NASA Contractor Report 172389

ICASE REPORT NO. 84-24

ICASE

ESTIMATION OF COEFFICIENTS AND BOUNDARY PARAMETERS IN HYPERBOLIC SYSTEMS

H. Thomas Banks

Katherine A. Murphy

Contract Nos. NAS1-16394, NAS1-17130

June 1984

INSTITUTE FOR COMPUTER APPLICATIONS IN SCIENCE AND ENGINEERING
NASA Langley Research Center, Hampton, Virginia 23665

Operated by the Universities Space Research Association

National Aeronautics and Space Administration

Langley Research Center
Hampton, Virginia 23665

ESTIMATION OF COEFFICIENTS AND BOUNDARY PARAMETERS IN HYPERBOLIC SYSTEMS

H. Thomas Banks
Lefschetz Center for Dynamical Systems
Brown University
and
Southern Methodist University

Katherine A. Murphy
Southern Methodist University

ABSTRACT

We consider semi-discrete Galerkin approximation schemes in connection with inverse problems for the estimation of spatially varying coefficients and boundary condition parameters in second order hyperbolic systems typical of those arising in 1-D surface seismic problems. Spline based algorithms are proposed for which theoretical convergence results along with a representative sample of numerical findings are given.

Research supported in part by NSF Grant MCS-8205355, AFOSR Contract 81-0198, and ARO Contract ARO-DAAG-29-83-K0029. Parts of the research were carried out while the first author was a visitor at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA which is operated under NASA Contract Nos. NAS1-16394 and NAS1-17130.

1. Introduction. In this paper we consider computational techniques for the following class of inverse problems: For the system

$$\rho(x) \frac{\partial^2 v}{\partial t^2} = \frac{\partial}{\partial x} \left(E(x) \frac{\partial v}{\partial x}\right) \qquad t > 0, \quad 0 \le x \le 1 , \tag{1.1}$$

$$\frac{\partial v}{\partial x}(t,0) + k_1 v(t,0) = s(t;\tilde{k}) \tag{1.2}$$

$$\frac{\partial v}{\partial t}(t,1) + k_2 \frac{\partial v}{\partial x}(t,1) = 0 \tag{1.3}$$

$$v(0,x) = \phi(x) , \quad v_t(0,x) = \psi(x) , \tag{1.4}$$

given observations $\{\hat{y}_{ij}\}$ for $\{v(t_i,x_j)\}$, choose, from some admissible set, "best" estimates for the parameters ρ, E, k_1, k_2, $\tilde{k}$. These problems are motivated by certain versions of the so-called "1-D Seismic Inversion Problem" (see, e.g. [1], [8]). Roughly speaking, one has an elastic medium (e.g., the earth) with density ρ and elastic modulus E. A perturbation of the system (explosions, or vibrating loads from specially designed trucks) near the surface ($x=0$) produces a source s for particle disturbances v that travel as elastic waves, being partially reflected due to the inhomogeneous nature of the medium. An important but difficult problem involves using the observed disturbances at the surface or at points along a "bore hole" to determine properties (represented by parameters in the system) of the medium. In the highly idealized 1-D "surface seismic" problem, one assumes that data are collected at the same point ($x=0$) where the original disturbance or "source" is located. In addition to this hypothesis which cannot be true, other unrealistic special assumptions are made about the nature of the traveling and reflected waves. Although the standard 1-D formulations are far from reality, exploration seismologists have developed techniques for processing actual field data (performing a series of experiments and "stacking" the data) so that the 1-D problems are generally accepted as useful and worthy subjects of investigation. Consequently, numerous papers (for some interesting references, see the bibliographies of [1], [8]) on the 1-D problems can be found in the research literature.

In many formulations of the seismic inverse problem, the medium is assumed to be the half-line $x > 0$ (with $x = 0$ the surface) while in others (especially some of those dealing with computational schemes) one finds the assumption of an artificial finite boundary (say at $x = 1$) at which no downgoing waves are reflected (an "absorbing" boundary). While there are several ways to approximate such a condition in 2 or 3 dimensional problems (see [12], [21]), for the 1-D formulation this condition is embodied in a simple boundary condition of the form (1.3); here $k_2 \approx \sqrt{E(1)/\rho(1)}$ and one can view this boundary condition as resulting from factoring the wave equation (1.1) at $x = 1$ and imposing the condition of "no upgoing waves" at $x = 1$.

Equation (1.1) is a 1-D version of the equations for an isotropic elastic medium while (1.2) represents an "elastic" boundary condition at the surface $x = 0$ (k_1 represents an elastic modulus for the restoring force produced by the medium).

As is the case in many inverse or "identification" problems, the problems described above tend to be ill-posed (including a computationally undesirable instability) unless careful restrictions are imposed on the admissible parameter class (for some discussions of these aspects, see [1], [10]). We shall not focus on this aspect here. Rather, the purpose of our presentation in this paper is to demonstrate the feasibility of a certain theoretical approach and certain approximations in developing computational schemes for problems in which there are i) unknown boundary parameters and ii) unknown spatially varying coefficients in the system equations.

We choose the "1-D seismic inverse problem" involving (1.1) - (1.4) as a test example to exhibit the efficacy of our ideas. However the technical features and notions we present are of importance in a number of other applications. There are rather easily motivated and fundamental problems in dealing with large elastic structures (large space structures - e.g. beam-like structures with tip bodies) that involve estimation of boundary condition parameters. In these cases the models are often hybrid models with distributed system (Euler-Bernoulli, Timoshenko) state equations and ordinary differential equation boundary conditions (see, for example, [2], [9], [18], [20]). A second class of problems for which the techniques introduced in this paper have immediate use are related to bioturbation [7], [13]. This is the mixing of lake and deep-sea sediments by burrowing activities of organisms. Understanding of this phenomenon is fundamental to geologists in interpreting geologic records contained in sediment core samples. The best models to date involve parabolic state equations (for a nonuniform "mixing chamber") with unknown parameters in the boundary conditions describing the flux into and out of the chamber.

In our approach here we employ the Trotter-Kato theorem to obtain theoretical convergence results (assuming regularity of parameter sets to guarantee existence of solutions to the inverse problems) for spline approximation schemes for the states. Boundary parameter estimation is treated directly via mappings that iteratively change the parameter-dependent spline basis elements into "conforming" elements (i.e., elements which satisfy the appropriate boundary conditions). We deal only with estimation of regular spatially-varying coefficients in (1.1), where again splines are used for parameters in a secondary approximation. Estimation of discontinuous coefficients (including location of the discontinuities) in problems such as those that are the focus of our attention in this paper can be effectively treated theoretically and numerically in a framework similar to that here using, for example, tau-Legendre state approximation schemes [4].

We turn then to the estimation problem for (1.1)-(1.4). It is theoretically and numerically advantageous to deal with homogeneous boundary conditions by transforming the problem so that the source term s in (1.2) appears in the initial data and in a term in the state equation. We make the transformation $u = v + G$ where (here "$\cdot$" represents differentiation with respect to t)

$$G(t,x;q) = -\left(\frac{1}{k_1}\right)s(t;\tilde{k}) + \left(\frac{1}{k_1 k_2}\right)x^2(x-1)\dot{s}(t;\tilde{k})$$

and obtain the system

$$q_1(x)\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(q_2(x)\frac{\partial u}{\partial x}\right) + F(t,x;q)$$

$$u_x(t,0) + q_3 u(t,0) = 0$$

(1.5)

$$u_t(t,1) + q_4 u_x(t,1) = 0$$

$$u(0,x) = \tilde{\phi}(x;q)\ ,\ u_t(0,x) = \tilde{\psi}(x;q).$$

Here the forcing function F is given by

$$F(t,x;q) \equiv q_1(x)\ \{-\left(\frac{1}{q_3}\right)\ddot{s}(t;\tilde{k}) + \left(\frac{1}{q_3 q_4}\right)x^2(x-1)\dddot{s}(t;\tilde{k})\}$$

$$-\frac{\partial}{\partial x}\{q_2(x)\left(\frac{1}{q_3 q_4}\right)(3x^2-2x)\dot{s}(t;\tilde{k})\}\ ,$$

where here and throughout we adopt the notation $q = (q_1, q_2, q_3, q_4, \tilde{k})$ with $q_1 \equiv \rho$, $q_2 \equiv E$, $q_3 = k_1$, and $q_4 = k_2$. The transformed initial conditions have the form

$$\tilde{\phi}(x;q) = \phi(x) - (\frac{1}{q_3})s(0;\tilde{k}) + (\frac{1}{q_3q_4})x^2(x-1)\dot{s}(0;\tilde{k})$$

$$\tilde{\psi}(x;q) = \psi(x) - (\frac{1}{q_3})\dot{s}(0;\tilde{k}) + (\frac{1}{q_3q_4})x^2(x-1)\ddot{s}(0;\tilde{k}) .$$

We assume henceforth that we have observations $\hat{y}_i = (\hat{y}_{i1}, \ldots, \hat{y}_{im})$, $i = 1, 2, \ldots, n$, corresponding to $w(t_i;q) = (u(t_i,x_1), \ldots, u(t_i,x_m))$ where u is the solution of (1.5). For a criterion in determining a best estimate $\hat{q}$ of the parameters we use a least-squares function

$$(1.6) \qquad J(q) = \sum_{i=1}^{n} |\hat{y}_i - w(t_i;q)|^2$$

which we seek to minimize as q ranges over some admissible parameter set Q. We remark that in the event our observations $\hat{\eta}_i = (\hat{\eta}_{i1}, \ldots, \hat{\eta}_{im})$ are for the original system (1.1)-(1.4), we may apply directly the theory and techniques of this paper by considering in place of (1.6) the criterion

$$(1.7) \qquad \tilde{J}(q) = \sum_{i=1}^{n} |\hat{\eta}_i + \tilde{G}(t_i;q) - w(t_i;q)|^2$$

where $\tilde{G}(t_i;q) \equiv (G(t_i,x_1;q), \ldots, G(t_i,x_m;q))$.

We make some standing assumptions to facilitate consideration of our problem in subsequent discussions. We shall search for q in a set $Q \subset C(0,1) \times H^1(0,1) \times R \times R \times R^k$ (we shall sometimes write Q as $Q_1 \times Q_2 \times Q_3 \times Q_4 \times Q_5$). We further assume that Q is compact in the $C \times H^1 \times R^{2+k}$ topology, and that there exist positive constants

$\underline{q}_i$, $\bar{q}_i$, i=1, 2, 3, 4 such that

$$\underline{q}_i \le q_i(x) \le \bar{q}_i \qquad \text{for } q_i \in Q_i \text{ , } i=1, 2,$$

$$\underline{q}_3 \le -q_3 \le \bar{q}_3 \qquad \text{for } q_3 \in Q_3 \text{ , and}$$

$$\underline{q}_4 \le q_4 \le \bar{q}_4 \qquad \text{for } q_4 \in Q_4 \text{ .}$$

Finally, we assume $\phi \in H^1(0,1)$, $\psi \in H^0(0,1)$, and $s(\cdot;\tilde{k}) \in H^3(0,T)$ for each $\tilde{k} \in Q_5$, where $t_i \in [0,T]$, $T < \infty$, and that $\tilde{k} \to s(\cdot;\tilde{k})$ is a continuous mapping from Q_5 to $H^3(0,T)$.

We turn next to the theoretical foundations of the approximation schemes we propose to use in solving our inverse problem of minimizing J over Q, subject to (1.5).

2. Abstract Formulation.

The object in this section is to lay the theoretical foundation for the problem. First, we shall write our partial differential equation as an abstract ordinary differential equation in a Hilbert space, then determine a set of approximating ordinary differential equations. Each of these abstract equations will have an associated identification problem; the original will be referred to as (ID), the N^{th} approximating problem will be referred to as (ID^N). We shall use the theory of semigroups to obtain existence and uniqueness of solutions to the differential equations. We can then fit our problem into the theoretical framework developed in [5], and deduce that, under conditions stated there (reiterated below for clarity), one can solve (ID^N) for each N, and these parameter estimates thus obtained will "lead to" a solution of (ID).

The equation (1.5) can be rewritten as a first order system, motivating the use of a product $(V(q) \times L^2(q))$ of two spaces to be our Hilbert space $X(q)$.

Define $V(q)$ to be $H^1(0,1)$ with inner product defined by $\langle v,w \rangle_{V(q)} = \int_0^1 q_2 DvDw dx - q_2(0)q_3 v(0)w(0)$. (D denotes the spatial differentiation operator $\frac{\partial}{\partial x}$). It can be readily shown that for any $q \epsilon Q$, $V(q)$ is a Hilbert space, and moreover, the assumptions made about Q imply that the $V(q)$ norm is uniformly equivalent to the H^1 norm as q ranges over Q. Let $V_B(q)$ contain those elements of $V(q)$ which satisfy the elastic boundary condition, i.e., $V_B(q) = \{v \epsilon V(q) \cap H^2(0,1) | Dv(0) + q_3 v(0) = 0\}$.

We define $L^2(q)$ to be $H^0(0,1)$ with inner product given by $\langle v,w \rangle_{0,q} = \int_0^1 q_1 vw dx$, and note that for each $q \epsilon Q$, $L^2(q)$ is a Hilbert space and its norm is uniformly equivalent to the standard H^0 norm as q ranges over Q.

