

# Quadratic Optimal Control of Regular Well-Posed Linear Systems, with Applications to Parabolic Equations

Olof J. Staffans  
Åbo Akademi University  
Department of Mathematics  
FIN-20500 Åbo, Finland  
Olof.Staffans@abo.fi  
<http://www.abo.fi/~staffans/>

September 24, 1997

**Abstract.** We study the nonstandard infinite horizon quadratic cost minimization problem in a particular subclass of Salamon's and Weiss' well-posed linear systems, more general than the Pritchard-Salamon class. In particular, instead of just working with controls that are  $L^2$  in time, we also investigate how the system behaves under the action of continuous controls. The additional structure has been modelled after existing results for parabolic equations, and it provides a unified framework for many of these results. We prove that in this setting all possible input/output maps (both open and closed loop; both original output and feedback output) are regular together with their adjoints in the sense of Weiss, and that the recently discovered correction term in the algebraic Riccati equation is absent in this case.

**AMS Subject Classification** 49J27, 49J20, 93A05.

**Keywords** Well-posed linear system, algebraic Riccati equation, parabolic equations.

# 1 Introduction

We study infinite-dimensional systems that can formally be written in the familiar form

$$\begin{aligned}x'(t) &= Ax(t) + Bu(t), \\y(t) &= Cx(t) + Du(t), \quad t \in \mathbf{R}^+, \\x(0) &= x_0.\end{aligned}\tag{1}$$

We interpret this equation in a weak “integral” sense that will be made precise in a moment. As for now, let us think about the state  $x$  as a continuous function of  $t$  in a Hilbert space  $H$ , and let us think of  $u$  and  $y$  are  $L^2$ -functions with values in two more Hilbert spaces  $U$  (the input space) and  $Y$  (the output space), respectively, which satisfy (1) in a generalized sense.

For this system we define a (slightly nonstandard) quadratic cost minimization problem as follows. We are given a weighting operator  $J = J^* \in \mathcal{L}(Y)$ , and, for each  $x_0 \in H$ , we minimize the cost

$$Q(x_0, u) = \langle y, Jy \rangle_{L^2(\mathbf{R}^+; Y)},\tag{2}$$

where  $y$  is the output of (1) defined in formula (24) below.<sup>1</sup> The minimization takes place over all those  $u \in L^2(\mathbf{R}^+; U)$  for which the corresponding output  $y$  belongs to  $L^2(\mathbf{R}^+; Y)$ .<sup>2</sup>

The problem that we have outlined above has been studied in, e.g., Staffans [1997abc] under minimal assumptions on  $A$ ,  $B$ , and  $C$ . The theory developed there was based on spectral factorization, and some parts of that theory required a somewhat restricting “regular spectral factorization assumption”. The purposes of the present work is to present a class of systems for which this regular spectral factorization assumption is always satisfied. The most important example belonging to this class is a system built around an analytic semigroup, with operators  $B$  and  $C$  that allow the “standard” amount of unboundedness that one finds in the literature for this class of systems.<sup>3</sup>

---

<sup>1</sup>As the examples discussed in Section 6 show, the standard quadratic cost minimization problem where  $D = 0$  and there is a direct cost on the input  $u$  is a special case of this problem.

<sup>2</sup>We shall actually most of the time limit the discussion to stable systems, in which case the condition  $y \in L^2(\mathbf{R}^+; Y)$  is redundant.

<sup>3</sup>That is,  $C(\alpha I - A)^{-\gamma}B$  is bounded for some  $\gamma < 1$  and all  $\alpha \in \rho(A)$ .

Our assumptions being “standard”, one asks the question as to what degree our conclusions are new, especially since the quadratic optimal control theory for parabolic equations appears to be in a fairly mature state. The major part of this theory up to the early 90’s is summarized in Bensoussan et al. [1992], Lasiecka and Triggiani [1991b], and some more recent developments are found in, e.g., Lasiecka et al. [1995 1997], Pandolfi [1997], Triggiani [1994]. However, none of these works say anything explicit about the input/output behavior of the resulting closed loop systems. This information may not be that important in the standard quadratic cost minimization problem, but it is crucial in, for example, the  $H^\infty$ -theory, where the original problem is to minimize the norm of a particular closed loop input/output map. Thus, we find it necessary to study the input/output behavior of the closed loop systems appearing in the references cited above.

However, this is not the only motivation, not even the main motivation for the present work. It appears to be possible, even straightforward, to add the missing statements about the input/output behavior of the closed loop systems to the references cited above by using their original technique. On the other hand, appropriate versions of these statements are already found in the theory based on spectral factorization developed in Mikkola [1997], Staffans [1997abcde], Weiss and Weiss [1997], and it is very tempting to try to combine the two separate theories, the one based on spectral factorization, and the one based on the analyticity of the semigroup, into one. Doing so we expand both theories. On one hand, we get a better understanding of the input/output behavior of systems built around an analytic semigroup under the action of  $L^2$ -controls (this behavior plays a very minor role in the existing theory for parabolic systems), and on the other hand, we gain information about the behavior of well-posed linear systems with a strong internal damping under the action of continuous controls (as opposed to  $L^2$ -controls). In particular, as we mentioned above, for this class of systems we are able to verify the troublesome regular spectral factor assumption which is needed for the Riccati equation theory.

We use the following notation:

$\mathcal{L}(U; Y)$ ,  $\mathcal{L}(U)$ : The set of bounded linear operators from  $U$  into  $Y$  or from  $U$  into itself, respectively.

$I$ : The identity operator.

$A^*$ : The (Hilbert space) adjoint of the operator  $A$ .

$A \geq 0$ :  $A$  is (selfadjoint and) positive definite.

- $A \gg 0$ :  $A \geq \epsilon I$  for some  $\epsilon > 0$ , hence  $A$  is invertible.  
 $\text{dom}(A)$ : The domain of the (unbounded) operator  $A$ .  
 $\text{range}(A)$ : The range of the operator  $A$ .  
 $\rho(A)$ : The resolvent set of the operator  $A$ .  
 $\mathbb{N}$ :  $\mathbb{N}$  is the set of positive integers.  
 $\mathbb{R}, \mathbb{R}^+, \mathbb{R}^-$ :  $\mathbb{R} := (-\infty, \infty)$ ,  $\mathbb{R}^+ := [0, \infty)$ , and  $\mathbb{R}^- := (-\infty, 0]$ .  
 $L^p(J; U)$ : The set of  $U$ -valued  $L^p$ -functions on the interval  $J$ .  
 $L^p_c(J; U)$ : Functions in  $L^p(J; U)$  whose support is bounded to the left.  
 $L^p(J; U; \omega)$ : The set of functions  $u$  for which  $(t \mapsto e^{-\omega t}u(t)) \in L^p(J; U)$ .  
 $C(J; U)$ : This is the set of  $U$ -valued continuous functions on  $J$ .  
 $C_c(J; U)$ : Functions in  $C(J; U)$  whose support is bounded to the left.  
 $C_{c*}(J; Y)$ : Functions in  $C(J; Y)$  whose support is bounded to the right.  
 $C_0(J; U)$ : Functions in  $C(J; U)$  vanishing at the left end-point of  $J$  (finite, or  $-\infty$ ).  
 $BC(J; U)$ : The set of  $U$ -valued bounded continuous functions on  $J$ .  
 $BC_0(J; U)$ : Functions in  $BC(J; U)$  vanishing at infinity. Here  $J = \mathbb{R}^-$  or  $J = \mathbb{R}^+$ .  
 $BC_0(J; U; \omega)$ : The set of functions  $u$  for which  $(t \mapsto e^{-\omega t}u(t)) \in BC_0(J; U)$ .  
 $H^\infty(U; Y; \omega)$ : The set of  $\mathcal{L}(U; Y)$ -valued  $H^\infty$  functions over the half-plane  $\Re z > \omega$ .  
 $\langle \cdot, \cdot \rangle_H$ : The inner product in the Hilbert space  $H$ .  
 $\tau(t)$ : The bilateral time shift operator  $\tau(t)u(s) := u(t + s)$  (this is a left-shift when  $t > 0$  and a right-shift when  $t < 0$ ).  
 $\pi_J$ :  $(\pi_J u)(s) = u(s)$  if  $s \in J$  and  $(\pi_J u)(s) = 0$  if  $s \notin J$ . Here  $J \subset \mathbb{R}$ .  
 $\pi_+, \pi_-$ :  $\pi_+ := \pi_{\mathbb{R}^+}$  and  $\pi_- := \pi_{\mathbb{R}^-}$ .

We extend a  $L^2$ -function  $u$  defined on a subinterval  $J$  of  $\mathbb{R}$  to the whole real line by requiring  $u$  to be zero outside of  $J$ , and we denote the extended function by  $\pi_J u$ . We use the same symbol  $\pi_J$  both for the embedding operator  $L^2(J) \rightarrow L^2(\mathbb{R})$  and for the corresponding projection operator  $L^2(\mathbb{R}) \rightarrow L^2(J)$ . With this interpretation,  $\pi_J L^2(\mathbb{R}; U) = L^2(J; U) \subset L^2(\mathbb{R}; U)$  for each interval  $J \subset \mathbb{R}$ .

**Acknowledgment.** In the fall of 1997 we had a stimulating discussion with Prof. Roberto Triggiani which enabled us to sharpen Corollary 1.

## 2 The Main Result

Let us begin by describing our assumptions on the operator  $A$  in (1). We let  $H$  be a Hilbert space, and let  $A$  generate a  $C_0$  semigroup (i.e., a strongly continuous semigroup)  $\mathcal{A}$  in  $H$ . We denote the domain of  $A$  by  $H_1$ . Choose an arbitrary number  $\alpha$  from the resolvent set of  $A$ . Then  $H_1 = (\alpha I - A)^{-1}H$ , and we can choose the norm in  $H_1$  to be  $\|x\|_{H_1} = \|(\alpha I - A)x\|_H$ . Let  $H_{-1}$  be the completion of  $H$  under the norm  $\|(\alpha I - A)^{-1}x\|_H$ . Then  $H_1 \subset H \subset H_{-1}$  with dense and continuous embeddings,  $(\alpha I - A)$  is an isomorphism of  $H_1$  onto  $H$ , and  $(\alpha I - A)$  extends to an isomorphism of  $H$  onto  $H_{-1}$ . We repeat the same construction with  $\mathcal{A}$  and  $A$  replaced by their adjoints  $\mathcal{A}^*$  and  $A^*$  to get two more Hilbert spaces  $H_1^* = \text{dom}(A^*)$  and  $H_{-1}^*$ , with  $H_1^* \subset H \subset H_{-1}^*$ . It is possible to identify  $H_{-1}^*$  with the dual of  $H_1$  and  $H_{-1}$  with the dual of  $H_1^*$  if we use  $H$  as the pivot space.

