AD-A117 068 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 21/1
ESTIMATION OF DELAYS AND OTHER PARAMETERS IN NONLINEAR FUNCTION--ETC(U)
DEC 81 H T BANKS, P L DANIEL DAAG29-79-C-0161
UNCLASSIFIED LCDS-82-2 AFOSR-TR-82-0533 NL

1 OF 1
AD A
117068

END
DATE
FILMED
08-82
DTIC

AFOSR-TR- 82-0533

2

AD A117068

Lefschetz Center for Dynamical Systems

Division of Applied Mathematics

Brown University Providence RI 02912

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

2

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 82-0533	2. GOVT ACCESSION NO AD-A117068	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ESTIMATION OF DELAYS AND OTHER PARAMETERS IN NONLINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
		6. PERFORMING ORG. REPORT NUMBER LCDS# 82-2
7. AUTHOR(s) H.T. BANKS AND P.L. DANIEL		8. CONTRACT OR GRANT NUMBER(s) AFOSR 81-0198
9. PERFORMING ORGANIZATION NAME AND ADDRESS LEFSCHETZ CENTER FOR DYNAMICAL SYSTEMS DIV. OF APPLIED MATH., BROWN UNIVERSITY PROVIDENCE RI 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE 61102F; 2304/A4-
11. CONTROLLING OFFICE NAME AND ADDRESS MATHEMATICAL & INFORMATION SCIENCES AIR FORCE OFFICE OF SCIENTIFIC RESEARCH BOLLING AFB DC 20332		12. REPORT DATE DECEMBER 1981
		13. NUMBER OF PAGES 44
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

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

DTIC ELECTE JUL 19 1982 S D B

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)

WE DISCUSS A SPLINE-BASED APPROXIMATION SCHEME FOR NONLINEAR NONAUTONOMOUS DELAY DIFFERNTIAL EQUATIONS. CONVERGENCE RESULTS (USING DISSPATIVE TYPE ESTIMATES ON THE UNDERLYING NONLINEAR OPERATORS) ARE GIVEN IN THE CONTEXT OF PARAMETER ESTIMATION PROBLEMS WHICH INCLUDE ESTIMATION OF MULTIPLE DELAYS AND INITIAL DATA AS WELL AS THE USUAL COEFFICIENT-TYPE PARAMETERS. A BRIEF SUMMARY OF SOME OF OUR RELEATED NUMERICAL FINDINGS IS ALSO GIVEN.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ESTIMATION OF DELAYS AND OTHER PARAMETERS IN NONLINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS

by

H. T. Banks and P. L. Daniel

December 1981 LCDS Report #82-2

ESTIMATION OF DELAYS AND OTHER PARAMETERS IN NONLINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS†

H. T. Banks*
Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, RI 02912

and

P. L. Daniel*
Department of Mathematics
Southern Methodist University
Dallas, Texas 75275
TX

December, 1981

*Part of this research was completed while both authors were visitors at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, Va., operated under NASA contracts No. NAS1-15810 and NAS1-16394.

†Work supported in part by the Air Force Office of Scientific Research under contract AFOSR 76-3092[illegible], 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-DAAG29-79-C-0161.

ESTIMATION OF DELAYS AND OTHER PARAMETERS IN NONLINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS

H. T. Banks and P. L. Daniel

ABSTRACT

We discuss a spline-based approximation scheme for nonlinear nonautonomous delay differential equations. Convergence results (using dissipative type estimates on the underlying nonlinear operators) are given in the context of parameter estimation problems which include estimation of multiple delays and initial data as well as the usual coefficient-type parameters. A brief summary of some of our related numerical findings is also given.

§1 Introduction

In [6] spline approximation ideas are developed in the context of numerical algorithms for solution of functional differential equations (FDE). The theoretical framework is based on a functional analytic formulation (in an appropriately chosen Hilbert space Z) of Ritz-Galerkin type ideas where one approximates on finite-dimensional subspaces Z^N the underlying linear solution semigroup T(t) (with infinitesimal generator A) for the FDE by linear semigroups $T^N(t)$ (with infinitesimal generators $A^N = P^N A P^N$, where P^N is the orthogonal projection of Z onto Z^N). These ideas were subsequently ([3], [4]) used in developing numerical schemes for parameter estimation and optimal control problems. The fundamental theoretical tool employed is the Trotter-Kato theorem (a functional analytic formulation of the Lax equivalence theorem: stability plus consistency yields convergence of approximation schemes) for linear semigroups. In this paper we present approximation results that subsume those in [6], [3], and [4] in that we develop schemes to estimate parameters that include multiple delays, coefficients and initial data for nonlinear nonautonomous FDE. Our theoretical arguments avoid the Trotter-Kato linear semigroup formulation altogether. Rather, we combine dissipative type estimates with use of the Gronwall inequality to develop a theory that not only allows for rather general nonlinearities but also accomodates with ease nonautonomous systems (both of which are features that the Trotter-Kato linear semigroup framework excludes). Of course, one could use an evolution operator analogue of the Trotter-Kato approximation theorem to obtain results for linear nonautonomous equations (see [7] for details), or a nonlinear Trotter-Kato type theorem for nonlinear autonomous FDE (see [13], [14]). Both of these separate approaches however are less direct than the one developed here when applied to parameter estimation problems.

While the approximation methods we develop can be used with great success to simply solve initial value problems for nonlinear nonautonomous FDE, the main focus of our treatment here is parameter identification or estimation. That our ideas can also be fruitfully employed in control problems is demonstrated in [9], [10] while application of the methods to estimation problems for certain partial differential equations can be found in [5].

The fundamental ideas (which were first presented for simple nonlinear autonomous, known delay, estimation problems in [1] and subsequently extended to treat nonautonomous, unknown delay, FDE problems in [10]) are really quite simple. However, the development of a theory for identification of the delays is a delicate matter since the "history space" for the delay system changes as one iteratively estimates the delays. This, unfortunately, results in a rather complicated presentation from the standpoint of technical notation regardless of the approach (e. g., see treatment of the linear autonomous system case in [4]).

Our presentation is as follows: In section 2 we describe a parameter estimation problem for FDE's and give an equivalent Hilbert space formulation involving an abstract nonlinear evolution equation. Section 3 contains a discussion of approximate estimation problems based on spline subspaces; general convergence results are given. We conclude with a final section in which we present representative numerical findings obtained using the approximation scheme proposed in section 3.

Most of the notation (e.g., H^p for Sobolev spaces, L_p for Lebesgue spaces, etc.) is quite standard and is in accordance with popular usage. The symbol $|\cdot|$ will be used in general to denote the norm in various spaces in instances where no confusion will result. However in some situations it is absolutely essential to distinguish special weighted norms. These special norms will be defined as they are used in the discussions below. For convenience of the reader, we have summarized these definitions in a brief appendix for quick reference.

Finally we wish to mention the motivation behind our efforts to develop the methods presented below. In [4] and [10] one finds brief descriptions of nonlinear delay equation estimation problems arising in the study of enzymatically active column reactors. Although such problems actually prompted the theoretical investigations that we report here, a discussion of the application of our methods to these problems would be quite lengthy and thus will be the subject of a separate report.

§2 Formulation of parameter estimation problems for nonlinear FDE

In the present section we describe the parameter estimation problem for a delay differential system and detail conditions under which solutions exist. Our approach then is to reformulate the FDE - governed identification problem as an abstract problem on an infinite-dimensional state space, concluding the section by establishing the equivalence between the FDE and the abstract state equation.

We consider the vector nonlinear delay equation

$$(2.1) \qquad \begin{cases} \dot{x}(t) = f(\alpha,\ r_\nu,\ t,\ x(t),\ x_t,\ x(t-r_1),\ \ldots,\ x(t-r_\nu)) + g(t), & a \le t \le b \\ (x(a),\ x_a) = (\eta,\ \phi) \end{cases}$$

where x_t denotes the R^n-valued function $\Theta \to x_t(\Theta) = x(t + \Theta)$, $- r \le \Theta \le 0$, and g is a general L_2^n (a, b) perturbation term. The equation depends on the parameters $\gamma = (\zeta,\ q)$, where $\zeta = (\eta,\ \phi)$ is the initial data in some set $S \subseteq W$, with

$$W \equiv \{(\psi(0),\ \psi) \epsilon R^n \times L_2^n(-\ r,\ 0) \,|\, \psi \epsilon H^1(-\ r,\ 0)\}.$$

The parameter $q = (\alpha,\ r_1,\ \ldots,\ r_\nu)$ is assumed to be in $Q = \mathcal{A} \times \mathcal{R}$ where α is a coefficient-type parameter in the set $\mathcal{A} \subseteq R^\mu$ and the discrete delays are chosen from the set

$$\mathcal{R} = \{(r_1,\ \ldots,\ r_\nu) \epsilon R^\nu \,|\, 0 = r_0 \le r_1 \le \ldots \le r_\nu \le r,\ r_\nu > 0\}$$

with $r > 0$ fixed and given throughout this paper.

To simplify notation we shall use $|\cdot|_{r_\nu}$ to denote the norm on $L_2^n(-\ r_\nu,\ 0)$ while we use $|\cdot|$ to denote the norm on $L_2^n(-\ r,\ 0)$ and $L_2^n(a,\ b)$. We make the following standing assumptions on f throughout the paper:

(H1) The mapping f satisfies a global Lipschitz condition on $R^n \times L_2^n(-r, 0) \times R^{n\nu}$ uniformly in $(\alpha, r_\nu) \in A \times [0, r]$. That is, there exists $m_1 \in L_2(a, b)$, $m_1 > 0$, such that for all $(\xi, \psi, w_1, \ldots, w_\nu), (\delta, \chi, y_1, \ldots, y_\nu) \in R^n \times L_2^n(-r, 0) \times R^{n\nu}$,

$$|f(\alpha, r_\nu, t, \xi, \psi, w_1, \ldots, w_\nu) - f(\alpha, r_\nu, t, \delta, \chi, y_1, \ldots, y_\nu)|$$
$$\leq m_1(t) \{|\xi - \delta| + |\psi - \chi|_{r_\nu} + \sum_{i=1}^{\nu} |w_i - y_i|\}$$

for all $(\alpha, r_\nu) \in A \times [0, r]$ and a.a. $t \in [a, b]$.

(H2) For each $(\alpha, r_\nu) \in A \times [0, r]$,

$f(\alpha, r_\nu, \cdot, \ldots, \cdot): [a, b] \times R^n \times L_2^n(-r, 0) \times R^{n\nu} \to R^n$

is differentiable, and

$t \to f(\alpha, r_\nu, t, \psi(0), \psi, \psi(-r_1), \ldots, \psi(-r_\nu))$ is in $H^1(a,b)$

for every $\psi \in C^n(-r,0) \equiv C([-r,0];R^n)$ and every $(\alpha, r_1, \ldots, r_\nu) \in A \times R$.

(H3) Given any $x \in C^n[a-r,b]$, the mapping

$\sigma \to f_t(\alpha, r_\nu, \sigma, x(\sigma), x_\sigma, x(\sigma-r_1), \ldots, x(\sigma-r_\nu))$ is in $L_2^n(a,b)$ for all $q \in Q$.