As described earlier, we take $X(q) = V(q) \times L^2(q)$ with inner product given by $\langle x,y \rangle_q = \langle x_1,y_1 \rangle_{V(q)} + \langle x_2,y_2 \rangle_{0,q}$ (where $x = (x_1,x_2)^T$ and $y = (y_1,y_2)^T$). It is clear from our remarks above that for $q \epsilon Q$, $X(q)$ is a Hilbert space, and the

X norm is uniformly equivalent to the $H^1 \times H^0$ norm as q ranges over Q. We can formally write (1.5) as an abstract equation in $X(q)$:

$$\dot{z}(t) = A(q)z(t) + G(t;q)$$

(2.1)

$$z(0) = z_0(q)$$

where we have identified $z(t) \in X(q)$ with $\begin{pmatrix} u(t,\cdot) \\ u_t(t,\cdot) \end{pmatrix}$. The boundary conditions are incorporated into the domain of $A(q)$ by defining $\text{dom}A(q) = \{\binom{u}{v} \in V_B(q) \times H^1(0,1) | v(1) + q_4Du(1) = 0\}$, and A is the unbounded linear operator given by

$$A(q) = \begin{pmatrix} 0 & I \\ (1/q_1)D(q_2D) & 0 \end{pmatrix} .$$

The function G and the initial condition are given by

$$G(t;q) = \begin{pmatrix} 0 \\ F(t,\cdot;q) \end{pmatrix} \text{ and } z_0(q) = \begin{pmatrix} \tilde{\phi}(\cdot;q) \\ \tilde{\psi}(\cdot;q) \end{pmatrix}.$$

It can be shown that for each $q \in Q$, $A(q)$ is the infinitesimal generator of a C_0-semigroup, $T(t;q)$ on $X(q)$, so that we have the existence of mild solutions to (2.1), given by

(2.2) $$z(t;q) = T(t;q)z_0(q) + \int_0^t T(t-s;q)G(s;q)ds$$

with $z(\cdot;q) \in C(0,T;X(q))$. In this context, the inverse problem can be stated as:

(ID) Given observations $\hat{y} = \{\hat{y}_i\}_{i=1}^n$, minimize $J(z(\cdot;q),\hat{y})$ over $q \in Q$ subject to $z(\cdot;q)$ satisfying (2.2).

Here, $J(q) \equiv J(z(\cdot;q),\hat{y}) = \sum_{i=1}^{n} |\hat{y}_i - \xi(t_i,q)|^2$ where $\xi(t_i,q) =$

$(z_1(t_i,x_1;q),\ \ldots,\ z_1(t_i,x_m;q))$ and z_1 denotes the first component of z.

To prove that for each q, A(q) generates a C_0-semigroup, one can use the Lumer-Phillips Theorem ([15], p.16). To employ this theorem, one must show the operator is dissipative, densely defined, and satisfies a certain range statement. To demonstrate the dissipativity of A(q), we take $f \in \mathrm{dom}A(q)$, $q \in Q$, and compute (with an integration by parts)

$$\langle A(q)f,f\rangle_q = \left\langle \begin{pmatrix} f_2 \\ (1/q_1)D(q_2Df_1) \end{pmatrix}, \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \right\rangle_q$$

$$= \langle f_2,f_1\rangle_{V(q)} + \langle (1/q_1)D(q_2Df_1),f_2\rangle_{0,q}$$

$$= \int_0^1 q_2Df_1Df_2dx - q_2(0)q_3f_1(0)f_2(0) + \int_0^1 D(q_2Df_1)f_2dx$$

$$= -\ q_2(0)q_3f_1(0)f_2(0) - q_2(0)Df_1(0)f_2(0) + q_2(1)Df_1(1)f_2(1)$$

$$= -\ q_2(1)q_4(Df_1(1))^2 \leq 0\ .$$

By relating domA(q) to other subsets (see [14] for details) which are known to be dense in $H^1 \times H^0$, one can easily argue that for each $q \in Q$, domA(q) is dense in X(q). One can also argue that $\mathcal{R}(\lambda - A(q)) = X(q)$ for some $\lambda > 0$, by demonstrating that given $\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in X(q)$, there exists $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathrm{dom}A(q)$ such that

$$\begin{pmatrix} \lambda u - v \\ -(1/q_1)D(q_2Du) + \lambda v \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} .$$

This is equivalent to solving the following two point boundary value problem:

$$- (1/q_1)D(q_2 Du) + \lambda^2 u = \lambda f_1 + f_2$$

$$Du(0) + q_3 u(0) = 0$$

$$\lambda u(1) + q_4 Du(1) = f_1(1)$$

for $u \in H^2(0,1)$, and setting $v(x) = \lambda u(x) - f_1(x)$.

If we let $y = u - (1/q_4)x^2(x-1)f_1(1)$ the above problem is transformed to an equivalent one with homogeneous boundary conditions:

$$(-1/q_1)D(q_2 Dy) + \lambda^2 y = F$$

$$Dy(0) + q_3 y(0) = 0$$

$$q_4 Dy(1) + \lambda y(1) = 0$$

where $F \in L^2(q)$. One can then use the theory of self-adjoint operators (again see [14]) to argue that a solution exists for any $F \in L^2(q)$.

We now turn to the approximation of our equation (2.1). We shall obtain a solution z^N to an approximating equation (to be discussed in detail below) in a finite dimensional subspace of $X(q)$, denoted $X^N(q)$. Specifically, let $S^3(\Delta^N)$ represent the standard subspace of C^2 cubic splines corresponding to the partition $\Delta^N = \left\{x_i\right\}_{i=0}^{N}$, $x_i = i/N$ (see pp. 78-81 of [16]); then, given $q \in Q$, we take $X^N(q)$ to be that subspace of $S^3(\Delta^N) \times S^3(\Delta^N)$ whose elements satisfy the boundary conditions corresponding to q (i.e., $X^N(q) \subset \mathrm{dom} A(q)$). Let B_j^N, $j = -1,\ldots,N+1$, be the B-spline basis elements for $S^3(\Delta^N)$. Then $X^N(q)$ is the $(2N+3)$-dimensional subspace spanned by the following set of basis functions:

$$\beta_1^N = \begin{pmatrix} \frac{4q_3}{N} B_{-1}^N + (3 - \frac{q_3}{N}) B_0^N \\ 0 \end{pmatrix}, \quad \beta_2^N = \begin{pmatrix} -\frac{4q_3}{N} B_1^N + (3 + \frac{q_3}{N}) B_0^N \\ 0 \end{pmatrix},$$

$$\beta_3^N = \begin{pmatrix} B_2^N \\ 0 \end{pmatrix}, \ldots, \quad \beta_{N-1}^N = \begin{pmatrix} B_{N-2}^N \\ 0 \end{pmatrix},$$

$$\beta_N^N = \begin{pmatrix} B_{N-1}^N \\ \frac{3Nq_4}{4} B_N^N \end{pmatrix}, \quad \beta_{N+1}^N = \begin{pmatrix} B_N^N \\ 0 \end{pmatrix}, \quad \beta_{N+2}^N = \begin{pmatrix} B_{N+1}^N \\ -\frac{3Nq_4}{4} B_N^N \end{pmatrix},$$

$$\beta_{N+3}^N = \begin{pmatrix} -1/(3Nq_4) B_{N+1}^N \\ B_{N+1}^N \end{pmatrix}, \quad \beta_{N+4}^N = \begin{pmatrix} -1/(3Nq_4) B_{N+1}^N \\ B_{N-1}^N \end{pmatrix},$$

$$\beta_{N+5}^N = \begin{pmatrix} 0 \\ B_{N-2}^N \end{pmatrix}, \ldots, \quad \beta_{2N+1}^N = \begin{pmatrix} 0 \\ B_2^N \end{pmatrix},$$

$$\beta_{2N+2}^N = \begin{pmatrix} 0 \\ \frac{-4q_3}{N} B_1^N + (3 + \frac{q_3}{N}) B_0^N \end{pmatrix}, \quad \beta_{2N+3}^N = \begin{pmatrix} 0 \\ \frac{4q_3}{N} B_{-1}^N + (3 - \frac{q_3}{N}) B_0^N \end{pmatrix}.$$

Let $P^N(q): X(q) \to X^N(q)$ denote the orthogonal projection of $X(q)$ onto $X^N(q)$, i.e., given $f \in X(q)$, $P^N(q)f$ is that element in $X^N(q)$ which satisfies $|P^N(q)f - f|_q \leq |g - f|_q$ for all $g \in X^N(q)$. For each $q \in Q$, we define an operator

$A^N(q)$ on $X(q)$ given by $A^N(q) = P^N(q)A(q)P^N(q)$, and then the approximating equation to (2.1) is written as:

$$\text{(2.3)}\qquad \begin{aligned} \dot{z}^N(t) &= A^N(q)z^N(t) + P^N(q)G(t;q) \\ z^N(0) &= P^N(q)z_0(q) \end{aligned}$$

where $z^N(t) \in X^N(q)$. Using the fact that $A(q)$ is closed, $P^N(q)$ is bounded, and the Closed Graph Theorem, one finds that $A^N(q)$ is bounded. The operator $A^N(q)$ inherits the dissipativity of $A(q)$, and therefore it follows that for each $q \in Q$, $A^N(q)$ is the infinitesimal generator of a C_0-semigroup of contractions $T^N(t;q)$ on $X(q)$. It is readily seen that $T^N(t;q)$ leaves $X^N(q)$ invariant. Thus, for each $q \in Q$ and each $N = 1,2, \ldots$, there exists a unique mild solution $z^N(\cdot;q) \in C(0,T;X^N(q))$ of (2.3), which can be expressed as

$$\text{(2.4)}\qquad z^N(t;q) = T^N(t;q)P^N(q)z_0(q) + \int_0^t T^N(t-s;q)P^N(q)G(s;q)ds.$$

The associated approximate identification problem is given by

(ID^N) Given observations $\hat{y} = \left\{\hat{y}_i\right\}_{i=1}^n$, minimize $J(z^N(\cdot;q),\hat{y})$ over $q \in Q$ subject to $z^N(\cdot;q)$ satisfying (2.4).

Here, $J^N(q) \equiv J(z^N(\cdot;q),\hat{y}) = \sum_{i=1}^n |\hat{y}_i - \xi^N(t_i,q)|^2$ where $\xi^N(t_i,q) =$

$(z_1^N(t_i,x_1;q), \ldots, z_1^N(t_i,x_m;q))$ and z_1^N denotes the first component of z^N.

Since $X^N(q)$ is finite dimensional, (2.3) is in fact a system of $2N+3$ ordinary differential equations, which can be solved using standard numerical packages. Similarly, there are numerical packages available to solve (ID^N), provided solutions exist and we have some computationally feasible representation for q_1 and q_2. A detailed description of our numerical implementation, including a discussion of possible representations of q_1 and q_2, will be deferred to subsequent sections. First, our concern is to determine under what conditions solutions of (ID^N) exist and how they relate to a solution of (ID). This is the subject of the next theorem, a slight modification of that given in [5 , p. 820].

Theorem 2.1. Assume Q is compact in the $C \times H^1 \times R^{2+k}$ topology. If $q \to z_0(q)$, $q \to P^N(q)f$, $q \to T^N(t;q)f$, $f \in X = X(q)$ are continuous in this same Q-topology, with the latter uniformly in $t \in [0,T]$, then

(i) There exists for each N a solution $\hat{q}^N$ of (ID^N) and the sequence $\{\hat{q}^N\}$ possesses a convergent subsequence $\hat{q}^{N_k} \to \hat{q}$.

(ii) If we further assume that, for any sequence $\{q^j\}$ in Q with $q^j \to \tilde{q}$, we have $|z^j(t;q^j) - z(t;\tilde{q})|_{q^j} \to 0$ as $j \to \infty$, uniformly in $t \in [0,T]$, then $\hat{q}$ is a solution of (ID).

The reader may, at first glance, find the convergence statement of (ii) suspect in that $z^j(t;q^j) \in X^j(q^j)$ and $z(t;\tilde{q}) \in X(\tilde{q})$, but this statement is meaningful in view of the following observation. In defining the spaces $V(q)$, $L^2(q)$, and $X(q)$, it was noted that $V(q)$, $L^2(q)$, and $X(q)$ are uniformly equivalent to H^1, H^0, and $H^1 \times H^0$, respectively, as q ranges over Q. This implies that the $X(q)$ are setwise equal as q ranges over Q. To be technically precise, we should use the canonical isomorphism when relating an element of $X(q^j)$ to its counterpart in $X(\tilde{q})$, but to simplify our presentation, we shall throughout abuse notation and omit the isomorphism.

It is easily seen from the form of $z_0(q)$ that $q \to z_0(q)$ is continuous. It is also true that for our $P^N(q)$, $T^N(t;q)$ we have $q \to P^N(q)f$ and $q \to T^N(t;q)f$ continuous; this will be readily seen from the matrix representations for our approximating scheme, and so further discussion is postponed until Section 5.

The next theorem gives sufficient conditions for the hypothesis of (ii) from Theorem 2.1 to hold.

Theorem 2.2. Let q^N, $\tilde{q}$ be arbitrary in Q such that $q^N \to \tilde{q}$ as $N \to \infty$ (recall convergence is in the $C \times H^1 \times R^{2+k}$ topology). Suppose that the projections $P^N(q)$ are such that $|(P^N(q^N) - I)f|_{q^N} \to 0$ as $N \to \infty$ for all $f \in X(\tilde{q})$, that $f \in X(\tilde{q})$ implies $|T^N(t;q^N)f - T(t;\tilde{q})f|_{q^N} \to 0$ as $N \to \infty$, uniformly in $t \in [0,T]$, and that $|z_0(q^N) - z_0(\tilde{q})|_{q^N} \to 0$ as $N \to \infty$. Then the mild solutions $z^N(t;q^N)$ of (2.3) converge to the mild solution $z(t;\tilde{q})$ of (2.1) uniformly in $t \in [0,T]$.

The proof of this theorem, which is based on a standard "variation-of-constants" representation for solutions z and z^N in terms of the semigroups T and T^N, essentially follows immediately from Theorem 3.1 of [5 , p. 823]. One only needs to verify that our spaces, operators, etc. satisfy the conditions required in [5].

It is clear from the continuity of $q \to z_0(q)$ that $|z_0(q^N) - z_0(\tilde{q})|_{q^N} \to 0$ as $q^N \to \tilde{q}$. It remains only to show the convergence of the projections and the semigroups. The main result of the next section is the convergence of the semigroups; the convergence of the projections is obtained as an intermediate proposition. In summary then, at the end of the next section, we will be able to deduce from Theorem 2.2 that $z^N(t;q^N)$ converges to $z(t;\tilde{q})$ whenever $q^N \to \tilde{q}$, and hence by Theorem 2.1 we are assured that the sequence of iterates $\{\hat{q}^N\}$ we obtain by solving (ID^N), has a subsequence which converges to a solution, $\hat{q}$, of (ID).

