NASA CR-172,155

NASA Contractor Report 172155

ICASE

NASA-CR-172155
19830021837

LEGENDRE-TAU APPROXIMATIONS FOR FUNCTIONAL DIFFERENTIAL EQUATIONS

Kazufumi Ito

and

Russell Teglas

Contract No. NAS1-17070
June 1983

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

Operated by the Universities Space Research Association

NASA
National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

NF02031

LEGENDRE-TAU APPROXIMATIONS

FOR FUNCTIONAL DIFFERENTIAL EQUATIONS

Kazufumi Ito
Institute for Computer Applications in Science and Engineering

and

Russell Teglas
Institute for Computer Applications in Science and Engineering

Abstract

In this paper we consider the numerical approximation of solutions to linear functional differential equations using the so-called Legendre-tau method. The functional differential equation is first reformulated as a partial differential equation with a non-local boundary condition involving time-differentiation. The approximate solution is then represented as a truncated Legendre series with time varying coefficients which satisfy a certain system of ordinary differential equations. The method is very easy to code and yields very accurate approximations. Convergence is established, various numerical examples are presented, and comparison between the latter and cubic spline approximations is made.

This research was supported by the National Aeronautics and Space Administration under the NASA Contract No. NAS1-17070 while the authors were in residence at ICASE, NASA Langley Research Center. Hampton, VA 23665.

N83-30108#

Introduction

In this paper we consider Legendre-tau approximations of solutions of functional differential equations (FDEs). The tau method, invented by Lanczos in 1938 [12], is one of several approximation techniques which are referred to as a spectral method [9]. Spectral methods have been used in numerical computations for a wide class of partial differential equations (PDEs). In this paper, we view the original FDE as a PDE $u_t = u_x$ with boundary conditions which involve time differentiation. The Legendre-tau method is based upon representing the approximate solution as a truncated series of Legendre polynomials. The evolution equation for the expansion coefficients is then determined by substituting the series into the above PDE and by imposing the boundary conditions. To our knowledge, the use of the tau method in approximating solutions of FDEs and the specific manner in which it is used (i.e., applying the tau method to a reformulated PDE with boundary conditions involving time evolution) are new. The idea of formulating FDEs as Cauchy problems on an appropriate Hilbert space is not new. Within this framework, Banks and Kappel [2] make use of approximation results from linear semigroup theory (in particular, the Trotter-Kato theorem) to establish the convergence of numerical schemes based upon splines.

In this paper our considerations are restricted to linear autonomous FDEs of retarded type. The use of our ideas for (i) nonautonomous, (ii) nonlinear, (iii) neutral-type or (iv) integro-differential systems will be discussed elsewhere. We add here that the tau method should prove useful in dealing with partial differential equations with boundary conditions which involve time differentiation. Our main goal is the application of the tau method to optimal control and parameter estimation problems. As will be discussed in section 7 the tau method may offer considerable improvements over other methods (e.g.. those discussed in [2], [3]) in many instances. One reason for this is that the semigroup $\{S(t) : t > 0\}$ associated with a retarded FDE has the property that the range of $S(t)$ is contained in $\mathcal{D}(A^k)$ for each $t > kr$ where A is the infinitesimal generator of $\{S(t) : t > 0\}$ with domain $\mathcal{D}(A)$ and r is the longest delay time appearing in the FDE. Thus, the regularity of solutions increases with time. In section 3, we will see that, in such a case, approximations by orthogonal polynomials are quite powerful.

The following is a brief summary of the contents of this paper. In section 2, we review the equivalence results between FDEs and abstract Cauchy problems on the product space $\mathbb{R}^n \times L_2$. In section 3, we recall various properties of Legendre polynomials including certain estimates which are needed to establish convergence. Section 4 is concerned with the developemnt of the numerical scheme based upon the Legendre-tau approximation for solving the class of FDEs under consideration. In order to establish the numerical convergence of such approximations to the actual solution in the $\mathbb{R}^n \times L_2$ norm, we first show, in section 5, that convergence holds in a stronger norm under the special assumptions that the initial data is sufficiently regular and that inhomogeneous forcing term f is identically zero. In section 6, we then extend our result to the inhomogeneous case wherein $f \in L_2^{loc}$ and show that the sequence obtained by truncating the last term in each of the Legendre-tau approximations converges to the actual solution in the $\mathbb{R}^n \times L_2$ norm whenever the initial data lies in $\mathbb{R}^n \times L_2$. Finally, in section 7, we present numerical results and compare these results for cubic spline approximations discussed in [2].

Throughout this paper the following notation will be used. $r > 0$ stands for the longest delay time appearing in the FDE. The Hilbert space of $\mathbb{R}^n$-valued square integrable functions on the interval [a.b] is denoted by

$L_2([a,b];\mathbb{R}^n)$. When the underlying space and interval can be understood from the context, we will abbreviate the notation and simply write L_2. $L_2^{loc}([0,\infty);\ \mathbb{R}^n)$, or L_2^{loc}, is the space of $\mathbb{R}^n$-valued locally square integrable functions on the semi-infinite interval $[0,\infty)$. H^k is the Sobolev space of $\mathbb{R}^n$-valued functions f on a compact interval with $f^{(k-1)}$ absolutely continuous and $f^{(k)} \in L_2$. The Banach space of $\mathbb{R}^n$-valued continuous functions on the interval $[-r,0]$ is denoted by C. We denote by Z the product space $\mathbb{R}^n \times L_2([-r,0];\mathbb{R}^n)$. Given an element $z \in Z$, $\eta \in \mathbb{R}^n$ and $\phi \in L_2$ denote the two coordinates of $z : z = (\eta\ \phi)$.The bracket $\langle\cdot,\cdot\rangle_H$ stands for the inner product in the Hilbert space H, and the subscript for the underlying Hilbert space will be omitted when understood from the context. $\|\cdot\|$ denotes the norm for elements of a Banach space and for operators between Banach spaces, while $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$.

If X and Y are Banach spaces, then the space of bounded operators from X to Y is denoted by $\mathcal{L}(X,Y)$. $\mathcal{D}(A)$ denotes the domain of a linear operator A . χ_I denotes the characteristic function of the interval I. Given a measurable function $x : [-r,\infty) \to \mathbb{R}^n$ and $t > 0$, the function $x_t : [-r,0] \to \mathbb{R}^n$ is defined by $x_t(\theta) = x(t+\theta)$, $\theta \in [-r,0]$. Finally, for any function ϕ of the independent variable θ, we shall use either $\dot{\phi}$ or $\frac{\partial \phi}{\partial \theta}$ to denote the derivative of ϕ with respect to θ.

2. <u>Linear Hereditary Differential Equations</u>

In this section, we state the type of equations to be considered and recall some results for such equations which are important for the discussion to follow.

Given $(\eta,\phi) \in Z$ and $f \in L_2^{loc}([0,\infty);\mathbb{R}^n)$, we consider the functional equation

$$x(t) = \eta + \int_0^t Dx_s\, ds + \int_0^t f(s)ds$$
$$x_0 = \phi, \tag{2.1}$$

where $D : \mathcal{D}(D) \subseteq L_2([-r,0];\mathbb{R}^n) \to \mathbb{R}^n$ has the form

$$D\phi = \int_{-r}^0 d\mu(\theta)\phi(\theta) \tag{2.2}$$

with μ a matrix-valued function of bounded variation on $[-r,0]$. As an example, consider

$$\mu(\theta) = \sum_{i=0}^m A_i\, \chi_{(\theta_i,0]}(\theta) + \int_{-r}^{\theta} A(s)ds. \tag{2.3}$$

where $-r = \theta_m < \cdots < \theta_0 = 0$ and A_i and $A(\cdot)$ are $n\times n$ matrices, the elements of the latter being integrable on $[-r,0]$. Then

$$Dx_t = \sum_{i=0}^m A_i x(t+\theta_i) + \int_{-r}^0 A(\theta)x(t+\theta)d\theta, \qquad t \geq 0. \tag{2.4}$$

For $(\eta,\phi) \in Z$ and $f \in L_2^{loc}$, there exists a unique solution x to (2.1) in $L_2([-r,T];\mathbb{R}^n)$ for any $T > 0$ such that

$$z(t) = \big(x(t),x_t\big) \in Z, \tag{2.5}$$

is a Z-valued continuous function which depends continuously on $(\eta,\phi) \in Z$ and $f \in L_2^{loc}$ for each $t \geq 0$ (see [3], [5], [8], e.g.). Note that $x(\cdot)$ is absolutely continuous on $[0,T]$ and satisfies the differential version of equation (2.1):

$$\frac{dx}{dt}(t) = Dx_t + f(t), \qquad t > 0 \tag{2.6}$$

$$x_0 = \phi.$$

For $t > 0$, define $S(t) : Z \to Z$ by $S(t)(\eta,\phi) = (x(t),x_t)$ where x is the homogeneous solution of (2.1) (i.e., $f \equiv 0$). Then $\{S(t) : t > 0\}$ forms a strongly continuous semigroup on Z. The following results are now standard if one deals with FDEs in the state space Z ([3], [4], [15]).