In addition to these spaces we shall need two more reflexive<sup>4</sup> Banach spaces  $W$  and  $V$ , satisfying<sup>5</sup>

$$H_1 \subset W \subset H \subset V \subset H_{-1} \text{ (continuous dense embeddings)}. \quad (3)$$

Thus, with  $H$  as pivot space,

$$H_1^* \subset V^* \subset H \subset W^* \subset H_{-1}^* \text{ (continuous dense embeddings)}. \quad (4)$$

We assume further that

$$\mathcal{A}(t)W \subset W, \ t \geq 0, \text{ and } \mathcal{A} \text{ is strongly continuous in } W, \quad (5)$$

$$\mathcal{A}(t)V \subset V, \ t \geq 0, \text{ and } \mathcal{A} \text{ is strongly continuous in } V. \quad (6)$$

Thus,  $\mathcal{A}$  is a  $C_0$  semigroup both in  $W$  and in  $V$ , in addition to being a  $C_0$  semigroup in  $H$ . We use the same letter  $A$  to represent any one of the generators of these semigroups.

Before proceeding, let us remark that the spaces  $W$  and  $V$  did not appear in our earlier work Staffans [1997abc] (although we there construct a space

---

<sup>4</sup>In some hyperbolic cases it would be desirable to remove the reflexivity assumptions on  $W$  and  $V$ , and the assumptions about the embeddings being dense.

<sup>5</sup>To fix the ideas, let us remark that we in the applications to a system based on a parabolic semigroup take  $W = (\alpha I - A)^{-\gamma_1}H$  and  $V = (\alpha I - A)^{\gamma_2}H$ , where  $\alpha \in \rho(A)$ ,  $0 \leq \gamma_1 < \frac{1}{2}$ , and  $0 \leq \gamma_2 < \frac{1}{2}$ . In particular, note that the total degree of unboundedness  $\gamma = \gamma_1 + \gamma_2$  is then less than one.

$W_B \subset W$  that in some respect plays the role of the present space  $W$ , and another space  $V_{(C,K)}^* \subset V^*$  that in some respect plays the role of  $V^*$ .<sup>6</sup> On the other hand, in much of the earlier literature on the parabolic equation (see e.g., Lasiecka and Triggiani [1991b]) the space  $H$  is absent, and  $W$  is used as the basic (pivot) space. This changes of pivot space changes the Riccati operator and all dual operators in a straightforward way; see Section 5. As we remarked earlier, the use of the space  $H$  could be avoided here too, at the expense of loosening the connection to Staffans [1997abc].

To this setting we add a control operator  $B$ , an observation operator  $C$ , and a feed-forward operator  $D$  with the following continuity properties:

$$B \in \mathcal{L}(U; V), \quad C \in \mathcal{L}(W; Y), \quad D \in \mathcal{L}(U; Y), \quad (7)$$

where  $U$  (the control space) and  $Y$  (the output space) are two more Hilbert spaces, with  $U$  separable. We furthermore define

$$(Lu)(t) := \int_0^t \mathcal{A}(t-s)Bu(s) ds, \quad u \in L_{\text{loc}}^1(\mathbf{R}^+; U), \quad t \in \mathbf{R}^+, \quad (8)$$

$$\mathcal{B}v := \int_{-\infty}^0 \mathcal{A}(-s)Bv(s) ds, \quad v \in L_c^1(\mathbf{R}^-; U), \quad (9)$$

$$(\mathcal{C}x_0)(t) := C\mathcal{A}(t)x_0, \quad x_0 \in W, \quad t \in \mathbf{R}^+. \quad (10)$$

Then, because of (5) and (7),

$$L: L_{\text{loc}}^1(\mathbf{R}^+; U) \rightarrow C_0(\mathbf{R}^+; V), \quad (11)$$

$$\mathcal{B}: L_c^1(\mathbf{R}^-; U) \rightarrow V, \quad (12)$$

$$\mathcal{C}: W \rightarrow C(\mathbf{R}^+; Y), \quad (13)$$

and

$$(Lu)(t) = \mathcal{B}\tau(t)\pi_+u, \quad (14)$$

where  $\tau$  is the bilateral time shift operator  $\tau(t)u(s) = u(s+t)$  and  $\pi_+$  is the cutoff operator  $(\pi_+u)(s) = u(s)$  for  $s > 0$ ,  $(\pi_+u)(s) = 0$  for  $s < 0$ .

We impose the following additional ‘‘admissibility’’ assumption on  $\mathcal{B}$  and  $\mathcal{C}$ :

$$\mathcal{B}: L_c^2(\mathbf{R}^-; U) \rightarrow H, \quad (15)$$

$$\mathcal{B}: C_c(\mathbf{R}^-; U) \rightarrow W, \quad (16)$$

$$\mathcal{C}: H \rightarrow L_{\text{loc}}^2(\mathbf{R}^+; Y). \quad (17)$$

---

<sup>6</sup>In those works the letter  $W$  was used as a synonym for  $H_1$ , and the letter  $V$  was used as a synonym for  $H_{-1}$ .

The precise interpretation of (17) is that for each  $T > 0$  there should exist a constant  $C_T$  such that  $\|\mathcal{C}x_0\|_{L^2([0,T];Y)} \leq C_T\|x_0\|_H$  for all  $x_0 \in W$ , and hence,  $\mathcal{C}$  has a unique extension to a continuous operator  $H \rightarrow L^2_{\text{loc}}(\mathbf{R}^+; Y)$  (that we still denote by the same letter  $\mathcal{C}$ ). As is well-known (and easy to see), the two conditions on  $\mathcal{B}$  can be replaced by the following equivalent conditions on  $L$ :

$$L: L^2_{\text{loc}}(\mathbf{R}^+; U) \rightarrow C_0(\mathbf{R}^+; H), \quad (18)$$

$$L: C_0(\mathbf{R}^+; U) \rightarrow C_0(\mathbf{R}^+; W). \quad (19)$$

The assumptions (7) and (16) make it possible to define the (time-invariant) input/output map  $\mathcal{D}$  of the system (1) as follows:

$$\begin{aligned} (\mathcal{D}u)(t) &:= C\mathcal{B}\tau(t)u + Du(t) \\ &= C \int_{-\infty}^t \mathcal{A}(t-s)Bu(s) ds + Du(t), \\ u &\in C_c(\mathbf{R}; U), \quad t \in \mathbf{R}. \end{aligned} \quad (20)$$

If  $u$  is supported on  $\mathbf{R}^+$ , then we can alternatively write this as

$$\begin{aligned} (\mathcal{D}u)(t) &= C(Lu)(t) + Du(t) = C \int_0^t \mathcal{A}(t-s)Bu(s) ds + Du(t), \\ u &\in C_0(\mathbf{R}^+; U), \quad t \in \mathbf{R}^+. \end{aligned} \quad (21)$$

Then it follows from (7) and (16) that

$$\mathcal{D}: C_c(\mathbf{R}; U) \rightarrow C_c(\mathbf{R}; Y). \quad (22)$$

As we shall see in a moment, this implies that  $\mathcal{D}$  can be extended to a continuous map

$$\mathcal{D}: L^2_c(\mathbf{R}; U) \rightarrow L^2_c(\mathbf{R}; Y), \quad (23)$$

that we still denote by  $\mathcal{D}$ .

Appealing to the standard variation of constants formula in the case of bounded control and observation operators and to the admissibility conditions (17) and (23) we define the solution  $x$  and the output  $y$  of (1) to be given by

$$\begin{aligned} x(t) &:= \mathcal{A}(t)x_0 + (Lu)(t) = \mathcal{A}(t)x_0 + \mathcal{B}\tau(t)\pi_+u, \\ y &:= \mathcal{C}x_0 + \mathcal{D}\pi_+u, \\ x_0 &\in H, \quad u \in L^2_{\text{loc}}(\mathbf{R}^+; U), \quad t \in \mathbf{R}^+. \end{aligned} \quad (24)$$

Then, by (8)–(23),

$$x_0 \in W \text{ and } u \in C_0(\mathbf{R}^+; U) \implies x \in C(\mathbf{R}^+; W) \text{ and } y \in C(\mathbf{R}^+; Y), \quad (25)$$

$$x_0 \in H \text{ and } u \in L^2_{\text{loc}}(\mathbf{R}^+; U) \implies x \in C(\mathbf{R}^+; H) \text{ and } y \in L^2_{\text{loc}}(\mathbf{R}^+; Y), \quad (26)$$

and if  $\mathcal{C}: V \rightarrow L^1_{\text{loc}}(\mathbf{R}^+; Y)$ , then

$$x_0 \in V \text{ and } u \in L^1_{\text{loc}}(\mathbf{R}^+; U) \implies x \in C(\mathbf{R}^+; V) \text{ and } y \in L^1_{\text{loc}}(\mathbf{R}^+; Y).$$

Let us summarize the various assumptions and claims that we have made so far into the following lemma:

**Lemma 1** *Let (3), (5), (7), (15), (16), and (17) hold in the sense explained above. Then so do (4), (11), (12), (13), (14), (18), (19), (22), (23), (25), and (26).*

In addition to the “standard” results listed above, we claim that the following more specific results are valid:

**Lemma 2** *Let conditions (3), (5), (7), (15), (16), and (17) hold. Then  $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$  is a regular well-posed linear system on  $(U, H, Y)$  in the sense of [Staffans 1997b, Definition 2.5], with generating operators  $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ , state  $x$ , and output  $y$ . Moreover,*

$$L: C(\mathbf{R}^+; U) \rightarrow C_0(\mathbf{R}^+; W), \quad (27)$$

$$\mathcal{D}\pi_+: C(\mathbf{R}^+; U) \rightarrow C(\mathbf{R}^+; Y), \quad (28)$$

$$(\alpha I - A)^{-1}BU \subset W, \quad \alpha \in \rho(A), \quad (29)$$

$$\lim_{\alpha \rightarrow \infty} \|(\alpha I - A)^{-1}Bu\|_W = 0, \quad u \in U, \quad (30)$$

the claim (25) can be strengthened to

$$x_0 \in W \text{ and } u \in C(\mathbf{R}^+; U) \implies x \in C(\mathbf{R}^+; W) \text{ and } y \in C(\mathbf{R}^+; Y) \quad (31)$$

(i.e., the assumption  $u(0) = 0$  has been removed), and for each  $x_0 \in W$  and  $u \in C(\mathbf{R}^+; U)$ , the state  $x$  and output  $y$  defined in (24) satisfy

$$y(t) = Cx(t) + Du(t), \quad t \in \mathbf{R}^+. \quad (32)$$



These lemmas are proved in Section 3, together with analogous lemmas for the dual system and for the stable case.

Let us remark that we have still not made any significant use of the space  $V$ , and we can without loss of generality take  $V = H_{-1}$  in Lemmas 1 and 2. The space  $V$ , or rather its dual space  $V^*$ , becomes important when we get to the dual system (see Section 3).

