

## ON THE STABILITY PROPERTIES OF SPLINE APPROXIMATIONS FOR RETARDED SYSTEMS\*

F. KAPPEL† AND D. SALAMON‡

**Abstract.** This paper studies the qualitative properties of the spline approximation scheme for retarded functional differential equations introduced by Kappel and Salamon [*SIAM J. Control Optim.*, 25 (1987), pp. 1082-1117]. It is shown that the approximating systems are stable for large  $N$  if the underlying retarded functional differential equation is stable. In this case the approximating equations are in some sense uniformly (with respect to the approximation index) stable in the vector component of the state but not so in the complete state.

**Key words.** retarded functional differential equations, approximation, splines, controllability

**AMS(MOS) subject classifications.** 34K35, 41A15, 93D15

**1. Introduction.** In [10] and [11] we have introduced a new spline approximation scheme for retarded functional differential equations. The aim of this paper is to study the qualitative properties of this approximation scheme with particular emphasis on the stability problem.

The fundamental convergence properties of this approximation scheme have been established in [11]. The central result is a convergence proof for both the original semigroup  $S(t)$  and its adjoint  $S^*(t)$  in the strong operator topology. Here lies the main advantage over the spline approximation scheme, developed earlier in [2], for which the adjoint semigroup is only approximated in the weak operator topology. In addition, we have observed a quite significant improvement in the convergence behaviour of our numerical computations, some of which are reported in [11].

The main result of this paper is that the approximating systems  $(\Sigma^N)$  are stable (stabilizable, detectable) for sufficiently large  $N$  provided the original system  $(\Sigma)$  is stable (stabilizable and detectable.) The proof consists of three parts. The first part is a convenient characterization of the stability, stabilizability, and detectability of the approximating systems in terms of a certain characteristic matrix  $\Delta^N(\lambda)$ . The second part is a convergence proof for these matrices  $\Delta^N(\lambda)$ . The third part establishes a priori bounds for the unstable eigenvalues of the approximating systems.

We also discuss the role of the structural operator  $F$  in the spline approximation scheme. Moreover, we prove that the approximating systems cannot be stable in a uniform sense with respect to  $N$  and illustrate this result with computations of the spectrum. In this respect the spline approximation differs from the averaging approximation scheme in [1] for which the uniform exponential stability property has been established in [19]. But if we take the output of the system to be the vector component of the state, then the approximating systems are in a sense uniformly output stable with respect to  $N$  if the hereditary system is stable. For simplicity of presentation we restrict ourselves to the single delay case. All results are true for equations with multiple commensurate delays and without distributed delay. Some results are also true for the general case. For details see [10].

\* Received by the editors December 23, 1986; accepted for publication (in revised form) May 13, 1988.  
This research was partly supported by Fonds zur Förderung der wissenschaftlichen Forschung (Austria)  
Project S3206.

† Institut für Mathematik, Karl-Franzens-Universität Graz, Elisabethstrasse 16, A-8010 Graz, Austria.

‡ Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom. The research of this author was partly supported by the National Science Foundation grant MCS-8210950.

## 2. Linear retarded control systems.

**2.1. Functional differential equations.** We consider the linear retarded functional differential equation (RFDE)

$$(2.1) \quad \dot{x}(t) = A_0 x(t) + A_1 x(t-h) + B_0 u(t), \quad y(t) = C_0 x(t),$$

where  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^l$ ,  $y(t) \in \mathbb{R}^m$ ,  $A_0, A_1 \in \mathbb{R}^{n \times n}$ ,  $B_0 \in \mathbb{R}^{n \times l}$ ,  $C_0 \in \mathbb{R}^{m \times n}$ , and  $h > 0$ . It is obvious that (2.1) admits a unique solution  $x(\cdot) \in L^2(-h, T; \mathbb{R}^n) \cap W^{1,2}(0, T; \mathbb{R}^n)$  for every input  $u(\cdot) \in L^2(0, T; \mathbb{R}^l)$  and every initial condition of the form

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

where  $\phi = (\phi^0, \phi^1) \in M^2 = \mathbb{R}^n \times L^2(-h, 0; \mathbb{R}^n)$  (see, for instance, [6], [9]). By  $X(t) \in \mathbb{R}^{n \times n}$ ,  $t \geq -h$ , we denote the fundamental matrix solution of (2.1), which corresponds to the initial condition  $X(0) = I$ ,  $X(\tau) = 0$ ,  $-h \leq \tau < 0$ , and the input  $u(t) \equiv 0$ . Its Laplace transform is given by  $\Delta(\lambda)^{-1}$ , where  $\Delta(\lambda) = \lambda I - A_0 - A_1 e^{-\lambda h}$ ,  $\lambda \in \mathbb{C}$ , is the characteristic matrix of (2.1).

**2.2. State space theory.** We consider two state concepts for (2.1). In the classical sense the state at time  $t \geq 0$  is defined to be the pair  $z(t) = (x(t), x_t) \in M^2$ , where  $x_t(\tau) = x(t+\tau)$  for  $-h \leq \tau \leq 0$ . This state defines a weak solution of the abstract Cauchy problem

$$(\Sigma) \quad \dot{z}(t) = Az(t) + Bu(t), \quad z(0) = \phi, \quad y(t) = Cz(t),$$

where  $B \in \mathcal{L}(\mathbb{R}^l, M^2)$  and  $C \in \mathcal{L}(M^2, \mathbb{R}^m)$  are defined by  $Bu = (B_0 u, 0)$  and  $C\phi = C_0 \phi^0$  for  $u \in \mathbb{R}^l$  and  $\phi \in M^2$ . The unbounded operator  $A: \text{dom } A \rightarrow M^2$  is given by

$$A\phi = (A_0 \phi^1(0) + A_1 \phi^1(-h), \phi^1), \quad \text{dom } A = \{\phi \in M^2 \mid \phi^1 \in W^{1,2}, \phi^0 = \phi^1(0)\}$$

and generates a strongly continuous semigroup  $S(t)$  of bounded linear operators on  $M^2$ . Therefore  $z(t) \in M^2$  is given by the variation-of-constants formula

$$(2.3) \quad z(t) = S(t)\phi + \int_0^t S(t-s)Bu(s) ds.$$

Now let  $S_T(t)$  denote the semigroup corresponding to the RFDE  $\dot{x}(t) = A_0^T x(t) + A_1^T x(t-h)$  so that its generator  $A_T$  is defined as  $A$  with  $A_0, A_1$  replaced by  $A_0^T, A_1^T$ . Then there is an alternative (dual) state concept for the RFDE (2.1) that relates the semigroups  $S(t)$  and  $S_T^*(t)$ . It can be defined in terms of the structural operator  $F \in \mathcal{L}(M^2)$  (for a normed linear space  $X$  we denote by  $\mathcal{L}(X)$  the space of all bounded linear operators  $X \rightarrow X$ ) given by

$$(2.4) \quad [F\phi]^0 = \phi^0, \quad [F\phi]^1(\sigma) = A_1 \phi(-h-\sigma), \quad -h \leq \sigma \leq 0$$

for  $\phi \in M^2$  (we define  $\phi^1(\tau) = 0$  for  $\tau \notin [-h, 0]$ ). It is a remarkable fact that for every weak solution  $z(t) \in M^2$  of the Cauchy problem  $(\Sigma)$  the function  $w(t) = Fz(t) \in M^2$  defines a weak solution of the abstract Cauchy problem

$$(\Sigma_T^*) \quad \dot{w}(t) = A_T^* w(t) + Bu(t), \quad w(0) = f \in M^2, \quad y(t) = Cw(t),$$

with  $f = F\phi$ .

Equivalently, the structural operator  $F$  satisfies the following equations:

$$(2.5) \quad FS(t) = S_T^*(t)F, \quad FB = B, \quad CF = C$$

for  $t \geq 0$ . In particular, for every solution  $x(t) \in \mathbb{R}^n$ ,  $t \geq -h$ , of (2.1) the function  $w(t) = F(x(t), x_t) \in M^2$  is given by

$$(2.6) \quad w(t) = S_T^*(t)F\phi + \int_0^t S_T^*(t-s)Bu(s) ds.$$

For more detailed discussion of these two state concepts and their relation see [6], [13], [16], and [18].

**2.3. Stability, stabilizability, and controllability.** System (2.1) is said to be *stable* if every solution  $x(t)$  of the free system ( $u(t) \equiv 0$ ) tends to zero as  $t$  goes to infinity. Equivalently,  $\det \Delta(\lambda) = 0$  implies that  $\operatorname{Re} \lambda < 0$  for  $\lambda \in \mathbb{C}$  (see, for instance, [9]). Note that  $\sigma(A) = \sigma(A_T^*) = \{\lambda \in \mathbb{C} \mid \det \Delta(\lambda) = 0\}$ . Moreover, system (2.1) is said to be *stabilizable* if

$$(2.7) \quad \operatorname{rank} [\Delta(\lambda), B_0] = n \quad \text{for } \operatorname{Re} \lambda \geq 0$$

and *detectable* if

$$(2.8) \quad \operatorname{rank} \begin{bmatrix} \Delta(\lambda) \\ C_0 \end{bmatrix} = n \quad \text{for } \operatorname{Re} \lambda \geq 0.$$

An abstract Cauchy problem is said to be *observable* if a nonzero initial state produces a nonzero output. Hence the Cauchy problem  $(\Sigma)$  is observable if and only if

$$y(t) = 0 \quad \text{for } t \geq 0 \quad \text{implies } x(t) = 0 \quad \text{for all } t \geq -h$$

for every solution of (2.1) and the Cauchy problem  $(\Sigma_T^*)$  is observable if and only if

$$y(t) = 0 \quad \text{for } t \geq -h \quad \text{implies } x(t) = 0 \quad \text{for all } t \geq -h$$

(see [17]). These two properties have been characterized as follows [13], [14], [17].

**THEOREM 2.1.** *System  $(\Sigma)$  is observable if and only if*

$$\operatorname{rank} \begin{bmatrix} \Delta(\lambda) \\ C_0 \end{bmatrix} = \operatorname{rank} A_1 = n \quad \text{for all } \lambda \in \mathbb{C};$$

$(\Sigma_T^*)$  is observable if and only if

$$\operatorname{rank} \begin{bmatrix} \Delta(\lambda) \\ C_0 \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A_1 \\ C_0 \end{bmatrix} = n \quad \text{for all } \lambda \in \mathbb{C}.$$

If (2.7) and (2.8) are satisfied, then there exist unique nonnegative, selfadjoint operators  $\Pi$ ,  $P \in L(M^2)$  satisfying  $\operatorname{range} \Pi \subset \operatorname{dom} A^*$ ,  $\operatorname{range} P \subset \operatorname{dom} A_T$ , and the algebraic Riccati operator equations

$$(2.9) \quad A^* \Pi \phi = \Pi A \phi - \Pi B B^* \Pi \phi + C^* C \phi = 0,$$

$$(2.10) \quad A_T P f + P A_T^* f - P B B^* P f + C^* C f = 0$$

for  $\phi \in \operatorname{dom} A$  and  $f \in \operatorname{dom} A_T^*$  (see [4], [20]). It follows from (2.5) that the solution operators  $\Pi$  of (2.9) and  $P$  of (2.10) satisfy the identity

$$(2.11) \quad \Pi = F^* P F.$$

This was first observed in [5] for the Riccati differential equation. Finally, we point out that  $\Pi$  is injective if and only if  $(\Sigma)$  is observable, and that  $P$  is injective if and only if  $(\Sigma_T^*)$  is observable.

