Explicit formulas for optimally robust controllers for delay systems[1]

Harry Dym[2], *Tryphon T. Georgiou*[3], *Malcolm C. Smith*[4]

Abstract

This paper considers systems whose transfer functions take the form of a strictly proper rational function times a delay. A closed form expression is presented for the controller which is optimally robust with respect to perturbations measured in the gap metric. The formula allows the $\mathcal{H}_\infty$ loop-shaping procedure of Glover-McFarlane to be carried out explicitly for this class of systems without the need to first find a rational approximation of the plant. The form of the controller involves a certain algebra of "pseudo-derivation" operators. These operators, and their matrix generalizations, play a central role in the derivation of the controller. A discussion of the main properties of these operators will be given. An example will be presented of a controller design to achieve disturbance attenuation and robust set-point following for a plant with two lightly damped poles and a non-trivial time delay. The performance is compared, and shown to be superior, to that of a Smith predictor.

Notation

$\mathbb{R}$ and $\mathbb{C}$ denote the real and complex fields. $\mathcal{H}_2$, $\mathcal{H}_\infty$ denote the standard Hardy spaces in the right half of the complex plane (RHP). $\mathcal{L}_p$ denotes the standard Lebesgue space and $\|\cdot\|_p$ its norm. The inner product in $\mathcal{L}_2$ is denoted by $\langle\cdot,\cdot\rangle$. For $x \in \mathcal{L}_2(-\infty,\infty)$ we denote the Laplace transform of x by $\hat{x}$. If $G(s)$ is a matrix function of s, $G^*(s) := \overline{G(-\overline{s})}^T$ is the conjugate of $G(s)$. If $G(s)$ is real rational, $G^*(s) = G(-s)^T$. $\mathcal{H}_2^n$, $\mathcal{L}_2^n$ etc. denote corresponding spaces of vector-valued functions. $(\mathcal{H}_2^n)^\perp$ denotes the orthogonal complement of $\mathcal{H}_2^n$ in $\mathcal{L}_2^n(-j\infty,j\infty)$. We denote by $\mathbf{\Pi}_+$ (resp. $\mathbf{\Pi}_-$) the orthogonal projection from

[1]This work was supported in part by the NSF, AFOSR, SERC and the Nuffield Foundation.
[2]Department of Theoretical Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel
[3]Dept. of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455, U.S.A.
[4]Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, U.K.

$\mathcal{L}_2^n(-j\infty, j\infty)$ onto $\mathcal{H}_2^n$ (resp. $(\mathcal{H}_2^n)^\perp$). Dimensions will be suppressed if they can be inferred from the context. The symbol spec(A) designates the eigenvalues of A.

1 Introduction

In this paper we consider the class of delay systems of the form $P(s) = e^{-s\tau}P_0(s)$, where $P_0(s)$ is a strictly proper rational function and $\tau > 0$ is a time-delay. We present a closed form formula for controllers which are optimally robust with respect to gap/normalized coprime factor uncertainty. This formula is a generalization of one given in [18] (see also [17]) for the case of a first order $P_0(s)$. The approach also builds on the work of [23] which developed state-space formulas for computing optimal H_∞-performance for such systems.

The derivation of the controllers is based on a synthesis of state-space techniques, associated with the rational part $P_0(s)$, along with the use of a certain commutative algebra of operators. These operators are defined as follows:

$$\partial_\alpha f := \frac{f(s) - f(\alpha)}{s - \alpha}$$

(for $\alpha \in \mathbb{C}$ and $f(s)$ any meromorphic function which is analytic at α), and they represent an analog of the "backward shift" operator in the continuous-time set-up. Some early use of these operators ([2] and [3]) was in the context of reproducing kernel Hilbert spaces and scattering theory—see [7], [8] for more recent accounts, numerous references, and applications to interpolation theory. The formula for the optimal controllers contains infinite-dimensional "distributed delays" resulting from the action of the ∂_α's (or a suitable matrix generalization ∂_H) on the delay element $e^{-s\tau}$.

The controller formula derived in this paper allows the $\mathcal{H}_\infty$ loop shaping design procedure of Glover-McFarlane to be carried out explicitly for the delay systems considered without the need to first find a rational approximation of the plant. We illustrate this fact by designing a controller to achieve disturbance attenuation and robust set-point following for a plant with two lightly damped poles and a non-trivial delay. This type of system is a suitable nominal model in a number of process control applications. We will show that our design has significantly better performance than a Smith predictor.

Research on the subject of $\mathcal{H}_\infty$ control for distributed parameter systems has been ongoing since the mid 1980's. One of the objectives of this research has been to develop explicit formulas which provide insight into the limitations on performance imposed by infinite-dimensional elements such as delays. An additional goal has been to understand the effect of such elements on the structure of the controller, thereby giving an alternative design approach to finite-dimensional approximation of the plant. Early work was concerned with computing the optimal performance for problems such as weighted sensitivity minimization [10], [13]. Subsequent efforts focused mainly on developing computational methodologies for increasingly general situations [14], [21], [27], [31] and for explicitly computing the controllers. The computation of the optimal performance involves finding the maximum singular value (norm) of a certain generalized Hankel or Sarason operator. The construction of the optimal controller requires computing the associated singular vectors. Extra care is needed in case the infimum is not achieved in the optimization problem [11] or in case the underlying operator is not compact. (We mention incidentally that neither of these two cases arises in the problem treated here.) The computation of singular values and vectors usually exploits special structures in the family of systems considered, such as the presence of general $\mathcal{H}_\infty$ inner factors multiplied by rational functions. In the case of delay systems, the computations were originally cast in the form of a certain two-point boundary value problem in the time domain [13], [23], [32] which in turn led to a solution in the form of a transcendental equation. Analogous computations can be carried out in the s-domain, or in the z-domain after a bilinear transformation, as in the skew-Toeplitz approach in [1] and [21]. The general methodology has been shown to apply to quite a general class of distributed parameter systems and provides a framework for obtaining a numerical solution [12], [22]. An alternate viewpoint for $\mathcal{H}_\infty$ control of distributed parameter systems, based on an infinite dimensional state space approach, has been considered in [5], [6] and the references therein.

The current work is part of an on-going effort to realize the original goals of the above research program. A formula for an $\mathcal{H}_\infty$ controller is derived which takes a particularly simple form and is easily programmable. Moreover, the controller minimizes a certain cost functional which has proved to be especially convenient for design. We would like to mention that, in some recent work [28], an algorithm has been outlined for the computation of suboptimal controllers of delay systems for robustness in the gap. At the moment, it is an interesting open problem to seek an explicit formula of the type given in this paper for the suboptimal case.

The paper is organized as follows. In Section 2 some background material is reviewed on the properties and computation of gap/normalized coprime factor optimally robust controllers. In Section 3 the singular value/vector equations of Partington-Glover, and the procedure to find the gap optimal controller from them, are summarized. In Section 4 the (scalar) operator ∂_α is introduced and the main properties of the operator, which will be used in the controller derivation, will be described. In Section 5 the general procedure for the computation of the optimal controller will be carried out for the case where $P(s) = e^{-s\tau}/(s-\sigma)$. The purpose of this section is to motivate the computational steps that are necessary in the higher order case. In Section 6 a generalization of the ∂_α operator to the matrix case will be presented together with its properties. In Section 7 the general controller formula is derived and, in Section 8, an illustrative example is presented.

