TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 349, Number 9, September 1997, Pages 3679–3715
S 0002-9947(97)01863-1

QUADRATIC OPTIMAL CONTROL OF STABLE WELL-POSED LINEAR SYSTEMS

OLOF J. STAFFANS

Abstract. We consider the infinite horizon quadratic cost minimization problem for a stable time-invariant well-posed linear system in the sense of Salamon and Weiss, and show that it can be reduced to a spectral factorization problem in the control space. More precisely, we show that the optimal solution of the quadratic cost minimization problem is of static state feedback type if and only if a certain spectral factorization problem has a solution. If both the system and the spectral factor are regular, then the feedback operator can be expressed in terms of the Riccati operator, and the Riccati operator is a positive self-adjoint solution of an algebraic Riccati equation. This Riccati equation is similar to the usual algebraic Riccati equation, but one of its coefficients varies depending on the subspace in which the equation is posed. Similar results are true for unstable systems, as we have proved elsewhere.

1. Introduction

This work treats the infinite horizon quadratic cost minimization problem for a time-invariant well-posed linear system in the sense of Salamon and Weiss, and extends the results presented in [37] to externally stable Salamon-Weiss systems (the unstable case is discussed in [40]). The main idea behind this work is the same as in [37], and it can be described briefly as follows.

It is well-known that it is possible to solve a certain canonical Wiener-Hopf factorization problem by a state space method. The problem is to factor $\mathcal{D}^*\mathcal{D}$, where $\mathcal{D}$ is a stable input/output map and $\mathcal{D}^*$ is its adjoint, in the form

$$\mathcal{X}^*\mathcal{X} = \mathcal{D}^*\mathcal{D},$$

where $\mathcal{X}$ is a stable input/output map from the control space U into itself with a stable inverse $\mathcal{X}^{-1}$. The solution $\mathcal{X}$ is called a *spectral factor* of $\mathcal{D}^*\mathcal{D}$. In the state space approach we first construct a realization Ψ of $\mathcal{D}$ (or alternatively, we assume that a system Ψ is given whose input/output map is $\mathcal{D}$). Then we pose an infinite horizon quadratic cost minimization problem in the state space H of Ψ, and solve the algebraic Riccati equation to get the Riccati operator Π describing the optimal cost. From the Riccati operator we get the feedback operator K, and this feedback operator is then used to construct the spectral factor $\mathcal{X}$.

Although the state space solution of the factorization problem is the most common one today, it is not the only one available, and it is not universally applicable.

Received by the editors January 30, 1995 and, in revised form, March 20, 1996.
1991 *Mathematics Subject Classification.* Primary 49J27, 93A05, 47B35.
Key words and phrases. Spectral factorization, inner-outer factorization, Wiener-Hopf factorization, algebraic Riccati equation, state feedback.

In particular, it requires a certain smoothness of the symbol of the input/output map $\mathcal{D}$. Without this smoothness every possible realization of $\mathcal{D}$ must have both an unbounded control operator B and an unbounded observation operator C (with respect to the same space), and there has been no general Riccati equation theory available for that case.

The spectral factorization problem always has a solution (see Lemma 18 below), even though it has not been known how to compute this factorization in general by a state space method. Furthermore, there exist some independent ways of computing the factorization; see, e.g., [16] and [51]. We take advantage of this fact and study the converse of the problem presented above, i.e., we use the spectral factorization to solve the quadratic cost minimization problem in feedback form. More specifically, in Theorem 27 we show that these two problems are equivalent in the sense that the factorization problem has a solution if and only if the solution of the quadratic cost minimization problem is of static state feedback type. Moreover, there is a simple relation between the spectral factor and the feedback solution, which means that they can easily be computed from each other.

In order to derive the Riccati equation solution to the optimal feedback problem we have to make one additional assumption, namely that both the input/output map $\mathcal{D}$ and the spectral factor $\mathcal{X}$ are regular together with their adjoints in the sense of [45] (a much stronger version of this assumption holds in the classical case). We show that in this case the Riccati operator of an arbitrary realization Ψ of $\mathcal{D}$ satisfies a certain algebraic Riccati equation, and that the feedback operator K can be computed from Π. However, this Riccati equation contains one unexpected feature (first discovered in [37]): the positive self-adjoint weighting operator in the quadratic term in the Riccati equation is $(X^*X)^{-1}$, where X is the feed-through operator of the spectral factor $\mathcal{X}$, instead of the expected $(D^*D)^{-1}$, where D is the feed-through operator of the input/output map $\mathcal{D}$. If the control and observation operators are bounded then $X^*X = D^*D$, so the new theory agrees with the old one in this respect. However, it is not true in general that $X^*X = D^*D$, and the computation of X is a nontrivial task. In some cases X can be computed from a discrete time Riccati equation [41], and there is also a fairly general method available for the case of a scalar control (see [18, Chapter VII]). An alternative formula for X^*X is given in [40].

As explained above, under certain regularity assumptions we prove that the Riccati operator (the optimal cost operator) necessarily satisfies a certain Riccati equation. In the finite dimensional case this is the "difficult" direction to prove, the converse direction being almost trivial. In our case the converse direction is still open, partly due to the fact that the Riccati equation that we get is much more complicated than the classical one. In some sense it resembles the discrete time Riccati equation; cf. [41]. Moreover, because of the required regularity hypothesis, it is still an unsettled question what the appropriate conditions are for our results to apply to boundary control problems for partial differential equations. For time delay systems the regularity problem is less severe; see the examples in [38].

We work formally in the time domain the whole time, but some of our proofs are adapted frequency domain proofs, with some added state space features. The key addition is the facorization of the Hankel operator induced by the input/output map as the product of the controllability and observability maps. This makes it possible to connect the state space and the frequency domain theories to each other. See Definition 1 and the paragraph following Definition 4.

The first preprints of this paper and [40] were circulated in the spring of 1995, and a preliminary version was presented in the summer of 1995 at the conference [36]. In the spring of 1996 we received the preprint [50] where some of the results presented here and in [40] are also found. That preprint uses quite different notations, it deals with weakly regular systems, it contains some additional examples (one of which is essentially the same as the one in [41]), and it carries the analysis further than we do before resorting to spectral factorization. In particular, [50, Proposition 10.5] contains one additional Riccati equation valid on the domain of the closed loop generator in the case where the observation operator is bounded.

There is a significant overlap between some of the results presented here (those that are not related to spectral factorization) and those in [17]. In that paper the output operator C is bounded, but the system need not be stable, and no regularity assumptions are imposed on the spectral factor (the spectral factor is not even mentioned). The well-posedness of the closed loop system (our first main result, presented in Theorem 27) is not studied in [17]. The Riccati equation given in [17, Theorem 2.2] looks different from our Riccati equation (it does not contain the nonstandard term), due to the fact that [17, Corollary 4.9] extends the operator $B^*\Pi$ in a different way than our Proposition 36. An example highlighting this difference is given in [47].

For a further discussion of the existing literature we refer the reader to [38], [40]. The recent literature includes (but is certainly not restricted to) [1], [2], [4], [5], [6], [7], [9], [12] [13], [14], [20], [21], [23], [24], [25], [26], [27], [28], [30], [31], [42], [48], [49], [51], and the other papers in our list of references.

We use the following notation:

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

I: The identity operator.

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

$\mathrm{dom}(A)$: The domain of the (unbounded) operator A.

$\mathrm{range}(A)$: The range of the operator A.

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

$L^2(J;U)$: The set of U-valued L^2-functions on J.

$W^{1,2}(J;U)$: The set of functions in $L^2(J;U)$ with a (distribution) derivative in $L^2(J;U)$.

$C(J;U)$, $C^1(J;U)$: $C(J;U)$ is the set of U-valued continuous functions on J, and $C^1(J;U)$ is the set of U-valued continuously differentiable functions on J.

$BC(J;U)$, $BC^1(J;U)$: $BC(J;U)$ is the set of U-valued bounded continuous functions on J, and $BC^1(J;U)$ is the set of U-valued bounded continuously differentiable functions on J with a bounded derivative.

$\langle\cdot,\cdot\rangle_H$: The inner product in the Hilbert space H.

$\tau(t)$: The bilateral time shift operator $\tau(t)u(s) = u(t+s)$ (this is a left-shift when $t > 0$ and a right-shift when $t < 0$).

π_J: $(\pi_J u)(s) = u(s)$ if $s \in J$ and $(\pi_J u)(s) = 0$ if $s \notin J$. Here J is a subset of $\mathbf{R}$. This operator is used both as a projection operator $L^2(\mathbf{R}) \to L^2(\mathbf{R})$ and as an embedding operator $L^2(J) \to L^2(\mathbf{R})$.

π_+, π_-: $\pi_+ = \pi_{\mathbf{R}^+}$ and $\pi_- = \pi_{\mathbf{R}^-}$.

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

2. Well-Posed Linear Systems and Time-Invariant Operators

In order to fix the notation and describe the basic setting we first give a brief presentation of the Salamon-Weiss class of time-invariant well-posed linear systems, including only those parts of the theory that we need. This theory has been developed in [33], [34], [35], [8], [11], and [43], [44], [45], [46] (and many other papers), and we refer the reader to these sources for additional reading. (Salamon calls these systems "well-posed semigroup control systems" and Weiss calls them "abstract linear systems".) Since we need only externally stable systems here, we restrict our discussion to this class of systems.

In the externally stable case it is possible to define a well-posed linear system in a slightly nonstandard way that turns out to be very convenient for our purposes. Usually the axioms of the system are formulated in terms of operators defined on local L^2-spaces, but in the stable case it is much more convenient to work with global L^2-spaces. We shall do so throughout, since it simplifies the formulation significantly; in particular, it leads to a much simpler dual theory. This setting is also the best one for H^∞-related work. Parts of this setting are found in [35].

In order to formulate the axioms satisfied by a well-posed linear system we need the "past time" projection operator π_-, the "future time" projection operator π_+, and the bilateral "time shift" group $\tau(t)$. These operate on functions u defined on $\mathbf{R} = (-\infty, \infty)$ in the following way:

$$
\begin{aligned}
(\pi_- u)(s) &= \begin{cases} u(s), & s \in \mathbf{R}^-, \\ 0, & s \in \mathbf{R}^+, \end{cases} \\
(\pi_+ u)(s) &= \begin{cases} u(s), & s \in \mathbf{R}^+, \\ 0, & s \in \mathbf{R}^-, \end{cases} \\
(\tau(t)u)(s) &= u(t+s), \quad t, s \in \mathbf{R}.
\end{aligned}
$$

Definition 1. Let U, H, and Y be Hilbert spaces. A (*causal*) *externally stable* (*time-invariant*) *well-posed linear system on* (U, H, Y) is a quadruple $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$, where $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ are bounded linear operators of the following type:

(i) $\mathcal{A}$ is a strongly continuous semigroup on H;
(ii) $\mathcal{B}\colon L^2(\mathbf{R}; U) \to H$ satisfies $\mathcal{A}(t)\mathcal{B}u = \mathcal{B}\tau(t)\pi_- u$ for all $u \in L^2(\mathbf{R}; U)$ and $t \in \mathbf{R}^+$;
(iii) $\mathcal{C}\colon H \to L^2(\mathbf{R}; Y)$ satisfies $\mathcal{C}\mathcal{A}(t)x = \pi_+\tau(t)\mathcal{C}x$ for all $x \in H$ and $t \in \mathbf{R}^+$;
(iv) $\mathcal{D}\colon L^2(\mathbf{R}; U) \to L^2(\mathbf{R}; Y)$ satisfies $\tau(t)\mathcal{D}u = \mathcal{D}\tau(t)u$, $\pi_-\mathcal{D}\pi_+ u = 0$, and $\pi_+\mathcal{D}\pi_- u = \mathcal{C}\mathcal{B}u$ for all $u \in L^2(\mathbf{R}; U)$ and $t \in \mathbf{R}$.

If, furthermore, $\sup_{t \in \mathbf{R}^+} \|\mathcal{A}(t)\| < \infty$, then Ψ is called *stable* rather than externally stable, and if $\mathcal{A}(t)x \to 0$ as $t \to \infty$ for all $x \in H$, then Ψ is called *strongly stable*.

The different components of Ψ are named as follows: U is the *input space*, H the *state space*, Y the *output space*, $\mathcal{A}$ the *semigroup*, $\mathcal{B}$ the *controllability* (or *reachability*) *map*, $\mathcal{C}$ the *observability map*, $\mathcal{D}$ the *input/output map*, $\mathcal{B}\mathcal{B}^*$ the *controllability gramian*, and $\mathcal{C}^*\mathcal{C}$ the *observability gramian* of Ψ.

Remark 2. The same axioms can be used to define an arbitrary (unstable) well-posed linear system. We simply replace all the L^2-spaces by the corresponding weighted L^2-spaces $L^2_\omega = \{ f \in L^2_{\mathrm{loc}} \mid (t \mapsto \mathrm{e}^{-\omega t} f(t)) \in L^2 \}$. Here ω is an arbitrary number larger than the exponential growth rate of $\mathcal{A}$. See [39, Section 2] for details.

The axioms listed above describe standard properties of the corresponding maps induced by exponentially stable systems with bounded control and observation operators. Whenever we refer to a stable "classical" system, we mean a system of the following type: we let A be the generator of an exponentially stable semigroup $\mathcal{A}$ on a Hilbert space H, let U and Y be Hilbert spaces, let $B \in \mathcal{L}(U;H)$, $C \in \mathcal{L}(H;Y)$, and $D \in \mathcal{L}(U;Y)$, and consider the system

$$
\begin{aligned}
x'(t) &= Ax(t) + Bu(t), \\
y(t) &= Cx(t) + Du(t), \qquad t \ge T, \\
x(T) &= x_T,
\end{aligned} \tag{1}
$$

where T is a given initial time and x_T a given initial value. We call u the control, x the state, y the output (or observation), A the generator, B the control operator, C the observation operator, and D the feed-through operator of this classical system. The state x is required to be a strong solution of (1), i.e., the state x and output y are given by

$$
\begin{aligned}
x(t) &= \mathcal{A}(t-T)x_T + \int_T^t \mathcal{A}(t-s)Bu(s)\,ds, \qquad t \ge T, \\
y(t) &= C\mathcal{A}(t-T)x_T + \int_T^t C\mathcal{A}(t-s)Bu(s)\,ds + Du(t), \qquad t \ge T.
\end{aligned} \tag{2}
$$

In this case we define $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ by

$$\mathcal{B}u = \int_{-\infty}^0 \mathcal{A}(-s)Bu(s)\,ds, \tag{3}$$

$$\mathcal{C}x = \left(t \mapsto C\mathcal{A}(t)x, \qquad t \in \mathbf{R}^+\right), \tag{4}$$

$$\mathcal{D}u = \left(t \mapsto \int_{-\infty}^t C\mathcal{A}(t-s)Bu(s)\,ds + Du(t), \qquad t \in \mathbf{R}\right). \tag{5}$$

Thus, $\mathcal{B}$ is the mapping from the control $u \in L^2(\mathbf{R}^-;U)$ to the final state $x(0) \in H$ (take $T = -\infty$, $x_T = 0$, and $t = 0$), $\mathcal{C}$ is the mapping from the initial state $x_0 \in H$ to the output $y \in L^2(\mathbf{R}^+;Y)$ (take $T = 0$ and $u = 0$), and $\mathcal{D}$ is the mapping from the control $u \in L^2(\mathbf{R};U)$ to the output $y \in L^2(\mathbf{R};Y)$ (take $T = -\infty$ and $x_T = 0$).

We shall also need the concept of an input/output map $\mathcal{D}$ without knowing that it is a part of a well-posed linear system. In this case we apply the following definition:

Definition 3. Let U and Y be two Hilbert spaces. A bounded linear operator $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ is called *time-invariant* iff it commutes with bilateral time shifts, i.e., $\tau(t)\mathcal{D}u = \mathcal{D}\tau(t)u$ for all $u \in L^2(\mathbf{R};U)$ and all $t \in \mathbf{R}$. The *Hankel operator induced by* $\mathcal{D}$ is the operator $\pi_+\mathcal{D}\pi_-$, and the *anti-Hankel operator induced* *by* $\mathcal{D}$ is the operator $\pi_-\mathcal{D}\pi_+$. *The Toeplitz operator induced by* $\mathcal{D}$ is the operator $\pi_+\mathcal{D}\pi_+$, and the *anti-Toeplitz operator induced by* $\mathcal{D}$ is the operator $\pi_-\mathcal{D}\pi_-$.

The word "causal" that we have included in the definition of a well-posed linear system relates to the fact that all the components of Ψ in Definition 1 are causal:

Definition 4. An operator $\mathcal{B}\colon L^2(\mathbf{R};U) \to H$ is *causal* [*anti-causal*] if $\mathcal{B}\pi_+ = 0$ [$\mathcal{B}\pi_- = 0$]. An operator $\mathcal{C}\colon H \to L^2(\mathbf{R};Y)$ is *causal* [*anti-causal*] if $\pi_-\mathcal{C} = 0$ [$\pi_+\mathcal{C} = 0$]. An time-invariant operator $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ is *causal* [*anti-causal*] if $\pi_-\mathcal{D}\pi_+ = 0$ [$\pi_+\mathcal{D}\pi_- = 0$], and it is *static* (or *memoryless*) if it is both causal and anti-causal.

Thus, a time-invariant operator is causal iff its anti-Hankel operator vanishes, it is anti-causal iff its Hankel operator vanishes, and it is static if both its Hankel operator and its anti-Hankel operator vanish. The condition imposed on the input/output map $\mathcal{D}$ in Definition 1 requires $\mathcal{D}$ to be a causal time-invariant operator whose Hankel operator is equal to $\mathcal{C}\mathcal{B}$. Intuitively, a causal controllability map $\mathcal{B}$ maps past inputs into the present state, a causal observability map $\mathcal{C}$ maps the present state into future outputs, and future inputs to a causal input/output map $\mathcal{D}$ have no effect on past outputs.

The following result is immediate:

Lemma 5. *Let* $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ *be time-invariant.*

(i) *The adjoint* $\mathcal{D}^*\colon L^2(\mathbf{R};Y) \to L^2(\mathbf{R};U)$ *of* $\mathcal{D}$ *is time-invariant.*
(ii) *If* $\mathcal{D}$ *is invertible, then the inverse* $\mathcal{D}^{-1}$ *is time-invariant (but the causality of* $\mathcal{D}$ *does not imply the causality of* $\mathcal{D}^{-1}$*).*
(iii) *If* $\mathcal{D}$ *is causal then* $\mathcal{D}\pi_+ = \pi_+\mathcal{D}\pi_+$ *and* $\pi_-\mathcal{D} = \pi_-\mathcal{D}\pi_-$*, and if* $\mathcal{D}$ *is anti-causal then* $\mathcal{D}\pi_- = \pi_-\mathcal{D}\pi_-$ *and* $\pi_+\mathcal{D} = \pi_+\mathcal{D}\pi_+$*.*

Below we use this lemma repeatedly without explicit reference.

(Different authors use different names for the same concepts. For example, [32] use the names "S-analytic" for a causal time-invariant operator and "S-constant" for a static time-invariant operator, and they identify the class of anti-causal time-invariant operators with the class of operators whose adjoints are S-analytic. Here S stands for the (unilateral) Laguerre shift on $L^2(\mathbf{R}^+;U)$; see [32, p. 16].)

Static time-invariant operators have a very simple characterization:

Lemma 6. *The set of bounded linear static time-invariant operators* $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ *can be identified with the set of bounded linear operators* $M \in \mathcal{L}(U;Y)$ *in the following sense: For every* $M \in \mathcal{L}(U;Y)$ *the multiplication operator* $u \mapsto Mu$ *induced by* M *is a static time-invariant operator mapping* $L^2(\mathbf{R};U)$ *into* $L^2(\mathbf{R};Y)$*, and conversely, every static time-invariant operator* $\mathcal{D}$ *has a representation of this type.*

This well-known result can be derived from, e.g., [32, Theorem 5.2C, p. 96] or [15, Chapter 1].