For a detailed discussion of the Riccati equation and its connection to optimal control theory see [4], [7], and [20].

### 3. Spline approximation.

**3.1. Notation and terminology.** Consider the finite-dimensional linear subspace

$$X^N = \left\{ \phi \in M^2 \mid \phi^1 = \sum_{j=0}^N e_j^N z_j, z_j \in \mathbb{R}^n \right\},$$

where the scalar functions  $e_j^N(\cdot) \in L^2(-h, 0)$  are defined by

$$e_0^N(\tau) = \begin{cases} \frac{N}{h}(\tau - t_1^N), & t_1^N \leq \tau < t_0^N, \\ 0 & \text{elsewhere,} \end{cases}$$

$$e_j^N(\tau) = \begin{cases} \frac{N}{h}(\tau - t_{j+1}^N), & t_{j+1}^N \leq \tau \leq t_j^N, \\ -\frac{N}{h}(\tau - t_{j-1}^N), & t_j^N \leq \tau \leq t_{j-1}^N, \\ 0 & \text{elsewhere,} \end{cases}$$

$$e_N^N(\tau) = \begin{cases} -\frac{N}{h}(\tau - t_{N-1}^N), & t_N^N \leq \tau \leq t_{N-1}^N, \\ 0 & \text{elsewhere,} \end{cases}$$

for  $j=1, \dots, N-1$  and meshpoints  $t_j^N = -jh/N$  for  $j=0, \dots, N$ . Note that the function component of every  $\phi \in X^N$  is a piecewise linear spline function on the interval  $[-h, 0]$ . The subspace  $X^N \subset M^2$  can be identified with the Euclidean space  $\mathbb{R}^{k(N)}$ ,  $k(N) = n + (N+1)n$ , via the embedding  $\iota^N : \mathbb{R}^{k(N)} \rightarrow M^2$  defined by

$$(3.1) \quad \iota^N z = \left( z_0, \sum_{j=0}^N e_j^N z_j \right)$$

for  $z = \text{col}(z_0, z_1) \in \mathbb{R}^{k(N)}$ , where  $z_0 \in \mathbb{R}^n$  and  $z_1 = \text{col}(z_{10}, \dots, z_{1N})$ ,  $z_{ij} \in \mathbb{R}^n$ ,  $j = 0, \dots, N$ . On  $\mathbb{R}^{k(N)}$  we will always consider the induced inner product

$$(3.2) \quad \langle w, z \rangle_N = w^T Q^N z = \langle \iota^N w, \iota^N z \rangle_{M^2},$$

where

$$Q^N = \begin{pmatrix} 1 & 0 \\ 0 & (h/N)q^N \end{pmatrix} \otimes I,$$

$$q^N = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & 0 & \cdots & 0 \\ \frac{1}{6} & \frac{2}{3} & & & \\ 0 & & \ddots & & 0 \\ & & & \frac{2}{3} & \frac{1}{6} \\ 0 & & & 0 & \frac{1}{6} \end{pmatrix} \in \mathbb{R}^{(N+1) \times (N+1)}.$$

Here  $I$  denotes the  $n \times n$  identity matrix. The corresponding vector and matrix norms will be denoted by  $\|\cdot\|_N$ . The adjoint operator  $\pi^N = (\iota^N)^* : M^2 \rightarrow \mathbb{R}^{k(N)}$  is then given by

$$\pi^N \phi = (Q^N)^{-1} z, \quad z_0 = \phi^0,$$

$$z_{1j} = \int_{-h}^0 e_j^N(\tau) \phi^1(\tau) d\tau, \quad j = 0, \dots, N,$$

and satisfies the identities

$$(3.3) \quad \pi^N \iota^N = id, \quad \iota^N \pi^N = p^N,$$

where  $p^N : M^2 \rightarrow X^N$  denotes the orthogonal projection. Let the matrices  $H^N \in \mathbb{R}^{k(N) \times k(N)}$ ,  $B^N \in \mathbb{R}^{k(N) \times l}$ , and  $C^N \in \mathbb{R}^{m \times k(N)}$  be defined by

$$H^N = \begin{pmatrix} A_0 & 0 & \cdots & 0A_1 \\ I & & & \\ 0 & & h^N \otimes I & \\ | & & & \\ 0 & & & \end{pmatrix}, \quad B^N = \begin{pmatrix} B_0 \\ 0 \\ | \\ 0 \end{pmatrix},$$

$$C^N = (C_0 \quad 0 \quad \cdots \quad 0),$$

$$h^N = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 & \cdots & 0 \\ \frac{1}{2} & 0 & & & \\ 0 & & & & \\ 0 & & & & \\ 0 & & & & \end{pmatrix} \in \mathbb{R}^{(N+1) \times (N+1)}.$$

Finally, we define

$$A^N = (Q^N)^{-1} H^N, \quad A_T^N = (Q^N)^{-1} H_T^N,$$

$$(A_T^N)^* = (Q^N)^{-1} (H_T^N)^T, \quad (A^N)^* = (Q^N)^{-1} (H^N)^T,$$

where the matrix  $H_T^N \in \mathbb{R}^{(k(N) \times k(N))}$  is obtained from  $H^N$  by transposing the matrices  $A_0, A_1$ .

Now we consider two control systems on the state space  $\mathbb{R}^{k(N)}$ :

$$(\Sigma^N) \quad \dot{z}^N(t) = A^N z^N(t) + B^N u(t), \quad z^N(0) = \pi^N \phi, \quad y^N(t) = C^N z^N(t),$$

$$(\Sigma_T^{N*}) \quad \dot{w}^N(t) = (A_T^N)^* w^N(t) + B^N u(t), \quad w^N(0) = \pi^N f, \quad y^N(t) = C^N w^N(t).$$

In [11] we establish the following convergence theorem.

**THEOREM 3.1.** (i) For every  $\phi \in M^2$  we have  $\phi = \lim_{N \rightarrow \infty} p^N \phi$ .

(ii)  $B = \iota^N B^N$ ,  $C = C^N \pi^N$  for every  $N \in \mathbb{N}$ .

(iii) There exist constants  $M \geq 1$ ,  $\omega \geq 0$  such that

$$\|e^{A^N t}\|_N \leq M e^{\omega t}, \quad \|e^{(A_T^N)^* t}\|_N \leq M e^{\omega t}$$

for every  $t \geq 0$  and every  $N \in \mathbb{N}$ .

(iv) For all  $\phi, f \in M^2$ ,

$$S(t)\phi = \lim_{N \rightarrow \infty} \iota^N e^{A^N t} \pi^N \phi, \quad S_T^*(t)f = \lim_{N \rightarrow \infty} \iota^N e^{(A_T^N)^* t} \pi^N f$$

and the limits are uniform on every compact time interval  $[0, T]$ .

In particular, this implies that for every  $\phi \in M^2$  and every input  $u(\cdot) \in L^2(0, T; \mathbb{R}^l)$  we have  $z(t) = \lim_{N \rightarrow \infty} \iota^N z^N(t)$  for  $0 \leq t \leq T$  (uniformly), where  $z(t)$  is the unique weak solution of  $(\Sigma)$  and  $z^N(t)$  satisfies  $(\Sigma^N)$ . In the same manner the solutions  $w^N(t)$  of  $(\Sigma_T^{N*})$  approximate the solution  $w(t)$  of  $(\Sigma_T^*)$ .

In the remainder of this section we will study the structural properties of the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ .

**3.2. The structural operator.** In § 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  $(\Sigma^N)$  and  $(\Sigma_T^N)$ . The first step in this direction is Lemma 3.2.

Suppose  $u(\cdot)$  is an arbitrary control in  $V_{(0,T)}[h]$  and its projection on the closed subspace  $\bar{\mathcal{U}}$  is denoted by  $u_p(\cdot)$ . For each  $\psi$  in  $X$  one has

$$\begin{aligned}(u_p(\cdot), (\mathcal{S}_T B)^*(\cdot)\psi)_{L^2(0,T;U)} &= (u(\cdot), (\mathcal{S}_T B)^*(\cdot)\psi)_{L^2(0,T;U)} \\ &= (h, \psi)_X = (G\varphi, \psi)_X \\ &= ((\mathcal{S}_T B)^*(\cdot)\varphi, (\mathcal{S}_T B)^*(\cdot)\psi)_{L^2(0,T;U)},\end{aligned}$$

or

$$(u_p(\cdot) - (\mathcal{S}_T B)^*(\cdot)\varphi, (\mathcal{S}_T B)^*(\cdot)\psi)_{L^2(0,T;U)} = 0.$$

Since  $\mathcal{U}$  is dense in  $\bar{\mathcal{U}}$  and  $[u_p(\cdot) - (\mathcal{S}_T B)^*(\cdot)\varphi] \in \mathcal{U}$ , we have

$$u_p(\cdot) = (\mathcal{S}_T B)^*(\cdot)\varphi \quad \text{for each } u(\cdot) \in V_{(0,T)}[h].$$

Therefore

$$\|(\mathcal{S}_T B)^*(\cdot)\varphi\|_{L^2(0,T;U)} = \|u_p(\cdot)\|_{L^2(0,T;U)} \leq \|u(\cdot)\|_{L^2(0,T;U)} \quad \text{for each } u(\cdot) \in V_{(0,T)}[h].$$

By uniqueness of the minimum norm optimal control of the linear system (1.5), (3.15) holds and  $u^*(\cdot) \in \mathcal{U}$ .

Since  $u_\varepsilon = u^* - \varepsilon(\varepsilon + \hat{G})^{-1}u^*$  (see (3.10)), (3.14) is equivalent to

$$(3.16) \quad \lim_{\varepsilon \rightarrow 0} \varepsilon(\varepsilon + \hat{G})^{-1}u^* = 0 \quad \text{in } L_2(0, T; U).$$

If we consider another family with parameter  $\varepsilon > 0$  of associated quadratic optimal control problems,

$$J_\varepsilon(v; \varphi) = \|(\mathcal{S}_T B)v - \varphi\|^2 + \varepsilon\|v(\cdot)\|_{L^2(0,T;U)}^2,$$

then  $J_\varepsilon(v; \varphi)$  takes its minimum at  $v = v_\varepsilon$  defined by

$$v_\varepsilon = (\varepsilon + \hat{G})^{-1}(\mathcal{S}_T B)^*\varphi = (\varepsilon + \hat{G})^{-1}u^*.$$

If we can show

$$(3.17) \quad \lim_{\varepsilon \rightarrow 0} \varepsilon\|v_\varepsilon(\cdot)\|_{L^2(0,T;U)}^2 = 0,$$

then

$$\begin{aligned}\lim_{\varepsilon \rightarrow 0} \varepsilon\|v_\varepsilon(\cdot)\|_{L^2(0,T;U)} &\leq \lim_{\varepsilon \rightarrow 0} \max\{\varepsilon; \varepsilon\|v_\varepsilon(\cdot)\|_{L^2(0,T;U)}\} \\ &\leq \lim_{\varepsilon \rightarrow 0} \max\{\varepsilon; \varepsilon\|v_\varepsilon(\cdot)\|_{L^2(0,T;U)}^2\} = 0.\end{aligned}$$

The last equation just is (3.16). The rest is to show (3.17) holds for  $\varphi \in X$ . (Notice, if  $\varphi \in \bar{K}_{(0,T)}$  then (3.17) holds. Here we may show (3.17) holds for any given  $\varphi \in X$ .)