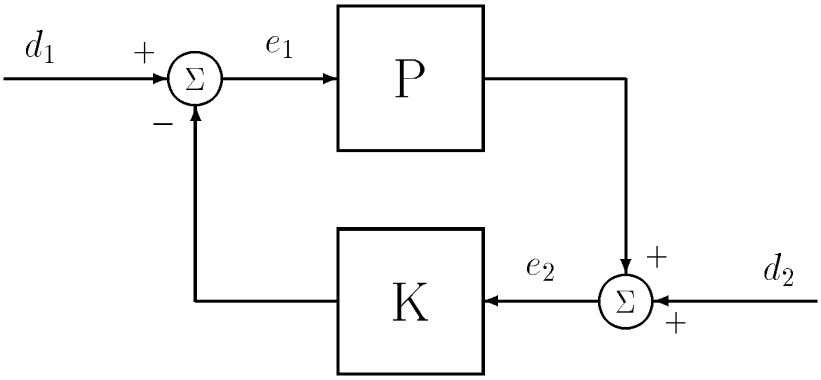

Figure 1: Standard feedback configuration

2 Background

We consider the feedback configuration of Fig. 1, denoted by $[P, K]$, where P and K are the transfer functions of the plant and the controller respectively (and are ratios of $\mathcal{H}_\infty$-functions). We say that $[P, K]$ is *stable* if the closed loop transfer functions, mapping d_i to e_j $(i, j = 1, 2)$, belong to $\mathcal{H}_\infty$. If $[P, K]$ is stable then the following quantity can be defined:

$$b_{P,K} := \left\| \begin{pmatrix} I \\ P \end{pmatrix} (I + KP)^{-1} \, (I \ , \ K) \right\|_\infty^{-1},$$

which is the gap metric robustness radius for the feedback system. That is, $b_{P,K}$ is the radius of the largest ball of uncertainty about P, measured in the gap metric, which is stabilized by K. For the definition of the gap metric, connections with normalized coprime factorizations, and associated results on robust stabilization the reader is referred to [16], [18].

It appears that $b_{P,K}$ is a useful quantity to maximize for a feedback system, especially if it is weighted appropriately. In fact, the maximization of $b_{WP,K}$, for some weighting function W chosen by the designer, is the basis of the Glover-McFarlane $\mathcal{H}_\infty$ loop shaping method [20]. This problem is also the same as the optimal robustness problem in the weighted gap metric (see [4], [15]). From the point of view of computations, it is sufficient to consider the unweighted case only. This will be our point of view until the example in the last section.

We now summarize some background results on the optimization of $b_{P,K}$. Let $P = N/M$, where $M, N \in \mathcal{H}_\infty$, is a normalized coprime factorization of P (i.e., $NN^* + MM^* = 1$) and let $F = (-N, \ M)$. We define the *optimal robustness radius*

$$b_{\mathrm{opt}}(P) = \sup_{K \text{ stblz}} b_{P,K}. \tag{1}$$

It can be shown that the following formula holds

$$b_{\mathrm{opt}}(P) = (1 - \|\mathbf{\Gamma}\|^2)^{1/2} \tag{2}$$

where $\mathbf{\Gamma}$ is the Hankel operator defined by $\mathbf{\Gamma} = \mathbf{\Pi}_- F^*|_{\mathcal{H}_2}$ (see [18], [16]).

If F is continuous on the compactified right half plane then it follows (from Hartman's Theorem) that the operator $\mathbf{\Gamma}$ is compact and hence that it achieves its norm $\|\mathbf{\Gamma}\|$, which is equal to the maximal singular value of the operator. Moreover if $0 < \lambda = \|\mathbf{\Gamma}\| < 1$ then there is a unique optimally robust controller which can be determined from λ and any corresponding singular vector. Specifically, let $\hat{x} \in \mathcal{H}_2$, $\hat{y} = (\hat{y}_1,\ \hat{y}_2) \in \mathcal{H}_2^{1\times 2}$ satisfy

$$\begin{aligned} \mathbf{\Gamma}\hat{x} &= \lambda \hat{y}^*, & (3)\\ \mathbf{\Gamma}^*\hat{y}^* &= \lambda \hat{x} & (4) \end{aligned}$$

and define $\hat{u} = \mathbf{\Pi}_+ F^* \lambda \hat{x}$. Then the optimal compensator has a transfer function $K = -\hat{u}_1/\hat{u}_2$. See [17] for a more detailed explanation of these facts and further references.

3 Computation of singular values and vectors

In this section we consider a single-input/single-output system with transfer function of the form $P(s) = e^{-s\tau} P_0(s)$ where $P_0(s)$ is a strictly proper rational function and τ is a time-delay. Since $e^{-s\tau}$ is inner we can find the normalized coprime factors of $P(s)$ by finding the normalized coprime factors of $P_0(s)$ using the state-space construction of [19], [29] and then multiplying the numerator by $e^{-s\tau}$. Let $P_0(s) = C(sI - A_0)^{-1}B$ be a minimal realization with state dimension n. We define R to be the stabilizing solution of the Riccati equation

$$A_0 R + R A_0 - R C^* C R + B B^* = 0 \tag{5}$$

and set $A = A_0 + EC$ where $E = -RC^*$. Then spec(A) is in the open left half plane, and

$$F = \begin{pmatrix} -N, & M \end{pmatrix} = \begin{pmatrix} 0, & I \end{pmatrix} + C(sI - A)^{-1} \begin{pmatrix} -Be^{-s\tau}, & E \end{pmatrix}. \tag{6}$$

We assume that $P(s) \not\equiv 0$ which implies that $\|\Gamma\| > 0$. Since $F(s)$ is right invertible over $\mathcal{H}_\infty$, which is equivalent to P being stabilizable, $\|\Gamma\| < 1$ (see [16, Sections III and VII]). Finally we note that, since $F(s)$ is continuous on the compactified right half plane, the optimal controller for $P(s)$ can be found from the singular values and vectors of $\mathbf{\Gamma}$.

