

AD-A142 963

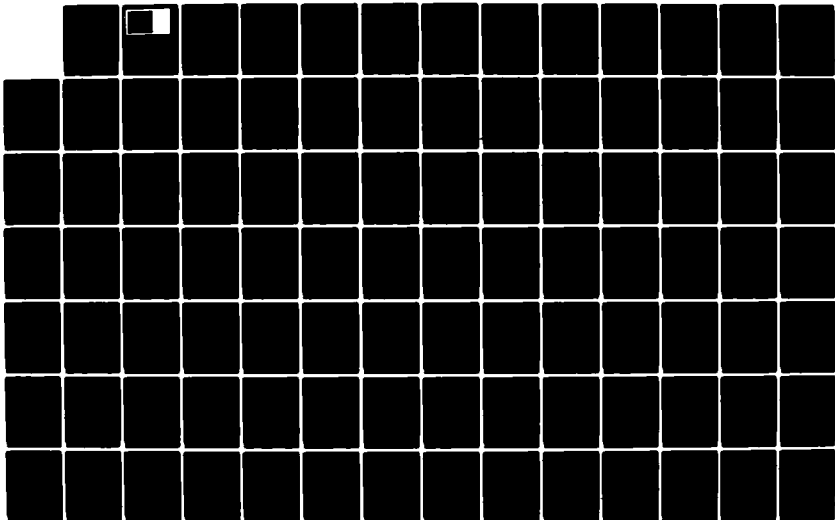
SPLINE APPROXIMATION FOR RETARDED SYSTEMS AND THE  
RICCATI EQUATION(U) WISCONSIN UNIV-MADISON MATHEMATICS  
RESEARCH CENTER F KAPPEL ET AL. APR 84 MRC-TSR-2680  
DAAG29-80-C-0041

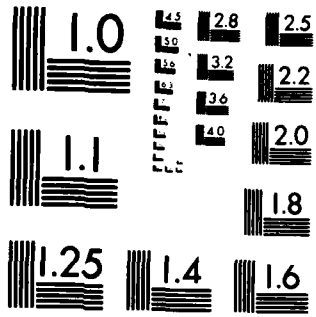
1/2

UNCLASSIFIED

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

2

AD-A142 963

MRC Technical Summary Report # 2680

SPLINE APPROXIMATION FOR  
RETARDED SYSTEMS AND THE  
RICCATI EQUATION

F. Kappel and D. Salamon

**Mathematics Research Center  
University of Wisconsin—Madison  
610 Walnut Street  
Madison, Wisconsin 53705**

April 1984

(Received April 3, 1984)

DTIC  
ELECTE  
JUL 1 0 1984  
A

**Approved for public release  
Distribution unlimited**

Sponsored by

U. S. Army Research Office  
P. O. Box 12211  
Research Triangle Park  
North Carolina 27709

National Science Foundation  
Washington, DC 20550

84 06 29 013

DTIC FILE COPY

UNIVERSITY OF WISCONSIN - MADISON  
MATHEMATICS RESEARCH CENTER

SPLINE APPROXIMATION FOR RETARDED  
SYSTEMS AND THE RICCATI EQUATION

F. Kappel<sup>1,2</sup> and D. Salamon<sup>2,3</sup>

Technical Summary Report # 2680

April 1984

ABSTRACT

The purpose of this paper is to introduce a new spline approximation scheme for retarded functional differential equations. The special feature of this approximation scheme is that it preserves the product space structure of retarded systems and approximates the adjoint semigroup in a strong sense. These facts guarantee the convergence of the solution operators to the differential Riccati equation in a strong sense. Numerical findings indicate a significant improvement in the convergence behaviour over both the averaging and the previous spline approximation scheme.

Furthermore, controllability and observability criteria are given for the approximating systems, which are shown to be stable respectively stabilizable for sufficiently large  $N$  provided that the underlying retarded system has the same property.

AMS (MOS) Subject Classifications: 34K35, 41A15, 93D15

Key Words: Retarded functional differential equations, Approximation, Splines, Riccati Equation.

Work Unit Number 3 - Numerical Analysis and Scientific Computing

---

<sup>1</sup>Institute for Mathematics, University of Graz, Elisabethstrasse 16, A-8010 Graz (Austria).

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

<sup>1</sup>This work was started while this author held a visiting professorship at the Forschungsschwerpunkt Dynamische Systeme, University of Bremen, in October 1982. He acknowledges financial support by the University of Bremen and the hospitality provided by the members of the Forschungsschwerpunkt.

<sup>2</sup>The authors acknowledge partial support by the Fonds zur Förderung der wissenschaftlichen Forschung (Austria) under Project No. P4534.

<sup>3</sup>This material is based upon work supported by the National Science Foundation under Grant No. MCS-8210950.

## SIGNIFICANCE AND EXPLANATION

For a large number of problems in engineering and biology appropriate mathematical models involve functional differential equations (FDE) which in turn can be reformulated as ordinary differential equations (ODE) in infinite dimensional state spaces. This paper is concerned with the approximation of these equations by a sequence of finite dimensional ODEs. A particular emphasis is placed on the approximate solution of the linear quadratic optimal control problem. The approximation scheme is based upon a projection of the underlying function space onto a spline subspace. The special feature of this scheme is that it preserves to product space structure of the FDE and approximates the optimal feedback law on the strong operator topology. A number of numerical examples indicate a significant improvement in the convergence behaviour over previously developed approximation schemes for FDEs.



A1

---

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

## C O N T E N T S

1.	INTRODUCTION	1
2.	STATE SPACE THEORY FOR LINEAR HEREDITARY CONTROL SYSTEMS	5
2.1	Linear hereditary control systems	5
2.2	Semigroups and state space description	7
2.3	The dual state concept	8
2.4	Stability, stabilizability and controllability	11
3.	THE LINEAR QUADRATIC CONTROL PROBLEM	14
3.1	Control systems in Hilbert space	14
3.2	Applications to hereditary systems	18
4.	APPROXIMATION	23
4.1	A general scheme	23
4.2	Approximation of the feedback law	31
4.3	A special spline scheme	37
5.	STRUCTURE AND STABILITY	61
5.1	The structural operator	61
5.2	Criteria for stability, stabilizability and controllability	64
5.3	Convergence of $\Delta^N(\lambda)$	72
5.4	Uniform bounds and nonuniform stability	79
6.	NUMERICAL RESULTS	85
6.1	Examples with finite final time	85
6.2	Examples for the infinite time horizon problem	91
	REFERENCES	102

SPLINE APPROXIMATION FOR RETARDED  
SYSTEMS AND THE RICCATI EQUATION

F. Kappel<sup>1,2</sup> and D. Salamon<sup>2,3</sup>

1. Introduction

In this paper we introduce a new spline approximation scheme for linear time invariant retarded functional differential equations (RFDEs) and establish a number of convergence results and structural properties for this scheme. In particular we show that the approximate feedback law and the solution of the operator Riccati equation, associated with the linear quadratic control problem for this class of systems, converge in the uniform operator topology.

The first step of the general approach is to transform the RFDE

$$\dot{x}(t) = Lx_t + B_0 u(t), \quad y(t) = C_0 x(t) \quad (1.1)$$

into an abstract Cauchy problem of the form

$$\frac{d}{dt} x(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t) \quad (1.2)$$

in the Hilbert space  $X = \mathbb{R}^n \times L^2[-h, 0; \mathbb{R}^n]$ ,  $h > 0$ , where  $A$  is the infinitesimal generator of the strongly continuous semigroup  $S(t)$  which is associated with the uncontrolled delay equation. For systems of the form (1.2) there exists a general theory of the linear quadratic control problem of minimizing the cost functional

---

<sup>1</sup>Institute for Mathematics, University of Graz, Elisabethstrasse 16, A-8010 Graz (Austria).

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

<sup>1</sup>This work was started while this author held a visiting professorship at the Forschungsschwerpunkt Dynamische Systeme, University of Bremen, in October 1982. He acknowledges financial support by the University of Bremen and the hospitality provided by the members of the Forschungsschwerpunkt.

<sup>2</sup>The authors acknowledge partial support by the Fonds zur Förderung der wissenschaftlichen Forschung (Austria) under Project No. P4534.

<sup>3</sup>This material is based upon work supported by the National Science Foundation under Grant No. MCS-8210950.

$$J(u) = \int_0^T (|y(t)|^2 + |u(t)|^2) dt \quad (1.3)$$

(see e.g. [9], [15], [20]). The optimal control can be characterized as a feedback law which is determined by an operator satisfying the differential Riccati equation (in the case  $T < \infty$ ) respectively the algebraic Riccati equation (in the case  $T = \infty$ ). These operator Riccati equations involve both the original generator  $A$  and its adjoint operator  $A^*$ . Therefore, in order to approximate the feedback law and the Riccati operator in the strong operator topology, we have to approximate both semigroups  $S(t)$  and  $S^*(t)$  in the strong operator topology (see [21]).

For the approximation of the semigroups we use a Galerkin type scheme, i.e. we define finite dimensional subspaces  $X^N$  of  $X$  and operators  $A^N$  on  $X^N$  which generate semigroups  $S^N(t)$  on  $X^N$ . The classical idea is to choose  $X^N \subset \text{dom } A$  and define  $A^N = p^N A p^N$  where  $p^N$  is the orthogonal projection of  $X$  onto  $X^N$ . Under appropriate consistency and stability hypotheses the convergence of  $S^N(t)$  to  $S(t)$  in the strong operator topology follows.

These ideas have been used by Banks and Kappel [7] for the development of a spline approximation scheme for RFDEs and have then been applied to problems of optimal control and parameter identification e.g. in [3], [6], [8], [26]. In particular, Kunisch [26] has established weak convergence results for the solution operators of the differential Riccati equations. Numerical findings in [8] indicate that these operators indeed do not converge strongly for the spline scheme developed in [7]. Furthermore, it can actually be shown that the adjoint semigroups  $S^N(t)^*$  cannot converge strongly in that scheme. The main reason for this is that the subspace  $X^N$  in [7] has been chosen to be contained in the domain of  $A$  which is different from the domain of  $A^*$ .

In order to overcome this unequal treatment of  $S(t)$  and  $S^*(t)$ , our first idea was to introduce two spline subspaces  $X^N \subset \text{dom } A$  and  $Y^N \subset \text{dom } A^*$  and to make use of the non orthogonal projections of  $X$  onto  $X^N$  along  $(Y^N)^\perp$ . Unfortunately, we found out after lots of



calculations that these projection operators did not converge strongly to the identity. The second (successful) idea was then to enlarge the subspace  $X^N$  such that it is neither contained in  $\text{dom } A$  nor in  $\text{dom } A^*$ , but contains sufficiently many elements of both domains. Of course, in this situation the approximating operators can no longer be defined by  $A^N = p^N A p^N$  but have to be defined directly instead (for details see Section 4.3). As a result we are able to establish the desired convergence of the solution operators of the Riccati equation in the uniform operator topology for the finite time horizon problem. Despite the fact that in the case of the infinite time horizon problem our scheme always did converge numerically, we were not able to prove this convergence following the approach presented in [21]. The reason is that we do not have the uniform (with respect to  $N$ ) exponential stability of the approximating semigroups for our scheme (compare Section 5.4). In this respect the spline approximation scheme differs from the averaging approximation scheme in [4] for which the uniform exponential stability property has been established in [37].

In two preliminary sections we collect some basic facts from the state space and control theory of retarded systems (Section 2) and give a short survey on the theory of the linear quadratic optimal control problem for abstract systems in Hilbert space and for RFDEs (Section 3). In Section 4.1 we present a general approximation scheme for abstract Cauchy problems in Banach space. In Section 4.2 we consider the problem of approximating the feedback law for the finite time horizon problem following the approach given in Gibson in [21]. The main part of this paper is Section 4.3 where we develop a special spline scheme and prove convergence results along the general ideas given in Section 4.1 and 4.2. We also give the explicit formulae for the matrices which are necessary for the implementation of our scheme. This scheme has remarkable qualitative properties which are presented in Section 5. First of all, the product space structure of the underlying RFDE (1.1) is preserved and there is a structural operator playing an important role for the approximating systems (Section 5.1).

Secondly, there exist convenient criteria for stability, controllability, observability, stabilizability and detectability of the approximating systems (Section 5.2). The main results of Section 5 are that the stability, stabilizability or detectability of the delay system imply the same properties for the approximating systems provided that  $N$  is sufficiently large and that the approximating systems cannot be stable in a uniform sense with respect to  $N$  (Section 5.4).

Finally, in Section 6 we present some of the many numerical calculations in order to demonstrate the good behaviour of our scheme and the significant improvement in the convergence property over both the averaging approximation scheme [4], [21] and the spline scheme in [7], [8].

## 2. State space theory for linear hereditary control systems

In the following we define the type of hereditary control systems to be considered in this paper (Section 2.1) and collect some well known facts on the state space description of retarded functional differential equations (RFDEs) in terms of semigroups and evolution equations (Section 2.2). Then we outline the basic duality relations (Section 2.3) and briefly review some of the existing results on the structural and control properties of hereditary control systems (Section 2.4).

### 2.1. Linear hereditary control systems

We consider the linear hereditary control system

$$\dot{x}(t) = Lx_t + B_0 u(t), \quad t \geq 0, \quad (2.1;1)$$

$$y(t) = C_0 x(t), \quad (2.1;2)$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^k$ ,  $y(t) \in \mathbb{R}^m$  and  $x_t$  is defined by  $x_t(s) = x(t+s)$  for  $-h \leq s \leq 0$ ,  $h > 0$ . Correspondingly  $B_0$  and  $C_0$  are real matrices of appropriate dimensions and  $L$  is a bounded linear functional  $C(-h, 0; \mathbb{R}^n) \rightarrow \mathbb{R}^n$  given by

$$\begin{aligned} L\phi &= \int_{-h}^0 [dn(\tau)]\phi(\tau) \\ &= \sum_{j=0}^p A_j \phi(-h_j) + \int_{-h}^0 A_{01}(\tau)\phi(\tau)d\tau, \quad \phi \in C(-h, 0; \mathbb{R}^n), \end{aligned}$$

where  $0 = h_0 < \dots < h_p = h$  and  $A_j \in \mathbb{R}^{n \times n}$ ,  $j = 0, \dots, p$ , as well as  $A_{01}(\cdot) \in L^2(-h, 0; \mathbb{R}^{n \times n})$ . Clearly, the function  $n: \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$  of bounded variation is of the form

$$\begin{aligned} n(\tau) &= -A_0 x(-\infty, 0)(\tau) - \sum_{j=1}^p A_j x(-\infty, -h_j)(\tau) \\ &\quad - \int_{\tau}^0 A_{01}(\sigma)d\sigma, \quad \tau \in \mathbb{R}, \end{aligned}$$

where  $\chi_I$  denotes the characteristic function of the interval  $I$ . A solution of (2.1;1) is a function  $x(\cdot) \in L^2_{loc}(-h, \infty; \mathbb{R}^n)$  which is absolutely continuous with  $L^2$ -derivative on every compact interval  $[0, T]$ ,  $T > 0$ , and satisfies (2.1;1) for almost all  $t \geq 0$ . It is well known that (2.1;1) admits a unique solution  $x(t) = x(t; \phi, u)$  for every input  $u(\cdot) \in L^2_{loc}(0, \infty; \mathbb{R}^l)$  and every initial condition

$$x(0) = \phi^0, \quad x(\tau) = \phi^1(\tau), \quad -h \leq \tau < 0, \quad (2.2)$$

where  $\phi = (\phi^0, \phi^1) \in M^2 = \mathbb{R}^n \times L^2(-h, 0; \mathbb{R}^n)$ . Moreover,  $x(\cdot; \phi, u)$  depends continuously on  $\phi$  and  $u$  on compact intervals, i.e. for any  $T > 0$  there exists a  $K > 0$  such that

$$\sup_{0 \leq t \leq T} |x(t; \phi, u)| \leq K(\|\phi\| + \|u\|_{L^2(0, T; \mathbb{R}^l)}),$$

where  $\|\phi\| = (|\phi^0|^2 + \|\phi^1\|_{L^2}^2)^{1/2}$  for  $\phi \in M^2$  (see e.g. [13], [19]).

The fundamental solution of (2.1;1) will be denoted by  $X(t)$  and is the  $n \times n$  matrix valued solution of (2.1;1) which corresponds to  $u \equiv 0$  and  $X(0) = I$ ,  $X(\tau) = 0$  for  $-h \leq \tau < 0$ . The Laplace-transform of  $X(\cdot)$  is given by  $\Delta^{-1}(\lambda)$ , where

$$\begin{aligned} \Delta(\lambda) &= \lambda I - L(e^{\lambda \cdot} I) \\ &= \lambda I - \sum_{j=0}^p A_j e^{-\lambda h j} - \int_{-h}^0 A_{01}(\tau) e^{\lambda \tau} d\tau, \quad \lambda \in \Phi, \end{aligned}$$

is the characteristic matrix of (2.1;1). Again it is well known that the forced motion of (2.1;1) (in case  $\phi = 0$ ) can be written as

$$x(t; 0, u) = \int_0^t X(t-\tau) B_0 u(\tau) d\tau, \quad t \geq 0. \quad (2.3)$$

## 2.2. Semigroups and state space description

In the theory of RFDEs two state space concepts are of importance which are actually dual to each other. Existence, uniqueness and continuous dependence results for solutions of RFDEs have motivated the "classical" definition of the state of system (2.1) to be the pair

$$w(t) = (x(t), x_t) \in M^2, \quad (2.4)$$

which completely describes the past history of the solution at time  $t \geq 0$ . The evolution of this state is governed by the variation-of-constants formula

$$w(t) = S(t)\phi + \int_0^t S(t-s)Bu(s)ds, \quad t \geq 0, \quad (2.5)$$

which is the infinite dimensional version of (2.3). The input operator  $B: \mathbb{R}^l \rightarrow M^2$  is given by

$$Bu = (B_0 u, 0) \in M^2, \quad u \in \mathbb{R}^l,$$

and the semigroup  $S(\cdot)$  corresponds to the free motion of the system, i.e.  $S(t): M^2 \rightarrow M^2$ ,  $t \geq 0$ , is defined by

$$S(t)\phi = (x(t; \phi, 0), x_t(\phi, 0)), \quad t \geq 0, \quad \phi \in M^2.$$

The infinitesimal generator of  $S(\cdot)$  is given by

$$\begin{aligned} \text{dom } A &= \{\phi \in M^2 \mid \phi^1 \in W^{1,2}, \phi^0 = \phi^1(0)\}, \\ A\phi &:= (L\phi^1, \dot{\phi}^1), \end{aligned} \quad (2.6)$$

where  $W^{1,2}$  denotes the Sobolev space  $W^{1,2}(-h, 0; \mathbb{R}^n)$ . The function  $w(t)$  as defined in (2.5) is a mild solution of the abstract system

$$\dot{w}(t) = Aw(t) + Bu(t), \quad (\Sigma)$$

$$y(t) = Cw(t), \quad w(0) = \phi.$$

The output operator  $C: M^2 \rightarrow \mathbb{R}^m$  is defined by  $C\phi = C_0\phi^0$ ,  $\phi \in M^2$ .

Analogously we introduce the semigroup  $S_T(\cdot)$  on  $M^2$  which corresponds to the transposed system

$$\dot{x}(t) = L^T x_t + C_0^T y(t), \quad (2.7;1)$$

$$u(t) = B_0^T x(t). \quad (2.7;2)$$

The infinitesimal generator of  $S_T(\cdot)$  is denoted by  $A_T$ . Corresponding to system (2.7) we have the abstract system

$$\dot{w}(t) = A_T w(t) + C^* y(t) \quad (\Sigma_T)$$

$$u(t) = B^* w(t).$$

Note, that  $S_T(\cdot)$  is not the adjoint semigroup to  $S(\cdot)$ . The duality relation between systems (2.1) and (2.7) involves another state concept which is due to Miller [31].

### 2.3. The dual state concept

Another state concept for system (2.1) - again in the state space  $M^2$  - can be obtained by viewing the  $L^2$ -component of the state introduced in the previous section as an additional forcing term instead as an initial function. To this end we rewrite system (2.1) in the following way

$$\dot{x}(t) = \int_{-t}^0 [dn(\tau)]x(t+\tau) + B_0 u(t) + f^1(-t), \quad (2.8;1)$$

$$x(0) = f^0,$$

$$y(t) = C_0 x(t), \quad (2.8;2)$$

where the pair  $f = (f^0, f^1) \in M^2$  is given by

$$\begin{aligned} f^0 &= \phi^0, \\ f^1(\sigma) &= \int_{-h}^{\sigma} [d\eta(\tau)] \phi^1(\tau - \sigma) \\ &= \sum_{j=1}^p A_j \phi^1(-h_j - \sigma) + \int_{-h}^{\sigma} A_{01}(\tau) \phi^1(\tau - \sigma) d\tau \end{aligned} \quad (2.9)$$

(we define  $\phi^1(\tau) = 0$  for  $\tau \notin [-h, 0]$ ). Now the initial state of system (2.8) is given by  $f \in M^2$ . Correspondingly the state at time  $t \geq 0$  is the pair

$$z(t) = (x(t), x^t) \in M^2, \quad (2.10)$$

where  $x^t \in L^2(-h, 0; \mathbb{R}^n)$  is defined by

$$\begin{aligned} x^t(\sigma) &= \int_{\sigma-t}^{\sigma} [d\eta(\tau)] x(t+\tau-\sigma) + f^1(\sigma-t) \\ &= \sum_{j=1}^p A_j x_t(-h_j - \sigma) + \int_{\sigma-t}^{\sigma} A_{01}(\tau) x(t+\tau-\sigma) d\tau + f^1(\sigma-t). \end{aligned} \quad (2.11)$$

Here for any solution  $x(t)$  of (2.8;1) we define  $x_t(\tau) = 0$  if  $\tau > 0$  or  $\tau < -t$ . The state  $z(t)$  determines the future behavior of the solution from  $t \geq 0$  on, i.e.  $x(t+s) = 0$  for  $s \geq 0$  if and only if  $z(t) = 0$ .

The evolution of  $z(t)$  is governed by the following variation-of-constants formula

$$z(t) = S_T^*(t)f + \int_0^t S_T^*(t-s)Bu(s)ds, \quad t \geq 0, \quad (2.12)$$

which means that  $z(t)$  is a mild solution of the system

$$\begin{aligned} \dot{z}(t) &= A_T^* z(t) + Bu(t), \quad z(0) = f, \\ y(t) &= Cz(t) \end{aligned} \quad (E_T^*)$$

([10],[14],[18]). Note that  $(\Sigma_T^*)$  is precisely the adjoint system to  $(\Sigma_T)$  which corresponds to the transposed system (2.7) in terms of the original state concept. The operator  $A_T^*$  is the infinitesimal generator of the semigroup  $S_T^*(\cdot)$  and can be described explicitly in the following way (see e.g. [18]):

Lemma 2.1. The operator  $A_T^*$  is given by

$$\begin{aligned} \text{dom } A_T^* &= \{f \in M^2 \mid f^1 + \sum_{j=1}^{p-1} A_j f^0 \chi_{[-h, -h_j]} \in W^{1,2}, f^1(-h) = A_p f^0\}, \\ [A_T^* f]^0 &= f^1(0) + A_0 f^0, \\ [A_T^* f]^1(\tau) &= A_{01}(\tau) f^0 - \frac{d}{d\tau} [f^1(\tau) + \sum_{j=1}^{p-1} A_j f^0 \chi_{[-h, -h_j]}(\tau)]. \end{aligned}$$

The relation between the two state concepts can be described by the so-called structural operator

$$F: M^2 \rightarrow M^2$$

which maps every initial state  $\phi \in M^2$  of system  $(\Sigma)$  to the corresponding initial state

$$F\phi = f \in M^2$$

of system  $(\Sigma_T^*)$  which is given by (2.9). This operator has been introduced by Bernier and Manitius [10]. The adjoint operator  $F^*$  is of the same form as  $F$  but with the transposed matrices. This means that  $F^*$  plays the same role for the description of the transposed system (2.7) as  $F$  does for the original system (2.1).

The operator  $F$  has the following important properties [10], [18]:



Theorem 2.2.

- (i)  $FS(t) = S_T^*(t)F, \quad t \geq 0.$   
(ii) If  $\phi \in \text{dom } A$ , then  $F\phi \in \text{dom } A_T^*$  and

$$A_T^*F\phi = FA\phi.$$

- (iii)  $FB = B$  and  $CF = C.$

2.4. Stability, stabilizability and controllability

System (2.1) is said to be *stable* if every solution  $x(t)$  of the free system (i.e.  $u(t) \equiv 0$ ) tends to zero as  $t$  goes to infinity. Equivalently, the semigroup  $S(\cdot)$  is exponentially stable, i.e.

$$\omega_0 = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \|S(t)\| = \sup \{ \text{Re } \lambda \mid \lambda \in \sigma(A) \} < 0$$

(see for instance [23]). The spectrum of  $A$  is given by  $\sigma(A) = \{ \lambda \in \mathbb{C} \mid \det \Delta(\lambda) = 0 \}$ . Note, that  $\sigma(A_T^*) = \sigma(A)$ . Clearly, the stability of system (2.1) is equivalent to the stability of the transposed system (2.7) and to the stability of system (2.8).

The control system (2.1) is said to be *stabilizable* if there exists a control law

$$\begin{aligned} u(t) &= K(x(t), x_t) \\ &= K_0 x(t) + \int_{-h}^0 K_1(\tau) x(t+\tau) d\tau, \end{aligned} \tag{2.13}$$

where  $K_0 \in \mathbb{R}^{l \times n}$ ,  $K_1(\cdot) \in L^2(-h, 0; \mathbb{R}^{l \times n})$ , such that the closed loop system (2.1), (2.13) is stable. We have the following important characterization (see [33], [35]).

Theorem 2.3. The following statements are equivalent:

- (i) System (2.1) is stabilizable.
- (ii) There exists a  $K \in L(M^2, \mathbb{R}^l)$  such that the operator  $A + BK$  generates an exponentially stable  $C_0$ -semigroup.
- (iii) There exists a  $K_T^* \in L(M^2, \mathbb{R}^l)$  such that the operator  $A_T^* + BK_T^*$  generates an exponentially stable semigroup.
- (iv)  $\text{rank } [\Delta(\lambda), B_0] = n$  for all  $\lambda \in \mathbb{C}$  with  $\text{Re } \lambda \geq 0$ .

The dual result is the following (see e.g. [10] or [35],[36]):

Theorem 2.4. The following statements are equivalent:

- (i) There exists a  $H \in L(\mathbb{R}^m, M^2)$  such that the operator  $A + HC$  generates an exponentially stable semigroup.
- (ii) There exists a  $H_T^* \in L(\mathbb{R}^m, M^2)$  such that the operator  $A_T^* + H_T^*C$  generates an exponentially stable semigroup.
- (iii)  $\text{rank } \begin{pmatrix} \Delta(\lambda) \\ C_0 \end{pmatrix} = n$  for all  $\lambda \in \mathbb{C}$  with  $\text{Re } \lambda \geq 0$ .

