Singular Linear Quadratic Optimal Control for Infinite-Dimensional Systems

Mark R. Opmeer and Olof J. Staffans

July 3, 2026

Abstract

We consider the infinite-horizon linear quadratic optimal control problem for time-invariant systems. The quadratic cost is the norm squared of the output and may be singular. The linear system may be infinite-dimensional and may be a differential–algebraic equation. We make the relatively mild assumption that the transfer function is left-invertible at some point in the open right half-plane. We further assume a finite cost condition. Under these assumptions we show that a unique optimal control exists, that the optimal cost operator satisfies an algebraic Riccati equation in Lur'e form and that the optimal control is given by a proportional–derivative feedback of the optimal state. Our proofs are based on reduction to the discrete-time case using the Cayley transform. Several examples to illustrate the results are provided.

1 Introduction

We consider the objective of, for a given initial condition, minimizing

$$\int_0^\infty \|y(t)\|^2 \, dt,$$

where y is the output of a linear time-invariant infinite-dimensional system.

To set the scene, we first briefly review the finite-dimensional case. In the most general finite-dimensional case, the system under consideration is (at least formally)

$$E\dot{x} = Ax + Bu, \qquad y = Cx + Du. \tag{1}$$

We first consider the special case $E = I$. In that case, the LQ problem is called *regular* if D has a left-inverse and *singular* if it does not. The regular case is well-studied and it is well-known that the optimal control is smooth and is given by proportional state feedback as $u = Fx$. In the singular case, the optimal control is smooth for positive time, but might also contain the Dirac delta (at zero) and its derivatives. Early work [7, 20, 4] considered the linear space of initial states for which the optimal control is smooth for all time and the instantaneous jump to this linear space induced by the Dirac delta and its derivatives. More recent work [8, 16, 19] showed that the optimal control is given by proportional-derivative feedback of the state: $u = F_1\dot{x} + F_0 x$ (contrary to just proportional state feedback as in the regular case). These ideas have been generalized to the case where E is not the identity or more generally is not invertible (the Differential-Algebraic Equation (DAE) case) [5, 18, 2, 17].

In the regular finite-dimensional case, one can restrict attention to the case where inputs, states and outputs are smooth (infinitely differentiable functions). In the singular finite-dimensional case, the class of impulsive-smooth distributions has instead been considered (consisting not only of smooth functions, but also the Dirac delta at zero and its derivatives).

In the infinite-dimensional situation, even in the regular case, the optimal control need not be smooth or even continuous. In the singular case, it is possible for the optimal control to contain the Dirac delta and its derivatives at times other than zero. Hence the impulsive-smooth framework used in the finite-dimensional case fundamentally breaks down in the infinite-dimensional situation.

We consider linear time-invariant infinite-dimensional systems which are formally given by

$$\begin{bmatrix} \dot{x}(t) \\ y(t) \\ x(t) \\ u(t) \end{bmatrix} \in V, \qquad V \subset \begin{bmatrix} \mathscr{X} \\ \mathscr{Y} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix},$$

where $\mathscr{X}$ is the (generalized) state space, $\mathscr{Y}$ is the output space, $\mathscr{U}$ is the (generalized) input space and V is a subspace. Finite-dimensional systems of the form (1) fit into the framework by choosing V as the graph of the (possibly multi-valued) operator $S := \left[\begin{smallmatrix} E^{-1}A & E^{-1}B \\ C & D \end{smallmatrix}\right]$. Formally Laplace transforming gives

$$\begin{bmatrix} s\hat{x}(s) - x^0 \\ \hat{y}(s) \\ \hat{x}(s) \\ \hat{u}(s) \end{bmatrix} \in V,$$

and for technical reasons it is easier to make rigorous this frequency domain formulation. Instead of impulsive-smooth distributions as is done in the finite-dimensional singular case, we therefore consider *frequency domain trajectories* of infinite-dimensional *input/state/output nodes* [1].

Since the basic framework (by necessity) is fundamentally different from the existing work mentioned above, our method of solution for the singular linear quadratic optimal control problem is also fundamentally different. As in our previous work on the regular infinite-dimensional problem [14, 15, 10], we will use the Cayley transform to reduce the problem to a discrete-time one. Under a very mild assumption (left-invertibility of the continuous-time transfer function at one particular point, which is much weaker than left-invertibility of D) this discrete-time problem is in fact regular and therefore relatively easily solvable.

Some of our results are new even in the finite-dimensional special case. In the finite-dimensional situation the main focus has been on the LQ problem with state stability, i.e. the minimization is over all controls for which $\lim_{t\to\infty} x(t) = 0$. We do not include a state stability constraint. Of course, for many systems of relevance the condition that the output is square integrable implies state stability, and the distinction between the two problems (with or without state stability) disappears for such systems. Additionally, many existing finite-dimensional results rely on controllability and observability assumptions, which we do not make.

In Section 2 we first revisit the discrete-time case, indicating the changes which need to be made to [13] to handle the relevant situation. In Section 3.1 we collect some material from [10] to make the statements in this article self-contained. Section 3.2 contains the main result (Theorem 15) and indicates the changes which need to be made to [10] in the singular situation. Section 4 contains several examples. We first consider a very elementary example to illustrate our results in this simplest of cases. We then consider the cart-pendulum example and an RLC circuit. Subsequently we consider an infinite-dimensional example chosen to highlight the key difference with the finite-dimensional case. Finally, we consider a differential–algebraic example and an example where the problem with state stability and the problem without state stability have different solutions. These last two examples additionally serve the purpose of highlighting some issues in the literature.

2 The discrete-time case

Since we will reduce the continuous-time case to the discrete-time case, we first consider the discrete-time case. The objective is, for a given $x^0 \in \mathscr{X}$, to minimize

$$\sum_{n=0}^{\infty} \|\mathbf{y}_n\|^2,$$

where

$$\mathbf{x}_{n+1} = \mathbf{A}\mathbf{x}_n + \mathbf{B}\mathbf{u}_n, \qquad \mathbf{x}_0 = x^0, \qquad \mathbf{y}_n = \mathbf{C}\mathbf{x}_n + \mathbf{D}\mathbf{u}_n.$$

Here $\mathbf{u}$ is a sequence (the control) with values in $\mathscr{U}$, $\mathbf{x}$ is a sequence (the state) with values in $\mathscr{X}$, $\mathbf{y}$ is a sequence (the output) with values in $\mathscr{Y}$; the spaces $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ are Hilbert spaces and $\mathbf{A} : \mathscr{X} \to \mathscr{X}$, $\mathbf{B} : \mathscr{U} \to \mathscr{X}$, $\mathbf{C} : \mathscr{X} \to \mathscr{Y}$ and $\mathbf{D} : \mathscr{U} \to \mathscr{Y}$ are bounded linear operators. We assume that $\mathbf{D}$ has a left-inverse which is a bounded linear operator.

In our earlier work [13] we considered the *standard* case where the objective functional instead is $\sum_{n=0}^{\infty} \|\mathbf{y}_n\|^2 + \|\mathbf{u}_n\|^2$. With one exception, the arguments used for that standard case generalize with trivial modifications to the situation considered here. The exception is the closedness of the affine space which the minimization happens over. We establish that in the next lemma.