3. Convergence Arguments.

This section will be devoted to establishing the result: For each convergent sequence $q^N \to \tilde{q}$ in Q, and for any $f \in X(\tilde{q})$, $|T^N(t;q^N)f - T(t;\tilde{q})f|_{q^N} \to 0$ as $N \to \infty$, uniformly in $t \in [0,T]$. As explained in the previous section, this convergence result is crucial in arguing that $z^N(t;q^N) \to z(t;\tilde{q})$ whenever $q^N \to \tilde{q}$, which in turn is necessary to ensure that our candidate (the limit of our approximating subsequence) is indeed a solution to our inverse problem.

We shall first prove a slightly different form of convergence of the semigroups using the following version of the Trotter-Kato Theorem [3].

<u>Theorem 3.1</u>. Let $(\mathcal{B},|\cdot|)$ and $(\mathcal{B}^N,|\cdot|_N)$, $N = 1, 2, \ldots$, be Banach spaces and let $\pi^N : \mathcal{B} \to \mathcal{B}^N$ be bounded linear operators. Further assume that $T(t)$ and $T^N(t)$ are C_0-semigroups on $\mathcal{B}$ and $\mathcal{B}^N$ with infinitesimal generators $\tilde{A}$ and $\tilde{A}^N$, respectively. If

(i) $\lim\limits_{N\to\infty} |\pi^N f|_N = |f|$ for all $f \in \mathcal{B}$,

(ii) there exist constants M, ω independent of N such that
$|T^N(t)|_N \leq Me^{\omega t}$, for $t \geq 0$,

(iii) there exists a set $\mathcal{D} \subset \mathcal{B}$, $\mathcal{D} \subset \operatorname{dom}(\tilde{A})$, with $\overline{(\lambda_0 - \tilde{A})\mathcal{D}} = \mathcal{B}$ for some $\lambda_0 > 0$, such that for all $f \in \mathcal{D}$ we have

$$|\tilde{A}^N \pi^N f - \pi^N \tilde{A} f|_N \to 0 \quad \text{as } N \to \infty,$$

then $|T^N(t)\pi^N f - \pi^N T(t) f|_N \to 0$ as $N \to \infty$, for all $f \in \mathcal{B}$, uniformly in t on compact intervals in $[0,\infty)$.

It will be a standing assumption throughout this section that $q^N \to \tilde{q}$ in Q with this convergence in the $C \times H^1 \times R^{2+k}$ topology. Let $\mathcal{B} = X(\tilde{q})$ with norm denoted by $|\cdot|_{\tilde{q}}$, $\mathcal{B}^N = X(q^N)$ with norm $|\cdot|_{q^N}$ for $N = 1, 2, \ldots$, $\tilde{A} = A(\tilde{q})$

with corresponding semigroup $T(t) = T(t;\tilde{q})$, and $\tilde{A}^N = A^N(q^N) = P^N(q^N)A(q^N)P^N(q^N)$ with corresponding semigroup $T^N(t) = T^N(t;q^N)$ (as described in Section 2). For each N, $\pi^N : X(\tilde{q}) \to X(q^N)$ will be a bounded linear operator which will map elements of $\mathrm{dom}A(\tilde{q})$ into elements of $\mathrm{dom}A(q^N)$. Define

$$g^N(x) = \exp([\tilde{q}_3 - q_3^N]x) - (x^2/2)[\tilde{q}_3 - q_3^N]\exp[\tilde{q}_3 - q_3^N];$$

given $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$, let π^N be defined by

$$\pi^N f = \begin{pmatrix} g^N f_1 \\ (q_4^N/\tilde{q}_4) g^N f_2 \end{pmatrix} .$$

The functions g^N are defined so that as $N \to \infty$, $g^N(x) \to 1$, and $D^j(g^N(x)) \to 0$ for any positive integer j, where in each case the convergence is uniform in $x \in [0,1]$.

A simple computation demonstrates that if $f \in \mathrm{dom}\, A(\tilde{q})$, then $\pi^N f \in \mathrm{dom}\, A(q^N)$. For each N, π^N is a bounded linear operator from $X(\tilde{q})$ to $X(q^N)$, but moreover, the set of operators $\{\pi^N\}$ is uniformly bounded. This statement can be proved using the assumptions on Q and the properties of g^N mentioned above. Similar comments apply to the proof of our first proposition.

Proposition 3.1. For any $f \in X(\tilde{q})$, $|\pi^N f - f|_{q^N} \to 0$ as $N \to \infty$.

In order to argue the convergence of the infinitesimal generators, we shall need error estimates for the spline approximations and their derivatives. These will be variations of estimates such as those found in [19], modified to take into account our q-dependent norm, and the presence of the operator π^N.

The following notation will be used throughout this section. Given a vector function f, we shall use f_i or $(f)_i$ to denote the i^{th} component of f. Given the scalar function h, $I^N h$ will denote the standard cubic spline interpolant of h (thus $I^N h \in S^3(\Delta^N)$). For a vector function $f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$, $I^N f$ will be the vector whose components are the spline interpolants of the components of f, i.e., $I^N f = \begin{pmatrix} I^N f_1 \\ I^N f_2 \end{pmatrix}$ and $I^N f \in S^3(\Delta^N) \times S^3(\Delta^N)$. The interpolant of f which satisfies the boundary conditions corresponding to q will be written as $I_B^N(q)f$. While $I^N f$ interpolates f_1 and f_2 at the values $\left\{i/N\right\}_{i=0}^N$ and the derivatives of f_1 and f_2 at 0 and 1, $I_B^N(q)f$ will interpolate f_1 and f_2 at the values $\left\{i/N\right\}_{i=0}^N$, and will additionally satisfy

$$[D(I_B^N(q)f)_i](0) + q_3[(I_B^N(q)f)_i](0) = 0, \text{ or equivalently,}$$

$$[D(I_B^N(q)f)_i](0) = -\, q_3 f_i(0) \qquad \text{for } i = 1, 2,$$

and

$$[(I_B^N(q)f)_2](1) + q_4[D(I_B^N(q)f)_1](1) = 0, \text{ or equivalently,}$$

$$[D(I_B^N(q)f)_1](1) = -\,(1/q_4)f_2(1).$$

We note that if f satisfies the boundary conditions involving q, then $I_B^N(q)f = I^N f$.

The first estimates involve cubic interpolants for scalar functions.

Lemma 3.1. If $h \in H^2$, then

$$|D^2(h-I^N h)|_0 \to 0 \qquad \text{as } N \to \infty ,$$

$$|D(h-I^N h)|_0 \le N^{-1}|D^2(h-I^N h)|_0 \le N^{-1}|D^2 h|_0 ,$$

$$|h - I^N h|_0 \le N^{-2}|D^2(h-I^N h)|_0 \le N^{-2}|D^2 h|_0 .$$

The convergence statement of this lemma follows immediately from the density of H^3 in H^2, the estimates of Theorem 6.9 of [19], and the first integral relation (4.15) of [19]. The estimates follow from (4.24) and (4.25), respectively, of [19] and the first integral relation.

One can use the results of Lemma 3.1 and the equivalence of the $X(q)$ and $H^1 \times H^0$ norms to derive similar statements for the interpolants in the $X(q)$ norm.

Lemma 3.2. If $f \in H^2 \times H^2$ and $q \in Q \subset C \times H^1 \times R^{2+k}$, then

$$|I^N f - f|_q \le K_1 N^{-1}(|D^2(f_1 - I^N f_1)|_0^2 + |D^2(f_2 - I^N f_2)|_0^2)^{1/2}$$

$$\le K_1 N^{-1}(|D^2 f_1|_0^2 + |D^2 f_2|_0^2)^{1/2}$$

$$|D(I^N f - f)|_q \le K_2(|D^2(f_1 - I^N f_1)|_0^2 + |D^2(f_2 - I^N f_2)|_0^2)^{1/2}$$

where K_1, K_2 are constants which are independent of f, q, and N.

Again, due to the equivalence of norms, the Schmidt inequality of [19, Thm. 1.5] can be modified and used component-wise to give a Schmidt type inequality in the $X(q)$ norm.

Lemma 3.3. If $f \in S^3(\Delta^N) \times S^3(\Delta^N)$ and $q \in Q$, then $|Df|_q \le K_3 N |f|_q$, where K_3 is a constant independent of f, N, and q.

The preceding estimates can be used to establish convergence properties for the canonical projections $P^N(q^N)$ where $q^N \to \tilde{q}$ in Q.

Proposition 3.2. If $f \in X(\tilde{q})$, then

$$|P^N(q^N)f - f|_{q^N} \to 0 \quad \text{as} \quad N \to \infty .$$

Proof. First consider $f \in \mathrm{dom}A(\tilde{q}) \cap (H^2 \times H^2)$. For such f, $\pi^N f \in \mathrm{dom}A(q^N) \cap (H^2 \times H^2)$

and $I_B^N(q^N)\pi^N f = I^N\pi^N f$. We use Lemma 3.2 in the triangle inequalities below to derive

$$|P^N(q^N)f - f|_{q^N} \le |P^N(q^N)[f - \pi^N f]|_{q^N} + |P^N(q^N)\pi^N f - \pi^N f|_{q^N} + |\pi^N f - f|_{q^N}$$

$$\le 2|\pi^N f - f|_{q^N} + |I_B^N(q^N)\pi^N f - \pi^N f|_{q^N}$$

$$= 2|\pi^N f - f|_{q^N} + |I^N\pi^N f - \pi^N f|_{q^N}$$

$$\le 2|\pi^N f - f|_{q^N} + K_1 N^{-1}(|D^2(\pi^N f)_1|_0^2 + |D^2(\pi^N f)_2|_0^2)^{1/2} \quad .$$

Thus we have $|P^N(q^N)f - f|_{q^N}$ bounded by terms which we can show converge to zero using Proposition 3.1 and the properties of g^N.

The $P^N(q^N)$ are uniformly bounded, and the set $\mathrm{dom}A(\tilde{q})\cap(H^2 \times H^2)$ is dense in $X(\tilde{q})$, hence one can use standard arguments to conclude that the statement of the proposition holds for all $f \in X(\tilde{q})$.

Proposition 3.3. For each $f \in X(\tilde{q})$, $|(P^N(q^N) - I)\pi^N f|_{q^N} \to 0$ as $N \to \infty$, and for each $f \in \mathrm{dom}A(\tilde{q})\cap(H^2 \times H^2)$, $|D[(P^N(q^N) - I)\pi^N f]|_{q^N} \to 0$ as $N \to \infty$.

Proof. The first statement is proved within the proof of Proposition 3.2; specifically, it was shown that $|P^N(q^N)\pi^N f - \pi^N f|_{q^N} \le K_1 N^{-1}(|D^2(\pi^N f)_1|_0^2 + |D^2(\pi^N f)_2|_0^2)^{1/2}$.

The proof of the second statement is obtained from the following triangle inequality (here we also use Lemmas 3.3, 3.2):

$$|D(P^N(q^N)\pi^N f - \pi^N f)|_{q^N} \le |D(P^N(q^N)\pi^N f - I^N(\pi^N f))|_{q^N} + |D(I^N(\pi^N f) - \pi^N f)|_{q^N}$$

$$\le K_3 N|P^N(q^N)\pi^N f - I^N(\pi^N f)|_{q^N} + |D(I^N(\pi^N f) - \pi^N f)|_{q^N}$$

$$\le K_3 N|(P^N(q^N) - I)\pi^N f|_{q^N} + K_3 N|\pi^N f - I^N\pi^N f|_{q^N} + |D[I^N(\pi^N f) - \pi^N f]|_{q^N}$$

$$\leq 2K_3 N|I^N \Pi^N f - \Pi^N f|_{q^N} + |D[I^N \Pi^N f - \Pi^N f]|_{q^N}$$

$$\leq (2K_1K_3 + K_2)(|D^2[(\Pi^N f)_1 - I^N(\Pi^N f)_1]|_0^2 + |D^2[(\Pi^N f)_2 - I^N(\Pi^N f)_2]|_0^2)^{1/2}$$

Thus the conclusion $|D[(P^N(q^N) - I)\Pi^N f]|_{q^N} \to 0$ as $N \to \infty$ follows from the observation that for $i = 1,2$

$$|D^2[I^N(\Pi^N f)_i - (\Pi^N f)_i]|_0 \leq |D^2[I^N((\Pi^N f)_i - f_i)]|_0$$

$$+ |D^2[I^N f_i - f_i]|_0 + |D^2[f_i - (\Pi^N f)_i]|_0$$

$$\leq 2|D^2[(\Pi^N f)_i - f_i]|_0 + |D^2[I^N f_i - f_i]|_0 \quad ,$$

with the latter terms approaching zero because of the properties of g^N and Lemma 3.1, respectively.

In later arguments, it will be helpful to have bounds (in the H^1 and H^0 norms) on one component of an element of X in terms of a bound (in the X(q) norm) on the entire element. Thus, we consider for $f \in X(q)$, $|f|_q^2 = |f_1|_{V(q)}^2 + |f_2|_{0,q}^2$ which is equivalent to $|Df_1|_0^2 + |f_1|_0^2 + |f_2|_0^2$, so that there exist constants k_1 and k_2 such that $|Df_1|_0^2 \leq k_1|f|_q^2$ and $|f_2|_0^2 \leq k_2|f|_q^2$. Similarly, $|Df|_q^2 = |Df_1|_{V(q)}^2 + |Df_2|_{0,q}^2$ which is equivalent to $|D^2f_1|_0^2 + |Df_1|_0^2 + |Df_2|_0^2$ so we infer the existence of constants k_3 and k_4 such that $|D^2f_1|_0^2 \leq k_3|Df|_q^2$ and $|Df_2|_0^2 \leq k_4|Df|_q^2$. For future reference, we combine and label these observations as

$$|Df_1|_0^2 \leq k_1|f|_q^2$$

$$(3.1) \qquad |D^2f_1|_0^2 \leq k_3|Df|_q^2$$

$$|f_2|_1^2 \leq k_2|f|_q^2 + k_4|Df|_q^2 \quad .$$

It is now possible to state and prove the following convergence theorem.

Theorem 3.2. Suppose $q^N \to \tilde{q}$ in Q (convergence is in the $C \times H^1 \times R^{2+k}$ topology). Then

$$|T^N(t;q^N)\pi^N f - \pi^N T(t;\tilde{q})f|_{q^N} \to 0 \quad \text{as } N \to \infty ,$$

for all $f \in X(\tilde{q})$, uniformly in t on compact intervals in $[0,\infty)$.

