NASA Contractor Report 181663

ICASE REPORT NO. 88-25

ICASE

ESTIMATION OF TIME- AND STATE-DEPENDENT DELAYS AND OTHER PARAMETERS IN FUNCTIONAL DIFFERENTIAL EQUATIONS

(NASA-CR-181663) ESTIMATION OF TIME- AND STATE-DEPENDENT DELAYS AND OTHER PARAMETERS IN FUNCTIONAL DIFFERENTIAL EQUATIONS Final Report (NASA) 53 p CSCL 12A N88-24164 Unclas G3/59 0145737

K. A. Murphy

Contract No. NAS1-18107
April 1988

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

Operated by the Universities Space Research Association

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

ESTIMATION OF TIME- AND STATE-DEPENDENT DELAYS AND OTHER PARAMETERS IN FUNCTIONAL DIFFERENTIAL EQUATIONS

K.A. Murphy
Department of Mathematics
University of North Carolina
Chapel Hill, NC 27514

Abstract. We develop a parameter estimation algorithm which can be used to estimate unknown time- or state-dependent delays and other parameters (eg., initial condition) appearing within a nonlinear nonautonomous functional differential equation. The original infinite dimensional differential equation is approximated using linear splines, which are allowed to move with the variable delay. The variable delays are approximated using linear splines as well. The approximation scheme produces a system of ordinary differential equations with nice computational properties. The unknown parameters are estimated within the approximating systems by minimizing a least-squares fit-to-data criterion. We prove convergence theorems for time-dependent delays and state-dependent delays within two classes, which say essentially that fitting the data by using approximations will, in the limit, provide a fit to the data using the original system. We present numerical test examples which illustrate the method for all types of delay.

This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-18107 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665. This research was also supported in part by the Air Force Office of Scientific Research under Contract No. AFOSR-86-0256.

N88-24164

1. Introduction. Delay and functional differential equations (FDEs) can be used to model a variety of phenomena in the physical and natural sciences. Our focus here is on models in which the delays are nonconstant: They are either time- or state-dependent.

A problem from electrodynamics involving state-dependent delays was considered by Driver (e.g., [7] and [8]). In [8] he derives and analyzes systems of equations of the general form:

$$\dot{y}_i(t) = f_i(t, y(t), y(g_2(t, y(t))), \ldots, y(g_m(t, y(t)))) \qquad (i=1,\ldots,n).$$

Such equations can be used (as in [7], for example) to describe the motion of two electrically charged particles. In such a case, the delays are state-dependent because the effect of one particle on another is not felt until the emitted electrical signal has traveled the distance between the two particles; the time lag depends on the distance, which, in turn, depends on the positions, which is (part of) the state of the system.

Another application area rich in FDE models is biology. In [10], for example, a model for an antigen-antibody interaction is discussed, involving a system of functional differential equations. The full model incorporates thresholding effects, which introduces two delays. The model equations are of the form $\dot{x}(t) = f(x(t), x(\omega(t)))$ where $x(t)$ represents the (vector) state of the system, and each retarding function ω satisfies an equation of the form:

$$\int_{\omega(t)}^{t} g(t-s, x(s))\,ds = \theta.$$

Here, the function ω depends explicitly only on time, but is also state-dependent through the above integral equation.

A second example from biology arises in the modeling of insect population growth. Consider an insect which passes through several larval stages (called "instars") during the course of its lifetime. In [17], a model for the population growth of such insects is developed under the assumption that the passage from one instar to the next is contingent upon a certain body weight gain. An alternative assumption (see, e.g., [13]) is that this passage is determined by the exposure of the insect to a sufficient number of degree-days. In this case, one can begin in an analogous way to Nisbet and Gurney in [17], but the final model developed is a system of delay differential equations with time-dependent delays (this model

development will be the subject of a future paper).

Many researchers have investigated the problem of approximating functional differential equations for the purpose of numerical solution (the "forward problem") and also the related problem of estimating unknown parameters, particularly delays, occuring within FDEs (the "inverse" problem). As an example of the former, we mention the work of Feldstein, et. al. (see, e.g., [9], [15], [16]). This work consists of developing accurate, high order methods for the forward problem; they consider, in particular, delay equations in which the delay is state-dependent. One approach to the numerical solution of delay equations is to begin with methods developed for ordinary differential equations, with the addition of an interpolation procedure for the approximation of solution values at lagged times. In the references cited above, efforts focus on determining locations of discontinuities in derivatives of the solution, and exploiting this information in order to retain a higher degree of accuracy in the numerical solution.

In this paper, we are concerned with the parameter estimation problem. In all but a few special cases, one expects to use an iterative (on the unknown parameters) method combined with some numerical approximation scheme for solving the differential equation. In the inverse problem, accuracy is important to some extent, however, as one is in general trying to fit field data, a high degree of accuracy is not as important as a fast and efficient solution algorithm. In the work of Banks, et. al., (see, e.g., [1], [2], [3]), one can find various approximation schemes (involving the use of splines and the "averaging" method); the differential equations are posed in an abstract, operator-theoretic setting, within which convergence results are obtained (i.e., parameters estimated using the approximations converge (subsequentially) to parameters which provide a fit for the original model equations). These results pertain to the estimation of multiple <u>constant</u> delays (and other parameters) in a system of functional differential equations.

In the present paper, we extend the ideas developed in [1] - [3] (primarily [3]), in order to devise an approximation scheme and prove convergence results for the estimation of a nonconstant delay in a nonautonomous, nonlinear functional differential equation. The generalization to a system of equations is straightforward, and is not considered here. We consider the estimation of only a single delay, although it is expected that the estimation of multiple delays should also be

a straightforward modification.

In our presentation here, we begin in section 2 by describing the type of model equations in which we are interested and state the parameter estimation problem precisely. In section 3, we formulate the differential equation abstractly and develop the approximation schemes for the states and the variable delay. We state and prove convergence theorems in section 4, discuss the numerical implementation of the scheme in section 5, and conclude with some numerical test examples in section 6.

2. The Parameter Estimation Problem. We formulate our parameter estimation problem as follows. We will assume that we are interested in modeling some phenomenon for which we have a general form of model equation, but some unknown parameters appear in the equations. We assume further that we have observed our system, and collected data. We then wish to determine the unknown parameters by fitting the model equation to the data. Here, our model equation is a functional differential equation of the general form:

$$\begin{aligned} &\dot{x}(t) = f(t,\gamma,x(t),x(t-\tau),x_t) && 0 \le t \le T \\ &x_0(s) = \phi(s) && -\tau|_{t=0} \le s \le 0 \end{aligned} \tag{2.1}$$

where $x(t) \in \mathbb{R}$, x_t denotes the function $x_t : s \to x_t(s) = x(t+s)$ where s varies over [-r, 0] (unless this interval is otherwise specified), and γ represents a vector of parameters occuring in the model equation (with perhaps some functional components). The unknowns in this equation might be any of γ, τ, or ϕ. If the delay is explicitly (and solely) time-dependent, it is the function $\tau : t \to \tau(t)$ which we assume is unknown and to be estimated. We shall also consider two classes of state-dependent delays. In the first class, we assume the delay is determined through a relationship of the form $\int_{t-\tau(t)}^{t} g(x(s))ds = \theta$; here it is the function g and the constant θ which we shall assume are unknown and to be estimated. In the second class, we assume the delay is of the form $\tau(t) = g(x(t))$, and it is the function $g : x \to g(x)$ which we wish to determine.

Suppose now that our data is given by $\{\chi_i\}_{i=1}^{m}$, where χ_i corresponds to the solution of (2.1), for some value of the parameters, evaluated at t_i, i.e., $x(t_i)$.

We would like to determine the unknowns in (2.1) such that the least squares fit-to-data criterion:

$$J(\rho) = \sum_{i=1}^{m} |x(t_i;\, \rho) - \chi_i|^2$$

is minimized. We use ρ to represent the unknown parameters and we write $x(t;\, \rho)$ to emphasize that the solution to (2.1) depends on these parameters; in the time-dependent delay case, $\rho = (\gamma, \tau, \phi)$, and in the state-dependent delay cases, $\rho = (\gamma, g, \theta, \phi)$ or $\rho = (\gamma, g, \phi)$.

We shall assume that the unknown parameters lie in some constraint set, denoted by $\Pi = \mathcal{G} \times \mathcal{T} \times \mathcal{J}$, for time-dependent delay, or $\Pi = \mathcal{G} \times \overline{\mathcal{T}}_g \times \mathcal{J}$ for state-dependent delays of the first class, or $\Pi = \mathcal{G} \times \mathcal{T}_g \times \mathcal{J}$ for the second class, where $\gamma \in \mathcal{G}$, $\tau \in \mathcal{T}$, $\phi \in \mathcal{J}$, $(g, \theta) \in \overline{\mathcal{T}}_g \equiv \mathcal{T}_g \times \Theta$ and $g \in \mathcal{T}_g$. In the state-dependent delay case, it will be convenient to define the constraint set $\mathcal{T}_g$ in such a way that the resulting τ, as a function of time, will belong to the same constraint set $\mathcal{T}$ which is defined for the time-dependent delay.

First consider the unknown parameters represented by γ. This will, in general, be a vector of unknowns, each component lying in $\mathbb{R}$, or in some function space. For definiteness, let us suppose that γ consists of a vector of ν constants and one functional unknown, so that γ belongs to $\mathbb{R}^{\nu} \times C(0,T)$. The set $\mathcal{G}$ will represent some appropriately chosen compact subset of $\mathbb{R}^{\nu} \times C(0,T)$, for example, we might take

$$\mathcal{G} = \{ (\gamma_1, \gamma_2, \ldots, \gamma_{\nu}, \gamma_0) \in \mathbb{R}^{\nu} \times H^1(0,T) \mid |\gamma_i| \le c_i \text{ for } i = 1, \ldots, \nu \text{ and } |\gamma_0|_{\infty} \le c_0,\ |\dot{\gamma}_0|_0 \le d_0 \}$$

Using the Ascoli Theorem, one can argue that such a choice satisfies

(HG) The set $\mathcal{G}$ is a compact subset of $\mathbb{R}^{\nu} \times C(0,T)$.

We define the sets associated with the delay as

$$\mathcal{T} = \{\tau \in W^{1,\infty}(0,T) \mid \tau_0 \le \tau \le r,\ |\dot{\tau}|_{\infty} \le \mu \quad \text{a.e. } t \in [0,T]\}$$

$$\mathcal{T}_g \equiv \{g \in W^{1,\infty}(\underline{x}, \bar{x}) \mid g_L \le g \le g^U,\ |g(x_1) - g(x_2)| \le \mu_g |x_1 - x_2|\} \quad \text{and}$$

$$\Theta \equiv \{\theta \in \mathbb{R} \mid \underline{\theta} \le \theta \le \bar{\theta}\}$$

where τ_0, r, g_L, g^U, $\underline{\theta}$, $\bar{\theta}$, μ, μ_g are positive constants which are either known or

chosen *a priori*. The constants $\underline{x}$ and $\bar{x}$ represent lower and upper bounds (perhaps only estimates) on the solution. Note that theoretically we could consider g as a function defined on $(-\infty,\infty)$, however when we consider approximating functions for g, it is necessary for our approximation schemes to consider functions defined on a finite interval. We can argue, using properties of our model equations and parameters that all solutions will be bounded, uniformly in the parameter sets. Thus, there do exist numbers $\underline{x}$ and $\bar{x}$ such that we can treat g as being defined only on the interval $[\underline{x},\bar{x}]$. We shall say more later about the choice of these numbers in practice. Again, one can argue using the Ascoli Theorem that each of the sets defined above satisfy

(HT) The set $\mathcal{T}$ is a compact subset of $C(0,T)$.
The set $\mathcal{T}_g$ is a compact subset of $C(\underline{x},\bar{x})$, and the set $\overline{\mathcal{T}}_g$ is a compact subset of $C(\underline{x},\bar{x})\times\mathbb{R}$.

Finally, we will assume the unknown initial data lies within a set $\mathcal{J}$, which is a subset of $H^1(-r,0)$. As we iterate on the unknown parameters (in particular, on τ), we will perform our computations using ϕ as a function defined on $[-\tau|_{t=0},0]$, and for a given $\tau|_{t=0}\le r$, the delay equation (2.1) is well-posed for ϕ defined on $[-\tau|_{t=0},0]$. However, we can consider ϕ as an element of $H^1(-r,0)$ by continuously extending it. We assume that

(HI) The set $\mathcal{J}$ is a compact subset of $C(-r,0)$.

We assume the function f is continuous in all its arguments and satisfies a Lipschitz condition of the following form:

(HF) There is a positive function $m(t)\in L_2(0,T)$ such that for any $\gamma\in\mathcal{G}$, and for any $(\xi_1,\varsigma_1,\eta_1)$, $(\xi_2,\varsigma_2,\eta_2)\in\mathbb{R}\times\mathbb{R}\times L_2(-r,0)$, it follows that
$$|f(t,\gamma,\xi_1,\varsigma_1,\eta_1)-f(t,\gamma,\xi_2,\varsigma_2,\eta_2)|\le m(t)\left\{|\xi_1-\xi_2|+|\varsigma_1-\varsigma_2|+|\eta_1-\eta_2|_0\right\}$$
where $|\cdot|_0$ denotes the norm in $L_2(-r,0)$.

Several authors have investigated questions of existence and uniqueness of

solutions of equation (2.1) (see for example, [8] and the references therein, [11], or [20]). They make varying assumptions on the function f and the delay(s). The assumptions we have made above are sufficient to prove the convergence arguments which are crucial to the parameter estimation problem; this is the problem of concern to us here.

We note that, as with any inverse problem, questions of well-posedness arise, i.e., one questions whether there is a unique solution to the parameter estimation problem (in fact, as will be seen in the final section, examples are easily constructed for which there is not a unique solution unless the problem is slightly reformulated), and how these solutions depend on the data. These are important questions but are outside the scope of this paper.

3. Abstract Formulation and Approximations. We shall begin by rewriting equation (2.1) in abstract form. This provides a framework for our approximation theory and facilitates our convergence arguments. For convenience in notation, let us define $\mathcal{Q} \equiv \mathcal{G} \times \mathcal{T}$ and let $q = (\gamma, \tau)$; it is to be understood that in the case of state-dependent delay, $\tau(\cdot): t \mapsto \tau(t)$ represents the delay obtained from the unknown parameters g, θ or g. Suppose $x(t)$ represents the solution of equation (2.1), defined for all $t \in [0,T]$, given $q \in \mathcal{Q}$ and initial condition $\phi \in \mathfrak{J}$. Just as in [2] and [3], for the constant delay problem, we shall take our state to be $\hat{z}(t) = \begin{bmatrix} x(t) \\ x_t(\cdot) \end{bmatrix}$ with state space $\mathbf{Z} = \mathbb{R} \times L_2(-r,0)$, and we define $\hat{\phi} = \begin{bmatrix} \phi(0) \\ \phi(\cdot) \end{bmatrix} = \begin{bmatrix} x(0) \\ x_0(\cdot) \end{bmatrix}$. Here we are considering $x_t(s)$ defined for all $-r \leq s \leq 0$. We could alternatively, and more economically, define our state space to be $X(\tau(t)) = \mathbb{R} \times L_2(-\tau(t),0)$. This is a state space which we shall use later, but it is more convenient to consider our "true solution" to be in the stationary state space, $\mathbf{Z}$. Given that x is a solution of (2.1), it follows that $\hat{z}$ is a solution of the abstract ordinary differential equation:

$$
\begin{aligned}
\dot{\hat{z}}(t) &= \mathcal{A}(t;q)\hat{z}(t) \\
\hat{z}(0) &= \hat{\phi}
\end{aligned}
\tag{3.1}
$$

where $\mathcal{A}$ is the nonlinear operator defined by $\mathcal{A}(t;q)\begin{bmatrix} z_0 \\ z(\cdot) \end{bmatrix} = \begin{bmatrix} F(t,q,z_0,z) \\ Dz \end{bmatrix}$ with

$\mathfrak{M} \equiv \mathrm{dom}\mathcal{A}(q) = \left\{ \begin{bmatrix} z_0 \\ z(\cdot) \end{bmatrix} \in \mathbb{R} \times H^1(-r,0) \mid z(0) = z_0 \right\}$; we are using D to represent the differentiation operator and we define $F(t,q,z_0,z) = f(t,\gamma,z_0,z(-\tau(t)),z)$.

It should be noted that if we had chosen to define $\hat{z}(t)$ in $X(\tau)$ rather than in $\mathcal{Z}$, this equation would have been written identically, except for a slight change in the definition of $\mathrm{dom}\mathcal{A}$. As it will be necessary to consider functions in both state spaces $\mathcal{Z}$ and $X(\tau)$, we now define extension and restriction operators. Fix $\tau \in \mathcal{T}$. Given $\hat{z} = \begin{bmatrix} z_0 \\ z(\cdot) \end{bmatrix} \in \mathcal{Z}$, let $\iota(\tau)\hat{z}$ be the element of $X(\tau)$ obtained by restricting the function $z \in L_2(-r,0)$ to be a function in $L_2(-\tau,0)$. Similarly, if $\hat{y} = \begin{bmatrix} y_0 \\ y(\cdot) \end{bmatrix} \in X(\tau)$, we define $\epsilon(\tau)\hat{y}$ to be the element of $\mathcal{Z}$ obtained by extending y in a constant, continuous way.

Equations (2.1) and (3.1) are equivalent in the sense that if x is a solution of equation (2.1), then $\hat{z}(t) = (x(t), x_t)^T$ (with x_t defined on $[-r,0]$) is a solution of equation (3.1), and, conversely, if $\hat{z}$ is a solution of (3.1), then $\iota(\tau)\hat{z}$ can be identified with $(x(t), x_t)^T$ with x_t defined on $[-\tau,0]$, where x is a solution of (2.1) (this equivalence can be argued rigorously just as in [3]). Thus, this equivalence can be used to argue that whenever $\hat{z}(0) \in \mathfrak{M}$, then $\hat{z}(t) \in \mathfrak{M}$ for all $t \in [0,T]$.

We shall use the following notation for norms. Absolute value will be denoted by $|\cdot|$, the norm in $L_2(-r,0)$ by $|\cdot|_0$, and the norm in $\mathcal{Z}$ by $\|\cdot\|_0$. Given $y \in L_2(-\tau(t),0)$, we define the norm by $|y|_\tau^2 = \int_{-\tau(t)}^{0} y^2 ds$. The norm on $X(\tau)$ is the usual cross product definition, and will be designated $\|\cdot\|_\tau$.

