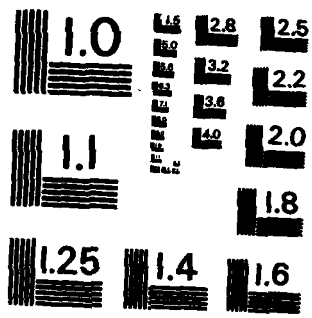
AD-A122 925 AN APPROXIMATION THEORY FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS WITH..(U) BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMICAL SYSTE.. 1/2

UNCLASSIFIED H T BANKS ET AL. 01 MAY 81 ARO-16835.27-MA F/G 12/1 NL

END
DATE
FILMED
2 83
DTIC

CONT


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

ADA122925

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 16835.27-MA	2. GOVT ACCESSION NO. AD-A122 92 N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) An Approximation Theory for Nonlinear Partial Differential Equations with Applications to Identification and Control		5. TYPE OF REPORT & PERIOD COVERED Reprint
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) H. T. Banks K. Kunisch		8. CONTRACT OR GRANT NUMBER(s) DAAG29 79 C 0161
9. PERFORMING ORGANIZATION NAME AND ADDRESS Brown University Providence, RI 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS N/A
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE Nov 82
		13. NUMBER OF PAGES 35
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)

Submitted for announcement only.

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)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SIAM J. CONTROL AND OPTIMIZATION
Vol. 20, No. 6, November 1982

0363-0129/82/2006-0006 $01.00/0

AN APPROXIMATION THEORY FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS WITH APPLICATIONS TO IDENTIFICATION AND CONTROL*

H. T. BANKS† AND K. KUNISCH‡

Abstract. Approximation results from linear semigroup theory are used to develop a general framework for convergence of approximation schemes in parameter estimation and optimal control problems for nonlinear partial differential equations. These ideas are used to establish theoretical convergence results for parameter identification using modal (eigenfunction) approximation techniques. Results from numerical investigations of these schemes for both hyperbolic and parabolic systems are given.

1. Introduction. When modeling real-world phenomena one often encounters a situation where a priori knowledge leads one to conjecture a certain type of model equation containing parameters which are unknown. In this paper we are primarily concerned with techniques for recovery of these unknown quantities from given data. In §§ 2 and 3 we present a quite general framework for approximation schemes for abstract nonlinear Cauchy problems. These approximation results are subsequently applied to modal techniques for identification and control problems in §§ 4 and 5, respectively. A summary of some of our numerical experience with parameter estimation problems using these techniques is given in § 6. The examples here were chosen so as to illustrate the feasibility and effectiveness of the method and to investigate possible inherent difficulties. We are quite confident that the ideas outlined here will be applicable in a variety of research areas where mathematical models for the phenomena under study are used. In a forthcoming monograph we shall discuss in more detail identification problems that arise in several areas of applications [35] including seismology [3], [10], [18], reservoir engineering [11], [17], [38], glaciology [16], physics [37], biology [4], [5], [29], [34] and large space structures technology. While our treatment here is restricted to constant unknown parameters, the theoretical ideas extend in large part to problems with unknown function parameters. Indeed, we are currently applying some of our techniques to specific problems from the areas mentioned above; in some cases these efforts involve identification of functions.

In this paper the general approximation results are used to carefully discuss modal approximation schemes for certain classes of partial differential equations (see §§ 4 and 5). Such schemes for specific identification and control problems are, of course, not new. Many discussions in the literature, however, are in the context of very specific examples and frequently no convergence proofs or evidence of numerical studies are supplied. Modal approximations have many advantages, including: they are readily discussed and understood in terms of classical spectral results; they are familiar to

* Received by the editors May 1, 1981, and in revised form January 11, 1982. This work was supported in part by the Air Force Office of Scientific Research under contract AFOSR 76-3092D, in part by the National Science Foundation under grant NSF-MCS 7905774-02, and in part by the U.S. Army Research Office under contract ARO-DAAG 29-79-C-0161.

† Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912. 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 NAS1-15810 and NAS1-16394.

‡ Institut für Mathematik, Technische Universität Graz, Kopernikusgasse 24, 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 and the Fulbright Commission.

and readily implemented by practicing engineers, and they give rise to a simple algebraic structure for the approximating ordinary differential equations. However, modal approximations do have some shortcomings, of which we mention several. First, in many practical problems it is very difficult to calculate the true natural modes. Secondly, for certain parabolic partial differential equations modal approximations by their very nature lead to stiff systems of approximating ordinary differential equations. Finally, one can encounter lack of "numerical identifiability" (i.e., the identification problems for the approximating ordinary differential equations yield parameter estimates that converge to different values for different sets of initial estimates) regardless of the well-posedness of the parameter estimation problem for the original partial differential equation model. With respect to the first difficulty pointed out here, we refer to Example 4.4, below, where we explain a "modal" approximation scheme for an identification problem which does not employ the natural modes of the system. For one solution of the latter problems, our experience indicates that for certain classes of parabolic problems spline-based approximation schemes can be more efficient. Details on this aspect of spline methods, along with a number of other features of these techniques, will be given in a separate manuscript currently in preparation.

The parameter identification and estimation problem has received a great amount of attention in the engineering literature and we refer to [1], [23], [31], [32] for review articles. In the future monograph alluded to above, we shall survey the research efforts from the engineering as well as from the mathematical literature. Much of the mathematical literature is concerned with the problem of identifiability, which, loosely speaking, is defined as the problem of injectivity of the map from the set of parameters to the set of outputs. Although this is a very important theoretical and practical question, it will not be a part of the discussion of the present paper.

We point out one important technical aspect that will become clearer in Examples 4.1 and 4.4 below. In general, the eigenfunctions of the model equation will depend on the parameters that are to be identified. For modal approximation schemes this is an extremely undesirable feature from the point of view of implementation, since in practical examples the representation of the operators in the approximating equations will involve a matrix of inner products of the eigenfunctions. It is, of course, desirable to have this matrix independent of the unknown parameters to avoid excessive numerical integrations when performing iterative searches on these parameters.

Our focus in this paper is on the development of semidiscrete approximation schemes for parameter identification and control problems which result in approximating problems governed by ordinary differential equations. Of course, full discretization methods (discretization in time as well as spatial coordinates, resulting in problems governed by difference equations) are of great importance and our investigations of a related theoretical framework, as well as detailed schemes for such an approach, will be reported elsewhere.

In summary, the emphasis of our presentation is twofold. First, we present a general theoretical framework, with unknown parameter-dependent spaces, which can be used to treat many types of problems (including estimation of function space parameters) and approximation schemes (see the remarks in § 7 below). As a concrete example of the use of this framework, we give a detailed treatment of "modal" approximation schemes, thereby putting on a sound theoretical foundation a class of methods that have been used in an ad hoc way by scientists and engineers for some time.

The notation used throughout the paper is quite standard. We employ the usual notation H^i for Sobolev spaces with "functions" and "derivatives" in L_2, and $|\cdot|$ to denote norms of elements, as well as those of operators. Only in cases where confusion may arise will we use subscripts to distinguish norms in various spaces.

2. The abstract identification problem and its approximation. We consider the abstract semilinear Cauchy problem

$$\tag{2.1} \begin{aligned} \dot{u}(t) &= A(q)u(t) + F(q, t, u(t)), \qquad t > 0, \\ u(0) &= u_0(q), \end{aligned}$$

where for each $q \in Q \subset R^k$, $A(q)$ is the infinitesimal generator of a linear C_0-semigroup $\{T(t;q)\}_{t\geq 0}$ on a real Hilbert space $X(q)$ with inner product $\langle\cdot,\cdot\rangle_q$ and norm $|\cdot|_q$ (denoted sometimes below by X, $\langle\cdot,\cdot\rangle$ and $|\cdot|$, respectively, when no loss of clarity results from suppression of the q). We shall, throughout our discussions, employ the concept of *mild solutions*, so that $t \to u(t;q)$ is called a *solution* of (2.1) if it satisfies

$$\tag{2.2} u(t;q) = T(t;q)u_0(q) + \int_0^t T(t-s;q)F(q,s,u(s;q))\,ds.$$

We note that for solutions u we have $t \to u(t;q)$ continuous. The conditions that we impose on F below will guarantee existence and uniqueness of mild solutions u of (2.1) on any given finite interval $[0, T]$. We shall in certain specific instances below, be required to discuss briefly the relationship between mild and strong (in a classical almost-everywhere sense) solutions of (2.1), but for more general results we refer the reader to [28].

Throughout our presentation we shall assume that $X(q)$ is a function space of R^n-valued "functions" (possibly one of the usual Lebesgue spaces of equivalence classes of functions) defined on the fixed interval $[0, 1]$; consequently, we shall also use the notation $u(t, x; q)$ or $u(t, \cdot\,; q)$ when discussing solutions of (2.1).

While we shall also discuss control-theoretic applications, much of our attention will be directed towards the problem of identifying the parameter q in (2.1) from observations of the system. Specifically, we assume that (2.1) models some physical, biological, economic, etc., system for which output measurements $\hat{y}$ are available. These measurements may be available in the form of continuous data $\hat{y}(t)$, $0 \leq t \leq T$, or discrete data $\hat{y}(t_i)$, $0 \leq t_1 < \cdots < t_r \leq T$. We then seek to find a "best" value for q in Q by minimizing an appropriately defined fit-to-data criterion. To be specific in our formulation here, we shall assume discrete time observations with values $\hat{y}(t_i)$ in an observation space $\mathscr{Y}$. All of the results of this paper are easily extended to the case of identification problems where one has continuous time data, but we shall not pursue such problems here. Assuming, then, that a criterion function $J: Q \times C(0, T; X(q)) \times \prod_{i=1}^r \mathscr{Y} \to R^1$ is defined, we formally state the identification problem:

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

Several traditional choices of fit-to-data criteria are included in our formulation; namely, we may consider either integral or pointwise (in a spatial sense) evaluation least-squares sums in the above formulation. In the case of integral evaluation we are given measurements $\hat{y}(t_i) \in L_2(0, 1; R^\nu)$ where $\nu \leq n$ and an output map $Y(t, x, q): R^n \to R^\nu$ on the "state" $u(t, x; q)$. The observation space is given by $\mathscr{Y} = L_2(0, 1; R^\nu)$ and the criterion is defined by

$$\tag{2.3} J(q, u(\cdot\,;q), \hat{y}) = \sum_{i=1}^r \int_0^1 |\hat{y}(t_i, x) - Y(t_i, x, q)u(t_i, x; q)|^2\,dx.$$

We assume that Y is continuous in q and sufficiently regular in x so that $x \to Y(t_i, x, q)u(t_i, x; q)$ is in $L_2(0, 1; R^\nu)$. For the choice of pointwise or spatially discrete measurements, we assume that we have observations $\hat{y}(t_i) \in \mathcal{Y} = \prod_{j=1}^{l} R^\nu$, corresponding to measurements of the output at points $\{x_j\}_{j=1}^{l}$ in $[0, 1]$ at time t_i. These observations represent measurements for $C(t_i, q)\xi(t_i, q)$ where $\xi(t_i, q) = \mathrm{col}\,(u(t_i, x_1; q), \cdots, u(t_i, x_l; q))$ and $C(t_i, q)$ is an $(\nu l) \times (nl)$-matrix depending continuously on q for each fixed t_i. The associated fit-to-data criterion is then defined by

$$J(q, u(\cdot\,; q), \hat{y}) = \sum_{i=1}^{r} |\hat{y}(t_i) - C(t_i, q)\xi(t_i, q)|^2. \tag{2.4}$$

The output maps Y and C introduced in (2.3) and (2.4) are necessitated by the fact that often in practice one can observe only some components (say ν) of the n-dimensional vectors $u(t, x; q)$, and that these observations may depend on the time at which they are made. We further note that the point evaluations at x_j used in defining $\xi(t_i, q)$ above may be meaningless without additional restrictions on the state space, the initial data and/or the right side of the equation in (2.1). A more detailed discussion of the problems arising from use of criteria such as (2.4) when dealing with mild solutions will be given in the context of Example 4.3 below.

We turn next to formulating a sequence (ID^N) of approximating problems on Hilbert spaces $X^N(q)$ for our original identification problem (ID). These problems involve "states" governed by ordinary differential equations and are (in the specific instances we shall propose) tractable using standard numerical procedures. We state first a series of hypotheses and definitions that will be needed at various points in the sequel.

(H1) For each $N = 1, 2, \cdots, X^N(q)$ is a closed linear subspace of $X(q)$, endowed with the $X(q)$ topology.

(H2) The spaces $X(q)$, $q \in Q \subset R^k$, are set-theoretically equal and uniformly topologically isomorphic so that there exists a constant $\mathcal{K} \geq 1$ such that $|v|_{\hat{q}} \leq \mathcal{K}|v|_q$ for all $v \in X = X(q)$ and $q, \hat{q} \in Q$.

(H3) For each $q \in Q$, $A(q)$ generates a linear C_0-semigroup $T(t; q)$ on $X(q)$.

(H4) The set Q is a compact subset of R^k.

(H5) (i) For each $q \in Q$, let $P^N(q): X(q) \to X^N(q)$ denote the canonical orthogonal projections along $X^N(q)^\perp$ and let $A^N(q): X(q) \to X^N(q)$ be defined by $A^N(q) = P^N(q)A(q)P^N(q)$. For each N, let $A^N(q)$ generate a linear C_0-semigroup on $X(q)$ denoted by $T^N(t; q)$.
(ii) For each N, there exist constants $\hat{M} = \hat{M}(N)$ and $\hat{\omega} = \hat{\omega}(N)$, independent of q, such that $|T^N(t; q)| \leq \hat{M} e^{\hat{\omega} t}$.

(H6) (i) For each continuous function $u: [0, T] \to X = X(q)$ (see (H2)), the map $t \to F(q, t, u(t))$ is measurable.
(ii) For each constant $M > 0$, there exists a function $k_1 = k_1(M)$ in $L_2(0, T)$ such that for any $q, \hat{q} \in Q$ we have

$$|F(q, t, u_1) - F(q, t, u_2)|_{\hat{q}} \leq k_1(t)|u_1 - u_2|_{\hat{q}}$$

for all $u_1, u_2 \in X$ with $|u_i|_{\hat{q}} \leq M$.
(iii) There exists a function k_2 in $L_2(0, T)$ such that

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

for all $v \in X$, $q, \hat{q} \in Q$.
(iv) For each $(t, v) \in [0, T] \times X$, the map $q \to F(q, t, v)$ is continuous. (Again, under (H2) recall that all the sets $X = X(q)$ are the same.)

(H7) The projections $P^N(q):X(q)\to X^N(q)$ are such that for any sequence $\{q^N\}$ in Q satisfying $q^N \to \bar{q} \in Q$, one has $|P^N(q^N)z - z|_{q^N} \to 0$ as $N\to\infty$ for all $z \in X(\bar{q})$.
(H8) For each convergent sequence $q^N \to \bar{q}$ in Q, there are constants $\mathcal{M}$, ω such that $|T^N(t;q^N)|_{q^N} \le \mathcal{M}e^{\omega t}$, $|T(t;\bar{q})|_{q^N} \le \mathcal{M}e^{\omega t}$ uniformly in $N = 1, 2, \cdots$.
(H9) For each convergent sequence $q^N \to \bar{q}$ in Q, one has for $z \in X(\bar{q})$, $|T^N(t;q^N)z - T(t;\bar{q})z|_{q^N} \to 0$ as $N\to\infty$, uniformly in $t\in[0,T]$.

The assumption (H2) *will be taken as a standing hypothesis for the remainder of our discussions.* For the approximating schemes we develop below, consistency will follow from (H7) while (H8) is a statement of stability. As we shall see, convergence of the schemes (which is (H9)) will follow from (H7), (H8) and the Trotter–Kato theorem.

Remark 2.1. It suffices, under the standing assumption (H2), that the following condition hold in place of (H6)(ii): For some fixed $q^*\in Q$ we have that for each $M>0$ there is a function k_1 such that for all $q\in Q$ the relation $|F(q,t,u_1)-F(q,t,u_2)|_{q^*}\le k_1(t)|u_1-u_2|_{q^*}$ for all $u_1, u_2\in X$ with $|u_i|_{q^*}\le M$. Indeed, it is easily seen that this condition, along with (H2), implies (H6)(ii). Similarly, we can in the presence of (H2) equivalently postulate in place of (H6)(iii) the conditions: For some fixed $q^*\in Q$ there exists a function k_2 such that $|F(q,t,v)|_{q^*}\le k_2(t)\{|v|_{q^*}+1\}$ for all $v\in X$, $q\in Q$. We further note that existence of a function $k_3\in L_2(0,T)$ such that $|F(q,t,0)|_{\bar{q}}\le k_3(t)$ for $q,\bar{q}\in Q$, a statement of the inequality of (H6)(iii) holding only for $|v|_{\bar{q}}$ sufficiently large (i.e., affine growth at ∞), along with (H6)(ii), are sufficient to imply (H6)(iii).

While the complete role played by the various hypotheses in our development will be clearer after our presentation, a few explanatory comments here might be helpful to readers. First, the desirability of the generality of allowing the underlying Hilbert space X for (2.1) to depend on q in such a way that (H2) obtains will not be apparent from the examples discussed here. (Rather, one must for this consider certain parabolic problems—see the comments in § 7.) However, in light of (H2) as a standing assumption, we are justified in suppressing the canonical isomorphism $\mathcal{I}^N: X(\bar{q})\to X(q^N)$ in writing $|P^N(q^N)z - z|_{q^N}\to 0$ in (H7) rather than the technically correct statement $|P^N(q^N)\mathcal{I}^N z - \mathcal{I}^N z|_{q^N}\to 0$. Similar observations are pertinent for the statement of (H9) as well as in numerous other places in our presentation where we suppress the $\mathcal{I}^N$ notation.

Condition (H4), while seemingly stringent, is an assumption often valid in practical problems where our theory might be useful. Since under (H3) $A(q)$ is closed, it follows from the closed graph theorem that $A^N(q)$ of (H5)(i) is, in fact, bounded and hence (H5)(i) follows immediately from (H3). It should be recognized that the form of the approximating operators defined in (H5) is a classical one (e.g., see [33, p. 369]) which has also recently been employed in the development of spline approximation techniques for delay differential equations [6]. The definition of $A^N(q)$ involves the implicit assumption that $X^N(q)\subset \mathrm{Dom}\,(A(q))$; since our goal here is the rigorous formulation of modal approximation schemes for (ID), this restriction poses no difficulties. However, it does prevent a straightforward inclusion of low-order finite-element methods for higher-order partial differential equations in our approximation framework.