Thanks to the preceding lemma we can apply the technique used in Staffans [1997abc] to study the quadratic cost minimization problem presented in the introduction, where we minimize the cost  $Q(x_0, u)$  in (2). A central role in this theory is played by the “optimal cost operator”  $\Pi$ , also called the *Riccati operator*:

**Definition 1** *If, for each  $x_0 \in H$ , the minimum of  $Q(x_0, u)$  is achieved for some  $u \in L^2(\mathbf{R}^+; U)$ , and if there exists an operator  $\Pi = \Pi^* \in \mathcal{L}(H)$  such that the optimal cost is given by*

$$\langle x_0, \Pi x_0 \rangle_H := \min_{u \in L^2(\mathbf{R}^+; U)} Q(x_0, u),$$

*then  $\Pi$  is called the Riccati operator of (1) with cost operator  $J$ .*

For simplicity, we most of the time suppose that the system is stable. In this connection stability means that for each  $x \in H$  and each  $u \in L^2(\mathbf{R}^+; U)$ , the output  $y$  belongs to  $L^2(\mathbf{R}^+; Y)$ . To achieve this we impose the following additional *stability* assumptions on  $\mathcal{C}$  and  $\mathcal{D}$ :

$$\mathcal{C}: H \rightarrow L^2(\mathbf{R}^+; Y), \tag{33}$$

$$\mathcal{D}: BC_0(\mathbf{R}; U) \rightarrow BC_0(\mathbf{R}; Y). \tag{34}$$

As we shall see, this implies that

$$\mathcal{D}: L^2(\mathbf{R}; U) \rightarrow L^2(\mathbf{R}; Y). \tag{35}$$

We remark that these conditions are implied by the earlier admissibility assumptions whenever  $\mathcal{A}$  is exponentially stable in both  $H$  and  $W$  (but they can be satisfied even when  $\mathcal{A}$  is unstable).

In order for the cost function  $Q$  to be bounded from below we impose the following standard *coercivity* condition:

**Definition 2** Let  $J = J^* \in \mathcal{L}(Y)$ . The operator  $\mathcal{D}\pi_+$  defined in (21) is  $J$ -coercive iff  $\pi_+ \mathcal{D}^* J \mathcal{D} \pi_+ \gg 0$  on  $L^2(\mathbf{R}^+; U)$ , i.e.,  $\langle \mathcal{D}\pi_+ u, J \mathcal{D}\pi_+ u \rangle_{L^2(\mathbf{R}^+; Y)} \geq \epsilon \|u\|_{L^2(\mathbf{R}^+; U)}^2$  for all  $u \in L^2(\mathbf{R}^+; U)$  and some  $\epsilon > 0$ . The system (1) is  $J$ -coercive iff its input/output map  $\mathcal{D}$  is  $J$ -coercive.

We recall the following basic result from Staffans [1997c]:

**Lemma 3** ([Staffans 1997a, Lemma 13 and Theorem 27] and [Staffans 1997c, Lemma 2.5]) Let  $J = J^* \in \mathcal{L}(Y)$  and let the system (1) be stable and  $J$ -coercive in the sense explained above. Then, for each  $x_0 \in H$ , there is a unique control  $u^{\text{opt}}(x_0) \in L^2(\mathbf{R}^+; U)$  that minimizes the cost function  $Q(x_0, u)$  in (2). This control is given by

$$u^{\text{opt}}(x_0) = -(\pi_+ \mathcal{D}^* J \mathcal{D} \pi_+)^{-1} \pi_+ \mathcal{D}^* J \mathcal{C} x_0 = -\mathcal{X}^{-1} S^{-1} \pi_+ \mathcal{N}^* J \mathcal{C} x_0,$$

where  $\mathcal{N} \mathcal{X}$  is an arbitrary  $(J, S)$ -inner-outer factorization of  $\mathcal{D}$  (cf. [Staffans 1997c, Lemma 2.4]). The corresponding state  $x^{\text{opt}}(x_0)$ , output  $y^{\text{opt}}(x_0)$ , and the minimum  $Q(x_0, u^{\text{opt}}(x_0))$  of the cost function are given by

$$\begin{aligned} x^{\text{opt}}(x_0) &= \mathcal{A} x_0 - \mathcal{B} \tau \pi_+ (\pi_+ \mathcal{D}^* J \mathcal{D} \pi_+)^{-1} \pi_+ \mathcal{D}^* J \mathcal{C} x_0 \\ &= \mathcal{A} x_0 - \mathcal{B} \mathcal{X}^{-1} \tau S^{-1} \pi_+ \mathcal{N}^* J \mathcal{C} x_0, \\ y^{\text{opt}}(x_0) &= (I - P) \mathcal{C} x_0, \\ Q(x_0, u^{\text{opt}}(x_0)) &= \langle x_0, \mathcal{C}^* J (I - P) \mathcal{C} x_0 \rangle_H, \end{aligned}$$

where

$$P := \mathcal{D} \pi_+ (\pi_+ \mathcal{D}^* J \mathcal{D} \pi_+)^{-1} \pi_+ \mathcal{D}^* J = I - \mathcal{N} S^{-1} \pi_+ \mathcal{N}^* J$$

is the projection onto the range of  $\mathcal{D}\pi_+$  along the null space of  $\pi_+ \mathcal{D}^* J$ . In particular,  $\Psi$  has a Riccati operator, namely

$$\Pi = \mathcal{C}^* J (I - P) \mathcal{C}$$

and  $y^{\text{opt}}(x_0)$  belongs to the null space of the projection  $P$ , i.e.,

$$\pi_+ \mathcal{D}^* J y^{\text{opt}}(x_0) = \pi_+ \mathcal{D}^* J (\mathcal{C} x_0 + \mathcal{D} \pi_+ u^{\text{opt}}(x_0)) = 0.$$

As shown in [Staffans 1997c, Theorem 2.6 and Remark 2.7], for a  $J$ -coercive system, there is an one-to-one connection between the set of all possible feedback solutions to the quadratic cost minimization problem and the set of all possible  $(J, S)$ -inner-outer factorizations of the input/output

map  $\mathcal{D}$  (or equivalently, of the transfer function of the system). Here  $S \in \mathcal{L}(U)$  is strictly positive. We shall not need that result here, so we refer the reader to [Staffans 1997c, Theorem 2.6] for details.

Under a further regularity assumption on the system, and in particular, on the outer factor in the inner-outer factorization mentioned in Lemma 3, it is possible to show that Riccati operator is the solution of a (possibly nonstandard) Riccati equation; see Staffans [1997c]. The converse is also true, if we are able to find a sufficiently regular stabilizing solution  $\Pi$  to this Riccati equation, then we have in fact found the Riccati operator of Definition 1, and this operator can then be used to construct both the feedback solution of the quadratic cost minimization problem and the corresponding  $(J, S)$ -inner-outer factorization of  $\mathcal{D}$ ; see Mikkola [1997].

The cited necessary and sufficient results for the existence of a stabilizing solution to the Riccati equation appear to give a “definite answer” to the general quadratic cost minimization problem, but in not quite the case. The remaining problem is related to the extra regularity assumptions used in these works. In this connection “regularity” means the following (cf. [Weiss 1994a, Theorem 5.8]):

**Definition 3** (i) *An input/output map  $\mathcal{D}: L_c^2(\mathbf{R}; U) \rightarrow L_c^2(\mathbf{R}; Y)$  is regular iff the strong limit  $Du_0 := \lim_{\lambda \rightarrow +\infty} \widehat{\mathcal{D}}(\lambda)u_0$  exists for every  $u_0 \in U$ ; here  $\lambda$  tends to  $+\infty$  along the positive real axis and  $\widehat{\mathcal{D}}$  is the transfer function (the distribution Laplace transform) of  $\mathcal{D}$ .*

(ii) *The operator  $D: U \rightarrow Y$  defined above is called the feed-through (or feed-forward) operator of  $\mathcal{D}$ .*

(iii) *A regular input/output map  $\mathcal{D}$  is strictly proper iff its feed-through operator vanishes.*

(iv) *We say that  $\mathcal{D}$  is regular together with its adjoint iff, in addition to (i), the strong limit  $\lim_{\lambda \rightarrow +\infty} \widehat{\mathcal{D}}^*(\lambda)y_0$  exists for every  $y_0 \in Y$ . (This limit is equal to  $D^*y_0$  whenever it exists.)*

(v) *The system  $\Psi = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$  is regular [together with its adjoint] iff its input/output map  $\mathcal{D}$  is regular [together with its adjoint].*

In Staffans [1997c] it was assumed throughout that all possible input/output maps that appear in the development of the theory are regular together with

their adjoints. In particular, this applies both to the original input/output map  $\mathcal{D}$ , and to the outer factor  $\mathcal{X}$  in the  $(J, S)$ -inner-outer factorization of  $\mathcal{D}$  mentioned in Lemma 3. In Weiss and Weiss [1997] and Mikkola [1997] this condition is weakened to weak regularity<sup>7</sup>, but even so we are left with an extra regularity assumption that can be difficult to verify. This is partly due to the very general setting used in Staffans [1997c], Weiss and Weiss [1997], Mikkola [1997]. In particular, that setting makes no use of the spaces  $W$  and  $V$  that we have introduced above; instead appropriate replacements for these spaces are constructed from the regularity assumption. Here we shall complement the basic theory presented in these works by adding assumptions on the system related directly to the spaces  $W$  and  $V$ , and use these extra assumptions to guarantee that the problem has enough regularity for the theory developed in Staffans [1997c], Weiss and Weiss [1997], Mikkola [1997] to apply. These assumptions (including all our previous assumptions related to  $W$  and  $V$ ) are analogous to those found in, e.g., Lasiecka and Triggiani [1991b].

Actually, we have already introduced most of the needed assumptions, and only make the following additions. We add the following admissibility and stability assumptions on the observability map  $\mathcal{C}$  and its adjoint:

$$\mathcal{C}: W \rightarrow BC_0(\mathbf{R}^+; Y), \quad (36)$$

$$\mathcal{C}^*: BC_0(\mathbf{R}^+; Y) \rightarrow V^*. \quad (37)$$

Observe that the former is a consequence of (7) if  $\mathcal{A}$  is strongly stable, and that the latter is implied by<sup>8</sup>

$$\mathcal{C}: V \rightarrow L^1(\mathbf{R}^+; Y); \quad (38)$$

a condition that we shall *not* require to be true.

We shall also need one extra condition on the input/output map  $\mathcal{D}$ , which is maybe the most restricting one in the whole setup. We assume (as is always done in the finite-dimensional case) that

$$D^*JD \gg 0, \quad (39)$$

---

<sup>7</sup>That is, the limits need only exist in the weak sense. In addition, these papers assume that the feed-through operator of the outer factor is invertible, a condition that is redundant in Staffans [1997c].

