JOURNAL OF DIFFERENTIAL EQUATIONS 18, 296–332 (1975)

Projection Series for Retarded Functional Differential Equations with Applications to Optimal Control Problems

H. T. BANKS*

Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912

AND

ANDRZEJ MANITIUS†

Institute of Automatics, Polytechnical University of Warsaw, Warsaw, Poland and Department of Computer, Information and Control Sciences, University of Minnesota, Minneapolis, Minnesota 55455

Received June 14, 1973

In this paper projection methods based on expansions of solutions of retarded function differential equations in terms of generalized eigenfunctions are considered. It is first shown that the projection series developed earlier by Hale and Shimanov and those considered by Bellman and Cooke are actually the same. Using extensions of the residue-type arguments of Bellman and Cooke, convergence results are then established for a class of perturbed systems. These results are applied to obtain approximations to optimal controls for certain infinite dimensional variational problems. Numerical results are presented for several examples.

1. INTRODUCTION

In this paper we consider the method of finite-dimensional projections for retarded functional differential equations (FDE) developed previously by Hale [11] and Shimanov [28]. These authors established that the state of linear autonomous FDEs can be decomposed into the sum of a projection

* This research was supported in part by the U.S. Air Force under contract AF-AFOSR-71-2078B and in part by the U.S. Army under contract DA-ARO-D-31-124-71-G95 F2.

† This research was performed in part while this author was a visiting staff member in the Center for Dynamical Systems, Brown University, and in part while he was a visiting staff member in the Center for Control Sciences, University of Minnesota, Minneapolis, with support in part from Honeywell, Inc. Present address: Centre de Recherches Mathématiques, Université de Montréal, Montréal, P.Q., Canada.

onto a finite-dimensional generalized eigenmanifold in the state space plus a residual term. A question of fundamental importance not dealt with in these papers is that of whether the finite-dimensional projection term converges to the infinite-dimensional state of the FDE as the eigenmanifold is extended to include the infinite set of all generalized eigenfunctions of the FDE. We treat this problem below and show that an affirmative answer can be given in certain cases. We then use these ideas to investigate a fairly wide class of optimal control problems with fixed initial and terminal values given in the infinite-dimensional state space. Such a problem falls into a class of problems with target set in function space discussed in [2] and [16]. However, the approach taken in this paper is not in the spirit of that of [2] and [16](in which a maximum principle and sufficiency results are developed), since here we project the original problem onto a finite-dimensional subspace where the "finite-dimensional" problem is then "solved." We then consider the sequence of solutions thus obtained and discuss convergence properties of this sequence relative to the solution of the original problem.

The projection method has been previously applied by other authors [20–22] to a class of control problems with no terminal constraints. Unfortunately, a crucial step in the proof for convergence of solutions of the finite-dimensional problems appears to be based on invalid arguments [19; see MR *33*, #2991].

In Section 2 we give a summary of the projection method of Hale and the decomposition results. The relationship of the series obtained from these projections to the well-known series expansions of Bellman and Cooke is rigorously established in Section 3, and the general question of convergence is considered in Section 4. Finally, Section 5 is devoted to a discussion of the applications of these results to the optimal control problems mentioned above.

For $[a, b]$ a subinterval of the real line and R^n euclidean (real or complex) n-space, we shall use $L_p([a, b], R^n)$ to denote the usual space of "functions" on $[a, b]$ with values in R^n whose pth power is Lebesgue integrable. The Sobolev space of absolutely continuous functions on $[a, b]$ into R^n with first derivatives in L_p will be denoted by $W_p^{(1)}([a, b], R^n)$. The symbol $|x|$ will denote the norm of x, where the norm is to be understood to be that of the space in which x lies.

2. Projection Results for Retarded FDE

This section will be devoted to a summary of projection ideas developed in detail by Hale [11, 12]. In addition, we shall discuss problems involving the convergence of series obtained using these projections. The notation adopted will be almost identical to that used by Hale.

Consider the system described by the linear stationary FDE

$$\dot{x}(t) = L(x_t) + f(t) \tag{2.1}$$

where L is a linear mapping $L: \mathscr{C} \to R^n$ given by

$$L(x_t) = \int_{-r}^{0} d\eta(\theta)\, x(t+\theta) \tag{2.2}$$

with η a matrix-valued function of bounded variation. Here, as in [11, 12], x_t will denote the complete state of the system in (2.1), i.e., $x_t(\theta) = x(t+\theta)$, $\theta \in [-r, 0]$ and $\mathscr{C} = C([-r, 0], R^n)$ is the Banach space of continuous functions on $[-r, 0]$ having values in R^n with the usual supremum norm. As is shown in [11], the infinitesimal generator $\mathscr{A}$ related to Eq. (2.1) has the form

$$\mathscr{A}\phi(\theta) = \begin{cases} d\phi(\theta)/d\theta, & -r \leqslant \theta < 0, \\ \int_{-r}^{0} d\eta(s)\, \phi(s), & \theta = 0. \end{cases} \tag{2.3}$$

The domain $\mathscr{D}(\mathscr{A})$ of the operator $\mathscr{A}$ is the subset in $C([-r, 0], R^n)$ consisting of functions that are continuous and have a continuous derivative on $[-r, 0]$ and satisfy $\dot{\phi}(0) = L(\phi)$. The spectrum of $\mathscr{A}$ consists only of point spectrum, i.e., points in the spectrum of $\mathscr{A}$ are eigenvalues of $\mathscr{A}$ (see [11] or [12, Lemma 20.1]). Any nonzero $\phi \in \mathscr{D}(\mathscr{A})$ satisfying the equation

$$(\lambda I - \mathscr{A})\phi = 0, \tag{2.4}$$

where I is the identity operator in $\mathscr{C}$, is called an eigenfunction of $\mathscr{A}$.

The knowledge of eigenvalues and eigenfunctions will be necessary later on. To find them let us rewrite (2.4) in the form

$$d\phi(\theta)/d\theta = \lambda\phi(\theta), \qquad -r \leqslant \theta < 0, \tag{2.5}$$

$$\int_{-r}^{0} d\eta(s)\, \phi(s) = \lambda\phi(0), \qquad \theta = 0. \tag{2.6}$$

From (2.5) we have $\phi(\theta) = \exp(\lambda\theta)a$, where $a \in R^n$. Substituting this into (2.6), we obtain

$$\left\{\lambda I - \int_{-r}^{0} d\eta(\theta) \exp(\lambda\theta)\right\} a = 0, \tag{2.7}$$

where I is here the identity operator in R^n. Define

$$\Delta(\lambda) = \lambda I - \int_{-r}^{0} d\eta(\theta) \exp(\lambda\theta). \tag{2.8}$$

Then the eigenvalues of $\mathscr{A}$ satisfy

$$\det \Delta(\lambda) = 0, \tag{2.9}$$

while the vector $a^j \in R^n$ corresponding to an eigenvalue λ_j can be found from

$$\Delta(\lambda_j)\, a^j = 0. \tag{2.10}$$

Then the eigenfunction ϕ_j corresponding to the eigenvalue λ_j is found to be

$$\phi_j(\theta) = \exp(\lambda_j \theta)\, a^j, \qquad \theta \in [-r, 0]. \tag{2.11}$$

The following general properties of eigenvalues λ_j of $\mathscr{A}$ are known [12]:

(i) they are of finite multiplicity;

(ii) there is a $\gamma > 0$ such that no eigenvalue has real part greater than γ;

(iii) when η is piecewise constant on $[-r, 0]$, as is the case for differential-difference equations, the λ_j are asymptotically distributed in curvilinear strips of type $|\operatorname{Re}(s + \mu_i \log s)| \leqslant c$, $i = 1,..., k$ [5].

Let $\mathscr{M}_{\lambda_j}$ denote (as in [11]) the smallest subspace of $C([-r, 0], R^n)$ containing the null spaces $\mathscr{N}(\mathscr{A} - \lambda_j I)^l$, $l = 1, 2,...$. It is shown in [11, 12] that $\mathscr{M}_{\lambda_j} = \mathscr{N}(\mathscr{A} - \lambda_j I)^k$, $k =$ ascent of $\mathscr{A} - \lambda_j I$, is finite-dimensional. Thus for every eigenvalue λ_j of $\mathscr{A}$ there exists a finite set of generalized eigenfunctions $\phi_j^1,..., \phi_j^{d_j}$ constituting a basis of $\mathscr{M}_{\lambda_j}$. Let Φ_{λ_j} denote a matrix-valued function $[-r, 0] \to R^{n \times d_j}$:

$$\Phi_{\lambda_j}(\theta) = [\phi_j^1(\theta),..., \phi_j^{d_j}(\theta)], \qquad \theta \in [-r, 0]. \tag{2.12}$$

Since $\mathscr{A}\mathscr{M}_{\lambda_j} \subset \mathscr{M}_{\lambda_j}$, there exists a $d_j \times d_j$ constant matrix B_{λ_j} whose only eigenvalue is λ_j, such that

$$\mathscr{A}\Phi_{\lambda_j} = \Phi_{\lambda_j} B_{\lambda_j}. \tag{2.13}$$

For $\phi \in C([-r, 0], R^n)$, $\psi \in C([0, r], R^{n*})$, where R^{n*} is the n-dimensional euclidean space of row vectors, define the bilinear functional $(\psi, \phi) \to \langle \psi, \phi \rangle$:

$$\langle \psi, \phi \rangle = \psi(0)\, \phi(0) - \int_{-r}^{0} \int_{0}^{\theta} \psi(\zeta - \theta)\, d\eta(\theta)\, \phi(\zeta)\, d\zeta. \tag{2.14}$$

Also, for $\psi \in \mathscr{D}(\mathscr{A}^*) \subset C([0, r], R^{n*})$(see [12]), let

$$\mathscr{A}^*\psi(\theta) = \begin{cases} -d\psi(\theta)/d\theta, & 0 < \theta \leqslant r, \\ \displaystyle\int_{-r}^{0} \psi(-s)\, d\eta(s), & \theta = 0. \end{cases}$$

It is shown in [11, 12] that the operator $\mathscr{A}^*$ has properties similar to the adjoint of $\mathscr{A}$. In particular, the following relation holds:

$$\langle \psi, \mathscr{A}\phi \rangle = \langle \mathscr{A}^*\psi, \phi \rangle, \qquad \text{for all} \quad \phi \in \mathscr{D}(\mathscr{A}), \quad \psi \in \mathscr{D}(\mathscr{A}^*). \tag{2.15}$$

Let Ψ_{λ_j} denote the matrix analogous to (2.12), with ψ_j^i defined on $[0, r]$ the generalized eigenfunctions of $\mathscr{A}^*$. The matrix functions Φ_{λ_j} and Ψ_{λ_j} may be chosen so that

$$\langle \Psi_{\lambda_j}, \Phi_{\lambda_j} \rangle = I \tag{2.16}$$

where I is the $d_j \times d_j$ identity matrix. Take the first N eigenvalues λ_j, $j = 1,..., N$, where the eigenvalues λ_j are ordered with respect to real parts, beginning with the eigenvalues with greatest real part. Let

$$\Phi^N = [\Phi_{\lambda_1},..., \Phi_{\lambda_N}] \tag{2.17}$$

be the basis chosen as above for the generalized eigenmanifold $\mathscr{M}^N$ associated with $\{\lambda_1,..., \lambda_N\}$, with corresponding basis

$$\Psi^N = \begin{bmatrix} \Psi_{\lambda_1} \\ \vdots \\ \Psi_{\lambda_N} \end{bmatrix} \tag{2.18}$$

for the generalized eigenmanifold for the adjoint. Let P^N denote the projection onto the subspace $\mathscr{M}^N$ given by

$$P^N\phi = \Phi^N\langle \Psi^N, \phi \rangle.$$

Thus for any $\phi \in \mathscr{C}$ we may write

$$\phi = \phi^{P^N} + \phi^{Q^N},$$

where $\phi^{P^N} \equiv P^N\phi$, $\phi^{Q^N} \equiv \phi - P^N\phi$. For any solution x_t of (2.1) we may write

$$x_t = x_t^{P^N} + x_t^{Q^N}, \tag{2.19}$$

where

$$x_t^{P^N} = \Phi^N\langle \Psi^N, x_t \rangle = \sum_{j=1}^{N} \Phi_{\lambda_j}\langle \Psi_{\lambda_j}, x_t \rangle. \tag{2.20}$$

Define the $d_j \times 1$ vector function y_{λ_j} by $y_{\lambda_j}(t) \equiv \langle \Psi_{\lambda_j}, x_t \rangle$, and the $\sum_{j=1}^N d_j \times 1$ vector function y^N by $y^N(t) \equiv \langle \Psi^N, x_t \rangle$. From (2.18) we have

$$y^N(t) = \begin{bmatrix} y_{\lambda_1}(t) \\ \vdots \\ y_{\lambda_N}(t) \end{bmatrix}. \tag{2.21}$$

Formula (2.20) can now be written as

$$x_t^{PN} = \Phi^N y^N(t) = \sum_{j=1}^{N} \Phi_{\lambda_j} y_{\lambda_j}(t). \tag{2.22}$$

Let $B^N = \text{quasidiag}\,\{B_{\lambda_1}, ..., B_{\lambda_N}\}$ so that $\mathscr{A}\Phi^N = \Phi^N B^N$. From results derived in [12, Section 24], we may write

$$y^N(t) = \exp(B^N t)\, y^N(0) + \int_0^t \exp(B^N(t-s))\, \Psi^N(0) f(s)\, ds \tag{2.23}$$

and

$$x_t^{QN} = T(t)\, x_0^{QN} + \int_0^t T(t-s)\, X_0^{QN} f(s)\, ds \tag{2.24}$$

where $T(t)$ is the semigroup of operators for the homogeneous form of (2.1), $X_Q^{QN} = X_0 - \Phi^N \Psi^N(0)$ and X_0 is defined by

$$X_0(\theta) = \begin{cases} 0, & -r \leqslant \theta < 0, \\ I, & \theta = 0, \end{cases}$$

with I the $n \times n$ identity matrix. Hence

$$X_0^{QN}(\theta) = \begin{cases} I - \Phi^N(0)\, \Psi^N(0), & \theta = 0, \\ -\Phi^N(\theta)\, \Psi^N(0), & -r \leqslant \theta < 0. \end{cases} \tag{2.25}$$

We thus see that the original functional differential equation is projected by the use of the projection P^N onto a finite-dimensional invariant subspace $\mathscr{M}^N$ of $\mathscr{C}$. The evolution of the system in this subspace in terms of "coordinates" $y^N(t)$ is described by (2.23) which is equivalent to

$$\begin{aligned} \dot{y}^N(t) &= B^N y^N(t) + \Psi^N(0) f(t), \\ y^N(0) &= \langle \Psi^N, x_0 \rangle. \end{aligned} \tag{2.26}$$

The behavior of the residual term x_t^{QN} is described by the integral Eq. (2.24). We remark that the formal use of a differential equation for x_t^{QN} (as well as for x_t in cases when $f(t) \not\equiv 0$) in some previous papers [22, 28] is apparently without justification, as is the use of a differential equation for $z(t)$ in [11] where $z_t = x_t^{QN}$.

