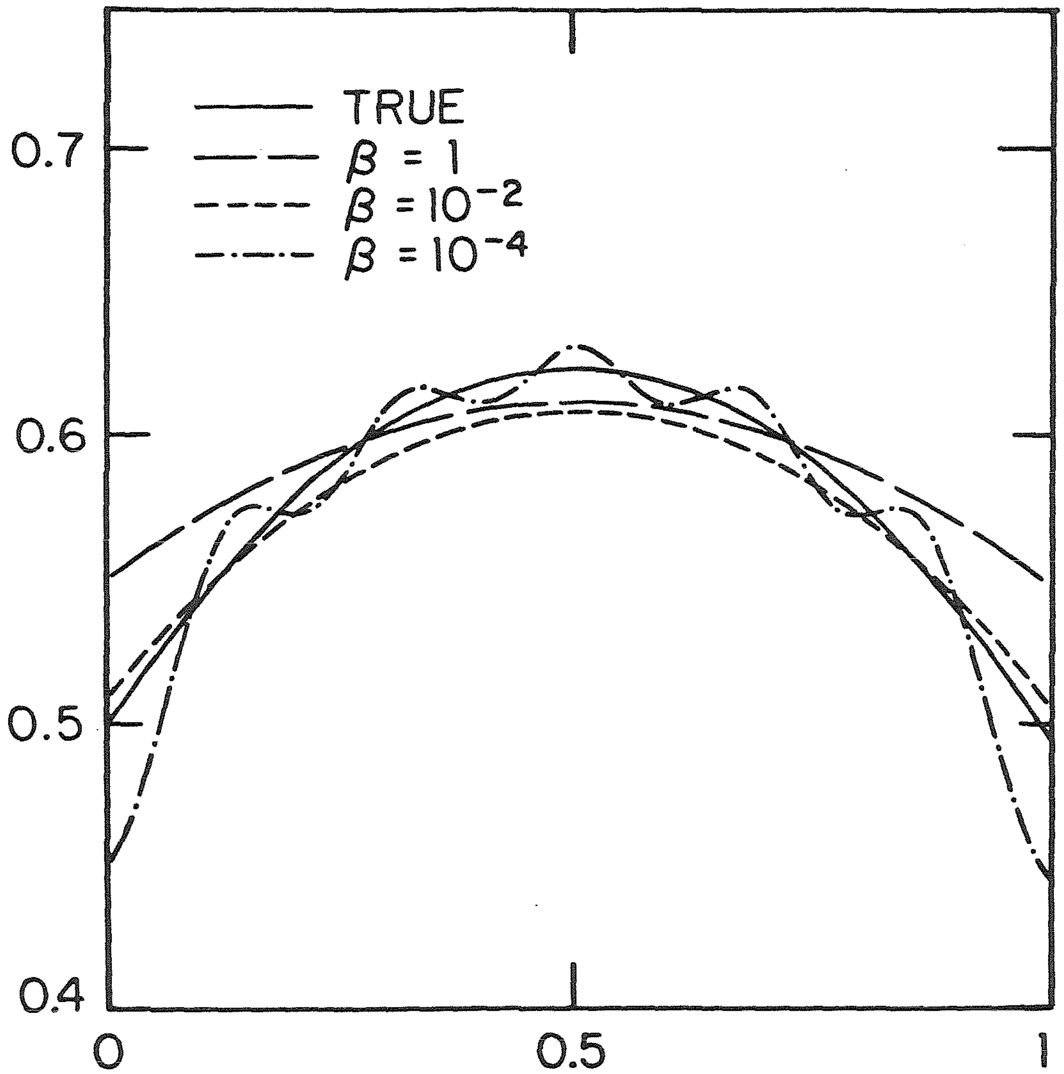


Figure 4

CHAPTER IV: IDENTIFICATION OF SPATIALLY-VARYING PARAMETERS IN DISTRIBUTED PARAMETER SYSTEMS BY DISCRETE REGULARIZATION

1. INTRODUCTION

Consider the following distributed parameter dynamic system:

$$
(1.1)\qquad
\begin{aligned}
&\frac{\partial u}{\partial t} + Au = f &&, \text{ in } \Omega \times \,]0,T[\\
&u(x,0) = u_o &&, \text{ in } \Omega\\
&B_j u = g_j &&, j = 0,\ldots, m-1, \text{ on } \Gamma \times \,]0,T[
\end{aligned}
$$

where $\Omega \subset \mathbb{R}^n$ with boundary Γ and $0 < T < \infty$ and where

$$Au = \sum_{|p|,|q|\le m} (-1)^{|p|} D^p\left(a_{pq}(x) D^q u\right)$$

$$B_j u = \sum_{|h|\le m_j} b_{jh}(x) D^h u \qquad , j = 0,\ldots, m-1$$

with $0 \le m_j$ = order of $B_j \le 2m-1$

The parameter identification problem associated with the above dynamic system can be stated as follows:

Assuming the input function f, the initial condition and the boundary condition(s) to be known, and given an observation of u, determine the system operator A, i.e. the parameters $a_{pq}(x)$.

A number of important physical identification problems fall within the above framework. For example, the partial differential equation

$$(1.2)\quad \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\alpha(x,y)\,\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\alpha(x,y)\,\frac{\partial u}{\partial y}\right) = f(x,y,t)$$

governs the temperature distribution in an inhomogeneous solid or the pressure distribution in a fluid-containing porous medium. In the case of fluid flow in a porous medium, α is termed the transmissivity. For models of petroleum reservoirs and subsurface aquifers the transmissivity is generally inaccessible to direct measurement, and its value must be inferred from measurements of the pressure u at wells.

The key difficulty in developing successful numerical techniques for identifying spatially-dependent parameters in partial differential equations is the fact that such problems are ill-posed in the sense of Hadamard ([4],[5]). In recent work of the authors ([3]), a general regularization identification approach has been developed to overcome this difficulty. In the present work we are concerned with the practical implementation of the regularization approach, i.e. the development of computationally efficient numerical algorithms. In §2 the abstract regularization theory is reviewed and applied to the identification of $\alpha(x,y)$ in (1.2) from point observation. In §3 a numerical method is proposed with a corresponding convergence theorem. The numerical method involves the use of finite-dimensional convergent approximations of Hilbert spaces. §4 presents a review of methods of construction of piecewise-polynomial approximations in the Sobolev spaces $H^m(\Omega)$, which can be used in all practical identification problems. Finally, §5 is concerned with the practical implementation of this discrete regularization approach by studying the identification problem associated with (1.2).

2. PARAMETER IDENTIFICATION BY REGULARIZATION-REVIEW OF BASIC RESULTS

This section consists of a brief necessary review of the general identification approach developed in previous work of the authors ([3]).

Let $\mathscr{A}$, U and F be Banach spaces. $\mathscr{A}$ represents a space of partial differential operators, U represents the space of solutions and F the space of right hand sides. Consider a system described by

$$\Psi(A,u) = f \tag{2.1}$$

where Ψ is a mapping, not necessarily linear, from $\mathscr{A} \times U$ into F. We assume:

(A1) Ψ is of C^k-class $(k \geq 1)$

(A2) There is an open subset $\mathscr{A}_c$ of $\mathscr{A}$ and an open subset U_c of U such that $\forall A \in \mathscr{A}_c$ (2.1) admits a unique solution $u \in U_c$.

(A3) $\forall A \in \mathscr{A}_c \quad \forall u \in U_c \ \frac{\partial \Psi}{\partial u}(A,u)$ is a linear homeomorphism of U onto F.

Thus, one can define an implicit function $u = \Phi(A)$ as the solution of (2.1). Φ is of C^k-class from $\mathscr{A}_c$ into U_c.

Furthermore, consider that A depends on a set of parameters λ belonging to the Banach space Λ. The set of physically admissible λ is Λ_{ad} . We assume:

(A4) $A;\ \Lambda \to \mathscr{A}$ is of C^k-class $(k \geq 1)$

(A5) Λ_{ad} is a norm-closed convex subset of Λ

(A6) $A(\Lambda_{ad}) \subseteq \mathscr{A}_c$

Now the identification problem can be posed as follows:

Knowing the mappings $\Psi;\ \mathscr{A} \times U \to F$ and $A;\ \Lambda \to \mathscr{A}$ and the element $f \in F$ and given an observation of u, find $\lambda \in \Lambda_{ad}$ to satisfy (2.1).

We need to be precise about the nature of the observation of u. Thus, consider a Hilbert space $\mathscr{H}$ (Observation Space). Also, consider an observation operator, not necessarily linear, $\mathscr{C};\ U \to \mathscr{H}$ and assume

(A7) $\mathscr{C}$ is of C^k-class ($k \geq 1$)

Thus, the identification problem can be viewed as solving in Λ_{ad} the (nonlinear) operator equation

$$(2.2) \qquad (\mathscr{C} \circ \Phi \circ A)(\lambda) = z_d$$

If the operator $\mathscr{C} \circ \Phi \circ A$; $\Lambda_{ad} \to \mathscr{H}$ has a unique inverse and the inverse is continuous, one can apply the least-squares method. It consists of minimizing over Λ_{ad} the functional

$$(2.3) \qquad J_{LS}(\lambda) = \| \mathscr{C}\left(\Phi(A(\lambda))\right) - z_d\|^2_{\mathscr{H}}$$

As mentioned in the introduction, the identification of spatially-varying parameters in distributed parameter systems is, as a rule, an ill-posed problem. In other words, the problem of solving (2.2) is ill-posed. Hence, minima of $J_{LS}(\lambda)$ over Λ_{ad} (if any) will not depend continuously on the data z_d.

In order to regularize the parameter λ, we introduce a more regular space $\mathscr{R}$, for which we assume:

(A8) $\mathscr{R}$ is a Hilbert space.

(A9) $\mathscr{R}$ is densely imbedded in Λ.

(A10) The imbedding operator from $\mathscr{R}$ into Λ is compact.

Define $\mathscr{R}_{ad} = \mathscr{R} \cap \Lambda_{ad}$. With (A5) and (A9) is readily follows that $\mathscr{R}_{ad}$ is a norm-closed convex subset of $\mathscr{R}$.

We now introduce the *stabilizing functional*

$$(2.4) \qquad J_S(\lambda) = \| \lambda \|^2_{\mathscr{R}}, \qquad \lambda \in \mathscr{R}_{ad}$$

and the *smoothing functional*

$$J_\beta(\lambda) = J_{LS}(\lambda) + \beta J_S(\lambda)$$

$$= \| \mathscr{C}\Big(\Phi(A(\lambda))\Big) - z_d\|^2_{\mathscr{H}} + \beta\|\lambda\|^2_{\mathscr{R}}, \quad \lambda \in \mathscr{R}_{ad} \tag{2.5}$$

where $\beta > 0$ is the *regularization parameter*. Identification by regularization proceeds as follows. Given $z_d \in \mathscr{H}$ and $\beta > 0$, find $\lambda_\beta \in \mathscr{R}_{ad}$ so as to minimize $J_\beta(\lambda)$.

PROPOSITION 2.1: *The functional* $J_\beta(\lambda)$ *is of* C^k*-class.*

Proof: Immediate consequence of (A1)-(A4), (A7) and (A9).

Remark 2.1: One can derive a formula for the first derivative of $J_\beta(\lambda)$ in terms of an adjoint state. See [3].

THEOREM 2.1: *The functional* $J_\beta(\lambda)$ *admits a global minimum on* $\mathscr{R}_{ad}$.

Proof: Let $\{\lambda_n\}$ be a minimizing sequence and $m = \inf_{\lambda \in \mathscr{R}_{ad}} J_\beta(\lambda)$. Without loss of generality we may assume that

$$\cdots \leq J_\beta(\lambda_{n+1}) \leq J_\beta(\lambda_n) \leq \cdots \leq J_\beta(\lambda_1)$$

Hence, for every $n \in N$,

$$\|\lambda_n\|^2_{\mathscr{R}} \leq \frac{1}{\beta} J_\beta(\lambda_n) \leq \frac{1}{\beta} J_\beta(\lambda_1)$$

i.e. $\{\lambda_n\}$ is norm-bounded in $\mathscr{R}$. Hence, there is a subsequence $\{\lambda_{n_k}\}$ that converges in the weak topology of $\mathscr{R}$ to some $\bar{\lambda} \in \mathscr{R}$. Since $\mathscr{R}_{ad}$ is norm-closed and convex , it is also weakly closed and hence $\bar{\lambda} \in \mathscr{R}_{ad}$. Due to (A10), $\{\lambda_{n_k}\}$ converges to $\bar{\lambda}$ in the norm topology of Λ as well. Finally, using the continuity of the functional $J_{LS}(\lambda)$ in the norm-topology of Λ and the weak lower semicontinuity of $J_S(\lambda)$ in $\mathscr{R}$, it is not difficult to see that $m = J_\beta(\bar{\lambda})$.

The next theorem establishes that minima of J_β depend continuously on the observation. This is the key result of the regularization approach. Roughly speaking, what the next theorem says is the following:

Let $\tilde{\lambda}$ be the "true" value of the parameter and $\tilde{z}_d = \mathscr{C}(\Phi(A(\lambda)))$, what we would have observed with a zero-error observation. Provided that

(i) $\tilde{\lambda}$ is the unique preimage of z_d

(ii) β is an appropriately chosen function of the observation error,

any minimum of $J_\beta(\lambda)$ converges (in the norm of Λ) to $\tilde{\lambda}$, as the observation error tends (in the norm of $\mathscr{H}$) to zero.

THEOREM 2.2: *For any* $\beta > 0$ *and* $z_d \in \mathscr{H}$, *denote by* $\lambda_\beta \in \mathscr{R}_{ad}$ *any minimum of* $J_\beta(\lambda)$ *on* $\mathscr{R}_{ad}$. *Also, denote by* T_{δ_1} *the class of functions that are nonnegative, nondecreasing and continuous on the interval* $[0,\delta_1]$. *Suppose*

$$\begin{cases} \tilde{z}_d \in \mathscr{H} \\ \exists \text{ a unique } \tilde{\lambda} \in \mathscr{R}_{ad} \text{ with } \tilde{z}_d = \mathscr{C}(\Phi(A(\tilde{\lambda}))) \end{cases}$$

Then $\forall \varepsilon > 0 \qquad \forall B_1, B_2 \in T_{\delta_1}$ *with*

$$\begin{cases} B_2(0) = 0 \\ \dfrac{\delta^2}{B_1(\delta)} \le B_2(\delta) \end{cases}$$

$\exists \delta_0(\varepsilon, B_1, B_2) \le \delta_1$ *such that* $\forall z_d \in \mathscr{H} \quad \forall \delta \le \delta_0$

$$\| z_d - \tilde{z}_d \|_{\mathscr{H}} \le \delta \Longrightarrow \| \lambda_\beta - \tilde{\lambda} \|_\Lambda \le \varepsilon$$

for all β *satisfying* $\dfrac{\delta^2}{B_1(\delta)} \le \beta \le B_2(\delta)$

Proof. See [3].

The regularization parameter β can be selected as a function of an upper bound δ on the observation error (i.e. $\| z_d - \tilde{z}_d \|_{\mathscr{H}} \le \delta$).