Hypothesis (H7) is a common requirement (e.g., see [24], [30]) in approximation theory, demanding that the sequence X^N of subspaces actually approximate the original state space X. Finally, (H6) is comprised of conditions on the nonlinearities in (2.1) that are sufficiently general to include many interesting problems of practical importance but yet are strong enough to guarantee global existence of solutions of (2.1) on fixed finite intervals. As the knowledgeable reader might expect, these conditions can be replaced by alternate and/or weaker, hypotheses, but only, in general, at the cost of additional tedium in the proofs below. We have tried to compromise between strong conditions that are easily stated and employed in the proofs and ones that are as general (and weak) as possible. Further comments on this matter will be made in § 3.

Before defining the approximating equations for (2.1), we define the projection of the nonlinearity F onto X^N by $F^N(q,t,v) = P^N(q)F(q,t,v)$ for each $(q,t,v) \in Q \times [0,T] \times X$. The approximating family of equations is then given by

$$\begin{aligned} \dot{v}(t) &= A^N(q)v(t) + F^N(q,t,v(t)), \qquad t>0, \\ v(0) &= P^N(q)u_0(q). \end{aligned} \tag{2.5}$$

Assuming existence of (mild) solutions to (2.5) (this will be established below), we denote (for a given q) these solutions by $u^N(t)$ or alternatively $u^N(t;q)$ or $u^N(t,x;q)$, depending on the context. We then define the approximate identification problems (ID^N) by:

(ID^N) Given observations $\hat{y} = \{\hat{y}(t_i)\}_{i=1}^r$ and a fit-to-data criterion J, minimize $J^N(q) = J(q, u^N(\cdot\,;q), \hat{y})$ over $q \in Q$ subject to $u^N(\cdot\,;q)$ satisfying (2.5).

We note that if (in addition to (H1)) $X^N(q)$ is finite-dimensional, then (2.5) can be equivalently interpreted in the strong sense and (ID^N) then becomes an optimization problem constrained by finite-dimensional ordinary differential equations.

In our discussions below, we shall denote by $\hat{q}^N$ any solution of (ID^N) so that it follows by definition that $J^N(\hat{q}^N) \le J^N(q)$ for all $q \in Q$.

PROPOSITION 2.1. *Assume that* (H2), (H3) *and* (H6) *obtain. Then for each* $q \in Q$ *there exists a unique* (*mild*) *solution* $u(\cdot\,;q) \in C(0,T;X(q))$ *of* (2.1). *If, in addition,* (H5) *holds, there exists, for each* $N = 1, 2, \cdots$, *a unique* (*mild*) *solution* $u^N(\cdot\,;q) \in C(0,T;X(q))$ *of* (2.5).

Proof. The proofs are completely standard and we only sketch the ideas for (2.1). Uniqueness follows immediately from (H6) and an application of Gronwall's inequality. Existence is established through the usual Picard iterate techniques. Define $v^0(t) = T(t;q)u_0(q)$ and for $j = 1, 2, \cdots$,

$$v^j(t) = T(t;q)u_0(q) + \int_0^t T(t-s;q)F(q,s,v^{j-1}(s))\,ds \tag{2.6}$$

for $t \in [0,T]$. From (H3) and (H6) it is easily seen that the iterates v^j are all well defined and $v^j\colon [0,T] \to X(q)$ is continuous. Moreover, $\{v^j\}_{j=0}^\infty$ is a bounded subset of $C(0,T;X)$. Employing (H6)(ii) and simple inductive arguments, one can establish that $\{v^j\}$ is Cauchy in $C(0,T;X)$. Passing to the limit in (2.6), one obtains the desired results. Existence of unique solutions of (2.5) is argued in an analogous manner by appealing to (H5) for appropriate boundedness.

THEOREM 2.1. *Assume hypotheses* (H1)–(H6) *hold and let* $J(\cdot,\cdot,\hat{y})\colon Q \times C(0,T;X) \to R^1$ *be continuous. Moreover, suppose* $q \to u_0(q)$, $q \to P^N(q)z$ *and* $q \to T^N(t;q)z$, $z \in X$, *are continuous, 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* $q^j \to q^0$, *we have* $|u^j(t;q^j) - u(t;q^0)|_{q^j} \to 0$ *as* $j \to \infty$ *uniformly in* $t \in [0,T]$, *then* $\hat{q}$ *is a solution of* (ID).

Proof. We show for fixed N that $q \to J^N(q) \equiv J(q, u^N(\cdot\,; q), \hat{y})$ is continuous on Q which, in the light of (H4), yields (i). First note that u^N satisfies

$$(2.7)\quad u^N(t;q) = T^N(t;q)P^N(q)u_0(q) + \int_0^t T^N(t-s;q)P^N(q)F(q,s,u^N(s;q))\,ds$$

on $[0, T]$. In view of (H5) and (H2) we find that there exist constants M and M_1 such that $|T^N(t;q)|_{\tilde q} \leq M$, $|P^N(q)|_{\tilde q} \leq M_1$ uniformly for $t \in [0, T]$, $q, \tilde{q} \in Q$. It follows from (H6)(iii) and (2.7) that

$$|u^N(t;q)|_q \leq M|P^N(q)u_0(q)|_q + M\int_0^t |P^N(q)F(q,s,u^N(s;q))|_q\,ds$$

$$\leq M|u_0(q)| + M\int_0^t k_2(s)\{|u^N(s;q)|_q + 1\}\,ds.$$

Since $q \to u_0(q)$ is continuous, this implies (via Gronwall's inequality) that $|u^N(t;q)|_q$ is uniformly bounded for $(t,q) \in [0, T] \times Q$. This in turn implies (by (H6)(iii)) that the mapping $s \to T^N(t-s;q)P^N(\tilde{q})F(\tilde{q},s,u^N(s;\tilde{q}))$ from $[0, T]$ to X is dominated by an integrable function uniformly in $q, \tilde{q}, \tilde{q} \in Q$. (This will permit us to invoke, below, the usual dominated convergence theorem.) Assuming that $q^j \to \tilde{q}$, q^j, $\tilde{q} \in Q$ are arbitrary, we obtain the following estimates:

$$\begin{aligned}
&|u^N(t;\tilde{q}) - u^N(t;q^j)| \\
&\quad\leq |T^N(t;\tilde{q})P^N(\tilde{q})u_0(\tilde{q}) - T^N(t;q^j)P^N(\tilde{q})u_0(\tilde{q})| \\
&\qquad + |T^N(t;q^j)P^N(\tilde{q})u_0(\tilde{q}) - T^N(t;q^j)P^N(q^j)u_0(\tilde{q})| \\
&\qquad + |T^N(t;q^j)P^N(q^j)u_0(\tilde{q}) - T^N(t;q^j)P^N(q^j)u_0(q^j)| \\
&\qquad + \int_0^t |\{T^N(t-s;\tilde{q}) - T^N(t-s;q^j)\}P^N(\tilde{q})F(\tilde{q},s,u^N(s;\tilde{q}))|\,ds \\
&\qquad + \int_0^t |T^N(t-s;q^j)(P^N(\tilde{q}) - P^N(q^j))F(\tilde{q},s,u^N(s;\tilde{q}))|\,ds \\
&\qquad + \int_0^t |T^N(t-s;q^j)P^N(q^j)\{F(\tilde{q},s,u^N(s;\tilde{q})) - F(q^j,s,u^N(s;q^j))\}|\,ds \\
&\quad = \rho_1(j) + \rho_2(j) + \rho_3(j) + \rho_4(j) + \rho_5(j) + \rho_6(j),
\end{aligned}$$

where the ρ_i's are defined as indicated (ρ_i the ith term), $i = 1, \cdots, 6$, and all norms are $|\cdot|_{\tilde q}$. We then have by hypothesis

$$\rho_1(j) = |\{T^N(t;\tilde{q}) - T^N(t;q^j)\}P^N(\tilde{q})u_0(\tilde{q})| \to 0 \quad \text{as } j \to \infty,$$

uniformly in $t \in [0, T]$. Also ρ_2 and $\rho_3 \to 0$ by the continuity assumptions on P^N and u_0 and the boundedness of $T^N(t;q^j)$. Dominated convergence implies that $\rho_4 \to 0$ and $\rho_5 \to 0$ as $j \to \infty$. Finally,

$$\rho_6(j) \leq MM_1\left\{\int_0^t |F(\tilde{q},s,u^N(s;\tilde{q})) - F(q^j,s,u^N(s;\tilde{q}))|\,ds\right.$$

$$\left. + \int_0^t |F(q^j,s,u^N(s;\tilde{q})) - F(q^j,s,u^N(s;q^j))|\,ds\right\}$$

$$\leq MM_1\rho_7(j) + \int_0^t k_1(s)|u^N(s;\tilde{q}) - u^N(s;q^j)|\,ds,$$

where $\rho_7(j)\to 0$ as $j\to\infty$ and k_1 depends on the uniform bounds for $u^N(t;\bar q)$, $u^N(t;q^j)$ (see (H6)(ii)). Thus we find

$$|u^N(t;\bar q)-u^N(t;q^j)|\le \varepsilon_j(t)+\int_0^t k_1(s)|u^N(s;\bar q)-u^N(s;q^j)|\,ds,$$

where $\varepsilon_j(t)\to 0$, uniformly in $t\in[0,T]$, as $j\to\infty$. Applying Gronwall's inequality, we have uniform (in t) continuity of $q\to u^N(t;q)$ on Q which implies the desired continuity of J^N.

Turning to (ii) and letting $\{\hat q^{N_j}\}$ be a convergent subsequence with $\hat q^{N_j}\to\hat q$, we first observe that $J^{N_j}(\hat q^{N_j})\le J^{N_j}(q)$ for all $q\in Q$. By hypothesis, $u^{N_j}(t;\hat q^{N_j})\to u(t;\hat q)$ and furthermore $u^{N_j}(t;q)\to u(t;q)$ for each $q\in Q$, with convergence in both cases uniform in t on $[0,T]$. This implies $J^{N_j}(\hat q^{N_j})\to J(\hat q,u(\cdot;\hat q),\hat y)$ and $J^{N_j}(q)\to J(q,u(\cdot;q),\hat y)$ as $j\to\infty$ and hence from the above inequality we obtain, by passing to the limit, $J(\hat q,u(\cdot;\hat q),\hat y)\le J(q,u(\cdot;q),\hat y)$ for all $q\in Q$. Thus $\hat q$ is a solution of (ID) and Theorem 2.1 is established.

Remark 2.2. If $v\to J(q,v,\hat y)$ as a mapping on $C(0,T;X)$ actually depends only on a finite number of values $v(t_i)$, $t_i\in[0,T]$, as, for example, in (2.3) or (2.4), then the hypotheses of Theorem 2.1 involving uniformity in t can be relaxed to statements holding only for each fixed $t\in[0,T]$. The above arguments remain unchanged except that the uniform (in t) convergence remarks are replaced with pointwise convergence statements (see especially the term $\rho_1(j)$ in the proof).

We conclude this section with a brief explanation of how (2.5) (or (2.7)) is to be used in actual computations. We adopt notation very similar to that found in [6] in the development of spline methods for delay systems. We assume that X^N is finite-dimensional and choose a basis independent of q (recall that (H2) is a standing hypothesis) by

$$\hat\beta^N=(\beta_1,\cdots,\beta_{d(N)})$$

where $d(N)=\dim X^N(q)$. From (2.7) under (H5) we see that the solution u^N of (2.5) satisfies $u^N(t;q)\in X^N$ for all t and hence there exists a representation $u^N(t;q)=\hat\beta^N w^N(t;q)$ with $w^N(t;q)=\mathrm{col}\,(w_1^N(t;q),\cdots,w_{d(N)}^N(t;q))\in R^{d(N)}$. We let $[A^N(q)]$ and $[F^N(q,t,w^N)]$ denote the matrix and coordinate representations relative to $\hat\beta^N$ of $A^N(q)$ and $P^N(q)F(q,t,u^N)$, respectively. The coordinate representation of (2.5) is then given by

$$\begin{aligned}&\dot w^N(t;q)=[A^N(q)]w^N(t;q)+[F^N(q,t,w^N(t;q))],\qquad t>0,\\ &w^N(0;q)=\gamma^N,\end{aligned}\tag{2.8}$$

where γ^N is defined through $P^N(q)u_0(q)=\hat\beta^N w^N(0;q)$. For any $z\in X^N$ the associated coordinate vector $\alpha^N\in R^{d(N)}$ in $P^N(q)z=\hat\beta^N\alpha^N$ is determined uniquely by the condition $(P^N(q)z-z)\perp X^N$, or equivalently $\langle\hat\beta^N,\hat\beta^N\rangle_q\alpha^N=\langle\hat\beta^N,z\rangle_q$. Thus we have

$$\alpha^N=(Q^N)^{-1}R^N z\tag{2.9}$$

where Q^N is the $d(N)\times d(N)$ matrix with elements $\langle\beta_i^N,\beta_j^N\rangle_q$ and $(R^N z)_i=\langle\beta_i^N,z\rangle_q$ for $i=1,2,\cdots,d(N)$. Arguing exactly as in [6, pp. 508–509], we therefore find

$$[A^N(q)]=(Q^N)^{-1}K^N,\tag{2.10}$$

where K^N is the $d(N)\times d(N)$ matrix with elements $K_{ij}^N=\langle\beta_i^N,A(q)\beta_j^N\rangle_q$, and

$$[F^N(q,t,w^N)]=(Q^N)^{-1}R^N F(q,t,u^N).\tag{2.11}$$

We thus arrive at the final form of the approximating system for (2.1) in $X^N(q)$ as

$$
\begin{aligned}
&Q^N \dot{w}^N(t) = K^N w^N(t) + R^N F(q, t, \hat{\beta}^N w^N(t)), \qquad t > 0, \\
&w^N(0) = (Q^N)^{-1} R^N u_0(q).
\end{aligned} \tag{2.12}
$$

3. Approximation theorems for abstract systems. In this section we shall focus our attention on the condition "$q^j \to \bar{q}$ implies $u^j(t; q^j) \to u(t; \bar{q})$" of Theorem 2.1(ii) and present results on the convergence of solutions of the approximating systems (2.7) to solutions of (2.1). We state and prove two theorems; the first is applicable to nonlinear parameter identification problems (§ 4) while the second will be used in connection with linear boundary control problems in § 5.

THEOREM 3.1. *Suppose hypotheses* (H1)–(H3) *and* (H5)–(H9) *hold and let* $q^N, \bar{q}$ *be arbitrary in* Q *such that* $q^N \to \bar{q}$. *Further, suppose that* $|u_0(q^N) - u_0(\bar{q})|_{q^N} \to 0$ *as* $N \to \infty$. *Then the mild solutions* $u^N(t; q^N)$ *of*

$$
\begin{aligned}
&\dot{u}^N(t) = A^N(q^N) u^N(t) + F^N(q^N, t, u^N(t)), \\
&u^N(0) = P^N(q^N) u_0(q^N)
\end{aligned} \tag{3.1}
$$

converge to the mild solution $u(t; \bar{q})$ *of* (2.1) *for each* $t \in [0, T]$. *If* $(t, v) \to F(\bar{q}, t, v)$ *is continuous on* $[0, T] \times X$, *then the convergence* $|u^N(t; q^N) - u(t; \bar{q})|_{q^N} \to 0$ *is uniform in* t *on* $[0, T]$.

Proof. Let $q^N \to \bar{q}$ be arbitrary as hypothesized. Recalling the proof of Theorem 2.1, we observe that one easily argues existence of a constant K such that $|u^N(t; q^N)|_{q^N} \le K$, $|u(t; \bar{q})|_{q^N} \le K$ for all N and $t \in [0, T]$. Further, we see that for $t \in [0, T]$ we have (where all norms are $|\cdot|_{q^N}$)

$$
\begin{aligned}
&|u^N(t; q^N) - u(t; \bar{q})| \\
&\quad \le |T^N(t; q^N) P^N(q^N) u_0(q^N) - T^N(t; q^N) P^N(q^N) u_0(\bar{q})| \\
&\qquad + |T^N(t; q^N) P^N(q^N) u_0(\bar{q}) - T^N(t; q^N) u_0(\bar{q})| \\
&\qquad + |T^N(t; q^N) u_0(\bar{q}) - T(t; \bar{q}) u_0(\bar{q})| \\
&\qquad + \int_0^t |T^N(t-s; q^N) P^N(q^N) \{F(q^N, s, u^N(s; q^N)) - F(q^N, s, u(s; \bar{q}))\}| \, ds \\
&\qquad + \int_0^t |T^N(t-s; q^N) P^N(q^N) \{F(q^N, s, u(s; \bar{q})) - F(\bar{q}, s, u(s; \bar{q}))\}| \, ds \\
&\qquad + \int_0^t |T^N(t-s; q^N) \{P^N(q^N) - I\} F(\bar{q}, s, u(s; \bar{q}))| \, ds \\
&\qquad + \int_0^t |\{T^N(t-s; q^N) - T(t-s; \bar{q})\} F(\bar{q}, s, u(s; \bar{q}))| \, ds \\
&\quad = \varepsilon_1(N) + \varepsilon_2(N) + \varepsilon_3(N) \\
&\qquad + \int_0^t |T^N(t-s; q^N) P^N(q^N) \{F(q^N, s, u^N(s; q^N)) - F(q^N, s, u(s; \bar{q}))\}| \, ds \\
&\qquad + \varepsilon_4(N) + \varepsilon_5(N) + \varepsilon_6(N).
\end{aligned}
$$

By (H8), our hypotheses and the definition of $P^N(q^N)$ in (H5), we find $|\varepsilon_1(N)| \le \mathcal{M} e^{\omega T} |u_0(q^N) - u_0(\bar{q})| \to 0$ as $N \to \infty$. Also, $|\varepsilon_2(N)| \le \mathcal{M} e^{\omega T} |(P^N(q^N) - I) u_0(\bar{q})| \to 0$ by

(H8) and (H7). That $|\varepsilon_3(N)| \to 0$ uniformly in t on $[0, T]$ follows directly from (H9). Moreover,

$$|\varepsilon_4(N)| \le \mathcal{M} e^{\omega T} \int_0^T |F(q^N, s, u(s; \bar{q})) - F(\bar{q}, s, u(s; \bar{q}))| \, ds \to 0$$

by (H6)(iv), (H6)(iii) and dominated convergence, while

$$|\varepsilon_5(N)| \le \mathcal{M} e^{\omega T} \int_0^T |\{P^N(q^N) - I\} F(\bar{q}, s, u(s; \bar{q}))| \, ds \to 0$$

by (H7), (H6)(iii) and dominated convergence. Finally,

$$|\varepsilon_6(N)| = \int_0^t |\{T^N(t-s; q^N) - T(t-s; \bar{q})\} F(\bar{q}, s, u(s; \bar{q}))| \, ds \to 0$$

by (H9) and dominated convergence ((H8) with (H6)(iii)) for each fixed $t \in [0, T]$.