Suppose  $\varphi$  is arbitrarily given in  $X$  and

$$\varphi = \bar{\varphi} + \varphi^\perp,$$

where  $\bar{\varphi} \in \bar{K}_{(0,T)}$  and  $\varphi^\perp \in K_{(0,T)}^\perp$ —the orthogonal complement of  $K_{(0,T)}$ . Since  $X = \bar{K}_{(0,T)} \oplus K_{(0,T)}^\perp$ , one has

$$\|\varphi\|^2 = \|\bar{\varphi}\|^2 + \|\varphi^\perp\|^2 \quad \text{for any } \varphi \in X.$$

Denote

$$\bar{v}_\varepsilon = (\varepsilon + \hat{G})^{-1}(\mathcal{S}_T B)^*\bar{\varphi} \quad \text{and} \quad v_\varepsilon^\perp = (\varepsilon + \hat{G})^{-1}(\mathcal{S}_T B)^*\varphi^\perp.$$

Then

$$v_\varepsilon = \bar{v}_\varepsilon + v_\varepsilon^\perp.$$

**3.3. Criteria for stability, stabilizability, and controllability.** We shall need the following facts on the real  $(N+1) \times (N+1)$ -matrix  $a^N = (q^N)^{-1}h^N$ .

LEMMA 3.5. (a) Let  $\|\cdot\|_N$  be the operator norm corresponding to the vector norm  $|x|_N^2 = \bar{x}^T q^N x$  on  $\mathbb{C}^{N+1}$ . Then  $\|e^{a^N t}\|_N \leq 1$  for  $N = 1, 2, \dots$  and  $t \geq 0$ .

(b) Let  $\mu \in \sigma(a^N)$  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  $\mu \in \sigma(a^N)$  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)  $\operatorname{Re} \mu < 0$  for every  $\mu \in \sigma(a^N)$ .

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

$$(3.7) \quad \begin{aligned} \operatorname{Re}(\bar{x}^T h^N x) &= (\operatorname{Re} x)^T h^N (\operatorname{Re} x) + (\operatorname{Im} x)^T h^N (\operatorname{Im} x) \\ &= -\frac{1}{2}|x_0|^2 - \frac{1}{2}|x_N|^2. \end{aligned}$$

Hence  $a^N$  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  $\exp(a^N t)$ ,  $t \geq 0$ , is a contraction semigroup on  $\mathbb{C}^{N+1}$  supplied with the norm  $|\cdot|_N$  (see, for instance, [15]).

(b) This follows from

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

and the fact that  $\mu = \pm 3$  is not an eigenvalue of  $a^N$ .

(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  $\mu \in \sigma(a^N)$  and  $\operatorname{Re} \mu \geq 0$ . Then there exists an  $x \in \mathbb{C}^{N+1}$ ,  $x \neq 0$ , such that  $(\mu q^N - h^N)x = 0$ . By (3.7) this implies

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

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

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

$$(3.8) \quad \alpha^N(\mu) = \text{col}(\alpha_0^N(\mu), \dots, \alpha_N^N(\mu))$$

as the unique solution of

$$(3.9) \quad (\mu q^N - h^N)\alpha^N(\mu) = \text{col}(1, 0, \dots, 0).$$

The complex  $n \times n$ -matrix

$$(3.10) \quad \Delta^N(\lambda) = \lambda I - A_0 - A_1 \alpha_N^N \left( \frac{\lambda h}{N} \right)$$

plays a role for the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  analogous to that of the characteristic matrix  $\Delta(\lambda)$  for the original systems  $(\Sigma)$  and  $(\Sigma_T^*)$ . In particular, it determines their input-output behaviour (see Proposition 3.7 below).

Next we characterize the resolvent operator  $(\lambda Q^N - H^N)^{-1}$  in terms of the matrices  $\Delta^N(\lambda)$ ,  $F^N$  and  $E^N(\lambda) \in \mathbb{C}^{k(N) \times n}$ ,  $T^N(\lambda) \in \mathbb{C}^{k(N) \times k(N)}$ . The latter are defined as follows:

$$(3.11) \quad E^N(\lambda) = \text{col} \left( 1, \alpha_0^N \left( \frac{\lambda h}{N} \right), \dots, \alpha_N^N \left( \frac{\lambda h}{N} \right) \right) \otimes I,$$

$$(3.12) \quad T^N(\lambda) = \begin{pmatrix} 0 & 0 \\ 0 & ((\lambda h/N) q^N - h^N)^{-1} \end{pmatrix} \otimes I.$$

**PROPOSITION 3.6.** *Let  $\lambda \in \mathbb{C}$  with  $\lambda h/N \notin \sigma(a^N)$  and  $x, z \in \mathbb{C}^{k(N)}$ , where  $x$  is written as  $x = (x_0, x_1)$  with  $x_0 \in \mathbb{R}^n$ , then*

(a)  $(\lambda Q^N - H^N)x = z$  if and only if

$$(3.13.1) \quad x = E^N(\lambda)x_0 + T^N(\lambda)z,$$

$$(3.13.2) \quad \Delta^N(\lambda)x_0 = E^N(\lambda)^T F^N z.$$

(b)  $(\lambda Q^N - (H_T^N)^T)x = z$  if and only if

$$(3.14.1) \quad x = F^N E^N(\lambda)x_0 + T^N(\lambda)^T z,$$

$$(3.14.2) \quad \Delta^N(\lambda)x_0 = E^N(\lambda)^T z.$$

*Proof.* We put  $z = (z_0, z_1)$ ,  $x_1 = (x_{10}, \dots, x_{1N})$ ,  $z_1 = (z_{10}, \dots, z_{1N})$ . It is easy to see that  $(\lambda Q^N - H^N)x = z$  if and only if

$$(3.15.1) \quad (\lambda I - A_0)x_0 - A_1 x_{1N} = z_0,$$

$$(3.15.2) \quad \left( \left( \frac{\lambda h}{N} q^N - h^N \right) \otimes I \right) x_1 = z_1 + \text{col}(x_0, 0, \dots, 0).$$

Observing that (see (3.9))

$$\left( \frac{\lambda h}{N} q^N - h^N \right)^{-1} = \begin{pmatrix} \alpha_0^N \left( \frac{\lambda h}{N} \right) & * & * \\ \vdots & \hline & \vdots \\ \alpha_N^N \left( \frac{\lambda h}{N} \right) & \cdots & \alpha_0^N \left( \frac{\lambda h}{N} \right) \end{pmatrix},$$

we get from (3.15.2):

$$x_1 = \left( \left( \frac{\lambda h}{N} q^N - h^N \right) \otimes I \right)^{-1} z_1 + E^N(\lambda)x_0,$$

which proves (3.13.1) and

$$x_{1N} = \alpha_N^N \left( \frac{\lambda h}{N} \right) x_0 + \sum_{k=0}^N \alpha_{N-k}^N \left( \frac{\lambda h}{N} \right) z_{1k}.$$

The last expression together with (3.15.1) establishes (3.13.2).

For (b) we observe that  $(\lambda Q^N - (H_T^N)^T)x = z$  is equivalent to

$$(3.16.1) \quad \lambda x_0 - A_0 x_0 - x_{10} = z_0,$$

$$(3.16.2) \quad \left( \left( \frac{\lambda h}{N} q^N - (h^N)^T \right) \otimes I \right) x_1 = z_1 + \text{col}(0, \dots, 0, A_1 x_0).$$

Observing

$$\left( \frac{\lambda h}{N} q^N - (h_T^N)^T \right)^{-1} = \begin{pmatrix} \alpha_0^N \left( \frac{\lambda h}{N} \right) & \cdots & \alpha_N^N \left( \frac{\lambda h}{N} \right) \\ * & \vdots & * \\ * & \vdots & * \\ & & \alpha_0^N \left( \frac{\lambda h}{N} \right) \end{pmatrix},$$

we get from (3.16.2):

$$x = F^N E^N(\lambda) x_0 + T^N(\lambda) z,$$

which proves (3.14.1), and

$$x_{10} = \sum_{k=0}^N \alpha_k^N \left( \frac{\lambda h}{N} \right) z_{1k} + A_1 \alpha_N^N \left( \frac{\lambda h}{N} \right) x_0,$$

which together with (3.16.1) gives (3.14.2).  $\square$

In particular, the previous proposition shows that

$$(\lambda Q^N - H^N)^{-1} = E^N(\lambda) \Delta^N(\lambda)^{-1} E^N(\lambda)^T F^N + T^N(\lambda),$$

$$(\lambda Q^N - (H_T^N)^T)^{-1} = F^N E^N(\lambda) \Delta^N(\lambda)^{-1} E^N(\lambda)^T + T^N(\lambda)^T,$$

and hence

$$(3.17) \quad (\lambda I - A^N)^{-1} = E^N(\lambda) \Delta^N(\lambda)^{-1} E^N(\lambda)^T Q^N F^N + T^N(\lambda) Q^N,$$

$$(3.18) \quad (\lambda I - (A_T^N)^*)^{-1} = F^N E^N(\lambda) \Delta^N(\lambda)^{-1} E^N(\lambda)^T Q^N + T^N(\lambda)^T Q^N$$

provided  $\lambda \notin \sigma(A^N)$  and  $\det \Delta^N(\lambda) \neq 0$ .

**PROPOSITION 3.7.** (a) *The left upper  $n \times n$  block  $X^N(t)$  in the matrix  $e^{A^N t}$  coincides with that of the matrix  $e^{(A^N)^* t}$ . Its Laplace transform is given by  $\Delta^N(\lambda)^{-1}$  (see (3.10) for the definition of  $\Delta^N(\lambda)$ ).*

(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.* Statement (a) is an immediate consequence of (3.17), (3.18), and the special form of the matrices  $E^N(\lambda)$ ,  $T^N(\lambda)$ ,  $f^N$ ,  $Q^N$ . Statements (b) and (c) follow directly from (a).  $\square$

The following characterization of stabilizability and detectability for the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^N)^*$  is precisely the analogue to (2.7) and (2.8).

**THEOREM 3.8.** (a) *For  $\lambda \in \mathbb{C}$  with  $\text{Re } \lambda \geq 0$  the following properties are equivalent:*

- (i)  $\lambda \in \sigma(A^N)$ ;
- (ii)  $\lambda \in \sigma((A_T^N)^*)$ ;
- (iii)  $\det \Delta^N(\lambda) = 0$ .

*In particular, 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  $\text{Re } \lambda \geq 0$ .*

(b) *The system  $(\Sigma^N)$  (or equivalently  $(\Sigma_T^N)^*$ ) is stabilizable if and only if*

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

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

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

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

*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  $\lambda(h/N) \notin \sigma(a^N)$  for  $\operatorname{Re} \lambda \geq 0$  (Lemma 3.5(d)). Therefore, according to Proposition 3.6(a),  $x = \operatorname{col}(x_0, x_1) \in \ker(\lambda Q - H^N)$  is equivalent to  $\Delta^N(\lambda)x_0 = 0$  and  $x = E^N(\lambda)x_0$ . This implies that detectability of  $(\Sigma^N)$  is equivalent to

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

for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ . When we use Proposition 3.6(b), it follows analogously that this condition is also equivalent to detectability of  $(\Sigma_T^{N*})$ .