Lemma 2.1:

(i) If A denotes the infinitesimal generator of $\{S(t) : t > 0\}$, then

$$\mathcal{D}(A) = \left\{(\eta,\phi) \in Z : \dot{\phi} \in L_2 \text{ and } \eta = \phi(0)\right\}$$

and for $\left(\phi(0),\phi\right) \in \mathcal{D}(A)$,

$$A\left(\phi(0),\phi\right) = (D\phi,\dot{\phi}).$$

(ii) If $z(0) = \left(\phi(0),\phi\right) \in \mathcal{D}(A)$, and $f \in L_2^{loc}$, then $z(t) = S(t)z(0) + \int_0^t S(t-s)\ f(s)ds$ satisfies the equation

$$\frac{dz}{dt}(t) = Az(t) + f(t), \qquad t > 0, \tag{2.7}$$

in Z where $\mathcal{B}: \mathbb{R}^n \to Z$ is defined by

$$\mathcal{B}x = (x,0) \in Z \qquad \text{for } x \in \mathbb{R}^n.$$

(iii) For $k > 2$,

$$\mathcal{D}(A^k) \subseteq \{(\phi(0),\phi) \in Z : \dot{\phi}(0) = D\phi \text{ and } \phi \in H^k\},$$

and is dense in Z.

3. Properties of Legendre Polynomials

In this section, we review some properties of Legendre polynomials ([11], [13], e.g.).

The Legendre polynomial of degree n, $p_n(x)$, $-1 < x < 1$, can be defined as the solution of the differential equation

$$\frac{d}{dx}\left((1-x^2)\frac{dp}{dx}(x)\right) + n(n+1)p(x) = 0, \tag{3.1}$$

which satisfies $p(1) = 1$. Thus, $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = \frac{1}{2}(3x^2-1)$, and so on. The Legendre polynomials $\{p_k\}_{k>0}$ satisfy the orthogonality relation

$$\int_{-1}^{1} p_m(x)p_n(x)dx = \frac{2}{2n+1}\delta_{mn}, \tag{3.2}$$

and they form a basis for $L_2(-1,1)$: any $f \in L_2(-1,1)$ can be written as

$$f = \sum_{n>0} f_n p_n$$

where

$$f_n = \frac{2n+1}{2}\int_{-1}^{1} f(x)p_n(x)dx$$

with

$$\|f\|_{L_2}^2 = \sum_{n>0} \frac{2}{2n+1} f_n^2 .$$

They possess the recursion formula

$$(n+1)p_{n+1}(x) = (2n+1)xp_n(x) - np_{n-1}(x). \tag{3.3}$$

From this, we have

$$p_n(\pm 1) = (\pm 1)^n, \qquad |p_n(x)| < 1, \qquad |x| < 1,$$

and

$$\dot{p}_n(\pm 1) = (\pm 1)^n n(n+1)/2.$$

If f is represented as

$$f = \sum_{n=0}^{N} f_n p_n,$$

then

$$\dot{f} = \sum_{n=0}^{N-1} b_n p_n \quad \text{where} \quad b_n \equiv (2n+1) \sum_{\substack{k=n+1 \\ k+n \text{ odd}}}^{N} f_k. \tag{3.4}$$

For any positive integer N, let P^N be the orthogonal projection of L_2 onto the subspace spanned by $\{p_k\}_{k=0}^{N}$. Then we have the following error estimates [6]:

Lemma 3.1: <u>For any real</u> $s > 0$, <u>there exists a constant</u> K <u>such that</u>

$$\|f - P^N f\|_{L_2} < KN^{-s} \|f\|_{H^s} .$$

Lemma 3.2: <u>For any real</u> s <u>and</u> σ <u>such that</u> $1 < s < \sigma$, <u>there exists a constant</u> K <u>such that</u>

$$\|f - P^N f\|_{H^s} < KN^{2s-\sigma-1/2} \|f\|_{H^\sigma} .$$

The next lemma gives an error estimate in the supremum norm:

Lemma 3.3: <u>For any positive integer</u> m <u>there exists a constant</u> K <u>such that</u>

$$|f(x) - P^N f(x)| < KN^{-2m+1} \|f\|_{H^{2m}} .$$

<u>Proof:</u> Let us denote by L the differential operator

$$(L f)(x) = \frac{d}{dx}\left((1-x^2)\, \frac{df}{dx}\,(x)\right).$$

From (3.1), we have

$$\int_{-1}^{1} f(x)p_n(x)d = \left(\frac{1}{n(n+1)}\right)^m \int_{-1}^{1} (L^m p_n)(x)f(x)dx, \qquad n > 1.$$

Since L is symmetric,

$$\int_{-1}^{1} f(x)p_n(x)dx = \left(\frac{1}{n(n+1)}\right)^m \int_{-1}^{1} (L^m f)(x)p_n(x)dx, \qquad n > 1.$$

It follows that if $\{f_n\}_{n>0}$ are the Legendre coefficients of f, then

$$f_n = \left(\frac{1}{n(n+1)}\right)^m \frac{2n+1}{2} \int_{-1}^{1} (L^m f)(x)p^n(x)dx \equiv \left(\frac{1}{n(n+1)}\right)^m g_n,$$

where $\{g_n\}_{n>0}$ are the Legendre coefficients of $L^m f$. Thus, for any $M > N$

and $|x| < 1$,

$$
\begin{aligned}
|P^M f(x) - P^N f(x)| &= |\sum_{n=N+1}^{M} f_n p_n(x)| \\
&< \sum_{n=N+1}^{M} |f_n| = \sum_{k=N+1}^{M} \left(\frac{1}{n(n+1)}\right)^m |g_n| \\
&< \left(\sum_{n=N+1}^{M} \left(\frac{1}{n(n+1)}\right)^{2m} \frac{2n+1}{2}\right)^{1/2} \left(\sum_{n=N+1}^{M} \frac{2}{2n+1} |g_n|^2\right)^{1/2} \\
&< \left(\int_N^\infty \left(\frac{1}{x(x+1)}\right)^{2m} \frac{2x+1}{2}\, dx\right)^{1/2} \| L^m f \|_{L_2} \\
&= (4m+2)^{-1/2} \left(\frac{1}{N(N+1)}\right)^{m-1/2} \| L^m f \|_{L_2}.
\end{aligned}
$$

Here we have used $|p_n(x)| < 1$ for $|x| < 1$ and

$$\left(\frac{1}{n(n+1)}\right)^{2m} \frac{2n+1}{2} < \int_{n-1}^{n} \left(\frac{1}{x(x+1)}\right)^{2m} \frac{2x+1}{2}\, dx.$$

It now follows that $\{P^N f\}_{N>0}$ is a Cauchy sequence in $C[-1,1]$ and hence that $P^N f$ converges uniformly to f. Letting $M \to \infty$ above, we obtain

$$|f(x) - P^N f(x)| < (4m+2)^{-1/2} \left(\frac{1}{N(N+1)}\right)^{m-1/2} \| L^m f \|_{L_2}.$$

Since L^m is a differential operator of order $2m$ with continuous coefficients on $[-1,1]$, there exists a constant C_m for each $m > 0$ such that

$$\| L^m f \| < C_m \| f \|_{H^{2m}},$$

which completes the proof.

If f is C^∞, then, from Lemma 3.1, the error $\|f-P^Nf\|$ decreases more rapidly than any power of $1/N$. This is usually referred to as "infinite order" approximation.

4. Legendre-Tau Approximation

In this section, we discuss the Legendre-tau approximation of solutions to (2.1). Throughout the rest of the paper, it is convenient to modify the Legendre polynomials by shifting their fundamental interval from $[-1,1]$ to $[-2,0]$. Thus, the p_k's will be understood to be orthogonal on $[-2,0]$ with $p_k(-2) = (-1)^k$ and $p_k(0) = 1$.