We note that the convergence in all of the terms above, except ε_6, is uniform in t on $[0, T]$. If, in addition, the continuity hypothesis of the theorem obtains, we find that $\{F(\bar{q}, s, u(s; \bar{q})) | s \in [0, T]\}$ is a compact subset of X and the convergence in the integrand of ε_6 is uniform in t and s; hence in this case $\varepsilon_6 \to 0$ uniformly in t also.

We have thus established the following estimate:

$$\begin{aligned} &|u^N(t; q^N) - u(t; \bar{q})| \\ &\le \sum_{i=1}^{6} \varepsilon_i + \mathcal{M} e^{\omega t} \int_0^t |F(q^N, s, u^N(s; q^N)) - F(q^N, s, u(s; \bar{q}))| \, ds \\ &\le \varepsilon^N(t) + \mathcal{M} e^{\omega T} \int_0^t k_1(s) |u^N(s; q^N) - u(s; \bar{q})| \, ds \end{aligned}$$

where $\varepsilon^N \to 0$ as $N \to \infty$. An application of Gronwall's inequality then yields that $|u^N(t; q^N) - u(t; \bar{q})| \to 0$ as $N \to \infty$, where the convergence is uniform in t under the added continuity hypothesis of the theorem.

COROLLARY 3.1. *Under the hypotheses of Theorem* 3.1, $u^N(t; q) \to u(t; q)$ *for each fixed* $q \in Q$, *uniformly in* t *on* $[0, T]$ *if, in addition,* $(t, v) \to F(q, t, v)$ *is continuous on* $[0, T] \times X$.

We turn next to convergence results needed for optimal control problems. Consider for fixed $q \in Q$ the system

$$\begin{aligned} \dot{u}(t) &= A(q)u(t) + G(q, t), \\ u(0) &= u_0, \end{aligned} \tag{3.2}$$

and the approximating system

$$\begin{aligned} \dot{u}^N(t) &= A^N(q)u^N(t) + P^N(q)G(q, t), \\ u^N(0) &= P^N(q)u_0, \end{aligned} \tag{3.3}$$

where G has the form $G(q, t) = \gamma(q)\sigma(t)$. We assume $\gamma(q) \in \Gamma$ where Γ is a subset of $\{\tilde{\gamma} = (\tilde{\gamma}_1, \cdots, \tilde{\gamma}_\mu) | \tilde{\gamma}_i : Q \to \hat{X} \subset X\}$ with $\hat{X}$ a given subset of X. We further assume $\sigma \in \Sigma$, Σ a given subset of $L_2(0, T; R^\mu)$.

THEOREM 3.2. *Assume* (H1)–(H3), (H5), (H7)–(H9). *Suppose moreover that* $\hat{X}$ *is compact and* Σ *is bounded. Then for each fixed* $q \in Q$, *mild solutions* u^N *of* (3.3) *converge to the mild solutions of* (3.2), *uniformly in* $\sigma \in \Sigma$, $\gamma \in \Gamma$ *and* $t \in [0, T]$.

Proof. We consider the estimates in the proof of Theorem 3.1 with $F(q, s, v) = G(q, s)$ *and* $q^N = \bar{q} = q$ fixed. Then

$$|u^N(t; q) - u(t; q)| \le \varepsilon_2(N) + \varepsilon_3(N) + \varepsilon_5(N) + \varepsilon_6(N)$$

($\varepsilon_1 = \varepsilon_4 = 0$), where ε_2 and ε_3 are as before and

$$\varepsilon_5 = \int_0^t |T^N(t-s; q)\{P^N(q) - I\}G(q, s)| \, ds,$$

$$\varepsilon_6 = \int_0^t |\{T^N(t-s; q) - T(t-s; q)\}G(q, s)| \, ds,$$

and, as usual, all X norms are $|\cdot|_q$. We have immediately (using (H7), (H8), (H9)) that ε_2 and $\varepsilon_3 \to 0$ as $N \to \infty$, uniformly in σ, γ and $t \in [0, T]$. Also,

$$|\varepsilon_5(N)| \le \mathcal{M} e^{\omega T} \int_0^T |(P^N(q) - I)\gamma(q)\sigma(s)| \, ds$$

$$\le \mathcal{M} e^{\omega T} \max_{1 \le i \le \mu} |(P^N(q) - I)\gamma_i(q)| \int_0^T |\sigma(s)| \, ds.$$

But since $\hat{\mathcal{X}}$ is compact and Σ is bounded, this latter term $\to 0$ as $N \to \infty$ uniformly in $\gamma \in \Gamma$, $\sigma \in \Sigma$ and $t \in [0, T]$. Finally,

$$|\varepsilon_6(N)| \le \max_{1 \le i \le \mu} \int_0^t |\{T^N(t-s; q) - T(t-s; q)\}\gamma_i(q)| |\sigma(s)| \, ds$$

$$\le \max_{1 \le i \le \mu} \left\{ \int_0^t |\{T^N(t-s; q) - T(t-s; q)\}\gamma_i(q)|^2 \, ds \right\}^{1/2} |\sigma|_{L_2(0,T)},$$

and this last estimate yields $\varepsilon_6(N) \to 0$ uniformly in $\gamma \in \Gamma$, $\sigma \in \Sigma$ and $t \in [0, T]$, again from the compactness of $\hat{\mathcal{X}}$, (H9), and the boundedness of Σ.

As we have previously noted, the main purpose of (H6) is to allow us to guarantee existence of solutions of (2.1) and (2.5) on fixed finite intervals $[0, T]$ (see Proposition 2.1). The condition (H6)(iii) is used in the proofs of Theorem 2.1 and 3.1 only to establish uniform bounds on the u^N. This permits us to employ the local Lipschitz condition (H6)(ii) and to appeal to the dominated convergence theorem in certain arguments. We have already noted that (H6)(iii) can be relaxed to "affine growth at ∞" (see Remark 2.1). With an alternative approach, one can relax this growth condition even further and still obtain the conclusions of Theorems 2.1 and 3.1 (with the other hypotheses remaining unchanged). Specifically, for N sufficiently large, the initial data and defining operators T^N, P^N and T for u^N and u, respectively, are close. Thus if one *assumes* (in place of (H6)(iii))

(A6) (i) For each $q \in Q$, there exists a solution $u(t; q)$ of (2.1) on $[0, T]$, and
(ii) There exists $k_3 \in L_2(0, T)$ such that $|F(q, t, 0)|_{\bar{q}} \le k_3(t)$, for $q, \bar{q} \in Q$ and $t \in [0, T]$,

it is rather tedious but not difficult to show that for N sufficiently large, all u^N defined by (2.7) exist on $[0, T]$ and lie in some bounded neighborhood of u, the solution of (2.1). (The arguments involve use of classical fixed point ideas to obtain solutions on some interval $[0, \delta_1]$ and then continuation to $[\delta_1, 2\delta_1], \cdots$, etc.) The condition (A6)(ii) can then be combined with (H6)(ii) to obtain domination of terms such as $F(q, s, u^N(s; q))$. Thus all of the arguments behind Theorems 2.1 and 3.1 remain valid, the hypotheses being changed only in that (A6)(i), (A6)(ii) replace (H6)(iii), and the conclusions changed only in that they can be obtained only for N sufficiently large (which of course is not important *theoretically* in approximation results such as those discussed in this paper).

Regarding the assumption (A6)(i), we note that there are various conditions one might impose on F to insure existence. For example, monotonicity hypotheses might be assumed so that $-(A+F)$ is maximal monotone and one could then appeal to standard existence results [9], [19]. In § 6 we present numerical results for our approximation scheme for an example (Example 6.5) in which (H6)(iii) is not satisfied, yet (A6)(i) and (A6)(ii) do hold. However, we shall not pursue any of the theoretical ideas here, since this is really not the focus of our presentation.

4. Examples: Parameter identification in hyperbolic and parabolic equations. We turn now to an application of the results developed in the preceding sections to identification problems for specific equations. A fundamental requirement in both Theorems 3.1 and 3.2 is that the conditions of (H7), (H8) and (H9) be verified. As we have indicated earlier, the convergence statement of (H9) can be obtained rather easily for our schemes from (H7), (H8) and some standard approximation results from linear semigroup theory. We state here, for our future reference, one version (due to Kurtz [24]) of these approximation theorems.

PROPOSITION 4.1. *Let* $(\mathcal{Z}, |\cdot|)$ *and* $(\mathcal{Z}^N, |\cdot|_N)$, $N = 1, 2, \cdots$, *be Banach spaces and let* $\pi^N : \mathcal{Z} \to \mathcal{Z}^N$ *be bounded linear operators. Assume further that* $\mathcal{T}(t)$ *and* $\mathcal{T}^N(t)$ *are linear* C_0*-semigroups on* $\mathcal{Z}$ *and* $\mathcal{Z}^N$ *with infinitesimal generators* $\mathcal{A}$ *and* $\mathcal{A}^N$, *respectively. If*

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

(ii) *there exist constants* $\tilde{M}$, $\tilde{\omega}$ *independent of N such that* $|\mathcal{T}^N(t)| \le \tilde{M} e^{\tilde{\omega} t}$, $t \ge 0$,

(iii) *there exists a set* $\mathcal{D} \subset \mathcal{Z}$ *such that* $\mathcal{D} \subset \mathrm{Dom}(\mathcal{A})$, $\bar{\mathcal{D}} = \mathcal{Z}$, *and* $\overline{(\lambda_0 - \mathcal{A})\mathcal{D}} = \mathcal{Z}$ *for some* $\lambda_0 > 0$,

(iv) *for* $z \in \mathcal{D}$ *we have* $\lim_{N\to\infty} |\mathcal{A}^N \pi^N z - \pi^N \mathcal{A} z|_N = 0$,

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

We note that the requirement in (iii) implies that $\mathcal{D}$ is a core [21, p. 166] of $\mathcal{A}$; this is easily seen using the fact that $\mathcal{A}$, being an infinitesimal generator, is closed and $(\lambda I - \mathcal{A})^{-1}$ is bounded for λ sufficiently large. The proposition then follows directly from Theorem 2.1 of [24] taken with subsequent remarks [24, p. 361] of that reference. Obviously, (iii) in our statement above could be replaced by the hypothesis that $\mathcal{D}$ be a core of $\mathcal{A}$. Further, we remark that the requirement $\bar{\mathcal{D}} = \mathcal{Z}$ is superfluous in (iii) if one verifies that $\overline{(\lambda_0 - \mathcal{A})\mathcal{D}} = \mathcal{Z}$ for $\lambda_0 \in \rho(\mathcal{A})$. In this case one can easily demonstrate directly that $\mathcal{D}$ is a core for the generator $\mathcal{A}$.

In the examples we discuss below, we shall use the notation A, F, A^N, P^N, F^N to denote the specific operators in each example, since this will facilitate reference back to the basic theorems of §§ 2 and 3 and should cause no confusion for readers. However, we shall adopt distinct notation for the various state spaces X, X^N within the context of each example.

Example 4.1. *Hyperbolic equations.* We consider the one-dimensional hyperbolic equation

$$v_{tt} = q_1 v_{xx} + q_2 v_t + q_3 v + f(q_4, t, x, v, v_t) \tag{4.1}$$

with initial and boundary conditions

$$(\mathrm{IC}_1)\qquad \begin{aligned} v(0,x)&=\sum_{i=1}^{m} q_4^i\phi_i(x),\\ v_t(0,x)&=\sum_{i=1}^{m} q_5^i\psi_i(x) \quad \text{for } 0\leq x\leq 1,\end{aligned}$$

$$(\mathrm{BC}_1)\qquad v(t,0)=v(t,1)=0 \quad \text{for } t>0,$$

where $v=v(t,x)$, $q_6\in R^m$ and the remaining q_i, q_i^j are scalars. The vector parameter q of § 2 thus has dimension $k=3m+3$. (If the output maps such as Y or C in the fit-to-data functions (2.3) and (2.4) depend explicitly on some parameters $q_7^1,\cdots,q_7^m$, we assume with no loss of generality that these have been embedded in the q_6 (or q_4 or q_5) vector.)

We remark that we do not formulate nontrivial boundary conditions, possibly depending on parameters, in (4.1)–(IC$_1$)–(BC$_1$); however, it is easily seen that by simple transformations such generalities actually can be included in our formulation above. Consider, for example,

$$(4.2)\qquad v_{tt}=q_1v_{xx}$$

with initial and boundary conditions

$$(\mathrm{IC}_2)\qquad \begin{aligned} v(0,x)&=q_4\phi(x),\\ v_t(0,x)&=q_5\psi(x),\end{aligned}$$

$$(\mathrm{BC}_2)\qquad v(t,0)=q_7b_1(t),\qquad v(t,1)=q_8b_2(t),$$

where b_1, b_2 are twice continuously differentiable functions. Employing the standard transformation $w(t,x)=v(t,x)-(1-x)q_7b_1(t)-xq_8b_2(t)$, we find that (4.2)–(IC$_2$)–(BC$_2$) can be reformulated as a special case of (4.1)–(IC$_1$)–(BC$_1$).

Returning to (4.1), we proceed in the usual manner to rewrite the equation with boundary and initial conditions as an abstract evolution equation. Let Δ denote the Laplacian operator $\partial^2/\partial x^2$ in $H^0=L_2(0,1;R)$; here and below the Sobolev spaces H^l consist of R^1-valued functions on $[0,1]$ taken with their usual inner products unless otherwise specified. It is well known that Δ with Dom $(\Delta)=H_0^1\cap H^2$ is a self-adjoint operator in H^0 satisfying $\langle -\Delta z, z\rangle\geq |z|^2$ for all $z\in$ Dom (Δ). We impose the following additional assumption on the coefficient q_1 in (4.1):

(HQ) There exist positive numbers q_1^L and q_1^U such that $q\in Q\subset R^k$ implies $q_1^L\leq q_1\leq q_1^U$.

For a given $q\in Q\subset R^{3m+3}$ we of course mean by q_1 the first coordinate of the vector $q=(q_1,\cdots,q_6)$ where $q_j=(q_j^1,\cdots,q_j^m)$, $j=4,5,6$.

Having thus assumed hypothesis (HQ) for a given fixed parameter restraint set Q, we endow the set H_0^1 with a family of inner products

$$\langle w,z\rangle_{q_1}=\int_0^1 q_1\frac{\partial w}{\partial x}\frac{\partial z}{\partial x}=\langle q_1w_x, z_x\rangle$$

where q ranges over Q. In view of (HQ), the space $(H_0^1,\langle\cdot,\cdot\rangle_{q_1})$ is, for each $q\in Q$, a Hilbert space which we denote by $V(q_1)$. The space $X(q)$ of § 2 is chosen for this example to be $X(q)=\mathscr{X}(q)=V(q_1)\times H^0$ with the usual product topology generated by $\langle(w_1,w_2),(z_1,z_2)\rangle_q=\langle w_1,z_1\rangle_{q_1}+\langle w_2,z_2\rangle$. Condition (H2) is obviously satisfied since for any $\tilde{q}, q \in Q$ we find $|z|_{\tilde{q}} \leq \mathcal{K}|z|_q$ for all $z \in \mathcal{H}$ where $\mathcal{K} = (q_1^U/q_1^L)^{1/2}$. Introducing the variable $w(t) = v_t$, we may rewrite (4.1)–(IC$_1$)–(BC$_1$) in $\mathcal{H}$ by

$$\frac{d}{dt}\begin{pmatrix} v(t) \\ w(t)\end{pmatrix} = A(q)\begin{pmatrix} v(t) \\ w(t)\end{pmatrix} + F(q, t, v(t), w(t)), \qquad t > 0,$$

$$(4.3) \qquad \begin{pmatrix} v(0) \\ w(0)\end{pmatrix} = \begin{pmatrix} \sum q_4^i \phi_i \\ \sum q_5^i \psi_i\end{pmatrix},$$

where $(\phi_i, \psi_i) \in \mathcal{H}$, Dom $(A(q)) = H_0^1 \cap H^2 \times H_0^1$,

$$A(q) = \begin{pmatrix} 0 & 1 \\ q_1\Delta + q_3 & q_2\end{pmatrix},$$

and

$$F(q, t, v(t), w(t)) = \begin{pmatrix} 0 \\ f(q_6, t, \cdot, v(t, \cdot), w(t, \cdot))\end{pmatrix}.$$

Before turning to a careful discussion of (4.3), we define the operators, etc., needed in formulating the *modal approximation scheme* associated with (4.1).

Since the operator $-\Delta$ is self-adjoint with compact resolvent, standard results in spectral theory and the theory of bilinear forms (see [15, p. 1331], [39, p. 343], [33, pp. 247–254]) are applicable. Defining $\hat{\Phi}_j(x) = (\sqrt{2}/j\pi)\sin j\pi x$ and $\Phi_j(x) = \sqrt{2}\sin j\pi x$, we find that $\{\hat{\Phi}_j\}_{j=1}^{\infty}$ and $\{\Phi_j\}_{j=1}^{\infty}$ constitute complete orthonormal sets (CONS) for $V(1)(=H_0^1)$ and H^0, respectively. The corresponding modal subspaces $X^N(q) = \mathcal{H}^N(q)$ of $\mathcal{H}(q)$ are then defined by $\mathcal{H}^N(q) = \operatorname{span}\{\beta_1^N, \cdots, \beta_{2N}^N\}$ where

$$\beta_j^N = \begin{pmatrix} \hat{\Phi}_j \\ 0\end{pmatrix}, \quad j = 1, \cdots, N, \qquad \beta_j^N = \begin{pmatrix} 0 \\ \Phi_{j-N}\end{pmatrix}, \quad j = N+1, \cdots, 2N.$$

It is easily seen that $\bigcup_{N=1}^{\infty}\{\{\beta_j^N\}_{j=1}^{2N}\}$ forms a CONS for $\mathcal{H}(q^*)$ where $q^* = (1, 0, \cdots, 0)$, and a complete orthogonal, but not normal, set for $\mathcal{H}(q)$, for $q = (q_1, \cdots, q_6)$ with $q_1 > 0$, $q_1 \neq 1$. We note also that Φ_j, $\hat{\Phi}_j$ are eigenfunctions of Δ corresponding to the eigenvalues $\lambda = -j^2\pi^2$, $j = 1, 2, \cdots$.

