

**NASA Contractor Report 181706**

**ICASE REPORT NO. 88-49**

# ICASE

**CONVERGENCE OF GALERKIN APPROXIMATIONS FOR OPERATOR RICCATI  
EQUATIONS -- A NONLINEAR EVOLUTION EQUATION APPROACH**

**I. G. Rosen**

**(NASA-CR-181706) CONVERGENCE OF GALERKIN  
APPROXIMATIONS FOR OPERATOR RICCATI  
EQUATIONS: A NONLINEAR EVOLUTION EQUATION  
APPROACH Final Report (NASA) 30 p CSCL 12A**

**N88-30386**

**Unclas  
G3/64 0164809**

**Contract No. NAS1-18107  
August 1988**

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

**Operated by the Universities Space Research Association**



**National Aeronautics and  
Space Administration**

**Langley Research Center  
Hampton, Virginia 23665**

# CONVERGENCE OF GALERKIN APPROXIMATIONS FOR OPERATOR RICCATI EQUATIONS - A NONLINEAR EVOLUTION EQUATION APPROACH

by

I.G. Rosen\*  
Department of Mathematics  
University of Southern California  
Los Angeles, CA 90089-1113 USA  
Telephone: (213) 743-2567

## Abstract

We develop an approximation and convergence theory for Galerkin approximations to infinite dimensional operator Riccati differential equations formulated in the space of Hilbert-Schmidt operators on a separable Hilbert space. We treat the Riccati equation as a nonlinear evolution equation with dynamics described by a nonlinear monotone perturbation of a strongly coercive linear operator. We prove a generic approximation result for quasi-autonomous nonlinear evolution systems involving accretive operators which we then use to demonstrate the Hilbert-Schmidt norm convergence of Galerkin approximations to the solution of the Riccati equation. We illustrate the application of our results in the context of a linear quadratic optimal control problem for a one dimensional heat equation.

---

\*This research supported in part by the United States Air force Office of Scientific Research under grant AFOSR-87-0356. Part of this research was carried out while the author was a visiting scientist and consultant at the Institute for Computer Applications in Science and Engineering at the NASA Langley Research Center, Hampton, VA, which is operated under NASA contract NAS1-18107.

## 1. Introduction

Recently a great deal of attention has been given to the analysis of, and the development of approximation theory for, infinite dimensional operator Riccati differential equations. This is due in part to the role they play in the characterization of the closed-loop feedback structure of the solution to linear-quadratic optimal control problems for distributed parameter systems (see [10]). Working from Temam's [14], [15] formulation (see also Barbu [3]) of a class of operator Riccati equations in the space of Hilbert-Schmidt operators on a separable Hilbert space, we develop an approximation and convergence theory for generic Galerkin type approximations. We make the usual assumption that our underlying Hilbert space admits a densely, continuously, and compactly embedded subspace. The only condition that we then require on our sequence of approximating Galerkin subspaces is the usual one that the corresponding sequence of orthogonal projections converge strongly to the identity in the stronger topology. We are able to obtain Hilbert-Schmidt operator norm convergence of the approximating Riccati operators, uniformly in time on compact time intervals.

Our effort here differs significantly from other recent treatments of the approximation theory for infinite dimensional operator Riccati equations (see, for example, [6]) in that i) we obtain Hilbert-Schmidt as opposed to strong operator convergence, and, more importantly, ii) our convergence theory is based directly upon, and involves only, the differential equation itself rather than equivalent integral equations. Of course in order to do this we must necessarily consider a somewhat more restrictive, but still sufficiently interesting from an applications point of view, class of problems. For example we require that the

linear part of the equation be strongly coercive on a space of Hilbert-Schmidt operators and that the non-homogeneous or quasi autonomous perturbation (the state penalization operator in the context of the LQ control problem) be Hilbert-Schmidt. Our treatment here is restricted to the quasi-autonomous or constant coefficient case. The temporally inhomogeneous problem requires a different approach; our results for the time dependent case will be reported elsewhere.

Our convergence theory is based upon a generic approximation result for nonlinear quasi-autonomous evolution equations in Banach space with dynamics described by accretive operators. In section 2 we prove a nonlinear analog of the well-known Trotter-Kato theorem on the approximation of linear semigroups - i.e. stability (uniform dissipativity) and consistency (strong operator convergence of the resolvent) yield convergence (see, for example, [11]). The result that we prove here is closely related to similar approximation results for nonlinear evolution systems which have appeared elsewhere in the literature (for example, [4] and [8]). However the theorem we prove below is one that is most appropriate for the particular class of problems that are of interest to us here. We follow Goldstein [8] and give a proof using an idea originally suggested by Kisynski [9] for the linear case wherein convergence is demonstrated via an application of an existence theorem to an evolution equation in an appropriately constructed sequence space.

In section 3 we briefly outline Temam's [15] formulation of the Riccati equation as a well posed nonlinear evolution equation in the space of Hilbert-Schmidt operators. Our approximation and convergence theory is developed and presented in section 4. In section 5 we illustrate the application of our results in the context of a linear quadratic optimal

control problem (i.e. the linear quadratic regulator problem) for a one dimensional heat equation.

## 2. An Abstract Approximation Result for Nonlinear Evolution Equations

Let  $X_0$  be a real Banach space with norm denoted by  $|\cdot|_0$ . Let  $X_0^*$  be its dual and let  $A_0 : X_0 \rightarrow 2^{X_0}$  be an, in general, multi-valued nonlinear closed accretive operator (i.e.  $\{(x, y) : x \in X_0, y \in A_0 x\}$  is a closed subset of  $X_0 \times X_0$ , and  $|x_1 - x_2|_0 \leq |x_1 + \lambda y_1 - (x_2 + \lambda y_2)|_0$  for all  $\lambda > 0$ ,  $x_i \in X_0$ , and  $y_i \in A_0 x_i$ ,  $i = 1, 2$ ) on  $X_0$ . Define the domain of  $A_0$  to be the set  $\text{Dom}(A_0) = \{x \in X_0 : A_0 x \neq \emptyset\}$ , and the range of  $A_0$  to be the set  $\mathcal{R}(A_0) = \bigcup_{x \in \text{Dom}(A_0)} A_0 x$ . Since  $A_0$  is accretive, it follows that for  $\lambda > 0$  the resolvent of  $A_0$  at  $\lambda$ ,  $J_0(\lambda) = (I + \lambda A_0)^{-1}$ , is a well defined single-valued nonexpansive operator (i.e.  $|J_0(\lambda)y_1 - J_0(\lambda)y_2|_0 \leq |y_1 - y_2|_0$ ,  $y_1, y_2 \in \text{Dom}(J_0(\lambda))$ ) defined on  $\text{Dom}(J_0(\lambda)) = \mathcal{R}(I + \lambda A_0)$ . Suppose that  $T > 0$ , let  $t \rightarrow f_0(t)$  be an  $X_0$ -valued map defined on  $[0, T]$  and let  $x_0^0 \in X_0$ . We consider the initial value problem in  $X_0$  given by

$$(2.1) \quad \dot{x}_0(t) + A_0 x_0(t) \ni f_0(t), \quad \text{a.e. } t \in (0, T),$$

$$(2.2) \quad x_0(0) = x_0^0.$$

We shall say that a function  $x_0(\cdot) : [0, T] \rightarrow X_0$  is a strong solution to the initial value problem (2.1), (2.2) if it is continuous on  $[0, T]$ , Lipschitz on every compact subinterval of  $(0, T)$ , differentiable almost everywhere on  $(0, T)$ , and satisfies (2.1) and (2.2). We shall call  $x_0$  an integral solution of (2.1), (2.2) if it is continuous on  $[0, T]$ , satisfies (2.2), and if the inequality

$$\frac{1}{2}|x_0(t) - x|_0^2 \leq \frac{1}{2}|x_0(s) - x|_0^2 + \int_s^t \langle f_0(\tau) - y, x_0(\tau) - x \rangle_0 d\tau$$

holds for all  $x \in \text{Dom}(A_0)$ ,  $y \in A_0x$ , and  $0 \leq s \leq t \leq T$  where the pairing  $\langle \cdot, \cdot \rangle_0$ :  $X_0 \times X_0 \rightarrow \mathbf{R}$  is given by  $\langle y, x \rangle_0 = \sup\{x^*(y), x^* \in \Phi_0(x)\}$  with  $\Phi_0 : X_0 \rightarrow 2^{X_0^*}$  denoting the duality map on  $X_0$  (see [3]). A strong solution of (2.1), (2.2) is of course also an integral solution. The following existence, uniqueness, and regularity results are given by Barbu [3].

**Theorem 2.1** Let  $\{X, |\cdot|\}$  be a real Banach space and suppose that  $A : X \rightarrow 2^X$  is a closed accretive operator on  $X$ . Let  $C$  be a closed convex cone in  $X$  such that  $\text{Dom}(A) \subset C$  and  $\mathcal{R}(I + \lambda A) \supset C$ ,  $\lambda > 0$ . If  $x^0 \in \overline{\text{Dom}(A)}$  and  $f \in L_1(0, T; X)$  with  $f(t) \in C$ , a.e.  $t \in (0, T)$ , then there exists a unique integral solution  $x$  to the initial value problem

$$(2.3) \quad \dot{x}(t) + Ax(t) \ni f(t), \quad \text{a.e. } t \in (0, T),$$

$$(2.4) \quad x(0) = x^0,$$

with  $x(t) \in \overline{\text{Dom}(A)}$  a.e.  $t \in (0, T)$ . Moreover, if  $y$  is the integral solution to (2.3), (2.4) with  $f$  replaced by  $g \in L_1(0, T; X)$ , then

$$|x(t) - y(t)|^2 \leq |x(s) - y(s)|^2 + 2 \int_s^t \langle f(\tau) - g(\tau), x(\tau) - y(\tau) \rangle d\tau$$

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

**Theorem 2.2** If in addition to the hypotheses and conditions of the previous theorem we have  $x^0 \in \text{Dom}(A)$  and  $f \in W^{1,1}(0, T; X)$ , then the initial value problem (2.3), (2.4) has a unique strong solution  $x$  with  $x \in W^{1,\infty}(0, T; X)$ ,  $x(t) \in \text{Dom}(A)$  a.e.  $t \in (0, T)$  and

$$|\dot{x}(t)| = |Ax(t) + f(t)| \leq |Ax^0 + f(0)| + 2 \int_0^t |f'(s)| ds, \quad \text{a.e. } t \in (0, T).$$

Furthermore, if  $X$  and  $X^*$  are uniformly convex, then  $x$  is everywhere differentiable from the right and

$$\frac{d^+x}{dt}(t) = (Ax(t) + f(t))^0, \quad 0 \leq t \leq T$$

where  $(Ax(t) + f(t))^0$  denotes the unique element of minimum norm in the set  $Ax(t) + f(t)$ .