The decomposition of the original equation into the system of Eqs. (2.26) and (2.24) has two essential properties:

(i) the system (2.26) is a finite-dimensional system of ordinary differential equations without time lag;

(ii) the operator $T(t)$ in (2.24) acts on $(I - P^N)\mathscr{C}$, so that [12, Theorem 24.1]:

$$| T(t)\, x_0^{Q^N} | \leqslant K \exp((\gamma - \delta)t) \,| x_0^{Q^N} |, \qquad x_0^{Q^N} \in (I - P^N)\mathscr{C}, \quad (2.27)$$

where γ is a constant such that $\operatorname{Re} \lambda_N > \gamma$, and $\delta > 0$, $K > 0$ are constants depending on γ (i.e., on λ_N). If, in particular, $\gamma \leqslant 0$, then (2.27) guarantees that $x_t^{Q^N}$ for the case $f(t) \equiv 0$ is uniformly bounded by an exponentially decreasing function of time.

A question of great interest is: for which $\phi \in \mathscr{C}$ do we have $\phi^{P^N} \to \phi$ as $N \to \infty$, or, in terms of (2.1), for which class of functions $\{f\}$ do we obtain $x_t^{P^N} \to x_t$, or equivalently $x_t^{Q^N} \to 0$ in $\mathscr{C}$?

From (2.24) we obtain

$$| x_t^{Q^N} | \leqslant | T(t)\, x_0^{Q^N} | + \int_0^t | T(t-s)\, X_0^{Q^N} f(s) |\, ds. \qquad (2.28)$$

An estimate for the first term is given by (2.27). The second term may be estimated by the use of the inequality [12, Section 24]:

$$| T(t)\, X_0^{Q^N} | \leqslant K \exp((\gamma - \delta)t), \qquad (2.29)$$

where K and γ are as in (2.27). It must be realized that K depends on λ_N and hence on N, and in fact estimates for K given in [12] in terms of the iterates of the map $T(t)$ show that $K(\lambda_N)$ increases as $N \to \infty$. Without further information on how fast $K(\lambda_N)$ increases with N one cannot use the estimates given in (2.27) and (2.29) to investigate the convergence of $x_t^{Q^N}$. Further evidence that the convergence can be difficult to ascertain is supplied by the following simple example. Consider the scalar system

$$\dot{x}(t) = x(t) + x(t-1) + f(t). \qquad (2.30)$$

The characteristic equation is

$$\Delta(\lambda) = \lambda - 1 - \exp(-\lambda) = 0. \qquad (2.31)$$

It is clear that the roots of $\Delta(\lambda) = 0$ are simple. Using the notation in (2.11) for eigenfunctions ϕ_{λ_j}, and representing ψ_{λ_j} by

$$\psi_{\lambda_j}(\theta) = \exp(-\lambda_j \theta)\, b^j,$$

where b^j (as well as a^j in (2.11)) is a scalar, we can, after some simple calculations, write the normalization condition (2.16) in the form

$$b^j(1 + \exp(-\lambda_j))\, a^j = 1.$$

If a^j is, for each j, arbitrarily chosen to be equal to 1, then b^j must satisfy $b^j = 1/(1 + \exp(-\lambda_j))$, which, in light of (2.31), yields

$$b^j = 1/\lambda_j . \tag{2.32}$$

Making note of the second term in (2.28), we compute

$$X_0^{Q^N}(\theta) f(s) = \begin{cases} \left\{1 - \sum_{j=1}^{N} \phi_{\lambda_j}(0)\, \psi_{\lambda_j}(0)\right\} f(s), & \theta = 0, \\ \left\{- \sum_{j=1}^{N} \phi_{\lambda_j}(\theta)\, \psi_{\lambda_j}(0)\right\} f(s), & -r \leqslant \theta < 0, \end{cases}$$

$$= \begin{cases} \left\{1 - \sum_{j=1}^{N} a^j b^j\right\} f(s), & \theta = 0, \\ \left\{- \sum_{j=1}^{N} \exp(\lambda_j \theta)\, a^j b^j\right\} f(s), & -r \leqslant \theta < 0 \end{cases}$$

and by (2.32)

$$X_0^{Q^N}(\theta) f(s) = \begin{cases} \left\{1 - \sum_{j=1}^{N} \dfrac{1}{\lambda_j}\right\} f(s), & \theta = 0, \quad (2.33) \\ \left\{- \sum_{j=1}^{N} \dfrac{1}{\lambda_j} \exp(\lambda_j \theta)\right\} f(s), & -r \leqslant \theta < 0. \quad (2.34) \end{cases}$$

It is known that the modulus of λ_j for the retarded system (2.30) grows with j nearly as an algebraic progression, and it is also known that $\sum_{j=1}^{N} 1/j = c + \ln N + \epsilon(N)$, where c is Euler's constant and $\epsilon(N) \to 0$ as $N \to \infty$. Thus the failure of $1 - \sum_{j=1}^{N} (1/\lambda_j)$ to converge casts serious doubt upon the usefulness of (2.28) in establishing $x_t^{Q^N} \to 0$. In view of the difficulties discussed above we turn to other methods of investigating the convergence question.

It has been known for some time that solutions of differential-difference equations can be written in terms of Fourier-type exponential series [5]. Since these series are nothing more than expansions in the generalized eigenfunctions $\phi(\theta) = (\theta^\nu/\nu!)\, e^{\lambda_j \theta}$ discussed above, one possible approach to the question of when $x_t^{Q^N} \to 0$ lies in showing the equivalence between the Fourier exponential series of Bellman and Cooke and that series obtained by use of the projection methods detailed above. We now proceed to do just that.

3. Equivalence of the Projection Series and the Bellman–Cooke Exponential Series

The exponential series expansion for solutions of certain classes of autonomous differential difference equations was obtained in [5] via inversion of the Laplace transform of the solution, and is an infinite sum of terms $p_j(t)\, e^{\lambda_j t}$, where the $p_j(t)$ are polynomials in t, n-vector-valued, and the λ_j are solutions of (2.9).

We shall now show the term-by-term equivalence of the exponential series and the projection series. Questions related to convergence of the exponential series will be treated in the next section. While the convergence results will be established only for differential-difference equations, the equivalence of terms will be proved for the general retarded FDE (2.1). By considering the Laplace transform of the solution $x(t)$ of Eq. (2.1), we shall first derive a general formula for $p_j(t)\, e^{\lambda_j t}$.

Assume that $f \in L_1([0, t_1], R^n)$ and extend f to the semiaxis $[0, \infty)$ by taking $f(t) \equiv 0$ for $t > t_1$. Then the Laplace transform of f exists. We shall assume that $x_0 = \xi$ is a continuous function (this assumption can be weakened to $\xi \in L_1$ if instead of the general FDE we consider a differential-difference equation). Since the solution of the homogeneous FDE is exponentially bounded [11, Corollary II.1], by applying the variation of constants formula [12, Section 16] to system (2.1) in $[0, t_1]$ and by considering the free motion of system (2.1) in $[t_1, \infty)$ we see that the solution of the nonhomogeneous Eq. (2.1) is exponentially bounded, $|x_t| \leqslant Ke^{\delta t}$, hence Laplace transformable. We may then consider the integral

$$I = \left| \int_0^\infty \left(\int_{-r}^0 d\eta(\theta)\, x(t+\theta) \right) e^{-st}\, dt \right| \tag{3.1}$$

$$\leqslant \int_0^\infty \operatorname*{Var}_{[-r,0]} |\eta(\cdot)| \max_{[-r,0]} |x(t+\theta)|\, e^{-\sigma t}\, dt$$

$$\leqslant \operatorname*{Var}_{[-r,0]} |\eta(\cdot)| \int_0^\infty Ke^{\delta t} e^{-\sigma t}\, dt$$

which converges for σ such that $\sigma = \operatorname{Re}(s) \geqslant \delta$.

By similar arguments it is not difficult to show that the integral

$$\int_{-r}^0 |d\eta(\cdot)| \int_{-r}^\infty |e^{-st}|\, |x(t+\theta)|\, dt$$

exists for $\sigma \geqslant \delta$. Therefore the Fubini theorem [9] can be applied to change the order of integrations in (3.1). An application of the Laplace transform to both sides of (2.1) yields the following expression for the Laplace transform of the solution $x(t)$:

$$\mathscr{L}[x](s) = \Delta^{-1}(s)\, q(s), \tag{3.2}$$

where

$$q(s) = \xi(0) - \int_{-r}^{0} \int_{0}^{\theta} e^{s(\theta-\tau)}\, d\eta(\theta)\, \xi(\tau)\, d\tau + \int_{0}^{t_1} e^{-s\tau} f(\tau)\, d\tau. \tag{3.3}$$

We observe that $s \to q(s)$ is analytic, which implies that the only singularities of $s \to e^{st} \Delta^{-1}(s)\, q(s)$ occur at $s = \lambda_j$, where λ_j are the roots of $\det \Delta(s) = 0$. Therefore the terms appearing in the exponential series will have the form

$$p_j(t)\, e^{\lambda_j t} = \operatorname{Res}\{e^{st} \Delta^{-1}(s)\, q(s)\}_{s=\lambda_j}. \tag{3.4}$$

We show that the term in (3.4) is equal to the corresponding term in the projection series. Denote by $x_t(\xi, f)$ the solution of (2.1) with initial data $x_0 = \xi$ and let λ_j be a solution of (2.9)(i.e., in the point spectrum of $\mathscr{A}$). Let $\mathscr{D}(S_{\lambda_j}) \equiv \{\phi \in \mathscr{C} \mid \phi = x_t(\xi, f)$ for some $t > 0$, $\xi \in \mathscr{C}$, $f \in L_1$ with f having compact support in $[0, \infty)\}$ and define, for $x_t(\xi, f) \in \mathscr{D}(S_{\lambda_j})$, the operator $S_{\lambda_j} : \mathscr{D}(S_{\lambda_j}) \subset \mathscr{C} \to \mathscr{C}$ by

$$S_{\lambda_j} x_t(\xi, f)(\theta) = \operatorname{Res}\{e^{\lambda(t+\theta)} \Delta^{-1}(\lambda)\, Q(\lambda, t; \xi, f)\}\, |_{\lambda=\lambda_j},$$

where

$$Q(\lambda, t; \xi, f) \equiv \xi(0) - \int_{-r}^{0} \int_{0}^{s} e^{\lambda(s-\tau)}\, d\eta(s)\, \xi(\tau)\, d\tau + \int_{0}^{t} e^{-\lambda\tau} f(\tau)\, d\tau.$$

The above arguments show that this operator is well-defined. Note that $(\xi, f) \to S_{\lambda_j} x_t(\xi, f)$ is a linear map.