System (2.1) is called *detectable* if the statements of the previous theorem are satisfied. A detailed discussion of the perturbed semigroups and the duality relations between feedback stabilization and dynamic observation in the product space framework can be found in [35].

System (2.1) is called *approximately controllable* if the reachable subspace

$$R = \left\{ \int_0^t S(t-s)Bu(s)ds \mid t \geq 0, u(\cdot) \in L^2(0,t;\mathbb{R}^l) \right\}$$

is dense in  $M^2$ ; it is called *strictly observable* if for all solutions of (2.1)  $y(t) = 0, t \geq 0$ , implies  $x(t) = 0, t \geq -h$ .

These two properties have been characterized by Manitius [27],[28] as follows.

Theorem 2.5. Let  $A_{01}(\tau) \equiv 0$ . Then system (2.1) is approximately controllable if and only if the following conditions are satisfied:

$$\text{rank } [\Delta(\lambda), B_0] = n \text{ for all } \lambda \in \mathbb{C},$$

$$\text{rank } [A_p, B_0] = n.$$

System (2.1) is strictly observable if and only if the following conditions are satisfied:

$$\text{rank } \begin{pmatrix} \Delta(\lambda) \\ C_0 \end{pmatrix} = n \text{ for all } \lambda \in \mathbb{C},$$

$$\text{rank } A_p = n.$$

### 3. The linear quadratic control problem

#### 3.1. Control systems in Hilbert spaces

Let us first deal with general linear control systems in Hilbert spaces  $X$ ,  $U$  and  $V$  described by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0, \\ y(t) &= Cx(t). \end{aligned} \tag{3.1}$$

We assume that  $B \in L(U, X)$ ,  $C \in L(X, V)$  and that  $A$  is the infinitesimal generator of a  $C_0$ -semigroup  $S(t)$  on  $X$ . System (3.1) will be understood in the sense of mild solutions, i.e. the trajectories of the system are given by

$$x(t) = S(t)x_0 + \int_0^t S(t-s)Bu(s)ds, \quad t \geq 0, \tag{3.2}$$

for any  $x_0 \in X$  and any input  $u(\cdot) \in L^2_{loc}(0, \infty; U)$ .

Let  $R: U \rightarrow U$  and  $G: X \rightarrow X$  be selfadjoint linear operators satisfying

$$\langle x, Gx \rangle \geq 0 \quad \text{for all } x \in X$$

and

$$\langle u, Ru \rangle \geq \epsilon \|u\|^2 \quad \text{for all } u \in U$$

with some  $\epsilon > 0$ . In this section we look at the control problem of minimizing the cost functional

$$\begin{aligned} J(u) &= \langle x(T), Gx(T) \rangle + \\ &+ \int_0^T [ \|Cx(t)\|^2 + \langle u(t), Ru(t) \rangle ] dt, \end{aligned} \tag{3.3}$$

where  $x(t)$  is given by (3.2) and  $T > 0$  is a fixed final time. The following result has been proved in [15]:

Theorem 3.1. For any  $x_0 \in X$  there exists a unique control function  $\bar{u}(\cdot) \in L^2(0, T; U)$  which minimizes the cost functional (3.3) under the constraint (3.2). The optimal control is of feedback form and is given by

$$\bar{u}(t) = -R^{-1}B^*P(t)\bar{x}(t), \quad t \geq 0 \quad (3.4)$$

where  $\bar{x}(t)$  is the mild solution of the Cauchy problem  $\dot{\bar{x}} = (A - BR^{-1}B^*P(t))\bar{x}$ ,  $\bar{x}(0) = x_0$ , and  $t \rightarrow P(t) \in L(X)$  is the unique operator valued function on  $[0, T]$  with the following properties:

- (i)  $P(t)$  is positive semidefinite for every  $t \in [0, T]$ .
- (ii) The function  $t \rightarrow P(t)x$  is continuous on  $[0, T]$  for every  $x \in X$ .
- (iii) The function  $t \rightarrow \langle x, P(t)y \rangle$  is continuously differentiable on  $[0, T]$  for all  $x, y \in \text{dom } A$  and satisfies the Riccati differential equation

$$\begin{aligned} \frac{d}{dt} \langle y, P(t)x \rangle + \langle Ay, P(t)x \rangle + \langle P(t)y, Ax \rangle \\ - \langle y, P(t)BR^{-1}B^*P(t)x \rangle + \langle Cy, Cx \rangle = 0, \end{aligned} \quad (3.5;1)$$

$$\langle y, P(T)x \rangle = \langle y, Gx \rangle. \quad (3.5;2)$$

Moreover, the optimal cost is given by

$$J(\bar{u}) = \langle x_0, P(0)x_0 \rangle.$$

It is easy to see (cf. [20]) that equations (3.5) can be written in the form

$$\begin{aligned} P(t)x &= S^*(T-t)GS(T-t)x \\ &+ \int_t^T S^*(\tau-t)[C^*C - P(\tau)BR^{-1}B^*P(\tau)]S(\tau-t)x d\tau, \end{aligned} \quad (3.6)$$

$0 \leq t \leq T, x \in X$ , or respectively,

$$P(t)x = S^*(T-t)G\phi(T,t)x + \int_t^T S^*(\tau-t)C^*C\phi(\tau,t)x d\tau, \quad (3.7)$$

$0 \leq t \leq T, x \in X$ , where  $\phi(\tau,t)$  is the evolution operator given by

$$\phi(\tau,t)x = S(\tau-t)x - \int_t^\tau S(\tau-\sigma)BR^{-1}B^*P(\sigma)\phi(\sigma,t)x d\sigma, \quad (3.8)$$

$0 \leq t \leq \tau \leq T, x \in X$ .

Let us now consider the problem of minimizing the cost functional

$$J(u) = \int_0^{\infty} [\|Cx(t)\|^2 + \langle u(t), Ru(t) \rangle] dt \quad (3.9)$$

where again  $x(t)$  is given by (3.2). For this situation the following result has been proved (see [15],[16],[41]; further references can be found in the survey paper [9]):

**Theorem 3.2.** a) The following statements are equivalent:

- (i) For any  $x_0 \in X$  there exists an input  $u(\cdot) \in L^2(0, \infty; U)$  such that the corresponding cost  $J(u)$  given by (3.9) and (3.2) is finite.
- (ii) There exists a positive semidefinite operator  $P \in L(X)$  satisfying the algebraic Riccati operator equation

$$\begin{aligned} \langle Ay, Px \rangle + \langle Py, Ax \rangle + \langle Cy, Cx \rangle \\ - \langle Py, BR^{-1}B^*Px \rangle = 0 \end{aligned} \quad (3.10)$$

for all  $x, y \in \text{dom } A$ .

b) If the statements under a) are valid, then there exists a unique optimal control  $\bar{u}(t)$  which is given by the feedback law

$$\bar{u}(t) = -R^{-1}B^*P\bar{x}(t), \quad t \geq 0, \quad (3.11)$$

where  $\bar{x}(t)$  is the mild solution of the Cauchy problem  $\dot{x} = (A - BR^{-1}B^*P)x$ ,  $x(0) = x_0$ , and  $P$  is the minimal solution of (3.10). Moreover, the optimal cost is given by

$$J(\bar{u}) = \langle x_0, Px_0 \rangle.$$

c) Suppose that the statements under a) are satisfied and let  $P$  be the minimal positive semidefinite solution of (3.10). Moreover, let  $P_T(t)$ ,  $0 \leq t \leq T$ , be the unique positive semidefinite solution of (3.5;1) with  $P_T(T) = 0$ . Then  $P$  is the strong limit of  $P_T(0)$  as  $T$  goes to infinity.

d) Suppose that there exists some  $H \in L(V, X)$  such that the operator  $A + HC$  generates an exponentially stable semigroup. Then there exists at most one positive semidefinite solution of (3.10). Moreover, if such a solution exists, then the closed loop semigroup generated by  $A - BR^{-1}B^*P$  is exponentially stable.

Finally note that any solution  $P$  of the algebraic Riccati equation is a stationary solution of the Riccati differential equation. Hence it follows from (3.6), (3.7) and (3.8) with  $G = P$  that the algebraic Riccati equation (3.10) is equivalent to

$$Px = S^*(t)PS(t)x + \int_0^t S^*(\tau)[C^*C - PBR^{-1}B^*P]S(\tau)x d\tau,$$

$t \geq 0$ ,  $x \in X$ , or, respectively, to

$$Px = S^*(t)PS_p(t)x + \int_0^t S^*(\tau)C^*CS_p(\tau)x d\tau, \quad (3.12)$$

$t \geq 0$ ,  $x \in X$ , where  $S_p(\cdot)$  is the closed loop semigroup generated by  $A - BR^{-1}B^*P$ .

### 3.2. Applications to hereditary systems

Let us first apply Theorem 3.1 to the systems  $(\Sigma)$  and  $(\Sigma_T^*)$ , which are associated to the system (2.1) in terms of the two state concepts introduced in Section 2. The cost functional for system (2.1) is given by

$$J(u) = x(T; \phi, u)^T G_0 x(T; \phi, u) + \int_0^T [ \|C_0 x(t; \phi, u)\|^2 + u(t)^T R u(t) ] dt, \quad (3.13)$$

where  $R \in \mathbb{R}^{\ell \times \ell}$  is positive definite and  $G_0 \in \mathbb{R}^{n \times n}$  is positive semidefinite. The operator  $G: M^2 \rightarrow M^2$  is defined by  $G\phi = (G_0 \phi^0, 0)$ ,  $\phi \in M^2$ . Then the cost functional for systems  $(\Sigma)$  and  $(\Sigma_T^*)$  is given by (3.3) (with  $G = G$ ,  $C = C$  and  $R = R$ , of course). According to Theorem 3.1 there exist two unique, positive semidefinite, strongly continuous families  $\Pi(\cdot)$  and  $P(\cdot)$  of operators in  $L(M^2)$  which satisfy the following Riccati differential equations:

$$\begin{aligned} \frac{d}{dt} \langle \psi, \Pi(t) \phi \rangle + \langle A \psi, \Pi(t) \phi \rangle + \langle \Pi(t) \psi, A \phi \rangle \\ - \langle \Pi(t) \psi, B R^{-1} B^* \Pi(t) \phi \rangle + \langle C \psi, C \phi \rangle = 0, \end{aligned} \quad (3.14)$$

$$\Pi(T) = G,$$

$\phi, \psi \in \text{dom } A$ ,  $0 \leq t \leq T$ , and, respectively,

$$\begin{aligned} \frac{d}{dt} \langle g, P(t) f \rangle + \langle A_T^* g, P(t) f \rangle + \langle P(t) g, A_T^* f \rangle \\ - \langle P(t) g, B P^{-1} B^* P(t) f \rangle + \langle C g, C f \rangle = 0, \end{aligned} \quad (3.15)$$

$$P(T) = G,$$

$f, g \in \text{dom } A_T^*$ ,  $0 \leq t \leq T$ .

The operators  $\Pi(t)$  and  $P(t)$  have the following properties:



Proposition 3.3. a)  $\Pi(t) = F^*P(t)F$ ,  $0 \leq t \leq T$ .

b)  $\text{range } P(t) \subset \text{dom } A_T$  for every  $t \leq T-h$ . If  $G_0 = 0$ , then  $\text{range } P(t) \subset \text{dom } A_T$  for every  $t \in [0, T]$ . Moreover, in this case the function  $t \rightarrow P(t)f$  is continuously differentiable on  $[0, T]$  for every  $f \in \text{dom } A_T^*$  and satisfies

$$\begin{aligned} \frac{d}{dt} P(t)f + A_T P(t)f + P(t)A_T^*f - \\ - P(t)BR^{-1}B^*P(t)f + C^*Cf = 0, \end{aligned} \quad (3.16)$$

$$P(T) = 0.$$

c)  $\text{range } \Pi(t) \subset \text{dom } A^*$  for every  $t \leq T-h$ . If  $G_0 = 0$ , then  $\text{range } \Pi(t) \subset \text{dom } A^*$  for every  $t \in [0, T]$ . Moreover, in this case the function  $t \rightarrow \Pi(t)\phi$  for every  $\phi \in \text{dom } A$  is continuously differentiable on  $[0, T]$  and satisfies

$$\begin{aligned} \frac{d}{dt} \Pi(t)\phi + A^*\Pi(t)\phi - \Pi(t)A\phi - \\ - \Pi(t)BR^{-1}B^*\Pi(t)\phi + C^*C\phi = 0, \end{aligned} \quad (3.17)$$

$$\Pi(T) = 0.$$

Proof. a) has been shown in [17]. However, it also follows immediately from Theorem 2.2 that  $\Pi(t) = F^*P(t)F$  defines a positive semidefinite solution of (3.14) if  $P(t)$  is a positive semidefinite solution of (3.15).

b) follows from (3.7) and (3.8) with  $P(t) = P(t)$ ,  $S(t) = S_T^*(t)$ ,  $B = B$ ,  $C = C$  and  $R = R$ . One has to observe the following facts:

(i)  $\text{range } S_T(t) \subset \text{dom } A_T$  for all  $t \geq h$ .

(ii)  $\int_0^t S_T(s)C^*u(s)ds \in \text{dom } A_T$  for all  $u \in L^2(0, T; \mathbb{R}^l)$ .

(iii) If  $z(s, t) \in M^2$  is continuous on  $\{(s, t) \mid 0 \leq t \leq s \leq T\}$  and continuously differentiable in the second variable, then the function

$$w(t) = \int_t^T S_T(s-t)z(s,t)ds, \quad 0 \leq t \leq T,$$

is continuously differentiable and satisfies

$$\frac{d}{dt} w(t) = -A_T w(t) + \int_t^T S_T(t-s) \frac{d}{dt} z(s,t)ds - z(t,t).$$

c) follows from a), b) and Theorem 2.2  $\square$

Let us now look at the structure of the operators  $\Pi(t)$  and  $P(t)$ . Due to the product space structure of the state space  $M^2$  we can write

$$\Pi(t) = \begin{bmatrix} \Pi_{00}(t) & \Pi_{01}(t) \\ \Pi_{10}(t) & \Pi_{11}(t) \end{bmatrix}, \quad P(t) = \begin{bmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{bmatrix},$$

where  $\Pi_{00}(t)$ ,  $P_{00}(t)$  are selfadjoint operators  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  which can be represented by symmetric matrices and  $\Pi_{11}(t)$ ,  $P_{11}(t)$  are selfadjoint operators  $L^2 \rightarrow L^2$ . The operators  $\Pi_{10}(t)$ ,  $P_{10}(t)$  can be represented by matrix-valued functions  $\Pi_{10}(t, \cdot)$ ,  $P_{10}(t, \cdot) \in L^2(-h, 0; \mathbb{R}^{n \times n})$ . The adjoint operators  $\Pi_{01}(t) = \Pi_{10}^*(t)$  and  $P_{01}(t) = P_{10}^*(t)$  from  $L^2 \rightarrow \mathbb{R}^n$  are given by

$$\begin{aligned} \Pi_{01}(t)\phi &= \int_{-h}^0 \Pi_{10}^T(t, \tau)\phi(\tau)d\tau, & P_{01}(t)\phi &= \int_{-h}^0 P_{10}^T(t, \tau)\phi(\tau)d\tau, \\ \phi &\in L^2. \end{aligned}$$

We are mainly interested in the matrices  $\Pi_{00}(t)$  and  $\Pi_{10}(t, \tau)$  which determine the optimal feedback law

$$\bar{u}(t) = -R^{-1}B_0^T \left[ \Pi_{00}(t)x(t) + \int_{-h}^0 \Pi_{10}^T(t, \tau)x(t+\tau)d\tau \right] \quad (3.18)$$

for system (2.1). Recall that  $B^*$  maps  $\phi \in M^2$  to  $B_0^T \phi^0 \in \mathbb{R}^l$ . By Proposition 3.3, a) we have the following relations between  $\Pi(t)$  and  $P(t)$ :

$$\Pi_{00}(t) = P_{00}(t), \quad (3.19)$$

$$\Pi_{10}(t, \sigma) = \sum_{j=1}^p A_j^T P_{10}(t, -h_j - \sigma) + \int_{-h}^{\sigma} A_{01}^T(\tau) P_{10}(t, \tau - \sigma) d\tau.$$

Hence the control law (3.18) can be written in the form

$$\begin{aligned} \bar{u}(t) = & -R^{-1} B_0^T [P_{00}(t)x(t) + \sum_{j=1}^p \int_{-h_j}^0 P_{10}^T(t, -h_j - \sigma) A_j x(t + \sigma) d\sigma \\ & + \int_{-h}^0 \int_{\tau}^0 P_{10}^T(t, \tau - \sigma) A_{01}(\tau) x(t + \sigma) d\sigma d\tau. \end{aligned} \quad (3.20)$$

Finally, note that  $P_{10}(t, \cdot) \in W^{1,2}(-h, 0; \mathbb{R}^{n \times n})$  and  $P_{00}(t) = P_{10}(t, 0)$  for  $t \leq T-h$  for all  $t \in [0, T]$  provided that  $G_0 = 0$  (Proposition 3.3, b)).

For the rest of this section we assume that system (2.1) is stabilizable and detectable, so that systems  $(\Sigma)$  and  $(\Sigma_T^*)$  satisfy the assumptions of Theorem 3.2. Hence there exist positive semi-definite operators  $\Pi, P \in L(M^2)$  satisfying the algebraic Riccati equations

$$A^* \Pi \phi + \Pi A \phi - \Pi B R^{-1} B^* \Pi \phi + C^* C \phi = 0, \quad (3.21)$$

$\phi \in \text{dom } A$ , and, respectively,

$$A_T^* P f + P A_T^* f - P B R^{-1} B^* P f + C^* C f = 0, \quad (3.22)$$

$f \in \text{dom } A_T^*$ . The equations can be written in this form since every solution  $P$  of (3.10) maps  $\text{dom } A$  into  $\text{dom } A^*$ . The relation between  $\Pi$  and  $P$  is as follows:

**Proposition 3.4.** a)  $\Pi = F^* P F$ .

b)  $\text{range } P \subset \text{dom } A_T$ .

c)  $\text{range } \Pi \subset \text{dom } A^*$ .

Proof. Statement a) follows again from the fact that for any solution P of (3.22) the operator  $\Pi = F^*PF$  defines a solution of (3.21) (see also [40]). Moreover, b) follows from (3.12) (for  $t \geq h$ ) and c) follows from a), b) and Theorem 2.2  $\square$

Again the operators  $\Pi$  and P can be written in block form

$$\Pi = \begin{pmatrix} \Pi_{00} & \Pi_{01} \\ \Pi_{10} & \Pi_{11} \end{pmatrix}, \quad P = \begin{pmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{pmatrix},$$

where  $\Pi_{01} = \Pi_{10}^*$  and  $P_{01} = P_{10}^*$  map  $L^2$  into  $\mathbb{R}^n$ . By Proposition 3.4 we have  $P_{10}(\cdot) \in W^{1,2}(-h, 0; \mathbb{R}^{n \times n})$  and

$$\Pi_{00} = P_{00} = P_{10}(0), \tag{3.23}$$

$$\Pi_{10}(\sigma) = \sum_{j=1}^p A_j^T P_{10}(-h_j - \sigma) + \int_{-h}^{\sigma} A_{01}^T(\tau) P_{10}(\tau - \sigma) d\tau.$$

Hence the optimal feedback law is of the form

$$\begin{aligned} u(t) &= -R^{-1}B^*\Pi(x(t), x_t) = -R^{-1}B^*PF(x(t), x_t) \\ &= -R^{-1}B_0^T \left[ \Pi_{00}x(t) + \int_{-h}^0 \Pi_{10}^T(\tau)x(t+\tau)d\tau \right] \tag{3.24} \\ &= -R^{-1}B_0^T \left[ P_{10}(0)x(t) + \sum_{j=1}^p \int_{-h_j}^0 P_{10}^T(-h_j - \sigma)A_j x(t+\sigma)d\sigma \right. \\ &\quad \left. + \int_{-h}^0 \int_{\tau}^0 P_{10}^T(\tau - \sigma)A_{01}(\tau)x(t+\sigma)d\sigma d\tau \right]. \end{aligned}$$

Finally note that the closed loop system (2.1), (3.24) is stable (Theorem 3.2).

## 4. Approximation

### 4.1. A general scheme

In this section we present a general approximation scheme for linear abstract Cauchy problems restricting ourselves to a situation which is of sufficient generality for our purposes.

Let  $X$  be a real Banach space with norm  $\|\cdot\|$ . Furthermore let  $A$  be the infinitesimal generator of the  $C_0$ -semigroup  $S(t)$ ,  $t \geq 0$ , on  $X$ . It is well known (see for instance [30], p. 278, or [32], p. 100) that  $x(t) = S(t)x_0$  for any  $x_0 \in \text{dom } A$  is the unique strong solution of the abstract Cauchy problem

$$\begin{aligned} \dot{x}(t) &= Ax(t), \quad t \geq 0, \\ x(0) &= x_0. \end{aligned} \tag{4.1}$$

By a strong solution  $x(t) = x(t; x_0)$  of (4.1) we mean a continuously differentiable function  $x: [0, \infty) \rightarrow X$  such that  $x(t) \in \text{dom } A$  for  $t \geq 0$  and (4.1) is satisfied. There exist constants  $M \geq 1$  and  $\omega \in \mathbb{R}$  such that

$$\|S(t)\| \leq Me^{\omega t}, \quad t \geq 0.$$

Our goal is to approximate the trajectories  $S(t)x_0$ ,  $x_0 \in X$ , by sequences of solutions to Cauchy problems in finite dimensional spaces. It is a standard idea to choose a sequence  $\{X^N\}$  of finite dimensional subspaces of  $X$  with corresponding projections

$$p^N: X \rightarrow X^N, \quad N = 1, 2, \dots,$$

and to define (in an appropriate way) a sequence  $\{A^N\}$  of linear operators

$$A^N: X^N \rightarrow X^N, \quad N = 1, 2, \dots$$

With  $A^N$  and  $x_0 \in X$  we associate the Cauchy problem

$$\begin{aligned} \dot{x}^N(t) &= A^N x^N(t), \quad t \geq 0 \\ x^N(0) &= p^N x_0, \end{aligned} \tag{4.2}$$

on  $X^N$ . We extend the definition of  $A^N$  to all of  $X$  by  $A^N x = A^N p^N x$  and define the  $C_0$ -semigroup  $S^N(t)$ ,  $t \geq 0$ , on  $X$  by

$$S^N(t)x_0 = e^{A^N t} x_0 = e^{A^N t} p^N x_0 + x_0 - p^N x_0, \quad t \geq 0, \quad x_0 \in X.$$

The following hypotheses will be used in order to guarantee the desired convergence  $S^N(t)p^N x_0 \rightarrow S(t)x_0$ :

- (H1)  $\lim_{N \rightarrow \infty} p^N x = x$  for all  $x \in X$ .  
 (H2) There exist constants  $\tilde{M} \geq 1$  and  $\tilde{\omega} \in \mathbb{R}$  such that

$$\|S^N(t)x\| \leq \tilde{M} e^{\tilde{\omega} t} \|x\|$$

for all  $t \geq 0$ ,  $x \in X^N$  and  $N = 1, 2, \dots$ .

- (H3) There exists a dense subset  $D \subset \text{dom } A$  which is invariant with respect to  $S(t)$ ,  $t \geq 0$ , such that

(i)  $\lim_{N \rightarrow \infty} A^N p^N x = Ax$  for all  $x \in D$

and

- (ii) for any  $x \in D$  there exists a function  $m(\cdot, x) \in L^1_{\text{loc}}(0, \infty; \mathbb{R})$  such that

$$\|A^N p^N S(t)x\| \leq m(t; x) \quad \text{a.e. on } [0, \infty)$$

for all  $N$ .

Hypothesis (H2) is equivalent to

(H2\*) For any  $N$  there exists a norm  $\|\cdot\|_N$  on  $X^N$  such that  
 (i) for some constant  $\tilde{M} \geq 1$

$$\|x\| \leq \|x\|_N \leq \tilde{M}\|x\|, \quad x \in X^N, \quad N = 1, 2, \dots,$$

and

(ii) for some constant  $\tilde{\omega} \in \mathbb{R}$  all operators  $A^N - \tilde{\omega}I$  are dissipative on  $(X^N, \|\cdot\|_N)$ , i.e. (cf. [32], Thm. 4.2)

$$\|(A^N - \mu I)x\|_N \geq (\mu - \tilde{\omega})\|x\|_N$$

for all  $x \in X^N$  and all  $\mu > \tilde{\omega}$ .

If (H2) is satisfied we define

$$\|x\|_N = \sup_{t \geq 0} \|S^N(t)x\| e^{-\tilde{\omega}t}, \quad x \in X^N, \quad N = 1, 2, \dots$$

It is easy to see that  $\|\cdot\|_N$  is a norm on  $X^N$  with

$$\|x\| \leq \|x\|_N \leq \tilde{M}\|x\|, \quad x \in X^N, \quad N = 1, 2, \dots$$

Moreover for the operator norm corresponding to  $\|\cdot\|_N$  we have

$$\|S^N(t)\|_N \leq e^{\tilde{\omega}t}, \quad t \geq 0, \quad N = 1, 2, \dots$$

Then (H2\*,ii) is an immediate consequence (cf. [32], Thm. 4.3).

Conversely, if (H2\*) is valid then by (H2\*,ii) we get

$$\|S^N(t)\|_N \leq e^{\tilde{\omega}t}, \quad t \geq 0, \quad N = 1, 2, \dots$$

Using (H2\*,i) we immediately get (H2).

We note some consequences of (H1) - (H3) which will be useful for the proof of our convergence result:

a) The projections  $p^N$  are uniformly bounded, i.e. there exists a constant  $\gamma > 0$  such that

$$\|p^N\| \leq \gamma, \quad N = 1, 2, \dots$$

b) If  $x: [0, T] \rightarrow X$ ,  $T > 0$ , is continuous then

$$\lim_{N \rightarrow \infty} p^N x(t) = x(t)$$

uniformly for  $t \in [0, T]$ .

Assertion a) follows from the uniform boundedness principle. If b) were not true, we could deduce the existence of a monotone sequence  $N_k \rightarrow \infty$  and a sequence  $\{t_k\}$  in  $[0, T]$  with  $t_k \rightarrow t_0 \in [0, T]$  such that  $\|p^{N_k} x(t_k) - x(t_k)\| \geq \alpha > 0$  for all  $k$ . Using continuity of  $x$  we immediately would get a contradiction to  $p^N x(t_0) \rightarrow x(t_0)$ .