For simplicity of exposition, we assume that $r = 2$. See the remark at the end of the section for the general case. If $z(t.\theta) = x(t+\theta)$, $\theta \in [-2,0]$, where x is the solution to (2.1) with initial data $(\eta,\phi) \in \mathcal{D}(A)$, then, according to Lemma 2.1, z satisfies

$$\frac{\partial z}{\partial t}(t,\theta) = \frac{\partial z}{\partial \theta}(t,\theta), \qquad \theta \in [-2,0] \tag{4.1}$$

$$\frac{dz}{dt}(t,0) = \int_{-2}^{0} d\mu(\theta) z(t,\theta) + f(t). \tag{4.2}$$

The approximate solution $z^N(t,\theta)$ is assumed to be expanded in a Legendre series:

$$z^N(t.\theta) = \sum_{k=0}^{N} a_k(t) p_k(\theta), \qquad a_k \in \mathbb{R}^n. \tag{4.3}$$

The tau approximation [9] of (4.1) and (4.2) is as follows. Note that from (3.3),

$$\frac{\partial z^N}{\partial \theta}(t,\theta) = \sum_{k=0}^{N-1} b_k(t) p_k(\theta),$$

where

$$b_k(t) \equiv (2k+1) \sum_{\substack{j=k+1 \\ j+k \text{ odd}}}^{N} a_j(t). \tag{4.4}$$

Equating $\frac{\partial z^N}{\partial t}$ with $\frac{\partial z^N}{\partial \theta}$ leads to the N equations

$$\frac{d}{dt} a_k(t) = b_k(t) \qquad 0 \leq k \leq N-1. \tag{4.5}$$

The essence of the tau method is that the boundary condition (4.2) is then imposed to determine an equation for a_N. From (4.2) we obtain

$$\frac{d}{dt}\Big(\sum_{k=0}^{N} a_k(t)\Big) = \int_{-2}^{0} d\mu(\theta) z^N(t,\theta) + f(t),$$

or

$$\frac{d}{dt} a_N(t) = -\sum_{k=0}^{N-1} b_k(t) + \int_{-2}^{0} d\mu(\theta) z^N(t+\theta) + f(t). \tag{4.6}$$

Hence, from (4.3) - (4.5), we obtain a system of ordinary differential equations for $a_0,\cdots,a_N$:

$$\frac{d}{dt}\alpha^N(t) = A^N \alpha^N(t) + B^N f(t), \tag{4.7}$$

where $\alpha^N = \mathrm{col}(a_0,\cdots,a_N)$ and

$$B^N = e_N \otimes I,$$

where $e_N \in \mathbb{R}^N$ is given by $e_N = \mathrm{col}(0,\cdots,0,1)$, I is the $n\times n$ identity matrix, and $\otimes$ denotes Kronecker product. For the case where N is even; A^N is given by

$$A^N = A_0^N + A_\mu^N,$$

where

$$A_0^N = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 3 & 0 & 3 & \cdots & 0 & 3 \\ 0 & 0 & 0 & 5 & 0 & \cdots & 5 & 0 \\ \vdots & \vdots & & & & & \vdots & \vdots \\ & & & & & & 2N-3 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 2N-1 \\ 0 & -1 & -3 & -6 & -10 & \cdots & & -\frac{N(N+1)}{2} \end{bmatrix} \otimes I \qquad (4.8)$$

and

$$A_\mu^N = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ D_0^N & D_1^N & \cdots & D_N^N \end{bmatrix} \qquad (4.9)$$

with

$$D_k^N = \int_{-2}^{0} d\mu(\theta) p_k, \qquad 0 < k < N. \qquad (4.10)$$

Let us introduce the projection operator L^N on Z. For any $z = (\eta, \phi) \in Z$, L^N is defined by

$$L^N z = \left(\eta, \sum_{k=0}^{N} a_k p_k\right), \qquad (4.11)$$

where, for $0 \leq k \leq N-1$,

$$a_k = \frac{2k+1}{2} \int_{-2}^{0} \phi(\theta) p_k(\theta) d\theta,$$

and

$$a_N = \eta - \sum_{k=0}^{N-1} a_k p_k(0) = \eta - \sum_{k=0}^{N-1} a_k.$$

The tau method can then be interpreted as follows. Let

$$z^N(t) = \left(z^N(t,0),\ z^N(t,\cdot)\right) \in Z$$

where

$$z^N(t.\theta) = \sum_{k=0}^{N} a_k(t) p_k(\theta)$$

as in (4.3). Then the approximate solution $z^N(t)$ satisfies

$$\frac{d}{dt} z^N(t) = L^N \mathcal{A}\, z^N(t) + L^N \mathcal{B}\, f(t), \tag{4.12}$$

in Z. Indeed,

$$L^N \mathcal{A}\, z^N(t) + L^N \mathcal{B}\, f(t) = \left(Dz^N(t,\cdot) + f(t),\ \sum_{k=0}^{N} b_k(t) p_k\right) \in Z,$$

with

$$b_N(t) = Dz^N(t,\cdot) + f(t) - \sum_{k=0}^{N-1} b_k(t)$$

$$= Dz^N(t,\cdot) + f(t) - \dot{z}^N(t,0).$$

Thus, the first component of (4.12) is equivalent to (4.6) and the second component is equivalent to (4.5). Alternatively, the approximate solution $z^N(t)$ is the exact solution to the modified equation

$$\frac{d}{dt} z^N(t) = \mathcal{A} z^N(t) + \mathcal{B} f(t) + \tau_N(t). \tag{4.13}$$

in Z where $\tau_N(t) = \big(0, b_N(t)p_N\big) \in Z$.

<u>Remark</u>: For the general case, i.e., $r \neq 2$, the matrix A^N appearing in (4.7) is replaced by

$$A^N = \frac{2}{r} A_0^N + \tilde{A}_\mu^N ,$$

where

$$\tilde{D}_k^N = \int_{-r}^{0} d\mu(\theta) p_k\Big(\frac{2\theta}{r}\Big) , \qquad 0 \leq k \leq N,$$

replaces D_k^N in (4.9). In particular, if the delay operator D is given by (2.3) and (2.4), then $\tilde{D}_k^N$ has the form

$$\tilde{D}_k^N = \sum_{i=0}^{m} A_i p_k\Big(\frac{2\theta_i}{r}\Big) + \int_{-r}^{0} A(\theta) p_k\Big(\frac{2\theta}{r}\Big) d\theta,$$

for $0 \leq k \leq N$.

5. <u>Convergence Proof</u>

In this section, we show that the approximate solution z^N converges to the true solution z of (2.1) in the case where the forcing term of $f \equiv 0$ and the initial data lies in $\mathcal{D}(\mathcal{A})$. The cases where $f \neq 0$ and the initial data lies merely in Z will be dealt with in the following section. The principal result is based upon the Trotter-Kato theorem (see Theorem 4.6 in

[14]).

Theorem 5.1: Let $S(t)$ and $S^N(t)$, $N \geq 1$, be C_0-semigroups acting on a Banach space X with infinitesimal generators A and A^N respectively. Assume that the following conditions are satisfied:

(i) (stability) There exists a constant ω such that $A - \omega I$ and $A^N - \omega I$ are dissipative on X.

(ii) (consistency) There exists a subset $\mathcal{D}$ contained in $\mathcal{D}(A) \cap \bigcap_{N=1}^{\infty} \mathcal{D}(A^N)$ which together with $(\lambda I - A)\mathcal{D}$ for some $\lambda > 0$ is dense in X and such that $A^N\phi \to A\phi$ for all $\phi \in \mathcal{D}$.

Then for all $\phi \in X$, $\|S^N(t)\phi - S(t)\phi\| \to 0$ uniformly on bounded t-intervals.

We begin by introducing the Hilbert space $X \equiv H^1[-r,0]$ with norm

$$\|\phi\|_1^2 \equiv |\phi(0)|^2 + \int_{-r}^{0} |\dot{\phi}(\theta)|^2 d\theta$$

and inner-product

$$\langle\phi,\psi\rangle_1 \equiv \langle\phi(0),\psi(0)\rangle_{\mathbb{R}^n} + \int_{-r}^{0} \langle\dot{\phi}(\theta),\dot{\psi}(\theta)\rangle_{\mathbb{R}^n} d\theta .$$

It is readily established that X is equivalent to $\mathcal{D}(A) \subset Z$ (see Lemma 2.1 for definition of $\mathcal{D}(A)$) with graph norm $\left(\|z\|^2_{\mathcal{D}(A)} \equiv \|z\|^2 + \|Az\|^2\right)$: there exist constants $0 < c \leq C < \infty$ such that

$$c\|\phi\|_1 \leq \|(\phi(0),\phi\|_{\mathcal{D}(A)} \leq C\|\phi\|_1 \tag{5.1}$$

Let us denote this isomorphism between X and $\mathcal{D}(A)$ by $E : \mathcal{D}(A) \to X$: $E(\phi(0),\phi) = \phi$. Since A generates a C_0-semigroup on $\mathcal{D}(A) \subset Z$ with graph norm, EAE^{-1} generates a C_0-semigroup, which we denote by $\{S(t) : t \geq 0\}$, on X. Moreover, it is easily seen that

$$EAE^{-1} = \frac{\partial}{\partial\theta}$$

and

$$\mathcal{D}(EAE^{-1}) = \{\phi \in X : \dot{\phi} \in X \text{ and } \dot{\phi}(0) = D\phi\}$$

For the sake of convenience, however, we will use the same symbols to denote corresponding elements in and operators on $\mathcal{D}(A)$ and X. except in one instance where greater clarity is required. Thus, on X. we have

$$A = \frac{\partial}{\partial\theta} \quad \text{and} \quad \mathcal{D}(A) = \{\phi \in X : \dot{\phi} \in X \text{ and } \dot{\phi}(0) = D\phi\}$$

