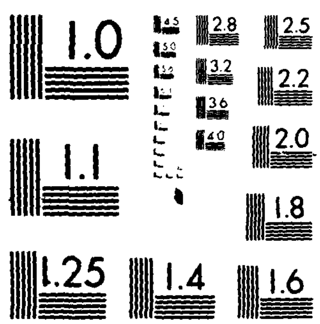
AD-A113 021 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 12/1
CUBIC SPLINE APPROXIMATION TECHNIQUES FOR PARAMETER ESTIMATION --ETC(U)
NOV 81 H T BANKS, J M CROWLEY, K KUNISCH AFOSR-81-0198
UNCLASSIFIED LCDS-81-25 AFOSR-TR-82-0242 NL

END
DATE
FILMED
4-82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

Lefschetz Center for Dynamical Systems

CUBIC SPLINE APPROXIMATION TECHNIQUES FOR PARAMETER ESTIMATION IN DISTRIBUTED SYSTEMS

by

H. T. Banks, J.M. Crowley, K. Kunisch

November, 1981

LCDS #81-25

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 82-0242	2. GOVT ACCESSION NO. AD-A113 691	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Cubic spline approximation techniques for parameter estimation in distributed systems		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) H. T. Banks, J. M. Crowley, and K. Kunisch		8. CONTRACT OR GRANT NUMBER(s) AFOSR 81-0198
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lefschetz Center for Dynamical Systems Div. of Applied Math., Brown University Providence, RI 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A4
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR, Bolling AFB, Washington, DC 20332		12. REPORT DATE November 1981 13. NUMBER OF PAGES 48
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Approximation schemes employing cubic splines in the context of a linear semigroup framework are developed for both parabolic and hyperbolic second order partial differential equation parameter estimation problems. Convergence results are established for problems with linear and nonlinear systems and a summary of numerical experiments with the techniques proposed is given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

CUBIC SPLINE APPROXIMATION TECHNIQUES

FOR

PARAMETER ESTIMATION IN DISTRIBUTED SYSTEMS

by

H. T. BANKS[*,†]

J. M. CROWLEY[*]

K. KUNISCH[+]

November, 1981

LCDS Technical Report No. 81-25

[*] Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, R.I. 02912.

This research was supported by part by the Air Force Office of Scientific Research under contract AFOSR 81-0198, in part by the National Science Foundation under grant NSF-MCS-79-05774-05, and in part by the U. S. Army Research Office under contract ARO-DAAG-29-79-C-0161.

[+] Technische Universität Graz, Institut für Mathematik, A-8010, Graz, Austria.

Part of this research was carried out while this author was a visitor at the Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University. This author also acknowledges support from the Steierm. Wissenschafts - und Forschungslandesfonds.

[†] Part of this research was carried out while this author was a visitor at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA. which is operated under NASA contracts No. NAS1-15810 and No. NAS1-16394.

Cubic Spline Approximation Techniques

for

Parameter Estimation in Distributed Systems

H. T. Banks, J. M. Crowley, and K. Kunisch

ABSTRACT

Approximation schemes employing cubic splines in the context of a linear semigroup framework are developed for both parabolic and and hyperbolic second order partial differential equation parameter estimation problems. Convergence results are established for problems with linear and nonlinear systems and a summary of numerical experiments with the techniques proposed is given.

1. Introduction.

In many problems of practical importance one is confronted with the task of estimating unknown parameters in mathematical models from certain observations of the underlying physical, biological, etc. phenomenon being modeled. This paper is devoted to the study of spline approximation techniques for the identification or estimation of constant parameters in partial differential equations. The results presented here are actually the outcome of investigations partly reported in [9] in which a theoretical convergence framework was developed and applied to treat "modal" approximation schemes for identification and control. Here we use the theoretical framework of [9] to treat spline based techniques that were developed and tested simultaneously with the "modal" methods of [9].

Two specific classes of problems are investigated below. In section 2 we treat an identification problem for a class of parabolic equations, whereas results for hyperbolic equations are given in section 4. Sections 3 and 5 are devoted to discussions of the implementation of the approximation schemes along with numerical examples involving parabolic and hyperbolic equations respectively. In both cases interesting technical questions arise. If the parabolic equations are studied as a system in the usual L_2 state spaces, then fit-to-data criteria involving only integral terms can be treated with relative ease, whereas fit-to-data criteria that employ point spatial evaluations (which are often used in practice) present some essential difficulties. In the case of hyperbolic equations, the semigroup-theoretic approach that is taken in this paper leads to a state space with inner product depending on one of the unknown parameters of the equation. For this reason it is desirable (in other problems this is essential - see the discussions regarding certain unknown function space parameters in [4] and [9]) to develop parameter dependent convergence results in the context of a parameter dependent state space framework.

In our presentation below a number of numerical examples are given to illustrate the theoretical convergence results. For a survey of the potential uses of our semigroup-theory based parameter estimation techniques to problems in a number of specific areas of applications (e.g., reservoir and seismic engineering, large space structures, transport models in physiology and population biology, elasticity and others) the interested reader is referred to [4] and a monograph that the authors currently have in preparation.

We defer until section 6 below a number of comments on other work related to the ideas presented here.

The notation used throughout this paper is rather standard and follows closely that explained in [9]. For norms of elements in Banach spaces we use $|\cdot|$, whereas $\|\cdot\|$ is used for operator norms. We shall occasionally use a subscript such as $|\cdot|_B$ to distinguish a norm in the space B from other norms. The standard practice of denoting the usual norm in $L_p(0,1)$, $p=2,\infty$, by $|\cdot|_p$ will be followed.

2. A class of parabolic partial differential equations.

As a first example for the techniques that were mentioned in the introduction we consider the heat equation

$$(2.1) \qquad u_t = (\frac{q_1}{k})D(p\,Du) + q_2 u + f(q_3,t,x,u), \quad \text{for} \quad t > 0,\ x \in [0,1],$$

with boundary and initial conditions,

(BC) $u(t,0) = u(t,1) = 0$, for $t > 0$,

(IC) $u(0,x) = q_4\varphi(x)$, for $x \in [0,1]$,

where $u = u(t,x) \in R$ and $D = \frac{\partial}{\partial x}$. We assume that (2.1)-(BC)-(IC) models a phenomenon for which we wish to "identify" or estimate the parameter vector $q = (q_1,\ldots,q_4) \in Q \subset R^4$ from known measurements $\hat{y} = \{\hat{y}(t_i)\}_{i=1}^r$ with $\hat{y}(t_i) \in R^m$ taken at times t_i in the fixed interval $[0,T]$ of observation. These observations $\hat{y}(t_i)$ correspond to values $C(t_i,q)\xi(t_i,q)$ in the mathematical model where $C(t_i,q)$ is a real $m \times \ell$ matrix which is continuous in q for each t_i, and

$$\xi(t_i,q) = \text{col}(u(t_i,x_1;q),\ldots,u(t_i,x_\ell;q)).$$

Subsequently we shall give conditions on f that guarantee the existence of a solution u of (2.1)-(BC)-(IC), which will also be denoted by $u(\cdot,\cdot;q)$ whenever we wish to emphasize dependence on q. Since the problem of determining a vector parameter $\tilde{q} \in Q$ such that $\hat{y}(t_i) = C(t_i,\tilde{q})\xi(t_i,\tilde{q})$ for all i would most often lead to an unreasonable and mathematically ill posed problem, we formulate the problem of finding an estimate $\bar{q}$ as an optimization problem in the following way:

$$\text{(ID)} \quad \text{Minimize} \quad J(q,u(\cdot,\cdot;q),\hat{y}) = \sum_{i=1}^{r} w_i |\hat{y}(t_i) - C(t_i,q)\xi(t_i,q)|^2$$

over $q \in Q \subset R^4$ subject to $u(\cdot,\cdot;q)$ satisfying

(2.1)-(BC)-(IC) on $[0,T] \times [0,1]$

In the above formulation of the cost functional the weights $w_i > 0$ can be used if needed to compensate for a priori known extreme behavior of the solution, as for example, exponential growth or decay. (For simplicity, we take $w_i = 1$ in all our discussions and examples in this paper.) The cost functional that is chosen here is but one of the several that are of practical relevance. The one that we chose for our presentation, however, exhibits many interesting technical difficulties (due to the use of point evaluations); in fact, our final convergence result, when used with this functional, requires restrictions on f (dependence on t, x is allowed, but no dependence on u is permitted). For the case of cost functionals involving distributed measurements as, for example, $\tilde{J}(q,u,\hat{y}) = \sum_{i=1}^{r} \int_0^1 |u(t_i,x;q) - \hat{y}(t_i,x)|^2 dx$, our approximation results are valid for equations containing quite general nonlinearities f (see [9]).

Before we formulate the identification problem in the Hilbert space $X = H^0$ with inner product $\langle \varphi,\psi \rangle = \int_0^1 k(x)\varphi(x)\psi(x)dx$ and associated norm $|\cdot|$, we summarize the hypotheses that will be needed throughout this section. We denote by H^i the usual Sobolev spaces over (0,1) as discussed in [1]. The function $F: Q \times [0,\infty) \times H^0 \to H^0$ which will be used below is defined to be the composition map $F(q,t,v) = f(q,t,\cdot,v(\cdot))$. Further, for a given parameter set Q, we let $Q_4 = \{q_4: q = (q_1,\ldots,q_4) \in Q\}$. Our hypotheses are listed for easy reference as follows:

(HQ) Q is a compact subset of R^4 and there exist positive numbers q_1^L and q_1^U such that $q \in Q \subset R^4$ implies

$q_1^L \leq q_1 \leq q_1^U$.

(HP) The functions p and k satisfy $p \in C^3[0,1]$, $k \in C[0,1]$, with $k(x) > 0$ and $p(x) > 0$ for all $x \in [0,1]$.

(HF) The nonlinear function F satisfies

(i) for each continuous function $u:[0,T] \to X$ and each $q \in Q$, the map $t \to F(q,t,u(t))$ is measureable,

(ii) for each constant $M > 0$ there exists a function $k_1 = k_1(M)$ in $L_2(0,T)$ such that for any $q \in Q$ we have

$$|F(q,t,u_1) - F(q,t,u_2)| \leq k_1(t)|u_1-u_2|$$

for all $u_1,u_2 \in X$ with $|u_i| \leq M$,

(iii) there exists a function k_2 in $L_2(0,T)$ such that

$$|F(q,t,v)| \leq k_2(t)\{|v| + 1\},$$

for all $v \in X$, $q \in Q$,

(iv) for each $(t,v) \in [0,T] \times X$ the map $q \to F(q,t,v)$ is continuous.

In [9] we have detailed conditions on the perturbation f that guarantee (HF). For our parabolic systems, statement of the conditions of (HF) in terms of f are rather obvious except in the case of (HF)(ii) which requires a global Lipschitz criterion for f with respect to u due to the fact that we are using the L_2 norm in the local statement of (HF)(ii) for F. We note that as a consequence of (HP) the spaces X and H^0 have equivalent norms.