(H4) The function f is continuous on $A \times [0, r] \times [a, b] \times R^n \times L_2^n(-r, 0) \times R^{n\nu}$.

<u>Remark 2.1</u> It follows immediately from (H1) and (H2) that f satisfies an affine growth condition; that is, for a given $x \in L_2^n(a - r, b)$,

(2.2) $$|f(\alpha, r_\nu, t, x(t), x_t, x(t - r_1), \ldots, x(t - r_\nu))|$$
$$\leq m_1(t) \{|x(t)| + |x_t|_{r_\nu} + \sum_{i=1}^{\nu} |x(t - r_i)|\} + m_2(\alpha, r_\nu, t),$$

where $m_2(\alpha, r_\nu, t) = |f(\alpha, r_\nu, t, 0, \ldots, 0)|$ is in $L_2(a, b)$. Quite standard arguments may then be employed to demonstrate that, for each $q \in Q$,

$t \to f(\alpha, r_\nu, t, x(t), x_t, x(t - r_1), \ldots, x(t - r_\nu))$ is in $L_1^n(a, b)$ and that the mapping depends only on the equivalence class of x; therefore there will be no difficulty associated with point evaluations of x appearing in f since we shall write (2.1) as an equivalent integral equation.

Before we direct our attention to the estimation of the parameters appearing in (2.1) we shall first state results guaranteeing the existence, uniqueness and continuous dependence of solutions to the state equation for each choice of parameters $(\zeta, q) \in S \times Q$.

Theorem 2.1 Let $\gamma = (\zeta, q) = (\eta, \phi, \alpha, r_1, \ldots, r_\nu)$ be given in $S \times Q$. There exists a unique solution x to (2.1) on the interval $[a - r, b]$ which depends continuously on $\{\eta, \phi, g\}$ in the $R^n \times L_2^n(-r,0) \times L_2^n(a,b)$ topology.

In the proof, which may be found in [10, p. 6] and will not be detailed here, one employs general uniform contraction principles (see [12, p. 7]) and relies heavily on hypothesis (H1) and the growth condition (2.2).

We turn now to an examination of (2.1) when the parameters, including the delays $r_1, \ldots, r_\nu$, and initial data (η, ϕ) are to be estimated. We will restrict our attention to parameters in the admissible initial data-parameter set $\Gamma = S \times Q$ where we assume throughout that Γ has the following property:

(H5) Q is compact in $R^{\mu+\nu}$ and $S \subseteq W$ is compact in the $R^n \times L_2^n(-r, 0)$ topology.

The identification problem consists of finding $\bar{\gamma} \in \Gamma$ which provides the best least squares fit of the parameter-dependent solution (of the model equations (2.1)) to observations of the output at discrete sample times. The problem, which could also be reformulated as a maximum likelihood estimation problem, may be formally stated as follows:

Given g and observations $\{\hat{u}_i\}$, $\hat{u}_i \in R^s$, at times $\{t_i\}$, i = 1, ..., M, find $\bar{\gamma}$ in Γ which minimizes

$$J(\gamma) = \frac{1}{2} \sum_{i=1}^{M} |C(q)x(t_i; \gamma) - \hat{u}_i|^2 \tag{2.3}$$

over all $\gamma = (\zeta, q)$ in Γ. Here C is a given s x n matrix continuous in q; and $u(t;\gamma) = C(q)x(t;\gamma)$ represents the "observable part" of $x(t;\gamma)$, the solution to (2.1) corresponding to γ.

Remark 2.2: For $\bar{\gamma} = (\bar{\eta}, \bar{\phi}, \bar{\alpha}, \bar{r}_1, \ldots, \bar{r}_\nu)$ the optimal parameter, it may happen that $\bar{r}_\nu < r$ so that we actually only need $\bar{\phi}$ defined on $[-\bar{r}_\nu, 0]$ to integrate the state equations (2.1) (and, in fact, the $\bar{\phi}$ we obtain in practice will be defined on that interval only). We will view $\bar{\phi}$ as a function on all of $[-r, 0]$ by making an arbitrary, but definite, continuous extension from $-\bar{r}_\nu$ back to $-r$, so that $(\bar{\eta}, \bar{\phi})$ is an element of S as required.

Remark 2.3: The compactness assumption on S will not be difficult to satisfy in practice since a sufficient condition for compactness is that all elements (η, ϕ) in S are such that η belongs to a compact set in R^n and ϕ is bounded in $H^1(-r, 0)$. An example of one such admissible initial data set is the set of all polynomials on $[-r, 0]$ of order $\leq k$ (k a nonnegative integer) with coefficients in a compact set.

§ 2.1 An abstract reformulation of the estimation problem

We next reformulate (2.1) as an abstract evolution equation in an infinite-dimensional state variable. Our approach involving use of the state space $R^n \times L_2^n(-r, 0)$ is quite standard and well-established in the FDE literature (see, for example, [2] and the references therein); however, the dependence here of operators and state spaces on unknown parameters requires that we make such definitions in this and the following section with a certain amount of care.

We will let $Z = R^n \times L_2^n(-r, 0)$ with norm $|\cdot|$ induced by the inner product $\langle(\xi, \psi), (\delta, \chi)\rangle \equiv \xi^T\delta + \int_{-r}^{0} \psi^T\chi$. For $(q, t) \in Q \times [a, b]$, $(\xi, \psi) \in R^n \times C^n(-r, 0)$ define $F(q, t, \xi, \psi) = f(\alpha, r_\nu, t, \xi, \psi, \psi(-r_1), \ldots, \psi(-r_\nu))$ and $A(q, t): W \to Z$ by

$$A(q, t)(\psi(0), \psi) = (F(q, t, \psi(0), \psi), \mathcal{D}\psi), \tag{2.4}$$

where $\mathcal{D}\psi$ denotes the $L_2^n(-r, 0)$ function that is the derivative of ψ. In addition, let $G(t) = (g(t), 0) \in Z$, for $t \in [a, b]$.

The equivalence of the FDE (2.1) to an abstract evolution equation is detailed in Theorem 2.2; before proceeding, however, we need two results that also will be called upon frequently in §3, so they are stated here as lemmas. Our first proposition is actually a restatement of the well-known result [8, p. 100] that $\frac{d}{dt}\frac{1}{2}|x(t)|^2 = \langle\dot{x}(t), x(t)\rangle$.

Lemma 2.1. If X is a Hilbert space and if $x: [a, b] \to X$ is given by $x(t) = x(a) + \int_a^t v(\sigma)d\sigma$, then

$$|x(t)|^2 = |x(a)|^2 + 2\int_a^t \langle x(\sigma), v(\sigma)\rangle d\sigma.$$

The second result describes how to construct an equivalent topology for Z so that the nonlinear operator A satisfies a dissipative-type inequality. The lemma, a nonlinear version of that found in [2, p. 186] and [6], greatly simplifies our calculations and is the foundation for our development without the use of semigroups.

Lemma 2.2. Let $q = (\alpha, r_1, \ldots, r_\nu) \in Q$ be given. For $y = (\xi, \psi)$, $z = (\delta, \chi) \in Z$ define a new inner product on Z by $\langle y, z\rangle_q \equiv \xi^T\delta + \int_{-r}^{0} \psi(\theta)\chi(\theta)\tilde{\rho}(q)(\theta)d\theta$ where $\tilde{\rho}(q)$ is given on $[-r, 0]$ by

$$\tilde{\rho}(q)(\theta) = \begin{cases} 1, & \theta \in [-r, -r_\nu] \\ 2, & \theta \in (-r_\nu, -r_{\nu-1}] \\ \vdots & \vdots \\ \nu+1, & \theta \in (-r_1, 0]. \end{cases} \tag{2.5}$$

Then,

$$\langle A(q, t)y - A(q, t)z, y - z\rangle_q \leq \omega(t)|y - z|_q^2$$

for all $q \in Q$, almost all $t \in [a, b]$ and all $y, z \in W$. The function $\omega > 0$ is in $L_1(a,b)$ and is given by $\omega(t) = \frac{3}{2} m_1(t) + \frac{\nu + 1}{2} + \frac{\nu}{2} m_1^2(t)$.

Proof: Let $y = (\psi(0), \psi)$, $z = (\chi(0), \chi) \in W$.

$$\begin{aligned} &\langle A(q, t)y - A(q, t)z, y - z\rangle_q = \\ &\quad [F(q, t, \psi(0), \psi) - F(q, t, \chi(0), \chi)]^T[\psi(0) - \chi(0)] \\ &\qquad + \int_{-r}^{0} (\mathcal{D}\psi - \mathcal{D}\chi)(\psi - \chi)(\theta)\tilde{\rho}(q)(\theta)d\theta \end{aligned}$$

where

$$\begin{aligned} &\int_{-r}^{0} (\mathcal{D}\psi - \mathcal{D}\chi)(\psi - \chi)(\theta)\tilde{\rho}(q)(\theta)d\theta \\ &\quad = \int_{-r}^{-r_\nu} \mathcal{D}\left(\frac{|\psi(\theta) - \chi(\theta)|^2}{2}\right) d\theta \\ &\qquad + \sum_{j=1}^{\nu} \int_{-r_j}^{-r_{j-1}} \mathcal{D}\left(\frac{|\psi(\theta) - \chi(\theta)|^2}{2}\right)(\nu + 2 - j)d\theta \end{aligned}$$

$$= \frac{|\psi(-r_\nu) - \chi(-r_\nu)|^2}{2} - \frac{|\psi(-r) - \chi(-r)|^2}{2} + \frac{|\psi(0) - \chi(0)|^2}{2} (\nu + 1)$$

$$- \sum_{j=1}^{\nu-1} \frac{|\psi(-r_j) - \chi(-r_j)|^2}{2} - \frac{|\psi(-r_\nu) - \chi(-r_\nu)|^2}{2} \cdot 2$$

$$\leq \frac{(\nu + 1)}{2} |\psi(0) - \chi(0)|^2 - \sum_{j=1}^{\nu} \frac{|\psi(-r_j) - \chi(-r_j)|^2}{2} .$$

Therefore, for almost all $t \in [a, b]$,

$$\langle A(q, t)y - A(q, t)z, y - z\rangle_q$$

$$\leq m_1(t)|\psi(0) - \chi(0)|^2 + m_1(t)|\psi - \chi|_{r_\nu} |\psi(0) - \chi(0)|$$

$$+ \sum_{j=1}^{\nu} |\psi(-r_j) - \chi(-r_j)| \quad (m_1(t)|\psi(0) - \chi(0)|)$$

$$+ \frac{\nu + 1}{2} |\psi(0) - \chi(0)|^2 - \frac{1}{2} \sum_{j=1}^{\nu} |\psi(-r_j) - \chi(-r_j)|^2$$

$$\leq \left(m_1(t) + \frac{m_1(t)}{2} + \frac{\nu m_1^2(t)}{2} + \frac{\nu + 1}{2}\right) |\psi(0) - \chi(0)|^2 + \frac{m_1(t)}{2} | \psi - \chi|_{r_\nu}^2$$

where we have used repeatedly the fact that $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$. It follows then that

$$\langle A(q, t)y - A(q, t)z, y - z\rangle_q \quad \leq \omega(t)|\psi(0) - \chi(0)|^2$$

$$+ \omega(t)|\psi - \chi|_{r_\nu}^2$$

$$\leq \omega(t)|y - z|_q^2 .$$

$\square$

It is clear that for all $q \in Q$, the norm $|\cdot|_q$ on Z induced by the $\tilde{\rho}(q)$ weighted inner product is equivalent to the usual Z norm since $1 \leq \tilde{\rho}(q) \leq \nu + 1$.