Corollary 7. *The Hankel operator of a causal [the anti-Hankel operator of an anti-causal] time-invariant operator* $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ *determines* $\mathcal{D}$ *uniquely, modulo a multiplication operator* $u \mapsto Mu$*, where* $M \in \mathcal{L}(U;Y)$*.*

Proof. If two different causal time-invariant operators have the same Hankel operator, then their difference is static. $\square$

Thus, the control and observation operators $\mathcal{B}$ and $\mathcal{C}$ of a well-posed linear system determine the input/output map $\mathcal{D}$ uniquely, except for an undetermined multiplication operator (which in the classical case is represented by the feed-through operator D in (1)). We remark that every bounded causal time-invariant operator $\mathcal{D}$ can be represented as the input/output map of a strongly stable well-posed linear system Ψ; see [35, Theorem 4.3] or [39, Definition 2.10]. Such a system Ψ is called a *realization* of $\mathcal{D}$.

The *adjoint* of the system Ψ can be defined in two different ways, that differ from each other through a reflection of the time axis. We shall follow the classical tradition and let the adjoint system evolve in the backward time direction. To get this system we use the standard inner products in $L^2(\mathbf{R}; U)$ and $L^2(\mathbf{R}; Y)$ (instead of inner products that reflect the time axis), replace all the operators in Definition 1 by their adjoints, and finally interchange $\mathcal{B}^*$ and $\mathcal{C}^*$. Doing so we get a system of the following type:

Definition 8. Let Y, H, and U be Hilbert spaces. An *anti-causal externally stable well-posed linear system* on (Y, H, U) is a quadruple $\Psi^* = \left[\begin{smallmatrix} \mathcal{A}^* & \mathcal{C}^* \\ \mathcal{B}^* & \mathcal{D}^* \end{smallmatrix}\right]$, where $\mathcal{A}^*$, $\mathcal{C}^*$, $\mathcal{B}^*$, and $\mathcal{D}^*$ are bounded linear operators of the following type:

(i) $\mathcal{A}^*$ is a strongly continuous semigroup on H;
(ii) $\mathcal{C}^* \colon L^2(\mathbf{R}; Y) \to H$ satisfies $\mathcal{A}^*(-s)\mathcal{C}^* y^* = \mathcal{C}^* \tau(s)\pi_+ y^*$ for all $y^* \in L^2(\mathbf{R}; Y)$ and $s \in \mathbf{R}^-$;
(iii) $\mathcal{B}^* \colon H \to L^2(\mathbf{R}; U)$ satisfies $\mathcal{B}^* \mathcal{A}^*(-s) x^* = \pi_- \tau(s) \mathcal{B}^* x^*$ for all $x^* \in H$ and $s \in \mathbf{R}^-$;
(iv) $\mathcal{D}^* \colon L^2(\mathbf{R}; Y) \to L^2(\mathbf{R}; U)$ satisfies $\tau(s)\mathcal{D}^* y^* = \mathcal{D}^* \tau(s) y^*$, $\pi_+ \mathcal{D}^* \pi_- y^* = 0$, and $\pi_- \mathcal{D}^* \pi_+ y^* = \mathcal{B}^* \mathcal{C}^* y^*$ for all $y^* \in L^2(\mathbf{R}; Y)$ and $s \in \mathbf{R}$.

If, furthermore, $\sup_{t \in \mathbf{R}^+} \|\mathcal{A}^*(t)\| < \infty$, then Ψ^* is called *stable* rather than externally stable, and if $\mathcal{A}^*(t)x^* \to 0$ as $t \to \infty$ for all $x^* \in H$, then Ψ^* is called *strongly stable*.

The different components of Ψ^* are named as follows: Y is the *input space*, H the *state space*, U the *output space*, $\mathcal{A}^*$ the *semigroup*, $\mathcal{C}^*$ the *controllability map*, $\mathcal{B}^*$ the *observability map*, $\mathcal{D}^*$ the *input/output map*, $\mathcal{C}^*\mathcal{C}$ the *controllability gramian*, and $\mathcal{B}\mathcal{B}^*$ the *observability gramian* of Ψ^*.

It is easy to see that the adjoint of the causal system Ψ in Definition 1 is an anti-causal system Ψ^* of the type described in Definition 8. The two key observations are that π_+ and π_- are self-adjoint, and that the adjoint of the left-shift $\tau(t)$ is the right-shift $\tau(-t)$. The reason for the name “anti-causal” is obvious: all the parts of the anti-causal system are anti-causal according to Definition 4. By reversing the time direction in which Ψ^* evolves we get a causal system of the type described in Definition 1 (i.e., let R be the reflection operator $Ru(t) = u(-t)$, and replace $\mathcal{C}^*$, $\mathcal{B}^*$, and $\mathcal{D}^*$ by $\mathcal{C}^* R$, $R\mathcal{B}^*$, and $R\mathcal{D}^* R$, respectively). Thus, all results proved for causal systems can be applied to anti-causal systems, too (with some trivial modifications due to the reflection of the time direction).

In the classical stable case (1) the adjoint system Ψ^* is induced by the (anti-stable) system

$$
\begin{aligned}
(x^*)'(s) &= -A^* x^*(s) - C^* y^*(s), \\
u^*(s) &= B^* x^*(s) + D^* y^*(s), \qquad s \le S, \\
x^*(S) &= x^*_S,
\end{aligned} \tag{6}
$$

whose strong solution x and output y are given by

$$
(7) \qquad \begin{aligned}
x^*(s) &= \mathcal{A}^*(S-s)x_S^* + \int_s^S \mathcal{A}^*(t-s)C^*y^*(t)\,dt, \qquad s \le S,\\
u^*(s) &= B^*\mathcal{A}^*(S-s)x_S^* + \int_s^S B^*\mathcal{A}^*(t-s)C^*y^*(t)\,dt + D^*y^*(s), \qquad s \le S.
\end{aligned}
$$

The different maps of Ψ^* are now given by

$$
(8) \qquad \mathcal{C}^*y^* = \int_0^\infty \mathcal{A}^*(t)C^*y^*(t)\,dt,
$$

$$
(9) \qquad \mathcal{B}^*x^* = \left(s \mapsto B^*\mathcal{A}^*(-s)x^*, \qquad s \in \mathbf{R}^-\right),
$$

$$
(10) \qquad \mathcal{D}^*y^* = \left(s \mapsto \int_s^\infty B^*\mathcal{A}^*(t-s)C^*y^*(t)\,dt + D^*y^*(s), \qquad s \in \mathbf{R}\right).
$$

Two of the central notions of a well-posed linear system are still missing, namely the notions of the *controlled state* and the *output*:

Definition 9. Let $\Psi = \left[\begin{smallmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$ be a causal externally stable well-posed linear system on (U,H,Y), and let $\Psi^* = \left[\begin{smallmatrix}\mathcal{A}^* & \mathcal{C}^*\\ \mathcal{B}^* & \mathcal{D}^*\end{smallmatrix}\right]$ be its anti-causal adjoint system on (Y,H,U). Let $x_0 \in H$, $x_0^* \in H$, $u \in L^2(\mathbf{R};U)$, and $y^* \in L^2(\mathbf{R};Y)$. In the time-invariant setting the *controlled state $x(t)$ at time $t \in \mathbf{R}$* and the *output y of Ψ with control u* are given by

$$
\begin{pmatrix} x(t)\\ y \end{pmatrix} = \begin{pmatrix} \mathcal{B}\tau(t)u\\ \mathcal{D}u \end{pmatrix},
$$

and the *controlled state $x^*(s)$ at time $s \in \mathbf{R}$* and the *output u^* of Ψ^* with control y^** are given by

$$
\begin{pmatrix} x^*(s)\\ u^* \end{pmatrix} = \begin{pmatrix} \mathcal{C}^*\tau(s)y^*\\ \mathcal{D}^*y^* \end{pmatrix}.
$$

In the initial value setting with initial time zero, the *controlled state $x(t)$ at time $t \in \mathbf{R}^+$* and the *output y of Ψ with initial value x_0 and control u* are given by

$$
\begin{pmatrix} x(t)\\ y \end{pmatrix} = \begin{pmatrix} \mathcal{A}(t) & \mathcal{B}\tau(t)\\ \mathcal{C} & \mathcal{D} \end{pmatrix}\begin{pmatrix} x_0\\ \pi_+u \end{pmatrix} = \begin{pmatrix} \mathcal{A}(t)x_0 + \mathcal{B}\tau(t)\pi_+u\\ \mathcal{C}x_0 + \mathcal{D}\pi_+u \end{pmatrix},
$$

and the *controlled state $x^*(s)$ at time $s \in \mathbf{R}^-$* and the *output u^* of Ψ^* with initial value x_0^* and control y^** are given by

$$
\begin{pmatrix} x^*(s)\\ u^* \end{pmatrix} = \begin{pmatrix} \mathcal{A}^*(-s) & \mathcal{C}^*\tau(s)\\ \mathcal{B}^* & \mathcal{D}^* \end{pmatrix}\begin{pmatrix} x_0^*\\ \pi_-y^* \end{pmatrix} = \begin{pmatrix} \mathcal{A}^*(-s)x_0^* + \mathcal{C}^*\tau(s)\pi_-y^*\\ \mathcal{B}^*x_0^* + \mathcal{D}^*\pi_-y^* \end{pmatrix}.
$$

Let us remark that the most common problem is the initial value problem for Ψ with initial time zero, and in many papers this is the only one that is treated. Frequently the adjoint system Ψ^* is also studied in an initial value setting with a positive initial time S; see, e.g., [34, Section 5]. This is in particular true in studies of the finite horizon quadratic cost minimization problem. See [39, Section 2] for the appropriate definition.

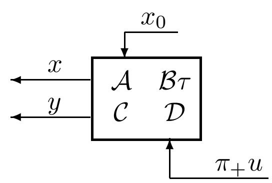

FIGURE 1. Input/State/Output Diagram of Ψ

In the case of the stable classical system (1) and its anti-stable adjoint (6), in the time-invariant setting the states of Ψ and Ψ^* are given by

$$x(t) = \int_{-\infty}^{t} \mathcal{A}(t-s)Bu(s)\,ds, \quad t \in \mathbf{R}, \tag{11}$$

$$x^*(s) = \int_{s}^{\infty} \mathcal{A}^*(t-s)C^*y^*(t)\,dt, \quad s \in \mathbf{R}, \tag{12}$$

and the outputs are given by (5) and (10). In the initial value setting with initial time zero the states and outputs of the same systems are given by

$$x(t) = \mathcal{A}(t)x_0 + \int_0^t \mathcal{A}(t-s)Bu(s)\,ds, \quad t \in \mathbf{R}^+, \tag{13}$$

$$y(t) = C\mathcal{A}(t)x_0 + \int_0^t C\mathcal{A}(t-s)Bu(s)\,ds + Du(t), \quad t \in \mathbf{R}^+, \tag{14}$$

$$x^*(s) = \mathcal{A}^*(-s)x_0^* + \int_s^0 \mathcal{A}^*(t-s)C^*y^*(t)\,dt, \quad s \in \mathbf{R}^-, \tag{15}$$

$$u^*(s) = B^*\mathcal{A}^*(-s)x_0^* + \int_s^0 B^*\mathcal{A}^*(t-s)C^*y^*(t)\,dt + D^*y^*(s), \quad s \in \mathbf{R}^-. \tag{16}$$

Compare these formulas to (2) and (7).

An important property of the two different settings is that if we "restrict the causal time-invariant problem to $\mathbf{R}^+$," then we get an initial value problem, or more precisely, on $\mathbf{R}^+$ the state and the output of the causal time-invariant problem are identical to the state and the output of the initial value problem with x_0 replaced by $x(0) = \mathcal{B}u$. A similar statement is true in the anti-causal case (replace $\mathbf{R}^+$ by $\mathbf{R}^-$ and x_0^* by $\mathcal{C}^*y^*$).

Remark 10. Because of Definition 9, we also use the alternative notation $\left[\begin{smallmatrix}\mathcal{A} & \mathcal{B}_{\mathcal{T}}\\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$ for the system $\left[\begin{smallmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$.

We use diagrams of the type drawn in Figure 1 to represent the relation between the state x, the output y, the initial value x_0, and the control u of Ψ in the initial value setting. In our diagrams we use the following conventions throughout:

(i) Initial states and controls enter at the top or bottom, and they are acted on by all the operators located in the column to which they are attached. In particular, note that x_0 is attached to the first column and u to the second.

(ii) Final states and outputs leave to the left or right, and they are the sums of all the elements in the row to which they are attached. In particular, note that x is attached to the top row, and y to the bottom row.

A similar diagram is used to describe the adjoint system Ψ^*. In the time-invariant setting we use the same diagram without the initial vector x_0 and the projection π_+.

3. Quadratic Cost Minimization

This study centers around four problems that are closely connected to each other, namely quadratic cost minimization, spectral factorization, inner-outer factorization, and various types of feedback. The first of these problems is presented here, the next two are presented in Section 4, and the feedback configurations in Section 5.

We formulate the *quadratic cost minimization problem* for a causal well-posed linear system Ψ in the initial value setting with initial time zero:

Definition 11. Let $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$ be an externally stable causal well-posed linear system on (U, H, Y). The *quadratic cost minimization problem* for Ψ consists of finding, for each $x_0 \in H$, the infimum over all $u \in L^2(\mathbf{R}^+; U)$ of the cost

$$Q(x_0, u) = \|y\|^2_{L^2(\mathbf{R}^+;Y)},$$

where $y = \mathcal{C}x_0 + \mathcal{D}\pi_+ u$ is the output of Ψ with initial value x_0 and control u.

As is well-known, the existence of a unique minimizing solution u_{opt} of the quadratic cost minimization problem is guaranteed if the input/output map $\mathcal{D}$ of Ψ is coercive:

Definition 12. A bounded linear operator $\mathcal{D}\colon L^2(\mathbf{R}; U) \to L^2(\mathbf{R}; Y)$ is *coercive* (or *bounded from below*) iff there exists some $\epsilon > 0$ such that

$$\|\mathcal{D}u\|^2_{L^2(\mathbf{R};Y)} \geq \epsilon \|u\|^2_{L^2(\mathbf{R};U)}$$

for all $u \in L^2(\mathbf{R}; U)$. Equivalently, $\mathcal{D}$ is *coercive* iff

$$\mathcal{D}^*\mathcal{D} \geq \epsilon I$$

for some $\epsilon > 0$. Equivalently (by the closed graph theorem), $\mathcal{D}$ is *coercive* iff $\mathcal{D}$ is one-to-one and has a closed range.

Lemma 13. *Let $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$ be an externally stable well-posed linear system on (U, H, Y).*

(i) *$\mathcal{D}$ is coercive if and only if the Toeplitz operator $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$, regarded as an operator mapping $L^2(\mathbf{R}^+; U)$ into itself, has a bounded inverse. We denote this inverse by $(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}$.*

(ii) *Suppose that $\mathcal{D}$ is coercive. Then, for each $x_0 \in H$, there is a unique control $u_{\mathrm{opt}} \in L^2(\mathbf{R}^+; U)$ that minimizes the cost function $Q(x_0, u)$ in Definition 11, namely*

$$u_{\mathrm{opt}} = -(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*\mathcal{C}x_0. \tag{17}$$

The corresponding state x_{opt}, the output y_{opt}, and the minimal value of the cost function Q are given by

$$x_{\mathrm{opt}} = \mathcal{A}x_0 - \mathcal{B}\tau\pi_+(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*\mathcal{C}x_0, \tag{18}$$

$$y_{\mathrm{opt}} = (I - P)\,\mathcal{C}x_0, \tag{19}$$

$$Q(x_0, u_{\mathrm{opt}}) = \langle x_0, \mathcal{C}^*\,(I - P)\,\mathcal{C}x_0\rangle_H, \tag{20}$$

where $P = \mathcal{D}\pi_+(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*$ is the orthogonal projection in $L^2(\mathbf{R}; Y)$ onto the range of $\mathcal{D}\pi_+$. Moreover,

$$\pi_+\mathcal{D}^*y_{\mathrm{opt}} = \pi_+\mathcal{D}^*\,(\mathcal{C}x_0 + \mathcal{D}\pi_+u_{\mathrm{opt}}) = 0, \tag{21}$$

that is, y_{opt} is orthogonal to the range of $\mathcal{D}\pi_+$.

Proof. (i) By definition, if $\mathcal{D}$ is coercive, then $\|\mathcal{D}u\|^2_{L^2(\mathbf{R};Y)} \ge \epsilon\|u\|^2_{L^2(\mathbf{R};U)}$ for some $\epsilon > 0$ and all $u \in L^2(\mathbf{R};U)$. In particular, this implies that $\|\mathcal{D}\pi_+u\|^2_{L^2(\mathbf{R};Y)} \ge \epsilon\|\pi_+u\|^2_{L^2(\mathbf{R};U)}$, and thus $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$ is positive and bounded away from zero on $L^2(\mathbf{R}^+;U)$, hence invertible.

Conversely, if $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$ is invertible on $L^2(\mathbf{R}^+;U)$, then it is strictly positive, and by the time-invariance of $\mathcal{D}$, there is some $\epsilon > 0$ such that for for all $u \in L^2(\mathbf{R};U)$,

$$\begin{aligned}
\|\mathcal{D}\pi_{[t,\infty)}u\|^2_{L^2(\mathbf{R};Y)} &= \|\mathcal{D}\tau(-t)\pi_+\tau(t)u\|^2_{L^2(\mathbf{R};Y)} \\
&= \|\tau(-t)\mathcal{D}\pi_+\tau(t)u\|^2_{L^2(\mathbf{R};Y)} \\
&= \|\mathcal{D}\pi_+\tau(t)u\|^2_{L^2(\mathbf{R};Y)} \\
&\ge \epsilon\|\pi_+\tau(t)u\|^2_{L^2(\mathbf{R};U)} \\
&= \epsilon\|u\|^2_{L^2([t,\infty);U)}.
\end{aligned}$$

Let $t \to -\infty$ to conclude that $\mathcal{D}^*\mathcal{D} \ge \epsilon I$.

(ii) If $\mathcal{D}$ is coercive, then the cost function $Q(x_0,u)$ is convex and coercive with respect to $u \in L^2(\mathbf{R}^+;U)$; hence there is a unique minimizing control $u_{\mathrm{opt}} \in L^2(\mathbf{R}^+;U)$. To show that the corresponding output y_{opt} satisfies (21) we argue as follows. Without loss of generality, let us suppose that U is a real Hilbert space (if not, then we replace the inner product in U by the real inner product $\Re\langle\cdot,\cdot\rangle$), and let us compute the Fréchet derivative of the cost function $Q(x_0,u)$ with respect to u at the optimal u_{opt}. For each variation $\eta \in L^2(\mathbf{R}^+;U)$, we have

$$\begin{aligned}
dQ(x_0,u_{\mathrm{opt}})\eta &= 2\left\langle \mathcal{C}x_0 + \mathcal{D}\pi_+u_{\mathrm{opt}}, \mathcal{D}\pi_+\eta\right\rangle_{L^2(\mathbf{R}^+;U)} \\
&= 2\left\langle y_{\mathrm{opt}}, \mathcal{D}\pi_+\eta\right\rangle_{L^2(\mathbf{R}^+;U)} \\
&= 2\left\langle \mathcal{D}^*y_{\mathrm{opt}}, \eta\right\rangle_{L^2(\mathbf{R}^+;U)}.
\end{aligned}$$

This is zero for all $\eta \in L^2(\mathbf{R}^+;U)$ iff (21) holds. Clearly, (17) follows from part (i) and (21). By substituting this value for u_{opt} into $x_{\mathrm{opt}} = \mathcal{A}x_0 + \mathcal{B}\tau\pi_+u_{\mathrm{opt}}$, $y_{\mathrm{opt}} = \mathcal{C}x_0+\mathcal{D}\pi_+u_{\mathrm{opt}}$, and $Q(x_0,u_{\mathrm{opt}})$ (and making a straightforward computation) we get (18), (19), and (20). ☐

Definition 14. Under the hypotheses of part (ii) of Lemma 13, define

$$\begin{aligned}
\mathcal{A}_{\circlearrowleft} &= \mathcal{A} - \mathcal{B}\tau\pi_+(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*\mathcal{C}, \\
\mathcal{C}_{\circlearrowleft} &= \left(I - \mathcal{D}\pi_+(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*\right)\mathcal{C}, \\
\mathcal{K}_{\circlearrowleft} &= -(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*\mathcal{C}, \\
\Pi &= \mathcal{C}^*\left(I - \mathcal{D}\pi_+(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}\pi_+\mathcal{D}^*\right)\mathcal{C}.
\end{aligned}$$

We call Π the *Riccati operator* of Ψ.