Without loss of generality, we can assume that $r = 2$. For any positive integer N, we define the finite dimensional subspace $X^N \subset X$ by

$$X^N = \{\phi \in X : \phi(\theta) = \sum_{k=0}^{N} a_k p_k(\theta) \ , \ a_k \in \mathbb{R}^n\}$$

and define the orthogonal projection Q^N of X onto X^N by

$$(Q^N\phi)(0) = \phi(0) \tag{5.2}$$

$$\frac{d}{d\theta}Q^N\phi = P^{N-1}\dot{\phi} \tag{5.3}$$

In fact, if $Q^N\phi = \sum_{k=0}^{N} b_k p_k$, then

$$\frac{d}{d\theta}Q^N\phi = \sum_{k=0}^{N-1} b_k\dot{p}_k,$$

so that, from (3.4), $\{b_k\}_{k=1}^N$ are uniquely determined by (5.2). b_0 is then determined by (5.1):

$$\left(Q^N\phi\right)(0) = \sum_{k=0}^{N} b_k = \phi(0) \Rightarrow b_0 = \phi(0) - \sum_{k=1}^{N} b_k$$

Let us define $\mathcal{A}^N\colon X \to X^N$ by

$$\mathcal{A}^N\phi = EL^N\mathcal{A}E^{-1}Q^N\phi \quad \text{for} \quad \phi \in X \tag{5.4}$$

where $\mathcal{A}$ is understood as in Lemma 2.1 and L^N is defined by (4.11). Returning to our above-mentioned convention, (4.12) can then be written as

$$\frac{d}{dt}z^N(t) = \mathcal{A}^Nz^N(t) + \mathcal{B}^Nf(t) \quad \text{in} \quad X \tag{5.5}$$

where, for $N \geq 1$, $\mathcal{B}^N \equiv L^N\mathcal{B}$.

Although it is our goal in this section to merely prove convergence for the case where $f \equiv 0$, it is simiple enough to include, in the latter part of the following Lemma, a statement concerning the more general inhomogeneous case which will be used in the next section.

Lemma 5.2 <u>(stability) There exists a constant</u> $\omega \in \mathbb{R}$ <u>such that, for all</u> $N \geq 1$, $\mathcal{A}^N - \omega I$ <u>is dissipative on</u> X, <u>i.e.</u>

$$\langle\mathcal{A}^N\phi,\phi\rangle_1 \leq \omega\|\phi\|_1^2 \quad \text{for} \quad \phi \in X$$

<u>Moreover, if</u> $z^N(t)$ <u>is the solution to (5.5) with</u> $z^N(0) = Q^N\phi$ for $\phi \in X$, then

$$\|z^N(t)\|_1^2 \le e^{2\omega t}\|z^N(0)\|_1^2 + \int_0^t e^{2\omega(t-s)}|f(s)|^2 ds \tag{5.6}$$

<u>Proof</u>: For any $\phi^N \in X^N$ and $f \in \mathbb{R}^n$, it follows from (5.4) and (4.13) that

$$\langle A^N\phi^N + {}^Nf,\phi^N\rangle_1 = \langle D\phi^N+f,\phi^N(0)\rangle + \int_{-2}^{0} \langle\ddot{\phi}^N(\theta)+b_N\dot{p}_N(\theta),\dot{\phi}^N(\theta)\rangle d\theta \tag{5.7}$$

where

$$b_N = D\phi^N + f - \dot{\phi}^N(0) \tag{5.8}$$

Let I denote the intergral term in (5.7). Then

$$I = \int_{-2}^{0}\frac{1}{2}\frac{d}{d\theta}|\dot{\phi}^N(\theta)|^2 d\theta + \int_{-2}^{0}\langle b_N,\dot{\phi}^N(\theta)\rangle\dot{p}_N(\theta)d\theta$$

$$= \frac{1}{2}|\dot{\phi}^N(0)|^2 - \frac{1}{2}|\dot{\phi}^N(-2)|^2 + \langle b_N,\dot{\phi}^N(0)-(-)^N\dot{\phi}^N(-2)\rangle$$

$$+ \int_{-2}^{0}\langle\ddot{\phi}^N(\theta),b_N\rangle p_N(\theta)d\theta \tag{5.9}$$

where we have used $p_N(0) = 1$ and $p_N(-2) = (-1)^N$. Note that the last term on the right-hand side vanishes because $\ddot{\phi}^N$ is a vector in $\mathbb{R}^n$ whose elements are polynomials of degree less than N and hence orthogonal to p_N. Thus, from (5.9),

$$I = \frac{1}{2}|\dot{\phi}^N(0)|^2 - \frac{1}{2}|\dot{\phi}^N(-2)|^2 + \langle D\phi^N+f-\dot{\phi}^N(0),\dot{\phi}^N(0)-(-)^N\dot{\phi}^N(-2)\rangle$$

$$= -\frac{1}{2}|\dot{\phi}^N(0)|^2 + (-)^N\langle\dot{\phi}^N(0),\dot{\phi}^N(-2)\rangle - \frac{1}{2}|\dot{\phi}^N(-2)|^2$$

$$+ \langle D\phi^N + f, \dot{\phi}^N(0) - (-)^N \dot{\phi}^N(-2)\rangle \leq \frac{1}{2}|D\phi^N + f|^2$$

where we have used the inequality $2\langle x,y\rangle_{\mathbb{R}^n} \leq |x|^2 + |y|^2$. From this and (5.7), we find

$$\langle A^N\phi^N + B^N f, \phi^N\rangle_1 \leq \langle D\phi^N + f, \phi^N(0)\rangle + \frac{1}{2}|D\phi^N + f|^2$$

$$= \langle D\phi^N, \phi^N(0)\rangle + \langle f, D\phi^N + \phi^N(0)\rangle + \frac{1}{2}|f|^2 + \frac{1}{2}|D\phi^N|^2$$

$$= \langle D\phi^N, \phi^N(0)\rangle + \frac{1}{2}\{|f|^2 + 2\langle f, D\phi^N + \phi^N(0)\rangle$$

$$+ |D\phi^N + \phi^N(0)|^2\} - \frac{1}{2}|D\phi^N + \phi^N(0)|^2 + \frac{1}{2}|D\phi^N|^2$$

$$= \frac{1}{2}|f + D\phi^N + \phi^N(0)|^2 - |\phi^N(0)|^2 \leq |D\phi^N + \phi^N(0)|^2 + |f|^2$$

Note that $D \in L(X, \mathbb{R}^n)$ with

$$|D\phi| \leq \beta\|\phi\|_1 \tag{5.10}$$

for some positive constant $\beta < \infty$. Thus

$$\langle A^N\phi^N + B^N f, \phi^N\rangle_1 \leq \omega\|\phi^N\|_1^2 + |f|^2 \tag{5.11}$$

with $\omega = (1+\beta)^2$

Since $Q^N L^N z = L^N z$ for all $z \in Z$ and Q^N is symmetric with respect to $\langle\cdot,\cdot\rangle_1$, we have, for all $\phi \in X$,

$$<A^N\phi,\phi>_1 = <Q^NA^NQ^N\phi,\phi>_1 = <A^NQ^N\phi,Q^N\phi>_1$$

It then follows from (5.10) that

$$<A^N\phi,\phi>_1 \leq \omega\|Q^N\phi\|_1^2 \leq \omega\|\phi\|_1^2$$

which implies that $A^N - \omega I$ is dissipative on X.

From (5.5) and (5.11),

$$\frac{1}{2}\frac{d}{dt}\|z^N(t)\|_1^2 = <A^Nz^N(t)+B^Nf(t),z^N(t)>_1 \leq \omega\|z^N(t)\|_1^2 + |f(t)|^2$$

A standard agrument using Gronwall's inequality then yields (5.6).

(Q.E.D.)

Lemma 5.3 (consistency) Let

$$\mathcal{D}^k = \{\phi \in X: \phi \in H^k \text{ and } \dot{\phi}(0)=D\phi\}$$

Then for $k \geq 2, \mathcal{D}^k$ and $(\lambda I-A)\mathcal{D}^k$ for $\lambda \in \mathbb{R}$ sufficiently large are dense in X and, for $k \geq 5, A^N\phi \to A\phi$ in X for $\phi \in \mathcal{D}^k$.

Proof: Since $\mathcal{D}^k$ densely contained in $\mathcal{D}(A)$ for $k \geq 2$, the denseness of $\mathcal{D}^k$ follows from that of $\mathcal{D}(A)$. Since $\mathcal{D}^2 = \mathcal{D}(A)$ and, for $\lambda\in\mathbb{R}$ sufficiently large

$$(\lambda I-A)\mathcal{D}(A) = X,$$

$(\lambda I-A)\mathcal{D}^k$ is dense in X for $k > 2$..