<sup>8</sup>This condition should be interpreted in the same way as condition (17). Note that condition (38) combined with (7) furthermore implies that  $\mathcal{D}: L^1(\mathbf{R}^+; U) \rightarrow L^1(\mathbf{R}^+; Y)$ .

define the operator  $\mathcal{E}$  by

$$\mathcal{E}u := u - (D^*JD)^{-1}\pi_+\mathcal{D}^*J\mathcal{D}\pi_+u, \quad u \in L^2(\mathbf{R}^+;U), \quad (40)$$

or equivalently,

$$\pi_+\mathcal{D}^*J\mathcal{D}\pi_+u = D^*JD(u - \mathcal{E}u),$$

and require that  $\mathcal{E}$  is “smoothing” in the following sense:

$$\mathcal{E}^n : L^2(\mathbf{R}^+;U) \rightarrow BC_0(\mathbf{R}^+;U) \text{ for som } n \in \mathbf{N}. \quad (41)$$

Let us summarize all our admissibility and stability assumptions into the following main hypothesis:

**Hypothesis 1** *Conditions (3), (5), (6), (7), (15), (16), (33), (34), (36), (37), (39), and (41) hold.*

The following is our main result:

**Theorem 1** *Let Hypothesis 1 hold, and let the system (1) be  $J$ -coercive. Then both the open and the closed loop systems in [Staffans 1997c, Theorem 2.6] are regular together with their adjoints, the feed-through operator of the outer factor  $\mathcal{X}$  in an arbitrary  $(J, S)$ -inner-outer factorization of  $\mathcal{D}$  is invertible, and the Riccati equation theory in [Staffans 1997c, Theorem 6.1] and Mikkola [1997] applies. Moreover, the correction term  $\lim_{\alpha \rightarrow \infty} \overline{B}^*\Pi(\alpha I - A)^{-1}Bu_0$  in [Staffans 1997c, Corollary 7.2] is zero, hence  $S = D^*JD$  if we normalize the feed-through operator of the outer factor  $\mathcal{X}$  to be the identity. In particular, this means that the following claims are true. The Riccati operator  $\Pi$  and the feedback operator  $K$  satisfy*

$$\Pi \in \mathcal{L}(W;V^*), \quad K \in \mathcal{L}(W;U), \quad (42)$$

$$K = -(D^*JD)^{-1}(B^*\Pi + D^*JC), \quad (43)$$

$$A^*\Pi + \Pi A = -C^*JC + K^*D^*JDK, \quad (44)$$

where (43) is valid in  $\mathcal{L}(W;U)$  and (44) is valid in  $\mathcal{L}(W;W^* + (\alpha I - A)^*V^*) \subset \mathcal{L}(W;H_{-1}^*)$ ; here  $\alpha \in \rho(A)$ . The normalized spectral factor  $\mathcal{X}$  is given by the “classical” formula

$$(\mathcal{X}\pi_+u)(t) = u(t) - K \int_0^t \mathcal{A}(t-s)Bu(s) ds, \quad u \in C(\mathbf{R}^+;U), \quad (45)$$

and it maps the function spaces  $C(\mathbf{R}^+; U)$ ,  $BC_0(\mathbf{R}^+; U)$ ,  $L^2_{\text{loc}}(\mathbf{R}^+; U)$ , and  $L^2(\mathbf{R}^+; U)$  continuously into themselves. The Laplace transform of this spectral factor and its inverse are given by

$$\widehat{\mathcal{X}}(s) = I - K(sI - A)^{-1}B, \quad \Re s > 0, \quad (46)$$

$$\widehat{\mathcal{X}}^{-1}(s) = I + K(sI - A_{\circ})^{-1}B, \quad \Re s > 0, \quad (47)$$

where  $A_{\circ} = A + BK$  is the generator of the closed loop semigroup  $\mathcal{A}_{\circ}$ . This semigroup is strongly continuous both in  $H$  and in  $W$ , and  $(\alpha I - A_{\circ})^{-1}H \subset W$  and  $(\alpha I - A_{\circ})^{-1}BU \subset W$  for all  $\alpha \in \rho(\mathcal{A}_{\circ})$ . The Riccati operator  $\Pi$  is the unique self-adjoint stabilizing solution (in the sense of Mikkola [1997]) of the Riccati equation

$$\begin{aligned} & \langle Ax_0, \Pi x_1 \rangle_H + \langle x_0, \Pi Ax_1 \rangle_H + \langle Cx_0, JCx_1 \rangle_Y \\ & = \langle (B^*\Pi + D^*JC)x_0, (D^*JD)^{-1}(B^*\Pi + D^*JC)x_1 \rangle_U, \quad (48) \\ & x_0, x_1 \in H_1, \end{aligned}$$

and this equation is, in fact, valid for all  $x_0, x_1 \in W \cap (\alpha I - A)^{-1}V$ , where  $\alpha \in \rho(A)$ .

In the parabolic examples given in Section 6 we shall throughout apply this theorem in the following form:

**Corollary 1** *Suppose that the semigroup  $\mathcal{A}$  generated by  $A$  is analytic and exponentially stable in  $H$ , that*

$$W = (-A)^{-\gamma_1}H, \quad V = (-A)^{\gamma_2}H \quad (49)$$

for some  $0 \leq \gamma_1 < \frac{1}{2}$  and  $0 \leq \gamma_2 < \frac{1}{2}$ , that (7) holds, and that the system is  $J$ -coercive. Then Hypothesis 1 holds and Theorem 1 applies. In this case the conclusion of Theorem 1 can be strengthened as follows. For all  $\epsilon > 0$ ,  $\Pi \in \mathcal{L}(W; (-A^*)^{-1+\gamma_1+\epsilon}H)$ , equation (44) is valid in  $\mathcal{L}(W; (-A^*)^{\gamma_1+\epsilon}H)$ , and equation (48) is valid for all  $x_0, x_1 \in (-A)^{-\gamma_1-\epsilon}H$ .<sup>9</sup> The closed loop semigroup  $\mathcal{A}_{\circ}$  is exponentially stable and analytic in  $(-A)^{\gamma}H$  for  $\gamma_2 - 1 \leq \gamma \leq 0$ , and it can be extended to an exponentially stable and analytic semigroup in  $(-A)^{\gamma}H$  for  $0 \leq \gamma \leq 1 - \gamma_1$ . In particular, it is analytic in  $H$  and  $W$ , and

---

<sup>9</sup>It is possible to get rid of the extra  $\epsilon$  in the case where  $A$  is normal; see Lasiecka and Triggiani [1991b].

it can be extended to an analytic semigroup in  $V$ . All possible input/output maps (open or closed loop, adjoint or not) are convolution operators with kernels that are analytic on  $(0, \infty)$  and belong to  $L^p(\mathbf{R}^+)$  for all  $1 \leq p < 1/(\gamma_1 + \gamma_2)$ .

### 3 Proofs.

*Proof of Lemma 1.* All of the claims in Lemma 1, except for (23), are either obvious or well-known from standard semigroup theory, so we leave their verification to the reader.

To prove (23) we first remark that  $\mathcal{D}$  can be (uniquely) extended to a bounded linear time-invariant operator

$$\mathcal{D}: BC_0(\mathbf{R}; U; \omega) \rightarrow BC_0(\mathbf{R}; Y; \omega), \quad (50)$$

where  $\omega$  is an arbitrary number bigger than the growth rate of  $\mathcal{A}$  in  $W$ . We have not been able to find this particular result in the literature, but its proof is essentially the same as the proofs given for the corresponding  $L^p$ -statements; see, e.g., [Salamon 1989, Lemma 2.1], [Weiss 1989a, Proposition 2.5], or [Weiss 1994a, Proposition 4.1]. We therefore leave the proof of (50) to the reader.

We claim that (50) implies that  $\mathcal{D}$  satisfies

$$\mathcal{D}: L^2(\mathbf{R}; U; \omega) \rightarrow L^2(\mathbf{R}; Y; \omega). \quad (51)$$

Clearly, if this is true, then (23) follows. By, e.g., [Staffans 1997b, Lemma 2.9], to prove this it suffices to show that the Laplace transform  $\widehat{\mathcal{D}}$  of  $\mathcal{D}$  satisfies

$$\widehat{\mathcal{D}} \in H^\infty(U; Y; \omega), \quad (52)$$

and this is done as follows. Take some arbitrary  $s \in \mathbf{C}$  with  $\Re s > \omega$  and  $u_0 \in U$ . Define  $u(t) = e^{st}u_0$ ,  $t \in \mathbf{R}$ . Then  $(\mathcal{D}u)(0) = \widehat{\mathcal{D}}(s)u_0$ . Thus, using also (50), we find that there is some constant  $M > 0$  such that

$$\begin{aligned} \|\widehat{\mathcal{D}}(s)u_0\|_Y &= \|(\mathcal{D}u)(0)\|_Y \leq \|\mathcal{D}u\|_{BC_0(\mathbf{R}^-; Y; \omega)} \\ &\leq M\|u\|_{BC_0(\mathbf{R}^-; U; \omega)} = M\|u_0\|_U. \end{aligned}$$

This shows that (52) holds.  $\square$

*Proof of Lemma 2.* As can be seen easily, all the algebraic properties listed in [Staffans 1997b, Definition 2.1](i)–(iv) hold for all  $x \in W$  and  $u \in C_c(\mathbf{R}; U)$  satisfying  $u(0) = 0$ . By density and continuity, the same conditions then hold for all  $x \in H$  and all  $u \in L_c^2(\mathbf{R}; U)$ . However, this implies that the same conditions hold for all  $u \in L^2(\mathbf{R}; U; \omega)$  where  $\omega$  is an arbitrary number bigger than the growth rate of  $\mathcal{A}$  in  $H$ ; see, e.g., [Salamon 1989, Lemma 2.1] or [Weiss 1989a, Proposition 2.5], [Weiss 1989b, Proposition 2.3], and [Weiss 1994a, Proposition 4.1]. Thus, we conclude that  $\begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$  is a well-posed linear system on  $(U, H, Y)$ . Comparing (24) with the corresponding formula in Staffans [1997b] we see that the state of this system (in the initial value setting) is  $x$ , and its output is  $y$ . It is also easy to show that the generators of  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  are  $A$ ,  $B$ , and  $C$ , respectively.

It still remains to prove the claim about the regularity of the system, the claim that the feed-through operator is  $D$ , and the additional claims (27)–(32).

For each  $u \in C^1(\mathbf{R}^+; U)$  and  $\alpha \in \rho(A)$ , we can integrate by parts and divide by  $\alpha I - A$  to get

$$\begin{aligned} (Lu)(t) &= \int_0^t \mathcal{A}(t-s)Bu(s) ds \\ &= (\alpha I - A)^{-1} \left( Bu(t) - \mathcal{A}(t)Bu(0) + \int_0^t \mathcal{A}(t-s) (\alpha Bu(s) - Bu'(s)) ds \right). \end{aligned}$$

In particular, by taking  $t = 1$  and  $u(s) = s^2 u_0$  for some fixed  $u_0 \in U$  we get

$$\begin{aligned} (\alpha I - A)^{-1} Bu_0 &= \int_0^1 \mathcal{A}(1-s)Bs^2 u_0 ds \\ &\quad - (\alpha I - A)^{-1} \int_0^1 \mathcal{A}(1-s)B(\alpha s^2 - 2s)u_0 ds, \quad (53) \end{aligned}$$

which together with (5) and (19) implies (29).

If we instead take  $u(s) \equiv u_0 \in U$  (a constant function), then we get

$$\begin{aligned} (Lu)(t) &= \int_0^t \mathcal{A}(t-s)Bu_0 ds \\ &= \left( I - \mathcal{A}(t) - \alpha \int_0^t \mathcal{A}(s) ds \right) (\alpha I - A)^{-1} Bu_0, \end{aligned}$$

and by (5), this tends to zero in  $W$  as  $t \rightarrow 0$ . This, together with (19) implies (27), which in turn together with (7) implies (28) and (31).