Thus, according to Lemma 13, $x_{\mathrm{opt}} = \mathcal{A}_{\circlearrowleft}x_0$, $y_{\mathrm{opt}} = \mathcal{C}_{\circlearrowleft}x_0$, $u_{\mathrm{opt}} = \mathcal{K}_{\circlearrowleft}x_0$, and $Q(x_0,u_{\mathrm{opt}}) = \langle x_0, \Pi x_0\rangle_H$.

Lemma 15. *Make the same assumption as in part (ii) of Lemma 13, and introduce the same notation as in Definition 14. Then the following claims are true:*

(i) *The operators* $\mathcal{A}_\circlearrowleft$, $\mathcal{C}_\circlearrowleft$, $\mathcal{K}_\circlearrowleft$, *and* Π *satisfy*

$$\begin{aligned}\mathcal{A}_\circlearrowleft &= \mathcal{A} + \mathcal{B}\tau\mathcal{K}_\circlearrowleft,\\ \mathcal{C}_\circlearrowleft &= \mathcal{C} + \mathcal{D}\mathcal{K}_\circlearrowleft,\\ \Pi &= \mathcal{C}_\circlearrowleft^*\mathcal{C}_\circlearrowleft = \mathcal{C}^*\mathcal{C}_\circlearrowleft = \mathcal{C}_\circlearrowleft^*\mathcal{C}.\end{aligned}$$

(ii) $\mathcal{A}_\circlearrowleft$ *is a strongly continuous semigroup on* H, *and* $\mathcal{C}_\circlearrowleft$ *and* $\mathcal{K}_\circlearrowleft$ *are admissible stable observability maps for* $\mathcal{A}_\circlearrowleft$ *in the sense that* $\mathcal{C}_\circlearrowleft \in \mathcal{L}(H; L^2(\mathbf{R}; Y))$, $\mathcal{K}_\circlearrowleft \in \mathcal{L}(H; L^2(\mathbf{R}; U))$, *and*

$$\begin{aligned}\mathcal{C}_\circlearrowleft\mathcal{A}_\circlearrowleft(t) &= \pi_+\tau(t)\mathcal{C}_\circlearrowleft,\\ \mathcal{K}_\circlearrowleft\mathcal{A}_\circlearrowleft(t) &= \pi_+\tau(t)\mathcal{K}_\circlearrowleft,\end{aligned}$$

for all $t \in \mathbf{R}^+$.

Proof. (i) This follows from Lemma 13 and Definition 14.

(ii) We fix some $t > 0$, and consider controls u of the type $u = u_{\mathrm{opt}} + \eta$, where $\eta \in L^2(\mathbf{R}^+; U)$ satisfies $\pi_{[0,t]}\eta = 0$ but is otherwise arbitrary. Then the state $x(t)$ at time t and the output y of Ψ are given by

$$\begin{aligned}x(t) &= x_{\mathrm{opt}}(t),\\ y(s) &= y_{\mathrm{opt}}(s), \qquad 0 \le s \le t,\\ \pi_+\tau(t)y &= \pi_+\tau(t)y_{\mathrm{opt}} + \mathcal{D}\tau(t)\eta\\ &= \mathcal{C}x_{\mathrm{opt}}(t) + \mathcal{D}\pi_+\tau(t)\left(u_{\mathrm{opt}} + \eta\right).\end{aligned}$$

We write the cost $Q(x_0, u)$ in the form

$$\begin{aligned}Q(x_0,u) &= \int_0^t \langle y_{\mathrm{opt}}(s), y_{\mathrm{opt}}(s)\rangle_Y \, ds + \int_t^\infty \langle y(s), y(s)\rangle_Y \, ds\\ &= \int_0^t \langle y_{\mathrm{opt}}(s), y_{\mathrm{opt}}(s)\rangle_Y \, ds + \int_0^\infty \langle (\tau(t)y)(s), (\tau(t)y)(s)\rangle_Y .\end{aligned}$$

Since u_{opt} is the minimizing control, the derivative of $Q(x_0, u)$ with respect to η must vanish at the point $\eta = 0$, and this implies that (cf. the proof of Lemma 13)

$$\pi_+\mathcal{D}^*\pi_+\tau(t)y_{\mathrm{opt}} = \pi_+\mathcal{D}^*\left(\mathcal{C}x_{\mathrm{opt}}(t) + \mathcal{D}\pi_+\tau(t)u_{\mathrm{opt}}\right) = 0,$$

i.e., (21) holds with x_0, y_{opt}, and u_{opt} replaced by $x_{\mathrm{opt}}(t)$, $\pi_+\tau(t)y_{\mathrm{opt}}$, and $\pi_+\tau(t)u_{\mathrm{opt}}$, respectively. From this equation we can solve $\pi_+\tau(t)u_{\mathrm{opt}}$, and get (cf. Lemma 13 and Definition 14)

$$\begin{aligned}\pi_+\tau(t)\mathcal{K}_\circlearrowleft x_0 &= \pi_+\tau(t)u_{\mathrm{opt}} = \mathcal{K}_\circlearrowleft x_{\mathrm{opt}}(t) = \mathcal{K}_\circlearrowleft\mathcal{A}_\circlearrowleft(t)x_0,\\ \pi_+\tau(t)\mathcal{C}_\circlearrowleft x_0 &= \pi_+\tau(t)y_{\mathrm{opt}} = \mathcal{C}_\circlearrowleft x_{\mathrm{opt}}(t) = \mathcal{C}_\circlearrowleft\mathcal{A}_\circlearrowleft(t)x_0,\\ \mathcal{A}_\circlearrowleft(s+t)x_0 &= x_{\mathrm{opt}}(s+t) = \mathcal{A}_\circlearrowleft(s)x_{\mathrm{opt}}(t) = \mathcal{A}_\circlearrowleft(s)\mathcal{A}_\circlearrowleft(t)x_0.\end{aligned}$$

Thus, $\mathcal{A}_\circlearrowleft$ is a semigroup, and $\mathcal{K}_\circlearrowleft$ and $\mathcal{C}_\circlearrowleft$ are admissible observability maps for $\mathcal{A}_\circlearrowleft$. The strong continuity of $\mathcal{A}_\circlearrowleft$ is immediate. □

A stronger version of this lemma was proved independently by Zwart [52].

4. Spectral Factorization and Inner-Outer Factorization

As we shall prove below, the quadratic cost minimization problem is closely related to the *spectral factorization problem*, where we factor $\mathcal{D}^*\mathcal{D}$ into two invertible factors, out of which one is causal, and the other is anti-causal:

Definition 16. Let $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ be bounded, linear, and time-invariant. A bounded linear time-invariant operator $\mathcal{X}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};U)$ is called a (*canonical and invertible*) *spectral factor* of $\mathcal{D}^*\mathcal{D}$ iff $\mathcal{X}$ is causal and has a bounded causal inverse (this inverse is necessarily time-invariant), and

$$\mathcal{X}^*\mathcal{X} = \mathcal{D}^*\mathcal{D}.$$

As is well-known, the spectral factorization problem can be reduced to an inner-outer factorization problem.

Definition 17. Let $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ be bounded, linear, causal, and time-invariant. An inner-outer factorization of $\mathcal{D}$ is a factorization of the form

$$\mathcal{D} = \mathcal{Y}\mathcal{X},$$

where both $\mathcal{Y}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ and $\mathcal{X}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};U)$ are bounded linear causal time-invariant operators, $\mathcal{Y}$ is an isometry, i.e., $\mathcal{Y}^*\mathcal{Y} = I$, and the range of $\mathcal{X}\pi_+$ is dense in $L^2(\mathbf{R}^+;U)$. We call $\mathcal{Y}$ an *inner factor* and $\mathcal{X}$ an *outer factor* of $\mathcal{D}$.

Note that $\mathcal{X}\pi_+$ is the Toeplitz operator induced by $\mathcal{X}$.

Lemma 18. *Let $\mathcal{D}\colon L^2(\mathbf{R};U) \to L^2(\mathbf{R};Y)$ be bounded, linear, causal, and time-invariant.*

(i) *If $\mathcal{X}$ is a spectral factor of $\mathcal{D}^*\mathcal{D}$, then $\mathcal{D}$ is coercive and $\mathcal{Y}\mathcal{X} = \left(\mathcal{D}\mathcal{X}^{-1}\right)\mathcal{X}$ is an inner-outer factorization of $\mathcal{D}$. Conversely, if $\mathcal{D}$ is coercive and $\mathcal{Y}\mathcal{X}$ is an inner-outer factorization of $\mathcal{D}$, then $\mathcal{X}$ is a spectral factor of $\mathcal{D}^*\mathcal{D}$.*

(ii) *$\mathcal{D}^*\mathcal{D}$ has a spectral factor $\mathcal{X}$ if and only if $\mathcal{D}$ is coercive.*

(iii) *The set of all spectral factors of $\mathcal{D}^*\mathcal{D}$ can be parametrized as $J\mathcal{X}$, where $\mathcal{X}$ is a fixed spectral factor and $J \in \mathcal{L}(U)$ is an arbitrary unitary operator.*

(iv) *The inverse of the Toeplitz operator $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$ in part (i) of Lemma 13 can be written in the form $(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1} = \mathcal{X}^{-1}\pi_+(\mathcal{X}^*)^{-1}$, where $\mathcal{X}$ is an arbitrary spectral factor of $\mathcal{D}^*\mathcal{D}$.*

Inner-outer factorizations (without any coercivity assumptions) can be parametrized in the same way; see [32, Theorem B, p. 101].

Proof of Lemma 18. (i) The first claim is obvious. Conversely, suppose that $\mathcal{D}$ is coercive, and that $\mathcal{Y}\mathcal{X}$ is an inner-outer factorization of $\mathcal{D}$. Then $\mathcal{D}^*\mathcal{D} = \mathcal{X}^*\mathcal{Y}^*\mathcal{Y}\mathcal{X} = \mathcal{X}^*\mathcal{X}$. Thus, in order to prove that $\mathcal{X}$ is a spectral factor of $\mathcal{D}^*\mathcal{D}$, it suffices to show that $\mathcal{X}$ has a bounded causal inverse.

Since $\mathcal{D}$ is coercive and $\mathcal{Y}$ is an isometry, also $\mathcal{X}$ is coercive; hence $\mathcal{X}$ is one-to-one and has a closed range. By the time-invariance of $\mathcal{X}$, we have $\mathcal{X}\tau(t)\pi_+ = \tau(t)\mathcal{X}\pi_+$ for all $t \in \mathbf{R}^+$, and this combined with the fact that the range of $\mathcal{X}\pi_+$ is dense in $L^2(\mathbf{R}^+;U)$ implies that the range of $\mathcal{X}$ is dense in $L^2(\mathbf{R};U)$. Thus, $\mathcal{X}$ is a bijection and, by the closed graph theorem, it has a bounded inverse $\mathcal{X}^{-1}$. To prove that this inverse is causal, it suffices to show that the range of $\mathcal{X}\pi_+$ is equal to $L^2(\mathbf{R}^+;U)$. But this follows from the fact that the range of $\mathcal{X}\pi_+$ is both dense in $L^2(\mathbf{R}^+;U)$ and (by the coercivity of $\mathcal{X}\pi_+$) closed in $L^2(\mathbf{R}^+;U)$.

(ii) If $\mathcal{D}^*\mathcal{D}$ has a spectral factor, then $\mathcal{D}^*\mathcal{D}$ is invertible, and $\mathcal{D}$ must be coercive. Conversely, if $\mathcal{D}$ is coercive, then it follows from [32, Theorem 3.4, p. 50 and Theorem 3.7, p. 54] that $\mathcal{D}^*\mathcal{D}$ has a factorization $\mathcal{X}^*\mathcal{X}$ where $\mathcal{X}$ is outer. By the argument given above $\mathcal{X}$ is invertible; hence $\mathcal{X}$ is a spectral factor of $\mathcal{D}^*\mathcal{D}$.

(iii) To prove the claim about the uniqueness of a spectral factor $\mathcal{X}$, suppose that both $\mathcal{X}$ and $\mathcal{Z}$ are spectral factors, i.e.,

$$\mathcal{X}^*\mathcal{X} = \mathcal{D}^*\mathcal{D} = \mathcal{Z}^*\mathcal{Z},$$

and that both $\mathcal{X}$ and $\mathcal{Z}$ are causal and have causal inverses. Then

$$\mathcal{X}\mathcal{Z}^{-1} = (\mathcal{X}^*)^{-1}\,\mathcal{Z}^*.$$

The left-hand side is causal and the right-hand side is anti-causal, so these operators are static and, by Lemma 6,

$$\mathcal{X}\mathcal{Z}^{-1} = (\mathcal{X}^*)^{-1}\,\mathcal{Z}^* = J$$

for some operator $J \in \mathcal{L}(U)$. Thus, $\mathcal{X} = J\mathcal{Z}$ and $\mathcal{Z}^* = \mathcal{X}^*J$; hence $\mathcal{Z} = J^*\mathcal{X} = J^*J\mathcal{Z}$. Since both $\mathcal{X}$ and J are invertible, this implies that J is unitary, i.e., $J^*J = JJ^* = I$.

(iv) Use the causality of $\mathcal{X}$ and anti-causality of $(\mathcal{X}^*)^{-1}$ to get

$$\begin{aligned}
\mathcal{X}^{-1}\pi_+(\mathcal{X}^*)^{-1}\pi_+\mathcal{D}^*\mathcal{D}\pi_+ &= \mathcal{X}^{-1}\pi_+(\mathcal{X}^*)^{-1}\pi_+\mathcal{X}^*\mathcal{X}\pi_+\\
&= \mathcal{X}^{-1}\pi_+(\mathcal{X}^*)^{-1}\mathcal{X}^*\mathcal{X}\pi_+\\
&= \mathcal{X}^{-1}\pi_+\mathcal{X}\pi_+\\
&= \mathcal{X}^{-1}\mathcal{X}\pi_+\\
&= \pi_+.
\end{aligned}$$

Thus, $\mathcal{X}^{-1}\pi_+(\mathcal{X}^*)^{-1}$ is a left inverse of $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$ on $L^2(\mathbf{R}^+;U)$. Since we know $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$ to be invertible, this operator must also be a right inverse of $\pi_+\mathcal{D}^*\mathcal{D}\pi_+$. □

5. Static Output Feedback, State Feedback, and Output Injection

To connect the quadratic cost minimization problem to the spectral factorization problem (or alternatively, to the inner-outer factorization problem), we need the notion of a state feedback, which can be reduced to the notion of a (static) *output feedback*. The idea is to feed a part $z = Ly$ of the output y of a well-posed linear system Ψ back into the input, as drawn in Figure 2. Here L is a bounded linear operator from the output space into the input space. Then, in the initial value setting with initial value x_0 and input v, we find that the effective input u, the state $x(t)$ at time $t \geq 0$, the output y, and the feedback control signal z satisfy the equations

$$\begin{aligned}
u &= z + \pi_+ v,\\
x(t) &= \mathcal{A}(t)x_0 + \mathcal{B}\tau(t)u,\\
y &= \mathcal{C}x_0 + \mathcal{D}u,\\
z &= Ly,
\end{aligned} \tag{22}$$

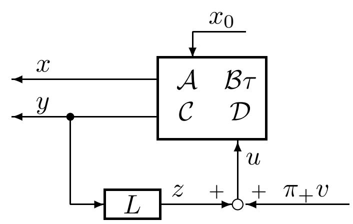

FIGURE 2. Static Output Feedback

which formally can be solved as

$$
\begin{aligned}
u &= (I - L\mathcal{D})^{-1} (L\mathcal{C}x_0 + \pi_+ v), \\
x(t) &= \left(\mathcal{A}(t) + \mathcal{B}\tau(t)L(I - \mathcal{D}L)^{-1}\mathcal{C}\right) x_0 + \mathcal{B}(I - L\mathcal{D})^{-1}\tau(t)\pi_+ v, \\
y &= (I - \mathcal{D}L)^{-1}(\mathcal{C}x_0 + \mathcal{D}\pi_+ v), \\
z &= (I - L\mathcal{D})^{-1} L(\mathcal{C}x_0 + \mathcal{D}\pi_+ v).
\end{aligned}
\tag{23}
$$

We say that the feedback operator L is admissible whenever these equations are valid:

Definition 19. Let $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$ be an externally stable causal [anti-causal] well-posed linear system on (U, H, Y). The operator $L \in \mathcal{L}(Y; U)$ is called an *admissible stable output feedback operator* for Ψ iff the time-invariant operator $I - L\mathcal{D}\colon L^2(\mathbf{R}; U) \to L^2(\mathbf{R}; U)$ has a bounded causal [anti-causal] inverse, or equivalently, iff the time-invariant operator $I - \mathcal{D}L\colon L^2(\mathbf{R}; Y) \to L^2(\mathbf{R}; Y)$ has a bounded causal [anti-causal] inverse.

Observe that the inverses are necessarily time-invariant whenever they exist. That the two different invertibility conditions are equivalent is well-known, and so are the facts that

$$
\begin{aligned}
(I - \mathcal{D}L)^{-1} &= I + \mathcal{D}(I - L\mathcal{D})^{-1} L, \\
(I - L\mathcal{D})^{-1} &= I + L(I - \mathcal{D}L)^{-1}\mathcal{D}, \\
(I - \mathcal{D}L)^{-1}\mathcal{D} &= \mathcal{D}(I - L\mathcal{D})^{-1}, \\
(I - L\mathcal{D})^{-1} L &= L(I - \mathcal{D}L)^{-1}.
\end{aligned}
$$

As Weiss [46, Section 6] proved, x and y in (23) can be interpreted as the state and output of another well-posed linear system:

Proposition 20. *Let $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$ be an externally stable causal well-posed linear system, and let $L \in \mathcal{L}(Y; U)$ be an admissible stable output feedback operator for Ψ.*

Then the system

$$\begin{aligned}\Psi_L = \begin{bmatrix}\mathcal{A}_L & \mathcal{B}_L\tau \\ \mathcal{C}_L & \mathcal{D}_L\end{bmatrix} &= \begin{bmatrix}\mathcal{A} + \mathcal{B}\tau L\,(I-\mathcal{D}L)^{-1}\,\mathcal{C} & \mathcal{B}\,(I-L\mathcal{D})^{-1}\,\tau \\ (I-\mathcal{D}L)^{-1}\,\mathcal{C} & \mathcal{D}\,(I-L\mathcal{D})^{-1}\end{bmatrix} \\ &= \begin{bmatrix}\mathcal{A} & \mathcal{B}\tau \\ \mathcal{C} & \mathcal{D}\end{bmatrix} + \begin{bmatrix}\mathcal{B}\tau \\ \mathcal{D}\end{bmatrix} L\,(I-\mathcal{D}L)^{-1}\begin{bmatrix}\mathcal{C} & \mathcal{D}\end{bmatrix} \\ &= \begin{bmatrix}\mathcal{A} & \mathcal{B}\tau \\ \mathcal{C} & \mathcal{D}\end{bmatrix} + \begin{bmatrix}\mathcal{B}\tau \\ \mathcal{D}\end{bmatrix} L\begin{bmatrix}\mathcal{C}_L & \mathcal{D}_L\end{bmatrix} \\ &= \begin{bmatrix}\mathcal{A} & \mathcal{B}\tau \\ \mathcal{C} & \mathcal{D}\end{bmatrix} + \begin{bmatrix}\mathcal{B}_L\tau \\ \mathcal{D}_L\end{bmatrix} L\begin{bmatrix}\mathcal{C} & \mathcal{D}\end{bmatrix}\end{aligned}$$

is another externally stable causal well-posed linear system on (U, H, Y). We call this system the closed loop system with output feedback operator L. In the initial value setting with initial time zero, initial value x_0, and control v, the controlled state $x(t)$ at time t and the output y of Ψ_L are given by (23). (An analogous result is valid in the anti-causal case.)