Theorem 4.1. Let (H1) - (H3) be satisfied for the sequences  $x^N$ ,  $p^N$ ,  $A^N$ ,  $N = 1, 2, \dots$ . Then for all  $x_0 \in X$

$$\lim_{N \rightarrow \infty} e^{A^N t} p^N x_0 = S(t)x_0 \quad (4.3)$$

uniformly for  $t$  in bounded intervals.

Proof. We first choose  $x_0 \in D$  and put

$$\Delta^N(t) = p^N x(t) - x^N(t), \quad t \geq 0, \quad N = 1, 2, \dots,$$

where  $x(t) = x(t; x_0)$  and  $x^N(t) = x^N(t; p^N x_0)$  are the solutions of (4.1) and (4.2), respectively. Since the derivative of  $\Delta^N(t)$



exists for all  $t > 0$ , the left-hand derivative  $\frac{d^-}{dt} \|\Delta^N(t)\|_N$  exists for all  $t > 0$  and is given by ([30], p. 228)

$$\frac{d^-}{dt} \|\Delta^N(t)\|_N = \tau_-(\Delta^N(t), \dot{\Delta}^N(t)), \quad t > 0,$$

where  $\tau_-(x, y) = \lim_{\tau \rightarrow 0^-} \frac{1}{\tau} (\|x + \tau y\|_N - \|x\|_N)$ .

Using (4.1), (4.2) and the estimates  $\tau_-(x, y_1 + y_2) \leq \tau_-(x, y_1) + \tau_+(x, y_2)$  and  $|\tau_+(x, y)| \leq \|y\|_N \leq \tilde{M}\|y\|$  we get

$$\begin{aligned} \frac{d^-}{dt} \|\Delta^N(t)\|_N &= \tau_-(\Delta^N(t), p^N A x(t) - A^N x^N(t)) \\ &= \tau_-(\Delta^N(t), p^N A x(t) - A^N p^N x(t) + A^N \Delta^N(t)) \\ &\leq \tilde{M}\gamma \|A x(t) - A^N p^N x(t)\| + \tau_-(\Delta^N(t), A^N \Delta^N(t)), \\ &t > 0. \end{aligned}$$

By (H2\*, ii) we have  $\tau_-(\Delta^N(t), A^N \Delta^N(t)) \leq \tilde{\omega} \|\Delta^N(t)\|_N$  for all  $N$  (see for instance [30], p. 244). Therefore for  $\tilde{\gamma} = \tilde{M}\gamma$

$$\frac{d^-}{dt} \|\Delta^N(t)\|_N \leq \tilde{\omega} \|\Delta^N(t)\|_N + \tilde{\gamma} \|A x(t) - A^N p^N x(t)\|, \quad t > 0,$$

which implies

$$\begin{aligned} \|\Delta^N(t)\|_N &\leq \|\Delta^N(t)\|_N \\ &\leq \tilde{\gamma} \int_0^t \|A x(\tau) - A^N p^N x(\tau)\| e^{\tilde{\omega}(t-\tau)} d\tau, \quad t \geq 0. \end{aligned} \tag{4.4}$$

for all  $N$ . Since  $D$  is invariant with respect to  $S(t)$  we have  $x(\tau) = S(\tau)x_0 \in D$  for  $\tau \geq 0$ . By (H3) we can use Lebesgue's dominated convergence theorem in order to get

$$\lim_{N \rightarrow \infty} \Delta^N(t) = 0$$

uniformly for  $t$  in bounded intervals provided  $x_0 \in D$ . The estimate  $\|x(t) - x^N(t)\| \leq \|x(t) - p^N x(t)\| + \|\Delta^N(t)\|$  together with b) from above proves

$$\lim_{N \rightarrow \infty} e^{A^N t} p^N x_0 = S(t)x_0$$

uniformly for  $t$  in bounded intervals for all  $x_0 \in D$ .

For arbitrary  $x_0 \in X$  we choose a sequence  $\{x_n\}$  in  $D$  with  $x_n \rightarrow x_0$ . Then the estimate

$$\begin{aligned} \|x(t; x_0) - x^N(t; p^N x_0)\| &\leq \|x(t; x_0) - x(t; x_n)\| \\ &\quad + \|x(t; x_n) - x^N(t; p^N x_n)\| + \|x^N(t; p^N x_n) - x^N(t; p^N x_0)\| \\ &\leq (\tilde{M}e^{\tilde{\omega}t} + Me^{\omega t})\|x_0 - x_n\| + \|x(t; x_n) - x^N(t; p^N x_n)\| \end{aligned}$$

proves (4.3) for all  $x_0 \in X$   $\square$

The methods in the proof of the previous theorem are well known in connection with numerical approximation of partial differential equations (see for instance the proof of the Lax-Richtmyer equivalence theorem in [22]). For delay equations this approach appears for the first time in [3], [6] and has later on been used in [25]. We consider the more general Banach space situation, because the proof is almost the same as in the Hilbert space case. We equally well could have used the Trotter Kato theorem [32]. We choose the classical Lax-Richtmyer idea, because with minor modifications the above proof also applies to the case of time varying coefficients.

Next we consider the nonhomogeneous problem

$$\dot{x}(t) = Ax(t) + Bu(t), \quad t \geq s, \quad (4.5;1)$$

$$x(s) = x_0 \in X, \quad (4.5;2)$$

where  $u \in L^2_{loc}(s, \infty; \mathbb{R}^l)$  and  $B$  is a linear operator  $\mathbb{R}^l \rightarrow X$ . The unique mild solution  $x(t) = x(t; s, x_0, u)$  of (4.5) is given by

$$x(t) = S(t-s)x_0 + \int_s^t S(t-\tau)Bu(\tau)d\tau, \quad t \geq s. \quad (4.6)$$

In addition to the approximating sequence  $X^N, p^N, A^N, N = 1, 2, \dots$ , introduced above let us assume that  $B^N, N = 1, 2, \dots$ , is a sequence of corresponding input operators  $\mathbb{R}^l \rightarrow X^N$ . Then we consider the approximating systems

$$\dot{x}^N(t) = A^N x^N(t) + B^N u(t), \quad t \geq s, \quad (4.7;1)$$

$$x^N(s) = p^N x_0, \quad x_0 \in X, \quad (4.7;2)$$

on  $X^N$  with the unique solution  $x^N(t) = x^N(t; s, p^N x_0, u)$  given by

$$x^N(t) = S^N(t-s)p^N x_0 + \int_s^t S^N(t-\tau)B^N u(\tau)d\tau, \quad t \geq s, \quad (4.8)$$

where  $S^N(t): X \rightarrow X$  is defined as above.

The following theorem and its proof are a slight modification of a result already established in [4].

**Theorem 4.2.** Assume that  $S^N(\cdot), N = 1, 2, \dots$ , and  $S(\cdot)$  are  $C_0$ -semigroups on  $X$  such that for constants  $M \geq 1, \omega \in \mathbb{R}$

$$\|S^N(t)\| \leq Me^{\omega t}, \quad t \geq 0, \quad N = 1, 2, \dots, \quad (4.9)$$

and for all  $x_0 \in X$

$$\lim_{N \rightarrow \infty} S^N(t) p^N x_0 = S(t) x_0 \quad (4.10)$$

uniformly on bounded  $t$ -intervals. Furthermore assume that

$$\lim_{N \rightarrow \infty} B^N \xi = B \xi \text{ for all } \xi \in \mathbb{R}^l. \quad (4.11)$$

Then for all  $x_0 \in X$  and  $T > 0$

$$\lim_{N \rightarrow \infty} x^N(t; s, p^N x_0, u) = x(t; s, x_0, u)$$

uniformly for  $0 \leq s \leq t \leq T$  and for  $u \in L^2(s, T; \mathbb{R}^l)$  with  $\|u\|_{L^2(s, T; \mathbb{R}^l)} \leq 1$ .

Proof. By (4.9), (4.10) and (4.11) we have  $S^N(t) B^N \xi \rightarrow S(t) B \xi$  for all  $\xi \in \mathbb{R}^l$  uniformly for  $t \in [0, T]$  (note, that (4.10) implies  $p^N \rightarrow I$  strongly) which shows

$$\|S^N(t) B^N - S(t) B\| \rightarrow 0$$

uniformly for  $t \in [0, T]$ . By (4.9) and an application of the dominated convergence theorem we get

$$\int_0^T \|S^N(t) B^N - S(t) B\|^2 dt \rightarrow 0.$$

Then the result follows from

$$\begin{aligned} & \|x(t; s, x_0, u) - x^N(t; s, p^N x_0, u)\| \\ & \leq \|S(t-s)x_0 - S^N(t-s)p^N x_0\| + \int_s^t \|S(t-\tau)B - S^N(t-\tau)B^N\| |u(\tau)| d\tau \\ & \leq \|S(t-s)x_0 - S^N(t-s)p^N x_0\| + \\ & \quad + \left( \int_0^T \|S(t)B - S^N(t)B^N\|^2 dt \right)^{1/2} \|u\|_{L^2(s, T; \mathbb{R}^l)} \quad \square \end{aligned}$$

## 4.2. Approximation of the feedback law

Throughout this section we assume that  $X$  is a Hilbert space. We restrict ourselves to the finite time control problem of minimizing the cost functional

$$\begin{aligned} J_s(u) = & \langle x(T), Gx(T) \rangle \\ & + \int_s^T [ \|Cx(t)\|^2 + u(t)^T Ru(t) ] dt \end{aligned} \quad (4.12)$$

associated with the Cauchy problem (4.5). We assume that the operators  $G: X \rightarrow X$ ,  $R: \mathbb{R}^k \rightarrow \mathbb{R}^k$ ,  $C: X \rightarrow \mathbb{R}^m$  are defined as in Section 3.1. As we have seen in that section (with obvious modifications for the case when the initial time  $s$  is not necessarily zero), the unique solution of this problem is given by the feedback law

$$\bar{u}_s(t) = -R^{-1}B^*P(t)\phi(t,s)x_0, \quad s \leq t \leq T, \quad (4.13)$$

where  $P(t): X \rightarrow X$  is the unique positive semidefinite solution of the Riccati differential equation (3.5) and  $\phi(t,s)$  is given by (3.8).

Correspondingly, we consider the sequence of control problems of minimizing

$$\begin{aligned} J_s^N(u) = & \langle x^N(T), Gx^N(T) \rangle \\ & + \int_s^T [ \|Cx^N(t)\|^2 + u(t)^T Ru(t) ] dt \end{aligned} \quad (4.14)$$

where  $x^N(t) = x^N(t; s, p^N x_0, u)$  is the unique solution of (4.7). The optimal control is given by the feedback law

$$\begin{aligned} \bar{u}_s^N(t) = & -R^{-1}(B^N)^*p^N(t)\phi^N(t,s)p^N x_0 \\ = & -R^{-1}(B^N)^*p^N(t)\phi^N(t,s)x_0, \quad s \leq t \leq T, \end{aligned} \quad (4.15)$$

where the strongly continuous, positive semidefinite operator  $P^N(t): X \rightarrow X$  and the strongly continuous evolution operator  $\phi^N(t,s): X \rightarrow X$  are defined by the equations

$$P^N(t)x = S^N(T-t) * p^N G p^N \phi^N(T,t)x + \int_t^T S^N(\tau-t) * p^N C * C p^N \phi^N(\tau,t)x d\tau, \quad t \leq T, \quad (4.16)$$

and

$$\phi^N(t,s)x = S^N(t-s)x - \int_s^t S^N(t-\tau) B^N R^{-1} (B^N)^* p^N(\tau) \phi^N(\tau,s)x d\tau, \quad t \geq s, \quad (4.17)$$

for  $x \in X$ . It follows immediately from (4.16) and the fact that  $P^N(t)$  is selfadjoint that

$$P^N(t) = p^N P^N(t) p^N, \quad t \leq T. \quad (4.18)$$

This in turn implies, by (4.17), that

$$p^N \phi^N(t,s) = \phi^N(t,s) p^N, \quad s \leq t \leq T. \quad (4.19)$$

Note that these two facts justify the second equation in (4.15). Moreover, the optimal cost of (4.14), (4.7) is given by

$$J_s^N(\bar{u}_s^N) = \langle x_0, P^N(s)x_0 \rangle. \quad (4.20)$$

We remark that  $P^N(t)$ , regarded as an operator on  $X^N$ , satisfies the following finite dimensional Riccati differential equation

$$\begin{aligned} & \frac{d}{dt} P^N(t) + (A^N) * P^N(t) + P^N(t) A^N \\ & - P^N(t) B^N R^{-1} (B^N) * P^N(t) + p^N C * C p^N = 0, \quad t \leq T, \quad (4.21) \\ & P^N(T) = p^N G p^N. \end{aligned}$$

Obviously, the most interesting question is how the original system (4.5) behaves when the optimal feedback control (4.13) is replaced by the approximate control law

$$\hat{u}_s^N(t) = -R^{-1} (B^N) * P^N(t) \hat{\phi}^N(t,s) x_0 \quad (4.22)$$

where  $\hat{\phi}^N(t,s)$  denotes the corresponding closed loop evolution operator on  $X$  and is defined by

$$\begin{aligned} \hat{\phi}^N(t,s)x &= S(t-s)x \\ & - \int_s^t S(t-\tau) B R^{-1} (B^N) * P^N(\tau) \hat{\phi}^N(\tau,s) x d\tau \end{aligned} \quad (4.23)$$

for  $x \in X$  and  $s \leq t \leq T$ . All the desired convergence results are contained in the next theorem which is a straight forward consequence of Theorems 6.1 - 6.3 in [21]. For the convenience of the reader we present the main ideas of the proof.

**Theorem 4.3.** Let us assume that

(i) there exist constants  $M \geq 1$ ,  $\omega \in \mathbb{R}$  such that

$$\|S^N(t)\| \leq M e^{\omega t}, \quad t \geq 0, \quad N = 1, 2, \dots,$$

(ii) for every  $x \in X$

$$\lim_{N \rightarrow \infty} S^N(t) p^N x = S(t)x, \quad \lim_{N \rightarrow \infty} S^N(t) * p^N x = S(t) * x$$

uniformly on  $[0, T]$  and

$$(iii) \quad \lim_{N \rightarrow \infty} B^N \xi = B \xi \text{ for every } \xi \in \mathbb{R}^k.$$

Then, for every  $x_0 \in X$ ,

$$(a) \quad \lim_{N \rightarrow \infty} J_s^N(\bar{u}_s^N) = \lim_{N \rightarrow \infty} J_s(\hat{u}_s^N) = J_s(\bar{u}_s),$$

$$(b) \quad \lim_{N \rightarrow \infty} \bar{u}_s^N(t) = \lim_{N \rightarrow \infty} \hat{u}_s^N(t) = \bar{u}_s(t),$$

$$(c) \quad \lim_{N \rightarrow \infty} \phi^N(t,s)x_0 = \lim_{N \rightarrow \infty} \hat{\phi}^N(t,s)x_0 = \phi(t,s)x_0,$$

$$(d) \quad \lim_{N \rightarrow \infty} P^N(s)x_0 = P(s)x_0$$

and the limits are uniform on the domain  $0 \leq s \leq t \leq T$ . If  $G: X \rightarrow X$  is a finite dimensional operator, then  $P^N(s)$  converges to  $P(s)$  in the uniform operator topology, uniformly on the interval  $[0, T]$ .

Proof. Let us introduce the operators  $F_s(t): L^2(s, T; \mathbb{R}^k) \rightarrow X$ ,  $G_s: X \rightarrow L^2(s, T; \mathbb{R}^k)$ ,  $R_s: L^2(s, T; \mathbb{R}^k) \rightarrow L^2(s, T; \mathbb{R}^k)$  by defining

$$F_s(t)u = \int_s^t S(t-\tau)Bu(\tau)d\tau,$$

$$G_s x = F_s(T)^* G S(T-s)x + \int_s^T F_s(\tau)^* C^* C S(\tau-s)x d\tau, \quad (4.24)$$

$$R_s u = F_s(T)^* G F_s(T)u + \int_s^T F_s(\tau)^* C^* C F_s(\tau)u d\tau + Ru$$

for  $u \in L^2(s, T; \mathbb{R}^k)$  and  $x \in X$ . Of course,  $Ru$  is defined by  $(Ru)(t) = Ru(t)$ ,  $s \leq t \leq T$ . Then it is easy to see that the Fréchet derivative of  $J_s$  with respect to  $u$  is given by  $J'_s(u) = 2R_s u + 2G_s x_0$ . Since the optimal control  $\bar{u}_s$  satisfies  $J'_s(\bar{u}_s) = 0$ , this implies

$$\bar{u}_s = -R_s^{-1} G_s x_0. \quad (4.25)$$

Analogously, we get



$$\bar{u}_s^N = - (R_s^N)^{-1} G_s^N p^N x_0 = - (R_s^N)^{-1} G_s^N x_0 \quad (4.26)$$

where  $R_s^N$ ,  $G_s^N$ ,  $F_s^N$  are defined as above with  $S(t)$ ,  $B$ ,  $C$ ,  $G$  replaced by  $S^N(t)$ ,  $B^N$ ,  $Cp^N$ ,  $p^N Gp^N$ , respectively. Combining these formulae with (3.8), (4.13) and (4.17), (4.15), we get

$$\phi(t,s)x_0 = S(t-s)x_0 - F_s(t)R_s^{-1}G_s x_0, \quad (4.27)$$

$$\phi^N(t,s)x_0 = S^N(t-s)x_0 - F_s^N(t)(R_s^N)^{-1}G_s^N x_0 \quad (4.28)$$

for every  $s \in [0, T]$  and every  $t \in [s, T]$ .

We have shown in Theorem 4.2 that  $F_s^N(t)$  converges to  $F_s(t)$  in the uniform operator topology, uniformly for  $0 \leq s \leq t \leq T$ . This implies that for every  $x \in X$

$$\lim_{N \rightarrow \infty} G_s^N x = G_s x \quad (4.29)$$

uniformly on  $[0, T]$  and moreover  $\|R_s^N - R_s\| \rightarrow 0$ , also uniformly on  $[0, T]$ . Choosing  $\epsilon > 0$  such that  $\xi^T R \xi \geq \epsilon |\xi|^2$  for  $\xi \in \mathbb{R}^l$ , we obtain

$$\|R_s^N u\| \geq \epsilon \|u\|, \quad u \in L^2(s, T; \mathbb{R}^l), \quad N = 1, 2, \dots$$

and hence

$$\lim_{N \rightarrow \infty} \|(R_s^N)^{-1} - R_s^{-1}\| = 0 \quad (4.30)$$

uniformly on  $[0, T]$ .

It follows immediately from (4.27 - 4.30) that  $\phi^N(t,s)$  converges strongly to  $\phi(t,s)$ . By (4.16) and (3.7), this implies the strong convergence of the Riccati operators  $P^N(s)$  to  $P(s)$ . Now the convergence result on  $\hat{\phi}^N(t,s)$  follows from the inequality

$$\begin{aligned}
& \|\phi(t,s)x - \hat{\phi}^N(t,s)x\| \\
& \leq \int_s^t \|S(t-\tau)BR^{-1}\| \|[(B^N)^*P^N(\tau) - B^*P(\tau)]\phi(\tau,s)x\| d\tau \\
& \quad + \int_0^t \|S(t-\tau)BR^{-1}(B^N)^*P^N(\tau)\| \|\hat{\phi}^N(\tau,s)x - \phi(\tau,s)x\| d\tau
\end{aligned}$$

and Gronwall's lemma. Thus we have established the statements (c) and (d). Statement (b) follows from (c) and (d), since the control functions  $\bar{u}_s$ ,  $\bar{u}_s^N$ ,  $\hat{u}_s^N$  are given by (4.13), (4.15), (4.22) respectively. Statement (a) is an immediate consequence of (b) and (d), since  $J_s^N(\bar{u}_s^N) = \langle x_0, P^N(s)x_0 \rangle$  and  $J_s(\bar{u}_s) = \langle x_0, P(s)x_0 \rangle$ . If  $G: X \rightarrow X$  is a finite dimensional operator, then the convergence of  $P^N(s)$  in the uniform operator topology can be established by analogous considerations as those in the proof of Theorem 4.2, again by the use of the formulae (4.16) and (3.7)  $\square$

### 4.3. A special spline scheme

In this section we develop a special scheme which satisfies all assumptions of Sections 4.1 and 4.2. For  $N = 1, 2, \dots$  we choose the meshpoints

$$t_{kj}^N = -h_{k-1} - j \frac{r_k}{N}, \quad k = 1, \dots, p; \quad j = 0, \dots, N,$$

where

$$r_k = h_k - h_{k-1}, \quad k = 1, \dots, p.$$

The sequence  $X^N$ ,  $N = 1, 2, \dots$ , of subspaces of  $M^2$  is defined by

$$X^N = \{ \phi \in M^2 \mid \phi = \hat{e}_0^N \alpha_0 + \sum_{k=1}^p \sum_{j=0}^N \hat{e}_{kj}^N \alpha_{kj}, \quad \alpha_0, \alpha_{kj} \in \mathbb{R}^n \}.$$

The "basis elements"  $\hat{e}_0^N, \hat{e}_{kj}^N$  are given by

$$\hat{e}_0^N = (I, 0), \quad \hat{e}_{kj}^N = (0, e_{kj}^N),$$

where

$$e_{k0}^N(\tau) = \begin{cases} \frac{N}{r_k}(\tau - t_{k1}^N)I & \text{for } \tau \in [t_{k1}^N, t_{k0}^N), \\ 0 & \text{elsewhere,} \end{cases}$$

$$e_{kj}^N(\tau) = \begin{cases} -\frac{N}{r_k}(\tau - t_{k,j-1}^N)I & \text{for } \tau \in [t_{kj}^N, t_{k,j-1}^N], \\ \frac{N}{r_k}(\tau - t_{k,j+1}^N)I & \text{for } \tau \in [t_{k,j+1}^N, t_{kj}^N], \\ 0 & \text{elsewhere,} \end{cases}$$

$j = 1, \dots, N-1$ , and

$$e_{kN}^N(\tau) = \begin{cases} -\frac{N}{r_k}(\tau - t_{k,N-1}^N)I & \text{for } \tau \in [t_{kN}^N, t_{k,N-1}^N], \\ 0 & \text{elsewhere.} \end{cases}$$

Here  $I$  denotes the  $n \times n$  identity matrix. The following diagram illustrates the definition of the basis elements  $e_{kj}^N$ .

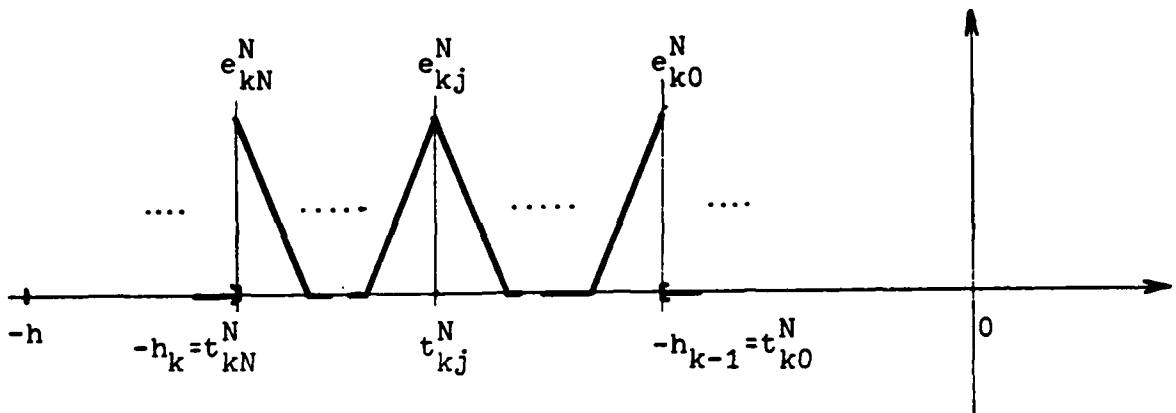


Figure 1

It is obvious that

$$\dim X^N = n((N+1)p+1).$$

We see that  $X^N$  is the subspace of all elements  $(\phi^0, \phi^1) \in M^2$  such that  $\phi^0$  is arbitrary in  $\mathbb{R}^n$  and  $\phi^1$  is a piecewise linear  $\mathbb{R}^n$ -valued function which is continuous except possibly at the delay points  $-h_k$  where jumps of arbitrary size can occur. By definition we can always assume that  $\phi^1$  is right-hand continuous on  $(-h, 0)$ . Thus  $X^N$  is neither a subspace of  $\text{dom } A$  nor one of  $\text{dom } A^*$ . Since  $X^N$  contains all  $(\phi^0, \phi^1)$  where  $\phi^0 \in \mathbb{R}^n$  and  $\phi^1$  is a spline of first order corresponding to the mesh  $\{t_{kj}^N\}$  it is clear that the orthogonal projections  $p^N: M^2 \rightarrow X^N$  satisfy hypothesis (H1) of Section 4.1.