Proof. The result is an immediate consequence of Theorem 3.1, once the hypotheses of that theorem have been shown to hold. Part (i) follows from Proposition 3.1, while part (ii) holds since $T^N(t;q)$ and $T(t;q)$ are contraction semigroups for each N and $q \in Q$. It remains only to verify (iii), for which we take $\mathcal{D}$ to be the set $\mathrm{dom}A(\tilde{q}) \cap (H^2 \times H^2)$. Let $f \in \mathcal{D}$. Then

$$\begin{aligned}
|A^N(q^N)\pi^N f - \pi^N A(\tilde{q})f|_{q^N} &= |P^N(q^N)A(q^N)P^N(q^N)\pi^N f - \pi^N A(\tilde{q})f|_{q^N} \\
&\leq |P^N(q^N)[A(q^N)P^N(q^N)\pi^N f - \pi^N A(\tilde{q})f]|_{q^N} + |P^N(q^N)\pi^N A(\tilde{q})f - \pi^N A(\tilde{q})f|_{q^N} \\
&\leq |A(q^N)P^N(q^N)\pi^N f - \pi^N A(\tilde{q})f|_{q^N} + |(P^N(q^N) - I)\pi^N A(\tilde{q})f|_{q^N} \\
&\equiv \varepsilon_1(N) + \varepsilon_2(N).
\end{aligned}$$

It follows directly from Proposition 3.3 that $\varepsilon_2(N) \to 0$ as $N \to \infty$. We must work harder to establish that $\varepsilon_1(N) \to 0$. We begin by breaking the norm into its two components and treat each separately. Thus

$$\begin{aligned}
[\varepsilon_1(N)]^2 &= \left| \begin{pmatrix} 0 & 1 \\ (1/q_1^N)D(q_2^N D) & 0 \end{pmatrix} (P^N(q^N)\pi^N f) - \pi^N \begin{pmatrix} 0 & 1 \\ (1/\tilde{q}_1)D(\tilde{q}_2 D) & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \right|^2_{q^N} \\
&= |(P^N(q^N)\pi^N f)_2 - g^N f_2|^2_{V(q^N)} \\
&\quad + |(1/q_1^N)D[q_2^N D(P^N(q^N)\pi^N f)_1] - (q_4^N/\tilde{q}_4)g^N(1/\tilde{q}_1)D[\tilde{q}_2 D f_1]|^2_{0,q^N}
\end{aligned}$$

$$\equiv [\delta_1(N)]^2 + [\delta_2(N)]^2 \; .$$

We first observe that

$$\delta_1(N) \le |(P^N(q^N)\pi^N f)_2 - (q_4^N/\tilde{q}_4)g^N f_2|_{V(q^N)} + |[(q_4^N/\tilde{q}_4) - 1]g^N f_2|_{V(q^N)}$$

$$= |(P^N(q^N)\pi^N f)_2 - (\pi^N f)_2|_{V(q^N)} + |((q_4^N/\tilde{q}_4) - 1)g^N f_2|_{V(q^N)} \quad .$$

It is more convenient, and due to the equivalence of the norms, it is sufficient, to establish the convergence in the H^1 norm. This can easily be done for the first term by invoking Proposition 3.3 and the inequalities (3.1). An argument can be made for the second term based on the properties of the g^N and the convergence $q^N \to \tilde{q}$.

We turn now to the estimation of $\delta_2(N)$. Using the equivalence of the $L^2(q^N)$ and H^0 norms, and the inequalities (3.1), we establish the following chain of inequalities:

$$\delta_2(N) = |\frac{1}{q_1^N} D[q_2^N D(P^N(q^N)\pi^N f)_1] - (\frac{q_4^N}{\tilde{q}_4} g^N) \frac{1}{\tilde{q}_1} D(\tilde{q}_2 D f_1)|_{0,q^N}$$

which is equivalent to

$$|\frac{1}{q_1^N} q_2^N D^2(P^N(q^N)\pi^N f)_1 + \frac{1}{q_1^N} D q_2^N D(P^N(q^N)\pi^N f)_1 - \frac{q_4^N}{\tilde{q}_4} g^N \frac{1}{\tilde{q}_1} \tilde{q}_2 D^2 f_1 - \frac{q_4^N}{\tilde{q}_4} g^N \frac{1}{\tilde{q}_1} D\tilde{q}_2 D f_1|_0$$

$$\le |\frac{q_2^N}{q_1^N} D^2(P^N(q^N)\pi^N f)_1 - \frac{\tilde{q}_2}{\tilde{q}_1}(\frac{q_4^N}{\tilde{q}_4} g^N) D^2 f_1|_0 + |\frac{Dq_2^N}{q_1^N} D(P^N(q^N)\pi^N f)_1 - \frac{D\tilde{q}_2}{\tilde{q}_1}(\frac{q_4^N}{\tilde{q}_4} g^N) D f_1|_0$$

$$\leq \left|\frac{q_2^N}{q_1^N} D^2(P^N(q^N)\Pi^N f)_1 - \frac{q_2^N}{q_1^N} D^2(\Pi^N f)_1\right|_0 + \left|\frac{q_2^N}{q_1^N} D^2(\Pi^N f)_1 - \frac{\tilde{q}_2}{\tilde{q}_1}\left(\frac{q_4^N}{\tilde{q}_4}\right) g^N D^2 f_1\right|_0 +$$

$$\left|\frac{Dq_2^N}{q_1^N} D(P^N(q^N)\Pi^N f)_1 - \frac{Dq_2^N}{q_1^N} D(\Pi^N f)_1\right|_0 + \left|\frac{Dq_2^N}{q_1^N} D(\Pi^N f)_1 - \frac{D\tilde{q}_2}{\tilde{q}_1}\left(\frac{q_4^N}{\tilde{q}_4} g^N\right) Df_1\right|_0$$

$$\leq \sqrt{k_3}\left|\frac{q_2^N}{q_1^N}\right|_\infty \left|D[(P^N(q^N)-I)\Pi^N f]\right|_{q^N} + \left|\frac{Dq_2^N}{q_1^N}\right|_0 \left|D((P^N(q^N)-I)\Pi^N f)_1\right|_\infty +$$

$$\left|\frac{q_2^N}{q_1^N} D^2(\Pi^N f)_1 - \frac{\tilde{q}_2}{\tilde{q}_1}\left(\frac{q_4^N}{\tilde{q}_4}\right) g^N D^2 f_1\right|_0 + \left|\frac{Dq_2^N}{q_1^N} D(\Pi^N f)_1 - \frac{D\tilde{q}_2}{\tilde{q}_1}\left(\frac{q_4^N}{\tilde{q}_4} g^N\right) Df_1\right|_0 .$$

We thus see that $\delta_2(N)$ can be bounded by four terms which go to zero as $N \to \infty$; the convergence of the first two terms is the result of Proposition 3.3 and the convergence of q^N to $\tilde{q}$, while the convergence of the second two can be argued using the properties of g^N and $q^N \to \tilde{q}$.

We can use this theorem, the convergence properties of the operators Π^N (Proposition 3.1), and the semigroup properties of T^N and T, to establish the final result we need, as a corollary.

Corollary 3.1. Suppose $q^N \to \tilde{q}$. Then

$$|T^N(t;q^N)f - T(t;\tilde{q})f|_{q^N} \to 0 \text{ as } N \to \infty$$

for all $f \in X(\tilde{q})$, uniformly in t on compact intervals in $[0,\infty)$.

We can now invoke the results (see Theorems 2.1 and 2.2) stated in Section 2 to conclude that $\hat{q}$ (obtained there as the limit of an approximating subsequence, $\{\hat{q}^{N_k}\}$) is a solution to the identification problem.

4. Parameter Approximation.

In Section 2, we pose the problem of minimizing $J^N(q)$ over Q. The arguments underlying Theorem 2.1 yield that (under certain assumptions) each N^{th} (approximate) problem has a solution $\hat{q}^N$, and for any convergent subsequence $\{\hat{q}^{N_k}\}$, with $\hat{q}^{N_k} \to \hat{q}$, we have $\hat{q}$ is a solution of the original identification problem. Recall, however, that q_1 and q_2 are functional coefficients, and hence each of the approximate optimization problems is in fact infinite dimensional in nature. In this section, we discuss some methods for approximating these infinite dimensional optimization problems by finite dimensional ones, thus providing numerically tractable problems. This, of course, results in a second, or parameter, approximation that must be considered.

In Section 5, we shall present the results of several numerical test examples. To reduce ill-posedness (see the comments in Section 1) we set $q_1 = \rho \equiv 1$ and search for $q_2 \equiv E$, q_3, q_4, $\tilde{k}$, with q_2 the only functional unknown. We therefore restrict our theoretical discussions here to this case. (We note however that in principle, our methods and ideas can be applied to the estimation of both ρ and E.)

An approach that one might take would be to assume an a priori parameterization for q_2. Thus the estimation of the unknown function becomes the estimation of a set of unknown constants appearing in the parameterization. The convergence theory developed thus far is directly applicable to this method. However, it would only yield results for best approximates (through the criterion on state observations) to q_2 <u>within the fixed a priori parameterization class</u>. Little can be said about convergence to a "best fit parameter" $\hat{q}_2$ from the original parameter set Q.

An alternate approach, which does not require qualitative (e.g., shape) assumptions about the parameter class, is to search for the unknown parameter in a sequence of sets Q^M which are finite dimensional approximations to the set Q. For example, one might search for the unknown parameter in sequences of classes of linear combinations of spline (or members of any other suitably chosen approximation family) basis elements.

We shall consider here two cases: Q^M as a set of linear spline interpolants, and Q^M as a set of cubic spline interpolants. For both cases we need to generalize the theory developed in Section 2, since we now have a "double index" (reflecting approximations for both the parameter and the state space) sequence of iterates, which we would like to argue converges to a solution of the original identification problem.

To be specific, let $Q = Q_2 \times Q_3 \times Q_4 \times Q_5 \subset H^1 \times R^{2+k}$, and assume we have a mapping $i^M : Q_2 \to H^1$. For I the identity map, define $\mathcal{I}^M = i^M \times (I)^{2+k}$, i.e., for $q \in Q$, we have $\mathcal{I}^M(q) = (i^M(q_2), q_3, q_4, q_5)$.

Let $Q^M = \mathcal{I}^M(Q)$. We assume

(4.1) The set $(Q^M)_2 \equiv i^M(Q_2)$ is compact in H^1.

(4.2) For $q_2 \in Q_2$, $i^M(q_2) \to q_2$ in H^1 as $M \to \infty$, and this convergence is uniform in $q_2 \in Q_2$.

The original set Q is assumed to be compact in $H^1 \times R^{2+k}$, so it follows from (4.1), the definition of $\mathcal{I}^M$, and Theorem 2.1 that for each N and M, a solution $\hat{q}_M^N$ exists to the problem of minimizing J^N over Q^M. From the definition $Q^M = \mathcal{I}^M(Q)$, we see that there exists $\bar{q}_M^N \in Q$ such that $\mathcal{I}^M(\bar{q}_M^N) = \hat{q}_M^N$ for each N and M. But the compactness of the original set Q then implies the existence of some subsequence $\{\bar{q}_{M_k}^{N_j}\}$ and an element $\hat{q} \in Q$ such that $\bar{q}_{M_k}^{N_j} \to \hat{q}$ in Q; moreover, this subsequence may be chosen so that both $N_j \to \infty$ and $M_k \to \infty$. The limit $\hat{q}$ is in fact a solution to the problem of minimizing J over Q; this claim is verified as follows: From the definition $\hat{q}_{M_k}^{N_j}$ we have

$$J^{N_j}(\hat{q}_{M_k}^{N_j}) \leq J^{N_j}(q) \quad , \qquad \text{for } q \in Q^{M_k} .$$

This implies

$$\text{(4.3)} \qquad J^{N_j}(\hat{q}_{M_k}^{N_j}) \leq J^{N_j}(\mathcal{I}^{M_k}(q)) \quad , \qquad \text{for } q \in Q .$$

But $|\hat{q}_{M_k}^{N_j} - \hat{q}\ | \leq |\mathcal{I}^{M_k}(\bar{q}_{M_k}^{N_j}) - \bar{q}_{M_k}^{N_j}| + |\bar{q}_{M_k}^{N_j} - \hat{q}\ |$, and thus $\hat{q}_{M_k}^{N_j} \to \hat{q}$ in Q as $N_j \to \infty$, $M_k \to \infty$ follows from (4.2), the definition of $\mathcal{I}^{M_k}$, and $\bar{q}_{M_k}^{N_j} \to \hat{q}$. If we take the limit in (4.3) as N_j, $M_k \to \infty$, we see that $J(\hat{q}\ ) \leq J(q)$ for $q \in Q$. Here we have used Theorem 2.2 with the observation that the convergence statement $z^N(t;q^N) \to z(t;\tilde{q})$ for any $q^N \to \tilde{q}$ is still valid if replaced by $z^N(t;q^j) \to z(t;\tilde{q})$ as j, $N \to \infty$, for any $q^j \to \tilde{q}$; this can be seen using a re-indexing argument. These remarks are summarized in the following theorem.

<u>Theorem 4.1.</u> Let $Q^M = \mathcal{I}^M(Q)$ where (4.1) and (4.2) are satisfied. Let $\hat{q}_M^N$ be a solution to the problem of minimizing J^N over Q^M. Then for any convergent subsequence $\{\hat{q}_{M_k}^{N_j}\}$ with N_j, $M_k \to \infty$ and $\hat{q}_{M_k}^{N_j} \to \hat{q}$, the limit $\hat{q}$ is a solution to the problem of minimizing J over Q.