In [3] the following methods have been discussed:

Method 1: When an a priori upper bound on $\| \tilde{\lambda} \|_{\mathscr{R}}$ is known, i.e. $\| \tilde{\lambda} \|_{\mathscr{R}} \leq \Delta$, one can choose $\beta(\delta) = (\delta/\Delta)^2$. (When $\mathscr{R}$ is a Sobolev space, $\| \; \|_{\mathscr{R}}$ is a measure of smoothness.)

Method 2: Choose $\beta(\delta)$ so that

$$\| \mathscr{C}(\Phi(A(\lambda_{\beta(\delta)}))) - z_d \|_{\mathscr{H}} = \delta$$

where $\lambda_{\beta(\delta)}$ minimizes

$$J_\beta(\lambda) = \| \mathscr{C}(\Phi(A(\lambda))) - z_d \|^2_{\mathscr{H}} + \beta(\delta) \| \lambda \|^2_{\mathscr{R}}$$

In order to apply the above abstract regularization identification theory to concrete examples, all that is needed is an appropriate PDE framework that will permit one to select physically meaningful function spaces Λ, $\mathscr{A}$, U, F, $\mathscr{H}$ and sets Λ_{ad}, $\mathscr{A}_c$, U_c so that assumptions (A1)-(A7) are satisfied. Then one can choose $\mathscr{R} = H^m(\Omega)$ (for sufficiently large m) and have (A8)-(A10) be satisfied. (In case that several parameters are to be identified, $\mathscr{R}$ will be a cartesian product of Sobolev spaces). Finally, one will need a convergent numerical method to carry out the minimization of the smoothing functional.

In [3] the selection of function spaces for the identification of second-order parabolic systems (general linear case and a nonlinear example) has been extensively discussed. Here we consider the special case of identification of spatially-varying diffusivity in the diffusion equation from point observations $z_{d_i}(t)$ of $u(x_i, y_i, t)$, $i = 1, \ldots, \mu$.

$$(2.6)\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial}{\partial x}\ \alpha(x,y)\ \dfrac{\partial u}{\partial x} + \dfrac{\partial}{\partial y}\ \alpha(x,y)\ \dfrac{\partial u}{\partial y} + f(x,y,t) & , \text{ in } Q \\ u(x,y,0) = u_o(x,y) & , \text{ in } \Omega \\ \dfrac{\partial u}{\partial \nu} = 0 & , \text{ in } \Sigma \end{cases}$$

where Ω is a bounded open subset of $\mathbb{R}^2$ with boundary Γ,

$Q = \Omega \times]0,T[$, $\Sigma = \Gamma \times]0,T[$ with $0 < T < \infty$.

Following [3], we can choose

$$\Lambda = C^1(\bar{\Omega})$$

$$\Lambda_{ad} = \{\alpha \in \Lambda / \alpha(x,y) \geq \alpha_o > 0 \qquad \forall (x,y) \in \Omega\}$$

$$\mathscr{A} = \left\{A \in \mathscr{L}(H^{2,1}(Q), L^2(Q)) \ / \ A = \frac{\partial}{\partial x}\left(\alpha(x,y)\ \frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left(\alpha(x,y)\ \frac{\partial}{\partial y}\right) \text{ where } \alpha \in C^1(\bar{\Omega})\right\}$$

$$\mathscr{A}_c = \left\{A \in \mathscr{A} / \ \{A - e^{i\theta}\ \frac{\partial^2}{\partial z^2}\ ,\ \frac{\partial}{\partial \nu}\} \text{ is a regular elliptic system on } \bar{\Omega} \times \mathbb{R} \quad \forall\ \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right\}$$

$$U = U_c = H^{2,1}(Q)$$

$$\mathscr{H} = (L^2(0,T))^{\mu}$$

$$\mathscr{R} = H^3(\Omega)$$

and assumptions (A1)-(A10) are satisfied. Thus, one can identify $\alpha(x,y)$ by minimizing the smoothing functional

$$(2.7) \quad J_\beta(\alpha) = \sum_{i=1}^{\mu} \int_0^T [u(x_i,y_i,t;\alpha) - z_{d_i}]^2\, dt + \beta \|\alpha\|^2_{H^3(\Omega)}$$

over the set

$$\mathscr{R}_{ad} = \{\alpha \in H^3(\Omega)/\alpha(x,y) \geq \alpha_o > 0 \qquad \forall (x,y) \in \Omega \}$$

3. DISCRETIZED MINIMIZATION OF THE SMOOTHING FUNCTIONAL

In this section we are concerned with the numerical minimization of the smoothing functional $J_\beta(\lambda)$. Since $J_\beta(\lambda)$ is differentiable, a natural approach would be to use classical Banach space gradient methods. This has been proposed in previous work of the authors ([3]) and implemented in a numerical example. Such an approach is computationally quite time-consuming, since it involves simultaneous solution of three coupled PDE's in each iteration: the state equation, the adjoint equation and an equation for the calculation of the gradient. In this work, we consider an alternate numerical approach that is computationally attractive. It involves minimization of $J_\beta(\lambda)$ over an appropriate finite-dimensional subspace of $\mathscr{R}$ (with sufficiently large dimension) to obtain an approximate minimum of $J_\beta(\lambda)$ over $\mathscr{R}$. We will first give the definition of a convergent approximation of a Hilbert space and then state and prove our main approximation theorem.

DEFINITION 3.1 ([1]): Let V be a separable Hilbert space. We define an *approximation* (V_N, p_N, r_N) associated with a parameter $N \in \mathbb{N}$ tending to infinity by the following:

(i) V_N is a Hilbert space

(ii) p_N is an isomorphism[1] from V_N onto its closed range P_N in V

(iii) r_N is a linear operator from V onto V_N. We name

V_N the *discrete space*

p_N the *prolongation*

r_N the *restriction*

P_N the *space of approximants*

(1) i.e. linear homeomorphism

DEFINITION 3.2 ([1]): The approximation (V_N, p_N, r_N) is said to be *convergent* if

$$\lim_{N\to\infty} \| v - p_N r_N v \|_V = 0 \qquad \forall v \in V$$

Remark 3.1: In most applications, V_N will be the finite-dimensional space $\mathbb{R}^N$. This justifies the nomenclature "discrete space."

Remark 3.2: The fact that p_N is an isomorphism from V_N onto its closed range in V implies that p_N is a left-invertible operator. In fact, in most practical situations, it makes intuitive sense to select r_N as a left inverse of p_N, i.e. $r_N p_N = 1$. In this case $p_N r_N$ is a projection operator from V onto P_N.

Remark 3.3: Sometimes it is convenient to consider convergent approximations (V_h, p_h, r_h) associated with a small parameter h converging to zero. The corresponding definitions are identical to 3.1 and 3.2.

THEOREM 3.1: *Let $(\mathscr{R}_N, p_N, r_N)$ be a convergent approximation of $\mathscr{R}$ and λ^* a minimizer of $J_\beta(\lambda)$ over $\mathscr{R}_{ad}$. Assume that $p_N r_N \lambda^* \in \mathscr{R}_{ad}$ for sufficiently large N. If λ_N^* minimizes $J_\beta(\lambda)$ over $\mathscr{R}_{ad} \cap P_N$, then*

$$\lim_{N\to\infty} J_\beta(\lambda_N^*) = J_\beta(\lambda^*)$$

and all weak limit points of $\{\lambda_N^*\}$, at least one of which exists, minimize $J_\beta(\lambda)$ over $\mathscr{R}_{ad}$.

Remark 3.4: As a consequence of assumption (A10), all $\mathscr{R}$-weak limit points of $\{\lambda_N^*\}$ will be Λ-strong limit points.

Remark 3.5: The assumption $p_N r_N \lambda^* \in \mathscr{R}_{ad}$ is a very weak one. It says only that the projection of λ^* on space of approximants P_N has to satisfy the constraints that characterize the set of admissible parameters.

Proof of Theorem 3.1: Existence of a minimum λ^* of $J_\beta(\lambda)$ over $\mathscr{R}_{ad}$ has been established by Theorem 2.1. Existence of a minimum λ_N^* of $J_\beta(\lambda)$ over $\mathscr{R}_{ad} \cap P_N$ can be established by using exactly the same argument and the fact that $\mathscr{R}_{ad} \cap P_N$ is a weakly closed set (both $\mathscr{R}_{ad}$ and P_N are closed and convex).

Now observe that

$$J_\beta(\lambda^*) = \inf_{\lambda \in \mathscr{R}_{ad}} J_\beta(\lambda) \leq \inf_{\bar{\lambda} \in \mathscr{R}_{ad} \cap P_N} J_\beta(\lambda) = J_\beta(\lambda_N^*)$$

and since $p_N r_N \lambda^* \in \mathscr{R}_{ad}$, we have

$$J_\beta(\lambda_N^*) \leq J_\beta(p_N r_N \lambda^*)$$

Defining

$$\delta_N = J_\beta(p_N r_N \lambda^*) - J_\beta(\lambda^*)$$

we can combine the above inequalities as follows:

$$J_\beta(\lambda^*) \leq J_\beta(\lambda_N^*) \leq J_\beta(\lambda^*) + \delta_N$$

Now due to the fact that $J_\beta(\lambda)$ is continuous is the norm-topology of $\mathscr{R}$ and that $\lim_{N\to\infty} \| \lambda - p_N r_N \lambda \|_{\mathscr{R}} = 0$ we conclude that $\lim_{N\to\infty} \delta_N = 0$.

So: $\quad J_\beta(\lambda^*) = \lim_{N\to\infty} J_\beta(\lambda_N^*)$

Now δ_N is a sequence of positive numbers converging to zero. Without loss of generality we can consider it to be decreasing

$$0 < \cdots \leq \delta_N \leq \cdots \leq \delta_2 \leq \delta_1$$

Hence

$$J_\beta(\lambda_N^*) \leq \delta_N + J_\beta(\lambda^*) \leq \delta_1 + J_\beta(\lambda^*)$$

$$\Longrightarrow \| \lambda_N^* \|_{\mathscr{R}} \leq \left[\frac{1}{\beta} (\delta_1 + J_\beta(\lambda^*))\right]^{\frac{1}{2}}$$

i.e. $\{\lambda_N^*\}$ is a norm-bounded sequence. Thus, it will have a weak limit point $\hat{\lambda}$, i.e. there will be a subsequence $\{\lambda_{N_i}^*\}$ such that

$$\lambda_{N_i}^* \xrightarrow{\text{weak top of } \mathscr{R}} \hat{\lambda}$$

Due to the weak closedness of the set $\mathscr{R}_{ad}$, all weak limit points $\hat{\lambda}$ of $\{\lambda_N^*\}$ will be in $\mathscr{R}_{ad}$. Furthermore, due to the weak lower semicontinuity of $J_\beta(\lambda)$ in $\mathscr{R}$,

$$J_\beta(\hat{\lambda}) \leq \liminf_{i\to\infty} J_\beta(\lambda_{N_i}^*) = J_\beta(\lambda^*)$$

Hence $\hat{\lambda}$ minimizes $J_\beta(\lambda)$ over $\mathscr{R}_{ad}$. This completes the proof.

It remains to indicate how to select convergent approximations for the space $\mathscr{R}$. Since for all practical purposes $\mathscr{R}$ will be a Sobolev space $H^m(\Omega)$, the next section is devoted to the construction of piecewise-polynomial convergent approximations of $H^m(\Omega)$.

4. CONVERGENT APPROXIMATIONS OF THE SOBOLEV SPACES $H^m(\Omega)$, $\Omega \subset \mathbb{R}^n$

In this section we present a systematic procedure of constructing convergent approximations of the Sobolev spaces $H^m(\Omega)$ in the sense of definitions 3.1 and 3.2, following [1, Chapters 4, 5]. At first, convergent approximations will be constructed for $H^m(\mathbb{R}^n)$. Using their prolongation and restriction operators, it will then be possible to construct convergent approximations for $H^m(\Omega)$, where $\Omega \subset \mathbb{R}^n$.

Let χ be the characteristic function of $[0,1[$ and denote by χ^{*m} its m-fold convolution, i.e.

$$\underbrace{\chi^{*m} = \chi * \chi * \cdots * \chi}_{\text{m times}}$$

Note that $\chi^{*m}(x)$ are B-spline functions. In fact, a straightforward computation gives:

$$\chi^{*2}(x) = \begin{cases} x, & \text{if } x \in [0,1] \\ 1 - x, & \text{if } x \in [1,2] \\ 0, & \text{otherwise} \end{cases}$$

$$\chi^{*3}(x) = \begin{cases} \frac{x^2}{2}, & \text{if } x \in [0,1] \\ \frac{1}{2} + (x-1) - (x-1)^2, & \text{if } x \in [1,2] \\ \frac{1}{2} - (x-2) + \frac{(x-2)^2}{2}, & \text{if } x \in [2,3] \\ 0, & \text{otherwise} \end{cases}$$

$$\chi^{*4}(x) = \begin{cases} \frac{x^3}{6} & , \text{ if } x \in [0,1] \\ \frac{1}{6} + \frac{x-1}{2} + \frac{(x-1)^2}{2} - \frac{(x-1)^3}{2} & , \text{ if } x \in [1,2] \\ \frac{4}{6} - (x-2)^2 + \frac{(x-2)^3}{2} & , \text{ if } x \in [2,3] \\ \frac{1}{6} - \frac{x-3}{2} + \frac{(x-3)^2}{2} - \frac{(x-3)^3}{6} & , \text{ if } x \in [3,4] \\ 0 & , \text{ otherwise} \end{cases}$$

In general,

$$\chi^{*(m+1)}(x) = \begin{cases} \sum_{j=0}^{m} a_m(0,j) \frac{x^j}{j!} & , \text{ if } x \in [0,1] \\ \sum_{j=0}^{m} a_m(1,j) \frac{(x-1)^j}{j!} & , \text{ if } x \in [1,2] \\ - - - - - - - - - - & \\ \sum_{j=0}^{m} a_m(k,j) \frac{(x-k)^j}{j!} & , \text{ if } x \in [k,k+1] \\ - - - - - - - - - - & \\ \sum_{j=0}^{m} a_m(m,j) \frac{(x-m)^j}{j!} & , \text{ if } x \in [m,m+1] \\ 0 & , \text{ otherwise} \end{cases}$$

where $a_m(k,j) = \sum_{i=0}^{k} (-1)^i \binom{m+1}{i} \frac{(k-i)^{m-j}}{(m-j)!}$

It is also straightforward to verify that the piecewise m-th degree polynomials $\chi^{*(m+1)}(x)$ are C^{m-1} functions.