We shall consider the semilinear equation

$$\frac{du(t)}{dt} = A(q)u(t) + F(q,t,u(t)) \quad \text{for} \quad t > 0, \tag{2.2}$$
$$u(0) = u_0(q) = q_4\phi \ ,$$

where $A(q)\psi = \frac{q_1}{k} D(pD\psi) + q_2\psi$, with $\mathrm{dom}\, A(q) = \{\psi \in H^2 : \psi(0) = \psi(1) = 0\}$. Under (HQ) it is well known that $A(q)$ is the infinitesimal generator of a linear C_0-semigroup $T(t;q)$ for each $q \in Q$. We point out that we have used the letter u in two different ways: In (2.1) $u = u(t,x) \in R$, whereas in (2.2) $u(t)$ or $u(t;q)$, if dependence on q is emphasized, is an element of X. This should not create any ambiguities, since the notion of solutions of (2.1)-(BC)-(IC) or (2.2) will be fixed throughout the paper to be that of <u>mild solutions</u>. We recall that $t \to u(t;q)$ is called a mild solution of (2.2) on $[0,T]$ if it satisfies

$$u(t;q) = T(t;q)u_0(q) + \int_0^t T(t-s;q)F(q,s,u(s;q))ds, \tag{2.3}$$

for $t \in [0,T]$. Under hypothesis (HF) one can easily demonstrate existence of a unique solution to (2.3)(e.g. see [9]). The relationship between mild and strong solutions of (2.2) and classical solutions of (2.1)-(BC)-(IC) has been the focus of many investigations and we only refer to [15] as one possible reference. We are now prepared to formulate the abstract identification problem associated with equation (2.2):

(IDA) Minimize $J(q,u(\cdot;q),\hat{y})$ over $q \in Q \subset R^4$ subject to $u(\cdot;q)$ satisfying (2.3) on $[0,T]$.

An unavoidable difficulty with parabolic problems now becomes apparent. We consider (2.2) as an equation in H^0 and a cost functional J which involves point evaluation, an operation that is not well-defined on H^0. Moreover, it will shortly become evident that "$q^N \to q^*$ implies $u^N(t;q^N) \to u(t;q^*)$", where $u^N(t;q)$ is an approximation to $u(t;q)$ in

X, is an analytical statement of central importance in our approach to the parameter identification problem. Convergence of $u^N(t;q^N)$ to $u(t;q^*)$ in X, however, clearly does not imply the desired convergence of the cost functionals $J(q^N,u^N(\cdot;q^N),\hat{y})$ to $J(q^*,u(\cdot;q^*),\hat{y})$. As a quick remedy to these difficulties, one might be tempted (this is not our remedy) to set up (2.1)-(BC)-(IC) in a state space with a stronger topology. This would involve a more complicated inner product, of course, and consequently lead to a more complex structure for the matrix representations of the approximating finite dimensional problems, a feature that is highly undesirable from the numerical point of view.

With regard to the point evaluations used in J, we note that as a consequence of the "smoothing effect" of parabolic systems, under (HF) one can use the theory of monotone operators to argue that $u(t;q) \in \text{dom}\,(\omega I - A(q))^{1/2} \subset H^1$, for some $\omega > 0$. Thus point evaluation can be justified; this is discussed in more detail in [9]. For a solution to the second of the above-mentioned difficulties see Theorem 2.1 below.

We now explain a Galerkin approach employing cubic spline subspaces to solve (IDA) iteratively. Given a value of N and a vector q, we seek an approximate solution to (2.2) in $X^N = \text{span}\{B_0^N,\ldots,B_N^N\}$ of the form

$$u^N(t;q) = \sum_{j=0}^{N} w_j^N(t)B_j^N = \sum_{j=0}^{N} w_j^N(t;q)B_j^N , \tag{2.4}$$

where $\{B_j^N\}$ is the set of cubic spline basis functions appropriately modified to be in dom $A(q)$. More precisely, let $\Delta^N = \{x_i\}_{i=0}^N$ with $x_i = \frac{i}{N}$, for $i = 0,\ldots,N$, and let $\tilde{B}_j^N$, $j = -1,\ldots,N+1$, denote the standard $C^2(0,1)$ basis elements for the cubic B-spline subspaces of dimension $N+3$ with respect to the grid Δ^N (see p. 208-209 of [17]). Then B_j^N is given by

$$B_j^N = \tilde{B}_j^N \ , \text{ for } \ 2 \leq j \leq N-2 \ ,$$

$$B_0^N = \tilde{B}_0^N - 4\tilde{B}_{-1}^N, \quad B_N^N = \tilde{B}_N^N - 4\tilde{B}_{N+1}^N \ ,$$

$$B_1^N = \tilde{B}_0^N - 4\tilde{B}_1^N \ , \ B_{N-1}^N = \tilde{B}_N^N - 4\tilde{B}_{N-1}^N \ .$$

Note then that $X^N = S_0^3(\Delta^N) = \{\phi \in S^3(\Delta^N) : \phi(0) = \phi(1) = 0\}$, where $S^3(\Delta^N) = \{\phi \in C^2(0,1) : \phi \text{ is a cubic polynomial on each interval } [x_i, x_{i+1}]\}$.

The approximate solutions to (2.2) are determined from requiring that for all $z \in X^N$, we have

$$\text{(2.5)} \qquad \begin{aligned} \langle \dot{u}^N(t), z \rangle &= \langle A(q)u^N(t), z \rangle + \langle F(q,t,u^N(t)), z \rangle , \\ \langle u^N(0), z \rangle &= \langle q_4\phi, z \rangle , \end{aligned}$$

or, equivalently,

$$\sum_{j=0}^{N} \langle \dot{w}_j^N(t) B_j^N, \ B_i^N \rangle = \sum_{j=0}^{N} \langle w_j^N(t) A(q) B_j^N, \ B_i^N \rangle + \langle F(q,t,u^N(t)), B_i^N \rangle ,$$

$$\sum_{j=0}^{N} \langle w_j^N(0) B_j^N, B_i^N \rangle = \langle q_4\phi, \ B_i^N \rangle .$$

This, in turn can be written in matrix form as

$$\text{(2.6)} \qquad \begin{aligned} Q^N \dot{w}^N(t) &= K^N w^N(t) + R^N(F(q,t,u^N(t))), \\ Q^N w^N(0) &= R^N(q_4\phi) \ , \end{aligned}$$

where

$$(Q^N)_{i,j} = \langle B_i^N, \ B_j^N \rangle,$$

$$(K^N)_{i,j} = \langle B_i^N, \ A(q) B_j^N \rangle,$$

$$(R^N\psi)_i = \langle \psi, \ B_i^N \rangle ,$$

and $w^N = \mathrm{col}(w_0^N, w_1^N, \ldots, w_N^N)$.

Alternatively, (2.6) can be viewed as the result of projecting the original problem (2.2) onto the finite dimensional subspaces X^N. In terms of projections, we can write the Galerkin equations in the form

$$(2.7)\qquad \begin{aligned} \dot{w}^N(t) &= A^N(q)w^N(t) + [P^N F(q,t,u^N(t))], \text{ for } t > 0, \\ w^N(0) &= [P^N(q_4\phi)], \end{aligned}$$

where $P^N: X \to X^N$ denotes the canonical orthogonal projection along $(X^N)^\perp$, $A^N(q)$ stands for the matrix representation of $\mathcal{A}^N = P^N\mathcal{A}(q)P^N$ on X^N and $[\cdot]$ denotes the coordinate vector for an element in X^N. Thus, $A^N = (Q^N)^{-1}K^N$ and $[P^N\phi] = (Q^N)^{-1}R^N(\phi)$ for $\phi \in X$.

Since $\mathcal{A}^N(q)$ is a bounded operator, it clearly generates a semigroup $T^N(t;q) = e^{\mathcal{A}^N(q)t}$. Moreover $u^N(t;q)$ will satisfy (2.5)(equivalently, (2.6)) on $[0,T]$ if and only if it satisfies

$$(2.8)\qquad u^N(t;q) = T^N(t;q)P^N u_0(q) + \int_0^t T^N(t-s;q)P^N F(q,s,u^N(s;q))ds.$$

Standard Picard iteration arguments yield that unique solutions $u^N(\cdot;q)$ of (2.8) exist under hypotheses (HF).

In light of the above discussions, we therefore formulate the approximate identification problems as:

(IDA^N) Minimize $J(q,u^N(\cdot;q),\hat{y}) = \sum_{i=1}^{r} |\hat{y}(t_i) - C(t_i,q)\xi^N(t_i,q)|^2$ subject to $u^N(\cdot;q)$ satisfying (2.8) on $[0,T]$, where $\xi^N(t_i,q) = \mathrm{col}(u^N(t_i,x_1;q),\ldots,u^N(t_i,x_\ell;q))$.

Using the continuous dependence with respect to q of $C(t_i,q)$ and $u^N(t_i;q)$ for each $i = 1,\ldots,r$, one can easily argue continuity of $q \to J(q,u^N(\cdot;q),\hat{y})$ and consequently, by (HQ), there exists for each

N a solution $\bar{q}^N$ of (IDA^N). By compactness of Q we can extract a convergent subsequence, again denoted by $\bar{q}^N$, with $\bar{q}^N \to \bar{q} \in Q$. In the remainder of this section we shall concentrate on proving that for any arbitrary sequence $\{q^N\} \subset Q$,

$$\text{(2.9)} \qquad \lim_{N\to\infty} |q^N - q^*| = 0 \text{ implies } \lim_{N\to\infty} |u^N(t;q^N) - u(t;q^*)|_\infty = 0$$

or (in the event one employs $\tilde{J}$ for the fit-to-data),

$$\text{(2.10)} \qquad \lim_{N\to\infty} |q^N - q^*| = 0 \text{ implies } \lim_{N\to\infty} |u^N(t;q^N) - u(t;q^*)|_X = 0$$

for each $t \in [0,T]$. We remark that for the problems considered in this paper we have $u^N(t;q^N) \in H^3$ and $u(t;q^*) \in H^1$ so that the norms in (2.9) and (2.10) can be employed without loss of meaning.