Lemma 1. *The subspace*

$$\left\{ \mathbf{y} \in \ell^2(\mathbb{N}_0; \mathscr{Y}) : \textit{ there exists a sequence } \mathbf{u} \textit{ such that } \quad \mathbf{y}_n = \sum_{k=0}^{n-1} \mathbf{C}\mathbf{A}^k\mathbf{B}\mathbf{u}_{n-1-k} + \mathbf{D}\mathbf{u}_n \right\},$$

is closed.

Proof. We first note that left-invertibility of $\mathbf{D}$ is equivalent to invertibility of $\mathbf{D}^*\mathbf{D}$ and that $(\mathbf{D}^*\mathbf{D})^{-1}\mathbf{D}^*$ is a left-inverse of $\mathbf{D}$. We also obtain from this left-invertiblility assumption that the image of $\mathbf{D}$ is closed.

Let $\mathbf{y}^m$ be a sequence in the indicated subspace with limit $\mathbf{y}$ and let $\mathbf{u}^m$ be a corresponding input sequence. Then for all $n \in \mathbb{N}_0$ we have

$$\mathbf{y}^m_n = \sum_{k=0}^{n-1} \mathbf{C}\mathbf{A}^k\mathbf{B}\mathbf{u}^m_{n-1-k} + \mathbf{D}\mathbf{u}^m_n \to \mathbf{y}_n, \qquad m \to \infty.$$

In particular for $n = 0$ we obtain $\mathbf{D}\mathbf{u}^m_0 \to \mathbf{y}_0$. Since the image of $\mathbf{D}$ is closed, it follows that $\mathbf{y}_0 = \mathbf{D}\mathbf{u}_0$ for some $\mathbf{u}_0 \in \mathscr{U}$. We further have

$$\mathbf{u}^m_0 = (\mathbf{D}^*\mathbf{D})^{-1}\mathbf{D}^*\mathbf{D}\mathbf{u}^m_0 \to (\mathbf{D}^*\mathbf{D})^{-1}\mathbf{D}^*\mathbf{y}_0 = (\mathbf{D}^*\mathbf{D})^{-1}\mathbf{D}^*\mathbf{D}\mathbf{u}_0 = \mathbf{u}_0.$$

For $n = 1$ we obtain $\mathbf{C}\mathbf{B}\mathbf{u}_0^m + \mathbf{D}\mathbf{u}_1^m \to \mathbf{y}_1$. Since as established above, $\mathbf{u}_0^m \to \mathbf{u}_0$, this gives $\mathbf{D}\mathbf{u}_1^m \to \mathbf{y}_1 - \mathbf{C}\mathbf{B}\mathbf{u}_0$. As above, this gives that there exists a $\mathbf{u}_1 \in \mathscr{U}$ such that $\mathbf{y}_1 - \mathbf{C}\mathbf{B}\mathbf{u}_0 = \mathbf{D}\mathbf{u}_1$ and

$$\mathbf{u}_1^m \to (\mathbf{D}^*\mathbf{D})^{-1}\mathbf{D}^*(\mathbf{y}_1 - \mathbf{C}\mathbf{B}\mathbf{u}_0) = \mathbf{u}_1.$$

Continuing like this, we obtain a sequence $\mathbf{u}$ such that $\mathbf{y}_n = \sum_{k=0}^{n-1} \mathbf{C}\mathbf{A}^k\mathbf{B}\mathbf{u}_{n-1-k} + \mathbf{D}\mathbf{u}_n$ for all $n \in \mathbb{N}_0$. Hence $\mathbf{y}$ belongs to the indicated subspace. □

With trivial modifications to the proofs in [13] we then obtain the following. Assume that for all $x^0 \in \mathscr{X}$ there exists a sequence $\mathbf{u}$ such that $\mathbf{y} \in \ell^2(\mathbb{N}_0; \mathscr{Y})$. Then for all $x^0 \in \mathscr{X}$ there exists a unique optimal control. There exists a bounded linear self-adjoint operator $X : \mathscr{X} \to \mathscr{X}$ such that the optimal cost is given by $\langle Xx^0, x^0 \rangle$. There exist bounded linear operators $\mathbf{K} : \mathscr{X} \to \mathscr{U}$ and $\mathbf{L} : \mathscr{U} \to \mathscr{U}$ where $\mathbf{L}$ has a bounded inverse such that the optimal control is given by

$$0 = \mathbf{K}\mathbf{x}_n + \mathbf{L}\mathbf{u}_n.$$

These operators are related by

$$\begin{aligned} \mathbf{A}^*X\mathbf{A} - X + \mathbf{C}^*\mathbf{C} &= \mathbf{K}^*\mathbf{K}, \\ \mathbf{B}^*X\mathbf{B} + \mathbf{D}^*\mathbf{D} &= \mathbf{L}^*\mathbf{L}, \\ \mathbf{B}^*X\mathbf{A} + \mathbf{D}^*\mathbf{C} &= \mathbf{L}^*\mathbf{K}. \end{aligned} \tag{2}$$

Moreover, X is the smallest nonnegative definite solution of (2).

Defining $\mathbf{F} := -\mathbf{L}^{-1}\mathbf{K}$, we have the closed-loop system

$$\mathbf{x}_{n+1} = (\mathbf{A} + \mathbf{B}\mathbf{F})\mathbf{x}_n, \qquad \mathbf{u}_n = \mathbf{F}\mathbf{x}_n, \qquad \mathbf{y}_n = (\mathbf{C} + \mathbf{D}\mathbf{F})\mathbf{x}_n.$$

This shows that the optimal control in fact belongs to the space $\ell_r^2(\mathbb{N}_0; \mathscr{U})$ for some $r > 0$. Here

$$\ell_r^2(\mathbb{N}_0; \mathscr{U}) = \left\{ \mathbf{u} : \mathbb{N}_0 \to \mathscr{U} : \sum_{k=0}^{\infty} \|r^{-k}\mathbf{u}_k\|^2 < \infty \right\}.$$

In particular, this means that the optimal control is Z-transformable (so that the corresponding state and output are also Z-transformable) and the Z-transforms satisfy (on some disc centered at the origin)

$$\begin{aligned} \hat{\mathbf{x}}(z) &= (I - z\mathbf{A})^{-1}x^0 + z(I - z\mathbf{A})^{-1}\mathbf{B}\hat{\mathbf{u}}(z), \\ \hat{\mathbf{y}}(z) &= \mathbf{C}(I - z\mathbf{A})^{-1}x^0 + \left[\mathbf{C}z(I - z\mathbf{A})^{-1}\mathbf{B} + \mathbf{D}\right]\hat{\mathbf{u}}(z). \end{aligned} \tag{3}$$

3 The continuous-time case

3.1 Preliminaries

We collect some definitions which also appeared in [10].

Definition 2. Let $\mathscr{U}$, $\mathscr{X}$ and $\mathscr{Y}$ be Hilbert spaces. An *i/s/o node* is a multi-valued operator from $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$ to $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{Y}\end{smallmatrix}\right]$. The graph of the i/s/o node S is (note that the components are the different way around than usual, this is to conform to the convention used in [1]):

$$\operatorname{gph}(S) = \left\{ \begin{bmatrix} Sq \\ q \end{bmatrix} : q \in \operatorname{dom}(S) \right\}.$$

The i/s/o node is called *closed* if S is a closed multi-valued operator (i.e. when $\mathrm{gph}(S)$ is a closed subspace) and *bounded* if S is a bounded single-valued operator with domain $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$.