We now formulate the parameter estimation problem in terms of the abstract equation (3.1) as follows:

$$\text{(P)} \qquad \begin{aligned} &\min_{\rho \in \Pi} J(\rho) = \sum_{i=1}^{m} |z_0(t_i;\rho) - \chi_i|^2 \qquad \text{subject to} \\ &\hat{z}(t;\rho) = \begin{bmatrix} z_0(t;\rho) \\ z(t,\cdot;\rho) \end{bmatrix} \text{ a solution of equation (3.1).} \end{aligned}$$

Problem (P) is clearly an infinite dimensional problem, in that the evaluation of z_0,

and hence of J, requires the solution of the infinite dimensional equation (3.1). Thus, we shall approximate the spaces $\mathbf{Z}$ and $X(\tau)$ by a sequence of finite dimensional approximating subspaces, and develop a sequence of corresponding approximating differential equations.

Our choice of approximation scheme is motivated by that of [2] and [3], but we must modify these ideas to accomodate the time-varying delay. For each N, consider $\tau \in \mathcal{T}$ and fix $t \in [0,T]$. Define the knot sequence $\{k_i(\tau)\}_{i=0}^N$ by $k_i = \frac{-i\tau(t)}{N}$. Let $\{B_i(\tau)\}_{i=0}^N$ be the linear spline basis elements, or "hat" functions (see, e.g., [18]) defined on the τ-dependent grid $\{k_i(\tau)\}_{i=0}^N$. Let $\hat{B}_i = \begin{pmatrix} B_i(0) \\ B_i(\cdot) \end{pmatrix}$ and $X^N(\tau) = \text{span}\{\hat{B}_i\}_{i=0}^N$. Notice that $X^N(\tau) \subset X(\tau)$, but $X(\tau)$ is not technically a subset of $\text{dom}\mathcal{A}(t;q)$; thus we define $\mathbf{X}^N(\tau) = \epsilon(\tau)X^N$, which is a subset of $\text{dom}\mathcal{A}(t;q)$. The sets $X^N(\tau)$ and $\mathbf{X}^N(\tau)$ are both finite dimensional (of the same dimension), and from a computational point of view, they are equivalent. We can think of $\mathbf{X}^N(\tau)$ as the span of $\{\hat{\mathfrak{B}}_i\}$, where we define $\hat{\mathfrak{B}}_i = \epsilon\hat{B}_i$. We define $\mathbf{X}^N(\tau)$ for theoretical reasons, but we perform our computations using elements of $X^N(\tau)$. We define the orthogonal projection (in the $X(\tau)$-topology) for each N and $\tau \in \mathcal{T}$ by $P^N(\tau(t)): X(\tau(t)) \to X^N(\tau(t))$. We further define the operator $\mathcal{P}^N(\tau) = \epsilon(\tau)P^N(\tau)\iota(\tau)$, which is essentially equivalent to P^N, except that it maps elements of $\mathbf{Z}$ to elements of $\mathbf{X}^N \subset \text{dom}\mathcal{A}(t;q)$. Let $\mathcal{A}^N(t;q) = \mathcal{P}^N(\tau)\mathcal{A}(t;q)$. The approximating equations for (3.1) are given by

$$
\begin{aligned}
\dot{\hat{z}}^N(t) &= \mathcal{A}^N(t;q)\hat{z}^N(t) \\
\hat{z}^N(0) &= \mathcal{P}^N(\tau(0))\hat{\phi}
\end{aligned}
\tag{3.2}
$$

where, for each t, $\hat{z}^N(t) \in \mathbf{X}^N(\tau(t))$.

One could formulate a corresponding parameter estimation problem, defined in terms of $\hat{z}^N$ in place of $\hat{z}$; there is still an aspect of infinite dimensionality, however, in that τ (or g), ϕ, and possibly γ are unknown functions. We describe an approximation scheme for τ or g (assuming for ease in presentation of our ideas that ϕ and γ are either simply unknown constants or have known *a priori* parametrizations involving only unknown constants; if this is not the case, then the ideas for τ, g presented here can be easily applied to these parameters as well), and then state our parameter estimation problem in a completely numerically

implementable form. The ideas we use here for approximating the delay are similar to those described in [4], [5], [6] and [12] for the secondary approximations of variable parameters in partial differential equations.

In the case where $\tau = \tau(t)$ is the unknown, we proceed as follows to define the secondary approximations. Given T, the final time of interest in the original delay differential equation, define the knot sequence $\{\kappa_i\}_{i=0}^M$ by $\kappa_i = \frac{iT}{M}$, and let $\{b_i\}_{i=0}^M$ be the hat functions defined for this grid. Let $S^M = \text{span}\{b_i\}$. We define a sequence of finite dimensional sets by $\mathcal{T}^M = \mathcal{T} \cap S^M$. Let I^M represent the interpolation operator from $\mathcal{T}$ to S^M (i.e., given $\tau \in \mathcal{T}$, $I^M\tau$ is the function in S^M which satisfies $(I^M\tau)(\kappa_i) = \tau(\kappa_i)$ for i=0,1,...,M). Then $\mathcal{T}^M$ can be characterized as $\mathcal{T}^M = I^M\mathcal{T}$. Given that $\tau \in \mathcal{T}$, it follows from standard results in the theory of linear splines (see, e.g., [19]) that $I^M\tau \to \tau$ in $C[0,T]$ as $M \to \infty$. Recall (HT) states that the set $\mathcal{T}$ forms a compact subset of $C[0,T]$, and each $\mathcal{T}^M \subset \mathcal{T}$ is also a compact subset of $C[0,T]$ (this follows from the fact that $\mathcal{T}^M$ is a closed subset of $\mathcal{T}$).

In either of the cases of state-dependent delay, one first needs to determine the interval $[\underline{x}, \bar{x}]$ on which to define the approximations. One might begin by inspecting the data, and then choosing the lower and upper bounds accordingly; it is the experience of the author (both for the estimation of delays and also with the estimation of unknown state-dependent parameters in partial differential equations, [6]) that, in practice, it can be determined in the course of the estimation procedure that one has either over- or underestimated the size of this interval. The computer program can be written so that the interval is automatically extended if the approximate state z_0^N takes on a value outside the original interval. If the original interval is chosen to be too large, then one sees convergence of the delay on the interior of the interval, and "chattering" behavior around the endpoint(s) where one is trying to estimate a dependence where none is there to be seen. From a theoretical standpoint, one could equivalently consider the approximating functions, g^M, as being defined on all of $(-\infty, \infty)$ by defining $g^M(x) = g^M(\underline{x})$ for $x \le \underline{x}$ and $g^M(x) = g^M(\bar{x})$ for $x \ge \bar{x}$.

Given the interval $[\underline{x}, \bar{x}]$, define the knot sequence $\{\kappa_i\}_{i=0}^M$ by $\kappa_i = i(\bar{x}-\underline{x})/M$, and let $\{b_i\}_{i=0}^M$ be the hat functions defined for this grid. Let $S^M = \text{span}\{b_i\}$, and $\mathcal{T}_g^M = \mathcal{T}_g \cap S^M$. As above, we then let I^M represent the interpolation operator from $\mathcal{T}_g$ to S^M and notice that $\mathcal{T}_g^M$ can be characterized as $\mathcal{T}_g^M = I^M\mathcal{T}_g$. We can argue

that for any $g \in \mathcal{T}_g$ $I^M g \to g$ in $C[\underline{x}, \bar{x}]$, and that $\mathcal{T}_g^M \subset \mathcal{T}_g$ is a compact subset of $C[\underline{x}, \bar{x}]$. We finally define $\bar{\mathcal{T}}_g^M \equiv \mathcal{T}_g^M \times \Theta$.

With $\mathcal{T}^M$, $\mathcal{T}_g^M$, $\bar{\mathcal{T}}_g^M$ defined as above, let $\Pi^M \equiv \mathcal{G} \times \mathcal{T}^M \times \mathcal{J}$, $\Pi^M \equiv \mathcal{G} \times \bar{\mathcal{T}}_g^M \times \mathcal{J}$, or $\Pi^M \equiv \mathcal{G} \times \mathcal{T}_g^M \times \mathcal{J}$. Note that by assumptions (HG) and (HI) and per the discussion above, the sets Π and Π^M are compact. We now state our approximate parameter estimation problem (for each N,M) as:

$$(P^{N,M}) \qquad \underset{\rho \in \Pi^M}{\text{Min}} \; J^N(\rho) = \sum_{i=1}^{m} |z_0^N(t_i;\rho) - \chi_i|^2 \qquad \text{subject to}$$

$$\hat{z}^N(t;\rho) = \begin{bmatrix} z_0^N(t;\rho) \\ z^N(t,\cdot;\rho) \end{bmatrix} \text{ a solution of equation (3.2).}$$

For each N,M, Problem $(P^{N,M})$ has a solution, $\rho^{N,M} \in \Pi^M$, since each Π^M is compact and J^N is a continuous function of $\rho \in \Pi^M$ (this can be seen from the matrix-vector representations of (3.2) developed in a later section). Moreover, each $\rho^{N,M}$ is in Π, also compact, and thus we conclude there is a convergent subsequence $\{\rho^{N_k,M_k}\}$ such that $\rho^{N_k,M_k} \to \rho^*$ with $N_k, M_k \to \infty$ and $\rho^* \in \Pi$. It is our claim that this limit ρ^* is a solution of Problem (P). In order to verify this claim, we must first show that for an arbitrary convergent sequence of parameters in Π, the states converge as well. This is established in the next section.

4. Convergence Arguments.

In this section we shall prove that a subsequence of the parameter estimates we generate by solving $(P^{N,M})$ converges to a solution of (P). As a first step, we shall demonstrate that convergence of an arbitrary sequence of parameters implies a corresponding convergence of solutions of the approximating differential equations to a solution of the original FDE.

Before we can prove this convergence result, we first establish some preliminary results in the form of lemmas. Throughout this section, we let $\mathcal{T} = \{\tau \in W^{1,\infty}(0,T) \mid 0 < \tau_0 \le \tau(t) \le r \;\; \forall t \in [0,T] \text{ and } |\dot{\tau}|_\infty \le \mu\}$.

Lemma 1. Suppose $\{\tau^M\}$ is an arbitrary collection of functions in $\mathcal{T}$ and $\hat{z}(t) \in \mathbf{Z}$ for each $t \in [0,T]$. Then $\|P^N(\tau^M)\iota(\tau^M)\hat{z}-\iota(\tau^M)\hat{z}\|_{\tau^M} \to 0$ as $N\to\infty$, for each $t \in [0,T]$.

Proof. Let us write $P^{N,M} = P^N(\tau^M)$ and $\iota^M = \iota(\tau^M)$. We begin by assuming that for fixed t, $\hat{z}(t) \in \mathfrak{M}$. We shall use the following notation: Let $\hat{z}(t) = \hat{z} = \begin{bmatrix} z_0 \\ z(\cdot) \end{bmatrix}$, (notice that since $\hat{z}\in\mathfrak{M}$, $z_0=z(0)$), and let $I^{N,M}\iota^M\hat{z} = \begin{bmatrix} z_I^{N,M}(0) \\ z_I^{N,M}(\cdot) \end{bmatrix}$ where $z_I^{N,M}$ denotes the spline interpolant of the function z (actually, of its restriction, considered as a function in $L_2(-\tau^M(t),0)$), using the grid formed by subdividing the interval $[-\tau^M,0]$ into N equal subintervals.

For fixed t, we have

$$\|P^{N,M}\iota^M\hat{z}-\iota^M\hat{z}\|^2_{\tau^M} \le \|I^{N,M}\iota^M\hat{z}-\iota^M\hat{z}\|^2_{\tau^M} = \left[z_I^{N,M}(0)-z(0)\right]^2 + \int_{-\tau^M(t)}^0 (z_I^{N,M}-z)^2\, ds$$

$$= \int_{-\tau^M(t)}^0 (z_I^{N,M}-z)^2\, ds.$$

Now, applying standard estimates (e.g., a slight modification of (2.14) in [19]) and using the fact that each τ^M belongs to $\mathcal{T}$, we obtain

$$\|P^{N,M}\iota^M\hat{z}-\iota^M\hat{z}\|^2_{\tau^M} \le \left[\frac{\tau^M(t)}{N\pi}\right]^2 \int_{-\tau^M(t)}^0 (Dz(t))^2\, ds$$

$$\le \left[\frac{r}{N\pi}\right]^2 \int_{-r}^0 (Dz(t))^2\, ds.$$

It is clear from the above estimate that $\|P^{N,M}\iota^M\hat{z}-\iota^M\hat{z}\|^2_{\tau^M} \to 0$ as $N \to \infty$ for $\hat{z} \in \mathfrak{M}$ and fixed t. Note that this convergence is independent of M as long as for each M, τ^M is in the set $\mathcal{T}$.

In fact the above convergence statement is also true for any $\hat{z}(t) \in \mathbf{Z}$. This follows from the fact that $\mathfrak{M}$ is dense in $\mathbf{Z}$ and that each $P^{N,M}\iota^M$ is linear and uniformly bounded (in N,M). □

We will often be interested in the $X(\tau)$-norm of functions that belong to the spaces $\mathbf{Z}$, $\epsilon(\tau)X(\tau)$, or $\epsilon(\tau)X^N(\tau)$ for some $\tau\in\mathcal{T}$. To be technically correct then, we should use the operator $\iota(\tau)$ to first map the function into the space $X(\tau)$. In order

to reduce the necessary notation, however, we will not write ι when it should be clear from the context that it is needed.

The next lemma tells how to relate the time derivative of a projected function to the projection of that function's derivative. As indicated in an earlier discussion, the fact that these two elements are not equal is one respect in which the time-varying delay problem differs from the constant delay problem.

<u>Lemma 2.</u> Assume $\tau \in \mathcal{T}$ and $\hat{z}(t) \in \mathfrak{M}$ for each fixed t. Recall that $\mathcal{P}^N = \mathcal{P}^N(\tau) = \epsilon(\tau) P^N(\tau)\iota(\tau)$, let $\hat{z}(t) = \begin{pmatrix} z_0 \\ z(\cdot) \end{pmatrix}$, and let us write $P^N(\tau)\iota(\tau)\hat{z} = \begin{pmatrix} P_0^N z \\ P^N z(\cdot) \end{pmatrix}$. Then for any $\hat{y} \in \mathbf{X}^N(\tau)$ with $\hat{y} = \begin{pmatrix} y(0) \\ y(\cdot) \end{pmatrix}$, we have for almost every $t \in [0,T]$:

$$\langle\langle \frac{d}{dt}(\mathcal{P}^N\hat{z}), \hat{y} \rangle\rangle_\tau = \langle\langle \mathcal{P}^N \frac{d\hat{z}}{dt}, \hat{y} \rangle\rangle_\tau + \begin{bmatrix} \dot{\tau} \\ \tau \end{bmatrix} \cdot \left[\int_{-\tau}^{0} s\left[Dz - D(P^N z)\right] y \, ds + \int_{-\tau}^{0} (z - P^N z) y \, ds \right].$$

<u>Proof.</u> We first note that we define $\mathcal{P}^N$ in terms of the extension operator $\epsilon(\tau)$ so that all functions will be defined on the full interval [-r,0]; thus, while $P^N\iota\hat{z} : [0,T] \to X(\tau(t))$, and thus has a "moving range-space", $\mathcal{P}^N\hat{z} : [0,T] \to \mathbf{X}^N \subset \mathbf{Z}$. We introduce the operator $\mathcal{P}^N$ in order to make precise the time derivative $\frac{d}{dt}(\mathcal{P}^N\hat{z})$. We recall that $\mathbf{X}^N = \epsilon X^N$, X^N is spanned by the set of basis functions $\{\hat{B}_i\}_{i=0}^N$ with $\hat{B}_i = \begin{pmatrix} B_i(0) \\ B_i(\cdot) \end{pmatrix}$ and we shall write $\hat{\mathfrak{B}}_i = \epsilon\hat{B}_i = \begin{pmatrix} \mathfrak{B}_i(0) \\ \mathfrak{B}_i(\cdot) \end{pmatrix} \in \mathbf{X}^N$. All of this is a theoretical technicality; as discussed earlier, from a computational standpoint P^N, X^N, $\hat{B}_i$ and $\mathcal{P}^N$, $\mathbf{X}^N$, $\hat{\mathfrak{B}}_i$ are equivalent.

For each t, $P^N\iota\hat{z}$ is characterized by

$$\langle\langle P^N\iota\hat{z}, \hat{B}_i \rangle\rangle_\tau = \langle\langle \hat{z}, \hat{B}_i \rangle\rangle_\tau \qquad \text{for } i=0,1,\ldots N$$

or, equivalently, by

$$(P_0^N z) B_i(0) + \int_{-\tau}^{0} (P^N z) B_i \, ds = z_0 B_i(0) + \int_{-\tau}^{0} z B_i \, ds \qquad \text{for } i=0,1,\ldots N.$$

In terms of the extensions defined above, we can write

$$(4.1) \qquad (P_0^N z) \mathfrak{B}_i(0) + \int_{-\tau}^{0} (\epsilon P^N z) \mathfrak{B}_i \, ds = z_0 \mathfrak{B}_i(0) + \int_{-\tau}^{0} z \mathfrak{B}_i \, ds \qquad \text{for } i=0,1,\ldots N.$$

We now differentiate both sides of equation (4.1) with respect to t to obtain

$$\frac{d}{dt}\left[P_0^N z\right]\mathcal{B}_i(0) + \int_{-\tau(t)}^{0}\left[\frac{\partial}{\partial t}(\epsilon P^N z)(s)\,\mathcal{B}_i(s) + (P^N z)(s)\frac{\partial \mathcal{B}_i}{\partial t}(s)\right]ds + \dot{\tau}(t)\left[(P^N z)\mathcal{B}_i\right]\Big|_{-\tau(t)}$$
$$= \frac{dz_0}{dt}\mathcal{B}_i(0) + \int_{-\tau(t)}^{0}\left[\frac{\partial z}{\partial t}(s)\,\mathcal{B}_i(s) + z(s)\frac{\partial \mathcal{B}_i}{\partial t}(s)\right]ds + \dot{\tau}(t)(z\mathcal{B}_i)\Big|_{-\tau(t)}.$$

Because $(\epsilon P^N z)(t)$ and $\mathcal{B}_i(t)$ belong to $W^{1,\infty}\left[(0,T);\, H^1(-r,0)\right]$, the partials $\frac{\partial(\epsilon P^N z)}{\partial t}$ and $\frac{\partial \mathcal{B}_i}{\partial t}$ are well defined (a.e. t). With a simple calculation, one can see that $\frac{\partial \mathcal{B}_i}{\partial t} = -\frac{\dot{\tau}}{\tau}s(D\mathcal{B}_i)$. After making this substitution in the above equation and integrating the corresponding terms by parts, we obtain

$$\text{(4.2)} \qquad \frac{d}{dt}\left[P_0^N z\right]\mathcal{B}_i(0) + \int_{-\tau}^{0}\frac{\partial(\epsilon P^N z)}{\partial t}\mathcal{B}_i\,ds + \frac{\dot{\tau}}{\tau}\left[\int_{-\tau}^{0} s\,D(P^N z)\,\mathcal{B}_i\,ds + \int_{-\tau}^{0}\mathcal{B}_i(P^N z)\,ds\right] =$$
$$\frac{dz_0}{dt}\mathcal{B}_i(0) + \int_{-\tau}^{0}\frac{\partial z}{\partial t}\mathcal{B}_i\,ds + \frac{\dot{\tau}}{\tau}\left[\int_{-\tau}^{0} s\,Dz\,\mathcal{B}_i\,ds + \int_{-\tau}^{0}\mathcal{B}_i\,z\,ds\right].$$

Because $\mathcal{B}_i = B_i$ for $s\in[-\tau,0]$, equation (4.2) is equivalent to

$$\langle\langle \frac{d}{dt}(\mathcal{P}^N\hat{z}), \hat{B}_i\rangle\rangle_\tau + \frac{\dot{\tau}}{\tau}\left[\int_{-\tau}^{0} s\,D(P^N z)\,B_i\,ds + \int_{-\tau}^{0} B_i(P^N z)\,ds\right] =$$
$$\langle\langle \mathcal{P}^N\frac{d\hat{z}}{dt}, \hat{B}_i\rangle\rangle_\tau + \frac{\dot{\tau}}{\tau}\left[\int_{-\tau}^{0} s\,Dz\,B_i\,ds + \int_{-\tau}^{0} B_i\,z\,ds\right].$$

Finally, since $\hat{y}\in\mathbf{X}^N(\tau)$ implies that $\hat{y}$ is a linear combination of $\{\hat{B}_i\}$, this equation can be rearranged to give the stated result. □

Our final lemma extends a standard convergence result for spline approximations to the situation here where we have a sequence of time-varying norms.