Let $P_{\lambda_j}^{CAN} : \mathscr{C} \to \mathscr{C}$ be the canonical projection (see [30, p. 306]) onto $\mathscr{M}_{\lambda_j}(\mathscr{A})$ along $\mathscr{R}(\mathscr{A} - \lambda_j I)^k$ and let $P_{\lambda_j}^{H-S} : \mathscr{C} \to \mathscr{C}$ be the projection of Hale-Shimanov onto $\mathscr{M}_{\lambda_j}(\mathscr{A})$ discussed above and given by $P_{\lambda_j}^{H-S}\phi = \Phi_{\lambda_j}\langle \Psi_{\lambda_j}, \phi\rangle$. If R_λ denotes the resolvent of $\mathscr{A} - \lambda I$, then we have [30]

$$P_{\lambda_j}^{CAN} = -\frac{1}{2\pi i} \int_{\Gamma_j} R_\lambda\, d\lambda,$$

where Γ_j is a contour enclosing the isolated singularity λ_j. Furthermore, by Lemmas 21.2 and 21.4 in [12] we observe that $P_{\lambda_j}^{H-S} = \Phi_{\lambda_j}\langle \Psi_{\lambda_j}, \cdot\rangle$ is the projection onto $\mathscr{M}_{\lambda_j}(\mathscr{A})$ along

$$\mathscr{Q} = \{\phi \in \mathscr{C} \mid \langle \Psi_{\lambda_j}, \phi\rangle = 0\} = \mathscr{R}(\mathscr{A} - \lambda_j I)^k$$

and thus we have

$$P_{\lambda_j}^{CAN} = P_{\lambda_j}^{H-S}.$$

We shall hereafter denote $P_{\lambda_j}^{CAN} = P_{\lambda_j}^{H-S}$ by simply P_{λ_j}.

THEOREM 3.1. *Assume $\xi \in \mathscr{C}$, $f \in L_1([0, t_1], R^n)$ with $f(t) \equiv 0$ for $t > t_1$. Then for $t \geqslant t_1$ we have*

$$P_{\lambda_j} x_t(\xi, f) = S_{\lambda_j} x_t(\xi, f).$$

Proof. We show first that $P_{\lambda_j} x_t(\xi, 0) = S_{\lambda_j} x_t(\xi, 0)$ for $\xi \in \mathscr{C}$ and $t > 0$. Recall that there exists a real number μ such that for $\lambda \in \rho(\mathscr{A})$, $\mathrm{Re}(\lambda) > \mu$, the resolvent R_λ is given by [9, p. 622]

$$R_\lambda \xi = -\int_0^\infty e^{-\lambda\tau} T(\tau)\, \xi\, d\tau$$

for any $\xi \in \mathscr{C}$, where $\{T(\tau)\}$ is the solution semigroup on $\mathscr{C}$. In particular, for $\mathrm{Re}(\lambda) > \mu$, $t > 0$ and $\xi \in \mathscr{C}$ we have

$$\begin{aligned}
-R_\lambda x_t(\xi, 0) &= -R_\lambda T(t)\xi = \int_0^\infty e^{-\lambda\tau} T(\tau)\, T(t)\, \xi\, d\tau \\
&= \int_0^\infty e^{-\lambda\tau} T(\tau + t)\, \xi\, d\tau = \int_t^\infty e^{-\lambda w} T(w) \{e^{\lambda t}\xi\}\, dw \\
&= \int_0^\infty e^{-\lambda w} T(w) \{e^{\lambda t}\xi\}\, dw - \int_0^t e^{-\lambda w} T(w) \{e^{\lambda t}\xi\}\, dw \\
&= -R_\lambda(e^{\lambda t}\xi) - \int_0^t e^{-\lambda w} e^{\lambda t} T(w)\, \xi\, dw.
\end{aligned}$$

Thus by analytic continuation we obtain

$$R_\lambda x_t(\xi, 0) = R_\lambda(e^{\lambda t}\xi) + \int_0^t e^{-\lambda w} e^{\lambda t} T(w)\, \xi\, dw$$

for all $\lambda \in \rho(\mathscr{A})$, $t > 0$ and $\xi \in \mathscr{C}$. This leads to

$$\begin{aligned}
P_{\lambda_j} x_t(\xi, 0) &= -\frac{1}{2\pi i} \int_{\Gamma_j} R_\lambda x_t(\xi, 0)\, d\lambda = -\mathrm{Res}\, \{R_\lambda x_t(\xi, 0)\}\, |_{\lambda=\lambda_j} \\
&= -\mathrm{Res} \left\{ R_\lambda(e^{\lambda t}\xi) + \int_0^t e^{-\lambda w} e^{\lambda t} T(w)\, \xi\, dw \right\} \Big|_{\lambda=\lambda_j} \\
&= -\mathrm{Res}\{R_\lambda(e^{\lambda t}\xi)\}\, |_{\lambda=\lambda_j}. \qquad (3.5)
\end{aligned}$$

But from [12, pp. 99–100] we see

$$(R_\lambda\psi)(\theta) = e^{\lambda\theta}\Delta^{-1}(\lambda)\left\{-\psi(0) + \int_{-r}^{0}\int_{0}^{s} e^{\lambda(s-\tau)}d\eta(s)\psi(\tau)d\tau\right\} + \int_{0}^{\theta} e^{\lambda(\theta-\tau)}\psi(\tau)d\tau \tag{3.6}$$

so that we find, using (3.5) with (3.6),

$$\begin{aligned} P_{\lambda_j}x_t(\xi, 0)(\theta) &= \operatorname{Res}\left\{e^{\lambda(t+\theta)}\,\Delta^{-1}(\lambda)\left[\xi(0) - \int_{-r}^{0}\int_{0}^{s} e^{\lambda(s-\tau)}\,d\eta(s)\,\xi(\tau)\,d\tau\right]\right\}\Big|_{\lambda=\lambda_j} \\ &= S_{\lambda_j}x_t(\xi, 0)(\theta), \end{aligned}$$

for any $t > 0$ and $\xi \in \mathscr{C}$.

Next observe that for f as in the hypotheses and $\sigma > 0$ we have

$$x_{t_1+\sigma}(0, f) = T(\sigma)\,x_{t_1}(0, f) = T(\sigma)\,\zeta_1 = x_\sigma(\zeta_1\,, 0)$$

where $\zeta_1 \equiv x_{t_1}(0, f)$. Then our above arguments yield

$$P_{\lambda_j}x_{t_1+\sigma}(0, f) = P_{\lambda_j}x_\sigma(\zeta_1\,, 0) = S_{\lambda_j}x_\sigma(\zeta_1\,, 0) = S_{\lambda_j}x_{t_1+\sigma}(0, f),$$

for $\sigma > 0$. Thus, for ξ, f as in the hypotheses we have $P_{\lambda_j}x_t(0, f) = S_{\lambda_j}x_t(0, f)$ for $t > t_1$, and $P_{\lambda_j}x_t(\xi, 0) = S_{\lambda_j}x_t(\xi, 0)$ for $t > 0$. The linearity of $(\xi, f) \to P_{\lambda_j}x_t(\xi, f)$ and $(\xi, f) \to S_{\lambda_j}x_t(\xi, f)$ then allows us to conclude

$$P_{\lambda_j}x_t(\xi, f) = S_{\lambda_j}x_t(\xi, f), \qquad \text{for} \quad t > t_1\,.$$

But it is easy to see that both $t \to P_{\lambda_j}x_t(\xi, f)$ and $t \to S_{\lambda_j}x_t(\xi, f)$ are continuous on $(0, \infty)$, whence the conclusion of the theorem follows.

4. Convergence Results

Having shown an equivalence between the terms in the expansions of Bellman and Cooke and the terms obtained using the projections due to Hale, we are now in a position to use the ideas and methods and, in some cases, already-proven results in [5] to establish convergence results for the sequence x_t^{pN} in a number of instances. We shall restrict our investigations to differential-difference equations, a subclass of the Eq. (2.1) where the measure η has only a finite number of atoms and no continuous part.

Consider equations with ν delays $h_1, \ldots, h_\nu$ (η then has atoms at $\theta = 0$, $\theta = -h_i$, $i = 1, \ldots, \nu$, $h_\nu = r$) and f as described in Section 3:

$$\dot{x}(t) = \sum_{i=0}^{\nu} A_i x(t - h_i) + f(t), \qquad t \in [0, t_1], \tag{4.1}$$

where $0 = h_0 < h_1 < \cdots < h_\nu$,

$$x_0 = \xi.$$

We shall assume in our subsequent discussions that $t_1 - h_\nu > 0$; i.e., the time interval on which we study the dynamical system in (4.1) is greater than that associated with the state of the system.

If we take the Laplace transform in this case, we obtain (3.2) with $q(s) = p(s) + \hat{f}(s)$, where

$$p(s) = \xi(0) + \sum_{i=1}^{\nu} A_i e^{-sh_i} \int_{-h_i}^{0} e^{-s\tau} \xi(\tau)\, d\tau \tag{4.2}$$

$$\hat{f}(s) = \int_0^{t_1} e^{-s\tau} f(\tau)\, d\tau$$

and

$$\Delta(s) = sI - \sum_{i=0}^{\nu} A_i e^{-sh_i}. \tag{4.3}$$

Our first results concern the case where the matrix A_ν is nonsingular.

THEOREM 4.1. *Let $\mathscr{F}$ be a bounded subset of $L_1([0, t_1], R^n)$. Assume det $A_\nu \neq 0$ and let $t \to x(t, \xi, f)$ denote the solution of (4.1) corresponding to $f \in \mathscr{F}$ and $\xi \in L_1([-h_\nu, 0], R^n)$(where we assume $\xi(0)$ is specified). Then*

$$x(t, \xi, f) = \lim_{l \to \infty} \sum_{\lambda_j \in C_l} p_j(t)\, e^{\lambda_j t}, \tag{4.4}$$

holds for $t > t_1 - h_\nu$, where $p_j(t)\, e^{\lambda_j t}$ is the residue (n-vector-valued) of the function $s \to e^{st} \Delta^{-1}(s)\{p(s) + \hat{f}(s)\}$ at the pole $s = \lambda_j$, and C_l are the contours described in [5, *p.* 100]. *Moreover, the convergence in* (4.4) *is uniform in $f \in \mathscr{F}$ and uniform in t on any interval $[a, b]$ with $t_1 - h_\nu < a < b < \infty$.*

Remark. Since our eventual concern here involves convergences of $x_{t_1}^{pN}$ in $\mathscr{C}$, the uniform (in t) convergence guaranteed by the theorem is of paramount interest. We point out that the statement and proof of the above theorem must be modified by replacing $t > t_1 - h_\nu$ type statements by $t > \max\{0, t_1 - h_\nu\}$ if one does not make the assumption that $t_1 - h_\nu > 0$.

Proof. The arguments are mostly modifications of those used in [5] to prove a similar theorem for the scalar case with one delay and with $f(t) \equiv 0$ (see [5 pp. 98–124] and the arguments for Theorems 4.1, 4.2 and 4.6). We shall therefore omit here those details for arguments which are the same as those in [5].

Taking the Laplace transform in (4.1) and then using the inversion formula, we have for $t > 0$

$$x(t) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{st}\,\Delta^{-1}(s)\{p(s)+\hat{f}(s)\}\,ds, \tag{4.5}$$

where γ is such that the roots of $\det \Delta(s) = 0$ all lie to the left of $\mathrm{Re}(s) = \gamma$. Using the fact that ξ, f are L_1, one can easily argue that $|p(s)|$ and $|\hat{f}(s)|$ are $O(1)$ on horizontal line segments $\sigma + i\mu$, $e_1 \leqslant \sigma \leqslant e_2$, as $\mu \to \infty$, where e_1 is such that all characteristic roots lie to the left of $\mathrm{Re}(s) = e_1$. Then, introducing the contours C_l as in [5], one proceeds to argue just as on pp. 103–104 of [5] that (4.5) can be written

$$x(t) = \lim_{l\to\infty}\left\{\sum_{C_l}\mathrm{Res} - \frac{1}{2\pi i}\int_{C_l-} e^{st}\,\Delta^{-1}(s)\{p(s)+\hat{f}(s)\}\,ds\right\}, \tag{4.6}$$

where here $\sum_{C_l}$ Res denotes the sum of the residues of

$$s \to e^{st}\,\Delta^{-1}(s)\{p(s)+\hat{f}(s)\}$$

in the contour C_l.

Thus, to obtain (4.4) from (4.6), it remains only to establish that

$$\lim_{l\to\infty}\int_{C_l-} e^{st}\,\Delta^{-1}(s)\{p(s)+\hat{f}(s)\}\,ds = 0, \tag{4.7}$$

and, of course, to point out for which values of t this convergence is uniform in t. Using the definitions of p and $\hat{f}$, we may write

$$\int_{C_l-} e^{st}\,\Delta^{-1}(s)\{p(s)+\hat{f}(s)\}\,ds = I_1(t)+I_2(t)+I_3(t),$$

where

$$I_1(t) \equiv \int_{C_l-} e^{st}\,\Delta^{-1}(s)\,ds\,\xi(0), \tag{4.8}$$

$$I_2(t) \equiv \sum_{i=1}^{\nu}\int_{-h_i}^{0}\left\{\int_{C_l-} e^{(t-\tau-h_i)s}\,\Delta^{-1}(s)\,ds\right\}A_i\xi(\tau)\,d\tau, \tag{4.9}$$

$$I_3(t) \equiv \int_0^{t_1}\left\{\int_{C_l-} e^{(t-\tau)s}\,\Delta^{-1}(s)\,ds\right\}f(\tau)\,d\tau. \tag{4.10}$$

We next recall Lemma 4.2 [5, p. 122] which states that for a scalar equation with a single delay with $\Delta(s) = s - A_0 - A_1e^{-sh_1}$ and $A_1 \neq 0$, the following relation holds:

$$\lim_{l\to\infty}\int_{C_l-} |e^{ts}\,\Delta^{-1}(s)|\,|ds| = 0, \tag{4.11}$$

for $t > -h_1$, and furthermore the convergence in (4.11) is uniform in t on $[a, b]$ whenever $-h_1 < a < b < \infty$. An analogue of this lemma for the n-vector equation, which we are going to prove, will be needed to conclude the proof of Theorem 4.1.