To get a better feeling for the feedback formula in Proposition 20 we recommend that the reader carry out the following exercise. If in the classical system (1) we replace u by $u = Ly + v$, then we get a new well-defined system of the same type iff $I - DL$ is invertible, or equivalently, iff $I - LD$ is invertible. In the new system the operators $\left[\begin{smallmatrix}A & B \\ C & D\end{smallmatrix}\right]$ have been replaced by

$$\begin{aligned}\begin{bmatrix}A_L & B_L \\ C_L & D_L\end{bmatrix} &= \begin{bmatrix}A + BL\,(I-DL)^{-1}\,C & B\,(I-LD)^{-1} \\ (I-DL)^{-1}\,C & D\,(I-LD)^{-1}\end{bmatrix} \\ &= \begin{bmatrix}A & B \\ C & D\end{bmatrix} + \begin{bmatrix}B \\ D\end{bmatrix} L\,(I-DL)^{-1}\begin{bmatrix}C & D\end{bmatrix} \\ &= \begin{bmatrix}A & B \\ C & D\end{bmatrix} + \begin{bmatrix}B \\ D\end{bmatrix} L\begin{bmatrix}C_L & D_L\end{bmatrix} \\ &= \begin{bmatrix}A & B \\ C & D\end{bmatrix} + \begin{bmatrix}B_L \\ D_L\end{bmatrix} L\begin{bmatrix}C & D\end{bmatrix}.\end{aligned} \tag{24}$$

This formula is probably not that familiar to most readers, since it is not common to allow a feed-through operator D inside a feedback loop, and the formula simplifies significantly when $D = 0$. In addition, the invertibility condition on $I - DL$ drops out when $D = 0$. However, there is a striking similarity between this formula and the one given in Proposition 20. A general well-posed linear system need not have a well-defined feed-through operator D (see Definition 35), so in the general case it is not possible to normalize D to be zero. Similar comments apply to our subsequent feedback formulas.

As the following lemma shows, stability and strong stability are preserved under stable feedback:

Lemma 21. *Let $\Psi = \left[\begin{smallmatrix}\mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$ be an externally stable causal well-posed linear system, and let $L \in \mathcal{L}(Y; U)$ be an admissible stable output feedback operator for Ψ. Then the closed loop system Ψ_L is stable iff Ψ is stable, i.e., $\sup_{t\in\mathbf{R}^+}\|\mathcal{A}(t)\| < \infty$ iff $\sup_{t\in\mathbf{R}^+}\|\mathcal{A}_L(t)\| < \infty$. Moreover Ψ_L is strongly stable iff Ψ is strongly stable, i.e., $\mathcal{A}(t)x \to 0$ as $t \to \infty$ for all $x \in H$ iff $\mathcal{A}_L(t)x \to 0$ as $t \to \infty$ for all $x \in H$.*

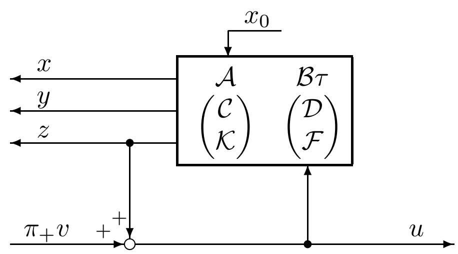

FIGURE 3. State Feedback Connection

Proof. The first claim follows from the fact that the difference $\mathcal{A}_L - \mathcal{A} = \mathcal{B}\tau L\,(I - \mathcal{D}L)^{-1}\,\mathcal{C}$ is bounded.

To prove the second claim it suffices to show that

$$\mathcal{B}\tau(t)L(I - \mathcal{D}L)^{-1}\mathcal{C}x \to 0 \text{ as } t \to \infty$$

for every $x \in H$. Fix $x \in H$, and split the expression above into

$$\begin{aligned}\mathcal{B}\tau(t)L(I - \mathcal{D}L)^{-1}\mathcal{C}x &= \mathcal{B}\tau(t-T)(\pi_+ + \pi_-)\tau(T)L(I - \mathcal{D}L)^{-1}\mathcal{C}x\\ &= \mathcal{B}\tau(t)\pi_{[T,\infty)}L(I - \mathcal{D}L)^{-1}\mathcal{C}x\\ &\quad + \mathcal{A}(t-T)\mathcal{B}\tau(T)L(I - \mathcal{D}L)^{-1}\mathcal{C}x.\end{aligned}$$

Here the first term tends to zero as $T \to \infty$, uniformly in $t \geq T$, and the second term tends to zero as $t \to \infty$ and T is fixed. □

As we mentioned above, a state feedback can be reduced to an output feedback as follows. The appropriate connection has been drawn in Figure 3.

Definition 22. Let $\Psi = \left[\begin{smallmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$ be an externally stable causal [anti-causal] well-posed linear system on (U, H, Y). The pair $\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix}$ is called an *admissible stable state feedback pair* for Ψ iff the extended system

$$\Psi_{\text{ext}} = \begin{bmatrix}\mathcal{A} & \mathcal{B}\\ \begin{pmatrix}\mathcal{C}\\ \mathcal{K}\end{pmatrix} & \begin{pmatrix}\mathcal{D}\\ \mathcal{F}\end{pmatrix}\end{bmatrix}$$

is an externally stable causal [anti-causal] well-posed linear system, and $\begin{pmatrix}0 & I\end{pmatrix}$ is an admissible stable output feedback operator for Ψ_{ext}, i.e., $I - \mathcal{F}$ has a bounded causal [anti-causal] inverse.

In the state feedback case with initial value x_0 and input v, we obtain the effective input u, the state $x(t)$, the output y, and the feedback control signal z of Ψ_{ext} from (22) and (23) by replacing $\mathcal{C}$, $\mathcal{D}$, L, and y by $\left(\begin{smallmatrix}\mathcal{C}\\ \mathcal{K}\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}\mathcal{D}\\ \mathcal{F}\end{smallmatrix}\right)$, $\begin{pmatrix}0 & I\end{pmatrix}$, and $\left(\begin{smallmatrix}y\\ z\end{smallmatrix}\right)$, respectively. This leads to the formulas

$$\begin{aligned}u &= z + \pi_+ v,\\ x(t) &= \mathcal{A}(t)x_0 + \mathcal{B}\tau(t)u,\\ y &= \mathcal{C}x_0 + \mathcal{D}u,\\ z &= \mathcal{K}x_0 + \mathcal{F}u,\end{aligned} \tag{25}$$

and

$$
\begin{aligned}
u &= (I-\mathcal{F})^{-1}(\mathcal{K}x_0 + \pi_+ v), \\
x(t) &= \Big(\mathcal{A}(t) + \mathcal{B}\tau(t)(I-\mathcal{F})^{-1}\mathcal{K}\Big)x_0 + \mathcal{B}(I-\mathcal{F})^{-1}\tau(t)\pi_+ v, \\
y &= \Big(\mathcal{C} + \mathcal{D}(I-\mathcal{F})^{-1}\mathcal{K}\Big)x_0 + \mathcal{D}(I-\mathcal{F})^{-1}\pi_+ v, \\
z &= (I-\mathcal{F})^{-1}(\mathcal{K}x_0 + \mathcal{F}\pi_+ v).
\end{aligned}
\tag{26}
$$

We get the closed loop state feedback system $\Psi_{\circlearrowleft}$ by making the same replacements in Proposition 20:

Proposition 23. *Let* $\Psi = \left[\begin{smallmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$ *be an externally stable causal well-posed linear system, and let* $\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix}$ *be an admissible stable state feedback pair for* Ψ. *Then the system*

$$
\begin{aligned}
\Psi_{\circlearrowleft} &= \begin{bmatrix} \mathcal{A}_{\circlearrowleft} & \mathcal{B}_{\circlearrowleft}\tau \\ \begin{pmatrix}\mathcal{C}_{\circlearrowleft}\\ \mathcal{K}_{\circlearrowleft}\end{pmatrix} & \begin{pmatrix}\mathcal{D}_{\circlearrowleft}\\ \mathcal{F}_{\circlearrowleft}\end{pmatrix}\end{bmatrix} \\
&= \begin{bmatrix} \mathcal{A} + \mathcal{B}\tau(I-\mathcal{F})^{-1}\mathcal{K} & \mathcal{B}(I-\mathcal{F})^{-1}\tau \\ \begin{pmatrix}\mathcal{C} + \mathcal{D}(I-\mathcal{F})^{-1}\mathcal{K}\\ (I-\mathcal{F})^{-1}\mathcal{K}\end{pmatrix} & \begin{pmatrix}\mathcal{D}(I-\mathcal{F})^{-1}\\ (I-\mathcal{F})^{-1} - I\end{pmatrix}\end{bmatrix} \\
&= \begin{bmatrix} \mathcal{A} & \mathcal{B}\tau \\ \begin{pmatrix}\mathcal{C}\\ \mathcal{K}\end{pmatrix} & \begin{pmatrix}\mathcal{D}\\ \mathcal{F}\end{pmatrix}\end{bmatrix} + \begin{bmatrix}\mathcal{B}\tau \\ \begin{pmatrix}\mathcal{D}\\ \mathcal{F}\end{pmatrix}\end{bmatrix}(I-\mathcal{F})^{-1}\begin{bmatrix}\mathcal{K} & \mathcal{F}\end{bmatrix}
\end{aligned}
$$

is another externally stable well-posed linear system on $(U, H, Y \times U)$. *We call this system the closed loop state feedback system with state feedback pair* $\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix}$. *In the initial value setting with initial time zero, initial value* x_0, *and control* v, *the controlled state* $x(t)$ *at time* t *and the outputs* y *and* z *of* $\Psi_{\circlearrowleft}$ *are given by (26). (An analogous result is valid in the anti-causal case.)*

Remark 24. We shall frequently regard the signal u in Figure 3 as an additional output of the closed loop system. This output has the same observability map $\mathcal{K}_{\circlearrowleft} = (I-\mathcal{F})^{-1}\mathcal{K}$ as the output z, and its input/output map is $\mathcal{F}_{\circlearrowleft} + I = (I-\mathcal{F})^{-1}$. A similar remark applies to the signal z^* in the adjoint output injection connection in Figure 4.

The adjoint of the state feedback system $\Psi_{\circlearrowleft}$ is the output injection system drawn in Figure 4.

Definition 25. Let $\Psi^* = \left[\begin{smallmatrix}\mathcal{A}^* & \mathcal{C}^*\\ \mathcal{B}^* & \mathcal{D}^*\end{smallmatrix}\right]$ be an externally stable anti-causal [causal] well-posed linear system on (Y, H, U). Then the pair $\left(\begin{smallmatrix}\mathcal{K}^*\\ \mathcal{F}^*\end{smallmatrix}\right)$ is called an *admissible stable output injection pair* for Ψ^* iff the extended system

$$
\Psi^*_{\mathrm{ext}} = \begin{bmatrix} \mathcal{A}^* & \begin{pmatrix}\mathcal{C}^* & \mathcal{K}^*\end{pmatrix} \\ \mathcal{B}^* & \begin{pmatrix}\mathcal{D}^* & \mathcal{F}^*\end{pmatrix}\end{bmatrix}
$$

is an externally stable anti-causal [causal] well-posed linear system, and $\left(\begin{smallmatrix}0\\ I\end{smallmatrix}\right)$ is an admissible stable output feedback operator for Ψ^*_{ext}, i.e., $I - \mathcal{F}^*$ has a bounded anti-causal [causal] inverse.

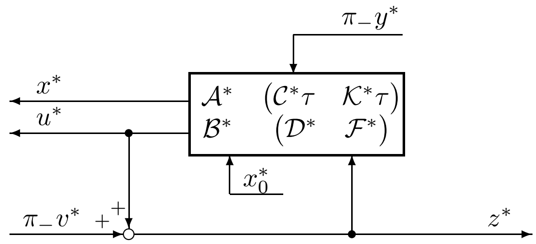

FIGURE 4. Adjoint Output Injection Connection

The appropriate formulas for the output injection connection are in the initial value setting

$$\begin{aligned} z^* &= u^* + \pi_- v^*, \\ x^*(s) &= \mathcal{A}^*(-s)x_0^* + \mathcal{C}^*\tau(s)\pi_- y^* + \mathcal{K}^*\tau(s)z^*, \\ u^* &= \mathcal{B}^* x_0^* + \mathcal{D}^*\pi_- y^* + \mathcal{F}^* z^*, \end{aligned} \tag{27}$$

and

$$\begin{aligned} z^* &= (I-\mathcal{F}^*)^{-1}\left(\mathcal{B}^* x_0^* + \mathcal{D}^*\pi_- y^* + \pi_- v^*\right), \\ x^*(s) &= \Big(\mathcal{A}^*(-s) + \mathcal{K}^*\tau(s)\,(I-\mathcal{F}^*)^{-1}\,\mathcal{B}^*\Big)\,x_0^* \\ &\quad + \Big(\mathcal{C}^* + \mathcal{K}^*\,(I-\mathcal{F}^*)^{-1}\,\mathcal{D}^*\Big)\,\tau(s)\pi_- y^* + \mathcal{K}^*\,(I-\mathcal{F}^*)^{-1}\,\tau(s)\pi_- v^*, \\ u^* &= (I-\mathcal{F}^*)^{-1}\left(\mathcal{B}^* x_0^* + \mathcal{D}^*\pi_- y^* + \mathcal{F}^*\pi_- v^*\right). \end{aligned} \tag{28}$$

We shall actually use this system in the time-invariant setting, for which slightly simpler formulas are valid (set $x_0^* = 0$ and delete π_-).

We get the closed loop output injection system $\Psi_{\circlearrowleft}^*$ by making the appropriate replacements in Proposition 20:

Proposition 26. *Let $\Psi^* = \left[\begin{smallmatrix} \mathcal{A}^* & \mathcal{C}^* \\ \mathcal{B}^* & \mathcal{D}^* \end{smallmatrix}\right]$ be an externally stable anti-causal well-posed linear system, and let $\left(\begin{smallmatrix} \mathcal{K}^* \\ \mathcal{F}^* \end{smallmatrix}\right)$ be an admissible stable output injection pair for Ψ. Then the system $\Psi_{\circlearrowleft}^*$ given by*

$$\begin{aligned} \Psi_{\circlearrowleft}^* &= \begin{bmatrix} \mathcal{A}_{\circlearrowleft}^* & \left(\mathcal{C}_{\circlearrowleft}^*\tau \quad \mathcal{K}_{\circlearrowleft}^*\tau\right) \\ \mathcal{B}_{\circlearrowleft}^* & \left(\mathcal{D}_{\circlearrowleft}^* \quad \mathcal{F}_{\circlearrowleft}^*\right) \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{A}^* + \mathcal{K}^*\tau^*\,(I-\mathcal{F}^*)^{-1}\,\mathcal{B}^* & \Big(\mathcal{C}^*\tau + \mathcal{K}^*\,(I-\mathcal{F}^*)^{-1}\,\mathcal{D}^*\tau \quad \mathcal{K}^*\,(I-\mathcal{F}^*)^{-1}\,\tau\Big) \\ (I-\mathcal{F}^*)^{-1}\,\mathcal{B}^* & \left((I-\mathcal{F}^*)^{-1}\,\mathcal{D}^* \qquad (I-\mathcal{F}^*)^{-1} - I\right) \end{bmatrix} \end{aligned}$$

is an externally stable anti-causal well-posed linear system. We call this system the closed loop output injection system with output injection pair $\left(\begin{smallmatrix} \mathcal{K}^* \\ \mathcal{F}^* \end{smallmatrix}\right)$. The controlled state $x^*(s)$ at time $s \le 0$ and output u^* of $\Psi_{\circlearrowleft}^*$ in the initial value setting with initial time zero, initial state x_0^*, and controls y^* and v^* are given by (28). (An analogous result is valid in the causal case.)

6. The Connection Between Quadratic Cost Minimization and Spectral Factorization

The following theorem is our first main result. It shows that the spectral factorization problem is equivalent to the problem of finding a state feedback solution to the quadratic cost minimization problem:

Theorem 27. *Let $\Psi = \left[\begin{smallmatrix}\mathcal{A} & \mathcal{B}\\ \mathcal{C} & \mathcal{D}\end{smallmatrix}\right]$ be an externally stable causal well-posed linear system on (U, H, Y), and let $\mathcal{D}$ be coercive (in the sense of Definition 12). For each $x_0 \in H$, let x_{opt}, u_{opt}, and y_{opt} be the optimal state, control, and output for the quadratic cost minimization problem (see Lemma 13). Let Π be the Riccati operator defined in Definition 14.*

(i) *If $\mathcal{Y}\mathcal{X}$ is an inner-outer factorization of $\mathcal{D}$ (in the sense of Definition 17), and if E is an arbitrary invertible operator in $\mathcal{L}(U)$, then*

$$\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix} = \begin{pmatrix}-\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C} & (I - E^{-1}\mathcal{X})\end{pmatrix}$$

is an admissible stable state feedback pair for Ψ, and

$$\begin{aligned}\begin{pmatrix}x_{\mathrm{opt}}\\ y_{\mathrm{opt}}\\ u_{\mathrm{opt}}\end{pmatrix} &= \begin{pmatrix}\mathcal{A}_\circlearrowleft\\ \mathcal{C}_\circlearrowleft\\ \mathcal{K}_\circlearrowleft\end{pmatrix} x_0 = \begin{pmatrix}\mathcal{A} + \mathcal{B}\mathcal{X}^{-1}\tau E\mathcal{K}\\ \mathcal{C} + \mathcal{Y}E\mathcal{K}\\ \mathcal{X}^{-1}E\mathcal{K}\end{pmatrix} x_0\\ &= \begin{pmatrix}\mathcal{A}\\ \mathcal{C}\\ 0\end{pmatrix} x_0 - \begin{pmatrix}\mathcal{B}\mathcal{X}^{-1}\tau\\ \mathcal{Y}\\ \mathcal{X}^{-1}\end{pmatrix} \pi_+\mathcal{Y}^*\mathcal{C}x_0\end{aligned}$$

is equal to the state and output of the closed loop system

$$\begin{aligned}\Psi_\circlearrowleft &= \begin{bmatrix}\mathcal{A}_\circlearrowleft & \mathcal{B}_\circlearrowleft\\ \begin{pmatrix}\mathcal{C}_\circlearrowleft\\ \mathcal{K}_\circlearrowleft\end{pmatrix} & \begin{pmatrix}\mathcal{D}_\circlearrowleft\\ \mathcal{F}_\circlearrowleft\end{pmatrix}\end{bmatrix} = \begin{bmatrix}\mathcal{A} + \mathcal{B}\mathcal{X}^{-1}\tau E\mathcal{K} & \mathcal{B}\mathcal{X}^{-1}E\\ \begin{pmatrix}\mathcal{C} + \mathcal{Y}E\mathcal{K}\\ \mathcal{X}^{-1}E\mathcal{K}\end{pmatrix} & \begin{pmatrix}\mathcal{Y}E\\ \mathcal{X}^{-1}E - I\end{pmatrix}\end{bmatrix}\\ &= \begin{bmatrix}\mathcal{A} - \mathcal{B}\mathcal{X}^{-1}\tau\pi_+\mathcal{Y}^*\mathcal{C} & \mathcal{B}\mathcal{X}^{-1}E\\ \begin{pmatrix}(I - \mathcal{Y}\pi_+\mathcal{Y}^*)\mathcal{C}\\ -\mathcal{X}^{-1}\pi_+\mathcal{Y}^*\mathcal{C}\end{pmatrix} & \begin{pmatrix}\mathcal{Y}E\\ \mathcal{X}^{-1}E - I\end{pmatrix}\end{bmatrix}\end{aligned} \tag{29}$$

with this feedback pair, initial value x_0, initial time zero, and zero control u. The Riccati operator Π of Ψ can be written in the following alternative forms:

$$\Pi = \mathcal{C}^*\mathcal{C} - \mathcal{K}^*E^*E\mathcal{K} = \mathcal{C}^*\left(I - \mathcal{Y}\pi_+\mathcal{Y}^*\right)\mathcal{C} = \mathcal{C}^*\mathcal{C}_\circlearrowleft = \mathcal{C}_\circlearrowleft^*\mathcal{C}_\circlearrowleft.$$

In particular, Π is the observability gramian of the closed loop system $\Psi_\circlearrowleft$.