Statement (b) follows from (c) by duality and statement (a) follows from (b) with  $B_0 = 0$ .  $\square$

**THEOREM 3.9.** (a) *Let  $\lambda \in \mathbb{C}$  be such that  $\lambda(h/N) \notin \sigma(a^N)$ . Then  $\lambda \in \sigma(A^N)$  if and only if  $\det \Delta^N(\lambda) = 0$ . If  $\lambda(h/N) \in \sigma(a^N)$ , then  $\lambda \in \sigma(A^N)$  if and only if  $\det A_1 = 0$ . Moreover,  $\sigma((A_T^N)^*) = \sigma(A^N)$ .*

(b) *System  $(\Sigma^N)$  is controllable if and only if*

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

$$\text{rank } [A_1, b_0] = n.$$

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

$$\text{rank} \begin{pmatrix} \Delta^N(\lambda) \\ C_0 \end{pmatrix} = n \quad \text{for all } \lambda \in \mathbb{C} \setminus \sigma\left(\frac{N}{h} a^N\right),$$

$$\text{rank } A_1 = n.$$

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

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

$$\text{rank } A_1 = n.$$

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

$$\text{rank} \begin{pmatrix} \Delta^N(\lambda) \\ C_0 \end{pmatrix} = n \quad \text{for all } \lambda \in \mathbb{C} \setminus \sigma\left(\frac{N}{h} a^N\right),$$

$$\text{rank} \begin{pmatrix} A_1 \\ C_0 \end{pmatrix} = n.$$

*Proof.* We first prove (c), i.e., we must show that  $\ker(\lambda I - A^N) \cap \ker C^N = \{0\}$  for all  $\lambda \in \mathbb{C}$ . For  $\lambda \notin \sigma((N/h)a^N)$ , or equivalently, for  $\lambda h/N \notin \sigma(a^N)$  we see as in the proof of Theorem 3.8(c) that this is equivalent to  $\text{rank} \begin{pmatrix} \Delta^N(\lambda) \\ C_0 \end{pmatrix} = n$ .

Now let  $\lambda \in \sigma((N/h)a^N)$ . We assume  $\text{rank } A_1 = n$  and take  $x = \text{col}(x_0, x_{10}, \dots, x_{1N}) \in \ker(\lambda I - A^N) \cap \ker C^N$ . Then (3.15.2) implies  $((\lambda h/N)q^N - h^N) \otimes I)x_1 = \text{col}(x_0, 0, \dots, 0)$  and therefore  $x_0 = 0$  by Lemma 3.5(c). This and (3.15.1) imply  $A_1 x_{1N} = 0$ , i.e.,  $x_{1N} = 0$ . Then we get from Lemma 3.5(b)  $x_1 = 0$ , and thus  $x = 0$ .

Conversely assume that  $\ker(\lambda I - A^N) \cap \ker C^N = \{0\}$  and take  $\xi \in \ker A_1$ . According to Lemma 3.5(b) there exists a vector  $x_1 = \text{col}(x_{10}, \dots, x_{1N}) \in \mathbb{C}^{(N+1)n}$  such that  $[(\lambda(h/N)q^N - h^N) \otimes I]x_1 = 0$  and  $x_{1N} = \xi$ . For  $x = \text{col}(0, x_1)$  equations (3.15) (with  $z_0 = 0$ ) imply  $(\lambda Q^N - H^N)x = 0$ . Obviously we have  $C^N x = 0$ . By assumption this implies  $x = 0$ , and thus  $\xi = 0$ . We conclude  $\text{rank } A_1 = n$ . This finishes the proof of statement (c).

It still holds that  $\lambda \in \sigma((N/h)a^N)$ . Assume that  $\ker A_1 \cap \ker c_0 = \{0\}$  and let  $x \in \mathbb{C}^{k(N)}$  satisfy  $(\lambda Q - (H_T^N)^T)x = 0$ ,  $C^N x = 0$ . Then (3.16.2) and Lemma 3.5(c) imply  $A_1 x_0 = 0$ . This together with  $C_0 x_0 = 0$  shows  $x_0 = 0$ . Hence it follows by (3.16.1) that  $x_{10} = 0$ . Finally we get from Lemma 3.5(b) that  $x_1 = 0$ , and thus  $x = 0$ , i.e.,  $\ker(\lambda I - (A^N)^*) \cap \ker C^N = \{0\}$ .

Conversely, suppose that  $\ker(\lambda I - (A_T^N)^*) \cap \ker C^N = \{0\}$  and let  $x_0 \in \ker A_1 \cap \ker C_0$ . By Lemma 3.5(b) there exists a vector  $\alpha = \text{col}(\alpha_0, \dots, \alpha_N) \in \mathbb{C}^{N+1}$  such that

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

We define  $x_1 = \text{col}(x_{10}, \dots, x_{1N}) \in \mathbb{C}^{n(N+1)}$  by

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

for  $k = 0, \dots, N$ . Then it follows from (3.16) that  $x = \text{col}(x_0, x_1) \in \ker(\lambda Q^N - (H_T^N)^T) \cap \ker C^N$ . By assumption this implies  $x = 0$ , and hence  $x_0 = 0$ , i.e.,  $\ker A_1 \cap \ker C_0 = \{0\}$ . 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  $C^N = 0$ , respectively  $C_0 = 0$ .  $\square$

**4. Stability.** It is our goal to prove that stability (respectively, stabilizability, or detectability) of the original system  $(\Sigma)$  implies the corresponding property for the approximating system  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  provided  $N$  is sufficiently large. The first step in this direction was the characterization of these properties in Theorem 3.8 using the matrices  $\Delta^N(\lambda)$ . The second step will be a convergence result for the characteristic matrices  $\Delta^N(\lambda)$ . As a third step we need a priori bounds for the unstable eigenvalues of the matrices  $A^N$ .

**4.1. Convergence of  $\Delta^N(\lambda)$ .** First we derive explicit formulas for the  $\alpha_k^N(\mu)$  (as defined in (3.8) and (3.9)) and use those to prove convergence of  $\Delta^N(\lambda)$  to  $\Delta(\lambda)$ . Let the rational functions  $d_k^N(\mu)$ ,  $k = 0, \dots, N$ , and the polynomials  $p_k(\mu)$ ,  $q_k(\mu)$ ,  $k = -1, \dots, N$ , respectively,  $k = 1, \dots, N$ , be defined recursively by

$$(4.1) \quad \begin{aligned} d_0^N(\mu) &= 2\mu + 3, \\ d_k^N(\mu) &= 4\mu + \frac{9 - \mu^2}{d_{k-1}^N(\mu)}, \quad k = 1, \dots, N-1, \\ d_N^N(\mu) &= 2\mu + 3 + \frac{9 - \mu^2}{d_{N-1}^N(\mu)}, \end{aligned}$$

$$(4.2) \quad \begin{aligned} p_{-1}(\mu) &= 1, & p_0(\mu) &= 2\mu + 3, \\ p_k(\mu) &= 4\mu p_{k-1}(\mu) + (9 - \mu^2)p_{k-2}(\mu), \\ q_k(\mu) &= (2\mu + 3)p_{k-1}(\mu) + (9 - \mu^2)p_{k-2}(\mu). \end{aligned}$$

The function  $w = w(\mu)$  is defined by

$$(4.3) \quad w(\mu) = (9 + 3\mu^2)^{1/2}, \quad \mu \in \mathbb{C},$$

taking the branch that is positive for  $\mu = i\theta$ ,  $\theta \in \mathbb{R}$ ,  $|\theta| < \sqrt{3}$ . Furthermore, we set

$$(4.4) \quad \gamma_0(\mu) = 2\mu + w(\mu), \quad \gamma_1(\mu) = 2\mu - w(\mu), \quad \mu \in \mathbb{C}.$$

Since  $w(\mu)$  is real for  $\mu = i\theta$ ,  $|\theta| \leq \sqrt{3}$ , we obtain

$$(4.5) \quad |\gamma_0(i\theta)| = |\gamma_1(i\theta)| = |3 - i\theta|, \quad |\theta| \leq \sqrt{3}, \quad \theta \in \mathbb{R}.$$

Therefore the function  $\delta = \delta(\theta)$  is well defined by

$$(4.6) \quad e^{i\delta(\theta)} = \frac{\gamma_1(i\theta)}{\gamma_0(i\theta)} = \frac{7\theta^2 - 9 + 4i\theta w(i\theta)}{9 + \theta^2}, \quad |\theta| \leq \sqrt{3}, \quad \theta \in \mathbb{R},$$

$$0 \leq \delta(\theta) \leq 2\pi.$$

LEMMA 4.1. (a) For  $k = 0, \dots, N$  and  $\mu \notin \sigma(a^N)$ ,

$$(4.7) \quad \alpha_k^N(\mu) = \frac{6(3-\mu)^k}{d_{N-k}^N(\mu) \cdots d_N^N(\mu)} = \frac{6(3-\mu)^k p_{N-k-1}(\mu)}{q_N(\mu)}.$$

(b)  $\det(\mu q^N - h^N) = (1/6^{N+1}) q_N(\mu)$  and

$$2 \cdot 6^{N+1} w \det(\mu q^N - h^N) = (3+w)^2 (\gamma_0)^N - (3-w)^2 (\gamma_1)^N$$

for all  $\mu \in \mathbb{C}$ .

(c) For  $\mu \notin \sigma(a^N)$  and  $\mu \neq \pm i\sqrt{3}$ ,

$$(4.8) \quad \alpha_k^N(\mu) = 6(3-\mu)^k \frac{(3+w)(\gamma_0)^{N-k} - (3-w)(\gamma_1)^{N-k}}{(3+w)^2 (\gamma_0)^N - (3-w)^2 (\gamma_1)^N},$$

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

$$(4.9) \quad |\alpha_{N-k}^N(i\theta)|^2 = 36 \frac{9(1 - \cos k\delta(\theta)) + w^2(i\theta)(1 + \cos k\delta(\theta))}{(9 + w^2(i\theta))^2(1 - \cos N\delta(\theta)) + 36w^2(i\theta)(1 + \cos N\delta(\theta))}$$

for  $k = 0, \dots, N$ ,  $|\theta| < \sqrt{3}$ ,  $\theta \in \mathbb{R}$ .

*Proof.* Suppose that the functions  $d_k = d_k^N(\mu)$ ,  $k = 0, \dots, N$ , are given by (4.1) and define

$$b_k = \frac{-3-\mu}{d_{k-1}}, \quad c_k = \frac{3-\mu}{d_{k-1}}, \quad k = 1, \dots, N.$$

Then it is easy to see that

$$6(\mu q^N - h^N) = \begin{pmatrix} 1 & -b_N & 0 & 0 \\ 0 & 1 & -b_{N-1} & 0 \\ 0 & 0 & 1 & -b_{N-2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} d_N & 0 & \cdots & 0 \\ 0 & d_{N-1} & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & d_0 \end{pmatrix} \\ \times \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -c_N & 1 & \cdots & 0 \\ 0 & -c_{N-1} & 1 & \cdots \\ 0 & 0 & -c_1 & 1 \end{pmatrix}.$$

It is not difficult to calculate the inverse matrices. Since  $\alpha^N(\mu)$  is the first column of  $(\mu q^N - h^N)^{-1}$  (see (3.9)) we conclude that

$$\alpha_k^N(\mu) = \frac{6c_{N-k+1} \cdots c_N}{d_N} = \frac{6(3-\mu)^k}{d_{N-k} \cdots d_N},$$

$k = 0, \dots, N$ .