Definition 3. For $\lambda \in \mathbb{C}$, the *formal i/s/o resolvent* of the i/s/o node S is the multi-valued operator $\widehat{\mathfrak{G}}(\lambda)$ from $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$ to $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{Y}\end{smallmatrix}\right]$ whose graph is given by

$$\mathrm{gph}(\widehat{\mathfrak{G}}(\lambda)) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathrm{gph}(S).$$

The *i/s/o resolvent set* $\rho(S)$ of S consists of those $\lambda \in \mathbb{C}$ for which $\widehat{\mathfrak{G}}(\lambda)$ is a bounded single-valued operator with domain $\left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{U}\end{smallmatrix}\right]$. The i/s/o node is called *resolvable* if $\rho(S)$ is non-empty and *future-resolvable* if $\rho(S) \cap \mathbb{C}_+ \neq \emptyset$. For $\lambda \in \rho(S)$ we have

$$\widehat{\mathfrak{G}}(\lambda) = \begin{bmatrix} \widehat{\mathfrak{A}}(\lambda) & \widehat{\mathfrak{B}}(\lambda) \\ \widehat{\mathfrak{C}}(\lambda) & \widehat{\mathfrak{D}}(\lambda) \end{bmatrix},$$

where $\widehat{\mathfrak{A}}$ is called the *state/state resolvent*, $\widehat{\mathfrak{B}}$ is called the *input/state resolvent*, $\widehat{\mathfrak{C}}$ is called the *state/output resolvent* and $\widehat{\mathfrak{D}}$ is called the *input/output resolvent.*

In the finite-dimensional case, the input/output resolvent and the transfer function can without any issue be identified with each other. In the infinite-dimensional case one has to be more careful about this [22].

Definition 4. Let Ω be a non-empty open subset of $\mathbb{C}$. A *frequency domain Ω trajectory* of an i/s/o node is a quadruple $(\hat{x}, \hat{y}, x^0, \hat{u})$ where $\hat{x}$, $\hat{y}$ and $\hat{u}$ are holomorphic functions defined on Ω with values in $\mathscr{X}$, $\mathscr{Y}$ and $\mathscr{U}$ respectively and $x^0 \in \mathscr{X}$ such that for all $\lambda \in \Omega$

$$\begin{bmatrix} \hat{x}(\lambda) \\ \hat{y}(\lambda) \\ x^0 \\ \hat{u}(\lambda) \end{bmatrix} \in \mathrm{gph}(\widehat{\mathfrak{G}}(\lambda)).$$

Remark 5. The above is [1, Definition 11.1.1]. By [1, Lemma 11.1.6], for a resolvable i/s/o node with $\Omega \subset \rho(S)$, for every $x^0 \in \mathscr{X}$ and every holomorphic $\mathscr{U}$-valued $\hat{u}$, there exist unique $\hat{x}$ and $\hat{y}$ such that the quadruple forms an Ω trajectory; namely

$$\begin{bmatrix} \hat{x}(\lambda) \\ \hat{y}(\lambda) \end{bmatrix} = \widehat{\mathfrak{G}}(\lambda) \begin{bmatrix} x^0 \\ \hat{u}(\lambda) \end{bmatrix}.$$

The following definition allows us to add an output to an i/s/o node. This is relevant in linear quadratic optimal control since the optimal control can be characterized by adding a certain output and subsequently putting that additional output equal to zero.

Definition 6. Let S be an i/s/o node, let $\mathscr{Y}_0$ be a Hilbert space and let $C = \begin{bmatrix} C_1 & C_0 \end{bmatrix} : \left[\begin{smallmatrix}\mathscr{X}\\ \mathscr{X}\end{smallmatrix}\right] \to \mathscr{Y}_0$ and $D = \begin{bmatrix} D_1 & D_0 \end{bmatrix} : \left[\begin{smallmatrix}\mathscr{Y}\\ \mathscr{U}\end{smallmatrix}\right] \to \mathscr{Y}_0$ be (bounded single-valued everywhere-defined) operators. The *nonstandard output extension* S^{ext} of S with observation extension C and feedthrough extension D is defined by

$$\operatorname{gph}(S^{\mathrm{ext}}) = \left\{ \begin{bmatrix} z \\ \begin{bmatrix} C_1 z + C_0 x + D_1 y + D_0 u \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \begin{bmatrix} \mathscr{Y}_0 \\ \mathscr{Y} \end{bmatrix} \\ \mathscr{X} \\ \mathscr{U} \end{bmatrix} : \begin{bmatrix} z \\ y \\ x \\ u \end{bmatrix} \in \operatorname{gph}(S) \right\}.$$

A *standard output extension* is a nonstandard output extension where $C_1 = 0$ and $D_1 = 0$.

Definition 6 is from [1, Definition 5.1.23 (ii)] and [1, Definition 5.1.33 (ii)]. If S is bounded, then a nonstandard output extension is equivalent to a standard output extension [1, Lemma 6.2.1 (vii)].

3.1.1 The internal Cayley transform

Definition 7. For $\alpha \in \mathbb{C}$ with $\mathrm{Re}(\alpha) > 0$, the *Cayley transform* of the i/s/o node S is the multi-valued operator S_d from $\left[\begin{smallmatrix} \mathscr{X} \\ \mathscr{U} \end{smallmatrix}\right]$ to $\left[\begin{smallmatrix} \mathscr{X} \\ \mathscr{Y} \end{smallmatrix}\right]$ whose graph is given by

$$\operatorname{gph}(S_d) = \begin{bmatrix} \frac{1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{\overline{\alpha}}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{-1}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 & \frac{\alpha}{\sqrt{2\mathrm{Re}(\alpha)}} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \operatorname{gph}(S).$$

If S is future-resolvable and $\alpha \in \rho(S) \cap \mathbb{C}_+$, then the Cayley transform with parameter α is a single-valued bounded operator with domain $\left[\begin{smallmatrix} \mathscr{X} \\ \mathscr{U} \end{smallmatrix}\right]$ and in particular it therefore can be written as

$$S_d = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

for (single-valued bounded everywhere-defined) operators $\mathbf{A} : \mathscr{X} \to \mathscr{X}$, $\mathbf{B} : \mathscr{U} \to \mathscr{X}$, $\mathbf{C} : \mathscr{X} \to \mathscr{Y}$ and $\mathbf{D} : \mathscr{U} \to \mathscr{Y}$.