<u>Theorem 2.2</u> For fixed $\gamma \in \Gamma$ let $y(t; \gamma, g) = (x(t; \gamma, g), x_t(\gamma, g))$, where x is the solution to (2.1) corresponding to $\gamma = (\eta, \phi, \alpha, r_1, \ldots, r_\nu)$ and $g \in L_2^n(a, b)$. Then $y(\gamma, g)$ is the unique solution on $[a, b]$ of

$$(2.6) \qquad z(t) = \zeta + \int_a^t \{A(q, \sigma)\, z(\sigma) + G(\sigma)\} d\sigma \, .$$

Furthermore, $y(t; \gamma, g) \in Z$ is continuous in $t \in [a, b]$, and uniformly continuous in $\{\zeta, g\} \in W \times L_2^n(a, b)$ (in the $Z \times L_2^n(a, b)$ topology) uniform in $t \in [a, b]$.

Proof: We shall sketch the proof of the theorem, which demonstrates that (i) the integrand in (2.6) is well-defined and integrable, (ii) the equality in (2.6) holds for $z(t) = y(t; \gamma, g)$, and (iii) the solution y is unique and continuously dependent.

To prove (i) we first must show that $y(t; \gamma, g) \in \text{dom}\,(A(q, t)) = W$ for each $t \in [a, b]$, or, since $x_t(0) = x(t)$, that $x_t \in H^1(-r, 0)$ for all t. Using the affine growth condition (2.2) on f and the continuity of x it is not difficult to show that $t \to f(\alpha, r_\nu, t, x(t), x_t, x(t - r_1), \ldots, x(t - r_\nu))$ is square-integrable on $[a, b]$ so that, using the fact that $\dot{x} = \dot{\phi}$ on $[a - r, a]$, we obtain $\dot{x} \in L_2^n\,(a - r, b)$. Standard estimates in [11, p. 254] may be invoked to demonstrate that

$$(2.7) \qquad \frac{1}{\varepsilon}\,(x_{t+\varepsilon} - x_t) \to (\dot{x})_t \quad \text{in} \quad L_2^n(-r, 0)$$

so that $\mathcal{D}\,(x_t) = (\dot{x})_t \in L_2^n\,(-r, 0)$, for each $t \in [a, b]$. Arguments similar to these may be used to show $A(t)y(t) + G(t)$ is integrable on $[a, b]$, concluding the proof of (i).

The argument that $y(t) = (x(t;\gamma, g), x_t(\gamma, g))$ satisfies equation (2.6) is trivial if this equation is examined componentwise: The R^n part of (2.6) is simply a restatement of (2.1) while the desired equality for the $L_2^n(-r, 0)$ component follows immediately from (2.7).

Finally, uniqueness and continuous dependence of solutions on $\{\zeta, g\}$ follow from arguments that will be repeated often throughout this paper and will be presented in detail here for the case of continuous dependence; uniqueness also follows from these arguments. Let z_1, z_2 denote solutions to (2.6) corresponding to $\{\zeta_1, g_1\}$, $\{\zeta_2, g_2\}$ respectively with $q \in Q$ fixed. Then, for $t \in [a, b]$,

$$z_1(t) - z_2(t) = \zeta_1 - \zeta_2 + \int_a^t \{A(\sigma)z_1(\sigma) - A(\sigma)z_2(\sigma) + (g_1(\sigma), 0) - (g_2(\sigma), 0)\}d\sigma$$

so that from Lemma 2.1,

$$\begin{aligned}
|z_1(t) - z_2(t)|_q^2 &\le |\zeta_1 - \zeta_2|_q^2 \\
&\quad + 2\int_a^t \langle A(\sigma)z_1(\sigma) - A(\sigma)z_2(\sigma), z_1(\sigma) - z_2(\sigma)\rangle_q \, d\sigma \\
&\quad + 2\int_a^t |(g_1(\sigma) - g_2(\sigma), 0)|_q \, |z_1(\sigma) - z_2(\sigma)|_q d\sigma \\
&\le |\zeta_1 - \zeta_2|_q^2 + 2\int_a^t \omega(\sigma)|z_1(\sigma) - z_2(\sigma)|_q^2 \, d\sigma \\
&\quad + \int_a^t |(g_1(\sigma) - g_2(\sigma), 0)|_q^2 + \int_a^t |z_1(\sigma) - z_2(\sigma)|_q^2 \, d\sigma \\
&\le |\zeta_1 - \zeta_2|^2(\nu + 1) + |g_1 - g_2|^2 + \int_a^b (2\omega(\sigma) + 1)|z_1(\sigma) - z_2(\sigma)|_q^2 \, d\sigma.
\end{aligned}$$

Gronwall's inequality may be used to obtain

$$|z_1(t) - z_2(t)|_q^2 \le (|\zeta_1 - \zeta_2|^2(\nu + 1) + |g_1 - g_2|^2)\cdot$$
$$\exp \int_a^b (2\omega(\sigma) + 1)d\sigma$$

from which continuous dependence (uniform in $t \in [a, b]$) follows at once. $\square$

We have demonstrated the equivalence between an FDE in $x(t) \in R^n$ and an abstract evolution equation (AEE) in the infinite-dimensional state variable $z(t)$. We remark that the infinite dimensionality of (2.6) is inherited from (2.1) in that in the latter the history of x on $[t - r_\nu, t)$ is required before x may be determined at t. Thus the computational difficulties encountered with (2.6) are not simply an undesirable feature of this reformulation of (2.1) but rather are a manifestation of the inherent infinite dimensionality of the underlying FDE.

In view of the established equivalence, the ID problem in (2.3) may be reformulated as an abstract ID problem, where we now wish to find $\bar{\gamma} \in \Gamma$ which minimizes

(2.8) $$J(\gamma) = \frac{1}{2} \sum_{i=1}^{M} |C(q)\pi_0 z(t_i; \gamma) - \hat{u}_i|^2$$

over all $\gamma \in \Gamma$ where $\pi_0 : Z \to R^n$ is defined by $\pi_0(\xi, \psi) = \xi$.

In the next section we investigate the problem of approximating the infinite-dimensional identification problem (2.8) by a sequence of finite-dimensional state space identification problems (where the state variable satisfies an ordinary differential equation (ODE) on a finite-dimensional state space X^N). Fundamental to this undertaking is the establishment of the convergence of solutions of the approximating systems on X^N to solutions of the original equation on Z. Although our formulation is a classical one of the Ritz type (involving orthogonal projections of an infinite-dimensional system onto a sequence of finite-dimensional subspaces) our problem is complicated by the fact that the state space changes for each choice of parameters $q = (\alpha, r_1, \ldots, r_\nu)$. This concept is explained in detail in [4, p. 800] and involves the idea that the natural state space for z(t) associated with the parameter choice $q = (\alpha, r_1, \ldots, r_\nu)$ is $X(q) = R^n \times L_2^n(-r_\nu, 0)$, where in general $X(q) \not\subseteq Z$. Since we would expect the (finite-dimensional) approximating spaces $X^N(q)$ associated with q to be subspaces of X(q), we obtain a sequence of spaces $\{X^N(q)\}$ where $X^N(q)$ is different for each choice of q and is usually not contained in Z.

§3 Approximate parameter estimation problem

Our focus in this section is the definition of finite-dimensional ODE-governed estimation problems which approximate the ID problem governed by the AEE (2.6), and their relationship to the original FDE-based ID problem. While we shall present the details for a scheme based on linear splines, arbitrary order spline approximations may be employed in a similar way with only slight modifications in the arguments detailed below (see the theory developed in [6] on which all of our development here is based).

For parameters $\gamma = (\zeta, q) \in \Gamma$ consider

$$z^N(t; \gamma) = P^N(q)\zeta + \int_a^t \{A^N(q, \sigma)z^N(\sigma; \gamma) + P^N(q)G(\sigma)\}d\sigma, \quad t \in [a, b], \tag{3.1}$$

where A^N and P^N are defined via the following q-dependent operators and spaces. For a given $q = (\alpha, r_1, \ldots, r_\nu)$ we define the Hilbert space X(q) as the set $R^n \times L_2^n(-r_\nu, 0)$ with inner product $\langle(\xi, \psi), (\delta, \chi)\rangle_{\rho,q} = \xi^T\delta + \int_{-r_\nu}^0 \psi(\theta)\chi(\theta)\rho(q)(\theta)d\theta$, where $\rho(q)(\theta) = \tilde{\rho}(q)(\theta) - 1$, with $\tilde{\rho}$ defined as in (2.5). We shall also use an equivalent topology on $X(q)$ given by the (unweighted) inner product $\langle(\xi, \psi), (\delta, \chi)\rangle_{X,q} = \xi^T\delta + \int_{-r_\nu}^{0} \psi(\theta)\chi(\theta)d\theta$. The operator $\mathcal{C}^+(q): X(q) \to Z$ is the "continuous extension" operator defined by $\mathcal{C}^+(q)(\xi, \psi) = (\xi, \tilde{\psi})$ where $\tilde{\psi} = \psi$ on $[-r_\nu, 0]$, $\tilde{\psi}(\theta) = \psi(-r_\nu)$, $\theta \in [-r, -r_\nu)$; $i(q) : Z \to X(q)$ is defined by $i(q)(\xi, \psi) = (\xi, \hat{\psi})$ where $\hat{\psi}$ is the restriction of ψ to $[-r_\nu, 0]$. The subspaces $X^N(q)$ of $X(q)$ are defined by $X^N(q) = \{(\psi(0), \psi) \mid \psi$ is a piecewise linear spline with knots at $t_j^N(q) = [-(j - (k-1)N)(r_k - r_{k-1})/N] - r_{k-1}$, $j = (k - 1)N + 1, \ldots, kN$, $k = 1, \ldots, \nu$; $t_0^N = 0\}$ and we denote by $\pi^N(q) : X(q) \to X^N(q)$ the canonical orthogonal projection of $X(q)$ onto $X^N(q)$ along $X^N(q)^\perp$. Finally, $P^N(q) : Z \to X^N(q)$ is defined by $P^N(q) = \pi^N(q)i(q)$ and $A^N(q, \sigma) : X(q) \to X^N(q)$ is given by $A^N(q, \sigma) = \pi^N(q)A(q, \sigma)\pi^N(q)$ where here $A(q, \sigma)$ is interpreted as an operator on $X(q)$ given by $A(q, \sigma)(\psi_q(0), \psi_q) = (F(q, \sigma, \psi_q(0), \overline{\psi}_q), \mathcal{D}\psi_q)$ for $(\psi_q(0), \psi_q) \in X(q)$ (with $\overline{\psi}_q$ the extension of ψ_q to all of $[-r, 0]$ defined by $\overline{\psi}_q \equiv 0$ outside of $[-r_\nu, 0]$).