Lemma 3. Suppose $\{\tau^M\}$ is an arbitrary collection of functions in $\mathcal{T}$ and assume $\hat{z}(t) \in \mathfrak{M}$ for each $t \in [0,T]$. Let us write $P^N(\tau^M)\iota(\tau^M)\hat{z}(t) = \begin{bmatrix} P_0^{N,M}z \\ P^{N,M}z(\cdot) \end{bmatrix}$ and $\hat{z}(t)=\begin{bmatrix} z_0 \\ z(\cdot) \end{bmatrix}$. Then for each fixed $t\in[0,T]$,

$$|D(P^{N,M}z - z)|_{\tau^M} \to 0 \text{ as } N \to \infty$$

and this convergence statement is independent of $\{\tau^M\}$.

Proof. Fix $t\in[0,T]$. Using the notation of Lemma 1, we let $I^{N,M}\iota(\tau^M)\hat{z}(t) = \begin{bmatrix} z_1^{N,M}(0) \\ z_1^{N,M}(\cdot) \end{bmatrix}$. Notice that $\hat{z}\in\mathfrak{M}$ means that $z\in H^1(-r,0)$ (and thus, $\iota^M z\in H^1(-\tau^M,0)$). With a triangle inequality and use of the Schmidt Inequality (Theorem 1.5 in [19]), one can show that

$$\begin{aligned} |D(P^{N,M}z-z)|_{\tau^M} &\le |D(P^{N,M}z-z_1^{N,M})|_{\tau^M} + |D(z_1^{N,M}-z)|_{\tau^M} \\ &\le k_1\left(\frac{N}{\tau_0}\right)|P^{N,M}z-z_1^{N,M}|_{\tau^M} + |D(z_1^{N,M}-z)|_{\tau^M} \\ &\le k_1\left(\frac{N}{\tau_0}\right)\left(|P^{N,M}z-z|_{\tau^M} + |z-z_1^{N,M}|_{\tau^M}\right) + |D(z_1^{N,M}-z)|_{\tau^M} \end{aligned}$$

where k_1 is a constant independent of N,τ^M or z, and τ_0 is as in the definition of the set $\mathcal{T}$. Now,

$$|P^{N,M}z-z|_{\tau^M} \le \|\mathcal{P}^{N,M}\hat{z}-\hat{z}\|_{\tau^M} \le \|I^{N,M}\hat{z}-\hat{z}\|_{\tau^M} = |z_1^{N,M}-z|_{\tau^M}$$

so that $\quad |D(P^{N,M}z-z)|_{\tau^M} \le 2k_1\left(\frac{N}{\tau_0}\right)|z_1^{N,M}-z|_{\tau^M} + |D(z_1^{N,M}-z)|_{\tau^M}$.

We can modifiy (2.16) of [19] slightly to show that for $z\in H^1[-\tau^M,0]$, $|z-z_1^{N,M}|_{\tau^M} \le k_2\left(\frac{\tau^M}{N}\right)|D(z-z_1^{N,M})|_{\tau^M} \le k_2\left(\frac{r}{N}\right)|D(z-z_1^{N,M})|_{\tau^M}$ where k_2 is a constant which is independent of N,τ^M and z. Hence there is some constant C such that

$$|D(P^{N,M}z-z)|_{\tau^M} \le C|D(z_1^{N,M}-z)|_{\tau^M}.$$

A slight modification of (2.17) of [19] yields $|D(w-w_1^{N,M})|_{\tau^M} \le \left(\frac{k_3}{N}\right)|D^2w|_{\tau^M} \le \left(\frac{k_3}{N}\right)|D^2w|_0$ whenever $w\in H^2(-r,0)$, in which case it is clear that the stated result of the Lemma holds. We can now use the First Integral Relation for linear splines ([19], equation (2.10)) and the dense inclusion of $H^2(-r,0)$ in $H^1(-r,0)$ to argue that the result holds whenever $z\in H^1(-r,0)$ and hence for any $\hat{z}\in\mathfrak{M}$. □

In order to prove our first theorem, we make an additional assumption about the unknown delay. We shall assume $\tau \in \mathcal{T}_c \equiv \mathcal{T} \cap \{\tau \mid (1-\dot{\tau}) \geq \delta > 0 \text{ a.e. } t \in [0,T]\}$. The assumption that the time derivative of the delay is bounded under and away from one is crucial to our convergence arguments. We would like to point out however, that this is not an unreasonably restrictive assumption in many modeling situations. Consider that this inequality is equivalent to constraining $\omega(t) \equiv t - \tau(t)$ to have positive derivative, implying that ω is monotonically increasing. Since ω gives the value of the lagged time, we can interpret this condition as follows: Whenever one event precedes another ($\omega(t_1) < \omega(t_2)$), then the effect of the first event must be felt by the state before the effect of the second event (since the monotonicity of ω implies $t_1 < t_2$). In both the models of [7] and [10], (when there is enough smoothness in the problem so that the delay is differentiable) this condition is satisfied.

We now consider the case of time-dependent unknown delay. We define the approximation set for $\mathcal{T}$ as $\mathcal{T}_c^M = \mathcal{T}_c \cap S^M$. The discussion applied to $\mathcal{T}^M$ in section 3 remains valid as applied to this set $\mathcal{T}_c^M$. Thus, the following statements are true:

The set $\mathcal{T}_c^M$ can be characterized as $\mathcal{T}_c^M = I^M \mathcal{T}_c$.

(HT_c) Given $\tau \in \mathcal{T}_c$, $I^M \tau \to \tau$ in $C[0,T]$ as $M \to \infty$.

Each of $\mathcal{T}_c^M$ and $\mathcal{T}_c$ is compact in $C[0,T]$.

<u>Theorem 1:</u> <u>Time-dependent Delay</u>. Assume $\{\rho^M\}$ is an arbitrary sequence in $\Pi_c^M \equiv (\mathcal{G}, \mathcal{T}_c^M, \mathcal{J})$, with $\rho^M \equiv (\gamma^M, \tau^M, \hat{\phi}^M)$ and $\rho^M \to \rho \in \Pi_c$ as $M \to \infty$; this convergence statement means that $\gamma^M \to \gamma$ in $\mathbb{R}^\nu \times C[0,T]$, $\tau^M \to \tau$ in $C[0,T]$, and $\hat{\phi}^M \to \hat{\phi}$ in $\mathbb{R} \times C[-r,0]$. Then, using the notation $\hat{z}(t;\rho) = \begin{bmatrix} z_0(\rho) \\ z(\cdot;\rho) \end{bmatrix} \in \mathbf{Z}$, where $\hat{z}$ is a solution of equation (3.1), and $\hat{z}^N(t;\rho^M) = \begin{bmatrix} z_0^N(\rho^M) \\ z^N(\cdot;\rho^M) \end{bmatrix} \in \mathbf{X}^N(\tau^M)$, where $\hat{z}^N$ is a solution of equation (3.2), it follows that $|z_0^N(\rho^M) - z_0(\rho)| \to 0$ as $N,M \to \infty$ for each $t \in [0,T]$.

<u>Proof</u>. We first observe that

$$|z_0^N(\rho^M) - z_0(\rho)| \leq \| \iota^M \hat{z}^N(t;\rho^M) - \iota^M \hat{z}(t;\rho) \|_{\mathcal{T}^M}$$
$$\leq \| \iota^M \hat{z}^N(t;\rho^M) - P^{N,M} \iota^M \hat{z}(t;\rho) \|_{\mathcal{T}^M} + \| P^{N,M} \iota^M \hat{z}(t;\rho) - \iota^M \hat{z}(t;\rho) \|_{\mathcal{T}^M} .$$

We apply Lemma 1 to the second term above and conclude that it goes to zero as $N, M \to \infty$, for each $t \in [0,T]$. We now turn to the first term. Here we will use the notation $\mathcal{P}^{N,M}\hat{z} = \mathcal{P}^{N}(\tau^{M})\hat{z} = \begin{pmatrix} P_0^{N,M}z \\ P^{N,M}z \end{pmatrix}$. A straightforward calculation demonstrates that (all equations should be interpreted for a.e. t):

$$(4.3) \qquad \frac{d}{dt}\left\| \iota^{M}\hat{z}^{N}(t;\rho^{M}) - P^{N,M}\iota^{M}\hat{z}(t;\rho) \right\|^{2}_{\tau^{M}(t)} = \frac{d}{dt}\left\| \iota^{M}\left[\hat{z}^{N}(t;\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(t;\rho)\right] \right\|^{2}_{\tau^{M}(t)}$$

$$= 2 \cdot \langle\langle \hat{z}^{N}(t;\rho^{M}) - \mathcal{P}^{N,M}z(t;\rho) \,,\, \frac{d}{dt}\left[\hat{z}^{N}(t;\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(t;\rho)\right] \rangle\rangle_{\tau^{M}(t)}$$

$$+ \dot{\tau}^{M}(t) \cdot \left[z^{N}(t;\rho^{M}) - P^{N,M}z(t;\rho)\right]^{2}\Big|_{s=-\tau^{M}(t)} .$$

From now on, in order to minimize notation, we will omit the explicit time dependence and simply write, e.g., $\hat{z}(\rho)$ instead of $\hat{z}(t;\rho)$. We use Lemma 2 and the fact that $\left[\hat{z}^{N}(\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(\rho)\right] \in \mathbf{X}^{N}(\tau^{M})$ to conclude that

$$\langle\langle \frac{d}{dt}\mathcal{P}^{N,M}\hat{z}(\rho) \,,\, (\hat{z}^{N}(\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(\rho)) \rangle\rangle_{\tau^{M}} = \langle\langle \mathcal{P}^{N,M}\frac{d\hat{z}(\rho)}{dt} \,,\, (\hat{z}^{N}(\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(\rho)) \rangle\rangle_{\tau^{M}}$$

$$+ \left(\frac{\dot{\tau}^{M}}{\tau^{M}}\right) \int_{-\tau^{M}}^{0} s\left[Dz(\rho) - DP^{N,M}z(\rho)\right]\left[z^{N}(\rho^{M}) - P^{N,M}z(\rho)\right] ds$$

$$+ \left(\frac{\dot{\tau}^{M}}{\tau^{M}}\right) \int_{-\tau^{M}}^{0} \left[z(\rho) - P^{N,M}z(\rho)\right]\left[z^{N}(\rho^{M}) - P^{N,M}z(\rho)\right] ds .$$

Let us write this as

$$\langle\langle \frac{d}{dt}\mathcal{P}^{N,M}\hat{z}(\rho) \,,\, (\hat{z}^{N}(\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(\rho)) \rangle\rangle_{\tau^{M}} = \langle\langle \mathcal{P}^{N,M}\frac{d\hat{z}(\rho)}{dt} \,,\, (\hat{z}^{N}(\rho^{M}) - \mathcal{P}^{N,M}\hat{z}(\rho)) \rangle\rangle_{\tau^{M}} + \Sigma_1$$

where $\Sigma_1 = \left(\frac{\dot{\tau}^{M}}{\tau^{M}}\right) \int_{-\tau^{M}}^{0} \left[s\{Dz(\rho) - DP^{N,M}z(\rho)\} + \{z(\rho) - P^{N,M}z(\rho)\}\right]\left[z^{N}(\rho^{M}) - P^{N,M}z(\rho)\right] ds$.

Then equation (4.3) is equivalent to

$$\frac{d}{dt}\| \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) \|^2_{\tau^M} = 2 \cdot \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \frac{d\hat{z}^N(\rho^M)}{dt} - \mathcal{P}^{N,M}\frac{d\hat{z}(\rho)}{dt}\rangle\rangle_{\tau^M}$$

$$- 2\Sigma_1 + \dot{\tau}^M \left[z^N(\rho^M) - P^{N,M}z(\rho)\right]^2\Big|_{s=-\tau^M} .$$

We now use the fact that $\left(\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right) \in \mathbf{X}^N(\tau^M) \subset \mathbf{Z}$ and equations (3.1) and (3.2) to rewrite the above equation as

$$\frac{d}{dt}\| \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) \|^2_{\tau^M} = 2 \cdot \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{P}^{N,M}\mathcal{A}(q^M)\hat{z}^N(\rho^M)\rangle\rangle_{\tau^M}$$

$$- 2 \cdot \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{P}^{N,M}\mathcal{A}(q)\hat{z}(\rho)\rangle\rangle_{\tau^M} + \Sigma_2 \tag{4.4}$$

$$= 2 \cdot \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{P}^{N,M}\left[\mathcal{A}(q^M)\hat{z}^N(\rho^M) - \mathcal{A}(q)\hat{z}(\rho)\right]\rangle\rangle_{\tau^M} + \Sigma_2$$

where $\Sigma_2 = -2\Sigma_1 + \dot{\tau}^M \left[z^N(\rho^M) - P^{N,M}z(\rho)\right]^2\Big|_{s=-\tau^M}$. Let us now consider

$$\langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{P}^{N,M}\left[\mathcal{A}(q^M)\hat{z}^N(\rho^M) - \mathcal{A}(q)\hat{z}(\rho)\right]\rangle\rangle_{\tau^M}$$

$$= \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{A}(q^M)\hat{z}^N(\rho^M) - \mathcal{A}(q^M)\mathcal{P}^{N,M}\hat{z}(\rho)\rangle\rangle_{\tau^M}$$

$$+ \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{A}(q^M)\mathcal{P}^{N,M}\hat{z}(\rho) - \mathcal{A}(q)\hat{z}(\rho)\rangle\rangle_{\tau^M} ,$$

which we will define as $T_1 + T_2$. We can estimate T_1 as follows:

$$T_1 = \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho) , \mathcal{A}(q^M)\hat{z}^N(\rho^M) - \mathcal{A}(q^M)\mathcal{P}^{N,M}\hat{z}(\rho)\rangle\rangle_{\tau^M}$$

$$= \left[z_0^N(\rho^M) - P_0^{N,M}z(\rho)\right]\left[F\left(t,q^M,z_0^N(\rho^M),z^N(\rho^M)\right) - F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right)\right]$$

$$+ \int_{-\tau^M}^{0} \left[z^N(\rho^M) - P^{N,M}z(\rho)\right]\left[D\left(z^N(\rho^M) - P^{N,M}z(\rho)\right)\right]ds$$

$$\leq |z_0^N(\rho^M) - P_0^{N,M}z(\rho)|\,|F\Big[t,q^M,z_0^N(\rho^M),z^N(\rho^M)\Big] - F\Big[t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\Big]|$$

$$+ \tfrac{1}{2}\int_{-\tau^M}^{0} D\Big[\Big(z^N(\rho^M) - P^{N,M}z(\rho)\Big)^2\Big]ds$$

$$\leq |z_0^N(\rho^M) - P_0^{N,M}z(\rho)|\,\Big\{m(t)\Big[|z_0^N(\rho^M) - P_0^{N,M}z(\rho)| + |z^N(\rho^M) - P^{N,M}z(\rho)|_{\tau^M}$$

$$+ |(z^N(\rho^M) - P^{N,M}z(\rho))\Big|_{s=-\tau^M}|\Big]\Big\} + \tfrac{1}{2}\Big[z^N(\rho^M) - P^{N,M}z(\rho)\Big]^2\Big|_{s=-\tau^M}^{s=0}$$

$$= m(t)|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|^2 + m(t)|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|\,|z^N(\rho^M) - P^{N,M}z(\rho)|_{\tau^M}$$

$$+ \left[\frac{m(t)}{\sqrt{c}}\,|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|\right]\left[\sqrt{c}\,|\Big[z^N(\rho^M) - P^{N,M}z(\rho)\Big]\Big|_{s=-\tau^M}|\right]$$

$$+ \tfrac{1}{2}\Big[z_0^N(\rho^M) - P_0^{N,M}z(\rho)\Big]^2 - \tfrac{1}{2}\Big[z^N(\rho^M) - P^{N,M}z(\rho)\Big]^2\Big|_{s=-\tau^M},$$

where we have used assumption (HF) in the last inequality and c is any positive constant. Thus,

$$T_1 \leq \Big[m(t)+\tfrac{1}{2}\Big]|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|^2$$

$$+ \frac{m^2(t)}{2}|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|^2 + \tfrac{1}{2}|z^N(\rho^M) - P^{N,M}z(\rho)|_{\tau^M}^2$$

$$+ \frac{m^2(t)}{2c}|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|^2 + \frac{c}{2}\Big[z^N(\rho^M) - P^{N,M}z(\rho)\Big]^2\Big|_{s=-\tau^M}$$

$$- \tfrac{1}{2}\Big[z^N(\rho^M) - P^{N,M}z(\rho)\Big]^2\Big|_{s=-\tau^M}.$$

The term T_2 satisfies:

$$T_2 = \langle\langle \hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\ ,\ \mathcal{A}(q^M)\mathcal{P}^{N,M}\hat{z}(\rho) - \mathcal{A}(q)\hat{z}(\rho)\rangle\rangle_{\tau^M}$$

$$\leq \tfrac{1}{2}\Big\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\Big\|_{\tau^M}^2 + \tfrac{1}{2}\Big\|\mathcal{A}(q^M)\mathcal{P}^{N,M}\hat{z}(\rho) - \mathcal{A}(q)\hat{z}(\rho)\Big\|_{\tau^M}^2.$$

