QUADRATIC OPTIMAL CONTROL THROUGH COPRIME AND SPECTRAL FACTORIZATIONS

OLOF J. STAFFANS
ÅBO AKADEMI UNIVERSITY
DEPARTMENT OF MATHEMATICS
FIN-20500 ÅBO, FINLAND
OLOF.STAFFANS@ABO.FI

Abstract. We consider the infinite horizon quadratic cost minimization problem for a linear time-invariant distributed parameter system with finitely may inputs and outputs. Our approach is to work in an input/output framework, and to reduce the problem to a symmetric Wiener-Hopf problem, that can be solved by means of a canonical factorization of the symbol. We have earlier solved the case where the system is stable, and this work is devoted to an extension of the theory to the unstable case. The extension is based on a right coprime factorization of the impulse response and on a preliminary stabilizing feedback, which makes it possible to reduce the unstable case to the stable one.

1. Introduction. This is a continuation of our earlier work [25] on the quadratic cost minimization problem for a linear time-invariant distributed parameter system with an impulse response of a certain type. In [25] we solved this problem for stable systems by means of spectral factorization, and here we extend that approach to unstable systems.

In order to provide some motivation for the approach that we use, let us begin by discussing the standard quadratic cost minimization problem for an infinite-dimensional time-invariant exponentially stabilizable and detectable system with bounded control and observation operators.[1] Suppose that we have such a system $\Sigma = (A, B, C, D)$ with finite-dimensional input space $U = \mathbf{R}^m$, (possibly) infinite-dimensional state space H, and finite-dimensional output space $Y = \mathbf{R}^n$, given by

$$\begin{aligned}
z'(t) &= Az(t) + Bu(t), \\
y(t) &= Cz(t) + Du(t), \quad t \in \mathbf{R}^+ = (0, \infty), \\
z(0) &= z_0.
\end{aligned} \tag{1}$$

All the space U, H, and Y are Hilbert spaces, A is the generator of a strongly continuous semigroup S on H, and $B \in \mathcal{L}(U; H)$, $C \in \mathcal{L}(H; Y)$, and $D \in \mathcal{L}(U; Y)$ are bounded linear operators. The object is to find the optimal control $u_{\text{opt}} \in L^2(\mathbf{R}^+; U)$ that minimizes the cost function

$$J(u) = \int_0^\infty |y(t)|^2 \mathrm{d}t. \tag{2}$$

Another (actually more frequently studied) possibility is to choose a different cost function, namely

$$J_R(u) = \int_0^\infty (|y(t)|^2 + |R^{1/2}u(t)|^2)\mathrm{d}t, \tag{3}$$

---

[1] See, for example, [1] or [6] for the appropriate definitions.

where R is a given positive semidefinite matrix. At the end of this introduction we shall return to this second possibility, and show that it can be regarded as a special case of (2).

We approach the problem described above in the following way. Since Σ is stabilizable, we can choose some stabilizing state feedback operator F, i.e., we can choose some bounded operator $F \in \mathcal{L}(H;U)$ such that $A + BF$ generates an exponentially stable semigroup S^F. Furthermore, we define a new auxiliary variable x by

$$(4) \qquad x(t) = u(t) - Fz(t), \quad t \in \mathbf{R}^+.$$

Then (1) can be rewritten in the form

$$(5) \qquad \begin{aligned} z'(t) &= (A + BF)z(t) + Bx(t), \quad t \in \mathbf{R}^+, \\ y(t) &= (C + DF)z(t) + Dx(t), \quad t \in \mathbf{R}^+, \\ u(t) &= Fz(t) + x(t), \qquad\qquad t \in \mathbf{R}^+, \\ z(0) &= z_0. \end{aligned}$$

According to the standard variation of constants formula,

$$(6) \qquad z(t) = S^F(t)z_0 + \int_0^t S^F(t-s)Bx(s)\mathrm{d}s, \quad t \in \mathbf{R}^+,$$

hence

$$(7) \quad y(t) = (C + DF)S^F(t)z_0 + \int_0^t (C + DF)S^F(t-s)Bx(s)\mathrm{d}s + Dx(t), \quad t \in \mathbf{R}^+,$$

and

$$(8) \qquad u(t) = FS^F(t)z_0 + \int_0^t FS^F(t-s)Bx(s)\mathrm{d}s + x(t), \quad t \in \mathbf{R}^+.$$

Define

$$(9) \qquad \begin{aligned} \nu(\mathrm{d}s) &= D\delta_0(\mathrm{d}s) + C^F S^F(s)B\mathrm{d}s, \\ \mu(\mathrm{d}s) &= I\delta_0(\mathrm{d}s) + FS^F(s)B\mathrm{d}s, \quad s \in \overline{\mathbf{R}}^+, \\ f(t) &= C^F S^F(t)z_0, \\ g(t) &= FS^F(t)z_0, \quad t \in \mathbf{R}^+, \end{aligned}$$

where $C^F = C + DF$, δ_0 represents a unit atom at zero, $\overline{\mathbf{R}}^+ = [0,\infty)$, and I is the identity matrix. Then ν and μ are matrix-valued measures that consist of an atom at zero plus a function in $L^1(\mathbf{R}^+)$, and f and g are L^2-functions (of appropriate dimensions). Moreover, equations (7) and (8) can be written in the form

$$(10) \qquad \begin{aligned} y(t) &= (\nu * x)(t) + f(t), \quad t \in \mathbf{R}^+, \\ (\mu * x)(t) &= u(t) - g(t), \qquad t \in \mathbf{R}^+. \end{aligned}$$

The convolutions are defined in the usual way, for example,

$$(\nu * x)(t) = \int_{[0,t]} \nu(\mathrm{d}s) x(t-s), \quad t \in \mathbf{R}^+,$$

for each function x locally in L^2 on $\mathbf{R}^+$. The measures ν and μ have one additional special property, namely, they are *right coprime* in the sense of Definition 2.1 below; this follows from, for example, [3, Lemma 2].[2]

In this work we study more general input/output relations of the type (10), without assuming anything about the underlying system. In fact, most of the time there is no need to assume a priori that there is any system $\Sigma = (A, B, C, D)$ underlying (10), and while we develop the basic theory we make no reference to the state space H, only to the finite-dimensional input and output spaces $U = \mathbf{R}^m$ and $Y = \mathbf{R}^n$. We work exclusively with the data given in (10), i.e., ν, μ, f, and g, under assumptions that will be explained in a moment.[3] The functions u and y are considered to be the control and observation, respectively, and the object is to find the control $u \in L^2(\mathbf{R}^+; \mathbf{R}^m)$ that minimizes either the cost function J defined in (2) or the cost function J_R defined in (3).

We make the following basic assumptions on the data in (10) (in addition to some other assumptions introduced later): The functions f and g are supposed to belong to $L^2(\mathbf{R}^+; \mathbf{R}^n)$ and $L^2(\mathbf{R}^+; \mathbf{R}^m)$, respectively. The matrix-valued measures ν and μ (of dimensions $n \times m$ and $m \times m$, respectively) are required to be of bounded variation on $\overline{\mathbf{R}}^+$ and to have no singular non-atomic part, and they should be right coprime in the sense of Definition 2.1. A measure without a singular non-atomic part can be split into a discrete part and an absolutely continuous part, e.g., ν can be written as

$$\nu = \sum_{i=0}^{\infty} N_i \delta_{a_i} + N,$$

where $N_i \delta_{a_i}$ represents an $n \times m$-dimensional atom of size N_i at the point $a_i \geq 0$, and $N \in L^1_{\mathrm{loc}}(\mathbf{R}^+; \mathbf{R}^{n \times m})$. That ν is of bounded total variation means that the sum $\sum_{i=0}^{\infty} \|N_i\|$ is finite, and that $N \in L^1(\mathbf{R}^+; \mathbf{R}^{n \times m})$. The space of scalar measures of bounded variation without a singular non-atomic part forms an extensively studied commutative Banach algebra. We shall refer to this algebra as the Beurling-Wiener-Pitt algebra after [2] and [37], and denote it by $\mathcal{A}(\overline{\mathbf{R}}^+)$.[4] Thus, ν and μ are required to belong to $\mathcal{A}(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$ and $\mathcal{A}(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$, respectively, and they should be right coprime in the algebra $\mathcal{A}(\overline{\mathbf{R}}^+)$ in the sense of Definition 2.1 (which is a special case of [33, Definition 1, p. 331]).

At this point, let us remark that we have chosen to work in the algebra $\mathcal{A}$ of impulse responses instead of in the more common algebra H^∞ of transfer functions

---

[2] This is where the detectability assumption on Σ is used.

[3] These assumptions are the following: f and g are L^2-functions on $\mathbf{R}^+$, ν and μ are measures on $\mathbf{R}^+$ that are right coprime, μ has an atom at zero equal to an identity matrix, and ν satisfies the Spectral Factorization Hypothesis 2.6.

[4] See [6, Appendix A.7.4], [7, p. 246], [13, Section 4.4], and [15, Sections 4.16–4.18] for more recent discussions of some measure algebras containing this one as a special case.

because of the fact that $\mathcal{A}$ has nicer spectral and coprime factorization properties than H^∞, and H^∞ is too large for some of our results to be true in H^∞. On the other hand, $\mathcal{A}$ is big enough to contain many of the standard examples; see Section 8. We shall return to the more general algebra H^∞ in [26, 27, 28, 29, 30].

According to [24, Theorem 1.2], in order for the observation y in (10) to be well-defined and to depend continuously on u in $L^2_{\text{loc}}(\mathbf{R}^+)$, it is necessary and sufficient that μ has an invertible atom at the origin.[5] We assume throughout that this is the case, and, without significant loss of generality, we take this atom to be the identity matrix.[6] In this case μ has a convolution inverse μ^{-1}, a measure supported on $\overline{\mathbf{R}}^+$, which is otherwise of the same type as μ except that its total variation may be infinite, cf. [13, Theorem 1.5, p. 114]. Eliminating the auxiliary variable x from (10) we get the input/output relation

$$(11) \qquad\qquad y = \nu * \mu^{-1} * (u - g) + f.$$

In particular, if we take the "transient" terms f and g to be zero, then we find that the impulse response of the system is given by

$$(12) \qquad\qquad \gamma = \nu * \mu^{-1}.$$

The formula above defines a right coprime factorization of γ in $\mathcal{A}(\overline{\mathbf{R}}^+)$ in the sense of [33, Definition 1, p. 331]. Thus, the class of impulse responses that we are able to handle is characterized by the fact that they have a right coprime factorization in $\mathcal{A}(\overline{\mathbf{R}}^+)$.

How do we then solve the quadratic cost minimization problem for the input/output relation (10) without using any state space methods? The basic idea of our solution, presented in Sections 2 and 3, is the following. The unstable problem that we consider here resembles the stable problem solved in [25], but there is one major complication: we cannot let u be an arbitrary function in $L^2(\mathbf{R}^+; \mathbf{R}^m)$, since not every such function u produces an output y in $L^2(\mathbf{R}^+, \mathbf{R}^n)$. Thus, in the unstable case, we are dealing with a constrained minimization problem and not with an unconstrained one. The key observation that helps us overcome this difficulty is that, although u cannot be chosen freely, the function x in (10) can be thought of as a free parameter in $L^2(\mathbf{R}^+; \mathbf{R}^m)$ in the sense that u will produce an output $y \in L^2(\mathbf{R}^+; \mathbf{R}^n)$ if and only if u is of the form $u = \mu * x + g$ for some function $x \in L^2(\mathbf{R}^+; \mathbf{R}^m)$. The first equation in (10) is independent of u, so it is possible to apply the theory developed in [25]. That theory is based on a Wiener-Hopf factorization, and in order to apply this theory we need the additional Spectral Factorization Hypothesis 2.6 given below.[7] That theory gives us the optimal x_{opt} and y_{opt}, after which we get the optimal u_{opt} from the second equation in (10). See Sections 2 and 3 for details.

---

[5] In the terminology of [14], this means that the "D-operator" $\mu*$ must be atomic at zero.

[6] Let D_μ be the atom of μ at zero, multiply the second equation in (10) by D_μ^{-1}, and replace $D_\mu^{-1}\mu$, $D_\mu^{-1}u$, and $D_\mu^{-1}g$ by μ, u, and g.

[7] This hypothesis is satisfied for the system $\Sigma = (A, B, C, D)$ with bounded control and observation operators if and only if the system is coercive; see Corollary 2.9 and Definition 2.10.

Does the solution outlined above then tell us anything about the quadratic cost minimization problem for the original system $\Sigma = (A, B, C, D)$, from which (10) was derived? Indeed, it does. This is to some extent true even in the case where the control and observation operators B and C are unbounded. For example, the argument that we gave above reducing the quadratic cost minimization problem for the system (1) with bounded control and observation operators B and C to the same problem for the input/output relation (10) extends immediately to the Pritchard-Salamon class discussed, e.g., in [19, 20, 32]. The only difference is that the functions $f(t) = C^F S^F(t) z_0$, $g(t) = F S^F(t) z_0$, and $C^F S^F(s) B$ are no longer continuous, instead they are defined just almost everywhere and belong to $L^2(\mathbf{R}^+)$ (and "decay exponentially" at infinity in an L^2-sense). This means that the resulting pair of equations (10) still can be treated by our method. It is even possible to go one step further and start with a stabilizable and detectable regular abstract linear system $\Sigma = (A, B, C, D)$ in the sense of [5, 34, 35].[8] In this case f and g will still belong to $L^2(\mathbf{R}^+)$, but the closed loop impulse responses ν and μ need no longer be measures. However, if they are, then our technique applies. See the examples in Section 8.

For the class of regular Pritchard-Salamon systems $\Sigma = (A, B, C, D)$ it is well-known that the optimal control u_{opt} is of feedback type, i.e., $u_{\mathrm{opt}}(t) = K z_{\mathrm{opt}}(t)$, $t \in \overline{\mathbf{R}}^+$, where K is a bounded linear feedback operator and $z(t)$ is the state of Σ at time t. Moreover, the closed loop system $\Sigma^K = (A^K, B, C^K, D)$ with this feedback operator is another system of Pritchard-Salamon type. See, for example, [19, 20, 32]. Not much has been known about to what extent this is true in the more general class of regular abstract linear systems. Here we contribute to the solution of this problem in the following way. In order to set the stage for the genuinely new results, we first use our knowledge about the quadratic cost minimization problem for the input/output relation (10) to develop a fairly complete Riccati equation theory for the class of systems $\Sigma = (A, B, C, D)$ with bounded control and observation operators. The results that we derive are well-known, but we believe that our proof is new. The same method can be extended to the class of Pritchard-Salamon systems with a minimal effort. We have not included this extension here, since the result is hardly interesting enough to motivate the introduction of the needed technical machinery. Indeed, the result is exactly what one expects it to be, with no surprises. In order to derive a truly interesting new result we have to allow much more unboundedness in the control and observation operators B and C (roughly of the same order of magnitude as in [17]). In order to be able to present this result with a minimal amount of technicalities, we have chosen not to treat an arbitrary regular abstract linear system Σ, but to start with a given impulse response γ of the general type that we described earlier, and to study a particular realization of γ, i.e., a particular system Σ with this impulse response.[9]

The explanation of how we can use our knowledge about the solution to the quadratic cost minimization problem for (10) in order to derive results about a stabilizable system $\Sigma = (A, B, C, D)$ (with bounded or unbounded control and observation oper-

---

[8] See also [1, Section 3.3] and [21, 22].
[9] Extensions to arbitrary systems Σ are given in [26, 27, 28].

ators B, and C) requires the introduce of some notations: We denote the mappings from z_0 to the functions $f(t) = C^F S^F(t)z_0$ and $g(t) = F S^F(t)z_0$ by

$$(13) \qquad \mathcal{C}z_0 = (t \mapsto C^F S^F(t)z_0), \qquad \mathcal{F}z_0 = (t \mapsto F S^F(t)z_0).$$

Then, in all the cases mentioned above, $\mathcal{C}$ and $\mathcal{F}$ are bounded linear maps from H into $L^2(\mathbf{R}^+; \mathbf{R}^n)$ and $L^2(\mathbf{R}^+; \mathbf{R}^m)$, respectively, and in the case of bounded observation and feedback operators they map H continuously into $BC(\overline{\mathbf{R}}^+; \mathbf{R}^n)$ and $BC(\overline{\mathbf{R}}^+; \mathbf{R}^m)$.[10] Moreover, they map $\mathcal{D}(A)$ (the domain of the generator A) into $W^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$ and $W^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$, and the functions f and g in (10) are given by

$$(14) \qquad f = \mathcal{C}z_0, \qquad g = \mathcal{F}z_0,$$

where z_0 is the initial state of the system.

In order to simplify the discussion below, let us assume that the control, observation, and feedback operators B, C, and F are bounded.[11] Let us suppose that the optimal control u_{opt} in the cost minimization problem for $\Sigma = (A, B, C, D)$ is of state feedback type, i.e., $u_{\mathrm{opt}}(t) = K z_{\mathrm{opt}}(t)$ for all $t \in \overline{\mathbf{R}}^+$. Since this is true for all $t \in \overline{\mathbf{R}}^+$, it is surely true for $t = 0$, i.e.,

$$(15) \qquad K z_0 = K z_{\mathrm{opt}}(0) = u_{\mathrm{opt}}(0).$$

From the formulae given in Section 3, we know how to compute $u_{\mathrm{opt}}(0)$. Let K_Φ be the operator that maps $\begin{pmatrix} f \\ g \end{pmatrix}$ into

$$(16) \qquad K_\Phi \begin{pmatrix} f \\ g \end{pmatrix} = u_{\mathrm{opt}}(0).$$

The operator K_Φ is unbounded on $L^2(\mathbf{R}^+; \mathbf{R}^n) \times L^2(\mathbf{R}^+; \mathbf{R}^m)$, but it is bounded on, for example, $BC(\overline{\mathbf{R}}^+; \mathbf{R}^n) \times BC(\overline{\mathbf{R}}^+; \mathbf{R}^m)$. From (14) and (15) we conclude that K must be the bounded linear operator given by

$$(17) \qquad K = K_\Phi \begin{pmatrix} \mathcal{C} \\ \mathcal{F} \end{pmatrix}.$$

We compute K_Φ in Section 4, and express K_Φ in terms of the solution of the Wiener-Hopf factorization problem (see formula (53)). This formula is very similar to the corresponding formula in [25].