LEMMA 4.1. *If $\Delta(s)$ is as given by* (4.3) *with det $A_\nu \neq 0$, then*

$$\lim_{l\to\infty} \int_{C_l-} | e^{ts} \Delta^{-1}(s)| \, | ds | = 0, \tag{4.12}$$

for $t > -h_\nu$; furthermore, the convergence in (4.12) *is uniform in t on $[a, b]$ whenever $-h_\nu < a < b < \infty$.*

Proof. We refer the reader to Chapter 12 of [5]. Define

$$G(s) = e^{sh_\nu} \Delta(s) = Ise^{sh_\nu} - A_0 e^{sh_\nu} - A_1 e^{s(h_\nu - h_1)} \cdots - A_\nu \tag{4.13}$$

and

$$g(s) = \det G(s). \tag{4.14}$$

Denote the elements of the matrix A_i by $a_i^{jk}, j, k = 1,..., n$. Each element of the main diagonal of the matrix $G(s)$ always contains a nonzero term se^{sh_ν} plus terms (possibly zero) $a_0^{kk} e^{sh_\nu} + a_1^{kk} e^{s(h_\nu - h_1)} + \cdots + a_\nu^{kk}$. If we represent the determinant $g(s)$ as a polynomial with respect to s and $e^{s\beta_j}$, where β_j are sums of terms like $h_\nu - h_i$, then we may observe the following:

(i) The maximal possible power of e^{sh_ν} is n.

(ii) The maximal power of s is n and such a term certainly is present, as it is the term $s^n e^{nsh_\nu}$ resulting from the product of all elements on the main diagonal.

(iii) Other terms will have a form $s^k e^{\beta_j s}$ with $\beta_j \geqslant kh_\nu$, $k = 0, 1,..., n-1$, the last inequality following from the fact that s always appears in product with e^{sh_ν} and is possibly multiplied by other terms of type $e^{s(h_\nu - h_i)}$ with $h_\nu - h_i \geqslant 0$.

(iv) Equation (4.13) implies that the coefficient corresponding to $s^0 e^{0s}$ is equal to $(-1)^n$ det A_ν (compare equation (12.10.6) in [5]), and, by nonsingularity of A_ν, is nonzero.

Let us write $g(s)$ in a form

$$g(s) = \sum_{j=0}^{p} p_j s^{m_j} (1 + \epsilon(s)) \, e^{\beta_j s}, \tag{4.15}$$

where $0 = \beta_0 \leqslant \beta_1 \leqslant \cdots \leqslant \beta_p$, and $\epsilon(s) \to 0$ as $|s| \to \infty$. The statements above yield the following implications:

$$\text{(i)} \quad \max_i \beta_i \leqslant nh_\nu, \tag{4.16}$$

$$\text{(ii)} \quad m_p = n,\ \beta_p = nh_\nu, \tag{4.17}$$

$$\text{(iii)} \quad \beta_j \geqslant m_j h_\nu, \tag{4.18}$$

$$\text{(iv)} \quad p_0 \neq 0,\ m_0 = 0. \tag{4.19}$$

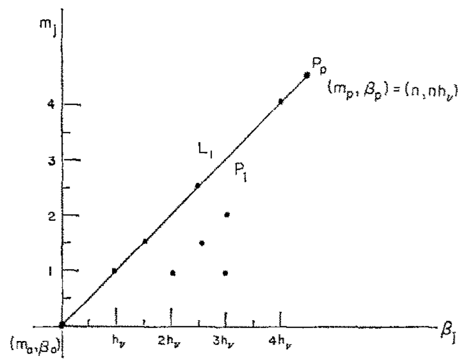

FIG. 1. Distribution diagram for the differential-difference equation with matrix A_ν nonsingular.

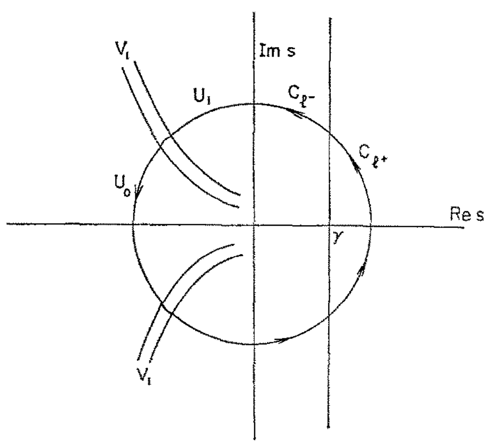

FIG. 2. The strips V_1, region U_0 and U_1 and the contours C_{l+}, C_{l-} in the complex plane.

If we now construct the polygonal distribution diagram defined in [5, Section 12.8], corresponding to the quasipolynomial (4.15), connecting the points P_j with coordinates (β_j, m_j)(see Fig. 1), then by (4.17) it will certainly contain the point $P_p = (n, nh_\nu)$ and, by (4.19), the point $P_0 = (0, 0)$. By upward convexity of the distribution diagram (see [5]), the only line segment L_i appearing in the distribution diagram will be the segment L_1 connecting the points P_0 and P_p, with a slope equal to $n/nh_\nu = 1/h_\nu$. Hence, by Theorem 12.10 of [5] we may conclude that the asymptotic zeros of $g(s)$ are contained in a single curvilinear strip $V_1 : |\operatorname{Re}(s + 1/h_\nu \log s)| \leqslant c_1$ (which may, however, contain more than one root chain). The strip V_1 divides the complex plane into two regions U_0 and U_1 (U_0 lying on the left of V_1) with different lower bounds for $g(s)$[5, Theorem 12.10, Part (a)]. By Theorem 12.9 of [5], within U_i and for $|s|$ large one term of $g(s)$ is of predominant order of magnitude, namely the one corresponding to the point of the distribution diagram at the righthand end of the segment L_i. Hence there are positive constants c, c_2 such that

$$|g(s)| \geqslant c\,|s^n e^{nsh_\nu}|, \qquad \text{for} \quad |s| \geqslant c_2, \quad s \in U_1 \tag{4.20}$$

and

$$|g(s)| \geqslant c\,|s^0 e^{0s}| = c, \qquad \text{for} \quad |s| \geqslant c_2, \quad s \in U_0. \tag{4.21}$$

In addition, by Theorem 12.10(c) of [5], c, c_2 can be chosen so that (4.21) holds in any subregion $\tilde{V}_1$ of V_1 in which s is uniformly bounded away from all zeros of $g(s)$ and $|s| \geqslant c_2$. Moreover, in $U_0 \cup V_1$ we have

$$\operatorname{Re}\left(s + \frac{1}{h_\nu} \log s\right) \leqslant c_1,$$

so that

$$|s e^{sh_\nu}| \leqslant e^{c_1 h_\nu},$$

or

$$|s| \leqslant |e^{(c_1 - s)h_\nu}|. \tag{4.22}$$

Considering the cofactors of the matrix $G(s)$, we see that they are quasipolynomials similar to (4.15), with $m_p = n - 1$, and, by (4.22),

$$|s^{m_j} e^{\beta_j s}| \leqslant |e^{(c_1 - s)h_\nu m_j} e^{\beta_j s}| = e^{c_1 m_j h_\nu}\,|e^{(\beta_j - m_j h_\nu)s}|.$$

Applying (4.18) (which is still valid for cofactor terms) along with Part (d) of Lemma 12.3 in [5], we conclude that all cofactors of $G(s)$ are $O(1)$ for $s \in U_0 \cup V_1$, $|s|$ sufficiently large. Combining this with (4.21), we find

$$|G^{-1}(s)| = O(1), \qquad s \in U_0 \cup \tilde{V}_1, \quad |s| \geqslant c_2,$$

or, since $\Delta^{-1}(s) = s^{sh_\nu} G^{-1}(s)$,

$$|\Delta^{-1}(s)| = O(|e^{sh_\nu}|), \qquad s \in U_0 \cup \tilde{V}_1, \qquad |s| \geqslant c_2. \tag{4.23}$$

Similar arguments using $| s^{m_j} e^{\beta_j s} | \leqslant | s |^{n-1} | e^{(n-1)sh_\nu} |$ for $s \in U_1$, $\mathrm{Re}(s) \leqslant 0$, in the cofactor terms together with (4.20) yield

$$| G^{-1}(s)| = O(| s |^{-1} | e^{-sh_\nu} |), \qquad s \in U_1, \quad \mathrm{Re}(s) \leqslant 0, \quad | s | \geqslant c_2,$$

or

$$| \Delta^{-1}(s)| = O(| s |^{-1}), \qquad s \in U_1, \quad \mathrm{Re}(s) \leqslant 0, \quad | s | \geqslant c_2. \qquad (4.24)$$

The above arguments show that there is a complete analogy between estimates (4.23), (4.24) for the n-vector case and the estimate (4.6.5) in [5] for the scalar case. Hence all further arguments for the convergence of the series of contour integrals given on pp. 122–123 in [5] can be repeated to complete the proof of Lemma 4.1.

Let us observe that the estimates derived above are sharper than those given in Theorem 12.14 of [5].

An application of Lemma 4.1 yields immediately that $I_1(t) \to 0$ as $l \to \infty$, uniformly in t on $[a, b]$ where $a > -h_\nu$. Likewise, the lemma guarantees that $I_2(t) \to 0$ as $l \to \infty$ for $t > 0$, the convergence being uniform on $[a, b]$ where $a > 0$. Also, $I_3(t) \to 0$ as $l \to \infty$ for $t > t_1 - h_\nu$, with uniform convergence on $[a, b]$, $a > t_1 - h_\nu$. Noting that I_3 is the only term which depends on f, we also observe that the convergence in (4.7) (and hence in (4.4)) is uniform with respect to f for $f \in \mathscr{F}$, $\mathscr{F}$ a bounded subset of L_1. This completes the proof of Theorem 4.1.

Investigating the term I_3 in (4.10) once again, we see that if f vanishes for $t > t_1 - \epsilon$, $\epsilon > 0$, and if $t_1 - h_\nu - \epsilon \geqslant 0$, then $I_3(t) \to 0$ as $l \to \infty$ for $t > t_1 - h_\nu - \epsilon$, with the convergence uniform in t on any interval $[a, b]$ with $a > t_1 - h_\nu - \epsilon$. This then yields

COROLLARY 4.1. *Suppose $\mathscr{F}_\epsilon$ is a bounded subset of $L_1([0, t_1], R^n)$ such that $f \in \mathscr{F}_\epsilon$ implies $f(t) = 0$ a.e. on $(t_1 - \epsilon, t_1)$, $\epsilon > 0$, where $t_1 - h_\nu - \epsilon \geqslant 0$. Then, if $\det A_\nu \neq 0$, (4.4) obtains for $t > t_1 - h_\nu - \epsilon$ with the convergence being uniform in $f \in \mathscr{F}_\epsilon$ and uniform in t on $[a, b]$ with $t_1 - h_\nu - \epsilon < a < b < \infty$. Equivalently, we may write, using the notation and results of Sections 2 and 3,*

$$x_t^{pN}(\xi, f) \to x_t(\xi, f), \qquad f \in \mathscr{F}_\epsilon,$$

for each $t > t_1 - \epsilon$, the convergence being uniform in $f \in \mathscr{F}_\epsilon$.

We have thus shown that for system (2.1) with η piecewise constant on $[-r, 0]$, and with A_ν nonsingular, we have $x_{t_1}^{pN}(\xi, f) \to x_{t_1}(\xi, f)$ for $t_1 - h_\nu - \epsilon \geqslant 0$, $\xi \in L_1([-h, 0], R^n)$, $f \in \mathscr{F}_\epsilon$, where $\mathscr{F}_\epsilon$ is defined in Corollary 4.1. At first glance the requirement that f vanish on $(t_1 - \epsilon, t_1)$ may appear rather puzzling and somewhat arbitrary. Since $\epsilon > 0$ is arbitrary, we suspect that this is in reality a condition on x at t_1. One observation that sheds some light on this hypothesis on f involves the "boundary condition" $\dot{\phi}(0) = L(\phi)$ required of a C^1 function ϕ in order for it to lie in the domain of $\mathscr{A}$, the infinitesimal generator (2.3) associated with Eq. (2.1). Recalling that $\mathscr{M}_{\lambda_j} \subset \mathscr{D}(\mathscr{A})$ from (3.5), we see that the convergence $x_{t_1}^{PN} \to x_{t_1}$ implies that x_{t_1} is the limit in $\mathscr{C}$ of functions all satisfying the boundary condition $\dot{\phi}(0) = L(\phi)$. Requiring that f vanish in a neighborhood of t_1 yields $\dot{x}(t_1) = L(x_{t_1})$, which is the same boundary condition for x_{t_1}. This boundary condition will also arise naturally in the discussion below of the convergence $\xi^{P^N} \to \xi$ for initial functions.