Now let $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ be a parameter with positive components associated with a grid of points $(j_1 h_1, \ldots, j_n h_n)$ where $j = (j_1, \ldots, j_n)$ ranges over $\mathbb{Z}^n$. We denote by $\ell^2(\mathbb{Z}^n)$ the space of square-summable sequences $v_h = \{v_h^j\}_{j \in \mathbb{Z}^n}$ and define the operator $p_h^m ; \ell^2(\mathbb{Z}^n) \to L^2(\mathbb{R})$ by

(4.1) $$p_h^m v_h = \sum_{j \in \mathbb{Z}^n} v_h^j \chi^{*(m+1)} \left(\frac{x_1}{h_1} - j_1\right) \cdots \chi^{*(m+1)}\left(\frac{x_m}{h_n} - j_n\right)$$

It can be easily seen that for every $v_h \in \ell^2(\mathbb{Z}^n)$, $p_h^m v_h \in H^k(\mathbb{R}^n)$ $\forall\, k \leq m$. Moreover, p_h^m is an isomorphism from $\ell^2(\mathbb{Z}^n)$ onto its closed range in $H^k(\mathbb{R}^n)$.

Now let ν be a bounded Radon measure such that (i) ν has compact support and (ii) $\int_{\mathbb{R}^n} d\nu = 1$. Denote by ν_h the measure defined by

$$\int_{\mathbb{R}^n} w(x_1, \ldots, x_n) d\nu_h = \int_{\mathbb{R}^n} w(h_1 x_1, \ldots, h_n x_n) d\nu \quad \forall \text{ continuous function } w$$

With such a measure ν, we associate the restriction $r_h \in \mathscr{L}(L^2(\mathbb{R}^n), \ell^2(\mathbb{Z}^n))$ defined by

$$(r_h v)^j = \int_{\mathbb{R}^n} v(x_1 + j_1 h_1, \ldots, x_n + j_n h_n) d\nu_h \tag{4.2}$$

For example, if $d\nu = \theta_{(0,1)}(x)dx$, where $\theta_{(0,1)}(x)$ is the characteristic function of the unit cube $[0,1[\times \cdots \times [0,1[$ and dx is the Lebesgue measure, then the corresponding restriction is given by

$$(r_h v)^j = \frac{1}{h_1 \ldots h_n} \int_{j_1 h_1}^{(j_1+1)h_1} \cdots \int_{j_n h_n}^{(j_n+1)h_n} v(x_1, \ldots, x_n) dx_1 \ldots dx_n$$

As a second example, we mention that if ν is the Dirac measure $\delta(0)$, then the corresponding restriction is given by

$$(r_h v)^j = v(j_1 h_1, \ldots, j_n h_n)$$

PROPOSITION 4.1 ([1]): *When* p_h^m *and* r_h *are defined by* (4.1) *and* (4.2), *the triples* $(\ell^2(\mathbb{Z}^n), p_h^m, r_h)$ *are convergent approximations of* $H^k(\mathbb{R}^n)$ $\forall\, k \leq m$:

$$\lim_{h \to 0} \|v - p_h^m r_n v\|_{H^k(\mathbb{R}^n)} = 0 \tag{4.3}$$

Moreover, if r_h *is a left inverse of* p_h^m *(i.e.* $r_h p_h^m = 1$*), the following estimate holds:*

$$(4.4) \qquad \|v - p_h^m r_n v\|_{H^k(\mathbb{R}^n)} \leq c\ |h|^{m+1-k}\ \|v\|_{H^{m+1}(\mathbb{R}^n)} \qquad \forall\ k \leq m$$

where $|h| = \max(h_1, \ldots, h_n)$ *and* c *is a constant independent of* h.

We can now proceed to define convergent approximations of $H^m(\Omega)$, where Ω is a bounded open subset of $\mathbb{R}^n$.

Let ρ be the restriction operator that assocites with a function $v \in H^m(\mathbb{R}^n)$ its restriction $\rho u = u)_\Omega$ to Ω. ρ is a bounded linear operator from $H^m(\mathbb{R}^n)$ *onto* $H^m(\Omega)$. Hence, there exists a continuous right inverse π of ρ:

$$(4.5) \qquad \begin{aligned} &\pi \in \mathscr{L}(H^m(\Omega),\ H^m(\mathbb{R}^n)) \\ &\rho\pi v = v \qquad \forall\ v \in H^m(\Omega) \end{aligned}$$

It is not difficult (see [2]) to construct such an extension operator π. Besides (4.5), one can select π to satisfy

$$\rho\, D^k \pi\, v = D^k v \qquad \forall\ k : |k| \leq m$$

Now we define the diserete space

$$(4.6) \qquad H_h(\Omega) = \text{space of finite sequences } v_h = \{v_h^j\}_{j \in G_h(\Omega)}$$

where $G_h(\Omega)$ is the finite grid

$$(4.7) \quad G_h(\Omega) = \{j \in \mathbb{Z}^n / \Omega \cap \text{supp}[\chi^{*(m+1)}(\frac{x_1}{h_1} - j_1) \cdots \chi^{*(m+1)}(\frac{x_n}{h_n} - j_n)] \neq \emptyset\}$$

Clearly, $H_h(\Omega)$ is the finite-dimensional space $\mathbb{R}^{N(h)}$ when N(h) is the number of multiintegers belonging to the finite grid $G_h(\Omega)$.

We define

$$(4.8)\qquad p^m_{h,\Omega} = \rho p^m_h$$

i.e. given $v_h \in H_h(\Omega)$

$$(4.9)\qquad p^m_{h,\Omega}\, v_h = \sum_{j \in G_h(\Omega)} v^j_h\, \chi^{*(m+1)}\left(\frac{x_1}{h_1} - j_1\right) \ldots \chi^{*(m+1)}\left(\frac{x_n}{h_n} - j_n\right) \text{ , where } x \in \Omega$$

It is easy to see that $p^m_{h,\Omega} \in \mathscr{L}(H_h(\Omega), H^m(\Omega))$ and that it is an isomorphism of $H_h(\Omega)$ onto its closed range in $H^m(\Omega)$. Furthermore, define

$$(4.10)\qquad r_{h,\Omega} = r_h \pi$$

i.e. given $v \in H^m(\Omega)$,

$$(4.11)\qquad r_{h,\Omega}\, v = \{r_h(\pi v)^j\}_{j \in G_h(\Omega)} \in H_h(\Omega)$$

Now Proposition 4.1 transforms into

PROPOSITION 4.2: *When* $p^m_{h,\Omega}$ *and* $r_{h,\Omega}$ *are defined by* (4.9) *and* (4.11), *the triples* $(H_h(\Omega), p^m_{h,\Omega}, r_{h,\Omega})$ *are convergent approximations of* $H^k(\Omega)$ $\forall\, k \leq m$:

$$(4.12)\qquad \lim_{h \to 0} \|v - p^m_{h,\Omega}\, r_{h,\Omega} v\|_{H^k(\Omega)} = 0$$

Moreover, if $r_{h,\Omega}$ *is a left inverse of* $p^m_{h,\Omega}$, *the following estimate holds:*

$$(4.13)\qquad \|v - p^m_{h,\Omega}\, r_{h,\Omega} v\|_{H^k(\Omega)} \leq c\, |h|^{m+1-k} \|v\|_{H^{m+1}(\Omega)} \quad \forall\, k \leq m$$

5. NUMERICAL EXAMPLE

To illustrate the theory developed in §3, we will now apply it to the identification of $\alpha(x,y)$ in (2.6) with $\Omega =]0,a[x]0,b[$ given the point observations

$$(5.1) \qquad z_{d_i}(t) = u(x_i,y_i,t) \qquad , \qquad i = 1, \ldots, \mu.$$

In order to minimize the smoothing functional $J_\beta(\alpha)$ given by (2.7), we will need to define a convergent approximation of the space $\mathscr{R} = H^3(\Omega)$. This can be done by applying the techniques of §4 as follows. Given two positive integers K,L let $h = (h_1,h_2) = (\frac{a}{K}, \frac{b}{L})$

$G_h(\Omega) = \{(j_1,j_2) \in \mathbb{Z}^2 \mid -3 \leq j_1 \leq K-1, -3 \leq j_2 \leq L-1\}$

$H_h(\Omega)$ = space of finite sequences $w = \{w_{j_1,j_2}\}_{(j_1,j_2)\in G_h(\Omega)}$.

Clearly, $H_h(\Omega) = \mathbb{R}^N$, where $N = (K+3)(L+3)$. Now define the prolongation operator

$$p^3_{h,\Omega} ; \mathbb{R}^N \to H^3(\Omega)$$

by

$$(5.2) \qquad p^3_{h,\Omega} w = \sum_{(j_1,j_2)\in G_h(\Omega)} w_{j_1,j_2} \chi^{*4}\left(\frac{x}{h_1} - j_1\right)\chi^{*4}\left(\frac{y}{h_2} - j_2\right)$$

$$= \sum_{j_1=-3}^{K-1} \sum_{j_2=-3}^{L-1} w_{j_1,j_2}\chi^{*4}\left(\frac{x}{h_1} - j_1\right)\chi^{*4}\left(\frac{y}{h_2} - j_2\right)$$

$$= \sum_{k=-1}^{K+1} \sum_{\ell=-1}^{L+1} w_{k\ell}\, B_k(x)\, B_\ell(y)$$

where $B_k(x) = \chi^{*4}\left(\frac{x}{h_1} - k + 2\right)$, $B_\ell(y) = \chi^{*4}\left(\frac{y}{h_2} - \ell + 2\right)$

Clearly, $p^3_{h,\Omega}$ coincides with the one defined by (4.9) when

$m = 3$, $\Omega =]0,a[x]0,b[$, $h = (\frac{a}{K}, \frac{b}{L})$. Note that the corresponding space of approximants $P^3_{h,\Omega} \equiv p^3_{h,\Omega}\mathbb{R}^N$ is simply the subspace of $H^3(\Omega)$ spanned by the functions $B_k(x)\ B_\ell(y)$, $-1 \leq k \leq K+1$, $-1 \leq \ell \leq L+1$.

Also, define the restriction operator $r_{h,\Omega}$ as the one that associates $\alpha \in H^3(\Omega) \rightarrow \{w_{k\ell}\}_{\substack{-1\leq k\leq K+1 \\ -1\leq \ell\leq L+1}} \equiv r_{h,\Omega}\alpha \in \mathbb{R}^N$ where $\{w_{k\ell}\}_{\substack{-1\leq k\leq K+1 \\ -1\leq \ell\leq L+1}}$ is the solution of the following system of $(K+3)(L+3) = N$ linear algebraic equations

$$\begin{aligned}
&\frac{1}{36} w_{k-1,\ell-1} + \frac{1}{9} w_{k,\ell-1} + \frac{1}{36} w_{k+1,\ell-1} \\
&+ \frac{1}{9} w_{k-1,\ell} + \frac{4}{9} w_{k,\ell} + \frac{1}{9} w_{k+1,\ell} \\
&+ \frac{1}{36} w_{k-1,\ell+1} + \frac{1}{9} w_{k,\ell+1} + \frac{1}{36} w_{k+1,\ell+1} = \alpha(kh_1,\ell h_2), \quad \begin{array}{l} 0 \leq k \leq K \\ 0 \leq \ell \leq L \end{array}
\end{aligned}$$

$$(5.3)\quad \begin{aligned}
&- \frac{1}{12h_1} w_{-1,\ell-1} + \frac{1}{12h_1} w_{1,\ell-1} \\
&- \frac{1}{3h_1} w_{-1,\ell} + \frac{1}{3h_1} w_{1,\ell} \\
&- \frac{1}{12h_1} w_{-1,\ell+1} + \frac{1}{12h_1} w_{1,\ell+1} = \frac{\partial\alpha}{\partial x}(0,\ell h_2), \quad 0 \leq \ell \leq L
\end{aligned}$$

$$\begin{aligned}
&- \frac{1}{12h_1} w_{K-1,\ell-1} + \frac{1}{12h_1} w_{K+1,\ell-1} \\
&- \frac{1}{3h_1} w_{K-1,\ell} + \frac{1}{3h_1} w_{K+1,\ell} \\
&- \frac{1}{12h_1} w_{K-1,\ell-1} + \frac{1}{12h_1} w_{K+1,\ell+1} = \frac{\partial\alpha}{\partial x}(a,\ell h_2), \quad 0 \leq \ell \leq L
\end{aligned}$$

$$-\frac{1}{12h_2}w_{k-1,-1} - \frac{1}{3h_2}w_{k,-1} - \frac{1}{12h_2}w_{k+1,-1}$$

$$+\frac{1}{12h_2}w_{k-1,1} + \frac{1}{3h_2}w_{k,1} + \frac{1}{12h_2}w_{k+1,1} = \frac{\partial\alpha}{\partial y}(kh_1,0),\ 0 \le k \le K$$

$$-\frac{1}{12h_2}w_{k-1,L-1} - \frac{1}{3h_2}w_{k,L-1} - \frac{1}{12h_2}w_{k+1,L-1}$$

$$+\frac{1}{12h_2}w_{k-1,L+1} + \frac{1}{3h_2}w_{k,L+1} + \frac{1}{12h_2}w_{k+1,L+1} = \frac{\partial\alpha}{\partial y}(kh_1,b),\ 0 \le k \le K$$

$$\frac{1}{4h_1h_2}w_{-1,-1} - \frac{1}{4h_1h_2}w_{1,-1} - \frac{1}{4h_1h_1}w_{-1,1} + \frac{1}{4h_1h_2}w_{1,1} = \frac{\partial^2_\alpha}{\partial x\partial y}(0,0)$$

$$\frac{1}{4h_1h_2}w_{k-1,-1} - \frac{1}{4h_1h_2}w_{k+1,-1} - \frac{1}{4h_1h_2}w_{k-1,1} + \frac{1}{4h_1h_2}w_{k+1,1} = \frac{\partial^2_\alpha}{\partial x\partial y}(a,0)$$

$$\frac{1}{4h_1h_2}w_{-1,L-1} - \frac{1}{4h_1h_2}w_{1,L-1} - \frac{1}{4h_1h_2}w_{-1,L+1} + \frac{1}{4h_1h_2}w_{1,L+1} = \frac{\partial^2_\alpha}{\partial x\partial y}(0,b)$$

$$\frac{1}{4h_1h_2}w_{k-1,L-1} - \frac{1}{4h_1h_2}w_{k+1,L-1} - \frac{1}{4h_1h_2}w_{k-1,L+1} + \frac{1}{4h_1h_2}w_{k+1,L+1} = \frac{\partial^2_\alpha}{\partial x\partial y}(a,b)$$

It is not difficult to see that the above restriction $r_{h,\Omega}$ is a bounded linear operator from $H^3(\Omega)$ onto $\mathbb{R}^N$ of the form (4.11). Furthermore, it is straightforward to verify that $r_{h,\Omega}$ is a left inverse of $p^3_{h,\Omega}$ i.e. $r_{h,\Omega}\, p^3_{h,\Omega}\, w = w \quad \forall\, w \in \mathbb{R}^N$.