(ii) *Conversely, if $\left(\begin{smallmatrix}y_{\mathrm{opt}}\\ u_{\mathrm{opt}}\end{smallmatrix}\right)$ is equal to the output of some externally stable closed loop state feedback extension $\Psi_\circlearrowleft$ of Ψ with initial value x_0, initial time 0, zero control u, and some admissible stable state feedback pair $\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix}$, then there exists an invertible operator $E \in \mathcal{L}(U)$ such that $\mathcal{Y}\mathcal{X}$ is an inner-outer factorization of $\mathcal{D}$, where $\mathcal{Y} = \mathcal{D}\left(I - \mathcal{F}\right)^{-1}E^{-1}$ and $\mathcal{X} = E\left(I - \mathcal{F}\right)$. Moreover, $\mathcal{K}$ is given by $\mathcal{K} = -\pi_+E^{-1}\mathcal{Y}^*\mathcal{C}$. Thus, every state feedback solution of the quadratic cost minimization problem is of the type described in part (i).*

Remark 28. The operator $\mathcal{Y}\pi_+\mathcal{Y}^*$ in the formulas above is equal to the projection operator P in Lemma 13. The operator $I - \mathcal{Y}\pi_+\mathcal{Y}^*$ can be interpreted as the Riccati operator of a particular realization of $\mathcal{D}$; see [37].

The explanation for the appearance of the extra undetermined operator E in Theorem 27 is the following. In the case of the classical system (1) a standard state feedback law has the form $u = Lx$ for some bounded linear operator L. The right-hand side of this feedback equation can be thought of as an extra output of the original system with the property that the feed-through term has been normalized to be zero. As we mentioned in connection with the formula for the output feedback connection, it is not in general possible to normalize the feed-through term of a well-posed linear system to be zero; to do this we need the extra regularity condition introduced in Definition 35. If we do not normalize the feed-through term in the feedback loop to be zero, the we must write the feedback equation in the form $u = Kx + Fu$ for some bounded linear operators K and F. This equation is equivalent to the earlier equation $u = Lx$ iff $I - F$ is invertible, and $L = (I-F)^{-1}K$. Define $E = (I-F)^{-1}$. Then the feedback equation $u = Kx + Fu$ becomes $u = E^{-1}Lx + (I - E^{-1})u$, which begins to resemble the actual formula used in Theorem 27. Notice, in particular, that the "effective" feedback is $L = EK$ instead of K.

For later reference, let us write out the expression for the extended open loop system Ψ_{ext} in Definition 22 with the particular state feedback pair $\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix} = \begin{pmatrix}-\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C} & (I - E^{-1}\mathcal{X})\end{pmatrix}$ given by Theorem 27, and also the expressions for the adjoints of Ψ_{ext} and $\Psi_{\circlearrowleft}$:

$$\Psi_{\text{ext}} = \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \begin{pmatrix}\mathcal{C}\\ \mathcal{K}\end{pmatrix} & \begin{pmatrix}\mathcal{D}\\ \mathcal{F}\end{pmatrix}\end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \begin{pmatrix}\mathcal{C}\\ -\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C}\end{pmatrix} & \begin{pmatrix}\mathcal{D}\\ I - E^{-1}\mathcal{X}\end{pmatrix}\end{bmatrix}, \tag{30}$$

$$\Psi_{\text{ext}}^* = \begin{bmatrix} \mathcal{A}^* & \begin{pmatrix}\mathcal{C}^* & \mathcal{K}^*\end{pmatrix} \\ \mathcal{B}^* & \begin{pmatrix}\mathcal{D}^* & \mathcal{F}^*\end{pmatrix}\end{bmatrix} = \begin{bmatrix} \mathcal{A}^* & \begin{pmatrix}\mathcal{C}^* & -\mathcal{C}^*\mathcal{Y}\left(E^*\right)^{-1}\pi_+\end{pmatrix} \\ \mathcal{B}^* & \begin{pmatrix}\mathcal{D}^* & I - \mathcal{X}^*\left(E^*\right)^{-1}\end{pmatrix}\end{bmatrix}, \tag{31}$$

$$\begin{aligned}\Psi_{\circlearrowleft}^* &= \begin{bmatrix} \mathcal{A}_{\circlearrowleft}^* & \begin{pmatrix}\mathcal{C}_{\circlearrowleft}^* & \mathcal{K}_{\circlearrowleft}^*\end{pmatrix} \\ \mathcal{B}_{\circlearrowleft}^* & \begin{pmatrix}\mathcal{D}_{\circlearrowleft}^* & \mathcal{F}_{\circlearrowleft}^*\end{pmatrix}\end{bmatrix} \\ &= \begin{bmatrix} \mathcal{A}^* - \mathcal{C}^*\mathcal{Y}\pi_+\tau^*\left(\mathcal{X}^*\right)^{-1}\mathcal{B}^* & \begin{pmatrix}\mathcal{C}^*\left(I - \mathcal{Y}\pi_+\mathcal{Y}^*\right) & -\mathcal{C}^*\mathcal{Y}\pi_+\left(\mathcal{X}^*\right)^{-1}\end{pmatrix} \\ E^*\left(\mathcal{X}^*\right)^{-1}\mathcal{B}^* & \begin{pmatrix}E^*\mathcal{Y}^* & E^*\left(\mathcal{X}^*\right)^{-1} - I\end{pmatrix}\end{bmatrix}.\end{aligned} \tag{32}$$

Proof of Theorem 27. (i) We suppose that $\mathcal{Y}\mathcal{X}$ is an inner-outer factorization of $\mathcal{D}$. Let $E \in \mathcal{L}(H)$ be invertible, and define

$$\begin{pmatrix}\mathcal{K} & \mathcal{F}\end{pmatrix} = \begin{pmatrix}-\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C} & (I - E^{-1}\mathcal{X})\end{pmatrix}.$$

We claim that the system Ψ_{ext} defined in (30) is an externally stable well-posed linear system on $(U, H, Y \times U)$. It is obvious that all the operators in this system are causal, bounded linear operators on their appropriate domains. The only non-obvious parts of the claim that Ψ_{ext} is well-posed are that

$$\mathcal{K}\mathcal{A}(t) = \pi_+\tau(t)\mathcal{K}, \qquad t \in \mathbf{R}^+,$$

and that the Hankel operator of $\mathcal{F}$ is equal to $\mathcal{KB}$. To compute $\mathcal{KA}(t)$, we recall that $E^{-1}\mathcal{Y}^*$ is time-invariant and anti-causal; hence

$$\begin{aligned}
\mathcal{KA}(t) &= -\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C}\mathcal{A}(t) \\
&= -\pi_+ E^{-1}\mathcal{Y}^*\pi_+\tau(t)\mathcal{C} \\
&= -\pi_+ E^{-1}\mathcal{Y}^*\tau(t)\mathcal{C} \\
&= -\pi_+\tau(t) E^{-1}\mathcal{Y}^*\mathcal{C} \\
&= -\pi_+\tau(t)\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C} \\
&= \pi_+\tau(t)\mathcal{K}.
\end{aligned}$$

In the computation of the Hankel operator of $\mathcal{F}$, we use the facts that $\mathcal{F}$ and $-E^{-1}\mathcal{X}$ have the same Hankel operator, that $\mathcal{X} = \mathcal{Y}^*\mathcal{Y}\mathcal{X} = \mathcal{Y}^*\mathcal{D}$, that $\pi_+ + \pi_- = I$, that $E^{-1}\mathcal{Y}^*$ is anti-causal, and that the Hankel operator of $\mathcal{D}$ is $\mathcal{CB}$:

$$\begin{aligned}
\pi_+\mathcal{F}\pi_- &= -\pi_+ E^{-1}\mathcal{X}\pi_- \\
&= -\pi_+ E^{-1}\mathcal{Y}^*\mathcal{D}\pi_- \\
&= -\pi_+ E^{-1}\mathcal{Y}^*\pi_+\mathcal{D}\pi_- \\
&= -\pi_+ E^{-1}\mathcal{Y}^*\mathcal{C}\mathcal{B} \\
&= \mathcal{KB}.
\end{aligned}$$

To complete the proof of (i) it suffices to insert the formula for $(\pi_+\mathcal{D}^*\mathcal{D}\pi_+)^{-1}$ given in part (iv) of Lemma 18 into the formulas of Lemma 13, and to compare the result to the output of the closed loop system $\Psi_{\circlearrowleft}$ with initial time zero, initial state x_0, and zero control u.

(ii) Conversely, suppose that $\left(\begin{smallmatrix} y_{\mathrm{opt}} \\ u_{\mathrm{opt}} \end{smallmatrix}\right)$ is equal to the output of some externally stable closed loop state feedback extension $\Psi_{\circlearrowleft}$ of Ψ with initial value x_0, initial time 0, zero control u, and some admissible stable state feedback pair $\begin{pmatrix} \mathcal{K} & \mathcal{F} \end{pmatrix}$. According to Proposition 23,

$$\begin{aligned}
y_{\mathrm{opt}} &= \mathcal{C}_{\circlearrowleft} x_0 = \left(\mathcal{C} + \mathcal{D}\,(I-\mathcal{F})^{-1}\,\mathcal{K}\right) x_0, \\
u_{\mathrm{opt}} &= \mathcal{K}_{\circlearrowleft} x_0 = (I-\mathcal{F})^{-1}\,\mathcal{K} x_0.
\end{aligned} \tag{33}$$

The first step of this proof is to show that it is possible to write $\mathcal{D}^*\mathcal{D}$ in the form $\mathcal{Z}^*\mathcal{W}$, where both $\mathcal{Z}$ and $\mathcal{W}$ are causal and $\mathcal{W}$ has a bounded causal inverse. Motivated by the formula $\mathcal{F} = I - E^{-1}\mathcal{X}$ in part (i), we take $E = I$, and choose $\mathcal{W}$ and $\mathcal{Z}^*$ to be

$$\mathcal{W} = I - \mathcal{F}, \qquad \mathcal{Z}^* = \mathcal{D}^*\mathcal{D}\mathcal{W}^{-1} = \mathcal{D}^*\mathcal{D}\,(I-\mathcal{F})^{-1}.$$

Since $\begin{pmatrix} \mathcal{K} & \mathcal{F} \end{pmatrix}$ is an admissible stable state feedback pair for Ψ, $\mathcal{W} = I - \mathcal{F}$ is causal and has a bounded causal inverse (which makes it possible to define $\mathcal{Z}^*$ in the way that we did above).

We claim that $\mathcal{Z}^*$ is anti-causal. To prove this we choose an arbitrary $u \in L^2(\mathbf{R}; U)$, and choose x_0 in the quadratic cost minimization problem to be $x_0 = \mathcal{B}_{\circlearrowleft} u$. Recall that the Hankel operator of $\mathcal{D}_{\circlearrowleft} = \mathcal{D}\,(I-\mathcal{F})^{-1}$ is $\mathcal{C}_{\circlearrowleft}\mathcal{B}_{\circlearrowleft}$ since $\Psi_{\circlearrowleft}$ is a well-posed system, and compute

$$\begin{aligned}
y_{\mathrm{opt}} &= \mathcal{C}_{\circlearrowleft} x_0 = \mathcal{C}_{\circlearrowleft}\mathcal{B}_{\circlearrowleft} u \\
&= \pi_+\mathcal{D}_{\circlearrowleft}\pi_- u = \pi_+\mathcal{D}\,(I-\mathcal{F})^{-1}\,\pi_- u.
\end{aligned}$$

By (21), $\pi_+\mathcal{D}^* y_{\rm opt} = 0$, and this together with the computation above and the anti-causality of $\mathcal{D}^*$ shows that $\pi_+\mathcal{Z}^*\pi_- = \pi_+\mathcal{D}^*\mathcal{D}(I-\mathcal{F})^{-1}\pi_- = 0$, i.e., $\mathcal{Z}^*$ is anti-causal.

By now we know that $\mathcal{Z}^*\mathcal{W} = \mathcal{D}^*\mathcal{D}$, that $\mathcal{Z}^*$ is anti-causal, that $\mathcal{W}$ is causal and invertible, and that the inverse of $\mathcal{W}$ is causal. Taking adjoints in the equation $\mathcal{Z}^*\mathcal{W} = \mathcal{D}^*\mathcal{D}$, we get

$$\mathcal{W}^*\mathcal{Z} = \mathcal{D}^*\mathcal{D} = \mathcal{Z}^*\mathcal{W};$$

hence

$$\mathcal{Z}\mathcal{W}^{-1} = (\mathcal{W}^*)^{-1}\mathcal{D}^*\mathcal{D}\mathcal{W}^{-1} = (\mathcal{W}^*)^{-1}\mathcal{Z}^*.$$

Since $\mathcal{Z}\mathcal{W}^{-1}$ is causal and $(\mathcal{W}^*)^{-1}\mathcal{Z}^*$ is anti-causal, the operator $(\mathcal{W}^*)^{-1}\mathcal{D}^*\mathcal{D}\mathcal{W}^{-1}$ is static. By Lemma 6, it is equal to a multiplication operator by an operator $M \in \mathcal{L}(U)$, i.e., $(\mathcal{W}^*)^{-1}\mathcal{D}^*\mathcal{D}\mathcal{W}^{-1} = M$ for some $M \in \mathcal{L}(U)$. The operator M must be positive and invertible since $(\mathcal{W}^*)^{-1}\mathcal{D}^*\mathcal{D}\mathcal{W}^{-1}$ is positive and invertible. Let $E \in \mathcal{L}(U)$ be an arbitrary invertible operator satisfying $E^*E = M$; for example, we may take E to be the positive square root of M. Then $(\mathcal{W}^*)^{-1}\mathcal{D}^*\mathcal{D}\mathcal{W}^{-1} = E^*E$, i.e.,

$$\mathcal{D}^*\mathcal{D} = \mathcal{W}^*E^*E\mathcal{W} = (I-\mathcal{F}^*)E^*E(I-\mathcal{F}).$$

Consequently, if we define

$$\mathcal{X} = E(I-\mathcal{F}),$$

then $\mathcal{X}$ is a spectral factor of $\mathcal{D}^*\mathcal{D}$, and $\mathcal{Y}\mathcal{X} = \left(\mathcal{D}(I-\mathcal{F})^{-1}E^{-1}\right)(E(I-\mathcal{F}))$ is an inner-outer factorization of $\mathcal{D}$.

To prove that $\mathcal{K}$ must be of the form $\mathcal{K} = -\pi_+E^{-1}\mathcal{Y}^*\mathcal{C}$, we observe that by (21) and (33),

$$\pi_+\mathcal{D}^*\left(\mathcal{C} + \mathcal{D}(I-\mathcal{F})^{-1}\mathcal{K}\right) = 0.$$

Replacing $\mathcal{D}^*\mathcal{D}$ by $\mathcal{X}^*\mathcal{X} = (I-\mathcal{F}^*)E^*E(I-\mathcal{F})$, we get

$$\pi_+\mathcal{X}^*E\mathcal{K} = -\pi_+\mathcal{D}^*\mathcal{C},$$

which we can solve for $\mathcal{K}$ (since $E^{-1}(\mathcal{X}^*)^{-1}$ is anti-causal) in the form

$$\mathcal{K} = -\pi_+E^{-1}(\mathcal{X}^*)^{-1}\mathcal{D}^*\mathcal{C} = -\pi_+E^{-1}\mathcal{Y}^*\mathcal{C}. \qquad \square$$

7. The Generators of a Well-Posed Linear system

By a result due to Salamon [35] (and apparently discovered independently by Weiss [43, 44]), every well-posed linear system Ψ has a well-defined (unbounded) control operator B and a well-defined (unbounded) observation operator C, and formulas (3), (4), (8), (9), (11), (12), (13), and (15) hold in a weak sense (see Remark 30 below). In order to present this result we need some additional definitions.

Let $\mathcal{A}$ be the semigroup of the well-posed linear system Ψ on (U, H, Y), i.e., $\mathcal{A}$ is a semigroup on H. Denote the generator of $\mathcal{A}$ by A, and denote the domain $\mathrm{dom}(A)$ of A by W. Choose an arbitrary number α from the resolvent set of A. Then $W = (\alpha I - A)^{-1}H$, and we can choose the norm in W to be $\|x\|_W = \|(\alpha I - A)x\|_H$. Let V be the completion of H under the norm $\left\|(\alpha I - A)^{-1}x\right\|_H$. Then

$$W \subset H \subset V$$

with dense and continuous embeddings, $(\alpha I - A)$ is an (isometric) isomorphism of W onto H, and $(\alpha I - A)$ extends to an (isometric) isomorphism of H onto V. The semigroup $\mathcal{A}$ can be restricted to a semigroup on W, and extended to a semigroup on V. The three semigroups that we get in this way are (isometrically) isomorphic, and we denote them all with the same letter $\mathcal{A}$.

We repeat the same construction with $\mathcal{A}$ and A replaced by their adjoints $\mathcal{A}^*$ and A^* to get two more spaces $V^* = \operatorname{dom}(A^*)$ and W^*, with

$$V^* \subset H \subset W^*.$$

It is possible to identify W^* with the dual of W and V with the dual of V^* if we use H as the pivot space.