For notational purposes we introduce the orthogonal projection  $p_1^N: L^2(-h, 0; \mathbb{R}^n) \rightarrow \text{span}(e_{10}^N, \dots, e_{pN}^N)$ , i.e. for  $\phi = (\phi^0, \phi^1) \in M^2$  we have  $p_1^N \phi = (\phi^0, p_1^N \phi^1)$ . Furthermore we put

$$\hat{E}^N = (\hat{e}_0^N, \hat{e}_{10}^N, \dots, \hat{e}_{pN}^N)$$

and denote by  $\alpha^N(\phi)$  the "coordinate vector"  $(\alpha_0^N, \alpha_{10}^N, \dots, \alpha_{pN}^N)^T \in \mathbb{R}^{n[(N+1)p+1]}$  of an element  $\phi \in X^N$ , i.e.

$$\phi = \hat{E}^N \alpha^N(\phi), \quad \phi \in X^N.$$

An easy calculation shows

$$\alpha^N(p_1^N \phi) = (Q^N)^{-1} d^N(\phi), \quad \phi \in M^2, \quad (4.31)$$

where

$$d^N(\phi) = \langle \hat{E}^N, \phi \rangle_{M^2} = \text{col}(\phi^0, \langle e_{10}^N, \phi^1 \rangle_{L^2}, \dots, \langle e_{pN}^N, \phi^1 \rangle_{L^2})$$

and

$$Q^N = \text{diag}(I, \frac{r_1}{N} q^N \otimes I, \dots, \frac{r_p}{N} q^N \otimes I).$$

The  $(N+1) \times (N+1)$  matrix  $q^N$  is given by

$$q^N = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & 0 & \dots & 0 \\ \frac{1}{6} & \frac{2}{3} & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \frac{1}{6} & \frac{1}{3} \end{pmatrix}.$$

For elements in  $X^N$  the inner product has the representation

$$\langle \phi, \psi \rangle_{M^2} = \alpha^N(\phi)^T Q^N \alpha^N(\psi), \quad \phi, \psi \in X^N. \quad (4.32)$$

We have seen that  $X^N$  is not contained in the domains of the operators  $A$  and  $A^*$ . However, these operators can be formally extended to all of  $X^N$  in the following way.

$$[A\phi]^0 = A_0\phi^0 + \sum_{k=1}^p A_k\phi^1(-h_k) + \int_{-h}^0 A_{01}(\tau)\phi^1(\tau)d\tau,$$

$$[A\phi]^1(\tau) = \frac{d^+}{d\tau}\phi^1(\tau) + \delta_0(\tau)(\phi^0 - \lim_{\tau \rightarrow 0} \phi^1(\tau)) \\ + \sum_{k=1}^{p-1} \delta_k(\tau)(\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^1(\tau)),$$

$$[A^*\psi]^0 = \lim_{\tau \rightarrow 0} \psi^1(\tau) + A_0^T\psi^0,$$

$$[A^*\psi]^1(\tau) = A_{01}^T(\tau)\psi^0 - \frac{d^+}{d\tau} \psi^1(\tau) \\ + \sum_{k=1}^{p-1} \delta_k(\tau)(A_k^T\psi^0 - \psi^1(-h_k) + \lim_{\tau \rightarrow -h_k} \psi^1(\tau)) \\ + \delta_p(\tau)(A_p^T\psi^0 - \psi^1(-h)),$$

for  $\phi, \psi \in X^N$  where  $\delta_k$  denotes the Dirac delta impulse at  $-h_k$ ,  $k = 0, 1, \dots, p$ . We will introduce the operators  $A^N$  and  $(A^N)^*$  by projecting these formal extensions formally back into the subspace  $X^N$ . Since the jumps of the function components of elements of  $X^N$  occur precisely at  $\tau = -h_k$ , we have two possible interpretations of  $\delta_k$  as a functional on  $X^N$ , namely the evaluation of either the right hand or the left hand limit at  $-h_k$ . Correspondingly we introduce the following two types of approximate delta impulses which can be obtained by a formal projection of  $\delta_k$  in one of these two ways. We define

$$\delta_{k,+}^N = \hat{E}^N \gamma_{k,+}^N, \quad k = 1, \dots, p,$$

$$\delta_{k,-}^N = \hat{E}^N \gamma_{k,-}^N, \quad k = 0, \dots, p-1,$$

where

$$Q^N \gamma_{k,+}^N = \text{col}(0, e_{10}^N(-h_k), \dots, e_{pN}^N(-h_k))$$

and

$$Q^N \gamma_{k,-}^N = \text{col}(0, \lim_{\tau \rightarrow -h_k} e_{10}^N(\tau), \dots, \lim_{\tau \rightarrow -h_k} e_{pN}^N(\tau)).$$

First observations are contained in

Lemma 4.4. a) For any  $x \in \mathbb{R}^n$  and  $\phi \in M^2$

$$\langle \delta_{k,+}^N, x, \phi \rangle_{M^2} = x^T (p_1^N \phi^1)(-h_k), \quad k = 1, \dots, p,$$

$$\langle \delta_{k,-}^N, x, \phi \rangle_{M^2} = x^T \lim_{\tau \rightarrow -h_k} (p_1^N \phi^1)(\tau), \quad k = 0, \dots, p-1.$$

b) For  $\delta_{k,\pm}^N$  as operators  $\mathbb{R}^n \rightarrow M^2$  we have

$$\|\delta_{k,+}^N\| \leq \left(\frac{6N}{r_k}\right)^{1/2}, \quad k = 1, \dots, p,$$

$$\|\delta_{k,-}^N\| \leq \left(\frac{6N}{r_{k+1}}\right)^{1/2}, \quad k = 0, \dots, p-1.$$

Proof. a) Using (4.31), (4.32) and the definition of  $\delta_{k,+}^N$  we get

$$\begin{aligned} \langle \delta_{k,+}^N, x, \phi \rangle_{M^2} &= \langle \hat{E}^N \gamma_{k,+}^N, x, p^N \phi \rangle_{M^2} = \\ &= (\gamma_{k,+}^N, x)^T Q^N \alpha^N (p^N \phi) \\ &= x^T \text{col}(0, e_{10}^N(-h_k), \dots, e_{pN}^N(-h_k)) \alpha^N (p^N \phi) \\ &= x^T (p_1^N \phi^1)(-h_k). \end{aligned}$$

The proof for  $\delta_{k,-}^N$  is analogous.

b) Using (4.32) and the definitions of  $\delta_{k,+}^N$ ,  $Q^N$  we get for any  $x \in \mathbb{R}^n$

$$\begin{aligned}
\|\delta_{k,+}^N x\|^2 &= (\gamma_{k,+}^N x)^T Q^N \gamma_{k,+}^N x \\
&= (0, \dots, 0, x^T) \frac{N}{r_k} (q^N \otimes I)^{-1} \text{col}(0, \dots, 0, x) \\
&\leq \frac{6N}{r_k} |x|^2, \quad k = 1, \dots, p,
\end{aligned}$$

where we have used  $\lambda_{\min}(q^N) \geq \frac{1}{6}$ . The estimate for  $\delta_{k,-}^N$  is analogous  $\square$

Definition 4.5. For any  $\phi = (\phi^0, \phi^1) \in X^N$  we define

$$\begin{aligned}
A^N \phi &= (A_0 \phi^0 + \sum_{k=1}^p A_k \phi^1(-h_k) + \int_{-h}^0 A_{01}(\tau) \phi^1(\tau) d\tau, p_1^N(\frac{d^+}{d\theta} \phi^1)) \\
&\quad + \delta_{0,-}^N (\phi^0 - \lim_{\tau \rightarrow 0} \phi^1(\tau)) + \sum_{k=1}^{p-1} \delta_{k,-}^N (\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^1(\tau)).
\end{aligned}$$

It is clear that  $A^N$  is a linear operator  $X^N \rightarrow X^N$ . The adjoint operators are given in

Lemma 4.6. The adjoint operator  $(A^N)^*$  is given by

$$\begin{aligned}
(A^N)^* \psi &= (\lim_{\tau \rightarrow 0} \psi^1(\tau) + A_0^T \psi^0, p_1^N(A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1)) \\
&\quad + \sum_{k=1}^{p-1} \delta_{k,+}^N (A_k^T \psi^0 + \lim_{\tau \rightarrow -h_k} \psi^1(\tau) - \psi^1(-h_k)) \\
&\quad + \delta_{p,+}^N (A_p^T \psi^0 - \psi^1(-h))
\end{aligned}$$

for  $\psi = (\psi^0, \psi^1) \in X^N$ .

Proof. By definition of the adjoint operator we get



$$\begin{aligned}
\langle (A^N)^+ \psi, \phi \rangle_{M^2} &= \langle \psi, A^N \phi \rangle_{M^2} \\
&= (\psi^0)^T [A_0 \phi^0 + \sum_{k=1}^p A_k \phi^1(-h_k) + \int_{-h}^0 A_{01}(\tau) \phi^1(\tau) d\tau] \\
&\quad + \langle \psi^1, p_1^N \left( \frac{d^+}{d\theta} \phi^1 \right) \rangle_{L^2} + \langle \psi, \delta_{0,-}^N (\phi^0 - \lim_{\tau \rightarrow 0} \phi^1(\tau)) \rangle_{M^2} \\
&\quad + \sum_{k=1}^{p-1} \langle \psi, \delta_{k,-}^N (\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^1(\tau)) \rangle_{M^2}
\end{aligned}$$

for any  $\phi = (\phi^0, \phi^1)$ ,  $\psi = (\psi^0, \psi^1)$  in  $X^N$ . By part a) of Lemma 4.4 we see

$$\langle \psi, \delta_{0,-}^N (\phi^0 - \lim_{\tau \rightarrow 0} \phi^1(\tau)) \rangle_{M^2} = [\lim_{\tau \rightarrow 0} \psi^1(\tau)]^T (\phi^0 - \lim_{\tau \rightarrow 0} \phi^1(\tau))$$

and

$$\begin{aligned}
\langle \psi, \delta_{k,-}^N (\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^1(\tau)) \rangle_{M^2} &= \\
&= [\lim_{\tau \rightarrow -h_k} \psi^1(\tau)]^T (\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^1(\tau)),
\end{aligned}$$

$k = 1, \dots, p-1$ . Furthermore,

$$\begin{aligned}
\langle \psi^1, p_1^N \left( \frac{d^+}{d\theta} \phi^1 \right) \rangle_{L^2} &= \langle \psi^1, \frac{d^+}{d\theta} \phi^1 \rangle_{L^2} = \sum_{k=1}^p \int_{-h_k}^{-h_{k-1}} \psi^1(\tau)^T \left( \frac{d}{d\theta} \phi^1 \right)(\tau) d\tau \\
&= \sum_{k=1}^p \lim_{\tau \rightarrow -h_{k-1}} \psi^1(\tau)^T \phi^1(\tau) - \sum_{k=1}^p \psi^1(-h_k)^T \phi^1(-h_k) \\
&\quad - \langle p_1^N \left( \frac{d^+}{d\theta} \psi^1 \right), \phi^1 \rangle_{L^2}
\end{aligned}$$

and

$$(\psi^0)^T \int_{-h}^0 A_{01}(\tau) \phi^1(\tau) d\tau = \langle A_{01}^T \psi^0, \phi^1 \rangle_{L^2} = \langle p_1^N (A_{01}^T \psi^0), \phi^1 \rangle_{L^2}.$$

Putting things together we get

$$\begin{aligned}
\langle (A^N)^* \psi, \phi \rangle_{M^2} &= [\lim_{\tau \rightarrow 0} \psi^1(\tau) + A_0^T \psi^0]^T \phi_0 + \\
&+ \sum_{k=1}^{p-1} [A_k^T \psi^0 + \lim_{\tau \rightarrow -h_k} \psi^1(\tau) - \psi^1(-h_k)]^T \phi^1(-h_k) \\
&+ [A_p^T \psi^0 - \psi^1(-h_p)]^T \phi^1(-h_p) + \langle p_1^N (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1), \phi^1 \rangle_{L^2}.
\end{aligned}$$

The result now follows by an application of part a) of Lemma 4.4  $\square$

We define the sets  $\tilde{D}$  and  $\tilde{D}^*$  by

$$\tilde{D} = \{ \phi \in M^2 \mid \phi^0 = \phi^1(0), \phi^1 \in W^{2,2}(-h, 0; \mathbb{R}^n) \}$$

and

$$\tilde{D}^* = \{ \psi \in M^2 \mid \psi^1(-h) = A_p^T \psi^0, \psi^1 + \sum_{k=1}^{p-1} A_k^T \psi^0 \chi_{[-h, -h_k]} \in W^{1,2}(-h, 0; \mathbb{R}^n)$$

$$\text{and } A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1 \in W^{1,2}(-h_k, -h_{k-1}; \mathbb{R}^n), k = 1, \dots, p \}.$$

Lemma 4.7. a)  $D(A^2) \subset \tilde{D}$  and  $D((A^*)^2) \subset \tilde{D}^*$ .

b) There exists a constant  $\gamma_0$  such that for any  $\phi \in \tilde{D}$  and  $N = 1, 2, \dots$

$$\| A_p^N \phi - A \phi \|_{M^2} \leq \frac{\gamma_0}{N} \| \phi^1 \|_{W^{2,2}}.$$

c) Assume that  $A_{01} \in W^{1,2}(-h_k, -h_{k-1}; \mathbb{R}^{n \times n})$  for  $k = 1, \dots, p$ . Then there exists a constant  $\delta_0$  such that for any  $\psi \in \tilde{D}^*$  and  $N = 1, 2, \dots$

$$\begin{aligned}
&\| (A^N)^* \psi - A^* \psi \|_{M^2} \\
&\leq \frac{\delta_0}{N} \max_{k=1, \dots, p} (\| A_{01}^T \psi^0 \|_{L^2(-h_k, -h_{k-1}; \mathbb{R}^n)} + \\
&\quad + \| \psi^1 \|_{W^{2,2}(-h_k, -h_{k-1}; \mathbb{R}^n)}).
\end{aligned}$$

Proof. a) This is clear by (2.6) and Lemma 2.1.

b) For  $\phi = (\phi^0, \phi^1) \in \bar{D}$  we put  $\phi^N = p_1^N \phi^1$ . Then by (2.6) and Definition 4.5 we get

$$\begin{aligned}
 & \|A_p^N \phi - A\phi\| \\
 & \leq \left\| \left( \sum_{k=1}^p A_k [\phi^N(-h_k) - \phi^1(-h_k)] \right. \right. \\
 & \quad \left. \left. + \int_{-h}^0 A_{01}(\tau) [\phi^N(\tau) - \phi^1(\tau)] d\tau, p_1^N \left( \frac{d^+}{d\theta} \phi^N \right) - \dot{\phi}^1 \right\| \\
 & \quad + \|\delta_{0,-}^N\| |\phi^0 - \lim_{\tau \rightarrow 0} \phi^N(\tau)| + \sum_{k=1}^{p-1} \|\delta_{k,-}^N\| |\phi^N(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^N(\tau)| \\
 & \leq \sum_{k=1}^p \|A_k\| \|\phi^N - \phi^1\|_{L^\infty} + \|A_{01}\|_{L^2} \|\phi^N - \phi^1\|_{L^2} \\
 & \quad + \|p_1^N \left( \frac{d^+}{d\theta} \phi^N - \dot{\phi}^1 \right)\|_{L^2} + \|p_1^N(\dot{\phi}^1) - \dot{\phi}^1\|_{L^2} \\
 & \quad + \left( \frac{6N}{\rho} \right)^{1/2} [|\phi^1(0) - \lim_{\tau \rightarrow 0} \phi^N(\tau)| + \sum_{k=1}^{p-1} (|\phi^N(-h_k) - \phi^1(-h_k)| \\
 & \quad + |\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^N(\tau)|)] \\
 & \leq \left[ \sum_{k=1}^p \|A_k\| + (2p-1) \left( \frac{6N}{\rho} \right)^{1/2} \|\phi^1 - \phi^N\|_{L^\infty} + \|A_{01}\|_{L^2} \|\phi^1 - \phi^N\|_{L^2} \right. \\
 & \quad \left. + \left\| \frac{d^+}{d\theta} \phi^N - \dot{\phi}^1 \right\|_{L^2} + \|p_1^N(\dot{\phi}^1) - \dot{\phi}^1\|_{L^2} \right].
 \end{aligned}$$

Here we have put  $\rho = \min_{k=1, \dots, p} r_k$ . Since  $\phi^N|_{[-h_k, -h_{k-1}]}$ ,  $k = 1, \dots, p$ , is the image of  $\phi^1|_{[-h_k, -h_{k-1}]}$  under the orthogonal projection of  $L^2(-h_k, -h_{k-1}; \mathbb{R}^n)$  onto  $\text{span}(e_{k0}^N|_{[-h_k, -h_{k-1}]}, \dots, e_{kN}^N|_{[-h_k, -h_{k-1}]})$ , we get from standard estimates (cf. [38], Theorem 6.5 and Exercise 6.1)

$$\|\phi^1 - \phi^N\|_{L^2} \leq \frac{\text{const.}}{N^2} \|\phi^1\|_{L^2},$$

$$\|\dot{\phi}^1 - \frac{d^+}{d\theta} \phi^N\|_{L^2} \leq \frac{\text{const.}}{N} \|\ddot{\phi}^1\|_{L^2}$$

and

$$\|\dot{\phi}^1 - p_1^N(\dot{\phi}^1)\|_{L^2} \leq \frac{\text{const.}}{N} \|\ddot{\phi}^1\|_{L^2},$$

where the constants are not dependent on  $\phi^1$  and  $N$ . On each subinterval  $[-h_k, -h_{k-1}]$   $\phi^N = \chi^N$ , where  $\chi^N$  is the cubic type I interpolating spline for  $\phi(\tau) = \int_{-h_k}^{\tau} \int_{-h_k}^{\theta} \phi^1(\sigma) d\sigma d\theta$  (see for instance [38], Proof of Theorem 6.6). Note, that interpolating cubic splines in [38] are type I splines (cf. [1]). From [24], Theorem 5.7.1 (with  $L = \frac{d^2}{d\theta}$  and  $m = 2$ ) or [12], p. 235 (with  $m = r = 2, q = \infty$ ) we get

$$\|\phi^1 - \phi^N\|_{L^\infty} \leq \text{const.} \left(\frac{1}{N}\right)^{3/2} \|\phi^1\|_{W^{2,2}},$$

where again the constant is not dependent on  $\phi^1$  and  $N$ .

c) As in b) we put  $\psi^N = p_1^N \psi^1$  for  $\psi \in \tilde{D}^*$ . Using Lemma 2.1, Lemma 4.6 and  $\psi \in D(A^*)$  we get

$$\begin{aligned} & \| (A^N)^* p^N \psi - A^* \psi \| \\ & \leq \| (\lim_{\tau \rightarrow 0} \psi^N(\tau) - \psi^1(0), p_1^N (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^N) \\ & \quad - A_{01}^T \psi^0 + \frac{d}{d\theta} (\psi^1 + \sum_{k=1}^{p-1} A_k^T \psi^0 \chi_{[-h, -h_k]})) \| \\ & \quad + \sum_{k=1}^{p-1} \|\delta_{k,+}^N\| |A_k^T \psi^0 + \lim_{\tau \rightarrow -h_k} \psi^N(\tau) - \psi^N(-h_k)| \\ & \quad + \|\delta_{p,+}^N\| |A_p^T \psi^0 - \psi^N(-h)| \\ & \leq [1 + (2p-1) \left(\frac{6N}{\rho}\right)^{1/2}] \|\psi^1 - \psi^N\|_{L^\infty} \\ & \quad + \| p_1^N (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1) - (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1) \|_{L^2} + \|\frac{d^+}{d\theta} (\psi^1 - \psi^N)\|_{L^2}. \end{aligned}$$

Since  $A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1 \in W^{1,2}(-h_k, -h_{k-1}; \mathbb{R}^n)$ ,  $k = 1, \dots, p$ , we get from [38], Exercise 6.1,

$$\begin{aligned} & \left\| p_1^N (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1) - (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1) \right\|_{L^2} \\ & \leq \frac{\text{const.}}{N} \max_{k=1, \dots, p} \left\| \ddot{A}_{01}^T \psi^0 - \ddot{\psi}^1 \right\|_{L^2(-h_k, -h_{k-1}; \mathbb{R}^n)} \end{aligned}$$

where the constant is not dependent on  $\psi$  and  $A_{01}$ .

Similarly as in b) we get

$$\left\| \psi^1 - \psi^N \right\|_{L^\infty} \leq \text{const.} \left(\frac{1}{N}\right)^{3/2} \max_{k=1, \dots, p} \left\| \psi^1 \right\|_{W^{2,2}(-h_k, -h_{k-1}; \mathbb{R}^n)}$$

and

$$\left\| \frac{d^+}{d\theta} (\psi^1 - \psi^N) \right\|_{L^2} \leq \frac{\text{const.}}{N} \max_{k=1, \dots, p} \left\| \ddot{\psi}^1 \right\|_{L^2(-h_k, -h_{k-1}; \mathbb{R}^n)}.$$

Note, that from the assumption on  $A_{01}$ , we get  $\psi^1|_{[-h_k, -h_{k-1}]} \in W^{2,2}(-h_k, -h_{k-1}; \mathbb{R}^n)$ ,  $k = 1, \dots, p$ . Putting things together we get the desired estimate  $\square$

An immediate consequence of Lemma 4.7 is

**Corollary 4.8.** a) If we take  $D = D(A^2)$  then hypothesis (H3,i) is satisfied for the sequence  $X^N, p^N, A^N$ ,  $N = 1, 2, \dots$  and the semigroup  $S(\cdot)$ .

b) If  $A_{01} \in W^{1,2}(-h_k, -h_{k-1}; \mathbb{R}^{n \times n})$ ,  $k = 1, \dots, p$ , and  $D = D((A^*)^2)$  then hypothesis (H3,i) is satisfied for the sequence  $X^N, p^N, (A^N)^*$ ,  $N = 1, 2, \dots$ , and the semigroup  $S^*(\cdot)$ .

The next lemma establishes (H3,ii) for  $X^N, p^N, A^N, S(\cdot)$  and  $X^N, p^N, (A^N)^*, S^*(\cdot)$ , respectively.

Lemma 4.9. a) There exist constants  $\tilde{M} \geq 1$  and  $\omega \in \mathbb{R}$  such that for all  $\phi \in D(A^2)$

$$\|A_p^N S(t)\phi\| \leq \tilde{M} e^{\omega t} |\phi|_2, \quad t \geq 0, \quad N = 1, 2, \dots,$$

where  $|\phi|_2 = \|\phi\| + \|A\phi\| + \|A^2\phi\|$ .

b) Assume  $A_{01} \in W^{1,2}(-h_k, -h_{k-1}; \mathbb{R}^{n \times n})$ ,  $k = 1, \dots, p$ . Then there exist constants  $M^* \geq 1$  and  $\omega \in \mathbb{R}$  such that for all  $\psi \in D((A^*)^2)$

$$\|(A^N)^* S^*(t)\psi\| \leq M^* e^{\omega t} |\psi|_2^*, \quad t \geq 0, \quad N = 1, 2, \dots,$$

where  $|\psi|_2^* = \|\psi\| + \|A^*\psi\| + \|(A^*)^2\psi\|$ .

Proof. a) Since  $S(\cdot)$  restricted to  $D(A^2)$  is a  $C_0$ -semigroup on  $D(A^2)$  equipped with the norm  $|\cdot|_2$ , we have

$$|S(t)\phi|_2 \leq M e^{\omega t} |\phi|_2, \quad t \geq 0, \quad \phi \in D(A^2),$$

with some constants  $M \geq 1$ ,  $\omega \in \mathbb{R}$ . From  $A^k \phi = (L(\frac{d^{k-1}}{d\theta^{k-1}} \phi^1), \frac{d^k}{d\theta^k} \phi^1)$ ,  $k = 1, 2, \dots$ ,  $\phi = (\phi^0, \phi^1) \in D(A^k)$  we see that

$$\|\phi^1\|_{W^{2,2}} \leq |\phi|_2, \quad \phi \in D(A^2).$$

Therefore, for  $\phi \in D(A^2)$ ,

$$\|(S(t)\phi)^1\|_{W^{2,2}} \leq M e^{\omega t} |\phi|_2, \quad t \geq 0,$$

and by Lemma 4.7, b)

$$\begin{aligned} \|A_p^N S(t)\phi\| &\leq \|AS(t)\phi\| + \|A_p^N S(t)\phi - AS(t)\phi\| \\ &\leq M e^{\omega t} \|A\phi\| + \frac{\gamma_0}{N} M e^{\omega t} |\phi|_2 \\ &\leq (M + \gamma_0) e^{\omega t} |\phi|_2 \end{aligned}$$

for  $t \geq 0$  and  $N = 1, 2, \dots$ .

b) As in part a) we have

$$|S^*(t)\psi|_2^* \leq Me^{\omega t} |\psi|_2^*$$

for  $t \geq 0$  and  $\psi \in D((A^*)^2)$ . From  $A^*\psi = (\psi^1(0) + A_{01}^T \psi^0, A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1)$ ,  $(A^*)^2 \psi = (\dots, A_{01}^T (\psi^1(0) + A_{01}^T \psi^0) - \frac{d^+}{d\theta} (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1))$  and the fact that on the intervals  $[-h_k, -h_{k-1}]$ ,  $k = 1, \dots, p$ , we have  $\frac{d^+}{d\theta} (A_{01}^T \psi^0 - \frac{d^+}{d\theta} \psi^1) = \dot{A}_{01}^T \psi^0 - \ddot{\psi}^1$  it is not difficult to see that for a constant  $\kappa > 0$  depending on  $A_{01}$  the following estimate is valid:

$$\max_{k=1, \dots, p} \|\psi^1\|_{W^{2,2}(-h_k, -h_{k-1}; \mathbb{R}^n)} \leq \kappa |\psi|_2^*, \quad \psi \in D((A^*)^2).$$

The rest of the proof is analogous to that for part a) but now using Lemma 4.7, c)  $\square$

Lemma 4.10. Hypothesis (H2) is valid for the sequences  $A^N$  and  $(A^N)^*$ ,  $N = 1, 2, \dots$ .