The argument above indicates that, if u_{opt} is of feedback type, then the feedback operator must be given by (17). It does not yet prove that u_{opt} is of feedback type, i.e., that $u_{\mathrm{opt}}(t) = K z_{\mathrm{opt}}(t)$ for $t \in \overline{\mathbf{R}}^+$. In Theorem 4.2 we prove (the well-known result) that for a coercive[12] system $\Sigma = (A, B, C, D)$ with bounded control and observation

---

[10] BC is the space of bounded continuous functions.

[11] An extended version of this argument applies in the unbounded case.

[12] See Definition 2.10.

operators B and C, the optimal u_{opt} is indeed of state feedback type. Our proof is very unorthodox in the sense that it makes no reference to the Riccati operator or to the Riccati equation; that reference is replaced by a reference to the Spectral Factorization Lemma 2.4.

So far we have said nothing about the Riccati operator Π of the system Σ. By using essentially the same argument as above we can compute this operator, too. Under our standing assumptions, it is not difficult to show that the optimal cost $J(u_{\mathrm{opt}})$ in the minimization problem (2) and (10) is a bounded quadratic function of the data $\begin{pmatrix} f \\ g \end{pmatrix}$; hence there is a unique positive self-adjoint operator Π_Φ such that the optimal cost can be written in the form

$$(18) \qquad \left\langle \begin{pmatrix} f \\ g \end{pmatrix}, \Pi_\Phi \begin{pmatrix} f \\ g \end{pmatrix} \right\rangle = J(u_{\mathrm{opt}}),$$

where the inner product is the usual inner product in $L^2(\mathbf{R}^+; \mathbf{R}^n) \times L^2(\mathbf{R}^+; \mathbf{R}^m)$. We define the Riccati operator of Σ by the analogous formula

$$\langle z_0, \Pi z_0 \rangle_H = J(u_{\mathrm{opt}}),$$

where H is the state space of Σ. Then, according to (14), Π will be given by[13]

$$(19) \qquad \Pi = (\, \mathcal{C}^* \quad \mathcal{F}^* \,) \, \Pi_\Phi \begin{pmatrix} \mathcal{C} \\ \mathcal{F} \end{pmatrix}.$$

We compute Π_Φ in Section 5, and give Π_Φ in formula (60) in terms of the solution of the spectral factorization problem. This formula turns out to be an almost exact copy of the corresponding formula in [25].

To connect our theory to the standard quadratic cost minimization theory we still have to connect the feedback operator K to the Riccati operator Π, and show that Π satisfies the algebraic Riccati equation. The two crucial equations are[14]

$$(20) \qquad K = - \left(D^* D \right)^{-1} \left(B^* \Pi + D^* C \right).$$

and

$$(21) \qquad \begin{aligned} \langle A z_1, \Pi z_0 \rangle_H &+ \langle z_1, \Pi A z_0 \rangle_H + \langle C z_1, C z_0 \rangle_{\mathbf{R}^n} \\ &= \langle K z_1, D^* D K z_0 \rangle_{\mathbf{R}^m}, \qquad z_0, \ z_1 \in \mathcal{D}(A). \end{aligned}$$

Clearly, by combining these two equations we get the standard algebraic Riccati equation

$$(22) \qquad \begin{aligned} \langle A z_1, \Pi z_0 \rangle_H &+ \langle z_1, \Pi A z_0 \rangle_H + \langle C z_1, C z_0 \rangle_{\mathbf{R}^n} \\ &= \left\langle \left(B^* \Pi + D^* C \right) z_1, \left(D^* D \right)^{-1} \left(B^* \Pi + D^* C \right) z_0 \right\rangle_{\mathbf{R}^m}, \\ &\qquad z_0, \ z_1 \in \mathcal{D}(A). \end{aligned}$$

---

[13] Actually, we shall see in Section 5, the optimal cost $J(u_{\mathrm{opt}})$ and the Riccati operator Π_Φ are independent of g, so this formula for Π does not depend on $\mathcal{F}$.

[14] In the case where the feed-through operator D is nonzero and coercive and the cost function is (2).

As we show at the end of Section 5, equations (20) and (21) can, indeed, be derived from our general results for (10) through some straightforward (but rather tedious) computations.

To get to the most interesting part of this work we have to treat a more general case of (10) that is not induced by a system of Pritchard-Salamon type. The purpose is to investigate to what extent equations (20)–(22) remain valid in this case. For simplicity, we do not treat an arbitrary regular abstract linear system, but rather a "canonical" realization of a given impulse response γ with a right coprime factorization $\gamma = \nu * \mu^{-1}$. This realization, presented in Section 6, is an extension of one of the four realizations developed in [25]. As always in the study of systems with an unbounded control operator, the term causing us the most trouble is the term $B^*\Pi$ in (20), and in order to simplify the treatment of this term as much a possible we have settled for a realization where this term is as simple as possible. That realization is based on an extended version of (10) where an initial function component has been added to x, namely[15]

$$
\begin{aligned}
x(t) &= \varphi(t), & t &\in \mathbf{R}^-, \\
y(t) &= (\nu * x)(t) + f(t), & t &\in \overline{\mathbf{R}}^+, \\
(\mu * x)(t) &= u(t) - g(t), & t &\in \overline{\mathbf{R}}^+.
\end{aligned}
\tag{23}
$$

Accordingly, we define the convolutions in a slightly different way as

$$
(\nu * x)(t) = \int_{\overline{\mathbf{R}}^+} \nu(\mathrm{d}s) x(t-s), \quad (\mu * x)(t) = \int_{\overline{\mathbf{R}}^+} \mu(\mathrm{d}s) x(t-s), \quad t \in \overline{\mathbf{R}}^+.
$$

Moreover, we replace the definitions of K_Φ and Π_Φ by

$$
K_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = u_{\mathrm{opt}}(0),
\tag{24}
$$

and

$$
\left\langle \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \right\rangle = J(u_{\mathrm{opt}}),
\tag{25}
$$

where the inner product is the usual inner product in $L^2(\mathbf{R}^-; \mathbf{R}^m) \times L^2(\mathbf{R}^+; \mathbf{R}^n) \times L^2(\mathbf{R}^+; \mathbf{R}^m)$.[16]

---

[15] We have, in addition, replaced $\mathbf{R}^+$ by $\overline{\mathbf{R}}^+$. This is due to the fact that we shall from time to time refer to this equation in cases were the functions are defined everywhere instead of almost everywhere, and in this case $\overline{\mathbf{R}}^+$ is the better choice.

[16] A reader who is not interested in the results presented in Sections 6–8 may throughout take the initial function φ to be zero, and stick with the old equations (10), (16), (18), and the earlier notion of a convolution (although $\mathbf{R}^+$ should be replaced by $\overline{\mathbf{R}}^+$ for the pointwise results). This simplifies some of the formulae in Sections 2–5.

Section 7 is devoted to a proof of the fact that formulae (20) and (21) remain almost valid for the realizations given in Section 6, this time with an unbounded operator K, provided the numerator ν in the right coprime factorization $\gamma = \nu * \mu^{-1}$ satisfies the Spectral Factorization Hypothesis 2.6. There is one extremely interesting difference: the matrix D^*D has been replace by a new matrix. According to (9), the matrix D is characterized by the fact that it is the feed-through matrix[17] of the numerator ν, and at the same time it is the feed-through matrix of the impulse response γ, since the feed-through matrix of the denominator μ is the identity matrix. In the situation described in Section 6 the feed-through matrix of μ is still normalized to be the identity operator, and the feed-through matrices D_γ and D_ν of γ and ν are the same. Let us denote this common matrix by $D = D_\gamma = D_\nu$. This is the matrix that we would expect to show up in (20) and (21). Indeed, it is the correct matrix in the term D^*C in (20), but in all other places it has been replaced by D_ξ, where ξ is the spectral factor of a symmetrized version of the numerator ν of γ. This means that the matrix D^*D must be replaced throughout by $D_\xi^*D_\xi$.[18] In the classical case these two matrices are the same, but this is not true in the case where the numerator ν contains some delayed atoms $N_i\delta_{a_i}$ with $a_i > 0$. This is exactly the same anomaly that we discovered in [25] in the stable case. See the examples in Section 8.

As we mentioned above, it is not standard to work with the cost function J in (2); the usual choice is to normalize the feed-through matrix D to be zero, and to replace J by the function J_R in (3). We have not done so here, since it leads to a slight loss of generality and to more complicated formulae.[19] Indeed, (3) is a special case of (2) in the following sense.[20] The argument is essentially the same for the the system (1) and the input/output relation (10), so below we give only the argument for (1).[21] Clearly $J_R(u)$ is equal to the L^2-norm of the vector $\begin{pmatrix} y \\ R^{1/2}u \end{pmatrix}$. By (7) and (8),

$$(26) \quad \begin{pmatrix} y(t) \\ R^{1/2}u(t) \end{pmatrix} = C_R^F S^F(t)z_0 + \int_0^t C_R^F S^F(t-s)Bx(s)\mathrm{d}s + D_R x(t), \quad t \in \overline{\mathbf{R}}^+,$$

where

$$(27) \quad D_R = \begin{pmatrix} D \\ R^{1/2} \end{pmatrix}, \quad C_R^F = C_R + D_F F = \begin{pmatrix} C + DF \\ R^{1/2}F \end{pmatrix}, \quad C_R = \begin{pmatrix} C \\ 0 \end{pmatrix}.$$

This leads to a system of the type (9) and (10), with Y replaced by $Y \times U$, D replaced by D^F, C replaced by C_R, and C^F replaced by $C_R^F = C_R + D_F F$. In order to keep

---

[17] According to Definition 2.7, the feed-through matrix of a measure is the atom that this measure has at the origin. The feed-through matrix is zero if the measure has no such atom.

[18] See Theorem 7.1.

[19] In the most general case it does not make sense to normalize the feed-through matrix D to be zero. See our discussion above on how we are sometimes forced to replace D^*D by a different operator.

[20] As the construction in [12, Section 5.2.3] shows, the converse is also true, at least in the finite-dimensional case. That construction is less appealing in the infinite-dimensional case since it does not preserve stability, i.e., it may convert a stable system into an unstable one. Moreover, it does not work for systems that are not regular, since such systems do not have well-defined feed-through operators.

[21] In the case of (10), replace ν and f by $\begin{pmatrix} \nu \\ R^{1/2}\mu \end{pmatrix}$ and $\begin{pmatrix} f \\ R^{1/2}g \end{pmatrix}$.

the notations simple, we shall throughout write D, C, and C^F, and we leave it to the reader to replace these operators by D_R, C_R, and C_R^F if the cost function is (3) instead of (2).

Finally, let us comment on how our results relate to some other recently obtained results.[22] A general comment that applies to all of these works is that they use almost exclusively the cost function J_R defined in (3) instead of our cost function J in (2), and, in addition, they usually normalize the feed-through operator D to be zero. Thus, in the comparison of those results with ours we should throughout replace D, C, and C^F by D_R, C_R, and C_R^F.

The most recent publications of Grabowski [10, 11] contain certain results that are quite close to some of ours. He uses spectral factorization to solve quadratic minimization problems, but he does not develop a general Riccati equation theory for systems with unbounded control and observation operators. We discuss Grabowski's example [11] in detail in Section 8.

Our results for systems with bounded control and observation operators extend those of Callier and Winkin [3, 4], in particular, Theorem 4.2 extends [4, Theorem 3] in several ways. There it is not proved that the optimal solution u_{opt} is of feedback type; that result is borrowed from the standard Riccati equation theory. Likewise, Callier and Winkin borrow (20) from the standard Riccati equation theory instead of proving this equation. They give a different formula [4, Formulae (19) and (21a)] (involving a Diophantine equation) for the restriction of the feedback operator K to the reachable subspace in terms of the spectral factorization, and they use a considerable amount of space (in particular, [4, Theorem 4]) to discuss how this restricted operator is related to the full feedback operator. Our formula (56) giving K directly in terms of the inverse of the spectral factor appears to be new.

In his thesis [36], Martin Weiss studies a spectral factorization problem of a more general type in the Pritchard-Salamon setting. In the Pritchard-Salamon case (that we did not include) our Theorems 4.2 and 5.2 become related to the sufficiency part of [36, Theorem 4.20].[23]

The recent book [1] of Bensoussan, Da Prato, Delfour, and Mitter uses the classical Riccati equation approach to the quadratic cost minimization problem for boundary control problems for partial differential equations, and whereas we can refer the reader to this book for results of that nature, it does not mention the spectral factorization problem, and it is not really a relevant reference for the major part of this work. The same statement applies to the books [16] by Lasiecka and Triggiani and [18] by Lions. In another more recent work [17] Lasiecka and Triggiani treat the quadratic cost minimization problem in an abstract setting for systems with very unbounded control operators B, but bounded observation operators C. The total amount of unboundedness in that paper (in B and C) is approximately the same as in our setting, but it is distributed in a different way (more unboundedness in B, and less in

---

[22] None of these have had any significant impact on our work, due to the fact that we got acquainted with them at a rather late stage.

[23] In [28] we extend the results presented here to cover the more general situation considered by Martin Weiss.

C). Their impulse responses are not always measures, and in that case our method does not apply. In spite of the great amount of unboundedness that Lasiecka and Triggiani are able to allow in the control operator B, they do not yet discover the fact that we consider to be the most interesting one, namely the need for the new weighting operator $(D_\xi^* D_\xi)^{-1}$ in the Riccati equation and in the equation connecting the feedback operator K to the Riccati operator. We believe that this is due to the fact that when they derive the Riccati equation they impose an extra regularity condition [17, Condition (H.5)] on the observation operator C.[24]

2. Coprimeness and Spectral Factorization. Let us begin with a definition of coprimeness:[25]

DEFINITION 2.1. *The measures $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$ and $\mu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ are right coprime in $M(\overline{\mathbf{R}}^+)$ iff the Bezout identity*

$$(28) \qquad \nu_1 * \nu + \mu_1 * \mu = I\delta_0$$

has a solution $\nu_1 \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times n})$ and $\mu_1 \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$. Here I is the identity matrix, and δ_0 is the Dirac delta.

As we mentioned earlier, the following result is the key observation, that enables us to reduce the unstable case to the stable case, solved in [25]:

LEMMA 2.2. *Let $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$ and $\mu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ be right coprime. Let $\varphi \in L^2(\mathbf{R}^-; \mathbf{R}^m)$, $f \in L^2(\mathbf{R}^+; \mathbf{R}^n)$, and $g \in L^2(\mathbf{R}^+; \mathbf{R}^m)$, and let x, y, and u be locally in L^2 and satisfy (23). Then $x \in L^2(\mathbf{R}^+; \mathbf{R}^m)$ if and only if both $y \in L^2(\mathbf{R}^+; \mathbf{R}^n)$ and $u \in L^2(\mathbf{R}^+; \mathbf{R}^m)$.*

Proof. We immediately observe that[26] $x = \overline{\pi}_+ x + \pi_- \varphi$. Since $\varphi \in L^2(\mathbf{R}^-; \mathbf{R}^m)$, the condition $x \in L^2(\mathbf{R}^+; \mathbf{R}^m)$ is therefore equivalent to $x \in L^2(\mathbf{R}; \mathbf{R}^m)$. In particular, then $y \in L^2(\mathbf{R}^+; \mathbf{R}^n)$ because $y = \nu * x + f$ on $\mathbf{R}^+$, and $u \in L^2(\mathbf{R}^+; \mathbf{R}^m)$ because $u = \mu * x + g$ on $\mathbf{R}^+$.[27] To prove the converse claim we choose some measures ν_1 and μ_1 satisfying the Bezout identity (28). We convolve the equation $\overline{\pi}_+ (\nu * (\overline{\pi}_+ x + \pi_- \varphi) + f - y) = 0$ with ν_1, convolve the equation $\overline{\pi}_+ (\mu * (\overline{\pi}_+ x + \pi_- \varphi) + g - u) = 0$ with μ_1, and add the resulting two equations to get

$$\overline{\pi}_+ x = \nu_1 * \overline{\pi}_+ (y - f - \nu * \pi_- \varphi) + \mu_1 * \overline{\pi}_+ (u - g - \mu * \pi_- \varphi).$$

This shows that $x \in L^2(\mathbf{R}^+; \mathbf{R}^m)$ whenever both $y \in L^2(\mathbf{R}^+; \mathbf{R}^n)$ and $u \in L^2(\mathbf{R}^+; \mathbf{R}^m)$. □

---

[24] Another difference in [17] is their standing coercivity hypothesis (H.4) that forces the output space Y to be infinite-dimensional.

[25] In the introduction we restricted ourselves to the class $\mathcal{A}(\overline{\mathbf{R}}^+)$ of measures without a singular non-atomic part. This is important if we want to apply Lemma 2.4, but elsewhere it does not matter if the measures include singular non-atomic parts or not. For this reason we shall in the sequel work in the class M of measures (with or without singular non-atomic parts). Lemma 2.4 is also true for measures with a sufficiently small singular non-atomic part, as can be shown with a perturbation argument.

[26] See Section 9 for a list of notations.

[27] The standard properties of convolutions may be found in many places, for example, in [13, Theorem 6.1, pp. 96–98] and in [25, Lemma 3.1].