<u>Remark 3.1</u> A^N is well-defined since $X^N(q)$, the range of $\pi^N(q)$, is contained in the domain of $A(q, t)$. Note also that $A^N(q, t)$ actually may be considered as an operator from Z into $X^N(q)$ if it is defined by $A^N(q, \sigma) = P^N(q)A(q, \sigma)P^N(q)$. For present uses though, $P^N(q)\zeta$ and $z^N(\sigma, \gamma)$ are in $X^N(q)$ so that viewing $A^N(q, \sigma)$ as an operator from $X^N(q)$ to itself yields (3.1) as an equation on $X^N(q)$, a finite-dimensional space since each of its elements is completely determined by its value at each of $\nu N + 1$ knots. Equation (3.1) is then equivalent to the ODE

$$
(3.2) \quad \begin{cases} \dot{z}^N(t;\gamma) = A^N(q, t)z^N(t;\gamma) + P^N(q)G(t), \quad t \in (a, b] \\ z^N(a;\gamma) = P^N(q)\zeta, \end{cases}
$$

which, as we shall show in the arguments that follow, approximates (2.1) in some sense.

When the parameter γ is unknown we may state an "approximate identification problem" P_N associated with (3.1) and (3.2):

Find $\bar{\gamma}^N = (\bar{\zeta}^N, \bar{q}^N) \in \Gamma$ so as to minimize

$$
J^N(\gamma) = \frac{1}{2} \sum_{i=1}^{M} |C(q)\pi_0 z^N(t;\gamma) - \hat{u}_i|^2
$$

over $\gamma \in \Gamma$, where g and observations $\hat{u}_i$ at times t_i, $i = 1, \ldots, M$ are given and $z^N(t;\gamma)$ *satisfies* (3.1).

We now establish the existence of a unique solution to (3.1) for each choice of γ, and, further, the existence of a solution $\bar{\gamma}^N$ to the N^{th} ID problem, P_N. First we must state an analog of Lemma 2.2 which demonstrates a type of dissipativeness for A^N.

<u>Lemma 3.1</u>. Let $q = (\alpha, r_1, \ldots, r_\nu) \in Q$ be given. Then

$$
\langle A^N(q, t)y^N - A^N(q, t)z^N, y^N - z^N \rangle_{\rho,q}
$$

$$
\leq \omega(t) \; |y^N - z^N|^2_{\rho,q}
$$

for all $y^N, z^N \in X^N(q)$ where ω, defined in Lemma 2.2, is independent of q and N.

Proof: Note first that for $y, z \in W(q) = \{(\psi(0), \psi) | \psi \in H^1(-r_\nu, 0)\}$ we may argue that

$$\langle A(q, t)y - A(q, t)z, y - z\rangle_{\rho,q} \leq \omega(t)|y - z|^2_{\rho,q}$$

using estimates similar to those used to prove Lemma 2.2, where $\omega(t)$ is independent of q, N. Then for $y^N, z^N \in X^N(q) \subseteq W(q)$,

$$\langle A^N(q, t)y^N - A^N(q, t)z^N, y^N - z^N\rangle_{\rho,q}$$

$$= \langle \pi^N(q)A(q, t)\pi^N(q)y^N - \pi^N(q)A(q, t)\pi^N(q)z^N, y^N - z^N\rangle_{\rho,q}$$

$$= \langle A(q, t)\pi^N(q)y^N - A(q, t)\pi^N(q)z^N, \pi^N(q)y^N - \pi^N(q)z^N\rangle_{\rho,q}$$

$$\leq \omega(t)|\pi^N(q)y^N - \pi^N(q)z^N|^2_{\rho,q}$$

$$\leq \omega(t)|y^N - z^N|^2_{\rho,q}$$

where we have used the properties of the (self-adjoint) orthogonal projection π^N and the dissipativeness of $A(q, t)$ on $W(q)$.

$\square$

Our next result demonstrates the existence of solutions to (3.1) as well as to the identification problem P_N. In addition, the proof sheds light on the numerical procedure used to solve (3.1).

Theorem 3.1 Let $g \in L_2^n(a, b)$ and $\gamma = (\zeta, q) \in \Gamma$ be given. Then there exists a unique solution $z^N(t; \gamma, g) \in X^N(q)$ to (3.1) on $[a, b]$ with the property that the map $\{i(q)\zeta, g\} \to z^N(t; (\zeta, q), g)$ is uniformly continuous on $X(q) \times L_2^n(a, b)$, uniformly in N and t. Finally, there exists a solution $\bar{\gamma}^N$ to the N^{th} identification problem P_N for each N = 1, 2, ...

Remark 3.2: The continuity with respect to initial data given in this theorem is actually "uniform in $q \in Q$" in the following sense: Given $\varepsilon > 0$, there exists $\delta > 0$ independent of q and N such that for $\zeta_1, \zeta_2 \in S$ and $q \in Q$ with $|\zeta_1 - \zeta_2|_{X,q} < \delta$, we have $|z^N(t; (\zeta_1, q), g) - z^N(t; (\zeta_2, q), g)|_{X,q} < \varepsilon$. This type of "uniformity in q" follows from the arguments given below for Theorem 3.1 and will be used in establishing the convergence results of Theorem 3.3.

Proof: We first argue existence, uniqueness and continuous dependence of solutions to (3.1). We shall do this using arguments similar to those in [1] and [6] (where $z^N(t)$ is written in terms of basis vectors for $X^N(q)$). Let $q \in Q$ be fixed, $q = (\alpha, r_1, \ldots, r_\nu)$, and let e_j^N denote the scalar first-order spline function on $[-r_\nu, 0]$ characterized by

$$e_j^N(t_i^N) = \delta_{ij}, \qquad i, j = 0, 1, \ldots, \nu N$$

where δ_{ij} is the Kronecker symbol and $t_i^N = t_i^N(q)$ are the knots defined for functions in $X^N(q)$, $i = 0, \ldots, \nu N$. Define

$$\hat{e}_j^N = (e_j^N(0), e_j^N), \quad j = 0, \ldots, \nu N, \text{ and}$$

$$\beta^N = (e_0^N, \ldots, e_{\nu N}^N) \otimes I, \text{ where}$$

I is the $n \times n$ identity matrix and $\otimes$ denotes the Kronecker product so that β^N is an $(n \times n(\nu N + 1))$-matrix-valued function on $[-r_\nu, 0]$. Represent by $\hat{\beta}^N$ the matrix-valued pair,

$$\hat{\beta}^N = (\beta^N(0), \beta^N).$$

From [6], $X^N(q) = \text{span}\{\hat{\beta}_j^N\}$, $j = 1, \ldots, n(\nu N + 1)$ where the basis vectors are given by $\hat{\beta}_j^N = (\beta_j^N(0), \beta_j^N)$, β_j^N the j^{th} column of β^N. It follows then that since $z^N(t) \in X^N(q)$, there exists $w^N(t) \in R^{n(\nu N + 1)}$ such that

$$z^N(t) = \hat{\beta}^N w^N(t)$$

$$= \sum_{j=0}^{\nu N} w_j^N(t)\hat{e}_j^N$$

$$= (w_0^N(t), \sum_{j=0}^{\nu N} w_j^N(t)e_j^N)$$

for $w_j^N(t) \in R^n$, $j = 0, \ldots, \nu N$. Furthermore, since $P^N G(t)$ and $P^N \zeta$ are vectors in $X^N(q)$, there exist $G^N(t)$, $\zeta^N \in R^{n(\nu N + 1)}$ such that

$$P^N G(t) = \hat{\beta}^N G^N(t)$$

and

$$P^N \zeta = \hat{\beta}^N \zeta^N$$

so that equation (3.2) may now be written in terms of $\hat{\beta}^N$ as

$$(3.3) \qquad \begin{cases} \hat{\beta}^N \dot{w}^N(t) = \mathcal{A}^N(q, t)\hat{\beta}^N w^N(t) + \hat{\beta}^N G^N(t), & t \in (a, b] \\ \hat{\beta}^N w^N(a) = \hat{\beta}^N \zeta^N . \end{cases}$$

Let $A^N(q, t)$ denote the representation of $\mathcal{A}^N(q, t)$ (restricted to $X^N(q)$) with respect to the basis of $X^N(q)$. Here $A^N(q, t)$ is nonlinear as opposed to the matrix (linear) version of the operator arising in [6]. As in [1] and [6], usual Galerkin calculations establish that the coefficients $w^N(t)$ in (3.3) satisfy

$$(3.4) \qquad \begin{cases} \dot{w}^N(t) = A^N(q, t)w^N(t) + G^N(t) , & t \in (a, b] \\ w^N(a) = \zeta^N. \end{cases}$$

We next establish a representation of $A^N(q, t)$ which will enable us to consider the existence and uniqueness of solutions to (3.4) as well as the realization of numerical solution techniques for the system. Note first that

$$A^N(q, t)z^N(t) = \pi^N(q)A(q, t)\pi^N(q)(w_0^N(t), \sum_{j=0}^{\nu N} w_j^N(t)e_j^N)$$

$$= \pi^N(q)\Big(f(\alpha, r_\nu, t, w_0^N(t), \sum_{j=0}^{\nu N} w_j^N(t)e_j^N, \sum_{j=0}^{\nu N} w_j^N(t)e_j^N(-r_1), \ldots, \sum_{j=0}^{\nu N} w_j^N(t)e_j^N(-r_\nu)), \sum_{j=0}^{\nu N} w_j^N(t)\mathcal{D}e_j^N\Big)$$

$$= \pi^N(q)(\tilde{f}(\alpha, r_\nu, t, w^N(t)), \sum_{j=0}^{\nu N} w_j^N(t)\mathcal{D}e_j^N)$$

where $\tilde{f} : A \times [0,r] \times R^{1+n(\nu N + 1)} \to R^n$ is defined by

$$\tilde{f}(\alpha, r_\nu, t, (v_0, \ldots, v_{\nu N})^T) =$$

$$f(\alpha, r_\nu, t, v_0, \sum_{j=0}^{\nu N} v_j e_j^N, \sum_{j=0}^{\nu N} v_j e_j^N(-r_1), \ldots, \sum_{j=0}^{\nu N} v_j e_j^N(-r_\nu))$$

for $v_j \in R^n$, $j = 0, \ldots, \nu N$, and can be shown to be globally Lipschitz in $(v_0, \ldots, v_{\nu N})^T \in R^{n(\nu N + 1)}$ since f satisfies such a condition. Thus, $\mathcal{A}^N(q, t)w^N(t) = \alpha^N(t)$, where $\alpha^N(t) \in R^{n(\nu N + 1)}$ is such that

$$\hat{\beta}^N\alpha^N(t) = A^N(q, t)\hat{\beta}^N w^N(t)$$
$$= \pi^N(q)\Big(\tilde{f}(\alpha, r_\nu, t, w^N(t)), \sum_{j=0}^{\nu N} w_j^N(t)\mathcal{D}e_j^N\Big).$$

It follows from [6, p. 508] that whenever $\pi^N(q)(\xi, \psi) = \hat{\beta}^N \delta^N$, $(\xi, \psi) \in X(q)$, $\delta^N \in R^{n(\nu N + 1)}$, we have

$$\delta^N = (Q^N)^{-1} h^N(\xi, \psi)$$

where the nonsingular matrix Q^N is given by $Q^N = (\beta^N(0))^T \beta^N(0) + \int_{-r_\nu}^{0} \beta^N(\theta)^T \beta^N(\theta) \rho(q)(\theta) d\theta$ and $h^N(\xi, \psi) = (\beta^N(0))^T \xi + \int_{-r_\nu}^{0} \beta^N(\theta)^T \psi(\theta) \rho(q)(\theta) d\theta$.