But $\left\| \mathcal{A}(q^M)\mathcal{P}^{N,M}\hat{z}(\rho) - \mathcal{A}(q)\hat{z}(\rho)\right\|^2_{\tau^M} = |F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)|^2$

$$+ |D\left(P^{N,M}z(\rho) - z(\rho)\right)|^2_{\tau^M}$$

so that we have

$$T_2 \leq \tfrac{1}{2}\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} + \tfrac{1}{2}|F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)|^2$$

$$+ \tfrac{1}{2}\,|D\left(P^{N,M}z(\rho) - z(\rho)\right)|^2_{\tau^M}$$

We now combine the above estimates with equation (4.4) to obtain

$$\text{(4.5)} \quad \frac{d}{dt}\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} \leq \left(2m(t) + 1 + (1 + \tfrac{1}{c})\,m^2(t)\right)|z_0^N(\rho^M) - P_0^{N,M}z(\rho)|^2$$

$$+ |z^N(\rho^M) - \mathcal{P}^{N,M}z(\rho)|^2_{\tau^M} + (c - 1)\left(z^N(\rho^M) - P^{N,M}z(\rho)\right)^2\Big|_{s=-\tau^M}$$

$$+ \left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} + |F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)|^2$$

$$+ |D\left(P^{N,M}z(\rho) - z(\rho)\right)|^2_{\tau^M} + \dot{\tau}^M\left[z^N(\rho^M) - P^{N,M}z(\rho)\right]^2\Big|_{s=-\tau^M} - 2\Sigma_1 .$$

We can obtain a bound on Σ_1 as follows:

$$2\cdot|\Sigma_1| \leq 2\cdot\left|\left(\frac{\dot{\tau}^M}{\tau^M}\right)\right|\int_{-\tau^M}^{0} |s\,\{Dz(\rho) - DP^{N,M}z(\rho)\} + \{z(\rho) - P^{N,M}z(\rho)\}|\,|z^N(\rho^M) - P^{N,M}z(\rho)|\,ds$$

$$\leq |\dot{\tau}^M|\left[|Dz(\rho) - DP^{N,M}z(\rho)|^2_{\tau^M} + |z^N(\rho^M) - P^{N,M}z(\rho)|^2_{\tau^M}\right]$$

$$+ \frac{|\dot{\tau}^M|}{|\tau^M|}\left[|z(\rho) - P^{N,M}z(\rho)|^2_{\tau^M} + |z^N(\rho^M) - P^{N,M}z(\rho)|^2_{\tau^M}\right].$$

If we define $\omega_c(t) = |2m(t) + 2 + (1 + \frac{1}{c})m^2(t) + \mu\left(1+\frac{1}{\tau_0}\right)|$ and

$$h^{N,M}(t) = (1+\mu)|Dz(\rho) - DP^{N,M}z(\rho)|^2_{\tau^M} + |F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)|^2$$
$$+ \frac{\mu}{\tau_0}|z(\rho) - P^{N,M}z(\rho)|^2_{\tau^M}$$

then we have a bound of the form

$$\text{(4.6)} \qquad \frac{d}{dt}\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} \le \omega_c(t)\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} + h^{N,M}(t)$$
$$+ \left(c - 1 + \dot{\tau}^M(t)\right)\left(z^N(\rho^M) - P^{N,M}z(\rho)\right)^2\Big|_{s=-\tau^M} .$$

By assumption, $\tau^M \in \mathcal{T}_c$ for all M, which means we can choose $c = \frac{\delta}{2}$ and we are assured that $\left(c - 1 + \dot{\tau}^M(t)\right) \le 0$. Let $\omega = \omega_c$ for this value of c. Then, equation (4.6) implies

$$\frac{d}{dt}\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} \le \omega(t)\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M} + h^{N,M}(t) .$$

We can now use the Gronwall Lemma to see that

$$\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M}(t) \le \left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M}(0)\left(e^{\int_0^t \omega(s)ds}\right) + \int_0^t e^{\int_s^t \omega(\sigma)d\sigma} h^{N,M}(s)ds$$
$$\le \left\|\hat{\phi}^M - \hat{\phi}\right\|_0^2 e^{\int_0^T \omega(s)ds} + \int_0^T e^{\int_0^T \omega(s)ds} h^{N,M}(s)ds \qquad \text{for every } t \in [0,T],$$

where we have used the following estimate for the initial data:

$$\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^M}(0) = \left\|\mathcal{P}^N(\tau^M(0))\hat{z}(0;\rho^M) - \mathcal{P}^N(\tau^M(0))\hat{z}(0;\rho)\right\|^2_{\tau^M(0)}$$
$$\le \left\|\hat{\phi}^M - \hat{\phi}\right\|^2_{\tau^M(0)} \le \left\|\hat{\phi}^M - \hat{\phi}\right\|_0^2 .$$

Our parameter convergence assumption and our assumption (HF) guarantee that

$\|\hat{\phi}^M - \hat{\phi}\|_0^2 \to 0$ as $M \to \infty$ and $\int_0^T \omega(s)ds < \infty$. Therefore, the final step is to argue that $\int_0^T h^{N,M}(s)ds \to 0$ as $N,M \to \infty$.

To this end, consider that

$$|F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)| \leq$$

$$|F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q^M,z_0(\rho),z(\rho)\right)| + |F\left(t,q^M,z_0(\rho),z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)|$$

$$\leq m(t)\left(|P_0^{N,M}z(\rho) - z_0(\rho)| + |P^{N,M}z(\rho) - z(\rho)|_{\tau^M} + |\left(P^{N,M}z(\rho) - z(\rho)\right)\Big|_{s=-\tau^M}|\right)$$

$$+ |F\left(t,q^M,z_0(\rho),z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)| .$$

Thus we can argue that

$$h^{N,M}(t) \leq C(t)\Big(|Dz(\rho) - DP^{N,M}z(\rho)|_{\tau^M}^2 + \|\mathcal{P}^{N,M}\hat{z}(\rho) - \hat{z}(\rho)\|_{\tau^M}^2$$

$$+ |F\left(t,q^M,z_0(\rho),z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)|^2\Big)$$

where $C(t) \in L_2(0,T)$ is independent of N and M (it does depend on $m(t)$, τ_0, μ) . Finally, Lemmas 1 and 3, the continuity of F in q (this follows from the assumptions about f and the fact that $\hat{z}(t;\rho) \in \mathfrak{M}$ implies that $z(t;\rho) \in C(-r,0)$), and the convergence $q^M \to q$ (following from $\rho^M \to \rho$) imply that $h^{N,M}(t) \to 0$ as $N,M \to \infty$ pointwise in t; the dominated convergence theorem can now be applied to deduce that $\int_0^T h^{N,M}(s)ds \to 0$ as $N,M \to \infty$, and this concludes the proof of the theorem. □

<u>Theorem</u> 2. Let $\{\rho^{N,M}\}$ with $\rho^{N,M} = (\gamma^{N,M},\tau^{N,M},\hat{\phi}^{N,M}) \in \mathcal{G} \times \mathcal{T}_C^M \times \mathcal{J} \subset \mathcal{G} \times \mathcal{T}_C \times \mathcal{J}$, be a sequence of solutions to $\{(P^{N,M})\}$. Then, given (HG), (HT_C), and (HI), there exist $\rho^* = (\gamma^*,\tau^*,\hat{\phi}^*) \in \mathcal{G} \times \mathcal{T}_C \times \mathcal{J}$ and a subsequence $\{\rho^{N_k,M_k}\}$ such that $\rho^{N_k,M_k} \to \rho^*$ with N_k, $M_k \to \infty$ (the meaning of this convergence statement is the same as that of Theorem 1), and ρ^* is a solution of (P).

Proof. The existence of the convergent subsequence with limit ρ^* in $\mathcal{G}\times\mathcal{T}_c\times\mathcal{J}$ follows from the compactness of the parameter sets. We would like to show that this limit is a solution of (P), i.e., that $J(\rho^*)\leq J(\rho)$ for any $\rho\in\mathcal{G}\times\mathcal{T}_c\times\mathcal{J}$. Consider any element, $\tilde{\rho}$, of $\mathcal{G}\times\mathcal{T}_c\times\mathcal{J}$. By definition, $J^{N_k}(\rho^{N_k,M_k})\leq J^{N_k}(\rho)$ for any $\rho\in\mathcal{G}\times\mathcal{T}_c^M\times\mathcal{J}$, and, in particular, $J^{N_k}(\rho^{N_k,M_k})\leq J^{N_k}(I^{M_k}\tilde{\rho})$, where we are using the shorthand notation $I^{M_k}\tilde{\rho}\equiv(\tilde{\gamma},I^{M_k}\tilde{\tau},\tilde{\tilde{\phi}})$. Let N_k, $M_k\to\infty$. Using the second statement of (HT_c), we can conclude that $I^{M_k}\tilde{\rho}\to\tilde{\rho}$. Thus, we can apply Theorem 1 to conclude that $J(\rho^*)\leq J(\tilde{\rho})$. This proves the theorem. □

For our second class of problems, we consider equations of the form (2.1), with the delay determined by $\int_{t-\tau(t)}^{t} g(x(s))ds = \theta$, where we shall assume (g,θ) belongs to the set $\overline{\mathcal{T}}_g$. By differentiating, we can write this in the equivalent form

$$\dot{\tau}(t) = 1 - \frac{g(x(t))}{g(x(t-\tau(t)))}\,; \qquad \tau(0) = \tau_i$$

where τ_i is determined by $\int_{-\tau_i}^{0} g(x_0(s))ds = \theta$ (we assume that x_0 and θ are such that τ_i exists). This (the integral form for the delay) is the type of model equation considered in [10] and [11] (see especially [11] for questions of existence and uniqueness of solutions). The model equations of [7] are of a similar nature; a system of delay equations with multiple delays is considered, where each delay satisfies an equation of the form $\dot{\tau}(t) = G(x(t),x(t-\tau(t)))$, where x represents a vector state (eg., see equation (4) of [7]). We do not address these equations directly, but note that the ideas developed here should extend, as the delays of the type in [7] also satisfy the conditions required for our methods.

With the differentiated form, it is more convenient to consider that τ_i is unknown (rather than θ), so now we shall assume that we estimate τ_i and recover θ from the above relationship. It is readily seen (using both of the equivalent integral and differential forms for τ) that if $(g,\theta)\in\overline{\mathcal{T}}_g$, and x and x_0 are any continuous functions, then τ and τ_i satisfy:

$$\underline{\tau} \leq \tau_i \leq \overline{\tau}$$

$$\underline{\tau} \leq \tau(t) \leq \overline{\tau}, \quad |\dot{\tau}|_\infty \leq \mu \quad \text{and} \quad \left[1 - \dot{\tau}(t)\right] \geq \delta \quad \text{for all } t\in[0,T]$$

where $\underline{\tau} \equiv \underline{\theta}/g^U$, $\bar{\tau} \equiv \bar{\theta}/g_L$, $\delta \equiv g_L/g^U$, and $\mu \equiv |1-\delta^{-1}|$; i.e., τ_i belongs to a compact subset of $\mathbb{R}$, and $\tau \in \mathcal{T}_c$. Using the notation of section 3, we write the above differential equation for τ as

$$\dot{\tau}(t) = 1 - \frac{g(z_0(t))}{g(z(t)|_{-\tau(t)})}; \qquad \tau(0) = \tau_i.$$

In the parameter estimation problem, we shall assume that g and τ_i are unknown, and represent our unknown delay terms by the pair $\bar{g} \equiv (g, \tau_i)$. We shall still use the notation $\bar{g} \in \bar{\mathcal{T}}_g$ (the compact set for θ being replaced by the corresponding compact set for τ_i). We must approximate the function g, as discussed in section 3. We are thus led to the consideration of a corresponding approximating equation for the delay:

$$\dot{\tau}^{N,M}(t) = \left[1 - \frac{g^M(z_0^N(t))}{g^M\left[z^N(t)|_{-\tau^{N,M}(t)}\right]}\right]; \qquad \tau^{N,M}(0) = \tau_i^M,$$

where $\bar{g}^M \equiv (g^M, \tau_i^M) \in \bar{\mathcal{T}}_g^M$ (the notation was established in section 3). Unfortunately, however, the sequence $\{\tau^{N,M}\}$ may not satisfy the constraints of $\mathcal{T}_c$, as there is no guarantee that $\tau^{N,M}$ remains positive for all $t \in [0,T]$. We therefore define

$$\lambda^{N,M}(t) = \frac{\tau^{N,M}(t)}{\max\left[\underline{\tau}, \tau^{N,M}(t)\right]}$$

where $\underline{\tau}$ is defined above, and

$$\dot{\tau}^{N,M}(t) = \left[1 - \frac{g^M(z_0^N(t))}{g^M\left[z^N(t)|_{-\tau^{N,M}(t)}\right]}\right]\lambda^{N,M}(t); \qquad \tau^{N,M}(0) = \tau_i^M. \tag{4.7}$$

Our approximation system, then, is the coupled pair of equations (3.2) (solved with $q^M = (\gamma^M, \tau^{N,M})$) and (4.7). That a solution of such a system exists (and is unique) is more easily seen in section 5, and we thus postpone that discussion, assuming for now that this is the case. It can be argued that $\tau^{N,M}$ a solution of (4.7) belongs to $\mathcal{T}_c$ for all N,M. It is clear that all $\tau^{N,M}$ satisfy $(1 - \dot{\tau}^{N,M}) \geq \delta$ (δ defined above) and it is not difficult to show that, for all $t \in [0,T]$, and all N,M

$$\underline{\tau}\, e^{-|\delta^{-1}-1|T/\underline{\tau}} \leq \tau^{N,M}(t) \leq \bar{\tau}\, e^{1/\underline{\tau}}.$$

In this case then, the above will define the numbers τ_0 and r used in the definition of $\mathcal{T}$. The functions $\{\lambda^{N,M}\}$ enjoy certain properties as well. For later reference, we collect the various results in the following lemma.

Lemma 4. Let $\{x^N\}$ and $\{y^N\}$ be any functions satisfying $x^N \in C[0,T]$, $y^N \in C([0,T]\times[-r,0])$, let $\lambda^{N,M}$ be defined as above, and let $\tau^{N,M}$ be the solution of

$$\dot{\tau}^{N,M}(t) = \left(1 - \frac{g^M(x^N(t))}{g^M\left(y^N(t,\tau^{N,M}(t))\right)}\right)\lambda^{N,M}(t); \qquad \tau^{N,M}(0) = \tau_i^M.$$

Then for all N,M :

$\tau^{N,M} \in \mathcal{T}_c$,

$|\lambda^{N,M}|_\infty \le 1$, and

$|1 - \lambda^{N,M}| \le \frac{1}{\underline{\tau}}|\tau - \tau^{N,M}|$ for each $t\in[0,T]$, and any $\tau\in\mathcal{T}_c$.

We are now ready to prove a convergence result for this class of state-dependent delays.

Theorem 3: State-dependent Delay, Type I. Assume $\{\rho^M\}$ is an arbitrary sequence in $\Pi^M \equiv (\mathcal{G}, \overline{\mathcal{T}}_g^M, \mathcal{J})$, with $\rho^M \equiv (\gamma^M, \bar{g}^M, \hat{\phi}^M)$, and $\rho^M \to \rho \in \Pi \equiv (\mathcal{G}, \overline{\mathcal{T}}_g, \mathcal{J})$ as $M \to \infty$; this convergence statement means that $\gamma^M \to \gamma$ in $\mathbb{R}^\nu \times C[0,T]$, $g^M \to g$ in $C[\underline{x}, \bar{x}]$, $\tau_i^M \to \tau_i$ in $\mathbb{R}$, and $\hat{\phi}^M \to \hat{\phi}$ in $\mathbb{R}\times C[-r,0]$. Then (using the same notation as that of Theorem 1), it follows that $|\tau^{N,M} - \tau| + |z_0^N(\rho^M) - z_0(\rho)| \to 0$ as $N,M \to \infty$ for each $t \in [0,T]$.