Thanks to Lemma 2.2, we can, for a moment, forget about the last equation in (23), and study the quadratic cost minimization problem for the first two equations in (23) in the same say as we did in [25], with x replacing u. Once the optimal x_{opt} has been found, we can compute the optimal u_{opt} from the last equation in (23). In the new setting, [25, Lemma 3.2] reads as follows:

LEMMA 2.3. *Let $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$ and $\mu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ be right coprime, let $\varphi \in L^2(\mathbf{R}^-; \mathbf{R}^m)$, $f \in L^2(\mathbf{R}^+; \mathbf{R}^n)$, $g \in L^2(\mathbf{R}^+; \mathbf{R}^m)$, and suppose that ν is coercive in the sense that there exists some $\eta > 0$ such that*

$$(29) \qquad (\hat{\nu}(\mathrm{i}\omega))^* \hat{\nu}(\mathrm{i}\omega) \geq \eta^2 I, \quad \omega \in \mathbf{R}.$$

Then the function $J(u)$ defined in (2), with y given by (23), achieves its minimum at some $u_{\mathrm{opt}} \in L^2(\mathbf{R}; \mathbf{R}^m)$. The minimizing function u_{opt} is unique, and at the minimum, the optimal output y_{opt} and optimal control u_{opt} are given by

$$(30) \qquad \begin{aligned} y_{\mathrm{opt}} &= \nu * x_{\mathrm{opt}} + f, \\ u_{\mathrm{opt}} &= \mu * x_{\mathrm{opt}} + g, \end{aligned}$$

where x_{opt} is the solution of the symmetric Wiener-Hopf equation[28]

$$(31) \qquad \begin{aligned} x_{\mathrm{opt}}(t) &= \varphi(t), & t \in \mathbf{R}^-, \\ (\check{\nu} * \nu * x_{\mathrm{opt}})(t) &= -(\check{\nu} * f)(t), & t \in \overline{\mathbf{R}}^+. \end{aligned}$$

Proof. [29] Define $W(x) = J(u)$, where $J(u)$ is the function defined in (2), and y is given by (23). As we observed above, thanks to Lemma 2.2, the (constrained) minimization of $J(u)$ with respect to u can be replaced by a free minimization of $W(x)$ with respect to $x \in L^2(\mathbf{R}^+; \mathbf{R}^m)$. In this minimization we must split the function x into two parts: the part $\pi_- x = \pi_- \varphi$ is given in advance, and the free variable is $\overline{\pi}_+ x$. Thus, let us rewrite the first two equations in (23) in the form

$$y = \nu * (\overline{\pi}_+ x + \pi_- \varphi) + f,$$

and minimize with respect to $\overline{\pi}_+ x$.

If the coercivity condition (29) holds, then the convolution operator $\overline{\pi}_+ x \mapsto \nu * \overline{\pi}_+ x$ is strictly coercive (i.e., $\|\nu * \overline{\pi}_+ x\|_{L^2(\mathbf{R}^+)} \geq \eta \|x\|_{L^2(\mathbf{R}^+)}$). The first claim is then obvious, since every coercive quadratic continuous function in a Hilbert space has a unique minimum. To deduce the necessary condition for the minimum we write $W(x)$ in the form

$$W(x) = \langle \nu * (\pi_- x + \overline{\pi}_+ x) + f, \nu * (\pi_- x + \overline{\pi}_+ x) + f \rangle_{L^2(\mathbf{R}^+)},$$

---

[28] The convolution operator $\check{\nu}*$ is the adjoint of the convolution operator $\nu*$. To obtain the measure $\check{\nu}$ one takes matrix adjoints and reflects the time axis, i.e., $\check{\nu}(E) = \nu(-E)^*$ for every Borel set $E \subset \mathbf{R}$; here the star represents the operation of replacing a matrix by its adjoint (complex conjugate transpose). Clearly, $\check{\nu}$ is supported on $\overline{\mathbf{R}}^- = (-\infty, 0]$, and the convolution operator $\check{\nu}*\nu*$ is selfadjoint.

[29] This proof is almost identical to the proof of [25, Lemma 3.2].

and differentiate $W(x)$ with respect to $\overline{\pi}_+ x$ to get for each variation $\eta \in L^2(\mathbf{R}^+; \mathbf{R}^m)$,

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} W(x_{\text{opt}} + \epsilon\eta)\Big|_{\epsilon=0} = 2\langle \nu * x_{\text{opt}} + f, \nu * \eta\rangle_{L^2(\mathbf{R}^+)} = 2\langle y_{\text{opt}}, \nu * \eta\rangle_{L^2(\mathbf{R}^+)}$$
$$= 2\langle \check{\nu} * y_{\text{opt}}, \eta\rangle_{L^2(\mathbf{R}^+)}.$$

This is zero for all η iff

$$(32) \qquad\qquad\qquad (\check{\nu} * y_{\text{opt}})(t) = 0, \qquad t \in \overline{\mathbf{R}}^+.$$

By substituting (23) into (32) we get (30) and (31). $\square$

The key step in our earlier solution of the Wiener-Hopf equation (31) was to find an *invertible*[30] solution $\xi \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ of the equation

$$(33) \qquad\qquad\qquad\qquad \check{\xi} * \xi = \check{\nu} * \nu.$$

LEMMA 2.4.   *Let $\nu \in \mathcal{A}(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$, and suppose that ν is coercive in the sense that (29) holds for some $\eta > 0$. Then there is a measure $\xi \in \mathcal{A}(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ with a convolution inverse ξ^{-1} that also belongs to $\mathcal{A}(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ satisfying (33) in (at least) the two following cases:*

1. *$m = 1$ (the control is scalar);*
2. *the atoms of ν are integrally dependent in the sense that they are located at points that are integer multiples of one fixed time T (in particular, this is true in Case 3 below).*

Moreover,

3. *if ν is the sum of an atom at zero plus a function in $L^1 \cap L^p(\mathbf{R}^+; \mathbf{R}^{n \times m})$ for some $p \in [1, \infty]$, then ξ and ξ^{-1} are of the same type (except that the dimensions are $m \times m$).*

Case 1 of Lemma 2.4 is due to Gohberg and Fel'dman (see, e.g., [8, Chapter VII]), Case 3 is due to Gohberg and Kreĭn [9], and Case 2 is due to Winkin [38]. It is an interesting open problem whether Lemma 2.4 is true when $m > 1$ and there are multiple, integrally independent delays; we conjecture that this is the case. As is well-known and easy to see, the coercivity of ν is a necessary condition for the existence of a factorization.[31]

The measure ξ in Lemma 2.4 is usually called a spectral factor:

DEFINITION 2.5.   *A measure $\xi \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ is called a (canonical) spectral factor of the measure $\check{\nu} * \nu$ if ξ has a convolution inverse $\xi^{-1} \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ and ξ satisfies (33).*

Throughout this work we make the following hypothesis:

HYPOTHESIS 2.6.   *The measure $\check{\nu} * \nu$ has a spectral factor $\xi \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$.*

In particular, this hypothesis will be true if ν is coercive and satisfies the conditions in Parts 1, 2, or 3 of Lemma 2.4.

---

[30] The measure ν itself will hardly ever be invertible. It cannot be invertible unless $n = m$, which is not true if we start off with the cost function (3) and reformulate the problem as we did at the end of Section 1.

[31] Take $\eta = \|\xi^{-1}\|^{-1}$.

The invertibility in $M(\overline{\mathbf{R}}^+; \mathbf{R}^{m \times m})$ of the spectral factor ξ implies that ξ has an invertible atom at zero. It will become important in the sequel to separate this atom at zero from the rest of ξ. We shall need to repeat the same construction for other measures, as well. Therefore, let us introduce the following terminology:

DEFINITION 2.7. *Write $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$ in the form $\nu(\mathrm{d}s) = D_\nu \delta_0(\mathrm{d}s) + \nu_+(\mathrm{d}s)$, where δ_0 is the Dirac delta, D_ν is a matrix, and ν_+ has no atom at zero. Then $D_\nu \delta_0$ is called the (instantaneous) feed-through part of ν, D_ν is called the (instantaneous) feed-through matrix of ν, and ν_+ is called the strictly causal part of ν.*

Compare this definition to the formulae in (9). Observe that ν_+ is allowed to contain delayed atoms, but no atom at zero.

As we mentioned above, it will be important to separate ν, μ, and ξ into

$$
\begin{aligned}
\nu(\mathrm{d}s) &= D_\nu \delta_0(\mathrm{d}s) + \nu_+(\mathrm{d}s), \\
\mu(\mathrm{d}s) &= I \delta_0(\mathrm{d}s) + \mu_+(\mathrm{d}s), \\
\xi(\mathrm{d}s) &= D_\xi \delta_0(\mathrm{d}s) + \xi_+(\mathrm{d}s).
\end{aligned}
\tag{34}
$$

Observe that we have normalized the feed-through matrix D_μ of μ to be the identity matrix. The feed-through matrix D_ξ is always invertible. As is well-known (see, e.g., [26, Lemma 3.5]), ξ is determined uniquely, modulo the multiplication from the left by a unitary matrix M. This ambiguity can be used to force the feed-through matrix D_ξ to be positive definite.[32] With this normalization, the spectral factor ξ is unique.

The coercivity condition (29) is not standard, in fact, in virtually all standard papers this condition is absent. This is due to the fact it is customary to work with the cost function J_R in (3) with a strictly positive definite R instead of our J in (2), and in this case the analogue of (29) is always satisfied.[33] One drawback with (29) is that this condition is phrased in terms of the numerator ν in the right coprime factorization $\gamma = \nu * \mu^{-1}$ of the impulse response γ, and not directly in terms of γ itself. However, this condition is equivalent to the corresponding condition on γ:

LEMMA 2.8. *Let ν and μ be right coprime, and extend the domain of definition of the Laplace transform $\hat{\gamma}$ of $\gamma = \nu * \mu^{-1}$ to all those $s \in \mathbf{C}$ with $\Re s \geq 0$ for which $\det \hat{\mu}(s) \neq 0$ by defining $\hat{\gamma}(s) = \hat{\nu}(s)(\hat{\mu}(s))^{-1}$. Then ν satisfies (29) if and only if the extended function $\hat{\gamma}$ satisfies*

$$
(\hat{\gamma}(\mathrm{i}\omega))^* \hat{\gamma}(\mathrm{i}\omega) \geq \epsilon^2 I, \quad \omega \in \mathbf{R}, \quad \det \hat{\mu}(\mathrm{i}\omega) \neq 0,
\tag{35}
$$

for some $\epsilon > 0$.

Proof. Suppose that ν satisfies (29). Since $\|\hat{\mu}(\mathrm{i}\omega)\| \leq \|\mu\|$ for all $\omega \in \mathbf{R}$, get from (29)

$$
(\hat{\nu}(\mathrm{i}\omega))^* \hat{\nu}(\mathrm{i}\omega) \geq \frac{\eta^2}{\|\mu\|^2} (\hat{\mu}(\mathrm{i}\omega))^* \hat{\mu}(\mathrm{i}\omega), \quad \omega \in \mathbf{R}.
$$

---

[32] Factor D_ξ into $D_\xi = MR$, where $M = D_\xi (D_\xi^* D_\xi)^{-1/2}$ is unitary and $R = (D_\xi^* D_\xi)^{1/2}$ is positive definite, and multiply ξ by M^{-1}.

[33] Use the fact that ν and μ are right coprime, choose some ν_1 and μ_1 as in Definition 2.1, and define $\eta = \left(\sup_{\omega \in \mathbf{R}} \|(\hat{\nu}_1(\mathrm{i}\omega) \quad \hat{\mu}_1(\mathrm{i}\omega) R^{-1/2})\| \right)^{-1}$.

Multiplying this formula by $(\hat{\mu}(i\omega))^{-1}$ from the right and by $((\hat{\mu}(i\omega))^*)^{-1}$ from the left we get (35) with $\epsilon = \eta/\|\mu\|$.

Conversely, suppose that (35) holds. Multiply (35) by $\hat{\mu}(i\omega)$ from the right and by $(\hat{\mu}(i\omega))^*$ from the left to get

$$(\hat{\nu}(i\omega))^*\hat{\nu}(i\omega) \geq \epsilon^2(\hat{\mu}(i\omega))^*\hat{\mu}(i\omega), \quad \omega \in \mathbf{R}, \quad \det\hat{\mu}(i\omega) \neq 0.$$

The functions on both sides of this inequality are continuous in ω, and $\det\hat{\mu}(i\omega)$ can vanish at most of a set of measure zero (it is the boundary function of a nontrivial H^∞-function), so the extra condition $\det\hat{\mu}(i\omega) \neq 0$ can be removed, leading to

$$(36) \qquad (\hat{\nu}(i\omega))^*\hat{\nu}(i\omega) \geq \epsilon^2(\hat{\mu}(i\omega))^*\hat{\mu}(i\omega), \quad \omega \in \mathbf{R}.$$

Since ν and μ are coprime, we can choose some ν_1 and μ_1 satisfying the Bezout identity (28). Define $\epsilon_1 = (\sup_{\omega\in\mathbf{R}}\|(\,\hat{\nu}_1(i\omega) \quad \hat{\mu}_1(i\omega)\,)\|)^{-1}$. Then

$$(\hat{\nu}(i\omega))^*\hat{\nu}(i\omega) + (\hat{\mu}(i\omega))^*\hat{\mu}(i\omega) \geq \epsilon_1^2 I, \quad \omega \in \mathbf{R}.$$

Multiplying this inequality by ϵ^2 and adding the result to (36) we get (29) with $\eta^2 = \epsilon_1^2\epsilon^2/(1+\epsilon^2)$.□

Applying this result to the system $\Sigma = (A, B, C, D)$ in (1) with ν and μ defined as in (9) we get the following corollary:

COROLLARY 2.9. *Let $\Sigma = (A, B, C, D)$ be an exponentially stabilizable and detectable system with bounded control and observation operators, and define ν and μ as in (9). Then the following three conditions are equivalent:*

 1. *The measure $\check{\nu} * \nu$ has a spectral factor ξ.*
 2. *Condition (29) holds for some $\eta > 0$.*
 3. *There is some $\epsilon > 0$ such that*

$$(37) \qquad (C(i\omega - A)^{-1}B + D)^*(C(i\omega - A)^{-1}B + D) \geq \epsilon^2 I$$

for all those $\omega \in \mathbf{R}$ for which $i\omega$ belongs to the resolvent set of A.

When these conditions hold, the spectral factor ξ is the sum of an invertible atom D_ξ at zero and a function $X \in L^1 \cap L^\infty(\mathbf{R}^+; \mathbf{R}^{m\times m})$.

We leave the easy proof of this corollary to the reader.

DEFINITION 2.10. *The system Σ in Corollary 2.9 is called coercive whenever the condition in Part 3 of that corollary is true.*

Observe that the extended system Σ_R that we get by replacing C and D by C_R and D_R defined in (27) is always coercive whenever R is strictly positive definite.

The rest of this section is devoted to a closer study of the system Σ in Corollary 2.9 under the assumption that Σ is coercive. From the definition (9) of ν we get after a short computation

$$(\check{\nu} * \nu)(ds) = D^*D\delta_0(ds) + Z(s)ds,$$

where Z is the function defined by

$$(38) \qquad Z(t) = \begin{cases} \left(D^*C^F + B^*\Gamma_{C^F}\right)S^F(t)B, & t > 0, \\ Z^*(-t) & t < 0, \end{cases}$$

and where in turn

$$(39) \qquad \Gamma_{CF} = \int_0^\infty (S^F)^*(s)(C^F)^* C^F S^F(s) \mathrm{d}s$$

is the observability Gramian of the stabilized system. By the coercivity of Σ and Corollary 2.9, a spectral factor ξ exists, and the strictly causal part of ξ is a function $X \in L^1 \cap L^\infty(\mathbf{R}^+; \mathbf{R}^{m \times m})$. The normalized positive definite feed-through matrix of ξ must be equal to the square root of the corresponding atom of $\check\nu * \nu$. Thus, for the system Σ in (1),

$$(40) \qquad \xi(\mathrm{d}s) = (D^*D)^{1/2} \delta_0(\mathrm{d}s) + X(s)\mathrm{d}s.$$

The inverse of ξ will be of the type

$$(41) \qquad \xi^{-1}(\mathrm{d}s) = (D^*D)^{-1/2} \delta_0(\mathrm{d}s) + Y(s)\mathrm{d}s,$$

where $Y \in L^1 \cap L^\infty(\mathbf{R}^+; \mathbf{R}^{m \times m})$. That ξ and ξ^{-1} are convolution inverses of each other means that the functions X and Y are interrelated through the formula

$$(42) \qquad \begin{aligned} (D^*D)^{1/2}Y &+ X (D^*D)^{-1/2} + X * Y \\ &= Y (D^*D)^{1/2} + (D^*D)^{-1/2} X + Y * X = 0. \end{aligned}$$

On the half-line $\mathbf{R}^+$ the equation $\xi = \check\xi^{-1} * \check\nu * \nu$ becomes, in terms of the functions X, Y, and Z,

$$(43) \qquad \begin{aligned} X(t) &= (D^*D)^{-1/2} Z(t) + \int_0^\infty Y^*(s)Z(t+s)\mathrm{d}s \\ &= (D^*D)^{-1/2} \left(D^*C^F + B^*\Gamma_{CF} \right) S^F(t)B \\ &\quad + \int_0^\infty Y^*(s) \left(D^*C^F + B^*\Gamma_{CF} \right) S^F(t+s)B\mathrm{d}s \\ &= QS^F(t)B, \quad t \in \mathbf{R}^+, \end{aligned}$$

where

$$(44) \qquad \begin{aligned} Q &= (D^*D)^{-1/2} \left(D^*C^F + B^*\Gamma_{CF} \right) \\ &\quad + \int_0^\infty Y^*(s) \left(D^*C^F + B^*\Gamma_{CF} \right) S^F(s)\mathrm{d}s. \end{aligned}$$

As we shall see later, the operator Q is closely related to the feedback operator K, that expresses the optimal control u_{opt} for the closed loop system as a function of the present state of the system.