Henceforth, we shall assume that there exists a closed convex cone  $C_0 \subset X_0$  for which  $\text{Dom}(A_0) \subset C_0$  and  $\mathcal{R}(I + \lambda A_0) \supset C_0, \lambda > 0$ , that  $x_0^0 \in \overline{\text{Dom}(A_0)}$  and that  $f_0 \in L_1(0, T; X_0)$  with  $f_0(t) \in C_0$  for a.e.  $t \in (0, T)$ . Consequently Theorem 2.1 applies and we are guaranteed that the initial value problem in  $X_0$ , (2.1), (2.2), admits a unique integral solution.

We prove an approximation result in the spirit of the well known Trotter-Kato theorem for the approximation of linear semigroups (see, for example, [11]), and the approximation theorems given by Crandall and Pazy [4] and Goldstein [8] in a nonlinear setting. Although the convergence theorem we shall prove here does not differ significantly from the ones given in [4] and [8], the latter results are stated inappropriately and are somewhat too restrictive for the application we intend to consider below. Following Goldstein [8], the proof we shall give here is based upon an idea first suggested by Kisynski [9] in the context of linear semigroup approximation. We argue convergence via an application of the existence and uniqueness result, Theorem 2.1.

For each  $n = 1, 2, \dots$  let  $X_n$  be a closed linear subspace of  $X_0$ , and let  $A_n : X_n \rightarrow 2^{X_n}$  be a closed accretive operator on  $X_n$ . As was the case above with  $A_0$ , for  $\lambda > 0$  the resolvent of  $A_n$ ,  $J_n(\lambda) = (I + \lambda A_n)^{-1}$  is a well defined, single-valued, nonexpansive operator defined on  $\text{Dom}(J_n(\lambda)) = \mathcal{R}(I + \lambda A_n)$ . We assume that there exists a closed convex cone  $C_n$  in  $X_n$

for which  $\text{Dom}(A_n) \subset C_n$  and  $\mathcal{R}(I + \lambda A_n) \supset C_n, \lambda > 0$ . We assume further that for each  $n = 1, 2, \dots, x_n^0$  is an element in  $\overline{\text{Dom}(A_n)}$  and that  $f_n$  is a function in  $L_1(0, T; X_n)$  with  $f_n(t) \in C_n$  for almost every  $t \in (0, T)$ . It then follows that Theorem 2.1 implies that the initial value problem in  $X_n$  given by

$$(2.5) \quad \dot{x}_n(t) + A_n x_n(t) \ni f_n(t), \quad \text{a.e. } t \in (0, T),$$

$$(2.6) \quad x_n(0) = x_n^0,$$

admits a unique integral solution  $x_n \in C(0, T; X_n)$ .

In our discussions below, we shall use the notation  $\lim_{n \rightarrow \infty} D_n \supset D_0$  where  $D_n \subset X_n$  and  $D_0 \subset X_0$ . By this we shall mean that for each  $z_0 \in D_0$  there exists a sequence  $\{z_n\}_{n=1}^{\infty}$  with  $z_n \in D_n$  and  $\lim_{n \rightarrow \infty} z_n = z_0$ . Our fundamental approximation and convergence result is given in the following theorem.

**Theorem 2.3** Suppose

- (i)  $\lim_{n \rightarrow \infty} C_n \supset C_0$
- (ii)  $\lim_{n \rightarrow \infty} f_n(t) = f_0(t)$  and there exists a  $g \in L_1(0, T)$  for which  $|f_n(t)| \leq g(t), n = 1, 2, \dots$ ,  
for almost every  $t \in (0, T)$
- (iii)  $\lim_{n \rightarrow \infty} x_n^0 = x_0^0$
- (iv)  $\lim_{n \rightarrow \infty} J_n(\lambda)y_n = J_0(\lambda)y_0$  for each  $\lambda > 0$  whenever  $y_n \in C_n$  and  $\lim_{n \rightarrow \infty} y_n = y_0 \in C_0$ .

Then  $\lim_{n \rightarrow \infty} x_n = x_0$  in  $C(0, T; X_0)$  where  $x_n$  for  $n = 1, 2, \dots$  and  $x_0$  are the unique integral solutions to the initial value problems (2.5), (2.6) and (2.1), (2.2) respectively.

**Proof:** If we define the linear space  $X$  over the reals by

$$X = \{u = \{u_n\}_{n=0}^{\infty} : u_n \in X_n, n = 0, 1, 2, \dots, \lim_{n \rightarrow \infty} u_n = u_0\}$$

and for  $u = \{u_n\}_{n=0}^\infty \in X$  set  $|u| = \sup_n |u_n|_0$ , then  $\{X, |\cdot|\}$  is a real Banach space. Let  $C$

be the closed convex cone in  $X$  given by  $C = \{u = \{u_n\}_{n=0}^\infty \in X : u_n \in C_n, n = 0, 1, 2, \dots\}$ .

Note that if  $C_0 \neq \phi$ , hypothesis (i) implies that  $C \neq \phi$ . Define the operator  $A : X \rightarrow 2^X$  by

$$Au = \begin{cases} \{v = \{v_n\}_{n=0}^\infty : v_n \in A_n u_n, \\ \quad n = 0, 1, 2, \dots, \lim_{n \rightarrow \infty} v_n = v_0\}, & u = \{u_n\}_{n=0}^\infty \in \text{Dom}(A) \\ \phi & u \notin \text{Dom}(A) \end{cases}$$

$\text{Dom}(A) = \{u = \{u_n\}_{n=0}^\infty \in X : u_n \in \text{Dom}(A_n), n = 0, 1, 2, \dots, \text{ and for each } n = 1, 2, \dots$

there exists a  $v_n \in A_n u_n$  for which  $\lim_{n \rightarrow \infty} v_n = v_0 \in A_0 u_0\}$ .

The operators  $A_n$  being closed and accretive implies that the operator  $A$  is closed and accretive as well. Indeed, for example, for  $\lambda > 0$ ,  $u^i \in \text{Dom}(A)$ ,  $i = 1, 2$ , and  $v^i \in Au^i$ ,  $i = 1, 2$ , we have

$$|u^1 - u^2| = \sup_n |u_n^1 - u_n^2|_0 \leq \sup_n |u_n^1 + \lambda v_n^1 - (u_n^2 + \lambda v_n^2)|_0 = |u^1 + \lambda v^1 - (u^2 + \lambda v^2)|.$$

The fact that  $A$  is closed can be argued analogously.

Clearly  $\text{Dom}(A) \subset C$  and it also follows that for  $\lambda > 0$   $\mathcal{R}(I + \lambda A) \in C$ . To see this let  $\lambda > 0$  and let  $v = \{v_n\}_{n=0}^\infty \in C$ . Then  $v_n \in C_n$  and  $C_n \subset \mathcal{R}(I + \lambda A_n)$ ,  $n = 0, 1, 2, \dots$  imply that we may define  $u_n = J_n(\lambda)v_n$ ,  $n = 0, 1, 2, \dots$  and set  $u = \{u_n\}_{n=0}^\infty$ . Now  $v \in C \subset X$  and hypothesis (iv) yield  $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} J_n(\lambda)v_n = J_0(\lambda)v_0 = u_0$ ; thus  $u \in X$ . Also, if we set  $w_n = (v_n - u_n)/\lambda$ ,  $n = 0, 1, 2, \dots$  then  $w = \{w_n\}_{n=0}^\infty \in X$ ,  $w_n \in A_n u_n$  and  $\lim_{n \rightarrow \infty} w_n = w_0 \in A_0 u_0$ . It follows that  $u_n \in \text{Dom}(A_n)$  and therefore that  $u \in \text{Dom}(A)$  with  $w \in Au$  and  $v \in (I + \lambda A)u$ . We conclude that  $v \in \mathcal{R}(I + \lambda A)$  and therefore that  $C \subset \mathcal{R}(I + \lambda A)$ .

Define  $z^0 = \{x_n^0\}_{n=0}^\infty$ . Then hypothesis (iii) implies that  $z \in X$ . Moreover, using (iv) it can be shown that  $\overline{\text{Dom}(A)} = \{u = \{u_n\}_{n=0}^\infty \in X : u_n \in \overline{\text{Dom}(A_n)}, n = 0, 1, 2, \dots\}$ . Thus  $z \in \overline{\text{Dom}(A)}$ . If we define the function  $f$  by  $f(t) = \{f_n(t)\}_{n=0}^\infty$  for almost every  $t \in (0, T)$ , then hypothesis (ii) implies that  $f \in L_1(0, T; X)$ , and  $f_n(t) \in C_n$ ,  $n = 0, 1, 2, \dots$  a.e.  $t \in (0, T)$  implies

that  $f(t) \in C$ , a.e.  $t \in (0, T)$ . It follows from the arguments given above and Theorem 2.1 that the initial value problem in  $X$

$$\dot{z} + Az(t) \ni f(t), \quad \text{a.e. } t \in (0, T)$$

$$z(0) = z^0$$

admits a unique integral solution  $z \in C(0, T; X)$  with  $z(t) \in \overline{\text{Dom}(A)}$  for almost every  $t \in (0, T)$ . In addition, inspection of the proof of Theorem III. 2.1 given in [3] (paying particular attention to how solutions are actually constructed) reveals that we must have  $z(t) = \{x_n(t)\}_{n=0}^\infty$  where for  $n = 1, 2, \dots$ ,  $x_n$  is the unique integral solution to the initial value problem (2.5), (2.6) in  $X_n$  and  $x_0$  is the unique integral solution to the initial value problem (2.1), (2.2) in  $X_0$ . Consequently, since  $z(t) \in X$ ,  $t \in [0, T]$ , it follows that  $\lim_{n \rightarrow \infty} x_n(t) = x_0(t)$  for each  $t \in [0, T]$ .

Finally, that the convergence of  $x_n(t)$  to  $x_0(t)$  is in fact uniform in  $t$  for  $t \in [0, T]$  can be argued as it was done by Goldstein in [8]. Let  $\epsilon > 0$  and  $t \in [0, T]$  be given. Then  $z = \{x_n\}_{n=0}^\infty \in C(0, T; X)$  implies that  $|z(t) - z(s)| < \epsilon/3$  for all  $0 \leq s \leq T$  with  $|t - s| < 2\delta = 2\delta(\epsilon)$ . Now  $\lim_{n \rightarrow \infty} x_n(s) = x_0(s)$  for each  $s \in [0, T]$  implies that  $|x_n(s) - x_0(s)|_0 < \epsilon/3$  for all  $n > N_s = N(s, \epsilon)$ . Thus

$$|x_n(t) - x_0(t)|_0 \leq |x_n(t) - x_n(s)|_0 + |x_n(s) - x_0(s)|_0 + |x_0(s) - x_0(t)|_0 < \epsilon,$$

for all  $0 \leq s \leq T$  with  $|t - s| < 2\delta$  and  $n > N_s$ . Let  $k$  be the greatest integer less than or equal to  $T/\delta$  and set  $N = \max\{N_0, N_\delta, N_{2\delta}, \dots, N_{k\delta}, N_T\}$ . Then  $|x_n(t) - x_0(t)|_0 < \epsilon$  for all  $n > N = N(\epsilon)$  and the proof of the theorem is complete.