For $\phi \in \mathcal{D}(A)$ and $Q^N\phi \equiv \phi^N$,

$$(A^N-A)\phi = \dot{\phi}^N - \dot{\phi} + \left(D\phi^N-\dot{\phi}^N(0)\right)p_N \tag{5.12}$$

in X for $\phi \in \mathcal{D}(A)$. From (5.10), the definition of Q^N and Lemmas 3.1 - 3.3,

$$|\dot{\phi}^N(0)-\dot{\phi}(0)| < KN^{-2m+1}\|\dot{\phi}\|_{H^{2m}}$$

$$\|\dot{\phi}^N-\dot{\phi}\| < KN^{-2m+3/2}\|\dot{\phi}\|_{H^{2m}}$$

and

$$|D(\phi^N-\phi)| < (1+\beta)KN^{-2m}\|\dot{\phi}\|_{H^{2m}}$$

Since $\dot{\phi}(0) = D\phi$ for $\phi \in \mathcal{D}(A)$,

$$|D\phi^N-\dot{\phi}^N(0)| < (1+\beta)KN^{-2m+1}\|\dot{\phi}\|_{H^{2m}}$$

Using the formula (3.4), it is easy to verify that $\|\dot{p}_N\| = \sqrt{N(N+1)}$. It now follows from (5.12) that

$$\|(A^N-A)\phi\|_1 < \|\dot{\phi}^N-\dot{\phi}\|_1 + |D\phi^N-\dot{\phi}^N(0)|\left(1+\sqrt{N(N+1)}\right) < (2+\beta)KN^{-2m+2}\|\dot{\phi}\|_{H^{2m}}$$

Hence, if $\phi \in \mathcal{D}^k$, $k > 5$,

$$\|(A^N-A)\phi\|_1 < (2+\beta)KN^{-2}\|\phi\|_{H^5} \to 0 \qquad \text{(Q.E.D.)}$$

<u>Remark</u>: Using the same "completing the square" argument as the one employed in the proof of Lemma 5.2, it is easily verified that $\dot{A} - \omega I$ is disipative on X as well.

Combining Theorem 5.1 with Lemmas 5.2 and 5.3, we have

Theorem 5.4: *If* $\{S^N(t):t\geq 0\}$ *denotes the semigroup on* X *generated by* A^N, $N \geq 1$, *then, for all* $\phi \in X$,

$$\|S^N(t)\phi - S(t)\phi\|_1 \to 0$$

uniformly on bounded t-intervals.

6. Convergence Proof (continued)

In this section, we prove the convergence of our scheme for the cases where $f \neq 0$ and where the initial data (η,ϕ) is merely assumed to lie in Z. Throughout this section, the operator $\mathcal{A}$ is defined as in Lemma 2.1.

Let us first consider the case where $f \in L_2^{loc}([0,\infty);\ \mathbb{R}^n)$ and $(\eta,\phi) \equiv 0$. The solution to (2.1) is given by

$$z(t) = \int_0^t S(t-s)\mathcal{B}f(s)ds \tag{6.1}$$

Recall that for any $f \in \mathbb{R}^n$, $\mathcal{B}f = (f,0) \in Z$. It is easily seen that

$$\mathcal{A}^{-1}\mathcal{B}f = \left(\Delta^{-1}f, p_0\Delta^{-1}f\right) \in Z \tag{6.2}$$

provided that $\Delta \equiv \int_{-r}^{0} d\mu(\theta)$ is invertible.

It is easily seen that the spectrum of $\mathcal{A}$ conicides with the zero set of the entire function

$$c(\lambda) = \det\left[\lambda I - \int_{-r}^{0} e^{\lambda\theta} d\mu(\theta)\right]$$

If $c(0) = 0$, i.e., if Δ is not invertible, we can choose any $\lambda \in \rho(\mathcal{A})$, i.e., $c(\lambda) \neq 0$, and consider $y(t) = e^{-\lambda t}x(t)$, $x(t)$ being the solution to the initial-value problem (2.6). Then $y(t)$ will be the solution to

$$\frac{d}{dt}y(t) = -\lambda y(t) + \int_{-r}^{0} e^{\lambda\theta} d\mu(\theta) y(t+\theta) + e^{-\lambda t}f(t) \equiv D_\lambda y_t + e^{-\lambda t}f(t)$$

with initial data $y(\theta) = e^{-\lambda\theta}\phi(\theta)$, $-r \leq \theta \leq 0$.

Clearly, this problem may be formulated as before on the product space Z; the corresponding generator $\mathcal{A}_\lambda$ will by construction satisfy $0 \in \rho(\mathcal{A}_\lambda)$:

$$c_\lambda(\omega) = \det\left[(\omega+\lambda)I - \int_{-r}^{0} e^{(\omega+\lambda)\theta} d\mu(\theta)\right]$$

$$c_\lambda(0) \equiv \det(\Delta_\lambda) \equiv c(\lambda) \neq 0$$

Thus, without loss of generality, we may consider (6.2) above.

Let us rewrite (6.1) as

$$z(t) = \int_0^t \mathcal{A}S(t-s)\mathcal{C}f(s)ds$$

where $\mathcal{C}f \equiv \mathcal{A}^{-1}\mathcal{B}f$. If f is continuously differentiable, it follows from [10, p.488] that

$$z(t) = S(t)\mathcal{C}f(0) - \mathcal{C}f(t) + \int_0^r S(t-s)\mathcal{C}\dot{f}(s)ds \tag{6.3}$$

From the definition of $\mathcal{A}^N$ and $\mathcal{B}^N$, it is easy to show that

$$(\mathcal{A}^N)^{-1}\mathcal{B}^N f = \mathcal{C}f \quad \text{for} \quad f \in \mathbb{R}^n.$$

(See Lemma 6.2 below). Applying the same argument to

$$z^N(t) = \int_0^t \mathcal{A}^N S^N(t-s)(\mathcal{A}^N)^{-1}\mathcal{B}f(s)ds$$

yields

$$z^N(t) = S^N(t)\mathcal{C}f(0) - \mathcal{C}f(t) + \int_0^t S^N(t-s)\mathcal{C}\dot{f}(s)ds \tag{6.4}$$

Since $\mathcal{C}f \in \mathcal{D}(A)$ for $f \in \mathbb{R}^n$, it then follows from Theoren 5.4 that

$$S^N(t)\mathcal{C}f \to S(t)\mathcal{C}f \text{ in } X \text{ for } f \in \mathbb{R}^n$$

and since $\dim(\mathbb{R}^n) < \infty$, this implies

$$\|\left(S^N(t)-S(t)\right)\mathcal{C}\|_{L(\mathbb{R}^n,X)} \to 0$$

Hence, from (6.3) and (6.4), if f is continuously differentiable,

$$\|z^N(t)-z(t)\|_1 \to 0 \tag{6.5}$$

uniformly on bounded t-intervals.

According to the stability result of Lemma 5.2 (c.f. (5.6)),

$$\|\int_0^t S^N(t-s)\mathcal{B}f(s)ds\|_1 \leq e^{\omega t}\|f\|_{L_2([0,t];\mathbb{R}^n)} \tag{6.6}$$

According to the remark made prior to the statement of Theorem 5.4, the above estimate holds true with S^N replaced by S:

$$\|\int_0^t S(t-s)\mathcal{B}f(s)ds\|_1 \leq e^{\omega t}\|f\|_{L^2([0,t];\mathbb{R}^n)}. \tag{6.7}$$

Since the space of continuously differentiable functions on $[0,t]$ is dense in $L^2(0,t)$ for all $t > 0$, a simple limit argument using (6.6) and (6.7) shows that (6.5) holds for any $f \in L_2^{loc}$ as well.

Thus, combining the above with Theorem 5.4, we have

Theorem 6.1: For any initial data $(\phi(0),\phi) \in Z$ with $\phi \in H^1$ and $f \in L_2^{loc}$, the approximate solution $z^N(t)$ to (4.12) converges strongly to $z(t)$ in X, uniformly on bounded t-intervals.

Note that Theorem 6.1 is not applicable when the initial data (η,ϕ) is merely assumed to lie in Z. Thus, to complete our study of convergence, we need to consider the case where $(\eta,\phi) \in Z$ and $f \equiv 0$. Let us define the orthogonal projection Π^N on Z by

$$\Pi^N(\eta,\phi) = \left(\eta,P^{N-1}\phi\right) \quad \text{for} \quad (\eta,\phi) \in Z. \tag{6.8}$$

Lemma 6.2: For $z = (\eta,\phi) \in Z$,

$$\left(A^N\right)^{-1}EL^N z = EA^{-1}\Pi^N z$$

Proof: Recall that the isomorphism $E:\mathcal{D}(A) \to X$ and that $A^N = EL^N A E^{-1} Q^N: X \to X^N \equiv Q^N X$; thus, there will in general be many solutions $w \in \mathcal{D}(A)$ to

$$A^N Ew = EL^N z \tag{6.9}$$

We shall interpret $\left(A^N\right)^{-1}EL^N z$ to be the unique solution of (6.9) lying in X^N. We must then show that a unique solution exists and has the indicated form.