We first consider the above results applied to the case where the Q^M are sets of linear spline interpolants. Let $S^1(\Delta^M)$ represent the subspace of piecewise linear splines corresponding to the partition $\Delta^M = \{x_i\}_{i=0}^M$, $x_i = \frac{i}{M}$, and let $i^M : H^1 \to S^1(\Delta^M)$ denote the standard linear spline interpolating operator. If, in addition to assuming Q_2 is compact in H^1, we assume Q_2 satisfies $Q_2 \subset \{q_2 \in H^2 \mid |D^2 q_2|_0 \le K\}$, then it is not difficult to show that (4.1) and (4.2) are true for Q^M and i^M as defined above. From a standard representation result for linear interpolating splines [19, p.12], we infer the continuity of the operator i^M as a mapping from H^1 to H^1, and the compactness of $(Q^M)_2 = i^M(Q_2)$ in H^1 follows immediately. To establish (4.2) we appeal to standard estimates such as (2.17) and (2.18) in [19]. Having verified (4.1) and (4.2), we now state

Theorem 4.2. Suppose $Q = Q_2 \times Q_3 \times Q_4 \times Q_5$ is a compact subset of $H^1 \times R^{2+k}$ with Q_2 additionally satisfying $Q_2 \subset \{q_2 \in H^2 \mid |D^2 q_2|_0 \le K\}$. Let $Q^M = \mathcal{I}^M(Q)$ where $\mathcal{I}^M \equiv i^M \times (I)^{2+k}$, and i^M is the linear spline interpolating operator. If $\hat{q}^N_M$ represents a solution obtained from minimizing J^N over Q^M, then for any subsequence $\{\hat{q}^{N_j}_{M_k}\}$ of $\{\hat{q}^N_M\}$ such that as $N_j, M_k \to \infty$, $\hat{q}^{N_j}_{M_k} \to \hat{q}$ in Q, we have that $\hat{q}$ is a minimizer for J over Q.

Under slightly stronger assumptions on the set Q, we can develop a similar convergence result using cubic spline approximations to q_2. Let $S^3(\Delta^M)$ be the subspace of C^2 cubic splines corresponding to the partition Δ^M, and let $i^M : C^1 \to S^3(\Delta^M)$ denote the standard cubic spline interpolating operator (see Sections 2 and 3 for details). We assume Q_2 is a compact subset of C^1 satisfying also $Q_2 \subset \{q_2 \in H^2 \mid |D^2 q_2|_0 \le K\}$. We again may use standard interpolating spline representations (see [19, p. 45]) to conclude that i^M is a continuous operator from C^1 to H^1, from whence it follows that $(Q^M)_2$ is compact in H^1. To verify (4.2), we again refer to (4.19) and (4.20) in

[19]. Thus we have

Theorem 4.3. Suppose $Q = Q_2 \times Q_3 \times Q_4 \times Q_5$ is a compact subset of $C^1 \times R^{2+k}$ with $Q_2 \subset \{q_2 \in H^2 \mid |D^2 q_2|_0 \leq K\}$. Let $Q^M = \mathcal{I}^M(Q)$ where $\mathcal{I}^M \equiv i^M \times (I)^{2+k}$, and i^M is the cubic spline interpolating operator. If $\hat{q}_M^N$ represents a solution obtained from minimizing J^N over Q^M, then there exists $\hat{q} \in Q$ which minimizes J over Q, and a subsequence $\left\{\hat{q}_{M_k}^{N_j}\right\}$ of $\left\{\hat{q}_M^N\right\}$ such that as $N_j, M_k \to \infty$, $\hat{q}_{M_k}^{N_j} \to \hat{q}$.

In the next section we present mumerical findings for double (state and parameter) approximation schemes such as those described here.

5. Numerical Implementation and Examples. Recall from Section 2 that the approximating identification problem is:

Given $\hat{y}$, minimize $J^N(q) = \sum_{i=1}^{n} |\hat{y}_i - \xi^N(t_i,q)|^2$ over $q \in Q$ (where ξ^N involves point evaluations, in space, of the first component of z^N) subject to $z^N(\cdot;q)$ satisfying the following ordinary differential equation:

$$\dot{z}^N(t) = A^N(q)z^N(t) + P^N(q)G(t;q)$$

$$z^N(0) = P^N(q)z_0(q).$$

(We continue our discussions in terms of the transformed system (1.5) and criterion (1.6) even though the numerical examples summarized in this section involve "data" for the original system (1.1)-(1.4) used in conjunction with the criterion (1.7).) Since $z^N \in X^N(q)$, z^N has a representation in terms of the basis elements of $X^N(q)$, $z^N(t;q) = \sum_{i=1}^{2N+3} w_i^N(t;q)\beta_i^N(x;q)$. If we let $[A^N(q)]$ and $[f^N]$ be the matrix and vector representations, respectively of $A^N(q)$ and $P^N(q)f$ (where f is an arbitrary function in $X(q)$) with respect to the basis elements of $X^N(q)$, and let $w^N(t;q) \equiv \text{col}(w_1^N(t;q), \ldots, w_{2N+3}^N(t;q))$, then $w^N(t;q)$ solves the following system of ordinary differential equations:

$$\dot{w}^N(t;q) = [A^N(q)]w^N(t;q) + [G^N(t;q)]$$

$$w^N(0;q) = [z_0^N(q)].$$

As in [5], this can be written more explicitly as:

$$(5.1)\quad \begin{aligned} Q^N \dot{w}^N(t;q) &= K^N w^N(t;q) + R^N G(t;q) \\ Q^N w^N(0;q) &= R^N z_0(q) \end{aligned}$$

where Q^N and K^N are matrices, with elements described by $(Q^N)_{i,j} = \langle \beta_i^N, \beta_j^N \rangle_q$, $(K^N)_{i,j} = \langle \beta_i^N, A(q)\beta_j^N \rangle_q$, and $(R^N f)_i = \langle \beta_i^N, f \rangle_q$ for $f \in X(q)$. Due to the form of the B-spline basis elements we have chosen (see Section 2), Q^N can be stored as a banded symmetric matrix; this banded, symmetric structure permits more efficient computations and requires less storage space. The matrix K^N has a similar sparse (although not symmetric) structure.

Each element of the matrices Q^N and K^N, and of the vector $R^N f$ depends continuously on q, therefore the representations $[A^N(q)]$ and $[f^N]$ are continuous in q. The basis elements for $X^N(q)$ depend linearly on q, and hence are continuous in q, which implies $q \to P^N(q)f$ and $q \to T^N(t;q)f$ (we note $T^N(t;q) = \exp(A^N(q)t)$ since $A^N(q)$ is a bounded operator) are continuous mappings (recall this was a necessary condition in Theorem 2.1).

We note that in the case where q_1 and q_2 are assumed to be constant, or to have a representation as, for example, a linear combination of spline elements, then the computations can be done more efficiently; in such cases, the numerical quadratures required to compute the inner products which form Q^N and K^N need be performed only once for each N. Then, to construct Q^N and K^N the appropriate multiples or linear combinations of these stored values are computed.

Many of the computations in the software package used to generate the following examples were done with IMSL subroutines (for example, the optimization, and the solution of the differential equation in (5.1)). Although much modification was necessary for the present application, the core of the package was developed by James Crowley [11]. The examples were computed either on an IBM VM/370, or a CDC 6600.

The optimization is done using a Levenberg-Marquardt algorithm. For fixed N, each iteration in the optimization is performed as follows. Given q, beginning at time zero ($t_1=0$), a Cholesky decomposition method is used to solve (5.1) for $\dot{w}^N(t;q)$ and $w^N(t_1;q)$; this is then integrated using Gear's method to obtain $w^N(t_2;q)$. We use the components of the vector $w^N(t_2;q)$ to recover $z_1^N(t_2;q)$ as the linear combination of the first components of the basis elements. The vector $\xi^N(t_2,q)$ is $z_1^N(t_2;q)$ evaluated at each of the spatial observation points. Using $w^N(t_2;q)$ as the initial value, (5.1) is solved again for $t \in [t_2,t_3]$, $\xi^N(t_3,q)$ is obtained, and this procedure is repeated until $\xi^N(t_i,q)$ has been evaluated at all times t_i; then $J^N(q)$ can be computed as the sum of the residuals, $|\hat{y}_i - \xi^N(t_i,q)|^2$. The data $\{\hat{y}_i\}$ is read in and stored at the beginning.

In the selection of examples to follow, the "data" has been generated with an independent finite difference scheme (an implicit method [17] was modified for our boundary conditions and the variable coefficient, $q_2(x)$) applied to the model with a priori chosen "true" values q^* of the parameters. In all examples, $q_1(x)$ is taken to be identically one (this is done to reduce ill-posedness, as mentioned in Section 1). We begin each example with an initial guess, and a value of N; we solve (ID^N), to get converged values, $\bar{q}^N$ (these are numerical approximations (to $\hat{q}^N$) that result from the Levenberg-Marquardt algorithm), which we then use as starting values for the next value of N. So, in Example 5.1 (below) we begin with $N=4$ and a guess q^0, and generate $\bar{q}^4$. We then start with $\bar{q}^4$ at $N=8$, and generate $\bar{q}^8$.

We remind the reader that the computations reported on below were carried out using "data" for the system (1.1)-(1.4) with criterion (1.7) and an appropriate approximate criterion for the N^{th} problem. (We have also successfully tested the methods on similar examples with the transformed system (1.5) and criterion (1.6), although, of course, this is not the typical formulation of the inverse problem for which data will be available.)

Example 5.1. For our first example we used "data" consisting of observations at $x=0$ and times $t = .25, .5, .75, \ldots, 2.0$. This is meant to simulate the situation in "surface seismic" experiments where only data at the surface are available. The source term was chosen as $s(t;\tilde{k}) = q_5(1-e^{-5t})e^{q_6 t}$, a function which rises to a peak quickly and then gradually diminishes to zero; again this attempts to mimic the situation in seismic experiments. We assume vanishing initial conditions and seek to estimate a constant elastic modulus q_2 as well as the boundary parameters q_3, q_4 and the source parameters $\tilde{k} = (q_5,q_6)$. True values along with our estimates are given in the results summarized in Table 5.1. Graphs comparing the true solution at the surface $u(t,0;q^*)$ with the approximate solution $u^N(t,0;\bar{q}^N)$ are shown in Figure 5.1. We also tested the method on this example using "data" for more spatial observations (data at $x=0$, $.5$, 1.0 and at $t = .5, 1.0, 1.5$) with our findings given in Table 5.2. Based on these computations and a number of other tests, we suggest that there appears to be little difficulty with our method in the case where only one spatial observation is available as long as a sufficient number of time observations are available.

TABLE 5.1

Initial Guess	Converged Values N=4	Converged Values N=8	True Values
$q_2^0 = 2.0$	$\bar{q}_2^4 = 2.96001$	$\bar{q}_2^8 = 3.0001$	$q_2^* = 3.0$
$q_3^0 = -1.0$	$\bar{q}_3^4 = -1.98861$	$\bar{q}_3^8 = -1.99012$	$q_3^* = -2.0$
$q_4^0 = 2.0$	$\bar{q}_4^4 = 0.97428$	$\bar{q}_4^8 = 1.00683$	$q_4^* = 1.0$
$q_5^0 = 1.5$	$\bar{q}_5^4 = 1.97135$	$\bar{q}_5^8 = 1.99809$	$q_5^* = 2.0$
$q_6^0 = -0.5$	$\bar{q}_6^4 = -0.98500$	$\bar{q}_6^8 = -1.00506$	$q_6^* = -1.0$
No. of Iterations[1]	11	2	
R.S.S.[2]	0.659×10^{-5}	0.119×10^{-5}	
CPU[3]	125.363	84.688	

[1]Number of iterations in the optimization algorithm.

[2]Residual sum of squares = $J^N(\bar{q}^N)$.

[3]The CPU time given in seconds.

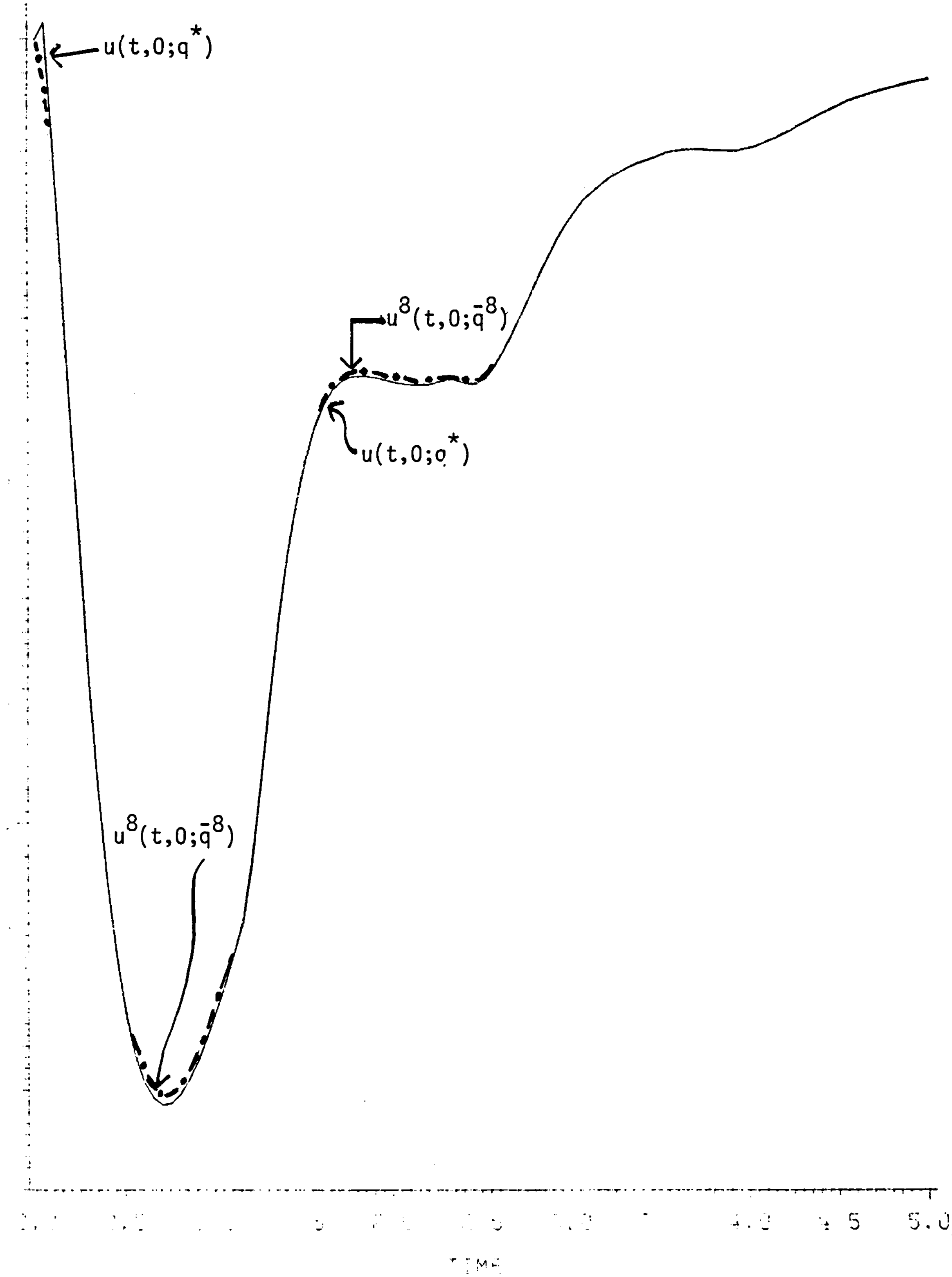

FIGURE 5.1

TABLE 5.2

Initial Guess	Converged Values N = 4	Converged Values N = 8	True Values
$q_2^0 = 2.0$	$\bar{q}_2^4 = 2.98515$	$\bar{q}_2^8 = 2.99378$	$q_2^* = 3.0$
$q_3^0 = -1.0$	$\bar{q}_3^4 = -1.92304$	$\bar{q}_3^8 = -2.01999$	$q_3^* = -2.0$
$q_4^0 = 2.0$	$\bar{q}_4^4 = 1.01302$	$\bar{q}_4^8 = 1.00285$	$q_4^* = 1.0$
$q_5^0 = 1.5$	$\bar{q}_5^4 = 1.97120$	$\bar{q}_5^8 = 2.00578$	$q_5^* = 2.0$
$q_6^0 = -0.5$	$\bar{q}_6^4 = -1.03296$	$\bar{q}_6^8 = -0.99172$	$q_6^* = -1.0$
No. of Iterations	12	5	
R.S.S.	0.235×10^{-5}	0.441×10^{-5}	
CPU	117.597	147.323	