If the polynomials  $p_k(\mu)$  and  $q_k(\mu)$  are given by (4.2) and the rational functions  $d_k^N(k)$  by (4.1), then we see by induction that

$$\begin{aligned} d_k^N(\mu) &= p_k(\mu)/p_{k-1}(\mu), \quad k = 0, \dots, N-1, \\ d_N^N(\mu) &= q_N(\mu)/p_{N-1}(\mu). \end{aligned}$$

This implies

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

This finishes the proof of (a) and also establishes the first part of (b).

To prove (c) we choose  $\mu \in \mathbb{C}$ ,  $\mu \neq \pm i\sqrt{3}$ . 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 (4.2). This implies that

$$(4.10) \quad p_k(\mu) = \frac{3+w}{2w} (\gamma_0)^{k+1} - \frac{3-w}{2w} (\gamma_1)^{k+1},$$

$k = -1, 0, \dots, N-1$ . Using  $\gamma_0\gamma_1 = \mu^2 - 9$ , we get from (4.2) and (4.10) that

$$\begin{aligned} (4.11) \quad q_N(\mu) &= (2\mu + 3)p_{N-1}(\mu) + (9 - \mu^2)p_{N-2}(\mu) \\ &= \frac{3+w}{2w} [(2\mu + 3)(\gamma_0)^N + (9 - \mu^2)(\gamma_0)^{N-1}] \\ &\quad - \frac{3+w}{2w} [(2\mu + 3)(\gamma_1)^N + (9 - \mu^2)(\gamma_1)^{N-1}] \\ &= \frac{(3+w)^2}{2w} (\gamma_0)^N - \frac{(3-w)^2}{2w} (\gamma_1)^N. \end{aligned}$$

The second part of (b) and (4.8) are immediate consequences of (4.10), (4.11).

To prove (4.9) we use (4.8) and observe  $\gamma_1(i\theta) = e^{i\delta(\theta)}\gamma_0(i\theta)$  and (4.5).  $\square$

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 4.2.**  $\Delta(\lambda) = \lim_{N \rightarrow \infty} \Delta^N(\lambda)$ ,  $\lambda \in \mathbb{C}$ , *the limit being uniform on bounded subsets of  $\mathbb{C}$ .*

*Proof.* Fix  $\delta \in (0, \sqrt{3})$ . Then  $w(\mu)$  as defined in (4.3) is continuous and  $|w(\mu) + 3| \geq 3$  on  $|\mu| \leq \delta$ . From  $w(\mu) - 3 = 3\mu^2/(w(\mu) + 3)$  we see that

$$(4.12) \quad |w(\mu) - 3| \leq |\mu|^2 \quad \text{if } |\mu| \leq \delta.$$

In the next step we prove that  $\alpha_N^N(\mu/N)$  converges for arbitrary  $c > 0$  uniformly to  $e^{-\mu}$  on  $|\mu| \leq c$  as  $N \rightarrow \infty$ . To this end we use formula (4.8) for  $k = N$  and obtain, with  $w = w(\mu/N)$ ,  $N \geq c\delta^{-1}$ ,

$$(4.13) \quad \alpha_N^N\left(\frac{\mu}{N}\right)^{-1} = \frac{(3+w)^2}{12w} \left(w + 2\frac{\mu}{N} \middle/ 3 - \frac{\mu}{N}\right)^N - \frac{(3-w)^2}{12w} \left(-w - 2\frac{\mu}{N} \middle/ 3 - \frac{\mu}{N}\right)^N.$$

From (4.12) and  $\lim_{N \rightarrow \infty} w(\mu/N) = 3$  uniformly on  $|\mu| \leq c$  we see that

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

uniformly on  $|\mu| \leq c$ . Moreover, we also obtain from (4.12):

$$\begin{aligned}\frac{w+2\mu/N}{3-\mu/N} &= 1 + \frac{\mu}{N} + \frac{w-3+(\mu/N)^2}{3-\mu/N} = 1 + \frac{\mu}{N} + O\left(\frac{1}{N^2}\right), \\ \frac{w-2\mu/N}{3-\mu/N} &= 1 - \frac{\mu}{3N} + \frac{w-3-\frac{1}{3}(\mu/N)^2}{3-\mu/N} = 1 - \frac{\mu}{3N} + O\left(\frac{1}{N^2}\right)\end{aligned}$$

as  $N \rightarrow \infty$  uniformly on  $|\mu| \leq c$ . These relations together with (4.13) show

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

uniformly on  $|\mu| \leq c$ . Finally, the theorem follows from (3.10).  $\square$

**4.2. Uniform bounds.** We first establish bounds for the  $\alpha_N^N(\mu)$  in  $\operatorname{Re} \mu \geq 0$ .

LEMMA 4.3. *The estimate  $|\alpha_N^N(\mu)| \leq 2$  is valid for all  $\mu \in \mathbb{C}$  with  $\operatorname{Re} \mu \geq 0$  and all  $N = 1, 2, \dots$ .*

*Proof.* Since, according to Lemmas 3.5(d) and 4.1(b), the polynomial  $q_N(\mu)$  is stable,  $\alpha_N^N(\mu)$  is a proper rational function without poles in  $\operatorname{Re} \mu \geq 0$  (cf. (4.7)). It follows from the maximum principle for analytic functions that  $|\alpha_N^N(\mu)|$  achieves its maximum value in  $\operatorname{Re} \mu \geq 0$  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  $N$ .

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

$$(4.14.1) \quad |d_k^N(\mu)| \geq |3 - \mu|, \quad k = 0, \dots, N-1,$$

$$(4.14.2) \quad |d_N^N(\mu)| \geq 3$$

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 (4.1), assuming that the estimate is already established for  $k$ :

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

This proves (4.14.1). To prove (4.14.2) we note that  $\operatorname{Re} d_k^N(\mu)$  is always positive (and decreasing with respect to  $k$ ) because

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

Therefore the last equation in (4.1) 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 (4.14.2). Now it follows from (4.14) and (4.7) that

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

It remains to consider  $\mu = i\theta$  with  $|\theta| < \sqrt{3}$ ,  $\theta \in \mathbb{R}$ .

Then we obtain from (4.9), with  $k = N$ ,

$$|\alpha_N^N(\mu)|^2 = 36w^2 \left/ \left( (9+w^2)^2 \frac{1-\cos N\delta}{2} + 36w^2 \frac{1+\cos N\delta}{2} \right) \right. \leq 1,$$

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

From Lemma 4.3 we obtain the following a priori bounds for the unstable eigenvalues of the matrices  $A^N$  and  $(A_T^N)^*$ .

**PROPOSITION 4.4.** *Let  $\omega = \|A_0\| + 2\|A_1\|$ . For every  $N \in \mathbb{N}$  and every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  and  $\det \Delta^N(\lambda) = 0$  we have  $|\lambda| \leq \omega$ .*

*Proof.* It follows from Lemma 4.3 that

$$\left\| A_0 + A_1 \alpha_N^N \left( \frac{\lambda h}{N} \right) \right\| \leq \|A_0\| + 2\|A_1\| = \omega$$

for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  and every  $N \in \mathbb{N}$ . Therefore we obtain from (3.10) that  $\|\lambda I - \Delta^N(\lambda)\| \leq \omega$  for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  and every  $N \in \mathbb{N}$ . Hence  $\det \Delta^N(\lambda) \neq 0$  for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  and  $|\lambda| > \omega$ . This proves the statement of the proposition.  $\square$

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

### 4.3. Stability, stabilizability, and detectability.

**THEOREM 4.5.** *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.* Suppose  $(\Sigma)$  is stable. Then  $\det \Delta(\lambda) \neq 0$  for every  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$ . Hence the uniform convergence result for  $\Delta^N(\lambda)$  on bounded domains (Theorem 4.2) shows that  $\det \Delta^N(\lambda) \neq 0$  for  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  and  $|\lambda| \leq \omega$  provided  $N$  is sufficiently large. If  $\omega > 0$  is large enough then we obtain from Proposition 4.4 that  $\det \Delta^N(\lambda) \neq 0$  for all  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  provided  $N$  is sufficiently large. Now the stability of  $(\Sigma^N)$  follows from Theorem 3.8(a). This proves (a). Statement (b) can be established analogously.  $\square$

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

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

for  $N$  sufficiently large. A result of this type would be needed to apply a result of Gibson [8] concerning the approximation of the solution to the algebraic Riccati equation. Moreover, the uniform stability has been stated as a conjecture in [3] for the spline approximation scheme developed in [2]. 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 [2].

**PROPOSITION 4.6.** *Suppose that there exist constants  $M \geq 1$  and  $\varepsilon_N > 0$  such that*

$$\|\exp(Na^N t)\|_N \leq M e^{-\varepsilon_N t}, \quad t \geq 0,$$

*for all  $N$ . Then  $\varepsilon_N = o(1/N^{1/2})$ . Here  $\|\cdot\|_N$  denotes the operator norm corresponding to the vector norm  $|x|_N^2 = (1/N)x^T q^N x$  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. (3.9)),

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

Therefore

$$(4.15) \quad \varepsilon_N \leq 6NM \left( \sum_{k=0}^N |\alpha_k^N(\mu)|^2 \right)^{-1/2}$$

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

Now let  $\mu = i\theta$  satisfy  $|\theta| < \sqrt{3}$ ,  $\theta \in \mathbb{R}$ . Since  $\delta(i\theta) \rightarrow 0$  as  $|\theta| \rightarrow \sqrt{3}$ , we can choose a sequence  $\theta_N \in \mathbb{R}$ ,  $|\theta_N| < \sqrt{3}$ , such that  $|\theta_N| \rightarrow \sqrt{3}$  and  $\delta_N = \delta(i\theta_N) = 2\pi/N$ . We put  $w_N = w(i\theta_N)$  and get, using (4.9) and  $\sum_{k=0}^N \cos(2k\pi/N) = 1$ ,

$$\begin{aligned} (4.16) \quad \sum_{k=0}^N \left| \alpha_{N-k}^N(i\theta_N) \right|^2 &= \frac{9}{w_N^2} \sum_{k=0}^N \left( 1 - \cos \frac{2k\pi}{N} \right) + \frac{1}{2} \sum_{k=0}^N \left( 1 + \cos \frac{2k\pi}{N} \right) \\ &\geq \frac{9}{2w_N^2} N. \end{aligned}$$