3. The Basic Solution. With the aid of the spectral factorization in Lemma 2.4 we are able to solve (31):

THEOREM 3.1. *Let $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$, let Hypothesis 2.6 hold, and let $\mathcal{B}$ be any one of the spaces (see the list of notations in Section 9)*

$$(45) \qquad L^p \text{ with } p \in [1, \infty], \ B^\infty, \ B_0^\infty, \ \text{or } BC_{\mathrm{right}}.$$

Then for each $\varphi \in \mathcal{B}(\mathbf{R}^-; \mathbf{R}^m)$ and $f \in \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$, equation (31) has a unique solution $x_{\mathrm{opt}} \in \mathcal{B}(\mathbf{R}; \mathbf{R}^m)$, given by

$$(46) \qquad \begin{aligned} x_{\mathrm{opt}} &= \xi^{-1} * (\pi_-(\xi * \varphi) - \overline{\pi}_+(\check{\xi}^{-1} * \check{\nu} * f)) \\ &= \pi_- \varphi - \xi^{-1} * \overline{\pi}_+(\xi * \pi_- \varphi + \check{\xi}^{-1} * \check{\nu} * f). \end{aligned}$$

If, in addition, $\varphi \in \mathcal{B}(\overline{\mathbf{R}}^-; \mathbf{R}^m)$ and $f \in \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$, where $\mathcal{B}$ is any one of the spaces

$$(47) \qquad\qquad\qquad BC, \; BUC, \; or \; BUC_0,$$

and if

$$(48) \qquad\qquad\qquad (\xi * \varphi)(0) = -(\check{\xi}^{-1} * \check{\nu} * f)(0),$$

then $x_{\mathrm{opt}} \in \mathcal{B}(\mathbf{R}; \mathbf{R}^m)$. If $\varphi \in \mathcal{B}(\mathbf{R}^-; \mathbf{R}^m)$ and $f \in \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$, where $\mathcal{B}$ is any one of the spaces listed in (47), and if we instead of (48) suppose that ν has no singular part (apart from a feed-through part), then $x_{\mathrm{opt}} \in \mathcal{B}(\mathbf{R}^-; \mathbf{R}^m) \cap \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$ (but x_{opt} may be discontinuous at zero). In all cases, the output y_{opt} defined in (30) belongs to $\mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$, and if furthermore $g \in \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$, then the function u_{opt} defined in (30) belongs to $\mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$.

Proof. [34] We solve (31) by arguing in the following way. We replace $\check{\nu} * \nu$ in (31) by $\check{\xi} * \xi$, and then convolve the second equation by $\check{\xi}^{-1}$ (which is supported on $\overline{\mathbf{R}}^-$) to get

$$(49) \qquad \begin{aligned} x_{\mathrm{opt}}(t) &= \varphi(t), & t \in \mathbf{R}^-, \\ (\xi * x_{\mathrm{opt}})(t) &= -(\check{\xi}^{-1} * \check{\nu} * f)(t), & t \in \overline{\mathbf{R}}^+. \end{aligned}$$

This is a delay equation for x_{opt}, with an initial function φ and a forcing term $-\check{\xi}^{-1} * \check{\nu} * f$. To solve this equation we argue as follows. We observe that $\xi * x_{\mathrm{opt}} = \xi * \varphi$ on $\mathbf{R}^-$. Thus

$$\xi * x_{\mathrm{opt}} = \pi_-(\xi * \varphi) - \overline{\pi}_+(\check{\xi}^{-1} * \check{\nu} * f),$$

where the equation is valid on all of $\mathbf{R}$. Convolve this equation by ξ^{-1} to get

$$(50) \qquad x_{\mathrm{opt}} = \xi^{-1} * (\pi_-(\xi * \varphi) - \overline{\pi}_+(\check{\xi}^{-1} * \check{\nu} * f)).$$

Taking into account that $\pi_- = I - \overline{\pi}_+$, and that $\pi_-(\xi * \varphi) = \pi_-(\xi * \pi_- \varphi)$, we get (46). All the additional claims follow from (50) and standard properties of convolution operators.   □

By separating ξ and ν into their feed-through parts and strictly causal parts we may alternatively write formula (46) as[35]

$$(51) \qquad \begin{aligned} x_{\mathrm{opt}} &= \pi_- \varphi - \xi^{-1} * \overline{\pi}_+ \left(\xi_+ \quad \check{\xi}^{-1} * \check{\nu} \right) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix} \\ &= \pi_- \varphi - \xi^{-1} * \overline{\pi}_+ \check{\xi}^{-1} * \check{\nu} * \left(\nu_+ \quad \delta_0 \right) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix}, \end{aligned}$$

---

[34] This proof is almost identical to the proof of [25, Theorem 3.3].

[35] Since the Hankel operator induced by δ_0 is zero, we have $\overline{\pi}_+(\xi * \pi_- \varphi) = \overline{\pi}_+(\xi_+ * \pi_- \varphi)$ and $\overline{\pi}_+(\nu * \pi_- \varphi) = \overline{\pi}_+(\nu_+ * \pi_- \varphi)$.

where the second version is derived from the first version and the fact that

$$
\begin{aligned}
\overline{\pi}_+(\check{\xi}^{-1} * \check{\nu} * \nu_+ * \pi_-\varphi) &= \overline{\pi}_+(\check{\xi}^{-1} * \check{\nu} * \nu * \pi_-\varphi) \\
&= \overline{\pi}_+(\xi * \pi_-\varphi) \\
&= \overline{\pi}_+(\xi_+ * \pi_-\varphi).
\end{aligned}
$$

(52)

4. The Feedback Operator. In order to compute the operator K_Φ defined (24) it suffices to evaluate (30) at zero, with x_{opt} given by (51). According to Theorem 3.1, if φ, f, and g are bounded and right-continuous, then the solution x_{opt} of (31) is right-continuous, and so is u_{opt}. By (30), (34) and (51)[36]

$$
\begin{aligned}
K_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= u_{\mathrm{opt}}(0) = x_{\mathrm{opt}}(0) + (\mu_+ * \varphi)(0) + g(0) \\
&= (\, \mu_+ - D_\xi^{-1}\xi_+ \quad -D_\xi^{-1}\check{\xi}^{-1} * \check{\nu} \quad \delta_0 \,) * \begin{pmatrix} \pi_-\varphi \\ f \\ g \end{pmatrix}(0). \\
&= (\, \mu_+ - D_\xi^{-1}\check{\xi}^{-1} * \check{\nu} * \nu_+ \quad -D_\xi^{-1}\check{\xi}^{-1} * \check{\nu} \quad \delta_0 \,) * \begin{pmatrix} \pi_-\varphi \\ f \\ g \end{pmatrix}(0).
\end{aligned}
$$

(53)

Thus, we have the following result:

PROPOSITION 4.1. *The operator K_Φ defined in (53) is a continuous linear operator from $BC_{\mathrm{right}}(\mathbf{R}^-; \mathbf{R}^m) \times BC_{\mathrm{right}}(\overline{\mathbf{R}}^+; \mathbf{R}^n) \times BC_{\mathrm{right}}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$ into $\mathbf{R}^n$.*

Observe that, because of the presence of the term $g(0)$, this operator is always unbounded on $L^2(\mathbf{R}^-; \mathbf{R}^m) \times L^2(\mathbf{R}^+; \mathbf{R}^n) \times L^2(\mathbf{R}^+; \mathbf{R}^m)$.

In the case where y is the output of the system $\Sigma = (A, B, C, D)$ in (1) with bounded control and observation operators, we can make the following computation.[37] We first recall that in this case $\varphi = 0$, hence the terms involving φ drop out. Moreover, $g(t) = FS^F(t)z_0$, hence $g(0) = Fz_0$. After a short computation, similar to the one leading to (38), we find that

$$
(\check{\nu} * f)(t) = (D^*C_F + B^*\Gamma_{C_F})S^F(t)z_0, \qquad t \in \mathbf{R}^+,
$$

where Γ_{C_F} is the observability Gramian defined in (39). Moreover, another equally short computation shows that

$$
(\check{\xi}^{-1} * \check{\nu} * f)(t) = QS^F(t)z_0, \qquad t \in \mathbf{R}^+,
$$

where Q is the operator defined in (44). In particular, taking $t = 0$ we get

$$
(\check{\xi}^{-1} * \check{\nu} * f)(0) = Qz_0.
$$

---

[36] The feed-through part of ξ^{-1} is $D_\xi^{-1}\delta_0$, and the strictly causal part of ξ^{-1} does not affect the value of (51) at zero.

[37] In the stable case this reduces to the corresponding computation in [25].

Substituting these values into (53), we find that the operator K defined in (15) is given by

$$K = F - (D^*D)^{-1/2}\, Q.$$

Thus,

$$(54) \qquad (\check{\xi}^{-1} * \check{\nu} * f)(t) = (D^*D)^{1/2}\, (F - K)S^F(t)z_0, \qquad t \in (0, \infty),$$

and (43) becomes

$$(55) \qquad X(t) = (D^*D)^{1/2}\, (F - K)S^F(t)B, \quad t \in (0, \infty).$$

If we rewrite (44) in terms of the operator K, then we get

$$
(56) \qquad
\begin{aligned}
Kz_0 = {} & Fz_0 - (D^*D)^{-1}\left(D^*C^F + B^*\Gamma_{C^F}\right) z_0 \\
& - (D^*D)^{-1/2}\int_0^\infty Y^*(s)\left(D^*C^F + B^*\Gamma_{C^F}\right) S^F(s)z_0\, \mathrm{d}s.
\end{aligned}
$$

The argument above shows that if u_{opt} is obtained via state feedback, i.e., if $u_{\mathrm{opt}}(t) = Kz(t)$ for $t \geq 0$, then K must be the operator defined above. It does not yet prove that $u_{\mathrm{opt}}(t) = Kz(t)$ for $t \geq 0$. However, by (40), (49), and (54),

$$(D^*D)^{1/2}\, x_{\mathrm{opt}} + X * x_{\mathrm{opt}} = (D^*D)^{1/2}\, (K - F)S^F(t)z_0,$$

which combined with (6) and (55) gives $x_{\mathrm{opt}} = (K - F)z_{\mathrm{opt}}$. This, together with (4) gives $u_{\mathrm{opt}} = Kz_{\mathrm{opt}}$. Thus, we have proved the following theorem:

THEOREM 4.2. *Let $\Sigma = (A, B, C, D)$ be a coercive exponentially stabilizable and detectable system with bounded control and observation operators. Then there is a unique u_{opt} that minimizes the cost function (2), where y is the output of the system (1). Moreover, the optimal u_{opt} can be written in feedback form $u_{\mathrm{opt}} = Kz_{\mathrm{opt}}$, where K is the operator given by (56).*

5. The Riccati Operator. Let us next turn to the Riccati operator Π_Φ defined in (25). The computation of this operator is very easy. It suffices to observe that we have rewritten the constrained minimization problem with respect to u as a free minimization problem with respect to x, hence the cost of the function g must be zero, and the cost of f and φ must be the same as in the stable case, if we replace the measure γ in [25] by ν. For the convenience of the reader, we repeat the derivation given in [25, Section 5] here. Define

$$(57) \qquad h = \overline{\pi}_+\left(\nu_+ * \pi_-\varphi + f\right) = \overline{\pi}_+\begin{pmatrix}\nu_+ & \delta_0\end{pmatrix} * \begin{pmatrix}\pi_-\varphi \\ f\end{pmatrix}.$$

Then the first two equations in (23) become

$$(58) \qquad y(t) = (\nu * \overline{\pi}_+ x)(t) + h(t), \qquad t \in \overline{\mathbf{R}}^+,$$

and the cost function $J(u)$ can be written in the form

$$J(u) = \langle y, y \rangle_{L^2(\mathbf{R}^+)} = \langle \nu * \overline{\pi}_+ x + h, \nu * \overline{\pi}_+ x + h \rangle_{L^2(\mathbf{R}^+)}$$
$$= \langle \nu * \overline{\pi}_+ x, \nu * \overline{\pi}_+ x + h \rangle_{L^2(\mathbf{R}^+)} + \langle h, \nu * \overline{\pi}_+ x + h \rangle_{L^2(\mathbf{R}^+)}.$$

However, because of (32) and (58), we have for the optimal x_{opt},

$$\langle \nu * \overline{\pi}_+ x_{\mathrm{opt}}, \nu * \overline{\pi}_+ x_{\mathrm{opt}} + h \rangle_{L^2(\mathbf{R}^+)} = \langle x_{\mathrm{opt}}, \check{\nu} * (\nu * \overline{\pi}_+ x_{\mathrm{opt}} + h) \rangle_{L^2(\mathbf{R}^+)} = 0.$$

Thus (recall also (33)),

$$J(x_{\mathrm{opt}}) = \langle h, h \rangle_{L^2(\mathbf{R}^+)} + \langle \check{\nu} * h, x_{\mathrm{opt}} \rangle_{L^2(\mathbf{R}^+)}$$
$$= \langle h, h \rangle_{L^2(\mathbf{R}^+)} - \langle \check{\nu} * \nu * \overline{\pi}_+ x_{\mathrm{opt}}, x_{\mathrm{opt}} \rangle_{L^2(\mathbf{R}^+)}$$
$$= \langle h, h \rangle_{L^2(\mathbf{R}^+)} - \langle \check{\xi} * \xi * \overline{\pi}_+ x_{\mathrm{opt}}, x_{\mathrm{opt}} \rangle_{L^2(\mathbf{R}^+)}$$
$$= \langle h, h \rangle_{L^2(\mathbf{R}^+)} - \langle \xi * \overline{\pi}_+ x_{\mathrm{opt}}, \xi * \overline{\pi}_+ x_{\mathrm{opt}} \rangle_{L^2(\mathbf{R}^+)}.$$

If we here substitute x_{opt} from (51) and h from (57), then we get

$$
\begin{aligned}
(59) \qquad J(x_{\mathrm{opt}}) &= \left\| (\nu_+ \quad \delta_0) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix} \right\|_{L^2(\mathbf{R}^+)}^2 - \left\| (\xi_+ \quad \check{\xi}^{-1} * \check{\nu}) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix} \right\|_{L^2(\mathbf{R}^+)}^2 \\
&= \left\| (\nu_+ \quad \delta_0) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix} \right\|_{L^2(\mathbf{R}^+)}^2 - \left\| \check{\xi}^{-1} * \check{\nu} * (\nu_+ \quad \delta_0) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix} \right\|_{L^2(\mathbf{R}^+)}^2 .
\end{aligned}
$$

From (59) we see that the Riccati operator Π_Φ can be written in the forms[38]

$$
\begin{aligned}
(60) \qquad \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= \begin{pmatrix} \pi_- \check{\nu}_+ \\ \delta_0 \\ 0 \end{pmatrix} * \overline{\pi}_+ (\nu_+ \quad \delta_0 \quad 0) * \begin{pmatrix} \pi_- \varphi \\ f \\ g \end{pmatrix} \\
&\quad - \begin{pmatrix} \pi_- \check{\xi}_+ \\ \nu * \xi^{-1} \\ 0 \end{pmatrix} * \overline{\pi}_+ (\xi_+ \quad \check{\xi}^{-1} * \check{\nu} \quad 0) * \begin{pmatrix} \pi_- \varphi \\ f \\ g \end{pmatrix} \\
&= \begin{pmatrix} \pi_- \check{\nu}_+ \\ \delta_0 \\ 0 \end{pmatrix} * \overline{\pi}_+ (\nu_+ \quad \delta_0 \quad 0) * \begin{pmatrix} \pi_- \varphi \\ f \\ g \end{pmatrix} \\
&\quad - \begin{pmatrix} \pi_- \check{\nu}_+ \\ \delta_0 \\ 0 \end{pmatrix} * \nu * \xi^{-1} * \overline{\pi}_+ \check{\xi}^{-1} * \check{\nu} * (\nu_+ \quad \delta_0 \quad 0) * \begin{pmatrix} \pi_- \varphi \\ f \\ g \end{pmatrix},
\end{aligned}
$$

and we have the following theorem:

THEOREM 5.1. *Let $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$ satisfy Hypothesis 2.6. Then the Riccati operator Π_Φ defined in (25) can be expressed in the different forms (60). In particular, it maps $\mathcal{B}(\mathbf{R}^-; \mathbf{R}^m) \times \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^n) \times \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$ continuously into $\mathcal{B}(\mathbf{R}^-; \mathbf{R}^m) \times \mathcal{B}(\overline{\mathbf{R}}^+; \mathbf{R}^n) \times \{ 0 \}$, where $\mathcal{B}$ is any one of the spaces listed in (45).*

---

[38] Compare this to [25, formula (42)].

Moreover, if ν has no singular part (apart from a feed-through part), then Π_Φ maps $\mathcal{B}(\mathbf{R}^-;\mathbf{R}^m) \times \mathcal{B}(\overline{\mathbf{R}}^+;\mathbf{R}^n) \times \mathcal{B}(\overline{\mathbf{R}}^+;\mathbf{R}^m)$ continuously into $\mathcal{B}(\mathbf{R}^-;\mathbf{R}^m) \times \mathcal{B}(\overline{\mathbf{R}}^+;\mathbf{R}^n) \times \{0\}$, where $\mathcal{B}$ is any one of the spaces listed in (47).

Observe that formulae (33) and (34) in [25] are still valid as well. For example, if we let P be the "restriction operator" that maps $\begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}$ into f, then in the particular case where φ is absent, the projected Riccati operator $\Pi_\Phi^P = P\Pi_\Phi P^*$ can be written in the form

$$(61) \qquad \Pi_\Phi^P f = P\Pi_\Phi P^* f = f - \nu * \xi^{-1} * \overline{\pi}_+ \check{\xi}^{-1} * \check{\nu} * f.$$

Of course, this operator has the same type of properties as those listed in Theorem 5.1 for the operator Π_Φ.