The subspaces $\mathcal{H}^N(q)$ and the corresponding orthogonal projections $P^N(q)$ (see (H5)) having been thus defined, the modal approximation operators $A^N(q)$ for $A(q)$ are determined as in (H5). The corresponding matrices Q^N and K^N of (2.10), which in this case are $2N \times 2N$-matrices, are readily seen to be given by

$$(4.4) \qquad Q^N = \operatorname{diag}(q_1, \cdots, q_1, 1, \cdots, 1),$$

where the q_1 and 1 each appear N times, and

$$(4.5) \qquad K^N = \begin{pmatrix} 0 & D_1^N \\ D_2^N & D_3^N\end{pmatrix},$$

where the D_j^N, $j = 1, 2, 3$, are $N \times N$ diagonal matrices defined by $D_1^N = \operatorname{diag}(\pi q_1, 2\pi q_1, \cdots, N\pi q_1)$, $D_2^N = \operatorname{diag}(q_3/\pi, q_3/2\pi, \cdots, q_3/N\pi) - D_1^N$, and $D_3^N = q_2 I$. Recalling (2.9), we observe that in this case the projection operators $P^N(q)$ *are actually independent of* q.

We are now in a position to verify that (H8) and (H9) obtain for the hyperbolic examples under consideration.

THEOREM 4.1. *Assume that* (HQ) *holds and let* $\{q^N\}$ *be an arbitrary sequence in* $Q \subset R^{3m+3}$ *converging to* $\tilde{q} \in Q$. *Then the operators* $A(\tilde{q})$ *and* $A^N(q^N)$ *of* (4.3) *and* *the associated modal approximations described above generate C_0-semigroups $T(t;\bar q)$ and $T^N(t;q^N)$ on $\mathscr{H}(\bar q)$ and $|T^N(t;q^N)z-T(t;\bar q)z|_{q^N}\to 0$ for $z\in\mathscr{H}(\bar q)$, uniformly in t on compact subsets of $[0,\infty)$. If we further assume* (H4) *(Q is compact), then there exists $\omega\in R$, independent of $q\in Q$ and N, such that $|T(t;q)|\le e^{\omega t}$ and $|T^N(t;q)|\le e^{\omega t}$ for all $q\in Q$, $t\ge 0$, $N=1,2,\cdots$.*

Proof. For any $q\in Q$, a straightforward calculation shows that

$$iA_0(q)=\begin{pmatrix}0 & i\\ iq_1\Delta & 0\end{pmatrix}$$

with $\mathrm{Dom}\,(A_0(q))=H_0^1\cap H^2\times H_0^1$ is a symmetric operator in $\mathscr{H}(q)$. Furthermore, $\mathrm{Dom}\,(A_0(q))$ is dense and A_0 is invertible. It follows [28, p. 97] that A_0 is skew adjoint and hence by Stone's theorem [26, p. 252], [41, p. 345] generates a unitary group on $\mathscr{H}(q)$. Defining the operator $\tilde A(q)$ on $\mathscr{H}(q)$ by $\tilde A(q)(z_1,z_2)=(0,q_3z_1+q_2z_2)$, it is easily seen that $\tilde A(q)$ is bounded. Indeed, using the fact that the H_0^1 norm is stronger than the H^0 norm, one finds $|\tilde A(q)(z_1,z_2)|_q\le c(q_1,q_2,q_3)|(z_1,z_2)|_q$ where the constant c is bounded above uniformly on Q if (HQ) and (H4) hold, say $c(q_1,q_2,q_3)\le\omega$.

We thus find that $A(q)=A_0(q)+\tilde A(q)$ is the perturbation of A_0 by a bounded operator and hence [21, p. 495], [30, p. 80] generates a C_0-semigroup $T(t;q)$ on $\mathscr{H}(q)$ satisfying

$$|T(t;q)|\le\exp\{c(q_1,q_2,q_3)t\},\tag{4.6}$$

where there exist $\omega>0$ independent of q such that $c(q_1,q_2,q_3)\le\omega$ in case (HQ) and (H4) obtain (or in case (HQ) holds and q lies in a bounded subset of Q).

As we have pointed out earlier (see the remarks in § 2), $A^N(q)$ is a bounded linear operator for each N and hence generates a C_0-semigroup on $\mathscr{H}(q)$. Assuming that (HQ) holds and q lies in a bounded subset of Q, we have that $A(q)$ is the infinitesimal generator of a C_0-semigroup satisfying $|T(t;q)|\le e^{\omega t}$ so that $A(q)-\omega I$ is the generator of a contraction semigroup and is hence dissipative [22, p. 90], [30, pp. 14–17]. That is,

$$\langle A(q)z,z\rangle\le\omega\langle z,z\rangle$$

for all $z\in\mathrm{Dom}\,(A(q))$. From the definition of A^N it follows that for $z\in\mathscr{H}(q)$

$$\langle A^N(q)z,z\rangle_q=\langle A(q)P^Nz,P^Nz\rangle_q\le\omega|P^Nz|_q^2\le\omega|z|_q^2,\tag{4.7}$$

since $P^N(q)$ is the orthogonal projection of $\mathscr{H}(q)$ onto $\mathscr{H}^N(q)$. Hence we find $|T^N(t;q)|\le e^{\omega t}$, as desired in the second conclusion of the theorem.

We make use of Proposition 4.1 to establish the convergence results of the theorem. We take for our discussions $\mathscr{X}=\mathscr{H}(\bar q)$ and $\mathscr{X}^N=\mathscr{H}(q^N)$, which of course satisfy the conditions of (H2). The above arguments yield immediately that (ii) of Proposition 4.1 holds for our family of semigroups $T^N(t;q^N)$. Letting $\pi^N=\mathscr{I}^N$ be the canonical isomorphism from $\mathscr{H}(\bar q)$ to $\mathscr{H}(q^N)$, we see immediately from the hypothesis $q^N\to\bar q$ and the definition of the norms in $\mathscr{H}(q)$ that $|\pi^Nz|_{q^N}\to|z|_{\bar q}$, so that (i) is satisfied.

We define $\mathscr{D}=\bigcup_{N=1}^{\infty}\mathscr{H}^N(\bar q)$ and have at once that $\mathscr{D}\subset\mathrm{Dom}\,(A(\bar q))$ and $\bar{\mathscr{D}}=\mathscr{H}(\bar q)$. From the definition of $A(\bar q)$, the fact that $\lambda I-A(\bar q)$ is invertible for λ sufficiently large and that the Φ_l, $\bar\Phi_l$ are eigenfunctions of Δ, it is easily argued that $(\lambda I-A(\bar q))\mathscr{D}=\mathscr{D}$ so that $\overline{(\lambda-A(\bar q))\mathscr{D}}=\mathscr{H}(\bar q)$; hence (iii) is satisfied.

Finally, suppressing the notation $\pi^N = \mathcal{I}^N$ for the canonical isomorphism (see our comments in § 2), we have for each $z \in \mathcal{D}$ and $N = N(z)$ sufficiently large (then $P^N z = z$)

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

where all norms are $|\cdot|_{q^N}$. Since $q^N \to \bar q$ by hypothesis, the form of $A(q)$ yields that the first term $\to 0$ as $N \to \infty$. From the completeness of the $\{\beta_i^N\}$ in $\mathcal{H}(q^*)$ (see our remarks above), the equivalence of norms and thus the strong convergence $P^N \to I$ in any of the norms, we obtain that the second term $\to 0$. This completes the proof of Theorem 4.1.

We return to the abstract nonlinear equation (4.3) to consider conditions on f of (4.1) so that F of (4.3) will satisfy (H6) and hence the results of §§ 2 and 3 will be applicable. Define $Q_6 = \{q_6 \in R^m | q \in Q\}$ where $Q \subset R^{3m+3}$ is given for (4.1). We impose the following hypotheses on f.

(H6*) The nonlinear function f: $Q_6 \times [0, T] \times [0, 1] \times R \times R \to R$ satisfies:

(i) For each $(q_6, v, w) \in Q_6 \times R^2$, the map $(t, x) \to f(q_6, t, x, v, w)$ is measurable.

(ii) For each constant $M > 0$, there exists a function $\tilde k_1 = \tilde k_1(M)$ in $L_2(0, T)$ such that for all $(q_6, t, x) \in Q_6 \times [0, T] \times [0, 1]$ we have

$$|f(q_6, t, x, v_1, w) - f(q_6, t, x, v_2, w)| \le \tilde k_1(t)|v_1 - v_2|$$

for all $(v_i, w) \in R^2$ with $|v_i| \le M$.

(iii) There exists a function $\tilde k_2$ in $L_2(0, T)$ such that

$$|f(q_6, t, x, v, w_1) - f(q_6, t, x, v, w_2)| \le \tilde k_2(t)|w_1 - w_2|$$

for all $(q_6, t, x) \in Q_6 \times [0, T] \times [0, 1]$, and $v, w_1, w_2 \in R$.

(iv) There exists a function $\tilde k_3$ in $L_2(0, T)$ such that

$$|f(q_6, t, x, v, 0)| \le \tilde k_3(t)\{|v| + 1\}$$

for all $(q_6, t, x, v) \in Q_6 \times [0, T] \times [0, 1] \times R$.

(v) For each $(t, x, v, w) \in [0, T] \times [0, 1] \times R^2$, the map $q_6 \to f(q_6, t, x, v, w)$ is continuous.

Employing rather standard arguments and results from analysis (e.g., see [15, Lem. 16(b), p. 196] in connection with (i)), it is quite straightforward to verify under (H2) that (H6*) for f implies (H6) for F in the example under consideration. We can therefore appeal to Proposition 2.1 to guarantee existence of a unique mild solution of (4.3) for any $q \in Q$.

Summarizing, we have shown that under (HQ), (H4) and (H6*) the conditions (H1)–(H9) hold for the abstract formulation (4.3) of (4.1)–(IC$_1$)–(BC$_1$) when considering the modal approximation scheme

$$\begin{aligned}&\frac{d}{dt}\begin{pmatrix} v^N(t)\\ w^N(t)\end{pmatrix} = A^N(q)\begin{pmatrix} v^N(t)\\ w^N(t)\end{pmatrix} + P^N F(q, t, v^N(t), w^N(t)),\\ &\begin{pmatrix} v^N(0)\\ w^N(0)\end{pmatrix} = P^N\begin{pmatrix}\sum q_4^i\phi_i\\ \sum q_5^i\psi_i\end{pmatrix}\end{aligned} \tag{4.8}$$

in $\mathcal{H}^N(q)$. The convergence of Theorems 3.1 and 3.2 is thus assured and we may, when an appropriate fit-to-data function is defined, appeal to Theorem 2.1 to obtain a solution of the associated identification problem for (4.1)–(IC_1)–(BC_1). For example, suppose we are given observations $\hat{y}(t_i) \in R^l$, $i = 1, \cdots, r$, for $(v(t_i, x_1), \cdots, v(t_i, x_l))$ in (4.1) where $0 \leq t_1 < \cdots < t_r \leq T$. Let $u(t; q) = (u_1(t, \cdot; q), u_2(t, \cdot; q))$ denote the unique mild solution of (4.3) where we observe that $u_1(t; q) \in H_0^1$ for each t, q, so that pointwise evaluation in $[0, 1]$ is a meaningful operation. Let $\hat{J}$ have the form given in (2.4) where now $\xi(t_i, q) = \mathrm{col}\,(u_1(t_i, x_1; q), \cdots, u_1(t_i, x_l; q))$ and $C(t_i, q)$ is an $l \times l$-matrix depending continuously on q in Q for each i. Then clearly this $\hat{J}$ satisfies the continuity requirements of Theorem 2.1. Furthermore, it is easily seen that the initial data in (4.3) depend continuously on q. For the modal approximations, recall that $P^N(q)$ is independent of q and finally note that the continuity of $q \to T^N(t; q)z$ follows directly from the forms of $A^N(q)$, Q^N, and K^N given explicitly in (2.10), (4.4) and (4.5). The conclusions of our deliberations for Example 4.1 may thus be stated:

THEOREM 4.2. *If* (HQ), (H4) *and* (H6*) *obtain, then the problem* (ID^N) *for* (4.8) *with $\hat{J}$ as defined above has, for each N, a solution $\bar{q}^N \in Q \subset R^{3m+3}$. Letting $\{\bar{q}^{N_j}\}$ be any subsequence of $\{\bar{q}^N\}$ converging to $\bar{q} \in Q$, then $\bar{q}$ is a solution of the problem* (ID) *for* (4.3) *and moreover for each $t \in [0, T]$,*

$$|(v^{N_j}(t; \bar{q}^{N_j}), w^{N_j}(t; \bar{q}^{N_j})) - (u_1(t; \bar{q}), u_2(t; \bar{q}))| \to 0$$

as $N_j \to \infty$ where (u_1, u_2) is the solution of (4.3) *and the norm is that of $H_0^1 \times H^0$.*

Remark 4.1. We remark that the dependence of the norm on q in the above treatment of hyperbolic systems is somewhat artificial. While one cannot effectively rescale the time variable to remove the q_1-dependence in problems where sampling times (observations) are important, one can rescale the state variables (use $w(t) = 1/\sqrt{q_1}v_t$ in place of the variable used in (4.3)) to avoid use of a weighted norm. We chose not to do that in our computations for Example 4.1. A preliminary consideration leads one to conjecture that such a rescaling does not result in simplification from a numerical viewpoint.

Example 4.2. *Parabolic equations* I. For our second class of examples we discuss scalar parabolic equations

$$v_t = \frac{q_1}{k}(pv_x)_x + q_2 v + f(q_4^1, \cdots, q_4^m, t, x, v) \tag{4.9}$$

for $t > 0$, $x \in [0, 1]$ with initial and boundary conditions

$$v(0, x) = \sum_{i=1}^{m} q_3^i \phi_i(x), \qquad 0 \leq x \leq 1, \tag{IC}$$

$$R_j v(t, \cdot) = 0, \qquad j = 1, 2. \tag{BC}$$

Here we assume that $\phi_i \in H^0$, $v(t, x)$ (or $v(t, x; q)$) is in R, and $q = (q_1, q_2, q_3, q_4) \in Q$ where $Q \subset R^{2m+2}$ and $q_j = (q_j^1, \cdots, q_j^m)$ for $j = 3, 4$. The operators R_1, R_2 defining the boundary conditions have domain H^2 and are given for $\psi \in H^2$ by

$$R_j\psi = \alpha_{j1}\psi(0) + \alpha_{j2}\psi'(0) + \alpha_{j3}\psi(1) + \alpha_{j4}\psi'(1), \tag{4.10}$$

where $\alpha_{ji} \in R$. We impose the following conditions on k and p in (4.9) and the α_{ji}:

(HP1) The functions p, p_x and k are continuous with $k(x) > 0$, $p(x) > 0$ for $0 \leq x \leq 1$.

(HP2) The matrix

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \end{pmatrix}$$

has rank 2 and we have $p(0)\{\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}\} = p(1)\{\alpha_{13}\alpha_{24} - \alpha_{14}\alpha_{23}\}$.

We shall also assume that our parameter set Q satisfies hypothesis (HQ) already given above.

To apply the results of §§ 2 and 3 we again rewrite our system as an abstract evolution equation. Although we consider here only a scalar equation, the results we present generalize immediately to coupled systems of parabolic equations. Indeed, we present numerical results in § 6 for examples of such systems which are of interest in biological applications (see [4]), among others.

We define the general Sturm–Liouville operator $A(q)$ (our operator here is the negative of the one found in the literature cited below) in H^0 by $\mathrm{Dom}\,(A(q)) = \{\psi \in H^2 | R_j\psi = 0,\ j = 1, 2\}$ and $A(q)\psi = k^{-1}(q_1 p\psi_x)_x + q_2\psi$. Then (4.9)–(IC)–(BC) can be rewritten as

$$
(4.11)\qquad
\begin{aligned}
\dot{v}(t) &= A(q)v(t) + F(q, t, v(t)), \qquad t > 0,\\
v(0) &= \sum_{i=1}^{m} q_3^i \phi_i,
\end{aligned}
$$

where $F(q, t, v(t)) = f(q_4, t, \cdot, v(t, \cdot))$ and the equation is taken in the state space $X(q) = \mathcal{H} = (H^0, \langle \cdot, \cdot \rangle)$ with $\langle \phi, \psi \rangle = \int_0^1 \phi(x)\psi(x)k(x)\,dx$. We note that, unlike in Example 4.1, here the state space actually doesn't depend on q.

Spectral results for the operator $A(q)$ are readily found in the literature—e.g., see [28, p. 182], [20, p. 126]. The hypotheses (HP1), (HP2) imply that $A(q)$ is self-adjoint and its spectrum consists of a countable number of real eigenvalues $\{\lambda_j(q)\}_{j=1}^{\infty}$, each of multiplicity less than or equal to 2, and, moreover, these eigenvalues can be ordered so that $-\infty < \cdots \leq \lambda_j \leq \lambda_{j-1} \leq \cdots \leq \lambda_1 < \infty$. The eigenfunctions $\{\Psi_j\}_{j=1}^{\infty}$ corresponding to $A(q^*)$ where $q^* = (1, 0, \cdots, 0)$ form a CONS in $\mathcal{H}$.

We further observe that the eigenvalues $\{\lambda_j(q)\}$ of $A(q)$ are bounded above uniformly in q on bounded subsets of Q. This is easily seen as follows: Let $\tilde{\lambda}_j$ be the eigenvalues of $A_0 = A(q^*)$ (i.e., $A_0\psi = k^{-1}(p\psi_x)_x$) with corresponding eigenfunctions Ψ_j. From our remarks above we have $\tilde{\lambda}_j \leq \tilde{\omega}$, $j = 1, 2, \cdots$, for some positive finite constant $\tilde{\omega}$. The eigenvalues for $A(q)$ are then found to be $\lambda_j(q) = q_1\tilde{\lambda}_j + q_2$ (with eigenfunction Ψ_j) so that we find $\lambda_j(q) \leq \omega$ on bounded subsets of Q, where ω is independent of q (but depends, of course, on the particular bounded subset of Q involved).