Let us turn next to a discussion of convergence of ξ^{P^N}, where ξ is the initial function for the system (4.1). The question of expansion of "initial functions" in uniformly convergent series of exponentials (or, equivalently, the question of $\xi^{P^N} \to \xi$ in $\mathscr{C}$) for special cases of (4.1) has been treated by other authors [19]. But arguments substantiating the results presented there appear to be in error [Math. Rev. *33*, no. 3, #2991, March 1967]. Suppose that for a given ξ on $[-r, 0]$ there exists $\tilde{\xi}$ defined on $[-2r - \epsilon, -r]$ so that the function given by

$$\bar{\xi}(t) = \begin{cases} \tilde{\xi}(t), & t \in [-2r - \epsilon, -r], \\ \xi(t), & t \in (-r, 0], \end{cases}$$

has the properties that $\bar{\xi} \in L_1([-2r - \epsilon, -r - \epsilon], R^n)$, $\bar{\xi}$ is absolutely continuous on $[-r - \epsilon, 0]$ and satisfies

$$\dot{\bar{\xi}}(t) = \sum_{i=0}^{\nu} A_i \bar{\xi}(t - h_i), \qquad \text{a.e. } t \text{ in } [-r - \epsilon, 0].$$

We remark that the above comments simply point out that $\xi^{P^N} \to \xi$ if a certain type of backwards continuation theorem holds. We shall present explicit conditions for such a backward continuation in Corollary 4.2. For a general discussion of continuation results the reader should consult [13, 14].

COROLLARY 4.2. *Suppose ξ is absolutely continuous on $[-r, 0]$ with $\xi \in W_1^{(2)}([-\epsilon, 0], R^n)$ for some $\epsilon > 0$. Further, suppose ξ satisfies $\dot{\xi}(0) = \sum_{i=0}^{\nu} A_i \xi(-h_i)$, where A_ν is nonsingular. Then $\xi^{P^N} \to \xi$ in $\mathscr{C}$, where the projections P^N are made relative to the homogeneous form of* (4.1).

Proof. We take, without loss of generality, $\epsilon < h_1$ and define an extension of ξ to $[-2r, 0]$ by

$$\xi(\tau) = A_\nu^{-1} \left\{ \dot{\xi}(\tau + r) - \sum_{i=0}^{\nu-1} A_i \xi(\tau + r - h_i) \right\}, \tag{4.25}$$

for $\tau \in [-2r, -r]$. We see that the extended ξ is continuous at $-r$ and in fact is absolutely continuous on $[-r - \epsilon, 0]$ and in L_1 on $(-2r, -r - \epsilon)$. Using (4.25) we can further extend ξ to $[-2r - \epsilon, 0]$ so that

$$\xi \in L_1([-2r - \epsilon, -r - \epsilon], R^n),$$

from which the results of the corollary then follow.

If the condition $\det A_\nu \neq 0$ is not satisfied, then $p_0 = 0$ and the distribution diagram indicates that the zeros of $\det \Delta(s)$ may be distributed in more than one strip V_i. The reasoning leading to estimates (4.23) and (4.24) is therefore not applicable. However, the order results for $\Delta^{-1}(s)$ described in Sections 12.12–12.14 of [5] can still be applied, yielding

$$\lim_{l\to\infty} \int_{C_l-} | e^{ts} \Delta^{-1}(s)| \, | ds | = 0$$

for $t > (n - 1) h_\nu$. Employing these estimates in arguments similar to those used in establishing Theorem 4.1, we see that in order to obtain convergence of $x_{t_1}^{PN}$ we need I_3 in (4.10) to approach zero uniformly on intervals $[a, b]$ with $a > t_1 - h_\nu - \epsilon$. This will be true if $f(\tau)$ vanishes for $\tau > t_1 - nh_\nu - \epsilon$. In considering terms for the n-vector system analogous to I_2 in (4.9), we also find the added restriction $t_1 > (n + 1) h_\nu$ for $x_{t_1}^{PN} \to x_{t_1}$ in $\mathscr{C}$.

If the n-vector system under consideration actually respresents an nth-order scalar equation, somewhat improved results can be given. Using the estimate (Sections 12.13, 12.14 of [5])

$$\lim_{l\to\infty} \int_{C_l-} | e^{ts} \Delta^{-1}(s)| \, | ds | = 0$$

for $t > 0$, one obtains $x_{t_1}^{PN} \to x_{t_1}$ if $f(\tau) = 0$ for $\tau > t_1 - h_\nu - \epsilon$ and $t_1 > 2h_\nu$.

We remark that the restrictions $t_1 > (n + 1) h_\nu$ and $t_1 > 2h_\nu$ associated with the I_2-type terms (4.9) can be removed by modification of the transform technique used in obtaining (3.2). This necessitates use of a different residue function $p(s)$ and also requires additional smoothness assumptions on the initial function ξ (so that an integration by parts can be performed and the inversion formula will have a larger region of validity). This modified residue function has the form (compare with (4.2)):

$$\tilde{p}(s) = \xi(-h_\nu) e^{sh_\nu} + \int_{-h_\nu}^{0} \dot{\xi}(t) e^{-st} \, dt - \sum_{i=0}^{\nu} A_i e^{-sh_i} \int_{-h_\nu}^{-h_i} \dot{\xi}(t) e^{-st} \, dt.$$

Furthermore, since the residue function is no longer of the form used in (3.3), the equivalency arguments used in Section 3 need modification. Since these modifications would in no way alleviate the more serious restrictions that

$f(\tau) = 0$ for $\tau > t_1 - nh - \epsilon$ ($\tau > t_1 - h_\nu - \epsilon$, respectively), we shall not pursue their development here.

Careful consideration of the above discussion leads one to suspect that even if one takes $f \equiv 0$ in (4.1), i.e., the homogeneous system, one might well have $x_{t_1}^{pN} \to x_{t_1}$ for $t_1 \geqslant nh_\nu$ in some cases where $\det A_\nu = 0$ while for $t_1 < nh_\nu$ this convergence statement does not obtain. That this is, in fact, the case for some n-dimensional systems can be seen from results of Henry [15]. We consider the case of the n-vector equation

$$\dot{x}(t) = A_0 x(t) + A_1 x(t - h_1), \tag{4.26}$$

with solution operator $T(t)$ (see (2.24) and [12, 15]). The adjoint system is given by

$$\dot{z}(t) = -z(t) A_0 - z(t + h_1) A_1 \tag{4.27}$$

and, as in [15], we denote by $T^*(t)$ the functional analytic adjoint of $T(t)$. (Note that while $T^*(t)$ is not the solution operator for (4.27), it is related to the solution operator $\tilde{T}(-t)$ of (4.27) in a very "nice" way [15].) Henry shows that the mapping $t \to \mathcal{N}(T^*(t))$ is nondecreasing and there is a real number δ, $0 \leqslant \delta \leqslant nh_1$, such that $t \to \mathcal{N}(T^*(t))$ is constant for $t \geqslant \delta$. Furthermore, it is easy to construct examples (similar to that in [15, p. 497] or that in [12, p. 36] with, for example $n = 3$ and

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

in (4.27)) so that $t \to \mathcal{N}(T^*(t))$ does not become constant until $t = nh_1$. That is, for such examples, there exists $\psi \in \mathcal{N}(T^*(nh_1))$ such that $\psi \notin \mathcal{N}(T^*(t))$ for $t < nh_1$ so that δ is precisely nh_1. It follows easily from Corollary 2 of [15] that $\overline{\mathscr{R}}(T(t))$, the closure of the range of $T(t)$, contains properly the set $\overline{\operatorname{span}}\{\mathscr{M}_\lambda \mid \lambda \in \sigma(\mathscr{A})\}$ for $t < nh_1$. In fact, one sees that

$$\mathscr{R}(T(t)) \not\subseteq \overline{\operatorname{span}}\{\mathscr{M}_\lambda \mid \lambda \in \sigma(\mathscr{A})\},$$

for $t < nh_1$ while

$$\mathscr{R}(T(t)) \subseteq \overline{\operatorname{span}}\{\mathscr{M}_\lambda \mid \lambda \in \sigma(\mathscr{A})\}$$

for $t \geqslant nh_1$.

We point out that for the unperturbed ($f \equiv 0$) Eq. (4.1), in addition to the results of Henry cited above, some convergence results are already available in [5] (see in particular Theorems 4.2, 4.6, 6.5 and 6.9) which, with proper modifications of the equivalency results of Section 3, could be used to ensure $x_{t_1}^{pN} \to x_{t_1}$ whenever x is a solution of the homogeneous form of (4.1). Other authors (see [5, Theorems 6.10, 6.11; 1]) have obtained convergence results for perturbed systems, but due to smoothness assumptions and/or restrictive regions of convergence these results do not appear to be directly applicable to the problems under discussion here.

Finally, in concluding our remarks on convergence results, we remind the reader that Pitt [25] has also investigated convergence of the exponential series considered above. But his convergence results in L_1 are not applicable in ensuring $x_{t_1}^{PN} \to x_{t_1}$ in $\mathscr{C}$.

5. Optimal Control Via Projection Methods

In this section we shall show how the projection method can be used to investigate optimal control problems for linear retarded FDE with fixed initial and terminal functions.

We consider the problem of minimizing a functional $J: L_2 \to R^1$ on

$$V = \{u \mid u \in \mathscr{U}, x_{t_1}(\xi, u) = \zeta\} \tag{5.1}$$

where $\mathscr{U}$ is a given closed linear subspace of $L_2([0, t_1], R^p)$, and $x(\xi, u)$ is the solution of the n-vector retarded FDE

$$\dot{x}(t) = L(x_t) + Du(t), \qquad t \in [0, t_1], \tag{5.2}$$

$$x_0 = \xi. \tag{5.3}$$

The functions ξ, ζ are given elements of some subset of $\mathscr{C}$, to be detailed below along with the choice of $\mathscr{U}$. We make the controllability assumptions that for the choice of ξ, ζ and $\mathscr{U}$ under consideration, there is at least one $u \in \mathscr{U} \ni x_{t_1}(\xi, u) = \zeta$. That is, V in (5.1) is not empty (see Remark 5.1 below). We also need some continuity and convexity hypotheses on the functional J.

We shall say that a functional J is *quasiconvex* (see [26]) if

$$J((1-\lambda)\, u + \lambda v) \leqslant \max\{J(u), J(v)\}, 0 \leqslant \lambda \leqslant 1, u, v \in L_2 .$$

(As usual, when strict inequality holds for every $\lambda \in (0, 1)$ and $u \neq v$, we say that J is *strictly quasiconvex*.) An equivalent definition [8] of quasiconvexity requires that $E_\alpha = \{v \mid J(v) \leqslant \alpha\}$ be convex for each $\alpha \in R^1$. Since lower semi-continuity (l.s.c.) of J is characterized by the sets E_α being closed [27], Mazur's theorem [9, p. 422] yields immediately that J l.s.c. and quasiconvex imply that J is weakly l.s.c. With these well-known observations in mind, we make the following assumptions on J:

H(i): J is strictly quasiconvex and l.s.c. on L_2 ;

H(ii): For $\mathscr{K} \subset L_2$, $J(v) \leqslant M$, $\forall v \in \mathscr{K}$, implies $\mathscr{K}$ bounded.

Using quite standard arguments, one obtains immediately the well-known results

THEOREM 5.1 *Under $H(i)$ and $H(ii)$, there exists a unique $e^* \in E$ such that $J(e^*) = \inf\{J(e) | e \in E\}$ for any given nonempty closed convex $E \subset L_2$.*

An easy application of this theorem yields the following theorem.

THEOREM 5.2. *Under* H(i), H(ii), *there exists a unique element u^* in V satisfying*

$$J(u^*) = \inf_{v \in V} J(v).$$

Proof. The mapping $u \to x_{t_1}(\xi, u)$ is a linear map of L_2 into $\mathscr{C}$. From the variation of constants formula

$$x_{t_1}(\xi, u)(\theta) = x_{t_1}(\xi, 0)(\theta) + \int_0^{t_1+\theta} X(t_1 + \theta, s)\, Du(s)\, ds$$

(where X is the appropriate "fundamental matrix" for (5.2)—see [12]), one sees easily that this map is also continuous. The assumptions on $\mathscr{U}$ then imply that V is closed, convex and nonempty in L_2, and hence Theorem 5.1 assures the desired result.