For $z \in Z$, let

$$\Pi^N z = \left(\eta, \sum_{k=0}^{N-1} b_k p_k\right) \equiv (\eta,\phi)$$

then

$$L^N z = \left(\eta, \sum_{k=0}^{N} b_k p_k\right) \quad \text{with} \quad b_N = \eta - \sum_{k=0}^{N-1} b_k = \left(\eta,\phi+(\eta-\dot{\phi}(0))p_N\right)$$

and it follows that $L^N\Pi^N z = L^N z$. A short calculation shows that $A^{-1}(\eta,\phi) = (\psi(0),\psi)$ where

$$\psi(\theta) = \Delta^{-1}\left(\eta + D\int_{\bullet}^{0}\phi(s)ds\right) - \int_{\theta}^{0}\phi(s)ds.$$

and thus, since ϕ is a polynomial of degree $N-1$, ψ has degree N. This shows that

$$EA^{-1}\Pi^N z \in X^N$$

Since $Q^N : X \to X^N$ is a projection,

$$\begin{aligned} A^N EA^{-1}\Pi^N z &= EL^N AE^{-1} Q^N EA^{-1}\Pi^N z \\ &= EL^N \Pi^N z \\ &= EL^N z, \end{aligned}$$

i.e., $Ew = EA^{-1}\Pi^N z \in X^N$ is a solution of (6.9).

To establish uniqueness, we must show that

$$A^N Ev = 0 \text{ and } Ev \in X^N \Rightarrow v = 0$$

Let $Ev = \sum_{k=0}^{N} a_k p_k \equiv \phi$ so that

$$v = E^{-1}\phi = (\phi(0),\phi) \in \mathcal{D}(A)$$

Thus, $Av = (D\phi,\dot{\phi})$, and so

$$A^N Ev = EL^N AE^{-1} Q^N Ev = EL^N Av = E\big(D\phi, \dot{\phi} + (D\phi - \dot{\phi}(0))p_N\big) = \dot{\phi} + (D\phi - \dot{\phi}(0))p_N$$

Since $\dot{\phi}$ is a polynomial of degree N-1, $A^N Ev = 0$ and the orthogonality of the Legendre polynomials yields the fact that $\phi \equiv \phi_0 =$ constant and $D\phi_0 = 0$. But

$$D\phi_0 = D\phi_0 p_0 = \Delta\phi_0$$

whence (c.f. (6.2)) $\phi_0 = 0$, or $v = 0$.

(Q.E.D.)

Theorem 6.3: <u>For any</u> $z = (\eta, \phi) \in Z$,

$$\| \Pi^N\big(E^{-1} S^N(t) EL^N z - S(t)z\big)\|_Z \to 0$$

<u>uniformly on bounded t-intervals.</u>

<u>Proof</u>: Let

$$d = d(t) = \Pi^N\big(E^{-1} S^N(t) EL^N z - S(t)z\big)$$

thus

$$d = \Pi^N E^{-1} A^N S^N(t) (A^N)^{-1} EL^N z - \Pi^N S(t) z$$

It follows from Lemma 6.2 that

$$d = \Pi^N E^{-1} A^N S^N(t) E^{-1} \Pi^N z - \Pi^N S(t) z = \Pi^N \big(E^{-1} A^N - AE^{-1}\big) S^N(t) E^{-1} \Pi^N z$$

$$+ \Pi^N AE^{-1}\big(S^N(t) - S(t)\big) E A^{-1} z + \Pi^N AE^{-1} S^N(t) E A^{-1}\big(\Pi^N z - z\big) \equiv d_1 + d_2 + d_3.$$

Since $S^N(t)E^{-1}\Pi^N z \in X^N$,

$$d_1 = \Pi^N(E^{-1}EL^N AE^{-1}Q^N - AE^{-1})S^N(t)EA^{-1}\Pi^N z$$

$$= \Pi^N(L^N - I)AE^{-1}S^N(t)EA^{-1}\Pi^N z = 0$$

because $\Pi^N L^N y = \Pi^N y$ for any $y \in Z$. Since Π^N is orthogonal on Z,

$$\|d_2\|_Z < \|AE^{-1}(S^N(t)-S(t))EA^{-1}z\|_Z < \|E^{-1}(S^N(t)-S(t))EA^{-1}z\|_{\mathcal{D}(A)}$$

$$< C\|(S^N(t)-S(t))EA^{-1}z\|_1$$

by (5.1). Thus, by Theorem 5.4, $\|d_2(t)\|_Z \to 0$ uniformly on bounded t-intervals. Likewise,

$$\|d_3\|_Z^2 < C^2\|S^N(t)EA^{-1}(\Pi^N z-z)\|_1^2 < C^2 \cdot \|S^N(t)\|_{L(X)}^2 \cdot \|EA^{-1}(\Pi^N z-z)\|_1^2$$

$$< C^2 e^{2\omega t} \cdot c^{-2}\|A^{-1}(\Pi^N z-z)\|_{\mathcal{D}(A)}^2$$

by Lemma 5.2 and the first inequality in (5.1). Thus

$$\|d_3\|_Z^2 < \left(\frac{C}{c}\right)^2 e^{2\omega t}\{\|A^{-1}(\Pi^N z-z)\|_Z^2 + \|\Pi^N z-z\|_Z^2\}$$

$$< \left(\frac{C}{c}\right)^2 e^{2\omega t} \cdot (\|A^{-1}\|_{L(Z)}^2 + 1)\|\Pi^N z-z\|_Z^2$$

and thus $\|d_3(t)\|_Z \to 0$ uniformly on bounced t-intervals as well. This completes the proof of Theorem 6.3.

7. Numerical Results and Conclusions

In this section we discuss some numerical examples which demonstrate the feasibility of the Legendre-tau approximation. For the purpose of comparison, we have also computed approximate solutions by using the cubic spline approximation (S_3) which is discussed in [2]. All computations were performed on a Control Data Corporation Cyber 170 model 730 at NASA Langley Research Center (LaRC) using software written in Fortran. The integration of the system of ordinary differential equations (ODEs) (4.6) was carried out by an IMSL routine (DVERK) employing the Runge-Kutta-Verner fifth and sixth order method.

For the Legendre-tau approximation, implementation of the algorithms is almost as easy for the averaging approximation (AV) which is discussed in [3] and the first-order spline approximation (S_1). In the following, we give the number of operations for each approximation to compute the right-hand side of (4.6) for the scalar system:

$$\frac{d}{dt} x(t) = ax(t) + bx(t-r) + f(t). \tag{7.1}$$

Table I

Method	Number of Additions	Number of Multiplications
SP	$2N + 4$	$N + 2$
AV	$N + 2$	$N + 2$
S_1	$3N + 4$	$3N + 5$
S_3	$11N + C_1$	$9N + C_2$

SP stands for the Legendre-tau approximation. For S_3, C_1 and C_2 are some positive constants (independent of N). Note that, for S_1 and S_3, the operations which are required in order to perform the Cholesky decomposition of Q_k^N (see, for the definition [2], p. 511) are not included. However these numbers may not reflect directly the CPU time required to solve the approximating system of ODEs. According to our calculations, for the same value of N, the Legendre-tau approximation is about four times as fast as S_3. In addition, no storage space is necessary for the matrix A^N appearing in (4.6) for system (7.1) because of its simple structure. For the general system (2.6) in $\mathbb{R}^n$, the last n rows of A^n need to be stored in order to integrate the approximating system of ODEs using the DVERK routine.

As will be evident from the numerical results for the examples presented below, both approximation methods (SP and S_3) behave about the same initially. The typical feature of the Legendre-tau method is that the relative error decreases with increasing time. This is not unexpected due to two facts: (i) the regularity of the solution to a problem which has smooth inhomgoeneous terms increases in time (as pointed out in the Introduction) and (ii) Lemma 3.1. For all examples, on the interval where the solution is infinitely differentiable, the rate of convergence seems to be infinite order. In contrast, spline approximation methods and AV yields finite-order rates of convergence, which is observed in our calculations for S_3.

In order to approximate the initial data, or, in the case when the system involves distributed delays, the expansion coefficients of certain functions need to be computed. For AV and spline approximations such computations are relatively easy, since each element has local support. In our calculations for S_3 we used a Gauss quadrature rule [7]. In contrast, the Legendre polynomials are supported on the whole interval [-2,0], and, for k large,

$p_k(x)$ is a rapidly oscillating function. A feasible algorithm for computing Legendre coefficients will be discussed in a forthcoming paper. However for AV and spline approximations, the expansion coefficients must be recomputed if N is changed. For the Legendre-tau approximation we can use the values which have already been computed for smaller N.

In the tables below we will use the following notation. δ^N_{sp} is one of the differences

$$|\sum_{k=0}^{N} (a_k^N)_j(t) - x_j(t)|, \qquad j=1,2,\cdots,n,$$

where $a_k^N(t)$ is the k^{th} segment of dimension n in the solution vector α^N of the approximating system of ODEs (4.6), and $x(t)$ is the true solution. Similarly, $\delta^N_{S_3}$ denotes one of the differences

$$|(\beta^N(0)w_j^N)(t) - x_j(t)|, \qquad j=1,2,\cdots,n,$$

where $w^N(t)$ is the coordinate vector of the approximate solution obtained by the cubic spline approximation method (S_3) (see, for the definition of $\beta^N(0)$, [2], p. 507). Note that $\dim w^N = n(N+3)$ and $\dim \alpha^N = n(N+1)$.