Once (2.9) (respectively (2.10)) is verified it follows immediately that $\bar{q}$ is a solution of (IDA) (respectively of (IDA) with J replaced by $\tilde{J}$) and that $\lim_N J(\bar{q}^N, u^N(\cdot;\bar{q}^N), \hat{y}) = J(\bar{q}, u(\cdot;\bar{q}), \hat{y})$ (respectively, $\lim_N \tilde{J}(\bar{q}^N, u^N(\cdot;\bar{q}^N), \hat{y}) = \tilde{J}(\bar{q}, u(\cdot;\bar{q}), \hat{y})$). Indeed $J(\bar{q}^N, u^N(\cdot;\bar{q}^N), \hat{y}) \le J(q, u^N(\cdot;q), \hat{y})$ for all $q \in Q$ and all N. Under (2.9) we have that $\lim_{N\to\infty} |u^N(t;\bar{q}^N) - u(t;\bar{q})|_\infty = 0$ and $\lim_{N\to\infty} |u^N(t;q) - u(t;q)|_\infty$ for each $t \in [0,T]$ and $q \in Q$. Consequently, taking limits in the above inequality we obtain $J(\bar{q}, u(\cdot;\bar{q}), \hat{y}) \le J(q, u(\cdot;q), \hat{y})$ for each $q \in Q$, so that $\bar{q}$ is a minimizer of $q \to J(q, u(\cdot;q), \hat{y})$.

Since $u(t;q)$ and $u^N(t;q)$ are solutions of (2.3) and (2.8) respectively, (2.10) will follow easily by the Gronwall lemma and Lebesgue's bounded convergence theorem, once we have shown

(Hi) $\|T^N(t;q)\| \le Me^{\omega t}$, with M and ω independent of N and q,

(Hii) $P^N \to I$ strongly in X, and

(Hiii) $T^N(t;q^N) \to T(t;q^*)$ strongly in X and uniformly in $t \in [0,T]$, for any convergent sequence $q^N \to q^*$.

For details in a more

general setting, we refer to [9]. Due to the special choice of $A^N = P^N A P^N$ we find that (Hi) trivially holds: indeed $\|T(t;q)\| \le e^{\omega t}$ with $\omega = \max_{q \in Q} q_2$. Then, since $\langle A^N \psi, \psi \rangle = \langle A P^N \psi, P^N \psi \rangle \le \omega |\psi|^2$ it immediately follows that (Hi) holds. The usefulness of this classical argument in the context of general approximation schemes for dissipative operators was apparently first noted in [7].The strong convergence of P^N to I is established with the aid of three lemmas. The first is a statement of the standard cubic spline-interpolation error bounds, the second a statement of the useful Schmidt inequality and finally in the third lemma we will establish (Hii) on a dense subset of X. That (Hii) obtains then follows from the uniform boundedness of the sequence of projections $\{P^N\}$.

<u>Lemma 2.1</u>. [20,p.54].If $z \in H^4$, and $I^N z$ denotes the cubic spline interpolant of z in $S^3(\Delta^N)$, then

$$|z - I^N z|_2 \le C_0 N^{-4} |D^4 z|_2 ,$$

$$|D(z - I^N z)|_2 \le C_1 N^{-3} |D^4 z|_2$$

$$|D^2(z - I^N z)|_2 \le C_2 N^{-2} |D^4 z|_2 ,$$

where the C_i are constants independent of z and N.

<u>Lemma 2.2</u>. [20,p.7]. If p_n is a polynomial of degree $n = 1,2,3$ on $[a,b]$, then

$$\int_a^b [Dp_n(x)]^2 dx \le k_n (b-a)^{-2} \int_a^b p_n(x)^2 dx ,$$

with k_n a constant.

Lemma 2.3. If $z \in H_0^1 \cap H^4$, then

$$|P^N z - z| \le \tilde{C}_0 N^{-4} |D^4 z|_2 ,$$

$$|D(P^N z - z)| \le \tilde{C}_1 N^{-3} |D^4 z|_2 ,$$

$$|D^2(P^N z - z)| \le \tilde{C}_2 N^{-2} |D^4 z|_2 ,$$

where the $\tilde{C}_i$ are constants.

Proof. Let $z \in H_0^1 \cap H^4$ and denote by $I^N z$ and $I_0^N z$ the interpolating cubic B-spline for z in $S^3(\Delta^N)$ and $S_0^3(\Delta^N)$ respectively. Using the boundary conditions we have $|P^N z - z| \le |I_0^N z - z| = |I^N z - z| \le \sqrt{\tilde{k}}\, C_0 N^{-4} |D^4 z|_2$, where $\tilde{k} = \max\{k(x) | 0 \le x \le 1\}$. This implies the first estimate with $\tilde{C}_0 = C_0 \left(\frac{\max k(x)}{\min k(x)}\right)^{1/2}$, the extremes being taken over $x \in [0,1]$. To verify the second estimate we make use of Lemmas 2.1 and 2.2:

$$|D(P^N z - z)|_X \le \sqrt{\tilde{k}}\ |D(P^N z - z)|_2 \le \sqrt{\tilde{k}}\, |D(P^N z - I_0^N z)|_2 + \sqrt{\tilde{k}}\, |D(I_0^N z - z)|_2$$

$$\le \sqrt{\tilde{k} k_3}\ N |P^N z - I_0^N z|_2 + \sqrt{\tilde{k}}\, |D(I^N z - z)|_2$$

$$\le \sqrt{\tilde{k} k_3}\ N(|P^N z - z|_2 + |z - I_0^N z|_2) + \sqrt{\tilde{k}}\, |D(I^N z - z)|_2$$

$$\le 2\sqrt{\tilde{k} k_3}\ N^{-3} \tilde{C}_0 |D^4 z|_2 + \sqrt{\tilde{k}}\, C_1 N^{-3} |D^4 z|_2 .$$

This establishes the convergence of the first derivative of the projected elements. The final estimate can be argued in a similar manner.

Having thus established (Hii) we next turn to (Hiii). It is convenient to use the following version of the Trotter-Kato theorem.

Proposition 2.1 [14]. 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. Assume further that $\mathcal{T}(t)$ and $\mathcal{T}^N(t)$ are linear C_0 - semigroups on $\mathcal{B}$ and $\mathcal{B}^N$ with infinitesimal generators $\tilde{A}$ and $\tilde{A}^N$ respectively. If

i) $\lim_{N\to\infty} |\pi^N z|_N = |z|$ for all $z \in \mathcal{B}$,

ii) there exist constants $\tilde{M},\tilde{\omega}$ independent of N such that

$\|\mathcal{T}^N(t)\|_N \leq \tilde{M}e^{\tilde{\omega}t}$, for $t \geq 0$,

iii) there exists a set $\mathcal{D} \subset \mathcal{B}$ such that $\mathcal{D} \subset \mathrm{dom}(\tilde{A})$, $\overline{\mathcal{D}} = \mathcal{B}$ and $\overline{(\lambda_0 - \tilde{A})\mathcal{D}} = \mathcal{B}$ for some $\lambda_0 > 0$ and for which for all $z \in \mathcal{D}$ we have $\lim_{N\to\infty} |\tilde{A}^N \pi^N z - \pi^N \tilde{A} z|_N = 0$,

then $\lim_{N\to\infty} |\mathcal{T}^N(t)\pi^N z - \pi^N \mathcal{T}(t)z|_N = 0$ for all $z \in \mathcal{B}$, uniformly in t on compact subsets of $[0,\infty)$.

Proposition 2.2. Let $\{q^N\} \subset Q$ be an arbitrary sequence satisfying $q^N \to q^*$ as $N \to \infty$. Then the semigroups $T^N(t;q^N)$ and $T(t;q^*)$ generated by $A^N(q^N)$ and $A(q^*)$ respectively, satisfy

$$\lim_{N\to\infty} |T^N(t;q^N)z - T(t;q^*)z| = 0$$

for each $z \in X$, uniformly on compact subsets of $[0,\infty)$.

Proof. Let us first recall some elementary facts concerning spectral properties of the self adjoint operator $A(q)$ (c.f. [10,p. 291-295] or [15]). The spectrum of $A(q)$ consists of a countable number of real eigenvalues $\{\lambda_j(q)\}_{j=1}^{\infty}$, each of multiplicity 1, which can be ordered so that $-\infty < \ldots < \lambda_j(q) < \lambda_{j-1}(q) < \ldots < \infty$ and which are uniformly bounded above as q varies in Q. The eigenfunctions $\{\Psi_j\}_{j=1}^{\infty}$ of $A(q^*)$ form a complete orthogonal set in X. To apply Proposition 2.1 we take $\mathcal{B} = X$, $\mathcal{B}^N = X$, π^N the identity operator for each N and $\tilde{A} = A(q^*)$, $\tilde{A}^N = A^N(q^N)$. (The observant reader will notice that we could equally well take $\mathcal{B} = X$, $\mathcal{B}^N = X^N$, $\pi^N = P^N$ with $\tilde{A} = A(q^*)$ and $\tilde{A}^N = P^N A(q^N)$, which in this case would be defined on X^N.) Of course, i) of Proposition 2.1 is satisfied trivially while the stability hypothesis ii) is the same as (Hi) which has already been established. To verify the consistency hypothesis iii) we need to choose the set $\mathcal{D}$. We define $\mathcal{D} = \bigcup_{N=1}^{\infty} \text{span}\{\Psi_1, \ldots, \Psi_N\}$. Clearly $\mathcal{D} \subset \text{dom}\, A(q^*)$ and, by completeness of the eigenfunctions, $\bar{\mathcal{D}} = \mathcal{B}$. We will show $(\lambda - A(q^*))\mathcal{D} = \mathcal{D}$ for any λ in the resolvent set of $A(q^*)$. Trivially $(\lambda - A)\mathcal{D} \subset \mathcal{D}$ holds. To show that $(\lambda - A)\mathcal{D} \supset \mathcal{D}$ take an arbitrary $\psi \in \mathcal{D}$. Without loss of generality let ψ be an eigenfunction, say Ψ_i, of $A(q^*)$. Then $(\lambda - A(q^*))(\frac{1}{\lambda - \lambda_i}\Psi_i) = \Psi_i$ and therefore $\mathcal{D} \subset (\lambda - A(q^*))\mathcal{D}$. It remains to show that $A^N(q^N)z \to A(q^*)z$ for any $z \in \mathcal{D}$. We first note that the smoothness assumed in (HP) can be used to easily argue that any Ψ_i (and hence any $z \in \mathcal{D}$) is in H^4. Thus for any $z \in \mathcal{D}$ we have $z \in H^4 \cap H_0^1$.

For fixed $z \in \mathcal{D}$, let z^N denote $P^N z$. Then

$$\begin{aligned} |A^N(q^N)z - A(q^*)z| &= |P^N A(q^N) P^N z - A(q^*)z| \\ &\le |P^N(A(q^N) - A(q^*))z^N| + |P^N A(q^*)(z^N - z)| + |(P^N - I)A(q^*)z| \qquad (2.11)\\ &\le |(A(q^N) - A(q^*))z^N| + |A(q^*)(z^N - z)| + |(P^N - I)A(q^*)z| . \end{aligned}$$

For the second term we have the estimate

$$|A(q^*)(z^N - z)| \le \sup_{x \in [0,1]} \left| \frac{q_1^*(Dp)(x)}{k(x)} \right| |D(z^N - z)|$$

$$+ \sup_{x \in [0,1]} \left| \frac{q_1^* p(x)}{k(x)} \right| |D^2(z^N - z)| + q_2^* |z^N - z| .$$

By Lemma 2.3 we therefore find that $|A(q^*)(z^N - z)|$ converges to zero like $O(N^{-2})$ as $N \to \infty$. Strong convergence of $P^N \to I$ implies that the third term in (2.11) converges to zero. Finally, we have $A(q^N)z^N = \frac{q_1^N}{k}(DpDz^N + pD^2z^N) + q_2^N z^N$, which converges to $\frac{q_1^*}{k}(DpDz + pD^2z) + q_2^* z$, by Lemma 2.3 and the fact that $q^N \to q^*$. This concludes the proof.