We define the approximating modal subspaces of $X(q) = \mathcal{H}$ by $\mathcal{H}^N = \mathrm{span}\,\{\Psi_1, \cdots, \Psi_N\}$ and let $P^N : \mathcal{H} \to \mathcal{H}^N$ denote the associated canonical orthogonal projections. As before, we determine the operator $A^N(q)$ and F^N by $A^N(q) = P^N A(q) P^N$ and $F^N = P^N F$. We have the following convergence results.

THEOREM 4.3. *Suppose that* (HQ), (HP1) *and* (HP2) *hold and* $q^N, \bar{q} \in Q \subset R^{2m+2}$ *are such that* $q^N \to \bar{q}$ *as* $N \to \infty$. *Then* $A(\bar{q})$ *and* $A^N(q^N)$ *generate* C_0-*semigroups* $T(t; \bar{q})$ *and* $T^N(t; q^N)$ *on* $\mathcal{H}$ *and* $|T^N(t; q^N)z - T(t; \bar{q})z| \to 0$ *for* $z \in \mathcal{H}$ *with the convergence uniform in* t *on compact subsets of* $[0, \infty)$. *Furthermore, if* (H4) *obtains, then there exists a constant* ω *independent of* N *and* q *such that* $|T(t; q)| \leq e^{\omega t}$ *and* $|T^N(t; q)| \leq e^{\omega t}$ *for* $t > 0$, $q \in \mathring{Q}$, *and* $N = 1, 2, \cdots$.

Proof. Let $\mathring{Q}$ be any bounded subset of Q. Then our remarks above imply existence of $\tilde{\omega} = \omega(\mathring{Q})$ such that the self-adjoint operator $A(q)$, $q \in \mathring{Q}$ has its spectrum lying in $(-\infty, \tilde{\omega})$. Hence (see [36, p. 349]) $A(q) - \tilde{\omega}I$ is dissipative, i.e., $\langle (A(q) - \tilde{\omega}I)z, z \rangle \leq 0$ for all $z \in \mathrm{Dom}\,(A(q))$ and $q \in \mathring{Q}$. For $\lambda \notin \sigma(A(q))$, we have [28, p. 180] that $A(q) - \lambda I$ has compact resolvent so that in particular we have $(A(q) - \lambda I)\,\mathrm{Dom}\,(A(q)) = \mathcal{H}$ for $\lambda > 0$ properly chosen. It follows immediately [30, p. 17], [2, p. 175], [22, p. 87] that $A(q) - \tilde{\omega}I$ is maximal dissipative and generates a C_0-semigroup of contractions on $\hat{\mathscr{X}}$. Arguing as in (4.7), we have that $A^N(q)-\tilde{\omega}I$ is dissipative, uniform in $q\in\tilde{Q}$ and $N=1,2,\cdots$ (it is maximal, being defined on all of $\hat{\mathscr{X}}$—see [22, p. 86]), and hence is also the generator of a C_0-semigroup of contractions.

The above remarks obviously apply if we choose $\tilde{Q}=\{q^N\}\cup\{\bar{q}\}$ where $q^N\to\bar{q}$ or if $\tilde{Q}=Q$ where Q itself is compact, i.e., (H4) holds. To obtain the convergence results of the theorem, take $\tilde{Q}=\{q^N\}\cup\{\bar{q}\}$ and use Proposition 4.1. Here $\mathscr{B}=\mathscr{B}^N=\hat{\mathscr{X}}$ and conditions (i) and (ii) of the proposition clearly are satisfied. Let $\mathscr{D}=\bigcup_{N=1}^{\infty}\hat{\mathscr{X}}^N$ so that $\mathscr{D}\subset\operatorname{Dom}(A(\bar{q}))$ and $\mathscr{D}$ is dense in $\hat{\mathscr{X}}$. From our remarks above, we have $A(\bar{q})\Psi_j=\lambda_j(\bar{q})\Psi_j$. For $\lambda>\max_j\lambda_j(\bar{q})$ we have $\lambda I-A(\bar{q})$ invertible where $(\lambda I-A(\bar{q}))\Psi_j=(\lambda-\lambda_j(\bar{q}))\Psi_j$, and it follows that $(\lambda I-A(\bar{q}))\mathscr{D}=\mathscr{D}$. Hence (iii) of Proposition 4.1 is satisfied. The arguments that (iv) is satisfied are exactly analogous to those used to complete the proof of Theorem 4.1, here the completeness of the $\{\Psi_j\}$ yielding $P^N\to I$ strongly in $\hat{\mathscr{X}}$. We thus have established the results of Theorem 4.3.

We turn finally to conditions on f in (4.9) that will insure that F of (4.11) satisfies (H6). Let $Q_4=\{q_4\in R^m|q\in Q\}$ where Q is a given subset of R^{2m+2}.

(H6**) The nonlinear function f: $Q_4\times[0,T]\times[0,1]\times R\to R$ satisfies:
- (i) For each $(q_4,v)\in Q_4\times R$, the map $(t,x)\to f(q_4,t,x,v)$ is measurable.
- (ii) There exists a function $\tilde{k}_1$ in $L_2(0,T)$ such that $|f(q_4,t,x,v_1)-f(q_4,t,x,v_2)|\le\tilde{k}_1(t)|v_1-v_2|$ for all $(q_4,t,x,v_i)\in Q_4\times[0,T]\times[0,1]\times R$.
- (iii) There exists a function $\tilde{k}_2$ in $L_2([0,T]\times[0,1])$ such that $|f(q_4,t,x,0)|\le\tilde{k}_2(t,x)$ for all $q_4\in Q_4$.
- (iv) For each (t,x,v) in $[0,T]\times[0,1]\times R$, the map $q_4\to f(q_4,t,x,v)$ is continuous on Q_4.

It is an easy exercise to verify that (H6**) for f implies (H6) for F (note that in this example the condition (H2) is trivially satisfied). We thus may invoke Theorem 3.1 for convergence of our modal approximations defined in $\hat{\mathscr{X}}^N$ by the equation

$$\begin{aligned}\dot{v}^N(t)&=A^N(q)v^N(t)+P^NF(q,t,v^N(t)),\\ v^N(0)&=P^N\sum_{i=1}^m q_3^i\phi_i.\end{aligned}\tag{4.12}$$

For these parabolic systems defined in H^0, the question of an appropriate cost functional is somewhat more delicate, since in general, point evaluation may not be meaningful. One possibility (a different approach will be discussed below) is to choose a cost functional J as in (2.3) where now $u=v$ is the mild solution of (4.11) and we might assume, for example, that $(x,q)\to Y(t_i,x,q)$ is continuous for each t_i. We are again in a position to employ Theorem 2.1 (taking into account the comments in Remark 2.2) to establish the following results for this example.

THEOREM 4.4. *Suppose* (HQ), (H4), (H6**) *and* (HP1), (HP2) *are satisfied. Then the problem* (ID^N) *for* (4.12) *with J as in* (2.3) *has a solution $\bar{q}^N\in Q\subset R^{2m+2}$ for each $N=1,2,\cdots$. If $\{\bar{q}^{N_j}\}$ is any subsequence of $\{\bar{q}^N\}$ converging to $\bar{q}\in Q$, then $\bar{q}$ is a solution of* (ID) *for* (4.11) *and for each $t\in[0,T]$ we have $|v^{N_j}(t;\bar{q}^{N_j})-v(t;\bar{q})|\to 0$, as $N_j\to\infty$, where v, v^{N_j} are the mild solutions of* (4.11), (4.12) *respectively and the norm is that of H^0.*

Example 4.3. *Parabolic equations* II. We consider again the parabolic equation (4.9) with initial condition (IC) and boundary condition (BC) but with slightly more restrictive conditions on the boundary operators R than those given in (4.10)–(HP2). We treat problems with boundary operators chosen from the standard Sturm-

Liouville operators (see [14, p. 145], [13, p. 291])

$$\text{(4.13)}\quad \begin{array}{lll} \text{(A)} & R_1\psi=\psi(0) \quad \text{or} \quad R_2\psi=\psi(1); \\ \text{(B)} & R_1\psi=\psi'(0) \quad \text{or} \quad R_2\psi=\psi'(1); \\ \text{(C)} & R_1\psi=\psi'(0)-\sigma_1\psi(0) \quad \text{or} \quad R_2\psi=\psi'(1)+\sigma_2\psi(1), \quad \sigma_1,\sigma_2>0; \\ \text{(D)} & R_1\psi=\psi(0)-\psi(1) \quad \text{and} \quad R_2\psi=p(0)\psi'(0)-p(1)\psi'(1). \end{array}$$

In the sequel, in referring to the standard Sturm–Liouville boundary conditions (SLBC), we shall mean any combination of choices for R_1 and R_2 from (4.13) (A)–(C) or the periodic boundary conditions arising from the choice of R_1, R_2 given in (4.13) (D).

Our main objective is to discuss the use of the simple pointwise fit-to-data criteria

$$J_1(q, v(\cdot\,;q), \hat{y}) = \sum_{i=1}^{r} |\hat{y}(t_i) - C(t_i, q)\xi(t_i, q)|^2 \tag{4.14}$$

as defined in § 2 (see (2.4) and the discussions thereof). When treating (4.9) in $\mathscr{X}$ as we did in Example 4.2, it is by no means clear that the associated ID problem with (4.14) is well posed. Indeed, one must first justify the pointwise evaluation (in the spatial coordinate) of v involved in defining ξ; assuming that this can be done, one must entertain a second difficulty in that the convergence (of Theorems 3.1 and 4.3) of $v^N(t;q^N)$ to $v(t;\bar{q})$ is in the $\mathscr{X}$ (i.e. H^0) norm. Since $J_1(\cdot,\cdot,\hat{y})$ is *not* continuous on $Q\times C(0,T;\mathscr{X})$, the results of Theorem 2.1 are not directly applicable.

We turn first to the difficulty raised by point evaluations in (4.14). In this regard we note that the mapping $w\to J_1(q,w,\hat{y})$ from $C(0,T;\mathscr{X}^N)$ to R, where J_1 is given in (4.14) and $\mathscr{X}^N$ are the modal subspaces defined in Example 4.2, is well defined for each $N=1,2,\cdots$, and, in particular, is well defined on solutions of (4.12). Hence the approximating problems (ID^N) associated with J_1 are well posed in any event. Justification of point evaluation for (ID) with J_1 depends heavily on the smoothing properties of (4.9) or, equivalently, (4.11). Roughly speaking, for $t>0$ the solution values $v(t;q)$ will be contained in Dom $(A(q))$ if only $t\to f(t)=F(q,t,v(t))$ is smooth enough [8, p. 192]. However, since we wish to avoid additional smoothness hypotheses, we choose a slightly more technical route to the same end.

Recalling the arguments for Theorem 4.3, we have that $A(q)-\omega I$ is self-adjoint and maximal dissipative where ω can be chosen independent of $q\in Q$. It therefore follows [9, p. 47] that $-A(q)+\omega I=\partial\phi^*(q)$ where $\partial\phi^*(q)$ denotes the subdifferential of the functional $\phi^*(q)$ is given by

$$\phi^*(q)(u)=\begin{cases}\frac{1}{2}|(\omega I-A(q))^{1/2}u|^2 & \text{if } u\in \text{Dom}\,(\omega I-A(q))^{1/2},\\ +\infty & \text{otherwise.}\end{cases} \tag{4.15}$$

Here $(\omega I-A(q))^{1/2}$ denotes the square root, which by standard results [21, p. 281] is known to exist. Under assumption (H6**) for f, we have that (H6) holds for $F(q,t,w)=f(q,t,\cdot,w)$ and in particular (see (H6)(iii)) the mapping $t\to F(q,t,v(t;q))$ is in $L_2(0,T;\mathscr{X})$ for $t\to v(t;q)$ the solution of (4.9). Thus (see [9, p. 72] and note that a weak solution in the sense of [9] is in fact the unique mild solution in our sense, which is, moreover, also a strong solution) it follows that the map $t\to\phi^*(q)(v(t;q))$ is in $L_1(0,T;\mathscr{X})$ and is absolutely continuous on all subintervals $[\delta,T]$, $\delta>0$, of $(0,T]$. In light of (4.15) this implies $v(t;q)\in\text{Dom}\,(\omega I-A(q))^{1/2}$ for all $t>0$ and we observe that this holds for arbitrary initial data in $\mathscr{X}$. It remains to note that Dom $(\omega I-A(q))^{1/2}\subset H^1$. To see this one only needs make standard arguments using elementary results on interpolating spaces—see [27, p. 9]. Briefly, by defining X (in the notation of [27]) as either H^1 with appropriate boundary conditions (in the event of B.C.'s using combinations from (A) and (B) or B.C.(D) from (4.13)) or H^1 with an appropriate energy norm (in the event elastic boundary conditions from (C) of (4.13) are involved) and $Y = \mathscr{X}$, one can make the identification $\Lambda = (\omega I - A(q))^{1/2}$ with Dom $(\Lambda) = X$.

We can therefore summarize by stating that conditions (HQ), (H4), (H6**), (HP1) in the case of (SLBC) are sufficient to allow point evaluation in J_1.

If we wish to relax the continuity hypothesis on J in Theorem 2.1, it is necessary to consider in more detail the fit-to-data criterion $J^N(q) = J(q, u^N(\cdot\,; q), \hat{y})$ of (ID^N). The following proposition will be useful in our deliberations; we state and prove it using the notation of Theorem 2.1.

PROPOSITION 4.2. *We suppose there exist maps $\mathscr{J}$ and $\mathscr{J}^N$, $N = 1, 2, \cdots$, from the compact set Q to R satisfying*:

(i) *for each $N = 1, 2, \cdots$, the map $q \to \mathscr{J}^N(q)$ is continuous on Q;*

(ii) *for any $q \in Q$ and any sequence $\{N_k\}$ with $N_k \to \infty$, there exists a subsequence N_{k_i} such that $\mathscr{J}^{N_{k_i}}(q) \to \mathscr{J}(q)$;*

(iii) *for any $q^N \to \bar{q}$, there exists a subsequence $\{q^{N_k}\}$ such that $\mathscr{J}^{N_k}(q^{N_k}) \to \mathscr{J}(\bar{q})$.*

Then for each N there exists $q^N \in Q$ that minimizes $\mathscr{J}^N$ over Q and, moreover, for any convergent subsequence $\{q^{N_k}\}$ of $\{q^N\}$ with $q^{N_k} \to \bar{q}$, $\mathscr{J}$ is a minimum over Q at $\bar{q}$.

Proof. Let q^N denote the minimizer (whose existence is guaranteed by (i) and the compactness of Q) of $\mathscr{J}^N$ so that $\mathscr{J}^N(q^N) \leq \mathscr{J}^N(q)$ for all $q \in Q$. Suppose $q^{N_k} \to \bar{q}$; then by (iii) (reindexing for convenience in notation) we have $\mathscr{J}^m(q^m) \to \mathscr{J}(\bar{q})$ for some subsequence $\{q^m\}$ of $\{q^N\}$. We use this to argue that $\mathscr{J}(\bar{q}) \leq \mathscr{J}(q)$ for $q \in Q$. If we assume that there exist $\hat{q} \in Q$ and $\bar{j}$ such that $\mathscr{J}^{m_j}(q^{m_j}) > \mathscr{J}(\hat{q})$ for $j \geq \bar{j}$, then by (ii) there is yet another subsequence of m_j, denoted by m_{j_i} such that $\mathscr{J}^{m_{j_i}}(\hat{q}) \to \mathscr{J}(\hat{q})$. Hence, for sufficiently large i we have $\mathscr{J}^{m_{j_i}}(q^{m_{j_i}}) > \mathscr{J}^{m_{j_i}}(\hat{q})$, which contradicts the definition of $q^{m_{j_i}}$ as a minimizer for $\mathscr{J}^{m_{j_i}}$.

To use Proposition 4.2 with our particular $J^N = \mathscr{J}^N$ defined using J_1, we must consider special cases for which hypotheses (i)–(iii) of that proposition are readily verified. We discuss the homogeneous ($f = 0$) version of (4.9) in $\mathscr{X}$ in this regard.

PROPOSITION 4.3. *Consider* (4.9) *with $f = 0$, initial conditions* (IC) *and boundary conditions* (SLBC) *from* (4.13). *Assume that* (HQ), (H4) *and* (HP1) *obtain. Then for the solutions $v^N(\cdot\,; q)$ of the approximating equations* ((4.12) *with $F = 0$) we have $\{v^N(q) | q \in Q\}$ is a relatively compact subset of $C(t^*, T; C(0, 1; R))$ for each $t^* \in (0, T]$.*

Since the proof of this result consists of checking compactness criteria for a specific subset of $C(t^*, T; Y)$, where Y is a Banach space, we shall only sketch the ideas involved, leaving the details to the interested and determined reader.

First recall that equation (4.12) is in $\mathscr{X}^N = \text{span}\{\Psi_1, \cdots, \Psi_N\}$ where $\{\Psi_i\}_{i=1}^{\infty}$ is the CONS of eigenfunctions of $A(q^*)$ with corresponding eigenvalues $\bar{\lambda}_i = \lambda_i(q^*)$ (see the discussion immediately preceding Theorem 4.3). Solutions v^N of (4.12) have the representation (see (2.8)–(2.12) and the associated discussions) $v^N(t; q) = \sum_{i=1}^N w_i^N(t)\Psi_i$ where $w_i^N(t) = w_i^N(0) \exp\{(\bar{\lambda}_i q_1 + q_2)t\}$. To verify relative compactness of the desired set, one can use the Ascoli theorem [25, p. 211] which requires equicontinuity of the set along with relative compactness of $\{v^N(t; q) | q \in Q, N = 1, 2, \cdots,\}$ in $C(0, 1; R)$ for each t in $[t^*, T]$. Use of the representation results along with standard estimates in Fourier analysis (Parseval, Cauchy–Schwarz, etc.) reduces the compactness criteria to the task of verifying

$$\text{(a)} \qquad \sum_{i=1}^{\infty} e^{\bar{\lambda}_i q_1^+ t^*} < \infty,$$

(b) $$\sum_{i=1}^{\infty} e^{2\lambda_i q_1^L t^*} |\Psi_i'|_{H^0} < \infty,$$

(c) $$\sup\{|\Psi_i(x)| \mid 0 \le x \le 1,\ i = 1, 2, \cdots\} \text{ is finite.}$$

Here q_1^L is the lower bound for q_1 (see (HQ)).