We may apply these results to obtain

$$\alpha^N(t) = (Q^N)^{-1} h^N(\tilde{f}(\alpha, r_\nu, t, w^N(t)), \sum_{j=0}^{\nu N} w_j^N(t) \mathcal{D} e_j^N)$$

$$= (Q^N)^{-1} \begin{pmatrix} \tilde{f}(\alpha, r_\nu, t, w^N(t)) \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (Q^N)^{-1} H_{12}^N w^N(t)$$

where H_{12}^N is given in [6] and [10] by

$$H_{12}^N = \begin{pmatrix} \langle \dot{e}_0^N, e_0^N \rangle & \cdots & \langle \dot{e}_{\nu N}^N, e_0^N \rangle \\ \vdots & & \vdots \\ \langle \dot{e}_0^N, e_{\nu N}^N \rangle & \cdots & \langle \dot{e}_{\nu N}^N, e_{\nu N}^N \rangle \end{pmatrix} \otimes I.$$

In this matrix $\langle \cdot, \cdot \rangle$ denotes the $\rho(q)$-weighted $L_2^1(-r_\nu, 0)$ inner product.

Similarly, $G^N(t)$ in (3.4) is given by $G^N(t) = (Q^N)^{-1} h^N(G(t))$, $h^N((g(t), 0)) = (g(t), 0, \ldots, 0)^T \in R^{n(\nu N + 1)}$, so that (3.4) may be rewritten as

$$(3.5) \quad \begin{cases} \dot{w}^N(t) = (Q^N)^{-1} (\tilde{f}(\alpha, r_\nu, t, w^N(t)) + g(t), 0 \ \ldots, 0)^T \\ \qquad + (Q^N)^{-1} H_{12}^N w^N(t), \quad t \in (a, b] \\ w^N(a) = \zeta^N, \end{cases}$$

an ODE in $w^N(t) = w^N(t; \gamma, g) \in R^{n(\nu N + 1)}$. Since $\hat{f}$ satisfies a global Lipschitz condition in $w^N(t)$, the form of (3.5) allows one to employ standard ODE theory to obtain the existence of a unique solution $w^N(t)$ on $[a,b]$. We can therefore conclude that

$$z^N(t) = \hat{\beta}^N w^N(t)$$

is the unique solution to (3.1) (and (3.2)) on $[a, b]$ for $\zeta \in Z$ given.

The proof of the continuous dependence on ζ and g as stated in the theorem is identical to the corresponding proof in Theorem 2.2 where dissipativeness for $A^N(q, t)$ is now used to show that whenever $|\zeta_1 - \zeta_2|_{X,q} < \delta$ and $|g_1 - g_2| < \delta$, $\delta = \delta(\varepsilon,\omega,a,b)$ independent of t, q, and N, we have for the corresponding solutions, $z_1^N(t; (\zeta_1, q), g_1)$ and $z_2^N(t; (\zeta_2, q), g_2)$,

$$\begin{aligned} &|z_1^N(t) - z_2^N(t)|_{X,q} \\ &\leq |z_1^N(t) - z_2^N(t)|_{\rho,q} \\ &< \varepsilon \end{aligned}$$

for all $t \in [a, b]$.

Finally, to establish existence of a solution $\bar{\gamma}^N$ to P_N, one argues continuity (for fixed N) of the map $\gamma = (\zeta, q) \rightarrow \pi_0 z^N(t; \gamma) = w_0^N(t, \gamma)$ and thus infers continuity of $\gamma \rightarrow J^N(\gamma)$ on the compact set Γ. But it is not difficult to see that the right side of (3.5) depends continuously on q as do the basis elements $\hat{e}_j^N(q)$. Continuous dependence results (with respect to parameters and initial data) from the theory of ordinary differential equations can then be invoked to obtain the desired conclusions.

$\square$

In view of the last result, we are assured of a solution $\bar{\gamma}^N$ to the N^{th} estimation problem P_N (which is a standard least squares problem governed by an ODE). Since an application of conventional optimization techniques requires a solution to (3.1) for each choice of γ, straightforward computational schemes may be devised to solve (3.5), the associated ODE in the "Fourier" coefficients $w^N(t)$. Although it may be relatively easy to solve the finite-dimensional problem P_N, the solution $\bar{\gamma}^N$ we find is meaningful only if $\bar{\gamma}^N$ approximates the solution $\bar{\gamma}$ to the original ID problem. Fundamental to the establishment of this fact (i.e., the convergence of $\bar{\gamma}^N$ to $\bar{\gamma}$ in some sense) is the demonstration that the sequence of state variables $\{z^N(t;\, \gamma^N,\, g)\}$ converges to $z(t;\, \tilde{\gamma},\, g)$ given <u>any</u> sequence $\{\gamma^N\}$ with $\gamma^N \to \tilde{\gamma} = (\tilde{\zeta},\, \tilde{q})$ in Γ. We shall first consider this problem for limits $\tilde{\gamma}$ and perturbing functions g such that $\{\tilde{\zeta},\, g\}$ lies in a smooth but dense subset of $W \times L_2^n(a,\, b)$ (which simplifies our calculations). We then extend the convergence results for all limits $\tilde{\zeta}$ and perturbations g such that $\{\tilde{\zeta},\, g\} \in W \times L_2^n(a,\, b)$.

§3.1 Convergence of state variables

We shall assume that a sequence of parameters $\{\gamma^N\}$ in Γ has been given, $\gamma^N = (\zeta^N,\, q^N) = (\eta^N,\, \phi^N,\, \alpha^N,\, r_1^N,\, \ldots,\, r_\nu^N)$, and that $\gamma^N \to \tilde{\gamma} = (\tilde{\zeta},\, \tilde{q}) = (\tilde{\eta},\, \tilde{\phi},\, \tilde{\alpha},\, \tilde{r}_1,\, \ldots,\, \tilde{r}_\nu)$, in the sense that (i) $q^N \to \tilde{q}$ in $R^{\mu+\nu}$ and (ii) $|i(q^N)(\tilde{\zeta} - \zeta^N)|_{\rho,q^N} \to 0$ as $N \to \infty$. We make the following standing assumptions on $\tilde{\gamma}$ and S:

(H6) There exists some $\delta_r > 0$ such that $|\tilde{r}_k - \tilde{r}_{k-1}| \geq \delta_r$, $k = 1, 2, \ldots, \nu$.

(H7) If $\zeta \in S$ then $C^+(q^N)i(q^N)\zeta \in S$ for all N.

Remark 3.3: We note that the set S involving all polynomials of order $\leq k$ on $[-r, 0]$ mentioned in Remark 2.3 does not, strictly speaking, satisfy (H7). However, the reader can easily see from the arguments below that a modification in defining the extension operator C^+ (rather than extend from $[-r_\nu, 0]$ to $[-r, 0]$ by constant values, extend any polynomial on $[-r_\nu, 0]$ to $[-r, 0]$ by simply extending the domain of definition of the polynomial) would allow the set S of Remark 2.3 to satisfy (H7) and not require any change in the convergence arguments to follow.

In what follows we will simplify notation by abbreviating $X^N \equiv X^N(q^N)$, $A^N(t) \equiv A^N(q^N, t)$, $A(t) \equiv A(\tilde{q}, t)$, $P^N \equiv P^N(q^N)$, $i^N \equiv i(q^N)$, $\pi^N \equiv \pi^N(q^N)$, $C_N^+ \equiv C^+(q^N)$, and $|\cdot|_N \equiv |\cdot|_{\rho,q^N}$. We shall also use $|\cdot|_N$ to denote the $L_2^n(-r_\nu^N, 0)$ norm weighted with $\rho(q^N)$. We remind the reader that $|\cdot|$ denotes either the Z or $L_2^n(-r, 0)$ norm while $|\cdot|_{r_\nu^N}$ denotes the unweighted $L_2^n(-r_\nu^N, 0)$ norm. When no confusion results we shall also write $z(t)$ instead of $z(t; \tilde{\gamma}, g)$ and $z^N(t)$ for $z^N(t; \gamma^N, g)$, the solutions to (2.6) and (3.1) associated with $\tilde{\gamma}$ and γ^N, respectively.

For q given in Q, define $\mathcal{I}(q) = \{\{\zeta, g\} \in W \times L_2^n(a, b) | \zeta = (\psi(0), \psi), \psi \in H^2(-r, 0), g \in H^1(a, b), \dot{\psi}(0) = F(q, a, \psi(0), \psi) + g(a)\}$ and define $\mathcal{S} = \{(\psi(0), \psi) \in Z | \psi \in H^2(-r, 0)\}$.

Lemma 3.2. For any $q \in Q$, $\mathcal{I}(q)$ is dense in $W \times L_2^n(a, b)$ (in the $Z \times L_2^n(a, b)$ topology). Furthermore, if $\{\zeta, g\} \in \mathcal{I}(q)$, then the solution $z(t;(\zeta,q),g) = (x(t;(\zeta,q),g), x_t((\zeta,q),g))$ to (2.6) corresponding to ζ, q, g, satisfies $z(t) \in \mathcal{S}$ for all $t \in [a, b]$.

Proof: Let $q \in Q$ and $\zeta = (\psi(0), \psi)$ be fixed in S and define $I(q, \zeta) = \{g \in L_2^n(a, b) | g \in H^1(a, b),\ g(a) = \dot{\psi}(0) - F(q, a, \psi(0), \psi)\}$. Then for $g \in L_2^n(a, b)$ given and $\varepsilon > 0$, standard arguments may be used to construct a $\hat{g}$ that is piecewise-$C^{(1)}$ satisfying $\hat{g}(a) = \dot{\psi}(0) - F(q, a, \psi(0), \psi)$ with $|g - \hat{g}| < \varepsilon$. That is, $I(q, \zeta)$ is dense in $L_2^n(a, b)$. Furthermore, for $\zeta \in S$, the pair $\{\zeta, g\}$ belongs to $I(q)$ whenever $g \in I(q, \zeta)$, so that

$$\bigcup_{\zeta \varepsilon S}[\{\zeta\} \times I(q, \zeta)] \subseteq I(q) \subseteq W \times L_2^n(a, b)$$

where the first set is dense in the last since S is dense in W. It follows that $I(q)$ is dense in $W \times L_2^n(a, b)$.