From (4.6) we get

$$\sin \delta(\theta) = \frac{4\theta w(i\theta)}{9 + \theta^2},$$

which shows that for positive constants  $c_1, c_2$ ,

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

This and (4.16) imply

$$\sum_{k=0}^N |\alpha_k^N(\mu_N)|^2 \geq \text{const. } N^3.$$

This last estimate and (4.15) show that

$$\varepsilon_N \leq \text{const. } \frac{1}{N^{1/2}}. \quad \square$$

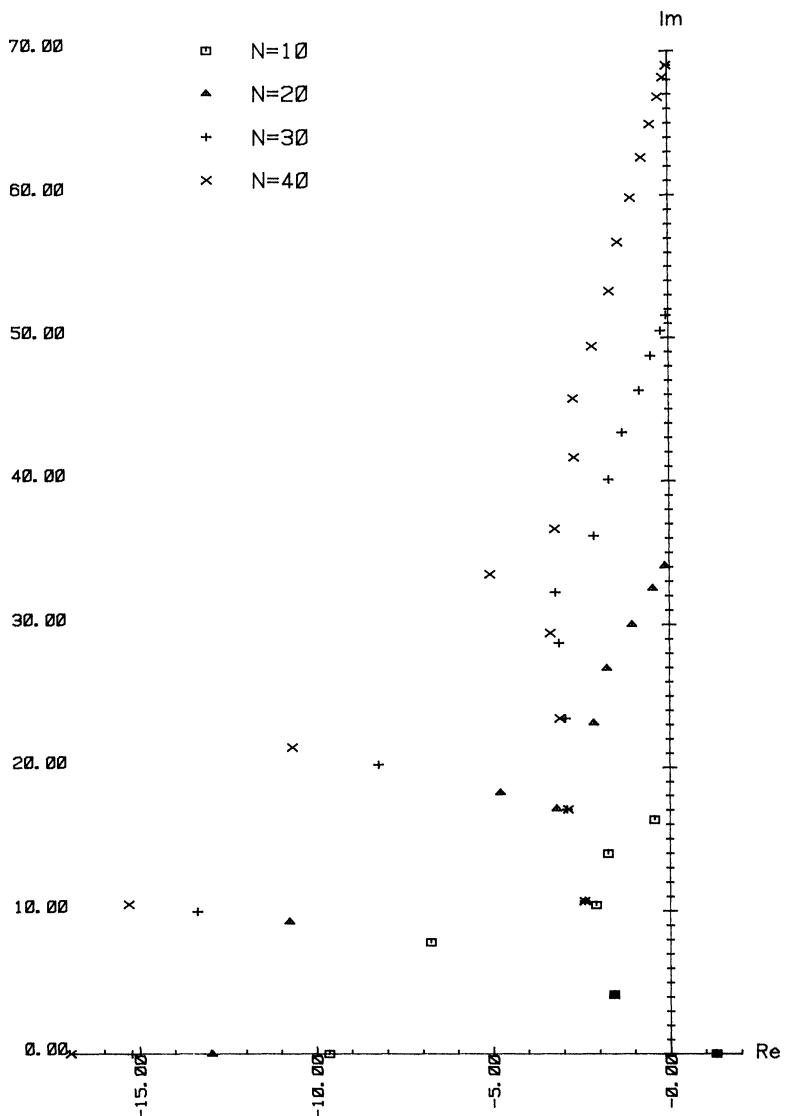


FIG. 1

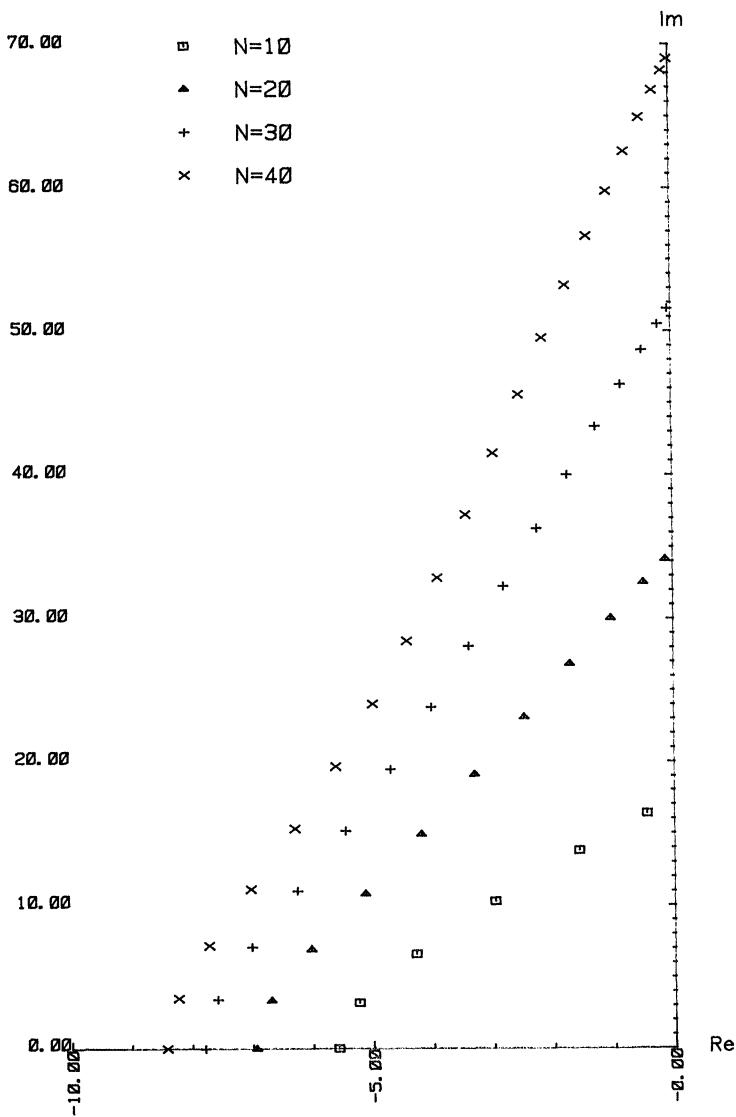


FIG. 2

The result above shows that exponential stability uniform with respect to  $N$  is, in general, impossible for our scheme. Numerical studies show that there is a sequence  $\lambda_N$ ,  $N = 1, 2, \dots$ , of eigenvalues  $\lambda_N \in \sigma((N/h)a^N)$  such that  $\operatorname{Re} \lambda_N \rightarrow 0$  and  $\operatorname{Im} \lambda_N \rightarrow \infty$ . In fact, the numerical results indicate  $\operatorname{Re} \lambda_N = O(1/N^2)$ . In the general case where  $\sigma((N/h)a^N)$  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  $\operatorname{Re} \lambda_N < 0$ ,  $\operatorname{Re} \lambda_N \rightarrow 0$  and  $\operatorname{Im} \lambda_N \rightarrow \infty$ . Figure 1 illustrates the location of the spectrum for the approximating systems in case of the scalar equation  $\dot{x}(t) = x(t) + x(t-1)$ . For comparison Fig. 2 illustrates the spectrum of  $Na^N$ .

**5. Uniform output stability.** Despite the negative result of Proposition 4.6 and the fact that some eigenvalues of the approximating systems approach the imaginary axis, we are still able to prove a uniform  $L^2$ -estimate for the  $\mathbb{R}^n$ -components

$$z_0^N(t; \phi) \in \mathbb{R}^n, \quad w_0^N(t; f) \in \mathbb{R}^n$$

of the unique solutions of  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  (with  $u \equiv 0$ ). We call this property the uniform output stability of the systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$ .

**THEOREM 5.1.** *Suppose system  $(\Sigma)$  is stable. Then the approximating systems  $(\Sigma^N)$  and  $(\Sigma_T^{N*})$  are uniformly output stable for  $N$  sufficiently large, i.e., there exists an  $N_0 \in \mathbb{N}$  and a constant  $c > 0$  such that for  $N \geq N_0$*

$$(5.1) \quad \int_0^\infty |z_0^N(t; \phi)|_{\mathbb{R}^n}^2 dt \leq c \|\phi\|_{M^2}^2 \quad \text{for all } \phi \in M^2,$$

$$(5.2) \quad \int_0^\infty |w_0^N(t; f)|_{\mathbb{R}^n}^2 dt \leq c \|f\|_{M^2}^2 \quad \text{for all } f \in M^2.$$

*Proof.* Choose  $N_0 \in \mathbb{N}$  such that  $\det \Delta^N(\lambda) \neq 0$  for all  $\lambda \in \mathbb{C}$  with  $\operatorname{Re} \lambda \geq 0$  and all  $N \geq N_0$  (Theorem 4.5). Then it follows from (3.17), (3.18), and Lemma 3.2 that the Fourier transforms of  $z_0^N(t; \phi)$  and  $w_0^N(t; f)$  (determined to be identically zero for  $t < 0$ ) are given by

$$\hat{z}_0^N(i\omega; \phi) = \Delta^N(i\omega)^{-1} E^N(i\omega)^T Q^N \pi^N F \phi, \quad \omega \in \mathbb{R},$$

$$\hat{w}_0^N(i\omega; \phi) = \Delta^N(i\omega)^{-1} E^N(i\omega)^T Q^N \pi^N f, \quad \omega \in \mathbb{R},$$

for  $N \geq N_0$ . Using Plancherel's theorem, we see that to prove (5.1) and (5.2) it is enough to show

$$\int_{-\infty}^\infty |\Delta^N(i\omega)^{-1} E^N(i\omega)^T Q^N z|_{\mathbb{R}^n}^2 dt \leq 2\pi c z^T Q^N z$$

for all  $z \in \mathbb{R}^{k(N)}$  and all  $N \geq N_0$ . The definition of  $E^N(\lambda)$  in (3.11), together with Lemma 4.3, show that it suffices to prove a uniform estimate of the form

$$(5.3) \quad \frac{1}{N} \int_{-\infty}^\infty \|\Delta^N(i\omega)^{-1}\|^2 \sum_{k=0}^N \left| \alpha_k^N \left( \frac{i\omega h}{N} \right) \right|^2 d\omega \leq c$$

for all  $N \geq N_0$  (with a possibly different constant  $c$ ). Of course, it is only necessary to consider  $\omega \geq 0$ .

Using (3.10) and Lemma 4.3 we immediately get the estimate

$$\|\Delta^N(i\omega)^{-1}\| \leq \frac{1}{|\omega| - c_0} \quad \text{for } |\omega| > c_0, \quad N = 1, 2, \dots,$$

where  $c_0 = \|A_0\| + 2\|A_1\|$ . By Theorem 4.2 and the stability assumption on  $(\Sigma)$  we obtain

$$(5.4) \quad \|\Delta^N(i\omega)^{-1}\|^2 \leq \frac{c_1}{1 + \omega^2} \quad \text{for all } \omega \in \mathbb{R}$$

and  $N$  sufficiently large, where  $c_1$  is not dependent on  $N$ .

Defining

$$f^N(\theta) = \frac{1}{N} \sum_{k=0}^N |\alpha_k^N(i\theta)|^2,$$

we find that for all  $N = 1, 2, \dots$

$$(5.5) \quad f^N(\theta) \leq \frac{N+1}{N} 4 \leq 8 \quad \text{for } \theta \geq \sqrt{3}$$

and for any  $\alpha \in (0, 1)$

$$(5.6) \quad f^N(\theta) \leq \frac{N+1}{N} \frac{2}{1-\alpha^2} \leq \frac{4}{1-\alpha^2} \quad \text{for } 0 \leq \theta \leq \alpha\sqrt{3}.$$

The estimate (5.5) is a straightforward consequence of (4.7) and estimates (4.14). To obtain (5.6) we can use the representation (4.9) and the estimates  $w^2(i\theta) = 9 - 3\theta^2 \geq 9(1 - \alpha^2)$ ,  $w^2(i\theta) \leq 9$ , and  $9 + w^2(i\theta) \geq 6w(i\theta)$  for  $0 \leq \theta \leq \alpha\sqrt{3}$ .