Example 5.2. In this example we compared the performance of our method on problems with "noisy data" with that on those without noise in the data. We used the same source term as that in Example 5.1, zero initial conditions, but a "true" parameterized elastic modulus $E(x) = 3/2 + 1/\pi \text{ Arctan } [q_{21}(x-q_{22})]$. Data for observations at x = 0.0, 0.5, 1.0 and t = .416, .832, 1.248, 1.664, 2.08, 2.496 were used. Results for the case of data without noise are summarized in Table 5.3, while findings employing data with a noise level of approximately 3% are given in Table 5.4. In both cases, the method converges nicely but as one might expect, the converged values of the parameters do not agree with the true parameters in the case of noisy data. In Figures 5.2, 5.3, 5.4 and 5.5, we graphically depicted the curves for $\bar{E}^N$ and E^* in several cases.

Example 5.3. In this example we illustrate the ideas discussed in Section 4 regarding parameter approximation in the set of linear and cubic splines. We do not assume an a priori shape for the elastic modulus E(x), the "true" value of which is given by $E^*(x) = 3/2 + \tanh [6(x-.5)]$. Rather we first search for E in the class of linear spline approximations to E^*. We then carry out the search using cubic splines. Initial conditions are $u(0,x) = e^x$, $u_t(0,x) = -3e^x$ and no source term was assumed (i.e., $s \equiv 0$). Data for observations at 3 spatial points (x = 0.0, 0.5, 1.0) and 6 time points (t = .16, .32, ..., 1.0) were used. Figure 5.6 depicts graphs of the true modulus E^*, the initial guess E^0, and the converged estimate $\bar{E}^4$ where we used linear splines (with 4 basis elements -- M = 3 in the notation of Section 4) to approximate E and cubic splines (N = 4) to approximate the state. At the same time we searched for the boundary parameters q_3, q_4 (true values $q_3^* = -1.0$, $q_4^* = 3.0$) and obtained converged

TABLE 5.3

Initial Guess	Converged Values N = 4	Converged Values N = 8	True Values
$q_{21}^{0} = 1.0$	$\bar{q}_{21}^{4} = 2.97352$	$\bar{q}_{21}^{8} = 2.99994$	$q_{21}^{*} = 3.0$
$q_{22}^{0} = 1.0$	$\bar{q}_{22}^{4} = 0.51115$	$\bar{q}_{22}^{8} = 0.50053$	$q_{22}^{*} = 0.5$
$q_{3}^{0} = -2.0$	$\bar{q}_{3}^{4} = -0.99892$	$\bar{q}_{3}^{8} = -1.00026$	$q_{3}^{*} = -1.0$
$q_{4}^{0} = 2.0$	$\bar{q}_{4}^{4} = 3.05138$	$\bar{q}_{4}^{8} = 3.01070$	$q_{4}^{*} = 3.0$
$q_{5}^{0} = 1.0$	$\bar{q}_{5}^{4} = 2.00322$	$\bar{q}_{5}^{8} = 2.00056$	$q_{5}^{*} = 2.0$
$q_{6}^{0} = -2.0$	$\bar{q}_{6}^{4} = -1.01163$	$\bar{q}_{6}^{8} = -1.00217$	$q_{6}^{*} = -1.0$
No. of Iterations	13	3	
R.S.S.	0.1025×10^{-3}	0.82859×10^{-5}	
CPU	269.696	196.335	

<u>TABLE 5.4</u>

(NOISY DATA)

Initial Guess	Converged Values N = 4	Converged Values N = 8	True Values
$q_{21}^{0} = 1.0$	$\bar{q}_{21}^{4} = 3.30536$	$\bar{q}_{21}^{8} = 3.29222$	$q_{21}^{*} = 3.0$
$q_{22}^{0} = 1.0$	$\bar{q}_{22}^{4} = 0.53802$	$\bar{q}_{22}^{8} = 0.53115$	$q_{22}^{*} = 0.5$
$q_{3}^{0} = -2.0$	$\bar{q}_{3}^{4} = -0.86648$	$\bar{q}_{3}^{8} = -0.86017$	$q_{3}^{*} = -1.0$
$q_{4}^{0} = 2.0$	$\bar{q}_{4}^{4} = 2.99610$	$\bar{q}_{4}^{8} = 2.96002$	$q_{4}^{*} = 3.0$
$q_{5}^{0} = 1.0$	$\bar{q}_{5}^{4} = 2.09207$	$\bar{q}_{5}^{8} = 2.09295$	$q_{5}^{*} = 2.0$
$q_{6}^{0} = -2.0$	$\bar{q}_{6}^{4} = -1.15602$	$\bar{q}_{6}^{8} = -1.15571$	$q_{6}^{*} = -1.0$
No. of Iterations	13	2	
R.S.S.	0.6509×10^{-3}	0.476×10^{-3}	
CPU	270.11	136.87	

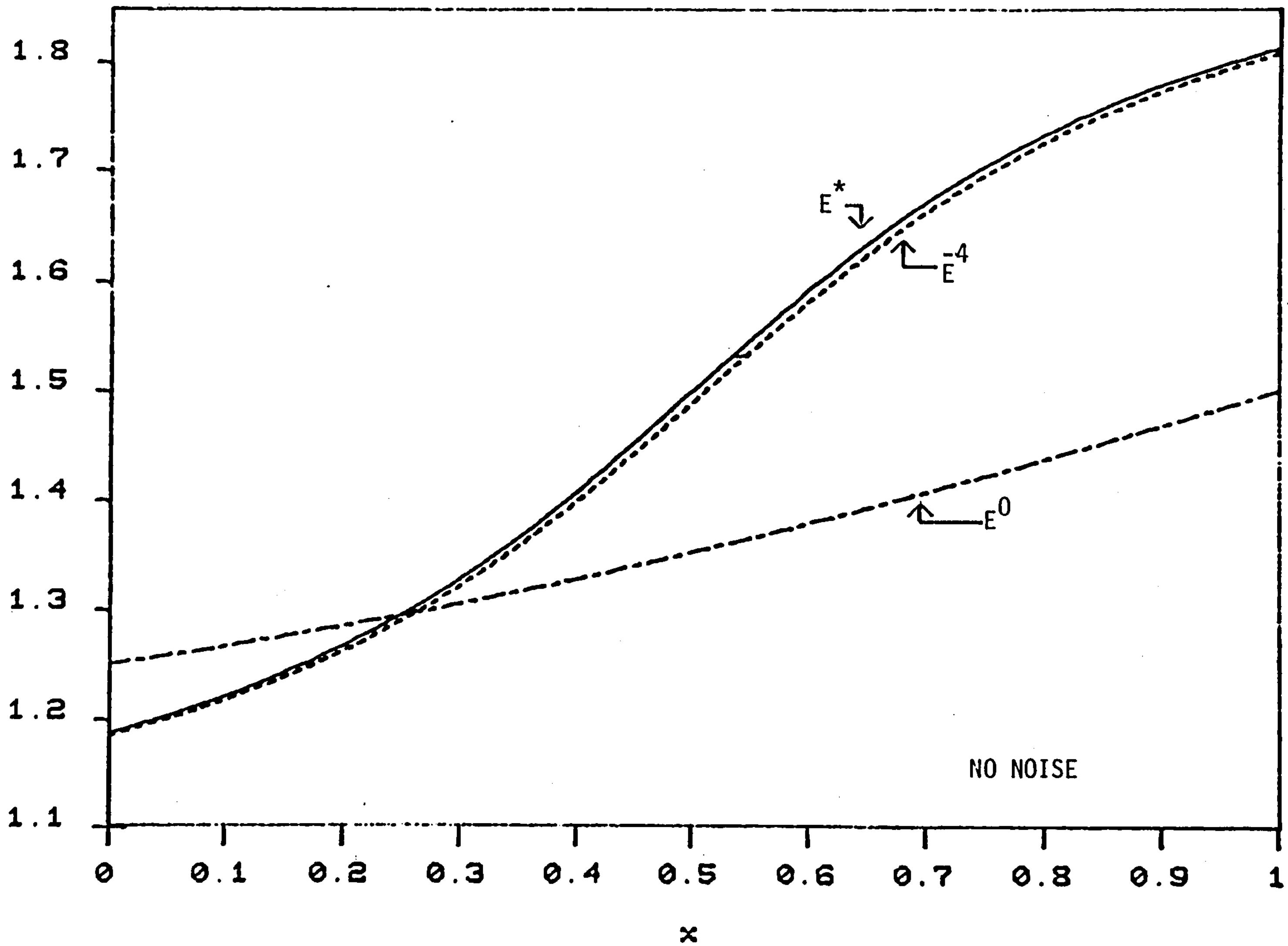

FIGURE 5.2

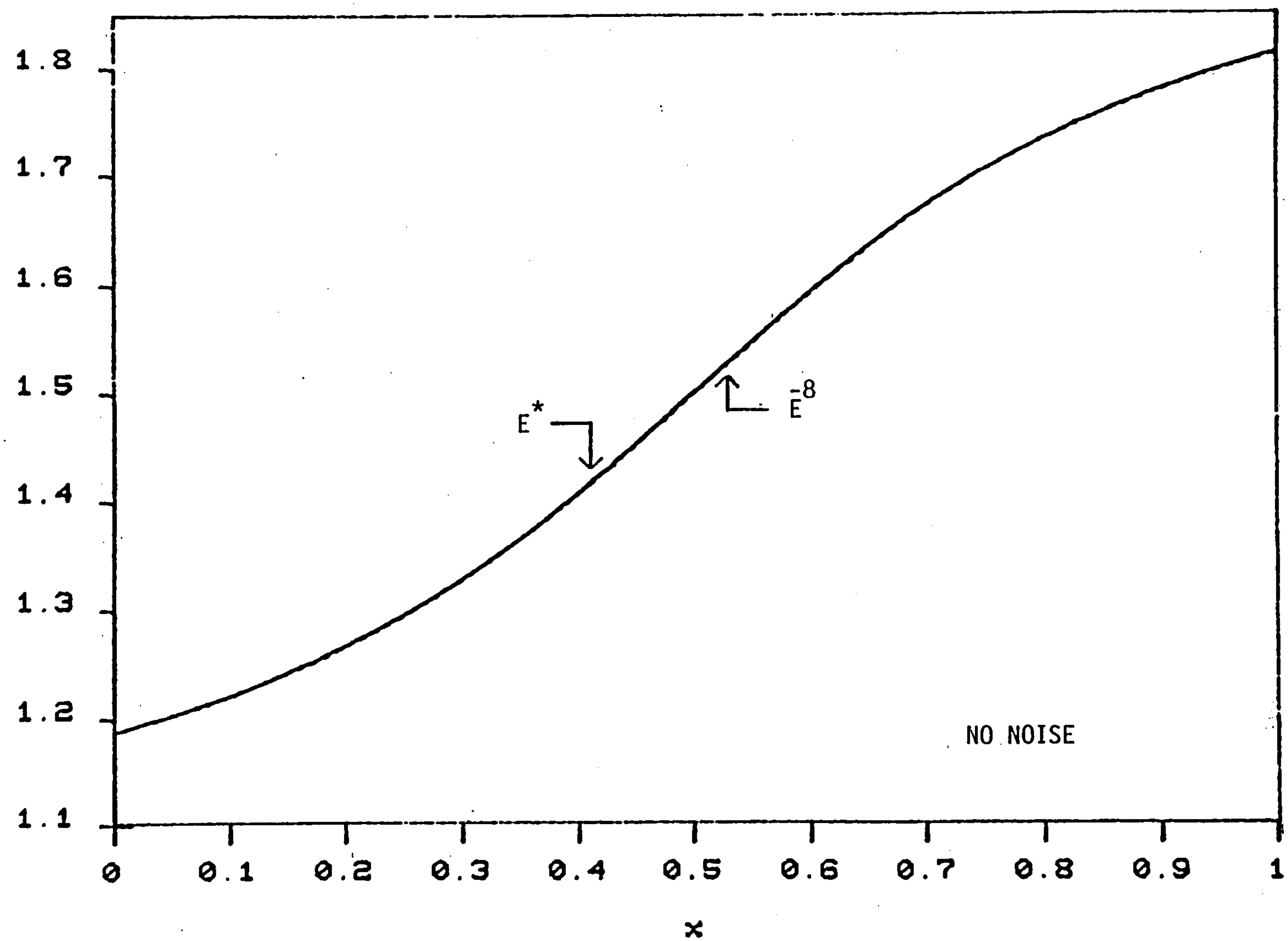
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
x
E^*
$\bar{E}^8$
NO NOISE