From standard Fourier results [15, p. 1332] we have $\langle Af, \Psi_i\rangle \to 0$ for $f \in \mathrm{Dom}\,(A(q^*))$ or that $\{A(q^*)\Psi_i\}$ is bounded in $\mathcal{H}$. Hence $\{\Psi_i\}$ is bounded in the graph $(A(q^*))$ norm and, consequently, after some calculations taking into account the (SLBC), in H^1. It follows that (c) obtains and that $\{\Psi_i'\}$ is H^0 bounded. Thus if (a) holds, we immediately obtain (b). We have, therefore, reduced the compactness criteria to a requirement on the rate at which $\lambda_i \to -\infty$. But asymptotic estimates for the eigenvalues of Sturm–Liouville problems are readily available. For example, for (SLBC) from combinations of (A), (B), (C) of (4.13) we have $\lambda_i \sim -i^2$ [14, p. 153], which yields (a). For the periodic boundary conditions (D) a slight modification of the arguments in [12, § 8.3] yield the same rate estimates. In some higher-dimensional problems, similar rates for the eigenvalues are also available [13, § VI.4.1].

With the compactness results of Proposition 4.3 we are now able to give well-posedness results for the special case under consideration.

PROPOSITION 4.4. *Suppose* (HQ), (H4), (HP1) *obtain and* $f = 0$ *in* (4.9)–(IC)–(SLBC). *Then* (ID^N) *with* J_1 *of* (4.14) *has a solution* $\bar q^N \in Q$ *for each* $N = 1, 2, \cdots$. *Moreover, there exists a subsequence* $\{\bar q^{N_k}\}$ *such that* $\bar q^{N_k} \to \bar q$ *with* $\bar q$ *a solution of* (ID) *with* J_1 *and* $|v^{N_k}(t;\bar q^{N_k}) - v(t;\bar q)|_C \to 0$ *as* $k \to \infty$, *uniformly on compact subsets of* $(0, T]$.

Proof. We first verify that (i)–(iii) of Proposition 4.2 hold with $\mathcal{J}^N(q) = J_1(q, v^N(\cdot\,;q), \hat y)$ and $\mathcal{J}(q) = J_1(q, v(\cdot\,;q), \hat y)$. Since the matrices K^N and Q^N of (2.10) are given by $Q^N = I$ and $K^N_{ij} = \langle \Psi_i, A(q)\Psi_j\rangle_{\mathcal{H}} = (q_1\lambda_i + q_2)\delta_{ij}$, continuity of the map

$$q \to \sum_{i=1} |C(t_i, q)\,\mathrm{col}\,(v^N(t_i, x_1; q), \cdots, v^N(t_i, x_l; q)) - \hat y(t_i)|^2$$

can be demonstrated by elementary arguments; this implies (i) and hence the existence of a minimum $\bar q^N$ for $\mathcal{J}^N$.

By compactness of Q there exists a convergent subsequence $\bar q^{N_k}$ with $\bar q^{N_k} \to \bar q \in Q$. Employing Proposition 4.3 with $t^* < t_1$, t^* otherwise arbitrarily chosen in $(0, T)$, we find that for a subsequence of $\{\bar q^{N_k}\}$, again denoted by $\{\bar q^{N_k}\}$, the sequence $\{v^{N_k}(\cdot\,;\bar q^{N_k})\}$ is Cauchy in $C(t^*, T; C(0, 1; R))$. Since we already have $|v^{N_k}(t;\bar q^{N_k}) - v(t;\bar q)|_{\mathcal{H}} \to 0$, uniformly in t on $[0, T]$, we obtain $|v^{N_k}(t;\bar q^{N_k}) - v(t;\bar q)|_C \to 0$, uniformly in t on $[t^*, T]$. It follows that (ii) and (iii) of Proposition 4.2 hold and thus the conclusions of the present proposition are established.

The special case considered above involved homogeneous ($f = 0$) equations. One can also obtain results in the case $f \neq 0$ by formulating the (ID) problem in a "smoother" space than H^0. We illustrate these ideas with an example involving a special case of (4.9). Consider

$$\text{(4.16)} \qquad v_t = q_1 v_{xx} + q_2 v + f(q_3, \cdots, q_\nu, t, \cdot, v)$$

with boundary conditions

$$\text{(DBC)} \qquad v(t, 0) = v(t, 1) = 0$$

and initial condition (IC) with $\phi_1 \in H_0^1$. We again rewrite (4.16)–(DBC)–(IC) as an abstract equation, but this time in the Hilbert space $\mathcal{H} = (H_0^1, \langle\cdot,\cdot\rangle_1)$ with inner product $\langle\phi, \psi\rangle_1 = \int_0^1 \phi'\psi'\,dx$ (e.g., see [16, pp. 94, 105]) and define the operator $A(q)$ by $\mathrm{Dom}\,(A(q)) = \{\phi \mid \phi \in H_0^1, \phi'' \in H_0^1\}$ and $A(q)\phi = q_1\phi'' + q_2\phi$. Assuming (HQ) and (H4), it is easy to verify existence of an ω independent of $q \in Q$ such that $A(q)-\omega I$ is dissipative and symmetric in $\hat{\mathcal{X}}$, that Dom $(A(q))$ is dense and $R(A(q)-\lambda I)=\hat{\mathcal{X}}$ for some λ, chosen sufficiently large and independent of $q \in Q$. In particular, $A(q)$ is therefore self-adjoint [28, p. 97]. The approximating subspaces $\hat{\mathcal{X}}^N$ are defined by $\hat{\mathcal{X}}^N = \text{span}\{\hat{\Phi}_1, \cdots, \hat{\Phi}_N\}$ where $\hat{\Phi}_j(x)=(\sqrt{2}/j\pi)\sin j\pi x$ as in Example 4.1. We recall that $\{\hat{\Phi}_j\}_1^\infty$ is a CONS in $\hat{\mathcal{X}}$. As before we let $P^N:\hat{\mathcal{X}} \to \hat{\mathcal{X}}^N$ denote the canonical orthogonal projections and $A^N(q)=P^NA(q)P^N$, $F^N(q,t,v)=P^NF(q,t,v)$ where $F: Q\times[0,T]\times\hat{\mathcal{X}} \to \hat{\mathcal{X}}$ is given by $F(q,t,v)=f(q,t,\cdot,v)$. Formulation in $\hat{\mathcal{X}}$ will require additional assumptions on f, to be detailed in (H6***) below. We then have:

THEOREM 4.5. *Suppose that* (HQ) *holds and let* $q^N, \bar{q} \in Q$ *be such that* $q^N \to \bar{q}$ *as* $N\to\infty$. *Then* $A(\bar{q})$ *and* $A^N(q^N)$ *generate* C_0*-semigroups* $T(t;\bar{q})$ *and* $T^N(t;q^N)$ *on* $\hat{\mathcal{X}}$ *and* $|T^N(t;q^N)z-T(t;\bar{q})z|_{\hat{\mathcal{X}}}\to 0$ *for each* $z\in\hat{\mathcal{X}}$, *with the limit uniform in* t *on compact subsets of* $[0,\infty)$. *Furthermore, if* Q *is compact, then there exists a constant* ω *independent of* N *and* q, *such that* $|T(t;q)|\le e^{\omega t}$ *and* $|T^N(t;q)|\le e^{\omega t}$ *for* $t>0$, $q\in Q$, *and* $N=1,2,\cdots$.

The proof of this theorem is quite similar to that of Theorem 4.3 and will therefore not be given here.

We shall simply list conditions (compare with (H6**)) on f that will insure that F satisfies (H6).

(H6***) The nonlinear function $f: Q_4\times[0,T]\times[0,1]\times R\to R$ satisfies:

(0) For each $v\in\hat{\mathcal{X}}$, the map $x\to f(q,t,x,v(x))$ is in $\hat{\mathcal{X}}$.

(i) For each $(q_4,v)\in Q_4\times R$, the maps $(t,x)\to f(q_4,t,x,v)$, $(t,x)\to f_x(q_4,t,x,v)$, and $(t,x)\to f_v(q_4,t,x,v)$ are measurable.

(ii) There exists a function $\hat{k}_1$ in $L_2(0,T)$ such that $|f_v(q_4,t,x,v)|\le\hat{k}_1(t)$ for all $(q_4,x,v)\in Q_4\times[0,1]\times R$; for each $M>0$ there exists $\hat{k}_1^M$ in $L_2(0,1)$ such that $|f_x(q_4,t,x,v_1)-f_x(q_4,t,x,v_2)|\le\hat{k}_1^M(t)|v_1-v_2|$ and $|f_v(q_4,t,x,v_1)-f_v(q_4,t,x,v_2)|\le\hat{k}_1^M(t)|v_1-v_2|$ for all $(q_4,x)\in Q_4\times[0,1]$ and v_1, v_2 with $|v_1|\le M$, $|v_2|\le M$.

(iii) There exists $\hat{k}_2$ in $L_2(0,T)$ such that $|f(q_4,t,x,0)|\le\hat{k}_2(t)$ and $|f_x(q_4,t,x,v)|\le\hat{k}_2(t)\{1+|v|\}$ for all $(q_4,x,v)\in Q_4\times[0,1]\times R$.

(iv) For each (t,x,v) in $[0,T]\times[0,1]\times R$, the maps $q_4\to f(q_4,t,x,v)$, $q_4\to f_x(q_4,t,x,v)$ and $q_4\to f_v(q_4,t,x,v)$ are continuous on Q_4.

The fit-to-data criterion (4.14) together with the state equation in $\hat{\mathcal{X}}$ are such that the map $(q,v)\to J_1(q,v,\hat{y})$ from $Q\times C(0,T;\hat{\mathcal{X}})\to R$ is continuous and, therefore, Theorems 3.1 and 2.1 are readily applicable. We leave a precise statement of the theorem for (4.16)–(IC)–(DBC) with (H6***), analogous to Thoorem 4.4, to the reader.

In concluding this discussion, we remark that numerical implementation of a scheme formulated as above in $\hat{\mathcal{X}}$ (the projections in defining the approximations in (2.8)–(2.12) are now in the H_0^1 inner product) is, of course, somewhat more tedious from a technical viewpoint than that for schemes such as those in Example 4.3 where the state space is H^0.

Example 4.4. *Diffusion-convection equations.* For the final example of this section, we return to the setting of Example 4.2 and indicate how, in (4.9), one might include convection (or advection) terms that are independent of the Sturm–Liouville operator $(pu_x)_x$. To illustrate the ideas, we, for ease in exposition only, take a simple linear example (nonlinearities of the type discussed previously present no essential difficulties) involving only diffusion and convection terms. Consider then

$$v_t = q_1 v_{xx} + q_2 v_x \tag{4.17}$$

with initial conditions (IC) and the Dirichlet boundary conditions

(DBC) $$v(t,0) = v(t,1) = 0.$$

As state space we choose H^0 with its usual inner product. We define the operator $A(q)$ with $\mathrm{Dom}\,(A(q)) = H^2 \cap H_0^1$ by $A(q)\psi = q_1\psi'' + q_2\psi'$. We then have $A(q)$ is dissipative since (again we assume (HQ) holds) for $z \in \mathrm{Dom}\,(A(q))$,

$$\begin{aligned}\langle A(q)z, z\rangle &= q_1\langle z_{xx}, z\rangle + q_2\langle z_x, z\rangle \\ &= -q_1|z_x|^2 + q_2\int_0^1 (z^2/2)_x\,dx \\ &= -q_1|z_x|^2 \leqq 0.\end{aligned}$$

Thus if we assume (H4), we find that $A(q)$ is dissipative uniformly in $q \in Q$.

In considering modal approximations, the question of existence of a complete set of eigenfunctions for the operator $A(q)$ arises naturally. Standard spectral results for nonself-adjoint operators allow one to answer this question in the affirmative. First, $A(q)$ is a relatively bounded perturbation of a discrete spectral operator and is itself a discrete spectral operator (see [15, Thm. XIX.4.16, p. 2347]—in this case the boundary conditions (DBC) are easily seen to satisfy the necessary regularity hypotheses—see [15, p. 2341–2344]). It follows that $\sigma(A(q))$ consists of point spectrum and that the eigenprojections $\{E_{\lambda_j}\}$ (see [15, p. 2292]) of the resolution of identity for $A(q)$ satisfy $\sum_{j=1}^N E_{\lambda_j} z \to z$ for $z \in H^0$ (see [15, Cor. XVIII.2.33, p. 2257], along with the properties of the projection operators—e.g., [15, Lem. XVIII.2.31, p. 2255]). One can easily argue for our example that the generalized eigenmanifolds are one-dimensional so that the eigenfunctions $\tilde{\Psi}_j(q) = \exp(-q_2x/2q_1)\sin j\pi x$ corresponding to the eigenvalues $\lambda_j(q) = -j^2\pi^2 q_1 - q_2^2/2q_1$ form a complete (but not orthogonal) set in H^0.

We thus also have that $\lambda \in \rho(A(q))$ if $\lambda > 0$ so that $(A(q) - I)\,\mathrm{Dom}\,(A(q)) = H^0$ for $\lambda > 0$ and hence $A(q)$ is maximal dissipative [22, p. 87], [30, p. 17]. The operator $A(q)$ generates a C_0-semigroup $T(t;q)$ satisfying $|T(t;q)| \leqq e^{\omega t}$ for $q \in Q$.

For a modal approximation scheme, it might be tempting at first thought to use the finite-dimensional subspaces $\tilde{X}^N(q) = \mathrm{span}\,\{\tilde{\Psi}_1(q), \cdots, \tilde{\Psi}_N(q)\}$, but of course this would prove rather difficult computationally in identification problems. Here we choose to use the basis elements $\Phi_j(x) = \sqrt{2}\sin j\pi x$ since we know $\{\Phi_j\}_1^\infty$ forms a CONS in H^0 and $\Phi_j \in \mathrm{Dom}\,(A(q))$. We thus define $H^N = \mathrm{span}\,\{\Phi_1, \cdots, \Phi_N\}$ and remind the reader that "modal" is something of a misnomer for this scheme (actually, we took a similar approach in Example 4.1 in choosing basis elements corresponding to $q^* = (1, 0, \cdots, 0)$ fixed).

As usual, we define $A^N(q) = P^N A(q) P^N$ where P^N are the orthogonal projectors $P^N z = \sum_{j=1}^N \langle z, \Phi_j\rangle\Phi_j$ onto H^N which converge strongly to the identity on H^0.

To develop approximation results similar to those given in Theorems 4.3 and 4.4, the essential effort remaining in our example is to verify the stability and consistency hypotheses (ii), (iii), (iv) of Proposition 4.1 with $\mathscr{A} = A(\bar{q})$ and $\mathscr{A}^N = A^N(q^N)$ where $q^N \to \bar{q}$ in Q. Stability (i.e., (ii)) is immediate while consistency is slightly more delicate. A natural choice (the one we have used in previous examples) for the set $\mathscr{D}$ is $\bigcup_{N=1}^\infty H^N$ since then (iv) is trivial to verify. However, it is not apparent to us that this choice of $\mathscr{D}$ satisfies (iii) of Proposition 4.1. We choose instead $\mathscr{D} = \bigcup_{N=1}^\infty \tilde{X}^N(\bar{q})$, where $\tilde{X}^N(\bar{q})$ is defined above in terms of the true modes $\tilde{\Psi}_j(\bar{q})$ for

$A(\bar{q})$. Since $(\lambda I - A(\bar{q}))\bar{\Psi}_j(\bar{q}) = (\lambda - \lambda_j)\bar{\Psi}_j(\bar{q})$ and the set $\{\bar{\Psi}_j(\bar{q})\}$ is complete, (iii) is easily established and it only remains to show $A^N(q^N)z \to A(\bar{q})z$ for $z \in \mathcal{D}$.

We first note that from

$$|A^N(q^N)z - A(\bar{q})z| \le |P^N(A(q^N)P^Nz - A(\bar{q})z)| + |P^NA(\bar{q})z - A(\bar{q})z|$$

and the strong convergence of P^N to I, it is sufficient to argue $A(q^N)P^Nz \to A(\bar{q})z$ for $z \in \mathcal{D}$. It suffices to argue this latter convergence for $z = \psi_k \equiv \bar{\Psi}_k(\bar{q})$ fixed. For this choice of z we find

$$\begin{aligned} A^N(q^N)P^N\psi_k &= A(q^N)\sum_{j=1}^N \langle\psi_k, \Phi_j\rangle\Phi_j = \sum_{j=1}^N \langle\psi_k, \Phi_j\rangle A(q^N)\Phi_j \\ &= \sum_{j=1}^N \langle\psi_k, \Phi_j\rangle\{q_1^N\Phi_j'' + q_2^N\Phi_j'\} \\ &= \sum_{j=1}^N \langle\psi_k, \Phi_j\rangle\{q_1^N(-j^2\pi^2)\Phi_j + j\pi q_2^N\chi_j\} \\ &= q_1^N\sum_{j=1}^N \langle\psi_k, -j^2\pi^2\Phi_j\rangle\Phi_j + q_2^N\sum_{j=1}^N \langle\psi_k, j\pi\Phi_j\rangle\chi_j \\ &= q_1^N\sum\langle\psi_k, \Phi_j''\rangle\Phi_j + q_2^N\sum\langle\psi_k, -\chi_j'\rangle\chi_j, \end{aligned}$$

where $\chi_j(x) = \sqrt{2}\cos j\pi x$ and we have used the facts that $\chi_j' = -j\pi\Phi_j$ and $\Phi_j'' = -j^2\pi^2\Phi_j$.

Integration by parts twice (using the fact that ψ_k and Φ_j are in H_0^1) yields

$$\langle\psi_k, \Phi_j''\rangle = \langle\psi_k'', \Phi_j\rangle$$

while a single integration by parts establishes (again use $\psi_k \in H_0^1$)

$$\langle\psi_k, -\chi_j'\rangle = \langle\psi_k', \chi_j\rangle.$$

We thus have

$$A^N(q^N)P^N\psi_k = q_1^N\sum_{j=1}^N \langle\psi_k'', \Phi_j\rangle\Phi_j + q_2^N\sum_{j=1}^N \langle\psi_k', \chi_j\rangle\chi_j,$$

where both $\{\Phi_j\}$ and $\{\chi_j\}$ constitute CONS in H^0. Since $q_1^N \to \bar{q}_1$, $q_2^N \to \bar{q}_2$ we thus obtain $A^N(q^N)P^N\psi_k \to \bar{q}_1\psi_k'' + \bar{q}_2\psi_k' = A(\bar{q})\psi_k$, as was desired.

