New Riccati equations for well-posed linear systems

Mark R. Opmeer and Ruth F. Curtain

Systems Control Lett., vol. 52 (2004), no. 5, pp 339-347

Abstract

We consider the classic problem of minimizing a quadratic cost functional for well-posed linear systems over the class of inputs that are square integrable and that produce a square integrable output. As is well-known, the minimum cost can be expressed in terms of a bounded nonnegative selfadjoint operator X that in the finite-dimensional case satisfies a Riccati equation. Unfortunately, the infinite-dimensional generalization of this Riccati equation is not always well-defined. We show that X always satisfies alternative Riccati equations, which are more suitable for algebraic and numerical computations.

1 Introduction

The problem of minimizing a quadratic cost functional is a classic problem that has recieved much attention in the control literature. We are interested in this problem for the class of well-posed linear systems. Although this has been extensively studied in the recent thesis of Mikkola [11], we obtain useful new results here by taking a different perspective on the control problem. Our aim is not to obtain a Riccati equation that most resembles the finite-dimensional one, but to obtain a Riccati equation that is useful for algebraic and numerical computations.

For clarity let us consider the finite-dimensional problem of minimizing the cost functional

$$Q(x_0, u) = \int_0^\infty \|u(t)\|^2 + \|y(t)\|^2 dt = \langle u, u\rangle_{L^2} + \langle y, y\rangle_{L^2} \tag{1}$$

over all inputs $u \in L^2(0, \infty; \mathbb{C}^m)$ such that $y \in L^2(0, \infty; \mathbb{C}^p)$ where u and y satisfy

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \tag{2}$$

$$y(t) = Cx(t) + Du(t), \tag{3}$$

and A, B, C, D are matrices of compatible dimensions. Clearly it will be necessary that for each $x_0 \in \mathbb{R}^n$, there exists an input $u \in L^2(0, \infty; \mathbb{C}^m)$ such that the corresponding output $y \in L^2(0,\infty;\mathbb{C}^p)$ and $Q(x_0,u) < \infty$. We call this *optimizability.* If this condition is satisfied, then for every initial state x_0 there exists a unique input $u_{min} \in L^2(0,\infty;\mathbb{C}^m)$ such that $Q(x_0,u_{min}) \leq Q(x_0,u)$ for all inputs u and the minimum cost is given by

$$Q(x_0,u_{min}) = \langle x_0, Xx_0\rangle, \tag{4}$$

where X is the smallest nonnegative selfadjoint solution to the Riccati equation

$$A^*X + XA + C^*C = (XB + C^*D)(I + D^*D)^{-1}(D^*C + B^*X). \tag{5}$$

The above solution has been generalized to the class of well-posed linear systems by Zwart [20] and for more general cost functionals in Staffans [13] and Mikkola [11]. While the optimizability condition guarantees the existence of a unique minimizing input and the minumum cost has the form (4), the generalization of the Riccati equation (5) is incomplete. The problem lies in giving a correct interpretation to all the unbounded operators in (5), which is not always possible. The thesis [11] provides an exhaustive study of sufficient conditions for the existence of various generalizations of the Riccati equation and it gives very useful properties of the closed-loop system. In particular, it studies the important question of when the optimal control can be expressed as an admissible (and stabilizing) feedback of the state. We have nothing to add to this monumental study. Our perspective is different. We obtain a number of new Riccati equations to replace (5), the most interesting one being

$$A_-^*X + XA_- + C_-^*C_- = (XB_- + C_-^*D_-)(I + D_-^*D_-)^{-1}(D_-^*C_- + B_-^*X), \tag{6}$$

where, $A_- = A^{-1}, B_- = A^{-1}B, C_- = -CA^{-1}, D_- = \mathfrak{G}(0)$ and $\mathfrak{G}$ is the characteristic function of the well-posed linear system. Equation (6) is always well-defined provided that $0 \in \rho(A)$. This is a mild assumption, and it guarantees that $A^{-1}, A^{-1}B$ and CA^{-1} are bounded operators. Equation (6) is the Riccati equation associated with the optimal control problem for the reciprocal system with generating operators $A^{-1}, A^{-1}B, -CA^{-1}, \mathfrak{G}(0)$.

We show that the optimal control problem for this reciprocal system has a solution if and only if the optimal control problem for the well-posed linear system has a solution. Moreover, the minimum cost for both is $\langle x_0, Xx_0\rangle$, where X is the minimal solution to (6). The earlier paper, Curtain [3], on the connection between the control problem for a well-posed linear system and that for its reciprocal system considered only *stable* well-posed linear systems. It gave a detailed analysis of the relationships between the Riccati equations for the well-posed linear system and for its reciprocal system that we do not attempt here. In Curtain [2] the connection between *all* solutions of the two Riccati equations was made for the particular case when B is *bounded.* In this paper we only consider the minimal solution.

Following Mikkola [11] and Staffans [13], we consider the problem of minimizing a more general quadratic functional than (1) for the class of well-posed linear systems. We show that this problem is equivalent to minimizing an analogous quadratic cost functional for a related system. This system may be the reciprocal one mentioned above or a system obtained via a Cayley transform. To each of these related systems corresponds a Riccati equation that has the same nonnegative selfadjoint operator X as a solution, where the minimum cost is $\langle x_0, Xx_0\rangle$. The reciprocal Riccati equation is a continuous-time Riccati equation, but the Cayley transforms result in discrete-time Riccati equations.