### 3. Operator Riccati Equations on Spaces of Hilbert-Schmidt Operators

In this section we briefly review and outline Temam's [15] results on operator Riccati equations set in the Hilbert space of Schmidt class operators on a separable Hilbert space. Let  $H$  be a real separable Hilbert space with inner product and associated induced norm denoted by  $(\cdot, \cdot)$  and  $|\cdot|$  respectively. Let  $V$  be another real separable Hilbert space with inner product  $\langle \cdot, \cdot \rangle$  and induced norm  $\|\cdot\|$ . We assume that  $V$  is densely and continuously embedded in  $H$  (i.e.  $\overline{V} = H$ , and there exists a positive constant  $\mu$  for which  $|\varphi| \leq \mu\|\varphi\|, \varphi \in V$ ). Identifying  $H$  with its dual,  $H^*$ , we have  $V \hookrightarrow H = H^* \hookrightarrow V^*$  with the embedding of  $H^*$  in  $V^*$  being dense and continuous as well. Denote the usual operator norm on  $V^*$  by  $\|\cdot\|_*$ . Let  $\gamma$  denote the canonical isomorphism (Riesz map) from  $V$  onto  $V^*$ . Then for  $\varphi, \psi \in V$  we have

$$(\gamma\varphi, \psi) = \langle \varphi, \psi \rangle$$

where  $(\cdot, \cdot)$  in the above expression denotes the usual extension of the  $H$  inner product to the duality pairing between  $V^*$  and  $V$ . We assume further that the embedding of  $V$  into  $H$  is compact. It then follows that  $\gamma^{-1} \in \mathcal{L}(V, V^*) \cap \mathcal{L}(H, V)$  and that  $\gamma^{-1}$  is self-adjoint, positive, and compact as a mapping from  $H$  into  $H$ . We note that  $V^*$  is in fact a Hilbert space with inner product  $\langle \cdot, \cdot \rangle_*$  given by

$$\langle \varphi, \psi \rangle_* \equiv \langle \gamma^{-1}\varphi, \gamma^{-1}\psi \rangle = (\varphi, \gamma^{-1}\psi), \quad \varphi, \psi \in V^*.$$

We have that  $\|\varphi\|_* = \sqrt{\langle \varphi, \varphi \rangle_*}, \quad \varphi \in V^*.$

Since  $\gamma^{-1}$  is self-adjoint, positive, and compact on  $H$  there exists an orthonormal basis,  $\{e_k\}_{k=1}^\infty$ , for  $H$  such that  $\gamma^{-1}e_k = \rho_k^{-2}e_k, k = 1, 2, \dots$  for some real numbers  $\rho_k, k = 1, 2, \dots$ . Consequently  $\{\rho_k^{-1}e_k\}_{k=1}^\infty$  and  $\{\rho_k e_k\}_{k=1}^\infty$  are orthonormal bases for  $V$  and  $V^*$  respectively.

For separable Hilbert spaces  $X$  and  $Y$ , let  $HS(X, Y)$  denote the Hilbert space of Hilbert-Schmidt operators from  $X$  into  $Y$ . Denote the corresponding inner product and induced norm by  $[\cdot, \cdot]_{HS(X, Y)}$  and  $|\cdot|_{HS(X, Y)}$  respectively. Set  $\mathcal{H} = HS(H, H)$  with

$$[\Phi, \Psi]_{\mathcal{H}} = [\Phi, \Psi]_{HS(H, H)} = \sum_{k=1}^{\infty} (\Phi e_k, \Psi e_k),$$

and  $|\Phi|_{\mathcal{H}} = |\Phi|_{HS(H, H)} = \sqrt{[\Phi, \Phi]_{\mathcal{H}}}$ ,  $\Phi, \Psi \in \mathcal{H}$ . Define the Hilbert space  $\mathcal{V}$  by  $\mathcal{V} = HS(V^*, H) \cap HS(H, V)$  with inner product

$$[\Phi, \Psi]_{\mathcal{V}} = [\Phi, \Psi]_{HS(V^*, H)} + [\Phi, \Psi]_{HS(H, V)} = \sum_{k=1}^{\infty} \rho_k^2 (\Phi e_k, \Psi e_k) + \sum_{k=1}^{\infty} \langle \Phi e_k, \Psi e_k \rangle,$$

$\|\Phi\|_{\mathcal{V}} = \sqrt{[\Phi, \Phi]_{\mathcal{V}}}$ ,  $\Phi, \Psi \in \mathcal{V}$ . It is not difficult to show that the dense and continuous inclusions  $HS(V^*, H) \subset HS(H, H) \subset HS(V, H)$ , and  $HS(H, V) \subset HS(H, H) \subset HS(H, V^*)$  hold. Also, it can be argued that  $HS(V^*, H)$  and  $HS(V, H)$ , and  $HS(H, V)$  and  $HS(H, V^*)$  are dual pairs with respect to the duality pairing

$$[\Phi, \Psi]_{\mathcal{H}} = \sum_{k=1}^{\infty} (\Phi e_k, \Psi e_k).$$

It follows that  $\mathcal{V}^* = HS(V, H) + HS(H, V^*)$  and, identifying  $\mathcal{H}$  with its dual  $\mathcal{H}^*$ , that  $\mathcal{V} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{V}^*$ , with the inclusions dense and continuous.

Let  $a(\cdot, \cdot) : V \times V \rightarrow \mathbf{R}$  be a bounded strongly  $V$ -elliptic bilinear form on  $V \times V$ . That is, there exist constants  $\alpha, \beta > 0$  for which  $a(\varphi, \varphi) \geq \alpha \|\varphi\|^2$  and  $|a(\varphi, \psi)| \leq \beta \|\varphi\| \|\psi\|$ ,  $\varphi, \psi \in V$ . (With minor modification our theory continues to hold if the form  $a(\cdot, \cdot)$  satisfies the coercive inequality  $a(\varphi, \varphi) + \lambda |\varphi|^2 \geq \alpha \|\varphi\|^2$ ,  $\varphi \in V$ , for some  $\lambda \in \mathbf{R}$ . To keep our presentation here as simple as possible, however, in our discussions below we shall treat only the case  $\lambda = 0$ .) Let  $\mathcal{A} \in \mathcal{L}(V, V^*)$  be the operator defined by  $(\mathcal{A}\phi, \psi) =$

$a(\phi, \psi)$ ,  $\phi, \psi \in V$ . Let  $a^*(\cdot, \cdot) : V \times V \rightarrow \mathbf{R}$  be the form which is adjoint to  $a(\cdot, \cdot)$ ; that is  $a^*(\phi, \psi) = a(\psi, \phi)$ ,  $\phi, \psi \in V$ . Then  $a^*(\varphi, \varphi) \geq \alpha \|\varphi\|^2$  and  $|a^*(\varphi, \psi)| \leq \beta \|\varphi\| \|\psi\|$ ,  $\varphi, \psi \in V$ . Let  $\mathcal{A} \in \mathcal{L}(V, V^*)$  be the operator defined by  $(\mathcal{A}^* \varphi, \psi) = a^*(\varphi, \psi) = a(\psi, \varphi)$ ,  $\varphi, \psi \in V$ . Let  $C_0$  be the closed convex cone in  $\mathcal{H}$  given by  $C_0 = \{\Phi \in \text{HS}(\mathbf{H}, \mathbf{H}) : \Phi = \Phi^*, \Phi \geq 0\}$ . Let  $T > 0$  and suppose that  $\Pi_0^0 \in C_0$  and that  $F_0 \in L_1(0, T; \mathbf{H})$  with  $F_0(t) \in C_0$  for almost every  $t \in (0, T)$  are given. Let  $\Phi \rightarrow \mathcal{F}_0(\Phi)$  be a single valued map defined for each  $\Phi \in C_0$  with range in  $\mathcal{H}$  which is continuous from  $\mathcal{H}$  into itself. Assume further that  $\mathcal{F}_0$  has the property that

$$[\mathcal{F}_0(\Phi) - \mathcal{F}_0(\Psi), \Phi - \Psi]_{\mathcal{H}} \geq 0,$$

for all  $\Phi, \Psi \in C_0$ . We note that if  $z \rightarrow \mathcal{F}_0(z)$  is a single valued complex function of the complex variable  $z$  with  $\mathcal{F}_0(0) = 0$  and which is analytic on the nonnegative real axis, then the mapping  $\Phi \rightarrow \mathcal{F}_0(\Phi)$  satisfies  $\mathcal{F}_0(\Phi) \in \mathcal{H}$  for all  $\Phi \in C_0$  and is continuous from  $\mathcal{H}$  into  $\mathcal{H}$  (see Dunford and Schwartz [5], Theorem XI.6.7.7). We seek a solution  $\Pi_0$  to the initial value problem

$$(3.1) \quad \dot{\Pi}_0(t) + \mathcal{A}^* \Pi_0(t) + \Pi_0(t) \mathcal{A} + \mathcal{F}_0(\Pi_0(t)) = F_0(t), \text{ a.e. } t \in (0, T)$$

$$(3.2) \quad \Pi_0(0) = \Pi_0^0$$

with  $\Pi_0(t) \in C_0$ , a.e.  $t \in (0, T)$ .

Note that when  $\mathcal{F}_0(\Phi) = \Phi^2$ , (3.1) becomes the standard quadratic Riccati equation. Elementary properties of Hilbert-Schmidt operators (see [14]) can be used to argue that for  $\Phi, \Psi \in C_0$  we have

$$[\Phi^2 - \Psi^2, \Phi - \Psi]_{\mathcal{H}} = [\Phi(\Phi - \Psi), \Phi - \Psi]_{\mathcal{H}} + [(\Phi - \Psi)\Psi, \Phi - \Psi]_{\mathcal{H}} \geq 0.$$

Essentially the same argument can be used for the case of  $\mathcal{F}_0(\Phi) = \Phi^n$  where  $n$  is any

positive integer. For more general quadratic terms, for example, the one that results in the case of the linear-quadratic optimal control problem,  $\mathcal{F}_0(\Phi) = \Phi M \Phi$  with  $M \in \mathcal{L}(H)$ ,  $M = M^*$ , and  $M \geq 0$ , once again a similar argument will work so long as  $\Phi M \geq 0$ ,  $\Phi \in C_0$  (see Proposition 2.2 in [15]). This will of course be true if  $\Phi M = M \Phi$  for all  $\Phi \in C_0$

**Remark 3.1** At this point a comment regarding the relationship between the operator  $\mathcal{A}^*$  and the adjoint of  $\mathcal{A}$  is in order. Define the operator  $A : \text{Dom}(A) \subset H \rightarrow H$  to be the restriction of the operator  $\mathcal{A}$  to the set  $\text{Dom}(A) = \{\varphi \in V : \mathcal{A}\varphi \in H\}$ . It can be shown (see [12]) that  $\overline{\text{Dom}(A)} = H$  and consequently that  $A$  admits an  $H$ -adjoint,  $A^* : \text{Dom}(A^*) \subset H \rightarrow H$ . The operator  $\mathcal{A}^*$  is the extension of  $A^*$  to an operator defined on all of  $V$  or, equivalently,  $A^*$  is the restriction of  $\mathcal{A}^*$  to the  $H$ -dense subset  $\text{Dom}(A^*) = \{\varphi \in V : \mathcal{A}^*\varphi \in H\}$ . We note also that  $-A$  is the infinitesimal generator of a uniformly exponentially stable semigroup,  $\{\mathcal{T}(t) : t \geq 0\}$ , of bounded linear operators on  $H$ . Similarly,  $-A^*$  is the infinitesimal generator of the adjoint semigroup  $\{\mathcal{T}(t)^* : t \geq 0\}$  on  $H$ . In addition, it can be argued that both of these semigroups admit respectively restrictions and extensions to analytic semigroups on  $V$  and  $V^*$  (see [1], [13]).