Example 1 (Banks-Kappel [2], Example 1)

In this example we study the equation for a damped oscillator with delayed restoring force and constant external force,

$$\frac{d^2}{dt^2} x(t) + \frac{d}{dt} x(t) + x(t-1) = 10,$$

with initial conditions

$$x(\theta) = \cos\theta, \qquad \frac{d}{dt}x(\theta) = -\sin\theta \qquad \text{for} \quad \theta \in [-1,0].$$

Rewriting the above equation as a first order system we have

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix} + \begin{bmatrix} 0 \\ 10 \end{bmatrix},$$

where $x_1(t) = x(t)$ and $x_2(t) = \frac{dx}{dt}(t)$.

Table II and Table III show the numerical results for $x(t)$ and $\frac{d}{dt}x(t)$. For this example SP achieves the same accuracy with smaller N that S_3 does. For SP, it appears that the rate of convergence is infinite-order.

Table II

t	x(t)	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	δ^4_{SP}	δ^8_{SP}
.25	1.2704759	0.00007	0.00022	0.000024	0.00011	0.0000047
.5	1.9936737	0.00171	0.00021	0.000022	0.00028	0.0000012
.75	3.0614837	0.00247	0.00019	0.000017	0.00011	0.0000176
1.0	4.3927203	0.00083	0.00024	0.000012	0.00027	0.0000106
1.25	5.9259310	0.00082	0.00032	0.000016	0.00020	0.0000016
1.5	7.6000709	0.00090	0.00005	0.000010	0.00007	0.0000009
1.75	9.3440157	0.00054	0.00025	0.000005	0.00006	0.0000002
2.0	11.0833011	0.00013	0.00003	0.000015	0.00003	0.

Table III

t	$\dot{x}(t)$	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	δ^4_{SP}	δ^8_{SP}
.25	2.06969	0.00946	0.00205	0.00048	0.00001	0.00001
.5	3.64428	0.00792	0.00021	0.00186	0.00171	0.00025
.75	4.84445	0.00073	0.00200	0.00045	0.00335	0.00014
1.0	5.76581	0.00790	0.00192	0.00046	0.00348	0.00031
1.25	6.45956	0.00625	0.00119	0.00039	0.00029	0.00008
1.5	6.45956	0.00185	0.00245	0.00062	0.00078	0.00002
1.75	7.0159957	0.00061	0.00003	0.00067	0.00075	0.
2.0	6.8497211	0.00180	0.00200	0.00059	0.00019	0.

Example 2 ([2], Example 4)

Next we use an example due to Popov for a degenerate system where we have $(1,-2,-1)^T x(t) = 0$ for $t \geq 2$ and all initial data $(\eta . \phi) \in Z$. The equation is

$$\frac{d}{dt} x(t) = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} x(t-1).$$

We choose the initial data:

$$\eta = \mathrm{col}(1,1,1) \quad \text{and} \quad \phi(\theta) \equiv 0 \in \mathbb{R}^3 \quad \text{for} \quad -1 \leq \theta \leq 0.$$

Note that $(\eta,\phi) \notin \mathcal{D}(A)$.

In Tables IV, V, and VI we give the numerical results. Here the initial function is not in the subspace Z^N for either method. With respect to $x_2(t)$ and $x_3(t)$ we can see that S_3 gives slightly better approximations than SP for $t < 1$. This is because x_2 and x_3 have a jump discontinuity in the derivative at $t = 1$. For SP, the convergence of the approximate solutions to the asymptotic solution (2,0,2) is quite rapid and seems to be infinite-order.

Table IV

t	$x_1(t)$	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	δ^4_{SP}	δ^8_{SP}	δ^{16}_{SP}
0.2	1.36	0.01211	0.00493	0.00247	0.00140	0.00041	0.00001
0.4	1.64	0.00889	0.00528	0.00001	0.00146	0.00158	0.00001
0.6	1.84	0.00445	0.00334	0.00190	0.00547	0.00098	0.00016
0.8	1.96	0.00980	0.00074	0.00146	0.00233	0.00111	0.00019
1.0	2.0	0.00020	0.00221	0.00054	0.00339	0.00033	0.00009
1.2	2.0	0.00725	0.00424	0.00175	0.00237	0.00002	0.00001
1.4	2.0	0.00160	0.00460	0.00085	0.00064	0.00003	3×10^{-7}
1.6	2.0	0.00527	0.00406	0.00130	0.00036	0.00003	13×10^{-7}
1.8	2.0	0.00331	0.00281	0.00242	0.00028	0.00001	5×10^{-7}
2.0	2.0	0.00256	0.00117	0.00130	0.00004	5×10^{-6}	1×10^{-7}
2.2	2.0	0.00306	0.00049	0.00096	0.00004	1×10^{-6}	9×10^{-9}
2.4	2.0	0.00531	0.00178	0.00214	0.00003	2×10^{-8}	2×10^{-9}
2.6	2.0	0.00212	0.00246	0.00115	5×10^{-6}	9×10^{-8}	1×10^{-9}
2.8	2.0	0.00055	0.00251	0.00092	3×10^{-6}	5×10^{-8}	3×10^{-9}
3.0	2.0	0.00117	0.00200	0.00192	2×10^{-6}	2×10^{-8}	0.

Table V

t	$x_2(t)$	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	δ^4_{SP}	δ^8_{SP}	δ^{16}_{SP}
0.2	0.8	0.00947	0.00536	0.00230	0.02139	0.00517	0.00295
0.4	0.6	0.01206	0.00451	0.00034	0.02201	0.00650	0.00302
0.6	0.4	0.00526	0.00260	0.00202	0.00021	0.00126	0.00119
0.8	0.2	0.00153	0.00015	0.00139	0.02790	0.00240	0.00039
1.0	0.	0.02133	0.01895	0.00949	0.04081	0.02072	0.01034
1.2	0.	0.01505	0.00168	0.00010	0.00135	0.00265	0.00053
1.4	0.	0.001369	0.00068	0.00094	0.00440	0.00085	0.00013
1.6	0.	0.00105	0.00078	0.00209	0.00071	0.00027	0.00003
1.8	0.	0.01149	0.00105	0.00193	0.00067	0.00006	3×10^{-6}
2.0	0.	0.00444	0.00123	0.00020	0.00042	4×10^{-6}	4×10^{-7}
2.2	0.	0.00636	0.00108	0.00173	0.00004	5×10^{-6}	3×10^{-7}
2.4	0.	0.00521	0.00066	0.00202	0.00007	3×10^{-6}	8×10^{-8}
2.6	0.	0.00205	0.00019	0.00073	0.00003	1×10^{-6}	1×10^{-8}
2.8	0.	0.00407	0.00021	0.00084	0.00005	3×10^{-7}	3×10^{-10}
3.0	0.	0.00039	0.00051	0.00119	0.00005	3×10^{-7}	6×10^{-10}

Table VI

t	$x_3(t)$	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	δ^4_{SP}	δ^8_{SP}	δ^{16}_{SP}
0.2	1.0	0.01131	0.00517	0.00235	0.04191	0.01044	0.00592
0.4	1.0	0.01278	0.00471	0.00022	0.04592	0.01267	0.00600
0.6	1.0	0.00628	0.00373	0.00202	0.00799	0.00181	0.00224
0.8	1.0	0.00472	0.00100	0.00164	0.04989	0.00105	0.00105
1.0	1.0	0.04871	0.03399	0.01807	0.08354	0.04269	0.02070
1.2	1.36	0.02598	0.00193	0.00149	0.00955	0.00603	0.00091
1.4	1.64	0.02266	0.00337	0.00228	0.00595	0.00250	0.00014
1.6	1.84	0.00015	0.00167	0.00248	0.00443	0.00143	0.00007
1.8	1.96	0.01932	0.00086	0.00153	0.00314	0.00117	0.00016
2.0	2.0	0.01007	0.00266	0.00049	0.00295	0.00057	0.00008
2.2	2.0	0.00901	0.00349	0.00223	0.00328	0.00002	0.00001
2.4	2.0	0.00127	0.00375	0.00198	0.00098	0.00007	1×10^{-6}
2.6	2.0	0.00071	0.00321	0.00069	0.00052	0.00005	3×10^{-6}
2.8	2.0	0.00714	0.00211	0.00049	0.00050	0.00002	9×10^{-7}
3.0	2.0	0.00465	0.00076	0.00060	0.00011	0.00001	1×10^{-7}

Example 3 ([2], Example 5)

We consider the scalar equation

$$\frac{d}{dt}\, x(t) = 5x(t) + x(t-1),$$

with initial function

$$x(\theta) = 5 \qquad \text{for} \quad \theta \in [-1,0].$$

The numerical results for this example can be found in Table VII and VIII. Again, we observe the quickness of the convergence of the approximate SP solution in this example.