Proof. We begin as with the proof of Theorem 1, arriving at the following analogue of equation (4.3):

$$\text{(4.8)} \qquad \frac{d}{dt}\left\| \iota^M \hat{z}^N(t;\rho^M) - P^{N,M}\iota^M \hat{z}(t;\rho) \right\|^2_{\tau^{N,M}(t)} + \frac{d}{dt}|\tau^{N,M}(t) - \tau(t)|^2$$

$$= 2 \cdot \langle\langle \hat{z}^N(t;\rho^M) - \mathcal{P}^{N,M} z(t;\rho)\,,\, \frac{d}{dt}\left[\hat{z}^N(t;\rho^M) - \mathcal{P}^{N,M}\hat{z}(t;\rho)\right] \rangle\rangle_{\tau^{N,M}(t)}$$

$$+ \dot{\tau}^{N,M}(t) \cdot \left[z^N(t;\rho^M) - P^{N,M} z(t;\rho)\right]^2\Big|_{s=-\tau^{N,M}(t)} + 2\cdot(\tau^{N,M} - \tau)(\dot{\tau}^{N,M} - \dot{\tau}).$$

Our analysis may now proceed exactly as in the proof of Theorem 1 up to equation (4.5), except we shall replace Σ_1 by $\hat{\Sigma}_1 = \Sigma_1 + (\tau^{N,M} - \tau)(\dot{\tau}^{N,M} - \dot{\tau})$. With this modification, the analogue of (4.5) is:

$$(4.9)\quad \frac{d}{dt}\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^{N,M}} + \left|\tau^{N,M} - \tau\right|^2\right)$$

$$\leq \left[2m(t) + 1 + (1 + \tfrac{1}{c})\, m^2(t)\right] |z_0^N(\rho^M) - P_0^{N,M} z(\rho)|^2$$

$$+ |z^N(\rho^M) - \mathcal{P}^{N,M} z(\rho)|^2_{\tau^{N,M}} + (c - 1)\left[z^N(\rho^M) - P^{N,M} z(\rho)\right]^2\Big|_{s=-\tau^{N,M}}$$

$$+ \left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\tau^{N,M}} + \left|F\left[t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right] - F\left[t,q,z_0(\rho),z(\rho)\right]\right|^2$$

$$+ \left|D\left[P^{N,M}z(\rho) - z(\rho)\right]\right|^2_{\tau^{N,M}} + \dot{\tau}^{N,M}\left[z^N(\rho^M) - P^{N,M}z(\rho)\right]^2\Big|_{s=-\tau^{N,M}} - 2\hat{\Sigma}_1 .$$

Now, in order to bound $\hat{\Sigma}_1$, we first consider the new delay term; using a triangle inequality and the statements of Lemma 4, we obtain:

$$(\tau^{N,M} - \tau)(\dot{\tau}^{N,M} - \dot{\tau}) = (\tau^{N,M} - \tau)\left\{\left[1 - \frac{g^M(z_0^N(\rho^M))}{g^M\left[z^N(\rho^M)\big|_{-\tau^{N,M}}\right]}\right]\lambda^{N,M} - \left[1 - \frac{g(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]}\right]\right\}$$

$$\leq \left|\tau^{N,M} - \tau\right|\left|\lambda^{N,M}\right|\left|\left[1 - \frac{g^M(z_0^N(\rho^M))}{g^M\left[z^N(\rho^M)\big|_{-\tau^{N,M}}\right]}\right] - \left[1 - \frac{g(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]}\right]\right|$$

$$+ \left|\tau^{N,M} - \tau\right|\left|(\lambda^{N,M} - 1)\left[1 - \frac{g(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]}\right]\right|.$$

Thus,

$$(\tau^{N,M} - \tau)(\dot{\tau}^{N,M} - \dot{\tau})$$

$$(4.10)\qquad \leq \left|\tau^{N,M} - \tau\right|\left|\frac{g(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]} - \frac{g^M(z_0^N(\rho^M))}{g^M\left[z^N(\rho^M)\big|_{-\tau^{N,M}}\right]}\right| + \frac{\left|1 - \delta^{-1}\right|}{\underline{\tau}}\left|\tau^{N,M} - \tau\right|^2$$

$$\equiv \sigma_1 + \sigma_2.$$

The term σ_1 satisfies

$$\sigma_1 \leq \left|\tau^{N,M} - \tau\right|\left(\left|\frac{g(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]} - \frac{g^M(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]}\right| + \left|\frac{g^M(z_0(\rho))}{g\left[z(\rho)\big|_{-\tau}\right]} - \frac{g^M(z_0^N(\rho^M))}{g\left[z(\rho)\big|_{-\tau}\right]}\right|\right.$$

$$\left. + \left|\frac{g^M(z_0^N(\rho^M))}{g\left[z(\rho)\big|_{-\tau}\right]} - \frac{g^M(z_0^N(\rho^M))}{g^M\left[z^N(\rho^M)\big|_{-\tau^{N,M}}\right]}\right|\right)$$

$$\leq \frac{1}{g_L^2}\left|\tau^{N,M}-\tau\right|^2+\frac{1}{2}\left|g(z_0(\rho))-g^M(z_0(\rho))\right|^2+\frac{1}{2}\mu_g^2\left|z_0(\rho)-z_0^N(\rho^M)\right|^2$$

$$+\frac{g^U}{g_L^2}\left|\tau^{N,M}-\tau\right|\left|g^M\left[z^N(\rho^M)\big|_{-\tau^{N,M}}\right]-g\left[z(\rho)\big|_{-\tau}\right]\right|$$

$$\leq \frac{1}{g_L^2}\left|\tau^{N,M}-\tau\right|^2+\frac{1}{2}\left|g(z_0(\rho))-g^M(z_0(\rho))\right|^2+\mu_g^2\left|z_0(\rho)-P_0^{N,M}z(\rho)\right|^2$$

$$+\mu_g^2\left|P_0^{N,M}z(\rho)-z_0^N(\rho^M)\right|^2+\frac{g^U}{g_L^2}\left|\tau^{N,M}-\tau\right|\left[\mu_g\left|z^N(\rho^M)\big|_{-\tau^{N,M}}-z(\rho)\big|_{-\tau^{N,M}}\right|\right.$$

$$\left.+\mu_g\left|z(\rho)\big|_{-\tau^{N,M}}-z(\rho)\big|_{-\tau}\right|+\left|g^M\left[z(\rho)\big|_{-\tau}\right]-g\left[z(\rho)\big|_{-\tau}\right]\right|\right]$$

$$\leq\left[\frac{1}{g_L^2}+\frac{1}{2}\left(\frac{g^U}{g_L^2}\mu_g\right)^2+\frac{1}{2}\left(\frac{g^U}{g_L^2}\right)^2\right]\left|\tau^{N,M}-\tau\right|^2+\mu_g^2\left|P_0^{N,M}z(\rho)-z_0^N(\rho^M)\right|^2$$

$$+\frac{\mu_g g^U}{g_L^2}\left|\tau^{N,M}-\tau\right|\left[\left|z^N(\rho^M)\big|_{-\tau^{N,M}}-P^{N,M}z(\rho)\big|_{-\tau^{N,M}}\right|+\left|P^{N,M}z(\rho)\big|_{-\tau^{N,M}}-z(\rho)\big|_{-\tau^{N,M}}\right|\right]$$

$$+\mu_g^2\left|z_0(\rho)-P_0^{N,M}z(\rho)\right|^2+\frac{1}{2}\left|g(z_0(\rho))-g^M(z_0(\rho))\right|^2+\frac{1}{2}\left|z(\rho)\big|_{-\tau^{N,M}}-z(\rho)\big|_{-\tau}\right|^2$$

$$+\frac{1}{2}\left|g^M\left[z(\rho)\big|_{-\tau}\right]-g\left[z(\rho)\big|_{-\tau}\right]\right|^2.$$

Since we may identify $\hat{z}(t;\rho)$ with $(x(t),x_t(\cdot))^T$, which is a solution of the original delay equation (with the parameter ρ), we see that $\left|z(\rho)\big|_{-\tau^{N,M}}-z(\rho)\big|_{-\tau}\right|\leq|Dz(\rho)|_\infty\left|\tau-\tau^{N,M}\right|\leq\sup_{-r\leq t\leq T}|\dot{x}(t)|\left|\tau-\tau^{N,M}\right|$, and therefore we can write (for some constants K_1, K_2 which depend on the various constants, g^U, g_L, μ_g, etc., but not on N,M):

$$\sigma_1\leq K_1\left|\tau^{N,M}-\tau\right|^2+\mu_g^2\left|P_0^{N,M}z(\rho)-z_0^N(\rho^M)\right|^2$$

$$+K_2\left|\tau^{N,M}-\tau\right|\left|z^N(\rho^M)\big|_{-\tau^{N,M}}-P^{N,M}z(\rho)\big|_{-\tau^{N,M}}\right|$$

$$+\frac{1}{2}\left|P^{N,M}z(\rho)\big|_{-\tau^{N,M}}-z(\rho)\big|_{-\tau^{N,M}}\right|^2+\mu_g^2\left|z_0(\rho)-P_0^{N,M}z(\rho)\right|^2$$

$$+\frac{1}{2}\left|g(z_0(\rho))-g^M(z_0(\rho))\right|^2+\frac{1}{2}\left|g^M\left[z(\rho)\big|_{-\tau}\right]-g\left[z(\rho)\big|_{-\tau}\right]\right|^2$$

and finally, for some positive constant $\tilde{c}$ (to be determined later) we have

$$\sigma_1 \le \left(K_1 + \frac{K_2^2}{2\tilde{c}}\right)\left|\tau^{N,M} - \tau\right|^2 + \frac{\tilde{c}}{2}\left|z^N(\rho^M)\big|_{-\tau^{N,M}} - P^{N,M}z(\rho)\big|_{-\tau^{N,M}}\right|^2$$
$$+ \mu_g^2\left|P_0^{N,M}z(\rho) - z_0^N(\rho^M)\right|^2 + \frac{1}{2}\left|P^{N,M}z(\rho)\big|_{-\tau^{N,M}} - z(\rho)\big|_{-\tau^{N,M}}\right|^2$$
$$+ \mu_g^2\left|z_0(\rho) - P_0^{N,M}z(\rho)\right|^2 + \frac{1}{2}\left|g(z_0(\rho)) - g^M(z_0(\rho))\right|^2$$
$$+ \frac{1}{2}\left|g^M\left[z(\rho)\big|_{-\tau}\right] - g\left[z(\rho)\big|_{-\tau}\right]\right|^2.$$

Combining the above estimate with equation (4.10) and the estimate for Σ_1 obtained in the proof of Theorem 1, we have now shown that

$$2\left|\hat{\Sigma}_1\right| \le \mu\left[|Dz(\rho) - DP^{N,M}z(\rho)|^2_{\tau^{N,M}} + |z^N(\rho^M) - P^{N,M}z(\rho)|^2_{\tau^{N,M}}\right]$$
$$+ \frac{\mu}{\tau_0}\left[|z(\rho) - P^{N,M}z(\rho)|^2_{\tau^{N,M}} + |z^N(\rho^M) - P^{N,M}z(\rho)|^2_{\tau^{N,M}}\right] + \tilde{K}_1\left|\tau^{N,M} - \tau\right|^2$$
$$+ \tilde{c}\left|\left(z^N(\rho^M) - P^{N,M}z(\rho)\right)\Big|_{-\tau^{N,M}}\right|^2 + 2\mu_g^2\left|P_0^{N,M}z(\rho) - z_0^N(\rho^M)\right|^2$$
$$+ \left|\left(P^{N,M}z(\rho) - z(\rho)\right)\Big|_{-\tau^{N,M}}\right|^2 + 2\mu_g^2\left|z_0(\rho) - P_0^{N,M}z(\rho)\right|^2$$
$$+ \left|g(z_0(\rho)) - g^M(z_0(\rho))\right|^2 + \left|g^M\left[z(\rho)\big|_{-\tau}\right] - g\left[z(\rho)\big|_{-\tau}\right]\right|^2.$$

Proceeding as in Theorem 1, we define

$$\omega_{\tilde{c}} = \left|2m(t) + 2 + \left(1 + \frac{1}{\tilde{c}}\right)m^2(t) + \mu\left(1 + \frac{1}{\tau_0}\right) + 2\mu_g^2 + \tilde{K}_1\right|$$

and

$$h^{N,M}(t) = (1+\mu)|Dz(\rho) - DP^{N,M}z(\rho)|^2_{\tau^{N,M}} + \left|\left(P^{N,M}z(\rho) - z(\rho)\right)\Big|_{-\tau^{N,M}}\right|^2$$
$$+ \left|F\left[t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right] - F\left[t,q,z_0(\rho),z(\rho)\right]\right|^2 + \frac{\mu}{\tau_0}\left|z(\rho) - P^{N,M}z(\rho)\right|^2_{\tau^{N,M}}$$
$$+ 2\mu_g^2\left|z_0(\rho) - P_0^{N,M}z(\rho)\right|^2 + \left|g(z_0(\rho)) - g^M(z_0(\rho))\right|^2 + \left|g^M\left[z(\rho)\big|_{-\tau}\right] - g\left[z(\rho)\big|_{-\tau}\right]\right|^2$$

and then equation (4.9) becomes

$$\frac{d}{dt}\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}} + \left|\tau^{N,M} - \tau\right|^2\right)$$

$$\leq \omega_{\tilde{c}}\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}} + \left|\tau^{N,M} - \tau\right|^2\right)$$

$$+ \left(c - 1 + \dot{\tau}^{N,M} + \tilde{c}\right)\left[z^N(\rho^M) - P^{N,M}z(\rho)\right]^2\Big|_{s=-\tau^{N,M}} + h^{N,M}(t).$$

Because each $\tau^{N,M}$ belongs to $\mathcal{T}_c$, we are assured of being able to choose c and $\tilde{c}$ such that $(c - 1 + \dot{\tau}^{N,M} + \tilde{c}) \leq 0$ for all N,M. Thus, we can write

$$\text{(4.11)} \qquad \frac{d}{dt}\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}} + \left|\tau^{N,M} - \tau\right|^2\right)$$

$$\leq \omega\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}} + \left|\tau^{N,M} - \tau\right|^2\right) + h^{N,M}(t).$$

An application of the Gronwall Lemma now gives:

$$\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}}(t) + \left|\tau^{N,M} - \tau\right|^2(t)$$

$$\leq \left(\left\|\hat{\phi}^M - \hat{\phi}\right\|^2_0 + \left|\tau_i^M - \tau_i\right|^2\right) e^{\int_0^T \omega(s)ds} + \int_0^T e^{\int_0^T \omega(s)ds} h^{N,M}(s)ds$$

for each $t \in [0,T]$. Given that $\hat{\phi}^M \to \hat{\phi}$ and $\tau_i^M \to \tau_i$, our final step is to argue that $h^{N,M} \to 0$ in $L_1(0,T)$. The function $h^{N,M}$ is composed of the same terms as in the time-dependent delay case, with some additional terms here. We can begin in the same way as in the proof of Theorem 1 to bound $h^{N,M}$, additionally considering the new terms. Thus we can derive

$$h^{N,M}(t) \leq \tilde{C}(t)\left(\left|Dz(\rho) - DP^{N,M}z(\rho)\right|^2_{\mathcal{T}^{N,M}} + \left\|\mathcal{P}^{N,M}\hat{z}(\rho) - \hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}}\right)$$

$$+ \left|F\left(t,q^M,z_0(\rho),z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)\right|^2$$

$$+ \left|g(z_0(\rho)) - g^M(z_0(\rho))\right|^2 + \left|g^M\left(z(\rho)|_{-\tau}\right) - g\left(z(\rho)|_{-\tau}\right)\right|^2.$$

The second term above satisfies

$$\left|F\left(t,q^M,z_0(\rho),z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)\right| =$$

$$\left|f[t,\gamma^M,z_0(\rho),z(\rho)|_{s=-\tau^{N,M}},z(\rho)] - f[t,\gamma,z_0(\rho),z(\rho)|_{s=-\tau},z(\rho)]\right|$$

$$\leq \left|f[t,\gamma^M,z_0(\rho),z(\rho)|_{s=-\tau^{N,M}},z(\rho)] - f[t,\gamma^M,z_0(\rho),z(\rho)|_{s=-\tau},z(\rho)]\right|$$

$$+ \left|f[t,\gamma^M,z_0(\rho),z(\rho)|_{s=-\tau},z(\rho)] - f[t,\gamma,z_0(\rho),z(\rho)|_{s=-\tau},z(\rho)]\right|$$

$$\leq m(t)\left|z(\rho)\big|_{s=-\tau^{N,M}} - z(\rho)\big|_{s=-\tau}\right|$$

$$+ \left|f[t,\gamma^M,z_0(\rho),z(\rho)\big|_{s=-\tau},z(\rho)] - f[t,\gamma,z_0(\rho),z(\rho)\big|_{s=-\tau},z(\rho)]\right|.$$

$$\leq m(t)\left(\sup_{-r\leq t\leq T}|\dot{x}(t)|\right)\left|\tau^{N,M}-\tau\right|$$

$$+ \left|f[t,\gamma^M,z_0(\rho),z(\rho)\big|_{-\tau},z(\rho)] - f[t,\gamma,z_0(\rho),z(\rho)\big|_{-\tau},z(\rho)]\right|.$$

$$\leq \hat{C}(t)\left|\tau^{N,M}-\tau\right| + \left|f[t,\gamma^M,z_0(\rho),z(\rho)\big|_{-\tau},z(\rho)] - f[t,\gamma,z_0(\rho),z(\rho)\big|_{-\tau},z(\rho)]\right|.$$

Finally, we shall redefine $\omega = \omega + 2\hat{C}^2$ in equation (4.10) (thereby removing the first term above from the definition of $h^{N,M}$), and then we write

$$h^{N,M}(t) \leq \tilde{C}(t)\left(|Dz(\rho)-DP^{N,M}z(\rho)|^2_{\mathcal{T}^{N,M}} + \left\|\mathcal{P}^{N,M}\hat{z}(\rho)-\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}}\right)$$

$$+ 2\left|f[t,\gamma^M,z_0(\rho),z(\rho)\big|_{-\tau},z(\rho)] - f[t,\gamma,z_0(\rho),z(\rho)\big|_{-\tau},z(\rho)]\right|^2$$

$$+ \left|g(z_0(\rho)) - g^M(z_0(\rho))\right|^2 + \left|g^M\left[z(\rho)\big|_{-\tau}\right] - g\left[z(\rho)\big|_{-\tau}\right]\right|^2.$$

We can now see that the desired convergence follows from Lemmas 1 and 3, the continuity of f and the convergence $\gamma^M \to \gamma$, and the convergence $g^M \to g$. This concludes the proof of the theorem. □

We can also prove an analogue of Theorem 2 for this state-dependent case. The proof follows almost word for word; we must only replace the statements of (HT_c) by the corresponding statements for the set $\mathcal{T}_g^M$ (these properties have been discussed in section 3, we compile them here for reference):

(HT_g) The set $\mathcal{T}_g^M$ can be characterized as $I^M\mathcal{T}_g$.
Given $g\in\mathcal{T}_g$, $I^M g \to g$ in $C[\underline{x},\bar{x}]$ as $M\to\infty$.
Each of $\mathcal{T}_g^M$ and $\mathcal{T}_g$ is compact in $C[\underline{x},\bar{x}]$; each of $\bar{\mathcal{T}}_g^M$ and $\bar{\mathcal{T}}_g$ is compact in $C[\underline{x},\bar{x}]\times\mathbb{R}$.

<u>Theorem 4.</u> Let $\{\rho^{N,M}\}$ with $\rho^{N,M} = (\gamma^{N,M},\bar{g}^{N,M},\hat{\phi}^{N,M}) \in \mathcal{G}\times\bar{\mathcal{T}}_g^M\times\mathcal{J} \subset \mathcal{G}\times\bar{\mathcal{T}}_g\times\mathcal{J}$, be a sequence of solutions to $\{(P^{N,M})\}$. Then, given (HG), (HT_g), and (HI), there exist $\rho^* = (\gamma^*,\bar{g}^*,\hat{\phi}^*) \in \mathcal{G}\times\bar{\mathcal{T}}_g\times\mathcal{J}$ and a subsequence $\{\rho^{N_k,M_k}\}$ such that $\rho^{N_k,M_k}\to\rho^*$ with N_k, $M_k\to\infty$ (the meaning of this convergence statement is the same as that of Theorem 3), and ρ^* is a solution of (P).