We now summarize the formulae of Partington and Glover [23] for the computation of the singular values of $\mathbf{\Gamma}$. The idea is to give a time-domain realization of the equations (3) and (4). These are the equations

$$w(\tau) = \int_\tau^\infty e^{A^*(u-\tau)} C^* x(u)\, du, \tag{7}$$

$$\begin{aligned}
\dot{w}(t) &= -A^*w(t) - C^*x(t) && \text{for } 0 \le t \le \tau, \\
\lambda y_1(t) &= -B^*w(\tau - t) && \text{for } 0 \le t \le \tau, \\
\lambda \begin{bmatrix} y_1(t+\tau) \\ y_2(t) \end{bmatrix} &= \begin{bmatrix} -B^* \\ E^* \end{bmatrix} e^{A^*t} w_0 && \text{for } t > 0.
\end{aligned}$$

and

$$\begin{aligned}
v(0) &= \int_0^\infty e^{Au} \left(-By_1(u+\tau) + Ey_2(u)\right) du, && (8) \\
\dot{v}(t) &= Av(t) - By_1(\tau - t) \qquad \text{for } 0 \le t \le \tau, \\
\lambda x(t) &= \begin{cases} Cv(t) & \text{for } 0 \le t \le \tau \\ Ce^{A(t-\tau)}v(\tau) & \text{for } t \ge \tau \end{cases} && (9)
\end{aligned}$$

These equations can be solved by considering a two-point boundary value problem on the interval $[0, \tau]$. On this interval the differential equations become

$$\frac{d}{dt}\begin{bmatrix} v \\ w \end{bmatrix} = H \begin{bmatrix} v \\ w \end{bmatrix}, \tag{10}$$

where

$$H = \begin{bmatrix} A & \lambda^{-1}BB^* \\ -\lambda^{-1}C^*C & -A^* \end{bmatrix}. \tag{11}$$

Thus

$$\begin{bmatrix} v(t) \\ w(t) \end{bmatrix} = e^{Ht} \begin{bmatrix} v(0) \\ w(0) \end{bmatrix}. \tag{12}$$

Next, by substituting for x and y in (8) and (7), we obtain the relations

$$\begin{aligned}
v(0) &= \lambda^{-1}Rw(0), && (13) \\
w(\tau) &= \lambda^{-1}Sv(\tau), && (14)
\end{aligned}$$

where S is the unique solution of the Lyapunov equation

$$A^*S + SA + C^*C = 0. \tag{15}$$

(Incidentally, observe that R satisfies the Lyapunov equation $AR + RA^* + BB^* + EE^* = 0$.) This leads to the equation

$$0 = L(\lambda)w(0),$$

where

$$L(\lambda) := \left\{ \begin{bmatrix} -S & \lambda I \end{bmatrix} \exp(H\tau) \begin{bmatrix} R \\ \lambda I \end{bmatrix} \right\}. \tag{16}$$

Proposition 1 [23]. A necessary and sufficient condition for λ to be a singular value of $\mathbf{\Gamma} = \mathbf{\Pi}_- F^*|_{\mathcal{H}_2}$ is that $\det L(\lambda) = 0$. $\Box$

It follows that, in order to find the optimal controller for $P(s)$, we can use the following procedure:

P1. Find the largest λ in $[0, 1]$ such that $\det L(\lambda) = 0$ and then find a non-zero vector w_0 such that $L(\lambda)w_0 = 0$. (By the assumptions on $P(s)$, $\lambda \neq 0, 1$.) Then find $v_0 = \lambda^{-1} R w_0$ and solve for $v(t)$ on $[0, \tau]$ using equation (12) with $v(0) = v_0$ and $w(0) = w_0$.

P2. Calculate $\lambda x(t)$ on $[0, \infty)$ from equation (9).

P3. Find $\hat{u} = \mathbf{\Pi}_+ F^* \lambda \hat{x}$, whence $K(s) = -\hat{u}_1/\hat{u}_2$ [17].

4 The operator ∂_α

Let $f(s)$ be analytic in $\Omega \subset \mathbb{C}$, and define

$$(\partial_\alpha f)(s) := \frac{f(s) - f(\alpha)}{s - \alpha}$$

for $\alpha \in \Omega$. Then ∂_α defines a linear operator and has a number of interesting properties (see, e.g., [7] and [8]). It is readily verified that it satisfies the *resolvent identity*

$$\partial_\alpha - \partial_\beta = (\alpha - \beta)\partial_\alpha \partial_\beta \tag{17}$$

for any pair of points α, $\beta \in \mathbb{C}$ which are points of analyticity for the class of functions acted upon. A direct consequence is that $\partial_\alpha \partial_\beta = \partial_\beta \partial_\alpha$. The following *pseudo-derivation* identities take the place of a product rule

$$(\partial_\alpha fg)(s) = (\partial_\alpha f)(s) g(s) + f(\alpha)(\partial_\alpha g)(s), \tag{18}$$

$$(\partial_\alpha fg)(s) = f(s)(\partial_\alpha g)(s) + (\partial_\alpha f)(s) g(\alpha). \tag{19}$$

(They differ from the usual product rule of differentiation by the fact that one of the two terms involves evaluation of one of the functions at the point α.) If $\alpha + \gamma \neq 0$ then

$$\partial_\alpha \frac{1}{s + \gamma} = \frac{-1/(\alpha + \gamma)}{(s + \gamma)}. \tag{20}$$

If $\alpha + \gamma \neq 0$ and g is analytic at α, then (18), (19), and (20) give the following two identities:

$$\partial_\alpha \frac{g}{s + \gamma} = \frac{1}{s + \gamma}\left(\partial_\alpha g - \frac{1}{\alpha + \gamma} g(\alpha)\right), \tag{21}$$

$$\partial_\alpha \frac{g}{s + \gamma} = \frac{1}{\alpha + \gamma}\left(\partial_\alpha - \frac{1}{s + \gamma}\right) g. \tag{22}$$

If $\mathcal{R}e(\gamma) > 0$, and $g \in \mathcal{H}_2$ then

$$\mathbf{\Pi}_+ \frac{1}{s-\gamma} g = \partial_\gamma g. \tag{23}$$

Finally we present the following result which will be needed in the controller derivations of Sections 5 and 7.

Proposition 2. Let $\mathcal{R}e(\gamma) > 0$ and $\alpha + \gamma \neq 0$. Suppose $F \in \mathcal{H}_\infty$ admits a representation of the form $F(s) = c + \int_0^\infty e^{-st} f(t)dt$ with c constant and $f(t) \in \mathcal{L}_1[0,\infty)$. Then $\mathbf{\Pi}_+ F^* \partial_\alpha e^{-s\tau}/(s+\gamma)^k$ is analytic in $\mathbb{C} \setminus \{-\gamma\}$, and

$$\mathbf{\Pi}_+ F^* \partial_\alpha \frac{e^{-s\tau}}{(s+\gamma)^k} = \partial_\alpha \mathbf{\Pi}_+ F^* \frac{e^{-s\tau}}{(s+\gamma)^k} \tag{24}$$

for $k = 1, 2, \ldots$.

Proof. We consider first the case $k = 1$. Under the given assumptions it is readily checked that

$$\begin{aligned}
\mathbf{\Pi}_+ F^* \partial_\alpha \frac{1}{s+\gamma} &= -\frac{1}{\alpha+\gamma} \mathbf{\Pi}_+ F^* \frac{1}{s+\gamma} \\
&= -\frac{1}{\alpha+\gamma} \frac{F^*(-\gamma)}{s+\gamma} \qquad (25) \\
&= \partial_\alpha \frac{F^*(-\gamma)}{s+\gamma} \\
&= \partial_\alpha \mathbf{\Pi}_+ F^* \frac{1}{s+\gamma}. \qquad (26)
\end{aligned}$$

Equation (25) shows that $\mathbf{\Pi}_+ F^* \partial_\alpha 1/(s+\gamma)$ is analytic in $\mathbb{C} \setminus \{-\gamma\}$. The proof now proceeds in two steps. We first introduce the natural projection operator, in the Wiener algebra on the line, defined as follows:

$$\mathbf{P} : \int_{-\infty}^\infty e^{-st} g(t)dt \to \int_0^\infty e^{-st} g(t)dt$$

with $g \in \mathcal{L}_1(\mathbb{R})$ and $s \in j\mathbb{R}$.