Required for the proof of the second part of the theorem is a verification that $\ddot{x} \in L_2(a - r, b)$ (since $\mathcal{D}^2(x_t) = (\ddot{x})_t$ for $t \in [a, b]$). If $\{\zeta, g\} \in I(q)$, $\zeta = (\eta, \phi)$, it follows that $x \in C^{(1)}[a - r, b]$ since: (1) $\dot{x}(t) = \dot{\phi}(t - a)$ for $t \in [a - r, a)$; (2) for $t \in (a, b)$, $\dot{x}(t) = f(\alpha, r_\nu, t, x(t), x_t, x(t - r_1), \ldots, x(t - r_\nu)) + g(t)$, which is continuous from assumption (H2) and the definition of $I(q)$; and (3) $\dot{x}(a^-) = \dot{\phi}(0) = F(q, a, \phi(0), \phi) + g(a) = \dot{x}(a^+)$. Further, the differentiability of f and g yields

$$\begin{aligned}\ddot{x}(t) = {} & f_\sigma(\alpha, r_\nu, t, x(t), x_t, \ldots) + f_\xi(\alpha, r_\nu, t, x(t), x_t, \ldots)\,\dot{x}(t) \\ & + f_\psi[\alpha, r_\nu, t, x(t), x_t, \ldots; \dot{x}_t] \\ & + \sum_{i=1}^{\nu} f_{y_i}(\alpha, r_\nu, t, x(t), x_t, \ldots)\dot{x}(t - r_i) + \dot{g}(t)\end{aligned}$$

for $t \in (a, b)$, where f_δ denotes the Frechet derivative of $f(\alpha, r_\nu, \sigma, \xi, \psi, y_1, \ldots, y_\nu)$ with respect to δ, $\delta = \sigma, \xi, \ldots, y_\nu$. The global Lipschitz condition on f ensures that these derivatives (excluding f_σ) are bounded, so that, for almost all $t \in (a, b)$,

$$|f_\xi(\alpha, r_\nu, t, x(t), x_t, \ldots)\dot{x}(t)| \leq m_1(t)|\dot{x}(t)|,$$

$$|f_\psi[\alpha, r_\nu, t, x(t), x_t, \ldots; \dot{x}_t]| \leq m_1(t)|\dot{x}_t|,$$

and similarly for f_{y_i}. Therefore,

$$|\ddot{x}(t)| \leq |f_\sigma(\alpha, r_\nu, t, x(t), x_t, \ldots)| + cm_1(t) + |\dot{g}(t)|$$

almost everywhere on (a, b), where m_1, $\dot{g} \in L_2^n(a, b)$, and c is a constant. Using (H3) we thus obtain that $\ddot{x} \in L_2^n(a,b)$ and it hence follows that $\ddot{x} \in L_2^n(a - r, b)$ since $\ddot{x}(t) = \ddot{\phi}(t - a)$, $t \in (a - r, a)$, $\ddot{\phi} \in L_2^n(- r, 0)$.

$\square$

Essential to our convergence proofs are certain standard estimates from the theory of spline approximations, in particular the Schmidt inequality and Theorem 2.5 from [16]. The inequalities are stated in the next lemma.

<u>Lemma 3.3</u>. Let $z = (\psi(0), \psi)$ be given in S, and denote by $(\psi^N(0), \psi^N)$ the element $P^N z$ of X^N. Then the following estimates may be obtained for N sufficiently large:

$$(3.6) \qquad |P^N z - z|_N \leq \frac{k_1}{N^2} |\mathcal{D}^2 \psi|$$

$$(3.7) \qquad |\mathcal{D}\psi^N - \mathcal{D}\psi|_N \leq \frac{k_2}{N} |\mathcal{D}^2 \psi|$$

$$(3.8) \qquad |\psi^N(\Theta) - \psi(\Theta)| \leq \left(\frac{k_1}{N^2} + \frac{r^{1/2} k_2}{N}\right) |\mathcal{D}^2 \psi|, \ \Theta \in [-r_\nu^N, 0]$$

where k_1 and k_2 are positive constants independent of N and q^N.

Proof:

$$|P^N z - z|_N = |\pi^N i^N z - i^N z|_N$$

$$\leq |z_I^N - i^N z|_N$$

$$= |\psi_I^N - \psi|_N$$

where $z_I^N = (\psi_I^N(0), \psi_I^N)$, ψ_I^N the interpolating spline for $\psi \in H^2(-r, 0)$ with knots $\{t_j^N\}$. From [4, (6.10)] we know

$$|\psi_I^N - \psi|_N \leq \frac{r^2 \nu^{1/2}}{\pi^2 N^2} |\mathcal{D}\psi|$$

so (3.6) obtains. The calculations for the estimates in (3.7), (3.8) are found in [4, pp. 814-15].

$\square$

These estimates may now be employed to show convergence of $z^N(t;\gamma^N, g)$ to $z(t;\tilde{\gamma}, g)$ (in the proper sense) when $z(t)\in S$; i.e., when $\{\tilde{\zeta}, g\}\in I(\tilde{q})$.

<u>Theorem 3.2</u>. Let $\{\gamma^N\}$ be arbitrary in Γ with $\gamma^N \to \tilde{\gamma}$, $\gamma^N = (\zeta^N, q^N)$, $\tilde{\gamma} = (\tilde{\zeta}, \tilde{q}) \in \Gamma$, where $\{\tilde{\zeta}, g\}\in I(\tilde{q})$, and let $z^N(t;\gamma^N, g)$, $z(t;\tilde{\gamma}, g)$ denote the solutions to (3.1) and (2.6) associated with γ^N and $\tilde{\gamma}$ respectively. Then

$$|z^N(t;\gamma^N, g) - P^N z(t;\tilde{\gamma}, g)|_N \to 0$$

as $N\to\infty$ uniformly in $t\in[a, b]$.

Proof: We have $\quad \Delta^N(t) \equiv z^N(t) - P^N z(t)$

$$= P^N\zeta^N + \int_a^t \{A^N(\sigma)z^N(\sigma) + P^N G(\sigma)\}d\sigma$$

$$- P^N\tilde{\zeta} - \int_a^t \{P^N A(\sigma)z(\sigma) + P^N G(\sigma)\}d\sigma$$

$$= \Delta^N(a) + \int_a^t \{A^N(\sigma)z^N(\sigma) - P^N A(\sigma)z(\sigma)\}d\sigma$$

so that from Lemma 2.1 we obtain

$$|\Delta^N(t)|_N^2 = |\Delta^N(a)|_N^2$$

$$+ 2\int_a^t \langle A^N(\sigma)z^N(\sigma) - A^N(\sigma)P^N z(\sigma), \Delta^N(\sigma)\rangle_N d\sigma$$

$$+ 2\int_a^t \langle A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma), \Delta^N(\sigma)\rangle_N d\sigma$$

$$\leq |\Delta^N(a)|_N^2 + 2\int_a^t \omega(\sigma)|\Delta^N(\sigma)|_N^2 d\sigma$$

$$+ 2\int_a^t |A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma)|_N |\Delta^N(\sigma)|_N d\sigma$$

$$\leq |\Delta^N(a)|_N^2 + \int_a^t (2\omega(\sigma) + 1)|\Delta^N(\sigma)|_N^2 d\sigma$$

$$+ \int_a^b |A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma)|_N^2 d\sigma.$$

Gronwall's inequality may be employed (since the L_1 function $2\omega + 1$ is positive and Δ^N is continuous in t) to obtain

$$|\Delta^N(t)|_N^2 \leq (\varepsilon_1(N) + \varepsilon_2(N)) \exp \int_a^b (2\omega(\sigma) + 1)d\sigma$$

where

$$\varepsilon_1(N) = |\Delta^N(a)|_N^2$$

$$\varepsilon_2(N) = \int_a^b |A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma)|_N^2 d\sigma.$$

It remains to show that $\varepsilon_i(N) \to 0$ as $N \to \infty$; that the convergence is uniform in t is readily seen. First,

$$\varepsilon_1(N) = |P^N\zeta^N - P^N\tilde{\zeta}|_N \leq |i^N\zeta^N - i^N\tilde{\zeta}|_N$$

converges to 0 as $N \to \infty$ from the definition of convergence of γ^N to $\tilde{\gamma}$. We will also obtain $\varepsilon_2(N) \to 0$ once we demonstrate the dominated convergence of

$$|A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma)|_N^2 \to 0.$$

Let $z(\sigma) = (y_\sigma(0), y_\sigma)$, $P^N z(\sigma) = (y_\sigma^N(0), y_\sigma^N)$, $y_\sigma^N \in L_2^n(-r_\nu^N, 0)$. Then

$$|A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma)|_N^2$$

$$= |\pi^N A(q^N, \sigma)\pi^N P^N z(\sigma) - \pi^N i^N A(\tilde{q}, \sigma)z(\sigma)|_N^2$$

$$= |\pi^N(F(q^N, \sigma, y_\sigma^N(0), y_\sigma^N), \mathcal{D}y_\sigma^N)$$

$$- \pi^N(F(\tilde{q}, \sigma, y_\sigma(0), y_\sigma), \mathcal{D}y_\sigma)|_N^2$$

$$\leq |F(q^N, \sigma, y_\sigma^N(0), y_\sigma^N) - F(\tilde{q}, \sigma, y_\sigma(0), y_\sigma)|^2$$

$$+ |\mathcal{D}y_\sigma^N - \mathcal{D}y_\sigma|_N^2$$

$$\equiv T_1^N(\sigma) + T_2^N(\sigma),$$

where y_σ^N must be extended to all of $[-r, 0]$ (by defining it to be zero off $[-r_\nu^N, 0]$) before F is evaluated at y_σ^N. From (3.7), $T_2^N(\sigma) \to 0$ as $N \to \infty$. Furthermore,

$$(T_1^N(\sigma))^{1/2} \leq |F(q^N, \sigma, y_\sigma^N(0), y_\sigma^N) - F(q^N, \sigma, y_\sigma(0), y_\sigma)|$$

$$+ |F(q^N, \sigma, y_\sigma(0), y_\sigma) - F(\tilde{q}, \sigma, y_\sigma(0), y_\sigma)|$$

$$\equiv \tau_1^N(\sigma) + \tau_2^N(\sigma)$$

where $\tau_2^N(\sigma) \to 0$ as $N \to \infty$ since quite standard arguments (recall (H4)) may be used to show that the map

$$q = (\alpha, r_1, \ldots, r_\nu) \to F(q, \sigma, \psi(0), \psi)$$

$$= f(\alpha, r_\nu, \sigma, \psi(0), \psi, \psi(-r_1), \ldots, \psi(-r_\nu))$$

is continuous whenever ψ is continuous. In addition, $\tau_1^N(\sigma)$ is $O(\frac{1}{N})$ (for almost all σ) from (3.6), (3.8) since, for almost all σ,

$$|F(q^N, \sigma, y_\sigma^N(0), y_\sigma^N) - F(q^N, \sigma, y_\sigma(0), y_\sigma)|$$

$$\leq m_1(\sigma)\{|y_\sigma^N(0) - y_\sigma(0)| + |y_\sigma^N - y_\sigma|_N + \sum_{i=1}^{\nu} |y_\sigma^N(-r_i^N) - y_\sigma(-r_i^N)|\}$$

and $|y_\sigma^N - y_\sigma|_N \leq |P^N z(\sigma) - z(\sigma)|_N$. Therefore, for almost all $\sigma \in [a, b]$, $T_1^N(\sigma) \to 0$ as $N \to \infty$, and the convergence (a.e.) to zero of the integrand of $\varepsilon_2(N)$ is assured. Dominated convergence follows from similar arguments:

$$|A^N(\sigma)P^N z(\sigma) - P^N A(\sigma)z(\sigma)|_N^2$$

$$\leq (\tau_1^N(\sigma) + \tau_2^N(\sigma))^2 + T_2^N(\sigma)$$

as before where, from (3.7), $\quad T_2^N(\sigma) \leq k_2^2|\mathcal{D}^2 y_\sigma|^2 \leq k_2^2 M_0 < \infty$,

$$M_0 = \sup_{\sigma\varepsilon[a,b]} |\mathcal{D}^2 y_\sigma|^2 = \sup_{\sigma\varepsilon[a,b]} \int_{\sigma-r}^{\sigma} |\ddot{y}(\theta)|^2 d\theta \leq \int_{a-r}^{b} |\ddot{y}(\theta)|^2 d\theta = |\ddot{y}|^2_{L_2^n(a-r,\, b)} < \infty$$