It remains to investigate the behaviour of  $f^N(\theta)$  at intervals of the form  $(\alpha\sqrt{3}, \sqrt{3})$ ,  $0 < \alpha < 1$ . There we cannot expect to have a bound for  $f^N(\theta)$  uniformly with respect to  $N$ . Formula (4.9) shows that we should expect difficulties for those  $\theta$  near  $\sqrt{3}$  such that  $N\delta(\theta)$  is close to an integer multiple of  $2\pi$ . This reflects the fact that the eigenvalues of  $a^N$  are closest to the imaginary axis near  $\pm iN\sqrt{3}$  (see Fig. 2), i.e., for  $\theta = \omega h/N$  close to  $\pm\sqrt{3}$ . In Fig. 3 we show the plot for  $f^N(\omega h/N)$ ,  $N = 10, 20, 30, 40$ ,  $h = 1$ , which illustrates the difficulties.

We first determine those parts of  $(\alpha\sqrt{3}, \sqrt{3})$ , where we still can find a uniform bound for  $f^N(\theta)$ . Since  $\alpha$  is not yet fixed we consider  $\theta \in [0, \alpha\sqrt{3}]$ .

**CLAIM 1.** *If  $\theta \in [0, \sqrt{3}]$  is such that  $0 \leq \delta(\theta) \leq \pi/3N$ , then*

$$(5.7) \quad f^N(\theta) \leq 8 \quad \text{for all } N.$$

*Proof.* From  $\frac{1}{2} \leq \cos N\delta(\theta) \leq \cos k\delta(\theta)$ ,  $k = 0, \dots, N$ , and (4.9) we get

$$\begin{aligned} |\alpha_{N-k}^N(i\theta)|^2 &\leq 4 \frac{9(1 - \cos N\delta(\theta)) + w^2(i\theta)(1 + \cos k\delta(\theta))}{9(1 - \cos N\delta(\theta)) + 4w^2(i\theta)(1 + \cos N\delta(\theta))} \\ &\leq 4 \frac{9(1 - \cos N\delta(\theta)) + 2w^2(i\theta)}{9(1 - \cos N\delta(\theta)) + 6w^2(i\theta)} \leq 4, \end{aligned}$$

for  $k = 0, \dots, N$ , which implies the result.  $\square$

**CLAIM 2.** *If  $\theta \in [0, \sqrt{3}]$  is such that  $|\delta(\theta) - 2\pi(\nu/N)| \geq \pi/3N$  for  $\nu = 0, \dots, [N/2]$ , then*

$$(5.8) \quad f^N(\theta) \leq 64 \quad \text{for all } N.$$

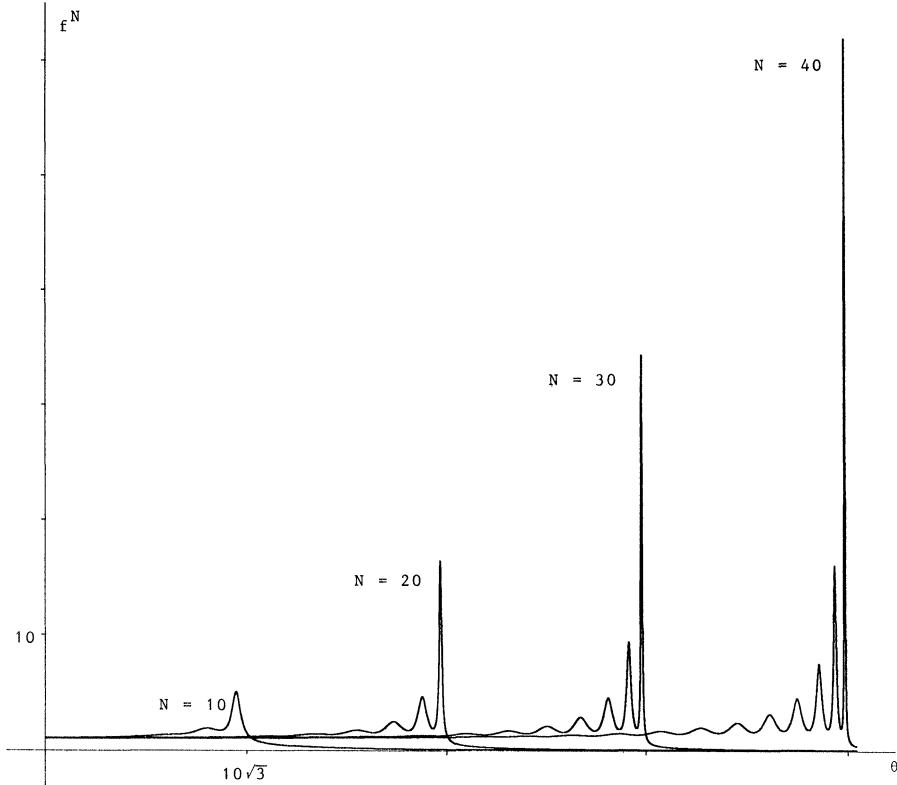


FIG. 3

*Proof.* From  $\cos N\delta(\theta) \leq \cos(\pi/3) = \frac{1}{2}$ ,  $w^2(i\theta) \leq 9$  and (4.9) we obtain

$$|\alpha_{N-k}^N(i\theta)|^2 \leq 36 \frac{2 \cdot 9 + 2 \cdot 9}{9^2(1 - \cos N\delta(\theta))} \leq 36, \quad k = 0, \dots, N.$$

For  $\nu, N \in \mathbb{N}$  and  $\alpha \in (0, 1)$  we define the intervals

$$I_\nu^N = \left\{ \theta \in [\alpha\sqrt{3}, \sqrt{3}] \mid |N\delta(\theta) - 2\pi\nu| < \frac{\pi}{3} \right\}.$$

We have  $I_\nu^N \neq \emptyset$  if and only if  $\nu = 0, \dots, \nu_1$ , where  $\nu_1 < N$  is determined by the conditions  $2\pi\nu_1 - \pi/3 < N\delta(\alpha\sqrt{3})$  and  $2\pi(\nu_1 + 1) - \pi/3 \geq N\delta(\alpha\sqrt{3})$ . Inequalities (5.5)–(5.8) imply that for any  $\alpha \in (0, 1)$  there exists a constant  $c = c(\alpha)$  independent of  $N$  such that

$$(5.9) \quad f^N(\theta) \leq c(\alpha), \quad \theta \in M := [0, \infty) \setminus \bigcup_{\nu=1}^{\nu_1} I_\nu^N$$

for all  $N \in \mathbb{N}$ . Let  $\tilde{M} = \{\omega \geq 0 \mid \omega h/N \in M\}$ . Then by (5.4) and (5.9)

$$(5.10) \quad \int_{\tilde{M}} \|\Delta^N(i\omega)^{-1}\|^2 f^N\left(\frac{\omega h}{N}\right) d\omega \leq c(\alpha) c_1 \frac{\pi}{2}.$$

Let  $\tilde{I}_\nu^N = \{\omega \geq 0 \mid (\omega h/N) \in I_\nu^N\}$ ,  $\nu, N \in \mathbb{N}$ . Then by (5.4) (note, that  $\omega > \alpha_0\sqrt{3}(N/h)$  for  $\omega \in \tilde{I}_\nu^N$ )

$$(5.11) \quad \begin{aligned} J &= \sum_{\nu=1}^{\nu_1} \int_{I_\nu^N} \|\Delta^N(i\omega)^{-1}\|^2 f^N\left(\frac{\omega h}{N}\right) d\omega \\ &\leq \frac{c_1}{1+3\alpha_0^2 N^2/h^2} \sum_{\nu=1}^{\nu_1} \frac{N}{h} \int_{I_\nu^N} f^N(\theta) d\theta \end{aligned}$$

for  $N$  sufficiently large. It remains to prove an estimate of the form

$$\sum_{\nu=1}^{\nu_1} \frac{N}{h} \int_{I_\nu^N} f^N(\theta) d\theta \leq c$$

for all  $N = 1, 2, \dots$ , where  $c$  is independent of  $N$ .

Now we fix  $\alpha = \alpha_0$  by imposing the condition

$$(5.12) \quad \delta(\alpha_0\sqrt{3}) = \frac{\pi}{2}.$$

Then there exists a constant  $c_2$  such that

$$(5.13) \quad \frac{1}{c_2} \delta(\theta) \leq w(i\theta) \leq c_2 \delta(\theta), \quad \theta \in [\alpha_0\sqrt{3}, \sqrt{3}].$$

*Proof of (5.13).* From (4.6) we see that  $\sin \delta(\theta) = 4\theta w(i\theta)/(9+\theta^2)$ . Therefore

$$(5.14) \quad \frac{4\alpha_0\sqrt{3}}{12} w(i\theta) \leq \sin \delta(\theta) \leq \frac{4\sqrt{3}}{9} w(i\theta), \quad \theta \in [\alpha_0\sqrt{3}, \sqrt{3}].$$

The monotonicity of  $\delta(\theta)$  and (5.11) imply  $0 \leq \delta(\theta) \leq \pi/2$ . Then  $(2/\pi)\delta(\theta) \leq \sin \delta(\theta) \leq \delta(\theta)$ , which together with (5.14) implies the result.  $\square$

**CLAIM 3.** *There exists a constant  $c_3$  independent of  $N$  and  $\nu$  such that*

$$(5.15) \quad |I_\nu^N| \leq c_3 \frac{\nu}{N^2}$$

for  $N = 1, 2, \dots$  and  $\nu = 1, \dots, \nu_1$ . Here  $|I_\nu^N|$  denotes the length of the interval  $I_\nu^N$ .

*Proof.* From (5.13) we get

$$(5.16) \quad -\frac{1}{w(i\theta)} \leq -\frac{1}{c_2 \delta(\theta)}, \quad \theta \in [\alpha_0\sqrt{3}, \sqrt{3}].$$

The definition (4.6) of  $\delta(\theta)$  together with (4.3) implies

$$\delta'(\theta) = -\frac{36}{(9+\theta^2)w(i\theta)}.$$

Using (5.16), for  $\theta \in I_\nu^N$  we obtain

$$\delta'(\theta) \leq -\frac{36}{9+\theta^2} \cdot \frac{1}{c_2 \delta(\theta)} \leq -\frac{3}{c_2 \delta(\theta)} \leq -\frac{1}{c_2 \pi} \frac{N}{\nu}.$$

Therefore

$$\frac{2\pi}{3N} = -\int_{I_\nu^N} \delta'(\theta) d\theta \geq \frac{1}{c_2 \pi} \frac{N}{\nu} |I_\nu^N|$$

for  $N = 1, 2, \dots$  and  $\nu = 1, \dots, \nu_1$ , which proves the result.  $\square$

**CLAIM 4.** *There exists a constant  $c_4$  independent of  $N$  and  $\nu$  such that*

$$(5.17) \quad f^N(\theta) \leq c_4 \left( \frac{N}{\nu} \right)^2$$

for  $\theta \in I_\nu^N$ ,  $N = 1, 2, \dots$  and  $\nu = 1, \dots, \nu_1$ .