STEP 1. Let $F_1(s) = F(s) - c = \int_0^\infty e^{-st} f(t)dt$. Then $\mathbf{P} F_1^* e^{-s\tau}$ is an entire function and

$$\mathbf{P} F_1^* \partial_\alpha e^{-s\tau} = \partial_\alpha \mathbf{P} F_1^* e^{-s\tau}$$

for every choice of $\alpha \in \mathbb{C}$.

Proof of STEP 1. The proof is an adaptation of [9, Lemma 4.4]. Observe that

$$\begin{aligned}
\mathbf{P} F_1^* e^{-s\tau} &= \mathbf{P} \int_{-\infty}^0 e^{-st} f^*(t)dt \, . \, e^{-s\tau} \\
&= \mathbf{P} \int_{-\infty}^\tau e^{-su} f^*(u-\tau)du
\end{aligned}$$

$$= \int_0^\tau e^{-su} f^*(u-\tau)du,$$

and hence $\mathbf{P}F_1^* e^{-s\tau}$ is entire since it is the Laplace transform of a function with support on the interval $[0,\tau]$. Further

$$\begin{aligned}
\partial_\alpha \mathbf{P}F_1^* e^{-s\tau} &= \int_0^\tau (\partial_\alpha e^{-su}) f^*(u-\tau)du \\
&= \int_0^\tau \int_0^u e^{-(s-\alpha)v} dv(-e^{\alpha u}) f^*(u-\tau)du \\
&= -\int_0^\tau \int_v^\tau e^{\alpha(v-u)} f^*(u-\tau)du\, dv \\
&= -\int_0^\tau \int_v^\tau e^{\alpha(x-\tau)} f^*(v-x)dx\, dv\,. \qquad (27)
\end{aligned}$$

On the other hand

$$\begin{aligned}
\mathbf{P}F_1^* \partial_\alpha e^{-s\tau} &= \mathbf{P}\int_0^\tau e^{-sx}\left(-e^{\alpha(x-\tau)}\right)dx \int_{-\infty}^0 e^{-st} f^*(t)dt \\
&= \mathbf{P}\int_0^\tau \left(-e^{\alpha(x-\tau)}\right) \int_{-\infty}^x e^{-sv} f^*(v-x)dv\, dx
\end{aligned}$$

which turns into the expression (27) after interchanging the order of integration and then taking the projection. $\bigtriangledown$

STEP 2. The function $\mathbf{\Pi}_+ F^* \partial_{-\gamma} e^{-s\tau}$ is an entire function, and

$$\mathbf{\Pi}_+ F^* \partial_\alpha \partial_{-\gamma} e^{-s\tau} = \partial_\alpha \mathbf{\Pi}_+ F^* \partial_{-\gamma} e^{-s\tau}$$

for every complex number α such that $\alpha + \gamma \neq 0$.

Proof of STEP 2. It follows using STEP 1 that $\mathbf{\Pi}_+ F_1^* \partial_{-\gamma} e^{-s\tau}$ is entire. Clearly $\mathbf{\Pi}_+ c^* \partial_{-\gamma} e^{-s\tau} = c^* \partial_{-\gamma} e^{-s\tau}$ is also entire. We now verify the formula for each of the components of $F(s) = c + F_1(s)$ separately. The first is easy:

$$\mathbf{\Pi}_+ c^* \partial_\alpha \partial_{-\gamma} e^{-s\tau} = c^* \partial_\alpha \partial_{-\gamma} e^{-s\tau} = \partial_\alpha \mathbf{\Pi}_+ c^* \partial_{-\gamma} e^{-s\tau}.$$

Next, for $\alpha \neq -\gamma$ we have

$$\begin{aligned}
\mathbf{\Pi}_+ F_1^* \partial_\alpha \partial_{-\gamma} e^{-s\tau} &= \mathbf{\Pi}_+ F_1^* \frac{\partial_\alpha - \partial_{-\gamma}}{\alpha+\gamma} e^{-s\tau} \\
&= \mathbf{P}F_1^* \frac{\partial_\alpha - \partial_{-\gamma}}{\alpha+\gamma} e^{-s\tau} \\
&= \frac{\partial_\alpha - \partial_{-\gamma}}{\alpha+\gamma} \mathbf{P}F_1^* e^{-s\tau} \\
&= \partial_\alpha \partial_{-\gamma} \mathbf{P}F_1^* e^{-s\tau} \\
&= \partial_\alpha \mathbf{P}F_1^* \partial_{-\gamma} e^{-s\tau} \\
&= \partial_\alpha \mathbf{\Pi}_+ F_1^* \partial_{-\gamma} e^{-s\tau},
\end{aligned}$$

using STEP 1. The operator $\mathbf{P}$ is needed here because $F_1^* e^{-s\tau}$ may not belong to $\mathcal{L}_2(j\mathbb{R})$ and hence $\mathbf{\Pi}_+$ cannot always be applied to it directly. $\bigtriangledown$

With the help of the preceding analysis, we can now prove the proposition for the case $k = 1$. For $\alpha \neq -\gamma$,

$$\begin{aligned}\mathbf{\Pi}_+ F^* \partial_\alpha \frac{e^{-s\tau}}{s+\gamma} &= \mathbf{\Pi}_+ F^* \partial_\alpha \left\{ \partial_{-\gamma} e^{-s\tau} + \frac{e^{\gamma\tau}}{s+\gamma} \right\} \\ &= \partial_\alpha \mathbf{\Pi}_+ F^* \left\{ \partial_{-\gamma} e^{-s\tau} + \frac{e^{\gamma\tau}}{s+\gamma} \right\},\end{aligned}$$

using (26) and STEP 2. This establishes the required equality.