Let $\mathcal{L}$ denote the Laplace transform and note that by the Paley–Wiener theorem for a Hilbert space $\mathscr{K}$ this is an isometric isomorphism between $L^2(\mathbb{R}_+; \mathscr{K})$ and the Hardy space $H^2(\mathbb{C}_+; \mathscr{K})$. The Z-transform $\mathcal{Z}$ maps a sequence $(h_n)_{n \in \mathbb{N}_0}$ to the corresponding formal power series $\sum_{n=0}^{\infty} h_n z^n$ and gives an isometric isomorphism between $\ell^2(\mathbb{N}_0; \mathscr{K})$ and the Hardy space of the disc $H^2(\mathbb{D}; \mathscr{K})$. For $\alpha \in \mathbb{C}$ with $\mathrm{Re}(\alpha) > 0$, the linear fractional transformation

$$\begin{aligned} (F_\alpha g)(z) &= \frac{\sqrt{\mathrm{Re}(2\alpha)}}{1+z}\, g\left(\frac{\alpha - \overline{\alpha} z}{1+z}\right), \\ (F_\alpha^{-1} f)(\lambda) &= \frac{\sqrt{\mathrm{Re}(2\alpha)}}{\overline{\alpha} + \lambda}\, f\left(\frac{\alpha - \lambda}{\overline{\alpha} + \lambda}\right), \end{aligned} \tag{4}$$

gives an isometric isomorphism between the Hardy spaces $H^2(\mathbb{C}_+; \mathscr{K})$ and $H^2(\mathbb{D}; \mathscr{K})$ (but we note that F_α more generally maps functions which are holomorphic at α to functions which are holomorphic at zero). For $\alpha \in \mathbb{C}$ with $\mathrm{Re}(\alpha) > 0$, the *Laguerre transform* T_α is defined by

$$T_\alpha : L^2(\mathbb{R}_+; \mathscr{K}) \to \ell^2(\mathbb{N}_0; \mathscr{K}), \qquad T_\alpha := \mathcal{Z}^{-1} F_\alpha \mathcal{L},$$

and is an isometric isomorphism.

3.2 Linear quadratic optimal control

Definition 8. Let S be a future-resolvable i/s/o node and let Ω be a non-empty open subset of $\rho(S) \cap \mathbb{C}_+$. An Ω trajectory $(\hat{x}, \hat{y}, x^0, \hat{u})$ is called *output-stable* if $\hat{y}$ is the restriction to Ω of the Laplace transform of some (necessarily unique) $y \in L^2(\mathbb{R}_+; \mathscr{Y})$.

We say that S satisfies the Ω *finite cost condition* if for all $x^0 \in \mathscr{X}$ the corresponding set of output-stable Ω trajectories is non-empty.

In the standard case where the cost function is $\|y\|^2_{L^2(\mathbb{R}_+;\mathscr{Y})} + \|u\|^2_{L^2(\mathbb{R}_+;\mathscr{U})}$, not only the output but also the control is square-integrable. Since square-integrable functions translate nicely under the Cayley transform [10, Lemma 20], in the standard case existence of an optimal control follows easily from the corresponding discrete-time result. In the singular case, where only square-integrability of the output is known, an additional assumption and an addition argument is needed. The following result motivates that additional assumption.

Lemma 9. *Let S be a future resolvable i/s/o node with input/output resolvent $\widehat{\mathfrak{D}}$ and let Ω be a non-empty connected open subset of $\rho(S) \cap \mathbb{C}_+$. Assume that there exists a holomorphic function $\widehat{\mathfrak{L}} : \Omega \to \mathscr{B}(\mathscr{Y}, \mathscr{U})$ such that $\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{D}}(\lambda) = I$ for all $\lambda \in \Omega$. Let $(\hat{x}, \hat{y}, x^0, \hat{u})$ be an Ω trajectory of S. Then for all $\lambda \in \Omega$*

$$\hat{u}(\lambda) = -\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{L}}(\lambda)\hat{y}(\lambda).$$

Proof. By definition of Ω trajectory we have for all $\lambda \in \Omega$

$$\hat{y}(\lambda) = \widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{D}}(\lambda)\hat{u}(\lambda).$$

Applying $\widehat{\mathfrak{L}}(\lambda)$ to both sides gives

$$\widehat{\mathfrak{L}}(\lambda)\hat{y}(\lambda) = \widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{C}}(\lambda)x^0 + \hat{u}(\lambda),$$

which re-arranges to the desired equality. □

Remark 10. The assumption in Lemma 9 on the existence of $\widehat{\mathfrak{L}}$ at first glance seems to be stronger than the earlier mentioned assumption that the input/output resolvent has a left-inverse at a given $\alpha \in \rho(S) \cap \mathbb{C}_+$. To some extent this is true and to some extent it is not. Let $L \in \mathscr{B}(\mathscr{Y}, \mathscr{U})$ be such that $L\widehat{\mathfrak{D}}(\alpha) = I$. Define F by $F(\lambda) := L\widehat{\mathfrak{D}}(\lambda)$. Then $F(\alpha) = I$, which is particularly invertible, so that by openness of the set of invertible operators in $\mathscr{B}(\mathscr{U})$ and continuity of F, we have that $F(\lambda)$ is invertible for all λ in some neighborhood of α. Define $\widehat{\mathfrak{L}}(\lambda) := F(\lambda)^{-1}L$. Then $\widehat{\mathfrak{L}}$ is holomorphic and

$$\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{D}}(\lambda) = F(\lambda)^{-1}L\widehat{\mathfrak{D}}(\lambda) = F(\lambda)^{-1}F(\lambda) = I.$$

Therefore we obtain existence of $\widehat{\mathfrak{L}}$ with the properties indicated in Lemma 9 as long as Ω is restricted to be a sufficiently small neighborhood of α. In this sense, the assumption in Lemma 9 is not stronger than the assumption that the iinput/output resolvent has a left-inverse at a given α. If we want to work on an a priori given Ω, then the assumption in Lemma 9 might be genuinely stronger. However, in the rational finite-dimensional case, it is easy to see by considering minors and determinants that $\widehat{\mathfrak{D}}(\alpha)$ having a left-inverse implies that $\widehat{\mathfrak{D}}$ has a left-inverse in the sense of rational matrices. This might still entail a restriction of an a priori given Ω, but would only mean the exclusion of finitely many points (the poles of the rational left-inverse) which has no significant consequence. See [20, Proposition 4] for more on left-inveribility in the sense of rational matrices.

Remark 11. If we want to consider time-domain trajectories, then a natural assumption is to consider Ω a right half-plane and to assume that the i/s/o resolvent $\widehat{\mathfrak{G}}$ is polynomially bounded on Ω (see [11]). Strengthening the assumption in Lemma 9 to $\widehat{\mathfrak{L}}$ being polynomially bounded on Ω then gives that for $\hat{y}$ the Laplace transform of $y \in L^2(\mathbb{R}_+; \mathscr{Y})$ we have by the formula in Lemma 9 that $\hat{u}$ is the Laplace transform of an exponentially bounded distribution and consequently that $\hat{x}$ is as well. The additional polynomial boundedness assumptions are always satisfied in the finite-dimensional rational case. From the above we see how our frequency domain framework specializes to a distributional framework in the finite-dimensional rational case (and in many infinite-dimensional situations as well).

The following result shows that continuous-time trajectories give rise to discrete-time trajectories preserving stability of the output.