Proof. We introduce an equivalent inner product on  $M^2$  by

$$\langle \phi, \psi \rangle_g = (\phi^0)^T \psi^0 + \int_{-h}^0 \phi^1(\tau)^T \psi^1(\tau) g(\tau) d\tau, \quad \phi, \psi \in M^2,$$

where  $g$  is right-hand continuous on  $[-h, 0]$  and

$$g(\tau) = p - k + 1 \quad \text{for } \tau \in (-h_k, -h_{k-1}), \quad k = 1, \dots, p.$$

It is clear that the corresponding norm  $\|\cdot\|_g$  on  $M^2$  is equivalent to the original norm. In fact we have

$$\|\phi\| \leq \|\phi\|_g \leq \sqrt{p} \|\phi\|, \quad \phi \in M^2.$$

Since  $(\phi^0, \phi^1 g) \in X^N$  for any  $\phi \in X^N$ , we get from Lemma 4.4, a) that

$$\langle \delta_{k, -x, \phi}^N \rangle_g = (p-k)x^T \lim_{\tau \rightarrow -h_k} \phi^1(\tau)$$

for  $k = 0, \dots, p-1$ ,  $x \in \mathbb{R}^n$  and  $\phi \in X^N$ . Using this equation and Definition 4.5 we get for  $\phi \in X^N$

$$\begin{aligned} \langle A^N \phi, \phi \rangle_g &= [A_0 \phi^0 + \sum_{k=1}^p A_k \phi^1(-h_k) + \int_{-h}^0 A_{01}(\tau) \phi^1(\tau) d\tau]^T \phi^0 \\ &+ \langle p_1^N(\frac{d^+}{d\theta} \phi^1), \phi^1 g \rangle_{L^2} + p[\phi^0 - \lim_{\tau \rightarrow 0} \phi^1(\tau)]^T \lim_{\tau \rightarrow 0} \phi^1(\tau) \\ &+ \sum_{k=1}^{p-1} (p-k)[\phi^1(-h_k) - \lim_{\tau \rightarrow -h_k} \phi^1(\tau)]^T \lim_{\tau \rightarrow -h_k} \phi^1(\tau). \end{aligned}$$

Obviously  $p_1^N(\phi^1 g) = \phi^1 g$  and hence

$$\begin{aligned} \langle p_1^N(\frac{d^+}{d\theta} \phi^1), \phi^1 g \rangle_{L^2} &= \langle \frac{d^+}{d\theta} \phi^1, \phi^1 g \rangle_{L^2} \\ &= \sum_{k=1}^p (p-k+1) \int_{-h_k}^{-h_{k-1}} \phi^1(\tau)^T \phi^1(\tau) d\tau \\ &= \frac{1}{2} \sum_{k=1}^p (p-k+1) [\lim_{\tau \rightarrow -h_{k-1}} |\phi^1(\tau)|^2 - |\phi^1(-h_k)|^2]. \end{aligned}$$

Using this and several times the inequality  $a\beta \leq \frac{1}{2} a^2 + \frac{1}{2} \beta^2$  we get for  $\phi \in X^N$



$$\begin{aligned}
\langle A^N \phi, \phi \rangle_{\mathcal{G}} &\leq (|A_0| + \frac{1}{2} \sum_{k=1}^p |A_k|^2 + \|A_{01}\|_{L^2}) \|\phi\|_{\mathcal{G}}^2 \\
&+ \frac{1}{2} \sum_{k=1}^p |\phi^1(-h_k)|^2 + \frac{p}{2} |\phi^0|^2 - \frac{p}{2} \lim_{\tau \rightarrow 0} |\phi^1(\tau)|^2 \\
&+ \frac{1}{2} \sum_{k=1}^p (p-k+1) \{ \lim_{\tau \rightarrow -h_{k-1}} |\phi^1(\tau)|^2 - |\phi^1(-h_k)|^2 \} \\
&+ \frac{1}{2} \sum_{k=2}^p (p-k+1) \{ |\phi^1(-h_{k-1})|^2 - \lim_{\tau \rightarrow -h_{k-1}} |\phi^1(\tau)|^2 \} \\
&\leq \omega \|\phi\|_{\mathcal{G}}^2 + \frac{1}{2} \sum_{k=1}^p |\phi^1(-h_k)|^2 - \frac{p}{2} |\phi^1(-h_1)|^2 \\
&+ \frac{1}{2} \sum_{k=2}^p (p-k+1) \{ |\phi^1(-h_{k-1})|^2 - |\phi^1(-h_k)|^2 \} \\
&= \omega \|\phi\|_{\mathcal{G}}^2
\end{aligned}$$

with  $\omega = \frac{p}{2} + |A_0| + \frac{1}{2} \sum_{k=1}^p |A_k|^2 + \|A_{01}\|_{L^2}$ . This proves (H2\*) with  $\|\cdot\|_N = \|\cdot\|_{\mathcal{G}}$  for all  $N$  (cf. [30], p. 244). Since (H2) and (H2\*) are equivalent and  $\|S^N(t)\| = \|S^N(t)^*\|$  the proof is finished  $\square$

Lemma 4.10 was the final step to show that Theorem 4.1 applies to the sequences  $X^N, p^N, A^N, N = 1, 2, \dots$ , and  $X^N, p^N, (A^N)^*$ ,  $N = 1, 2, \dots$ , defined in this section. The corresponding input and output operators are given by  $B^N = p^N B = B$  and  $C^N = C p^N = C$ , since  $\mathbb{R}^n \times \{0\} \subset X^N$  for every  $N \in \mathbb{N}$ . Hence the approximating systems are described by the ordinary differential equations

$$\begin{aligned}
\dot{w}^N(t) &= A^N w^N(t) + Bu(t) \\
y^N(t) &= Cw^N(t), \quad t \geq 0, \\
w^N(0) &= p^N \phi, \quad \phi \in M^2,
\end{aligned} \tag{E^N}$$

on the subspaces  $X^N$ .

In an analogous way we can define the operators  $A_T^N: M^2 \rightarrow M^2$  by taking the transposed matrices  $A_0^T, \dots, A_p^T$  and  $A_{01}^T(\cdot)$  in Definition 4.5. Obviously all the results of this section can also be applied to the operators  $A_T^N$ . Therefore we obtain the sequence

$$\begin{aligned} \dot{z}^N(t) &= (A_T^N)^* z^N(t) + Bu(t), \\ y^N(t) &= Cz^N(t), \quad t \geq 0, \\ z^N(0) &= p^N f, \quad f \in M^2, \end{aligned} \quad (\Sigma_T^{N*})$$

of control systems on  $X^N$  approximating the Cauchy problem  $(\Sigma_T^*)$ .

Now let us assume that  $R \in \mathbb{R}^{l \times l}$  is positive definite,  $G_0 \in \mathbb{R}^{n \times n}$  is positive semidefinite and  $G: M^2 \rightarrow M^2$  is defined by  $G\phi = (G_0\phi^0, 0)$  for  $\phi \in M^2$  (compare Section 3.2). Then we consider the control problems of minimizing the cost functional

$$J^N(u) = \langle z^N(T), Gz^N(T) \rangle + \int_0^T (|y^N(t)|^2 + u(t)^T Ru(t)) dt \quad (4.33)$$

subject to  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ , respectively. The corresponding Riccati operators are

$$\Pi^N(t) = p^N \Pi^N(t) p^N, \quad P^N(t) = p^N P^N(t) p^N, \quad 0 \leq t \leq T,$$

and satisfy the following Riccati differential equations on  $X^N$

$$\begin{aligned} \frac{d}{dt} \Pi^N(t) + (A^N)^* \Pi^N(t) + \Pi^N(t) A^N \\ - \Pi^N(t) B R^{-1} B^* \Pi^N(t) + C^* C = 0, \quad 0 \leq t \leq T, \end{aligned} \quad (4.34)$$

$$\Pi^N(T) = G,$$

and

$$\begin{aligned} & \frac{d}{dt} P^N(t) + A_T^N P^N(t) + P^N(t) (A_T^N)^* \\ & - P^N(t) B R^{-1} B^* P^N(t) + C^* C = 0, \quad 0 \leq t \leq T, \end{aligned} \quad (4.35)$$

$$P^N(T) = G.$$

These two Riccati equations have an interesting interconnection. In Section 5.1 we shall see that - for a large class of systems - there exists an operator  $F^N: X^N \rightarrow X^N$  which maps the solutions of  $(\Sigma^N)$  onto those of  $(\Sigma_T^{N*})$  or, equivalently, satisfies

$$(A_T^{N*}) F^N = F^N A^N, \quad B = F^N B, \quad C = C F^N.$$

Under these conditions it is easy to see that

$$\Pi^N(t) = (F^N)^* P^N(t) F^N, \quad 0 \leq t \leq T. \quad (4.36)$$

It follows from the results of this section that Theorem 4.3 can be applied to the systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ . More precisely, we have the following theorem which may be considered as the main result of this paper.

**Theorem 4.11.** Assume that  $A_{01} \in W^{1,2}(-h_k, -h_{k-1}; \mathbb{R}^{n \times n})$  for  $k = 1, \dots, p$ . Then

$$\lim_{N \rightarrow \infty} \|\Pi(t) - \Pi^N(t)\| = 0 = \lim_{N \rightarrow \infty} \|P(t) - P^N(t)\|$$

uniformly for  $0 \leq t \leq T$ .

For implementation of the scheme we have to calculate matrix representations for the operators  $A^N$ ,  $(A^N)^*$  and  $B$  (as an operator  $\mathbb{R}^l \rightarrow X^N$ ) with respect to the basis  $\hat{E}^N$ . Formula (4.31) shows how to calculate the coordinate vector of  $p^N_\phi$  for  $\phi \in M^2$ .

Define the  $(N+1) \times (N+1)$ -matrix  $h^N$  by

$$h^N = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

and the  $[p(N+1)+1]n \times [p(N+1)+1]n$ -matrix  $H^N$  by

$$H^N = \left( \begin{array}{c|ccc|ccc| \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ A_0 & A_{10}^N & \dots & A_{1N}^N + A_1 & A_{20}^N & \dots & A_{2N}^N + A_2 & \dots & \dots & A_{p0}^N & \dots & A_{pN}^N + A_p \\ \hline I & & & & & & & & & & & \\ \hline & h^N \otimes I & & & & & & & & & & \\ \hline & & & I & & & & & & & & \\ & & & & h^N \otimes I & & & & & & & \\ & & & & & & I & & & & & \\ & & & & & & & & I & & & \\ \hline & 0 & & & & & & & & I & & h^N \otimes I \end{array} \right)$$

where  $I$  is the  $n \times n$  identity matrix and

$$A_{kj}^N = \int_{-h}^0 A_{01}(\tau) e_{kj}^N(\tau) d\tau, \quad k = 1, \dots, p, \quad j = 0, \dots, N.$$

Lemma 4.12. a) The matrix representation  $[A^N]$  of  $A^N$  is given by

$$[A^N] = (Q^N)^{-1}H^N, \quad N = 1, 2, \dots$$

b) The matrix representation  $[(A^N)^*]$  of  $(A^N)^*$  is given by

$$[(A^N)^*] = (Q^N)^{-1}(H^N)^T.$$

c) The matrix representation of  $B$  as an operator  $\mathbb{R}^l \rightarrow X^N$  is given by

$$[B^N] = \text{col}(B_0, 0, \dots, 0) \in \mathbb{R}^{[(N+1)p+1] \times l}.$$

Proof.  $[A^N]$  is characterized by

$$\alpha^N(A^N \phi) = [A^N] \alpha^N(\phi), \quad \phi \in X^N.$$

On the other hand we get from (4.31) and  $\phi = \hat{E}^N \alpha^N(\phi)$

$$\begin{aligned} \alpha^N(A^N \phi) &= (Q^N)^{-1} d^N(A^N \phi) = (Q^N)^{-1} \langle \hat{E}^N, A^N \phi \rangle \\ &= (Q^N)^{-1} \langle \hat{E}^N, A^N \hat{E}^N \rangle \alpha^N(\phi), \end{aligned}$$

i.e.

$$[A^N] = (Q^N)^{-1} \langle \hat{E}^N, A^N \hat{E}^N \rangle.$$

We only have to show

$$\begin{aligned}
H^N &= \langle \hat{E}^N, A^N \hat{E}^N \rangle \\
&= \begin{pmatrix} \langle \hat{e}_0^N, A^N \hat{e}_0^N \rangle & \langle \hat{e}_0^N, A^N \hat{e}_{10}^N \rangle & \dots & \langle \hat{e}_0^N, A^N \hat{e}_{pN}^N \rangle \\ \langle \hat{e}_{10}^N, A^N \hat{e}_0^N \rangle & \langle \hat{e}_{10}^N, A^N \hat{e}_{10}^N \rangle & \dots & \langle \hat{e}_{10}^N, A^N \hat{e}_{pN}^N \rangle \\ \vdots & \vdots & & \vdots \\ \langle \hat{e}_{pN}^N, A^N \hat{e}_0^N \rangle & \langle \hat{e}_{pN}^N, A^N \hat{e}_{10}^N \rangle & \dots & \langle \hat{e}_{pN}^N, A^N \hat{e}_{pN}^N \rangle \end{pmatrix}
\end{aligned}$$

From Definition 4.5 we get

$$\begin{aligned}
A^N \hat{e}_0^N &= (A_0, 0) + \delta_{0,-}^N, \\
A^N \hat{e}_{k0}^N &= (A_{k0}^N, p_1^N(\dot{e}_{k0}^N)) - \delta_{k-1,-}^N, \quad k = 1, \dots, p, \\
A^N \hat{e}_{kj}^N &= (A_{kj}^N, p_1^N(\dot{e}_{kj}^N)), \quad k = 1, \dots, p, \quad j = 1, \dots, N-1, \\
A^N \hat{e}_{kN}^N &= (A_{kN}^N + A_{k,p}^N(\dot{e}_{kN}^N)) + \delta_{k,-}^N, \quad k = 1, \dots, p-1, \\
A^N \hat{e}_{pN}^N &= (A_{pN}^N + A_{p,p}^N(\dot{e}_{pN}^N)).
\end{aligned}$$

Observing

$$\begin{aligned}
\langle e_{kj, p_1}^N(\dot{e}_{mi}^N) \rangle_{L^2} &= \langle e_{kj, \dot{e}_{mi}^N}^N \rangle_{L^2}, \\
\langle \hat{e}_{kj, \delta_{i,-}^N}^N \rangle &= \lim_{\tau \rightarrow -h_i} e_{kj}^N(\tau) \quad (\text{cf. Lemma 4.4, a})
\end{aligned}$$

and the definition of  $\hat{e}_0^N$  and  $\hat{e}_{kj}^N$ , we get the desired result through straight-forward calculation.

In order to prove the representation for  $[(A^N)^*]$  we use (4.32) and get

$$\begin{aligned}
\alpha^N(\phi)^T Q^N [(A^N)^*] \alpha^N(\psi) &= \alpha^N(\phi)^T Q^N \alpha^N((A^N)^* \psi) \\
&= \langle \phi, (A^N)^* \psi \rangle = \langle A^N \phi, \psi \rangle \\
&= \alpha^N(A^N \phi)^T Q^N \alpha^N(\psi) = \alpha^N(\phi)^T [A^N]^T Q^N \alpha^N(\psi) \\
&= \alpha^N(\phi)^T (H^N)^T \alpha^N(\psi)
\end{aligned}$$

i.e.

$$[(A^N)^*] = (Q^N)^{-1} (H^N)^T.$$

Finally we have for any  $u \in \mathbb{R}^k$

$$\begin{aligned}
\alpha^N(Bu) &= \alpha^N((B_0 u, 0)) = (Q^N)^{-1} d^N((B_0 u, 0)) \\
&= (Q^N)^{-1} \text{col}(B_0 u, 0, \dots, 0) \\
&= \text{col}(B_0, 0, \dots, 0) u
\end{aligned}$$

which proves the given form of  $[B^N]$   $\square$

In order to make use of Theorem 4.3 we have to solve the Riccati differential equation (4.34) in the subspaces  $X^N$ . If in the optimal feedback law (3.18) for the delay system (2.1) we use  $\Pi^N(t)$  instead of  $\Pi(t)$  we get suboptimal controls  $\hat{u}^N(t)$  which by Theorem 4.3,b) converge to the optimal control  $\bar{u}(t)$ . The corresponding solution  $\hat{x}^N(t) \in \mathbb{R}^n$  of the RFDE (2.1) satisfies

$$\frac{d}{dt} \hat{x}^N(t) = L(\hat{x}_t^N) + B_0 \hat{u}^N(t),$$

where  $\hat{u}^N(t)$  is given by the feedback law

$$\hat{u}^N(t) := -R^{-1} B^* \Pi^N(t) p^N(\hat{x}^N(t), \hat{x}_t^N), \quad t \geq 0. \quad (4.37)$$

Taking matrix representations for the operators involved equation (4.34) takes the following form:

$$\begin{aligned} \frac{d}{dt} [\pi^N(t)] + [(A^N)^*] [\pi^N(t)] + [\pi^N(t)] [A^N] \\ - [\pi^N(t)] [B^N] R^{-1} [B^N]^T [\pi^N(t)] + [C^N]^T [C^N] \end{aligned} \quad (4.38)$$

$$= 0, \quad 0 \leq t \leq T,$$

$$[\pi^N(T)] = [G^N].$$

Here  $[\pi^N(t)]$ ,  $[C^N]$  and  $[G^N]$  denote the matrix representations of the restrictions of  $\pi^N(t)$ ,  $C$  and  $G$  to  $X^N$  considered as operators  $X^N \rightarrow X^N$ , respectively. From the definition of  $C$  and  $G$  we immediately see

$$[C^N] = (C_0, 0, \dots, 0) \in \mathbb{R}^{m \times [(N+1)p+1]n}$$

and

$$[G^N] = \begin{pmatrix} G_0 & 0 & \text{---} & 0 \\ 0 & 0 & \text{---} & 0 \\ | & | & & | \\ 0 & 0 & \text{---} & 0 \end{pmatrix} \in \mathbb{R}^{[(N+1)p+1]n \times [(N+1)p+1]n}.$$

The transformation

$$\Gamma^N(t) = Q^N [\pi^N(T-t)], \quad 0 \leq t \leq T,$$

puts problem (4.38) into the form

$$\begin{aligned} \frac{d}{dt} \Gamma^N &= [A^N]^T \Gamma^N + \Gamma^N [A^N] \\ &- \Gamma^N [B^N] R^{-1} [B^N]^T \Gamma^N + [C^N]^T [C^N], \end{aligned} \quad (4.39)$$

$$0 \leq t \leq T,$$



$$\Gamma^N(0) = [G^N].$$

Note, that  $[(A^N)^*] = (Q^N)^{-1} [A^N]^T Q^N$ . Equation (4.39) is the standard Riccati matrix differential equation ( $\Pi^N(t) = \Pi^N(t)^*$  implies  $\Gamma^N(t)^T = \Gamma^N(t)$ ) and can be solved numerically by standard methods. In many cases a method developed by Casti and Kailath (see for instance [34], p. 304 ff.) can be used advantageously. In the case  $p = 1$  and  $A_{01} = 0$ , for instance, we define

$$W_0 = A_0^T G_0 + G_0 A_0 - G_0 B_0 R^{-1} B_0^T G_0 + C_0^T C_0$$

and the  $2n \times (N+2)n$  matrices

$$F_1^N = \begin{pmatrix} W_0 & 0 & \text{---} & 0 & G_0 A_1 \\ A_1^T G_0 & 0 & \text{---} & \text{---} & 0 \end{pmatrix},$$

$$F_2^N = \begin{pmatrix} I & 0 & \text{---} & \text{---} & 0 \\ 0 & \text{---} & \text{---} & 0 & I \end{pmatrix}.$$

Note that  $(F_1^N)^T F_2^N = \dot{\Gamma}^N(0)$ . Then

$$\Gamma^N(t) = [G^N] + \int_0^t L_1^N(\tau)^T L_2^N(\tau) d\tau, \quad (4.40)$$

where

$$\frac{d}{dt} L_i^N(t) = L_i^N(t) ([A^N] - [B^N] R^{-1} [B^N]^T \Gamma^N(t))$$

$$L_i^N(0) = F_i^N, \quad i = 1, 2. \quad (4.41)$$

Note, that this is a system of  $4n^2(N+2)$  differential equations compared to the  $n^2(N+2)^2$  differential equations of system (4.39) (in case  $p = 1$  and  $A_{01} \equiv 0$ ).

Finally, let us rewrite the approximate control law (4.37) in terms of the above matrices. For this sake let us introduce the real  $n \times n$ -matrices  $\Pi_0^N(t), \Pi_{10}^N(t), \dots, \Pi_{pN}^N(t)$  by

$$[\Pi^N(t)] = \begin{pmatrix} \Pi_0^N(t) & * & \dots & * \\ \Pi_{10}^N(t) & & & \\ \vdots & & & \\ \Pi_{pN}^N(t) & * & \dots & * \end{pmatrix} \quad (4.42)$$

and define

$$\Pi_1^N(t, \tau) = \sum_{k=1}^p \sum_{j=0}^N \Pi_{kj}^N(t) e_{kj}^T(\tau) \quad (4.43)$$

for  $-h \leq \tau \leq 0$  and  $0 \leq t \leq T$ . Moreover, recall that  $[\Pi^N(t)]^T Q^N = Q^N [\Pi^N(t)]$ . Then the control law (4.37) takes the following form

$$\begin{aligned} \hat{u}^N(t) &= \\ &= -R^{-1} [B^N]^T [\Pi^N(t)] (Q^N)^{-1} d^N((\hat{x}^N(t), \hat{x}_t^N)) \\ &= -R^{-1} (B_0^T, 0, \dots, 0) [\Pi^N(t)]^T d^N((\hat{x}^N(t), \hat{x}_t^N)) \\ &= -R^{-1} B_0^T \{ \Pi_0^N(t) \hat{x}^N(t) + \\ &\quad + \sum_{k=1}^p \sum_{j=0}^N \Pi_{kj}^N(t) \int_{-h}^0 e_{kj}^N(\tau) \hat{x}^N(t+\tau) d\tau \} \\ &= -R^{-1} B_0^T \{ \Pi_0^N(t) \hat{x}^N(t) + \int_{-h}^0 \Pi_1^N(t, \tau) \hat{x}^N(t+\tau) d\tau \} \end{aligned} \quad (4.44)$$

## 5. Structure and stability

### 5.1. The structural operator

In Section 2 we have seen that the structural operator  $F: M^2 \rightarrow M^2$  plays an important role for the state space description of retarded systems. In this section we introduce an analogous operator for the description of the approximating systems  $(E^N)$  and  $(E_T^N)$ . The first step in this direction is

Lemma 5.1. Suppose that  $r_1 = \dots = r_p$  and  $A_{01}(\tau) \equiv 0$ . Then  $X^N$  is invariant under the operator  $F: M^{2p} \rightarrow M^{2p}$  introduced in Section 2.1. Moreover for every  $\phi \in M^2$

$$p^N F \phi = F^N p^N \phi,$$

where  $F^N: X^N \rightarrow X^N$  has the following matrix representation with respect to the basis  $\hat{E}^N$ :

$$[F^N] = \begin{pmatrix} I & 0 & \dots & 0 \\ 0 & a_1^n & \dots & a_p^n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_p^N & \dots & 0 \end{pmatrix}, \quad a_j^N = \begin{pmatrix} 0 & \dots & 0 & A_j \\ \vdots & \ddots & \vdots & \vdots \\ C & \dots & \dots & \dots \\ A_j & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{(N+1)n \times (N+1)n}$$

Proof. It is easy to see that under the assumptions of the lemma we have (cf. (2.9))

$$F \hat{e}_0^N = \hat{e}_0^N$$

and

$$F \hat{e}_{k,j}^N = \sum_{i=k}^p A_i \hat{e}_{i-k+1, N-j}^N, \quad k = 1, \dots, p, \quad j = 0, \dots, N \quad \square$$



- (i)  $F^N e^{A^N t} = e^{(A_T^N)^* t} F^N, \quad t \geq 0.$   
(ii)  $F^N A^N = (A_T^N)^* F^N.$   
(iii)  $F^N B^N = B^{N T}, \quad C^N F^N = C^N.$  Here  $B^N$  and  $C^N$  are defined by  $B^N \xi = B \xi,$   
 $\xi \in \mathbb{R}^2,$  and  $C^N = C|X^N.$

Proof. Recall that  $[A^N] = (Q^N)^{-1} H^N$  and  $[(A_T^N)^*] = (Q^N)^{-1} (H_T^N)^T.$   
Hence Lemma 5.2 shows

$$\begin{aligned} [F^N][A^N] &= [F^N](Q^N)^{-1} H^N = (Q^N)^{-1} [F^N] H^N \\ &= (Q^N)^{-1} (H_T^N)^T [F^N] = [(A_T^N)^*][F^N]. \end{aligned}$$

This proves (ii). Statement (i) follows directly from (ii) and (iii) is trivial  $\square$

The above results indicate that - to a certain extent - the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  show the same structural relation as the original systems  $(\Sigma)$  and  $(\Sigma_T^*).$  In particular, if  $r_1 = \dots = r_p$  and  $A_{10}(\tau) \equiv 0,$  then for every solution

$$w^N(t) = e^{A^N t} w^N(0) + \int_0^t e^{A^N(t-\tau)} B^N u(\tau) d\tau, \quad t \geq 0,$$

of  $(\Sigma^N)$  the function

$$z^N(t) = F^N w^N(t)$$

is the solution of  $(\Sigma_T^{N*})$  with initial value  $F^N w^N(0).$  The consequences of statement (ii) in Corollary 5.3 for the Riccati equations corresponding to  $(\Sigma^N)$  and  $(\Sigma_T^{N*}),$  respectively, have been discussed in Section 4.3 (see (4.36)).

## 5.2. Criteria for stability, stabilizability and controllability

In this section we examine some basic structural properties of the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ . We shall need the following facts on the real  $(N+1) \times (N+1)$ -matrix  $(q^N)^{-1}h^N$ .

Lemma 5.4. a) Let  $\|\cdot\|_N$  be the operator norm which corresponds to the vector norm  $|x|_N^2 = \bar{x}^T q^N x$  on  $\mathbb{C}^{N+1}$ . Then

$$\|e^{(q^N)^{-1}h^N t}\|_N \leq 1$$

for  $N = 1, 2, \dots$  and  $t \geq 0$ .

b) Let  $\mu \in \sigma((q^N)^{-1}h^N)$  and  $x = \text{col}(x_0, \dots, x_N) \in \mathbb{C}^{N+1}$ ,  $x \neq 0$ , such that either  $(\mu q^N - h^N)x = 0$  or  $(\mu q^N - (h^N)^T)x = 0$ . Then  $x_0 \neq 0$  and  $x_N \neq 0$ .

c) Let  $\mu \in \sigma((q^N)^{-1}h^N)$  and  $x = \text{col}(1, 0, \dots, 0)$  or  $x = \text{col}(0, \dots, 0, 1)$ . Then  $x \notin \text{range}(\mu q^N - h^N)$  and  $x \notin \text{range}(\mu q^N - (h^N)^T)$ .

d)  $\text{Re } \mu < 0$  for every  $\mu \in \sigma((q^N)^{-1}h^N)$ .

Proof. a) For every  $x \in \mathbb{C}^{N+1}$  the following equation holds:

$$\begin{aligned} \text{Re}(\bar{x}^T h^N x) &= (\text{Re } x)^T h^N (\text{Re } x) + (\text{Im } x)^T h^N (\text{Im } x) \\ &= -\frac{1}{2}|x_0|^2 - \frac{1}{2}|x_N|^2. \end{aligned} \quad (5.1)$$

Hence  $(q^N)^{-1}h^N$  is a dissipative operator on  $\mathbb{C}^{N+1}$  with respect to the inner product

$$\langle y, x \rangle_N = \bar{y}^T q^N x, \quad x, y \in \mathbb{C}^{N+1}.$$

Therefore  $\exp((q^N)^{-1}h^N t)$ ,  $t \geq 0$ , is a contraction semigroup on  $\mathbb{C}^{N+1}$  supplied with the norm  $|\cdot|_N$  (see for instance [32]).

b) This follows from

$$\mu q^N - h^N = \begin{bmatrix} \frac{\mu}{3} + \frac{1}{2} & \frac{\mu}{6} + \frac{1}{2} & 0 & 0 \\ \frac{\mu}{6} - \frac{1}{2} & \frac{2\mu}{3} & 0 & 0 \\ 0 & 0 & \frac{2\mu}{3} & \frac{\mu}{6} + \frac{1}{2} \\ 0 & 0 & \frac{\mu}{6} - \frac{1}{2} & \frac{\mu}{3} + \frac{1}{2} \end{bmatrix}$$

and the fact that  $\mu = \pm 3$  is not an eigenvalue of  $(q^N)^{-1}h^N$ .

c)  $x \in \text{range}(\mu q^N - h^N)$  would imply  $x \perp \ker(\mu q^N - (h^N)^T)$  which is impossible by b).

d) Assume that  $\mu \in \sigma((q^N)^{-1}h^N)$  and  $\text{Re } \mu \geq 0$ . Then there exists an  $x \in \mathbb{C}^{N+1}$ ,  $x \neq 0$ , such that  $(\mu q^N - h^N)x = 0$ . By (5.1) this implies

$$\begin{aligned} 0 &\leq (\text{Re } \mu) \bar{x}^T q x = \text{Re}(\bar{x}^T h^N x) \\ &= -\frac{1}{2}|x_0|^2 - \frac{1}{2}|x_N|^2. \end{aligned}$$

Hence  $x_0 = x_N = 0$  and therefore  $x = 0$  by b) in contradiction to  $x \neq 0$   $\square$

For every  $\mu \in \mathbb{C}$  which is not in the spectrum of  $(q^N)^{-1}h^N$  (in particular for every  $\mu$  in the closed right half plane) we introduce the vector

$$a^N(\mu) = \text{col}(a_0^N(\mu), \dots, a_N^N(\mu))$$

as the unique solution of

$$(\mu q^N - h^N) a^N(\mu) = \text{col}(1, 0, \dots, 0). \quad (5.2)$$

Moreover, we define

$$\alpha_{j,k}^N(\lambda) = \alpha_k^N(\lambda \frac{r_j}{N}) \alpha_N^N(\lambda \frac{r_{j-1}}{N}) \dots \alpha_N^N(\lambda \frac{r_1}{N}), \quad (5.3)$$

$j = 1, \dots, p, \quad k = 0, \dots, N.$  Then the complex  $n \times n$ -matrix

$$\Delta^N(\lambda) = \lambda I - A_0 - \sum_{j=1}^p [A_j \alpha_{j,N}^N(\lambda) + \sum_{k=0}^N A_{j,k}^N \alpha_{j,k}^N(\lambda)] \quad (5.4)$$

plays an analogous role for the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  as the characteristic matrix  $\Delta(\lambda)$  does for the original systems  $(\Sigma)$  and  $(\Sigma_T^*)$ . In particular it determines their input-output behaviour.