Thus, we now have a finite-dimensional convergent approximation $(\mathbb{R}^N,\ p^3_{h,\Omega},\ r_{h,\Omega})$ of the space $\mathscr{R} = H^3(\Omega)$. Theorem 3.1 suggests minimizing $J_\beta(\alpha)$ over the corresponding space of approximants i.e. the subspace of $H^3(\Omega)$ spanned by $B_k(x)\ B_\ell(y)$, $-1 \le k \le K+1$, $-1 \le \ell \le L+1$. So one can minimize

(5.4) $$J_\beta(w) = J_{Ls}(w) + \beta J_s(W)$$

$$= \sum_{i=1}^{\mu} \int_0^T [u(x_i,y_i,t;w) - z_{d_i}(t)]^2 dt + \beta \left\| \sum_{k=-1}^{K+1} \sum_{\ell=-1}^{L+1} w_{k\ell} B_k(x) B_\ell(y) \right\|^2_{H^3(\Omega)}$$

where $u(x\ ,y\ ,t\ ;w)$ is the solution of

$$
(5.5)\quad \begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{\partial}{\partial x}\left(\sum_{k=-1}^{K+1}\sum_{\ell=-1}^{L+1} w_{k\ell}\, B_k(x)\, B_\ell(y)\, \dfrac{\partial u}{\partial x}\right) + \\ \qquad \dfrac{\partial}{\partial y}\left(\sum_{k=-1}^{K+1}\sum_{\ell=-1}^{L+1} w_{k\ell}\, B_k(x)\, B_\ell(y)\, \dfrac{\partial u}{\partial y}\right) + f(x,y,t) \\ u(x,y,0) = u_0(x,y) \\ \dfrac{\partial u}{\partial x}(0,y,t) = \dfrac{\partial u}{\partial x}(a,y,t) = 0 \\ \dfrac{\partial u}{\partial y}(x,0,t) = \dfrac{\partial u}{\partial y}(x,b,t) = 0 \end{cases}
$$

The minimization of $J_\beta(w)$ can be conveniently carried out via a Newton-like method as follows:

(1) Make an initial guess $\alpha^{(0)}(x,y)$

(2) Calculate $w^{(0)} = r_{h,\Omega}\,\alpha^{(0)}$ i.e. solve the system (5.3) for $\alpha = \alpha^{(0)}$

(3) Calculate $J_s^{(0)} = \left\| \sum_{k=-1}^{K+1}\sum_{\ell=-1}^{L+1} w_{k\ell}^{(0)}\, B_k(x)\, B_\ell(y) \right\|^2_{H^3(\Omega)}$ as well as its gradient $\underset{\sim}{g}_s^{(0)}$ and its Hessian $\underset{\approx}{G}_s^{(0)}$.

(4) Solve (5.5) for $w = w^{(0)}$ and calculate

$$J_{Ls}^{(0)} = \sum_{i=1}^{\mu} \int_0^T [u(x_i,y_i,t;w^{(0)}) - z_{d_i}(t)]^2\, dt$$

as well as its gradient $\underset{\sim}{g}_{Ls}^{(0)}$ and approximate Hessian $\underset{\approx}{G}_{Ls}^{(0)}$.

(5) Set

$$w^{(1)} = w^{(0)} - \gamma(\underset{\approx}{G}_{Ls}^{(0)} + \beta \underset{\approx}{G}_s^{(0)})^{-1}\,(\underset{\sim}{g}_{Ls}^{(0)} + \beta \underset{\sim}{g}_s^{(0)})$$

where γ is a step length parameter chosen by one-dimensional line search (e.g. golden section search).

(6) Calculate $J_s^{(1)} = \| \sum_{k=-1}^{K+1} \sum_{\ell=-1}^{L+1} w_{k\ell}^{(1)} B_k(x) B_\ell(y) \|_{H^3(\Omega)}^2$ as well as its gradient $\underset{\sim}{g}_s^{(1)}$ and its Hessian $\underset{\approx}{G}_s^{(1)}$.

(7) Solve (5.5) for $w = w^{(1)}$ and calculate

$$J_{Ls}^{(1)} = \sum_{i=1}^{\mu} \int_0^T [u(x_i,y_i,t;w^{(1)}) - z_{d_i}(t)]^2 dt$$

as well as its gradient $\underset{\sim}{g}_{Ls}^{(1)}$ and an approximate Hessian $\underset{\approx}{G}_{Ls}^{(1)}$.

(8) If $\left|(J_{Ls}^{(0)} + \beta J_s^{(0)}) - (J_{Ls}^{(1)} + \beta J_s^{(1)})\right| < $ Tolerance, Stop.

Otherwise, set

$$\begin{cases} J_{Ls}^{(0)} = J_{Ls}^{(1)}, & \underset{\sim}{g}_{Ls}^{(0)} = \underset{\sim}{g}_{Ls}^{(1)}, \quad \underset{\approx}{G}_{Ls}^{(0)} = \underset{\approx}{G}_{Ls}^{(1)} \\ J_s^{(0)} = J_s^{(1)}, & \underset{\sim}{g}_s^{(0)} = \underset{\sim}{g}_s^{(1)}, \quad \underset{\approx}{G}_s^{(0)} = \underset{\approx}{G}_s^{(1)} \end{cases}$$

and go to step (5).

Note that the stabilizing functional $J_s(w)$, its gradient $\underset{\sim}{g}_s$ and its Hessian $\underset{\approx}{G}_s$ can be calculated *exactly*, since $J_s(w)$ is a quadratic form. Appendix A gives the appropriate formulae. In Appendix B a procedure for estimating the gradient $\underset{\sim}{g}_{Ls}$ and the approximate Hessian $\underset{\approx}{G}_{Ls}$ is suggested based on the so-called Gauss-Newton approach ([6]).

In our simulation study, we have considered (2.6) with spatial domain $\Omega =]0,12[x]0,9[$, time domain $]0,T[=]0,10[$, input $f(x,t) = 0$ and initial condition $u_0(x,y) = 10y + 10$. The objective was to identify $\alpha(x,y)$ given the observation $z_{d_i}(t) = u(x_i,y_i,t)$ at $\mu = 30$ distinct points of Ω.

At first (2.6) was solved numerically with

$$(5.6) \qquad \alpha(x,y) = 7\left[\exp\left\{-\frac{(x-4)^2+(y-5.5)^2}{7}\right\} + \exp\left\{-\frac{(x-10)^2+(y-5.5)^2}{7}\right\}\right]$$

as "true" α, using the alternating direction implicit scheme with a 13 x 10 spatial grid and timestep $\Delta t = 0.25$. Then data $z_{d_{ij}}$ were generated by adding to

$$u(x_i,y_i,t_j), \quad t_j = 0.5, 1, 1.5, \ldots, 10$$

normally distributed random numbers with zero mean and standard deriva-tion 0.1.

For the approximate minimization of the smoothing functional we have used the bicubic spline approximation of $H^3(\Omega)$ defined by (5.2) and (5.3) with $K = 4$ and $L = 3$. Hence, $h_1 = h_2 = 3$, $N = 42$ and $P^3_{h,\Omega} w = \sum_{k=-1}^{5} \sum_{\ell=-1}^{4} w_{k\ell} B_k(x) B_\ell(y)$.

The smoothing functional

$$(5.7) \qquad J_\beta(w) = \sum_{i=1}^{30} \sum_{j=1}^{20} [u(x_i,y_i,t_j;w) - z_{d_{ij}}]^2 + \beta\left\|\sum_{k=-1}^{5} \sum_{\ell=-1}^{4} w_{k\ell} B_k(x) B_\ell(y)\right\|^2_{H^3(\Omega)}$$

was minimized using the Newton-like algorithm described previously. The test for stopping the iterations was

$$|J_\beta^{(0)} - J_\beta^{(1)}| < 10^{-2}$$

Two alternate norms for $H^3(\Omega)$ have been used

$$(5.8) \qquad \|u\|^2_{H^3} = \int_\Omega \left\{ u^2 + \left(\frac{\partial^3 u}{\partial x^3}\right)^2 + 3\left(\frac{\partial^3 u}{\partial x^2 \partial y}\right)^2 + 3\left(\frac{\partial^3 u}{\partial x \partial y^2}\right)^2 + \left(\frac{\partial^3 u}{\partial y^3}\right)^2 \right\} dxdy$$

and

$$(5.9) \qquad |||u|||^2_{H^3} = \int_\Omega \left\{ u^2 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial^2 u}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 u}{\partial y^2}\right)^2 + \right.$$

$$\left. + \left(\frac{\partial^3 u}{\partial x^3}\right)^2 + 3\left(\frac{\partial^3 u}{\partial x^2 \partial y}\right)^2 + 3\left(\frac{\partial^3 u}{\partial x \partial y^2}\right)^2 + \left(\frac{\partial^3 u}{\partial y^3}\right)^2 \right\} dxdy$$

A close initial guess was first used; it was generated by adding to $r_{h,\Omega}\alpha^{true}$ normally distributed random numbers with zero mean and standard derivation 0.2. Next, the flat initial guess $\alpha(x,y) = 1.7$ was used. Table 1 summarizes the results of these simulations for various values of the regularization parameter β. The corresponding estimated surfaces are depicted in Figures 2-15. The true $\alpha(x,y)$ given by (5.7) is depicted in Figure 1.

Upon a visual examination of Figures 2-15 one clearly observes that as β increases, the "humps" tend to get smoothed out. On the other hand, as β decreases, the estimates become less and less smooth. It is remarkable that despite the low level of discretization ($K = 4$, $L = 3$), anomalies do show up at the edge of the estimated surface for $\beta = 0, 10^{-4}$, as a result of ill-conditioning. It is seen from Table 1 that J_{LS} and βJ_S are of the same order of magnitude when $\beta = 10^{-2}$ or $\beta = 10^{-1}$; to these values of β there correspond estimated surfaces that are neither too flat nor too anomalous.

It is noteworthy that

$$\sum_{i=1}^{30} \sum_{j=1}^{20} [u(x_i, y_i, t_j;\ p_{h,\Omega} r_{h,\Omega} \alpha^{true}) - z_{d_{ij}}]^2 = 126.1$$

and

$$\| p_{h,\Omega} r_{h,\Omega} \alpha^{true} \|_{H^3} = 1185.6$$

$$||| p_{h,\Omega} r_{h,\Omega} \alpha^{true} |||_{H^3} = 1509.1$$

and thus

$$\delta^2 = \begin{pmatrix} \text{Squared derivation of predicted } u(x_i, y_i, t) \\ \text{corresponding to } p_{h,\Omega} r_{h,\Omega} \alpha^{true} \text{ from} \\ \text{observed } u(x_i, y_i, t) \end{pmatrix} = 0(10^2)$$

$$\Delta^2 = (\text{Smoothness of } p_{h,\Omega} r_{h,\Omega} \alpha^{true}) = 0(10^3)$$

which suggests [2] $\beta = (\frac{\delta}{\Delta})^2 = 0(10^{-1})$

---

[2] δ and Δ include here the effect of the discretization on the observation and smoothness respectively. If the level of discretization N were sufficiently large, we could have used δ and Δ in the sense of Method 1 of §2.

Comparison of Figures 2-6 and Figures 7-11, in conjunction with Table 1, indicates the effect of the initial guess on the estimated surfaces. It is seen that J_{Ls}^{final} is rather insensitive to the initial guess; on the other hand J_s^{final} seems to be significantly influenced by the initial guess for low values of β $(0, 10^{-4})$ as a result of ill-conditioning.

For a given value of β, the estimates obtained with $J_s = |||\ |||^2_{H^3}$, are smoother and make a poorer match with the observation than the ones obtained with $J_s = \|\ \|^2_{H^3}$; this is due to the fact that $|||\ |||_{H^3} > \|\ \|_{H^3}$ and thus the smoothing term in J_β is given more weight when $J_s = |||\ |||^2_{H^3}$. However, the effect is not significant in our simulation results. (See Table 1 and compare Figures 7-11 to Figures 12-15).

APPENDIX A: EXACT EVALUATION OF THE STABILIZING FUNCTIONAL, ITS GRADIENT AND ITS HESSIAN.

Consider

$$\text{(A.1)} \qquad J_s(w) = \Big\| \sum_{k=-1}^{K+1} \sum_{\ell=-1}^{L+1} w_{k\ell}\, B_k(x)\, B_\ell(y) \Big\|^2_{H^3(\Omega)}$$

where $\Omega =]0,a[\times]0,b[$.

The most frequently used norms for $H^3(\Omega)$ are of the form

$$\text{(A.2)} \qquad \|u\|^2_{H^3(\Omega)} = \zeta_0 \int_0^a\int_0^b u^2\, dxdy + \zeta_1 \int_0^a\int_0^b \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \right] dxdy$$

$$+ \zeta_2 \int_0^a\int_0^b \left[\left(\frac{\partial^2 u}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 u}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 u}{\partial y^2}\right)^2 \right] dxdy$$

$$+ \zeta_3 \int_0^a\int_0^b \left[\left(\frac{\partial^3 u}{\partial x^3}\right)^2 + 3\left(\frac{\partial^3 u}{\partial x \partial y^2}\right)^2 + \left(\frac{\partial^3 u}{\partial y^3}\right)^2 \right] dxdy$$

with $\zeta_0 > 0$, $\zeta_1 \geq 0$, $\zeta_2 \geq 0$, $\zeta_3 > 0$. Although the exact evaluation of J_s, its gradient $\underset{\sim}{g}_s$ and its Hessian $\underset{\approx}{G}_s$ is possible for other norms as well, the formulae given here are restricted to norms of the form (A.2).

Define the following matrices:

$$\underset{\approx}{A}^{0x} \quad : \quad A^{0x}_{kk'} \quad = \quad \int_0^a B_k(x)B_{k'}(x)dx$$

$$\underset{\approx}{A}^{0y} \quad : \quad A^{0y}_{\ell\ell'} \quad = \quad \int_0^b B_\ell(y)B_{\ell'}(y)dy$$

$$\underset{\approx}{A}^{1x} \quad : \quad A^{1x}_{kk'} \quad = \quad \int_0^a B'_k(x)B'_{k'}(x)dx$$

$$\underset{\approx}{A}^{1y} \quad : \quad A^{1y}_{\ell\ell'} \quad = \quad \int_0^b B'_\ell(y)B'_{\ell'}(y)dy$$

$$\underset{\approx}{A}^{2x} \quad : \quad A^{2x}_{kk'} \quad = \quad \int_0^a B''_k(x)B''_{k'}(x)dx$$

$$\underset{\approx}{A}^{2y} \quad : \quad A^{2y}_{\ell\ell'} \quad = \quad \int_0^b B''_\ell(y)B''_{\ell'}(y)dy$$

$$\underset{\approx}{A}^{3x} \quad : \quad A^{3x}_{kk'} \quad = \quad \int_0^a B'''_k(x)B'''_{k'}(x)dx$$

$$\underset{\approx}{A}^{3y} \quad : \quad A^{3y}_{\ell\ell'} \quad = \quad \int_0^b B'''_\ell(y)B'''_{\ell'}(y)dy$$

The entries of these matrices are given in Tables 2-5.