The advantage of these new Riccati equations is that they are always well-defined and all the operators are bounded. Consequently algebraic and numerical computation for X becomes relatively straightforward. This opens the way to obtaining explicit solutions to various J-spectral factorization problems for well-posed linear systems using a technique developed by Curtain and Sasane [6]. We believe that the new Riccati equations will provide an alternative approach for obtaining numerical approximations of the Riccati equation and for $\mathbf{H}^\infty$ control problems.

We emphasize that we solve one particular optimization problem over the class of L^2 inputs that produce L^2 outputs. In Mikkola [11] other types of problems are also considered, for example, over the class of strongly or exponentially stabilizing inputs. However, by combining our result with his results it is possible to obtain information concerning strongly or exponentially stabilizing solutions.

The concept of a well-posed linear system is reviewed in Section 2 and in Section 3 we review optimal control theory for this class of systems. In Section 4 we briefly review infinite-dimensional discrete-time systems. Section 5 contains new results on Cayley transforms for well-posed linear systems. In Section 6 we review the concept of a reciprocal system and prove that the optimal cost operator of a well-posed linear system equals the optimal cost operator of its reciprocal system. This gives us the Riccati equation (6) for the optimal cost operator of the well-posed linear system.

2 Well-posed linear systems

In this section we review the concept of well-posed linear system (see Staffans [12], Mikkola [11] and for a survey [18]). We first introduce some notation. Let $\mathbb{R}^- = (-\infty, 0)$, $\mathbb{R}^+ = [0, \infty)$, and for any function u defined on $\mathbb{R}$ denote

$$(\tau^t u)(s) = u(t+s) \quad \forall\ t, s \in \mathbb{R},$$

$$(\pi_- u)(s) = \begin{cases} u(s) & \forall s \in \mathbb{R}^-, \\ 0 & \forall s \in \mathbb{R}^+, \end{cases}$$

$$(\pi_+ u)(s) = \begin{cases} u(s) & \forall s \in \mathbb{R}^+, \\ 0 & \forall s \in \mathbb{R}^-. \end{cases}$$

All Hilbert spaces in this article are assumed to be separable. For each Hilbert space U, each $\sigma \in \mathbb{R}$ and each interval $I \subset \mathbb{R}$ we let $L^2_\sigma(I;U)$ be the weighted L^2 space

$$L^2_\sigma(I;U) := \{u \in L^2_{\text{loc}}(I;U) : (t \mapsto \mathrm{e}^{-\sigma t} u(t)) \in L^2(I;U)\}.$$

This is a Hilbert space with the natural norm $\|e^{-\sigma \cdot}u(.)\|_{L^2(I;U)}$.

Definition 2.1. Let U, H and Y be Hilbert spaces. A quadruple $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ is called a well-posed linear system on (U, H, Y) if there exists a $\sigma \in \mathbb{R}$ such that

1. $t \mapsto \mathfrak{A}(t)$ is a strongly continuous semigroup of bounded operators on H;
2. $\mathfrak{B} : L^2_\sigma(\mathbb{R}^-; U) \to H$ is bounded and satisfies $\mathfrak{B}\tau^t = \mathfrak{A}(t)\mathfrak{B}$ for all $t \in \mathbb{R}^+$;
3. $\mathfrak{C} : H \to L^2_\sigma(\mathbb{R}^+; Y)$ is bounded and satisfies $\mathfrak{C}\mathfrak{A}(t) = \pi_+\tau^t\mathfrak{C}$ for all $t \in \mathbb{R}^+$;
4. $\mathfrak{D} : L^2_\sigma(\mathbb{R}; U) \to L^2_\sigma(\mathbb{R}; Y)$ is bounded and satisfies $\tau^t\mathfrak{D} = \mathfrak{D}\tau^t$ for all $t \in \mathbb{R}$, $\pi_-\mathfrak{D}\pi_+ = 0$, and $\pi_+\mathfrak{D}\pi_- = \mathfrak{C}\mathfrak{B}$.

$\mathfrak{B}$ is called the input map, $\mathfrak{C}$ the output map and $\mathfrak{D}$ the input-output map. Given an initial condition $x_0 \in H$ and an input $u \in L^2_\sigma(\mathbb{R}^+; U)$ the state $x(t) \in H$ at time $t \in \mathbb{R}^+$ and the output $y \in L^2_\sigma(\mathbb{R}^+; Y)$ of Σ are given by

$$\begin{bmatrix} x(t) \\ y \end{bmatrix} = \begin{bmatrix} \mathfrak{A}(t) & \mathfrak{B}\pi_-\tau^t \\ \mathfrak{C} & \mathfrak{D} \end{bmatrix} \begin{bmatrix} x_0 \\ u \end{bmatrix} = \begin{bmatrix} \mathfrak{A}(t)x_0 + \mathfrak{B}\pi_-\tau^t u \\ \mathfrak{C}x_0 + \mathfrak{D}u \end{bmatrix}. \tag{7}$$