In the case of the system Σ in (1) we define the Riccati operator Π through the formula

$$J(u) = \langle z_0, \Pi z_0 \rangle_H,$$

where H is the state space of Σ, and get from (9), (39), (54), and (61)

$$
\begin{aligned}
(62) \qquad \Pi z_0 &= \int_0^\infty S^{F^*}(s)\left((C^F)^* C^F - (K-F)^* D^* D(K-F)\right) S^F(s) z_0 \mathrm{d}s, \\
&= \Gamma_{C^F} z_0 - \int_0^\infty S^{F^*}(s)(K-F)^* D^* D(K-F) S^F(s) z_0 \mathrm{d}s.
\end{aligned}
$$

In particular, from this equation we can conclude that Π satisfies a Lyapunov equation, arguing as follows.[39] Take z_0 and z_1 in $\mathcal{D}(A)$, observe that for all $t \geq 0$,

$$
\int_0^\infty \langle S^F(s+t)z_1, \left((C^F)^* C^F - (K-F)^* D^* D(K-F)\right) S^F(s+t)z_0 \rangle_H \mathrm{d}s
$$
$$
= \int_t^\infty \langle S^F(s)z_1, \left((C^F)^* C^F - (K-F)^* D^* D(K-F)\right) S^F(s)z_0 \rangle_H \mathrm{d}s,
$$

differentiate both sides of this equation with respect to t, take $t = 0$ and use (62) to get the Lyapunov equation

$$
\begin{aligned}
(63) \qquad \langle (A+BF)z_1, \Pi z_0 \rangle_H &+ \langle z_1, \Pi(A+BF)z_0 \rangle_H + \langle C^F z_1, C^F z_0 \rangle_{\mathbf{R}^n} \\
&= \langle (K-F)z_1, D^* D(K-F)z_0 \rangle_{\mathbf{R}^m}, \\
& z_0, \ z_1 \in \mathcal{D}(A).
\end{aligned}
$$

To derive (20) we compute,[40] using (62), (55), (56), a change of integration variable,

---

[39] This argument reduces to the corresponding one in [25, Section 5] in the stable case.

[40] This argument, too, reduces to the corresponding one in [25, Section 5] in the stable case. To get the computation given here from the one in [25, Section 5] it suffices to replace z_0, C, K, and S in [25] by z_0, C^F, $K-F$, and S^F, respectively.

(42), and once again (56)

$$
\begin{aligned}
B^*\Pi z_0 &= B^*\Gamma_{C^F} z_0 - \int_0^\infty B^* S^{F*}(s)(K-F)^*\,(D^*D)\,(K-F)S^F(s)z_0\,\mathrm{d}s \\
&= B^*\Gamma_{C^F} z_0 + \int_0^\infty X^*(s)\,(D^*D)^{1/2}\,(K-F)S^F(s)z_0\,\mathrm{d}s \\
&= B^*\Gamma_{C^F} z_0 - \int_0^\infty X^*(s)\,(D^*D)^{-1/2}\left(D^*C^F + B^*\Gamma_{C^F}\right)z_0\,\mathrm{d}s \\
&\quad - \int_0^\infty \int_0^\infty X^*(s)Y^*(t)\left(D^*C^F + B^*\Gamma_{C^F}\right)S^F(s+t)z_0\,\mathrm{d}s\mathrm{d}t \\
&= B^*\Gamma_{C^F} z_0 - \int_0^\infty X^*(s)\,(D^*D)^{-1/2}\left(D^*C^F + B^*\Gamma_{C^F}\right)z_0\,\mathrm{d}s \\
&\quad - \int_0^\infty \int_0^v X^*(s)Y^*(v-s)\mathrm{d}s\left(D^*C^F + B^*\Gamma_{C^F}\right)S^F(v)z_0\,\mathrm{d}v \\
&= B^*\Gamma_{C^F} z_0 - \int_0^\infty \left(X^*(v)\,(D^*D)^{-1/2} + (X^* * Y^*)(v)\right)\left(D^*C^F + B^*\Gamma_{C^F}\right)S^F(v)z_0\,\mathrm{d}v \\
&= B^*\Gamma_{C^F} z_0 + (D^*D)^{1/2}\int_0^\infty Y^*(v)\left(D^*C^F + B^*\Gamma_{C^F}\right)S^F(v)z_0\,\mathrm{d}v \\
&= B^*\Gamma_{C^F} z_0 - (D^*D)\,(K-F)z_0 - \left(D^*C^F + B^*\Gamma_{C^F}\right)z_0 \\
&= -(D^*D)\,(K-F)z_0 - D^*C^F z_0.
\end{aligned}
$$

Thus, we find that K is connected to Π through the formula

$$
(64) \qquad\qquad K = F - (D^*D)^{-1}\left(B^*\Pi + D^*C^F\right).
$$

If we here replace C^F by its definition $C^F = C + DF$, then we get the formula (20) (which is independent of F). Combining (63) and (64) we get

$$
(65) \quad
\begin{aligned}
\langle(A+BF)z_1, \Pi z_0\rangle_H &+ \langle z_1, \Pi(A+BF)z_0\rangle_H + \langle C^F z_1, C^F z_0\rangle_{\mathbf{R}^n} \\
&= \left\langle (B^*\Pi + D^*C^F)z_1, (D^*D)^{-1}(B^*\Pi + D^*C^F)z_0\right\rangle_{\mathbf{R}^m}, \\
&\quad z_0,\ z_1 \in \mathcal{D}(A).
\end{aligned}
$$

If we again replace C^F by its definition $C^F = C + DF$, then we get (after a lengthy but straightforward algebraic manipulation) the standard Riccati equation (22) for the Riccati operator Π. Observe that the original preliminary feedback F has disappeared from this equation, too (as it must do, since the Riccati operator does not depend on F). We can use (20) to replace $(D^*D)^{-1}(B^*\Pi + D^*C)$ in this equation by $-K$, and get the standard F-independent version (21) of (63).

Thus, we have proved the following theorem:

THEOREM 5.2. *Under the assumptions of Theorem 4.2, the feedback operator K in that theorem is connected to the Riccati operator Π through the formula (20), and the Riccati operator Π is a positive self-adjoint solution of the Riccati equation (22).*

Of course, this theorem is well-known, but we believe our proof to be new.

6. A Realization Based on a Left Shift Semigroup. Continuing on the same lines as in [25, Section 6], we shall next give a particular realization Φ of the input/output relation (23). This realization is an extension of the realizations Ξ developed in [25, Section 6]. The presentation is rather brief, and we refer the reader to [13, Chapter 8], [22], [23], [24], [25], [35], and [34] for additional reading.

The basis for our construction is the delay equation (23), where we consider u, f, and g to be given, x to be the unknown solution, and y to be the output that depends on the control u and on the solution x. According to [24, Theorem 1.2], this equation is well-posed if and only if μ has an invertible atom at zero, and as we have already mentioned several times, we shall normalize this atom to be the identity matrix.

The basic semigroup S_Φ around which we build Φ is the standard combined initial and forcing function semigroup in $\mathcal{I} \times \mathcal{G} = L^2(\mathbf{R}^-; \mathbf{R}^m) \times L^2(\mathbf{R}^+; \mathbf{R}^m)$ induced by the equation[41]

$$(66) \qquad \begin{aligned} (\mu * x)(t) &= -g(t), \quad t \in \mathbf{R}^+, \\ x(t) &= \varphi(t), \qquad t \in \mathbf{R}^-. \end{aligned}$$

The idea is to solve (66), to shift x and g to the left by an equal amount, and to restrict the shifted x to $\mathbf{R}^-$ and the shifted g to $\mathbf{R}^+$ to get a new initial function and new forcing function for the same equation. This semigroup is combined with the left-shift semigroup in $\mathcal{F} = L^2(\mathbf{R}^+; \mathbf{R}^n)$ that shifts the function f to the left, and restricts the shifted function to $\mathbf{R}^+$. Thus, the action of S_Φ can be described by

$$S_\Phi(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = \begin{pmatrix} \tau_t \pi_- x \\ \overline{\pi}_+ \tau_t f \\ \overline{\pi}_+ \tau_t g \end{pmatrix},$$

where x is the solution of (66). The generator of this semigroup is the operator A_Φ that differentiates each of the functions $(\varphi, f, g) \in \mathcal{I} \times \mathcal{F} \times \mathcal{G}$, with domain

$$\mathcal{D}(A_\Phi) = \left\{ \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in \mathcal{I}_1 \times \mathcal{F}_1 \times \mathcal{G}_1 \;\middle|\; \mu * \varphi(0) = -g(0) \right\};$$

here $\mathcal{I}_1 = W^{1,2}(\overline{\mathbf{R}}^-; \mathbf{R}^m)$, $\mathcal{F}_1 = W^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$, and $\mathcal{G}_1 = W^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$. According to (23), for nonzero u, the controlled state z_Φ at time $t > 0$ should be the same as the uncontrolled state, except that $-g(s)$ has been replaced by $u(s) - g(s)$ for $0 \leq s < t$. Thus, the controlled state at time $t > 0$ is given by

$$z_\Phi(t, \varphi, f, g, u) = \begin{pmatrix} \tau_t \pi_- x \\ \overline{\pi}_+ \tau_t f \\ \overline{\pi}_+ \tau_t g \end{pmatrix},$$

where x is the solution of (23). In order to compute the control operator we remark that the first equation in (23) can be written in the form (see also (34)),

$$(67) \qquad x(t) = u(t) - g(t) - (\mu_+ * x)(t), \quad t \in \mathbf{R}^+.$$

----

[41] This semigroup is described in, e.g., [24, Theorem 3.1].

This means that the appropriate control operator is the boundary input[42]

$$
(68) \qquad\qquad B_\Phi u = \begin{pmatrix} \delta_0 \\ 0 \\ 0 \end{pmatrix} u,
$$

where, as usual, δ_0 is the Dirac delta. Note that the control has no influence on the second and third components of the state, only on the x-component.[43]

In our definition of the feed-through operator D and the observation operator C_Φ we follow the same conventions as in [25], i.e., we let D represent the instantaneous feed-through, and incorporate all delayed terms into C_Φ. Since the second equation in (23) can be written in the form

$$
\begin{aligned}
y(t) &= D_\nu x(t) + (\nu_+ * x)(t) + f(t), \\
&= D_\nu(u(t) - g(t) - (\mu_+ * x)(t)) + (\nu_+ * x)(t) + f(t),
\end{aligned}
$$

we find that the appropriate feed-through and observation operators[44] are

$$
(69) \qquad\qquad D = D_\nu,
$$

$$
\begin{aligned}
(70) \qquad C_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= D(-g(0) - (\mu_+ * \varphi)(0)) + (\nu_+ * \varphi)(0) + f(0) \\
&= \begin{pmatrix} \nu_+ - D\mu_+ & \delta_0 & -D\delta_0 \end{pmatrix} * \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}(0).
\end{aligned}
$$

Let us choose u to be the optimal control u_{opt} given by Lemma 2.3 and Theorem 3.1. Then the evolution of Φ is described by the delay equation (49). As in [25], we find that if we define S_Φ^K by

$$
S_\Phi^K(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = \begin{pmatrix} \pi_- \tau_t x_{\mathrm{opt}} \\ \overline{\pi}_+ \tau_t f \\ \overline{\pi}_+ \tau_t g \end{pmatrix},
$$

then S_Φ^K is a strongly continuous semigroup on $\mathcal{I} \times \mathcal{F} \times \mathcal{G}$. The generator A_Φ^K of S_Φ^K is again the differentiation operator, this time with domain

$$
\mathcal{D}(A_\Phi^K) = \left\{ \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in \mathcal{I}_1 \times \mathcal{F}_1 \times \mathcal{G}_1 \;\middle|\; (\xi * \varphi)(0) + (\check{\xi}^{-1} * \check{\nu} * f)(0) = 0 \right\}.
$$

---

[42] This is a typical example of a boundary control system in the sense of [21], since the control operator is strictly unbounded (the range of B_Φ lies in the space of measures, and its intersection with $\mathcal{I} \times \mathcal{F} \times \mathcal{G}$ is the zero function).

[43] The last two components of the state are completely uncontrollable. It is possible to construct another realization where these components have been removed; see the realization Ξ^- in [25, Section 6].

[44] Due to the presence of the term $f(0)$, the observation operator C_Φ is always strictly unbounded on $\mathcal{I} \times \mathcal{F} \times \mathcal{G}$. This means that Φ is at the same time both a boundary control process and a point observation process in the sense of Salamon [21].

The free evolution of S_Φ^K is exactly the same as the controlled evolution of the system Φ with control u_{opt}. According to (30) and (34),

$$u_{\mathrm{opt}} = \mu_+ * x_{\mathrm{opt}} + x_{\mathrm{opt}} + g,$$

which combined with (49) and (34) gives (for right-continuous data)

$$
\begin{aligned}
u_{\mathrm{opt}}(t) &= (\mu_+ * x_{\mathrm{opt}})(t) - D_\xi^{-1}(\xi_+ * x_{\mathrm{opt}})(t) - D_\xi^{-1}(\check{\xi}^{-1} * \check{\nu} * f)(t) + g(t) \\
&= (\,\mu_+ - D_\xi^{-1}\xi_+ \quad -D_\xi^{-1}\check{\xi}^{-1} * \check{\nu} \quad \delta_0\,) * \begin{pmatrix} \pi_- \tau_t x_{\mathrm{opt}} \\ \overline{\pi}_+ \tau_t f \\ \overline{\pi}_+ \tau_t g \end{pmatrix}(0), \\
&= K_\Phi \begin{pmatrix} \pi_- \tau_t x_{\mathrm{opt}} \\ \overline{\pi}_+ \tau_t f \\ \overline{\pi}_+ \tau_t g \end{pmatrix} = K_\Phi z_\Phi(t, \varphi, f, g, u_{\mathrm{opt}}),
\end{aligned}
$$

where K_Φ is the operator defined in (53). But this is nothing but a state feedback.[45] Thus, we find that for the system Φ, the optimal control is given by a state feedback, with the feedback operator K_Φ.

The derivation of (20) and (21) that we gave earlier was based on a preliminary stabilizing feedback. In the next section we shall derive similar formulae for the realization Φ using a similar method, i.e., we first stabilize the system, and then optimize. In order to make the formulae as simple as possible, we choose the stabilizing feedback operator F_Φ to be the operator

$$(71) \qquad F_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = (\,\mu_+ \quad 0 \quad \delta_0\,) * \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}(0) = (\mu_+ * \varphi)(0) + g(0).$$

In other words, we replace the control u by the new control

$$v = u - F_\Phi z_\Phi = u - \mu_+ * x - g.$$

Equation (23) then simplifies into

$$
\begin{aligned}
x(t) &= \varphi(t), & t &\in \mathbf{R}^-, \\
y(t) &= (\nu * x)(t) + f(t), & t &\in \overline{\mathbf{R}}^+, \\
x(t) &= v(t), & t &\in \overline{\mathbf{R}}^+.
\end{aligned}
$$

The stabilized semigroup S_Φ^F that we get in this way is the standard left-shift on $\mathcal{I} \times \mathcal{F} \times \mathcal{G}$ of the initial function φ and the forcing functions f and g, and its generator A_Φ^F

---

[45] An operator K_Φ is called an admissible state feedback operator for a regular abstract linear system $\Phi = (A_\Phi, B_\Phi, C_\Phi, D_\Phi)$ if both the system that one gets from Φ by replacing C_Φ and D_Φ by $\begin{pmatrix} C_\Phi \\ K_\Phi \end{pmatrix}$ and $\begin{pmatrix} D_\Phi \\ 0 \end{pmatrix}$, and the system that one gets by closing the feedback loop from the output corresponding to K_Φ into the input are regular abstract linear systems. See [5, Conditions (C1) and (C2)].

is the differentiation operator on $W_0^{1,2}(\overline{\mathbf{R}}^-;\mathbf{R}^m) \times W^{1,2}(\overline{\mathbf{R}}^+;\mathbf{R}^n) \times W^{1,2}(\overline{\mathbf{R}}^+;\mathbf{R}^m)$.[46] Thus,

$$
(72) \qquad S_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = \begin{pmatrix} \tau_t \pi_- \varphi \\ \overline{\pi}_+ \tau_t f \\ \overline{\pi}_+ \tau_t g \end{pmatrix}.
$$

This is the same semigroup S_Ξ that we used as one of the basic unperturbed semigroups in [25, Section 5], except for the additional g-component that has been completely decoupled from the rest of the system. By (70) and (71),

$$
(73) \qquad C_\Phi^F \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = (C_\Phi + DF_\Phi) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = \begin{pmatrix} \nu_+ & \delta_0 \end{pmatrix} * \begin{pmatrix} \varphi \\ f \end{pmatrix}(0),
$$

and by (53) and (71), the operator $K_\Phi - F_\Phi$ is given by

$$
(74) \qquad (K_\Phi - F_\Phi) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = -D_\xi^{-1} \begin{pmatrix} \xi_+ & \check{\xi}^{-1} * \check{\nu} \end{pmatrix} * \begin{pmatrix} \varphi \\ f \end{pmatrix}(0).
$$

In particular, we observe that both the present control operator B_Φ given by (68) and the present observation operator C_Φ^F in (73) are the same as the operators B_Ξ in [25, Formula (54)] and C_Ξ in [25, Formula (55)].[47] Thus, we can apply the theory developed in [25, Section 7] to the feedback stabilized system Φ^F. See the next section for details.

For later reference, let us remark that domains of the generators A_Φ, A_Φ^K, and A_Φ^F can be written in the forms

$$
(75) \qquad
\begin{aligned}
\mathcal{D}(A_\Phi) &= \left\{ \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in \mathcal{I}_1 \times \mathcal{F}_1 \times \mathcal{G}_1 \;\middle|\; \varphi(0) = -F_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \right\}, \\
\mathcal{D}(A_\Phi^K) &= \left\{ \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in \mathcal{I}_1 \times \mathcal{F}_1 \times \mathcal{G}_1 \;\middle|\; \varphi(0) = (K_\Phi - F_\Phi) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \right\}, \\
\mathcal{D}(A_\Phi^F) &= \left\{ \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in \mathcal{I}_1 \times \mathcal{F}_1 \times \mathcal{G}_1 \;\middle|\; \varphi(0) = 0 \right\}.
\end{aligned}
$$

By combining the argument above with the theory of regular abstract linear systems[48] and general properties of convolution operators[49] we can summarize the findings of this section into the following theorem:

---

[46] This system is not exponentially stable, only weakly asymptotically stable. However, as in [25], it is possible to base the presentation on a shift in an exponentially weighted L^2-space instead, and in this setting it is possible to make Φ^F exponentially stable.

[47] Replace γ in [25] by ν, and ignore the extra decoupled g-component.

[48] See, for example, [21], [22], [35], and [34].

[49] See, for example [13, Theorem 6.1, pp. 96–96] and [25, Lemma 3.1].