An appropriate reformulation of (3.1), (3.2) will allow an application Theorems 2.1 and 2.2. Define the operator  $L_0 \in \mathcal{L}(V, V^*)$  by  $L_0\Phi = \mathcal{A}^*\Phi + \Phi\mathcal{A}$ , for  $\Phi \in V$ . It is not difficult to argue that  $L_0$  is strongly  $V$ -elliptic; that is, there exists a constant  $\omega > 0$  for which

$$(3.3) \quad [L_0\Phi, \Phi]_{\mathcal{H}} \geq \omega \|\Phi\|_V^2, \quad \Phi \in V,$$

and therefore that the set  $\text{Dom}(L_0) = \{\Phi \in V; L_0\Phi \in \mathcal{H}\}$  is dense in  $\mathcal{H}$  (see [13]). Define

the operator  $A_0 : \text{Dom}(A_0) \subset \mathcal{H} \rightarrow \mathcal{H}$  by

$$A_0 \Phi = L_0 \Phi + \mathcal{F}_0(\Phi), \quad \Phi \in \text{Dom}(A_0) = \text{Dom}(L_0) \cap C_0.$$

It follows that  $A_0$  is a closed operator on  $\mathcal{H}$  and that it is strongly  $\mathcal{V}$ -monotone. That is, for  $\Phi, \Psi \in \text{Dom}(A_0)$  we have

$$[A_0 \Phi - A_0 \Psi, \Phi - \Psi]_{\mathcal{H}} = [L_0(\Phi - \Psi), \Phi - \Psi]_{\mathcal{H}} + [\mathcal{F}_0(\Phi) - \mathcal{F}_0(\Psi), \Phi - \Psi]_{\mathcal{H}} \geq \omega \|\Phi - \Psi\|_{\mathcal{V}}^2.$$

From this it can be argued at once that  $A_0$  is accretive on its domain and that  $\mathcal{R}(I + \lambda A_0) \supset C_0$  for all  $\lambda > 0$ .

We rewrite (3.1), (3.2) as the initial value problem in  $\mathcal{H}$  given by

$$(3.4) \quad \dot{\Pi}_0(t) + A_0 \Pi_0(t) = F_0(t), \quad \text{a.e. } t \in (0, T)$$

$$(3.5) \quad \Pi_0(0) = \Pi_0^0$$

Now  $\overline{\text{Dom}(A_0)} = \overline{\text{Dom}(L_0)} \cap C_0 = C_0$  and recall that it was assumed that  $\Pi_0^0 \in C_0$ . Also it was assumed that  $F_0 \in L_1(0, T; \mathcal{H})$  with  $F_0(t) \in C_0$ , a.e.  $t \in (0, T)$ . Consequently Theorem 2.1 yields the existence of a unique integral solution  $\Pi_0$  to the initial value problem (3.1), (3.2) with  $\Pi_0(t) \in C_0$  for almost every  $t \in (0, T)$ . If it is further assumed that  $\Pi_0^0 \in \text{Dom}(A_0)$  (i.e. that  $\Pi_0^0 \in C_0$  and that  $\mathcal{A}^* \Pi_0^0 + \Pi_0^0 \mathcal{A} \in \text{HS}(H, H)$ ) and that  $F_0 \in W^{1,1}(0, T; \mathcal{H})$  with  $F_0(t) \in C_0$ , a.e.  $t \in (0, T)$ , then Theorem 2.2 implies the existence of unique strong solution  $\Pi_0 \in W^{1,\infty}(0, T; H)$  with  $\mathcal{A}^* \Pi_0(t) + \Pi_0(t) \mathcal{A} \in \text{HS}(H, H)$ , a.e.  $t \in (0, T)$ . Using density it can be argued further that if  $\Pi_0^0 \in C_0$  and  $F_0 \in L_2(0, T; \mathcal{V}^*)$  with  $F_0(t)^* = F_0(t)$ ,  $F_0(t) \geq 0$ , a.e.  $t \in (0, T)$ , then there exists a unique solution  $\Pi_0 \in L_2(0, T; \mathcal{V}) \cap C(0, T; \mathcal{H}) \cap H^1(0, T; \mathcal{V}^*)$  with  $\Pi_0(t) \in C_0$  for every  $t \in [0, T]$  (of course in this setting,  $\Pi_0$  is a solution of (3.4), (3.5) in a distributional or  $\mathcal{V}^*$  sense).

#### 4. Approximation Theory

For each  $n = 1, 2, \dots$  let  $H_n$  be a finite dimensional subspace of  $H$  with  $H_n \subset V$  for all  $n$ . Let  $P_n : H \rightarrow H_n$  denote the orthogonal projection of  $H$  onto  $H_n$  with respect to the standard inner product,  $(\cdot, \cdot)$ , on  $H$ . We shall require the assumption that

$$(4.1) \quad \lim_{n \rightarrow \infty} \|P_n \varphi - \varphi\| = 0, \quad \varphi \in V.$$

Note that (4.1) implies that  $\lim_{n \rightarrow \infty} |P_n \varphi - \varphi| = 0$ ,  $\varphi \in H$  and that  $P_n$  is uniformly bounded in  $n$  in both  $\mathcal{L}(H)$  and  $\mathcal{L}(V)$  with respect to the uniform operator topologies.

It will be necessary for us to define an extension of the operator  $P_n$  to a bounded operator from  $V^*$  into  $V^*$ . We do this as follows. For  $\varphi \in V^*$  define  $P_n \varphi \in H_n$  to be the representer in  $H_n$  of the functional  $\varphi$  restricted to a functional on  $H_n$ . That is  $P_n \varphi = \varphi_n$  where  $\varphi_n$  is the unique element in  $H_n$  (guaranteed to exist by the Riesz representation theorem applied to the Hilbert space  $\{H_n, (\cdot, \cdot)\}$ ) which satisfies  $(\varphi, \theta_n) = (\varphi_n, \theta_n)$ ,  $\theta_n \in H_n$ . Since  $H_n \subset V \subset V^*$ ,  $P_n$  can be considered to be a linear map of  $V^*$  into itself. Indeed  $P_n \varphi = \varphi_n$  is the continuous linear functional on  $V$  given by  $(P_n \varphi, \theta) = (\varphi_n, \theta)$ ,  $\theta \in V$ . The definition of the orthogonal projection of  $\varphi \in H$  onto  $H_n$ ,  $(P_n \varphi - \varphi, \theta_n) = 0$ ,  $\theta_n \in H_n$ , reveals that we have in fact defined an extension of  $P_n$  to an operator on  $V^*$  with range  $H_n$  considered as a subspace of  $V^*$ . Using the fact that for  $\varphi \in V^*$  and  $\psi \in V$  we have  $(P_n \varphi, \psi) = (P_n \varphi, \psi) - (P_n \varphi, \psi - P_n \psi) = (P_n \varphi, P_n \psi) = (\varphi, P_n \psi)$ , and that the  $P_n$  are uniformly bounded in  $\mathcal{L}(V)$ , it is not difficult to show that the  $P_n$  are uniformly bounded in  $\mathcal{L}(V^*)$ . It is worth noting that the extension of  $P_n$  defined above agrees with the operator that would be obtained if we were to extend  $P_n$  in the usual way by considering it as a bounded operator defined on the dense subset  $H$  of the Banach space  $V^*$ .

It follows from assumption (4.1) that  $\lim_{n \rightarrow \infty} \|P_n \varphi - \varphi\|_* = 0$ ,  $\varphi \in V^*$ . Indeed, for  $\varphi \in H$  we have  $\|P_n \varphi - \varphi\|_* \leq \mu |P_n \varphi - \varphi| \rightarrow 0$  as  $n \rightarrow \infty$ . Since  $H$  is dense in  $V^*$ , and the  $P_n$  are uniformly bounded in  $\mathcal{L}(V^*)$ , it follows that  $\|P_n \varphi - \varphi\|_* \rightarrow 0$  as  $n \rightarrow \infty$  for  $\varphi \in V^*$ . For  $\varphi, \psi \in V^*$ , we have

$$\langle P_n \varphi, \psi \rangle_* = (P_n \varphi, \gamma^{-1} \psi) = (\varphi, P_n \gamma^{-1} \psi) = (\varphi, \gamma^{-1} \gamma P_n \gamma^{-1} \psi) = \langle \varphi, \gamma P_n \gamma^{-1} \psi \rangle_*.$$

Thus  $P_n^* \in \mathcal{L}(V^*)$ , the adjoint of  $P_n$  considered as a bounded operator on  $V^*$ , is given by  $P_n^* = \gamma P_n \gamma^{-1}$ . Assumption (4.1) yields

$$\|P_n^* \varphi - \varphi\|_* = \|\gamma^{-1} P_n^* \varphi - \gamma^{-1} \varphi\| = \|\gamma^{-1} \gamma P_n \gamma^{-1} \varphi - \gamma^{-1} \varphi\| = \|P_n \gamma^{-1} \varphi - \gamma^{-1} \varphi\| \rightarrow 0$$

as  $n \rightarrow \infty$ , for each  $\varphi \in V^*$ .

For each  $n = 1, 2, \dots$  define the finite dimensional subspace  $\mathcal{H}_n$  of  $\mathcal{H}$  by

$$\mathcal{H}_n = \{\Phi_n P_n : \Phi_n \in \mathcal{L}(H_n)\}.$$

Note that  $\mathcal{H}_n$  is clearly a subspace of both  $\mathcal{H}$  and  $\mathcal{V}$  since  $H_n$  finite dimensional implies that all operators in  $\mathcal{H}_n$  are of finite rank. Define the closed convex cone  $C_n \subset \mathcal{H}_n$  by  $C_n = \{\Phi_n P_n \in \mathcal{H}_n : \Phi_n = \Phi_n^*, \Phi_n \geq 0\}$ . Using the fact that  $\mathcal{H}_n \subset \mathcal{V}$  and  $C_n \subset C_0$ , we define the operator  $A_n : \text{Dom}(A_n) \subset \mathcal{H}_n \rightarrow \mathcal{H}_n$  by

$$(4.2) \quad A_n(\Phi_n P_n) = \{A_0(\Phi_n P_n)\}|_{\mathcal{H}_n}, \Phi_n P_n \in \text{Dom}(A_n) = C_n.$$

That is  $A_n(\Phi_n P_n)$  is the element  $A_0(\Phi_n P_n)$  in  $V^*$  restricted to a linear functional on  $\mathcal{H}_n$ .