For our final case, we consider equation (2.1) with a delay of the form $\tau(t) = g(x(t))$, where the function g is assumed unknown and is to be estimated. Using our abstract notation, this is written as $\tau(t) = g(z_0(t))$. In order to prove a convergence result, we reformulate our delay in terms of a differential equation. Thus, as long as we consider $g \in W^{1,\infty}(\underline{x},\bar{x})$ and $z_0 \in W^{1,\infty}(0,T)$ the unknown delay satisfies an equation of the form

$$\dot{\tau} = Dg(z_0(t))\dot{z}_0(t) \; ; \qquad \tau(0) = g(\phi(0)),$$

which is equivalently written as

$$\dot{\tau} = Dg(z_0(t))\, f\Big[t, \gamma, z_0(t), z(t)\big|_{-\tau(t)}, z(t)\Big] \; ; \qquad \tau(0) = g(\phi(0)).$$

For this form of the delay, we must make additional assumptions both about f and the unknown function g in order to prove a convergence theorem.

($\hat{H}F$) There exists a constant $\mathcal{F}$ such that for any $\gamma \in \mathcal{G}$, and any $(\xi,\varsigma,\eta) \in \mathbb{R}\times\mathbb{R}\times L_2(-r,0)$, it follows that $|f(t,\gamma,\xi,\varsigma,\eta)| \leq \mathcal{F}$.

Let $(\hat{z},\tau)_g$ represent a solution pair of the above equation with a particular $g \in \mathcal{T}_g$, coupled with equation (3.1) for the state.

($\hat{H}T_g$) Assume we search for g in the set $\tilde{\mathcal{T}}_g = \mathcal{T}_g \cap \{g \in W^{2,\infty}(\underline{x},\bar{x}) \mid |D^2 g|_\infty \leq \tilde{\mu}_g$ and $(\hat{z},\tau)_g$ satisfies $\tau(\cdot) \equiv g o z_0 \in \mathcal{T}_c\}$.

It is no longer clear in this case of state-dependent delays how to restrict the set of unknown functions in order to guarantee that $\tau \in \mathcal{T}_c$. This particular formulation of delay is perhaps not so natural. We note that the previous formulation was motivated by specific scientific phenomena ([7], [10]), whereas the present formulation comes from a more mathematical perspective (eg., our example problem in the final section comes from [16], where no underlying model has been considered). Nonetheless, it is still possible to obtain some results for this case.

Let us assume we begin with some compact (in $W^{1,\infty}$) subset of $\tilde{\mathcal{T}}_g$ (without changing notation - i.e., now let us consider that $\tilde{\mathcal{T}}_g$ is compact), and define our approximation subset as $\mathcal{T}_g^M \equiv I^M \tilde{\mathcal{T}}_g$; this set $\mathcal{T}_g^M$ is also compact, however the compactness is no longer easy to characterize, as was the case for the previous

types of delays. We could begin, as before, by considering the analogue of the above equation with g replaced by g^M and z_0 replaced by z_0^N, however, we could not then guarantee that the resultant $\tau^{N,M}$ belongs to $\mathcal{T}_C$. Thus we will instead consider the following equation to determine the approximating delays:

$$
\begin{aligned}
&\dot{\tau}^{N,M}(t) = \max\Big[(1-\delta),\ Dg^M(z_0^N(t)) \cdot f(t,\gamma^M, z_0^N(t), z^N(t)|_{-\tau^{N,M}(t)}, z^N(t)) \cdot \lambda^{N,M}(t)\Big] \\
&\tau^{N,M}(0) = g^M(z_0^N(0)),
\end{aligned}
\tag{4.12}
$$

where $\lambda^{N,M}$ is defined in the same way as in the previous class of delays. Our approximating equations are then the coupled system (3.2) and (4.12), and again it can be argued (assuming for now that the system has a unique solution) that $\tau^{N,M}$ a solution of (4.12), and $\lambda^{N,M}$ will satisfy the conclusions of Lemma 4.

<u>Theorem 5:</u> <u>State-dependent Delay, Type II.</u> Assume $\{\rho^M\}$ is an arbitrary sequence in $\Pi^M \equiv (\mathcal{G}, \mathcal{T}_g^M, \mathcal{J})$, with $\rho^M \equiv (\gamma^M, g^M, \hat{\phi}^M)$, and $\rho^M \to \rho \in \Pi \equiv (\mathcal{G}, \tilde{\mathcal{T}}_g, \mathcal{J})$ as $M \to \infty$; this convergence statement means that $\gamma^M \to \gamma$ in $\mathbb{R}^\nu \times C[0,T]$, $g^M \to g$ in $C[\underline{x}, \bar{x}]$, $Dg^M \to Dg$ in $L^\infty[\underline{x}, \bar{x}]$, and $\hat{\phi}^M \to \hat{\phi}$ in $\mathbb{R} \times C[-r,0]$. Then (using the same notation as that of Theorem 1), it follows that $|\tau^{N,M} - \tau| + |z_0^N(\rho^M) - z_0(\rho)| \to 0$ as $N,M \to \infty$ for each $t \in [0,T]$.

<u>Proof.</u> The proof follows the same outline as those of Theorems 1 and 3, so we shall only present the differences. Consider first that

$$
\begin{aligned}
(\tau^{N,M} - \tau)(\dot{\tau}^{N,M} - \dot{\tau}) &\le |\tau^{N,M} - \tau| \Big\{ \Big| f\big(t,\gamma,z_0,z|_{-\tau},z\big) \Big| \big| Dg(z_0) - Dg^M(z_0^N) \big| \\
&\quad + \big| Dg^M(z_0^N) \big| \Big| f\big(t,\gamma,z_0,z|_{-\tau},z\big) - f\big(t,\gamma^M,z_0^N,z^N|_{-\tau^{N,M}},z^N\big) \cdot \lambda^{N,M} \Big| \Big\} \\
&\le \mathcal{F} |\tau^{N,M} - \tau| \big| Dg(z_0) - Dg(z_0^N) \big| + \mathcal{F} |\tau^{N,M} - \tau| \big| Dg(z_0^N) - Dg^M(z_0^N) \big| \\
&\quad + K |\tau^{N,M} - \tau| \ \Big\{ \mathcal{F} \frac{1}{\underline{\tau}} |\tau - \tau^{N,M}| \\
&\qquad + \Big| f\big(t,\gamma,z_0,z)|_{-\tau},z\big) - f\big(t,\gamma^M,z_0^N,z^N|_{-\tau^{N,M}},z^N\big) \Big| \Big\}
\end{aligned}
$$

where we have used the properties of $\lambda^{N,M}$ and ($\hat{\text{HF}}$), and K is a constant which does not depend on N,M. It then follows (using the fact that $g \in \tilde{\mathcal{T}}_g$ and (HF)) that

$$(\tau^{N,M}-\tau)(\dot\tau^{N,M}-\dot\tau)\le$$

$$\Big(\frac{\mathcal{F}^2\,\tilde\mu_g^2}{2}+\frac{\mathcal{F}^2}{2}+\frac{\mathcal{F}K}{\underline{\tau}}+\frac{K^2}{2}\Big)\left|\tau^{N,M}-\tau\right|^2+\frac{1}{2}\left|z_0-z_0^N\right|^2$$

$$+\frac{1}{2}\left|Dg(z_0^N)-Dg^M(z_0^N)\right|^2+\frac{1}{2}\left|f\big(t,\gamma,z_0,z|_{-\tau},z\big)-f\big(t,\gamma^M,z_0,z|_{-\tau},z\big)\right|^2$$

$$+K\left|\tau^{N,M}-\tau\right|m(t)\Big\{\left|z_0-z_0^N\right|+\left|z|_{-\tau}-z^N|_{-\tau^{N,M}}\right|+\left|z-z^N\right|_{\tau^{N,M}}\Big\}$$

$$\le\omega_1(t)\Big(\left|\tau^{N,M}-\tau\right|^2+\left\|\hat z^N(\rho^M)-\mathcal{P}^{N,M}\hat z(\rho)\right\|^2_{\tau^{N,M}}\Big)+C_1\left\|P^{N,M}\hat z-\hat z\right\|^2_{\tau^{N,M}}$$

$$+\frac{1}{2}\left|Dg(z_0^N)-Dg^M(z_0^N)\right|^2+\frac{1}{2}\left|f\big(t,\gamma,z_0,z|_{-\tau},z\big)-f\big(t,\gamma^M,z_0,z|_{-\tau},z\big)\right|^2$$

$$+m(t)\,K\left|\tau^{N,M}-\tau\right|\left|z|_{-\tau}-z^N|_{-\tau^{N,M}}\right|$$

where $\omega_1\in L_2[0,T]$, $C_1\in\mathbb{R}$ do not depend on N,M. With several triangle inequalities, and letting $\tilde c$ be some positive constant to be determined later, we see that

$$(\tau^{N,M}-\tau)(\dot\tau^{N,M}-\dot\tau)\le$$

$$\omega_1(t)\Big(\left|\tau^{N,M}-\tau\right|^2+\left\|\hat z^N(\rho^M)-\mathcal{P}^{N,M}\hat z(\rho)\right\|^2_{\tau^{N,M}}\Big)+C_1\left\|P^{N,M}\hat z-\hat z\right\|^2_{\tau^{N,M}}$$

$$+\frac{1}{2}\left|Dg(z_0^N)-Dg^M(z_0^N)\right|^2+\frac{1}{2}\left|f\big(t,\gamma,z_0,z|_{-\tau},z\big)-f\big(t,\gamma^M,z_0,z|_{-\tau},z\big)\right|^2$$

$$+m^2K^2\left|\tau^{N,M}-\tau\right|^2+\frac{1}{2}\left|z|_{-\tau}-z|_{-\tau^{N,M}}\right|^2+\frac{1}{2}\left|(z-P^{N,M}z)|_{-\tau^{N,M}}\right|^2$$

$$+\frac{m^2K^2}{\tilde c}\left|\tau^{N,M}-\tau\right|^2+\frac{\tilde c}{2}\left|(P^{N,M}z-z^N)|_{-\tau^{N,M}}\right|^2.$$

We can bound the term $\left|z|_{-\tau}-z|_{-\tau^{N,M}}\right|^2$ exactly as in the arguments of Theorem 3, so that we shall finally write

$$(\tau^{N,M}-\tau)(\dot\tau^{N,M}-\dot\tau)\le$$

$$\tilde\omega_1(t)\Big(\left|\tau^{N,M}-\tau\right|^2+\left\|\hat z^N(\rho^M)-\mathcal{P}^{N,M}\hat z(\rho)\right\|^2_{\tau^{N,M}}\Big)+C_1\left\|P^{N,M}\hat z-\hat z\right\|^2_{\tau^{N,M}}$$

$$+\frac{1}{2}\left|Dg(z_0^N)-Dg^M(z_0^N)\right|^2+\frac{1}{2}\left|f\big(t,\gamma,z_0,z|_{-\tau},z\big)-f\big(t,\gamma^M,z_0,z|_{-\tau},z\big)\right|^2$$

$$+\frac{1}{2}\left|(z-P^{N,M}z)|_{-\tau^{N,M}}\right|^2+\frac{\tilde c}{2}\left|(P^{N,M}z-z^N)|_{-\tau^{N,M}}\right|^2.$$

We can now follow the proof of Theorem 3: We have

$$
\begin{aligned}
2\left|\hat{\Sigma}_1\right| \le\ & \mu|Dz(\rho)-DP^{N,M}z(\rho)|^2_{\mathcal{T}^{N,M}} + |z^N(\rho^M)-P^{N,M}z(\rho)|^2_{\mathcal{T}^{N,M}} \\
& + \frac{\mu}{\tau_0}\left[|z(\rho)-P^{N,M}z(\rho)|^2_{\mathcal{T}^{N,M}} + |z^N(\rho^M)-P^{N,M}z(\rho)|^2_{\mathcal{T}^{N,M}}\right] \\
& + 2\tilde{\omega}_1(t)\left(\left|\tau^{N,M}-\tau\right|^2 + \left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}}\right) + 2C_1\left\|P^{N,M}\hat{z}-\hat{z}\right\|^2_{\mathcal{T}^{N,M}} \\
& + \left|Dg(z_0^N) - Dg^M(z_0^N)\right|^2 + \left|f\left(t,\gamma,z_0,z|_{-\mathcal{T}},z\right) - f\left(t,\gamma^M,z_0,z|_{-\mathcal{T}},z\right)\right|^2 \\
& + \left|(z-P^{N,M}z)|_{-\mathcal{T}^{N,M}}\right|^2 + \tilde{c}\left|(P^{N,M}z-z^N)|_{-\mathcal{T}^{N,M}}\right|^2.
\end{aligned}
$$

We define $\omega_{\tilde{c}} = |2m(t) + 2 + (1+\frac{1}{\tilde{c}})m^2(t) + \mu\left(1+\frac{1}{\tau_0}\right) + 2\tilde{\omega}_1(t)|$ and

$$
\begin{aligned}
h^{N,M}(t) =\ & (1+\mu)|Dz(\rho)-DP^{N,M}z(\rho)|^2_{\mathcal{T}^{N,M}} + \left|\left(P^{N,M}z(\rho)-z(\rho)\right)\right|_{-\mathcal{T}^{N,M}}\Bigg|^2 \\
& + \left|F\left(t,q^M,P_0^{N,M}z(\rho),P^{N,M}z(\rho)\right) - F\left(t,q,z_0(\rho),z(\rho)\right)\right|^2 + \frac{\mu}{\tau_0}\left|z(\rho)-P^{N,M}z(\rho)\right|^2_{\mathcal{T}^{N,M}} \\
& + \left|f\left(t,\gamma,z_0,z|_{-\mathcal{T}},z\right) - f\left(t,\gamma^M,z_0,z|_{-\mathcal{T}},z\right)\right|^2 + \left|Dg(z_0^N) - Dg^M(z_0^N)\right|^2 \\
& + 2C_1\left\|P^{N,M}\hat{z}-\hat{z}\right\|^2_{\mathcal{T}^{N,M}},
\end{aligned}
$$

and, choosing our constants appropriately, we conclude that

$$
\begin{aligned}
& \frac{d}{dt}\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}} + \left|\tau^{N,M}-\tau\right|^2\right) \\
& \qquad \le \omega\left(\left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}} + \left|\tau^{N,M}-\tau\right|^2\right) + h^{N,M}(t).
\end{aligned}
$$

The Gronwall Lemma now gives

$$
\begin{aligned}
& \left\|\hat{z}^N(\rho^M) - \mathcal{P}^{N,M}\hat{z}(\rho)\right\|^2_{\mathcal{T}^{N,M}}(t) + \left|\tau^{N,M}-\tau\right|^2(t) \\
& \qquad \le \left(\left\|\hat{\phi}^M-\hat{\phi}\right\|_0^2 + \left|g^M(P_0^{N,M}\hat{\phi}^M) - g(\phi(0))\right|^2\right) e^{\int_0^T\omega(s)ds} + \int_0^T e^{\int_0^T\omega(s)ds} h^{N,M}(s)ds.
\end{aligned}
$$

We can use the fact that $g^M, g \in \tilde{\mathcal{T}}_g$ and our assumptions that the parameters converge to show that

$$
\left|g^M(P_0^{N,M}\hat{\phi}^M) - g(\phi(0))\right| \le k_1\left\|\hat{\phi}-\hat{\phi}^M\right\|_0 + k_2\left\|P^{N,M}\hat{\phi}-\hat{\phi}\right\|_{\mathcal{T}^{N,M}} + \left|g^M(\phi^M(0)) - g(\phi^M(0))\right|
$$

and now the statement of the theorem follows from the argument that $h^{N,M} \to 0$ as $N,M \to \infty$ in $L_1(0,T)$ (this argument is now essentially the same as that for the proof of Theorem 3) and the assumed convergence of the parameters. □

Theorem 6. Let $\{\rho^{N,M}\}$ with $\rho^{N,M} = (\gamma^{N,M}, g^{N,M}, \hat{\phi}^{N,M}) \in \mathcal{G} \times \mathcal{T}_g^M \times \mathcal{J} \subset \mathcal{G} \times \mathcal{T}_g \times \mathcal{J}$, be a sequence of solutions to $\{(P^{N,M})\}$. Then, given (HG), ($\hat{H}T_g$), (HI), and ($\hat{H}F$), there exist $\rho^* = (\gamma^*, g^*, \hat{\phi}^*) \in \mathcal{G} \times \mathcal{T}_g \times \mathcal{J}$ and a subsequence $\{\rho^{N_k,M_k}\}$ such that $\rho^{N_k,M_k} \to \rho^*$ with N_k, $M_k \to \infty$ (the meaning of this convergence statement is the same as that of Theorem 5), and ρ^* is a solution of (P).

Proof. As in the proofs of Theorems 2 and 4, the existence of the convergent subsequence is a consequence of compactness. That the limit ρ^* is a solution of (P) can be argued as follows. By definition of $\mathcal{T}_g^M$ we can write $g^{N,M} = I^M \tilde{g}^{N,M}$ with $\tilde{g}^{N,M} \in \tilde{\mathcal{T}}_g$. The compactness of $\tilde{\mathcal{T}}_g$ implies there exists a subsequence $\tilde{g}^{N_k,M_k}$ with $\tilde{g}^{N_k,M_k} \to g^* \in \tilde{\mathcal{T}}_g$ (this convergence is in the $W^{1,\infty}$ topology). We can now see that

$$\left| g^{N_k,M_k} - g^* \right|_{1,\infty} = \left| I^{M_k} \tilde{g}^{N_k,M_k} - g^* \right|_{1,\infty} \le \left| I^{M_k} \tilde{g}^{N_k,M_k} - \tilde{g}^{N_k,M_k} \right|_{1,\infty} + \left| \tilde{g}^{N_k,M_k} - g^* \right|_{1,\infty}$$

where $|\cdot|_{1,\infty}$ denotes the norm in $W^{1,\infty}$. Because all $g \in \tilde{\mathcal{T}}_g$ have a uniform bound on $|D^2 g|_\infty$, we can use equation (2.21) of [19] and the fact that $\tilde{g}^{N_k,M_k} \to g^*$ to conclude that $g^{N_k,M_k} \to g^*$. By definition $J^{N_k}(\rho^{N_k,M_k}) \le J^{N_k}(I^{M_k}\rho)$ for any $\rho \in \Pi$, where we are using the notation $I^{M_k}\rho = (\gamma, I^{M_k} g, \hat{\phi})$. We now obtain a proof of this theorem by letting N_k and $M_k \to \infty$, and applying Theorem 5.