THEOREM 6.1. *The system Φ is a regular abstract linear system with impulse response $\gamma = \nu * \mu^{-1}$. It is well-posed in the L^2-setting described above, but it is also well-posed if we throughout replace L^2 by L^p, $1 \le p < \infty$. The evolution of the system (23) with u replaced by u_{opt} corresponds to the evolution of the feedback perturbed version S_Φ^K of S_Φ, with the admissible feedback operator K_Φ. If we denote the space of bounded right-continuous functions in the state space by $\mathcal{W}$, then this spaces contains the domain of the generators of both the original semigroup and the perturbed semigroup, and it is invariant under both these semigroups. Moreover, both the observation operator C_Φ and the feedback operator K_Φ are continuous on $\mathcal{W}$. The ranges of the adjoints of the observation and feedback operators are contained in the space of measures on $\overline{\mathbf{R}}^+$ and $\overline{\mathbf{R}}^-$. Also the operator F_Φ is an admissible feedback operator, and it has the same properties that we listed above for the feedback operator K_Φ.*

7. The Equations Connecting K and Π. In Section 5 we obtained two standard equations (20) and (21) that connect K and Π. We want to extend these equations to the case of a general input/output relation of the type (23), with the state space representation given in Section 6. As we already mentioned above, we shall proceed in the same way as before, and start with a preliminary stabilizing feedback. The feedback operator that we use is the operator F_Φ defined in (71). This leads to the stabilized system Φ^F, to which we can apply the theory in [25, Section 7].

We begin with the analogue of formula (21). Let $\begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in \mathcal{W}$. For each $t \ge 0$, we may apply both C_Φ^F and $K_\Phi - F_\Phi$ to $S_\Phi^F(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}$, and by (72), (73), and (74),

$$D_\xi (K_\Phi - F_\Phi) S_\Phi^F(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = -\,(\, \xi_+ \quad \check{\xi}^{-1} * \check{\gamma} \,) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix}(t),$$

$$C_\Phi^F S_\Phi^F(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} = (\, \nu_+ \quad \delta_0 \,) * \begin{pmatrix} \pi_- \varphi \\ f \end{pmatrix}(t).$$

Comparing this to (60) we find that

$$(76) \quad \left\langle \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \right\rangle$$
$$= \int_0^\infty \left| C_\Phi^F S_\Phi^F(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \right|^2 \mathrm{d}t - \int_0^\infty \left| D_\xi (K_\Phi - F_\Phi) S_\Phi^F(t) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \right|^2 \mathrm{d}t.$$

The same argument that brings us from (62) to (63) can be used to turn this equation

into an analogue of (63), namely

$$
\left\langle A_\Phi^F \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, \Pi_\Phi \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathcal{I} \times \mathcal{F} \times \mathcal{G}} + \left\langle \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, \Pi_\Phi A_\Phi^F \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathcal{I} \times \mathcal{F} \times \mathcal{G}}
$$

$$
+ \left\langle C_\Phi^F \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, C_\Phi^F \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathbf{R}^n}
$$

(77)

$$
= \left\langle (K_\Phi - F_\Phi) \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, D_\xi^* D_\xi (K_\Phi - F_\Phi) \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathbf{R}^m},
$$

$$
\begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix}, \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix} \in \mathcal{D}(A_\Phi^F).
$$

This equation is identical to (63) if ν contains no discrete delays, because then $D^* D = D_\xi^* D_\xi = D_\nu^* D_\nu$.

Let us next turn to the extension of (20) to the realization Φ. As usual in these situations, the interpretation of the term $B_\Phi^* \Pi_\Phi$ poses the greatest problem. The operator B_Φ^* is the observation operator of the adjoint $(\Phi^F)^* = ((A_\Phi^F)^*, (C_\Phi^F)^*, B_\Phi^*, D^*)$ of the system Φ^F. This is a system which is analogous to Φ^F; the main change is that the time direction has been reversed. The adjoint $(S_\Phi^F)^*$ of the semigroup S_Φ^F is the right-shift on $\mathcal{I} \times \mathcal{F} \times \mathcal{G}$, and the domain of its generator is $\mathcal{D}((A_\Phi^F)^*) = W^{1,2}(\overline{\mathbf{R}}^-; \mathbf{R}^m) \times W_0^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^n) \times W_0^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$. By (68), for $\begin{pmatrix} \varphi^* \\ f^* \\ g^* \end{pmatrix} \in \mathcal{D}((A_\Phi^F)^*)$, we have $B_\Phi^* \begin{pmatrix} \varphi^* \\ f^* \\ g^* \end{pmatrix} = \varphi^*(0)$. However, we want to apply B_Φ^* to $\Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}$, and, because of the projections operators in (60), Π_Φ does not map $\mathcal{D}(A_\Phi^F)$ into $\mathcal{D}((A_\Phi^F)^*)$ (unless $\nu_+ \in L^2(\mathbf{R}^+; \mathbf{R}^{n \times m})$). In particular, the first component of $\Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}$ in (60) is not continuous at zero, even if we require φ and f to be bounded and continuous (and it does not help to add smoothness assumptions on φ and f). As a matter of fact, we have written (60) in such a way that a formal evaluation of the first component in (60) at zero gives a plain zero. Since the adjoint semigroup is a right-shift instead of a left-shift, the obvious solution is to replace π_- by $\overline{\pi}_-$ and $\overline{\pi}_+$ by π_+ in order to make all the components of (60) left-continuous instead of right-continuous. An equivalent solution is to redefine B_Φ^* to be given by the left-hand limit

(78)
$$
B_\Phi^* \begin{pmatrix} \varphi^* \\ f^* \\ g^* \end{pmatrix} = \varphi^*(0-),
$$

whenever this limit exists. This is the way that we proceeded in [25, Section 7], and

this is the way that we shall proceed here, too.[50]

By repeating the computation leading to [25, Equation (60)][51] we get

$$
\begin{aligned}
B_\Phi^* \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= D_\xi^* \left(\xi_+ \quad \check{\xi}^{-1} * \check{\nu} \right) * \begin{pmatrix} \varphi \\ f \end{pmatrix}(0) \\
&\quad - D_\nu^* \left(\nu_+ \quad \delta_0 \right) * \begin{pmatrix} \varphi \\ f \end{pmatrix}(0) \\
&\quad + (D_\xi^* D_\xi - D_\nu^* D_\nu)\varphi(0) \\
&= -D_\xi^* D_\xi (K_\Phi - F_\Phi - B_\Phi^*) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \\
&\quad - D^*(C_\Phi^F + D B_\Phi^*) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}.
\end{aligned}
\tag{79}
$$

In the computation leading to this equation we require φ, f and g to be just bounded and continuous. If we strengthen this assumption, and require the data to belong to the domain of the generators of any one of the semigroups S_Φ, S_Φ^F, or S_Φ^K, then (79) simplifies into (see (75))

$$
\begin{aligned}
B_\Phi^* \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= - \left(D_\xi^* D_\xi K_\Phi + D^* C_\Phi \right) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, & \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in D(A_\Phi), \\
B_\Phi^* \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= - \left(D_\xi^* D_\xi (K_\Phi - F_\Phi) + D^* C_\Phi^F \right) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, & \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in D(A_\Phi^F), \\
B_\Phi^* \Pi_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= - D^*(C_\Phi + D K_\Phi) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, & \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in D(A_\Phi^K).
\end{aligned}
\tag{80}
$$

From the first two of these equations we may solve K_Φ to get

$$
\begin{aligned}
K_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= -(D_\xi^* D_\xi)^{-1} \left(B_\Phi^* \Pi_\Phi + D^* C_\Phi \right) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, & \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in D(A_\Phi), \\
K_\Phi \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} &= F_\Phi - (D_\xi^* D_\xi)^{-1} \left(B_\Phi^* \Pi_\Phi + D^* C_\Phi^F \right) \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix}, & \begin{pmatrix} \varphi \\ f \\ g \end{pmatrix} \in D(A_\Phi^F).
\end{aligned}
\tag{81}
$$

Observe how these equations agree with [25, equations (61) and (62)] on one hand, and with equation (64) on the other hand. Of course, the latter formula is the one

---

[50] The simplicity of this formula is the main reason why we introduced the extra initial function φ in the first place. It is possible to eliminate φ from the realization, but that leads to a much more complicated formula. The same comment applies to the formulae for the domains of the generators given in (75).

[51] We leave the verification of this computation to the reader.

that we should use to eliminate K_Φ from (77). Doing so we get the Riccati equation

$$
\left\langle A_\Phi^F \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, \Pi_\Phi \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathcal{I} \times \mathcal{F} \times \mathcal{G}} + \left\langle \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, \Pi_\Phi A_\Phi^F \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathcal{I} \times \mathcal{F} \times \mathcal{G}}
$$

$$
+ \left\langle C_\Phi^F \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, C_\Phi^F \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathbf{R}^n}
$$

(82)
$$
= \left\langle \left(B_\Phi^* \Pi_\Phi + D^* C_\Phi^F \right) \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix}, (D_\xi^* D_\xi)^{-1} \left(B_\Phi^* \Pi_\Phi + D^* C_\Phi^F \right) \begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix} \right\rangle_{\mathbf{R}^m},
$$

$$
\begin{pmatrix} \varphi_0 \\ f_0 \\ g_0 \end{pmatrix}, \begin{pmatrix} \varphi_1 \\ f_1 \\ g_1 \end{pmatrix} \in \mathcal{D}(A_\Phi^F).
$$

This formula is the same as (65), with the appropriate substitutions.

Thus, we have the following theorem:

THEOREM 7.1. *The Riccati operator Π_Φ and the feedback operator K_Φ satisfy equations (81) and (82). These equations are the same as the standard equations (64) and (65) valid for the system Σ with bounded control and observation operators, except for the fact that the matrix $D^* D$ has been replaced throughout by $D_\xi^* D_\xi$, where D_ξ is the feed-through matrix of the spectral factor ξ of $\check{\nu} * \nu$. If ν has no singular part (apart from a feed-through part), then $D_\xi^* D_\xi = D^* D$.*

8. Two Examples. Below we present two examples to which our theory applies. Both of these are fairly simple in the sense that it is possible to find an exact analytic solution, and the results that we derive are more or less known from before.[52] To simplify the comparison with earlier work we here use the cost function J_R defined in (3) instead of the function J in (2).

We begin by considering the example (1) with the cost function J_R defined in (3) and with one additional input delay.[53] The defining equations become in this case

(83)
$$
\begin{aligned}
z'(t) &= Az(t) + Bu(t - T), \\
y(t) &= Cz(t) + Du(t - T), \quad t \in [0, \infty), \\
z(0) &= z_0, \\
u(t) &= \psi(t), \quad t \in [-T, 0),
\end{aligned}
$$

where $T > 0$ is a constant delay, $\psi \in L^2((-T, 0); \mathbf{R}^m)$, and the rest of the setting is the same as in (1). This example has been studied extensively in different connections, see, e.g., [5], [31], and the references mentioned therein. This system has a simple realization $\Psi = (A_\Psi, B_\Psi, C_\Psi, D_\Psi)$ as a regular linear system. It is the serial connection of a delay line (a left-shift on $(-T, 0)$ of u, with a control operator δ_0 similar to the one

---

[52]We have not seen the full solution of the first example in print before, and we present a more complete physical interpretation of the second example than Grabowski does in [11].

[53]An output delay is even simpler to handle.

in (68) and an observation operator $u(-T)$), connected in series with the system Σ in (1). We refer the reader to [5, Example 4.1] for a closer description of this realization. For our purposes it suffices to know that its state space H_Ψ is $L^2((-T, 0); \mathbf{R}^m) \times H$, its input space is $U = \mathbf{R}^m$, its output space is $Y = \mathbf{R}^n$, and that the generating operators are

$$\mathcal{D}(A_\Psi) = \{ \psi \in W^{1,2}([-T, 0]; \mathbf{R}^m) \mid \psi(0) = 0 \} \times H,$$

(84)
$$A_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} = \begin{pmatrix} \psi' \\ B\psi(-T) + Az_0 \end{pmatrix}, \quad B_\Psi u = \begin{pmatrix} \delta_0 u \\ 0 \end{pmatrix},$$

$$C_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} = D\psi(-T) + Cz_0, \qquad D_\Psi = 0.$$

It is possible to derive a system of the type (23) in the following way. We start by choosing some stabilizing state feedback operator F for the system (1), i.e., we choose an operator $F \in \mathcal{L}(H; U)$ such that $A + BF$ generates an exponentially stable semigroup S^F. For the moment it does not really matter how this operator F is chosen, but looking ahead, it will simplify the final formulae if we choose F to be the optimal feedback operator K for the system (1) with the cost function (3). In other words, we take $F = K$, where K is given by (56) with the replacements listed in (27). This time we define the auxiliary variable x by

(85)
$$x(t) = u(t) - Kz(t + T), \quad t \geq -T.$$

Then (83) can be rewritten in the form

(86)
$$\begin{aligned}
z'(t) &= (A + BK)z(t) + Bx(t - T), \quad t \in [0, \infty), \\
y(t) &= (C + DK)z(t) + Dx(t - T), \quad t \in [0, \infty), \\
u(t) &= Kz(t + T) + x(t), \qquad\qquad t \in [-T, \infty), \\
z(0) &= z_0, \\
x(t) &= \varphi(t), \quad t \in [-T, 0),
\end{aligned}$$

provided φ is chosen as explained below. The restriction of the function z to the interval $[0, T]$ depends only on z_0 and the initial functions ψ and φ, and the further evolution of the two systems depends only on $z(T)$ and the restrictions of x and u to $[0, \infty)$. Solving z from (83) we get

(87)
$$z(t) = S(t)z_0 + \int_{-T}^{t-T} S(t - T - s)B\psi(s)\mathrm{d}s, \qquad t \in [0, T].$$

If we, instead, solve z from (86), then we get

(88)
$$z(t) = S^K(t)z_0 + \int_{-T}^{t-T} S^K(t - T - s)B\varphi(s)\mathrm{d}s, \qquad t \in [0, T].$$

By using either (87) or (88) in (85) we find that the systems (83) and (86) become identical if we choose ψ and φ to satisfy

(89)
$$\begin{aligned}
\varphi(t) &= \psi(t) - K\left(S(t + T)z_0 + \int_{-T}^{t} S(t - s)B\psi(s)\mathrm{d}s \right), \qquad t \in [-T, 0], \\
\psi(t) &= \varphi(t) + K\left(S^K(t + T)z_0 + \int_{-T}^{t} S^K(t - s)B\psi(s)\mathrm{d}s \right), \quad t \in [-T, 0].
\end{aligned}$$

In particular, taking $t = T$ in (87) and (88) we get

$$
\begin{aligned}
z(T) &= S(T)z_0 + \int_{-T}^{0} S(-s)B\psi(s)\mathrm{d}s \\
&= S^K(T)z_0 + \int_{-T}^{0} S^K(-s)B\varphi(s)\mathrm{d}s.
\end{aligned}
\tag{90}
$$

Whereas it was fairly obvious how to realize (83) as a regular linear system, it is less obvious how to realize (86) (the third equation is the difficult one.) However, this can be done. In order to fit this case into our earlier formalism we extend φ to all of $\mathbf{R}^-$ by defining $\varphi(t) = 0$ for $t < -T$. A computation similar to the one in the introduction shows that, by defining (with $C^K = C + DK$))

$$
\begin{aligned}
\nu(\mathrm{d}s) &= D\delta_0(\mathrm{d}s) + C^K S^K(s)B\mathrm{d}s, \\
\mu(\mathrm{d}s) &= I\delta_0(\mathrm{d}s) + KS^K(s)B\mathrm{d}s, \quad s \in \overline{\mathbf{R}}^+, \\
f(t) &= C^K S^K(t)z_0, \\
g(t) &= KS^K(t+T)z_0, \quad t \in \mathbf{R}^+, \\
\nu_T &= \delta_T * \nu, \\
\tilde{\nu} &= \begin{pmatrix} \nu_T \\ R^{1/2}\mu \end{pmatrix}, \quad \tilde{y} = \begin{pmatrix} y \\ R^{1/2}u \end{pmatrix}, \quad \tilde{f} = \begin{pmatrix} f \\ R^{1/2}g \end{pmatrix},
\end{aligned}
\tag{91}
$$

we can turn (86) into a system of the type (23), namely

$$
\begin{aligned}
x(t) &= \varphi(t), & t \in \mathbf{R}^-, \\
\tilde{y}(t) &= (\tilde{\nu} * x)(t) + \tilde{f}(t), & t \in \overline{\mathbf{R}}^+, \\
(\mu * x)(t) &= u(t) - g(t), & t \in \overline{\mathbf{R}}^+.
\end{aligned}
\tag{92}
$$

It is well-known that ν_T and μ are right coprime,[54] hence so are $\tilde{\nu}$ and μ. The objective is to minimize the L^2-norm of $\tilde{y}$ on $\mathbf{R}^+$. The only difference compared to (23) is that we have replaced y, ν, and f by $\tilde{y}$, $\tilde{\nu}$, and $\tilde{f}$, respectively. We shall denote the realization Φ developed in Section 6 with these replacements by $\widetilde{\Phi}$.