Referring to (53), we observe that both  $\int_0^1 \mathcal{A}(1-s)Bs u_0 ds$  and  $\int_0^1 \mathcal{A}(1-s)Bs^2 u_0 ds$  belong to  $W$ . Therefore, by letting  $\alpha \rightarrow \infty$  in (53) and using the fact that  $\alpha(\alpha I - A)^{-1}w \rightarrow w$  in  $W$  for each  $w \in W$  we get (30). This condition, in turn, implies that the system is regular, and that the feed-through operator of  $\mathcal{D}$  is  $D$ . The claim (32) follows from, e.g., [Weiss 1994a, Remark 6.2].  $\square$

Since we shall need part of the preceding argument later, too, let us separate it into a lemma of its own:

**Lemma 4** *Let  $\Psi = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$  be a well-posed linear system on  $(U, H, Y)$  with generators  $\begin{bmatrix} A & B \\ C & * \end{bmatrix}$ . Suppose that  $H_1 \subset W \subset H$  with continuous and dense injections, that  $C \in \mathcal{L}(W, Y)$  (i.e.,  $C$  has a unique continuous extension to this space), that  $(\alpha I - A)^{-1}BU \subset W$  for all  $\alpha \in \rho(A)$ , and that, for each  $w \in W$ ,  $\alpha(\alpha I - A)^{-1}w \rightarrow w$  in  $W$  as  $\alpha \rightarrow +\infty$  (which is true, in particular, when (5) holds). Then  $\Psi$  is regular,  $(\alpha I - A)^{-1}Bu \rightarrow 0$  in  $W$  for each  $u \in U$  as  $\alpha \rightarrow \infty$ , and the transfer function of  $\Psi$  is given by*

$$\widehat{\mathcal{D}}(s) = C(sI - A)^{-1}B + D$$

for all  $s \in \mathbf{C}$  with sufficiently large real part ( $D$  is defined in Definition 3).

*Proof.* By the resolvent identity, for each  $\alpha, \beta \in \rho(A)$  and each  $u \in U$  (cf. [Salamon 1989, pp. 148–149])

$$\widehat{\mathcal{D}}(\alpha)u = \widehat{\mathcal{D}}(\beta)u + (\beta - \alpha)C(\alpha I - A)^{-1}(\beta I - A)^{-1}Bu.$$

As  $C \in \mathcal{L}(W; Y)$ ,  $(\beta I - A)^{-1}Bu \in W$ , and  $\alpha(\alpha I - A)^{-1}w \rightarrow w$  in  $W$  for all  $w \in W$ , the limit of the right hand side as  $\alpha \rightarrow \infty$  exists in  $Y$  and is equal to  $\widehat{\mathcal{D}}(\beta)u - C(\beta I - A)^{-1}Bu$ . Thus the system is regular, and  $\widehat{\mathcal{D}}(\beta) - C(\beta I - A)^{-1}B = D$ .

To show that  $(\alpha I - A)^{-1}Bu \rightarrow 0$  in  $W$  for all  $u \in U$  we argue in the same way, starting from the identity

$$(\alpha I - A)^{-1}Bu = (\beta I - A)^{-1}Bu + (\beta - \alpha)(\alpha I - A)^{-1}(\beta I - A)^{-1}Bu. \quad \square$$

Lemmas 1 and 2 have the following counterpart for the dual system:

**Lemma 5** *In addition to (3), (5), (6), (7), (15), (16), (17), suppose that*

$$\mathcal{C}^* : C_{c^*}(\mathbf{R}^+; Y) \rightarrow V^*. \quad (54)$$

Then

$$\mathcal{D}^* : C_{c^*}(\mathbf{R}; Y) \rightarrow C_{c^*}(\mathbf{R}; U), \quad (55)$$

$$\mathcal{D}^* : L_{c^*}^2(\mathbf{R}; Y) \rightarrow L_{c^*}^2(\mathbf{R}; U), \quad (56)$$

and  $\mathcal{D}^*$  is regular. Moreover,  $\mathcal{B}^*$ ,  $\mathcal{C}^*$  and  $\mathcal{D}^*$  are given by

$$(\mathcal{B}^* x_0^*)(s) := B^* \mathcal{A}^*(-s) x_0^*, \quad x_0^* \in V^*, \quad s \in \mathbf{R}^-, \quad (57)$$

$$\mathcal{C}^* y^* := \int_0^\infty \mathcal{A}^*(t) C^* y^*(t) dt, \quad y^* \in L_{c^*}^1(\mathbf{R}^+; Y), \quad (58)$$

$$\begin{aligned} (\mathcal{D}^* y^*)(s) &:= B^* \mathcal{C}^* \tau(s) y^* + D^* y^*(s) \\ &= B^* \int_s^\infty \mathcal{A}^*(t-s) C^* y^*(t) dt + D^* y^*(s), \quad (59) \\ &y^* \in C_{c^*}(\mathbf{R}; Y), \quad s \in \mathbf{R}. \end{aligned}$$

*Proof of Lemma 5.* The proofs of these two lemmas are virtually identical to the proofs of Lemmas 1 and 2. For the main part of the proof it suffices to apply exactly the same arguments, with the replacements  $\mathcal{A} \rightarrow \mathcal{A}^*$ ,  $H_1 \rightarrow H_1^*$ ,  $W \rightarrow V^*$ ,  $U \rightarrow Y$ ,  $Y \rightarrow U$ ,  $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \rightarrow \begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix}$ , etc., and to revert the direction of time. In particular, we define  $\mathcal{B}^*$ ,  $\mathcal{C}^*$ , and  $\mathcal{D}^*$  by (57)–(59), and observe that (15) and (17) are equivalent to

$$\begin{aligned} \mathcal{B}^* : H &\rightarrow L_{\text{loc}}^2(\mathbf{R}^-; U), \\ \mathcal{C}^* : L_{c^*}^2(\mathbf{R}^+; Y) &\rightarrow H. \end{aligned}$$

There is only one potential problem with this approach: How do we know that the system that we construct in this way is indeed the adjoint of the earlier constructed system  $\Psi$ ? Fortunately, this follows from the fact that the new system has the same generators  $\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix}$  as the adjoint system  $\Psi^*$ , and every (causal or anti-causal) regular well-posed linear system is uniquely determined by its generators, cf. [Weiss 1994a, Theorem 2.3].  $\square$

In the proof of Theorem 1 we shall also need the following stable version of Lemmas 1–5.

**Lemma 6** *In addition to the assumptions of Lemma 2, let (33), (34), and (36) hold. Then (35) holds, and*

$$x_0 \in W \text{ and } u \in BC_0(\mathbf{R}^+; U) \implies y \in BC_0(\mathbf{R}^+; Y), \quad (60)$$

$$x_0 \in H \text{ and } u \in L^2(\mathbf{R}^+; U) \implies y \in L^2(\mathbf{R}^+; Y). \quad (61)$$

In particular,

$$\mathcal{D}\pi_+ : BC_0(\mathbf{R}^+; U) \rightarrow BC_0(\mathbf{R}^+; Y). \quad (62)$$

If, in addition both the assumption of Lemma 5 and (37) holds, then

$$\mathcal{D}^* : BC_0(\mathbf{R}; Y) \rightarrow BC_0(\mathbf{R}; U). \quad (63)$$

The proof of this lemma is similar to the proofs of Lemmas 1 and 2, and it is left to the reader.

For the convenience of the reader, let us cite the following result from Staffans [1997c], which is a key ingredient in the proof of Theorem 1:

**Theorem 2** ([Staffans 1997c, Theorem 7.1]) *Let Hypothesis 1 hold, and let the system (1) be  $J$ -coercive. Denote the generating operators of  $\Psi$  by the same letters as the corresponding operators [Staffans 1997a, Section 7], and let  $\widehat{\mathcal{D}}$  and  $\widehat{\mathcal{F}}$  be the transfer functions (i.e., distribution Laplace transforms) of  $\mathcal{D}$  and  $\mathcal{F}$  [Staffans 1997b, Lemma 2.9]. Let  $x_0 \in H$  and  $u_0 \in U$  satisfy  $Ax_0 + Bu_0 \in H$ .*

*If  $\alpha \in \mathbf{C}$  has real part bigger than the growth rate of  $\Psi$ , then the vectors  $y_0 \in Y$  and  $w_0 \in U$  defined by*

$$y_0 = C(\alpha I - A)^{-1}(\alpha x_0 - Ax_0 - Bu_0) + \widehat{\mathcal{D}}(\alpha)u_0, \quad (64)$$

$$w_0 = -K(\alpha I - A)^{-1}(\alpha x_0 - Ax_0 - Bu_0) + (I - \widehat{\mathcal{F}}(\alpha))u_0 \quad (65)$$

*are independent of  $\alpha$ . Moreover,*

$$A^*\Pi x_0 + C^*Jy_0 + K^*Sw_0 = -\Pi(Ax_0 + Bu_0) \in H, \quad (66)$$

*and, for all  $\beta \in \mathbf{C}$  with real part bigger than the growth rate of  $\Psi$ ,*

$$\begin{aligned} & (I - \widehat{\mathcal{F}}(\beta))^*Sw_0 \\ &= B^*(\overline{\beta}I - A^*)^{-1}\Pi(\overline{\beta}x_0 + Ax_0 + Bu_0) + (\widehat{\mathcal{D}}(\beta))^*Jy_0. \end{aligned} \quad (67)$$

We are finally ready to prove Theorem 1.

*Proof of Theorem 1.* The idea behind this proof is to show that all possible signals appearing in the system are continuous functions of time whenever  $x_0 \in W$ .<sup>10</sup> To do this it suffices to show that

$$u^{\text{opt}} \in BC_0(\mathbf{R}^+; U), \quad (68)$$

---

<sup>10</sup>Here we follow the route outlined in Lasiecka and Triggiani [1991b].

because if this is true, then it follows from (31) and (60) that

$$x^{\text{opt}} \in C(\mathbf{R}^+; W), \quad y^{\text{opt}} \in BC_0(\mathbf{R}^+; Y). \quad (69)$$

Thus, let us prove (68). For this we use the explicit expression for  $u^{\text{opt}}$  given in Lemma 3, writing in the form (cf. (40))

$$u^{\text{opt}} = \mathcal{E}u^{\text{opt}} - (D^*JD)^{-1}\pi_+\mathcal{D}^*J\mathcal{C}x_0.$$

By (36) and (63),  $\pi_+\mathcal{D}^*J\mathcal{C}x_0 \in BC_0(\mathbf{R}^+; U)$ . Furthermore, we know that  $u^{\text{opt}} \in L^2(\mathbf{R}^+; U)$ . By iterating the preceding equation  $n - 1$  times we get

$$u^{\text{opt}} = \mathcal{E}^n u^{\text{opt}} - \sum_{k=0}^{n-1} \mathcal{E}^k (D^*JD)^{-1}\pi_+\mathcal{D}^*J\mathcal{C}x_0.$$

The first term on the right-hand side belongs to  $BC_0(\mathbf{R}^+; U)$  because of (41), and other terms belong to  $BC_0(\mathbf{R}^+; U)$  because of (40), (62), and (63). Thus, (68) and (69) hold.

Now let us turn our attention to the Riccati operator  $\Pi$ . It follows from the fact that

$$\Pi x_0 = \mathcal{C}^*Jy^{\text{opt}} = \mathcal{C}^*J\mathcal{C}_\circ x_0$$

and from (37) and (69) that  $\Pi \in \mathcal{L}(W; V^*)$ . Clearly, this combined with (7) and (30) implies that, for all  $u_o \in U$

$$\lim_{\beta \rightarrow \infty} B^*\Pi(\beta I - A)^{-1}Bu_0 = 0.$$

Note that this means that the correction term mentioned in the statement of Theorem 1 is zero.