Table VII

t	x(t)	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	$\delta^{32}_{S_3}$	$\delta^{64}_{S_3}$
0.2	15.309691	0.03	0.00486	0.00028	0.000621	0.000066
0.4	43.334337	0.05	0.00403	0.00229	0.000243	0.000077
0.6	119.513222	0.26	0.01267	0.00080	0.000056	0.000107
0.8	326.588900	0.78	0.01900	0.00136	0.000357	0.000079
1.0	889.478955	2.39	0.09568	0.00206	0.000012	0.000123
1.2	2420.772761	7.28	0.25894	0.00031	0.000845	0.000149
1.4	6588.865818	22.02	0.77234	0.00316	0.001704	0.000036
1.6	17934.153211	65.58	0.71924	0.01228	0.004029	0.000051
1.8	48815.256906	194.12	6.70537	0.03713	0.013520	0.001419
2.0	132871.377933	570.95	19.58949	0.10230	0.039730	0.004005

Table VIII

t	x(t)	δ^4_{SP}	δ^8_{SP}	δ^{16}_{SP}	δ^{32}_{SP}	δ^{64}_{SP}
0.2	15.309691	0.04	0.00823	0.00021	0.000291	0.000046
0.4	43.334337	6×10^{-5}	0.00316	0.00076	0.000368	0.000058
0.6	119.513222	0.22	0.01232	0.00196	0.000193	0.000042
0.8	326.588900	0.60	0.01626	0.00285	0.000452	0.000066
1.0	889.478955	1.67	0.01392	0.00220	0.000334	0.000049
1.2	2420.772761	5.22	0.00216	0.00007	0.000032	0.000003
1.4	6588.865818	15.56	0.00108	0.00004	2×10^{-7}	4×10^{-7}
1.6	17934.153211	46.07	0.00075	0.00003	0.000002	4×10^{-8}
1.8	48815.256906	135.58	0.00002	0.00002	0.000003	2×10^{-7}
2.0	132871.377933	396.72	0.00069	0.00009	0.000014	0.000002

Example 4

Here we deal with the equation which has multiple point delays

$$\frac{d}{dt} x(t) = x(t) + 2x(t - 1/2) + x(t-1)$$

with initial function

$$x(\theta) = 1 \qquad \text{for} \quad \theta \in [-1,0].$$

The numerical results in Table IX show that both methods work equally well.

Table IX

t	x(t)	$\delta^4_{S_3}$	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	δ^4_{SP}	δ^8_{SP}	δ^{16}_{SP}
0.2	1.885611	0.002409	0.000311	0.000086	0.006696	0.000193	0.000183
0.4	2.967299	0.001213	0.000022	0.000403	0.024093	0.000916	0.000748
0.6	4.331245	0.004842	0.001300	0.000565	0.021832	0.000727	0.000472
0.8	6.342954	0.010185	0.004028	0.000170	0.006117	0.001556	0.000436
1.0	9.278242	0.005778	0.003518	0.000356	0.011526	0.001310	0.000254
1.2	13.563777	0.011080	0.001466	0.000741	0.008807	0.000168	0.000027
1.4	19.903791	0.010571	0.000014	0.000361	0.007846	0.000388	0.000024
1.6	29.212354	0.002034	0.000492	0.000141	0.004891	0.000145	0.000025
1.8	42.845032	0.015807	0.001791	0.000343	0.015039	0.000077	0.000002
2.0	62.841170	0.001850	0.001460	0.000224	0.021896	0.000072	0.000013

Table X

t	x(t)	$\delta^8_{S_3}$	$\delta^{16}_{S_3}$	$\delta^{32}_{S_3}$	δ^8_{SP}	δ^{16}_{SP}	δ^{32}_{SP}
0.2	0.304159	0.000447	0.000099	0.000051	0.001490	0.000040	0.000105
0.4	-0.463775	0.000060	0.000124	0.000024	0.001717	0.000151	0.000089
0.6	-1.173353	0.000424	0.000006	0.000014	0.003747	0.000519	0.000058
0.8	-1.796286	0.000113	0.000018	0.000043	0.004891	0.000624	0.000178
1.0	-2.307740	0.001049	0.000400	0.000101	0.003346	0.000082	0.000115
1.2	-2.497698	0.002418	0.000151	0.000138	0.001002	0.000014	0.000048
1.4	-2.121921	0.001361	0.000205	0.000073	0.000652	0.000013	0.000003
1.6	-1.236591	0.000108	0.000266	0.000022	0.000533	0.000027	0.000002
1.8	0.051303	0.001261	0.000063	0.000019	0.000482	0.000037	0.000003
2.0	1.606226	0.001956	0.000227	0.000014	0.000693	0.000092	0.000002

Example 5

This example is used in [1] for the identification problem which has a distributed delay and a discontinuous forcing function. This equation is

$$\frac{d}{dt}\,x(t) = -3x(t) - \int_{-1}^{0} x(t+\theta)d\theta + u_{.1}(t),$$

with initial function

$$x(\theta) \equiv 1 \qquad \text{for} \quad \theta \in [-1,0],$$

where $u_{.1}(t)$ is defined by

$$u_{.1}(t) = \begin{cases} 1 & \text{on } [0,.1] \\ 0 & \text{otherwise} \end{cases} \quad .$$

In Table X we give the numerical results. The approximations by S_3 seem to be slightly better initially than by SP. This is because the solution $x_t(\theta)$ has a discontinuous derivative with respect to θ for $t < 1.1$. The superiority of SP on the interval where the solution is infinitely differentiable is again observed for $t > 1.2$.

From the discussion and numerical results presented here we can conclude the following. For the Legendre-tau approximation, the effort in implementing the algorithm is as much as for AV. But the accuracy of approximation is as good as that which can be obtained using high order spline approximation. When the solution to a problem is infinitely differentiable, the rate of convergence is faster than any finite power of $(1/N)$. In addition, from our calculations, our method appears to be about four times as fast as the cubic spline approximation method. These properties are much more evident for examples (not presented in this paper) where optimal control and the approximation of eigenvalues were considered.

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. Differential Equations, 34 (1979), 496-522.

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

[4] C. Bernier and A. Manitius, On semigroup in $\mathbb{R}^n \times L^p$ corresponding to differential equations with delays, Canada J. Math. XXX (1980), 969-978.

[5] J. G. Borisovic and A. S. Turbabin, On the Cauchy problem for linear nonhomogeneous differential equations with retarded argument, Soviet Math. Dokl. 10 (1969), 401-405.

[6] C. Canuto and A. Quarteroni, Approximation results for orthogonal polynomials in Sobolev spaces, Math. Comp. 38 (1982), 67-86.

[7] P. J. Davis and P. Rabinowitz, Methods of Numerical Integration, Academic Press, 1975.

[8] M. C. Delfour and S. K. Mitter, Hereditary differential systems with constant delays: I. General case, J. Diff. Equations 12 (1972), 213-235.

[9] D. Gottlieb and S. A. Orszag, Numerical Analysis of Spectral Methods: Theory and Applications, Regional Conference Series in Applied Math, SIAM, 1977.

[10] T. Kato, Perturbation Theory for Linear Operators, Springer, 132, 1966.

[11] D. L. Kreider, R. G. Kuller, D. R. Ostberg, F. W. Perkins, An Introduction to Linear Analysis, Addison-Wesley, 1966.

[12] C. Lanczos, Applied Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1956.

[13] N. N. Lebedev, Special Functions and their Applications, Dover, New York, 1972.

[14] A. Pazy,Semigroup of Linear Operators and Applications to Partial Differential Equations, Mathematics Department Lecture Notes, Vol. 10, University of Maryland, College Park, MD, 1974.

[15] R. B. Vinter, On the evolution of the state of linear differential delay equations in M^2: Properties of the generator, J. Inst. Math. Appl. 21 (1978), 13-23.

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR-172155		

4. Title and Subtitle	5. Report Date
Legendre-Tau Approximations for Functional Differential Equations	June 1983
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Kazufumi Ito and Russell Teglas	83-17
	10. Work Unit No.
9. Performing Organization Name and Address	
Institute for Computer Applications in Science and Engineering NASA Langley Research Center Hampton, VA 23665	11. Contract or Grant No. NAS1-17070
	13. Type of Report and Period Covered Contractor Report
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546	14. Sponsoring Agency Code

15. Supplementary Notes

Langely Technical Monitor: Robert H. Tolson
Final Report

16. Abstract

In this paper we consider the numerical approximation of solutions to linear functional differential equations using the so-called Legendre-tau method. The functional differential equation is first reformulated as a partial differential equation with a non-local boundary condition involving time-differentiation. The approximate solution is then represented as a truncated Legendre series with time varying coefficients which satisfy a certain system of ordinary differential equations. The method is very easy to code and yields very accurate approximations. Convergence is established, various numerical examples are presented, and comparison between the latter and cubic spline approximations is made.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
tau method retarded systems Legendre polynomials numerical convergence	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	45	A03

For sale by the National Technical Information Service, Springfield, Virginia 22161

End of Document