The crucial part in our solution was to find a spectral factor of the measure $\check{\tilde{\nu}} * \tilde{\nu}$. The Laplace transform of this measure, restricted to the imaginary axis (with $s = \mathrm{i}\omega$) is

$$
\begin{aligned}
\begin{pmatrix} \hat{\nu}_T(\mathrm{i}\omega)^* & \hat{\mu}(\mathrm{i}\omega)^* R^{1/2} \end{pmatrix} \begin{pmatrix} \hat{\nu}_T(\mathrm{i}\omega) \\ R^{1/2}\hat{\mu}(\mathrm{i}\omega) \end{pmatrix} &= \begin{pmatrix} \mathrm{e}^{\mathrm{i}\omega T}\hat{\nu}(\mathrm{i}\omega)^* & \hat{\mu}(\mathrm{i}\omega)^* R^{1/2} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\omega T}\hat{\nu}(\mathrm{i}\omega) \\ R^{1/2}\hat{\mu}(\mathrm{i}\omega) \end{pmatrix} \\
&= \hat{\nu}(\mathrm{i}\omega)^*\hat{\nu}(\mathrm{i}\omega) + \hat{\mu}(\mathrm{i}\omega)^* R\hat{\mu}(\mathrm{i}\omega).
\end{aligned}
$$

But this is exactly the same function that we have to factor in the case where there is no delay, and we are in the case (1) discussed earlier. For this factorization we may

---

[54] See, for example, [5, Example 4.1] or [31, p. 932].

use the formulae in Sections 3 and 4, with the replacements listed in (27) with $F = K$. In particular, we conclude that

$$D_\xi^* D_\xi = D^* D + R, \tag{93}$$

and from (55) we find that the strictly causal part ξ_+ of ξ is zero (since we chose $K = F$; this is where this particular choice of K pays off). Thus,

$$\xi = (D^* D + R)^{1/2} \delta_0, \qquad \xi^{-1} = (D^* D + R)^{-1/2} \delta_0.$$

By comparing (93) to $D_{\tilde\nu}^* D_{\tilde\nu}$ we find that

$$D_\xi^* D_\xi = D^* D + R \neq R = \begin{pmatrix} 0 & R^{1/2} \end{pmatrix} \begin{pmatrix} 0 \\ R^{1/2} \end{pmatrix} = D_{\tilde\nu}^* D_{\tilde\nu},$$

unless $D = 0$, and we have found an example where $D_\xi^* D_\xi \neq D_{\tilde\nu}^* D_{\tilde\nu}$.

The rest of the computations are straightforward, so let us only list the final results. It turns out that for the data in (91),

$$\pi_+ \check{\tilde\nu} * \tilde f = 0,$$

hence

$$
\begin{aligned}
x_{\mathrm{opt}}(t) &= 0, & t &\in [0, \infty), \\
u_{\mathrm{opt}} &= K z_{\mathrm{opt}}(t+T), & t &\in [0, \infty), \\
z_{\mathrm{opt}}(t) &= S^K(t-T) z(T), & t &\in [T, \infty), \\
z_{\mathrm{opt}}(t+T) &= S^K(T) z_{\mathrm{opt}}(t), & t &\in [T, \infty),
\end{aligned}
$$

where $z(T)$ is given by (90). By rewriting the optimal feedback operator $K_{\widetilde\Phi}$ in terms of the original data z_0 and ψ in (83) and calling the rewritten operator K_Ψ we get

$$K_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} = K_{\widetilde\Phi} \begin{pmatrix} \varphi \\ \tilde f \\ g \end{pmatrix} = K z(T), \tag{94}$$

or, equivalently,

$$
\begin{aligned}
K_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} &= K \left(S(T) z_0 + \int_0^T S(s) B \psi(-s) \, \mathrm{d}s \right) \\
&= K \left(S^K(T) z_0 + \int_0^T S^K(s) B \varphi(-s) \, \mathrm{d}s \right),
\end{aligned}
\tag{95}
$$

where φ is connected to ψ through the relation (89). Evidently, in spite of the unboundedness of $K_{\widetilde\Phi}$, the operator K_Ψ in (95) is bounded. Using the same type of arguments as we did in Section 6, we find that if we use this operator as a state feedback operator for the system Ψ, then we get a perturbed system $\Psi^K = (A_\Psi^K, B_\Psi, C_\Psi^K, D_\Psi)$, with

$$\mathcal{D}(A_\Psi^K) = \left\{ \begin{pmatrix} \psi \\ z_0 \end{pmatrix} \in W^{1,2}([-T, 0]; \mathbf{R}^m) \times H \;\middle|\; \psi(0) = K_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} \right\}. \tag{96}$$

The easiest way to describe the optimal cost Π_Ψ is to write it in the form

$$\left\langle \begin{pmatrix} \psi \\ z_0 \end{pmatrix}, \Pi_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} \right\rangle = \int_0^T \|y(t)\|^2 \mathrm{d}t + \langle z(T), \Pi z(T) \rangle,$$

where $y(t) = Cz(t) + D\psi(t-T) = C^K z(t) + D\varphi(t-T)$ and $z(t)$ is given by (87) and (88), and where Π is the Riccati operator for the non-delayed system Σ with cost function (3). A lengthy but straightforward computation shows that, for continuous initial functions ψ,

$$
\begin{aligned}
B^*\Pi_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} &= (B^*\Pi + D^*C)z(T) + D^*D\psi(0) \\
&= -(D^*D + R)Kz(T) + D^*D\psi(0) \\
&= -(D^*D + R)K_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} + D^*D\psi(0).
\end{aligned}
$$

For $\begin{pmatrix} \psi \\ z_0 \end{pmatrix} \in \mathcal{D}(A_\Psi)$ we have $\psi(0) = 0$, and for $\begin{pmatrix} \psi \\ z_0 \end{pmatrix} \in \mathcal{D}(A_\Psi^K)$ we have $\psi(0) = K_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix}$. Thus, we find that

$$
\begin{aligned}
B^*\Pi_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} &= -(D^*D + R)K_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix}, \quad & \begin{pmatrix} \psi \\ z_0 \end{pmatrix} \in \mathcal{D}(A_\Psi), \\
B^*\Pi_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix} &= -RK_\Psi \begin{pmatrix} \psi \\ z_0 \end{pmatrix}, & \begin{pmatrix} \psi \\ z_0 \end{pmatrix} \in \mathcal{D}(A_\Psi^K),
\end{aligned}
$$

which is in agreement with (80) since[55]

$$\begin{pmatrix} D_\Psi^* & R^{1/2} \end{pmatrix} \begin{pmatrix} D_\Psi \\ R^{1/2} \end{pmatrix} = R, \qquad \begin{pmatrix} D_\Psi^* & R^{1/2} \end{pmatrix} \begin{pmatrix} C_\Psi \\ 0 \end{pmatrix} = 0.$$

The formulae given above for the optimal feedback K_Ψ in (94) can be interpreted as shown in Figure 1. The compensator in the upper feedback loop is initiated at time $t = -T$ (or earlier) to zero, the switch on the left is in the middle position, and the initial function ψ enters during the time interval $[-T, 0]$ and initiates both the system Ψ and the compensator. At time $t = 0$ the system Σ that is part of Ψ is initiated to the value z_0, and the switch is moved to its top position, activating the compensator. From time $t = T$ on the signal produced by the proportional feedback in the lower feedback loop will be identical to the signal produced by the compensator in the upper feedback loop, so the switch can be moved to its bottom position at any instant after the time $t = T$ without influencing the output of the controlled system.

We observe that the initialization phase in Figure 1 can be rather critical. If Σ is unstable and T large, then errors tend to be greatly magnified during initialization due to the instability of the system. However, this is not just a fault of the particular connection drawn in Figure 1. The real problem is that, although the feedback operator K_Ψ is bounded, its norm will be very large if Σ is unstable and T is large
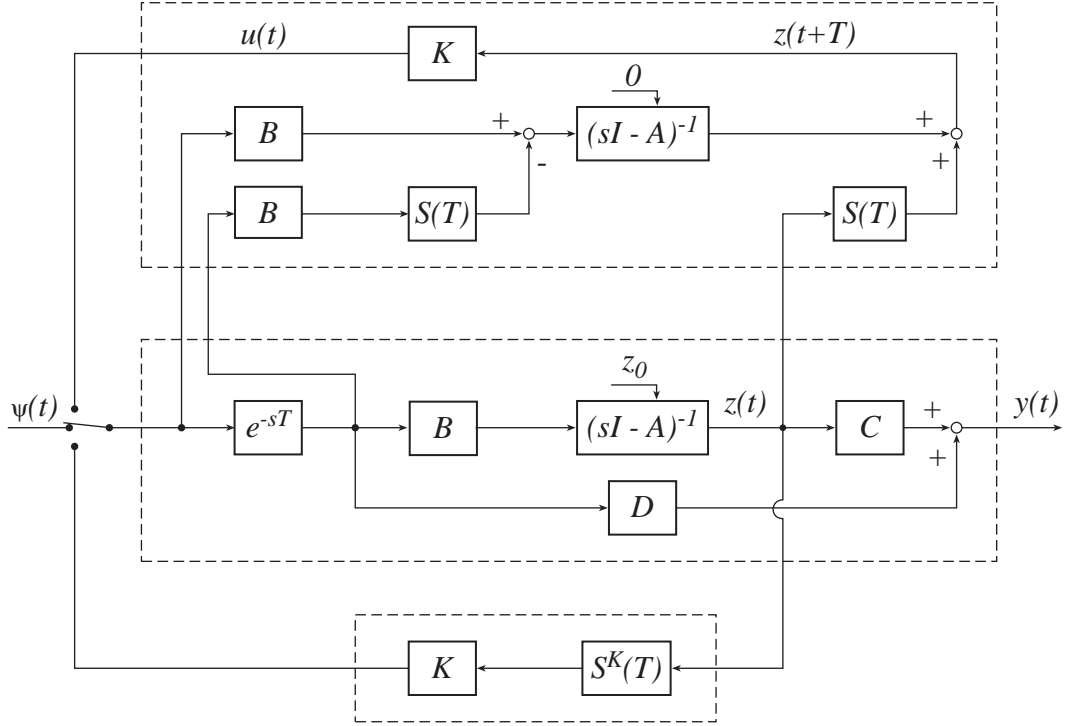
Fig. 1. *Optimal Compensator for System* Ψ

(because the norm of the mapping from the initial function ψ to $z(T)$ will be large), and the problem is simply numerically rather ill-posed.

Our second example is the same example that Grabowski presents in [11]. It is a controlled $RLCG$ transmission line of length one without distortion, i.e., $R/L = G/C$, driven by a control voltage u at one end, and loaded by a resistance R_1 at the opposite end.[56] If we let $i(x, t)$ represent the current and $v(x, t)$ the voltage of the line at the point $x \in [0, 1]$ at time $t \in [0, \infty)$, then these satisfy the equations

(97)
$$
\begin{aligned}
C\,\frac{\partial v(x, t)}{\partial t} &= -\frac{\partial i(x, t)}{\partial x} - G\,v(x, t), \\
L\,\frac{\partial i(x, t)}{\partial t} &= -\frac{\partial v(x, t)}{\partial x} - R\,i(x, t), \quad x \in [0, 1]; \quad t \in [0, \infty), \\
i(1, t)R_1 &= v(1, t), \\
v(0, t) &= u(t), \\
y(t) &= v(1, t), \qquad\qquad\qquad t \in [0, \infty).
\end{aligned}
$$

The objective is to minimize the cost function J_R in (3), with R replaced by one. The

---

[55] Recall formulae (27) and (84).

[56] The following derivation of the state equations is essentially the same as in [11]. When comparing our formulae with those in [11], one should replace our notations u, T, $e(t)$, Z, and β by Grabowski's w, r, $2x_2(t)$, z, and $1/\rho$, respectively. The physical explanation of the signal e that we present is not found in [11].

pair of d'Alembert solutions of (97) are

$$(98) \quad
\begin{aligned}
i(x,t) &= \frac{e^{-\alpha t}}{2Z} \left[\phi(x - t/T) - \psi(x + t/T) \right], \\
v(x,t) &= \frac{e^{-\alpha t}}{2} \left[\phi(x - t/T) + \psi(x + t/T) \right], \quad x \in [0,1]; \quad t \in [0, \infty),
\end{aligned}$$

where $\alpha = R/L = G/C$ is the decay rate, $T = \sqrt{LC}$ is the time that it takes a wave to travel from one end of the line to the other ($1/T$ is the wave speed), and $Z = \sqrt{L/C} = \sqrt{R/G}$ is the wave impedance of the line. Observe that $\phi(s)$ is defined for $s \leq 1$ and $\psi(s)$ is defined for $s \leq 0$. Substitute the boundary conditions in (97) into these equations to get

$$(99) \quad
\begin{aligned}
u(t) &= \frac{e^{-\alpha t}}{2} [\phi(-t/T) + \psi(t/T)], \\
\psi(1 + t/T) &= \kappa \phi(1 - t/T), \\
y(t) &= \frac{e^{-\alpha t}}{2} [\phi(1 - t/T) + \psi(1 + t/T)] \\
&= \frac{e^{-\alpha t}}{2} (1 + \kappa) \phi(1 - t/T), \qquad\qquad t \geq 0
\end{aligned}$$

where κ is the reflection coefficient at the output end, i.e., $\kappa = (R_1 - Z)/(R_1 + Z)$. We introduce the two new variables

$$(100) \quad
\begin{aligned}
e(t) &= e^{-\alpha t} \phi(-t/T), \quad t \geq -T, \\
w(t) &= e^{-\alpha t} \psi(t/T) \qquad t \geq 0.
\end{aligned}$$

Then

$$(101) \quad
\begin{aligned}
e(t) &= v(0,t) + Z i(0,t) = u(t) + Z i(0,t), \\
w(t) &= v(0,t) - Z i(0,t) = u(t) - Z i(0,t), \qquad t \geq 0,
\end{aligned}$$

hence

$$(102) \quad
\begin{aligned}
e(t) + w(t) &= 2u(t), \\
e(t) - w(t) &= 2Z i(0,t), \qquad t \geq 0.
\end{aligned}$$

It follows from the first equation in (99) and from (100) that $w(t) = \kappa \beta^2 e(t - 2T)$ for $t \geq T$, where $\beta = e^{-\alpha T}$ is the attenuation of the line. This identity can be extended to the interval $0 \leq t < T$ if we choose $e(t)$ in the interval $[-2T, T)$ to satisfy $\kappa e(t) = e^{-\alpha t} \psi(2 + t/T)$.[57] We denote the restriction of e to the interval $[-2T, 0]$ by ϵ, and arrive at the delay equation

$$(103) \quad
\begin{aligned}
e(t) &= \epsilon(t), & t &= [-2T, 0], \\
y(t) &= \tfrac{1}{2}(1 + \kappa)\beta e(t - T), & t &\in [0, \infty), \\
e(t) &= 2u(t) - w(t), & t &\in [0, \infty), \\
w(t) &= \kappa \beta^2 e(t - 2T), & t &\in [0, \infty).
\end{aligned}$$

---

[57]We get this identity for free if the original wave equation holds for all $t \geq -T$.

We extend ϵ to all of $\mathbf{R}^-$ by defining $\epsilon(t) = 0$ for $t < -2T$, and get an equation that is almost of the type (23); the only difference is the factor 2 in front of $u(t)$. To get rid of this factor we can, for example, define

$$(104) \qquad\qquad x = \tfrac{1}{2}e, \qquad \varphi = \tfrac{1}{2}\epsilon,$$

and get the system

$$(105) \qquad \begin{aligned} x(t) &= \varphi(t), & t &= (-\infty, 0], \\ y(t) &= (1+\kappa)\beta x(t-T), & t &\in [0, \infty), \\ x(t) &= u(t) - \kappa\beta^2 x(t-2T), & t &\in [0, \infty), \end{aligned}$$

which is a system of the type (23) with

$$(106) \qquad \begin{aligned} \nu &= (1+\kappa)\beta\delta_T, \\ \mu &= \delta_0 + \kappa\beta^2\delta_{2T}, \\ f &= g = 0. \end{aligned}$$

The system (103) is (exponentially) stable, unless $\kappa\beta^2 = \pm1$.[58] The trivial solution for the case $\kappa = -1$ is $u = 0$, so let us exclude this case from the following discussion, and in the sequel take $\kappa > -1$. Let us immediately observe that our basic assumption about the coprimeness of ν and μ is satisfied in the stable case, due to the fact that in this case $\mu^{-1} \in M(\mathbf{R}^+; \mathbf{R})$. It is also true that ν and μ are coprime in the case $\kappa = \beta = 1$, as can be easily seen (but not in the case $\kappa = -1$ and $\beta = 1$).

Recall that we this time use the cost function (3) with $R = 1$ instead of (2). Thus, the measure that should be factorized in not $\check{\nu} * \nu$ but $\check{\tilde{\nu}} * \tilde{\nu}$, where $\tilde{\nu} = \begin{pmatrix} \nu \\ \mu \end{pmatrix}$. A short computation shows that

$$|\hat{\tilde{\nu}}(\mathrm{i}\omega)|^2 = a + 2b\cos(2T\omega), \qquad \omega \in (-\infty, \infty),$$

where

$$a = \left(1 + (1+\kappa)^2\beta^2 + \kappa^2\beta^4\right), \qquad b = \kappa\beta^2.$$

Observe, in particular, that $a > 1 + |b|^2$, hence $a > 2|b|$. The spectral factor ξ of $\check{\tilde{\nu}} * \tilde{\nu}$ can be found by direct inspection: it suffices to take

$$\xi = p\delta_0 + q\delta_{2T},$$

with $p > |q|$ (in order to guarantee the invertibility of ξ), and to choose the coefficients p and q to satisfy

$$p^2 + q^2 = a, \qquad pq = b, \qquad p > |q|.$$

---

[58] This corresponds to the case where there is no internal damping in the line ($R = G = 0$ and $\beta = 1$) and the output end is either open ($R_1 = \infty$ and $\kappa = 1$) or shunted ($R_1 = 0$ and $\kappa = -1$).