At this point we turn to our attention to Theorem 2. By the regularity of  $\mathcal{D}$  established in Lemma 2, we have (in the notations of Theorem 2)

$$y_0 = Cx_0 + Du_0$$

(to see this, let  $\alpha \rightarrow \infty$  in (64) and use (30)).

We next want to let  $\bar{\beta} = \beta \rightarrow \infty$  in (67), so let us examine the terms in this equation. From Lemma 5 we know that  $\mathcal{D}^*$  is regular, i.e.,

$$\lim_{\beta \rightarrow \infty} (\widehat{\mathcal{D}}(\beta))^*Jy_0 = D^*Jy_0 = D^*J(Cx_0 + Du_0).$$

The term  $Ax_0 + Bu_0$  belongs to  $H$ , so  $\Pi(Ax_0 + Bu_0) \in H$ , and  $(\beta I - A^*)^{-1}\Pi(Ax_0 + Bu_0) \rightarrow 0$  in  $H_1^* \subset V^*$  as  $\beta \rightarrow \infty$ , hence

$$\lim_{\beta \rightarrow \infty} B^*(\beta I - A^*)^{-1}\Pi(Ax_0 + Bu_0) = 0.$$

It follows from (29) and [Staffans 1997a, Lemma 32] that  $x_0 \in W$ , hence  $\Pi x_0 \in V^*$ . As  $\beta(\beta I - A^*)^{-1}$  tends strongly to the identity in  $V^*$  as  $\beta \rightarrow \infty$  (this follows from (6)), we find that

$$\lim_{\beta \rightarrow \infty} B^*(\beta I - A^*)^{-1}\Pi\beta x_0 = B^*\Pi x_0.$$

Thus, combining these three estimates with (67) we conclude that the limit

$$\lim_{\beta \rightarrow \infty} (I - \widehat{\mathcal{F}}(\beta))^* S w_0 = B^*\Pi x_0 + D^*J(Cx_0 + Du_0) \quad (70)$$

exists for all quadruples  $(x_0, u_0, y_0, w_0)$  of the type mentioned in Theorem 2.

We claim that (70) implies that  $\mathcal{X}^*$  is regular, arguing as follows. In the statement of Theorem 2  $w_0$  and  $y_0$  are functions of  $x_0$  and  $u_0$ , but it is possible to use  $w_0$  as an independent variable and let the other parameters depend on  $w_0$ ; for example, we can let  $u_\cup(t) = w(t)$  be an arbitrary function in  $W^{1,2}(\mathbf{R}^-; U)$  with compact support and with  $w(0) = w_0$ , and let  $x_0, y_0$ , and  $u_0$  be given by  $x_0 = x(0)$ ,  $y_0 = y(0)$ , and  $u_0 = u(0)$ , where  $x, y$  and  $u$  are the corresponding signals in the time invariant setting of [Staffans 1997c, Figure 2.1] (see that paper for details). Thus,  $S$  being invertible, (70) implies that the limit

$$\lim_{\beta \rightarrow \infty} (\widehat{\mathcal{X}}(\beta))^* u_0 = \lim_{\beta \rightarrow \infty} (I - \widehat{\mathcal{F}}(\beta))^* u_0$$

exists for all  $u_0 \in U$ , i.e.,  $\mathcal{X}^*$  is regular. Moreover, if we denote the feed-through operator of  $\mathcal{X}^*$  by  $X^*$ , then

$$X^* S w_0 = (B^*\Pi + D^*JC)x_0 + D^*JDu_0.$$

From this equation it is possible to solve  $w_0$  for the following reason. We know that  $\mathcal{X}^*$  is regular and that  $\widehat{\mathcal{X}}$  is invertible in  $H^\infty(U; U; 0)$ . This implies that  $X^*$  has a left inverse<sup>11</sup>  $M^* = (X X^*)^{-1}X$ . Thus,

$$w_0 = S^{-1}M^*((B^*\Pi + D^*JC)x_0 + D^*JDu_0). \quad (71)$$

---

<sup>11</sup>This is true since  $\langle u_0, X X^* u_0 \rangle = \langle X^* u_0, X^* u_0 \rangle = \lim_{\beta \rightarrow \infty} \langle \widehat{\mathcal{X}}^*(\beta) u_0, \widehat{\mathcal{X}}^*(\beta) u_0 \rangle \geq \inf_{\Re s > 0} \langle \widehat{\mathcal{X}}^*(s) u_0, \widehat{\mathcal{X}}^*(s) u_0 \rangle \geq \epsilon^2 \|u_0\|^2$  for all  $u_0 \in U$ ; here  $1/\epsilon = \sup_{\Re s > 0} \|(\widehat{\mathcal{X}}^*(s))^{-1}\|_{\mathcal{L}(U)} < \infty$ . This is the adjoint version of [Weiss 1994b, Proposition 4.6].

At this point, let us take  $x_0 \in H_1$  and  $u_0 = 0$ . Then (65) and (71) give

$$Kx_0 = -S^{-1}M^*(B^*\Pi + D^*JC)x_0, \quad x_0 \in H_1.$$

But this implies that  $K \in \mathcal{L}(W;U)$  (since  $H_1$  is dense in  $W$  and the operators on the right hand side belong to  $\mathcal{L}(W;U)$ ).

The regularity on  $\mathcal{F}$  now follows from Lemma 4, and this implies that also  $\mathcal{X} = I - \mathcal{F}$  is regular. Thus, both  $\mathcal{X}$  and  $\mathcal{X}^*$  are regular.

The rest of the claims in Theorem 1 follow from [Staffans 1997c, Theorem 6.1] and Mikkola [1997], except that these papers only tell us that (44) is valid in  $\mathcal{L}(H_1;H_{-1}^*)$  and that (48) is valid for all  $x_0, x_1 \in H_1$ . To derive the stronger statement given here it suffices to observe that  $H_1$  is dense in  $W$ , that the first operator in (44) belongs to  $\mathcal{L}(W;(\alpha I - A^*)V^*)$ , and that the operators on the right hand side of (44) belong to  $\mathcal{L}(W;W^*)$ ; hence the remaining operator  $\Pi A$  in (44) must belong to  $\mathcal{L}(W;W^* + (\alpha I - A)^*V^*)$ .

□

*Proof of Corollary 1.* Clearly, conditions (3), (5), and (6) hold with this choice of  $W$  and  $V$ .

The proofs of (15), (16), (33), (34), (36), and (37) are straightforward. They are all consequences of well-known properties of convolutions operators (in particular, of Young's inequality) and of the simple fact that for all  $\gamma$ ,  $0 \leq \gamma < 1$ , and all  $\alpha \in \rho(A)$ ,

$$\|(\alpha I - A)^\gamma \mathcal{A}(t)\| \leq Ct^{-\gamma} e^{-\epsilon t} \quad (72)$$

for some  $C > 0$ ,  $\epsilon > 0$ , and all  $t > 0$ ; here the norm represents the operator norm in any one of the three basic spaces  $W$ ,  $H$ , and  $V$ . In particular, this means that for each fixed  $x \in W$ , the function  $t \mapsto (\alpha I - A)^\gamma \mathcal{A}(t)x$  belongs to  $L^p(\mathbf{R}^+;W)$  for all  $p < 1/\gamma$ , and the same claims are true with  $W$  replaced by  $H$  and by  $V$ . We leave the proofs of these six claims to the reader.

To prove the positivity condition (39) it suffices to observe that the  $J$ -coercivity implies that  $\widehat{\mathcal{D}}^*(j\omega)J\widehat{\mathcal{D}}(j\omega) \geq \epsilon I$  for some  $\epsilon > 0$  and all  $\omega \in \mathbf{R}$ , and to let  $\omega \rightarrow \infty$ .

To verify (41) we make repeated use Young's inequality and the fact that for each  $u \in U$ ,  $\mathcal{E}u \in L^p(\mathbf{R};Y)$  for some  $p > 1$ . This argument is known under the name "boot-strap argument", see, e.g., [Lasićka et al. 1995, pp. 556–557] (the same argument appears already in [Lasićka and Triggiani 1983, pp. 52–53]).

We conclude that Hypothesis 1 holds, and that Theorem 1 applies.

To prove that the closed loop semigroup  $\mathcal{A}_\zeta$  is analytic in the given spaces we first use [Weiss 1994b, Theorem 7.2] to show that its generator is given by  $A_\zeta = A + BK$  where  $BK \in \mathcal{L}(W; V)$ . This, together with, e.g., [Lunardi 1995, Propositions 2.2.15 and 2.4.1] implies that  $A_\zeta = A + BK$  generates an analytic semigroup in  $V$ , and that  $(\alpha I - A_\zeta)^{-1}V = (\alpha I - A)^{-1}V$  for all  $\alpha \in \rho(A_\zeta) \cap \rho(A)$ . Thus  $(\alpha I - A_\zeta)^{-\gamma}V = (\alpha I - A)^{-\gamma}V$  for all  $0 \leq \gamma \leq 1$ , and  $A_\zeta$  generates analytic semigroups in  $(-A)^{-\gamma}V$  for all  $0 \leq \gamma \leq 1$ . By repeating the same argument with  $A_\zeta$  replaced by its adjoint  $A_\zeta^* = A^* + K^*B^*$  we conclude that  $A_\zeta^*$  generates analytic semigroups in  $(-A^*)^{-\gamma}W^*$  for all  $0 \leq \gamma \leq 1$ . Together these two claims imply the claim that  $A_\zeta$  generates analytic semigroups in  $(-A)^\gamma H$  for all  $\gamma_2 - 1 \leq \gamma \leq 1 - \gamma_1$ .

The exponential stability of  $\mathcal{A}_\zeta$  follows from [Datko 1970, Corollary, p. 615] and the fact that  $x^{\text{opt}} = \mathcal{A}_\zeta x_0 \in L^2(\mathbf{R}^+; H)$  for all  $x_0 \in H$ ; this in turn follows from (72), Young's inequality, and the fact that  $u^{\text{opt}} \in L^2(\mathbf{R}^+; U)$ .

The proof of the final claim about the regularity of the input/output maps is also based on the estimate (72). We leave this easy proof to the reader.