Since  $\mathcal{H}_n$  is a finite dimensional Hilbert space, the Riesz representation theorem implies that  $A_n(\Phi_n P_n) = \Psi_n P_n \in \mathcal{H}_n$  where  $\Psi_n P_n$  is the unique element in  $\mathcal{H}_n$  which satisfies

$$[A_0(\Phi_n P_n), \Theta_n P_n]_{\mathcal{H}} = [\Psi_n P_n, \Theta_n P_n]_{\mathcal{H}}, \Theta_n P_n \in \mathcal{H}_n.$$

The operators  $A_n$  defined in (4.2) are in effect the standard Galerkin approximation to  $A_0$ . Furthermore, the definition (4.2) leads to the same approximation to  $A_0$  that would be obtained via the more conventional procedure wherein the operator  $\mathcal{A}$  in the definition of  $L_0$  is replaced by its Galerkin approximation. Indeed, for each  $n = 1, 2, \dots$  define the operator  $\mathcal{A}_n \in \mathcal{L}(\mathcal{H}_n)$  by  $\mathcal{A}_n \varphi_n = \mathcal{A} \varphi_n|_{H_n}$ ,  $\varphi_n \in H_n$ . That is,  $\mathcal{A}_n \varphi_n = \psi_n$  where  $\psi_n$  is that element in  $H_n$  (once again whose existence and uniqueness is guaranteed by the Riesz representation theorem) which satisfies  $(\mathcal{A} \varphi_n, \theta_n) = (\psi_n, \theta_n)$ ,  $\theta_n \in H_n$ . Noting that since  $P_n$  is the orthogonal projection of  $H$  onto  $H_n$ , it is not difficult to argue that  $[\Phi P_n, \Theta_n]_{\mathcal{H}} = [\Phi, \Theta_n]_{\mathcal{H}}$  for all  $\Phi \in V^*$  and  $\Theta_n \in \mathcal{H}_n$ , it then follows that for  $\Phi_n P_n$  and  $\Psi_n P_n \in \mathcal{H}_n$  we have

$$\begin{aligned}
[A_n(\Phi_n P_n), \Psi_n P_n]_{\mathcal{H}} &= [A_0(\Phi_n P_n), \Psi_n P_n]_{\mathcal{H}} = [A_0(\Phi_n P_n)P_n, \Psi_n P_n]_{\mathcal{H}} \\
&= [\mathcal{A}^* \Phi_n P_n + \Phi_n P_n \mathcal{A} P_n + \mathcal{F}_0(\Phi_n P_n)P_n, \Psi_n P_n]_{\mathcal{H}} \\
&= \sum_{k=1}^{\infty} \{(\mathcal{A}^* \Phi_n P_n e_k, \Psi_n P_n e_k) + (\Phi_n P_n \mathcal{A} P_n e_k, \Psi_n P_n e_k) + (\mathcal{F}_0(\Phi_n P_n)P_n e_k, \Psi_n P_n e_k)\} \\
&= \sum_{k=1}^{\infty} \{(\mathcal{A}^* \Phi_n P_n e_k, \Psi_n P_n e_k) + (\mathcal{A} P_n e_k, \Phi_n^* \Psi_n P_n e_k) + (\mathcal{F}_0(\Phi_n P_n)P_n e_k, \Psi_n P_n e_k)\} \\
&= \sum_{k=1}^{\infty} \{(\mathcal{A}_n^* \Phi_n P_n e_k, \Psi_n P_n e_k) + (\mathcal{A}_n P_n e_k, \Phi_n^* \Psi_n P_n e_k) + (P_n \mathcal{F}_0(\Phi_n P_n)P_n e_k, \Psi_n P_n e_k)\} \\
&= \sum_{k=1}^{\infty} (\{\mathcal{A}_n^* \Phi_n + \Phi_n \mathcal{A}_n + P_n \mathcal{F}_0(\Phi_n P_n)\} P_n e_k, \Psi_n P_n e_k);
\end{aligned}$$

or

$$A_n(\Phi_n P_n) = \{\mathcal{A}_n^* \Phi_n + \Phi_n \mathcal{A}_n + \mathcal{F}_n(\Phi_n P_n)\} P_n,$$

where  $\mathcal{F}_n(\Phi) = P_n \mathcal{F}_0(\Phi)$ ,  $\Phi \in C_0$ . In particular when  $\mathcal{F}_0(\Phi) = \Phi^2$ , we have  $A_n(\Phi_n P_n) = \{\mathcal{A}_n^* \Phi_n + \Phi_n \mathcal{A}_n + \Phi_n^2\} P_n$ .

For each  $n = 1, 2, \dots$  define  $\Pi_n^0 \in C_n$  by  $\Pi_n^0 = P_n \Pi_0^0 P_n$  and  $F_n \in L_1(0, T; \mathcal{H}_n)$  by  $F_n(t) = P_n F_0(t) P_n$ , for almost every  $t \in (0, T)$ . We consider the problem of finding a

solution  $\Pi_n$  to the initial value problem in  $\mathcal{H}_n$  given by

$$(4.3) \quad \dot{\Pi}_n(t) + A_n \Pi_n(t) = F_n(t), \quad \text{a.e. } t \in (0, T)$$

$$(4.4) \quad \Pi_n(0) = \Pi_n^0$$

with  $\Pi_n(t) \in C_n$  for almost every  $t \in (0, T)$ . The definition of the operator  $A_n$  together with the properties of the operator  $A_0$  yield that for each  $n = 1, 2, \dots$   $A_n$  is closed and strongly  $\mathcal{V}$ -monotone on its domain. Hence  $A_n$  is accretive with  $\mathcal{R}(I + \lambda A_n) \supset C_n$  for all  $\lambda > 0$ . Since  $\overline{\text{Dom}(A_n)} = \overline{C_n} = C_n$ ,  $\Pi_n^0 \in C_n$ , and  $F_n \in L_1(0, T; \mathcal{H}_n)$  with  $F_n(t) \in C_n$  for almost every  $t \in (0, T)$ , Theorem 2.1 yields the existence of a unique integral solution  $\Pi_n$  to the initial value problem (4.3), (4.4) with  $\Pi_n(t) \in C_n$ , a.e.  $t \in (0, T)$ .

We shall argue the convergence of  $\Pi_n$  to  $\Pi_0$  as  $n \rightarrow \infty$  in  $C(0, T; \mathcal{H})$  (i.e. the Hilbert Schmidt norm convergence of  $\Pi_n(t)$  to  $\Pi_0(t)$ , uniformly in  $t$  for  $t \in [0, T]$ ) via an application of Theorem 2.3. In order to do this we shall require some preliminary lemmas. The first lemma below is a technical lemma which can also be found in [7]. For completeness we state it here and have included its rather brief proof.

**Lemma 4.1** If  $\{a_i\}_{i=1}^{\infty}$  is an absolutely summable sequence of real numbers, then there exist sequences  $\{b_i\}_{i=1}^{\infty}$  and  $\{c_i\}_{i=1}^{\infty}$  such that  $\lim_{i \rightarrow \infty} b_i = 0$ ,  $\{c_i\}_{i=1}^{\infty}$  is absolutely summable, and  $a_i = b_i c_i$ ,  $i = 1, 2, \dots$ .

**Proof:** Let  $\alpha = \sum_{i=1}^{\infty} |a_i|$  and for  $j = 0, 1, 2, \dots$  define the nonnegative integers  $k_j$  as follows.

Let  $k_0 = 0$  and let  $k_j$  denote the first index for which

$$\sum_{i=1}^{k_j} |a_i| > \alpha - \frac{1}{j^3}, \quad j = 1, 2, \dots$$

Set  $b_i = 1/j$  and  $c_i = ja_i$ , for  $i = k_{j-1} + 1, \dots, k_j$ ,  $j = 1, 2, \dots$ . Then  $b_i c_i = a_i$ ,  $i = 1, 2, \dots$ ,  $\lim_{i \rightarrow \infty} b_i = 0$ , and  $\sum_{i=1}^{\infty} |c_i| = \sum_{j=1}^{\infty} j \sum_{k=k_{j-1}+1}^{k_j} |a_k| \leq \alpha + \sum_{j=1}^{\infty} \frac{1}{j^2} < \infty$ .

**Lemma 4.2** Let  $X$  and  $Y$  be real separable Hilbert spaces with inner products denoted by  $\langle \cdot, \cdot \rangle_X$  and  $\langle \cdot, \cdot \rangle_Y$ , respectively. If  $\Phi \in \text{HS}(X, Y)$  then  $\Phi$  can be factored as  $\Phi = \Phi^1 \Phi^2$  with  $\Phi^1 \in \mathcal{L}(Y)$  compact and  $\Phi^2 \in \text{HS}(X, Y)$ .

**Proof:** Let  $\{x_i\}_{i=1}^{\infty}$  and  $\{y_i\}_{i=1}^{\infty}$  be orthonormal bases for  $X$  and  $Y$  respectively. Since  $\Phi \in \text{HS}(X, Y)$  it has a representation in the form of an infinite matrix  $\Phi \leftrightarrow [\varphi_{ij}] = [\langle y_i, \Phi x_j \rangle_Y]$  with  $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \varphi_{ij}^2 < \infty$ . For  $i = 1, 2, \dots$  set  $a_i = \sum_{j=1}^{\infty} \varphi_{ij}^2$ . Since the sequence  $\{a_i\}_{i=1}^{\infty}$  is absolutely summable, we can apply Lemma 4.1 and obtain sequences  $\{b_i\}_{i=1}^{\infty}$  and  $\{c_i\}_{i=1}^{\infty}$  for which  $a_i = b_i c_i$ ,  $i = 1, 2, \dots$ ,  $\lim_{i \rightarrow \infty} b_i = 0$ , and  $\sum_{i=1}^{\infty} c_i = \sum_{i=1}^{\infty} |c_i| < \infty$ . Define  $\Phi^1 \in \mathcal{L}(Y)$  and  $\Phi^2 \in \mathcal{L}(X, Y)$  by  $\Phi^1 y = \sum_{i=1}^{\infty} \sqrt{b_i} \langle y, y_i \rangle_Y y_i$ ,  $y \in Y$ , and  $\Phi^2 x = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{\varphi_{ij}}{\sqrt{b_i}} \langle x, x_j \rangle_X y_i$ ,  $x \in X$ . Then  $\Phi^1 \Phi^2 = \Phi$ , and since  $\lim_{i \rightarrow \infty} \sqrt{b_i} = 0$  and  $\sum_{i,j} \left( \frac{\varphi_{ij}}{\sqrt{b_i}} \right)^2 = \sum_{i=1}^{\infty} \frac{1}{b_i} \sum_{j=1}^{\infty} \varphi_{ij}^2 = \sum_{i=1}^{\infty} \frac{a_i}{b_i} = \sum_{i=1}^{\infty} c_i < \infty$ , it follows that  $\Phi^1$  is compact and  $\Phi^2 \in \text{HS}(X, Y)$ .