To complete the proof we note that the result for the case of $k = 1$ can be applied to:

$$g_n(s) = \frac{1}{(s+\gamma)(s+\gamma+\frac{1}{n})\dots(s+\gamma+\frac{k-1}{n})},$$

for n sufficiently large, to give:

$$\mathbf{\Pi}_+ F^* \partial_\alpha e^{-s\tau} g_n(s) = \partial_\alpha \mathbf{\Pi}_+ F^* e^{-s\tau} g_n(s).$$

By letting $n \to \infty$ we can check in turn that

$$\begin{aligned} g_n(s) &\to g(s) = \frac{1}{(s+\gamma)^k} \text{ in } \mathcal{L}_2(j\mathbb{R}), \\ \partial_\alpha g_n(s) &\to \partial_\alpha g(s) \text{ in } \mathcal{L}_2(j\mathbb{R}), \text{ and} \\ \mathbf{\Pi}_+ F^* e^{-s\tau} g_n(s) &\to \mathbf{\Pi}_+ F^* e^{-s\tau} g(s) \end{aligned}$$

pointwise in $\mathbb{C} \setminus \{-\gamma\}$. Thus

$$\begin{aligned} \mathbf{\Pi}_+ F^* \partial_\alpha e^{-s\tau} g(s) &= \lim_{n\to\infty} \mathbf{\Pi}_+ F^* \partial_\alpha e^{-s\tau} g_n(s) \\ &= \lim_{n\to\infty} \partial_\alpha \mathbf{\Pi}_+ F^* e^{-s\tau} g_n(s) \\ &= \partial_\alpha \mathbf{\Pi}_+ F^* e^{-s\tau} g(s). \quad \square \end{aligned}$$

5 First-order case.

In the present section we consider the simplest example of a plant having transfer function of the form "time-delay $\times$ rational function": $P(s) = e^{-s\tau}/(s-\sigma)$. An expression for the optimal controller was derived in [17]. A simplified formula involving the operator ∂_α was given in [18]. Below we present a new derivation of the optimal controller making explicit use of certain properties of the operator ∂_α. The computations illustrate several key steps which motivate analogous steps in the general case.

Proposition 3. Let $P(s) = e^{-s\tau}/(s-\sigma)$ where σ is any real constant and $\tau > 0$.

(a) The optimal robustness radius for gap ball uncertainty is given by

$$b_{\rm opt}(P) = (1 - \lambda_{\rm max}^2)^{1/2}, \tag{28}$$

where $\lambda_{\rm max}$ is the largest solution of the following transcendental equation:

$$e^{2\mu\tau} = \frac{\sigma + \mu}{\sigma - \mu} \left(\frac{\gamma - \mu}{\gamma + \mu} \right)^2, \tag{29}$$

where $\mu = \mu(\lambda) = \sqrt{\gamma^2 - \lambda^{-2}}$ and $\gamma = \sqrt{1 + \sigma^2}$.

(b) The transfer function of the optimal controller is given by:

$$K_{\rm opt}(s) = \frac{1}{\sqrt{\sigma^2 - \mu^2} - \frac{\gamma^2 - \mu^2}{2\mu} \left((\sigma + \mu)\partial_\mu - (\sigma - \mu)\partial_{-\mu} \right) e^{-s\tau}}, \tag{30}$$

where $\mu = \sqrt{\gamma^2 - \lambda_{\rm max}^{-2}}$, in the case $\mu \neq 0$. [A general formula covering also the case $\mu = 0$ is given in Theorem 1.]

Proof. Setting $B = C = 1$, $A_0 = \sigma$ gives $A = -\gamma$, $R = \sigma + \gamma$ and $S = 1/2\gamma$, whence

$$F(s) = (-\frac{e^{-s\tau}}{s + \gamma}, \ \frac{s - \sigma}{s + \gamma}).$$

Then

$$H = \begin{pmatrix} -\gamma & \lambda^{-1} \\ -\lambda^{-1} & \gamma \end{pmatrix}$$

with eigenvalues $\pm\mu$ where $\mu = \sqrt{\gamma^2 - \lambda^{-2}}$. Note that μ can be complex but $\pm\mu \neq \gamma$. An eigenvalue-eigenvector decomposition of H leads to

$$e^{Ht} = \begin{pmatrix} \lambda^{-1} & \lambda^{-1} \\ \gamma + \mu & \gamma - \mu \end{pmatrix} \begin{pmatrix} e^{\mu t} & 0 \\ 0 & e^{-\mu t} \end{pmatrix} \begin{pmatrix} \gamma - \mu & -\lambda^{-1} \\ -(\gamma + \mu) & \lambda^{-1} \end{pmatrix} \frac{-\lambda}{2\mu}.$$

By substituting into (16) it follows that

$$\begin{aligned} L(\lambda) \ = \ 0 \ \Leftrightarrow \ e^{2\mu\tau} \ &= \ \frac{-\lambda^{-2} + 2\gamma(\gamma - \mu)}{-\lambda^{-2} + 2\gamma(\gamma + \mu)} \cdot \frac{(\sigma + \gamma)(\gamma + \mu) - 1}{(\sigma + \gamma)(\gamma - \mu) - 1} \\ &= \ \frac{(\gamma - \mu)^2}{(\gamma + \mu)^2} \cdot \frac{(\sigma + \gamma)(\sigma + \mu)}{(\sigma + \gamma)(\sigma - \mu)}. \end{aligned}$$

This gives (29). Also (28) follows from (2). Setting $w_0 = 1$ gives $v_0 = \lambda^{-1} R$, which leads to

$$\begin{aligned} v(t) \ &= \ \begin{pmatrix} 1 & 0 \end{pmatrix} e^{Ht} \begin{pmatrix} \lambda^{-1} R \\ 1 \end{pmatrix} \\ &= \ \alpha e^{\mu t} + \beta e^{-\mu t} \end{aligned}$$

on $[0,\tau]$, where

$$\begin{aligned}\alpha &= \frac{-(\sigma+\gamma)(\sigma-\mu)}{2\mu\lambda}, \\ \beta &= \frac{(\sigma+\gamma)(\sigma+\mu)}{2\mu\lambda}.\end{aligned}$$

(The calculation of α and β is eased by judicious use of the fact that $1=\gamma^2-\sigma^2$.) Thus

$$\lambda x(t) = \begin{cases} \alpha e^{\mu t}+\beta e^{-\mu t} & \text{on } [0,\tau] \\ e^{-\gamma(t-\tau)}(\alpha e^{\mu\tau}+\beta e^{-\mu\tau}) & \text{on } [\tau,\infty). \end{cases}$$

Hence

$$\begin{aligned}\lambda\hat{x} &= \int_0^\infty \lambda x(t) e^{-st}dt \\ &= -\alpha e^{\mu\tau}\left(\partial_\mu - \frac{1}{s+\gamma}\right)e^{-s\tau} - \beta e^{-\mu\tau}\left(\partial_{-\mu} - \frac{1}{s+\gamma}\right)e^{-s\tau} \\ &= \left(-\alpha(\gamma+\mu)e^{\mu\tau}\partial_\mu - \beta(\gamma-\mu)e^{-\mu\tau}\partial_{-\mu}\right)\frac{e^{-s\tau}}{s+\gamma}\end{aligned}$$

where the first step uses direct integration and the second step follows from (22). Next

$$\begin{aligned}\mathbf{\Pi}_+F^*\lambda\hat{x} &= \mathbf{\Pi}_+\begin{pmatrix} e^{s\tau}/(s-\gamma) \\ (s+\sigma)/(s-\gamma)\end{pmatrix}\lambda\hat{x} \\ &= \left(-\alpha(\gamma+\mu)e^{\mu\tau}\partial_\mu - \beta(\gamma-\mu)e^{-\mu\tau}\partial_{-\mu}\right)\mathbf{\Pi}_+\begin{pmatrix} e^{s\tau}/(s-\gamma) \\ (s+\sigma)/(s-\gamma)\end{pmatrix}\frac{e^{-s\tau}}{s+\gamma} \\ &= \left(-\alpha(\gamma+\mu)e^{\mu\tau}\partial_\mu - \beta(\gamma-\mu)e^{-\mu\tau}\partial_{-\mu}\right)\begin{pmatrix} -1/2\gamma(s+\gamma) \\ (1+(\sigma+\gamma)\partial_\gamma)\,e^{-s\tau}/(s+\gamma)\end{pmatrix}\end{aligned}$$