Proposition 29. *Let $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$ be a causal and $\Psi^* = \left[\begin{smallmatrix} \mathcal{A}^* & \mathcal{C}^* \\ \mathcal{B}^* & \mathcal{D}^* \end{smallmatrix}\right]$ an anti-causal externally stable well-posed linear system on (U,H,Y), respectively on (Y,H,U). These systems have (unique) generating operators $\left[\begin{smallmatrix} A & B \\ C & ? \end{smallmatrix}\right]$, respectively $\left[\begin{smallmatrix} A^* & C^* \\ B^* & ? \end{smallmatrix}\right]$, with the following properties (the question mark stands for a missing entry):*

- (i) *$A \in \mathcal{L}(W;H) \cap \mathcal{L}(H;V)$ is the generator of $\mathcal{A}$, $A^* \in \mathcal{L}(V^*;H) \cap \mathcal{L}(H;W^*)$ is the generator of $\mathcal{A}^*$, $B \in \mathcal{L}(U;V)$, $C^* \in \mathcal{L}(Y;W^*)$, $C \in \mathcal{L}(W;Y)$, and $B^* \in \mathcal{L}(V^*;U)$. Moreover, for every sufficiently large real number α, $\alpha I - A$ has an inverse in $\mathcal{L}(H;W) \cap \mathcal{L}(V;H)$, and $\alpha I - A^*$ has an inverse in $\mathcal{L}(H;V^*) \cap \mathcal{L}(W^*;H)$.*
- (ii) *$\mathcal{C} \in \mathcal{L}(W;W^{1,2}(\mathbf{R}^+;Y))$, $\mathcal{B}^* \in \mathcal{L}(V^*;W^{1,2}(\mathbf{R}^+;U))$, and (4) and (9) hold for every $x \in W$ and $x^* \in V^*$. Moreover, for such x and x^*, $(\mathcal{C}x)' = \mathcal{C}Ax$ and $(\mathcal{B}^*x^*)' = -\mathcal{B}^*A^*x^*$.*
- (iii) *In the time-invariant setting, if $u \in W^{1,2}(\mathbf{R};U)$, then the controlled state $x = \mathcal{B}\tau u$ and the output $y = \mathcal{D}u$ of Ψ satisfy $x \in BC^1(\mathbf{R};H)$, $y \in W^{1,2}(\mathbf{R};Y)$, $x' = \mathcal{B}\tau u' = Ax + Bu \in BC(\mathbf{R};H)$, $y' = \mathcal{D}u'$, and $x(t) \to 0$ and $x'(t) \to 0$ in H as $t \to -\infty$. Analogously, in the time-invariant setting, if $y^* \in W^{1,2}(\mathbf{R};Y)$, then the controlled state $x^* = \mathcal{C}^*\tau y^*$ and the output $u^* = \mathcal{D}^*y^*$ of Ψ^* satisfy $x^* \in BC^1(\mathbf{R};H)$, $u^* \in W^{1,2}(\mathbf{R};U)$, $(x^*)' = \mathcal{C}^*\tau\,(y^*)' = -A^*x^* - C^*y^* \in BC(\mathbf{R};H)$, $(u^*)' = \mathcal{D}^*\,(y^*)'$, and $x^*(s) \to 0$ and $(x^*)'(s) \to 0$ in H as $s \to \infty$.*
- (iv) *In the initial value setting, if $u \in W^{1,2}(\mathbf{R}^+;U)$, $x_0 \in H$, and $Ax_0 + Bu(0) \in H$, then the controlled state $x = \mathcal{A}x_0 + \mathcal{B}\tau\pi_+u$ and the output $y = \mathcal{C}x_0 + \mathcal{D}\pi_+u$ of Ψ satisfy*

$$x \in C^1(\mathbf{R}^+;H), \qquad y \in W^{1,2}(\mathbf{R}^+;Y),$$
$$x' = \mathcal{A}(Ax_0 + Bu(0)) + \mathcal{B}\tau u' = Ax + Bu \in C(\mathbf{R}^+;H),$$

$x(t)$ is given by (13), and $y' = \mathcal{C}\,(Ax_0 + Bu(0)) + \mathcal{D}u'$. Analogously, in the initial value setting, if $y^* \in W^{1,2}(\mathbf{R}^-;Y)$, $x_0^* \in H$, and $A^*x_0^* + C^*y^*(0) \in H$, then the controlled state $x^*(s) = \mathcal{A}^*(-s)x_0^* + \mathcal{C}^*\tau(s)\pi_-y^*$ and the output $u^* = \mathcal{B}^*x_0^* + \mathcal{D}^*\pi_-y^*$ of Ψ^* satisfy

$$x^* \in C^1(\mathbf{R}^-;H), \qquad u^* \in W^{1,2}(\mathbf{R}^+;U),$$
$$\begin{aligned}(x^*)'(s) &= -\mathcal{A}^*(-s)(A^*x_0^* + C^*y^*(0)) + \mathcal{C}^*\tau(s)\,(y^*)' \\ &= -A^*x^* - C^*y^* \in C(\mathbf{R}^-;H),\end{aligned}$$

$x^*(s)$ is given by (15), and $(u^*)' = -\mathcal{B}^*\,(A^*x_0^* + C^*y^*(0)) + \mathcal{D}^*\,(y^*)'$.

Outline of Proof. Most of these claims follow directly from [34, Lemmas 2.3 and 2.5] and [35, Theorem 3.1 and Lemma 3.2]. The key step is to prove that $B \in \mathcal{L}(U;V)$ and $C \in \mathcal{L}(W;Y)$ and that (4) and (13) hold in the appropriate spaces, because that enables us to regard $\left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ 0 & 0 \end{smallmatrix}\right]$ as a classical system on $(U, V, ?)$ and $\left[\begin{smallmatrix} \mathcal{A} & 0 \\ \mathcal{C} & 0 \end{smallmatrix}\right]$ as a classical system on $(?, W, Y)$ (the question marks represent irrelevant spaces). This makes it is possible to use standard semigroup theory as presented in, e.g., [29]. To show that $x(t) \to 0$ in H as $t \to -\infty$ in (iii) is suffices to observe that

$$\begin{aligned} \|x(t)\|_H &= \|\mathcal{B}\tau(t)u\|_H \\ &= \|\mathcal{B}\pi_-\tau(t)u\|_H \\ &= \|\mathcal{B}\tau(t)\tau(-t)\pi_-\tau(t)u\|_H \\ &= \left\|\mathcal{B}\tau(t)\pi_{(-\infty,t]}u\right\|_H \\ &\le \|\mathcal{B}\|_{\mathcal{L}(L^2(\mathbf{R};U);H)} \left\|\pi_{(-\infty,t]}u\right\|_{L^2(\mathbf{R};U)}, \end{aligned}$$

which tends to zero as $t \to -\infty$. A similar proof with u replaced by u' shows that $x'(t) \to 0$ in H as $t \to -\infty$. To deduce the statements about the adjoint system Ψ^* it suffices to reflect the time axis and to apply the corresponding result for Ψ. □

Remark 30. According to Proposition 29, (4) holds for each $x \in W$, and (13) holds as an equation in H, provided that the integral is interpreted as a Bochner integral in V. To derive a weak version of (11) we argue as follows. Fix some $u \in L^2(\mathbf{R};U)$, and define $u_T = \pi_{[T,\infty)}u$ and $x_T = \mathcal{B}\tau u_T$. Then $u - u_T = \pi_{(-\infty,T)}u \to 0$ in $L^2(\mathbf{R};U)$ as $T \to -\infty$, so $x_T(t) \to x(t) = \mathcal{B}\tau(t)u$ in H, uniformly in t. Moreover, $x_T(t) = 0$ for $t \le T$. Since we can regard $\left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ 0 & 0 \end{smallmatrix}\right]$ as a classical system on $(U, V, ?)$ (the output space is irrelevant), it follows from (2) that

$$x_T(t) = \int_T^t \mathcal{A}(t-s)Bu(s)\,ds, \qquad t \ge T.$$

We conclude that (11) holds in the sense that

$$x(t) = \lim_{T\to-\infty} \int_T^t \mathcal{A}(t-s)Bu(s)\,ds, \quad t \in \mathbf{R},$$

where the integral is computed in V, but the limit is taken in H and is uniform in t (we interpret the integral as zero if $t < T$). In particular, taking $t = 0$ we find that

$$\mathcal{B}u = \lim_{T\to-\infty} \int_T^0 \mathcal{A}(-s)Bu(s)\,ds,$$

where the integral is computed in V but the limit is taken in H. Analogous interpretations are valid for the adjoint equations (8), (9), (12), and (15).

Proposition 29 is not quite sufficient for our purposes. We need a small enhancement, which involves some additional Hilbert spaces. According to Proposition 29, the values of the states x and x^* and their derivatives lie in H, but we need to know that the states actually lie in some smaller spaces. These smaller spaces are not uniquely determined. We choose to follow Salamon and use the smallest possible spaces, instead of the larger spaces introduced by G. Weiss. These smallest possible spaces are "range spaces" of certain mappings, constructed as follows:

Lemma 31. *Let $E \in \mathcal{L}(X;Y)$, where X and Y are Banach spaces. For each $y \in \operatorname{range}(E)$, let $\|y\|_Z = \inf\{\,\|x\|_X \mid y = Ex\,\}$. This defines a norm on* range(E) *that makes* range(E) *a Banach space Z, and this norm is (not necessarily strictly) stronger than the norm of Y. The operator E induces an isometric isomorphism of the quotient $X/\ker(E)$ onto Z, and E is continuous and open from X onto Z. If X is a Hilbert space, then so is Z, and then E is an isometric isomorphism of $\ker(E)^\perp$ onto Z.*

This result is standard but difficult to find in the literature. The idea is to factor E into $E = F\pi$, where $\pi\colon X \to X/\ker(E)$ is the (continuous and open) quotient map, and $F\colon X/\ker(E)$ is injective [3, Problem F, p. 266], and to let the norm in $\operatorname{range}(E) = \operatorname{range}(F)$ be the one induced by F. This makes F an isometric isomorphism of $X/\ker(E)$ onto Z.

Let Ψ be a well-posed linear system on (U, H, Y) with generating operators $\left[\begin{smallmatrix} A & B \\ C & ? \end{smallmatrix}\right]$. Choose an arbitrary number α from the resolvent sets of A. We denote the range space (constructed as in Lemma 31) of the mapping $(\alpha I - A)^{-1}\begin{pmatrix} I & B\end{pmatrix} : H \times U \to H$ by W_B, and the corresponding space that we get by replacing Ψ by its adjoint by V_C^*. These spaces are independent of the particular value of α.

The elements of W_B and V_C^* can be characterized in the following way:

Lemma 32. *The following conditions are equivalent:*

(i) $x \in W_B$;
(ii) $x = (\alpha I - A)^{-1}(x_1 + Bu)$ *for some* $x_1 \in H$ *and* $u \in U$;
(iii) $x = x_2 + (\alpha I - A)^{-1}Bu$ *for some* $x_2 \in W$ *and* $u \in U$;
(iv) $x \in H$ *and* $Ax + Bu \in H$ *for some* $u \in U$.

Likewise, the following conditions are also equivalent:

(v) $x^* \in V_C^*$;
(vi) $x^* = (\alpha I - A^*)^{-1}(x_1^* + C^*y^*)$ *for some* $x_1^* \in H$ *and* $y^* \in Y$;
(vii) $x^* = x_2^* + (\alpha I - A^*)^{-1}C^*y^*$ *for some* $x_2^* \in V^*$ *and* $y^* \in Y$;
(viii) $x^* \in H$ *and* $A^*x^* + C^*y^* \in H$ *for some* $y^* \in Y$.

Proof. The equivalence of (i) and (ii) follows from the definition of W_B. Take $x_2 = (\alpha I - A)^{-1}x_1$ to show that (ii) and (iii) are equivalent. If x is of the form (ii), then $x \in H$ and $Ax + Bu = -x_1 + \alpha(\alpha I - A)^{-1}(x_1 + Bu) \in H$, so (iv) holds. Conversely, if (iv) holds, then (ii) is satisfied with $x_1 = \alpha x - (Ax + Bu)$. The equivalence of (v)-(viii) is proved analogously. $\square$

It is easy to see that

$$W \subset W_B \subset H \text{ and } V^* \subset V_C^* \subset H,$$

with continuous injections.

Obviously, W_B and V_C^* are dense in H since W and V^* are dense in H. However, W need not be dense in W_B, and V^* need not be dense in V_C^*. In particular, in the realizations discussed in [37] and [38] W is a closed subspace of W_B and V^* is a closed subspace of V_C^*. For example, the realization that is denoted by Ξ^- in [37] is Salamon's exactly controllable realization which we get by taking H to be $H = L^2(\mathbf{R}^-;U)$, $\mathcal{A}(t)$ to be the unilateral left-shift $\tau(t)\pi_-$ on $L^2(\mathbf{R}^-;U)$, $\mathcal{B}u = \pi_- u$, and $\mathcal{C}x = \pi_+\mathcal{D}x$ [39, Definition 2.10]. For this realization $W = W_0^{1,2}(\mathbf{R}^-;U)$ and $W_B = W^{1,2}(\mathbf{R}^-;U)$, so W is a closed subspace of W_B. Salamon introduces the space W_B in [34] (denoting it by Z) in the special case where the mapping

$(\alpha I - A)^{-1}\begin{pmatrix} I & B \end{pmatrix}$ is one-to-one (a case that he refers to as an abstract boundary control system; the exactly controllable realization is of this type).

Proposition 29 implies the following result:

Proposition 33. *The conclusion of Proposition 29 can be strengthened as follows:*

(i) *In part (iii), $x \in BC(\mathbf{R}; W_B)$, $x^* \in BC(\mathbf{R}; V_C^*)$, $x(t) \to 0$ in W_B as $t \to -\infty$, and $x^*(s) \to 0$ in V_C^* as $s \to +\infty$.*

(ii) *In part (iv), $x \in C(\mathbf{R}^+; W_B)$ and $x^* \in C(\mathbf{R}^+; V_C^*)$.*

Proof. Use the fact that $x' = Ax + Bu$ to write x in the form

$$x = (\alpha I - A)^{-1}(\alpha x - x' + Bu).$$

Since x and x' are continuous in H and u is continuous in U, this implies that x is continuous in W_B. The other claims are proved in a similar way. ☐

Remark 34. By Lemma 32, the spaces W_B and V_C^* can be characterized as the spaces of permitted initial values x_0 and x_0^* in part (iv) of Proposition 29. By Proposition 33, they are invariant in the sense that the states x and x^* in parts (iii) and (iv) of Proposition 29 stay in these spaces for all time. Moreover, by Proposition 37 below, they are large enough to contain the domains of all state feedback perturbations of A and A^*.

In general, a well-posed linear system need not have a well-defined feed-through operator D, so equations (5), (10), (14), and (16) cannot possibly be true in general. This fact motivated G. Weiss to introduce the class of *regular* systems. In [45, Theorem 5.8] Weiss gives eight equivalent characterizations of regularity, one of which is the following:

Definition 35. A causal time-invariant operator $\mathcal{D}\colon L^2(\mathbf{R}; U) \to L^2(\mathbf{R}; Y)$ is *regular* iff, for every $u_0 \in U$, the strong Cesàro mean of order one

$$Du_0 = \lim_{t \to 0+} \frac{1}{t} \int_0^t (\mathcal{D}\chi_{\mathbf{R}^+} u_0)(s)\, ds$$

exists (here $\chi_{\mathbf{R}^+}$ stands for the characteristic function of $\mathbf{R}^+$, and $\mathcal{D}\chi_{\mathbf{R}^+}u_0$ is the step response corresponding to the constant input u_0). An anti-causal time-invariant operator $\mathcal{D}^*\colon L^2(\mathbf{R}; Y) \to L^2(\mathbf{R}; U)$ is *regular* iff, for every $y_0^* \in Y$, the strong Cesàro mean of order one

$$D^* y_0^* = \lim_{s \to 0-} \frac{1}{s} \int_0^s (\mathcal{D}^* \chi_{\mathbf{R}^-} y_0^*)(t)\, dt$$

exists (here $\chi_{\mathbf{R}^-}$ stands for the characteristic function of $\mathbf{R}^-$, and $\mathcal{D}^*\chi_{\mathbf{R}^+}y_0^*$ is the step response corresponding to the constant input y_0^*). The operators $D\colon U \to Y$ and $D^*\colon Y \to U$ defined above are called the feed-through operators of $\mathcal{D}$ and $\mathcal{D}^*$. The systems Ψ and Ψ^* are regular iff their input/output maps are regular.

Thus, a system is regular iff its step response has a one-sided limit in the strong C_1-sense at the origin. It is known [46, p. 38] that the adjoint of a regular input/output map need not be regular.

Let us rewrite equations (5), (10), (14), and (16) in terms of the controlled states $x(t)$ and $x^*(s)$ of Ψ and Ψ^* in the common forms

$$\begin{aligned} y(t) &= Cx(t) + Du(t), \\ u^*(s) &= B^*x^*(s) + D^*y^*(s). \end{aligned}$$

Then, as Weiss proved, for regular systems these equations remain valid in a weak sense:

Proposition 36. *Let Ψ and Ψ^* be well-posed linear systems on (U, H, Y), respectively, (Y, H, U), with generating operators $\left[\begin{smallmatrix} A & B \\ C & ? \end{smallmatrix}\right]$, respectively $\left[\begin{smallmatrix} A^* & C^* \\ B^* & ? \end{smallmatrix}\right]$.*

(i) *Ψ is regular if and only if the strong Abel limit*

$$\overline{C}x = \lim_{\lambda \to +\infty} C\lambda(\lambda I - A)^{-1}x$$

exists for every $x \in W_B$, and Ψ^* is regular if and only if the strong Abel limit

$$\overline{B}^* x^* = \lim_{\lambda \to +\infty} B^*\lambda(\lambda I - A^*)^{-1}x^*$$

exists for every $x^* \in V_C^*$. The operators $\overline{C}$ and $\overline{B}^*$ defined in this way are bounded linear operators from W_B into Y and from V_C^* into U, respectively, and they are extensions of the operators $C \in \mathcal{L}(W;Y)$ and $B^* \in \mathcal{L}(V^*;U)$.

(ii) *In the regular case, denote the feed-through operators of Ψ and Ψ^* by D and D^*, respectively. Then the conclusion of Proposition 29 can be strengthened as follows: Both in parts (iii) and (iv), $y = \overline{C}x + Du$ and $u^* = \overline{B}^* x^* + D^* y^*$.*

Proof. Part (i) follows easily from the equivalence of (1) and (4) in [45, Theorem 5.8], and Part (ii) from [45, Remark 6.2]. $\square$

Let us remark that our operator $\overline{C}$ is a restriction of the operator that Weiss denotes by $\widetilde{C}_L$ in [45] and by C_Λ in [46], and also a restriction of Weiss' operator C_L. He does not explicitly study $\overline{B}^*$ and the time-invariant setting.

Since W and V^* need not be dense in W_B, respectively V_C^*, the extensions $\overline{C}$ and $\overline{B}^*$ are not uniquely determined by Weiss C and B^*. Salamon uses a different extension K of C in [34, Section 2.2]. As shown in [40], some of our final results can be extended to the non-regular case through the use of Salamon's extension of C and the corresponding extension of B^* instead of our extensions $\overline{C}$ and $\overline{B}^*$ (induced by Weiss' extensions).

In this work we make no explicit use of the notion of the transfer function of the input/output map of a well-posed linear system, and we refer the reader to [39], [45], [46], and [50] for discussions of these functions. In our setting, the transfer function of $\mathcal{D}$ is an H^∞-function over the right half-plane, and the transfer function of $\mathcal{D}^*$ is an H^∞-function over the left half-plane (in Weiss' setting, both of these are defined on right-half planes).

8. The Generators of the Closed Loop System

Consider the closed loop output feedback system Ψ_L introduced in Proposition 20. We denote the generating operators of this system by $\left[\begin{smallmatrix} A_L & B_L \\ C_L & ? \end{smallmatrix}\right]$, and in the regular case by $\left[\begin{smallmatrix} A_L & B_L \\ C_L & D_L \end{smallmatrix}\right]$. In the classical case we know that these generators are given by (24), and Salamon and Weiss have extended this relation to general well-posed closed loop linear output feedback systems.

The generator A_L of the closed loop semigroup S_L was computed by Salamon in [34, Section 4]. His description of the domain $\mathrm{dom}(A_L)$ of A_L was incomplete in the case where the observation operator is unbounded and the control space U infinite-dimensional, due to the fact that Salamon never makes the connection between the well-posedness of the closed loop system and the invertibility of the loop gain transfer function in some right-half plane that Weiss makes in [46, Proposition 3.6].

However, by combining [34, Lemma 4.4(iii)] with [46, Proposition 3.6] we get a satisfactory result (presented in [46] in the regular case). We shall not need the exact result here, but we shall need the following consequence of this result, that can also be derived from [46, Proposition 6.6].

Proposition 37. *Denote the generators of the closed loop system Ψ_L in Proposition 20 by $\left[\begin{smallmatrix} A_L & B_L \\ C_L & ? \end{smallmatrix}\right]$, and let W_{B_L} and $V^*_{C_L}$ be the closed loop versions of W_B and V^*_C, respectively. Then $W_{B_L} = W_B$ and $V^*_C = V^*_{C_L}$; hence $\operatorname{dom}(A_L) \subset W_B$ and $\operatorname{dom}(A^*_L) \subset V^*_C$.*

Thus, in the sequel we shall throughout replace W_{B_L} and $V^*_{C_L}$ by W_B and V^*_C, respectively.