It is easy to see that $J_s(w)$ is a quadratic form. Furthermore, its hessian $\underset{\approx}{G_s} = \left(\frac{\partial^2 J_s}{\partial w_{k\ell}\partial w_{k'\ell'}}\right)$ is a symmetric matrix with entries

(A.3)
$$\frac{\partial^2 J_s}{\partial w_{k\ell}\partial w_{k'\ell'}} = 2\Big[\varsigma_0 A^{0x}_{kk'}A^{0y}_{\ell\ell'} + \varsigma_1(A^{1x}_{kk'}A^{0y}_{\ell\ell'} + A^{0x}_{kk'}A^{1y}_{\ell\ell'}) + \varsigma_2(A^{2x}_{kk'}A^{0y}_{\ell\ell'} + 2A^{1x}_{kk'}A^{1y}_{\ell\ell'} + A^{0x}_{kk'}A^{2y}_{\ell\ell'}) + \varsigma_3(A^{3x}_{kk'}A^{0y}_{\ell\ell'} + 3A^{2x}_{kk'}A^{1y}_{\ell\ell'} + 3A^{1x}_{kk'}A^{2y}_{\ell\ell'} + A^{0x}_{kk'}A^{3y}_{\ell\ell'})\Big]$$

Thus

(A.4) $$\underset{\sim}{g_s} = \underset{\approx}{G_s}\,\underset{\sim}{w} = \underset{\sim}{w}^T\underset{\approx}{G_s}$$

(A.5) $$J_s = \frac{1}{2}\,\underset{\sim}{w}^T\,\underset{\approx}{G_s}\,\underset{\sim}{w}$$

APPENDIX B : CALCULATION OF THE GRADIENT AND THE APPROXIMATE HESSIAN OF $J_{Ls}(w)$

Consider

(B.1) $$J_{Ls}(w) = \sum_{i=1}^{\mu} \int_0^T [u(x_i,y_i,t;w) - z_{d_i}(t)]^2 dt$$

Since (5.6) is solved numerically, $u(x_i,y_i,t;w)$ is not known at all times but rather at $t = t_j$, where $t_j = \Delta t, 2\Delta t, \ldots, T$. Thus, in practice, one will minimize

(B.2) $$\hat{J}_{Ls}(w) = \sum_{i=1}^{\mu} \sum_{j=1}^{T/\Delta t} [u(x_i,y_i,t_j;w) - z_{d_i}(t_j)]^2$$

$$\approx \frac{1}{\Delta t} \sum_{i=1}^{\mu} \int_0^T [u(x_i,y_i,t;w) - z_{d_i}(t)]^2 dt$$

The gradient $\underset{\sim}{g_s} = (\frac{\partial \hat{J}_{Ls}}{\partial w_{k\ell}})$ and the approximate hessian $\underset{\approx}{G_s} = (\frac{\partial^2 \hat{J}_{Ls}}{\partial w_{k\ell} \partial w_{k'\ell'}})$ can be calculated via the formulae([6]):

(B.3) $$\frac{\partial \hat{J}_{Ls}}{\partial w_{k\ell}} = 2 \sum_{i=1}^{\mu} \sum_{j=1}^{T/\Delta t} [u(x_i,y_i,t_j;w) - z_{d_i}(t_j)] \frac{\partial u(x_i,y_i,t_j;w)}{\partial w_{k\ell}}$$

(B.4) $$\frac{\partial^2 \hat{J}_{Ls}}{\partial w_{k\ell} \partial w_{k'\ell'}} = 2 \sum_{i=1}^{\mu} \sum_{j=1}^{T/\Delta t} \frac{\partial u(x_i,y_i,t_j;w)}{\partial w_{k\ell}} \frac{\partial u(x_i,y_i,t_j;w)}{\partial w_{k'\ell'}}$$

The derivatives $\frac{\partial u(x_i,y_i,t_j,w)}{\partial w_{k\ell}}$ can be approximated by the finite-difference expression

$$\frac{u(x_i,y_i,t_j;w+\Delta w_{k\ell}) - u(x_i,y_i,t_j;w)}{\Delta w_{k\ell}}$$

which requires that the basic program be run once for the base case and then once for each $w_{k\ell}$ where the value $w_{k\ell} + \Delta w_{k\ell}$ is used.

REFERENCES

[1] J. P. AUBIN, *Approximation of Elliptic Boundary - Value Problems,* Wiley-Interscience, New York, 1972.

[2] J. P. AUBIN, *Applied Functional Analysis,* Wiley, New York, 1979.

[3] See Chapter III.

[4] M. M. LAVRENTIEV, V. G. ROMANOV AND V. G. VASILIEV, *Multidimentional Inverse Problems for Differential Equations,* Springer-Verlag, Berlin, 1970.

[5] J. L. LIONS, *Some Aspects of Modeling Problems in Distributed Parameter Systems,* Proceedings of the IFIP Working Conference held in Rome 1976, A. Ruberti, ed., Lecture Notes in Control and Information Sciences Vol. 1, Springer-Verlag, Berlin, 1978, pp. 11-41.

[6] J. H. SEINFELD AND L. LAPIDUS, *Mathematical Methods in Chemical Engineering Volume III: Process Modeling, Estimation and Identification,* Prentice Hall, Englewood Cliffs, N.J., 1974.

Table 1. Simulation Results

Case	J_β initial	J_β final	J_{Ls} final	J_s final	# of iterations
Close Initial Guess(*), $J_s = \| \|^2_{H^3}$					
$\beta = 0$	126.1	29.9	29.9	1155.9	12
$\beta = 10^{-4}$	126.2	30.8	30.6	1156.9	12
$\beta = 10^{-2}$	137.9	45.9	35.6	1029.7	11
$\beta = 10^{-1}$	224.6	130.5	50.6	798.4	6
$\beta = 0.5$	718.9	379.0	82.3	593.4	17
Flat Initial Guess(**), $J_s = \| \|^2_{H^3}$					
$\beta = 0$	17,671.5	33.4	33.4	1270.9	26
$\beta = 10^{-4}$	17,671.5	34.2	34.1	1465.4	22
$\beta = 10^{-2}$	17,674.6	47.8	36.5	1129.4	23
$\beta = 10^{-1}$	17,702.7	150.1	74.1	759.1	15
$\beta = 0.5$	17,827.5	372.3	80.3	584.1	21
Flat Initial Guess(**), $J_s = \|\| \|\|^2_{H^3}$					
$\beta = 0$	17,671.5	33.4	33.4	1790.0	26
$\beta = 10^{-4}$	17,671.5	35.1	34.9	2165.8	22
$\beta = 10^{-2}$	17,674.6	58.4	43.7	1463.3	16
$\beta = 10^{-1}$	17,702.7	159.7	67.6	921.7	20
$\beta = 0.5$	17,827.5	471.6	100.8	741.6	18

(*) $J_{Ls}^{\text{initial}} = 126.1$, $J_s^{\text{initial}} = 1185.6$

(**) $J_{Ls}^{\text{initial}} = 17{,}671.5$, $J_s^{\text{initial}} = 312.1$

Table 2 : Entries of the matrices $\underset{\approx}{A}^{0x}$ and $\underset{\approx}{A}^{oy}$

$\int_0^a [B_{-1}(x)]^2 dx$	$= \frac{h_1}{252}$	$\int_0^b [B_{-1}(y)]^2 dy$	$= \frac{h_2}{252}$
$\int_0^a [B_0(x)]^2 dx$	$= \frac{151\, h_1}{630}$	$\int_0^b [B_0(y)]^2 dy$	$= \frac{151\, h_2}{630}$
$\int_0^a [B_1(x)]^2 dx$	$= \frac{599\, h_1}{1260}$	$\int_0^b [B_1(y)]^2 dy$	$= \frac{599\, h_2}{1260}$
$\int_0^a [B_k(x)]^2 dx$	$= \frac{151\, h_1}{315}$, $2 \leq k \leq K-2$	$\int_0^b [B_\ell(y)]^2 dy$	$= \frac{151\, h_2}{315}$, $2 \leq \ell \leq L-2$
$\int_0^a [B_{K-1}(x)]^2 dx$	$= \frac{599\, h_1}{1260}$	$\int_0^b [B_{L-1}(y)]^2 dy$	$= \frac{599\, h_2}{1260}$
$\int_0^a [B_K(x)]^2 dx$	$= \frac{151\, h_1}{630}$	$\int_0^b [B_L(y)]^2 dy$	$= \frac{151\, h_2}{630}$
$\int_0^a [B_{K+1}(x)]^2 dx$	$= \frac{h_1}{252}$	$\int_0^b [B_{L+1}(y)]^2 dy$	$= \frac{h_2}{252}$
$\int_0^a B_{-1}(x)B_0(x)dx$	$= \frac{43\, h_1}{1680}$	$\int_0^b B_{-1}(y)B_0(y)dy$	$= \frac{43\, h_2}{1680}$
$\int_0^a B_0(x)B_1(x)dx$	$= \frac{59\, h_1}{280}$	$\int_0^b B_0(y)B_1(y)dy$	$= \frac{59\, h_2}{280}$
$\int_0^a B_k(x)B_{k+1}(x)dx$	$= \frac{397\, h_1}{1680}$, $1 \leq k \leq K-2$	$\int_0^b B_\ell(y)B_{\ell+1}(y)dy$	$= \frac{397\, h_2}{1680}$, $1 \leq \ell \leq L-2$
$\int_0^a B_{K-1}(x)B_K(x)dx$	$= \frac{59\, h_1}{280}$	$\int_0^b B_{L-1}(y)B_L(y)dy$	$= \frac{59\, h_2}{280}$
$\int_0^a B_K(x)B_{K+1}(x)dx$	$= \frac{43\, h_1}{1680}$	$\int_0^b B_L(y)B_{L+1}(y)dy$	$= \frac{43\, h_2}{1680}$
$\int_0^a B_{-1}(x)B_1(x)dx$	$= \frac{h_1}{84}$	$\int_0^b B_{-1}(y)B_1(y)dy$	$= \frac{h_2}{84}$
$\int_0^a B_k(x)B_{k+2}(x)dx$	$= \frac{h_1}{42}$, $0 \leq k \leq K-2$	$\int_0^b B_\ell(y)B_{\ell+2}(y)dy$	$= \frac{h_2}{42}$, $0 \leq \ell \leq L-2$
$\int_0^a B_{K-1}(x)B_{K+1}(x)dx$	$= \frac{h_1}{84}$	$\int_0^b B_{L-1}(y)B_{L+1}(y)dy$	$= \frac{h_2}{84}$
$\int_0^a B_k(x)B_{k+3}(x)dx$	$= \frac{h_1}{5040}$, $-1 \leq k \leq K-2$	$\int_0^b B_\ell(y)B_{\ell+3}(y)dy$	$= \frac{h_2}{5040}$, $-1 \leq \ell \leq L-2$
$\int_0^a B_k(x)B_{k'}(x)dx$	$= 0$, if $\lvert k-k' \rvert > 3$	$\int_0^b B_\ell(y)B_{\ell'}(y)dy$	$= 0$, if $\lvert \ell-\ell' \rvert > 3$

Table 3 : Entries of the matrices $\underset{\approx}{A}^{1x}$ and $\underset{\approx}{A}^{1y}$

$\int_0^a [B'_{-1}(x)]^2 dx = \frac{1}{20 h_1}$	$\int_0^b [B'_{-1}(y)]^2 dy = \frac{1}{20 h_2}$
$\int_0^a [B'_0(x)']^2 = \frac{1}{3 h_1}$	$\int_0^b [B'_0(y)]^2 dy = \frac{1}{3 h_2}$
$\int_0^a [B'_1(x)']^2 dx = \frac{37}{60 h_1}$	$\int_0^b [B'_1(y)]^2 dy = \frac{37}{60 h_2}$
$\int_0^a [B'_k(x)]^2 dx = \frac{2}{3 h_1}$, $2 \le k \le K-2$	$\int_0^b [B'_\ell(y)]^2 dy = \frac{2}{3 h_2}$, $2 \le \ell \le L-2$
$\int_0^a [B'_{K-1}(x)]^2 dx = \frac{37}{60 h_1}$	$\int_0^b [B'_{L-1}(y)]^2 dy = \frac{37}{60 h_2}$
$\int_0^a [B'_K(x)]^2 dx = \frac{1}{3 h_1}$	$\int_0^b [B'_L(y)]^2 dy = \frac{1}{3 h_2}$
$\int_0^a [B'_{K+1}(x)]^2 dx = \frac{1}{20 h_1}$	$\int_0^b [B'_{L+1}(y)]^2 dy = \frac{1}{20 h_2}$
$\int_0^a B'_{-1}(x)B'_0(x)dx = \frac{7}{120 h_1}$	$\int_0^b B'_{-1}(y)B'_0(y)dy = \frac{7}{120 h_2}$
$\int_0^a B'_0(x)B'_1(x)dx = -\frac{11}{60 h_1}$	$\int_0^b B'_0(y)B'_1(y)dy = -\frac{11}{60 h_2}$
$\int_0^a B'_k(x)B'_{k+1}(x)dx = -\frac{1}{8 h_1}$, $1 \le k \le K-2$	$\int_0^b B'_\ell(y)B'_{\ell+1}(y)dy = -\frac{1}{8 h_2}$, $1 \le \ell \le L-2$
$\int_0^a B'_{K-1}(x)B'_K(x)dx = -\frac{11}{60 h_1}$	$\int_0^b B'_{L-1}(y)B'_L(y)dy = -\frac{11}{60 h_2}$
$\int_0^a B'_K(x)B'_{K+1}(x)dx = \frac{7}{120 h_1}$	$\int_0^b B'_L(y)B'_{L+1}(y)dy = \frac{7}{120 h_2}$
$\int_0^a B'_{-1}(x)B'_1(x)dx = -\frac{1}{10 h_1}$	$\int_0^b B'_{-1}(y)B'_1(y)dy = -\frac{1}{10 h_2}$
$\int_0^a B'_k(x)B'_{k+2}(x)dx = -\frac{1}{5 h_1}$, $0 \le k \le K-2$	$\int_0^b B'_\ell(y)B'_{\ell+2}(y)dy = -\frac{1}{5 h_2}$, $0 \le \ell \le L-2$
$\int_0^a B'_{K-1}(x)B'_{K+1}(x)dx = -\frac{1}{10 h_1}$	$\int_0^b B'_{L-1}(y)B'_{L+1}(y)dy = -\frac{1}{10 h_2}$
$\int_0^a B'_k(x)B'_{k+3}(x)dx = -\frac{1}{120 h_1}$, $1 \le k \le K-2$	$\int_0^b B'_\ell(y)B'_{\ell+3}(y)dy = -\frac{1}{120 h_2}$, $-1 \le \ell \le L-2$
$\int_0^a B'_k(x)B'_{k'}(x)dx = 0$, if $\lvert k-k' \rvert > 3$	$\int_0^b B'_\ell(y)B'_{\ell'}(y)dy = 0$, if $\lvert \ell-\ell' \rvert > 3$

Table 4 : Entries of the matrices $\underset{\approx}{A}^{2x}$ and $\underset{\approx}{A}^{2y}$