*Proof.* Let  $\theta \in I_\nu^N$ . Then  $\delta(\theta) > (2\pi\nu - \pi/3)/N$  and by (5.13) we obtain

$$w(i\theta) > \frac{1}{c_2} \cdot \frac{1}{N} \left( 2\pi\nu - \frac{\pi}{3} \right) \geq \frac{5\pi}{3c_2} \cdot \frac{\nu}{N}.$$

This together with (4.9),  $w^2(i\theta) \leq 9$ , and  $\cos N\delta(\theta) > \frac{1}{2}$  for  $\theta \in I_\nu^N$  implies

$$|\alpha_{N-k}^N(i\theta)|^2 \leq \frac{18}{(1 + \cos N\delta(\theta))w^2(i\theta)} \leq \frac{108c_2^2}{25\pi^2} \cdot \frac{N^2}{\nu^2}.$$

Then the result follows immediately.  $\square$

Using (5.11), (5.15), and (5.17) we get the following estimate for  $J$ :

$$\begin{aligned} J &\leq \frac{c_1 c_3 c_4}{1 + 3\alpha_0^2 N^2/h^2} \frac{N}{h} \sum_{\nu=1}^{\nu_1} \left( \frac{N}{\nu} \right)^2 \frac{\nu}{N^2} \\ &\leq \frac{c_1 c_3 c_4}{h} \frac{N}{1 + 3\alpha_0^2 N^2/h^2} \sum_{\nu=1}^{\nu_1} \frac{1}{\nu} \leq \frac{c_1 c_2 c_3}{h} \frac{N(1 + \ln N)}{1 + 3\alpha_0^2 N^2/h^2} \end{aligned}$$

for  $N$  sufficiently large. This together with (5.10) establishes (5.3). Thus the proof of Theorem 5.1 is finished.

*Remarks.* (1) Uniform output stability in general does not imply a uniform (with respect to  $N$ ) exponential decay for the  $\mathbb{R}^n$ -components of solutions of the approximating equations. If we are willing to accept the existence of eigenvalues  $\lambda_N$  for the approximating equations with  $\operatorname{Re} \lambda_N \rightarrow 0$  as  $N \rightarrow \infty$  also in case  $\det A_1 \neq 0$  (as is demonstrated numerically in Fig. 1 but not proved in this paper), it is sufficient to show that in case  $\det A_1 \neq 0$  any eigenvector for the approximating system has a nonzero  $\mathbb{C}^n$ -component. To prove this, assume  $y^N = \operatorname{col}(y_0^N, y_1^N)$  with  $y_0^N \in \mathbb{C}^n$ ,  $y_1^N \in \mathbb{C}^{(N+1)n}$  is an eigenvector of  $A^N$  corresponding to the eigenvalue  $\lambda_N$ . Assume  $y_0^N = 0$ . Then  $A^N y^N = \lambda_N y^N$  is equivalent to

$$(5.18) \quad (0 \longrightarrow 0 A_1) y_1^N = 0 \quad \text{and} \quad \frac{N}{h} (a^N \otimes I) y_1^N = \lambda_N y_1^N.$$

The second equation implies  $y_1^N = x \otimes v$ , where  $a^N x = (\lambda h/N)x$ ,  $x = (x_0, \dots, x_N) \in \mathbb{C}^{N+1} \setminus \{0\}$  and  $v \in \mathbb{C}^n \setminus \{0\}$ . By Lemma 3.5(b) we have  $x_N \neq 0$ . The first equation in (5.18) implies  $A_1(x_N v) = x_N A_1 v = 0$ , a contradiction to  $\det A_1 \neq 0$ .

(2) It is interesting to state a consequence of uniform output stability for the eigenvectors of the approximating equations. Assume  $(\Sigma)$  is stable so that (5.1) is true, and let  $y^N = (y_0^N, y_1^N)$ ,  $y_0^N \in \mathbb{C}^n$ ,  $y_1^N \in \mathbb{C}^{(N+1)n}$  be an eigenvector of  $A^N$  corresponding to an eigenvalue  $\lambda_N$ . Then  $z_0^N(t; \iota^N y^N) = y_0^N e^{\lambda_N t}$ ,  $t \geq 0$ , and therefore

$$\int_0^\infty |z_0^N(t; \iota^N y^N)|_{\mathbb{C}^n}^2 dt = \frac{1}{2|\operatorname{Re} \lambda_N|} |y_0^N|_{\mathbb{C}^n}^2.$$

For  $N \geq N_0$ , (5.1) implies

$$|y_0^N|_{\mathbb{C}^n}^2 \leq \frac{2c|\operatorname{Re} \lambda_N|}{1 - 2c|\operatorname{Re} \lambda_N|} \|(\iota^N y^N)\|_{L^2}^2$$

provided  $|\operatorname{Re} \lambda_N| < 1/2c$ . Therefore if  $|\operatorname{Re} \lambda_N| \rightarrow 0$  as  $N \rightarrow \infty$  then also  $y_0^N \rightarrow 0$  in  $\mathbb{C}^n$ .

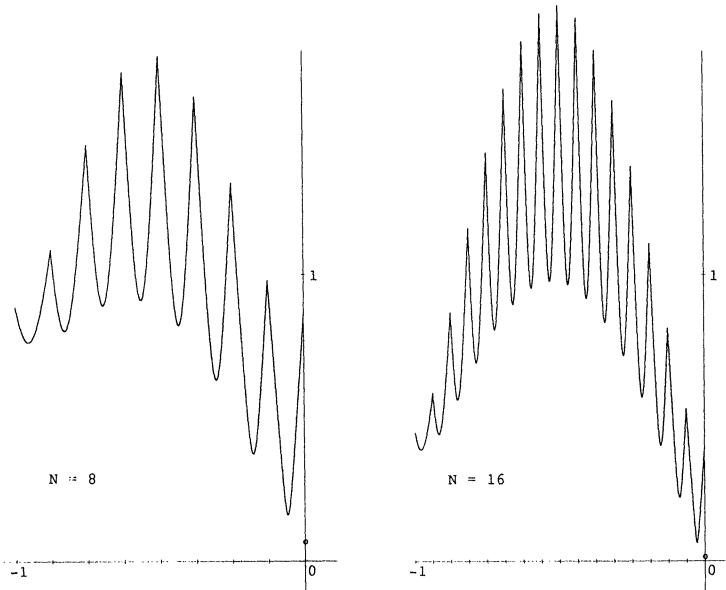


FIG. 4

Figure 4 shows, for  $N = 8, 16$ ,  $|(\iota^N y^N)^1(\theta)|$ ,  $-1 \leq \theta \leq 0$ , and  $|(\iota^N y^N)^0|$  for the normalized eigenvector  $y^N$  (i.e.,  $\|\iota^N y\|_{M^2} = 1$ ) of  $A^N$  corresponding to the eigenvalue  $\lambda_N$  with the smallest real part (and at the same time largest imaginary part) in case of the scalar equation  $\dot{x}(t) = -2x(t) + x(t-1)$ . The eigenvalues  $\lambda_N$  and  $|(\iota^N y^N)^0|$ ,  $N = 4, 8, 16$ , are given by

$$\begin{aligned} \lambda_4 &= -2.9294 + 5.1788i, \quad \lambda_8 = -0.7218 + 12.7848i, \quad \lambda_{16} = -0.1874 + 27.0462i, \\ |(\iota^4 y^4)^0| &= 0.3536, \quad |(\iota^8 y^8)^0| = 0.0690, \quad |(\iota^{16} y^{16})^0| = 0.0167. \end{aligned}$$

(3) Uniform output stability is sufficient to prove convergence of the approximating Riccati operators in the case of the infinite time horizon problem observed numerically in [11]. This will be shown in a forthcoming paper [12].

**Acknowledgment.** We thank W. Prager for the computations concerning Figs. 1-4.

#### REFERENCES

- [1] H. T. BANKS AND J. A. BURNS, *Heredity control problems: numerical methods based on averaging approximation*, SIAM J. Control Optim., 16 (1978), pp. 169-208.
- [2] H. T. BANKS AND F. KAPPEL, *Spline approximations for functional differential equations*, J. Differential Equations, 34 (1979), pp. 496-522.
- [3] 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. Statist. Comput., 5 (1984), pp. 830-855.
- [4] R. F. CURTAIN AND A. J. PRITCHARD, *Infinite Dimensional Linear Systems Theory*, Lecture Notes in Control and Information Sciences 8, Springer-Verlag, Berlin, New York, 1978.
- [5] M. C. DELFOUR, E. B. LEE, AND A. MANITIUS, *F-reduction of the operator Riccati equation*, Automatica, 14 (1978), pp. 385-395.
- [6a] 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), pp. 466-490.
- [6b] ———, *The structural operator F and its role in the theory of retarded systems, Part II*, J. Math. Anal. Appl., 74 (1980), pp. 359-381.

- [7] J. S. GIBSON, *The Riccati integral equations for optimal control problems in Hilbert spaces*, SIAM J. Control Optim., 17 (1979), pp. 537–565.
- [8] ———, *Linear quadratic optimal control of hereditary differential systems: Infinite dimensional Riccati equation and numerical approximations*, SIAM J. Control Optim., 21 (1983), pp. 95–139.
- [9] J. K. HALE, *Theory of Functional Differential Equations*, Springer-Verlag, Berlin, New York, 1977.
- [10] F. KAPPEL AND D. SALAMON, *On the stability properties of spline approximations for retarded systems*, Technical Report No. 78-1986, Institute for Mathematics, University of Graz, Graz, Austria, 1986.
- [11] ———, *Spline approximation for retarded systems and the Riccati equation*, SIAM J. Control Optim., 25 (1987), pp. 1082–1117.
- [12] ———, *An approximation theorem for the algebraic Riccati equation*, SIAM J. Control Optim., submitted.
- [13] A. MANITIUS, *Completeness and F-completeness of eigenfunctions associated with retarded functional differential equations*, J. Differential Equations, 35 (1980), pp. 1–29.
- [14] ———, *Necessary and sufficient conditions of approximate controllability for general linear retarded systems*, SIAM J. Control Optim., 19 (1981), pp. 516–532.
- [15] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, Berlin, New York, 1983.
- [16] D. SALAMON, *On dynamic observation and state feedback for time delay systems*, in Evolution Equations and their Applications, F. Kappel and W. Schappacher, eds., Research Notes in Mathematics 68, Pitman, London, 1982, pp. 202–219.
- [17] ———, *On controllability and observability of time delay systems*, IEEE Trans. Automat. Control, 29 (1984), pp. 432–439.
- [18] ———, *Control and Observation of Neutral Systems*, Research Notes in Mathematics 91, Pitman, London, 1984.
- [19] ———, *Structure and stability of finite dimensional approximations for functional differential equations*, SIAM J. Control Optim., 23 (1985), pp. 928–951.
- [20] J. ZABCZYK, *Remarks on the algebraic Riccati equation in Hilbert space*, Appl. Math. Optim., 2 (1976), pp. 251–258.