(we have made use of Lemma 3.2 to assert that $z(\sigma) \in S$ for all σ; i.e., $y \in H^2(a - r, b)$). The Lipschitz condition on f and estimates (3.6), (3.8) may be used to show

$$\tau_1^N(\sigma) \leq cm_1(\sigma)M_0$$

for a constant $c > 0$ and almost all σ. Finally,

$$\tau_2^N(\sigma) = |F(q^N, \sigma, y(\sigma), y_\sigma) - F(\tilde{q}, \sigma, y(\sigma), y_\sigma)|$$

$$\leq 2 \sup_{(q,\sigma)\varepsilon Q\times[a,b]} |F(q, \sigma, y(\sigma), y_\sigma)|$$

where y was determined by a fixed $\tilde{q} \varepsilon Q$ and is thus independent of q and is continuous. Again the continuity of $F(q, \sigma, y(\sigma), y_\sigma)$ in (q, σ) may be easily established, (q, σ) in the compact set $Q \times [a, b]$, so that there is some (q^*, σ^*) in $Q \times [a, b]$ such that

$$\tau_2^N(\sigma) \leq 2 \ |F(q^*, \sigma^*, y(\sigma^*), y_{\sigma^*})|$$

$$\equiv M_1 < \infty.$$

It follows then that, for almost all $\sigma \varepsilon [a, b]$,

$$|A^N(\sigma)P^N z(\sigma) - P^N A(\sigma) z(\sigma)|_N^2$$

$$\leq (cm_1(\sigma)M_0 + M_1)^2 + k_2^2 M_0 \equiv h(\sigma)$$

where $h \varepsilon L_1(a, b)$ (since $m_1 \varepsilon L_2(a, b)$). The theorem thus obtains. $\square$

We now turn to the main state variable convergence result for arbitrary $\{\tilde{\zeta}, g\} \varepsilon W \times L_2^n(a, b)$; it contains the key arguments needed to prove Theorem 3.4 below which describes how solutions $\bar{\gamma}^N$ (to P_N) converge to $\bar{\gamma}$, a solution to the original parameter estimation problem.

<u>Theorem 3.3</u>. Suppose $\gamma^N \to \tilde{\gamma}$ where $\gamma^N = (\zeta^N, q^N)$ and $\tilde{\gamma} = (\tilde{\zeta}, \tilde{q})$ are arbitrary in Γ. Then for any $g \varepsilon L_2^n(a, b)$,

$$\pi_0 z^N(t;\ \gamma^N,\ g) \to \pi_0 z(t;\ \tilde{\gamma},\ g)$$

as $N \to \infty$ uniformly in $t \in [a, b]$.

Proof:

$$|\pi_0 z^N(t;\ \gamma^N,\ g) - \pi_0 z(t;\ \tilde{\gamma},\ g)|$$

$$\leq |\pi_0 z^N(t;\ \gamma^N,\ g) - \pi_0 P^N z(t;\ \tilde{\gamma},\ g)|$$

$$+ |\pi_0 P^N z(t;\ \tilde{\gamma},\ g) - \pi_0 z(t;\ \tilde{\gamma},\ g)|$$

$$\equiv T_1^N(t) + T_2^N(t)$$

where $T_2^N(t) \to 0$ as $N \to \infty$ uniformly in $t \in [a, b]$ from the convergence $\pi_0 P^N z \to \pi_0 z$, $z \in Z$, demonstrated in [4, p. 814]. (Uniformity here is due to the fact that $z \in \{z(t;\ \tilde{\gamma},\ g) | t \in [a, b]\}$, a compact set in Z). Further, since $I(\tilde{q})$ is dense in $W \times L_2^n(a, b)$, a pair $\{\hat{\zeta},\ \hat{g}\}$ may be chosen in $I(\tilde{q})$ arbitrarily close to $\{\tilde{\zeta},\ g\}$ so that, given that

$$T_1^N(t) \leq |z^N(t;\ (\zeta^N,\ q^N),\ g) - P^N z(t;\ (\tilde{\zeta},\ \tilde{q}),\ g)|_N$$

$$\leq |z^N(t;\ (\zeta^N,\ q^N),\ g) - z^N(t;\ (C_N^+ i^N \hat{\zeta},\ q^N),\ \hat{g})|_N$$

$$+ |z^N(t;\ (C_N^+ i^N \hat{\zeta},\ q^N),\ \hat{g}) - P^N z(t;\ (\hat{\zeta},\ \tilde{q}),\ \hat{g})|_N$$

$$+ |P^N z(t;\ (\hat{\zeta},\ \tilde{q}),\ \hat{g}) - P^N z(t;\ (\tilde{\zeta},\ \tilde{q}),\ g)|_N\ ,$$

the first and third terms may be made as small as desired from the continuous dependence of z^N, z on $\{i^N\zeta,\ g\} \in X(q^N) \times L_2^n(a, b)$ and $\{\zeta,\ g\} \in Z \times L_2^n(a, b)$ respectively, uniform in N and t (we may use this result for the first term since $|\zeta^N - C_N^+ i^N \hat{\zeta}|_{X,q} = |i^N \zeta^N - i^N \hat{\zeta}|_{X,q} \leq |i^N \zeta^N - i^N \tilde{\zeta}|_N + |i^N \tilde{\zeta} - i^N \hat{\zeta}|_N$ is arbitrarily small from the convergence of ζ^N to $\tilde{\zeta}$). Finally, the middle term goes to 0 uniformly in $t \in [a, b]$ as $N \to \infty$ since $\{\hat{\zeta}, \hat{g}\} \in I(\tilde{q})$ and the parameters $(C_N^+ i_N \hat{\zeta}, q^N)$ involved converge to $(\hat{\zeta}, \tilde{q})$ in the sense required ($|\tilde{q} - q^N| \to 0$ and $|i_N \hat{\zeta} - i_N(C_N^+ i_N \hat{\zeta})|_N = 0$), so that we are quaranteed the uniform convergence of $\pi_0 z^N(t; \gamma^N, g)$ to $\pi_0 z(t; \tilde{\gamma}, g)$.

$\square$

§3.2 Convergence of Parameters

Our attention to this point has been focused on the convergence of solutions z^N (to (3.1)) to the solution z (to (2.6)), once the convergence of any sequence of parameters has been established. In reality, though, we have yet to determine that any sequence of *solutions* $\{\bar{\gamma}^N\}$ *to* P_N *is in fact convergent;* even then, we must prove that the limiting value $\bar{\gamma}$ is indeed a solution to the original parameter identification problem. The result we now state addresses this question and indicates when an approximate ID problem P_N may be used to compute numerical solutions for the original problem.

Theorem 3.4. Let $\{\bar{\gamma}^N\}$, $\bar{\gamma}^N \in \Gamma$, be a sequence of solutions to the approximate parameter estimation problems P_N. Then there exists $\bar{\gamma} \in \Gamma$ and a subsequence $\{\bar{\gamma}^{N_k}\}$ such that $\bar{\gamma}^{N_k} \to \bar{\gamma}$ and, if $\bar{\gamma}$ and S satisfy hypotheses (H6) and (H7), $\bar{\gamma}$ is a solution to the original parameter identification problem.

Proof: From (H7), the sequence $\{C_N^+ i_N \bar{\zeta}^N\}$ belongs to S and S is compact in the Z topology, so that a subsequence satisfies $|C_{N_k}^+ i_{N_k} \bar{\zeta}^{N_k} - \bar{\zeta}| \to 0$ for some $\bar{\zeta} \in S$. The compactness of Q guarantees the convergence of a subsequence of $\{\bar{q}^{N_k}\}$,

$\bar{q}^{N_{k_j}} \to \bar{q}$ for some $\bar{q} \in Q$. Relabelling as $\bar{\gamma}^{N_k}$, we have a sequence $\bar{\gamma}^{N_k} = (\bar{\zeta}^{N_k}, \bar{q}^{N_k})$ in Γ that converges to $\bar{\gamma} = (\bar{\zeta}, \bar{q})$ in the required sense because $\bar{q}^{N_k} \to \bar{q}$ and

$$|i_{N_k} \bar{\zeta}^{N_k} - i_{N_k} \bar{\zeta}|_{N_k} \le \nu^{1/2} |i_{N_k} C^+_{N_k} i_{N_k} \bar{\zeta}^{N_k} - i_{N_k} \bar{\zeta}|_{X,q}$$

$$\le \nu^{1/2} |C^+_{N_k} i_{N_k} \zeta^{N_k} - \bar{\zeta}| \to 0.$$

It remains to show that $\bar{\gamma}$ is a solution to the original ID problem. We have (see (2.8) and J^N in the definition of P_N)

$$J(\bar{\gamma}) = \frac{1}{2} \sum_{i=1}^{M} |C(\bar{q})\pi_0 z(t_i; \bar{\gamma}, g) - \hat{u}_i|^2$$

$$= \lim_{N_k \to \infty} J^{N_k}(\bar{\gamma}^{N_k})$$

$$\le \lim_{N_k \to \infty} J^{N_k}(\gamma)$$

where the continuity of C and the convergence of $\pi_0 z^{N_k}(t; \bar{\gamma}^{N_k}, g)$ to $\pi_0 z(t; \bar{\gamma}, g)$ is used to obtain the second expression and the final inequality holds for any $\gamma \in \Gamma$ since $\bar{\gamma}^{N_k}$ is a solution to P_{N_k}. The convergence of $\pi_0 z^{N_k}(t; \gamma, g)$ to $\pi_0 z(t; \gamma, g)$ for any $\gamma \in \Gamma$ also follows from Theorem 3.3, so it follows that $J^{N_k}(\gamma) \to J(\gamma)$ as $N_k \to \infty$, or that

$$J(\bar{\gamma}) \le J(\gamma)$$

for any $\gamma \in \Gamma$. Thus $\bar{\gamma}$ is a solution to the original identification problem.