In the regular case the closed loop generating operators have been determined by Weiss [46]; see, in particular, Theorem 4.7, Proposition 4.8, and Section 7 in that paper. From there we extract the following results:

Proposition 38. *Let Ψ and Ψ_L be the open and closed loop systems in Proposition 20, with admissible output feedback operator L.*

(i) *If both Ψ and Ψ^* are regular, then both Ψ_L and Ψ^*_L are regular, $I - DL$ is invertible in $\mathcal{L}(Y)$, and $I - LD$ is invertible in $\mathcal{L}(U)$.*

(ii) *Suppose that Ψ and Ψ^* are regular. Denote the generators of Ψ and Ψ_L by $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} A_L & B_L \\ C_L & D_L \end{smallmatrix}\right]$, respectively. Then the relation*

$$\begin{bmatrix} A_L & B_L \\ \overline{C}_L & D_L \end{bmatrix} = \begin{bmatrix} A + BL\,(I-DL)^{-1}\,\overline{C} & B\,(I-LD)^{-1} \\ (I-DL)^{-1}\,\overline{C} & D\,(I-LD)^{-1} \end{bmatrix}$$

holds in the following sense: the equation for A_L is valid on $\operatorname{dom}(A_L) \subset W_{B_L} = W_B$, the equation for $\overline{C}_L$ is valid on $W_{B_L} = W_B$, the equation for D_L is valid on U, and the equation for B_L should be interpreted as $\overline{B}^*_L = (I - D^*L^*)^{-1}\overline{B}^*$, which is valid on $V^*_{C_L} = V^*_C$.

Another different interpretation of the formula $B_L = B(I - LD)^{-1}$ is given in [46, Section 7].

Applying the preceding result with $L = (0\ I)$ to the system Ψ_{ext} in (30), we get the following result:

Proposition 39. *In addition to the hypothesis of Theorem 27, suppose that both the extended system Ψ_{ext} given by (30) and its adjoint system Ψ^*_{ext} given by (31) are regular, i.e., $\mathcal{D}$, $\mathcal{D}^*$, $\mathcal{F}$, and $\mathcal{F}^*$ are regular. Denote the generating operators of Ψ_{ext} by $\left[\begin{smallmatrix} A & B \\ \left(\begin{smallmatrix} C \\ K \end{smallmatrix}\right) & \left(\begin{smallmatrix} D \\ F \end{smallmatrix}\right) \end{smallmatrix}\right]$ and the generating operators of $\Psi_\circlearrowleft$ by $\left[\begin{smallmatrix} A_\circlearrowleft & B_\circlearrowleft \\ \left(\begin{smallmatrix} C_\circlearrowleft \\ K_\circlearrowleft \end{smallmatrix}\right) & \left(\begin{smallmatrix} D_\circlearrowleft \\ F_\circlearrowleft \end{smallmatrix}\right) \end{smallmatrix}\right]$ (the regularity of $\Psi_\circlearrowleft$ follows from Proposition 38.)*

(i) *Under these assumptions the outer factor $\mathcal{X}$ and the inner factor $\mathcal{Y}$ are regular together with their adjoints. If we denote the feed-through operator of $\mathcal{X}$ by X and the feed-through operator of $\mathcal{Y}$ by Y, then $X = E(I - F)$, X is invertible, and $D = YX$.*

(ii) *Construct the range space $V^*_{(C,K)}$ in the same way as the space V^*_C was constructed, except that Ψ is replaced by Ψ_{ext}. Then $V^*_C \subset V^*_{(C,K)}$, and the relation*

$$\begin{bmatrix} A_\circlearrowleft & B_\circlearrowleft \\ \begin{pmatrix} \overline{C}_\circlearrowleft \\ \overline{K}_\circlearrowleft \end{pmatrix} & \begin{pmatrix} D_\circlearrowleft \\ F_\circlearrowleft \end{pmatrix} \end{bmatrix} = \begin{bmatrix} A + BX^{-1}E\overline{K} & BX^{-1}E \\ \begin{pmatrix} \overline{C} + YE\overline{K} \\ X^{-1}E\overline{K} \end{pmatrix} & \begin{pmatrix} YE \\ X^{-1}E - I \end{pmatrix} \end{bmatrix}$$

holds in the following sense: the equation for $A_{\circlearrowleft}$ *is valid on* $\operatorname{dom}(A_{\circlearrowleft}) \subset W_B$*, the equations for* $\overline{C}_{\circlearrowleft}$ *and* $\overline{K}_{\circlearrowleft}$ *are valid on* W_B*, the equations for* $D_{\circlearrowleft}$ *and* $F_{\circlearrowleft}$ *are valid on* U*, the equation for* $F_{\circlearrowleft}$ *is valid on* Y*, and the equation for* $B_{\circlearrowleft}$ *should be interpreted as* $\overline{B}^*_{\circlearrowleft} = E^*(X^*)^{-1}\overline{B}^*$*, which is valid on* $V^*_{(C,K)}$*.*

This follows directly from Theorem 27 and Proposition 38. It is also true that $\operatorname{dom}(A^*_{\circlearrowleft}) \subset V^*_K \subset V^*_{(C,K)}$, with continuous inclusions where V^*_K is the range space constructed in the same way as V^*_C and $V^*_{(C,K)}$ were constructed, but with Ψ replaced by the system $\left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{K} & \mathcal{F} \end{smallmatrix}\right]$. That this inclusion is true follows from the fact that the closed loop semigroup $\mathcal{A}_{\circlearrowleft}$ depends only on $\mathcal{A}$, $\mathcal{B}$, $\mathcal{K}$, and $\mathcal{F}$, and not explicitly on $\mathcal{C}$.

We have the following corollary to Theorem 27 and Proposition 38:

Corollary 40. *Let the assumption of Theorem 27 hold. If* Ψ *is regular and* $\mathcal{D}^*\mathcal{D}$ *has a regular spectral factor* $\mathcal{X}$*, then every spectral factor of* $\mathcal{D}^*\mathcal{D}$ *is regular, and this is the case if and only if the extended system* Ψ_{ext} *in (30) is regular. If* Ψ^* *is regular and* $\mathcal{D}^*\mathcal{D}$ *has a spectral factor* $\mathcal{X}$ *with a regular adjoint* $\mathcal{X}^*$*, then every spectral factor of* $\mathcal{D}^*\mathcal{D}$ *has a regular adjoint, and this is the case if and only if the adjoint extended system* Ψ^*_{ext} *in (31) is regular. If both* Ψ *and* Ψ^* *are regular and* $\mathcal{D}^*\mathcal{D}$ *has a regular spectral factor* $\mathcal{X}$ *with a regular adjoint* $\mathcal{X}^*$*, then both the closed loop systems* $\Psi_{\circlearrowleft}$ *and* $\Psi^*_{\circlearrowleft}$ *are regular, the feed-through operator* X *of* $\mathcal{X}$ *is invertible, and there is a unique spectral factor of* $\mathcal{D}^*\mathcal{D}$ *with a positive self-adjoint feed-through operator. Moreover, in this case there is a unique feedback pair* $(\mathcal{K}, \mathcal{F})$ *in Theorem 27 for which the feed-through operator of* $\mathcal{F}$ *is zero, namely*

$$\begin{pmatrix} \mathcal{K} & \mathcal{F} \end{pmatrix} = \begin{pmatrix} -\pi_+ X^{-1}\mathcal{Y}^*\mathcal{C} & \left(I - X^{-1}\mathcal{X}\right) \end{pmatrix},$$

where $\mathcal{Y}\mathcal{X}$ *is an arbitrary inner-outer factorization of* $\mathcal{D}$*, and* X *is the feed-through operator of* $\mathcal{X}$*.*

Proof. Most of these claims follow immediately from Theorem 27 and Proposition 38. To get the unique spectral factor that has a positive self-adjoint feed-through operator we take an arbitrary spectral factor $\mathcal{X}$ with feed-through operator X, and multiply it to the left by the unitary operator $\left(X^*X\right)^{1/2}X^{-1}$, where $\left(X^*X\right)^{1/2}$ is the unique positive self-adjoint square root of X^*X. To get the unique feedback pair $(\mathcal{K}, \mathcal{F})$ for which the feed-through part of $\mathcal{F}$ is zero we take an arbitrary spectral factor $\mathcal{X}$ with feed-through operator X, and choose the operator E in Theorem 27 to be $E = X$. □

We remark that the case treated in [37] was regular, and there we throughout normalized $\mathcal{X}$ to have a positive self-adjoint feed-through operator, and took E in Theorem 27 to be the feed-through operator of $\mathcal{X}$.

9. The Open and Closed Loop Lyapunov Equations

In the sequel we suppose throughout that $\mathcal{D}$ is coercive. We claim that the feedback operator can be expressed in terms of the Riccati operator, and conversely, that the Riccati operator can be computed from the feedback operator through a Lyapunov equation. By combining these two facts we get an algebraic Riccati equation for the Riccati operator. Since the proof of the fact that the Riccati operator can be computed from the feedback operator through a Lyapunov equation is quite elementary and does not require any regularity assumptions, we start with this part.

It is possible to develop (at least) three different Lyapunov equations, one for the open loop system Ψ_{ext}, one for the closed loop system $\Psi_{\circlearrowleft}$, and a mixture of these two:

Theorem 41. *Let $\Psi = \left[\begin{smallmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{smallmatrix}\right]$ be a causal externally stable well-posed linear system on (U, H, Y) with a coercive input/output map $\mathcal{D}$. Let E be the operator in Theorem 27, denote the generating operators of the system Ψ_{ext} in (30) by $\left[\begin{smallmatrix} A & B \\ \left(\begin{smallmatrix} C \\ K \end{smallmatrix}\right) & \left(\begin{smallmatrix} ? \\ ? \end{smallmatrix}\right) \end{smallmatrix}\right]$, and denote the generating operators of the system $\Psi_{\circlearrowleft}$ in (29) by $\left[\begin{smallmatrix} A_{\circlearrowleft} & B_{\circlearrowleft} \\ \left(\begin{smallmatrix} C_{\circlearrowleft} \\ K_{\circlearrowleft} \end{smallmatrix}\right) & \left(\begin{smallmatrix} ? \\ ? \end{smallmatrix}\right) \end{smallmatrix}\right]$. With this notation, the Riccati operator Π of Ψ satisfies the open loop Lyapunov equation*

$$\langle Ax_1, \Pi x_0\rangle_H + \langle x_1, \Pi A x_0\rangle_H = -\langle Cx_1, Cx_0\rangle_Y + \langle EKx_1, EKx_0\rangle_U, \qquad x_0, x_1 \in \operatorname{dom}(A),$$

the closed loop Lyapunov equation

$$\langle A_{\circlearrowleft}x_1, \Pi x_0\rangle_H + \langle x_1, \Pi A_{\circlearrowleft} x_0\rangle_H = -\langle C_{\circlearrowleft}x_1, C_{\circlearrowleft}x_0\rangle_Y, \quad x_0, x_1 \in \operatorname{dom}(A_{\circlearrowleft}),$$

and the mixed Lyapunov equation

$$\langle A_{\circlearrowleft}x_1, \Pi x_0\rangle_H + \langle x_1, \Pi A x_0\rangle_H = -\langle C_{\circlearrowleft}x_1, Cx_0\rangle_Y, \quad x_0 \in \operatorname{dom}(A),\ x_1 \in \operatorname{dom}(A_{\circlearrowleft}).$$

If Ψ is strongly stable, then each of these equations determines Π uniquely (as a function of the other operators).

Proof. By Theorem 27 and Proposition 29, we have for all x_0 and x_1 in $W = \operatorname{dom}(A)$ and for all $t \in \mathbf{R}^+$,

$$\begin{aligned}
\langle \mathcal{A}(t)x_1, \Pi\mathcal{A}(t)x_0\rangle_H &= \langle \mathcal{C}\mathcal{A}(t)x_1, \mathcal{C}\mathcal{A}(t)x_0\rangle_{L^2(\mathbf{R}^+;Y)} \\
&\quad - \langle E\mathcal{K}\mathcal{A}(t)x_1, E\mathcal{K}\mathcal{A}(t)x_0\rangle_{L^2(\mathbf{R}^+;U)} \\
&= \int_0^\infty \langle C\mathcal{A}(s+t)x_1, C\mathcal{A}(s+t)x_0\rangle_Y \, ds \\
&\quad - \int_0^\infty \langle EK\mathcal{A}(s+t)x_1, EK\mathcal{A}(s+t)x_0\rangle_U \, ds \\
&= \int_t^\infty \langle C\mathcal{A}(s)x_1, C\mathcal{A}(s)x_0\rangle_Y \, ds \\
&\quad - \int_t^\infty \langle EK\mathcal{A}(s)x_1, EK\mathcal{A}(s)x_0\rangle_U \, ds.
\end{aligned}$$

Differentiating both sides of this equation with respect to t and taking $t = 0$, we get the open loop Lyapunov equation. To get the closed loop Lyapunov equation we argue in the same way, except that we start from the equation $\mathcal{A}_{\circlearrowleft}^*\Pi\mathcal{A}_{\circlearrowleft} = \mathcal{A}_{\circlearrowleft}^*\mathcal{C}_{\circlearrowleft}^*\mathcal{C}_{\circlearrowleft}\mathcal{A}_{\circlearrowleft}$. Finally, the mixed Lyapunov equation is also deduced in the same way, starting from the fact that $\mathcal{A}_{\circlearrowleft}^*\Pi\mathcal{A} = \mathcal{A}_{\circlearrowleft}^*\mathcal{C}_{\circlearrowleft}^*\mathcal{C}\mathcal{A}$.

To prove that the open loop Lyapunov equation determines Π uniquely we argue as follows. Take x_0 and x_1 in W. Then $\langle \mathcal{A}x_1, \Pi\mathcal{A}x_0\rangle$ is continuously differentiable on $\mathbf{R}^+$, with the derivate $\langle A\mathcal{A}x_1, \Pi A\mathcal{A}x_0\rangle$. Thus, if Π satisfies the open loop

Lyapunov equation, then for each $t > 0$,

$$\begin{aligned}\langle x_1, \Pi x_0\rangle_H - \langle \mathcal{A}(t)x_1, \Pi\mathcal{A}(t)x_0\rangle_H &= \int_0^t \langle C\mathcal{A}(s)x_1, C\mathcal{A}(s)x_0\rangle_Y \; ds \\ &\quad - \int_0^t \langle EK\mathcal{A}(s)x_1, EK\mathcal{A}(s)x_0\rangle_U \; ds \\ &= \int_0^t \langle (\mathcal{C}x_1)(s), (\mathcal{C}x_1)(s)x_0\rangle_Y \; ds \\ &\quad - \int_0^t \langle E(\mathcal{K}x_1)(s), E(\mathcal{K}x_1)(s)\rangle_U \; ds.\end{aligned}$$

Let $t \to \infty$. Because of the strong stability of Ψ, $\mathcal{A}(t)x_0 \to 0$ and $\mathcal{A}(t)x_1 \to 0$ in H, and we conclude that

$$\langle x_1, \Pi x_0\rangle_H = \langle \mathcal{C}x_1, \mathcal{C}x_0\rangle_{L^2(\mathbf{R}^+;Y)} - \langle E\mathcal{K}x_1, E\mathcal{K}x_0\rangle_{L^2(\mathbf{R}^+;U)}.$$

This being true for all x_0 and x_1 in W, we must have $\Pi = \mathcal{C}^*\mathcal{C} - \mathcal{K}^*E^*E\mathcal{K}$.

The proofs that also the closed loop and the mixed Lyapunov equations determine Π uniquely are similar (cf. Lemma 21). $\square$

The uniqueness proofs given above have been adapted from [19, Theorems 3 and 4].

Remark 42. As is well known, in the classical case the different Lyapunov equations given in Theorem 41 are equivalent; we can pass from one to another by using the appropriate version of (24) and the connection between the Riccati operator Π and the feedback operator K displayed in Theorem 44. A crucial role in these formulas is played by the feed-through operator D of the original system. Since a general well-posed linear system does not have a well-defined feed-through operator, it is difficult if not impossible to pass from one of the Lyapunov equations given above to another in the general case. Note that the closed loop and mixed Lyapunov equations are quite elementary, since they are based on the identities $\Pi = \mathcal{C}_\circlearrowleft^*\mathcal{C}_\circlearrowleft$ and $\Pi = \mathcal{C}_\circlearrowleft^*\mathcal{C}$ proved in Lemma 15, but the validity of the open loop Lyapunov equation is a deeper result.

Corollary 43. *Make the same same assumption and introduce the same notation as in Theorem 41.*

(i) *The Riccati operator* Π *satisfies the equations*

$$\begin{aligned}\Pi A x &= -\left(A^*\Pi + C^*C - K^*E^*EK\right)x = -\left(A_\circlearrowleft^*\Pi + C_\circlearrowleft^*C\right)x, \quad x \in \operatorname{dom}(A),\\ \Pi A_\circlearrowleft x &= -\left(A^*\Pi + C^*C_\circlearrowleft\right)x = -\left(A_\circlearrowleft^*\Pi + C_\circlearrowleft^*C_\circlearrowleft\right)x, \quad x \in \operatorname{dom}(A_\circlearrowleft).\end{aligned}$$

In particular, the operators $A^*\Pi + C^*C - K^*E^*EK$ *and* $A_\circlearrowleft^*\Pi + C_\circlearrowleft^*C$ *map* $\operatorname{dom}(A)$ *continuously into* H, *and the operators* $A^*\Pi + C^*C_\circlearrowleft$ *and* $A_\circlearrowleft^*\Pi + C_\circlearrowleft^*C_\circlearrowleft$ *map* $\operatorname{dom}(A_\circlearrowleft)$ *continuously into* H.

(ii) *Construct the range space* $V_{(C,K)}^*$ *in the same way as the space* V_C^* *was constructed, except that* Ψ *is replaced by* Ψ_{ext}. *Then* $\Pi \in \mathcal{L}(\operatorname{dom}(A); V_{(C,K)}^*) \cap \mathcal{L}(\operatorname{dom}(A_\circlearrowleft); V_C^*)$.

(iii)

$$\begin{gathered}\Pi \in \mathcal{L}(\mathrm{dom}(A); \mathrm{dom}(A^*)) \quad \textit{iff} \quad C^*C - K^*E^*EK \in \mathcal{L}(\mathrm{dom}(A); H),\\ \Pi \in \mathcal{L}(\mathrm{dom}(A); \mathrm{dom}(A_\circlearrowleft^*)) \quad \textit{iff} \quad C_\circlearrowleft^* C \in \mathcal{L}(\mathrm{dom}(A); H),\\ \Pi \in \mathcal{L}(\mathrm{dom}(A_\circlearrowleft); \mathrm{dom}(A^*)) \quad \textit{iff} \quad C^* C_\circlearrowleft \in \mathcal{L}(\mathrm{dom}(A_\circlearrowleft); H),\\ \Pi \in \mathcal{L}(\mathrm{dom}(A_\circlearrowleft); \mathrm{dom}(A_\circlearrowleft^*)) \quad \textit{iff} \quad C_\circlearrowleft^* C_\circlearrowleft \in \mathcal{L}(\mathrm{dom}(A_\circlearrowleft); H).\end{gathered}$$

(iv) $\Pi \in \mathcal{L}(W_B; V_{C,K}^*)$.

Proof. The first part follows immediately from Theorem 41. To prove parts (ii) and (iii) it suffices to choose some α in the resolvent set of A^* or $A_\circlearrowleft^*$, to add $\alpha\Pi x_0$ to both the equations, and to solve the resulting equations for Πx_0. Claim (iv) is proved in [40]. □

10. The Algebraic Riccati Equation

In order to turn the Lyapunov equations developed in the previous section into algebraic Riccati equations we still need to show that in some sense $B^*\Pi = -D^*(C + DK)$. Actually, as we shall see, this equation is not quite correct in the sense that in some cases we shall have to replace the feed-through operator D of $\mathcal{D}$ by the feed-through operator X of the outer factor $\mathcal{X}$ of $\mathcal{D}$, a phenomenon that was first discovered in [37].