Proposition 12. *Let S be a future resolvable i/s/o node and let Ω be a non-empty connected open subset of $\rho(S) \cap \mathbb{C}_+$. Let $(\hat{x}, \hat{y}, x^0, \hat{u})$ be an output-stable Ω trajectory of S. Let $\alpha \in \Omega$ and let S_d be the Cayley transform of S with parameter α. Let $\mathbf{y}$ be the Laguerre transform with parameter α of y. Then there exist $\mathbf{u}$ and $\mathbf{x}$ such that $(\mathbf{x}, \mathbf{y}, x^0, \mathbf{u})$ is a Z-transformable trajectory of S_d.*

Proof. We have for $\lambda \in \Omega$ that

$$\hat{x}(\lambda) = \widehat{\mathfrak{A}}(\lambda)x^0 + \widehat{\mathfrak{B}}(\lambda)\hat{u}(\lambda), \qquad \hat{y}(\lambda) = \widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{D}}(\lambda)\hat{u}(\lambda).$$

Define, using the linear fractional transformation (4), $Y := F_\alpha \hat{y}$ and $U := F_\alpha \hat{u}$. Then Y is the Z-transform of $\mathbf{y}$ and U is the Z-transform of some sequence $\mathbf{u}$. Furthermore we have that on a neighborhood of zero (utilizing the i/s/o resolvent of S_d)

$$\hat{\mathbf{y}}(z) = \widehat{\mathfrak{C}}_d(z)x^0 + \widehat{\mathfrak{D}}_d(z)\hat{\mathbf{u}}(z).$$

We then define $\mathbf{x}$ as the state sequence of S_d corresponding to the initial state x^0 and the control sequence $\mathbf{u}$. ☐

The following result shows that, under a left-invertibility assumption, discrete-time trajectories give rise to continuous-time trajectories preserving stability of the output.

Proposition 13. *Let S be a future resolvable i/s/o node with input/output resolvent $\widehat{\mathfrak{D}}$ and let Ω be a non-empty connected open subset of $\rho(S) \cap \mathbb{C}_+$. Assume that there exists a holomorphic function $\widehat{\mathfrak{L}} : \Omega \to \mathscr{B}(\mathscr{Y}, \mathscr{U})$ such that $\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{D}}(\lambda) = I$ for all $\lambda \in \Omega$.*

Let $\alpha \in \Omega$ and let S_d be the Cayley transform of S with parameter α. Let $(\mathbf{x}, \mathbf{y}, x^0, \mathbf{u})$ be a Z-transformable trajectory of S_d with $\mathbf{y} \in \ell^2(\mathbb{N}_0; \mathscr{Y})$. Let y be the inverse Laguerre transform with parameter α of $\mathbf{y}$. Let $\hat{y}$ be the restriction to Ω of the Laplace transform of y. Then there exist $\hat{u}$ and $\hat{x}$ such that $(\hat{x}, \hat{y}, x^0, \hat{u})$ is an output-stable Ω trajectory of S.

Proof. Let $\hat{\mathbf{u}}$ be the Z-transform of $\mathbf{u}$ (defined on a neighborhood of zero) and let $U := F_\alpha^{-1}\hat{\mathbf{u}}$ (defined on a neighborhood of α). On a neighborhood N of α (contained in Ω) we then have by transforming the frequency-domain discrete-time output equation using the linear fractional transformation

$$\hat{y}(\lambda) = \widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{D}}(\lambda)U(\lambda).$$

By applying $\widehat{\mathfrak{L}}(\lambda)$ to both sides and re-arranging we then have on N

$$U(\lambda) = -\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{L}}(\lambda)\hat{y}(\lambda).$$

Define for $\lambda \in \Omega$

$$\hat{u}(\lambda) = -\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{L}}(\lambda)\hat{y}(\lambda).$$

Then $\hat{u}$ coincides with U on N and $\hat{u}$ is holomorphic on Ω. By the Identity Theorem we have

$$\hat{y}(\lambda) = \widehat{\mathfrak{C}}(\lambda)x^0 + \widehat{\mathfrak{D}}(\lambda)\hat{u}(\lambda),$$

since both sides are holomorphic on the connected open set Ω and agree on the open subset N.

Define $\hat{x}$ through

$$\hat{x}(\lambda) = \widehat{\mathfrak{A}}(\lambda)x^0 + \widehat{\mathfrak{B}}(\lambda)\hat{u}(\lambda).$$

Then $(\hat{x}, \hat{y}, x^0, \hat{u})$ is indeed an Ω trajectory of S and it is clearly output-stable. ☐

Remark 14. It follows from Proposition 12 and Proposition 13 that, under the assumption of the existence of $\widehat{\mathfrak{L}}$, minimizing the norm of y over all output-stable Ω trajectories of S and minimizing the norm of $\mathbf{y}$ over all Z-transformable trajectories of S_d is equivalent (in the sense that the costs are the same and the optimal outputs are Laguerre transforms). Since Z-transformability follows from optimality, and stability of the output is implicit in the minimization, we have that minimizing over all Ω trajectories of S and over all trajectories of S_d is equivalent.

Theorem 15. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty connected open subset of $\rho(S) \cap \mathbb{C}_+$. Assume that there exists a holomorphic function $\widehat{\mathfrak{L}} : \Omega \to \mathscr{B}(\mathscr{Y}, \mathscr{U})$ such that $\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{D}}(\lambda) = I$ for all $\lambda \in \Omega$. If the Ω finite cost condition is satisfied, then for every $x^0 \in \mathscr{X}$ there exists a unique optimal control, there exists a bounded single-valued everywhere-defined self-adjoint operator $X : \mathscr{X} \to \mathscr{X}$ such that the optimal cost is given by $\langle Xx^0, x^0\rangle$, S has a non-standard output extension S^{ext} with a feedthrough extension which is standard and has invertible standard part and is such that*

$$\langle z, Xx\rangle + \langle Xx, z\rangle + \|y\|^2 = \|w\|^2 \text{ for all } \begin{bmatrix} z \\ \begin{bmatrix} w \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} \in \operatorname{gph}(S^{\text{ext}}); \tag{5}$$

holds and the optimal control is characterized by putting the additional output in S^{ext} equal to zero. Moreover, X is the smallest nonnegative definite solution of (5).

Proof. The proof is the same as that of [10, Theorem 22]. We note that [10, Theorem 22] only concludes that the feedthrough extension has left-invertible standard part, but in fact it follows from the proof that it is invertible rather than just left-invertible. That X is the the smallest nonnegative definite solution of (5) is not explicitly stated in [10, Theorem 22], but it follows since by [10, Proposition 21], solutions of the discrete-time and continuous-time equations coincide and X is the smallest nonnegative definite solution of the discrete-time equation. ☐