Finite-dimensional well-posed linear systems can be generated by a quadruple of matrices A, B, C, D of compatible dimensions as

$$\mathfrak{A}(t) := e^{At} \quad \mathfrak{B}u := \int_{-\infty}^0 \mathfrak{A}(-s)Bu(s)ds \quad \mathfrak{C}x := (t \mapsto C\mathfrak{A}(t)x \quad t \geq 0) \tag{8}$$

$$\mathfrak{D}u := \left(t \mapsto \int_{-\infty}^t C\mathfrak{A}(t-s)Bu(s)ds + Du(t) \quad t \in \mathbb{R}\right).$$

The state and output defined by (7) then satisfy the equations (2) and (3), respectively. If A is the infinitesimal generator of a strongly continuous semigroup $\mathfrak{A}$ on H and $B \in \mathcal{L}(U, H)$, $C \in \mathcal{L}(H, Y)$, $D \in \mathcal{L}(U, Y)$, then (8) also defines a well-posed linear system. We say that A, B, C, D are the *generating operators* of the well-posed linear system. In fact, any well-posed linear system has uniquely defined generating operators A, B, C where A is the infinitesimal generator of the strongly continuous semigroup $\mathfrak{A}$ and B and C are in general unbounded operators (see Weiss [15], [16]). In general a feedthrough operator 'D' may not exist. A well-posed linear system for which the feedthrough operator does exist is called a *regular linear system* (see Weiss [17]).

We note that if Σ is a well-posed linear system with growth index σ, then it defines a well-posed linear system with growth index σ' for any $\sigma' > \sigma$. This implies that for any well-posed linear system the state and output can be defined for an input in $L^2(\mathbb{R}^+; U)$.

We will also need to consider well-posed linear systems in the frequency domain. Let $\mathcal{L}$ be the Laplace transform defined by

$$(\mathcal{L}u)(s) := \int_0^\infty e^{-st}u(t)dt.$$

We remind the reader that the Laplace transform is a unitary map from $L^2_\sigma(\mathbb{R}^+;U)$ onto the Hardy space $\mathbf{H}^2(\mathbb{C}^+_\sigma;U)$, where $\mathbb{C}^+_\sigma$ is the right half-plane consisting of those complex numbers that have real part larger than σ.

It is well-known that the map $\mathfrak{D}$ is uniquely determined by its Toeplitz operator $\mathfrak{D}_+ := \mathfrak{D}\pi_+$. We define the map $\widehat{\mathfrak{D}} : \mathbf{H}^2(\mathbb{C}^+_\sigma;U) \to \mathbf{H}^2(\mathbb{C}^+_\sigma;Y)$ by $\widehat{\mathfrak{D}} = \mathcal{L}\mathfrak{D}_+\mathcal{L}^{-1}$. Then obviously $\widehat{\mathfrak{D}}$ uniquely determines $\mathfrak{D}$. It is also well-known that there exists a function G in the Hardy space $\mathbf{H}^\infty(\mathbb{C}^+_\sigma;\mathcal{L}(U,Y))$ such that $\widehat{\mathfrak{D}}$ is multiplication by G. This G is called the *transfer function* of $\mathfrak{D}$ or of the well-posed linear system. We define $\widehat{\mathfrak{C}} : H \to \mathbf{H}^2(\mathbb{C}^+_\sigma;Y)$ by $\widehat{\mathfrak{C}} = \mathcal{L}\mathfrak{C}$. Note that the function $\widehat{\mathfrak{C}}$ is given by the formula

$$\widehat{\mathfrak{C}}(s) = C(sI-A)^{-1} \quad \text{for } s \in \mathbb{C}^+_{\bar{\sigma}}, \tag{9}$$

and the transfer function G satisfies (see Weiss [17, formula 4.13])

$$G(s) - G(\beta) = (\beta - s)C(sI-A)^{-1}(\beta I - A)^{-1}B \quad \text{for } s, \beta \in \mathbb{C}^+_{\bar{\sigma}}, \tag{10}$$

where $\bar{\sigma}$ is the maximum of σ and the growth bound of the semi-group $\mathfrak{A}$. As in Staffans and Weiss [14] we define the *characteristic function* $\mathfrak{G}$ of the well-posed linear system on $\rho(A)$ by fixing a $\beta \in \mathbb{C}^+_{\bar{\sigma}}$ and then defining $\mathfrak{G}(s)$ by

$$\mathfrak{G}(s) - G(\beta) = (\beta - s)C(sI-A)^{-1}(\beta I - A)^{-1}B \quad \text{for } s \in \rho(A). \tag{11}$$

We note that the transfer function and the characteristic function of a well-posed linear system are equal on some right half-plane, but need not be equal everywhere (this is explained in detail in [21]). The output $y \in L^2_\sigma(0,\infty;Y)$ of a well-posed linear system is related to the initial state x_0 and input $u \in L^2_\sigma(0,\infty;U)$ by

$$\hat{y}(s) = C(sI-A)^{-1}x_0 + G(s)\hat{u}(s), \quad s \in \mathbb{C}^+_{\bar{\sigma}},$$

where again $\bar{\sigma}$ is the maximum of σ and the growth bound of the semi-group $\mathfrak{A}$.