$\int_0^a [B''_{-1}(x)]^2 dx = \frac{1}{3h_1^3}$	$\int_0^b [B''_{-1}(y)]^2 dy = \frac{1}{3h_2^3}$
$\int_0^a [B''_0(x)]^2 dx = \frac{4}{3h_1^3}$	$\int_0^b [B''_0(y)]^2 dy = \frac{4}{3h_2^3}$
$\int_0^a [B''_1(x)]^2 dx = \frac{7}{3h_1^3}$	$\int_0^b [B''_1(y)]^2 dy = \frac{7}{3h_2^3}$
$\int_0^a [B''_k(x)]^2 dx = \frac{8}{3h_1^3}$, $2\leq k\leq K-2$	$\int_0^b [B''_\ell(y)]^2 dy = \frac{8}{3h_2^3}$, $2\leq \ell\leq L-2$
$\int_0^a [B''_{K-1}(x)]^2 dx = \frac{7}{3h_1^3}$	$\int_0^b [B''_{L-1}(y)]^2 dy = \frac{7}{3h_2^3}$
$\int_0^a [B''_K(x)]^2 dx = \frac{4}{3h_1^3}$	$\int_0^b [B''_L(y)]^2 dy = \frac{4}{3h_2^3}$
$\int_0^a [B''_{K+1}(x)]^2 dx = \frac{1}{3h_1^3}$	$\int_0^b [B''_{L+1}(y)]^2 dy = \frac{1}{3h_2^3}$
$\int_0^a B''_{-1}(x)B''_0(x)dx = -\frac{1}{2h_1^3}$	$\int_0^b B''_{-1}(y)B''_0(y)dy = -\frac{1}{2h_2^3}$
$\int_0^a B''_0(x)B''_1(x)dx = -\frac{1}{h_1^3}$	$\int_0^b B''_0(y)B''_1(y)dy = -\frac{1}{h_2^3}$
$\int_0^a B''_k(x)B''_{k+1}(x)dx = -\frac{3}{2h_1^3}$, $1\leq k\leq K-1$	$\int_0^b B''_\ell(y)B''_{\ell+1}(y)dy = -\frac{3}{2h_2^3}$, $1\leq \ell\leq L-1$
$\int_0^a B''_{K-1}(x)B''_K(x)dx = -\frac{1}{h_1^3}$	$\int_0^b B''_{L-1}(y)B''_L(y)dy = -\frac{1}{h_2^3}$
$\int_0^a B''_K(x)B''_{K+1}(x)dx = -\frac{1}{2h_1^3}$	$\int_0^b B''_L(y)B''_{L+1}(y)dy = -\frac{1}{2h_2^3}$
$\int_0^a B''_{-1}(x)B''_1(x)dx = 0$	$\int_0^b B''_{-1}(y)B''_1(y)dy = 0$
$\int_0^a B''_k(x)B''_{k+2}(x)dx = 0$, $0\leq k\leq K-2$	$\int_0^b B''_\ell(y)B''_{\ell+2}(y)dy = 0$, $0\leq \ell\leq L-2$
$\int_0^a B''_{K-1}(x)B''_{K+1}(x)dx = 0$	$\int_0^b B''_{L-1}(y)B''_{L+1}(y)dy = 0$
$\int_0^a B''_k(x)B''_{k+3}(x)dx = \frac{1}{6h_1^3}$, $-1\leq k\leq K-2$	$\int_0^b B''_\ell(y)B''_{\ell+3}(y)dy = \frac{1}{6h_2^3}$, $1\leq \ell\leq L-2$
$\int_0^a B''_k(x)B''_{k'}(x)dx = 0$, if $\lvert k-k'\rvert > 3$	$\int_0^b B''_\ell(y)B''_{\ell'}(y)dy = 0$, if $\lvert \ell-\ell'\rvert > 3$

Table 5 : Entries of the matrices $\underset{\approx}{A}^{3x}$ and $\underset{\approx}{A}^{3y}$

$\int_0^a [B'''_{-1}(x)]^2 dx = \frac{1}{h_1^5}$	$\int_0^b [B'''_{-1}(y)]^2 dy = \frac{1}{h_2^5}$
$\int_0^a [B'''_0(x)]^2 dx = \frac{10}{h_1^5}$	$\int_0^b [B'''_0(y)]^2 dy = \frac{10}{h_2^5}$
$\int_0^a [B'''_1(x)]^2 dx = \frac{19}{h_1^5}$	$\int_0^b [B'''_1(y)]^2 = \frac{19}{h_2^5}$
$\int_0^a [B'''_k(x)]^2 dx = \frac{20}{h_1^5},\ 2\le k\le K-2$	$\int_0^b [B'''_\ell(y)]^2 dy = \frac{20}{h_2^5},\ 2\le \ell\le L-2$
$\int_0^a [B'''_{K-1}(x)]^2 dx = \frac{19}{h_1^5}$	$\int_0^b [B'''_{L-1}(y)]^2 dy = \frac{19}{h_2^5}$
$\int_0^a [B'''_K(x)]^2 dx = \frac{10}{h_1^5}$	$\int_0^b [B'''_L(y)]^2 dy = \frac{10}{h_2^5}$
$\int_0^a [B'''_{K+1}(x)]^2 dx = \frac{1}{h_1^5}$	$\int_0^b [B'''_{L+1}(y)]^2 dy = \frac{1}{h_2^5}$
$\int_0^a B'''_{-1}(x)B'''_0(x)dx = -\frac{3}{h_1^5}$	$\int_0^b B'''_{-1}(y)B'''_0(y)dy = -\frac{3}{h_2^5}$
$\int_0^a B'''_0(x)B'''_1(x)dx = -\frac{12}{h_1^5}$	$\int_0^b B'''_0(y)B'''_1(y)dy = -\frac{12}{h_2^5}$
$\int_0^a B'''_k(x)B'''_{k+1}(x)dx = -\frac{15}{h_1^5},\ 1\le k\le K-2$	$\int_0^b B'''_\ell(y)B'''_{\ell+1}(y)dy = -\frac{15}{h_2^5},\ 1\le \ell\le L-2$
$\int_0^a B'''_{K-1}(x)B'''_K(x)dx = -\frac{12}{h_1^5}$	$\int_0^b B'''_{L-1}(y)B'''_L(y)dt = -\frac{12}{h_2^5}$
$\int_0^a B'''_K(x)B'''_{K+1}(x)dx = -\frac{3}{h_1^5}$	$\int_0^b B'''_L(y)B'''_{L+1}(y)dy = -\frac{3}{h_2^5}$
$\int_0^a B'''_{-1}(x)B'''_1(x)dx = \frac{3}{h_1^5}$	$\int_0^b B'''_{-1}(y)B'''_1(y)dy = \frac{3}{h_2^5}$
$\int_0^a B'''_k(x)B'''_{k+2}(x)dx = \frac{6}{h_1^5},\ 0\le k\le K-2$	$\int_0^b B'''_\ell(y)B'''_{\ell+2}(y)dy = \frac{6}{h_2^5},\ 0\le \ell\le L-2$
$\int_0^a B'''_{K-1}(x)B'''_{K+1}(x)dx = \frac{3}{h_1^5}$	$\int_0^b B'''_{L-1}(y)B'''_{L+1}(y)dy = \frac{3}{h_2^5}$
$\int_0^a B'''_k(x)B'''_{k+3}(x)dx = -\frac{1}{h_1^5},\ 1<k<K-2$	$\int_0^b B'''_\ell(y)B'''_{\ell+3}(y)dy = -\frac{1}{h_2^5},\ 1<\ell<L-2$
$\int_0^a B'''_k(x)B'''_{k'}(x)dx = 0,\ \text{if } \lvert k-k'\rvert > 3$	$\int_0^b B'''_\ell(y)B'''_{\ell'}(y)dy = 0,\ \text{if } \lvert \ell-\ell'\rvert > 3$

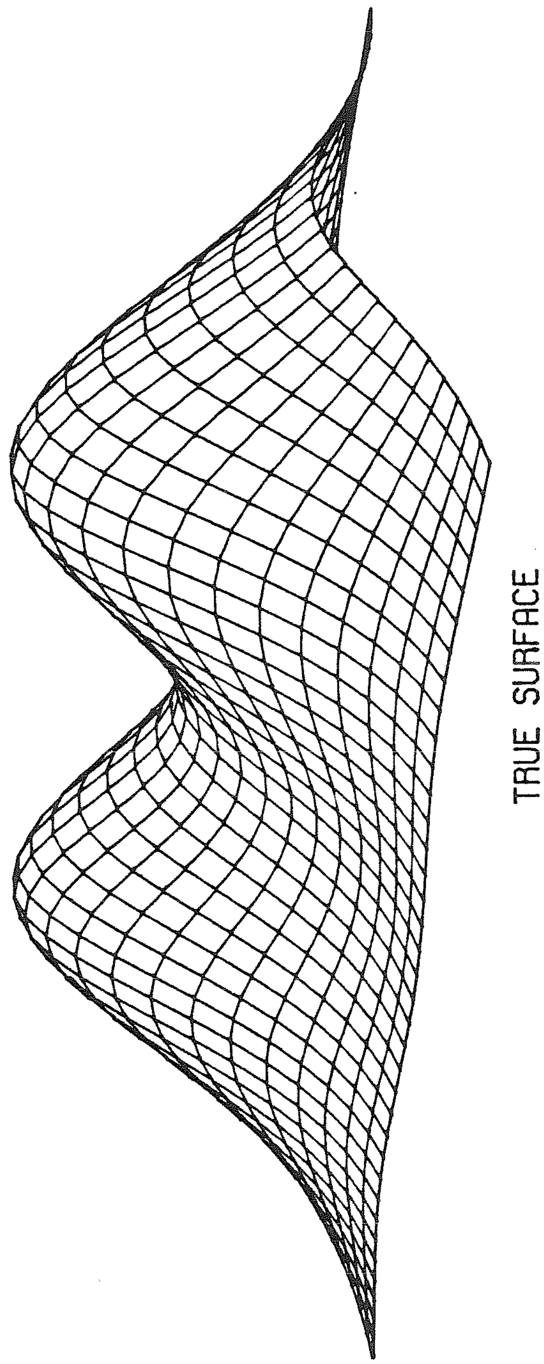
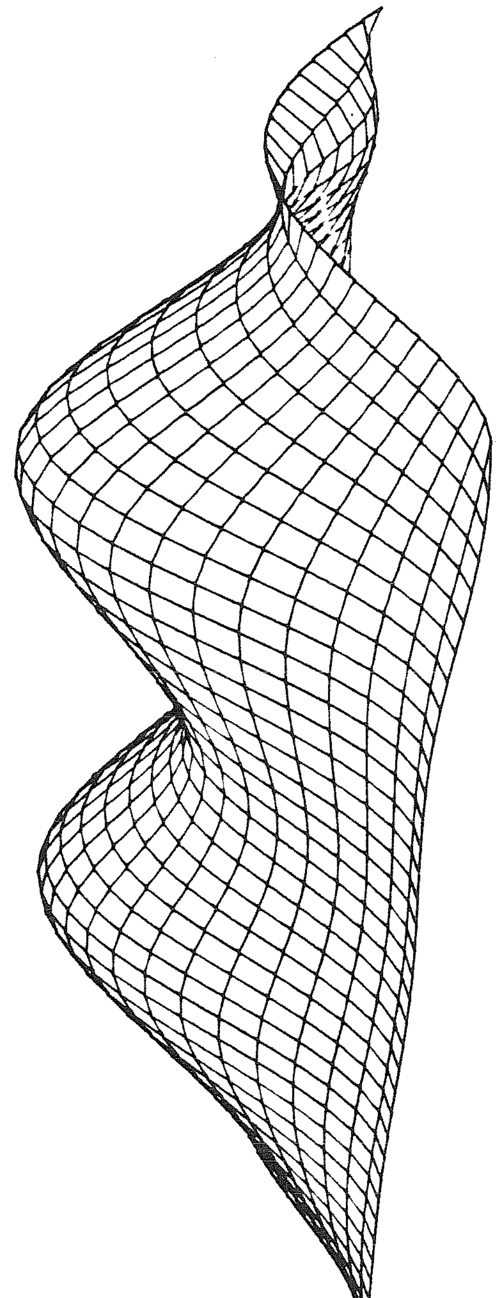