### **Lemma 4.3**

$$(a) \quad \lim_{n \rightarrow \infty} |P_n \Phi P_n - \Phi|_{\mathcal{H}} = 0, \quad \Phi \in \mathcal{H}.$$

$$(b) \quad \lim_{n \rightarrow \infty} \|P_n \Phi P_n - \Phi\|_{\mathcal{V}} = 0, \quad \Phi \in \mathcal{V}.$$

**Proof:** (a). We consider  $P_n$  to be an element in  $\mathcal{L}(H)$ . Then for  $\Phi \in \mathcal{H} = \text{HS}(H, H)$  we have

$$|P_n \Phi P_n - \Phi|_{\mathcal{H}} \leq |P_n \Phi P_n - P_n \Phi|_{\mathcal{H}} + |P_n \Phi - \Phi|_{\mathcal{H}}$$

$$\begin{aligned}
&\leq |\Phi P_n - \Phi|_{\mathcal{H}} + |P_n \Phi - \Phi|_{\mathcal{H}} \\
&= |(\Phi P_n)^* - \Phi^*|_{\mathcal{H}} + |P_n \Phi - \Phi|_{\mathcal{H}} \\
&= |P_n \Phi^* - \Phi^*|_{\mathcal{H}} + |P_n \Phi - \Phi|_{\mathcal{H}}.
\end{aligned}$$

Since  $\Phi, \Phi^* \in \text{HS}(H, H)$ , we can apply Lemma 4.2 with  $X = Y = H$  to obtain  $\Phi = \Phi^1 \Phi^2$  and  $\Phi^* = (\Phi^*)^1 (\Phi^*)^2$  with  $\Phi^1, (\Phi^*)^1 \in \mathcal{L}(H)$  compact and  $\Phi^2, (\Phi^*)^2 \in \text{HS}(H, H)$ . Then

$$|P_n \Phi - \Phi|_{\mathcal{H}} = |(P_n - I) \Phi^1 \Phi^2|_{\text{HS}(H, H)} \leq |(P_n - I) \Phi^1|_{\mathcal{L}(H)} |\Phi^2|_{\text{HS}(H, H)}.$$

The fact that  $P_n$  converges strongly to the identity on  $H$  and that  $\Phi^1 \in \mathcal{L}(H)$  is compact imply  $|P_n \Phi - \Phi|_{\mathcal{H}} \rightarrow 0$  as  $n \rightarrow \infty$ . Similar estimates for  $|P_n \Phi^* - \Phi^*|_{\mathcal{H}}$  yield the desired result.

(b). Let  $\Phi \in \mathcal{V} = \text{HS}(V^*, H) \cap \text{HS}(H, V)$  and consider  $\|P_n \Phi P_n - \Phi\|_{\text{HS}(V^*, H)}$ .

Since  $|P_n| = 1$  it follows that

$$\begin{aligned}
\|P_n \Phi P_n - \Phi\|_{\text{HS}(V^*, H)} &\leq \|P_n \Phi P_n - P_n \Phi\|_{\text{HS}(V^*, H)} + \|P_n \Phi - \Phi\|_{\text{HS}(V^*, H)} \\
&\leq \|\Phi P_n - \Phi\|_{\text{HS}(V^*, H)} + \|P_n \Phi - \Phi\|_{\text{HS}(V^*, H)}.
\end{aligned}$$

Now  $\Phi \in \text{HS}(V^*, H)$  implies that  $\Phi^* \in \text{HS}(H, V^*)$  and that  $(\Phi P_n)^* = P_n^* \Phi^* \in \text{HS}(H, V^*)$  where  $P_n^*$  denotes the adjoint of  $P_n$  as an element of  $\mathcal{L}(V^*)$ . Then

$$(4.5) \quad \|P_n \Phi P_n - \Phi\|_{\text{HS}(V^*, H)} \leq \|P_n^* \Phi^* - \Phi^*\|_{\text{HS}(H, V^*)} + \|P_n \Phi - \Phi\|_{\text{HS}(V^*, H)}.$$

An applicaiton of Lemma 4.2 with  $X = H$  and  $Y = V^*$  and the strong convergence of  $P_n^*$  to the identity on  $V^*$  yield that the first term on the right hand side of the estimate (4.5) above tends to zero as  $n \rightarrow \infty$ . Similarly, Lemma 4.2 with  $X = V^*$  and  $Y = H$  and the strong convergence of  $P_n$  to the identity on  $H$  implies that the second term tends to zero as  $n \rightarrow \infty$  as well. A similar argument can be used to show that  $\|P_n \Phi P_n - \Phi\|_{\text{HS}(H, V)} \rightarrow 0$  as  $n \rightarrow \infty$  and the lemma is proved.

**Theorem 4.1** If  $\Pi_0 \in C(0, T; \mathcal{H})$  is the unique integral solution to the initial value problem (3.1), (3.2) (or, equivalently, (3.4), (3.5)) and  $\Pi_n \in C(0, T; \mathcal{H}_n)$  is the unique integral solution to the initial value problem (4.3), (4.4), then  $\Pi_n$  converges to  $\Pi_0$  in  $C(0, T; \mathcal{H})$  as  $n \rightarrow \infty$ . That is,  $\lim_{n \rightarrow \infty} |\Pi_n(t) - \Pi_0(t)|_{\mathcal{H}} = 0$  with the convergence uniform in  $t$  for  $t \in [0, T]$ .

**Proof** The desired result will follow immediately from Theorem 2.3 once we have verified that the hypotheses (i)-(iv) given in the statement of that theorem hold. If  $\Phi_0 \in C_0$  then  $\Phi_n = P_n \Phi_0 P_n \in C_n$  and Lemma 4.3 implies that  $\lim_{n \rightarrow \infty} |\Phi_n - \Phi_0|_{\mathcal{H}} = 0$ . Thus  $\lim_{n \rightarrow \infty} C_n \supset C_0$ . Lemma 4.3 also implies that  $\lim_{n \rightarrow \infty} |F_n(t) - F_0(t)|_{\mathcal{H}} = \lim_{n \rightarrow \infty} |P_n F_0(t) P_n - F_0|_{\mathcal{H}} = 0$  for almost every  $t \in (0, T)$ . Properties of Hilbert-Schmidt operators and the fact that  $P_n$  is an orthogonal projection yield  $|F_n(t)|_{\mathcal{H}} = |P_n F_0(t) P_n|_{\mathcal{H}} \leq |F_0(t)|_{\mathcal{H}} \in L_1(0, T)$ , for a.e.  $t \in (0, T)$ . Consequently hypothesis (ii) is satisfied. Once again from Lemma 4.3 we obtain  $\lim_{n \rightarrow \infty} |\Pi_n^0 - \Pi_0^0|_{\mathcal{H}} = \lim_{n \rightarrow \infty} |P_n \Pi_0^0 P_n - \Pi_0^0|_{\mathcal{H}} = 0$ . The verification of hypothesis (iv) is all that remains.

Let  $\lambda > 0$  and let  $\Phi_n \in C_n$ ,  $n = 0, 1, 2, \dots$  with  $\lim_{n \rightarrow \infty} |\Phi_n - \Phi_0|_{\mathcal{H}} = 0$ . Set  $\Psi_n = J_n(\lambda) \Phi_n$ ,  $n = 0, 1, 2, \dots$  where  $J_n(\lambda) = (I + \lambda A_n)^{-1}$ . Then recalling (3.3) we have

$$\begin{aligned}
\lambda \omega \|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}}^2 &\leq \lambda [L_0 \{\Psi_n - P_n \Psi_0 P_n\}, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \\
&= [(I + \lambda A_n) \Psi_n - (I + \lambda A_0) \Psi_0 P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} - [\Psi_n - \Psi_0 P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \\
&\quad + \lambda [L_0 \{\Psi_0 - P_n \Psi_0 P_n\}, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} - \lambda [\mathcal{F}_n(\Psi_n) P_n - \mathcal{F}_0(\Psi_0) P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \\
&= [\Phi_n - P_n \Phi_0 P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} - [\Psi_n - P_n \Psi_0 P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \\
&\quad + \lambda [L_0 \{\Psi_0 - P_n \Psi_0 P_n\} P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} - \lambda [\mathcal{F}_0(\Psi_n) P_n - \mathcal{F}_0(P_n \Psi_0 P_n) P_n, \\
&\quad \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} + \lambda [\mathcal{F}_0(\Psi_0) P_n - \mathcal{F}_0(P_n \Psi_0 P_n) P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}}
\end{aligned}$$

$$\begin{aligned}
&\leq [\Phi_n - P_n \Phi_0 P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} + \lambda [L_0 \{\Psi_0 - P_n \Psi_0 P_n\} P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \\
&+ \lambda [\mathcal{F}_0(\Psi_0) P_n - \mathcal{F}_0(P_n \Psi_0 P_n) P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \\
&\leq \|\Phi_n - P_n \Phi_0 P_n\|_{\mathcal{V}^*} \|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}} + \lambda \|L_0 \{\Psi_0 - P_n \Psi_0 P_n\} P_n\|_{\mathcal{V}^*} \cdot \\
&\|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}} + \lambda \|\mathcal{F}_0(\Psi_0) P_n - \mathcal{F}_0(P_n \Psi_0 P_n) P_n\|_{\mathcal{V}^*} \|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}}.
\end{aligned}$$

where in the above estimate we have used the facts that  $[\mathcal{F}_0(\Psi_n) - \mathcal{F}_0(P_n \Psi_0 P_n) P_n, \Psi_n - P_n \Psi_0 P_n] \geq 0$  and  $P_n$  nonnegative self-adjoint and compact (being of finite rank) imply  $[\mathcal{F}_0(\Psi_n) P_n - \mathcal{F}_0(P_n \Psi_0 P_n) P_n, \Psi_n - P_n \Psi_0 P_n]_{\mathcal{H}} \geq 0$ . Thus

$$\begin{aligned}
\|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}} &\leq \frac{k}{\lambda \omega} |\Phi_n - P_n \Phi_0 P_n|_{\mathcal{H}} \\
&+ (1/\omega) \{ \|\mathcal{A}^*\|_{\mathcal{L}(\mathcal{V}, \mathcal{V}^*)} |P_n|_{\mathcal{L}(\mathcal{H})} + \|\mathcal{A}\|_{\mathcal{L}(\mathcal{V}, \mathcal{V}^*)} |P_n|_{\mathcal{L}(\mathcal{V})} \} \|\Psi - P_n \Psi_0 P_n\|_{\mathcal{V}} \\
&+ \frac{k}{\omega} |\mathcal{F}_0(\Psi_0) - \mathcal{F}_0(P_n \Psi_0 P_n)|_{\mathcal{H}}
\end{aligned}$$

for some positive constant  $k$ . Now

$$|\Phi_n - P_n \Phi_0 P_n|_{\mathcal{H}} \leq |\Phi_n - \Phi_0|_{\mathcal{H}} + |\Phi_0 - P_n \Phi_0 P_n|_{\mathcal{H}}$$

and

$$\|\mathcal{A}^*\|_{\mathcal{L}(\mathcal{V}, \mathcal{V}^*)} |P_n|_{\mathcal{L}(\mathcal{H})} + \|\mathcal{A}\|_{\mathcal{L}(\mathcal{V}, \mathcal{V}^*)} |P_n|_{\mathcal{L}(\mathcal{V})} \leq \beta(1 + |P_n|_{\mathcal{L}(\mathcal{V})})$$

which is uniformly bounded in  $n$ . The assumption that  $\lim_{n \rightarrow \infty} \Phi_n = \Phi_0$ , Lemma 4.3, and