If S is bounded and we write $S = \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right]$ and we write the non-standard output extension with a feedthrough extension which is standard as $K_1 z + K_0 x + L_0 u$, then the Riccati equation (5) can be written as

$$\langle Ax+Bu, Xx\rangle + \langle Xx, Ax+Bu\rangle + \|Cx+Du\|^2 = \|(K_1A+K_0)x + (K_1B+L_0)u\|^2, \text{ for all } x \in \mathscr{X},\ u \in \mathscr{U},$$

which is equivalent to

$$\begin{aligned} A^*X + XA + C^*C &= K^*K, \\ D^*D &= L^*L, \\ B^*X + D^*C &= L^*K, \end{aligned} \tag{6}$$

where

$$K = K_1 A + K_0, \qquad L = K_1 B + L_0. \tag{7}$$

In (6) we have the Lur'e form of an algebraic Riccati equation in the unknowns X, K and L. An interesting aspect is the decomposition (7) of K and L into K_1, K_0 and L_0. This decomposition is less unique than K and L are. Since S is bounded, it is always possible to choose the observation extension to be standard (i.e. $K_1 = 0$), which gives $K_0 = K$ and $L_0 = L$ and defines the optimal manifold $Kx(t) + Lu(t) = 0$ (which does not involves $\dot{x}$, but may not be solvable for u). Another interesting special decomposition exists when $D = 0$ (which happens in cheap control and gives $L = 0$) and B has a left-inverse B^+. In that case, we can choose $K_1 = B^+$ and we obtain $K_0 = K - B^+A$ and $L_0 = -I$. This gives $u(t) = B^+\dot{x}(t) + (K - B^+A)x(t)$, which gives an easy explicit proportional-derivative feedback formula for the optimal control. We note that by Theorem 15 such a proportional-derivative feedback formula always exists under the left-invertibility condition, but the case where $D = 0$ and B is left-invertible gives a particularly easy formula. See [19, Section 4] for a generalization of this formula to the case $D \neq 0$.

Similarly as in [13, 10], it is possible to consider the situation where not every initial state has a corresponding input for which the output is square integrable. The relevant definition and theorem are as follows.

Definition 16. Let S be a future-resolvable i/s/o node and let Ω be a non-empty open subset of $\rho(S) \cap \mathbb{C}_+$ and let $\widehat{\mathfrak{B}}$ be the input/state resolvent.

We say that S satisfies the Ω *input finite cost condition* if for all $\lambda \in \Omega$, for all initial states in $\operatorname{im}(\widehat{\mathfrak{B}}(\lambda))$ the corresponding set of output-stable Ω trajectories is non-empty.

Theorem 17. *Let S be a future-resolvable i/s/o node and let Ω be a non-empty connected open subset of $\rho(S) \cap \mathbb{C}_+$. Assume that there exists a holomorphic function $\widehat{\mathfrak{L}} : \Omega \to \mathscr{B}(\mathscr{Y}, \mathscr{U})$ such that $\widehat{\mathfrak{L}}(\lambda)\widehat{\mathfrak{D}}(\lambda) = I$ for all $\lambda \in \Omega$. If the Ω input finite cost condition is satisfied, then for every initial condition for which the set of output-stable Ω trajectories is nonempty, there exists a unique optimal control, there exists a closed symmetric sesquilinear form q on $\mathscr{X}$ such that the optimal cost is given by $q[x^0, x^0]$, S has a non-standard output extension S^{ext} with a feedthrough extension which is standard and has invertible standard part and is such that*

$$q[z,x] + q[x,z] + \|y\|^2 = \|w\|^2 \text{ for all } \begin{bmatrix} z \\ \begin{bmatrix} w \\ y \end{bmatrix} \\ x \\ u \end{bmatrix} \in \operatorname{gph}(S^{\mathrm{ext}}); \tag{8}$$

holds and the optimal control is characterized by putting the additional output in S^{ext} equal to zero. Moreover, q is the smallest nonnegative definite solution of (8).

Proof. This follows as in [10, Section 9.2] with the same modifications as considered for Theorem 15. ☐

4 Examples

4.1 The most elementary example

Example 18. Consider the finite-dimensional example $A = 0$, $B = 1$, $C = 1$, $D = 0$, i.e. the minimization of

$$\int_0^\infty |y(t)|^2 \, dt, \qquad \dot{y} = u.$$

In frequency domain this is $\hat{y}(\lambda) = \frac{1}{\lambda}x^0 + \frac{1}{\lambda}\hat{u}(\lambda)$. We have $\widehat{\mathfrak{D}}(\lambda) = \frac{1}{\lambda}$, so that $\widehat{\mathfrak{L}}(\lambda) = \lambda$ and we can choose $\Omega = \mathbb{C}_+$. The optimal output is $y = 0$ and the optimal control in frequency domain is $\hat{u}(s) = -x^0$, which in time-domain corresponds to $u = -x^0\delta$. A nonstandard output extension which gives rise to this is (here $X = 0$, $K = 1$, $L = 0$ and $B^+ = 1$)

$$\dot{x} + x - u = 0.$$

4.2 The cart-pendulum example

Consider the cart-pendulum example from [3, Section 5.7] and more particularly, the linearization about pendulum down. A state space representation is (here x_1 is the deviation of the carts horizontal position from the desired one, $x_3 := \ell\theta$ where ℓ is the length of the pendulum and θ is the angle and u is a force applied to the cart)

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{mg}{M\ell} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{(m+M)g}{M\ell} & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1}{M} \\ 0 \\ \frac{1}{M} \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where g is the gravitational constant, M is the mass of the cart and m is the mass at the tip of the pendulum. This output, leading to a singular optimal control problem, is relevant for the study of the effect of underactuation regardless of any additional control constraints/cost [21]. We have

$$D + C(I - A)^{-1}B = \begin{bmatrix} g + \ell \\ \ell \end{bmatrix} \frac{1}{(m + M)g + M\ell},$$

so that the transfer function evaluated at 1 is left-invertible.

The equation (6) can be solved by hand and we obtain

$$X = \begin{bmatrix} X_1 & X_2 & -X_1 & -X_2 \\ X_2 & X_3 & -X_2 & -X_3 \\ -X_1 & -X_2 & X_1 & X_2 \\ -X_2 & -X_3 & X_2 & X_3 \end{bmatrix},$$

$$K = \begin{bmatrix} 1 & X_1 & \frac{g}{\ell}X_2 & -X_1 \end{bmatrix}, \qquad L = 0,$$

where

$$X_1 = \sqrt{\frac{2\ell}{g}(\sqrt{2}-1)}, \qquad X_2 = \frac{\ell}{g}(\sqrt{2}-1),$$
$$X_3 = \sqrt{2}\frac{\ell}{g}\sqrt{\frac{2\ell}{g}(\sqrt{2}-1)}.$$

Note that the solution is independent of the values of the masses.

The relevant optimal manifold is given by $Kx(t) = 0$. With $K_1 = \begin{bmatrix} 0 & M & 0 & 0 \end{bmatrix}$ which is a left-inverse of B, we obtain $L_0 = -I$ and $K_0 = K - K_1 A$ and the corresponding feedback formula for the optimal control is

$$\begin{aligned} u &= K_1\dot{x} + (K - K_1 A)x \\ &= M\dot{x}_2 + \begin{bmatrix} 1 & X_1 & \frac{g}{\ell}(m + X_2) & -X_1 \end{bmatrix} x. \end{aligned}$$

4.3 A circuit example

We consider an RLC circuit as in Figure 1.