Figure 1



CLOSE INITIAL GUESS , $\beta = 0$

Figure 2

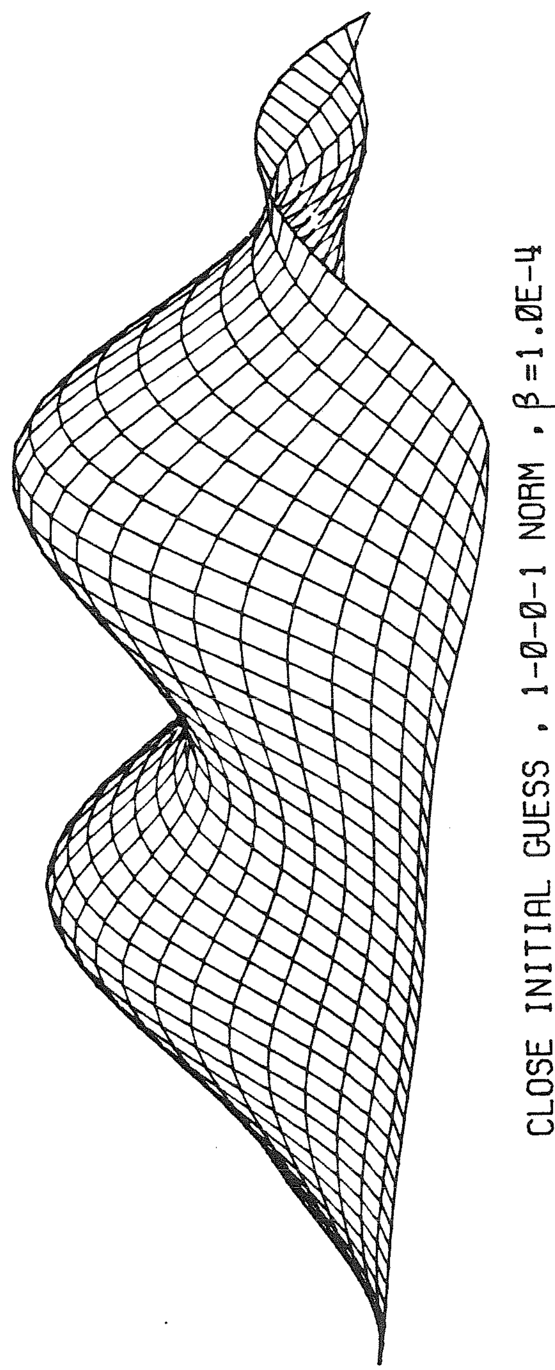


Figure 3

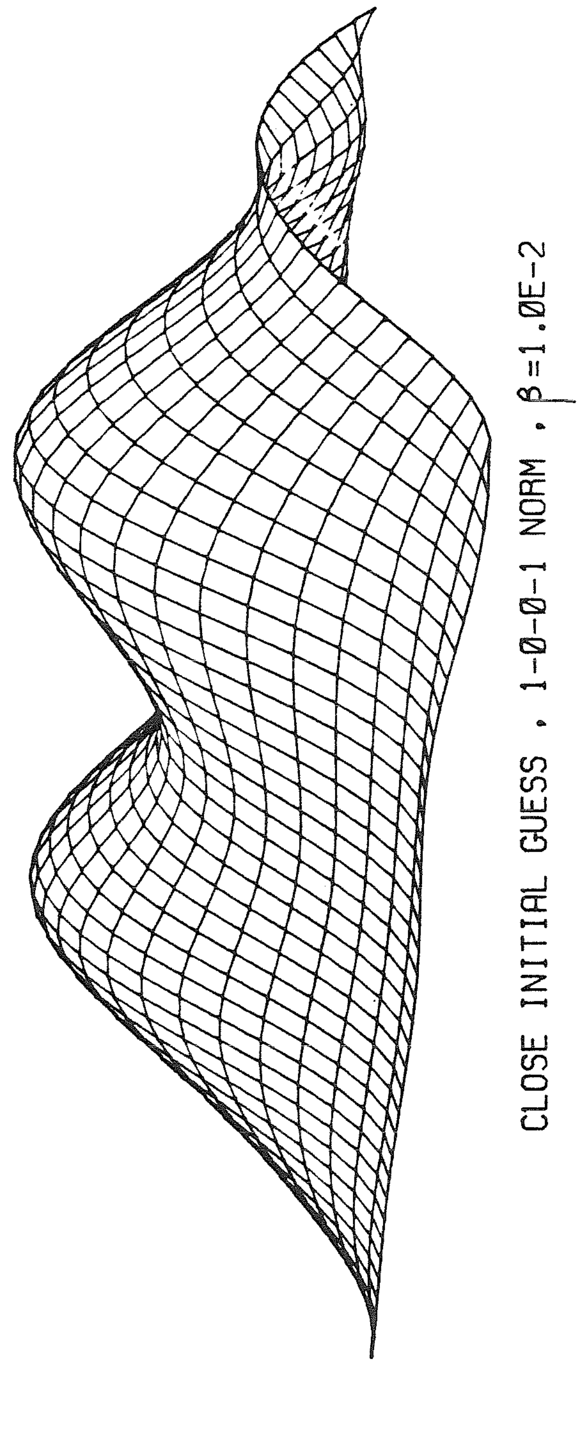


Figure 4

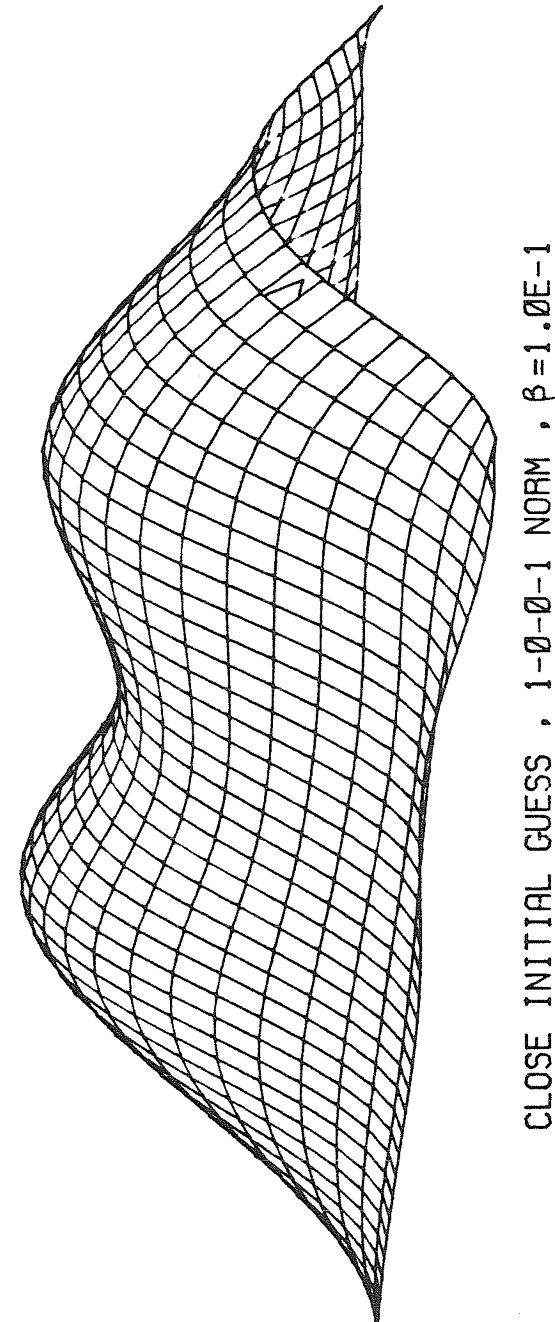
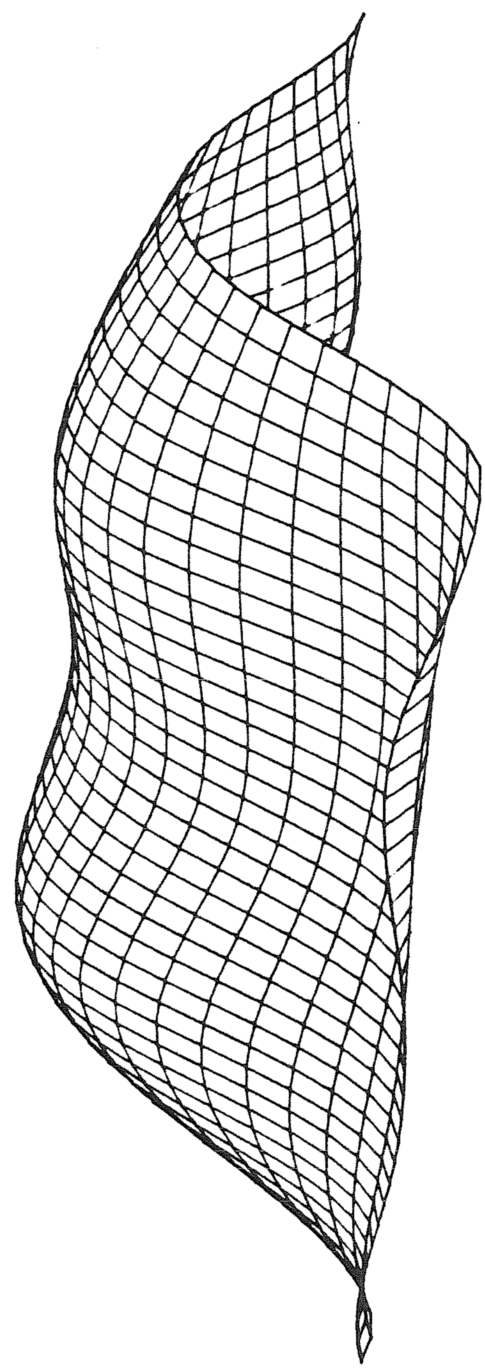