$\lim_{n \rightarrow \infty} |\Psi_0 - P_n \Psi_0 P_n|_{\mathcal{H}} = 0$  together with the continuity of the map  $\Theta \rightarrow \mathcal{F}_0(\Theta)$  from  $\mathcal{H}$

into itself yield  $\lim_{n \rightarrow \infty} \|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}} = 0$ . Consequently the estimate

$$\begin{aligned}
\lim_{n \rightarrow \infty} |\Psi_n - \Psi_0|_{\mathcal{H}} &\leq K \lim_{n \rightarrow \infty} \|\Psi_n - \Psi_0\|_{\mathcal{V}} \\
&\leq K \lim_{n \rightarrow \infty} \|\Psi_n - P_n \Psi_0 P_n\|_{\mathcal{V}} + K \lim_{n \rightarrow \infty} \|P_n \Psi_0 P_n - \Psi_0\|_{\mathcal{V}}
\end{aligned}$$

together with Lemma 4.3 yield the desired result.

**Remark 4.1** Although we are unable to demonstrate that the hypotheses of the existence result, Theorem 2.1, are satisfied when the initial value problem (3.4), (3.5) is considered in the space  $\mathcal{V}$ , we can show that hypotheses (i)-(iv) of Theorem 2.3 are in fact satisfied in the stronger  $\mathcal{V}$  - topology. More precisely if  $\Pi_0^0 \in \mathcal{V}$  and  $F_0 \in L_1(0, T; \mathcal{V})$ , then Lemma 4.3 is sufficient to obtain the convergence of  $\Pi_n$  to  $\Pi_0$  in  $C(0, T; \mathcal{V})$ . In the case of a linear dynamical system (i.e. when  $A_0 \in \mathcal{L}(\mathcal{V}, \mathcal{V}^*)$ ) it is in fact possible to apply Theorem 2.1 in the space  $\mathcal{V}$  and therefore obtain a complete approximation theory in  $\mathcal{V}$  (see [1]).

## 5. An Example

We illustrate the application of our approximation results with a linear-quadratic optimal control problem (see [10]) involving a one dimensional heat equation (extension to higher dimensions is straightforward). Let  $H = L_2(0, 1)$  endowed with the usual inner product,  $(\varphi, \psi) = \int_0^1 \varphi(\eta)\psi(\eta)d\eta$ . Consider the optimal control problem of finding  $\bar{u} \in L_2((0, T) \times (0, 1))$  which minimizes the quadratic performance index

$$J(u) = (\mathcal{G}x(T, \cdot), x(T, \cdot)) + \int_0^T (Qx(t, \cdot), x(t, \cdot)) + r(u(t, \cdot), u(t, \cdot))dt$$

subject to the linear distributed dynamical system

$$(5.1) \quad \frac{\partial x}{\partial t}(t, \eta) - \frac{\partial}{\partial \eta} a(\eta) \frac{\partial x}{\partial \eta}(t, \eta) = bu(t, \eta), \quad t > 0, \quad 0 < \eta < 1,$$

$$(5.2) \quad x(t, 0) = 0, \quad x(t, 1) = 0, \quad t > 0$$

$$(5.3) \quad x(0, \eta) = x^0(\eta), \quad 0 < \eta < 1,$$

where  $a \in L_\infty(0, 1)$  with  $a(\eta) \geq \alpha > 0$ , a.e.  $\eta \in (0, 1)$ ,  $b, r \in \mathbf{R}$  with  $r > 0$ ,  $x^0 \in L_2(0, 1)$ ,

$\mathcal{G}, Q \in C_0 \equiv \{\Phi \in HS(L_2(0, 1), L_2(0, 1)), \Phi = \Phi^*, \Phi \geq 0\}$  Note that  $\mathcal{G}, Q \in C_0$  implies

that  $(\mathcal{G}\varphi)(\eta) = \int_0^1 g(\eta, \xi)\varphi(\xi)d\xi$  and  $(\mathcal{Q}\varphi)(\eta) = \int_0^1 q(\eta, \xi)\varphi(\xi)d\xi$  with  $g, q \in L_2((0, 1) \times (0, 1))$ ,  $g(\eta, \xi) = g(\xi, \eta)$ ,  $g(\eta, \xi) \geq 0$ ,  $q(\eta, \xi) = q(\xi, \eta)$ ,  $q(\eta, \xi) \geq 0$ , a.e.  $(\eta, \xi) \in (0, 1) \times (0, 1)$ .

If we set  $V = H_0^1(0, 1)$  endowed with the usual inner product,  $\langle \varphi, \psi \rangle = \int_0^1 D\varphi(\eta)D\psi(\eta)d\eta$ , and corresponding induced norm,  $\|\cdot\|$ , then  $V^* = H^{-1}(0, 1)$ . Define the operator  $\mathcal{A} \in \mathcal{L}(V, V^*)$  by  $(\mathcal{A}\varphi)(\psi) = (aD\varphi, D\psi)$ ,  $\varphi, \psi \in V$ . Then  $(\mathcal{A}\varphi, \varphi) \geq \alpha\|\varphi\|^2$ ,  $\varphi \in V$  and the restriction  $A$  of the operator  $\mathcal{A}$  to the set  $\text{Dom}(A) = \{\varphi \in V : \mathcal{A}\varphi \in H\} (= H^2(0, 1) \cap H_0^1(0, 1)$  when  $a$  is sufficiently smooth) is positive, self-adjoint and the infinitesimal generator of a uniformly exponentially stable analytic (parabolic) semi-group of bounded, self-adjoint linear operators  $\{\mathcal{T}(t) : t \geq 0\}$  on  $H$ . Let  $U = L_2(0, 1)$  and define the operator  $\mathcal{B} \in \mathcal{L}(U, H)$  by  $(\mathcal{B}v)(\eta) = bv(\eta)$ ,  $v \in U, 0 < \eta < 1$ , and define  $\mathcal{R} \in \mathcal{L}(U)$  by  $(\mathcal{R}v)(\eta) = rv(\eta)$ ,  $v \in U, 0 < \eta < 1$ .

The solution to the initial value problem (5.1) - (5.3) is given by

$$x(t, \cdot) = \mathcal{T}(t)x^0 + \int_0^t \mathcal{T}(t-s)\mathcal{B}u(s)ds, \quad t > 0.$$

where for  $u \in L_2((0, T) \times (0, 1))$  we have used the shorthand notation  $t \rightarrow u(t)$  to denote the function  $t \rightarrow u(t, \cdot) \in L_2(0, T; U)$ . The solution to the optimal control problem is given in closed-loop, linear state feedback form (see [10]) by

$$(5.4) \quad \bar{u}(t) = -\mathcal{R}^{-1}\mathcal{B}^*\Pi_0(T-t)x(t, \cdot) = -(b/r)\Pi_0(T-t)x(t, \cdot), \quad \text{a.e. } t \in (0, T),$$

where  $\Pi_0$  is the nonnegative self-adjoint solution to the Riccati differential equation initial value problem

$$(5.5) \quad \dot{\Pi}_0(t) + \mathcal{A}\Pi_0(t) + \Pi_0(t)\mathcal{A} + \Pi_0(t)\mathcal{B}\mathcal{R}^{-1}\mathcal{B}^*\Pi_0(t) = \mathcal{Q} \quad \text{a.e. } t \in (0, T)$$

$$(5.6) \quad \Pi_0(0) = \mathcal{G}.$$

If we set  $M = \mathcal{B}\mathcal{R}^{-1}\mathcal{B}^* = (b^2/r)I \in \mathcal{L}(H)$ , then  $M = M^*$ ,  $M \geq 0$  and, since  $\Phi M = M\Phi$ ,  $\Phi \in C_0$ ,  $\Phi M \geq 0$ ,  $\Phi \in C_0$ . Thus Temam's [15] theory (with  $F_0(t) = \mathcal{Q}$ ,  $t \in [0, T]$ ,  $\mathcal{F}_0(\Phi) = (b^2/r)\Phi^2$ , and  $\Pi_0^0 = \mathcal{G}$ ) presented in section 3 above yields the existence of a unique integral solution  $\Pi_0 \in C(0, T; \mathcal{H}S(H, H))$  to the initial value problem (5.5), (5.6) with  $\Pi_0 \in L_2(0, T; \mathcal{H}S(H, V) \cap \mathcal{H}S(V^*, H))$ ,  $\dot{\Pi}_0 \in L_2(0, T; \mathcal{H}S(H, V^*) + \mathcal{H}S(V, H))$  and  $\Pi_0(t) \in C_0$ ,  $t \in [0, T]$ . (If  $\mathcal{G}$  is such that  $\mathcal{A}\mathcal{G} + \mathcal{G}\mathcal{A} \in \mathcal{H}S(H, H)$  - for example if  $\mathcal{G} = 0$  - then  $\Pi_0$  will be a strong solution with  $\Pi_0 \in W^{1,\infty}(0, T; \mathcal{H}S(H, H))$ ). Since  $\mathcal{H} = \mathcal{H}S(L_2(0, 1), L_2(0, 1))$  is isometrically isomorphic to  $L_2((0, 1) \times (0, 1))$ , it follows that there exists a  $\pi_0 \in C(0, T; L_2((0, 1) \times (0, 1)))$  with  $\pi_0(t, \eta, \xi) = \pi_0(t; \xi, \eta)$ ,  $\pi_0(t, \eta, \xi) \geq 0$ , a.e.  $(\eta, \xi) \in (0, 1) \times (0, 1)$ ,  $t \in [0, T]$  such that the solution to the optimal control problem, (5.4), is given by

$$\bar{u}(t, \eta) = -\frac{b}{r} \int_0^1 \pi_0(T-t; \eta, \xi) x(t, \xi) d\xi,$$

for almost every  $(t, \eta) \in (0, T) \times (0, 1)$ .