Using the notation of Section 2, we let $y^N(t) = \langle \Psi^N, x_t \rangle$ where x is the solution of (5.2), (5.3). Denoting by $y^N(t; u)$ the solution of

$$\dot{y}^N(t) = B^N y^N(t) + \Psi^N(0)\, Du(t), \qquad t \in [0, t_1], \tag{5.4}$$

$$y^N(0) = \langle \Psi^N, x_0(\xi, u) \rangle = \langle \Psi^N, \xi \rangle, \tag{5.5}$$

we project the original problem described above onto the associated eigenmanifolds. The finite-dimensional problem obtained is that of minimizing J on the set

$$V^N = \{u \mid u \in \mathscr{U}, y^N(t_1; u) = \langle \Psi^N, \zeta \rangle\}. \tag{5.6}$$

Let us first consider the case where the eigenvalues λ_j are simple and $p = 1$ (scalar controls). Then we have

$$y^N(t) = \operatorname{col}(y_{\lambda_1}(t), \ldots, y_{\lambda_N}(t)),$$

where each y_{λ_i} is a scalar function. Denoting by k_j the scalar given by $\psi_{\lambda_j}(0)D$, we may write in place of (5.4), (5.5) for $j = 1, 2, \ldots, N$:

$$\dot{y}_{\lambda_j}(t) = \lambda_j y_{\lambda_j}(t) + k_j u(t), \tag{5.7}$$

$$y_{\lambda_j}(0) = \langle \psi_{\lambda_j}, x_0 \rangle = \langle \psi_{\lambda_j}, \xi \rangle. \tag{5.8}$$

The solution of (5.7) is given by

$$y_{\lambda_j}(t) = \exp(\lambda_j t)\, y_{\lambda_j}(0) + \int_0^t \exp(\lambda_j(t-\tau))\, k_j u(\tau)\, d\tau, \qquad j = 1,\dots, N.$$

The boundary conditions in (5.8) then reduce to the system of moment equations

$$\int_0^{t_1} \exp(-\lambda_j \tau)\, k_j u(\tau)\, d\tau = l_j \tag{5.9}$$

where

$$l_j = \langle \psi_{\lambda_j}, \zeta \rangle \exp(-\lambda_j t_1) - \langle \psi_{\lambda_j}, \xi \rangle. \tag{5.10}$$

Assume that the eigenvalues λ_j have been ordered so that the first ν eigenvalues are real and the remaining ones are complex pairwise conjugate. Then $k_1, \dots, k_\nu$ will also be real, and $k_{\nu+1}, k_{\nu+2}, \dots$ will be complex pairwise conjugate; the same property will be valid with respect to $l_1, \dots, l_\nu, l_{\nu+1}, \dots$. Then the system (5.9) can be transformed to a real form by multiplying it from the left by a quasidiagonal $N \times N$ matrix M defined by

$$M = \text{quasidiag}\left\{ I_\nu, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1/2i & -1/2i \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1/2i & -1/2i \end{bmatrix}, \dots \right\},$$

where I_ν is the identity matrix of dimension ν. Clearly M is invertible. Define

$$\begin{aligned} F(\tau) &= \text{diag}\{\exp(-\lambda_1\tau), \exp(-\lambda_2\tau), \dots, \exp(-\lambda_N\tau)\} \\ k &= \text{col}(k_1, k_2, \dots, k_N), \\ l &= \text{col}(l_1, \dots, l_N). \end{aligned}$$

Then (5.9) can be rewritten as

$$\int_0^{t_1} F(\tau)\, k u(\tau)\, d\tau = l.$$

Multiplying this equation from the left by M, one has

$$\int_0^{t_1} g(\tau)\, u(\tau)\, d\tau = c, \tag{5.11}$$

where $g(\tau) = MF(\tau)k$, $c = Ml$. Letting $g = \text{col}(g_1, \dots, g_N)$ and $c = \text{col}(c_1, \dots, c_N)$, the set V^N defined in (5.6) can also be written as

$$V^N = \{u \in \mathscr{U} \mid \langle g_j, u \rangle_2 = c_j, j = 1,\dots, N\} \tag{5.12}$$

where $\langle\ ,\ \rangle_2$ denotes the inner product in the Hilbert space L_2. We observe that $\mathscr{U}$ itself is a Hilbert space with the inner product $\langle\ ,\ \rangle_2$ and thus we see that the finite-dimensional problem reduces to a classical constrained optimization problem involving a convex (quasiconvex) functional on a Hilbert space with a finite number of constraints of equality or "moment" type (see [18, Chapter 9]). That the constraints in (5.12) are consistent ($V^N \neq \phi$) for every N follows from the assumption that V given in (5.1) is not empty.

All the considerations above can clearly be repeated for the more general case of multiple eigenvalues λ_j and for vector-valued control functions. Let k_j denote the $d_j \times p$ matrix (where d_j is the multiplicity of λ_j and u is $p \times 1$) defined by $k_j = \Psi_{\lambda_j}^T(0)D$. Upon use of (2.23) one obtains the analogues of (5.9), (5.10)

$$\int_0^{t_1} \exp(-B_{\lambda_j}\tau)\, k_j u(\tau)\, d\tau = l_j\,, \tag{5.13}$$

where

$$l_j = \langle \Psi_{\lambda_j}, \zeta\rangle \exp(-B_{\lambda_j} t_1) - \langle \Psi_{\lambda_j}, \xi\rangle.$$

Here each l_j is actually a $d_j \times 1$ vector. Defining

$$\begin{aligned} F(\tau) &= \text{quasidiag}\{\exp(-B_{\lambda_1}\tau),\ldots, \exp(-B_{\lambda_N}\tau)\},\\ k &= (k_1{}^T,\ldots, k_N{}^T)^T,\\ l &= (l_1{}^T,\ldots, l_N{}^T)^T \end{aligned}$$

and an appropriate $\sum_1^N d_j \times \sum_1^N d_j$ matrix M, we are led once again to moment equations

$$\int_0^{t_1} g_i(\tau)\, u(\tau)\, d\tau = c_i\,, \qquad i = 1,\ldots, \sum_1^N d_j\,,$$

where $g(\tau) = MF(\tau)k$, $c = Ml$. The formulation is then as before with the controls u in $\mathscr{U} \subset L_2([0, t_1], R^p)$ being $p \times 1$ vector functions and the constraint functions g_i lying in $L_2([0, t_1], R^{p*})$ being $1 \times p$ vector functions.

We next turn to a discussion of the behavior of the optimal controls $\bar{u}^N$ obtained from the finite-dimensional problems; i.e., we denote by $\bar{u}^N$ the unique solution (see Theorem 5.1) of minimizing J on V^N. We wish, of course, to ascertain that $J(\bar{u}^N) \to J(u^*)$ and, if possible, to show that $\bar{u}^N \to u^*$ in some sense.

Lemma 5.1. *Let, V, V^N be defined as in* (5.1) *and* (5.6), *respectively. Then $u \in V$ implies $u \in V^N$ for every N.*

Proof. Recalling that $y^N(t; u)$ denotes the solution of (5.4), (5.5), and also $y^N(t; u) = \langle \Psi^N, x_t(\xi, u)\rangle$, we see that for $u \in V$,

$$P^N x_{t_1}(\xi, u) = P^N \zeta = \Phi^N \langle \Psi^N, \zeta\rangle,$$

while

$$P^N x_{t_1}(\xi, u) = \Phi^N y^N(t_1; u)$$

so that $y^N(t_1 ; u) = \langle \Psi^N, \zeta\rangle (\Phi^N$ is a basis for $\sum_1^N \mathscr{M}_{\lambda_j}(\mathscr{A}))$ and thus $u \in V^N$.

From the above lemma and the assumption that V is nonempty, we conclude that there is a $u \in V$ such that $J(\bar{u}^N) \leqslant J(u)$ for every N. Hence, by H(ii) and the weak compactness of bounded subsets of L_2, there are $\bar{u} \in L_2$ and a subsequence $\{\bar{u}^{N_\nu}\}$ such that $\bar{u}^{N_\nu}$ converges weakly to $\bar{u}$ in L_2. Furthermore, $\bar{u} \in \mathscr{U}$ since $\mathscr{U}$ is convex and closed in L_2 and hence weakly closed.

We can in fact prove $\bar{u} \in V^N$ for each N. From the nature of the projections made in section 2, it follows immediately that $V^{N+1} \subset V^N$ and hence $J(\bar{u}^N) \leqslant J(\bar{u}^{N+1})$ for each N. Thus, for $N_m > N_\nu$ we have

$$y^{N_\nu}(t_1; \bar{u}^{N_m}) = \langle \Psi^{N_\nu}, \zeta\rangle.$$

But from the usual variation of parameters representation for (5.4), (5.5) and the fact that $\bar{u}^{N_m} \rightharpoonup \bar{u}$, we obtain

$$y^{N_\nu}(t_1; \bar{u}^{N_m}) \to y^{N_\nu}(t_1; \bar{u})$$

as $N_m \to \infty$, for each N_ν. It follows then that $y^{N_\nu}(t_1 ; \bar{u}) = \langle \Psi^{N_\nu}, \zeta\rangle$. Thus $\bar{u} \in V^{N_\nu}$. But $V^{N+1} \subset V^N$ and $\bar{u} \in V^{N_\nu}$ for a sequence N_ν with $N_\nu \to \infty$ yields that $\bar{u} \in V^N$ for each N.

From the preceding arguments we see that $J(\bar{u}^N) \leqslant J(\bar{u})$ for each N and thus weak lower semicontinuity of J yields

$$J(\bar{u}) \leqslant \underline{\lim}\, J(\bar{u}^{N_\nu}) \leqslant \overline{\lim}\, J(\bar{u}^{N_\nu}) \leqslant J(\bar{u})$$

so that $\lim J(\bar{u}^{N_\nu})$ exists and equals $J(\bar{u})$. In fact, since $J(\bar{u}^N) \leqslant J(\bar{u}^{N+1})$, we obtain actually

$$\lim J(\bar{u}^N) = J(\bar{u}).$$

If we further assume that $\mathscr{U}$ has been chosen so that the convergence results of Section 4 are valid (i.e., $\mathscr{U}$ is chosen so that $u \in \mathscr{U}$ implies $P^N x_{t_1}(\xi, u) \to x_{t_1}(\xi, u)$), the fact that $\bar{u} \in \mathscr{U}$ and $u^* \in \mathscr{U}$ with $x_{t_1}(\xi, u^*) = \zeta$ yields

$$x_{t_1}(\xi, \bar{u}) = \zeta \tag{5.14}$$

by letting $N \to \infty$ in the equality

$$P^N x_{t_1}(\xi, \bar{u}) = \Phi^N y^N(t_1; \bar{u}) = \Phi^N \langle \Psi^N, \zeta\rangle = P^N \zeta.$$

From (5.14) it follows at once that $\bar{u} \in V$. Now if u is any control in V, an application of lemma 5.1 yields $J(\bar{u}^N) \leqslant J(u)$ for every N. Thus, since $\lim J(\bar{u}^N) = J(\bar{u})$, we see that $J(\bar{u}) \leqslant J(u)$, $u \in V$, and by the uniqueness of u^* in Theorem 5.2, we find $\bar{u} = u^*$ and hence $J(\bar{u}^N) \to J(u^*)$.

Furthermore, we have seen above that $\bar{u}^{N_\nu} \rightharpoonup \bar{u} = u^*$. In fact, $\bar{u}^N \rightharpoonup u^*$. For if $\{\bar{u}^K\}$ is any subsequence of $\{\bar{u}^N\}$, $\{\bar{u}^K\}$ has a weakly convergent subsequence $\{\bar{u}^{K_m}\}$ converging weakly to some $\tilde{u}$ in L_2. By arguments exactly like those used above, one shows that $\tilde{u} = u^*$ and thus $\bar{u}^{K_m} \rightharpoonup u^*$ in L_2. Hence any subsequence $\{\bar{u}^K\}$ of $\{\bar{u}^N\}$ has in turn a subsequence $\{\bar{u}^{K_m}\}$ converging weakly in L_2 to u^* so that $\bar{u}^N \rightharpoonup u^*$ in L_2. Therefore we have established

THEOREM 5.3. *Under H(i), H(ii), and the assumption that the system* (5.2), (5.3) *and $\mathscr{U}$ are such that the convergence results of section* 4 *hold* (*i.e., see Theorem* 4.1), *one has $J(\bar{u}^N) \to J(u^*)$ and $\bar{u}^N \rightharpoonup u^*$ in L_2.*

Remark 5.1. We observe that a choice of $\mathscr{U}$ in the above problems in now quite obvious. For example, if (5.2) is a scalar differential difference equation or a vector system with A_ν nonsingular as given in (4.1) with $f = Du$ and $x_0 = \xi$, $\xi \in \mathscr{C}$, then an appropriate choice of $\mathscr{U}$ which satisfies the necessary hypotheses is

$$\mathscr{U} = \mathscr{U}_\epsilon \equiv \{u \in L_2 \mid u(t) = 0 \text{ a.e. } t \text{ in } (t_1 - \epsilon, t_1)\},$$

where $\epsilon > 0$.

While we do not wish to go into a lengthy discussion of controllability here, it is perhaps appropriate to make some comments on the assumption that V given in (5.1) is nonempty whenever $\mathscr{U} = \mathscr{U}_\epsilon$. We consider for simplicty the system (4.1) with $\nu = 1$ and $f = Du$,

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-h) + Du(t) \tag{5.16}$$

and let ϵ be as in the definition of $\mathscr{U}_\epsilon$. Let ξ be fixed and $\zeta \in W_2^{(1)}([-h, 0], R^n)$. Assume

H(a): $\dot{\zeta}(\theta) - A_0\zeta(\theta) \in \mathscr{R}(A_1)$ for $\theta \in [-\epsilon, 0]$ with, in particular, $\dot{\zeta}(0) = A_0\zeta(0) + A_1\zeta(-h)$.

(If A_1 is nonsingular then only the latter assumption in H(a) need be made.) Using the pseudoinverse of A_1 (see [3]), one can then extend x, where $x_{t_1} = \zeta$, backwards to obtain a function $\tilde{x}$ in $W_2^{(1)}([t_1 - h - \epsilon, t_1 - h], R^n)$. Defining a function w by

$$w(t) = \begin{cases} \tilde{x}(t), & t \in [t_1 - h - \epsilon, t_1 - h], \\ \zeta(t - t_1), & t \in (t_1 - h, t_1 - \epsilon], \end{cases}$$

we make a further assumption:

H(b): $w_{t_1-\epsilon}$ is attainable from ξ by (5.16) using controls in $L_2([0, t_1-\epsilon], R^p)$; i.e., for some $\tilde{u}$ in $L_2([0, t_1-\epsilon], R^p)$ we have $x_{t_1-\epsilon}(\xi, \tilde{u}) = w_{t_1-\epsilon}$.