3 Optimal control

In this section we study the following optimal control problem that was studied in Staffans [13] and Mikkola [11].

Problem 3.1. Let $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ be a well-posed linear system, and let $J = J^* \in \mathcal{L}(Y)$. The (nonstandard) quadratic cost minimization problem for the pair (Σ, J) consists of finding, for each $x_0 \in H$, the infimum over all $u \in L^2(\mathbb{R}^+;U)$ such that $y \in L^2(\mathbb{R}^+;U)$, of the cost

$$Q(x_0,u) = \langle y, Jy\rangle_{L^2(\mathbb{R}^+;Y)}, \tag{12}$$

where y is the output of Σ with initial value $x_0 \in H$ and input $u \in L^2(\mathbb{R}^+;U)$ according to (7). If there exists an operator $X = X^* \in \mathcal{L}(H)$ such that the optimal cost is given by

$$\inf_{\{u \in L^2(\mathbb{R}^+;U) : y = \mathfrak{C}x + \mathfrak{D}u \in L^2(\mathbb{R}^+;Y)\}} Q(x_0,u) = \langle x_0, Xx_0\rangle_H,$$

then X is called the optimal cost operator of the pair (Σ, J).

The next concept plays a crucial role in the optimal control problem.

Definition 3.2. Let $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ be a well-posed linear system and let $x \in H$. The set $\mathcal{U}(x)$ of admissible inputs is defined by

$$\mathcal{U}(x) = \{u \in L^2(\mathbb{R}^+; U) : y = \mathfrak{C}x + \mathfrak{D}u \in L^2(\mathbb{R}^+; Y)\}.$$

Σ is called *optimizable* if for all $x \in H$ we have $\mathcal{U}(x) \neq \emptyset$.

To formulate a sufficient condition under which this problem has a solution we first investigate $\mathcal{U}(0)$. In Mikkola [11, Lemma 8.4.11c] it is shown that $\mathcal{U}(0)$ is a Hilbert space under the norm

$$\|u\|_{\mathcal{U}} := \sqrt{\|u\|^2_{L^2(\mathbb{R}^+;U)} + \|\mathfrak{D}u\|^2_{L^2(\mathbb{R}^+;Y)}},$$

and that $\mathfrak{D}$ (and thus $\mathfrak{D}_+$) restricts to an operator $\mathfrak{D}_{\mathcal{U}} \in \mathcal{L}(\mathcal{U}(0), L^2(\mathbb{R}^+; Y))$. The operator $\mathfrak{D}$ is called *positively J-coercive* if $\mathfrak{D}^*_{\mathcal{U}} J \mathfrak{D}_{\mathcal{U}}$ is positive and boundedly invertible. We note that $\mathfrak{D}$ is positively J-coercive iff there exists an $\varepsilon > 0$ such that for all $u \in \mathcal{U}(0)$ we have

$$\langle \mathfrak{D}u, J\mathfrak{D}u\rangle \geq \varepsilon(\|u\|_2^2 + \|\mathfrak{D}u\|_2^2) \tag{13}$$

(see Mikkola [11, Lemma 8.4.11]). This condition was used by Staffans in [13] as a definition, but, unfortunately, he called it J-coercive and Mikkola uses J-coercive for a slightly more general concept. We will use the terminology of Mikkola.

In the stable case (i.e., $\mathfrak{D}$ restricts to the bounded operator $\mathfrak{D}_+$ from $L^2(\mathbb{R}^+; U)$ to $L^2(\mathbb{R}^+; Y)$) we have $\mathcal{U}(0) = L^2(\mathbb{R}^+; Y)$, the norms on $\mathcal{U}(0)$ and $L^2(\mathbb{R}^+; U)$ are equivalent and $\mathfrak{D}^*_{\mathcal{U}} J \mathfrak{D}_{\mathcal{U}} = \pi_+ \mathfrak{D}^* J \mathfrak{D} \pi_+$ is the Popov-Toeplitz operator associated with $\mathfrak{D}$ and J. The positively J-coercive control problem in the stable case was studied in Weiss and Weiss [19].

The following lemma follows from Mikkola [11, Theorem 8.4.3, Corollary 8.1.8 and Theorem 8.1.10].

Lemma 3.3. *Let $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ be a well-posed linear system. If $\mathfrak{D}$ is positively J-coercive and Σ is optimizable, then, for every initial condition $x_0 \in H$, the cost functional (12) has a minimum that is achieved for a unique input u_{min} and the optimal cost operator exists.*

We note that the linear quadratic regulator problem is a special case of Problem 3.1. The LQR problem for a well-posed linear system $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ is the same as Problem 3.1 for the system

$$\begin{bmatrix} \mathfrak{A} & \mathfrak{B} \\ \begin{bmatrix} \mathfrak{C} \\ 0 \end{bmatrix} & \begin{bmatrix} \mathfrak{D} \\ I \end{bmatrix} \end{bmatrix}$$

with the cost operator $J = I$. Positive J-coercivity is automatic in this case since we can take $\varepsilon = \frac{1}{2}$ in (13). In a similar manner, many optimal control problems can be put in the framework of Problem 3.1.