where we have used Proposition 2 in the first step (which is applicable because $\pm\mu+\gamma\neq 0$), and equations (20) and (23) in the second step. Using (20), the first element of $\mathbf{\Pi}_+F^*\lambda\hat{x}$ becomes

$$\frac{-1}{2\gamma(s+\gamma)}(\alpha e^{\mu\tau}+\beta e^{-\mu\tau}) = \frac{-(\sigma+\gamma)\sqrt{\sigma^2-\mu^2}}{\lambda(\gamma^2-\mu^2)(s+\gamma)} \tag{31}$$

where the equality follows by substituting for α, β and $e^{\mu\tau} = \sqrt{\frac{\sigma+\mu}{\sigma-\mu}}\left(\frac{\gamma-\mu}{\gamma+\mu}\right)$. The second element of $\mathbf{\Pi}_+F^*\lambda\hat{x}$ is equal to

$$\begin{aligned}\Delta = &\left\{(-\alpha(\gamma+\mu)e^{\mu\tau})\left(1+\frac{\sigma+\gamma}{\mu-\gamma}\right)\partial_\mu + \left(-\beta(\gamma-\mu)e^{-\mu\tau}\right)\left(1+\frac{\sigma+\gamma}{-\mu-\gamma}\right)\partial_{-\mu}\right. \\ &\left. +(\sigma+\gamma)\left(\frac{\alpha(\gamma+\mu)}{\mu-\gamma}e^{\mu\tau}+\frac{\beta(\gamma-\mu)}{-\mu-\gamma}e^{-\mu\tau}\right)\partial_\gamma\right\}\frac{e^{-s\tau}}{s+\gamma}\end{aligned}$$

after using (17). In fact, by substituting for α, β and $e^{\mu\tau}$, the coefficient of ∂_γ in Δ is easily seen to be zero. Next, using similar substitutions, we get

$$\begin{aligned}
\Delta &= \frac{(\sigma+\gamma)\sqrt{\sigma^2-\mu^2}}{2\mu\lambda}\left\{-(\sigma+\mu)\partial_\mu+(\sigma-\mu)\partial_{-\mu}\right\}\frac{e^{-s\tau}}{s+\gamma} \\
&= \frac{(\sigma+\gamma)\sqrt{\sigma^2-\mu^2}}{2\mu\lambda(s+\gamma)}\left\{-(\sigma+\mu)\partial_\mu+(\sigma-\mu)\partial_{-\mu}\right\}e^{-s\tau} \\
&\quad + \frac{(\sigma+\gamma)\sqrt{\sigma^2-\mu^2}}{2\mu\lambda(s+\gamma)}\frac{\sqrt{\sigma^2-\mu^2}}{\gamma^2-\mu^2}2\mu,
\end{aligned}$$

where the second step uses (21). Taking the quotient of (31) and the above expression for Δ gives the required result. $\square$

6 The matrix ∂-operator

To facilitate the controller computations in higher order cases we will need a matrix generalization of the operator ∂_α and its properties.

Let $A \in \mathbb{C}^{n\times n}$ and let $F(s)$ be an $n \times r$ matrix whose elements are analytic in an open simply connected region $\Omega \subset \mathbb{C}$ for which $\mathrm{spec}(A) \subset \Omega$. Then we define:

$$(\partial_A F)(s) = \frac{1}{2\pi j}\int_{\gamma_1}\frac{1}{z-s}(zI-A)^{-1}F(z)dz, \tag{32}$$

for s inside γ_1, where γ_1 is any simple closed curve (with positive orientation) inside Ω which encloses $\mathrm{spec}(A)$. We also define $(\partial_A F)(s)$ by the same formula if either A or $F(s)$ is scalar. The above definition (32) is independent of the choice of γ_1 and can be evaluated explicitly in terms of the Jordan decomposition of A and the values of F and its derivatives at the points of the spectrum of A, as in (39) below. The present contour integral formulation is, however, often easier to work with. In particular, it follows from Cauchy's theorem that (32) yields

$$(\partial_\alpha F)(s) = \frac{F(s)-F(\alpha)}{s-\alpha}$$

for $s, \alpha \in \Omega$, when $A = \alpha$ is a scalar. This is consistent with the definition given in Section 4. On the other hand, if F is equal to a scalar function f and A is either a matrix or a scalar, it is standard to define

$$f(A) = \frac{1}{2\pi j}\int_{\gamma_1}(zI-A)^{-1}f(z)dz,$$

where γ_1 is any simple closed curve inside Ω which encloses $\mathrm{spec}(A)$ (see e.g., [25, Chapter

VI.3]). Using the identity

$$(zI - sI)^{-1} - (zI - A)^{-1} = (zI - A)^{-1}(sI - A)(zI - sI)^{-1},$$

we see that

$$\begin{aligned}(\partial_A f)(s) &= (sI - A)^{-1}\left\{\frac{1}{2\pi j}\int_{\gamma_1}\frac{1}{z-s}f(z)dzI - \frac{1}{2\pi j}\int_{\gamma_1}(zI - A)^{-1}f(z)dz\right\} \\ &= (sI - A)^{-1}(f(s)I - f(A)).\end{aligned}$$

Let $A \in \mathbb{C}^{n\times n}$ and let $F(s)$ be an $n \times r$ matrix whose elements are analytic in an open simply connected region $\Omega \subset \mathbb{C}$. Take $\alpha \in \Omega$ with $\alpha \notin \mathrm{spec}(A)$. Then we can show that the following generalizations of (20–22) hold:

$$\begin{aligned}\partial_\alpha (sI - A)^{-1} &= -(sI - A)^{-1}(\alpha I - A)^{-1} \\ &= -(\alpha I - A)^{-1}(sI - A)^{-1}, && (33)\\ \partial_\alpha \left((sI - A)^{-1}F(s)\right) &= (\alpha I - A)^{-1}\left(\partial_\alpha F(s) - (sI - A)^{-1}F(s)\right), && (34)\\ \partial_\alpha \left((sI - A)^{-1}F(s)\right) &= (sI - A)^{-1}\left(\partial_\alpha F(s) - (\alpha I - A)^{-1}F(\alpha)\right). && (35)\end{aligned}$$