5. Numerical Implementation.

The abstract equation (3.2) is equivalent to a system of ordinary differential equations, as we shall now demonstrate. Consider a fixed parameter $\rho \in \Pi$ (actually, Π^M). As the resulting system of differential equations is very similar in form, whether we are considering the time- or state-dependent delay case, let us write $\tau(t)$ with the understanding that if the delay is state-dependent, this represents the solution of the appropriate differential equation involving the state and the unknown parameter(s). At any time $t \in [0,T]$, $\hat{z}^N(t)$ belongs to the (N+1)-dimensional space $\mathbf{X}^N(\tau)$, which is equivalent (numerically) to $X^N(\tau)$; in the following, we will write $\hat{z}^N$, but it should be understood that we mean $\iota(\tau)\hat{z}^N$. Given that $X^N(\tau) = \mathrm{span}\{\hat{B}_i\}$, we can write $\hat{z}^N = \sum_{i=0}^{N} w_i(t)\hat{B}_i(s;\tau(t)) = \begin{bmatrix} w_0(t) \\ \sum_{i=0}^{N} w_i B_i \end{bmatrix}$,

and $\dot{\hat{z}}^N = \begin{bmatrix} \dot{w}_0(t) \\ \sum_{i=0}^{N}\left(\dot{w}_i B_i + w_i \frac{\partial B_i}{\partial t}\right) \end{bmatrix}$; here we have used the properties of the splines that $B_0(0)=1$ and $B_i(0)=0$ for $i \neq 0$, and $\frac{\partial B_0(0)}{\partial t}=0$. Since $\hat{z}^N(t)$, $\dot{\hat{z}}^N(t) \in X^N$, then $P^N \hat{z}^N(t) = \hat{z}^N(t)$ and $P^N \dot{\hat{z}}^N(t) = \dot{\hat{z}}^N(t)$. Thus equation (3.2) is equivalent to

$$\begin{aligned} &\langle\langle \dot{\hat{z}}^N(t), \hat{B}_i \rangle\rangle_{\mathcal{T}(t)} = \langle\langle \mathcal{A}^N(t;q)\hat{z}^N(t), \hat{B}_i \rangle\rangle_{\mathcal{T}(t)} \\ &\langle\langle \hat{z}^N(0), \hat{B}_i \rangle\rangle_{\mathcal{T}(0)} = \langle\langle \hat{\phi}, \hat{B}_i \rangle\rangle_{\mathcal{T}(0)} \end{aligned} \qquad \text{for } i=0,1,\ldots,N$$

and, using the representations above, we can write:

$$\begin{aligned} &Q(t)\,\underline{\dot{w}}(t) = H(t)\,\underline{w}(t) + T(t)\,\underline{w}(t) + \mathbb{F}(t,\underline{w}) \\ &Q(0)\,\underline{w}(0) = \mathbb{C} \end{aligned} \tag{5.1}$$

where $\underline{w} = (w_0, w_1, \ldots, w_N)^T$, Q, H, and T are $(N+1) \times (N+1)$ matrices with entries $Q_{i,j} = \langle\langle \hat{B}_i, \hat{B}_j \rangle\rangle_{\mathcal{T}} = B_i(0)B_j(0) + \int_{-\mathcal{T}(t)}^{0} B_i B_j ds$, $H_{i,j} = \langle B_i, DB_j \rangle_{\mathcal{T}} = \int_{-\mathcal{T}(t)}^{0} B_i DB_j ds$, and $T_{i,j} = -\langle B_i, \frac{\partial B_j}{\partial t} \rangle_{\mathcal{T}} = -\int_{-\mathcal{T}(t)}^{0} B_i \frac{\partial B_j}{\partial t} ds$, for $i,j = 0,1,\ldots,N$, and $\mathbb{F}$ and $\mathbb{C}$ are vectors in $\mathbb{R}^{N+1}$ with $\mathbb{F}(t,\underline{w}) = \left[F(t,q,w_0(t),\sum w_k(t)B_k),\ 0,\ \ldots,\ 0\right]^T$ and $\mathbb{C}_i = \langle\langle \hat{\phi}, \hat{B}_i \rangle\rangle_{\mathcal{T}(0)}$. It may appear at first that as we iterate on the unknown parameters, we must recompute the entries of our matrices each time; in fact, this is not the case. Let us define the "reference" basis functions $\{\tilde{B}_i\}$ by $\tilde{B}_i(x) = B_i\left(\frac{s}{-\mathcal{T}(t)}\right)$; while each B_i is defined on the interval $[-\tau(t),0]$, each $\tilde{B}_i$ is defined on the fixed interval $[0,1]$. These reference elements are the hat functions defined on a uniform subdivision of $[0,1]$ into N subintervals. Using these reference elements, we can rewrite the above expressions for our matrix elements as $Q_{i,j} = \tilde{B}_i(0)\tilde{B}_j(0) + \tau(t)\int_0^1 \tilde{B}_i \tilde{B}_j dx$, $H_{i,j} = -\int_0^1 \tilde{B}_i \mathfrak{D}\tilde{B}_j dx$, and $T_{i,j} = \dot{\tau}(t)\int_0^1 x\tilde{B}_i \mathfrak{D}\tilde{B}_j dx$. Notice the appearance of $\dot{\tau}(t)$. In the case of time-dependent delay, this represents a straightforward derivative of the unknown parameter. In the state-dependent cases, this represents the quantity

$\left(1 - \frac{g(w_0(t))}{g(w_N(t))}\right)\lambda$ or $\max\left(Dg(w_0(t))\cdot F(t,q,w_0(t),\sum w_k(t)B_k)\cdot\lambda\,,\,(1-\delta)\right)$. We now rewrite equation (5.1) as

$$
\begin{aligned}
&\tau(t)\tilde{Q}(t)\dot{\underline{w}}(t) = H\underline{w}(t) + \dot{\tau}(t)\tilde{T}(t)\underline{w}(t) + \mathbb{F}(t,\underline{w}) \\
&\tau(0)\tilde{Q}(0)\underline{w}(0) = \mathbb{C}
\end{aligned}
\tag{5.2}
$$

where H and $\tilde{T}$ are constant matrices which are three-banded due to the nature of the linear splines, and do not depend on the unknown parameters. The matrix $\tilde{Q}$, also three-banded, depends on τ in a simple way: $\tilde{Q}_{0,0} = \frac{1}{\tau(t)} + \int_0^1 \tilde{B}_0^2\,dx$, and no other entry of $\tilde{Q}$ depends on τ.

At this point, we would like to return to two outstanding questions from sections 3 and 4, namely our claims that J^N varies continuously in the parameters, and that our approximating systems have unique solutions.

Consider N and M fixed. We show that J^N is continuous on Π^M by showing the vector $\underline{w}(t)$ is continuous for each t. We can demonstrate this continuity by verifying that the differential equations (5.2) depend continuously on the parameters. The continuity on the parameters $\hat{\phi}$ and γ is clear from the simple dependence: The parameter γ appears only in $\mathbb{F}$ (recall f is continuous in γ by assumption) and the initial condition appears through a projection. The parameter τ also appears in a simple (multiplicative) way except that the derivative $\dot{\tau}$ also appears. For the time-dependent case, we consider only $\tau \in \mathcal{T}^M$, i.e., τ is a linear spline. For this special type of function, $|\dot{\tau}_1 - \dot{\tau}_2|_\infty$ can be bounded by $|\tau_1 - \tau_2|_\infty$ and thus, it follows directly that J^N varies continuously with τ in the C[0,T] topology. In the two cases of state-dependent delay, we consider the appended system, $(\dot{\underline{w}},\dot{\tau})$, and claim that this system of differential equations is continuous in $\bar{g}$ (for $\bar{g}$ varying in $C[\underline{x},\bar{x}]\times\mathbb{R}$) in the first class of state-dependent delays, or in g (for g varying in $W^{1,\infty}[\underline{x},\bar{x}]$) in the second class.

The existence and uniqueness of solutions follows from the fact that the right hand side of equation (5.2) (modified so that $\dot{\underline{w}}$, $\underline{w}(0)$ are alone on the left) combined with the right hand sides of the equations for $\dot{\tau}$ in the cases of state-

dependent delay, are Lipschitz continuous in $\underline{w}$, or $(\underline{w},\tau)$. The verification of this continuity is straightforward.

We now turn to some computational concerns. The time-consuming computations (at least on a conventional computer, as opposed to a supercomputer) are the numerical quadratures required for the inner product evaluations. Using the reference elements defined above, these quadratures can be evaluated and stored once for any fixed N. As we iterate on the parameters, these quadratures are reused. If, as is the case in the examples we present in the next section, we are only attempting to estimate an unknown delay function, then this method is very efficient (in this case, the elements of $\mathbb{C}$ are also computed only once for a given N).

The simple dependence of the matrices Q and H on the delay was acknowledged in [2], and exploited in [3] (the use of the reference elements to avoid recomputing quadratures provides a significant computational savings in the estimation of constant delays as well as in the estimation of variable delays), however, the explicit use of these reference elements was first carefully stated and formalized in the context of estimation of unknown variable coefficients in partial differential equations in [12].

Focusing on the unknown τ, we make some observations regarding the implementation with the constraint set $\mathcal{T}_c^M$. Any function $\tau\in\mathcal{T}_c^M$ has a representation of the form $\tau = \sum_{i=0}^{M}\eta_i b_i(t)$. Thus, the estimation of $\tau\in\mathcal{T}_c^M$ is equivalent to the estimation of $\underline{\eta} = (\eta_0,\eta_1,\ldots,\eta_M)$ in $\mathbb{R}^{M+1}$. The constraints on τ are readily converted to constraints on $\underline{\eta}$. The requirement that $\tau_0 \le \tau(t) \le r$ $\forall t\in[0,T]$ translates to $\tau_0 \le \eta_i \le r$ for each $i = 0,1,\ldots,M$. The requirement that $-\mu \le \dot{\tau}(t) \le 1-\delta$ becomes $\frac{-\mu T}{M} \le (\eta_{i+1}-\eta_i) \le \frac{(1-\delta)T}{M}$ for each $i = 0,1,\ldots,M$.

Similarly, for the estimation of a state-dependent delay of the first class, our unknown is $g\in\mathcal{T}_g^M$, having the form $g = \sum_{i=0}^{M}\eta_i b_i(x)$. The corresponding constraints on $\underline{\eta}$ are $g_L \le \eta_i \le g^U$ and $\frac{-\mu_g(\bar{x}-\underline{x})}{M} \le (\eta_{i+1}-\eta_i) \le \frac{\mu_g(\bar{x}-\underline{x})}{M}$ for each $i = 0,1,\ldots,M$.

As noted earlier, the constraints on the parameter set for the second class of

state-dependent delay are not easily characterized, and therefore we cannot translate them into corresponding constraints on the approximations. We can, however, proceed with our estimation procedure without implementing all the constraints, and may still be successful (see, eg., example 6, below and [14]).

As a final comment, we note that in order to prove our convergence theorems for both cases of state-dependent delay, we impose constraints on the approximate delays by multiplying the differential equations which they satisfy by $\lambda^{N,M}$; this is done in order to assure that each $\tau^{N,M}$ belongs to $\mathcal{T}_c$. Our convergence theorems then guarantee that the approximate delays, $\tau^{N,M}$, converge to the "true" delay τ, which satisfies the constraints of $\mathcal{T}_c$ in the first instance (i.e., no externally imposed constraints are used). We thus conclude that, for large enough N and M, the constraints are not actually used, i.e., for large enough N and M, $\lambda^{N,M} \equiv 1$, and therefore can be left out. This is not the case for the constraint in the second class of delays, used to keep $\dot{\tau}^{N,M} \leq 1 - \delta$. This constraint must, in theory, be imposed (in our numerical example, however, we have not imposed it and were still able to succesfully identify the parameter - this may well be due to the fact that our test example is "nice", i.e., it has an exact solution, which may be unique).

The translation of the constraints of $\mathcal{T}_c$ and $\mathcal{T}_g$ into implementable form within $\mathcal{T}_c^M$ and $\mathcal{T}_g^M$ is clearly possible when these sets are based on linear splines. It is not clear how this could be done with higher order splines. For the state approximation, however, one could use higher order splines with few changes in the numerical scheme described above. The matrices would have more bands and evaluation of z_0^N would not be quite as simple, but one could still make use of reference elements to avoid excessive recomputations of quadratures.

6. Numerical Test Examples. In all examples, we estimate one unknown delay, and in some an unknown constant appearing in the function f. We use a Levenberg-Marquardt algorithm (the subroutine ZXSSQ in the IMSL library) to solve the minimization problem $(P^{N,M})$. For each iterate of q, we solve system (5.2) using DGEAR (also from IMSL, a stiff ordinary differential equation solver, based on Gear's method). The evaluation of J^N requires only $w_0(t_i)$. While $\mathcal{T}$, and hence also $\mathcal{T}^M$ carry derivative constraints, we did not impose these (this makes it easier to

implement the scheme for testing purpses); all of our examples work without the constraints, probably because they have exact solutions (and in many cases, solutions may well be unique).

In each case, we have a known "true" delay and a corresponding analytical solution to equation (2.1). We generate data by evaluating our solution at various time values. We choose a value for N and M and make an initial guess for τ^M, $\bar{g}^M$, or g^M (and possibly γ); since we do not presume to have any knowledge of the shape of the unknown functions, we always guess a constant. For a true problem, we might run our algorithm for many values of N and M. In most of our test examples, we get a very good match with N=30 and fairly low values of M. Thus, in many cases we only report the results for one such run.

We present a representative sample of test problems below, chosen to highlight the capabilities and interesting limitations of the method discussed here. Further examples can be found in [14].

<u>Example 1.</u> This work was motivated by problems in biology, where the time-varying delay changes in a periodic fashion. The delay of our first example was chosen to be simple (piecewise linear) but exhibits this oscillatory behavior.

Our model equation is

$$\dot{x}(t) = b\,x(t-\tau(t)) - d\,x(t) \qquad \text{for } 0 \le t \le 1.5$$
$$x_0(s) = e^{-s} \qquad \text{for } -\tau(0) \le s \le 0.$$

We have chosen the true $\tau(t)$ to be the function which increases linearly from $(t,\tau(t)) = (0,0.25)$ to $(0.5,0.5)$, then decreases linearly to $(1.0,0.25)$, then increases again to $(1.5,0.5)$ (see, e.g., "True tau" in Figure 1A). We also have an unknown $\gamma = (b,d)$ with true value $(4.0,2.0)$. An analytical solution can be found with the method of steps, and was evaluated at 15 t-values between 0 and 1.5 to produce the "data".

1A. We have approximated the states with linear splines using N=30, and the delay function with linear splines and M=3. We made an initial guess of $\tau \equiv 0.2$ and $\gamma = (1.0,3.0)$. Our best fit for $\tau^{30,3}$ is graphed in Figure 1A. We estimated $\gamma^{30,3} = (3.906,1.941)$. Our residual is $J^{30}(\rho^{30,3}) = 0.360\times10^{-6}$.

1B. We repeated the same example problem, but after adding random noise to

the data. The noisy data $\tilde{\chi}_i$ is obtained from the original data χ_i by setting

$$\tilde{\chi}_i = N\sigma|\chi_i| + \chi_i$$

where N represents a random number from a standard normal distribution (zero mean and unit variance), and σ is the level of noise. In this example, we chose $\sigma = 0.025$ (i.e., we have introduced noise at a level of 2.5%). We have estimated only the delay, beginning with the initial guess of 0.3. The best estimate, $\tau^{30,3}$, for this noisy data is plotted in Figure 1B. Our residual in this case is $J^{30}(\tau^{30,3}) = 0.170\times10^{-2}$.

1C. Here we have used the same example problem, with the "pure" data, but now with M=6. The choice of M=3 was a best case in the sense that the nodes in the true τ coincided with the nodes of our approximation scheme. Clearly we have no hope of success for estimation of such a τ if our nodes do not coincide, but a better test of our method uses more than enough nodes, thus allowing the possibility of estimating a choppier τ than the actual. As can be seen from Figure 1C, overestimating the number of nodes does not cause over-oscillation in the estimate for τ. Our residual is $J^{30}(\tau^{30,6}) = 0.487\times10^{-8}$.

<u>Example 2.</u> In this example, our true delay exhibits oscillatory behavior similar to that of Example 1, but is not piecewise linear. We first chose the "true" delay to be $\tau(t) = \frac{1}{2} + \sin\left[\frac{1}{2}t\right]$, and a true solution of $x(t) = 2 - (t-1)^2$. We then chose our model dynamics to be

$$\dot{x}(t) = 2x(t-\tau(t)) - x(t) + g(t) \qquad \text{for } 0 \le t \le 7.0$$

$$x_0(s) = s + 1 \qquad \text{for } -\tau(0) \le s \le 0$$

where we compute g(t) so that the above equation holds. We evaluated our true solution at 15 time values for our data. We approximated the states with linear splines with N=30, and the delay with linear splines and M=7. We used an initial guess for $\tau \equiv 1.0$. Our estimate, along with the true function and initial guess, is graphed in Figure 2. Our residual for this example is $J^{30}(\tau^{30,7}) = 0.141\times10^{-3}$.

<u>Example 3.</u> In [9], [15], and [16], several examples of state-dependent delay differential equations, with analytical solutions are presented. We have tested our algorithm using modifications of these examples (in some cases we have shifted the time scale, in others we have rewritten their state-dependent delay as a time-

varying delay, etc.). One such modification (of Example 5.1 in [9]) is:

$$\dot{x}(t) = \frac{1}{2\sqrt{t+1}}\, x(t-\tau(t)) \qquad \text{for } 0 \le t \le 2$$

$$x_0(s) = 1 \qquad \text{for } -\tau(0) \le s \le 0.$$

The true τ is given by:

$$\tau(t) = \begin{cases} t - \sqrt{t+1} + \sqrt{2} & \text{for } 0 \le t \le 1 \\ \frac{3}{4}(t-1) - (1-\frac{1}{2}\sqrt{2})\sqrt{t+1} + \sqrt{2} & \text{for } 1 \le t \le 2. \end{cases}$$

We tried to estimate $\tau(t)$, but found our optimization algorithm would never converge. In fact, the algorithm would match τ for later values of t, but would wander with values of τ for low values of t. In terms of our spline approximations, $\tau^M = \sum_{i=0}^{M} \eta_i b_i$, we would see $\eta_1, \eta_2, \ldots, \eta_M$ all approach values close to the interpolation values for the true τ, but the algorithm could not locate a "good" value for η_0. For all candidates for η_0 that the algorithm tried, we would see a small residual. This suggests nonuniqeness. This example illustrates an important (yet obvious, after the fact) limitation of any parameter estimation routine. With a constant initial function, any τ will provide a correct fit to data for early times; i.e., there is no unique solution to the inverse problem, in the sense that a whole class of functions will fit the data. Let τ^* represent the true τ; any function $\tilde{\tau}$ satisfying $t-\tilde{\tau}(t) \le 0$ for all t for which $t-\tau^*(t) \le 0$ can match the data exactly, since even though (for these times) $t-\tilde{\tau}(t) \ne t-\tau^*(t)$, as long as both values are negative, the same constant value of x is used.