To compute $B^*\Pi$ we study some coupled systems, where the output of the original open loop system Ψ or the closed loop state feedback system $\Psi_\circlearrowleft$ is used as the input of the adjoint closed loop system $\Psi_\circlearrowleft^*$ or open loop system Ψ^*, respectively. The systems Ψ and $\Psi_\circlearrowleft$ are considered in the initial value setting with initial time zero, and the adjoint systems $\Psi_\circlearrowleft^*$ and Ψ^* are considered in the time-invariant setting on $\mathbf{R}^+$. We start with the connection drawn in Figure 5, which will tell us how $B^*\Pi$ behaves on $\mathrm{dom}(A)$.

Theorem 44. *Make the same assumption and introduce the same notation as in Proposition 39. Then*

$$EKx = -(X^*)^{-1}\left(\overline{B}^*\Pi + D^*C\right)x, \qquad x \in \mathrm{dom}(A).$$

Proof. Let us first remark that $\overline{B}^*\Pi$ is well-defined, since, by Proposition 39 and Corollary 43, $\overline{B}^* \in \mathcal{L}(V_{(C,K)}^*; U)$ and $\Pi \in \mathcal{L}(W; V_{(C,K)}^*)$.

Consider the connection in Figure 5. We take $u = 0$ and $x_0 \in W$, and define

$$x(t) = \mathcal{A}(t)x_0, \quad y(t) = (\mathcal{C}x_0)(t) = C\mathcal{A}(t)x_0, \qquad t \in \mathbf{R}^+.$$

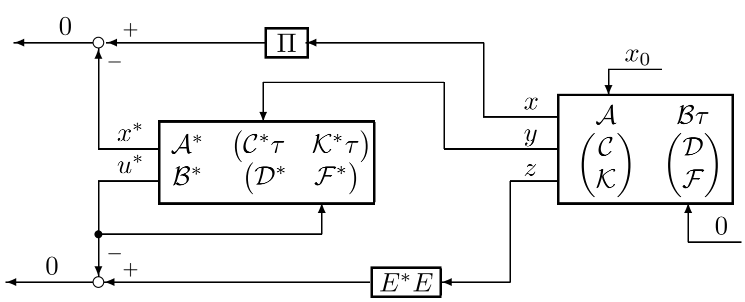

Figure 5. Primal-Dual Connection with Dual Feedback

Then, by Proposition 29, $x \in C^1(\mathbf{R}^+; H)$ and $y \in W^{1,2}(\mathbf{R}^+; Y)$. We extend y to an arbitrary function in $W^{1,2}(\mathbf{R}; Y)$ (for example, define $y(t) = y(-t)$ for $t < 0$), and define

$$x^* = \mathcal{C}_\circlearrowleft^* \tau y = \mathcal{C}^* \left(I - \mathcal{Y}\pi_+\mathcal{Y}^*\right) \tau y, \qquad u^* = \mathcal{D}_\circlearrowleft^* y = E^* \mathcal{Y}^* y.$$

By the same proposition, $x^* \in BC^1(\mathbf{R}; H)$ and $u^* \in W^{1,2}(\mathbf{R}; U)$. For nonnegative t we have (use the anti-causality of $\mathcal{C}_\circlearrowleft^*$ and $\mathcal{Y}^*$)

$$\begin{aligned} y(t) &= Cx(t), \\ x^*(t) &= \mathcal{C}_\circlearrowleft^* \tau(t) \mathcal{C} x_0 = \mathcal{C}_\circlearrowleft^* \mathcal{C}\mathcal{A}(t) x_0 = \Pi \mathcal{A}(t) x_0 = \Pi x(t), \\ u^*(t) &= E^* \left(\mathcal{Y}^* \mathcal{C} y\right)(t) = -E^* E \left(\mathcal{K} y\right)(t) = -E^* E K \mathcal{A}(t) x_0 = -E^* E K x(t). \end{aligned}$$

By Propositions 36 and 39, for all $t \in \mathbf{R}$,

$$u^*(t) = \overline{B}_\circlearrowleft^* x^*(t) + D_\circlearrowleft^* y(t) = E^* (X^*)^{-1} \left(\overline{B}^* x^*(t) + D^* y(t)\right).$$

For nonnegative t we can combine this with the preceding equations and use the invertibility of E^* to get

$$EKx(t) = -(X^*)^{-1} \left(\overline{B}^* \Pi x(t) + D^* C\right) x(t).$$

In particular, taking $t = 0$ we find that $EKx_0 = -(X^*)^{-1} \left(\overline{B}^* \Pi + D^* C\right) x_0$. □

Corollary 45. *In the regular case treated in Theorem 44, the Riccati operator satisfies the open loop algebraic Riccati Equation*

$$\begin{aligned} &\langle Ax_1, \Pi x_0 \rangle_H + \langle x_1, \Pi A x_0 \rangle_H \\ &\quad = - \langle Cx_1, Cx_0 \rangle_Y + \left\langle \left(\overline{B}^* \Pi + D^* C\right) x_1, (X^* X)^{-1} \left(\overline{B}^* \Pi + D^* C\right) x_0 \right\rangle_U, \\ &\qquad x_0, x_1 \in \operatorname{dom}(A), \end{aligned}$$

Proof. Combine Theorems 41 and 44. □

Theorem 44 tells us how $\overline{B}^* \Pi$ maps $W = \operatorname{dom}(A)$. It is an interesting fact that it is possible to define $\overline{B}^* \Pi$ also on $W_\circlearrowleft = \operatorname{dom}(A_\circlearrowleft)$, but that the resulting formula differs significantly from the one in Theorem 44:

Theorem 46. *With the assumption and notation of Proposition 39, the Riccati operator* Π *satisfies*

$$\overline{B}^* \Pi x + D^* C_\circlearrowleft x = \overline{B}^* \Pi x + D^* \left(\overline{C} + D X^{-1} E \overline{K}\right) x = 0, \qquad x \in \operatorname{dom}(A_\circlearrowleft).$$

In particular, if $D^* D$ *is invertible, then*

$$E\overline{K} x_1 = -X \left(D^* D\right)^{-1} \left(\overline{B}^* \Pi + D^* \overline{C}\right) x_1, \qquad x_1 \in \operatorname{dom}(A_\circlearrowleft).$$

In the classical case, $\operatorname{dom}(A) = \operatorname{dom}(A_\circlearrowleft)$ and $X^* X = D^* D$, so in that case this theorem is identical to Theorem 44. However, they differ from each other in the general case. In particular, although $X^* X$ is always invertible, $D^* D$ need not be so. See the examples in [38, Section 8] and [41].

Proof. Let us consider the connection in Figure 6. We take $u = 0$ and $x_0 \in \operatorname{dom}(A_\circlearrowleft)$, and define

$$x(t) = \mathcal{A}_\circlearrowleft(t) x_0, \quad y(t) = (\mathcal{C}_\circlearrowleft x_0)(t) = C_\circlearrowleft \mathcal{A}_\circlearrowleft(t) x_0, \qquad t \in \mathbf{R}^+.$$

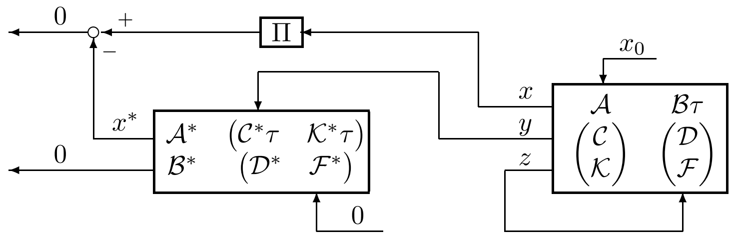

FIGURE 6. Primal-Dual Connection with Primal Feedback

Then, by Proposition 29, $x \in C^1(\mathbf{R}; H)$ and $y \in W^{1,2}(\mathbf{R}; Y)$, and we can extend y to a function in $W^{1,2}(\mathbf{R}; Y)$. Define

$$x^* = \mathcal{C}^*\tau y, \qquad u^* = \mathcal{D}^* y.$$

By the same proposition, $x^* \in BC^1(\mathbf{R}; H)$ and $u^* \in W^{1,2}(\mathbf{R}; U)$. For $t \in \mathbf{R}^+$ we have, by Proposition 39, by the anti-causality of $\mathcal{C}^*$ and $\mathcal{D}^* = (\mathcal{X}^*)^{-1}\mathcal{Y}^*$, and by the fact that $\mathcal{Y}^*\mathcal{Y} = I$,

$$\begin{aligned}
y(t) &= C_\circlearrowleft x(t) = \left(\overline{C} + DX^{-1}E\overline{K}\right)x(t),\\
x^*(t) &= \mathcal{C}^*\tau(t)\mathcal{C}_\circlearrowleft x_0 = \mathcal{C}^*\mathcal{C}_\circlearrowleft\mathcal{A}_\circlearrowleft(t)x_0 = \Pi\mathcal{A}_\circlearrowleft(t)x_0 = \Pi x(t),\\
\pi_+u^* &= \pi_+(\mathcal{X}^*)^{-1}\mathcal{Y}^*(I - \mathcal{Y}\pi_+\mathcal{Y}^*)\mathcal{C}x_0 = \pi_+(\mathcal{X}^*)^{-1}(\pi_+\mathcal{Y}^* - (\mathcal{Y}^*\mathcal{Y})\pi_+\mathcal{Y}^*) = 0.
\end{aligned}$$

By Propositions 36 and 39, for all $t \in \mathbf{R}$,

$$u^*(t) = \overline{B}^* x^*(t) + D^* y(t),$$

which for nonnegative t we may combine with the equations above to get

$$0 = \overline{B}^*\Pi x(t) + D^*\left(\overline{C} + DX^{-1}E\overline{K}\right)x(t).$$

Take $t = 0$ to get the claim of Theorem 46. □

The diagrams in Figures 5 and 6 are simplified versions of the more complete diagrams in [40].

Remark 47. Most of the results in this paper remain true if we throughout replace the algebra of time-invariant bounded linear operators from $L^2(\mathbf{R}; U)$ into $L^2(\mathbf{R}; Y)$ by some subalgebra, for example, by the algebra studied in [37]. The main exception is that spectral factorizations and inner-outer factorizations need not exist. In particular, Theorem 27 remains valid in that setting, too.

REFERENCES

1. D. Z. Arov and M. A. Nudelman. Passive linear stationary dynamical scattering systems with continuus time. *Integral Equations and Operator Theory*, 24:1–45, 1996. MR **96k:**47016
2. A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and Control of Infinite Dimensional Systems*, volumes 1 and 2. Birkhäuser, Boston, Basel, Berlin, 1992, 1993. MR **94b:**49003; MR **94m:**49001
3. A. Brown and C. Pearcy. *Introduction to Operator Theorey I: Elements of Functional Analysys*, volume 55 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin and New York, 1977. MR **58:**23463
4. F. M. Callier and J. Winkin. The spectral factorization problem for SISO distributed systems. In R. F. Curtain, editor, *Modelling, Robustness and Sensitivity Reduction in Control Theory*, volume F34 of *NATO ASI Series*, pages 463–489, Berlin and New York, 1987. Springer-Verlag. MR **88h:**93003

5. F. M. Callier and J. Winkin. Spectral factorization and LQ-optimal regulation for multivariable distributed systems. *International Journal of Control*, 52:55–75, 1990. MR **91h:**93039
6. F. M. Callier and J. Winkin. LQ-optimal control of infinite-dimensional systems by spectral factorization. *Automatica*, 28:757–770, 1992. MR **94b:**93072
7. L. E. Carpenter. Cascade analysis and synthesis of transfer functions of intinite dimensional linear systems. Doctoral dissertation, Virginia Polytechnic Inst. and State Univ., 1992.
8. R. F. Curtain. Well-posedness of infinite-dimensional linear systems in time and frequency domain. Technical report, University of Groningen, 1988.
9. R. F. Curtain. Representations of infinite-dimensional systems. In H. Nijmeijer and J. M. Schumacher, editors, *Three Decades of Mathematical System Theory. A Collection of Survays at the Occasion of the 50th Birthday of Jan C. Willems*, volume 135 of *Lecture Notes in Control and Information Sciences*, pages 101–128, Berlin and New York, 1989. Springer-Verlag. MR **90i:**93002
10. R. F. Curtain, editor. *Analysis and Optimization of Systems: State and Frequency Domain Approaches for Infinite-Dimensional Systems*, volume 185 of *Lecture Notes in Control and Information Sciences*, Berlin and New York, 1993. Springer-Verlag. MR **93i:**93005
11. R. F. Curtain and G. Weiss. Well posedness of triplets of operators (in the sense of linear systems theory). In *Control and Optimization of Distributed Parameter Systems*, pages 401–416, Basel, 1989. Birkhäuser-Verlag. MR **91d:**93027
12. R. F. Curtain, G. Weiss, and M. Weiss. Coprime factorization for regular linear systems. *Automatica*, 32: 1513–1531, 1996.
13. R. F. Curtain and H. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory*. Springer-Verlag, New York, 1995. MR **96i:**93001
14. G. Da Prato and M. C. Delfour. Unbounded solutions to the linear quadratic control problem. *SIAM J. Control and Optimization*, 30:31–48, 1992. MR **92k:**49069
15. A. Feintuch and R. Saeks. *System Theory: A Hilbert Space Approach*. Pure and Applied Mathematics. Academic Press, New York London, 1982. MR **84e:**93003
16. D. S. Flamm and K. Klipec. Numerical methods for H^∞ control of distributed parameter systems. In [10], pages 598–609. MR **93i:**93005
17. F. Flandoli, I. Lasiecka, and R. Triggiani. Algebraic Riccati equations with non-smoothing observation arising in hyperbolic and Euler-Bernulli boundary control problems. *Annali Mat. pure appl.*, CLIII:307–382, 1988. MR **90f:**49003
18. I. C. Gohberg and I. A. Fel′dman. *Convolution Equations and Projection Methods for their Solutions*, volume 41 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, Rhode Island, 1974. MR **50:**8149
19. P. Grabowski. On the spectral-Lyapunov approach to parametric optimization of distributed-parameter systems. *IMA Journal of Mathematical Control & Information*, 7:317–338, 1991. MR **92a:**49034
20. P. Grabowski. The LQ controller synthesis problem. *IMA Journal of Mathematical Control & Information*, 10:131–148, 1993. MR **93h:**93038
21. P. Grabowski. The LQ controller problem: an example. *IMA Journal of Mathematical Control & Information*, 11:355–368, 1994. MR **95j:**93015
22. J. W. Helton. Systems with infinite-dimensional state space: the Hilbert space approach. *Proceedings of the IEEE*, 64:145–160, 1976. MR **54:**4760
23. V. Kučera. New results in state estimation and regulation. *Automatica*, 17:754–748, 1981. MR **82j:**49014
24. I. Lasiecka. Riccati equations arising from boundary and point control problems. In [10], pages 23–45. MR **94b:**93069
25. I. Lasiecka and R. Triggiani. *Differential and Algebraic Riccati Equations with Applications to Boundary/Point Control Problems: Continuous Theory and Approximation Theory*, volume 164 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1991. MR **92k:**93009
26. I. Lasiecka and R. Triggiani. Algebraic Riccati equations arising from systems with unbounded input-output solution operator: applications to boundary control problems for wave and plate equations. *Nonlinear Analysis, Theory, Methods & Applications*, 20:659–693, 1993. MR **94g:**49087
27. J.-C. Louis and D. Wexler. The Hilbert space regulator problem and operator Riccati equations under stabilizability. *Annales del la Société Scientifique de Bruxelles*, 105:137–165, 1991. MR **93h:**93059

28. L. Pandolfi. The standard regulator problem for systems with input delays. An approach through singular control theory. *Applied Mathematics & Optimization*, 31:119–136, 1995. MR **95j:**49006
29. A. Pazy. *Semi-Groups of Linear Operators and Applications to Partial Differential Equations.* Springer-Verlag, Berlin, 1983. MR **85g:**47061
30. 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*, 2:335–362, 1985.
31. A. J. Pritchard and D. Salamon. The linear quadratic control problem for infinite dimensional systems with unbounded input and output operators. *SIAM Journal on Control and Optimization*, 25:121–144, 1987. MR **88a:**98041
32. M. Rosenblum and J. Rovnyak. *Hardy Classes and Operator Theory.* Oxford University Press, New York, Oxford, 1985. MR **87c:**47001
33. D. Salamon. *Control and Observation of Neutral Systems.* Pitman Publishing Ltd, London, 1984. MR **85e:**93002
34. D. Salamon. Infinite dimensional linear systems with unbounded control and observation: a functional analytic approach. *Transactions of the American Mathematical Society*, 300:383–431, 1987. MR **88d:**93024
35. D. Salamon. Realization theory in Hilbert space. *Mathematical Systems Theory*, 21:147–164, 1989. MR **89k:**93038
36. O. J. Staffans. Quadratic optimal control of stable abstract linear systems. In *The IFIP Conference on Modelling and Optimization of Distributed Parameter Systems with Applications to Engineering*, Warsaw, Poland, 1995. Chapman & Hall, pp. 167–174. CMP 96#72
37. O. J. Staffans. Quadratic optimal control of stable systems through spectral factorization. *Mathematics of Control, Signals, and Systems*, 8: 167–197, 1995.CMP 96#7
38. O. J. Staffans. Quadratic optimal control through coprime and spectral factorizations. Submitted, 1996a.
39. O. J. Staffans. Coprime factorizations for well-posed linear systems. To appear in SIAM Journal on Control and Optimization, 1996b.
40. O. J. Staffans. Quadratic optimal control of well-posed linear systems. To appear in SIAM Journal on Control and Optimization, 1996c.
41. O. J. Staffans. On the discrete and continuous time infinite-dimensional Riccati equations. *Systems and Control Letters*, 29: 131–138, 1996. CMP 19#5.
42. B. van Keulen. *H_∞-Control for Distributed Parameter Systems: A State Space Approach.* Birkhäuser Verlag, Basel Boston Berlin, 1993. MR **95b:**93002
43. G. Weiss. Admissibility of unbounded control operators. *SIAM Journal on Control and Optimization*, 27:527–545, 1989. MR **90c:**93060
44. G. Weiss. Admissible observation operators for linear semigroups. *Israel Journal of Mathematics*, 65:17–43, 1989. MR **90g:**47082
45. G. Weiss. Transfer functions of regular linear systems. Part I: Characterizations of regularity. *Transactions of American Mathematical Society*, 342:827–854, 1994a. MR **94f:**93074
46. G. Weiss. Regular linear systems with feedback. *Mathematics of Control, Signals, and Systems*, 7:23–57, 1994. MR **96i:**93046
47. G. Weiss and H. Zwart. An example in LQ optimal control. Preprint, 1996.
48. M. Weiss. Riccati equations in Hilbert space: A Popov function approach. Doctoral dissertation, Rijksuniversiteit Groningen, 1994.
49. M. Weiss. Riccati equation theory for Pritchard-Salamon systems: a Popov function approach. *IMA Journal of Mathematical Control & Information*, 14: 1–37, 1997.
50. M. Weiss and G. Weiss. Optimal control of stable weakly regular linear systems. Preprint, 1996.
51. J. Winkin. Spectral factorization and feedback control for infinite-dimensional control systems. Doctoral dissertation, Facultées Universitaires Notre-Dame de la Paix à Namur, 1989.
52. H. Zwart. Linear quadratic optimal control for abstract linear systems. In *The IFIP Conference on Modelling and Optimization of Distributed Parameter Systems with Applications to Engineering*, Warsaw, Poland, 1995. Chapman & Hall, pp. 175–182. CMP 96#12

Department of Mathematics, Åbo Akademi University, FIN-20500 Åbo, Finland
E-mail address: `Olof.Staffans@abo.fi`