For the matrix ∂ operator, a version of the *resolvent identity* still holds. Let $A, B \in \mathbb{C}^{n\times n}$ with $AB = BA$ then

$$\partial_A - \partial_B = (A - B)\partial_A\partial_B \tag{36}$$

where the operators act on a matrix $F(s)$ whose elements are analytic in an open simply connected region $\Omega \subset \mathbb{C}$ containing $\mathrm{spec}(A) \cup \mathrm{spec}(B)$. To see this, take two simple closed curves γ_1, γ_2 inside Ω, both containing $\mathrm{spec}(A) \cup \mathrm{spec}(B)$, and with γ_1 strictly inside γ_2. Then, for s inside γ_1,

$$\begin{aligned}(\partial_A\partial_B F)(s) &= \frac{-1}{4\pi^2}\int_{\gamma_1}\frac{1}{z-s}(zI - A)^{-1}\int_{\gamma_2}\frac{1}{w-z}(wI - B)^{-1}F(w)\,dw\,dz \\ &= \frac{-1}{4\pi^2}\int_{\gamma_2}\left\{\int_{\gamma_1}\frac{1}{w-z}\cdot\frac{1}{z-s}(zI - A)^{-1}dz\right\}(wI - B)^{-1}F(w)dw \\ &= \frac{1}{2\pi j}\int_{\gamma_2}\frac{1}{w-s}(wI - A)^{-1}(wI - B)^{-1}F(w)dw,\end{aligned}$$

where the last step uses the fact that w is outside of γ_1. The required result now follows from the identity:

$$(A - B)(wI - A)^{-1}(wI - B)^{-1} = (wI - A)^{-1} - (wI - B)^{-1}.$$

Let $f = f_+ + f_- \in \mathcal{L}_2(-j\infty, j\infty)$, where $f_+ \in \mathcal{H}_2$ and $f_- \in \mathcal{H}_2^\perp$. Then

$$f_+(s) = \langle f_+(jt), \frac{1}{2\pi(jt + s^*)}\rangle = \langle f(jt), \frac{1}{2\pi(jt + s^*)}\rangle$$

for $\mathcal{R}e(s) > 0$. Now consider $L \in \mathbb{C}^{n\times n}$ where spec(L) is in the open right half plane and let $g \in \mathcal{H}_2^n$. Then

$$\begin{aligned}
\mathbf{\Pi}_+(sI-L)^{-1}g(s) &= \int_{-\infty}^{\infty} \frac{1}{2\pi(jt+s^*)^*}(jtI-L)^{-1}g(jt)dt \\
&= \frac{-1}{2\pi j}\int_{-j\infty}^{j\infty} \frac{1}{z-s}(zI-L)^{-1}g(z)dz \\
&= (\partial_L g)(s). \qquad (37)
\end{aligned}$$

Let $L = PJP^{-1}$ be a Jordan decomposition with k Jordan blocks of size n_i and corresponding eigenvalues μ_i, and accordingly let

$$P = (V_1, \ldots, V_k) \ \text{ and } \ P^{-1} = \begin{pmatrix} U_1^* \\ \vdots \\ U_k^* \end{pmatrix}.$$

Then

$$(sI-L)^{-1} = \sum_{i=1}^{k} V_i \left\{ \sum_{\ell=1}^{n_i} \frac{1}{(s-\mu_i)^\ell} Z_{n_i}^{\ell-1} \right\} U_i^*, \qquad (38)$$

where Z_m is a $m \times m$ matrix with 1's on the first superdiagonal and zeros elsewhere. If γ_1 is a simple closed curve containing s and μ, and γ_R is a circle of radius R about zero, then

$$\frac{1}{2\pi j}\int_{\gamma_1} \frac{1}{(z-s)}\frac{1}{(z-\mu)^\ell}dz \ = \ \lim_{R\to\infty} \frac{1}{2\pi j}\int_{\gamma_R} \frac{1}{(z-s)}\frac{1}{(z-\mu)^\ell}dz \ = \ 0$$

for $\ell \geq 1$. Thus

$$\begin{aligned}
\frac{1}{2\pi j}\int_{\gamma_1} \frac{1}{(z-s)}\frac{1}{(z-\mu)^\ell} f(z)dz &= \frac{1}{2\pi j}\int_{\gamma_1} \frac{1}{(z-s)}\frac{1}{(z-\mu)^{\ell-1}} \left(\partial_\mu f\right)(z)dz \\
&= \frac{1}{2\pi j}\int_{\gamma_1} \frac{1}{(z-s)}\frac{1}{(z-\mu)} \left(\partial_\mu^{\ell-1} f\right)(z)dz \\
&= \left(\partial_\mu^{\ell} f\right)(s).
\end{aligned}$$

Hence, from (38),

$$\left(\partial_L F\right)(s) = \sum_{i=1}^{k}\sum_{\ell=1}^{n_i} V_i Z_{n_i}^{\ell-1} U_i^* \left(\partial_{\mu_i}^{\ell} F\right)(s). \qquad (39)$$

7 General controller formula

In this section we will generalize the formula of Proposition 3 to the case where $P(s) = e^{-s\tau}P_0(s)$, and $P_0(s) = C(sI-A_0)^{-1}B$ is (a minimal realization) of degree n. We recall the notation of Section 3, namely $A = A_0 + EC$, $E = -RC^*$, with R, S, H defined as in (5), (15), (11).

In order to carry out the derivations of this section, a mild genericity assumption will be needed on $P_0(s)$, namely that the reflections of the zeros of $P_0(s)$ about the imaginary axis are disjoint from the poles of $P_0(s)$. As shown in the lemma below, this is equivalent to the spectrum of H and A being disjoint, which is the assumption we will need in the computations below. A discussion on the validity of the formula when the assumption fails is given in Remark 2.

Lemma 1. Let A, B, C and $P_0(s)$ be defined as above, and let H be defined as in (11). Let $\{\zeta_i : i = 1, \ldots, k\}$ be the zeros of $P_0(s)$ and let $\{\pi_i : i = 1, \ldots, \ell\}$ be the poles of $P_0(s)$ (since $P_0(s)$ is a scalar function these two sets are disjoint). Then the following statements are equivalent:

(a) $\{\zeta_i : i = 1, \ldots, k\} \cap \{-\overline{\pi_i} : i = 1, \ldots, \ell\} = \emptyset$.

(b) (A, B) and (A, E) are controllable.

(c) $\mathrm{spec}(H) \cap \mathrm{spec}(A) = \emptyset$.

Proof. Define $\{\zeta_i^+ : i = 1, \ldots, k^+\}$ and $\{\zeta_i^- : i = 1, \ldots, k^-\}$ to be the set of zeros of $P_0(s)$ in $\mathcal{R}e(s) > 0$ and $\mathcal{R}e(s) < 0$ respectively. Similarly, $\{\pi_i^+ : i = 1, \ldots, \ell^+\}$ and $\{\pi_i^- : i = 1, \ldots, \ell^-\}$ define the set of poles of $P_0(s)$ in the same regions. We now introduce the following subsidiary statements:

(a_1) $\{\zeta_i^- : i = 1, \ldots, k^-\} \cap \{-\overline{\pi_i^+} : i = 1, \ldots, \ell^+\} = \emptyset$.

(a_2) $\{\zeta_i^+ : i = 1, \ldots, k^+\} \cap \{-\overline{\pi_i^-} : i = 1, \ldots, \ell^-\} = \emptyset$.

(b_1) (A, B) is controllable.

(b_2) (A, E) is controllable.

(b_2') $C(-\overline{\alpha} I - A)^{-1} B \neq 0$ for all $\alpha \in \mathrm{spec}(A)$.