The only thing left to verify is that the Riccati equation has the additional smoothness property; i.e., that  $\Pi \in \mathcal{L}(W; (-A^*)^{-1+\gamma_1+\epsilon}H)$  for all  $\epsilon > 0$ , because the claims about the exact sense in which (44) and (48) are valid follow once this claim has been proved. To get this property we apply Theorem 1 with a different choice of  $V$ , i.e., we replace  $V$  by  $\tilde{V} = (-A)^{1-\gamma_1-\epsilon}H$ , where  $0 < \epsilon \leq 1 - \gamma_1 - \gamma_2$  (we can always decrease the value of  $\epsilon$  without loss of generality). As  $V \subset \tilde{V} \subset H_{-1}$ , only one condition in Hypothesis 1 becomes stronger when  $V$  is replaced by  $\tilde{V}$ , namely (37), but the same argument that establishes (37) for the original space  $V$  is valid for  $\tilde{V}$ , too. Thus, we conclude that Theorem 1 applies with  $V$  replaced by  $\tilde{V}$ , and we get the extra smoothness of  $\Pi$ .  $\square$

## 4 Comments on the Unstable Case.

According to [Staffans 1997c, Theorems 4.4 and 6.1], most of the basic results on which the proof of Theorem 1 are based remain valid for unstable systems which are jointly stabilizable and detectable. This means that Theorem 1, too, can in principle be extended to certain unstable but jointly stabilizable and detectable systems. The idea of the proof for this case is to first stabilize the system, and to then apply Theorem 1 to the stabilized system. For this to be possible we need the stabilized system to have the same type of

smoothness properties as the original system, i.e., the assumptions listed in Hypothesis 1 should be satisfied by the stabilized system. This limits the set of permitted preliminary feedbacks: They should have the same type of continuity properties as the original system, i.e., among others we need to assume that the preliminary feedback operator  $K^1$  and the corresponding input/output map  $\mathcal{F}^1$  satisfy

$$K^1 \in \mathcal{L}(W; U), \quad (I - \mathcal{F}^1)^{-1} \pi_+ : C(\mathbf{R}^+; U) \rightarrow C(\mathbf{R}^+; U).$$

Since our results for this case are still far from complete, we leave the study of the unstable system to a later time.<sup>12</sup>

## 5 Change of Pivot Space.

In much of the earlier literature on parabolic equations our basic state space  $H$  is absent, and its role as pivot space is taken over by some other space  $Z \subset W$ , usually  $Z = W$ . Fortunately, in the applications of the theory that we shall present below, there is a simple relationship between  $H$  and  $Z$ , and this makes it possible to pass from one formulation to the other without difficulty.

In the rest of this section we suppose that the semigroup  $\mathcal{A}$  is *analytic* in  $H$ , and that

$$Z = (\alpha I - A)^{-\beta} H, \tag{73}$$

where  $\alpha \in \rho(A)$ ,  $\beta \in \mathbf{R}$ , and  $(\alpha I - A)^{-\beta}$  represents the usual fractional power of  $(\alpha I - A)$ . If we use  $Z$  as pivot space instead of  $H$ , then the definition of all the adjoint operators change, and so does the definition of the Riccati operator. Let us continue to denote adjoints with respect to the space  $H$  by  $*$ , and let us use  $\#$  to denote adjoints with respect to  $Z$ . Moreover, let us denote the Riccati operator in  $Z$  by  $\Pi_Z$ , i.e.,

$$\langle x_0, \Pi_Z x_0 \rangle_Z := \min_{u \in L^2(\mathbf{R}^+; U)} Q(x_0, u). \tag{74}$$

---

<sup>12</sup>Among others, the sufficiency result presented in Mikkola [1997] applies so far only to the stable case. An extension of Mikkola [1997] to the unstable case is being worked out at the moment.



**Proposition 1** *Suppose that (73) holds, and introduce the notations explained above. Define  $\tilde{A} := \alpha I - A$ , and use the abbreviation  $\tilde{A}^{\beta*} := (\tilde{A}^*)^\beta$  and  $\tilde{A}^{-\beta*} := (\tilde{A}^*)^{-\beta}$ . Then the following identities hold:  $x$*

$$\begin{aligned}
\langle x, y \rangle_Z &= \left\langle \tilde{A}^\beta x, \tilde{A}^\beta y \right\rangle_H, \\
A^\# &= \tilde{A}^{-\beta} A^* \tilde{A}^\beta, \\
B^\# &= B^* \tilde{A}^{\beta*} \tilde{A}^\beta = B^* \tilde{A}^\beta \tilde{A}^{\beta\#}, \\
C^\# &= \tilde{A}^{-\beta} \tilde{A}^{-\beta*} C^* = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta} C^*, \\
D^\# &= D^*, \\
K^\# &= \tilde{A}^{-\beta} \tilde{A}^{-\beta*} K^* = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta} K^*, \\
\Pi_Z &= \tilde{A}^{-\beta} \tilde{A}^{-\beta*} \Pi = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta} \Pi, \\
B^* \Pi &= B^\# \Pi_Z, \\
V^\# &= \tilde{A}^{-\beta} \tilde{A}^{-\beta*} V^* = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta} V^*, \\
H^\# &= \tilde{A}^{-\beta} \tilde{A}^{-\beta*} H = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta} H = \tilde{A}^{-\beta\#} Z, \\
W^\# &= \tilde{A}^{-\beta} \tilde{A}^{-\beta*} W^* = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta} W^*.
\end{aligned}$$

*In particular, the connection between  $K$  and  $\Pi$  given in Theorem 1 and stays the same in both settings, and the Riccati equation is multiplied by the invertible operator  $\tilde{A}^{-\beta} \tilde{A}^{-\beta*} = \tilde{A}^{-\beta\#} \tilde{A}^{-\beta}$  and it takes its values in  $\tilde{A}^{-\beta} \tilde{A}^{-\beta*} H_{-1}^*$  instead of in  $H_{-1}^*$ . In the new setting the continuity properties of the adjoint operators and the Riccati operator become*

$$\begin{aligned}
B^\# &\in \mathcal{L}(V^\#; U), \\
C^\# &\in \mathcal{L}(Y; W^\#), \\
K^\# &\in \mathcal{L}(U; W^\#), \\
\Pi_Z &\in \mathcal{L}(W; V^\#) \cap \mathcal{L}(H; H^\#).
\end{aligned}$$

*If, furthermore*

$$W = \tilde{A}^\delta Z = (\alpha I - A)^\delta Z, \quad V = \tilde{A}^{\delta+\gamma} Z = (\alpha I - A)^{\delta+\gamma} Z,$$

*for some constants  $\gamma$  and  $\delta$ , then*

$$W^\# = \tilde{A}^{-\delta\#} Z, \quad V^\# = \tilde{A}^{-(\delta+\gamma)\#} Z,$$

and the continuity requirements on  $B^\#$  and  $\Pi_Z$  in Corollary 1 become

$$B^\# \tilde{A}^{-(\delta+\gamma)\#} \in \mathcal{L}(Z, U), \quad \tilde{A}^{(1+\delta-\epsilon)\#} \Pi_Z \tilde{A}^{-\delta} \in \mathcal{L}(Z; Z),$$

where  $\gamma = \gamma_1 + \gamma_2$  and  $\epsilon > 0$  is arbitrary.

We leave the straightforward proof of this proposition to the reader.

Thus, generally speaking, the input operator  $B$  becomes more unbounded in this setting, the output and feedback operators  $C$  and  $K$  become bounded, and the Riccati operator gets an additional “smoothing” property, but the nature of the problem does not change.<sup>13</sup>

## 6 Comparison to Earlier Work.

Let us begin by comparing our Theorem 1 with the theory for the parabolic equation presented in Lasiecka and Triggiani [1991b]. In this comparison we take (the left hand side refers to concepts used in this work, and the right hand side gives the corresponding entry in [Lasiecka and Triggiani 1991b, Sections 2 and 5.1], and  $\alpha \in \rho(A)$ ):  $x := y$  (the state),  $y := \begin{bmatrix} Ry \\ u \end{bmatrix}$  (the output),  $C := \begin{bmatrix} R \\ 0 \end{bmatrix}$  (the observation operator),  $D := \begin{bmatrix} 0 \\ I \end{bmatrix}$  (the feed-through operator),  $J := \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$  (the weighting operator),  $C^*JC := R^*R$ ,  $D^*JD := I$ ,  $D^*JC := 0$ ,  $Y := \begin{bmatrix} Y \\ U \end{bmatrix}$  (the output space),  $Z = W = (\alpha I - A)^{-\beta}H = (\alpha I - A)^{-\gamma}V := Y$  (the pivot space),  $\beta := \gamma/2$  (relation between  $W$  and  $H$ ), and  $\delta := 0$  (since  $Z = W$ ) (the operator  $G$  in Lasiecka and Triggiani [1991b] is a finite horizon weight that has no counterpart in our work).

Comparing our Corollary 1 with the theory for the parabolic equation presented in Lasiecka and Triggiani [1991b] we find many similarities, but also differences. With the choice of parameters described above the setting in Lasiecka and Triggiani [1991b] is the one presented Proposition 1, except for the facts that there the system need not be stable, and that no explicit use is made of our basic pivot space  $H$ . In the stable case the basic conclusions are essentially the same, and our additional admissibility assumptions on

---

<sup>13</sup>The setting described in Proposition 1 agrees with the parabolic setting in Lasiecka and Triggiani [1991b] in the stable case. We remark that the continuity properties in our original setting resemble those that are used for the hyperbolic case in Lasiecka and Triggiani [1991b]. Thus, in our setting the sharp distinction between the parabolic case and the hyperbolic case made in Lasiecka and Triggiani [1991b] disappears. We regard this as one of the major advantages with our setting.

$B$  and  $C$  related to the space  $H$  are satisfied in all the parabolic examples presented in Lasiecka and Triggiani [1991b]. However, a major difference is that Lasiecka and Triggiani [1991b] gives a much more complete treatment of the unstable case. In particular, it is shown that the so called “finite cost condition” implies that the system is stabilizable (which in principle makes it possible to apply the result outlined in Section 4 in the unstable case). On the other hand, our cost function  $Q$  is slightly more general (for example, it can be used in the setting of the bounded real and positive real lemmas), and we are able to say something about the input/output behavior of the open and closed loop systems, to which no attention is paid in Lasiecka and Triggiani [1991b]. In particular, we show that the closed loop system is a regular well-posed linear system. Moreover, our Theorem 1 can easily be extended to cover the full information  $H^\infty$  problem as well.<sup>14</sup>

Although this work is primarily aimed at the parabolic case, it is interesting to observe that our Theorem 1 is not that distant from the results given in Lasiecka and Triggiani [1991b] for *hyperbolic* equations. In the “second class (first form)” (originally published in Flandoli et al. [1988]) it is assumed that [Lasiecka and Triggiani 1991b, condition (H.2)] holds; that condition is equivalent to our condition (15) if we take  $Y = H = W$ . This implies that the input/output map  $\mathcal{D}: L_c^2(\mathbf{R}; U) \rightarrow C_c(\mathbf{R}, Y)$  is more “smoothing” than in our condition (28). The crucial assumptions missing in [Lasiecka and Triggiani 1991b, Theorem 5.2] are our (37) and (41), and in that theorem  $u^{\text{opt}}(x_0)$  need not be continuous for all  $x_0 \in H = W$ .

In spirit our Theorem 1 is even closer to the finite horizon result described in [Lasiecka and Triggiani 1991b, Section 3.3] (originally in Da Prato et al. [1986]) which gives sufficient conditions for the existence and *uniqueness* of a solution of the *differential Riccati* equation in the same general class. We again take  $H = W$ . Our Theorem 1 is not directly comparable to [Lasiecka and Triggiani 1991b, Theorem 3.3] because of the (finite horizon and the) fact that there the assumptions on the adjoint operators  $\mathcal{D}^*$  and  $\mathcal{C}^*$  are formulated in a different way. In the setting of [Lasiecka and Triggiani 1991b, Theorem 3.3], these operators are always followed in all the important formulas by (in our notation) the operator  $J\mathcal{C}$ , and our assumptions (37) and (63) are replaced by assumptions on (in our notation)  $\mathcal{D}^*J\mathcal{C}$  and  $\mathcal{C}^*J\mathcal{C}$ .<sup>15</sup> These assumptions are approximately of the same strength as our assumptions on

---

<sup>14</sup>It suffices to replace the references to Staffans [1997c] by the corresponding references to Staffans [1997de].