This set of equations have a unique solution, namely

(107) $\qquad p = \frac{1}{2}\left(\sqrt{a+2b} + \sqrt{a-2b}\right), \qquad q = \frac{1}{2}\left(\sqrt{a+2b} - \sqrt{a-2b}\right).$

We remark that both p and $|q|$ are nonnegative solutions of the equation

$$p^2 + b^2/p^2 = a,$$

with $p > |q|$. For $p = 1$ the left hand side becomes $1 + b^2$, which is less that a. Thus,

(108) $\qquad\qquad\qquad\qquad p > 1, \qquad |q| < |b| = |\kappa|\beta^2 \le 1.$

When we apply the realization developed in Section 6 to (103) we can omit the forcing function components f and g, since these are zero all the time, and we have to replace ν by $\tilde{\nu}$ due to the different cost function (3). Let us denote the system that we get in this way by $\widetilde{\Phi}$. Then formulae (53), (68), (69), (70), (71), and (73) give [59]

$$B_{\widetilde{\Phi}} = \delta_0, \quad D_{\tilde{\nu}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad D_\xi = p, \qquad 1 = D_{\tilde{\nu}}^* D_{\tilde{\nu}} \ne D_\xi^* D_\xi = p^2,$$

(109) $C_{\widetilde{\Phi}}\varphi = \begin{pmatrix} (1+\kappa)\beta\varphi(-T) \\ 0 \end{pmatrix}, \quad F_{\widetilde{\Phi}}\varphi = \kappa\beta^2\varphi(-2T),$

$\qquad C_\Phi^F\varphi = \begin{pmatrix} (1+\kappa)\beta\varphi(-T) \\ \kappa\beta^2\varphi(-2T) \end{pmatrix}, \quad K_{\widetilde{\Phi}}\varphi = \left(\kappa\beta^2 - q/p\right)\varphi(-2T) = \dfrac{p^2-1}{p^2}F_{\widetilde{\Phi}}\varphi,$

From (60) we get after a short computation

(110) $\qquad\qquad\qquad \left(\Pi_{\widetilde{\Phi}}\varphi\right)(t) = \begin{cases} (p^2-1)\varphi(t), & \text{if } -T < t \le 0, \\ q^2(p^2-1)\varphi(t), & \text{if } -2T < t \le -T, \\ 0, & \text{otherwise.} \end{cases}$

Moreover, (75) gives

(111) $\qquad\qquad \begin{aligned} \varphi(0) &= -F_{\widetilde{\Phi}}\varphi = -\kappa\beta^2\varphi(-2T), & \varphi &\in \mathcal{D}(A_{\widetilde{\Phi}}), \\ \varphi(0) &= 0, & \varphi &\in \mathcal{D}(A_\Phi^F), \\ \varphi(0) &= (K_{\widetilde{\Phi}} - F_{\widetilde{\Phi}})\varphi = -(q/p)\varphi(-2T), & \varphi &\in \mathcal{D}(A_\Phi^K), \end{aligned}$

and (79) becomes

(112) $\qquad\qquad \begin{aligned} B_{\widetilde{\Phi}}^* \Pi_{\widetilde{\Phi}}\varphi &= -p^2 K_{\widetilde{\Phi}}\varphi, & \varphi &\in \mathcal{D}(A_{\widetilde{\Phi}}), \\ B_{\widetilde{\Phi}}^* \Pi_{\widetilde{\Phi}}\varphi &= 0, & \varphi &\in \mathcal{D}(A_\Phi^F), \\ B_{\widetilde{\Phi}}^* \Pi_{\widetilde{\Phi}}\varphi &= -K_{\widetilde{\Phi}}\varphi, & \varphi &\in \mathcal{D}(A_\Phi^K), \end{aligned}$

---

[59] We remark that our feedback operator differs from the one derived in [11], due to the fact that Grabowski uses a nonzero feed-through operator in the feedback loop. In the standard setting such a feed-through operator is never present. See the discussion in [26].
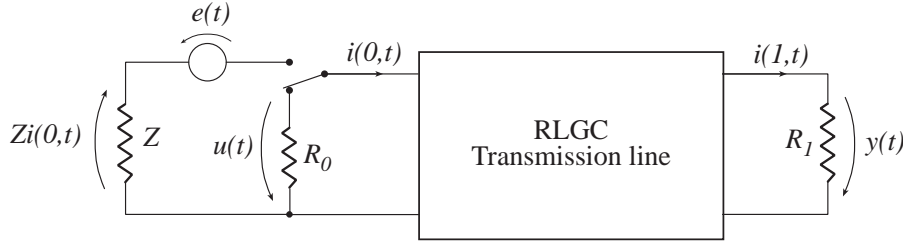
Fig. 2. *Optimal Controller for the Transmission Line*

The formulae that we have developed above can be interpreted as shown in Figure 2. To get a physical interpretation of the variable e it suffices to take a closer look at the first equation in (101). For $t < 0$, e can be interpreted as the voltage of a signal generator with internal resistance Z matching the wave impedance of the line used to transmit a signal into the line. During this stage the switch in Figure 2 is kept in its top position. Because of the matching terminating resistance at the left end of the line, there are no reflections at this end (although the signal may be reflected at the output end), and the output voltage y follows the input e with no distortion, but with the attenuation of $\frac{1}{2}(1 + \kappa)\beta$ and a time delay of T time units. At time zero the signal generator is switched off, and it is desired to minimize the spill-over energy $\left(\int_0^\infty \left(\|u(t)\|^2 + \|y(t)\|^2\right) \mathrm{d}t\right)^{1/2}$.

We can use (103), (104), and (109) to write the optimal control $u(t) = K_{\widetilde{\Phi}} x(t)$ in the form

$$u(t) = \frac{p^2 - 1}{2p^2} w(t) = \frac{p^2 - 1}{2p^2}(u(t) - Zi(0, t)),$$

from which we find that

$$(113) \qquad\qquad\qquad u(t) = -R_0 i(0, t),$$

where $R_0 = \frac{p^2 - 1}{p^2 + 1} Z$. Clearly, this is the formula for the voltage over a resistor of size R_0 carrying a current $i(0, t)$. Thus, the optimal controller is simply a terminating resistance of size R_0 at the left end, as drawn in Figure 2, where the switch is moved to its lower position at time $t = 0$. We remark that $R_0 < Z$, and that the reflection coefficient at the input end is $-1/p^2$. Note that the intuitive choice $R_0 = Z$, which leads to the extinction of the signals u and y in the finite time $2T$, is not the optimal one. It would have been optimal if we had not put any cost on the input voltage u; only on the output voltage y.

The optimal cost of the input signal is

$$(114) \qquad J(u_{\mathrm{opt}}) = \frac{1}{2}\left\{ q^2(p^2 - 1) \int_{-2T}^{T} \|e(t)\|^2 \mathrm{d}t + (p^2 - 1) \int_{-T}^{0} \|e(t)\|^2 \mathrm{d}t \right\}^{1/2}.$$

There is an obvious explanation for the different weights of the two time intervals $(-2T, T)$ and $(-T, 0)$: at time $t = 0$ the signal that entered the transmission line

during the time interval $(-2T, T)$ has been reflected and travels to the left, whereas the signal that entered during the time interval $(-T, 0)$ is still traveling to the right. The part of the signal that entered before time $t = -2T$ is no longer present in the system, so the cost of this part is zero. If $\kappa = 0$ (hence $q = 0$), then there is no reflection at the output end of the line, and the cost of the part of e that entered during the time interval $(-2T, -T)$ is zero, too.

Let us finally remark that, although the structure of the optimal solution is very simple, the formula (107) for the crucial number p is nontrivial. In particular, this formula could not have been deduced from the standard continuous time Riccati equation theory.

9. Notations. Above we have used the following notations:

$\mathbf{R}$, $\mathbf{R}^+$, $\overline{\mathbf{R}}^+$, $\mathbf{R}^-$, $\overline{\mathbf{R}}^-$: $\mathbf{R} = (-\infty, \infty)$, $\mathbf{R}^+ = (0, \infty)$, $\overline{\mathbf{R}}^+ = [0, \infty)$, $\mathbf{R}^- = (-\infty, 0)$, and $\overline{\mathbf{R}}^- = (-\infty, 0]$.

$\mathcal{L}(U)$, $\mathcal{L}(U; V)$: The Banach space of bounded linear operators mapping U into itself or U into V, respectively, with the operator norm.

$L^p(J; \mathbf{R}^n)$: The Banach space of $\mathbf{R}^n$-valued L^p-functions on the interval J, with $1 \leq p \leq \infty$, with the usual norm.

$B^\infty(J; \mathbf{R}^n)$: The Banach space of $\mathbf{R}^n$-valued bounded Borel measurable functions on the interval J, with the sup-norm.

$B_0^\infty(J; \mathbf{R}^n)$: The subspace of $B^\infty(J; \mathbf{R}^n)$ of functions tending to zero at infinity.

$BC(J; \mathbf{R}^n)$, $BC_{\text{left}}(J; \mathbf{R}^n)$, $BC_{\text{right}}(J; \mathbf{R}^n)$: The Banach spaces of bounded and continuous, or left-continuous, or right-continuous $\mathbf{R}^n$-valued functions on the interval J, with the sup-norm.

$BUC(J; \mathbf{R}^n)$: The Banach space of bounded and uniformly continuous $\mathbf{R}^n$-valued functions on the interval J, with the sup-norm.

$BUC_0(J; \mathbf{R}^n)$: The subspace of $BUC(J; \mathbf{R}^n)$ of functions tending to zero at infinity.

$M(J; \mathbf{R}^{n \times m})$: The set of $n \times m$-dimensional matrix-valued measures of bounded variation on the interval J, with the total variation norm.

$\mathcal{A}(J; \mathbf{R}^{n \times m})$: The subset of $M(J; \mathbf{R}^{n \times m})$ of measures without a singular non-atomic part.

$\mathcal{I}$, $\mathcal{F}$, $\mathcal{G}$: $\mathcal{I} = L^2(\mathbf{R}^-; \mathbf{R}^m)$, $\mathcal{F} = L^2(\mathbf{R}^+; \mathbf{R}^n)$, and $\mathcal{G} = L^2(\mathbf{R}^+; \mathbf{R}^m)$.

$\mathcal{I}_1$, $\mathcal{F}_1$, $\mathcal{G}_1$: $\mathcal{I}_1 = W^{1,2}(\overline{\mathbf{R}}^-; \mathbf{R}^m)$, $\mathcal{F}_1 = W^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^n)$, and $\mathcal{G}_1 = W^{1,2}(\overline{\mathbf{R}}^+; \mathbf{R}^m)$.

I: The identity operator.

A^*: The (Hilbert space) adjoint of the operator A (i.e., the complex conjugate transpose of A, if A is a matrix).

$\mathcal{D}(\mathcal{A})$: The domain of the unbounded operator $\mathcal{A}$.

$\check{\nu}$: The measure obtained from the measure ν through a reflection of the time axis, combined with the passing to the matrix adjoint, i.e., $\check{\nu}(E) = \nu(-E)^*$ for each Borel set E.

$\hat{\nu}$: The Laplace (Stieltjes) transform of the measure $\nu \in M(\overline{\mathbf{R}}^+; \mathbf{R}^{n \times m})$.

$f * g$:   The convolution of f and g, where f and g are either functions or measures.

δ_0:   The unit atom at zero (the Dirac delta).

μ^{-1}:   The convolution inverse of the measure μ, i.e., $\mu * \mu^{-1} = \mu^{-1} * \mu = I\delta_0$.

τ_t: The translation operator $\tau_t f(s) = f(t + s)$ (this is a left-shift when $t > 0$).

χ_J: The characteristic function of $J \subset \mathbf{R}$.

π_J: The (projection) operator that maps a function f defined on $\mathbf{R}$ into $\pi_J f = \chi_J f$.

$\pi_+, \overline{\pi}_+$: $\pi_+ = \pi_{(0,\infty)}$ and $\overline{\pi}_+ = \pi_{[0,\infty)}$. These operators are the same in L^p, but they differ from each other in B^∞.

$\pi_-, \overline{\pi}_-$: $\pi_- = \pi_{(-\infty,0)}$ and $\overline{\pi}_- = \pi_{(-\infty,0]}$. These operators are the same in L^p, but they differ from each other in B^∞.

We throughout extend functions and measures defined on a subinterval J of $\mathbf{R}$ to the whole real line by requiring them to be zero outside of J. Thus, with this interpretation, $L^p(J; \mathbf{R}^n) \subset L^p(\mathbf{R}; \mathbf{R}^n)$, $BUC_0(J; \mathbf{R}^n) \subset B^\infty(\mathbf{R}; \mathbf{R}^n)$, and $M(J; \mathbf{R}^{n \times m}) \subset M(\mathbf{R}; \mathbf{R}^{n \times m})$. (Observe, in particular, that an extended function may be discontinuous at the end-points of J, even if the original function is continuous on J.)

REFERENCES

[1] A. BENSOUSSAN, G. DA PRATO, M. C. DELFOUR, AND S. K. MITTER, *Representation and Control of Infinite Dimensional Systems*, vol. 1 and 2, Birkhäuser, Boston, Basel, Berlin, 1992.

[2] A. BEURLING, *Sur les intégrals de Fourier absolument convergentes et leur application à une transformation fonctionnelle*, Neuvieme Congrès des Mathématique Scandinaves, (1938), pp. 345–366.

[3] F. M. CALLIER AND J. WINKIN, *Spectral factorization and LQ-optimal regulation for multivariable distributed systems*, International Journal of Control, 52 (1990), pp. 55–75.

[4] ———, *LQ-optimal control of infinite-dimensional systems by spectral factorization*, Automatica, 28 (1992), pp. 757–770.

[5] R. F. CURTAIN, G. WEISS, AND M. WEISS, *Coprime factorization for regular linear systems*, Automatica, (1996). To appear.

[6] R. F. CURTAIN AND H. ZWART, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Springer-Verlag, New York, 1995.

[7] C. A. DESOER AND M. VIDYASAGAR, *Feedback Systems: Input-Output Properties*, Electrical Science Series, Academic Press, New York San Francisco London, 1975.

[8] I. C. GOHBERG AND I. A. FEL'DMAN, *Convolution Equations and Projection Methods for their Solutions*, vol. 41 of Translations of Mathematical Monographs, American Mathematical Society, Providence, Rhode Island, 1974.

[9] I. C. GOHBERG AND M. G. KREĬN, *Systems of integral equations on a half line with kernels depending on the difference of arguments*, Uspehi Mathem. Nauk SSSR, 13 (1958), pp. 3–72. Translated as: American Mathematical Society Translations, series 2, vol. 14 (1960), 217–287.

[10] P. GRABOWSKI, *The LQ controller synthesis problem*, IMA Journal of Mathematical Control & Information, 10 (1993), pp. 131–148.

[11] ———, *The LQ controller problem: an example*, IMA Journal of Mathematical Control & Information, 11 (1994), pp. 355–368.

[12] M. Green and D. L. L. Limebeer, *Linear Robust Control*, Prentice Hall, Englewood Cliffs, New Jersey, 1995.

[13] G. Gripenberg, S.-O. Londen, and O. J. Staffans, *Volterra Integral and Functional Equations*, Cambridge University Press, Cambridge and New York, 1990.

[14] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, Berlin and New York, 1993.

[15] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, American Mathematical Society, Providence, Rhode Island, revised ed., 1957.

[16] I. Lasiecka and R. Triggiani, *Differential and Algebraic Riccati Equations with Applications to Boundary/Point Control Problems: Continuous Theory and Approximation Theory*, vol. 164 of Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 1991.

[17] ———, *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 (1993), pp. 659–693.

[18] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, vol. 170 of Die Grundlehren der mathematishen Wissenschaften in Einzeldarstellungen, Springer-Verlag, Berlin, 1971.

[19] A. J. Pritchard and D. Salamon, *The linear-quadratic control problem for retarded systems with delays in control and observation*, IMA Journal of Mathematical Control & Information, 2 (1985), pp. 335–362.

[20] ———, *The linear quadratic control problem for infinite dimensional systems with unbounded input and output operators*, SIAM Journal on Control and Optimization, 25 (1987), pp. 121–144.

[21] D. Salamon, *Infinite dimensional linear systems with unbounded control and observation: a functional analytic approach*, Transactions of the American Mathematical Society, 300 (1987), pp. 383–431.

[22] ———, *Realization theory in Hilbert space*, Mathematical Systems Theory, 21 (1989), pp. 147–164.

[23] O. J. Staffans, *Extended initial and forcing function semigroups generated by a functional equation*, SIAM Journal on Mathematical Analysis, 16 (1985), pp. 1034–1048.

[24] ———, *A neutral FDE is similar to the product of an ODE and a shift*, Journal on Mathematical Analysis and Applications, 192 (1995), pp. 627–654.

[25] ———, *Quadratic optimal control of stable systems through spectral factorization*, Mathematics of Control, Signals, and Systems, 8 (1995), pp. 167–197.

[26] ———, *Quadratic optimal control of stable well-posed linear systems*. To appear in Transactions of American Mathematical Society, 1996b.

[27] ———, *Coprime factorizations for well-posed linear systems*. Submitted to SIAM Journal on Control and Optimization, 1996c.

[28] ———, *Quadratic optimal control of well-posed linear systems*. Submitted to SIAM Journal on Control and Optimization, 1996d.

[29] ———, *On the discrete and continuous time infinite-dimensional Riccati equations*. To appear in Systems and Control Letters, 1996g.

[30] ———, *Feedback/feed-forward solutions of the distributed full information stable H^∞ minimax problem*. Preprint, 1996h.

[31] O. Toker and H. Özbay, *Gap metric problems for MIMO delay systems: parametization of all suboptimal controllers*, Automatica, 31 (1995), pp. 931–940.

[32] B. van Keulen, *H_∞-Control for Distributed Parameter Systems: A State Space Approach*, Birkhäuser Verlag, Basel Boston Berlin, 1993.

[33] M. Vidyasagar, *Control System Synthesis: a Factorization Approach*, The MIT Press, Cambridge, Massachusetts, 1985.

[34] G. Weiss, *Transfer functions of regular linear systems. Part I: Characterizations of regularity*, Transactions of American Mathematical Society, 342 (1994a), pp. 827–854.

[35] ———, *Regular linear systems with feedback*, Mathematics of Control, Signals, and Systems, 7 (1994b), pp. 23–57.

[36] M. WEISS, *Riccati equations in Hilbert space: A Popov function approach*, doctoral dissertation, Rijksuniversiteit Gronigen, 1994.

[37] N. WIENER AND H. R. PITT, *On absolutely convergent Fourier-Stieltes transforms*, Duke Mathematical Journal, 4 (1938), pp. 420–436.

[38] J. WINKIN, *Spectral factorization and feedback control for infinite-dimensional control systems*, doctoral dissertation, Faculteés Universitaires Notre-Dame de la Paix á Namur, 1989.