Proposition 5.5. a) The left upper  $n \times n$  block  $X^N(t)$  in the matrix  $e^{[A^N]t}$  coincides with that of the matrix  $e^{[(A_T^N)^*]t}$  and its Laplace transform is given by  $\Delta^N(\lambda)^{-1}$ .

b) Let  $w^N(t) = \text{col}(w_0^N(t), \dots)$  and  $z^N(t) = \text{col}(z_0^N(t), \dots)$  be the unique solutions of  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ , respectively, with initial state zero. Then

$$w_0^N(t) = z_0^N(t) = \int_0^t X^N(t-s) B_0 u(s) ds, \quad t \geq 0.$$

c) The transfer matrices of  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  coincide and are given by

$$G^N(\lambda) = C_0 \Delta^N(\lambda)^{-1} B_0.$$

Proof. a) It is sufficient to show that the left upper  $n \times n$  block in  $(\lambda I - [A^N])^{-1}$  respectively in  $(\lambda I - [(A_T^N)^*])^{-1}$  is given by  $\Delta^N(\lambda)^{-1}$ . Let  $\lambda \in \mathbb{C}$  be such that  $\lambda \frac{r_j}{N} \notin \sigma((Q^N)^{-1} h^N)$  for  $j = 1, \dots, p$  and choose  $z = \text{col}(z_0, 0, \dots, 0) \in \mathbb{C}^{n+p(N+1)n}$  where  $z_0 \in \mathbb{C}^n$ . Then  $(\lambda I - [A^N])x = z$  if and only if  $(\lambda Q^N - H^N)x = z$  or equivalently

$$(\lambda I - A_0)x_0 - \sum_{j=1}^p [A_j x_{j,N} + \sum_{k=0}^N A_{j,k}^N x_{j,k}] = z_0, \quad (5.5;1)$$

$$((\lambda \frac{r_j}{N} Q^N - h^N) \otimes I)x_j = \text{col}(x_{j-1,N}, 0, \dots, 0), \quad (5.5;2)$$

$j = 1, \dots, p.$



Here  $x = \text{col}(x_0, x_1, \dots, x_p)$ , where  $x_0 \in \mathbb{C}^n$ ,  $x_j = \text{col}(x_{j,0}, \dots, x_{j,N}) \in \mathbb{C}^{(N+1)n}$  with  $x_{j,k} \in \mathbb{C}^n$  and  $x_{0,N} = x_0$ . By definition of the  $\alpha_{j,k}^N(\lambda)$  equations (5.5) are equivalent to

$$\Delta^N(\lambda)x_0 = z_0, \quad (5.6;1)$$

$$x_{j,k} = \alpha_{j,k}^N(\lambda)x_0, \quad j = 1, \dots, p, \quad k = 0, \dots, N. \quad (5.6;2)$$

This proves the statement on the matrix  $(\lambda I - [A^N])^{-1}$ . Furthermore,  $(\lambda I - [(A_T^N)^*])x = z$  if and only if  $(\lambda Q^N - (H_T^N)^T)x = z$  or equivalently

$$(\lambda I - A_0)x_0 - x_{1,0} = z_0, \quad (5.7;1)$$

$$((\lambda \frac{r_i}{N} q^N - (h^N)^T) \otimes I)x_j = \quad (5.7;2)$$

$$\text{col}(A_{j,0}^N x_0, \dots, A_{j,N-1}^N x_0, (A_j + A_{j,N}^N)x_0 + x_{j+1,0})$$

$$j = 1, \dots, p,$$

where  $x_{p+1,0} := 0$ . Since the first row of the matrix  $(\mu q^N - (h^N)^T)^{-1}$  is given by  $(\alpha_0^N(\mu), \dots, \alpha_N^N(\mu))$  we see that (5.7) is equivalent to (5.7;2) and

$$\Delta^N(\lambda)x_0 = z_0, \quad (5.8;1)$$

$$x_{j,0} = \sum_{i=j}^p [A_i \alpha_N^N(\lambda \frac{r_i}{N}) + \sum_{k=0}^N A_{i,k}^N \alpha_k^N(\lambda \frac{r_i}{N})] x_0 \alpha_N^N(\lambda \frac{r_{i-1}}{N}) \dots \alpha_N^N(\lambda \frac{r_j}{N}), \quad (5.8;2)$$

$$j = 1, \dots, p.$$

This finishes the proof of a).

Statements b) and c) follow directly from a)  $\square$

The following characterization of stabilizability and detectability for the approximating systems  $(\Gamma^N)$  and  $(\Gamma_T^N)$  is precisely the analogon to Theorems 2.3 and 2.4.

**Theorem 5.6.** a) The matrix  $[A^N]$  (or equivalently  $[(A_T^N)^*]$ ) is stable if and only if  $\det \Delta^N(\lambda) \neq 0$  for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ .  
 b) The system  $(\Sigma^N)$  (or equivalently  $(\Sigma_T^{N*})$ ) is stabilizable if and only if

$$\operatorname{rank} [\Delta^N(\lambda), B_0] = n$$

for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ .

c) The system  $(\Sigma^N)$  (or equivalently  $(\Sigma_T^{N*})$ ) is detectable if and only if

$$\operatorname{rank} \begin{bmatrix} \Delta^N(\lambda) \\ C_0 \end{bmatrix} = n$$

for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ .

Proof. It is well known from finite dimensional linear system theory that  $(\Sigma^N)$  is detectable if and only if  $\ker(\lambda I - [A^N]) \cap \ker[C^N] = \{0\}$  or equivalently

$$\ker(\lambda Q^N - H^N) \cap \ker[C^N] = 0$$

for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ . But  $\lambda \frac{r_j}{N} \notin \sigma((Q^N)^{-1} h^N)$  for  $\operatorname{Re} \lambda \geq 0$  (Lemma 5.4,d) and  $x = \operatorname{col}(x_0, x_1, \dots, x_p) \in \ker(\lambda Q^N - H^N)$  is equivalent to  $\Delta^N(\lambda)x_0 = 0$  and (5.6;2). Therefore detectability of  $(\Sigma^N)$  is equivalent to

$$\ker \Delta^N(\lambda) \cap \ker C_0 = \{0\}$$

for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ . Analogously it follows from (5.8) that this condition is also equivalent to detectability of  $(\Sigma_T^{N*})$ .

Statement b) follows from c) by duality and statement a) is a special case of c) (put  $C_0 = 0$ )  $\square$

In the special situation considered in Section 5.1 we can also characterize controllability and observability of systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  (compare Theorem 2.5):

**Theorem 5.7.** Suppose that  $r_1 = \dots = r_p = r$  and that  $A_{01}(\tau) \equiv 0$ .  
a) Let  $\lambda \in \mathbb{C}$  be such that  $\lambda \frac{r}{N} \notin \sigma((q^N)^{-1}h^N)$ . Then  $\lambda \in \sigma(A^N)$  if and only if  $\det \Delta^N(\lambda) = 0$ . If  $\lambda \frac{r}{N} \in \sigma((q^N)^{-1}h^N)$ , then  $\lambda \in \sigma(A^N)$  if and only if  $\det A_p = 0$ . Moreover,  $\sigma((A_T^N)^*) = \sigma(A^N)$ .  
b) System  $(\Sigma^{Np})$  is controllable if and only if

$$\text{rank}[\Delta^N(\lambda), B_0] = n \text{ for all } \lambda \in \mathbb{C} \setminus \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$$

and

$$\text{rank}[A_p, B_0] = n.$$

c) System  $(\Sigma^N)$  is observable if and only if

$$\text{rank} \begin{bmatrix} \Delta^N(\lambda) \\ C_0 \end{bmatrix} = n \text{ for all } \lambda \in \mathbb{C} \setminus \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$$

and

$$\text{rank } A_p = n.$$

d) System  $(\Sigma_T^{N*})$  is controllable if and only if

$$\text{rank}[\Delta^N(\lambda), B_0] = n \text{ for all } \lambda \in \mathbb{C} \setminus \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$$

and

$$\text{rank } A_p = n.$$

e) System  $(\Sigma_T^{N*})$  is observable if and only if

$$\text{rank} \begin{bmatrix} \Delta^N(\lambda) \\ C_0 \end{bmatrix} = n \text{ for all } \lambda \in \mathbb{C} \setminus \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$$

and

$$\text{rank} \begin{pmatrix} A_p \\ C_0 \end{pmatrix} = n.$$

Proof. Let  $\lambda \in \mathbb{C} \setminus \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$  or equivalently  $\lambda \frac{r}{N} \notin \sigma\left((q^N)^{-1}h^N\right)$ . Then (5.6) and (5.8) imply that both,  $\ker(\lambda I - A^N) \cap \ker C^N = \{0\}$  and  $\ker(\lambda I - (A_T^N)^*) \cap \ker C^N = \{0\}$ , are equivalent to  $\ker \Delta^N(\lambda) \cap \ker C_0 = \{0\}$ , i.e. to  $\text{rank} \begin{pmatrix} \Delta^N(\lambda) \\ C_0 \end{pmatrix} = n$ .

Now let  $\lambda \in \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$ . Since  $A_{01} = 0$ , we have  $A_{j,k}^N = 0$ ,  $j = 1, \dots, p$ ,  $k = 0, \dots, N$ . Assume first that  $\text{rank } A_p = n$  and let  $x \in \mathbb{C}^{n+p(N+1)n}$  satisfy  $(\lambda q^N - h^N)x = 0$ ,  $[C^N]x = 0$ . Then  $x$  satisfies (5.5) with  $z_0 = 0$ . Lemma 5.4,c) implies  $x_0 = x_{j,N} = 0$ ,  $j = 1, \dots, p-1$ . By (5.5;1) we have  $A_p x_{p,N} = 0$ , i.e. also  $x_{p,N} = 0$ . Now it follows from Lemma 5.4,b) that  $x_j = 0$  for  $j = 1, \dots, p$  and thus  $x = 0$ .

Conversely, assume that  $\ker(\lambda I - A^N) \cap \ker C^N = \{0\}$  and take  $\xi \in \ker A_p$ . Then Lemma 5.4,b) implies that there exists a vector  $x_p = \text{col}(x_{p,0}, \dots, x_{p,N}) \in \mathbb{C}^{(N+1)n}$  such that  $(\left(\lambda \frac{r}{N}q^N - h^N\right) \oplus I)x_p = 0$  and  $x_{p,N} = \xi$ . If we define  $x_0 = 0$  and  $x_j = 0$  for  $j = 1, \dots, p-1$ , then it follows for  $x = \text{col}(x_0, x_1, \dots, x_p)$  from (5.5) (with  $z_0 = 0$ ) that  $(\lambda q^N - h^N)x = 0$ ,  $C^N x = 0$ . By assumption this implies  $x = 0$  and thus  $\xi = 0$ . We conclude  $\text{rank } A_p = n$ . This finishes the proof of statement c).

Still let  $\lambda \in \sigma\left(\left(\frac{r}{N}q^N\right)^{-1}h^N\right)$ . Assume that  $\ker A_p \cap \ker C_0 = \{0\}$  and let  $x \in \mathbb{C}^{n+p(N+1)n}$  satisfy  $[\lambda q^N - (h_T^N)^T]x = 0$ ,  $[C^N]x = 0$ . Then (5.7;2) with  $j = p$  and Lemma 5.4,c) imply  $A_p x_0 = 0$ . This together with  $C_0 x_0 = 0$  shows  $x_0 = 0$ . Hence it follows by repeated use of (5.7;2) and Lemma 5.4,c) that  $x_{j,0} = 0$ ,  $j = 1, \dots, p$ . Finally we get from Lemma 5.4,b) that  $x_j = 0$ ,  $j = 1, \dots, p$ , and thus  $x = 0$ , i.e.  $\ker(\lambda I - (A_T^N)^*) \cap \ker C^{Nj} = \{0\}$ .

Conversely, suppose that  $\ker(\lambda I - (A_T^N)^*) \cap \ker C^N = \{0\}$  and let  $x_0 \in \ker A_p \cap \ker C_0$ . By Lemma 5.4,b) there exists a vector  $\alpha = \text{col}(\alpha_0, \dots, \alpha_N) \in \mathbb{C}^{N+1}$  such that

$$(\lambda \frac{r}{N} q^N - (h^N)^T) \alpha = 0 \quad \text{and} \quad \alpha_0 = 1.$$

We define  $x_j = \text{col}(x_{j,0}, \dots, x_{j,N}) \in \mathbb{C}^{n(N+1)}$  by

$$x_{1,k} = \alpha_k (\lambda I - A_0) x_0$$

$$x_{j,k} = -\alpha_k A_{j-1} x_0$$

for  $j = 2, \dots, p$  and  $k = 0, \dots, N$ . Then it follows from (5.7) that  $x = \text{col}(x_0, x_1, \dots, x_p) \in \ker(\lambda Q^N - (H^N)^T) \cap \ker[C^N]$ . By assumption this implies  $x = 0$  and hence  $x_0 = 0$ , i.e.  $\ker A_p \cap \ker C_0 = \{0\}$ . Thus statement e) is proved.

Statements b) and d) follow from e) and c) by duality. The proof of statement a) is the same as for c) with  $C^N = 0$  resp.  $C_0 = 0$   $\square$

Remark. Using similar consideration as above we can establish the following partial generalizations of statement a) in Theorem

5.7: a) Suppose  $\lambda \frac{r}{N} k \notin \sigma((q^N)^{-1} h^N)$  for  $k = 1, \dots, p$ . Then  $\lambda \in \sigma(A^N)$  if and only if  $\det \Delta^N(\lambda) = 0$ .

b) Assume  $A_{01}(\tau) \equiv 0$ . If  $\det A_p = 0$  then  $\lambda \frac{r}{N} p \in \sigma((q^N)^{-1} h^N)$  implies  $\lambda \in \sigma(A^N)$ .

We close this section with a very special result that guarantees stability of  $(\Sigma^N)$  for all  $N$ .

Proposition 5.8. Let  $p = 1$  and assume  $A_{01}(\tau) \equiv 0$ . If  $\xi^T A_0 \xi < -\frac{1}{2} |\xi|^2 - \frac{1}{2} |A_1^T \xi|^2$  for all nonzero  $\xi \in \mathbb{R}^n$ , then  $\text{Re } \lambda < 0$  for every  $\lambda \in \sigma(A^N)$  and every  $N = 1, 2, \dots$ .

Proof. Let  $\lambda \in \sigma(A^N)$  for some  $N$  and suppose that  $\text{Re } \lambda \geq 0$ . Then there exists a nonzero  $x \in \mathbb{C}^{n(N+2)}$  such that  $(\lambda Q^N - H^N)x = 0$ . This implies (observing the special form of  $H^N$ )

$$\begin{aligned}
0 &\leq (\operatorname{Re} \lambda) \bar{x}^T Q^N x = \operatorname{Re}(\bar{x}^T H^N x) \\
&= \operatorname{Re}(\bar{x}_0^T A_0 x_0) + \operatorname{Re}(\bar{x}_0^T A_1 x_{1,N}) + \operatorname{Re} \bar{x}_0^T x_{1,0} - \frac{1}{2} |x_{1,0}|^2 - \frac{1}{2} |x_{1,N}|^2 \\
&\leq \operatorname{Re}(\bar{x}_0^T A_0 x_0) + \frac{1}{2} |x_0|^2 + \frac{1}{2} |A_1^T x_0|^2 < 0
\end{aligned}$$

if  $x_0 \neq 0$ . This contradiction shows  $x_0 = 0$ . Therefore  $\lambda \frac{h}{N}$  is an eigenvalue of  $(q^N)^{-1} h^N$ , i.e.  $\operatorname{Re} \lambda < 0$  by Lemma 5.4,d). This contradiction proves the result  $\square$

### 5.3. Convergence of $\Delta^N(\lambda)$

The results of the previous section illustrate the important role which the coefficients  $\alpha_k^N(\mu)$  play for the structural properties of the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ . So far we have treated them only as implicit parameters which are given by equation (5.2). In the following we derive explicit formulas for the  $\alpha_k^N(\mu)$  and use those in order to prove convergence of  $\Delta^N(\lambda)$  to  $\Delta(\lambda)$ .

**Lemma 5.9.** a) Let the rational functions  $d_k^N(\mu)$ ,  $k = 0, \dots, N$ , be defined recursively by

$$\begin{aligned}
d_0^N(\mu) &= 2\mu + 3, \\
d_k^N(\mu) &= 4\mu + \frac{9-\mu^2}{d_{k-1}^N(\mu)}, \quad k = 1, \dots, N-1, \\
d_N^N(\mu) &= 2\mu + 3 + \frac{9-\mu^2}{d_{N-1}^N(\mu)}.
\end{aligned} \tag{5.9}$$

Then  $\alpha_k^N(\mu)$  is given by

$$\alpha_k^N(\mu) = \frac{6(3-\mu)^k}{d_{N-k}^N(\mu) \cdots d_N^N(\mu)} \tag{5.10}$$

for  $k = 0, \dots, N$ .

b) Let the polynomials  $p_k(\mu)$ ,  $q_k(\mu)$  of degree  $k+1$  be defined recursively by

$$\begin{aligned} p_{-1}(\mu) &= 1, & p_0(\mu) &= 2\mu + 3 \\ p_k(\mu) &= 4\mu p_{k-1}(\mu) + (9-\mu^2)p_{k-2}(\mu), \\ q_k(\mu) &= (2\mu+3)p_{k-1}(\mu) + (9-\mu^2)p_{k-2}(\mu), \end{aligned} \quad (5.11)$$

$k = 1, 2, \dots$ . Then these polynomials are stable for all  $k$  and, moreover,  $\mu \in \sigma((q^N)^{-1}h^N)$  if and only if  $q_N(\mu) = 0$ . If  $\mu \notin \sigma((q^N)^{-1}h^N)$ , then

$$a_k^N(\mu) = \frac{6(3-\mu)^k p_{N-k-1}(\mu)}{q_N(\mu)} \quad (5.12)$$

for  $k = 0, \dots, N$ .

c) Let  $\mu \neq \pm i\sqrt{3}$ ,  $\mu \in \mathbb{C}$ , be given and let  $w$  satisfy

$$w^2 = 9 + 3\mu^2.$$

Furthermore put

$$\gamma_0 = 2\mu + w, \quad \gamma_1 = 2\mu - w.$$

Then  $\mu \in \sigma((q^N)^{-1}h^N)$  if and only if

$$(3+w)^2(\gamma_0)^N - (3-w)^2(\gamma_1)^N = 0.$$

If  $\mu \notin \sigma((q^N)^{-1}h^N)$  then

$$a_k^N(\mu) = 6(3-\mu)^k \frac{(3+w)(\gamma_0)^{N-k} - (3-w)(\gamma_1)^{N-k}}{(3+w)^2(\gamma_0)^N - (3-w)^2(\gamma_1)^N} \quad (5.13)$$

for  $k = 0, \dots, N$ .

Proof. a) Suppose that the functions  $d_k = d_k^N(\mu)$ ,  $k = 0, \dots, N$ , are given by (5.9) and define

$$b_k = \frac{-\beta - \mu}{d_{k-1}}, \quad c_k = \frac{\beta - \mu}{d_{k-1}}, \quad k = 1, \dots, N.$$

Then it is easy to see that

$$6(\mu q^N - h^N) = \begin{pmatrix} 1 & -b_N & 0 & \dots & 0 \\ 0 & & & & 0 \\ & & & & \\ & & & & -b_1 \\ 0 & \dots & 0 & & 1 \end{pmatrix} \begin{pmatrix} d_N & 0 & \dots & 0 \\ 0 & & & \\ & & & \\ & & & \\ 0 & \dots & 0 & & d_0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ -c_N & & & \\ & & & \\ & & & \\ 0 & \dots & 0 & & -c_1 \\ & & & & 1 \end{pmatrix}.$$

It is not difficult to calculate the inverse matrices. Since  $a^N(\mu)$  is the first column of  $(\mu q^N - h^N)^{-1}$  we conclude that

$$a_k^N(\mu) = \frac{6c_{N-k+1} \dots c_N}{d_N} = \frac{6(\beta - \mu)^k}{d_{N-k} \dots d_N},$$

$k = 0, \dots, N$ .

b) If the polynomials  $p_k(\mu)$  and  $q_k(\mu)$  are given by (5.11) and the rational function  $d_k^N(\mu)$  by (5.9), then we see by induction that

$$d_k^N(\mu) = p_k(\mu)/p_{k-1}(\mu), \quad k = 0, \dots, N-1,$$

and

$$d_N^N(\mu) = q_N(\mu)/p_{N-1}(\mu).$$

This implies

$$\begin{aligned} p_k(\mu) &= d_0^N(\mu) \dots d_k^N(\mu), \quad k = 0, \dots, N-1, \\ q_N(\mu) &= d_0^N(\mu) \dots d_N^N(\mu) = \det[6(\mu q^N - h^N)]. \end{aligned} \quad (5.14)$$



Hence it follows from Lemma 5.4,d) that  $q_N(\mu)$  is a stable polynomial.

The considerations in the proof of a) together with (5.14) show that  $p_k(\mu)$  is the determinant of the matrix

$$\begin{pmatrix} 4\mu & \mu+3 & 0 & \dots & 0 \\ \mu-3 & 4\mu & & & \\ 0 & & 4\mu & & \\ \vdots & & & \ddots & \\ 0 & & & & \mu-3 & 2\mu+3 \end{pmatrix} \in \mathbb{C}^{(k+1) \times (k+1)} .$$

Hence it follows from analogous arguments as in the proof of Lemma 5.4,d) that  $p_k(\mu)$  is stable. Finally, equation (5.12) is an immediate consequence of (5.10) and (5.14).

c) Choose  $\mu \in \mathbb{C}$ ,  $\mu \neq \pm i\sqrt{3}$ , and define  $w$ ,  $\gamma_0$  and  $\gamma_1$  as in statement c) of the lemma. Then  $\gamma_0 \neq \gamma_1$  and

$$\gamma_i^2 - 4\mu\gamma_i - (9-\mu^2) = 0, \quad i = 0,1.$$

Hence  $\gamma_0$  and  $\gamma_1$  are the characteristic roots of the difference equation in (5.11). This implies that

$$p_k(\mu) = \frac{3+w}{2w}(\gamma_0)^{k+1} - \frac{3-w}{2w}(\gamma_1)^{k+1}, \quad (5.15)$$

$k = -1, 0, \dots, N-1$ . Using  $\gamma_0\gamma_1 = \mu^2 - 9$  we get from (5.11) and (5.15)

$$\begin{aligned} q_N(\mu) &= (2\mu+3)p_{N-1}(\mu) + (9-\mu^2)p_{N-2}(\mu) \\ &= \frac{3+w}{2w}[(2\mu+3)(\gamma_0)^N + (9-\mu^2)(\gamma_0)^{N-1}] \\ &\quad - \frac{3-w}{2w}[(2\mu+3)(\gamma_1)^N + (9-\mu^2)(\gamma_1)^{N-1}] \\ &= \frac{(3+w)^2}{2w}(\gamma_0)^N - \frac{(3-w)^2}{2w}(\gamma_1)^N. \end{aligned} \quad (5.16)$$

This formula shows that  $\mu$  is in the spectrum of  $(q^N)^{-1}h^N$  if and only if  $(3+w)^2(\gamma_0)^N - (3-w)^2(\gamma_1)^N = 0$ . Moreover, it follows from (5.12), (5.15) and (5.16) that  $\alpha_k^N(\mu)$  is given by (5.13) if  $\mu \notin \sigma((q^N)^{-1}h^N)$  and  $\mu \neq \pm i\sqrt{3}$   $\square$

The explicit formulas in the previous lemma allow us to prove that the matrices  $\Delta^N(\lambda)$  actually converge to the characteristic matrix  $\Delta(\lambda)$  of the delay system.

Theorem 5.10. Let  $A_{01}(\tau) \equiv 0$ . Then

$$\Delta(\lambda) = \lim_{N \rightarrow \infty} \Delta^N(\lambda), \quad \lambda \in \mathbb{C},$$

the limit being uniform on bounded subsets of  $\mathbb{C}$ .