$\square$

§4 Numerical Results

In this concluding section we present a sample of numerical findings obtained using the spline approximation estimation schemes discussed above. The test examples we investigated were chosen with certain types of applications and/or difficulties in mind. Example 4.1 deals with a nonlinear pendulum (small oscillations are not assumed) with damping through a linear feedback on the velocity., i.e., $U(\dot{x}) = k\dot{x}$. We assume the existence of accuator delays in effecting the feedback laws. (Delayed damping and delayed restoring forces are quite common in mechanical systems -- see Chapter 21 of [15].) A possible application is associated with the design of a damped "pendulum" to "track" a given course or program $\hat{x}(t)$. Example 4.2 involves a nonlinear nonautonomous multiple delay equation in which the nonlinearity is of the Michaelis-Menten, Briggs-Haldane velocity approximation type. Such nonlinearities occur in biological applications in which enzyme mediated reactions must be modeled. Our third example concerns a linear multiple delay system with unknown coefficients such as might arise in multi-compartment transport models, while Example 4.4 contains a nonlinearity that is only locally Lipschitz and thus it does not satisfy the hypotheses detailed above. It is interesting (although not at all surprising) to observe that the methods under investigation also perform admirably when applied to examples of this type.

The computations reported below were performed on the IBM 370/158 at Brown University. The goal of our numerical efforts was to test convergence properties of the estimation algorithm on selected examples. This was done in the following manner. "True" values of the parameters to be estimated were chosen and an independent method was used to integrate the systems with these values. These "exact" solutions or these solutions with random noise added were used as observed "data" (a number of "sample" data points were chosen) and the spline-based methods were employed with a least squares criterion. For a given N, an IMSL package (ZXSSQ) for the Levenberg-Marquardt method was used to iteratively find the corresponding parameters.

Example 4.1. (Nonlinear pendulum with delayed damping).

We consider the system

$$\ddot{x}(t) + k\dot{x}(t - r) + (g/\ell)\sin x(t) = 0 \quad , \quad 0 \le t \le 7 ,$$

$$x(\theta) = 1, \quad \theta \le 0,$$

$$\dot{x}(\theta) = 0, \quad \theta \le 0.$$

"Data" consisting of 28 sample points at times in [0, 7] were generated for "true" values $\bar{r} = 2$, $\bar{k} = 4$, and $\overline{g/\ell} = 9.81$. Several different estimation problems were investigated.

(a) We seek to estimate r with $k = \bar{k}$, $g/\ell = \overline{g/\ell}$ given (start-up value: $r^o = 2.5$). We denote by $\bar{r}^N$ the "converged" values for r corresponding to a fixed value N of the approximation index.

N	$\bar{r}^N$
2	2.429
4	2.412
8	1.908
16	2.003
32	2.002

(b) We estimate r, g/ℓ with $k = \bar{k}$ given (start-up values: $r^o = 2.2$, $(g/\ell)^o = 8.6$). For N = 16, we obtained $\bar{r}^{16} = 2.002$, $\overline{g/\ell}^{16} = 9.84$.

(c) We estimate r, k with $g/\ell = \overline{g/\ell}$ given (start-up values: $r^o = 2.5$, $k^o = 8.0$). For N = 16, we obtained $\bar{r}^{16} = 1.999$, $\bar{k}^{16} = 3.977$.

Example 4.2.

The nonlinear nonautonomous multiple delay equation for consideration is

$$\dot{x}(t) = -tx(t) + 2x(t - r_1) + \frac{3x(t - r_2)}{K + x(t - r_2)}, \quad 0 \le t \le 4,$$

$$x(\theta) = \begin{cases} -m\theta, & -2 \le \theta \le 0, \\ 20 + m\theta, & -4 \le \theta \le -2. \end{cases}$$

"Data" were generated for 16 sampling times in [0, 4] using true values $\bar{r}_1 = 1$, $\bar{r}_2 = 2$, $\bar{K} = 10$, $\bar{m} = 5$. The following problems were studied and results obtained.

(a) We estimate r_1, r_2 with $K = \bar{K}$, $m = \bar{m}$ (start-up values: $r_1^o = .5$, $r_2^o = 2.5$).

N	$\bar{r}_1^N$	$\bar{r}_2^N$
2	1.055	1.600
4	1.013	1.896
8	1.007	1.943
16	.9995	2.003
32	.9998	2.003

(b) We estimate K for $r_1 = \bar{r}_1$, $r_2 = \bar{r}_2$, $m = \bar{m}$ (start-up value: $K^o = .05$).

N	$\bar{K}^N$
2	8.345
4	9.706
8	9.816
16	10.027
32	9.9998

(c) We estimate m for $r_1 = \bar{r}_1$, $r_2 = \bar{r}_2$, $K = \bar{K}$ (start-up value: $m^o = -4.0$).

N	$\bar{m}^N$
2	5.114
4	5.028
8	5.014
16	4.998
32	4.999

(d) We repeat the calculations of (c) except we corrupt the data with random noise (Gaussian with zero mean and standard deviation $\sigma = .1$).

N	$\bar{m}^N$
2	5.059
4	4.973
8	4.956
16	4.940
32	4.940

<u>Example 4.3</u>

We consider next the linear multiple delay example

$$\dot{x}(t) = -\frac{1}{2}x(t) + \beta x(t - r_1) + x(t - r_2), \quad 0 \le t \le 3,$$

$$x(\theta) = \alpha\theta^2 - 3\theta, \quad -4 \le \theta \le 0 .$$

"True" values of $\bar{\beta} = 3$, $\bar{r}_1 = 1$, $\bar{r}_2 = 2$, $\bar{\alpha} = -.75$ were used to produce 24 data points on the interval [0, 3].

(a) We estimate α for $\beta = \bar{\beta}$, $r_1 = \bar{r}_1$, $r_2 = \bar{r}_2$, (start-up value: $\alpha^o = 5.0$).

N	$\bar{\alpha}^N$
2	-.661
4	-.724
8	-.742
16	-.748
32	-.749

(b) We estimate r_1, r_2, β with $\alpha = \bar{\alpha}$ (start-up values: $r_1^o = 1.3$, $r_2^o = 1.7$, $\beta^o = 3.5$).

N	$\bar{r}_1^N$	$\bar{r}_2^N$	$\bar{\beta}^N$
2	1.1233	1.600	3.1642
4	1.0028	1.957	3.0323
8	.9993	2.009	3.0064
16	.9996	2.005	3.0007
32	.9998	2.002	3.0000

(c) We repeat the calculations of (b) with data that has been corrupted by noise.

N	$\bar{r}_1^N$	$\bar{r}_2^N$	$\bar{\beta}^N$
2	1.096	1.600	3.152
4	.9998	1.970	3.023
8	.9940	2.024	2.994
16	.9934	2.025	2.987
32	.9941	2.023	2.987

Example 4.4.

As our final example, we present a multiple delay equation with nonlinearity satisfying only a local Lipschitz condition.

$$\dot{x}(t) = -1.5\, x(t) - 1.25\, x(t - r_1) + cx(t - r_2) \sin x(t - r_2),\quad 0 \le t \le 5,$$

$$x(\theta) = 10\theta + 1, \quad \theta \le 0.$$

True values were $\bar{c} = 1$, $\bar{r}_1 = 1$, $\bar{r}_2 = 2$, and data were generated corresponding to 20 sampling times in [0, 5]. We estimated r_1, r_2, c with start-up values of $r_1^o = 1.4$, $r_2^o = 2.2$, $c^o = .2$.

N	$\bar{r}_1^N$	$\bar{r}_2^N$	$\bar{c}^N$
2	1.0814	1.9863	1.0606
4	1.0537	1.9900	.9757
8	.9998	1.9906	.9745
16	.9992	1.9993	.9981
32	.9996	1.9995	.9986

§5 Appendix

$\lvert\cdot\rvert$	standard norm on R^n, $L_2^n(-r, 0)$, or more generally $L_2^n(a, b)$, or on $Z = R^n \times L_2^n(-r, 0)$
$\lvert\cdot\rvert_{r_\nu}$	standard norm on $L_2^n(-r_\nu, 0)$
$\lvert\cdot\rvert_q$	$\tilde{\rho}(q)$ weighted norm on Z
$\lvert\cdot\rvert_{X,q}$	standard norm on $X(q) = R^n \times L_2^n(-r_\nu, 0)$
$\lvert\cdot\rvert_{\rho,q}$	$\rho(q)$ weighted norm on X(q)
$\lvert\cdot\rvert_N$	$\rho(q^N)$ weighted norm on either $L_2^n(-r_\nu^N, 0)$ or $X(q^N)$

References

[1] H. T. Banks, Identification of nonlinear delay systems using spline methods, Proc. International Conf. on Nonlinear Phenomena in the Math Sciences, (Univ. Texas, Arlington, June 16-20, 1980), Academic Press, to appear.

[2] H. T. Banks and J. A. Burns, Hereditary control problems: numerical methods based on averaging approximations, SIAM J. Control and Optimization, 16 (1978), 169-208.

[3] H. T. Banks, J. A. Burns and E. M. Cliff, A comparison of numerical methods for identification and optimization problems involving control systems with delays, Brown Univ. LCDS Tech. Rep. 79-7, Providence, RI, 1979.

[4] H. T. Banks, J. A. Burns and E. M. Cliff, Parameter estimation and identification for systems with delays, SIAM J. Control and Optimization, 19 (1981), 791-828.

[5] H. T. Banks and P. L. Daniel, Parameter estimation of nonlinear non-autonomous distributed systems, Proc. 20th IEEE Conf. on Decision and Control, (San Diego, Dec. 16-18, 1981), 228-232.

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

[7] H. T. Banks and I. G. Rosen, Spline approximations for linear nonautonomous delay systems, ICASE Rep. No. 81-33, Oct. 1981, NASA Langley Research Center, Hampton, Va., to appear in J. Math. Anal. Appl.

[8] V. Barbu, Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff, Leyden, 1976.

[9] P. L. Daniel, Spline approximations for nonlinear hereditary control systems, to appear.

[10] P. L. Daniel, Spline-based approximation methods for the identification and control of nonlinear functional differential equations, Ph.D. Thesis, Brown Univ., Providence, R.I., June, 1981.

[11] L. M. Graves, The Theory of Functions of Real Variables, McGraw-Hill, New York, 1956.

[12] J. K. Hale, Ordinary Differential Equations, Interscience, New York, 1969.

[13] F. Kappel, An approximation scheme for delay equations, Proc. International Conf. on Nonlinear Phenomena in the Math Sciences, (Univ. Texas, Arlington, June 16-20, 1980), Academic Press, to appear.

[14] F. Kappel, Spline approximation for autonomous nonlinear functional differential equations, to appear.

[15] N. Minorsky, Nonlinear Oscillations, Van Nostrand, 1962.

[16] M. H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

FILMED

8-8