We have thus outlined arguments to establish (2.10).

Proposition 2.3. If (HQ), (HP) and (HF) hold, then for each convergent sequence $q^N \to q^*$ it follows that $\lim_{N\to\infty} |u^N(t;q^N) - u(t;q^*)|_X = 0$ for each $t \in [0,T]$.

We further desire convergence in $C(0,1)$ to obtain the needed pointwise (in the spatial variable) convergence to use with functionals of type J. Define $\tilde{H}_0^1 = (H_0^1, \langle\langle\cdot,\cdot\rangle\rangle)$, where $\langle\langle\varphi,\psi\rangle\rangle = \int_0^1 D\varphi(x)D\psi(x)p(x)dx$; then $\tilde{H}_0^1$ is a Hilbert space and the associated norm is equivalent to the H_0^1 norm given by $|\varphi|^2_{H_0^1} = \int_0^1 |D\varphi(x)|^2 dx$. Further we denote by $\tilde{X}^N$ the set X^N endowed with the $\tilde{H}_0^1$ topology. Let $A_R^N(q)\colon \tilde{X}^N \to \tilde{X}^N$ be given by $A_R^N(q) = P^N A(q)$. Due to the finite dimensionality of $\tilde{X}^N$ the operator $A_R^N(q)$ is a bounded linear operator which generates a semigroup $T_R^N(t;q)$ on $\tilde{X}^N$. To calculate its exponential bound, we consider for $v \in \tilde{X}^N$ the following estimate. Letting $\langle\cdot,\cdot\rangle_2$ denote the inner product in H^0, then we find

$$\langle\langle v, A_R^N(q)v\rangle\rangle = \langle pDv, DP^N A(q)v\rangle_2$$

$$= \langle pDv, DP^N(\frac{q_1}{k} D(pDv)) + DP^N q_2 v\rangle_2$$

$$= q_2\langle pDv, Dv\rangle_2 - q_1\langle D(pDv), P^N \frac{1}{k} D(pDv)\rangle_2$$

$$= q_2 \langle\langle v,v \rangle\rangle - q_1 \langle k \frac{1}{k} D(pDv),\ P^N \frac{1}{k} D(pDv) \rangle_2$$

$$= q_2 \langle\langle v,v \rangle\rangle - q_1 \langle \frac{1}{k} P^N D(pDv),\ P^N \frac{1}{k} D(pDv) \rangle_X \leq q_2 \langle\langle v,v \rangle\rangle ,$$

since $q_1 > 0$. Consequently (see [11,p.85,90]) we have

$$|T_R^N(t;q)z|_{\tilde{X}^N} \leq e^{q_2 t} |z|_{\tilde{X}^N}$$

for all $z \in X^N$. This last estimate implies (in light of the weighting of the $\tilde{H}_0^1$ norm by p)

$$|T^N(t;q)P^N z|_{H_0^1} \leq M e^{q_2 t} |P^N z|_{H_0^1} \tag{2.12}$$

for all $z \in H^0$, where $M^2 = \max\{p(x) | x \in [0,1]\} / \min\{p(x) | x \in [0,1]\}$.

We next verify that $|P^N z|_{H_0^1}$ is uniformly bounded in N for $z \in H_0^1 \cap H^2$. In fact $|P^N z|_{H_0^1} = |DP^N z|_2 \leq |D(P^N z - I^N z)|_2 + |D(I^N z - z)|_2 + |Dz|_2$, where $I^N z$ is the interpolating cubic spline in $S^3(\Delta^N)$. An application of the Schmidt inequality (Lemma 2.2) implies that

$$|P^N z|_{H_0^1} \leq 2\sqrt{k_3}\, N |I^N z - z|_2 + \hat{C}_1 N^{-1} |D^2 z|_2 + |Dz|_2 ,$$

with $\hat{C}_1$ a constant. Here we have also used a well known (see [20,p.53]) spline interpolation result to estimate $|D(I^N z - z)|_2$. Finally for $z \in \text{dom}\, A(q) = H_0^1 \cap H^2$ we obtain (again refer to [20, p. 53])

$$|P^N z|_{H_0^1} \leq \frac{\hat{C}_2}{N} |D^2 z|_2 + |Dz|_2 , \tag{2.13}$$

where $\hat{C}_2$ is a constant independent of N. We summarize the above discussion, in particular (2.12) and (2.13), in

Lemma 2.4. If (HQ) and (HP) hold and $z \in H_0^1 \cap H^2$, then for each $T > 0$ the set $\{T^N(t;q)P^N z | t \in [0,T],\ N=1,\ldots\}$ is a bounded subset of H_0^1.

The final convergence result of this section is concerned with convergence in the $|\cdot|_\infty$- norm as indicated in (2.9). Our results require severe restrictions on the function F in (2.2).

Theorem 2.1 a) Let (HQ) and (HP) hold and let $q^N \to q^*$ in Q. Then $\lim_{N\to\infty} |T^N(t;q^N)P^N z - T(t;q^*)z|_\infty = 0$ for each $t > 0$ and each $z \in H_0^1 \cap H^2$.
b) Suppose that the map F of (2.2) does not depend on u, i.e. $F = F(q,t)$. Suppose further that $t \to F(q,t)$ from $[0,T]$ to $H_0^1 \cap H^2$ is measurable for each $q \in Q$ and that $q \to F(q,t)$ is continuous from Q to H^2 for each $t \in [0,T]$ and finally for some $k_3 \in L_2(0,T)$ we have $|F(q,t)| + |D^2F(q,t)| \le k_3(t)$, for all $q \in Q$, $t \in [0,T]$. Then $q^N \to q^*$ implies $u^N(t;q^N) \to u(t;q^*)$ in $C(0,1)$ for each $t \in [0,T]$ and each $u_0(q) = q_4\varphi \in H_0^1 \cap H^2$.

Proof. To verify a) let $z,y \in H_0^1 \cap H^2$. Integrating by parts and using the fact that $T(t;q^*)z$ and $T^N(t;q^N)P^N z$ are in $H_0^1 \cap H^2$ we find

$$\begin{aligned}
&|\langle T^N(t;q^N)P^N z - T(t;q^*)z,y\rangle_{H_0^1}| \\
&= |\langle D(T^N(t;q^N)P^N z - T(t;q^*)z),\ Dy\rangle_2| \\
&= |\langle T^N(t;q^N)P^N z - T(t;q^*)z, D^2y\rangle_2| .
\end{aligned}$$

Proposition 2.2 and Lemma 2.4 together with the last estimate and the fact that dom $A(q)$ is dense in H_0^1 imply that $T^N(t;q^N)P^N z$ converges weakly to $T(t;q^*)z$ in H_0^1, uniformly in $t \in [0,T]$. Since H_0^1 is compactly embedded in $C(0,1)$ by the Sobolev embedding theorem [1 ,p.144] it follows that $T^N(t;q^N)P^N z \to T(t;q^*)z$ in $C(0,1)$; since $T > 0$ was arbitrary, this holds for all $t \geq 0$.

Turning next to b) of the theorem, we note first that (HF) is satisfied under the present conditions on F. For each N we have

$$(2.14)\qquad u^N(t;q^N) = T^N(t;q^N)P^N u_0(q^N) + \int_0^t T^N(t-s;q^N)P^N F(q^N,s)ds.$$

Since $|T^N(t;q)P^N z|_{H_0^1} \leq K(|D^2 z|_2 + |Dz|_2)$ for a constant K independent of $N=1,2,\ldots$, $q \in Q$, and $t \in [0,T]$ by (2.12) and (2.13), the integral in (2.14) clearly exists as a Bochner integral in $C(0,1)$. Moreover, for each t we get $T^N(t;q^N)P^N u_0(q^N) \to T(t;q^*)u_0(q^*)$ and $T^N(t-s;q^N)P^N F(q^N,s) \to T(t-s;q^*)F(q^*,s)$ for almost every $s \in [0,t]$, where the convergence is in the $C(0,1)$ norm. Since the functions $s \to T^N(t-s;q^N)P^N F(q^N,s)$ from $[0,t]$ to H_0^1 are bounded by the integrable function $Kk_3(\cdot)(\min_{x\in[0,1]} k(x))^{-1/2}$, we may use Lebesgue's bounded convergence theorem when taking the limit as $N \to \infty$ in (2.14) to find that $u^N(t;q^N) \to u(t;q^*)$ in $C(0,1)$ with $u(t;q^*) = T(t;q^*)u_0(q^*) + \int_0^t T(t-s;q^*)F(q^*,s)ds$.

<u>Remark 2.1.</u> In our discussion here we have considered only problems with Dirichlet boundary conditions. Our ideas are easily extended to treat Neumann or mixed boundary conditions as well. We sketch briefly some of the minor changes required in the above presentation. First one certainly must use different basis elements (recall how the B_i^N were constructed from the standard basis elements $\tilde{B}_j^N$) so as to ensure $X^N \subset \mathrm{dom}\, A(q)$. This in turn requires that one establish an analogue of Lemma 2.3 for the associated natural projections P^N (through use of appropriately defined interpolating splines - we remind the reader that

there are numerous types of interpolating cubic splines - see [20,Chapter 4],[21]). Finally, minor details in some of the above arguments must be modified. For example, in the proof of Proposition 2.2, the eigenvalues have multiplicity ≤ 2 (see Example 4.2 of [9]), not necessarily equal 1. Some of the integration by parts arguments (e.g. in establishing the bound for $A_R^N(q)$ in Proposition 2.3) also require modification.

Remark 2.2. If a pure "convection" type term $q_5 u_x$ also appears in the right side of (2.1), then one can use the theory of discrete spectral operators (see Example 4.4 of [9]) to establish Proposition 2.3 for this case in much the same manner as argued above. The analogue of Theorem 2.1 appears to be much more difficult to obtain however.

3. Implementation and numerical examples: parabolic equations.

In this section we discuss questions related to the implementation of the ideas developed in section 2 and present some numerical results. All of the results given below were obtained using the fit-to-data criterion J involving spatial point evaluations. While we were able to establish above the stronger convergence results (in $C(0,1)$) only in the case of linear equations, the reader will see from the results presented below that the methods also perform quite well when one uses the point evaluation criterion J with nonlinear parabolic equations.