Figure 5



CLOSE INITIAL GUESS . 1-Ø-Ø-1 NORM . β=Ø.5

Figure 6

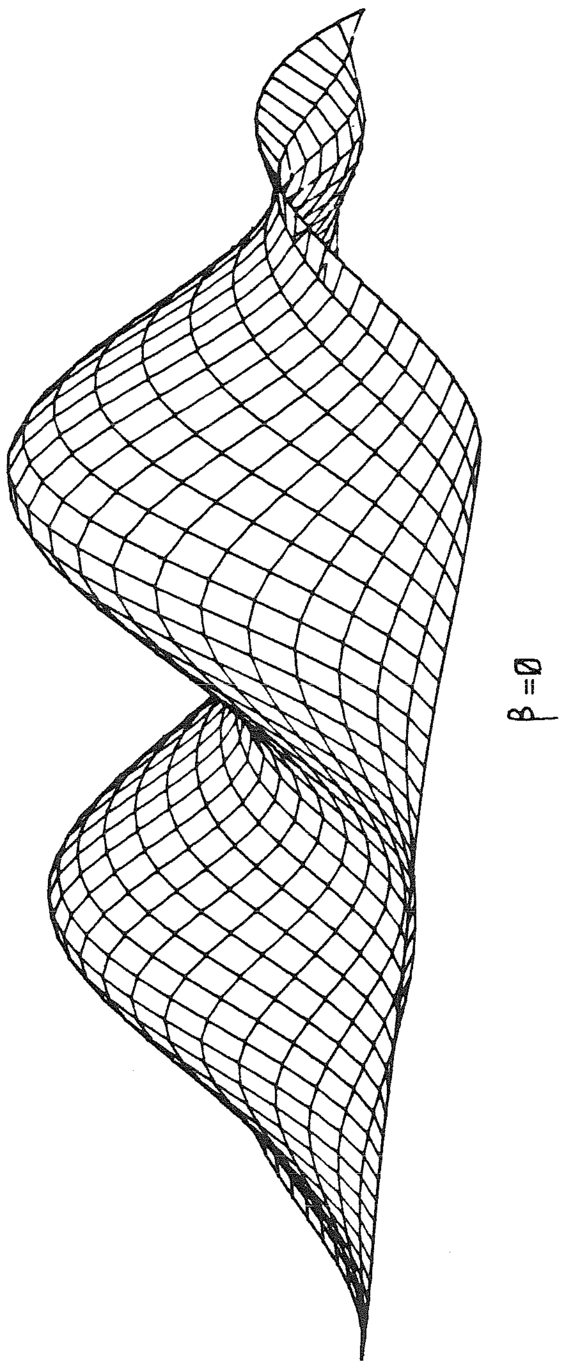


Figure 7

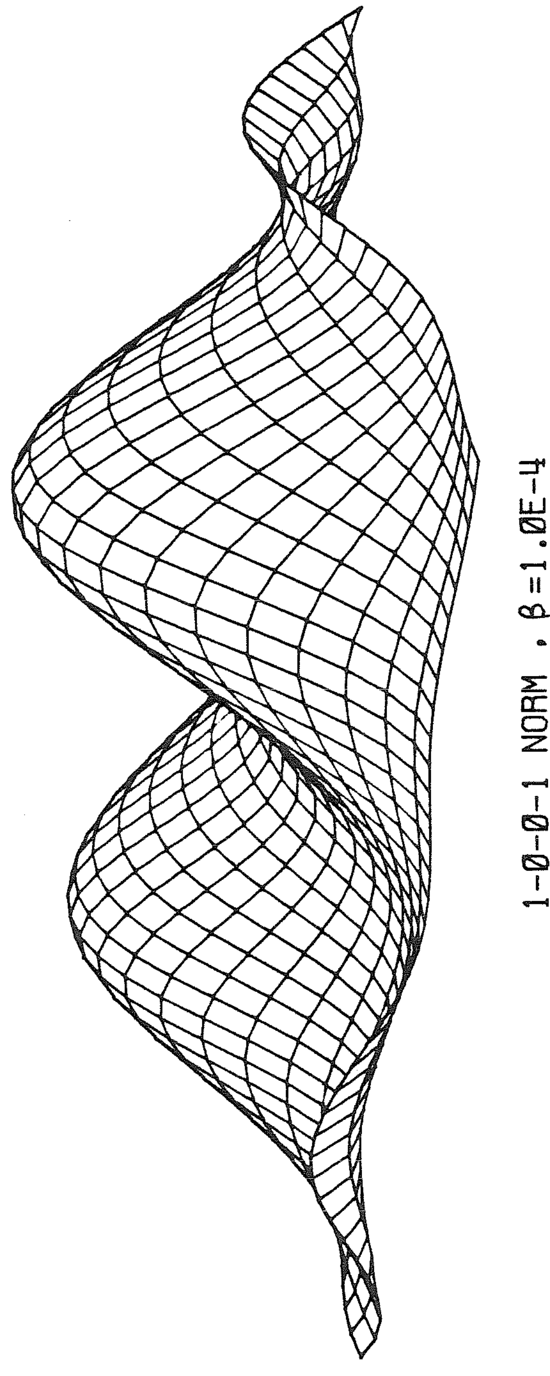


Figure 8

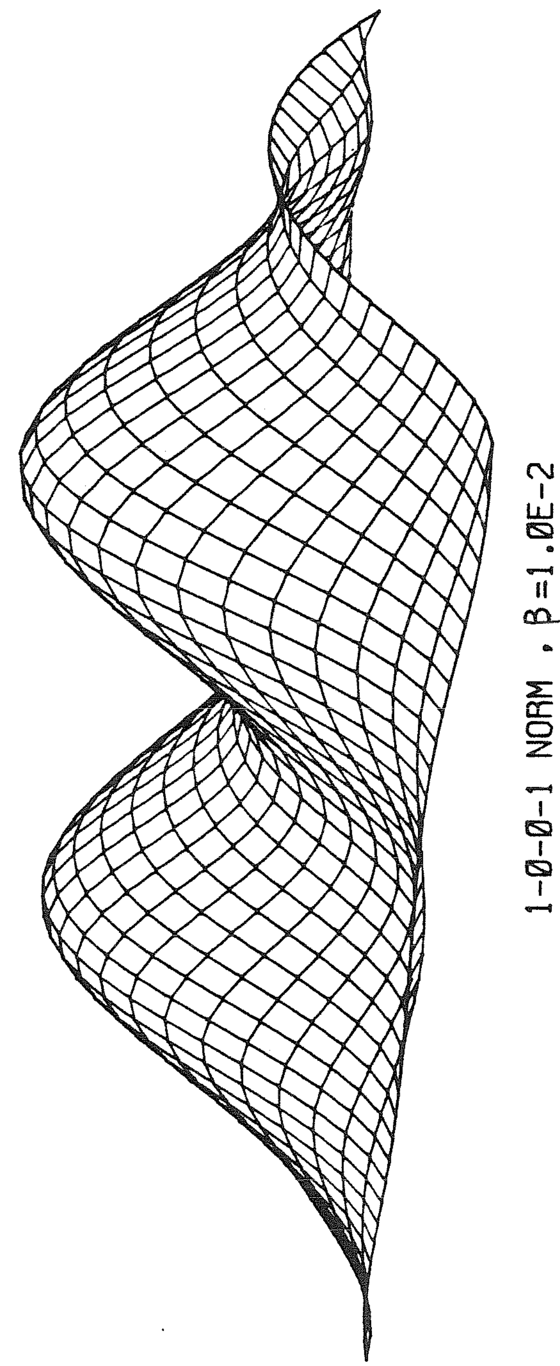


Figure 9

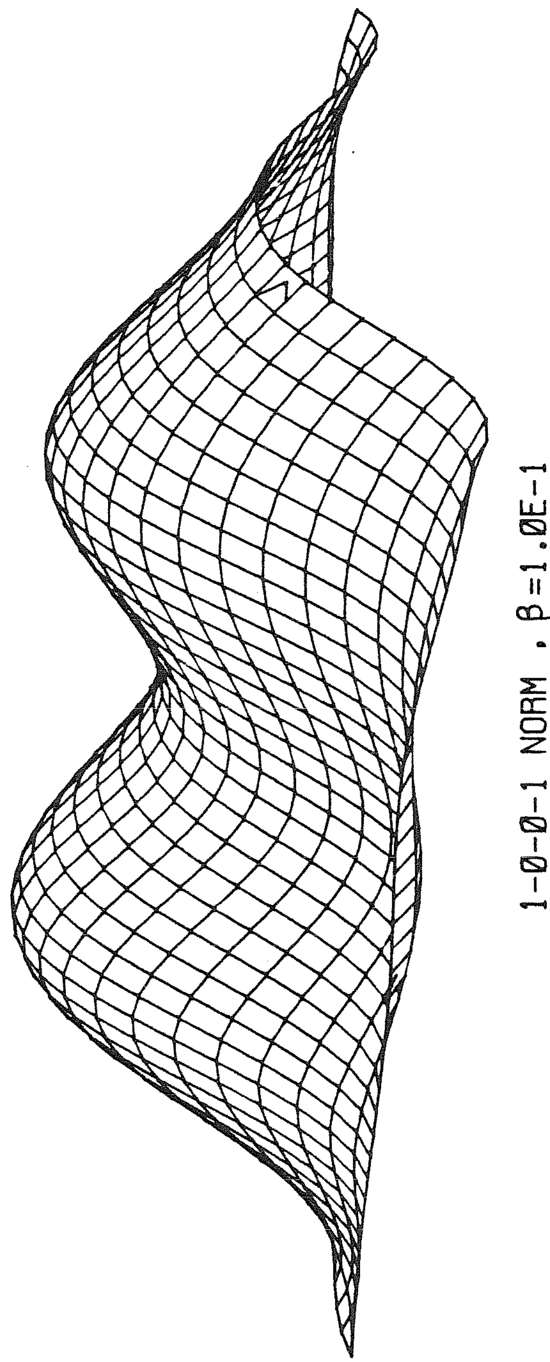


Figure 10

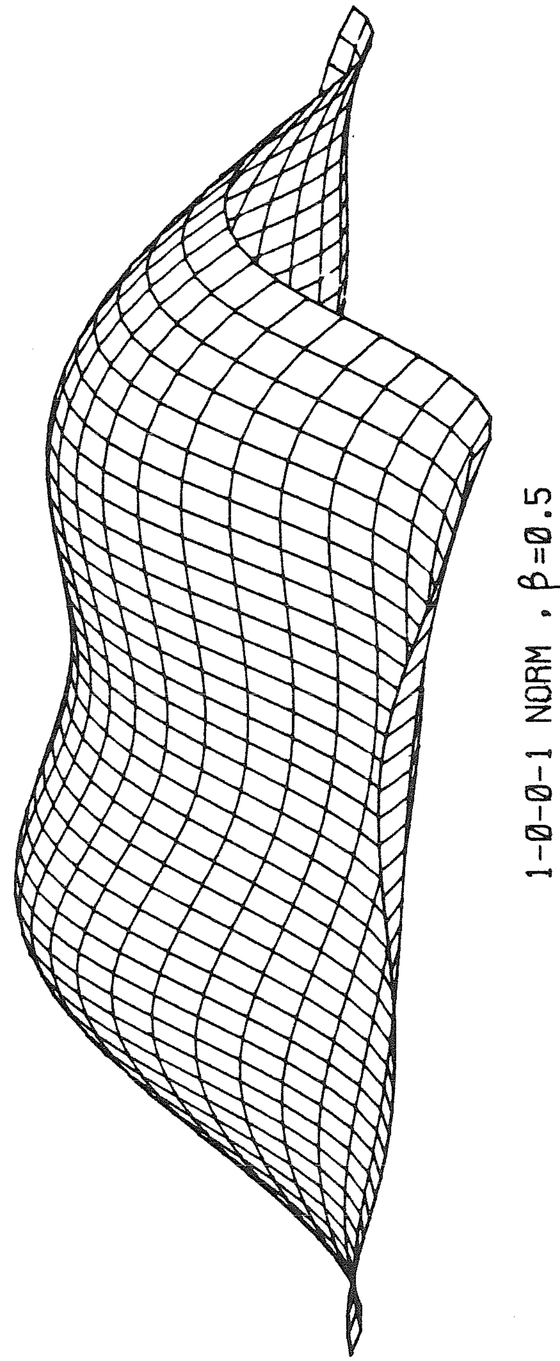


Figure 11

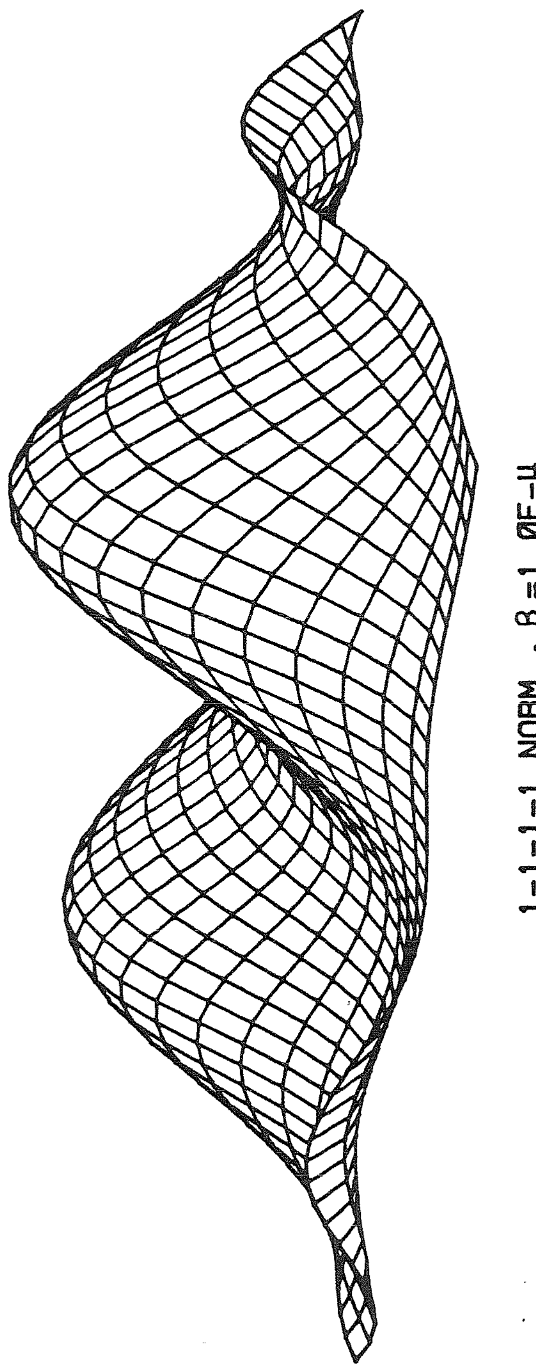


Figure 12

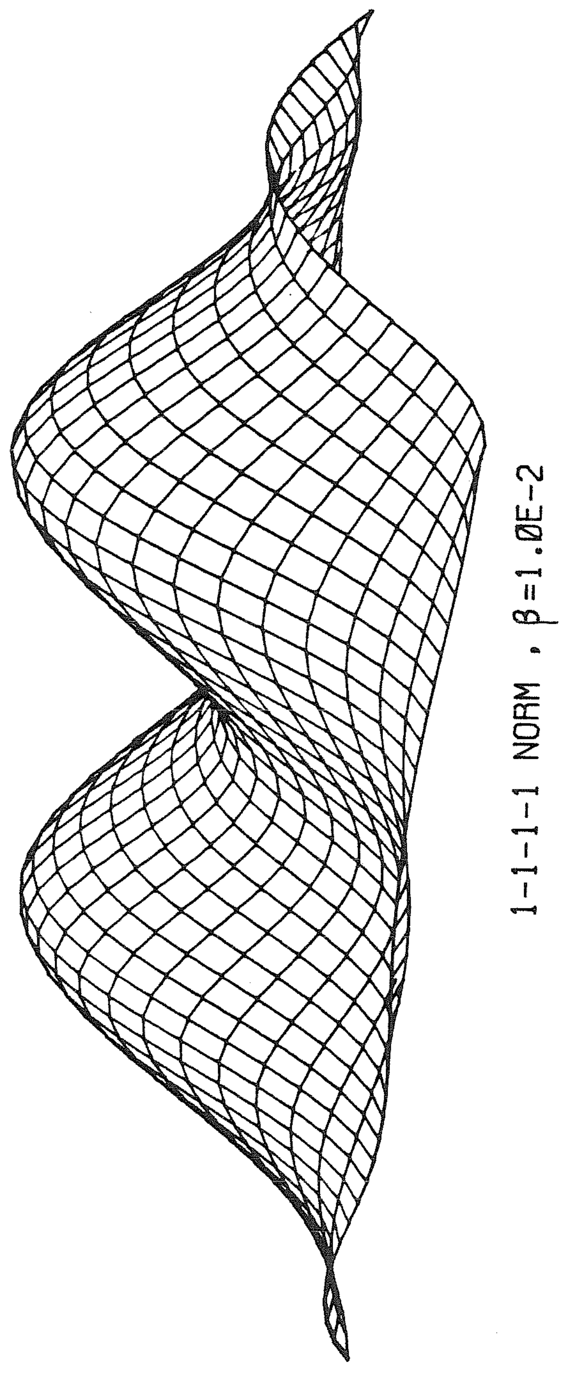


Figure 13

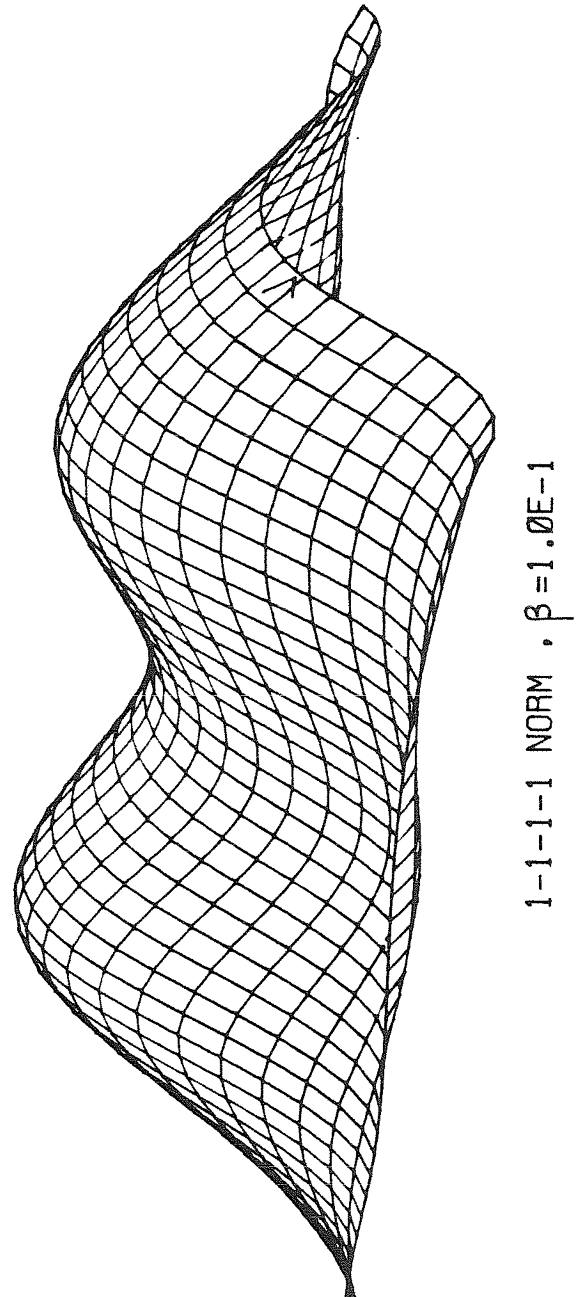


Figure 14

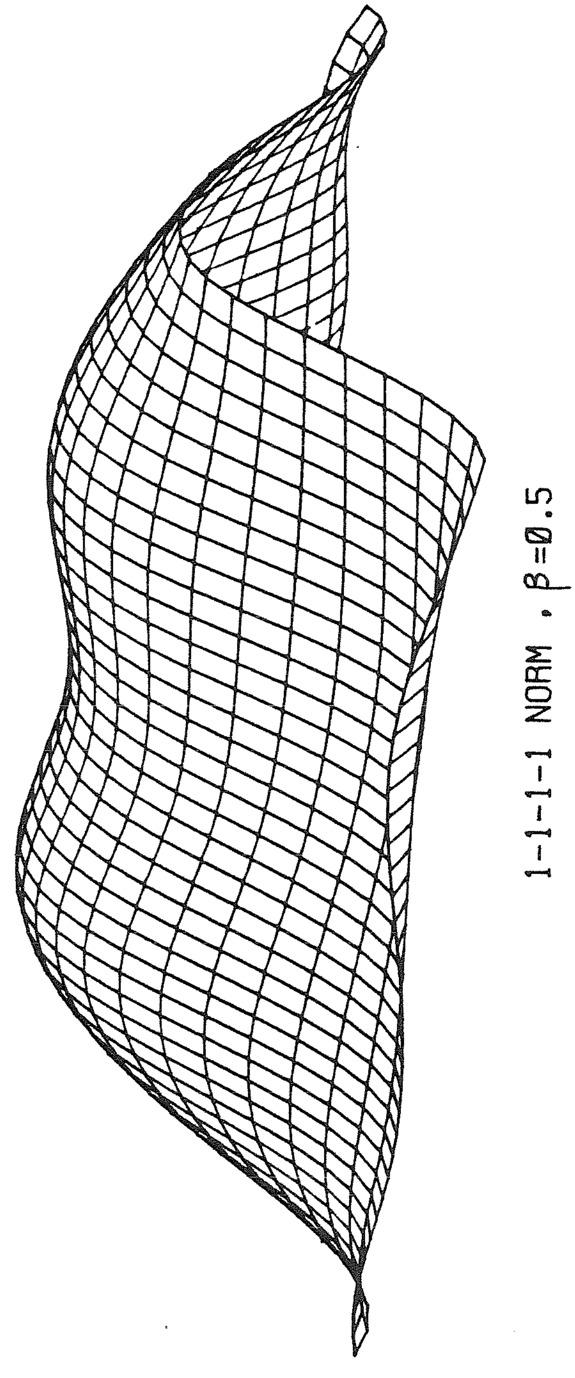


Figure 15

CONCLUSIONS

There are two basic problems in the area of identification of spatially-varying parameters in distributed parameter systems. The first is the so-called identifiability problem, which faces the question of whether measurements provide enough information to determine the unknown parameters uniquely. Since the unknown parameters, as a rule, do not depend continuously on the measurement of the state, the second problem concerns the construction of stable approximate solutions of the identification problem.

In Chapter II the first problem was addressed. It was demonstrated that identifiability from point observation (which is the only physically implementable type of observation) can be studied within the framework of inverse Sturm Liouville theory. The well-known Gel'fand-Levitan theory has been extended for inverse Sturm-Liouville problems which are not in normal form. As a consequence, identifiability and non-identifiability results have been obtained for the first time in the one-dimensional heat equation (with the spatially-varying conductivity as the unknown parameter).

In Chapters III and IV the second problem was addressed. The concept of regularization, which is widely used in solving linear Fredholm integral equations, has been extended for the identification of parameters in distributed parameter systems. A general regularization identification approach has been developed for the first time; it has been proved that this approach (under certain assumptions) generates stable approximate solutions to parameter identification problems.

The regularization identification approach has been proved to be applicable to a very broad class of parabolic systems. The numerical implementation of the regularization approach has been studied via two alternative numerical algorithms: classical Banach space gradient algorithm (Chapter III) and a discretized regularization algorithm (Chapter IV). Numerical experiments on the identification of spatially-varying diffusivity in the diffusion equation have demonstrated the power of the regularization identification approach.