We consider a linear spline based approximation scheme. For each  $n = 2, 3, \dots$  let  $H_n = \text{span}\{\varphi_n^j\}_{j=1}^{n-1}$  where for  $j = 1, 2, \dots, n-1$ ,  $\varphi_n^j$  is the  $j$ th linear spline ("hat") function on  $[0, 1]$  defined with respect to the uniform mesh  $\{0, 1/n, 2/n, \dots, 1\}$ . That is

$$\varphi_n^j(\eta) = \begin{cases} 0 & 0 \leq \eta \leq \frac{j-1}{n} \\ n\eta - j + 1 & \frac{j-1}{n} \leq \eta \leq \frac{j}{n} \\ j + 1 - n\eta & \frac{j}{n} \leq \eta \leq \frac{j+1}{n} \\ 0 & \frac{j+1}{n} \leq \eta \leq 1. \end{cases}$$

Let  $P_n : H \rightarrow H_n$  denote the orthogonal projection of  $H$  onto  $H_n$  with respect to the  $(\cdot, \cdot)$  inner product and define  $\mathcal{A}_n \in \mathcal{L}(H_n)$  to be the Galerkin approximation to  $\mathcal{A}$ . More

precisely, we set  $\mathcal{A}_n \varphi_n = \psi_n$ ,  $\varphi_n, \psi_n \in H_n$ , where  $\psi_n$  is the unique element in  $H_n$  which satisfies  $(\mathcal{A} \varphi_n, \theta_n) = (\psi_n, \theta_n)$ ,  $\theta_n \in H_n$ . Let  $\mathcal{Q}_n = P_n \mathcal{Q} \in \mathcal{L}(H_n)$  and  $\mathcal{G}_n = P_n \mathcal{G} \in \mathcal{L}(H_n)$ .

Using the properties of interpolatory splines and density arguments it is not difficult to show that  $\lim_{n \rightarrow \infty} \|P_n \varphi - \varphi\| = 0$ ,  $\varphi \in H_0^1(0,1) = V$  and consequently that the assumption (4.1) is satisfied. It follows therefore, that the approximation theory developed in section 4 yields that the solution  $\Pi_n$  to the finite dimensional Riccati differential equation initial value problem

$$(5.7) \quad \dot{\Pi}_n(t) + \mathcal{A}_n \Pi_n(t) + \Pi_n(t) \mathcal{A}_n + (b^2/r) \Pi_n(t)^2 = \mathcal{Q}_n, \quad \text{a.e. } t \in (0, T),$$

$$(5.8) \quad \Pi_n(0) = \mathcal{G}_n$$

with  $\Pi_n(t)$  nonnegative and self-adjoint, satisfies  $\lim_{n \rightarrow \infty} |\Pi_n(t) P_n - \Pi_0(t)|_{HS(H,H)} = 0$  uniformly in  $t$  for  $t \in [0, T]$ .

Since the basis  $\{\varphi_n^j\}_{j=1}^{n-1}$  is not orthonormal, simply replacing the operators in (5.7), (5.8) with their matrix representations will not lead to the familiar symmetric matrix Riccati differential equation. For a linear operator  $L_n$  with domain and/or range in  $H_n$ , we denote its matrix representation with respect to the basis  $\{\varphi_n^j\}_{j=1}^{n-1}$  by  $L_N$ . Define  $\Phi_n : [0, 1] \rightarrow \mathbb{R}^{n-1}$  by  $\Phi_n(\eta) = (\varphi_n^1(\eta), \dots, \varphi_n^{n-1}(\eta))^T$  and set  $\mathcal{M}_N = (\Phi_n, \Phi_n^T) = \int_0^1 \Phi_n(\eta) \Phi_n(\eta)^T d\eta$ . Then  $\mathcal{A}_N = \mathcal{M}_N^{-1}(a D \Phi_n, D \Phi_n^T)$ ,  $G_N = \mathcal{M}_N^{-1}(\mathcal{G} \Phi_n, \Phi_n^T)$  and  $\mathcal{Q}_N = \mathcal{M}_N^{-1}(\mathcal{Q} \Phi_n, \Phi_n^T)$ . If we let  $\tilde{\mathcal{G}}_N = \mathcal{M}_N \mathcal{G}_N$ ,  $\tilde{\mathcal{Q}}_N = \mathcal{M}_N \mathcal{Q}_N$  and  $\tilde{\Pi}_N(t) = \mathcal{M}_N \Pi_n(t)$ , then  $\tilde{\Pi}_N$  is the solution to the initial value problem

$$\dot{\tilde{\Pi}}_N(t) + \mathcal{A}_N^T \tilde{\Pi}_N(t) + \tilde{\Pi}_N(t) \mathcal{A}_N + (b^2/r) \tilde{\Pi}_N(t) \mathcal{M}_N^{-1} \tilde{\Pi}_N(t) = \tilde{\mathcal{Q}}_N, \quad t \in (0, T)$$

$$\tilde{\Pi}_N(0) = \tilde{\mathcal{G}}_N$$

with  $\tilde{\Pi}_N(t)$  nonnegative and self-adjoint. The approximating solution to the optimal control problem then takes the form

$$\bar{u}_n(t, \eta) = -\frac{b}{r} \int_0^1 \pi_n(T-t; \eta, \xi) x(t, \xi) d\xi, \quad \text{a.e. } (t, \eta) \in (0, T) \times (0, 1)$$

where  $\pi_n \in C(0, T; L_2((0, 1) \times (0, 1)))$  is given by

$$\pi_n(t; \eta, \xi) = \Phi_n(\eta)^T \mathcal{M}_N^{-1} \tilde{\Pi}_N(t) \mathcal{M}_N^{-1} \Phi_n(\xi),$$

for  $t \in [0, T]$ ,  $(\eta, \xi) \in [0, 1] \times [0, 1]$ . Our convergence theory yields  $\lim_{n \rightarrow \infty} \pi_n = \pi_0$  in  $L_2((0, 1) \times (0, 1))$  uniformly in  $t$  for  $t \in [0, T]$ . That is

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, T]} \int_0^1 \int_0^1 |\pi_n(t; \eta, \xi) - \pi_0(t; \eta, \xi)|^2 d\eta d\xi = 0.$$

We note that some interesting open questions which are related to our effort here remain. For example, a Hilbert-Schmidt existence, approximation, and convergence theory for the case in which the operator  $\mathcal{A}$  (and therefore the operator  $A_0$ ) is time varying. The standard existence and approximation results in the spirit of Theorem 2.1-2.3 for nonlinear systems involving time dependent operators are typically too restrictive to be of any use in the Hilbert space framework in which the problem would be formulated. An approach such as the one which was taken in [2] in the context of approximation methods for inverse problems for nonautonomous nonlinear distributed systems which was based upon Barbu's [3] Theorem III. 4.2 may be more appropriate.

Finally a Hilbert-Schmidt approximation theory for the steady state, or algebraic operator Riccati equation can also be developed (see [12]). (The algebraic Riccati equation arises in the context of the linear-quadratic optimal control problem on the infinite time interval).

## References

- [1] Banks, H.T. and K. Ito, A unified framework for approximation and inverse problems for distributed parameter systems, *Control-Theory and Advanced Technology*, 4, (1988) pp. 73-90.
- [2] Banks, H.T., S. Reich, and I.G. Rosen, Galerkin approximation for inverse problems for nonautonomous nonlinear distributed systems, ICASE Report 88-38, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA, 1988, and *Applied Mathematics and Optimization*, submitted.
- [3] Barbu, V., Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff International, Leyden, The Netherlands, 1976.
- [4] Crandall, M.G. and A. Pazy, Nonlinear evolution equations in Banach space, *Israel J. Math.* 11(1972), pp. 57-94.
- [5] Dunford, N. and J.T. Schwartz, Linear Operators, Part II, Spectral Theory, Self-Adjoint Operators in Hilbert Space, Wiley-Interscience, New York, 1963.
- [6] Gibson, J.S., The Riccati integral equations for optimal control problems on Hilbert spaces, *SIAM J. Contr. and Opt.*, 17(1979), pp. 537-565.
- [7] Gibson, J.S. and I.G. Rosen, Numerical approximation for the infinite-dimensional discrete-time optimal linear-quadratic regulator problem, *SIAM J. Contr. and Opt.*, 26(1988), pp. 428-452.
- [8] Goldstein, J.A., Approximation of nonlinear semigroups and evolution equations, *J. Math. Soc. Japan*, 24(1972), pp. 558-573.
- [9] Kisynski, J., A proof of the Trotter-Kato theorem on approximation of semigroups, *Colloq. Math.* 18(1967), pp. 181-184.
- [10] Lions, J.L., Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, New York, 1971.
- [11] Pazy, A., Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer-Verlag, New York, 1971.
- [12] Rosen, I.G., On Hilbert-Schmidt norm convergence of Galerkin approximations to the solution of infinite dimensional operator algebraic Riccati equations, *Proceedings of the Fourth International Conference on the Control and Identification of Distributed Parameter Systems*, held in Vorau Austria, July 10-16, 1988, Birkhauser Verlag, submitted.
- [13] Tanabe, H., Equations of Evolution, Pitman, London, 1979.
- [14] Teman, R., Etude directe d'une équation d'évolution du type de Riccati associée à des opérateurs non bornés, *C.R. Acad. Sc. Paris, Serie A.*, 268(1969), pp. 1335-1338.
- [15] Teman, R., Sur l'équation de Riccati associée à des opérateurs non bornés, en dimension infinie, *J. Functional Analysis*, 7(1971), pp. 85-115.



## Report Documentation Page

1. Report No. NASA CR-181706 ICASE Report No. 88-49		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle  CONVERGENCE OF GALERKIN APPROXIMATIONS FOR OPERATOR RICCATI EQUATIONS -- A NONLINEAR EVOLUTION EQUATION APPROACH				5. Report Date  August 1988	
				6. Performing Organization Code	
7. Author(s)  I. G. Rosen				8. Performing Organization Report No.  88-49	
9. Performing Organization Name and Address  Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23665-5225				10. Work Unit No.  505-90-21-01	
				11. Contract or Grant No.  NAS1-18107	
12. Sponsoring Agency Name and Address  National Aeronautics and Space Administration Langley Research Center Hampton, VA 23665-5225				13. Type of Report and Period Covered  Contractor Report	
				14. Sponsoring Agency Code	
15. Supplementary Notes  Langley Technical Monitor: Submitted to SIAM J. Numer. Richard W. Barnwell Anal.  Final Report					
16. Abstract  We develop an approximation and convergence theory for Galerkin approximations to infinite dimensional operator Riccati differential equations formulated in the space of Hilbert-Schmidt operators on a separable Hilbert space. We treat the Riccati equation as a nonlinear evolution equation with dynamics described by a nonlinear monotone perturbation of a strongly coercive linear operator. We prove a generic approximation result for quasi-autonomous nonlinear evolution systems involving accretive operators which we then use to demonstrate the Hilbert-Schmidt norm convergence of Galerkin approximations to the solution of the Riccati equation. We illustrate the application of our results in the context of a linear quadratic optimal control problem for a one dimensional heat equation.					
17. Key Words (Suggested by Author(s))  Riccati differential equation, Hilbert-Schmidt operator, Galerkin approximation, nonlinear evolution equation			18. Distribution Statement  64 - Numerical Analysis 66 - Systems Analysis  Unclassified - unlimited		
19. Security Classif. (of this report)  Unclassified	20. Security Classif. (of this page)  Unclassified		21. No. of pages  29	22. Price  A03	