FIGURE 5.3

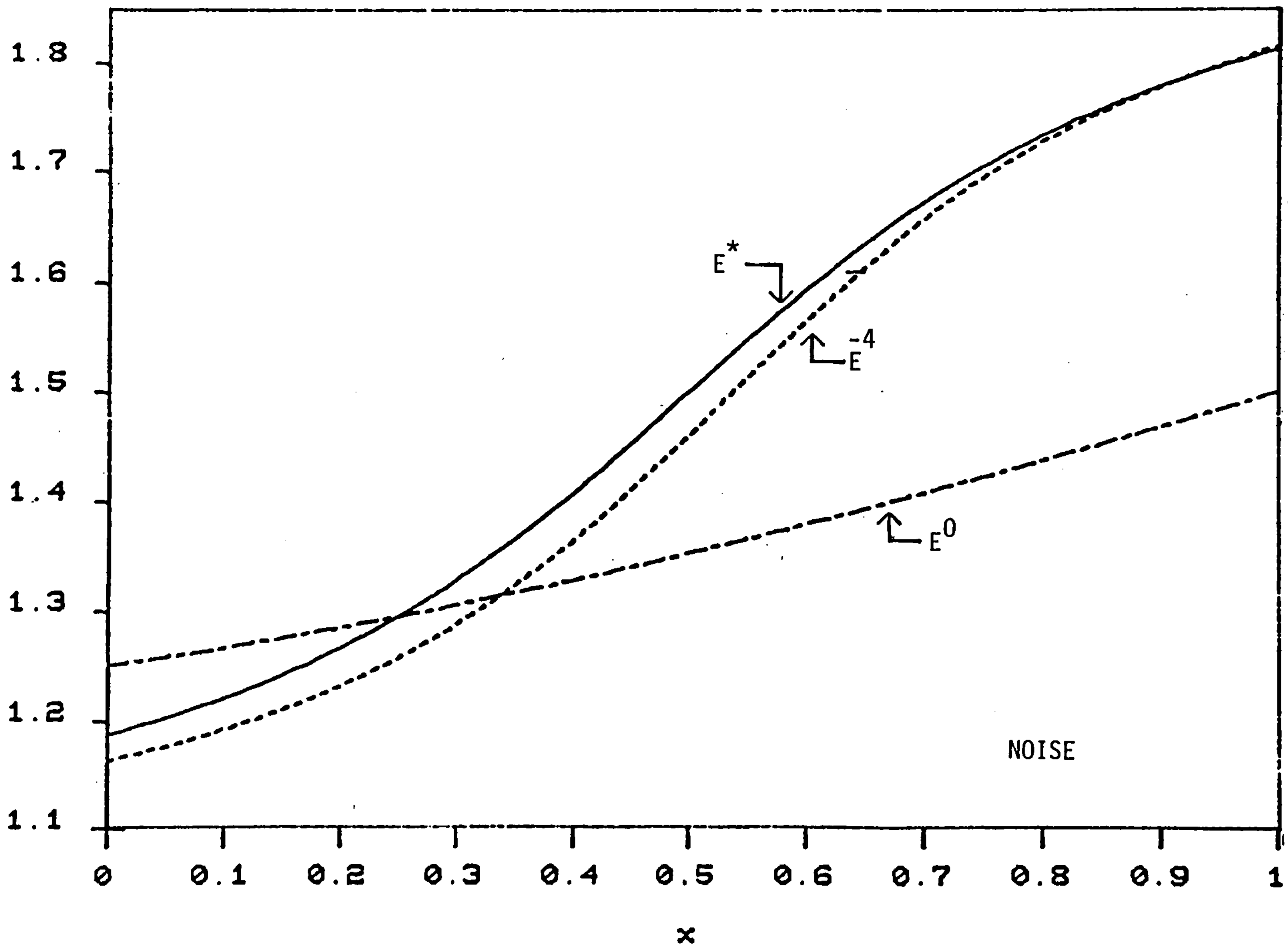

FIGURE 5.4

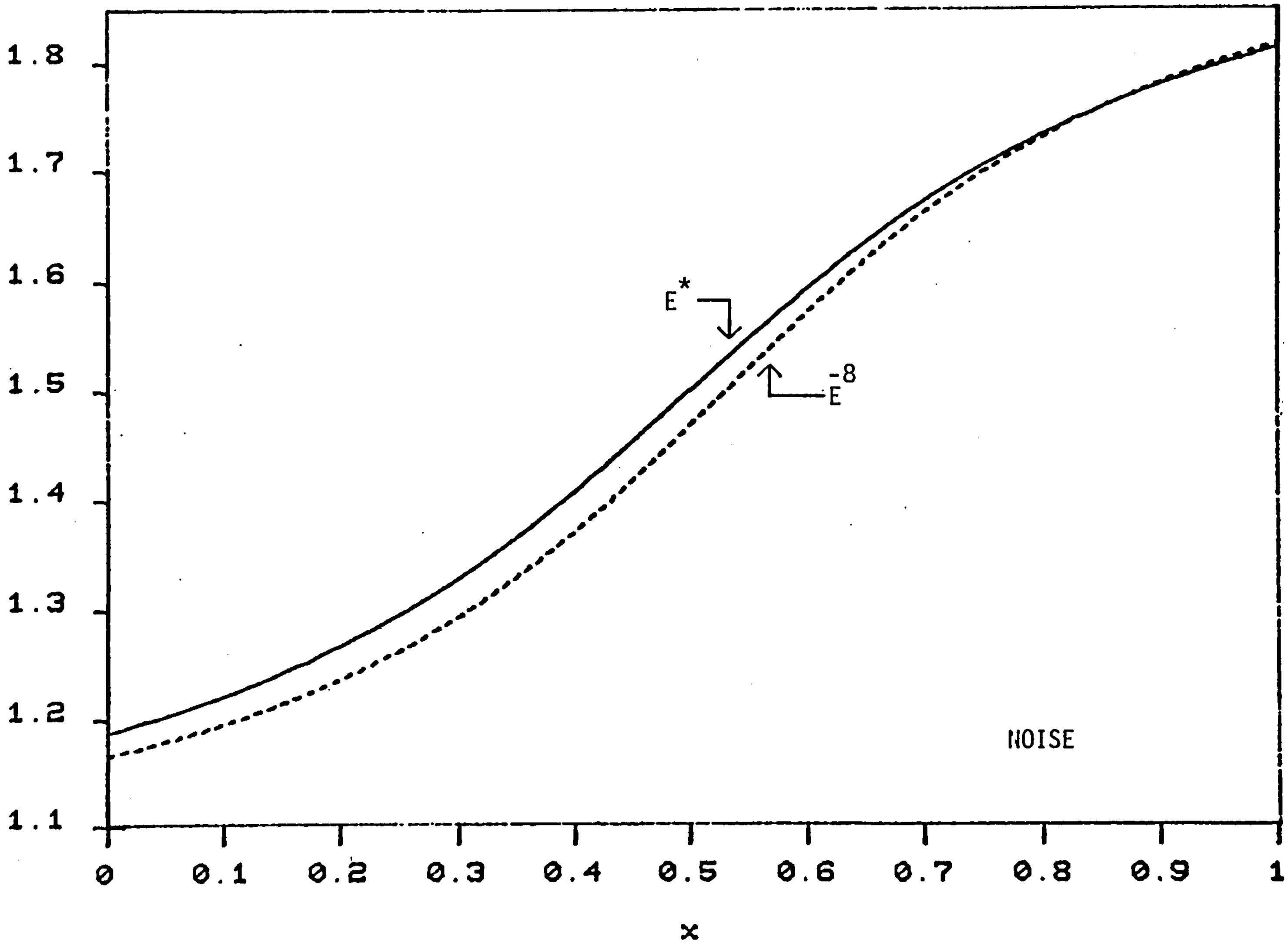

FIGURE 5.5

estimates $\bar{q}_3^4 = -1.05425$, $\bar{q}_4^4 = 3.3576$ with a CPU time of 38 seconds and R.S.S. = 0.255×10^{-2}. Figure 5.7 contains graphs similar to those in Fig. 5.6 except $N = 16$ was used in the state approximations. Boundary parameter estimates corresponding to $\bar{E}^{16}$ were $\bar{q}_3^{16} = -1.10063$, $\bar{q}_4^{16} = 3.07049$ with CPU time of 118 seconds and R.S.S. = 0.472×10^{-4}. The error (in the H^0 norm) in estimating E^* in each case was calculated to be $|E^* - \bar{E}^4| = .081$ and $|E^* - \bar{E}^{16}| = .030$.

We carried out similar calculations for the same example in which we employed cubic splines (M=1 in the notation of Section 4, i.e. 4 basis elements) for the parameter approximations. The graphs of E^*, E^0 and $\bar{E}^{16}$ are compared in Figure 5.8. In this second test we did not search on the boundary parameters q_3, q_4 but rather held them fixed at their "true" values. The error at the converged parameter was $|E^* - \bar{E}^{16}| = .109$, with R.S.S. = 0.293×10^{-2} and a CPU time of 178 seconds.