Under these two assumptions H(a), H(b), extending $\tilde{u}$ to $[0, t_1]$ by taking $\tilde{u}(t) = 0$, $t \in (t_1-\epsilon, t_1]$, we obtain $x_{t_1}(\xi, \tilde{u}) = \zeta$ with $\tilde{u}$ in $\mathscr{U}_\epsilon$ as defined above.

We present some situations where H(a), H(b) are satisfied:

(i) If $\zeta = 0$, then H(a) is obviously true and H(b) obtains if and only if (5.16) is null functional controllable with respect to L_2 controls on $[0, t_1-\epsilon]$. Sufficient conditions (which involve computable criteria) for this latter requirement can be found in [3].

(ii) If A_1 is nonsingular and ζ is such that $\dot{\zeta}(0) = A_0\zeta(0) + A_1\zeta(-h)$, then H(a) holds. If D is nonsingular, H(b) is satisfied. Note that this includes all scalar systems $\dot{x}(t) = a_0x(t) + a_1x(t-h) + du(t)$ where $a_1 \neq 0$, $d \neq 0$ and $\dot{\zeta}(0) = a_0\zeta(0) + a_1\zeta(-h)$.

Similar remarks can be made to show that the assumption $V \neq \varnothing$ is often satisfied for nth-order scalar equations with the appropriately defined $\mathscr{U}_\epsilon$ (relative to the desired convergence results discussed above). For a more detailed discussion of conditions for functional controllability (which, in many cases, can be modified to apply to the present situation), we refer the reader to [3, 4] and [10].

Remark 5.2. As we shall see below, the finite constraint problems obtained above may, in some instances, be analytically solvable for the controls $\bar{u}^N$. Even in cases where this is not possible, these problems may often be amenable to standard numerical techniques (see [18, p. 297 ff.]) leading to approximations for the $\bar{u}^N$ which are in turn approximations (in the sense of Theorem 5.3 and Theorem 5.4 below) for the solution of the originally posed infinite-dimensional problem.

Remark 5.3. We point out that if additional constraints are placed on the controls, say $|u|_2 \leqslant l$, or $|u(t)| \leqslant l$ a.e., so that in the above formulation we take $\mathscr{U}$ modified by the addition of these constraints, the preceding arguments and results can be carried through with little change (i.e., it is not crucial to the above developments that the set of admissible controls $\mathscr{U}$ be a linear subspace of L_2; a closed and convex subset will suffice). We note that in this case the set $\mathscr{U}$ would be bounded (as would V, V^N) so that the assumption H(ii) could be omitted.

As might be expected, under further assumptions on J, we can improve the results in the above theorem to obtain $\bar{u}^N \rightarrow u^*$ in L_2. Following Poljak [26] (see also [8]), we say that J is *strongly convex* if $\exists\, \delta > 0$ such that

$$J\left(\frac{u+v}{2}\right) \leqslant \tfrac{1}{2}J(u) + \tfrac{1}{2}J(v) - \tfrac{1}{4}\delta \mid u - v \mid^2,$$

for every $u, v \in L_2$. In [26], Poljak gives conditions on f for $J(u) = \int f(x, u, t)$ to be strongly convex. We shall discuss below a class of problems of interest for which one can easily verify strong convexity of the cost functional J. In fact, one can readily prove directly for this class of problems that $\bar{u}^N \rightarrow u^*$ in L_2. We shall, for the sake of generality and completeness, give the simple proof here for problems with strongly convex payoffs.

THEOREM 5.4. *In addition to the assumptions of Theorem* 5.3, *suppose that J is strongly convex. Then $\bar{u}^N \rightarrow u^*$ in L_2.*

Proof. We have

$$0 \leqslant \tfrac{1}{4}\delta \mid \bar{u}^N - u^* \mid^2 \leqslant \tfrac{1}{2}J(\bar{u}^N) + \tfrac{1}{2}J(u^*) - J\left(\frac{\bar{u}^N + u^*}{2}\right)$$

where, by Theorem 5.3, $\bar{u}^N \rightharpoonup u^*$ and $J(\bar{u}^N) \rightarrow J(u^*)$. Since J weakly l.s.c. implies $-J$ weakly upper semicontinuous, we have

$$\overline{\lim} \left\{-J\left(\frac{\bar{u}^N + u^*}{2}\right)\right\} \leqslant -J\left(\frac{u^* + u^*}{2}\right) = -J(u^*)$$

and hence

$$0 \leqslant \tfrac{1}{4}\delta\, \overline{\lim} \mid \bar{u}^N - u^* \mid^2 \leqslant \tfrac{1}{2}J(u^*) + \tfrac{1}{2}J(u^*) - J(u^*) = 0$$

or

$$\mid \bar{u}^N - u^* \mid \rightarrow 0 \text{ as } N \rightarrow \infty.$$

We turn now to the class of problems where J has the form

$$J(u) = \mathscr{B}(u, u) + \mathscr{L}(u) + \mathscr{K} \tag{5.15}$$

where $u, v \rightarrow \mathscr{B}(u, v)$ is a symmetric continuous bilinear functional satisfying $\mathscr{B}(u, u) \geqslant \delta \mid u \mid^2$ for some $\delta > 0$, $u \rightarrow \mathscr{L}(u)$ is a continuous linear functional, and $\mathscr{K}$ is a constant. It is readily seen that $u \rightarrow \mathscr{B}(u, u)$ is a strictly convex, l.s.c. mapping [17, p. 7] from which it follows that J given by (5.15) satisfies H(i). Furthermore, since $\mid \mathscr{L}(u) \mid \leqslant m \mid u \mid$, we have

$$J(u) \geqslant \delta \mid u \mid^2 - m \mid u \mid + \mathscr{K}$$

so that H(ii) is satisfied. Finally, using the identities

$$\mathscr{B}\left(\frac{u+v}{2}, \frac{u+v}{2}\right) + \mathscr{B}\left(\frac{u-v}{2}, \frac{u-v}{2}\right) = \tfrac{1}{2}\mathscr{B}(u, u) + \tfrac{1}{2}\mathscr{B}(v, v)$$

and

$$\mathscr{L}\left(\frac{u+v}{2}\right) = \tfrac{1}{2}\mathscr{L}(u) + \tfrac{1}{2}\mathscr{L}(v)$$

together with the inequality

$$-\mathscr{B}(u, u) \leqslant -\delta \mid u \mid^2,$$

one obtains

$$\begin{aligned} J\left(\frac{u+v}{2}\right) &= \mathscr{B}\left(\frac{u+v}{2}, \frac{u+v}{2}\right) + \mathscr{L}\left(\frac{u+v}{2}\right) + \mathscr{K} \\ &= \tfrac{1}{2}J(u) + \tfrac{1}{2}J(v) - \mathscr{B}\left(\frac{u-v}{2}, \frac{u-v}{2}\right) \\ &\leqslant \tfrac{1}{2}J(u) + \tfrac{1}{2}J(v) - \tfrac{1}{4}\delta \mid u - v \mid^2, \end{aligned}$$

so that J is strongly convex. The results of Theorems 5.3 and 5.4 are thus seen to be valid for problems having cost functionals (5.15). We mention next an often-studied class of problems having such a cost functional.

Let $\mathscr{W}$ and $\mathscr{R}$ be symmetric $n \times n$ and $p \times p$ matrices respectively with $\mathscr{W} \geqslant 0$, $\mathscr{R} > 0$ and define

$$J(u) = \int_0^{t_1} \{x^T(t)\, \mathscr{W} x(t) + u^T(t)\, \mathscr{R} u(t)\}\, dt,$$

where x is the solution of (5.2), (5.3). Using the variation of parameters representation referred to earlier, one obtains (with straightforward manipulations) that J has the form (5.15) where

$$\begin{aligned} \mathscr{B}(u, v) \equiv \int_0^{t_1} \Big\{ &\Big[\int_0^t X(t-s)\, Du(s)\, ds\Big]^T \\ &\times \mathscr{W} \Big[\int_0^t X(t-\tau)\, Dv(\tau)\, d\tau\Big] + u^T(t)\, \mathscr{R} v(t)\Big\}\, dt, \end{aligned}$$

$$\mathscr{L}(u) \equiv \int_0^{t_1} 2x^T(t, \xi, 0)\, \mathscr{W} \Big\{\int_0^t X(t-s)\, Du(s)\, ds\Big\}\, dt$$

and

$$\mathscr{K} \equiv \int_0^{t_1} x^T(t, \xi, 0)\, \mathscr{W} x(t, \xi, 0)\, dt,$$

with $\mathscr{B}$ satisfying $\mathscr{B}(u, u) \geqslant \delta \mid u \mid^2$, $\delta > 0$ since $\mathscr{R} > 0$.

A free-endpoint optimal control problem for linear FDE with integral quadratic cost functional over an infinite time interval ($t_1 = \infty$) has been considered by projection methods in [20–22]. However, as mentioned in Section 1, those results appear to be based on formal arguments and erroneous convergence results.

We consider the special case of these integral convex cost functional problems where $\mathscr{W} = 0$ and $\mathscr{R}$ is the $p \times p$ identity matrix; i.e., $J(u) = | u |^2$. As it turns out, this is equivalent to the minimum norm problem, $\tilde{J}(u) = | u |$. That is, u^* such that $| u^* |^2 = \inf\{| u |^2 \mid u \in V\}$ is the unique minimal norm element in the closed convex set V [9, p. 74]. Similarly, the finite-dimensional problems of minimizing J on V^N reduce to those of seeking the minimal norm element in V^N.

We apply well-known Hilbert space projection results to the problem of finding the element of minimal norm in V^N for the case of simple eigenvalues and scalar controls (the multiple eigenvalue and p-vector control situation can be handled in the same manner). We shall assume that for every N, the g_j of (5.12), $j = 1,..., N$ are also in $\mathscr{U}' \simeq \mathscr{U}$, where $\mathscr{U}'$ is the dual of $\mathscr{U}$. That we can, without loss of generality, do so is obvious from the problem formulation and our choice of $\mathscr{U}$ as discussed in Remark 5.1 above. Let $\tilde{u} \in V^N$, i.e., $\langle g_j , \tilde{u}\rangle_2 = c_j , j = 1,..., N$. Then we may write

$$V^N = \{u \in \mathscr{U} \mid \langle g_j , u - \tilde{u}\rangle_2 = 0, j = 1,..., N\}.$$

Defining $\mathscr{N}^N$ by

$$\mathscr{N}^N = \{v \in \mathscr{U} \mid \langle g_j , v\rangle_2 = 0, j = 1,..., N\},$$

we easily see that

$$V^N = \tilde{u} + \mathscr{N}^N,$$

so that the linear variety V^N is a translate of the closed linear subspace $\mathscr{N}^N$. Application of a version of the projection theorem [18, Theorem 3, p. 64] yields the existence of a unique element $\bar{u}^N$ of minimal norm in V^N with $\bar{u}^N \perp \mathscr{N}^N$. Letting $\mathscr{Z}^N$ denote the span of $\{g_1 ,..., g_N\}$ in $\mathscr{U}$, we see that the orthogonality requirement becomes the alignment condition $\bar{u}^N \in \mathscr{Z}^N$ or

$$\bar{u}^N = \sum_{j=1}^{N} e_j g_j . \tag{5.17}$$

This alignment condition, together with the constraints in (5.12), gives a computable expression for $\bar{u}^N$. If $\{g_1 ,..., g_N\}$ is a linearly independent set in $\mathscr{U}$, then substitution of (5.17) into these constraints yields

$$\sum_{j=1}^{N} \langle g_i , g_j\rangle_2 e_j = c_i , \qquad i = 1,..., N$$

or

$$He = c,$$

where $e = \text{col}(e_1, ..., e_N)$, and H is the nonsingular matrix with elements $h_{ij} = \langle g_i, g_j \rangle_2$. We then obtain

$$\bar{u}^N = g^T H^{-1} c.$$

Defining

$$\Gamma \equiv \int_0^{t_1} F(\tau)\, kk^* F^*(\tau)\, d\tau$$

(where here * denotes the conjugate transpose), we see that

$$M\Gamma M^* = \int_0^{t_1} MF(\tau)\, kk^* F^*(\tau)\, M^*\, d\tau$$
$$= \langle g, g^T \rangle_2 = H$$

so that (5.17) may be written

$$\bar{u}^N = k^* F^* \Gamma^{-1} l. \tag{5.18}$$

We remark that the matrix Γ is actually the familiar controllability matrix for the system (5.7). The assumption that $\{g_1, ..., g_N\}$ is a linearly independent set is thus easily seen to be equivalent to the assumption that (5.7) is controllable in the usual sense.

If the set $\{g_1, ..., g_N\}$ does not constitute a linearly independent set in $\mathscr{U}$, then a moment's reflection upon the above arguments yields a similar scheme for computing $\bar{u}^N$ using a maximal linearly independent subset of $\{g_1, ..., g_N\}$, i.e., a set $\{g_{N_j}\}_{j=1}^q$, $q < N$, that is linearly independent and spans $\mathscr{Z}^N$.