We tried the following experiment to verify the above discussion. We chose M=4, set $\eta_0 = I^M\tau(0)$ and $\eta_1 = I^M\tau(.5)$, and held these fixed. We then tried to estimate τ on the interval [.5,2], i.e., we searched for η_2, η_3, and η_4. With initial guess $\eta_i = 0.5$ for each i = 2,3,4, we were able to succesfully estimate τ, with a residual of $J^{30} = 0.345 \times 10^{-8}$. See Figure 3 for plots of τ (true, initial, and estimated).

<u>Example 4.</u> We repeat the model dynamics of the above example, however, having learned the problems a constant initial function creates, we modified the equation accordingly. Thus, our model equation is:

$$\dot{x}(t) = \frac{1}{2\sqrt{t+1}} x(t-\tau(t)) + g(t) \qquad \text{for } 0 \le t \le 2$$

$$x_0(s) = s + 1 \qquad \text{for } -\tau(0) \le s \le 0.$$

We chose a true solution, then evaluated g(t) so that the above equation holds. Our true τ, motivated by an interesting delay found in [16], is given by:

$$\tau(t) = \begin{cases} t-(t-1)^3 & \text{for } 0 \le t \le 1 \\ t-1+\frac{1}{\sqrt{2(t-1)^3+1}} & \text{for } 1 \le t \le 2. \end{cases}$$

With data at 14 time values between 0 and 2, N=30 and M=6, and an initial guess for τ of $\tau \equiv 1$, we obtained the estimate pictured in Figure 4. Our residual for this example was $J^{30}(\tau^{30,6}) = 0.311 \times 10^{-6}$.

<u>Example 5.</u> In this example we estimate a delay which is state-dependent of the first class discussed in the body of the paper. The model equation is:

$$\dot{x}(t) = \tfrac{1}{4}\exp\left[-x(t-\tau(t))\right] \qquad \text{for } 0 \le t \le 6.0$$

$$x_0(s) = \ln(\tfrac{1}{2}s + 3) \qquad \text{for } -\tau(0) \le s \le 0$$

$$\dot{\tau}(t) = 1 - \frac{g(x(t))}{g(x(t-\tau(t)))} \qquad \text{for } 0 \le t \le 6.0$$

$$\tau(0) = \tau_i$$

where we have chosen the true parameters $g(x) = \exp(-x)$, and $\tau_i = 3$. An analytical solution for this example is given by $x(t) = \ln\left[\tfrac{1}{2}t + 3\right]$ and $\tau(t) = \tfrac{1}{2}t + 3$. Our data is collected at 15 values of time between 0 and 6. We have estimated both τ_i and g; beginning with initial values of $\tau_i^0 = 1.0$ and $g^0 \equiv 0.1$, we obtained estimated values of $\tau_i^{30,3} = 3.007$ and g as depicted in Figure 5. Our residual is $J^{30}(\bar{g}^{30,3}) = 0.130 \times 10^{-7}$.

<u>Example 6.</u> Our final example is a nonlinear delay differential equation in which we estimate a state-dependent delay of the second type. We use a slight

modification of Example 1.1 of [16]. The model equation is:

$$\dot{x}(t) = \frac{1}{t+e} x(t) x(t - \tau(x(t))) \qquad \text{for } 0 \leq t \leq 2.3$$

$$x_0(s) = s + e \qquad \text{for } -\tau(e) \leq s \leq 0$$

and our true τ is given by $\tau(t) = g(x(t))$ with $g(x) = (e-1)\ln x$. We use data for 15 values of time between 0 and 2.3. As in the previous examples, we use linear splines to approximate the states, with N=30. The analytical solution is given by $x(t) = e^{(t+e)/e}$. For this example, we choose [e,e+4] as the state interval and approximate g here using M=5. Our initial guess for g was $g^0 \equiv 2$. In Figure 6, we have plotted the initial guess, and the true and approximate functions, g. Our residual is $J^{30}(g^{30,5}) = 0.275 \times 10^{-6}$.

Acknowledgments. The author is indebted to Joe Pedit for invaluable assistance with the figures. The continued support and encouragement of H.T. Banks is always appreciated.

This paper is dedicated to the memory of P. Roy and D.L. Mantey.

References.

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

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

[3] H.T. Banks and P.K. Daniel Lamm, Estimation of delays and other parameters in nonlinear functional differential equations, SIAM J. Control Optim., 21 (1983), 895-915.

[4] H.T. Banks and K.A. Murphy, Estimation of coefficients and boundary parameters in hyperbolic systems, SIAM J. Control Optim., 24 (1986), 926-950.

[5] H.T. Banks and K.A. Murphy, Quantitative modeling of growth and dispersal in population models, *in* Proc. International Symp. on Math. Biol., Kyoto, Nov. 10-15, 1985 (in press).

[6] H.T. Banks and K.A. Murphy, Estimation of nonlinearities in parabolic models for growth, predation and dispersal of populations, J. Math. Anal. Appl., to appear.

[7] R.D. Driver, A two-body problem of classical elecrodynamics: the one-dimensional case, Annals of Physics, 21 (1963), 122-142.

[8] R.D. Driver, Existence theory for a delay-differential system, Contributions to Diff. Eqns., 1 (1961), 317-336.

[9] A. Feldstein and K.W. Neves, High order methods for state-dependent delay differential equations with nonsmooth solutions, SIAM J. Numer. Anal., 21 (1984), 844-863.

[10] J.A. Gatica and P. Waltman, A threshold model of antigen antibody dynamics with fading memory, *In* Nonlinear Phenomena in Mathematical Sciences (V. Lakshmikantham, ed.), Academic Press, New York, NY, 1982, 425-439.

[11] J.A. Gatica and P. Waltman, Existence and uniqueness of solutions of a functional differential equation modeling thresholds, Nonlinear Anal., Theory, Methods & Appl., 8 (1984), 1215-1222.

[12] P.K. Lamm, Estimation of discontinuous coefficients in parabolic systems: applications to reservoir simulation, SIAM J. Control Optim., to appear.

[13] P.E. Lutz, Effects of temperature and photoperiod on larval development in *lestes eurinus* (odonata: lestidae), Ecology, 49 (1968), 637-644.

[14] K.A. Murphy, Time-dependent approximation schemes for some problems of parameter estimation in distributed systems, *In* Proc. 24th IEEE Conf. on Decision and Control, Los Angeles, Dec. 9-11, 1987, 165-169.

[15] K.W. Neves, Automatic integration of functional differential equations: an approach, ACM Trans. Math. Software, 1 (1975), 357-368.

[16] K.W. Neves and A. Feldstein, Characterization of jump discontinuities for state dependent delay differential equations, J. Math. Anal. Appl., 56 (1976), 689-707.

[17] R.M. Nisbet and W.S.C. Gurney, The systematic formulation of population models for insects with dynamically varying instar duration, Theor. Pop. Biol., 23 (1983), 114-135.

[18] P.M. Prenter, Splines and Variational Methods, Wiley-Interscience, New York, NY, 1975.

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

[20] G.F. Webb and M. Badii, Representation of solutions of functional differential equations with time-dependent delays, Houston J. Math., 6 (1980), 577-594.

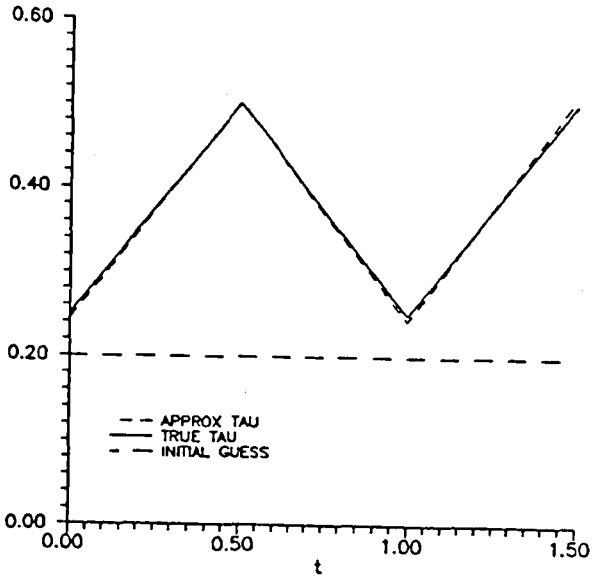

FIGURE 1A. Estimation of Tau, M=3

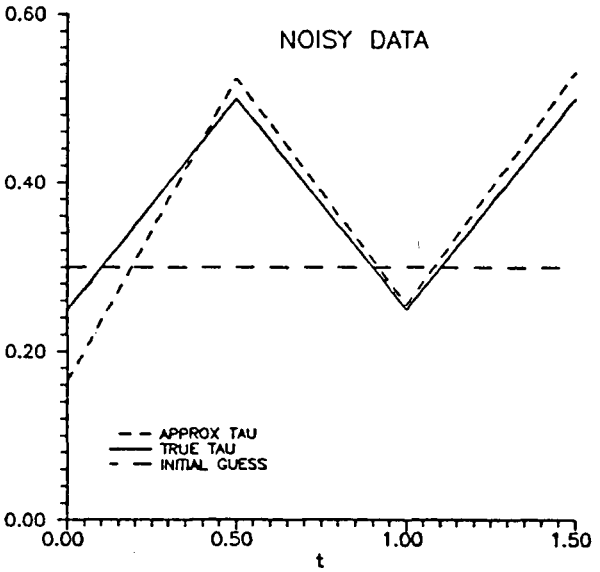

FIGURE 1B. Estimation of Tau, M=3

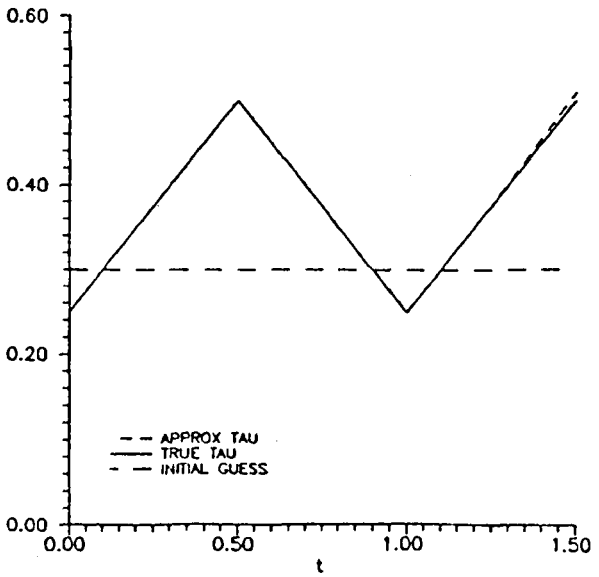

FIGURE 1C. Estimation of Tau, M=6

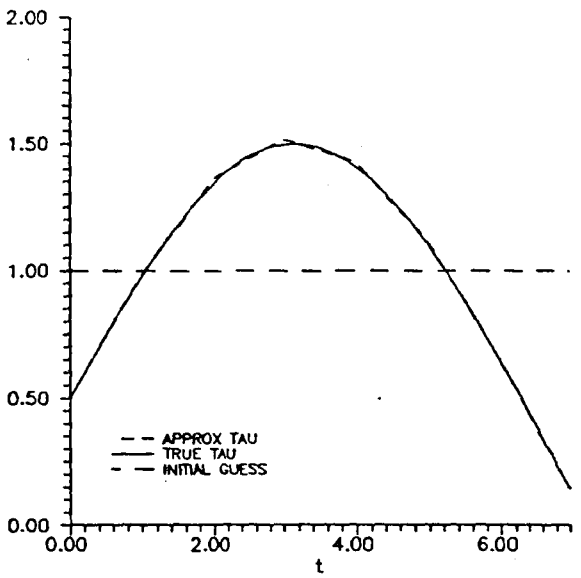

FIGURE 2. Estimation of Tau, M=7

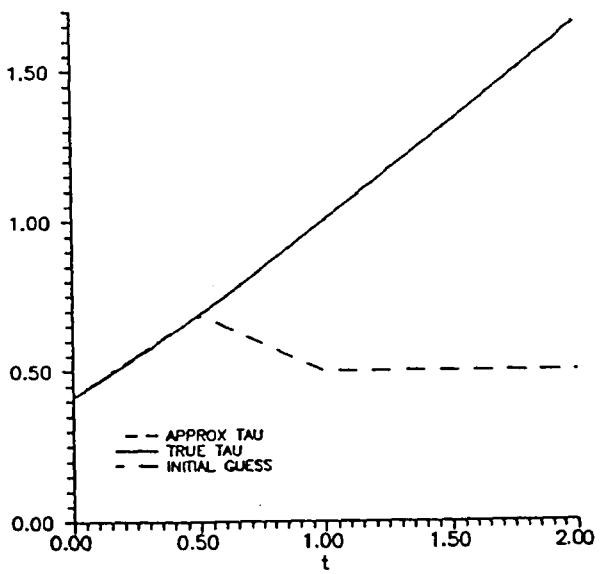

FIGURE 3. Estimation of Tau, M=4

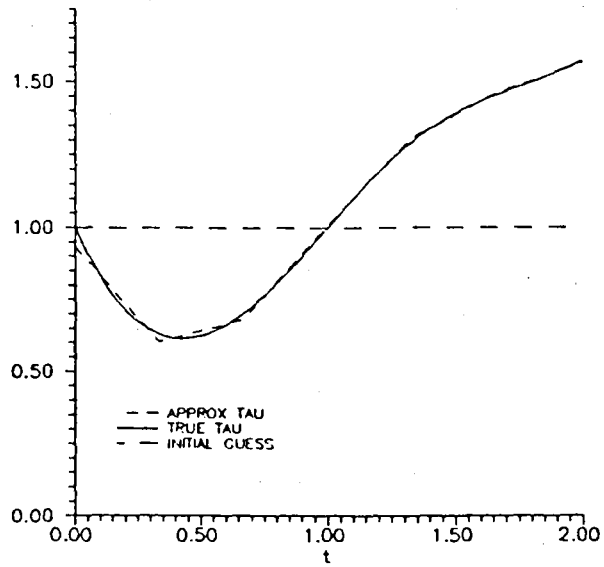

FIGURE 4. Estimation of Tau, M=6

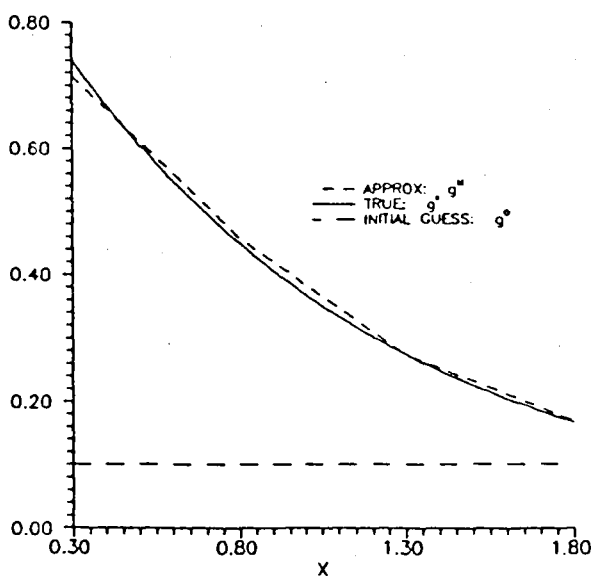

FIGURE 5. Estimation of g, M=3

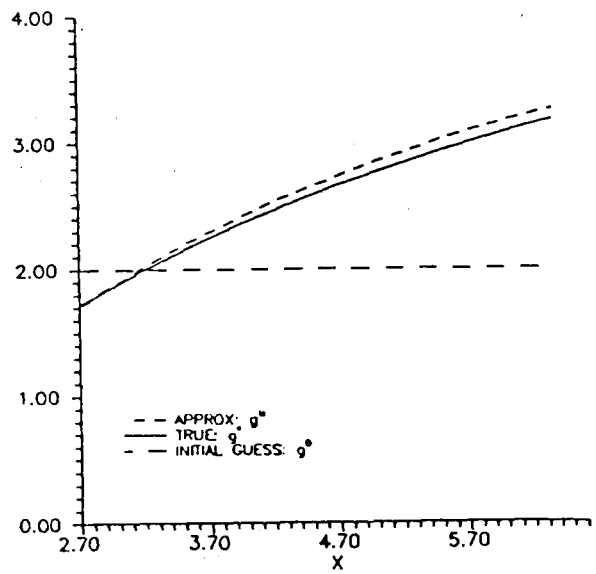

FIGURE 6. Estimation of g, M=5

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-181663 ICASE Report No. 88-25		

4. Title and Subtitle	5. Report Date
ESTIMATION OF TIME- AND STATE-DEPENDENT DELAYS AND OTHER PARAMETERS IN FUNCTIONAL DIFFERENTIAL EQUATIONS	April 1988
	6. Performing Organization Code
7. Author(s) K. A. Murphy	8. Performing Organization Report No. 88-25
	10. Work Unit No. 505-90-21-01
9. Performing Organization Name and Address Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23665-5225	11. Contract or Grant No. NAS1-18107
	13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Langley Research Center Hampton, VA 23665-5225	14. Sponsoring Agency Code

15. Supplementary Notes

Langley Technical Monitor:
Richard W. Barnwell

Final Report

Submitted to SIAM J. Applied Math.

16. Abstract

We develop a parameter estimation algorithm which can be used to estimate unknown time- or state-dependent delays and other parameters (e.g., initial condition) appearing within a nonlinear nonautonomous functional differential equation. The original infinite dimensional differential equation is approximated using linear splines, which are allowed to move with the variable delay. The variable delays are approximated using linear splines as well. The approximation scheme produces a system of ordinary differential equations with nice computational properties. The unknown parameters are estimated within the approximating systems by minimizing a least-squares fit-to-data criterion. We prove convergence theorems for time-dependent delays and state-dependent delays within two classes, which say essentially that fitting the data by using approximations will, in the limit, provide a fit to the data using the original system. We present numerical test examples which illustrate the method for all types of delay.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
parameter estimation, functional differential equations, variable delays, approximations	59 - Mathematical and Computer Sciences (General) 64 - Numerical Analysis Unclassified - unlimited

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	51	A04

NASA FORM 1626 OCT 86

NASA-Langley, 1988