The theorems for these approximation ideas for the (ID) and (ID^N) problems with (4.17)–(IC)–(DBC) are so similar in statement to Theorems 4.3 and 4.4 that we shall not prolong our discussion by giving a precise statement here.

With regard to implementation of this scheme, we point out that $A(q)$ does not leave the subspaces H^N invariant and hence the matrix representation of $A^N = P^NAP^N$ (see (2.8)–(2.10)) is not a simple diagonal matrix. However, for equations such as (4.17), it is rather easily seen that (2.10) is given by

$$[A^N(q)]_{ij} = \begin{cases} -q_1 i^2\pi^2 & \text{for } i = j, \\ 0 & \text{for } i \ne j \text{ and } i+j \text{ even}, \\ 2jq_2\left[\dfrac{2i}{i^2-j^2}\right] & \text{for } i \ne j \text{ and } i+j \text{ odd}. \end{cases}$$

While this is not a simple matrix, it does allow a rather straightforward implementation of the scheme in actual computations.

5. A boundary control problem. The theory developed in §§ 2 and 3 can also be applied to optimal control problems governed by partial differential equations. We shall demonstrate this by means of a specific example. Consider as a special case of (4.1)–(IC$_1$)–(BC$_1$) the problem

$$\tilde{v}_{tt} = \tilde{v}_{xx} \tag{5.1}$$

for $t>0$, $x\in[0,1]$ with initial and boundary conditions

$$\tilde{v}(0,x)=\phi(x), \qquad \tilde{v}_t(0,x)=\psi(x), \tag{IC$_3$}$$

$$\tilde{v}(t,0)=s_1(t), \qquad \tilde{v}(t,1)=s_2(t), \tag{BC$_3$}$$

where the boundary control functions s_i are chosen in $\mathcal{S}=\{s \mid s\in H^2(0,T;R),\ s(0)=s'(0)=0\}$ and $(\phi,\psi)\in H_0^1\times H^0$. The transformation $v(t,x)=\tilde{v}(t,x)-(1-x)s_1(t)-xs_2(t)$ applied to (5.1)–(IC$_3$)–(BC$_3$) leads to

$$v_{tt}=v_{xx}-(1-x)(s_1)_{tt}-x(s_2)_{tt}, \tag{5.2}$$

$$v(0,x)=\phi(x), \qquad v_t(0,x)=\psi(x), \tag{IC$_4$}$$

$$v(t,0)=v(t,1)=0. \tag{BC$_4$}$$

We let $w=v_t$ and reformulate (5.2)–(IC$_4$)–(BC$_4$) as in Example 4.1 in the Hilbert space $\mathcal{X}=H_0^1\times H^0$ with the usual inner product. This leads to a special case of (4.3) given by

$$\frac{d}{dt}\begin{pmatrix} v(t)\\ w(t)\end{pmatrix}=A\begin{pmatrix} v(t)\\ w(t)\end{pmatrix}+\gamma\sigma(t), \qquad \begin{pmatrix} v(0)\\ w(0)\end{pmatrix}=\begin{pmatrix}\phi\\ \psi\end{pmatrix}, \tag{5.3}$$

where

$$A=\begin{pmatrix} 0 & I\\ \Delta & 0\end{pmatrix}, \quad \gamma=\gamma(x)=\begin{pmatrix} 0 & 0\\ -x & x-1\end{pmatrix}, \quad \sigma(t)=\operatorname{col}((s_2)_{tt},(s_1)_{tt}), \quad (\phi,\psi)\in\mathcal{X}.$$

The finite-dimensional subspaces $\mathcal{X}^N=\mathcal{X}^N(q^*)$, $q^*=(1,0,\cdots,0)$, are chosen as in Example 4.1 and again we take $A^N=P^NAP^N$, where $P^N:\mathcal{X}\to\mathcal{X}^N$ denote the canonical orthogonal projections. For the convenience of the reader we repeat the family of approximating equations given by

$$\frac{d}{dt}\begin{pmatrix} v^N(t)\\ w^N(t)\end{pmatrix}=A^N\begin{pmatrix} v^N(t)\\ w^N(t)\end{pmatrix}+P^N\gamma\sigma(t), \qquad \begin{pmatrix} v^N(0)\\ w^N(0)\end{pmatrix}=P^N\begin{pmatrix}\phi\\ \psi\end{pmatrix}. \tag{5.4}$$

In the light of Theorems 3.2 and 4.1 (with $q^N=q^*$ for all N), the solutions $(v^N(t;\sigma), w^N(t;\sigma))$ and $(v(t;\sigma), w(t;\sigma))$ of (5.4) and (5.3), respectively, satisfy $\lim_N (v^N(t;\sigma), w^N(t;\sigma))=(v(t;\sigma), w(t;\sigma))$ in $\mathcal{X}$ uniformly in $t\in[0,T]$, for any $T>0$ and uniformly in σ, as σ varies in bounded subsets Σ of $L_2(0,T;R^2)$. We shall also need the following technical result.

LEMMA 5.1. *The operator $\mathcal{F}: L_2(0,T;R^2)\to C(0,T;\mathcal{X})$ defined via $(\mathcal{F}\sigma)(t)=\int_0^t T(t-\tau)\gamma\sigma(\tau)\,d\tau$ is compact.*

Defining the maps $(\mathcal{F}^N\sigma)(t)=\int_0^t T^N(t-\tau)P^N\gamma\sigma(\tau)\,d\tau$ and using the convergence of the semigroups $T^N(t)$ to $T(t)$, generated by A^N and A, respectively, it is easily seen that $\mathcal{F}^N\to\mathcal{F}$ in the operator norm topology. The proof is completed once one argues that the maps $\mathcal{F}^N$ themselves are compact.

The above remarks provide the technical tools that can be applied to a variety of optimal control problems, one of which will be outlined below. For a more complete discussion concerning approximation of optimal control problems for infinite-dimensional systems by sequences of optimization problems for finite-dimensional systems, we refer to [7] and the references given there.

We let Σ_{ad} be any fixed closed convex subset of $L_2(0, T; R^2)$ (possibly L_2 itself) and choose nonnegative continuous functions $g_0: \mathcal{H} \to R$, g_1: $C(0, T; \mathcal{H}) \to R$ and $g_2: L_2(0, T; R^2) \to R$. The functions g_i define the cost functional $\hat{J}: \Sigma_{ad} \to R$ by

$$(5.5) \qquad \hat{J}(\sigma) = g_0((v(T;\sigma), w(T;\sigma))) + g_1((v(\cdot;\sigma), w(\cdot;\sigma))) + g_2(\sigma).$$

The optimal boundary value control problem associated with (5.1)–(IC$_3$)–(BC$_3$), (5.5) is then taken to be:

$$(\mathcal{P}) \qquad \text{minimize } \hat{J} \text{ over } \Sigma_{ad}.$$

Suppose that a solution $\hat{\sigma} = \mathrm{col}\,(\hat{\sigma}_1, \hat{\sigma}_2) \in \Sigma_{ad}$ of $(\mathcal{P})$ is found; this will uniquely determine boundary controls $\hat{s}_1$ and $\hat{s}_2$ in $\mathcal{S}$. The approximate optimization problems are defined by

$$(\mathcal{P}^N) \qquad \text{minimize } \hat{J}^N \text{ over } \Sigma_{ad},$$

where

$$(5.6) \qquad \hat{J}^N(\sigma) = g_0((v^N(T;\sigma), w^N(T;\sigma))) + g_1((v^N(\cdot;\sigma), w^N(\cdot;\sigma))) + g_2(\sigma).$$

Notice that (5.6) is an optimization problem associated with an ordinary differential equation. We shall need two standard assumptions on the functions g_i:

(G1) The continuous functions g_i are convex,
(G2) $g_2(\sigma) \to \infty$ as $|\sigma| \to \infty$.

As a consequence of (G1), the maps $\sigma \to \hat{J}(\sigma)$ and $\sigma \to \hat{J}^N(\sigma)$ are convex, which together with (G2) implies the existence of solutions of $(\mathcal{P})$ and $(\mathcal{P}^N)$; these solutions are in addition unique if one of the g_i is strictly convex. Let σ^N denote a solution of $(\mathcal{P}^N)$. Then by (G2) it follows that $\{\sigma^N\}$ must be a bounded subset of $L_2(0, T; R^2)$. Indeed, the assumption $|\sigma^{N_k}| \to \infty$ for some subsequence $\{N_k\}$ contradicts the inequalities $g_2(\sigma^{N_k}) \leq \hat{J}^{N_k}(\sigma^{N_k}) \leq \hat{J}^{N_k}(\sigma) \to \hat{J}(\sigma) < \infty$ for all $\sigma \in L_2(0, T; R^2)$. The convergence of $\hat{J}^{N_k}(\sigma) \to \hat{J}(\sigma)$ is a consequence of $(v^N(t;\sigma), w^N(t;\sigma)) \to (v(t;\sigma), w(t;\sigma))$ uniformly in $t \in [0, T]$ and (G1). Since Σ_{ad} is convex and closed it is weakly closed so that there exists a weakly convergent subsequence $\{\sigma^{N_k}\}$ of $\{\sigma^N\}$ with σ^{N_k} converging weakly to some $\hat{\sigma} \in \Sigma_{ad}$. By Theorem 3.2, Lemma 5.1 and the estimates

$$\begin{aligned} &|(v^{N_k}(t;\sigma^{N_k}), w^{N_k}(t;\sigma^{N_k})) - (v(t;\hat{\sigma}), w(t;\hat{\sigma}))| \\ &\quad \leq |(v^{N_k}(t;\sigma^{N_k}), w^{N_k}(t;\sigma^{N_k})) - (v(t;\sigma^{N_k}), w(t;\sigma^{N_k}))| \\ &\qquad + |(v(t;\sigma^{N_k}), w(t;\sigma^{N_k})) - (v(t;\hat{\sigma}), w(t;\hat{\sigma}))|, \end{aligned}$$

it follows that

$$(v^{N_k}(t;\sigma^{N_k}), w^{N_k}(t;\sigma^{N_k})) \to (v(t;\hat{\sigma}), w(t;\hat{\sigma}))$$

in $\mathcal{H}$ uniformly in $t \in [0, T]$. Since convexity and continuity together imply weak lower semicontinuity, we obtain the following string of inequalities:

$$\begin{aligned} \hat{J}(\hat{\sigma}) &\leq \liminf \{g_0((v^{N_k}(T;\sigma^{N_k}), w^{N_k}(T;\sigma^{N_k}))) \\ &\qquad + g_1((v^{N_k}(\cdot;\sigma^{N_k}), w^{N_k}(\cdot;\sigma^{N_k}))) + g_2(\sigma^{N_k})\} \\ &= \liminf \hat{J}^{N_k}(\sigma^{N_k}) \leq \limsup \hat{J}^{N_k}(\sigma) = \hat{J}(\sigma) \end{aligned}$$

for every $\sigma \in \Sigma_{ad}$. This implies that $\hat{\sigma}$ is a solution of $(\mathcal{P})$. Further standard arguments can be used to show that strict convexity of $\hat{J}$ implies that σ^N itself converges weakly to the unique solution $\bar{\sigma}$ of $(\mathcal{P})$ and that σ^N converges strongly in $L_2(0, T; R^2)$ to $\bar{\sigma}$ if $\hat{J}$ is strongly convex (see [7]). We finally summarize some of the above discussion in

THEOREM 5.1. *Suppose that* (G1) *and* (G2) *hold. If* $\{\sigma^N\}$ *denotes a sequence of solutions of* $(\mathcal{P}^N)$, *then there exists a subsequence* $\{\sigma^{N_k}\}$ *converging weakly to a solution* $\hat{\sigma}$ *of* $(\mathcal{P})$. *Furthermore,* $\hat{J}^{N_k}(\sigma^{N_k}) \to \hat{J}(\hat{\sigma})$ *and* $(v^{N_k}(t; \sigma^{N_k}), w^{N_k}(t; \sigma^{N_k})) \to (v(t; \hat{\sigma}), w(t; \hat{\sigma}))$ *uniformly in* $t \in [0, T]$. *Moreover,* $\hat{\sigma}$ *determines uniquely boundary controls* $\hat{s}_1$, $\hat{s}_2$ *in* $\mathcal{S}$.

6. Numerical examples. In this section we briefly summarize our numerical findings when applying the modal approximation algorithms to some of the identification problems that were outlined in § 4. The aim here is to demonstrate the feasibility of the method for both hyperbolic and parabolic systems. As it turns out, modal approximations appear to be very well suited for hyperbolic systems, while for certain identification problems for parabolic systems we encountered some essential difficulties which one should take into consideration before attempting any practical use of the method for this type of equation. This will be explained further below. In developing our software packages, no great attention was given to maximizing efficiency in implementing the algorithms, or to minimizing computer time. The ordinary differential equations (see (2.12)) that arise were integrated by a simple fourth-order Runge–Kutta method (with step size varying from one example to the next from .0125 to .05), and the coefficients of the nonlinearity and the initial data (see (2.9) and (2.11)) were computed by employing Simpson's rule. The minimization problem arising in the identification problem for the approximating ordinary differential equations was numerically solved by using an IMSL package (ZXSSQ) employing the Levenberg–Marquardt algorithm. The "exact" solutions, which were used for the "data" $\hat{y}$ in the fit-to-data criterion J, were generated by a Crank–Nicolson algorithm whenever solutions in closed form were not available. These solutions were generated with fixed known values of the parameters in the equations; these values will be referred to in the sequel as the "true" parameter values.

In the examples below, a fit-to-data criterion of the type (2.4) with $C(t_i, q) = I$ was used throughout. Further, we usually (except in Example 6.5) let $T = 2$ and chose t_i and x_j equally spaced in $[0, 2]$ and $[0, 1]$, respectively, so that $|t_i - t_{i-1}| = 0.2$ and $|x_j - x_{j-1}| = 0.25$.

Example 6.1. Here we return to Example 4.1 and consider the linear one-dimensional hyperbolic equation, which we repeat for convenience:

$$\begin{aligned}
&v_{tt} = q_1 v_{xx} + q_2 v_t + q_3 v && \text{for } t > 0,\\
&v(0, x) = q_4 x(1 - x) && \text{for } 0 \leq x \leq 1,\\
&v_t(0, x) = q_5 \hat{\psi}(x) && \text{for } 0 \leq x \leq 1,\\
&v(t, 0) = v(t, 1) = 0, && \text{for } t > 0,
\end{aligned}$$

where $\hat{\psi}(x) = 2x$ for $x \in [0, .5]$ and $\hat{\psi}(x) = 2(1 - x)$ for $x \in [.5, 1]$. Below, we present numerical results which are typical of those obtained in making numerous runs with this example. The startup values $q^{N,0}$ for use in the Levenberg–Marquardt algorithm are recorded in the bottom row of the tables, whereas the next-to-last row contains the true parameter values. The tables contain only those parameters on which a search was performed, whereas the remaining parameters were assumed known and therefore were held fixed at the true values.

In a first run (Table 1) we assumed that $q_3 = 1$ and $q_5 = 0$ were known and we were searching for $q = (q_1, q_2, q_4)$ with the true parameter values chosen to be $\hat{q} = (1.414, -1, 2)$ and with startup values $q^{N,0} = (1, 0, 1)$.

TABLE 1

N	q_1^N	q_2^N	q_4^N
4	1.4103	−0.9961	1.9978
8	1.4126	−1.0021	2.0031
16	1.4129	−0.9968	2.0005
32	1.4129	−0.9992	2.0000
true value	1.414	−1	2
$q^{N,0}$	1	0	1

A feature of interest for these models used with the Levenberg–Marquardt algorithm is the range of convergence for the parameter q. For this specific example, we carried out computations keeping two of the parameters q_1, q_2, q_4 in addition to q_3, q_5 fixed while identifying one of q_1, q_2, q_4. It was observed that for $q_1^{N,0}$, $q_2^{N,0}$, $q_4^{N,0}$ taken in the ranges $1 \leq q_1^{N,0} \leq 5$, $-5 \leq q_2^{N,0} \leq 0$, $.5 \leq q_4^{N,0} \leq 5$, respectively, rapid convergence was still obtained. (The actual range of convergence may be much larger; these were merely the ranges of values we tested.)

In a second run (Table 2), $q_1 = 1.414$ and $q_5 = 1$ were assumed to be known and the search was performed on $q = (q_2, q_3, q_4)$ with true value $\hat{q} = (-5, 4, 2)$ and startup values $q^{N,0} = (0, 0, 1)$.

TABLE 2

N	q_2^N	q_3^N	q_4^N
4	−4.9930	4.0242	1.9997
8	−4.9803	4.0372	2.0025
true value	−5	4	2
$q^{N,0}$	0	0	1

Example 6.2. This is the nonlinear example (again a special case of Example 4.1):

$$
\begin{aligned}
&v_{tt} = q_1 v_{xx} + v + q_6(1+v)^{-1} && \text{for } t > 0,\\
&v(0,x) = q_4 x(1-x) && \text{for } 0 \leq x \leq 1,\\
&v_t(0,x) = 0 && \text{for } 0 \leq x \leq 1,\\
&v(t,0) = v(t,1) = 0 && \text{for } t > 0.
\end{aligned}
$$

We chose the true model parameters $\hat{q} = (1.414, 2, 1)$, whereas the startup values were taken to be $q^{N,0} = (1, 1, 0)$. For the numerical solutions we refer to Table 3.

Example 6.3. This is another nonlinear equation of the form

$$
\begin{aligned}
&v_{tt} = q_1 v_{xx} + q_3 v + q_6 v^3 && \text{for } t > 0,\\
&v(0,x) = q_4 x(1-x) && \text{for } 0 \leq x \leq 1,\\
&v_t(0,x) = 0 && \text{for } 0 \leq x \leq 1,\\
&v(t,0) = v(t,1) = 0 && \text{for } t > 0.
\end{aligned}
$$

TABLE 3

N	q_1^N	q_4^N	q_6^N
4	1.4141	1.9990	0.9735
8	1.4148	2.0013	0.9790
16	1.4152	2.0009	0.9788
true value	1.414	2	1
$q^{N,0}$	1	1	0

Although this nonlinearity does not satisfy (H6*) of Example 4.1, we report in Table 4 on calculations carried out in the subspaces $\mathscr{H}^N$ of $\mathscr{H}$. It is clear that the algorithm is converging in this case; indeed, one can relax the assumptions (H6*) so as to prove convergence for such nonlinearities; see the discussion involving (A6)(i), (ii) in § 3.