The algorithm for carrying out the identification of the unknown parameter involves two major tasks. The first task is: given N and $u^N(\cdot;q)$, find a $\bar{q}^N$ which minimizes $J(q,u^N,\hat{y})$. This task was carried out by the standard Levenberg-Marquardt algorithm (available in the IMSL library, routine ZXSSQ) and we shall therefore not discuss this part of the implementation. (For a discussion of the Levenberg-Marquardt as well as related algorithms, see [3].)

The Levenberg-Marquardt algorithm requires that $J^N(q) \equiv J(q,u^N,\hat{y})$ be evaluated for fixed N at a sequence of iterates q^j. Thus we need the approximations $u^N(\cdot;q^j)$ and computing these is the second task. By (2.4) the values of $u^N(\cdot;q^j)$ are obtained by solving (2.6) which is rather easily done since the matrices Q^N and K^N appearing there are seven-banded and symmetric, and hence they can be stored as three subdiagonals and the diagonal.

Numerical experiments were carried out for the general example

$$u_t = q_1 u_{xx} + q_2 u + q_4 f(u), \text{ for } 0 \le x \le 1, \text{ and } t > 0,$$

$$u(0,x) = q_3 \varphi(x), \text{ for } 0 \le x \le 1,$$

$$u(t,0) = u(t,1) = 0, \text{ for } t > 0,$$

so that k and p of the previous section are chosen identically 1. In this case $A = q_1 D^2 + q_2 I$ and thus denoting by A_1^N and A_2^N the matrices with elements

$$(A_1^N)_{i,j} = \langle DB_i^N, DB_j^N \rangle_2 ,$$
$$(A_2^N)_{i,j} = \langle B_i^N, B_j^N \rangle_2 ,$$

we have $Q^N = A_2^N$ and $K^N = -q_1 A_1^N + q_2 A_2^N$. In our implementation, the matrices A_1^N, A_2^N were calculated analytically and stored exactly.

The initial values and the nonlinear term require numerical quadrature. To compute $(R^N\psi)_i = \langle \psi, B_i^N \rangle_2 = \int_0^1 \psi(x) B_i^N(x)dx$, a composite two-point Gauss-Legendre rule was employed to evaluate the integral. The same quadrature rule was applied to

$$(R^N F)_i = \langle F(q,t,u^N(t;q)),\ B_i^N \rangle_2 = \int_0^1 f(q,t,x, \sum_{j=0}^{N} w_j^N(t;q) B_j^N(x)) B_i^N(x)dx.$$

The integration of the system of ordinary differential equations (2.6) was carried out by an IMSL routine (DGEAR) employing Gear's variable order, variable step method. For the parabolic equations, the stiffly stable backward difference methods of the routine were used. In most parabolic examples, the equations (2.6) are moderately stiff and while a standard Runge-Kutta scheme can be used effectively, it is more efficient to use Gear's stiffly stable method (in our computations with this method, local error tolerances were set at 10^{-5}).

Finally we discuss the "inversion" of Q^N, which is needed not only in the integration of (2.6), (in which one actually solves for $Q^N w^N(t)$ even though w^N is used in the nonlinearity), but also in the case where $f \equiv 0$ to get $u^N(t;q)$ from $Q^N w^N(t)$. The computation of

$y = (Q^N)^{-1}z$ was carried out by first using the Cholesky algorithm to decompose $Q^N = L_N L_N^T$, where L_N is lower triangular (this was done only once for a given N and the corresponding L_N was stored) and then backsolving the equations $L_N x = z$, $L_N^T y = x$, which involve only triangular matrices. We remark that the banded structure of Q^N is preserved in L_N.

To demonstrate the feasibility of the spline approximation schemes for identification problems, we generated solutions to equations with parameters fixed at given values (called "true" values below) and then attempted to identify these parameters from the "data" consisting of values of the generated solutions.(In some cases random noise was added to the solution values, but tests revealed that this did not affect the efficiency of the schemes.) The generated numerical solutions (with the fixed parameter values) which were used for the data $\hat{y}$ in the fit-to-data criterion J were computed by an independent finite difference method when closed form solutions (e.g. in the case of nonlinear examples) were not readily available.

In the examples below, the observations consisted of the generated solution values $\hat{u}(t_i,x_j)$ at three spatial points (x_j = 0.25, 0.5,0.75) sampled at ten times (t_i = 0.2, 0.4,...,2.0). The matrices $C(t_i,q)$ in the fit-to-data criterion J are taken to be the 3×3 identity matrix.

The first numerical example is of special importance: the modal approximation schemes investigated earlier in [9] failed to perform satisfactorily when we attempted to estimate two of the parameters (q_1,q_2) simultaneously. The cause of this "numerical unidentifiability" can be seen to be a feature of the modal approximation itself and, as we shall see below, does not arise with use of our spline approximations. Indeed, in our numerical investigations we were unable to find an example in our class of parabolic systems for which the spline schemes failed in attempts to estimate multiple parameters.

Example 3.1. We consider the linear initial-boundary value problem

$$u_t = q_1 u_{xx} + q_2 u ,$$
$$u(0,x) = \varphi(x)$$
$$u(t,0) = u(t,1) = 0$$

where $\varphi(x) = 2x$ for $0 \leq x \leq 0.5$ and $\varphi(x) = 2(1-x)$ for $0.5 \leq x \leq 1$, (see [9 ,Ex. 6.4]). Here and below $q^{N,0}$ denotes the start-up value for the Levenberg-Marquardt optimization routine, for the N-th approximate identification problem (IDA^N). Table 1 depicts the results obtained when q_1, q_2 are sought and $q_3 = 1$ is assumed to be known.

TABLE 1

N	$\bar{q}_1^N$	$\bar{q}_2^N$
4	.1020	.8195
6	.1000	.8004
true value	.1	.8
$q^{N,0}$	.25	.25

As we indicated above, use of the modal approximations of [9] did not produce good results when simultaneously estimating q_1 and q_2 in this example. The computational findings as summarized in Table 2 reveal that no such difficulty is associated with estimation of several parameters when one uses the cubic spline approximations described in section 2.

TABLE 2

N	$\bar{q}_1^N$	$\bar{q}_2^N$	$\bar{q}_3^N$
4	.4886	2.7856	5.1011
5	.4905	1.7779	5.3249
6	.5153	2.0594	5.1827
10	.5033	1.9884	5.0949
true value	.5	2.0	5.0
$q^{N,0}$	.25	1.0	1.0

Example 3.2 Next we consider the nonlinear I-BV problem

$$u_t = q_1 u_{xx} - q_4 u^3$$

$$u(0,x) = q_3 \varphi(x),$$

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

where φ is chosen as in Example 3.1. The numerical results that were found are given in Table 3.

TABLE 3

N	$\bar{q}_1^N$	$\bar{q}_3^N$	$\bar{q}_4^N$
4	.4978	5.2979	1.2354
8	.4989	5.1215	1.1055
16	.4992	5.0651	1.0653
true value	.5	5.0	1.0
$q^{N,0}$	.1	1.0	0

Example 3.3. As our final parabolic example, we consider one with a different nonlinear equation from that of the previous example.

$$u_t = q_1 u_{xx} + q_4 \frac{1}{1+u} ,$$

$$u(0,x) = q_3 \varphi(x) ,$$

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

where φ is chosen as in the previous examples. The numerical findings are recorded in Table 4.

TABLE 4

N	$\bar{q}_1^N$	$\bar{q}_3^N$	$\bar{q}_4^N$
4	.5309	5.2963	1.0781
8	.5126	5.1208	1.0316
16	.5067	5.0648	1.0170
true value	.5	5.0	1.0
$q^{N,0}$	.1	1.0	0

4. A class of hyperbolic partial differential equations.

In this section we consider identification problems for the one dimensional nonlinear hyperbolic equation

(4.1) $u_{tt} = q_1 D^2 u + q_2 u_t + q_3 u + f(q_6, t, x, u, u_t)$, for $t > 0$, $x \in [0,1]$,

with boundary and initial conditions,

(BC) $u(t,0) = u(t,1) = 0$, for $t > 0$

(IC) $u(0,x) = q_4 \varphi(x)$,

$u_t(0,x) = q_5 \psi(x)$, for $x \in [0,1]$.

We could allow equally well for multiple unknown parameters in the initial conditions, as for example $u(0,x) = \sum_{i=1}^{p} q_4^i \varphi_i(x)$.

Following the procedure of section 2, we rewrite (4.1)-(BC)-(IC) as an abstract evolution equation in an appropriate Hilbert space. Standard results imply that D^2 in $H^0 = L_2(0,1;R)$ with dom $(D^2) = H_0^1 \cap H^2$ is a selfadjoint operator satisfying $\langle -D^2 z, z\rangle_2 \geq |z|_2^2$ for every $z \in \mathrm{dom}(D^2)$. We again assume that (HQ) holds, except now modified in that $Q \subset R^6$. With (HQ) holding we define $V(q_1) = (H_0^1, \langle\cdot,\cdot\rangle_{V(q_1)})$ where we endow H_0^1 with the topology defined by the inner product

$$\langle w,z\rangle_{V(q_1)} = \int_0^1 q_1 Dw(x)Dz(x)dx = \langle q_1 Dw, Dz\rangle_2 .$$

Clearly $V(q_1)$ is a Hilbert space. Define $X(q) = V(q_1) \times H^0$ with the usual product topology generated by $\langle (w_1,w_2),(z_1,z_2)\rangle = \langle w_1,z_1\rangle_{V(q_1)} + \langle w_2,z_2\rangle_2$. Once again, $\langle\cdot,\cdot\rangle$ and $|\cdot|$ (or $\langle\cdot,\cdot\rangle_q$ and $|\cdot|_q$ whenever dependence on the vector q must be emphasized) denote the inner product and associated norm in X throughout this section. We may now rewrite (4.1) in X (taking $v = u_t$) as

$$(4.2)\quad \begin{aligned} \begin{pmatrix}\dot u(t)\\ \dot v(t)\end{pmatrix} &= A(q)\begin{pmatrix}u(t)\\ v(t)\end{pmatrix} + F(q,t,u(t),v(t)),\ \text{for } t>0,\\ \begin{pmatrix}u(0)\\ v(0)\end{pmatrix} &= \begin{pmatrix}q_4\ \varphi\\ q_5\ \psi\end{pmatrix}, \end{aligned}$$

where $(\varphi,\psi) \in X$, dom $A(q) = H_0^1 \cap H^2 \times H_0^1$, $A(q) = \begin{pmatrix} 0 & 1 \\ q_1D^2+q_3 & q_2 \end{pmatrix}$

and F is the operator given by $F(q,t,u(t),v(t)) = \begin{pmatrix} 0 \\ f(q,t,\cdot,u(t,\cdot),v(t,\cdot)) \end{pmatrix}$.