4 Discrete-time systems

In this section we very briefly review some definitions on discrete-time systems. A *discrete-time system* is a quadruple of bounded operators $(A, B, C, D) \in \mathcal{L}(H) \times \mathcal{L}(U, H) \times \mathcal{L}(H, Y) \times \mathcal{L}(U, Y)$, where U, H, Y are Hilbert spaces. For an input u and initial condition x_0 the state x and output y of the system are defined by

$$x_{n+1} = Ax_n + Bu_n, \quad x(0) = x_0, \quad y_n = Cx_n + Du_n. \tag{14}$$

In the frequency domain this gives for z with $|z|$ smaller than the spectral radius of A

$$\hat{y}(z) = C(I - zA)^{-1}x_0 + \big(C(I - zA)^{-1}zB + D\big)\,\hat{u}(z),$$

where $\hat{u}$ and $\hat{y}$ are the Z-transforms of u and y, respectively. The Z-transform of a sequence h is given by

$$\hat{h}(z) = \sum_{j\in\mathbb{Z}} h_j z^j,$$

for those $z \in \mathbb{C}$ for which the sum converges absolutely. The concepts of the quadratic cost minimization problem, admissible inputs, optimizability, positive J-coercivity and other notions introduced earlier for continuous-time systems have their natural discrete-time counterparts. We define the characteristic function of a discrete-time system by $\mathfrak{G}(z) := C(I - zA)^{-1}zB + D$ for $1/z \in \rho(A)$. Note that the generating operators of a discrete-time system are *bounded*, which makes the theory of discrete-time systems easier than the theory of well-posed linear systems.

5 The Cayley transforms

We first define the Cayley transforms of a well-posed linear system following Staffans [12, Section 12.3] (see also Staffans and Weiss [14]).

Definition 5.1. Let U, H and Y be Hilbert spaces and let $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ be a well-posed linear system with generating operators A, B, C. Let $\alpha \in \rho(A) \cap \mathbb{C}_0^+$. Let Σ_d be the discrete-time system with generating operators

$$A_d = (\bar{\alpha}I + A)(\alpha I - A)^{-1}, \quad B_d = \sqrt{2\mathrm{Re}\ \alpha}\ (\alpha I - A)^{-1}B, \tag{15}$$

$$C_d = \sqrt{2\mathrm{Re}\ \alpha}\ C(\alpha I - A)^{-1}, \quad D_d = \mathfrak{G}(\alpha).$$

The discrete-time system Σ_d is called the *Cayley transform* (with parameter α) of Σ.

Note that usually the Cayley transform with parameter $\alpha = 1$ is used. The following lemma is a combination of results from [2] and [5] and the proof consists of algebraic manipulations with Riccati equations with *bounded* coefficients.

Lemma 5.2. *Let Σ be a regular linear system with bounded generating operators A, B, C, D and let Σ_d be its Cayley transform with parameter α. Then a nonnegative selfadjoint operator X is solution of the Riccati equation*

$$A^*X + XA + C^*JC = (XB + C^*JD)(D^*JD)^{-1}(D^*JC + B^*X) \tag{16}$$

if and only if it is a solution of the discrete-time Riccati equation

$$\begin{aligned} &A_d^*XA_d - X + C_d^*JC_d \\ &= (A_d^*XB_d + C_d^*JD_d)(D_d^*JD_d + B_d^*XB_d)^{-1}(B_d^*XA_d + D_d^*JC_d), \end{aligned} \tag{17}$$

of Σ_d.

In the case of the linear quadratic regulator problem it is well-known that both in continuous- (with bounded generating operators) and in discrete-time the optimal cost operator is the smallest nonnegative solution of the appropriate Riccati equation. This together with Lemma 5.2 gives the following.

Corollary 5.3. *Let Σ be a regular linear system with bounded generating operators A, B, C, D and let Σ_d be its Cayley transform with parameter α. Then the optimal cost operator for the linear quadratic regulator problem for Σ equals the optimal cost operator for the linear quadratic regulator problem for Σ_d.*

The goal of the remainder of this section is to show that this generalizes to well-posed linear systems if we choose the parameter α in the Cayley transform suitably. The approach we use is based on the frequency domain relation between a well-posed linear system and its Cayley transform. The following lemma follows from easy algebraic manipulations.

Lemma 5.4. *Let Σ be a well-posed linear system and let Σ_d be its Cayley transform with parameter α. Then*

$$C(sI - A)^{-1} = \frac{1+z}{\sqrt{2\text{Re }\alpha}}\ C_d(I - zA_d)^{-1}, \quad \mathfrak{G}(s) = \mathfrak{G}_d(z), \quad \forall s \in \rho(A)$$

where $z = (\alpha - s)/(\overline{\alpha} + s)$.