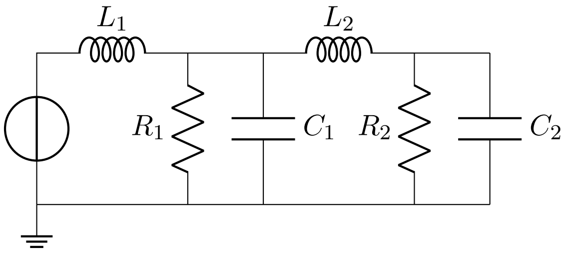

Figure 1: RLC circuit considered

As initial objective could be to maximize $\int_0^\infty I_0(t)V_0(t)\,dt$, where I_0 and V_0 are the current and voltage at the source. This maximizes the energy extracted from the initial energy stored in the capacitors and inductors. A related problem is to minimize the dissipated energy $\int_0^\infty \frac{1}{R_1}V_1(t)^2 + \frac{1}{R_2}V_2(t)^2\,dt$ where V_1 and V_2 are the voltages across the resistors (and across the capacitors). More generally, such a relation between an indefinite optimal control problem and a norm minimization problem is investigated in [6]. We consider V_0 as the control, the voltages through the capacitors and the currents through the inductors as states and obtain the system:

$$x = \begin{bmatrix} I_1 \\ V_1 \\ I_2 \\ V_2 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & \frac{-1}{L_1} & 0 & 0 \\ \frac{1}{C_1} & \frac{-1}{R_1C_1} & \frac{-1}{C_1} & 0 \\ 0 & \frac{1}{L_2} & 0 & \frac{-1}{L_2} \\ 0 & 0 & \frac{1}{C_2} & \frac{-1}{R_2C_2} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{L_1} \\ 0 \\ 0 \\ 0 \end{bmatrix},$$
$$C = \begin{bmatrix} 0 & \frac{1}{\sqrt{R_1}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{R_2}} \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We have that the transfer function in zero equals

$$-CA^{-1}B + D = \begin{bmatrix} \frac{1}{\sqrt{R_1}} \\ \frac{1}{\sqrt{R_1}} \end{bmatrix},$$

which is left-invertible (from which it follows that the transfer function is left-invertible as a rational matrix function).

The relevant equations can be solved by hand and result in the following solution. Let X_1 be the unique positive solution of (uniqueness following from Descartes' rule of signs, existence from the Intermediate Value Theorem)

$$\frac{C_2^2 R_1^3}{4L_2^6} x^4 + \frac{C_2 R_1^2}{R_2 L_2^4} x^3 + \left(\frac{C_2 R_1}{L_2^3} + \frac{R_1}{R_2^2 L_2^2} \right) x^2 + \frac{2}{R_2 L_2} x - \frac{1}{R_2} = 0,$$

let

$$X_2 = \frac{C_2^2 R_1^2}{2L_2^4} X_1^3 + \frac{C_2 R_1}{2R_2 L_2^2} X_1^2 + \frac{C_2}{L_2} X_1, \qquad X_0 = \frac{C_2 R_1}{2L_2^2} X_1^2,$$

then

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & X_1 & X_0 \\ 0 & 0 & X_0 & X_2 \end{bmatrix},$$

$$K = \begin{bmatrix} 0 & \frac{1}{\sqrt{R_1}} & \frac{\sqrt{R_1}}{L_2} X_1 & \frac{\sqrt{R_1}}{L_2} X_0 \end{bmatrix}, \qquad L = 0.$$

The optimal manifold therefore is

$$V_1(t) + \frac{R_1}{L_2} X_1 I_2(t) + \frac{R_1}{L_2} X_0 V_2(t) = 0.$$

Utilizing $K_1 = \begin{bmatrix} L_1 & 0 & 0 & 0 \end{bmatrix}$ (a left-inverse of B), one possible feedback formula for the control is

$$u(t) = L_1 \dot{I}_1(t) - \frac{R_1}{L_2} X_1 I_2(t) - \frac{R_1}{L_2} X_0 V_2(t).$$

4.4 An infinite-dimensional example

We re-consider an example from [12]. Define with the state space $\mathscr{X} = L^2(0,2)$, one-dimensional input space $\mathscr{U}$ and output space $\mathscr{Y} = L^2(0,1) \times \mathscr{U}$ for $R \geq 0$

$$S \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} x' \\ \begin{bmatrix} x|_{(0,1)} \\ \sqrt{R}\, u \end{bmatrix} \end{bmatrix},$$

$$\operatorname{dom}(S) = \left\{ \begin{bmatrix} x \\ u \end{bmatrix} \in \begin{bmatrix} \mathscr{X} \\ \mathscr{U} \end{bmatrix} : x|_{(0,1)}' \in L^2(0,1), \right.$$
$$\left. x|_{(1,2)}' \in L^2(1,2),\ x(1^-) - x(1^+) = u,\ x(2) = 0 \right\}.$$

The input/output resolvent $\widehat{\mathfrak{D}}(\lambda) : \mathscr{U} \to L^2(0,1)$ can be computed to be

$$\widehat{\mathfrak{D}}(\lambda)u = \xi \mapsto \mathrm{e}^{\lambda(\xi-1)}u.$$

A possible left-inverse $\widehat{\mathfrak{L}}(\lambda) : L^2(0,1) \to \mathscr{U}$ is given by

$$\widehat{\mathfrak{L}}(\lambda)y = \frac{\lambda}{1-\mathrm{e}^{-\lambda}} \int_0^1 y(\xi)\, d\xi,$$

and we can choose $\Omega = \mathbb{C}_+$.

The relevant computations in [12] give that

$$u^{\mathrm{opt}}(t) = \begin{cases} \frac{-1}{1+R} x^0(t+1) & t \in (0,1) \\ 0 & t > 1. \end{cases}$$

We note that the optimal control need not be continuous (and in fact on the interval $(0,1)$ can be an arbitrary square integrable function) even in the regular $(R > 0)$ situation. In [12] the feedback formula

$$u^{\mathrm{opt}} = \frac{-1}{1+R} x(1^+),$$

is derived. We note that we can write

$$x(1^+) = -\int_1^2 x'(\xi)\, d\xi,$$

which gives that the optimal feedback can alternatively be written as (consistent with nonstandard output extensions as considered in this article)

$$\begin{gathered} 0 = K_1 z + K_0 x + L_0 u, \quad K_1 z = \frac{-1}{\sqrt{1+R}} \int_1^2 z(\xi)\, d\xi, \\ K_0 = 0, \qquad L_0 u = \sqrt{1+R}\; u, \end{gathered}$$

which gives

$$u^{\mathrm{opt}}(t) = \frac{1}{1+R} \int_1^2 \dot{x}(t,\xi)\, d\xi.$$

This means that rather than the proportional feedback with an unbounded operator $u^{\mathrm{opt}} = \frac{-1}{1+R} x(1^+)$, we alternatively have a derivative feedback with a bounded operator.

4.5 A differential-algebraic equation example

Consider a mass-spring system with u a force, x_1 momentum, x_2 position, m the mass and a unit spring constant:

$$\dot{x}_1 = -x_2 + u, \qquad m\dot{x}_2 = x_1, \qquad y = x_2.$$

In frequency domain this is

$$s\hat{x}_1 - x_1^0 = -\hat{x}_2 + \hat{u}, \qquad ms\hat{x}_2 - mx_2^0 = \hat{x}_1, \qquad \hat{y} = \hat{x}_2.$$

It is not difficult to see that the optimal manifold is $\hat{x}_2 = 0$ and the optimal control satisfies

$$\hat{u} = -msx_2^0 - x_1^0,$$

i.e. in time-domain is $u = -m\delta_0' x_2^0 - \delta x_1^0$. In the limit of a vanishing mass ($m \downarrow 0$), this gives $x_2 = 0$ and $u = -\delta x_1^0$.