Proof. Fix  $\delta > 0$  and define  $w(\mu)$  by

$$w(\mu)^2 = 9 + 3\mu^2, \quad \operatorname{Re} w(\mu) \geq 0$$

for any  $\mu \in \mathbb{C}$  with  $|\mu| \leq \delta$ . Then the map  $\mu \rightarrow w(\mu)$  is continuous and  $w(\mu) + 3$  is non-zero on  $|\mu| \leq \delta$ . From  $w(\mu) - 3 = \frac{3\mu^2}{w(\mu)+3}$  we see that

$$|w(\mu) - 3| \leq |\mu|^2 \quad \text{if } |\mu| \leq \delta. \quad (5.17)$$

In the next step we shall prove that  $\alpha_N^N(\frac{\mu}{N})$  converges uniformly to  $e^{-\mu}$  on  $|\mu| \leq c$  as  $N \rightarrow \infty$ . To this end we use formula (5.13) for  $k = N$  and obtain with  $w = w(\frac{\mu}{N})$ ,  $N \geq c\delta^{-1}$

$$\alpha_N^N(\frac{\mu}{N})^{-1} = \frac{(3+w)^2}{12w} \left(\frac{w+2\frac{\mu}{N}}{3-\frac{\mu}{N}}\right)^N - \frac{(3-w)^2}{12w} \left(-\frac{w-2\frac{\mu}{N}}{3-\frac{\mu}{N}}\right)^N. \quad (5.18)$$

From (5.17) and  $\lim_{N \rightarrow \infty} w(\frac{\mu}{N}) = 3$  uniformly on  $|\mu| \leq c$  we see that

$$\lim_{N \rightarrow \infty} \frac{(3+w(\frac{\mu}{N}))^2}{12w(\frac{\mu}{N})} = 1 \quad \text{and} \quad \lim_{N \rightarrow \infty} \frac{(3-w(\frac{\mu}{N}))^2}{12w(\frac{\mu}{N})} = 0$$

uniformly on  $|\mu| \leq c$ . Moreover, we obtain from (5.17) also

$$\frac{w+2\frac{\mu}{N}}{3-\frac{\mu}{N}} = 1 + \frac{\mu}{N} + \frac{w-3+(\frac{\mu}{N})^2}{3-\frac{\mu}{N}} = 1 + \frac{\mu}{N} + o(\frac{1}{N^2})$$

and

$$\frac{w-2\frac{\mu}{N}}{3-\frac{\mu}{N}} = 1 - \frac{\mu}{3N} + \frac{w-3-\frac{1}{3}(\frac{\mu}{N})^2}{3-\frac{\mu}{N}} = 1 + \frac{\mu}{N} + o(\frac{1}{N^2})$$

as  $N \rightarrow \infty$  uniformly on  $|\mu| \leq c$ . These relations together with (5.18) show

$$\lim_{N \rightarrow \infty} \alpha_N^N (\frac{\mu}{N})^{-1} = e^\mu$$

uniformly on  $|\mu| \leq c$ . Finally, the theorem follows from equations (5.3) and (5.4)  $\square$

It is our goal to prove that stability (resp. stabilizability or detectability) of the original system  $(\Sigma)$  implies the corresponding property for the approximating system  $(\Sigma^N)$  provided  $N$  is sufficiently large. The first step in this direction is the characterization of these properties in Theorem 5.6 using the matrices  $\Delta^N(\lambda)$ . The second step is the convergence result for  $\Delta^N(\lambda)$  in Theorem 5.10. In addition to these results we need a priori bounds for the unstable eigenvalues of the matrices  $[A^N]$ . This problem will be considered in the next section.

#### 5.4. Uniform bounds and nonuniform stability

We first establish uniform bounds for the  $\alpha_N^N(\mu)$  in  $\text{Re } \lambda \geq 0$ .

Lemma 5.11. The estimate

$$|\alpha_N^N(\mu)| < 2$$

is valid for all  $\mu \in \mathbb{C}$  with  $\text{Re } \mu \geq 0$  and all  $N = 1, 2, \dots$ .

Proof. Since  $\alpha_N^N(\mu)$  is a proper rational function without poles in  $\text{Re } \mu \geq 0$  (cf. Lemma 5.9, b), it follows from the maximum principle for analytic functions that  $|\alpha_N^N(\mu)|$  achieves its maximum value in  $\text{Re } \mu \geq 0$  on the imaginary axis. Therefore we only have to prove  $|\alpha_N^N(i\omega)| \leq 2$  for all  $\omega \in \mathbb{R}$  and all  $N$ .

We first consider  $\mu \in i\mathbb{R}$  with  $|\mu| \geq \sqrt{3}$ . In this case we have

$$|d_k^N(\mu)| \geq |3 - \mu|, \quad k = 0, \dots, N-1, \quad (5.19;1)$$

and

$$|d_N^N(\mu)| \geq 3 \quad (5.19;2)$$

for all  $N$ . The first estimate is obviously satisfied for  $k = 0$ .  
Using

$$2|\mu| \geq (9 + |\mu|^2)^{1/2} = |3 - \mu|$$

we obtain from (5.9) assuming that the estimate is already established for  $k$

$$\begin{aligned}
|\operatorname{Im} d_{k+1}^N(\mu)| &= \left| 4 \operatorname{Im} \mu - \frac{9+|\mu|^2}{|d_k^N(\mu)|^2} \operatorname{Im} d_k^N(\mu) \right| \\
&\geq 4|\mu| - \frac{9+|\mu|^2}{|d_k^N(\mu)|} \geq 4|\mu| - (9+|\mu|^2)^{1/2} \\
&\geq (9+|\mu|^2)^{1/2}, \quad k = 0, \dots, N-2.
\end{aligned}$$

This proves (5.19;1). In order to prove (5.19;2) we note that  $\operatorname{Re} d_k^N(\mu)$  is always positive (and decreasing with respect to  $k$ ) since

$$\operatorname{Re} d_{k+1}^N(\mu) = \frac{9+|\mu|^2}{|d_k^N(\mu)|^2} \operatorname{Re} d_k^N(\mu), \quad k = 0, \dots, N-2.$$

Therefore the last equation in (5.9) implies

$$\operatorname{Re} d_N^N(\mu) = 3 + \frac{9+|\mu|^2}{|d_{N-1}^N(\mu)|^2} \operatorname{Re} d_{N-1}^N(\mu) \geq 3$$

which proves (5.19;2). Now it follows from (5.19) and (5.10) that

$$|\alpha_N^N(\mu)| = 2 \frac{|3-\mu|}{|d_0^N(\mu)|} \dots \frac{|3-\mu|}{|d_{N-1}^N(\mu)|} \cdot \frac{3}{|d_N^N(\mu)|} \leq 2.$$

It remains to consider  $\mu \in i\mathbb{R}$  with  $|\mu| < \sqrt{3}$ . Let the complex numbers  $w, \gamma_0, \gamma_1$  be defined as in Lemma 5.9,c). Then

$$w = (9 - 3|\mu|^2)^{1/2} \in \mathbb{R}$$

and hence

$$|\gamma_0| = |\gamma_1| = |3-\mu|. \quad (5.20)$$

This implies that there exists a  $\rho \in \mathbb{R}$  such that

$$\left(\frac{\gamma_1}{\gamma_0}\right)^N = e^{i\rho} \quad (5.21)$$

Using (5.20) and (5.21) we get from (5.13) that

$$\begin{aligned} |\alpha_N^N(\mu)|^2 &= 36 \left| \frac{3-\mu}{\gamma_0} \right|^N \cdot \frac{4w^2}{|(3+w)^2 - (3-w)^2 (\gamma_1/\gamma_0)^N|^2} \\ &= 144w^2 |(9+w^2)(1-e^{i\rho}) + 6w(1+e^{i\rho})|^{-2}. \end{aligned}$$

The identities

$$\frac{1+e^{i\rho}}{1-e^{i\rho}} = i \frac{\sin \rho}{1-\cos \rho} \quad \text{and} \quad |1-e^{i\rho}|^2 = 2-2\cos \rho$$

imply for any  $a, b \in \mathbb{R}$

$$\begin{aligned} &|a(1-e^{i\rho}) + b(1+e^{i\rho})|^2 \\ &= 2a^2(1-\cos \rho) + 2b^2(1+\cos \rho). \end{aligned}$$

Therefore

$$\begin{aligned} |\alpha_N^N(\mu)|^2 &= 72w^2 [(9+w^2)^2(1-\cos \rho) + 36w^2(1+\cos \rho)]^{-1} \\ &= \frac{36w^2}{(9+w^2)^2 \frac{1-\cos \rho}{2} + 36w^2 \frac{1+\cos \rho}{2}} \leq 1, \end{aligned}$$

because  $6w \leq 9+w^2$

Now we are in the position to prove the desired result on stabilizability and detectability for the approximating systems  $(\Sigma^N)$ .

**Theorem 5.12.** Suppose that  $A_{01}(\tau) \equiv 0$ . Then the following statements are true:

- a) If system  $(\Sigma)$  is stable, then there exists an  $N_0$  such that system  $(\Sigma^N)$  is stable for every  $N \geq N_0$ .
- b) If system  $(\Sigma)$  is stabilizable (respectively detectable)

then there exists an  $N_0$  such that system  $(\Sigma^N)$  is stabilizable (respectively detectable) for every  $N \geq N_0$ .

Proof. a) Suppose that  $(\Sigma)$  is stable. Then  $\det \Delta(\lambda) \neq 0$  on  $\text{Re } \lambda \geq 0$ . It follows from Lemma 5.11 and equation (5.3) that

$$\|A_0 + \sum_{j=1}^p A_j \alpha_{j,N}^N(\lambda)\| \leq \sum_{j=0}^p 2^j \|A_j\| =: \omega$$

for every  $\lambda \in \mathbb{C}$  with  $\text{Re } \lambda \geq 0$ . Since  $A_{j,k}^N = 0$  for  $j = 1, \dots, p$  and  $k = 0, \dots, N$ , we obtain from (5.4) that  $\det \Delta^N(\lambda) \neq 0$  for every  $N$  and every  $\lambda \in \mathbb{C}$  with  $\text{Re } \lambda \geq 0$  and  $|\lambda| > \omega$ . Finally, the uniform convergence result on bounded domains (Theorem 5.10) shows that  $\det \Delta^N(\lambda) \neq 0$  for every  $\lambda \in \mathbb{C}$  with  $\text{Re } \lambda \geq 0$  and  $|\lambda| \leq \omega$  provided  $N$  is sufficiently large. Hence by Theorem 5.6, a) system  $(\Sigma^N)$  is stable provided  $N$  is sufficiently large. This proves a). Statement b) can be established analogously  $\square$

One might now ask the question whether stability of system  $(\Sigma)$  implies stability of the approximating systems  $(\Sigma^N)$  uniformly with respect to  $N$ , i.e. the existence of constants  $M \geq 1$ ,  $\epsilon > 0$  such that

$$\|e^{[A^N]t}\|_N \leq M e^{-\epsilon t}, \quad t \geq 0,$$

for  $N$  sufficiently large (here  $\|\cdot\|_N$  denotes the operator norm corresponding to the vector norm  $|x|_N^2 = x^T Q^N x$  on  $\mathbb{R}^{n+p(N+1)n}$ ). A result of this type would be needed in order to apply a result of Gibson [21] concerning the approximation of the solution to the algebraic Riccati equation. Moreover, the uniform stability has been recently stated as a conjecture [8] for the spline approximation scheme developed in [7]. Our result below shows that such a conjecture is definitely wrong for the approximation scheme developed in this paper. This also indicates that it is wrong for the spline approximation scheme in [7].

Proposition 5.13. Suppose that there exist constants  $M \geq 1$  and  $\epsilon_N > 0$  such that

$$\|\exp((\frac{1}{N}q^N)^{-1}h^N t)\|_N \leq Me^{-\epsilon_N t}, \quad t \geq 0, \quad (5.22)$$

for all  $N$ . Then

$$\epsilon_N = O(\frac{1}{N^{1/2}}).$$

Here  $\|\cdot\|_N$  denotes the operator norm corresponding to the vector norm  $|x|^2 = \frac{1}{N}x^T q^N x$  on  $\mathbb{R}^{N+1}$ .

Proof. First note that

$$\frac{1}{6}x^T x \leq x^T q^N x \leq x^T x, \quad x \in \mathbb{R}^{N+1},$$

and therefore

$$x^T x \leq x^T (q^N)^{-1} x \leq 6x^T x, \quad x \in \mathbb{R}^{N+1}.$$

This implies for  $x_0 = \text{col}(1, 0, \dots, 0)$  and  $\mu \in i\mathbb{R}$  (cf. equation (5.2))

$$\begin{aligned} \sum_{k=0}^N |a_k^N(\mu)|^2 &\leq 6N |a^N(\mu)|_N^2 = 6N |(\mu q^N - h^N)^{-1} x_0|_N^2 \\ &= 6N \left| \int_0^\infty e^{-\mu t} \exp((\frac{1}{N}q^N)^{-1}h^N \frac{t}{N}) (q^N)^{-1} x_0 dt \right|_N^2 \\ &\leq 6NM^2 |(q^N)^{-1} x_0|_N^2 \left( \int_0^\infty e^{-\epsilon_N \frac{t}{N}} dt \right)^2 \\ &= \frac{6N^3 M^2}{\epsilon_N^2} |(q^N)^{-1} x_0|_N^2 = \frac{6N^2 M^2}{\epsilon_N^2} x_0^T (q^N)^{-1} x_0 \\ &\leq \frac{36N^2 M^2}{\epsilon_N^2}. \end{aligned}$$



Therefore

$$\epsilon_N \leq 6NM \left( \sum_{k=0}^N |a_k^N(\mu)|^2 \right)^{-1/2} \quad (5.23)$$

for all  $\mu \in i\mathbb{R}$ .

Now let  $\mu \in i\mathbb{R}$  satisfy  $|\mu| < \sqrt{3}$  and define the complex numbers  $w, \gamma_0, \gamma_1$  as in Lemma 5.9,c). Then  $w = \sqrt{9 - 3|\mu|^2} \in \mathbb{R}$  and  $|\gamma_0| = |\gamma_1| = |3 - \mu|$ . Hence there exists a  $\delta > 0$  such that

$$\frac{\gamma_1}{\gamma_0} = e^{i\delta}.$$

This together with (5.13) yields

$$\begin{aligned} |a_{N-k}^N(\mu)|^2 &= 36 \frac{|(3+w) - (3-w)e^{ik\delta}|^2}{|(3+w)^2 - (3-w)^2 e^{iN\delta}|^2} \\ &= 36 \frac{9(1 - \cos k\delta) + w^2(1 + \cos k\delta)}{(9+w)^2(1 - \cos N\delta) + 36w^2(1 + \cos N\delta)} \end{aligned}$$

for all  $\mu \in i\mathbb{R}$  with  $|\mu| < \sqrt{3}$ . Since  $\delta = \delta(\mu)$  and  $\delta \rightarrow 0$  as  $\mu \rightarrow i\sqrt{3}$ ,  $|\mu| < \sqrt{3}$ , we can choose a sequence  $\mu_N \in i\mathbb{R}$ ,  $|\mu_N| < \sqrt{3}$ , such that  $\mu_N \rightarrow i\sqrt{3}$  and  $\delta_N = \delta(\mu_N) = \frac{2\pi}{N}$ . The identity

$$\sin \delta = \operatorname{Im} \frac{\gamma_1}{\gamma_0} = \operatorname{Im} \frac{2\mu - w}{2\mu + w} = w \frac{4 \operatorname{Im} \mu}{w^2 + 4|\mu|^2}$$

shows that for positive constants  $c_1, c_2$

$$\frac{c_1}{N} \leq w_N \leq \frac{c_2}{N} \text{ for all } N.$$

From these facts we get

$$\begin{aligned}
\sum_{k=0}^N |\alpha_k^N(\mu_N)|^2 &= \frac{9}{2w_N^2} \sum_{k=0}^N (1 - \cos \frac{2k\pi}{N}) + \frac{1}{2} \sum_{k=0}^N (1 + \cos \frac{2k\pi}{N}) \\
&= \frac{9}{2w_N^2} (N-1) \\
&\geq \text{const. } N^3,
\end{aligned}$$

where we have used  $\sum_{k=0}^N \cos \frac{2k\pi}{N} \leq 2$ . This last estimate and (5.23) show that

$$\epsilon_N \leq \text{const. } \frac{1}{N^{1/2}} \alpha$$

The above result shows that exponential stability uniform with respect to  $N$  is in general impossible for our scheme - at least if  $A_p = 0$  and  $A_{01}(\tau) \equiv 0$ . Numerical studies show that there is a sequence  $\lambda_N$ ,  $N = 1, 2, \dots$ , of eigenvalues  $\lambda_N \in \sigma((\frac{1}{N}q^N)^{-1}h^N)$  such that  $\text{Re } \lambda_N \rightarrow 0$  and  $\text{Im } \lambda_N \rightarrow \infty$ . In fact, the numerical results indicate  $\text{Re } \lambda_N = O(\frac{1}{N^2})$  for this sequence. In the general case where  $\sigma((\frac{r_j}{N}q^N)^{-1}h^N)$  is not part of  $\sigma([A^N])$  numerical studies still show the existence of a sequence  $\lambda_N$ ,  $N = 1, 2, \dots$ , such that  $\lambda_N \in \sigma([A^N])$  with  $\text{Re } \lambda_N < 0$ ,  $\text{Re } \lambda_N \rightarrow 0$  and  $\text{Im } \lambda_N \rightarrow \infty$ .

## 6. Numerical results

The spline algorithm developed in Section 4.3 of this paper was applied to a large number of examples. In this section we present the numerical findings for some of those examples. The numerical results confirm the theoretical results in case of the finite time horizon problem. Despite the fact that we cannot prove convergence of the scheme for the infinite time horizon problem using the theory developed in [21] (as pointed out in Section 5.4) the scheme performs also very well for this class of problems. This is shown by some examples which already have been considered in the literature [ 8 ].

### 6.1. Examples with finite final time

Here we present examples where the true solutions of the control problems are available.

The suboptimal feedback law (4.37), which is governed by the matrices  $\pi_0^N(t), \pi_1^N(t), \dots, \pi_{pN}^N(t)$ , is calculated using the algorithm presented at the end of Section 4.3 (cf. (4.39) - (4.43)). The controls  $\hat{u}^N(t)$  and corresponding trajectories  $\hat{x}^N(t)$  were calculated by integration of

$$\frac{d}{dt} \hat{x}^N(t) = L(\hat{x}_t^N) - B_0 \hat{u}^N(t),$$

where  $\hat{u}^N(t)$  is given by (4.43). This delay system was solved by a modified Runge-Kutta procedure.

Example 6.1. This is the problem of minimizing

$$J(u) = \frac{3}{2}x(3)^2 + \frac{1}{2} \int_0^3 u(t)^2 dt$$

subject to

$$\dot{x}(t) = x(t-1) + u(t), \quad 0 \leq t \leq T = 3,$$

$$\phi^0 = \phi^1(0), \quad \phi^1(t) \equiv 1.$$

For this example we have  $n = 1$ ,  $A_{01} \equiv 0$ ,  $p = 1$ ,  $A_0 = 0$ ,  $A_1 = B_0 = 1$ ,  $C_0 = 0$ ,  $G_0 = \frac{3}{2}$  and  $R = \frac{1}{2}$ . The optimal controls, trajectories and costs were calculated in [5] using the maximum principle and are given by

$$\bar{u}(t) = \begin{cases} -\frac{\delta}{2}[(t-2)^2 + 3], & 0 \leq t \leq 1, \\ \delta(t-3) & , 1 \leq t \leq 2, \\ -\delta & , 2 \leq t \leq 3, \end{cases}$$

$$\bar{x}(t) = \begin{cases} 1+t-\delta\left[\frac{3t}{2}+\frac{1}{6}(t-2)^3+\frac{4}{3}\right], & 0 \leq t \leq 1, \\ \frac{3+t^2}{2}-\delta\left[4+\frac{3}{4}(t-1)^2+\frac{1}{24}(t-3)^4+\frac{4}{3}(t-1)-\frac{1}{2}(t-3)^2\right], & 1 \leq t \leq 2, \\ -\delta\left[\frac{547}{120}+5(t-2)+\frac{3}{2}(t-2)^2+\frac{1}{4}(t-2)^3-\frac{1}{6}(t-4)^3+\frac{1}{120}(t-4)^5\right], & 2 \leq t \leq 3, \end{cases}$$

$$J(\bar{u}) = \frac{329}{60} \delta^2$$

where  $\delta = \frac{185}{329}$ .

The numerical results we obtained are presented in Tables 6.1 and 6.2.

$t \backslash \hat{u}$	$\hat{u}^4(t)$	$\hat{u}^8(t)$	$\hat{u}^{16}(t)$	$\bar{u}(t)$
0	-1.9694	-1.9676	-1.9679	-1.9681
0.25	-1.7049	-1.7049	-1.7043	-1.7045
0.5	-1.4740	-1.4758	-1.4760	-1.4761
0.75	-1.2817	-1.2824	-1.2828	-1.2828
1.0	-1.1267	-1.1252	-1.1250	-1.1246
1.25	-0.9882	-0.9832	-0.9846	-0.9840
1.5	-0.8410	-0.8448	-0.8445	-0.8435
1.75	-0.6922	-0.7002	-0.7050	-0.7029
2.0	-0.5885	-0.5769	-0.5704	-0.5623
2.25	-0.5572	-0.5611	-0.5623	
2.5	-0.5620	-0.5623	-0.5623	
2.75	-0.5623	-0.5623	-0.5623	
3.0	-0.5621	-0.5622	-0.5623	-0.5623
$J(\hat{u})$	1.7338	1.7338	1.7338	1.7338

Table 6.1

$t \backslash \hat{x}$	$\hat{x}^4(t)$	$\hat{x}^8(t)$	$\hat{x}^{16}(t)$	$\bar{x}(t)$
0.25	0.7914	0.7916	0.7916	0.7917
0.5	0.6448	0.6448	0.6448	0.6448
0.75	0.5511	0.5506	0.5507	0.5507
1.0	0.5007	0.5005	0.5005	0.5005
1.25	0.4589	0.4593	0.4595	0.4595
1.5	0.4083	0.4096	0.4094	0.4094
1.75	0.3655	0.3642	0.3646	0.3646
2.0	0.3375	0.3372	0.3371	0.3370
2.25	0.3159	0.3167	0.3168	0.3168
2.5	0.2845	0.2848	0.2848	0.2849
2.75	0.2403	0.2407	0.2408	0.2408
3.0	0.1874	0.1874	0.1874	0.1874

Table 6.2

We observe that the error  $|\hat{u}^N(t) - \bar{u}(t)|$  is larger around  $t = 1$  and  $t = 2$  compared to other points in  $[0,3]$  because there  $\bar{u}(t)$  has jumps in the derivative whereas  $\hat{u}^N(t)$  is continuously differentiable on  $[0,3]$ . In Table 6.2 we didn't include the values for  $t = 0$  because always  $\hat{x}^N(0) = \bar{x}(0)$  for our algorithm.

Example 6.2. We have to minimize

$$J(u) = \frac{1}{2}(x_1(2)^2 + x_2(2)^2) + \frac{1}{2} \int_0^2 (u_1(t)^2 + u_2(t)^2) dt$$

subject to

$$\dot{x}(t) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x(t-1) + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} u(t), \quad 0 \leq t \leq 2$$

$$\phi^0 = \phi^1(0), \quad \phi^1(t) \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

We have  $n = 2$ ,  $p = 1$ ,  $A_{01} = 0$ ,  $A_0 = 0$ ,  $A_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ ,  $B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ,  $C_0 = 0$  and  $G_0 = R = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . Again the solution of this problem was obtained in [5] and is given by

$$\bar{u}_1(t) = \begin{cases} \mu + \delta(1-t), & 0 \leq t \leq 1, \\ \mu, & 1 \leq t \leq 2, \end{cases} \quad \hat{u}_2(t) = \delta, \quad 0 \leq t \leq 2,$$

$$\bar{x}_1(t) = \begin{cases} 1 + \mu t - \frac{\delta}{2}(t-1)^2 + \frac{\delta}{2}, & 0 \leq t \leq 1, \\ 1 + \mu t - \frac{\delta}{2}, & 1 \leq t \leq 2, \end{cases}$$

$$\bar{x}_2(t) = \begin{cases} 1 + (1+\delta)t, & 0 \leq t \leq 1, \\ 2 + (1+\frac{3}{2}\delta)(t-1) + \frac{\mu}{2}(t-1)^2 - \frac{\mu}{6}(t-2)^3 + \frac{5}{6}\delta, & 1 \leq t \leq 2, \end{cases}$$

and

$$J(\bar{u}) = \frac{3}{2}\mu^2 + \frac{1}{2}\mu\delta + \frac{5}{3}\delta^2,$$

$t \backslash \hat{u}_1$	$\hat{u}_1^4(t)$	$\hat{u}_1^8(t)$	$\hat{u}_1^{16}(t)$	$\bar{u}_1(t)$
0	-1.0608	-1.0601	-1.0603	-1.0598
0.25	-0.8495	-0.8407	-0.8428	-0.8419
0.5	-0.6197	-0.6260	-0.6256	-0.6239
0.75	-0.3890	-0.4018	-0.4092	-0.4060
1.0	-0.2284	-0.2107	-0.2007	-0.1880
1.25	-0.1799	-0.1861	-0.1881	
1.5	-0.1873	-0.1880	-0.1880	
1.75	-0.1879	-0.1880	-0.1880	
2.0	-0.1878	-0.1880	-0.1880	-0.1880
$t \backslash \hat{u}_2$	$\hat{u}_2^4(t)$	$\hat{u}_2^8(t)$	$\hat{u}_2^{16}(t)$	$\bar{u}_2(t)$
0	-0.8713	-0.8717	-0.8718	-0.8718
0.25	-0.8713	-0.8717		
0.5	-0.8712	-0.8717		
0.75	-0.8715	-0.8718		
1.0	-0.8718			
1.25	-0.8718			
1.5	-0.8719			
1.75	-0.8718			
2.0	-0.8718	-0.8718	-0.8718	-0.8718
$J(\hat{u})$	1.4018	1.4017	1.4017	1.4017

Table 6.3

$t \backslash \hat{x}_1$	$\hat{x}_1^4(t)$	$\hat{x}_1^8(t)$	$\hat{x}_1^{16}(t)$	$\bar{x}_1(t)$
0.25	0.76100	0.76192	0.76229	0.76229
0.5	0.57691	0.57927	0.57908	0.57906
0.75	0.45142	0.44963	0.45038	0.45032
1.0	0.37648	0.37640	0.37621	0.37607
1.25	0.32751	0.32893	0.32905	0.32906
1.5	0.28184	0.28198	0.28204	0.28504
1.75	0.23474	0.23498	0.23503	0.23504
2.0	0.18781	0.18799	0.18803	0.18803
$t \backslash \hat{x}_2$	$\hat{x}_2^4(t)$	$\hat{x}_2^8(t)$	$\hat{x}_2^{16}(t)$	$\bar{x}_2(t)$
0.25	1.03216	1.03207	1.03205	1.03205
0.5	1.06433	1.06414	1.06411	1.06410
0.75	1.09647	1.09619	1.09616	1.09615
1.0	1.12855	1.12823	1.12820	1.12821
1.25	1.12961	1.12939	1.12940	1.12941
1.5	1.07768	1.07797	1.07799	1.07799
1.75	0.98704	0.98747	0.98757	0.98758
2.0	0.87175	0.87177	0.87179	0.87180

Table 6.4



## 6.2. Examples for the infinite time horizon problem

Despite the fact we cannot prove convergence of our scheme for the infinite time horizon following the ideas of Gibson [21] the algorithm seems to behave very well also in this case. The following examples show that it should be possible to prove convergence of the operators  $\Pi^N$  to  $\Pi$  in the uniform operator topology. Here  $\Pi^N$  restricted to  $X^N$  (on  $(X^N)^\perp$  we have  $\Pi^N = 0$ ) is the solution of the algebraic Riccati equation

$$(A^N)^* \Pi^N + \Pi^N A^N - \Pi^N B R^{-1} B^* \Pi^N + C^* C = 0 \quad (6.1)$$

and  $\Pi$  is the solution of (3.21). Taking matrix representations for the operators equation (6.1) takes the following form (cf. also (4.38)):

$$\begin{aligned} [(A^N)^*][\Pi^N] + [\Pi^N][A^N] - [\Pi^N][B^N]R^{-1}[B^N]^T[\Pi^N] \\ + [C^N]^T[C^N] = 0. \end{aligned}$$

Analogously as in case of equation (4.38) we define

$$\Gamma^N = Q^N[\Pi^N]$$

and get the Riccati matrix equation

$$\begin{aligned} [A^N]^T \Gamma^N + \Gamma^N [A^N] \\ - \Gamma^N [B^N] R^{-1} [B^N]^T \Gamma^N + [C^N]^T [C^N] = 0. \end{aligned} \quad (6.2)$$

Using  $\Pi^N$  instead of  $\Pi$  in the feedback law (3.24) we get by analogous computations as in (4.44)

$$\hat{u}^N(t) = -R^{-1} B_0^T \left\{ \Pi_0^N \hat{x}^N(t) + \int_{-h}^0 \Pi_1^N(\tau) \hat{x}(t+\tau) d\tau \right\}, t \geq 0, \quad (6.3)$$

where  $\pi_1^N(\tau) = \sum_{k=1}^p \sum_{j=0}^N (\pi_{kj}^N)^T e_{kj}(\tau)$  and

$$[\pi^N] = \begin{pmatrix} \pi_0^N & * & \dots & * \\ \pi_{10}^N & & & \\ \vdots & & & \\ \pi_{pN}^N & * & & * \end{pmatrix}$$

(compare (4.42) and (4.43)). The numerical results clearly show  $\|\pi_p^N - \pi\| \rightarrow 0$  as  $N \rightarrow \infty$ . Thus there should be a way to prove this convergence without having uniform exponential stability for the approximating problem.