As in section 3 we consider the identification problem

(IDA) Minimize $J(q,u(\cdot;q),\hat{y})$ over $q \in Q \subset R^6$ subject to $u(\cdot;q)$ satisfying for $t \in [0,T]$

$$(4.3)\quad \begin{pmatrix}u(t;q)\\ v(t;q)\end{pmatrix} = T(t;q)\begin{pmatrix}u(0)\\ v(0)\end{pmatrix} + \int_0^t T(t-s;q)F(q,s,u(s;q),v(s;q))ds.$$

Here $T(t;q)$ denotes the semigroup in $X(q)$ generated by $A(q)$ (see [9]). We shall assume that (HF) with u replaced by (u,v) holds throughout this section; this suffices to ensure existence of a unique solution of (4.3) on $[0,T]$. Again, conditions on $f(q_6,t,x,u,u_t)$ that imply (HF) can be found in [9]. Since the first component $u(t)$ of the solution vector $(u(t),v(t))$ lies in H_0^1, point evaluations of $u(t)$ clearly pose no difficulties in this example. Moreover, a parameter dependent convergence result in the norm of the state space analogous to (2.10) will, in this case, be sufficient to prove that solutions $\bar{q}^N$ of the approximating identification problems and the associated fit-to-data term $J^N(\bar{q}^N)$ converge, respectively, to a solution $\bar{q}$ of the original problem (IDA) and its associated term $J(\bar{q})$.

An interesting aspect of this example is that now the norm of the state space depends on the parameter vector q. The desirability for allowing for such generality results from the fact that the operator $A(q)$ will be dissipative in this weighted inner product. In fact, with (HQ) obtaining, one can easily establish the existence of a constant ω independent of $q \in Q$ such that

$$\langle A(q)z,z\rangle_q \leq \omega|z|_q^2 \tag{4.4}$$

for all $z \in X$ (See [9] . One can also rescale the state variable to avoid the state space dependence on q, but this does not lead to any simplification in the computational scheme obtained.)

We now define the state approximation scheme. As in the previous example, define $S_0^3(\Delta^N) = \text{span}\{B_0^N,\ldots,B_N^N\}$. For the approximating subspaces we take $X^N = S_0^3(\Delta^N) \times S_0^3(\Delta^N) \subset H_0^1 \cap H^2 \times H_0^1 = \text{dom}(A(q))$. In terms of basis elements, $X^N(q) = \text{span}\{\beta_1^N,\ldots,\beta_{2N+2}^N\}$, where

$$\beta_i^N = \begin{cases} (B_{i-1}^N,0)^T, & 1 \leq i \leq N+1 \\ (0,B_{i-N-2}^N)^T, & N+2 \leq i \leq 2N+2 . \end{cases}$$

We again seek an approximation to the state of the form

$$(u^N(t;q),\ v^N(t;q))^T = \sum_{i=1}^{2N+2} w_i^N(t)\beta_i^N ,$$

by using a Galerkin approach. Proceeding as before we arrive at a semidiscrete approximation to (4.2) by a system of 2N+2 ordinary differential equations:

$$Q^N\dot{w}^N(t) = K^Nw^N(t) + R^NF(q,t,u^N(t),v^N(t))$$
$$Q^Nw^N(0) = R^N((q_4\varphi,q_5\psi)^T),$$

where $w^N(t) = \text{col}\,(w_1^N(t),\ldots,w_{2N+2}^N(t))$, and

$$(Q^N)_{i,j} = \left\langle \beta_i^N, \beta_j^N \right\rangle_q$$
$$(K^N)_{i,j} = \left\langle \beta_i^N, A(q)\beta_j^N \right\rangle_q$$
$$(R^N\psi)_i = \left\langle \beta_i^N, \psi \right\rangle_q \quad .$$

we again find it more convenient to write this in terms of the orthogonal projections $P^N(q):X(q) \to X^N(q)$, and defining $A^N(q) = P^N(q)A(q)P^N(q)$, we obtain

$$\text{(4.5)} \qquad \begin{aligned} \dot{w}^N(t) &= A^N(q)w^N(t) + [P^N F(q,t,u^N(t),v^N(t))] \\ w^N(0) &= [P^N(q_4\varphi,q_5\psi)^T], \end{aligned}$$

where $A^N(q)$ is the matrix representation of $A^N(q)$ and $[\cdot]$ denotes the coordinate vector of an element in X^N. Note here that as a consequence of the parameter dependent norm the projections also depend on q.

As before we solve (IDA) iteratively by solving the sequence of approximating problems (IDA^N) with (2.8) replaced by (4.5). Using considerations similar to those of section 2 it is easily seen that solutions $\bar{q}^N$ of the approximating problems exist and that any limit $\bar{q}$ of a convergent subsequence of $\bar{q}^N$ is a solution of (IDA), provided that for any arbitrary sequence q^N in Q we have

$$\text{(4.6)} \quad \lim_{N\to\infty} |q^N-q^*| = 0 \ \text{ implies } \ \lim_{N\to\infty} |(u^N(t;q^N),v^N(t;q^N)) - (u(t;q^*),v(t;q^*))|_X = 0,$$

for each $t \in [0,T]$. Employing the mild forms of equations (4.2) and (4.5) satisfied by $(u(t;q^*),v(t;q^*))$ and $(u^N(t;q^N),v^N(t;q^N))$ respectively, one can show as before that (4.6) holds provided that conditions (Hi)-(Hiii) explained in section 2 can be verified. Since (Hi) will be discussed together with (Hiii), we immediately turn to (Hii). In the calculations below we will freely use the fact that under (HQ) all the $X(q)$ norms are equivalent as q varies in Q.

To show that $P^N \to I$ strongly in X, we will need to modify the previous spline approximation Lemma 2.3 to obtain $V(q_1)$ norm estimates. First we state a spline interpolation estimate in the $V(q_1)$ norm; (see Lemma 2.1, of which this is simply a restatement of the last two inequalities).

Lemma 4.1 Let $I^N z$ be the cubic spline interpolant (from $S^3(\Delta^N)$) to $z \in H^4$. Then

$$|I^N z - z|_{V(q_1)} \leq q_1^{1/2} C_1 N^{-3} |D^4 z|_2 ,$$

$$|D(I^N z - z)|_{V(q_1)} \leq q_1^{1/2} C_2 N^{-2} |D^4 z|_2 .$$

Next we establish a $V(q_1)$ norm Schmidt inequality for splines.

Lemma 4.2 Let $s \in S^3(\Delta^N)$. Then $|s|_{V(q_1)} \leq \sqrt{k_3}\, N\, q_1^{1/2} |s|_2$, and $|Ds|_{V(q_1)} \leq \sqrt{k_2}\, N q_1^{1/2} |Ds|_2 = \sqrt{k_2}\, N |s|_{V(q_1)}$.

Proof. Take $s \in S^3(\Delta^N)$. Then

$$\int_0^1 |Ds(x)|^2 dx = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} |Ds(x)|^2 dx \leq \sum_{i=1}^{N} k_3 N^2 \int_{x_{i-1}}^{x_i} |s(x)|^2 dx$$
$$= k_3 N^2 \int_0^1 |s(x)|^2 dx,$$

where we have used Lemma 2.2. This yields the first inequality. Since Ds is a polynomial of degree 2 on $[x_{i-1}, x_i]$, a similar argument provides the second estimate.

Finally we have the desired projection estimates in the $V(q_1)$ norm.

Lemma 4.3. Let $P_1^N : H_0^1 \to S_0^3(\Delta^N)$ be the orthogonal projection in the $V(q_1)$ norm. Then for $z \in H^4 \cap H_0^1$ we have

$$|P_1^N z - z|_{V(q_1)} \le q_1^{1/2} C_1 N^{-3} |D^4 z|_2 \text{ , and}$$

$$|D(P_1^N z - z)|_{V(q_1)} \le q_1^{1/2}(2\sqrt{k_2}C_1 + C_2)N^{-2}|D^4 z|_2 \text{ .}$$

Proof. The first estimate follows directly from

$$|P_1^N z - z|_{V(q_1)} \le |I^N z - z|_{V(q_1)} \le q_1^{1/2} C_1 N^{-3} |D^4 z|_2 \text{ ,}$$

since the interpolate $I^N z$ in $S_0^3(\Delta^N)$ is the same as that in $S^3(\Delta^N)$ whenever $z \in H^4 \cap H_0^1$ - see Lemma 4.1. To obtain the second estimate, we first write

$$|D(P_1^N z - z)|_{V(q_1)} \le |D(P_1^N z - I^N z)|_{V(q_1)} + |D(I^N z - z)|_{V(q_1)} \text{ .}$$

The desired estimate on the second term follows directly from the second estimate in Lemma 4.1. To estimate the first term, we use the Schmidt inequality (from Lemma 4.2) along with Lemma 4.1 to obtain

$$|D(P_1^N z - I^N z)|_{V(q_1)} \le \sqrt{k_2} N |P_1^N z - I^N z|_{V(q_1)}$$

$$\le \sqrt{k_2} N [|P_1^N z - z|_{V(q_1)} + |z - I^N z|_{V(q_1)}]$$

$$\le 2\sqrt{k_2} N |I^N z - z|_{V(q_1)} \le 2\sqrt{k_2} N^{-2} C_1 q_1^{1/2} |D^4 z|_2 \text{ .}$$

Proposition 4.1. Assuming (HQ), the projections P^N converge strongly to I in X.

Proof. Let $z = (z_1, z_2) \in H^4 \cap H_0^1 \times H^4 \cap H_0^1$. Then $P^N z \equiv (z_1^N, z_2^N) \equiv (P_1^N z_1, P_2^N z_2)$, where P_1^N is the projection of the first coordinate into $S_0^3(\Delta^N)$ in the $V(q_1)$ norm and P_2^N is the projection of the second coordinate into $S_0^3(\Delta^N)$ in the H^0 norm. Then

$$|P^N z - z|^2 = |P_1^N z_1 - z_1|^2_{V(q_1)} + |P_2^N z_2 - z_2|_2^2$$
$$\leq (q_1^{1/2} C_1 N^{-3} |D^4 z_1|_2)^2 + (\tilde{C}_0 N^{-4} |D^4 z_2|_2)^2 .$$

This estimate together with the boundedness of P^N imply that $P^N z \to z$ for any $z \in X$.

Finally we discuss the parameter dependent convergence of the linear semigroups.

Theorem 4.1. Let (HQ) hold. Then the semigroups $T(t;q)$ and $T^N(t;q)$ generated by $A(q)$ and $A^N(q)$ satisfy exponential bounds $||T(t;q)|| \leq e^{\omega t}$ and $||T^N(t;q)|| \leq e^{\omega t}$ for some real ω independent of N and q. Moreover for any sequence $\{q^N\}$ converging to q^* in Q we have $|T^N(t;q^N)z - T(t;q^*)z| \to 0$ uniformly on $[0,T]$ for each $z \in X$.