Now instead of in the solution, let $m \downarrow 0$ in the equations. This gives the DAE

$$\dot{x}_1 = -x_2 + u, \qquad 0 = x_1, \qquad y = x_2.$$

If $x_1^0 \neq 0$, then no classical solutions exist. In frequency domain we have

$$s\hat{x}_1 - x_1^0 = -\hat{x}_2 + \hat{u}, \qquad \hat{x}_1 = 0, \qquad \hat{y} = \hat{x}_2.$$

It is not difficult to see that $\hat{x}_1 = \hat{x}_2 = 0$ and $\hat{u} = -x_1^0$ is optimal.

We conclude that considering frequency domain trajectories gives the correct answer whereas infimizing over classical trajectories in the DAE would result in the (inconsistent) conclusion that for $x_1^0 \neq 0$, the infimum is infinity.

Remark 19. The above example shows that [17, Theorem 3.3] is incorrect. As shown, the linear quadratic optimal control problem for the DAE when considered over frequency domain trajectories (or equivalently since we are in the finite-dimensional case: over impulsive-smooth trajectories as in [17]) has a solution. It can be verified that (with the definitions as in [17]) the system is stabilizable and [17, Assumption 2.4] is satisfied. It can also be shown that the system is not impulse controllable. This shows that [17, Theorem 3.3] is incorrect. The issue with the proof of [17, Theorem 3.3] is the solution formula [17, (6)] which is incorrect (this is a well-known error in the literature; see [9]).

4.6 The difference with the state stability case

Example 20. Consider the case where $\mathscr{X}$, $\mathscr{U}$ and $\mathscr{Y}$ are one-dimensional and $A = 1$, $B = 1$, $C = 0$, $D = 1$, i.e. the minimization of $\int_0^\infty |u(t)|^2 \, dt$ subject to $x(0) = x^0$ and $\dot{x} = x + u$ (note that this is in fact a regular problem). It is easily seen that (6) has the solutions $X = 0$ and $X = 2$ with corresponding $K = 0$, $L = 1$ and $K = 2$ and $L = 1$ respectively. The smallest nonnegative solution is $X = 0$, which leads to $u = 0$, which clearly indeed gives the optimal cost. The largest solution $X = 2$ instead leads to $u = -2x$ and, as is well-known (since this is the regular case), gives the optimal cost for the problem with state stability.

Remark 21. In [19] the singular linear quadratic optimal control problem ([19, Problem 1]) is *defined* in terms of the matrix from [19, Lemma 1], which is the *largest* (rank-minimizing) nonnegative definite solution of (6). In the case of Example 20, therefore $X = 2$ would be the matrix chosen. In that case [19, Problem 1(iii)] is simply state stability. Hence [19, Problem 1] for Example 20 is simply the problem with state stability and strictly speaking, [19] does not provide a solution to the problem without state stability for Example 20 (and examples like it where the smallest solution and the largest rank-minimizing solution do not coincide), contrary to what seems to be claimed in the abstract and conclusions of [19].

References

[1] Damir Z. Arov and Olof J. Staffans. *Linear state/signal systems*, volume 183 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2022.

[2] Chayan Bhawal, Jan Heiland, and Peter Benner. PD controllers to solve single-input, index-1 DAE based LQR problems. In *2020 European Control Conference (ECC)*, pages 1795–1800, 2020.

[3] John C. Doyle, Bruce A. Francis, and Allen R. Tannenbaum. *Feedback control theory*. Macmillan Publishing Company, New York, 1992.

[4] Ton Geerts. All optimal controls for the singular linear-quadratic problem without stability; a new interpretation of the optimal cost. *Linear Algebra Appl.*, 116:135–181, 1989.

[5] Ton Geerts. Linear-quadratic control with and without stability subject to general implicit continuous-time systems: coordinate-free interpretations of the optimal costs in terms of dissipation inequality and linear matrix inequality; existence and uniqueness of optimal controls and state trajectories. *Linear Algebra Appl.*, 203/204:607–658, 1994.

[6] Anthony Hastir and Timo Reis. Optimal control of infinite-dimensional dissipative systems, 2026.

[7] M. L. J. Hautus and L. M. Silverman. System structure and singular control. *Linear Algebra Appl.*, 50:369–402, 1983.

[8] Rachel Kalpana Kalaimani, Madhu N. Belur, and Debraj Chakraborty. Singular LQ control, optimal PD controller and inadmissible initial conditions. *IEEE Trans. Automat. Control*, 58(10):2603–2608, 2013.

[9] Peter C. Müller. Remark on the solution of linear time-invariant descriptor systems. *Proceedings in Applied Mathematics and Mechanics*, 5(1):175–176, 2005.

[10] Mark Opmeer and Olof Staffans. On linear quadratic optimal control and algebraic Riccati equations for infinite-dimensional differential-algebraic equations. *DAE Panel*, 2, Dec. 2024.

[11] Mark R. Opmeer. Infinite-dimensional linear systems: a distributional approach. *Proc. London Math. Soc. (3)*, 91(3):738–760, 2005.

[12] Mark R. Opmeer. The algebraic Riccati equation for infinite-dimensional systems. In *Proceedings of the MTNS*, 2014.

[13] Mark R. Opmeer and Olof J. Staffans. Optimal state feedback input-output stabilization of infinite-dimensional discrete time-invariant linear systems. *Complex Anal. Oper. Theory*, 2(3):479–510, 2008.

[14] Mark R. Opmeer and Olof J. Staffans. Optimal control on the doubly infinite continuous time axis and coprime factorizations. *SIAM J. Control Optim.*, 52(3):1958–2007, 2014.

[15] Mark R. Opmeer and Olof J. Staffans. Optimal control on the doubly infinite time axis for well-posed linear systems. *SIAM J. Control Optim.*, 57(3):1985–2015, 2019.

[16] Imrul Qais, Chayan Bhawal, and Debasattam Pal. Optimal singular LQR problem: a PD feedback solution. *SIAM J. Control Optim.*, 61(4):2655–2681, 2023.

[17] Imrul Qais, Chayan Bhawal, and Debasattam Pal. Singular LQR optimal control of DAEs: a feedback solution. *Automatica J. IFAC*, 173:Paper No. 112060, 8, 2025.

[18] Timo Reis and Matthias Voigt. Linear-quadratic optimal control of differential-algebraic systems: the infinite time horizon problem with zero terminal state. *SIAM J. Control Optim.*, 57(3):1567–1596, 2019.

[19] Jovan Stefanovski. Impulsive and impulse-free PD solutions of singular LQ control. *Systems Control Lett.*, 196:Paper No. 106024, 11, 2025.

[20] J. C. Willems, A. Kitapci, and L. M. Silverman. Singular optimal control: a geometric approach. *SIAM J. Control Optim.*, 24(2):323–337, 1986.

[21] A. R. Woodyatt, M. M. Seron, J. S. Freudenberg, and R. H. Middleton. Cheap control tracking performance for non-right-invertible systems. *Internat. J. Robust Nonlinear Control*, 12(15):1253–1273, 2002.

[22] Hans Zwart. Transfer functions for infinite-dimensional systems. *Systems Control Lett.*, 52(3-4):247–255, 2004.