<sup>15</sup>Recall that our operator  $\mathcal{C}^*J\mathcal{C}$  corresponds to the operator  $R^*R$  in Lasiecka and

$\mathcal{D}^*$  and  $\mathcal{C}^*$  under the extra restriction that  $H = W$ . In particular, in the setting of [Lasiiecka and Triggiani 1991b, Theorem 3.3], the finite horizon version of (41) holds with  $n = 1$ .

Lasiiecka and Triggiani [1991b] also treats another hyperbolic case, i.e., the “second class (second form)” studied in [Lasiiecka and Triggiani 1991b, Section 4] (originally in Lasiiecka and Triggiani [1991a]). The basic assumption for that class imply (in our notation) that  $\mathcal{D}: L_c^2(\mathbf{R}; U) \rightarrow C_c(\mathbf{R}; Y)$ ,  $\mathcal{D}^*: L_{c^*}^1(\mathbf{R}^+; Y) \rightarrow L_{c^*}^2(\mathbf{R}^+; U)$ . That setting is studied only in the finite horizon case. Our conditions (37) and (41) are still missing in [Lasiiecka and Triggiani 1991b, Theorem 4.1], and there  $u^{\text{opt}}(x_0)$  need not be continuous for all  $x_0 \in H = W$ . On the other hand, by taking the set  $\mathcal{U}$  in [Lasiiecka and Triggiani 1991b, Theorem 4.2] to be the space of continuous functions we get a result which resembles a finite horizon version of our Theorem 1.

There is a number of other recent results for hyperbolic equations to which our Theorem 1 does not apply, e.g, Lasiiecka and Triggiani [1993]. On the other hand, it does apply to the examples studied in a sequence of related papers Lasiiecka et al. [1995 1997], Triggiani [1994]. For simplicity, let us just discuss the most recent of these paper, i.e., Lasiiecka et al. [1997], and compare it with our Corollary 1. In both cases the semigroup ( $= \mathcal{A}$  in our notation) is exponentially stable. For simplicity (and without significant loss of generality) we assume that the constants  $\gamma$  and  $\bar{\gamma}$  in Lasiiecka et al. [1997] satisfy  $\bar{\gamma} \leq 1 + \gamma$ . We take (again the left hand side refers to concepts used in this work, and the right hand side gives the corresponding entry in Lasiiecka et al. [1997], and  $\alpha \in \rho(A)$ ):  $x := z - B_1 u$  (the state),  $y := \begin{bmatrix} Rz \\ u \end{bmatrix}$  (the output),  $B := AB_1 + B_0$  (the control operator),  $C := \begin{bmatrix} R \\ 0 \end{bmatrix}$  (the observation operator),  $D := \begin{bmatrix} RB_1 \\ I \end{bmatrix}$  (the feed-through operator),  $J := \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$  (the weighting operator),  $C^*JC := R^*R$ ,  $D^*JD := I + B_1^*R^*RB_1$ ,  $D^*JC := B_1^*R^*R$ ,  $B^*\Pi := [B_0^* + B_1^*A^*]P$ ,  $K := G$  (the optimal feedback operator),  $Y = \begin{bmatrix} W \\ U \end{bmatrix}$  (the output space; here  $W$  is our space  $W$ ),  $(\alpha I - A)^{-\delta}W = (\alpha I - A)^{-\beta}H = (\alpha I - A)^{-(\delta+\gamma)}V := Y = (\alpha I - A)^\gamma Z$  (the pivot space),  $\delta := \bar{\gamma}$  (relation between the pivot space and  $W$ ),  $\beta := (\bar{\gamma} + \gamma + 1)/2$  (relation between the pivot space and  $H$ ), and  $\gamma := 1 + \gamma - \bar{\gamma}$  (relation between  $W$  and  $V$ ). In particular, the space  $Z$  in Lasiiecka et al. [1997] relates to our space  $W$  as  $W := (\alpha I - A)^{\gamma-\bar{\gamma}}Z \subset W$ . Note that we have chosen the constants  $\beta$ ,  $\gamma$ , and  $\delta$  in such a way that (7) holds and  $\delta + \gamma - \beta = \beta - \delta = \gamma/2 < 1/2$ , hence Corollary 1 applies.

---

Triggiani [1991b].

Comparing the conclusion of Corollary 1 to the result given in Lasiecka et al. [1997] we observe the following. The statements about the smoothness of the Riccati operator and the optimal  $u^{\text{opt}}$ ,  $x^{\text{opt}}$ , and  $y^{\text{opt}}$  given in Lasiecka et al. [1997] are the same as those given here. We say a little bit more about the input/output behavior of the closed loop system and about the spectral factors, but apart from this, the conclusion of Theorem 1 is essentially equivalent to those conclusions of Lasiecka et al. [1997] that refer to “the first algebraic Riccati equation (ARE1)”.

In addition, Lasiecka et al. [1997] develop another algebraic Riccati equation (ARE2).<sup>16</sup> That equation, too, has a simple interpretation in terms of Theorem 1: it is a feedback solution which corresponds to an inner-outer factorization for which the feed-through term  $X$  of the outer factor  $\mathcal{X}$  has not been normalized to be the identity operator, but instead it has been chosen to be equal to  $X := (I + GB_1)^{-1}$ . It is possible to do so if and only if this inverse exists, and a (highly nontrivial) proof of its existence is given in Lasiecka et al. [1997] under an extra compactness assumption. This changes the sensitivity operator  $S$  from the earlier  $S = D^*JD := I + B_1^*R^*RB_1$  into

$$\begin{aligned} S &= (X^{-1})^*D^*JDX^{-1} \\ &:= (I + GB_1)^*(I + B_1^*R^*RB_1)(I + GB_1) = I - B_0^*PB_1 - B_1^*PB_0. \end{aligned}$$

Observe, in particular, that the Riccati equation that we get in this way is nonstandard in the sense that the sensitivity operator  $S$  now becomes a function of the Riccati operator  $\Pi$ , hence it is not known in advance.<sup>17</sup> As observed in Lasiecka et al. [1997], this particular feedback can be interpreted as a feedback from the variable  $z = x + B_1u$ , which is used as “state” variable in Lasiecka et al. [1997]. This feedback is equivalent to the one in Corollary 1 in the absence of an external input to the closed loop system, but the controllability map and the input/output maps of the closed loop system change, and so does the cost of a nonzero input to the closed loop system (because of the change in the sensitivity operator). For details, see [Staffans 1997c, Lemma 2.4, Theorem 2.6(iii), and Proposition 4.8].

---

<sup>16</sup>This is the main result of that paper.

<sup>17</sup>Although it is easy to modify Corollary 1 in such a way that the feed-through operator of  $\mathcal{X}$  is allowed to differ from the identity, even the modified version does not apply to this case because of the coupling between  $S$  and  $\Pi$ .

## References

- A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and Control of Infinite Dimensional Systems*, volume 1 and 2. Birkhäuser, Boston, Basel, Berlin, 1992.
- G. Da Prato, I. Lasiecka, and R. Triggiani. A direct study of the Riccati equation arising in hyperbolic boundary control problems. *J. Differential Equations*, 64:26–47, 1986.
- R. Datko. Extending a theorem of A. M. Liapunov to Hilbert space. *Journal of Mathematical Analysis and Applications*, 32:610–616, 1970.
- F. Flandoli. Algebraic Riccati equation arising in boundary control problems. *SIAM Journal on Control and Optimization*, 25:612–636, 1987.
- F. Flandoli, I. Lasiecka, and R. Triggiani. Algebraic Riccati equations with non-smoothing observation arising in hyperbolic and Euler-Bernulli boundary control problems. *Annali di Matematica Pura ed Applicata*, 153:307–382, 1988.
- I. Lasiecka, D. Lukes, and L. Pandolfi. Input dynamics and nonstandard Riccati equations with applications to boundary control of damped wave and plate equations. *J. Optimization Theory Appl.*, 84:549–574, 1995.
- I. Lasiecka, L. Pandolfi, and R. Triggiani. A singular control approach to highly damped second-order equations and applications. Preprint, 1997.
- I. Lasiecka and R. Triggiani. Dirichlet boundary control problem for parabolic equations with quadratic cost: analyticity and Riccati’s feedback synthesis. *SIAM Journal on Control and Optimization*, 21:41–67, 1983.
- I. Lasiecka and R. Triggiani. Differential Riccati equations with unbounded coefficients: applications to boundary control/boundary observation hyperbolic problems. *Nonlinear Analysis, Theory, Methods & Applications*, 17:655–682, 1991a.
- I. Lasiecka and R. Triggiani. *Differential and Algebraic Riccati Equations with Applications to Boundary/Point Control Problems: Continuous Theory and Approximation Theory*, volume 164 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1991b.

- I. Lasiecka and R. Triggiani. Algebraic Riccati equations arising from systems with unbounded input-output solution operator: applications to boundary control problems for wave and plate equations. *Nonlinear Analysis, Theory, Methods & Applications*, 20:659–693, 1993.
- A. Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Equations*. Birkhäuser Verlag, Basel Boston Berlin, 1995.
- K. Mikkola. On the stable  $H^2$  and  $H^\infty$  infinite-dimensional regulator problems and their algebraic Riccati equations. Technical Report A383, Institute of Mathematics, Helsinki University of Technology, Espoo, Finland, 1997.
- L. Pandolfi. Dissipativity and Lur’e problem for parabolic boundary control systems. Manuscript, 1997.
- D. Salamon. Realization theory in Hilbert space. *Mathematical Systems Theory*, 21:147–164, 1989.
- O. J. Staffans. Quadratic optimal control of stable well-posed linear systems. To appear in Transactions of American Mathematical Society, 1997a.
- O. J. Staffans. Coprime factorizations and well-posed linear systems. To appear in SIAM Journal on Control and Optimization, 1997b.
- O. J. Staffans. Quadratic optimal control of well-posed linear systems. To appear in SIAM Journal on Control and Optimization, 1997c.
- O. J. Staffans. Feedback representations of critical controls for well-posed linear systems. Submitted, 1997d.
- O. J. Staffans. On the distributed stable full information  $H^\infty$  minimax problem. Submitted, 1997e.
- R. Triggiani. Optimal quadratic boundary control problem for wave- and plate-like equations with high internal damping: An abstract approach. In *Control of partial differential equations. IFIP WG 7.2 Conference in Trento, Italy, January 4–9, 1993*, volume 165 of *Lecture Notes in Pure and Applied Mathematics*, pages 215–270, New York, 1994. Marcel Dekker, Inc.
- G. Weiss. Admissibility of unbounded control operators. *SIAM Journal on Control and Optimization*, 27:527–545, 1989a.

- G. Weiss. Admissible observation operators for linear semigroups. *Israel Journal of Mathematics*, 65:17–43, 1989b.
- G. Weiss. Transfer functions of regular linear systems. Part I: Characterizations of regularity. *Transactions of American Mathematical Society*, 342: 827–854, 1994a.
- G. Weiss. Regular linear systems with feedback. *Mathematics of Control, Signals, and Systems*, 7:23–57, 1994b.
- M. Weiss and G. Weiss. Optimal control of stable weakly regular linear systems. To appear in *Mathematics of Control, Signals, and Systems*, 1997.