In the remainder of this article we will take the parameter α to be real. We now study the mapping $s \mapsto z = (\alpha - s)/(\alpha + s)$. It is easy to see that if $\alpha > r \geq 0$ it maps the right half-plane $\mathbb{C}_r^+$ bijectively onto the disc $\mathbb{D}_r^\alpha$ with center $-r/(\alpha + r)$ and radius $\alpha/(\alpha + r)$. Since α maps to zero we have $0 \in \mathbb{D}_r^\alpha$. The mapping induces a unitary transformation between $\mathbf{H}^2$ of the right half-plane $\mathbb{C}_0^+$ and $\mathbf{H}^2$ of the unit disc by

$$(\mathcal{H}_d g)(z) = \frac{\sqrt{2\alpha}}{1+z}\ \ g\left(\alpha\ \frac{1-z}{1+z}\right),$$

with its inverse given by

$$(\mathcal{H}_d^{-1} f)(s) = \frac{\sqrt{2\alpha}}{\alpha + s}\ \ f\left(\frac{\alpha - s}{\alpha + s}\right).$$

To show that this transformation is unitary we use that the norm of a $\mathbf{H}^2$ function is equal to the L^2 norm of its boundary function (see Duren [9], Hoffmann [10]). That the boundary function of g (defined on the imaginary axis) and the boundary function of $\mathcal{H}_d g$ (defined on the unit circle) have the same L^2 norm follows from the change of variables formula. The above implies that $\mathcal{H}_d$ induces an isometric isomorphism $\mathcal{T}_d$ between $L^2(0,\infty;S)$ and $l^2(\mathbb{N};S)$ for a Hilbert space S. The following theorem and its counterpart Theorem 5.6 are the crucial parts of this article, they give the connection between the set of admissible inputs of a well-posed linear system and the set of admissible inputs of its Cayley transform with a suitable parameter.

Theorem 5.5. *Let Σ be a well-posed linear system. Let Σ_d be its Cayley transform with parameter α, where α is larger than the growth bound of the semigroup of Σ and larger than zero. Let y be the output of Σ for the initial state x_0 and input $u \in \mathcal{U}(x_0)$. Then $y_d := \mathcal{T}_d y$ is the output of Σ_d for the initial state x_0 and input $u_d := \mathcal{T}_d u$.*

Proof. Define r to be a number larger than the growth bound of the semigroup of Σ and larger than zero, but smaller than α. Since $y \in L^2(0,\infty;Y)$ we have for z in the unit disc

$$\hat{y}_d(z) = \frac{\sqrt{2\alpha}}{1+z} \;\; \hat{y}\left(\alpha \; \frac{1-z}{1+z}\right).$$

If $z \in \mathbb{D}_r^\alpha$ then $s := \alpha(1-z)/(1+z) \in \mathbb{C}_r^+$ and since r is larger than the growth bound of the semigroup of Σ and larger than zero we have for $z \in \mathbb{D}_r^\alpha$

$$\hat{y}\,(s) = C\,(sI - A)^{-1}\,x_0 + G\,(s)\,\hat{u}\,(s)\,.$$

Since on $\mathbb{C}_r^+$ the transfer function and the characteristic function are equal, we have for $z \in \mathbb{D}_r^\alpha$

$$\hat{y}\,(s) = C\,(sI - A)^{-1}\,x_0 + \mathfrak{G}\,(s)\,\hat{u}\,(s)\,.$$

We now use Lemma 5.4 to conclude that for $z \in \mathbb{D}_r^\alpha$

$$\hat{y}_d(z) = \frac{\sqrt{2\alpha}}{1+z} \;\; \hat{y}\,(s) = C_d(I - zA_d)^{-1}x_0 + \mathfrak{G}_d(z)\hat{u}_d(z).$$

Here we have also used that since $u \in L^2(0,\infty;U)$ for z in the unit disc there holds

$$\hat{u}_d(z) = \frac{\sqrt{2\alpha}}{1+z} \;\; \hat{u}\,(s)\,.$$

Since $\mathbb{D}_r^\alpha$ is a connected subset of $\rho(A_d)$ containing zero, the transfer function and the characteristic function are equal on $\mathbb{D}_r^\alpha$ and we thus have

$$\hat{y}_d(z) = C_d(I - zA_d)^{-1}x_0 + G_d(z)\hat{u}_d(z).$$

This shows that y_d is indeed the output of the system Σ_d for the initial state x_0 and input u_d. $\square$ $\square$

Theorem 5.5 has the following counterpart, which can be proven similarly.

Theorem 5.6. *Let Σ be a well-posed linear system. Let Σ_d be its Cayley transform with parameter α, where α is larger than the growth bound of the semigroup of Σ and larger than zero. Let y_d be the output of Σ_d for the initial state x_0 and input $u_d \in \mathcal{U}_d(x_0)$. Then $y := \mathcal{T}_d^{-1} y_d$ is the output of Σ for the initial state x_0 and input $u := \mathcal{T}_d^{-1} u_d$.*

From Theorems 5.5 and 5.6 and the fact that Cayley transforms are isometric isomorphisms between stable continuous-time and discrete-time signals, it follows that $\mathcal{U}(x_0)$ and $\mathcal{U}_d(x_0)$ are isomorphic under the Cayley transforms. Thus a well-posed linear system is optimizable if and only if its Cayley transform is. Moreover, for all initial states $x_0 \in H$ and inputs $u \in \mathcal{U}(x_0)$ we have

$$Q(x_0, u) = Q_d(x_0, u_d).$$

This proves the following theorem.