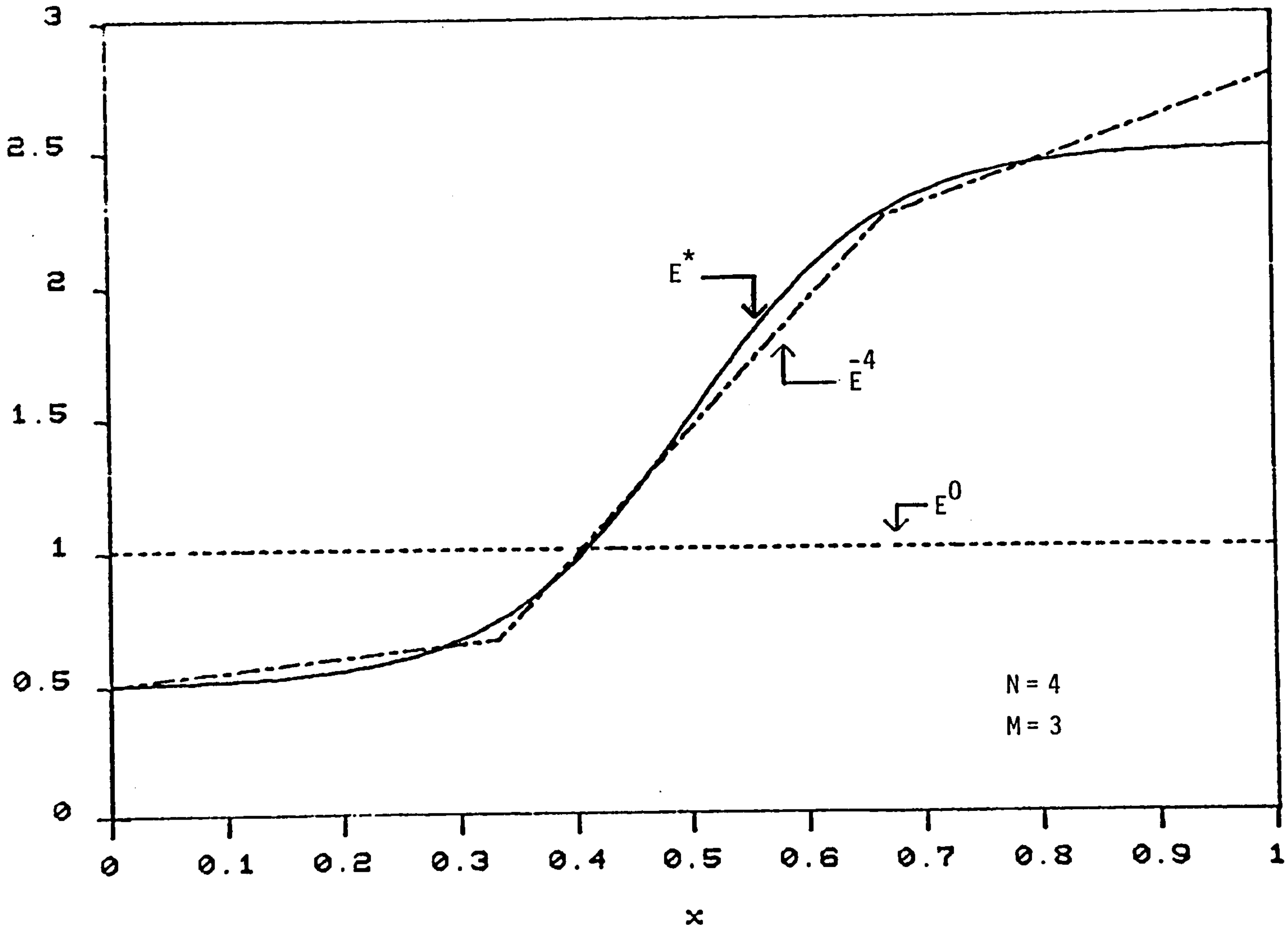

FIGURE 5.6

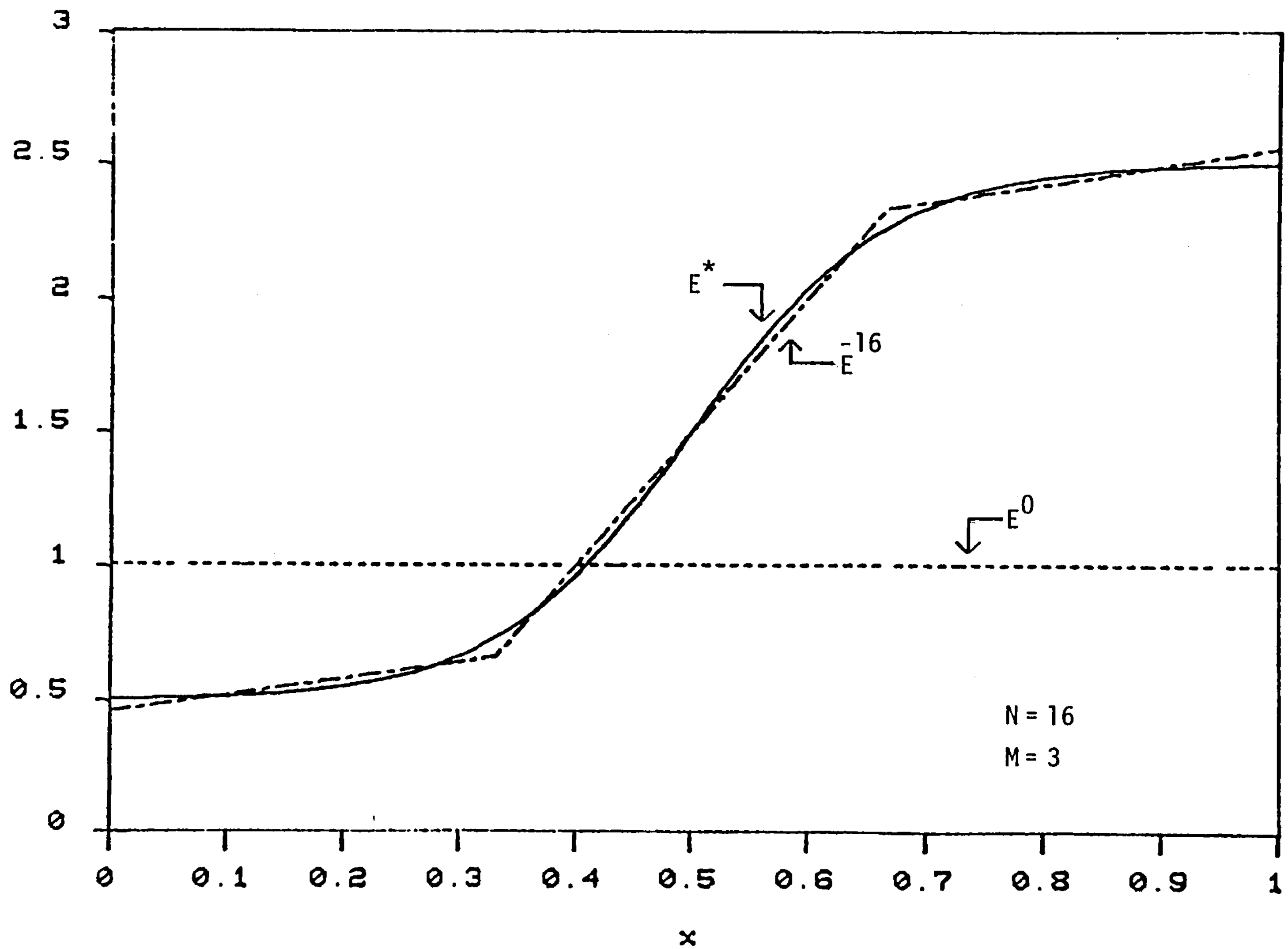

FIGURE 5.7

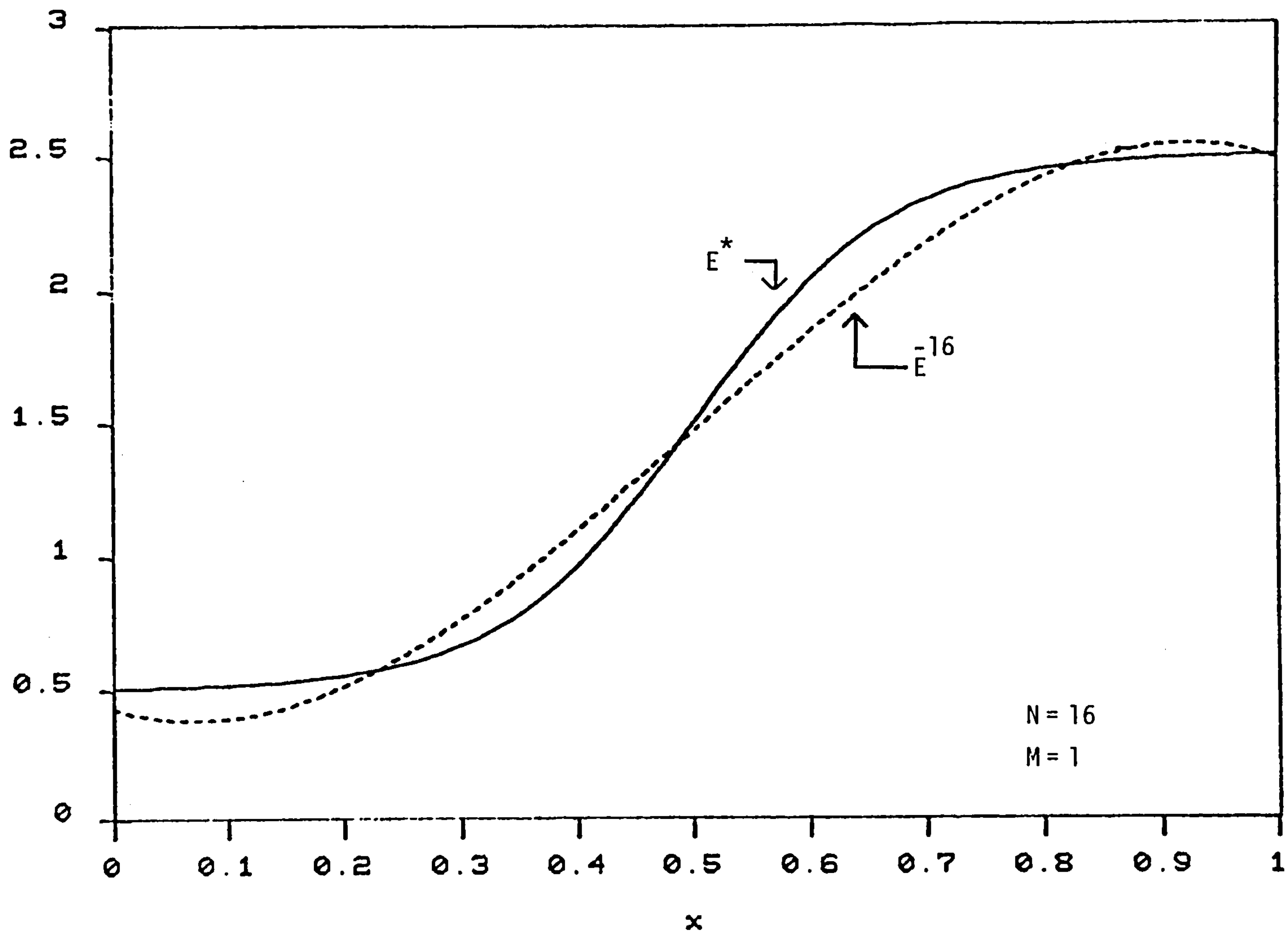

FIGURE 5.8

6. Concluding Remarks.

We have presented in this paper both theoretical and numerical results using some of our ideas involving spline approximations for inverse or parameter estimation problems for hyperbolic systems. Among the novel features is the capability of estimating variable coefficients and boundary parameters with methods that are both theoretically sound and readily implementable. Our techniques (reported on earlier, [6]) involve the use of parameter dependent basis elements for the approximation subspaces in a Galerkin type semi-discrete scheme.

While we have focused on 1-dimensional space domain problems here, our ideas are in principle applicable to problems in 2 and 3 dimensional domains. We have devoted some thought to such problems in connection with use of basis elements that are tensor products of 1-D elements. These ideas offer some promise, given the parallelism that would be inherent in the resulting algorithms and given the emerging technology related to supercomputers and array processors. However, there are other ideas that also offer great promise; in particular, there are those involving spectral methods such as the tau-Legendre for which we have reported preliminary findings in [4]. A fundamental difference between these techniques and those proposed in this paper is that in the tau-Legendre one does not require the approximation subspace basis elements to satisfy the boundary conditions. Instead the boundary conditions are essentially imposed as side constraints adjoined to the Galerkin type differential equations. This can offer significant computational advantages, especially in higher dimensional domain problems. We are currently pursuing investigations of these ideas.

In closing we remark that the theoretical results presented above only guarantee convergence of subsequences $\left\{\hat{q}^{N_k}\right\}$ to a minimizer $\hat{q}$ for J. But for the class of problems investigated here and for a number of other types of inverse problems we have studied, we have in practice only observed (numerically) convergence of the original sequence $\left\{\hat{q}^N\right\}$. This has been our experience even in examples with noisy data and may be due in many cases to the fact that the original problem of minimizing J over Q has a unique solution $\hat{q}$. In this situation, elementary and quite standard arguments can be employed to actually establish convergence of $\left\{\hat{q}^N\right\}$ itself to $\hat{q}$.

Acknowledgement.

The authors would like to express their sincere appreciation to G. Moeckel (Mobil Oil Co.), R. Ewing (U. Wyoming), and K. Kunisch (U. Graz) for stimulating discussions during the course of some of the work reported above. They are also grateful for the support and hospitality received during their visit at Southern Methodist University where a substantial portion of the investigations reported on here were carried out.

References

1. A. Bamberger, G. Chavent, and P. Lailly, About the stability of the inverse problem in 1-D wave equations-application to the interpretation of seismic profiles, Appl. Math. Optim. 5 (1979), p. 1-47.

2. H. T. Banks and J. M. Crowley, Parameter estimation for distributed systems arising in elasticity, Proc. Symposium on Engineering Sciences and Mechanics, (National Cheng Kung University, Tainan, Taiwan, Dec. 28-31, 1981), p. 158-177; (LCDS Tech. Rep. 81-24, November, 1981, Brown University).

3. H. T. Banks, J. M. Crowley and K. Kunisch, Cubic spline approximation techniques for parameter estimation in distributed systems. (LCDS Tech. Rep. 81-25, November, 1981, Brown University); IEEE Trans. Auto. Control., AC-28 (1983), p. 773-786.

4. H. T. Banks, K. Ito and K. A. Murphy, Computational methods for estimation of parameters in hyperbolic systems, in Conf. on Inverse Scattering: Theory and Application (J. B. Bednar, et.al, eds.), SIAM, Philadelphia, 1983, p. 181-193.

5. H. T. Banks and K. Kunisch, An approximation thoery for nonlinear partial differential equations with applications to identification and control, SIAM J. Control and Optimization, 20 (1982), p. 815-849.

6. H. T. Banks and K. A. Murphy, Inverse problems for hyperbolic systems with unknown boundary parameters, in Control Theory for Distributed Parameter Systems and Applications (F. Kappel, et.al., eds.), Springer-Verlag, Berlin, 1983, p. 35-44.

7. H. T. Banks and I. G. Rosen, Fully discrete approximation methods for the estimation of parabolic systems and boundary parameters, LCDS Technical Report No. 84-19, Brown University, 1984

8. K. P. Bube and R. Burridge, The one-dimensional inverse problem of reflection seismology, SIAM Rev. 25 (1983), p. 497-559.

9. J. A. Burns and E. M. Cliff, An approximation technique for the control and identification of hybrid systems, in Dynamics and Control of Large Flexible Spacecraft, 3rd VPISU/AIAA Symp., 1981, p. 269-284.

10. G. Chavent, About the stability of the optimal control solution of inverse problems, in Inverse and Improperly Posed Problems in Differential Equations, (G. Anger, ed.), Akademic-Verlag, Berlin, 1979, p. 45-58.

11. J. M. Crowley, Numerical methods of parameter identification for problems arising in elasticity, Ph.D. Thesis, Brown University, May, 1982.

12. B. Engquist and A. Majda, Absorbing boundary conditions for the numerical simulation of waves, Math. Comp. 31 (1977), p. 629-651.

13. N. L. Guinasso and D. R. Schink, Quantitative estimates of biological mixing rates in abyssal sediments, J. Geophys. Res. 80 (1975), p. 3032-3043.

14. K. A. Murphy, A spline-based approximation method for inverse problems for a hyperbolic system including unknown boundary parameters, Ph.D. Thesis, Brown University, May, 1983.

15. A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, App. Math Sciences, Vol. 44, Springer, New York, 1983.

16. P. M. Prenter, Splines and Variational Methods, New York: Wiley-Interscience, 1975.

17. R. D. Richtmyer and K. W. Morton, Difference Methods for Initial-Value Problems, New York: Wiley-Interscience, 1967.

18. I. G. Rosen, Approximation methods for the identification of hybrid systems involving the vibration of beams with tip bodies, CSDL-P-1893, Charles Stark Draper Laboratory, Cambridge, MA.

19. M. H. Schultz, Spline Analysis, Englewood Cliffs, N.J.: Prentice-Hall, 1973.

20. J. Storch and S. Gates, Planar dynamics of a uniform beam with rigid bodies affixed to the ends, CSDL-R-1629, Draper Labs, Cambridge, May, 1983.

21. E. Turkel, Numerical methods for large-scale time-dependent partial differential equations, in Comp. Fluid Dynamics (W. Kollmann, ed.), Hemisphere Publ., Washington, 1980, p. 127-262.

1. Report No. NASA CR-172389 ICASE Report No. 84-24	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Estimation of coefficients and boundary parameters in hyperbolic systems		5. Report Date June 1984
		6. Performing Organization Code
7. Author(s) H. Thomas Banks and Katherine A. Murphy		8. Performing Organization Report No. 84-24
9. Performing Organization Name and Address Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23666		10. Work Unit No.
		11. Contract or Grant No. NAS1-16394, NAS1-17130
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		13. Type of Report and Period Covered Contractor Report
		14. Sponsoring Agency Code 505-31-83-01

15. Supplementary Notes

Additional Support: NSF Grant MCS-8205355, AFOSR Contract 81-0198, and ARO Contract ARO-DAAG-29-83-K0029.
Langley Technical Monitor: Robert H. Tolson
Final Report

16. Abstract

We consider semi-discrete Galerkin approximation schemes in connection with inverse problems for the estimation of spatially varying coefficients and boundary condition parameters in second order hyperbolic systems typical of those arising in 1-D surface seismic problems. Spline based algorithms are proposed for which theoretical convergence results along with a representative sample of numerical findings are given.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
parameter estimation hyperbolic systems approximation techniques	42 Geosciences (General) 64 Numerical Analysis Unclassified - Unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	53	A04