Remark 5.4. We point out that, for the case of simple eigenvalues, controllability of (5.7) in the usual sense is implied by controllability of (5.2) in a certain functional sense. (We have *not*, in fact, in this paper made such an assumption on (5.2), and make the comments in this remark only as an aside to the reader.) Suppose that we assume (5.2) controllable in the sense that given any ξ, ζ in $\overline{\text{span}}\{\mathscr{M}_\lambda \mid \lambda \in \sigma(\mathscr{A})\}$, there exists $u \in \mathscr{U}$ such that $x_{t_1}(\xi, u) = \zeta$. Taking i arbitrary but fixed, $1 \leqslant i \leqslant N$, we choose $\xi = \phi_{\lambda_i}$ and $\zeta = \alpha\phi_{\lambda_i}$ where α is a scalar such that $\alpha e^{-\lambda_i t_1} - 1 \neq 0$. Then by (5.10) and the orthonormality condition (2.16) we have

$$l_i = \{\alpha e^{-\lambda_i t_1} - 1\} \langle \psi_{\lambda_i}, \phi_{\lambda_i} \rangle$$
$$= \{\alpha e^{-\lambda_i t_1} - 1\} \neq 0$$

which, by (5.9), implies $k_i \neq 0$. Hence, an assumption of function space controllability in the above sense for (5.2) implies $k_i \neq 0$, $i = 1,..., N$. But for simple eigenvalues, this latter condition is easily seen to be equivalent to controllability of (5.7) in the usual (Euclidean) sense.

The above results show that there is a complete analogy between the known solutions to the terminal control problem with minimal control energy for ordinary differential systems (see [29, Section II.C]) and the solution to the same problem in a finite-dimensional subspace of the state space of the FDE. The formulae (5.18) and (5.10) show that the control $\bar{u}^N$ is linear with respect to projection of initial and terminal functions.

We remark that the simplicity of solution to the minimum-norm moment problems encountered above is due to the formulation of the problem so that the control space $\mathscr{U}$ is a Hilbert space. Were we to formulate the problem for controls in some Banach space, we would then use the Hahn-Banach theorem in lieu of the projection theorem in Hilbert space. This, however, requires that the control space be the dual of some space, for example $\mathscr{U} = L_\infty = L_1^*$, and that the constraint functions g_i be considered as elements of that space. For a complete discussion of this we refer the reader to [24], [18, Section 5.8, 5.9] and to the references noted by these authors.

Finally, we wish to remark that the minimum-norm problem for delayed systems discussed above was also considered by I. Lasiecka in her thesis [16a] where she used finite-dimensional projection methods and pointed out the difficulties involved in proving convergence of $x_t^{QN} \to 0$ via use of (2.24) and estimates (2.27), (2.28) and (2.29).

The above discussions indicate that these projection methods for optimal control problems (5.1)–(5.3) offer a somewhat satisfactory approach from a theoretical viewpoint. While we are not yet in a position to make broad claims concerning the practical usefulness of these methods, our initial computational investigations, as documented in the following two examples, are encouraging. To test these methods we have chosen two examples for which we can obtain the exact solutions via other methods (see [2, 16]).

EXAMPLE 5.1. Minimize

$$J(u) = \int_0^2 u^2(t)\, dt,$$

subject to

$$\dot{x}(t) = \frac{1}{(2)^{1/2}}\, x(t) + \frac{1}{(2)^{1/2}}\, x(t-1) + u(t), \qquad t \in [0, 2],$$

$$x_0 = \xi = 1, \qquad x_2 = \zeta = 0.$$

The exact solution, obtained using calculus of variations arguments, is given by

$$u^*(t) = \begin{cases} \frac{1}{2}\{\delta_1(1 - 1/(2)^{1/2})\, e^t + \delta_2(-1 - 1/(2)^{1/2})\, e^{-t} - 1/(2)^{1/2}\}, & t \in [0, 1] \\ -1/2(2)^{1/2}\{\delta_1 e^{t-1} + \delta_2 e^{-(t-1)} - 1\}, & t \in [1, 2], \end{cases}$$

where

$$\delta_1 = \frac{3 - e}{1 - e^2}, \qquad \delta_2 = \frac{(1 - 3e)e}{1 - e^2},$$

with

$$\begin{aligned} J(u^*) &= \frac{1}{4(e^2 - 1)}\{(11 + 6(2)^{1/2})\, e^2 - 12e + (9 - 6(2)^{1/2})\} \\ &\approx 4.38. \end{aligned}$$

The characteristic equation (2.9), $\Delta(\lambda) = \lambda - (2)^{1/2} - 1/(2)^{1/2}\, e^{-\lambda} = 0$, has only simple roots, one of which is real, the others occurring in conjugate pairs. Newton's method, with asymptotic estimates as given in [5] as starting values, yields approximate ($|\Delta(\lambda_j)| < 10^{-8}$) root values

$$\begin{aligned} \lambda_1 &= .974 \\ \lambda_2, \lambda_3 &= -1.939 \pm 4.144i \\ \lambda_4, \lambda_5 &= -2.765 \pm 10.68i \\ \lambda_6, \lambda_7 &= -3.208 \pm 17.05i \\ \lambda_8, \lambda_9 &= -3.514 \pm 23.38i \\ \lambda_{10}, \lambda_{11} &= -3.748 \pm 29.69i, \end{aligned}$$
etc.

Using the notation above (see (5.4)–(5.6) through Lemma 5.1), our computations yield:

N	$J(\bar{u}^N)$	Upper bound for $\lvert x_{t_1}(\xi, \bar{u}^N) - \zeta \rvert$
3	4.284	0.13
5	4.313	0.09
7	4.328	0.07
11	4.343	0.05
21	4.357	0.03
41	4.365	0.02
61	4.369	0.007

EXAMPLE 5.2. Minimize

$$J(u) = \int_0^2 u^2(t)\, dt$$

subject to

$$\dot{x}(t) = -x(t-1) + u(t), \qquad t \in [0, 2],$$
$$x_0 = \xi = 1, \qquad x_2 = \zeta = 0.$$

The exact solution (see [16]) is given by

$$u^*(t) = \begin{cases} \delta(e^t + e^{2-t}) + 1, & t \in [0, 1] \\ \delta(e^{t-1} - e^{2-(t-1)}), & t \in [1, 2], \end{cases}$$

where $\delta = 1/(1 - e^2)$, with

$$J(u^*) = -2\delta \simeq .313.$$

The characteristic equation, $\Delta(\lambda) = \lambda + e^{-\lambda} = 0$, has no real roots and the roots, occurring in conjugate pairs, are all simple. Approximate root values are:

$$\begin{aligned} \lambda_1, \lambda_2 &= -.3181 \pm 1.337i \\ \lambda_3, \lambda_4 &= -2.062 \pm 7.588i \\ \lambda_5, \lambda_6 &= -2.653 \pm 13.95i \\ \lambda_7, \lambda_8 &= -3.020 \pm 20.27i \\ \lambda_9, \lambda_{10} &= -3.287 \pm 26.58i \\ \lambda_{11}, \lambda_{12} &= -3.498 \pm 32.88i, \end{aligned}$$

etc.

Computations yield:

N	$J(\bar{u}^N)$	Upper bound for $\lvert x_{t_1}(\xi, \bar{u}^N) - \zeta \rvert$
2	.2059	0.155
4	.2281	0.114
6	.2429	0.09
8	.2531	0.07
10	.2605	0.06
20	.2799	0.04
100	.3035	0.01

For both of the above examples, it appears that the convergence for both $J(\bar{u}^N)$ and $x_{t_1}(\xi, \bar{u}^N) - \zeta$ is reasonable. Our computations showed that the error in satisfying the terminal boundary condition (as measured by the bound on the supremum of $x_{t_1}(\xi, \bar{u}^N)(\theta) - \zeta(\theta)$, $\theta \in [-r, 0]$, given in column 3) is always largest at $\theta = -r$. This can be related to the manner in which $P^N\phi$

"converges" to ϕ in the event that ϕ is not required to satisfy the condition $\dot{\phi}(0) = L(\phi)$(in our numerical computations we did not force the condition $u^*(t) = 0$ on $(t_1 - \epsilon, t_1)$).

While the above two examples involve scalar systems, our preliminary investigations of vector systems yield similar computational behavior to that catalogued above. A more detailed and complete treatment of the numerical aspects of these methods will appear in a future paper.

The above methods constitute only one suggested possible application of the results developed in sections 3 and 4, and it may prove to be more efficient to treat some control problems of this type by adopting standare numerical techniques (e.g., gradient methods, penalty function techniques). One point in favor of the above methods is that many of the computations are independent of the length of the interval $[t_0, t_1]$. Other methods may depend on numerical integration of complicated delay equations and their adjoints which may introduce large errors in approximate controls if the time intervals involved are long.

Acknowlegiment

The authors would like to thank Professors M. Jacobs, F. Kappel, and J. Hale for stimulating comments and observations during various stages of the research reported above.

References

1. C. H. Anderson, The linear differential-difference equation with constant coefficients, *J. Math. Anal. Appl.* **40** (1972), 122–130.
2. H. T. Banks and G. A. Kent, Control of functional differential equations to target sets in function space, *SIAM J. Control* **10** (1972), 567–593.
3. H. T. Banks and M. Q. Jacobs, An attainable sets approach to optimal control of functional differential equations with function space terminal conditions, *J. Differential Equations* **13** (1973), 127–149.
4. H. T. Banks, M. Q. Jacobs and C. E. Langenhop, Characterization of the controlled states in $W_2^{(1)}$ of linear hereditary systems, *SIAM J. Control* **12** (1974), 611–649.
5. R. Bellman and K. L. Cooke, "Differential-Difference Equations," Academic Press, New York, 1963.
6. R. H. Cameron and W. T. Martin, An unsymmetric Fubini theorem, *Bull. Amer. Math. Soc.* **47** (1941), 121–125.
7. M. D. Canon, C. D. Cullum, and E. Polak, "Theory of Optimal Control and Mathematical Programming," McGraw-Hill, New York, 1970.
8. J. W. Daniel, "The Approximate Minimization of Functionals," Prentice-Hall, Englewood Cliffs, NJ, 1971.
9. N. Dunford and J. T. Schwartz, "Linear Operators, Part I," Interscience, New York, 1958.

10. R. GABASOV AND F. KIRILLOVA, "Qualitative Theory of Optimal Processes," Nauka Press, Moscow, 1971.
11. J. K. HALE, Linear functional differential equations with constant coefficients, *Contributions to Differential Equations* 2 (1963), 291–319.
12. J. K. HALE, "Functional Differential Equations," Springer-Verlag, New York, 1971.
13. J. K. HALE, Forward and backward continuation for neutral functional differential equations, *J. Differential Equations* 9 (1971), 168–181.
14. S. P. HASTINGS, Backward existence and uniqueness for retarded functional differential equations, *J. Differential Equations* 5 (1969), 441–451.
15. D. HENRY, Small solutions of linear autonomous functional differential equations, *J. Differential Equations* 8 (1970), 494–501.
16. M. Q. JACOBS AND T. J. KAO, An optimum settling problem for time-lag systems, *J. Math. Anal. Appl.* 40 (1972), 1–21.
16a. I. LASIECKA, On the application of the method of moments to optimal control of systems with delays, M.Sc. thesis, Institute of Automatics, Polytechnic University of Warsaw, 1971.
17. J. L. LIONS, "Optimal Control of Systems Governed by Partial Differential Equations," Springer-Verlag, New York, 1971.
18. D. G. LUENBERGER, "Optimization by Vector Space Methods," Wiley, New York, 1969.
19. E. M. MARKUSHIN, On the convergence of a series of exponentials, Ural. Gos. Univ., *Mat. Zap.* 5 (tetrad 2) (1965), 71–73.
20. E. M. MARKUSHIN AND S. N. SHIMANOV, The convergence of the optimal control of a countable system of differential equations, *Differentsial'nye Uravneniya* 2 (1966), 314–323.
21. E. M. MARKUSHIN AND S. N. SHIMANOV, Approximate solution of the problem of analytical design of a controller for a time-lag equation, *Differentsial'nye Uravneniya* 2 (1966), 1018–1026.
22. E. M. MARKUSHIN AND S. N. SHIMANOV, Approximate solution of analytical regulator problem for systems with delay, *Automation and Remote Control* 29 (1968), 367–374.
23. A. MOSTOWSKI AND M. STARK, "Introduction to Higher Algebra," Pergamon Press, Oxford, 1964.
24. L. W. NEUSTADT, Optimization, a moment problem, and nonlinear programming, *Siam J. Control* 2 (1964), 33–53.
25. H. R. PITT, On a class of linear integro-differential equations, *Cambridge Phil. Soc. Proc.* 43 (1947), 153–163.
26. B. T. POLJAK, Existence theorems and convergence of minimizing sequences in extremum problems with restrictions, *Dokl. Akad. Nauk SSSR* 166 (1966), 287–290.
27. H. L. ROYDEN, "Real Analysis," Second edition, Macmillan, New York, 1968.
28. S. N. SHIMANOV, On the theory of linear differential equations with retardations, *Differentsial'nye Uravneniya* 1 (1965), 102–116.
29. J. M. SWIGER, Application of the theory of minimum normed operators to optimum control system problems, *Advances in Control Systems* 3 (1966), 151–218.
30. A. E. TAYLOR, "Introduction to Functional Analysis," Wiley, New York, 1958.