Theorem 5.7. *Let Σ be a well-posed linear system. Let Σ_d be its Cayley transform with parameter α, where α is larger than the growth bound of the semigroup of Σ and larger than zero. Then, if Σ is positively J-coercive and optimizable, then so is Σ_d and the optimal cost operators of Σ and Σ_d are equal. Moreover, the optimal inputs are related by $u_{d,\min} = \mathcal{T}_d u_{\min}$.*

Theorem 5.7 gives us a Riccati equation for the optimal cost operator of the well-posed linear system: the optimal cost operator is a nonnegative selfadjoint solution of the (discrete-time) Riccati equation corresponding to the Cayley transform of Σ with a suitable parameter α.

6 Reciprocal systems

In this section we define the reciprocal system of a well-posed linear system as introduced in Curtain [4].

Definition 6.1. Let $\Sigma = \left[\begin{smallmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{smallmatrix}\right]$ be a well-posed linear system with generating operators A, B, C and characteristic function $\mathfrak{G}$. Assume that $0 \in \rho(A)$. Define

$$A_- = A^{-1}, \quad B_- = A^{-1}B, \quad C_- = -CA^{-1}, \quad D_- = \mathfrak{G}(0). \tag{18}$$

The well-posed linear system generated by A_-, B_-, C_-, D_- is called the reciprocal system of Σ and is denoted by $\Sigma_- = \left[\begin{smallmatrix} \mathfrak{A}_- & \mathfrak{B}_- \\ \mathfrak{C}_- & \mathfrak{D}_- \end{smallmatrix}\right]$.

Many system-theoretic properties of well-posed linear systems carry over to the reciprocal system and vice versa (see Curtain [2]-[4]). The following was noted in Curtain [2], the proof consists of elementary algebraic manipulations.

Lemma 6.2. *Let Σ be a well-posed linear system with $0 \in \rho(A)$, let Σ_d be its Cayley transform with parameter α , let Σ_- be the reciprocal system of Σ and let $\Sigma_{-,d}$ be the Cayley transform with parameter $1/\alpha \in \rho(A^{-1})$ of Σ_-. Then $A_{-,d} = -A_d, B_{-,d} = -B_d, C_{-,d} = C_d, D_{-,d} = D_d$.*

This lemma gives the following.

Lemma 6.3. *Let Σ be a positively J-coercive and optimizable well-posed linear system with $0 \in \rho(A)$. Let Σ_d be its Cayley transform with parameter α, where α is larger than the growth bound of the semigroup of Σ and larger than zero. Let Σ_- be the reciprocal system of Σ and let $\Sigma_{-,d}$ be the Cayley transform with parameter $1/\alpha \in \rho(A^{-1})$ of Σ_-. Then Σ_d and $\Sigma_{-,d}$ are positively J-coercive and optimizable and they have the same optimal cost operator.*

Proof. That Σ_d is positively J-coercive and optimizable follows from Theorem 5.7. It is easily seen that if y is the output of Σ_d for initial condition x_0 and input u, then $\tilde{y}$ defined by $\tilde{y}_n := (-1)^n y_n$ is the output of $\Sigma_{-,d}$ for the same initial condition and input $\tilde{u}_n := (-1)^n u_n$. This shows that $\Sigma_{-,d}$ is positively J-coercive and optimizable and that the optimal cost operators of Σ_d and $\Sigma_{-,d}$ are equal. □ □

Corollary 5.3, Theorem 5.7 and Lemma 6.3 give us the following.

Lemma 6.4. *Let Σ be an optimizable well-posed linear system with $0 \in \rho(A)$ and let Σ_- be its reciprocal system. Then Σ_- is optimizable and Σ and Σ_- have the same optimal cost operator for the linear quadratic regulator problem.*

Proof. Theorem 5.7 shows that Σ and its Cayley transform with parameter α, suitably chosen, have the same optimal cost operator. Corollary 5.3 shows that Σ_- and its Cayley transform with parameter $1/\alpha$ have the same optimal cost operator. Lemma 6.3 shows that the two Cayley transforms have the same optimal cost operator. This shows that Σ and Σ_- have the same optimal cost operator. Note that $1/\alpha$ may be in the wrong region of the complex plane, so we cannot use Theorem 5.7 to conclude that Σ_- and its Cayley transform have the same optimal cost operator. □ □

Since the generating operators of the reciprocal system are bounded, its optimal cost operator satisfies a Riccati equation.

Corollary 6.5. *Let Σ be an optimizable well-posed linear system with $0 \in \rho(A)$. Then its optimal cost operator for the linear quadratic regulator problem satisfies*

$$A_-^* X + X A_- + C_-^* C_- = (X B_- + C_-^* D_-)(I + D_-^* D_-)^{-1}(D_-^* C_- + B_-^* X),$$

where $A_- = A^{-1}, B_- = A^{-1}B, C_- = -CA^{-1}, D_- = \mathfrak{G}(0)$ and $\mathfrak{G}$ is the characteristic function of the well-posed linear system.