For the examples (6.2) was solved using the Newton-Kleinman algorithm as presented in [34], for instance. The Ljapunov matrix equation which has to be solved in each step of the Newton-Kleinman algorithm was solved using the quadratically convergent procedure given by R.A. Smith [39] (see also [34], p. 297). The approximating controls  $\hat{u}^N(t)$  and corresponding trajectories  $\hat{x}^N(t)$  were calculated as for the examples in Section 6.1. The following examples were already considered in [8] where the approximation was done using the spline scheme developed in [7].

Example 6.3. This is Example 4.1 in [8] and considers the minimization of

$$J(u) = \int_0^{\infty} [x(t)^2 + u(t)^2] dt$$

subject to

$$\dot{x}(t) = x(t) + x(t-1) + u(t), \quad t \geq 0,$$

$$\phi^0 = 0, \quad \phi^1(t) = \sin \pi t, \quad -1 \leq t \leq 0.$$

AD-A142 963

SPLINE APPROXIMATION FOR RETARDED SYSTEMS AND THE  
RICCATI EQUATION(U) WISCONSIN UNIV-MADISON MATHEMATICS  
RESEARCH CENTER F KAPPEL ET AL. APR 84 MRC-TSR-2680  
DAAG29-80-C-0041 F/G 12/1

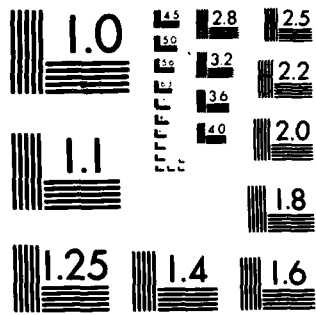
2/2

UNCLASSIFIED

NL



END  
DATE  
FILMED  
8 84  
DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

In this case we have  $n = p = 1$ ,  $A_0 = A_1 = B_0 = C_0 = R = 1$ . In Table 6.5 we give the values for  $J(\hat{u}^N)$  and the optimal costs  $J^N = \langle \Pi_p^N(\phi^0, \phi^1), p^N(\phi^0, \phi^1) \rangle$  for the approximating systems  $(\Sigma^N)$  with cost functional (4.33).

N	$J(\hat{u}^N)$	$J^N$
4	0.321439	0.321430
8	0.321439	0.321432
16	0.321439	0.321430

Table 6.5

In Table 6.6 we show the values for  $\Pi_0^N$  and  $\Pi_{1N}^N$ . Since  $\text{range } \Pi \subset \text{dom } A^*$  (cf. Proposition 3.4,b)), we have  $A_1^T \Pi_{00} \phi^0 = (\Pi_{10} \phi^0)(-1)$  for all  $\phi^0 \in \mathbb{R}^n$ . Therefore we should have

$$\Pi_0^N - \Pi_{1N}^N \rightarrow 0 \text{ as } N \rightarrow \infty.$$

N	$\Pi_0^N$	$\Pi_{1N}^N$
4	2.80886	2.77538
8	2.809328	2.80096
16	2.809390	2.80729
32	2.809396	2.80887

Table 6.6

In Table 6.7 we give the values for  $\Pi_1^N(\tau)$ , which governs the distributed feedback in (6.3), at the knots  $-\frac{j}{N}$ ,  $j = 0, \dots, N$ , for  $N = 4, 8, 16$  and  $32$ . We clearly see that  $\Pi_1^N(\tau)$  converges uniformly on  $-1 \leq \tau \leq 0$  as  $N \rightarrow \infty$ .

j	$\pi_1^4(-\frac{j}{32})$	$\pi_1^8(-\frac{j}{32})$	$\pi_1^{16}(-\frac{j}{32})$	$\pi_1^{32}(-\frac{j}{32})$
0	0.63598	0.63683	0.63694	0.63696
1	-	-	-	0.66132
2	-	-	0.68698	0.68764
3	-	-	-	0.71558
4	-	0.74240	0.74512	0.74553
5	-	-	-	0.77722
6	-	-	0.81040	0.81114
7	-	-	-	0.84695
8	0.87064	0.88269	0.88474	0.88517
9	-	-	-	0.92547
10	-	-	0.96750	0.96839
11	-	-	-	1.01361
12	-	1.05757	1.06113	1.06165
13	-	-	-	1.11225
14	-	-	1.16491	1.16591
15	-	-	-	1.22238
16	1.26664	1.27879	1.28151	1.28218
17	-	-	-	1.34508
18	-	-	1.41048	1.41162
19	-	-	-	1.48157
20	-	1.54972	1.55463	1.55547
21	-	-	-	1.63314
22	-	-	1.71381	1.71512
23	-	-	-	1.80125
24	1.86588	1.88693	1.89104	1.89209
25	-	-	-	1.98748
26	-	-	2.08649	2.08802
27	-	-	-	2.19358
28	-	2.29692	2.30348	2.30477
29	-	-	-	2.42147
30	-	-	2.54253	2.54432
31	-	-	-	2.67323
32	2.77538	2.80096	2.80729	2.80887

Table 6.7

With respect to  $\pi_0^N$  the behavior of the scheme presented in [8] is comparable to the performance of our scheme. But with respect to approximation of the feedback kernel our scheme behaves much better (compare Figures 4.1 - 4.4 in [8]). Our scheme seems also to be more accurate as far as approximation of  $J(\bar{u})$  by  $J^N(\hat{u}^N)$  and  $J^N$  is concerned. In Tables 6.8 and 6.9 we present the values for  $\hat{u}^N(t)$  and  $\hat{x}^N(t)$  on  $0 \leq t \leq 4$  and  $0 \leq t \leq 3$ , respectively, where  $\hat{u}^N(t)$  is given by (6.3) and  $\hat{x}^N(t)$  solves  $\frac{d}{dt}\hat{x}^N(t) = L(\hat{x}_t^N) + B_0\hat{u}^N(t)$ ,  $(\hat{x}(0), \hat{x}_0^N) = (\phi^0, \phi^1)$ .

$t \backslash \hat{u}^N$	$\hat{u}^4(t)$	$\hat{u}^8(t)$	$\hat{u}^{16}(t)$
0	0.86836	0.86817	0.86816
0.25	0.64894	0.64891	0.64891
0.5	0.49650	0.49657	0.49658
0.75	0.35400	0.36400	0.36401
1.0	0.24627	0.24618	0.24618
1.25	0.16154	0.16146	0.16146
1.5	0.10999	0.10993	0.10993
1.75	0.08024	0.08021	0.08021
2.0	0.06015	0.06015	0.06015
2.25	0.04348	0.04347	0.04347
2.5	0.02983	0.02982	0.02982
2.75	0.01996	0.01995	0.01995
3.0	0.01373	0.01372	0.01372
3.25	0.00991	0.00991	0.00991
3.5	0.00729	0.00729	0.00729
3.75	0.00524	0.00523	0.00523
4.0	0.00362	0.00362	0.00362

Table 6.8

$t \backslash \hat{x}^N$	$\hat{x}^4(t)$	$\hat{x}^8(t)$	$\hat{x}^{16}(t)$
0.25	0.11259	0.11258	0.11258
0.5	0.05332	0.05331	0.05332
0.75	-0.06628	-0.06626	-0.06626
1.0	-0.10850	-0.10846	-0.10846
1.25	-0.06160	-0.06158	-0.06158
1.5	-0.01397	-0.01397	-0.01397
1.75	0.00753	0.00752	0.00752
2.0	0.00178	0.00178	0.00178
2.25	-0.00784	-0.00784	-0.00784
2.5	-0.01030	-0.01029	-0.01029
2.75	-0.00646	-0.00646	-0.00646
3.0	-0.00178	-0.00178	-0.00178

Table 6.9

**Example 6.4.** This is a model for the Mach number control loop for the National Transonic Facility at NASA's Langley Research Center. For details see [2] or [8]. The problem is to minimize

$$J(u) = \int_0^{\infty} [x^T(t)C_0^T C_0 x(t) + u^2(t)] dt$$

subject to

$$\begin{aligned} \dot{x}(t) = & \begin{bmatrix} -a & 0 & 0 \\ 0 & 0_2 & 1 \\ 0 & -\omega_2 & -2\xi\omega \end{bmatrix} x(t) + \begin{bmatrix} 0 & ka & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x(t-0.33) \\ & + \begin{bmatrix} 0 \\ 0_2 \\ \omega_2 \end{bmatrix} u(t), \quad t \geq 0, \end{aligned} \quad (6.4)$$

$$\phi^0 = \text{col}(-0.1, 8.547, 0) \equiv \phi^1(t), \quad -0.33 \leq t \leq 0.$$



We have  $C_0 = (100, 0, 0)$ ,  $n = 3$ ,  $p = 1$ ,  $k = -0.0117$ ,  $\xi = 0.8$ ,  $\omega = 6.0$  and  $\frac{1}{a} = 1.964$ . Because of the simple structure of this problem it is possible to calculate the true solution following an idea contained in [29]. If we put for  $t \geq 0$

$$y_1(t) = x_1(t+h), \quad h = 0.33,$$

$$y_2(t) = x_2(t),$$

$$y_3(t) = x_3(t),$$

we obtain by a simple calculation

$$\frac{d}{dt} \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} -a & ka & 0 \\ 0 & 0 & 1 \\ 0 & -\omega^2 & -2\xi\omega \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \omega^2 \end{pmatrix} u(t). \quad (6.5)$$

The cost functional takes the form

$$J(u) = 10^4 \int_0^h x_1(t)^2 dt + \int_0^\infty [10^4 y_1(t)^2 + u(t)^2] dt, \quad (6.6)$$

where

$$x_1(t) = e^{-at} \phi_1^0 + ak \int_0^t e^{-a(t-\tau)} \phi_2^1(\tau-h) d\tau, \quad (6.7)$$

$$0 \leq t \leq h,$$

is not dependent on  $u(t)$  on the interval  $[0, h]$ . Therefore minimizing  $J(u)$  subject to (6.4) is equivalent to minimizing

$$\tilde{J}(u) = \int_0^\infty [10^4 y_1(t)^2 + u(t)^2] dt$$

subject to (6.5) with initial data

$$y_1(0) = x_1(h), \quad y_2(0) = x_2(0), \quad y_3(0) = x_3(0). \quad (6.8)$$

The solution of the latter problem is given by the feedback law

$$\bar{u}(t) = - (0, 0, \omega^2) \bar{\pi}_0 \bar{y}(t),$$

where  $\bar{y}(t)$  is the solution of (6.5) with  $u(t) = \bar{u}(t)$  and initial data (6.8).  $\bar{\pi}_0$  is the solution of the algebraic Riccati equation

$$\bar{A}^T \bar{\pi}_0 + \bar{\pi}_0 \bar{A} - \bar{\pi}_0 B_0 B_0^T \bar{\pi}_0 + C_0^T C_0 = 0, \quad (6.9)$$

where

$$\bar{A} = \begin{pmatrix} -a & ak & 0 \\ 0 & 0 & 1 \\ 0 & -\omega^2 & -2\xi\omega \end{pmatrix}.$$

Equation (6.9) was solved numerically to give

$$\bar{\pi}_0 = \begin{pmatrix} 8220.51099 & -11.61086 & -1.12107 \\ -11.61086 & 0.01851 & 0.00186 \\ -1.12107 & 0.00186 & 0.00019 \end{pmatrix}.$$

The optimal costs for the original problem are given by

$$\begin{aligned} J(\bar{u}) &= \bar{J}(\bar{u}) + 10^4 \int_0^h \bar{x}_1(t)^2 dt \\ &= \bar{y}^T(0) \bar{\pi}_0 \bar{y}(0) + 10^4 \int_0^h \bar{x}_1(t)^2 dt. \end{aligned} \quad (6.10)$$

Using (6.7) it is easy to calculate  $J(\bar{u})$ . In Table 6.10 we give the values for  $J(\bar{u})$ ,  $J(\hat{u}^N)$  and  $J^N = \langle \Pi^N p^N(\phi^0, \phi^1), p^N(\phi^0, \phi^1) \rangle$ .

N	$J(\hat{u}^N)$	$J^N$
4	136.39587	136.40499
8	136.40094	136.40509
16	136.40250	136.40521
$J(\bar{u})$	136.40490	

Table 6.10

Using (6.10), (6.8) and (6.7) we get by some calculations for general initial data  $(\phi^0, \phi^1)$

$$\begin{aligned}
 J(\bar{u}) = & (\phi^0)^T \left\{ \begin{pmatrix} e^{-ah} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} e^{-ah} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 10^4 \frac{1-e^{-2ah}}{2a} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\} \phi^0 \\
 & + 2ak(\phi^0)^T \begin{pmatrix} e^{-ah} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 e^{a\tau} \phi^1(\tau) d\tau \\
 & + 10^4 k e^{-ah} (\phi^0)^T \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 (e^{-a\tau} - e^{a\tau}) \phi^1(\tau) d\tau \\
 & + a^2 k^2 \int_{-h}^0 e^{a\tau} (\phi^1(\tau))^T d\tau \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 e^{a\tau} \phi^1(\tau) d\tau \\
 & + 10^4 a^2 k^2 \langle \phi_2^1(\sigma), \int_{-h}^0 e^{-a(\tau-\sigma)} \int_{-h}^{\tau} e^{-a(\tau-\sigma)} \phi_2^1(\sigma) d\sigma d\tau \rangle_{L^2} .
 \end{aligned}$$

On the other hand we have

$$J(\bar{u}) = (\phi^0)^T \pi_{00} \phi^0 + 2(\phi^0)^T \pi_{01} \phi^1 + \langle \phi^1, \pi_{11} \phi^1 \rangle_{L^2} .$$

Comparison immediately gives

$$\begin{aligned}
 \pi_{00} = & \begin{pmatrix} e^{-ah} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} e^{-ah} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + 10^4 \frac{1-e^{-2ah}}{2a} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \\
 \pi_{01} \phi^1 = & ak \begin{pmatrix} e^{-ah} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 e^{a\tau} \phi^1(\tau) d\tau \\
 & + 10^4 k e^{-ah} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 \frac{e^{-a\tau} - e^{a\tau}}{2} \phi^1(\tau) d\tau
 \end{aligned}$$

or equivalently,  $(\pi_{01}^* \phi^0)(\theta) = \pi_1(\theta) \phi^0$  with

$$\pi_1(\theta) = ak \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} e^{-a\theta} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} e^{a\theta} + 10^4 ke^{-a\theta} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \frac{e^{-a\theta} - e^{a\theta}}{2},$$

$-h \leq \theta \leq 0$ . Furthermore, we get after some calculation

$$\begin{aligned} (\pi_{11}^{\phi^1})(\theta) &= a^2 k^2 e^{a\theta} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \bar{\pi}_0 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 e^{a\tau} \phi^1(\tau) d\tau \\ &+ ak^2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \int_{-h}^0 \frac{e^{-a|\tau-\theta|} - e^{a(\tau+\theta)}}{2} \phi^1(\tau) d\tau, \quad -h \leq \theta \leq 0. \end{aligned}$$

In Table 6.11 we present the values for  $\pi_0^N$  and  $\pi_{00}$  and in Table 6.12 we show the values for the second row of  $\pi_1^N(-j\frac{h}{4})$  and  $\pi_1(-j\frac{h}{4})$  for  $j = 0, \dots, 4$ . The other rows of these matrices are always zero.

N	$\pi_0^N$
4	$\begin{pmatrix} 8677.02417 & -9.81502 & -0.94768 \\ -9.81502 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{pmatrix}$
8	$\begin{pmatrix} 8677.02698 & -9.81505 & -0.94768 \\ -9.81505 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{pmatrix}$
16	$\begin{pmatrix} 8677.03516 & -9.81506 & -0.94768 \\ -9.81506 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{pmatrix}$
$\pi_{00}$	$\begin{pmatrix} 8677.02405 & -9.81505 & -0.94768 \\ -9.81505 & 0.01851 & 0.00186 \\ -0.94768 & 0.00186 & 0.00019 \end{pmatrix}$

Table 6.11

j	$\pi_1^4(-\frac{jh}{4})$		
0	-41.39697	0.06916	0.00668
1	-43.83789	0.06652	0.00640
2	-46.37943	0.06334	0.00613
3	-48.97898	0.06118	0.00590
4	-51.69006	0.05828	0.00563
j	$\pi_1^8(-\frac{jh}{4})$		
0	-41.39721	0.06917	0.00668
1	-43.84998	0.06626	0.00640
2	-46.38019	0.06355	0.00614
3	-48.99226	0.06093	0.00588
4	-51.69080	0.05843	0.00564
j	$\pi_1^{16}(-\frac{jh}{4})$		
0	-41.39727	0.06917	0.00668
1	-43.85012	0.06631	0.00640
2	-46.38036	0.06358	0.00614
3	-48.99246	0.06097	0.00589
4	-51.69102	0.05846	0.00564
j	$\pi_1(-\frac{jh}{4})$		
0	-41.39721	0.06917	0.00668
1	-43.85008	0.06632	0.00641
2	-46.38034	0.06360	0.00614
3	-48.99246	0.06098	0.00589
4	-51.69103	0.05847	0.00565

Table 6.12

## References.

- [1] J.H. Ahlberg, E.N. Nilsen and J.C. Walsh, The Theory for Splines and Their Applications, Academic Press, New York 1967.
- [2] E.S. Armstrong and J.S. Tripp, An application of multivariable design techniques to the control of the National Transonic Facility, NASA Technical Paper 1887, NASA Langley Research Center, Hampton, Virginia 1981.
- [3] H.T. Banks, Identification of nonlinear delay systems using spline methods, Proc. Int. Conf. Nonl. Phenomena in Math. Sci., Arlington, Texas, June 1980.
- [4] H.T. Banks and J.A. Burns, Hereditary control problems: Numerical methods based on averaging approximation, SIAM J. Control Optimization 16 (1978), 169-208.
- [5] H.T. Banks, J.A. Burns, E.M. Cliff and P.R. Thrift, Numerical solutions of hereditary control problems via an approximation technique, Brown University, LCDS Technical Report 15-6, 1975, 1975, Providence, R.I.
- [6] H.T. Banks and P.L. Daniel, Parameter Estimation of nonlinear nonautonomous distributed systems, Proc. 20th IEEE Conf. Decision and Control, San Diego, December 1981, 228-232.
- [7] H.T. Banks and F. Kappel, Spline approximation for functional differential equations, J. Differential Equations 34 (1979), 496-522.
- [8] H.T. Banks, G.I. Rosen and K. Ito, A spline based technique for computing Riccati operators and feedback controls in regulator problems for delay equations, SIAM J. Sci. Stat. Comput., to appear.
- [9] A. Bensoussan, M.C. Delfour and S.K. Mitter, The linear quadratic optimal control problem for infinite dimensional systems over an infinite horizon; survey and examples, Electronic Systems Laboratory Reports M.I.T., Cambridge, Massachusetts, 1977
- [10] C. Bernier and A. Manitius, On semigroups in  $\mathbb{R}^n \times L^p$  corresponding to differential equations with delays, Can. J. Math. 30 (1978), 897-914.
- [11] K.P.M. Bhat and H.N. Koivo, An observer theory for time delay systems, IEEE Trans. Autom. Control AC-21 (1976), 266-269.
- [12] K. Boehmer, Spline-Funktionen, Theorie und Anwendungen, Teubner, Stuttgart 1974.

- [13] J. Borisovič and A.S. Turbabin, On the Cauchy problem for linear nonhomogeneous differential equations with retarded argument, Soviet Math. Doklady 10 (1969), 401-405.
- [14] J.A. Burns and T.L. Herdman, Adjoint semigroup theory for a class of functional differential equations, SIAM J. Math. Anal. 5 (1976), 729-745.
- [15] R.F. Curtain and A.J. Pritchard, Infinite Dimensional Linear Systems Theory, Springer, Berlin 1978.
- [16] R. Datko, Unconstrained control problems with quadratic cost, SIAM J. Control 11 (1973), 32-52.
- [17] M.C. Delfour, E.B. Lee and A. Manitius, F-reduction of the operator Riccati equation, Automatica 14 (1978), 385-395.
- [18] M.C. Delfour and A. Manitius, The structural operator F and its role in the theory of retarded systems, Part I: J. Math. Anal. Appl. 73 (1980), 466-490, Part II: J. Math. Anal. Appl. 74 (1980), 359-381.
- [19] M.C. Delfour and S.K. Mitter, Hereditary differential systems with constant delays I: General case, J. Diff. Equations 12 (1972), 213-235.
- [20] J.S. Gibson, The Riccati integral equations for optimal control problems in Hilbert spaces, SIAM J. Control Opt. 17 (1979), 537-565.
- [21] J.S. Gibson, Linear quadratic optimal control of hereditary differential systems: infinite dimensional Riccati equations and numerical approximations, SIAM J. Control Opt. 21 (1983), 95-139.
- [22] D. Gottlieb and S.A. Orszag, Numerical Analysis of Spectral Methods: Theory and Applications, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM 1977.
- [23] J.K. Hale, Theory of Functional Differential Equations, Springer, New York 1977.
- [24] A.S.B. Holland and B.N. Sahney, The General Problem of Approximation and Spline Functions, Krieger, Huntington, N.Y., 1979.
- [25] F. Kappel, Finite dimensional approximation of systems with infinite dimensional state space, EQUADIFF 82, H.W. Knobloch and K. Schmitt, eds., pp. 287-299, LNM 1017, Springer, Berlin 1983.

- [26] K. Kunisch, Approximation schemes for the linear quadratic optimal control problem associated with delay equations, SIAM J. Control Opt. 20 (1982), 506-540.
- [27] A. Manitius, Completeness and F-completeness of eigenfunctions associated with retarded functional differential equations, J. Diff. Equations 35 (1980), 1-29.
- [28] A. Manitius, Necessary and sufficient conditions of approximate controllability for general linear retarded systems, SIAM J. Control Opt. 19 (1981), 516-532.
- [29] A. Manitius and H. Tran, Numerical simulation of linear and nonlinear feedback controllers for a wind tunnel model involving a time delay, Rensselaer Polytechnic Institute, Preprint 1983.
- [30] R.H. Martin, Jr., Nonlinear Operators and Differential Equations in Banach spaces, J. Wiley, New York, 1976.
- [31] R.K. Miller, Linear Volterra integro-differential equations as semigroups, Funkcialaj Ekvacioj 17 (1974), 39-55.
- [32] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer, New York 1983.
- [33] L. Pandolfi, Feedback stabilization of functional differential equations, Bollettino U.M.I. 12 (1975), 626-635.
- [34] D.L. Russell, Mathematics of Finite Dimensional Control Systems, Theory and Design, Marcel Dekker, New York 1979.
- [35] D. Salamon, On dynamic observation and state feedback for time delay systems, "Evolution Equations and Their Applications", F. Kappel and W. Schappacher, eds., pp. 202-219, RNM 68, Pitman, London 1982.
- [36] D. Salamon, Control and Observation of Neutral Systems, RNM 91, Pitman, London 1984.
- [37] D. Salamon, Structure and stability of finite dimensional approximations for functional differential equations, MRC, University of Wisconsin-Madison, TSR# 2586, 1983.
- [38] M.H. Schultz, Spline Analysis, Prentice Hall, Englewood Cliffs, N.Y. 1973.
- [39] R.A. Smith, Matrix equation  $XA + BX = C$ , SIAM J. Appl. Math. 16 (1968), 198-201.



- [40] R.B. Vinter and R.H. Kwong, The infinite time quadratic control problem for linear systems with state and control delays: an evolution equation approach, SIAM J. Control Opt. 19 (1981), 139-153.
- [41] J. Zabczyk, Remarks on the algebraic Riccati equation in Hilbert space, Appl. Math. Opt. 2 (1976), 251-258.

FK/DS/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER # 2680	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Spline Approximation for Retarded Systems and the Riccati Equation		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) F. Kappel and D. Salamon		8. CONTRACT OR GRANT NUMBER(s) MCS-8210950 DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 3 - Numerical Analysis and Scientific Computing
11. CONTROLLING OFFICE NAME AND ADDRESS  See Item 18 below		12. REPORT DATE April 1984
		13. NUMBER OF PAGES 105
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)  UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES U. S. Army Research Office National Science Foundation P. O. Box 12211 Washington, DC 20550 Research Triangle Park North Carolina 27709		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Retarded functional differential equations, Approximation, Splines, Riccati Equation		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The purpose of this paper is to introduce a new spline approximation scheme for retarded functional differential equations. The special feature of this approximation scheme is that it preserves the product space structure of retarded systems and approximates the adjoint semigroup in a strong sense. These facts guarantee the convergence of the solution operators to the differential Riccati equation in a strong sense. Numerical findings indicate a significant improvement in the convergence behaviour over both the averaging and the previous spline approximation scheme.		

ABSTRACT (continued)

Furthermore, controllability and observability criteria are given for the approximating systems, which are shown to be stable respectively stabilizable for sufficiently large  $N$  provided that the underlying retarded system has the same property.