Proof. The bound $||T(t;q)|| \leq e^{\omega t}$ is, of course, a consequence of (4.4). As in section 2, $A^N(q)$ clearly generates a semigroup $T^N(t;q)$ whose exponential bound is seen directly from

$$\langle A^N(q)z,z\rangle_q \leq \langle A(q)P^N(q)z,\ P^N(q)z\rangle_q \leq \omega|P^N(q)z|_q^2 \leq \omega|z|_q^2 .$$

Here we used the fact that $P^N(q):X(q) \to X^N(q)$ is the orthogonal projection.

We now turn to Proposition 2.1 to establish the convergence result of this theorem. Let $\mathcal{B} = X(q^*)$, $\mathcal{B}^N = X^N(q^N)$ and $\tilde{A} = A(q^*)$, $\tilde{A}^N = A^N(q^N)$. Then ii) of Proposition 2.1 holds for the family of semigroups $T^N(t;q^N)$.

Letting $\pi^N: X(q^*) \to X(q^N)$ be the canonical isomorphism between $X(q^*)$ and $X(q^N)$, it follows immediately that $|\pi^N z| \to |z|$ from the hypothesis that $q^N \to q^*$ and thus i) of Proposition 2.1 is also satisfied.

To define $\mathcal{D}$, we note that $\tilde{\Phi}_j(x) = (\sqrt{2}/j\pi)\sin(j\pi x)$ and $\Phi_j(x) = \sqrt{2}\sin(j\pi x)$, $j = 1,2,\ldots$, form complete orthonormal sets for $V(1)$ and H^0 respectively. Let $\gamma_j^N(x) = (\tilde{\Phi}_j, 0)^T$, for $j = 1,\ldots,N$, and $\gamma_j^N = (0, \Phi_{j-N})^T$, for $j = N+1,\ldots,2N$. Then $\bigcup_{N=1}^{\infty}\{\{\gamma_j^N\}_{j=1}^{2N}\}$ forms a complete orthonormal set for $X(\hat{q})$, where $\hat{q} = (1,0,\ldots,0)$, and a complete orthogonal set for $X(q)$ for q arbitrary. We take $\mathcal{D}^N = \text{span}\{\gamma_1^N,\ldots,\gamma_{2N}^N\}$ and $\mathcal{D} = \bigcup_{N=1}^{\infty}\mathcal{D}^N$. For this choice of $\mathcal{D}$ it clearly follows that $\mathcal{D} \subset \text{dom}\, A(q^*)$ and $\overline{\mathcal{D}} = \mathcal{B}$. Also, from the definition of $A(q)$ it is easily argued that for $\lambda > 0$ sufficiently large $(\lambda I - A(q^*))\mathcal{D} = \mathcal{D}$, so that $(\lambda I - A(q^*))\mathcal{D}$ is dense in $\mathcal{B}$.

Finally, to establish the convergence part of the consistency hypothesis iii) in Proposition 2.1 (suppressing the notation π^N for the canonical isomorphism), we see that for each $z = (z_1, z_2) \in \mathcal{D}$, once again (see(2.11))

$$
\begin{aligned}
|A^N(q^N)z - A(q^*)z| &= |P^N(q^N)A(q^N)P^N(q^N)z - A(q^*)z| \\
&\le |(A(q^N) - A(q^*))P^N(q^N)z| + |A(q^*)(P^N(q^N)z - z)| + |(P^N(q^N) - I)A(q^*)z|.
\end{aligned}
\tag{4.7}
$$

These three terms will now be estimated separately. First we write explicitly the second term as

$$
\begin{aligned}
A(q^*)(P^N(q^N)z - z) &= \begin{pmatrix} 0 & 1 \\ q_1^* D^2 + q_3^* & q_2^* \end{pmatrix} (P^N(q^N)z - z) \\
&= \begin{pmatrix} z_2^N - z_2 \\ q_1^* D^2(z_1^N - z_1) + q_3^*(z_1^N - z_1) + q_2^*(z_2^N - z_2) \end{pmatrix}.
\end{aligned}
$$

Thus

$$|A(q^*)(P^N(q^N)z - z)|^2 = |z_2^N - z_2|^2_{V(q_1^*)} + |q_1^* D^2(z_1^N - z_1) + q_3^*(z_1^N - z_1) + q_2^*(z_2^N - z_2)|_2^2$$

$$\leq q_1^*|D(z_2^N - z_2)|_2^2 + \{q_1^*|D^2(z_1^N - z_1)|_2 + |q_3^*||z_1^N - z_1|_2 + |q_2^*||z_2^N - z_2|_2\}^2 .$$

Using the fact that $\sqrt{q_1^*}\,|D^2(z_1^N - z_1)|_2 = |D(z_1^N - z_1)|_{V(q_1^*)}$, we see that this last estimate together with Lemma 2.3 and Lemma 4.3 imply that $|A(q^*)(P^N(q^N)z - z)| \to 0$ as $N \to \infty$.

For the last term in (4.7) we have convergence from $P^N(q^N) \to I$ strongly on X. Finally convergence of the first term on the right in (4.7) can be seen from

$$(A(q^N) - A(q^*))P^N(q^N)z = \begin{pmatrix} 0 \\ (q_1^N - q_1^*)D^2 z_1^N + (q_3^N - q_3^*)z_1^N + (q_2^N - q_2^*)z_2^N \end{pmatrix},$$

and the fact that $D^2 z_1^N \to D^2 z_1$, $z_1^N \to z_1$, and $z_2^N \to z_2$ in H^0 and $q^N \to q^*$.

Thus Proposition 2.1 is applicable to establish the desired convergence in the theorem.

5. Implementation and Examples: the Hyperbolic Case.

This section is devoted to a discussion of the computer implementation of the theory developed in the previous section and to the documentation of some of our numerical findings. As in the parabolic case, the optimization part of the algorithm was carried out using the IMSL implementation of the Levenberg-Marquardt algorithm. For given values of N and q the state approximations must be obtained by solving (4.5). Again the matrices involved in this equation have a structure which permits efficient solution of the system of the approximating ordinary differential equations. The required matrices for $A^N(q) = (Q^N)^{-1}K^N$ are given by

$$Q^N = \begin{pmatrix} Q_1^N & 0 \\ 0 & Q_2^N \end{pmatrix} \quad , \quad K^N = \begin{pmatrix} 0 & K_1^N \\ K_2^N & K_3^N \end{pmatrix} ,$$

where $(Q_1^N)_{i,j} = (K_1^N)_{i,j} = \langle B_i^N, B_j^N \rangle_{V(q_1)}$, $(Q_2^N)_{i,j} = \langle B_i^N, B_j^N \rangle_2$, $(K_2^N)_{i,j} = \langle q_1 D^2 B_i^N + q_3 B_i^N, B_j^N \rangle_2$, and $K_3^N = q_3 Q_2^N$. Equivalently, if we let $(A_1^N)_{i,j} = \langle DB_i^N, DB_j^N \rangle_2$ and $(A_2^N)_{i,j} = \langle B_i^N, B_j^N \rangle_2$, these matrices become

$$Q^N = \begin{pmatrix} q_1 A_1^N & 0 \\ 0 & A_2^N \end{pmatrix}$$

and

$$K^N = \begin{pmatrix} 0 & q_1 A_1^N \\ -q_1 A_1^N + q_3 A_2^N & q_2 A_2^N \end{pmatrix} .$$

The matrices A_1^N and A_2^N are exactly those which were used in the parabolic case and hence they were computed analytically and stored. The "inversion" of Q^N in computing $(Q^N)^{-1}K^N w$ is carried out as follows: We have

$$(Q^N)^{-1}K^N = \begin{pmatrix} q_1^{-1}(A_1^N)^{-1} & 0 \\ 0 & (A_2^N)^{-1} \end{pmatrix} \begin{pmatrix} 0 & q_1 A_1^N \\ -q_1 A_1^N + q_3 A_2^N & q_2 A_2^N \end{pmatrix}$$

$$= \begin{pmatrix} 0 & I \\ -q_1 (A_2^N)^{-1} A_1^N + q_3 I & q_2 I \end{pmatrix} .$$

Given $w = \text{col}(w_1, w_2)$ where w_1, w_2 are vectors of length $N+1$, we compute $\tilde{w} = \text{col}(\tilde{w}_1, \tilde{w}_2) = (Q^N)^{-1}K^N w$ by first computing $y = A_1^N w_1$. Then, since A_2^N is stored as its Cholesky factor L_N, where $A_2^N = L_N L_N^T$, we compute $(A_2^N)^{-1} A_1^N w_1$ by solving $L_N z = y$, $L_N^T r = z$, for r. Finally we have $\tilde{w} = \text{col}(w_2, -q_1 r + q_3 w_1 + q_2 w_2)$.

The projection of the initial data and the nonlinear term again require numerical quadrature. Recall that the first coordinate of the projection map is projection with respect to the $V(q_1)$ inner product. Given the initial data $u(0,x) = q_4 \varphi(x)$, $v(0,x) = q_5 \psi(x)$, we compute

$$(Q^N)^{-1} R^N \begin{pmatrix} q_4 \varphi \\ q_5 \psi \end{pmatrix} = \begin{pmatrix} (Q_1^N)^{-1} & 0 \\ 0 & (Q_2^N)^{-1} \end{pmatrix} \begin{pmatrix} q_4 \langle \varphi, B_i^N \rangle_{V(q_1)} \\ q_5 \langle \psi, B_i^N \rangle_2 \end{pmatrix}$$

$$= \begin{pmatrix} -q_4 (A_1^N)^{-1} \langle \varphi, D^2 B_i^N \rangle_2 \\ q_5 (A_2^N)^{-1} \langle \psi, B_i^N \rangle_2 \end{pmatrix} ,$$

where $\langle \varphi, D^2 B_i^N \rangle_2$ and $\langle \psi, B_i^N \rangle_2$ are computed via a composite

Gauss-Legendre rule and the "inversions" of A_1^N, A_2^N are carried out through the Cholesky decomposition. The projection of the nonlinear term is computed in the same manner. We point out here that for this class of hyperbolic equations the projection operator $P^N(q):X(q) \to X^N(q)$ actually does not depend on q, as can be seen from the above calculations.

The approximating ordinary differential equations were again solved using the IMSL package DGEAR, but for the hyperbolic examples the Adams multistep option of that package was chosen. "Data" were generated as in the parabolic case above (using an independent finite difference scheme to generate solutions corresponding to "true" parameter values). Except where otherwise noted, the observations used in the following examples consisted of the solution values $\hat{u}(t_i,x_j)$ at three spatial points (x_j = .25,.5,.75) sampled at ten times (t_i = .2,.4,...,2.0).

Example 5.1. The first example for hyperbolic equations involves the linear initial-boundary value problem

$$u_{tt} = q_1 u_{xx} + q_2 u_t ,$$
$$u(t,0) = u(t,1) = 0 ,$$
$$u(0,x) = 2x(1-x),$$
$$u_t(0,x) = 0 .$$

The numerical results can be found in Table 5, where the same notation as in section 3 is used.