TABLE 4

N	q_1^N	q_2^N	q_4^N	q_6^N
4	1.3835	0.6774	1.9999	1.2368
8	1.4107	0.9875	2.0001	0.8973
16	1.4138	0.9983	2.0001	1.0016
true value	1.414	1	2	1
$q^{N,0}$	1	0	1	0

We turn now to some special cases of the parabolic problem (4.9)–(IC)–(BC). As pointed out earlier, parabolic equations can be more formidable than hyperbolic ones to handle via modal approximations. The difficulties are more than just a simple lack of identifiability (however this concept is defined), which, of course, can lead to substantial numerical embarrassment. Indeed, parabolic equations can lead to stiff systems of approximating ordinary differential equations. The reader can quickly convince himself of this fact by taking Dirichlet boundary conditions and putting $p=k=1$ and $f=q_2=0$. In our computational pursuits we did not make an effort to use specific numerical methods for the stiff systems that can arise, but we simply decreased the step size in the Runge–Kutta algorithm to effect numerical stability. A perhaps more reasonable approach to avoiding these difficulties due to modal approximations is to take a completely different approximation scheme, say for example spline-based methods. We have pursued this idea successfully for parabolic systems and the details of those investigations will be reported elsewhere.

The fit-to-data criterion is chosen to be (4.14) with $C(t_i, q)=I$ in all the scalar examples below. In the two-dimensional system of Example 6.7, we used the obvious analogue of (4.14) for a coupled system of equations.

Example 6.4. We consider the linear equation

$$
\begin{aligned}
&v_t = q_1 v_{xx} + q_2 v && \text{for } t>0,\\
&v(0,x)=\hat{\phi}(x) && \text{for } 0\le x\le 1,\\
&v(t,0)=v(t,1)=0 && \text{for } t>0,
\end{aligned}
$$

where $\hat{\phi}$ is the "hat"-function defined in Example 6.1. The modal approximation scheme failed to identify q_1 and q_2 simultaneously, although it did identify each of them individually so long as the other one was fixed. This is by no means surprising; the exact solution of the above problem has the explicit representation $v(t,x)=\sum_{j=1}^{\infty} v_j(t)\sin j\pi x$, where $v_j(t)=v_j(0)\exp((q_2-q_1(j\pi)^2)t)$ and $v_j(0)$, $j=1,2,\cdots$, are the Fourier coefficients of the sine series for $\hat{\phi}$. At time $t=0$, the values of q_1, q_2 have no influence and at the following times $t=.2, .4, \cdots$ the decay of the exponential term in addition to the decreasing magnitude of the coefficients $v_j(0)$ cause successive terms to contribute less to the criterion J. Moreover, in this example, $v_{2j}(0)=0$ for $j=1, 2, \cdots$, so that the criterion uses essentially only one mode to fit the model to the data. The results from the search on both parameters simultaneously are presented in Table 5.

TABLE 5

N	$\bar{q}_1^N$	$\bar{q}_2^N$
4	0.0236	0.2313
8	0.0335	0.3289
16	0.0336	0.3296
true value	0.1	0.986
$q^{N,0}$	0.25	0.25

Keeping $q_2=.2$ fixed and searching for q_1, when the true parameter value is $\hat{q}_1=.1$ and $q_1^{N,0}=0.25$, we find $\bar{q}_1^4=0.09999$. Similarly, when $q_1=.1$ is kept fixed and q_2 is to be identified, with $\hat{q}_2=.986$ and $q_2^{N,0}=.25$, the algorithm yields $\bar{q}_2^4=.986004$.

Example 6.5. We next consider the nonlinear parabolic equation

$$\begin{aligned} &v_t=q_1v_{xx}-q_4v^3 && \text{for } t>0,\\ &v(0,x)=q_3\hat{\psi} && \text{for } 0\leq x\leq 1,\\ &v(t,0)=v(t,1)=0 && \text{for } t>0. \end{aligned}$$

It is well known that for $q_4\geq 0$ the above system has a global solution and we are therefore again in a situation where hypotheses (A6)(i), (ii) of § 3 must be used in any theoretical considerations of convergence. Our findings for this example are given in Table 6. Here we choose $T=1$, while keeping the increments between the "data" points the same as before ($\Delta t=.2$ and $\Delta x=.25$).

TABLE 6

N	$\bar{q}_1^N$	$\bar{q}_3^N$	$\bar{q}_4^N$
2	.5030	4.8271	.8539
4	.4976	5.3001	1.2370
8	.4985	5.1774	1.1482
16	.5021	5.0845	1.0443
true value	.5	5	1
$q^{N,0}$	.25	1	0

Example 6.6. We consider

$$\begin{aligned} &v_t=q_1v_{xx}+2q_4(1+v)^{-1} && \text{for } t>0,\\ &v(0,x)=q_3\hat{\psi} && \text{for } 0\leq x\leq 1,\\ &v(t,0)=v(t,1)=0 && \text{for } t>0. \end{aligned}$$

In this and the next example we solved the approximating identification problem both without and with noise. When noise was added, then the Crank–Nicolson data which were used in the fit-to-data criterion were perturbed by Gaussian noise with zero mean and variance $\sigma^2=.01$. It is accurate to report that in these two examples the scheme behaves in a stable manner under the influence of noise. In Tables 7 and 8 below, a blank indicates that this parameter was kept fixed at the true parameter value. The estimates obtained for this example are recorded in Table 7.

TABLE 7

	N	$\bar{q}_1^N$	$\bar{q}_3^N$	$\bar{q}_4^N$
no noise; search on q_3, q_4	4		5.2275	1.9254
	8		5.1374	1.9741
	16		5.0668	1.9845
noise $\sigma^2 = .01$; search on q_3, q_4	4		5.2362	1.9335
	8		5.1459	1.9813
	16		5.0749	1.9917
no noise; search on q_1, q_3, q_4	4	0.2472	5.2846	2.5221
	8	0.2301	5.1706	2.3584
	16	0.2150	5.0823	2.1746
noise $\sigma^2 = .01$; search on q_1, q_3, q_4	4	0.2443	5.2903	2.4941
	8	0.2272	5.1760	2.3301
	16	0.2120	5.0873	2.1442
	true value	0.2	5	2
	$q^{N,0}$	0.1	1	0

Example 6.7. As a final example we consider the coupled parabolic system

$$\begin{aligned}
&v_t = q_1 v_{xx} + 2(1 + q_4 w + v)^{-1}, && \\
&w_t = q_2 w_{xx} && \text{for } t > 0, \\
&v(0, x) = \hat{\psi}(x) && \text{for } 0 \leq x \leq 1, \\
&w(0, x) = \hat{\phi}(x) && \text{for } 0 \leq x \leq 1, \\
&v(t, 0) = v(t, 1) = w(t, 0) = w(t, 1) = 0 && \text{for } t > 0,
\end{aligned}$$

for which the numerical results are given in Table 8.

TABLE 8

	N	$\bar{q}_1^N$	$\bar{q}_2^N$	$\bar{q}_4^N$
no noise; search on q_1, q_4	4	.2011		1.9933
	8	.1987		2.0226
noise $\sigma^2 = .01$; search on q_1, q_4	4	.1982		2.1105
	8	.1960		2.1246
no noise; search on q_2, q_4	4		.0500	2.0514
	8		.0498	1.9551
noise $\sigma^2 = .01$; search on q_2, q_4	4		.0522	2.0349
	8		.0520	1.9385
no noise; search on q_1, q_2, q_4	4	.2011	.0500	1.9931
	8	.1988	.0499	2.0187
noise $\sigma^2 = .01$; search on q_1, q_2, q_4	4	.1973	.0522	2.1776
	8	.1949	.0521	2.2066
	true value	.2	.05	2
	$q^{N,0}$	.1	.1	0

7. Concluding remarks. The contributions of the discussions in this paper are twofold. First, we have developed a general approximation framework in the context of semigroups that allows treatment of identification and control problems for a wide class of distributed parameter systems. Our second contribution is a development, using this framework, of "modal" approximation schemes in the spirit of those often proposed in the engineering literature. In addition to providing a solid theoretical foundation for such schemes, we have systematically tested them numerically on a number of examples and reported some of our findings. One result of these investigations has been our efforts to develop alternate schemes. The approximation framework can be used efficiently to develop a class of schemes based on spline or "finite-element" ideas. A discussion of our findings in this regard will appear in a manuscript that is currently in preparation.

We close with several further remarks that we have added in the final version of this paper, partly in response to referees' queries and partly as a result of our subsequent efforts and findings in related investigations. First, as we noted in Remark 4.1, the generality of our theoretical framework (q dependent spaces, norms, etc.) is not essential to treat Example 4.1 or, indeed, any of the specific examples discussed above. However, if one considers a parabolic system as in Example 4.2 for which the function k is parameter dependent, the q dependence of the appropriate inner product is essential. In fact, such problems arise naturally in estimation questions for porous media problems, where one of the parameters to be estimated is the function k (the field porosity) itself. A treatment using the theoretical framework developed above in connection with cubic spline approximations is outlined for such problems in [42].

With regard to general spline approximation schemes, we have, since this paper was first written, completed certain efforts on spline-based techniques (referred to several times above) in the context of the theoretical framework given above. Second-order parabolic and hyperbolic systems [43], as well as higher-order equations arising in elasticity [44], have been treated and our findings have been most positive from both computational and theoretical viewpoints.

Acknowledgment. We would like to express our sincere appreciation to James Crowley who, in addition to developing and managing the software packages extremely well, was also helpful in various discussions on the practical use of the algorithms employed in connection with the numerical results of § 6.

REFERENCES

[1] K. J. ÅSTROM AND P. EYKHOFF, *System identification—A survey*, Automatica, 7 (1971), pp. 123–162.
[2] A. V. BALAKRISHNAN, *Applied Functional Analysis*, Springer, New York, 1976.
[3] A. BAMBERGER, G. CHAVENT AND P. LAILLY, *About the stability of the inverse problem in 1-D wave equations—Application to the interpolation of seismic problems*, Appl. Math. Opt., 5 (1979), pp. 1–47.
[4] H. T. BANKS, *Modeling and Control in the Biomedical Sciences*, Lecture Notes in Biomathematics 6, Springer, New York, 1975.
[5] ———, *Parameter identification techniques for physiological control systems*, in Lectures in Applied Mathematics Vol. 19, American Mathematical Society, Providence, RI, 1981, pp. 361–383.
[6] H. T. BANKS AND F. KAPPEL, *Spline approximations for functional differential equations*, J. Differential Equations, 34 (1979), pp. 496–522.

[7] H. T. BANKS AND A. MANITIUS, *Projection series for functional differential equations with applications to optimal control problems*, J. Differential Equations, 18 (1975), pp. 296–332.
[8] V. BARBU, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Noordhoff, Bucuresti, 1976.
[9] H. BREZIS, *Operateurs maximaux monotones*, North-Holland, Amsterdam, 1973.
[10] R. BURRIDGE, *The Gelfand–Levitan, the Marchenko, and the Gopinath–Sondhi integral equations of inverse scattering theory, regarded in the context of inverse impulse-response problems*, Wave Motion, 2 (1980), pp. 305–323.
[11] G. CHAVENT, M. DUPUY AND P. LEMONNIER, *History matching by use of optimal theory*, Soc. Petroleum Eng. J., 15 (1975), pp. 74–86.
[12] A. E. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.
[13] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, Vol. I, Springer, New York, 1953.
[14] J. W. DETTMAN, *Mathematical Methods in Physics and Engineering*, McGraw-Hill, New York, 1962.
[15] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators*, Vols. I–III, John Wiley, New York, 1957, 1963, 1971.
[16] R. E. EWING AND R. S. FALK, *On some ill-posed problems arising in glaciology*, preprint.
[17] G. R. GAVALAS AND J. H. SEINFELD, *Reservoirs with spatially varying properties: Estimation of volume from late transient pressure data*, Soc. Petroleum Eng. J., 13 (1973), pp. 335–342.
[18] F. GRANT AND G. WEST, *Interpretation Theory in Applied Geophysics*, McGraw-Hill, New York, 1965.
[19] A. HARAUX, *Nonlinear Evolution Equations—Global Behavior of Solutions*, Lecture Notes in Mathematics 841, Springer-Verlag, Berlin, 1981.
[20] G. HELLWIG, *Differential Operators of Mathematical Physics*, Addison-Wesley, Reading, MA, 1964.
[21] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.
[22] S. G. KREIN, *Linear Differential Equations in Banach Space*, Transl. of Math. Monographs 29, American Mathematical Society, Providence, RI, 1971.
[23] C. S. KUBRUSLY, *Distributed parameter system identification, a survey*, Int. J. Control, 26 (1977), pp. 509–535.
[24] T. G. KURTZ, *Extensions of Trotter's operator semigroup approximation theorem*, J. Functional Anal., 3 (1969), pp. 354–375.
[25] S. LANG, *Analysis II*, Addison-Wesley, Reading, MA, 1969.
[26] P. D. LAX AND R. S. PHILLIPS, *Scattering Theory*, Academic Press, New York, 1967.
[27] J. L. LIONS AND E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications*, Vol. I, Springer, New York, 1972.
[28] R. H. MARTIN, *Nonlinear Operators and Differential Equations in Banach Spaces*, John Wiley & Sons, New York, 1976.
[29] A. OKUBO, *Diffusion and Ecological Problems: Mathematical Models*, Biomathematics, 10, Springer, New York, 1980.
[30] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Lecture Notes 10, University of Maryland, College Park, 1974.
[31] M. P. POLIS AND R. E. GOODSON, *Parameter identification in distributed systems: A synthesizing overview*, Proc. IEEE, 64 (1976), pp. 43–61.
[32] W. H. RAY, *Some recent applications of distributed parameter systems theory—A survey*, Automatica, 14 (1978), pp. 281–287.
[33] F. RIESZ AND B. SZ.-NAGY, *Functional Analysis*, F. Ungar, New York, 1955.
[34] G. A. ROSENBERG, W. T. KYNER AND E. ESTRADA, *Bulk flow of brain interstitial fluid under normal and hyperosmolar conditions*, Amer. J. Physiol., 238 (1980), pp. F42–F49.
[35] A. RUBERTI, ed., *Distributed Parameters Systems: Modelling and Identification*, Lecture Notes in Control and Information Sciences 1, Springer, Berlin, 1978.
[36] W. RUDIN, *Functional Analysis*, McGraw-Hill, New York, 1973.
[37] P. C. SABATIER, ed., *Applied Inverse Problems*, Lecture Notes in Physics 85, Springer, Berlin, 1978.
[38] P. C. SHAH, G. R. GAVALAS AND J. H. SEINFELD, *Error analysis in history matching: The optimum level of parametrization*, Soc. Petroleum Eng. J., 18 (1978), pp. 219–228.
[39] A. E. TAYLOR, *Introduction to Functional Analysis*, John Wiley, New York, 1967.
[40] J. A. WALKER, *Dynamical Systems and Evolution Equations, Theory and Applications*, Plenum Press, New York, 1980.
[41] K. YOSIDA, *Functional Analysis*, Springer, Berlin, 1971.

[42] H. T. BANKS, *A survey of some problems and recent results for parameter estimation and optimal control in delay and distributed parameter systems*, LCDS Rep. 81-19, Brown Univ., July, 1981; to appear in Proc. Conf. on Volterra and Functional Differential Equations, Blacksburg, VA, June 10–13, 1981.

[43] H. T. BANKS, J. M. CROWLEY AND K. KUNISCH, *Cubic spline approximation techniques for parameter estimation in distributed systems*, LCDS Rep. 81–25, Brown Univ., Nov., 1981.

[44] H. T. BANKS AND J. M. CROWLEY, *Parameter estimation for distributed systems arising in elasticity*, LCDS Rep. 81-24, Brown Univ., Nov. 1981; to appear in Proc. Symposium on Engineering Sciences and Mechanics, National Cheng Kung University, Tainan, Taiwan, Dec. 28–31, 1981.

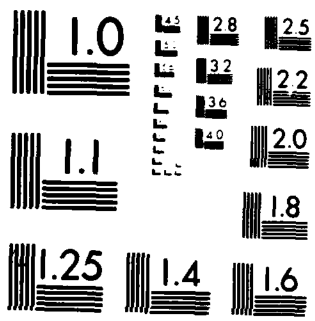

D
E
ED

CON

AD-A122 925 AN APPROXIMATION THEORY FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS WITH..(U) BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMICAL SYSTE.. 2/2

UNCLASSIFIED H T BANKS ET AL. 01 MAY 81 ARO-16835.27-MA F/G 12/1 NL

END
DATE
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

SUPPLEMENTARY

INFORMATION

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) UNCLASSIFIED

AD-A123 925

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 82-1097	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) AN APPROXIMATION THEORY FOR NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS WITH APPLICATIONS TO IDENTIFICATION AND CONTROL	5. TYPE OF REPORT & PERIOD COVERED Reprint
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) H.T. Banks and K. Kunisch	8. CONTRACT OR GRANT NUMBER(s) AFOSR-76-3092
9. PERFORMING ORGANIZATION NAME AND ADDRESS LCDS - Box F, Brown University, Providence, RI 02912	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A4
11. CONTROLLING OFFICE NAME AND ADDRESS Mathematical and Information Sciences Directorate, Air Force Office of Scientific Research, Bolling AFB, DC 20332	12. REPORT DATE November 1982
	13. NUMBER OF PAGES 35
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.

DTIC ELECTE MAR 1 1983 S D H

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

18. SUPPLEMENTARY NOTES

SIAM J. CONTROL AND OPTIMIZATION, Vol. 20, No. 6, November 1982

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 results from linear semigroup theory are used to develop a general framework for convergence of approximation schemes in parameter estimation and optimal control problems for nonlinear partial differential equations. These ideas are used to establish theoretical convergence results for parameter identification using model (eigenfunction) approximation techniques. Results from numerical investigations of these schemes for both hyperbolic and parabolic systems are given.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LME

8