We remark that for the general problem (Problem 3.1) we can also obtain a reciprocal Riccati equation as in Corollary 6.5 under a certain assumption on the spectrum of A. To specify this assumption we define Ω as the component of $\rho(A) \cap \mathbb{C}_0^+$ that contains a right half-plane. The generalization of Corollary 6.5 to the general case of Problem 3.1 is then true under the extra condition that $0 \in \Omega$. The proof follows as in Section 5 for the Cayley transform, but now using the fact that $s \mapsto 1/s$ defines a unitary transformation on $\mathbf{H}^2$ of the right half-plane.

7 Conclusions

We examined the classic problem of minimizing a quadratic cost functional for well-posed linear systems. We have derived Riccati equations for the optimal cost operator where the coefficients are bounded operators. Notice that the reciprocal Riccati equation retains the form of continuous-time Riccati equations, but the Cayley transform results in discrete-time Riccati equations. We believe that the reciprocal transformation used in this paper can be profitably used to translate other problems for well-posed linear systems into a problem with bounded generators, for example, J-spectral factorization problems as in [6], [7], H^∞ control problems and numerical algorithms for finding the optimal cost operator. Finally, we remark that the approach taken here for well-posed linear systems has an obvious extension to the slightly more general class of nodes.

Acknowledgement 1. We are indebted to the detailed study of this problem in Mikkola [11], but we obtained our insight into the problem while reading the paper [20] by Zwart. We are grateful to Kalle Mikkola for his useful comments concerning this paper, in particular, his suggestion to introduce the term characteristic function.

References

[1] D.Z. Arov and M.A.Nudelman, Passive Linear Stationary Dynamical Scattering Systems with Continuous Time, *Integral Equations and Operator Theory*, **24** (1996) 1–45.

[2] R.F. Curtain, Regular linear systems and their reciprocals: applications to Riccati equations, *Systems and Control Letters* **49** (2003) 81–89.

[3] R.F. Curtain, Riccati equations for well-posed regular linear systems: the generic case, *Siam Journal on Control and Optimization* (to appear), available at http://www.math.rug.nl/∼curtain

[4] R.F. Curtain, Reciprocals of well-posed linear systems: a survey, in: *Proc. MTNS 2002* (South Bend, Indiana, 2002) available at http://www.nd.edu/mtns

[5] R.F. Curtain and J.C. Oostveen, Bilinear transformations between discrete- and continuous-time infinite-dimensional linear systems, in: *Proc. 4th International Symposium on Methods and Models in Automation and Robots* (Poland, 1997) 861–870.

[6] R.F. Curtain and A.J. Sasane, Explicit formulae for J-spectral factors for regular linear systems, in: *Proc. MTNS 2002* (South Bend, Indiana, 2002) available at http://www.nd.edu/mtns

[7] R.F. Curtain and A.J. Sasane, Hankel norm approximation for well-posed linear systems, *Systems and Control Letters*, **48** (2003) 407–414.

[8] R.F. Curtain and H.J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory* (Springer-Verlag, New York, 1995).

[9] P.L. Duren, *Theory of H^p Spaces* (New York, Academic Press, 1970).

[10] K. Hoffman, *Banach Spaces of Analytic Functions*, (Englewood Cliffs, Prentice-Hall, 1962).

[11] K. Mikkola, *Infinite-Dimensional Linear Systems, Optimal Control and Riccati Equations*, PhD thesis, Helsinki University of Technology, 2002, Available at http://www.math.hut.fi/reports

[12] O.J. Staffans, *Well-Posed Linear Systems*, (Cambridge, Cambridge University Press, 2004). (Manuscript avaliable at http://www.abo.fi/∼staffans/.)

[13] O.J. Staffans, Quadratic optimal control of well-posed linear systems, *SIAM Journal on Control and Optimization*, **37** (1998) 131–164.

[14] O. J. Staffans and G. Weiss, Transfer functions of regular linear systems Part II: The system operator and the Lax-Phillips semigroup, *Transactions of the American Mathematical Society*, **354** (2002) 3229–3262.

[15] G. Weiss, Admissibility of unbounded control operators, *SIAM Journal on Control and Optimization*, **27** (1989) 527–545.

[16] G. Weiss, Admissible observation operators for linear semigroups, *Israel Journal of Mathematics*, **65** (1989) 17–43.

[17] G. Weiss, Transfer functions of regular linear systems part 1: Characterizations of regularity. *Transactions of the American Mathematical Society*, **342** (1994) 827–854.

[18] G. Weiss, O.J. Staffans and M. Tucsnak, Well-posed linear systems—a survey with emphasis on conservative systems, *Int. J. Appl. Math. Comput. Sci.*, **11** (2001) 7–33.

[19] M. Weiss and G. Weiss, Optimal control of stable weakly regular linear systems. *Mathematics of Control, Signals and Systems*, **10** (1997) 287–330.

[20] H. J. Zwart, Linear quadratic optimal control for abstract linear systems, in *Modelling and optimization of distributed parameter systems* (Warsaw, 1995), 175–182.

[21] H. J. Zwart, Transfer functions for infinite-dimensional systems. *Systems and control letters*, to appear.