TABLE 5

N	$\bar{q}_1^N$	$\bar{q}_2^N$
4	1.9967	-.9984
true value	2.0	-1.0
$q^{N,0}$	1.4	.4

Example 5.2. As another linear example we consider

$$u_{tt} = q_1 u_{xx} + q_3 u ,$$
$$u(t,0) = u(t,1) = 0 ,$$
$$u(0,x) = 2x(1-x),$$
$$u_t(0,x) = 0 .$$

As in the previous example, the approximations are very accurate even for low values of N; see Table 6.

TABLE 6

N	$\bar{q}_1^N$	$\bar{q}_3^N$
4	2.0130	-.8684
8	2.0004	-.9987
true value	2.0	-1.0
$q^{N,0}$	1.4	0

Example 5.3. Consider the nonlinear hyperbolic equation

$$u_{tt} = q_1 u_{xx} - u_t - q_6 u^3,$$
$$u(t,0) = u(t,1) = 0,$$
$$u(0,x) = 2x(1-x),$$
$$u_t(0,x) = 0.$$

Our numerical results are given in Table 7. Note that the estimates are very good at N=8, while numerical error in implementing the scheme (e.g. inverting large matrices) is apparent at N=16.

TABLE 7

N	$\bar{q}_1^N$	$\bar{q}_6^N$
4	1.9894	2.0538
8	1.9992	.9410
16	1.9979	1.0833
true value	2.0	1.0
$q^{N,0}$	1.0	.5

Example 5.4. Here we consider the nonlinear problem

$$u_{tt} = q_1 u_{xx} - u_t + q_6 \frac{u}{1+u},$$

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

$$u(0,x) = 2x(1-x),$$

$$u_t(0,x) = 0.$$

The observations for the results detailed in Table 8 consisted of 3 spatial points and 10 time samples, as in the previous examples. This example was also run using only two time samples ($t = .2,.4$) at the 3 spatial points and rapid convergence of the estimated parameters to the true values was also observed in this case.

TABLE 8

N	$\bar{q}_1^N$	$\bar{q}_6^N$
4	2.0084	3.0866
8	2.0003	3.0125
true value	2.0	3.0
$q^{N,0}$	1.0	1.0

Example 5.5. A third type of nonlinearity was used in the example

$$u_{tt} = q_1 u_{xx} + q_2 u_t + q_6 \sin u$$

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

$$u(0,x) = 2x(1-x),$$

$$u_t(0,x) = 0.$$

The reader will find the results in Table 9.

TABLE 9

N	$\bar{q}_1^N$	$\bar{q}_2^N$	$\bar{q}_6^N$
4	1.9549	-.9994	2.5559
8	2.0023	-.9992	3.0351
16	1.9984	-.9998	2.9969
true value	2.0	-1.0	3.0
$q^{N,0}$	1.0	0	1.0

Example 5.6. As a final example we present in Table 10 the results for

$$u_{tt} = q_1 u_{xx} + q_3 u + q_6 \frac{1}{1+u} \, ,$$

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

$$u(0,x) = 2x(1-x)$$

$$u_t(0,x) = 0.$$

TABLE 10

N	$\bar{q}_1^N$	$\bar{q}_3^N$	$\bar{q}_6^N$
4	1.9653	.6509	2.9892
8	2.0053	1.0516	2.9988
16	1.9999	.9971	3.0013
true value	2.0	1.0	3.0
$q^{N,0}$	1.0	0	1.0

6. Concluding Remarks.

In this paper we have shown that spline approximations may be profitably used to develop schemes for estimation of unknown parameters in initial-boundary value problems for second order partial differential equations. The practical utility of our ideas is supported by our computational experience with a large number of examples, a summary of which is also given in this paper. The use of spline functions in the context of parameter identification problems is not new; see, for instance, references found in the survey articles [2], [12],[16],[18]. It is, however, our belief that ours is the first presentation of a complete theoretical treatment (i.e. convergence proofs for the parameters, optimal states and optimal fit-to-data values) for spline-based methods for a large class of equations along with reports on a careful numerical testing of the methods on examples.

The fundamental ideas - involving use of a semigroup theoretic framework for the approximation of identification and control problems governed by partial differential equations - which are the basis of the convergence results of the present paper were first announced in [8]. A complete presentation of details of the approximation framework was later given in [9]. In fact the present paper is a companion paper to [9] detailing our work on spline approximations that was performed simultaneously with our efforts on modal approximations; a report on the latter is contained in [9] as an application of the general framework developed there.

The essential step in our considerations above and in [9] in approximation of the abstract differential equation is the approximation of the infinitesimal generator A by $P^N A P^N$. The importance of this classical approximation of an unbounded operator by a sequence of bounded operators to problems in system theory using semigroup methods (in particular the Trotter-Kato theorem - Prop. 2.1 above) was first, we believe, pointed out in [7], where the $P^N A P^N$ scheme was used with spline approximations for functional differential equations.

In this paper we chose two specific classes of examples to illustrate use of our spline approximation ideas. Particular boundary conditions and fit-to-data criteria were selected. However, given the general results of [9] and the technical developments discussed above, it should be obvious that many rather easy but important generalizations are possible. We mention a few of these which we have already investigated in applying our ideas to a number of specific problems. First, methods for problems with mixed boundary conditions involving unknown parameters can be readily obtained - see [5] for a partial discussion of some of our results in this direction. Furthermore, we have previously discussed elsewhere [4] how one might use our ideas to develop techniques for estimation of unknown variable coefficients in partial differential equations. Indeed we have already successfully used spline methods along the lines developed above to estimate spatially varying transport coefficients in insect dispersal models from field data (in collaboration with P. Kareiva). (These efforts are detailed in separate manuscripts currently in preparation.) In related efforts a rather different approach (not involving linear semigroup approximation results) is taken in [6] where spline methods for estimation of time varying coefficients are developed.

The framework of [9] and the present paper can also be used to treat higher order equations such as those arising in the study of large space structures and other elastic bodies (see [4],[5]). In particular we have in this spirit developed methods employing cubic and quintic spline approximations for equations such as those arising in both the Euler-Bernoulli and Timoshenko theories for beams. The ideas and some of our numerical findings are reported in [5].

In a modification of the theory presented here (not a $P^N A P^N$ type scheme) but using essentially the same semigroup approximation framework, we have developed methods that permit one to use lower order splines (e.g., piecewise linear splines for second order equations) in place of the cubic splines employed in this paper. A preliminary discussion of theoretical and computational results for second order equations can be found in [13]. A complete discussion of these ideas for second and higher order equations along with numerical results will be published in a manuscript currently in preparation.

All of the specific results in this paper and those mentioned above to appear elsewhere entail spline approximations for problems involving one dimensional spatial domains ($x \in R^1$). However, we have investigated extensions of our ideas to treat problems in higher dimensional domains ($x \in R^n$) and, not surprisingly, there appears to be little difficulty in making the needed extensions of the theory (e.g., see some of the estimates for approximation using product basis elements in Chapter 6 of [20]). We are currently considering the computational feasibility of the spline methods for such problems, but expect once again to find that they perform well.

Finally we note that the results in this paper and the framework of [9] concern development of schemes that involve semi-discretization in that one approximates the original partial differential equation by a sequence of ordinary differential equations. One can also develop a theoretical framework to combine spline approximations in the spatial coordinates with time discretizations resulting in approximation by a sequence of difference equations. In particular the theory developed in [19] involving full discretization methods for functional differential equations can be used to develop analogous techniques for partial differential equations. We are currently pursuing investigations of these ideas.

REFERENCES

[1] R. A. Adams, Sobolev Spaces, Academic Press, N.Y., 1975.

[2] K. J. Aström and P. Eykholf, System identification - A survey, Automatica, 7(1971), 123-162.

[3] H. T. Banks, Parameter identification techniques for physiological control systems, in Lectures in Applied Math (F. Hoppensteadt, ed.), Vol. 19 (1980), p. 361-383.

[4] H. T. Banks, A survey of some problems and recent results for parameter estimation and optimal control in delay and distributed parameter systems, LCDS Tech. Rep. 81-19, July 1981, to appear in Proc. Conf. on Volterra and Functional Differential Equations, VPI, Blacksburg, Va., June 10-13, 1981.

[5] H. T. Banks and J. M. Crowley, Parameter estimation for distributed systems arising in elasticity, LCDS Tech. Rep. 81-24, Nov. 1981; to appear in Proc. NCKU/AAS Symposium on Engineering Sciences and Mechanics, Tainan, Taiwan, Dec. 28-31, 1981.

[6] H. T. Banks and P. L. Daniel, Parameter estimation of nonlinear nonautonomous distributed systems, Proc. 20th IEEE Conf. on Decision and Control, San Diego, Dec. 1981.

[7] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, J. Differential Equations 34 (1979), 496-522.

[8] H. T. Banks and K. Kunisch, Parameter estimation techniques for nonlinear distributed parameter systems, to appear in Proc. Internat. Conference on Nonlinear Phenomena in the Mathematical Sciences, (U.Texas-Arlington, June 1980).

[9] H. T. Banks and K. Kunisch, An approximation theory for nonlinear partial differential equations with applications to identification and control, LCDS - Tech. Rep. 81-7, Brown University, April 1981; to appear in SIAM J. Control and Optimization.

[10] R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. 1, Interscience, N.Y., 1953.

[11] S. G. Krein, Linear Differential Equations in Banach Spaces, Transl. Math. Monographs, AMS, Providence, 1971.

[12] C. S. Kubrusly, Distributed parameter system identification, a survey, Int. J. Control, 26 (1977), 509-535.

[13] K. Kunisch, Identification and estimation in abstract Cauchy problems, to appear in Proc. Special Semester on Control, (Fall, 1980), S. Banach Center, Warsaw.

[14] T. G. Kurtz, Extensions of Trotter's operator semigroup approximation theorem, J. Functional Anal., 3 (1969), 354-375.

[15] R. H. Martin, Nonlinear Operators and Differential Equations in Banach Spaces, John Wiley & Sons, N.Y., 1976.

[16] M. P. Polis and R. E. Goodson, Parameter identification in distributed systems: a synthesizing overview, Proc. IEEE, 64 (1976), 45-60.

[17] P. M. Prenter, Splines and Variational Methods, Wiley-Interscience, N.Y., 1975.

[18] W. H. Ray, Some recent applications of distributed parameter systems theory - a survey, Automatica 14 (1978), 281-287.

[19] I. G. Rosen, A discrete approximation framework for hereditary systems, J. Differential Equations 40 (1981), 377-449.

[20] M. H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, 1973.

[21] M. H. Schultz and R. S. Varga, L-spines, Numer. Math., 10 (1967) 345-369.

FILMED

8