It is clear that (a) $\Leftrightarrow$ (a_1) and (a_2), since imaginary axis poles or zeros of $P_0(s)$ can only belong to one of the sets in (a). The lemma will be proved by establishing the following equivalences: (a_1) $\Leftrightarrow$ (b_1); (a_2) $\Leftrightarrow$ (b_2) $\Leftrightarrow$ (b_2'); (b_1), (b_2') $\Leftrightarrow$ (c).

Let $P_0(s) = p(s)/q(s)$ where $p(s)$, $q(s)$ are coprime polynomials and $q(s)$ is monic. Let $p(s)/r(s)$, $q(s)/r(s)$ be normalized coprime factors of $P_0(s)$ over $\mathcal{H}_\infty$ (so $r(s)$ is a monic polynomial of degree n).

(a_1) $\Leftrightarrow$ (b_1). Assume that (A, B) is not controllable. Then for some α ($\mathcal{R}e(\alpha) < 0$), $r(\alpha) = p(\alpha) = 0$. Since $p^* p + q^* q = r^* r$ then $q^*(\alpha) q(\alpha) = 0$. Therefore $q^*(\alpha) = \overline{q(-\overline{\alpha})} = 0$ since $p(s)$ and $q(s)$ are coprime. Thus (a_1) fails. Conversely, if (a_1) fails then for some α

($\mathcal{R}e(\alpha) < 0$), $p(\alpha) = q(-\overline{\alpha}) = 0$. This means that $r^*(\alpha)r(\alpha) = 0$ which implies $r(\alpha) = 0$, since $r(s)$ is Hurwitz. Therefore, $p(s)$ and $r(s)$ have a common factor. Since (C, A_0) is observable, so is (C, A), which means that (A, B) is not controllable.

$(\mathrm{a}_2) \Leftrightarrow (\mathrm{b}_2)$. The proof is the same as for $(\mathrm{a}_1) \Leftrightarrow (\mathrm{b}_1)$ but with p and q interchanged and B and E interchanged.

$(\mathrm{a}_2) \Leftrightarrow (\mathrm{b}_2')$. Suppose (b_2') fails. It follows that there exists an α ($\mathcal{R}e(\alpha) < 0$) such that $r(\alpha) = p(-\overline{\alpha}) = 0$. Since $p^*p + q^*q = r^*r$, it follows that $q^*(\alpha)q(\alpha) = 0$ and hence, since $p(s)$ and $q(s)$ are coprime, that $q^*(\alpha) \neq 0$ and $q(\alpha) = 0$. Thus (a_2) fails. Conversely, if (a_2) fails, then we can write $P_0(s) = \frac{s-\overline{\alpha}}{s+\alpha} P_1(s)$ where $\overline{\alpha}$ is a zero of $p(s)$, $-\alpha$ is a zero of $q(s)$, and hence $P_1(s)$ has degree $n-1$. We now find a normalized coprime factorization $P_1(s) = N_1(s)M_1(s)^{-1}$ over $\mathcal{H}_\infty$. Then $P_0(s) = \left(\frac{s-\overline{\alpha}}{s+\alpha} N_1(s)\right) M_1(s)^{-1}$ is a normalized coprime factorization. Therefore, since the degree of $\left(\frac{s-\overline{\alpha}}{s+\alpha} N_1(s),\ \ M_1(s)\right)$ must be n, $-\alpha$ is in the spectrum of the matrix A. Finally, since normalized coprime factorizations are unique modulo a unitary constant then (b_2') fails.

$(\mathrm{b}_1), (\mathrm{b}_2') \Leftrightarrow (\mathrm{c})$. Let $\alpha \in \mathrm{spec}(A)$. Then $\alpha I + A^*$ is non-singular since A is Hurwitz. Thus

$$
\begin{aligned}
\det(\alpha I - H) &= \det \begin{pmatrix} \alpha I - A & -\lambda^{-1} B B^* \\ \lambda^{-1} C^* C & \alpha I + A^* \end{pmatrix} \\
&= \det\left(\alpha I - A + \lambda^{-2} B B^* (\alpha I + A^*)^{-1} C^* C\right) \det\left(\alpha I + A^*\right). \qquad (40)
\end{aligned}
$$

Clearly

$$
\begin{aligned}
\det(\alpha I - H) \neq 0 \quad &\Rightarrow \quad B^*(\alpha I + A^*)^{-1} C^* \neq 0 \\
&\Leftrightarrow \quad C(-\bar{\alpha} I - A)^{-1} B \neq 0.
\end{aligned}
$$

Thus (c) $\Rightarrow$ (b_2'). Also, if (A, B) is not controllable, then there exists an $\alpha \in \mathrm{spec}(A)$ such that $v^T(\alpha I - A) = 0, v^T B = 0$ for some $v \neq 0$. This implies by (40) that $\alpha \in \mathrm{spec}(H)$. Thus (c) $\Rightarrow$ (b_1). Conversely, let $(\mathrm{b}_1), (\mathrm{b}_2')$ hold. Take $\alpha \in \mathrm{spec}(A)$ and note that $\beta := \lambda^2 B^*(\alpha I + A^*)^{-1} C^* \neq 0$. We now claim that $\alpha I - A + \beta BC$ is non-singular. To see this, note that $C(sI - A)^{-1}B = p(s)/q(s) = p(s)/\det(sI - A)$ (since $q(s)$ is monic) and that $p(s)$ and $\det(sI - A)$ have no common factor (since (A, B) is controllable and (C, A) is observable). Since

$$
\det(1 + \beta C(sI - A)^{-1}B) = \det(sI - A)^{-1} \det(sI - A + \beta BC),
$$

it follows that $\det(sI - A + \beta BC) = \det(sI - A) + \beta p(s)$, and hence that $\det(\alpha I - A + \beta BC) = \beta p(\alpha) \neq 0$. Therefore, $\alpha \notin \mathrm{spec}(H)$ using (40). This proves (c). □

We now follow the steps P1.–P3. outlined in Section 3. Let λ be the largest solution of $\det L(\lambda) = 0$ in $[0, 1]$ and let $w_0 \neq 0$ solve $L(\lambda) w_0 = 0$. Then

$$
v(t) = (I,\ \ 0) e^{Ht} \begin{pmatrix} \lambda^{-1} R \\ I \end{pmatrix} w_0
$$

on $[0,\tau]$. This gives

$$\lambda x(t) = \begin{cases} (C,\ \ 0)e^{Ht}\begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0 & \text{on } [0,\tau], \\ Ce^{A(t-\tau)}(I,\ \ 0)e^{H\tau}\begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0 & \text{on } [\tau,\infty), \end{cases}$$

by (9). Direct calculation shows that

$$\int_0^\tau e^{-st}e^{Ht}dt = -\partial_H(e^{-s\tau})e^{H\tau}$$

and

$$\int_\tau^\infty e^{-st}e^{A(t-\tau)}dt = e^{-s\tau}(sI-A)^{-1}.$$

Thus

$$\lambda\hat{x} = C\left(e^{-s\tau}(sI-A)^{-1}(I,\ \ 0) - (I,\ \ 0)\partial_H(e^{-s\tau})\right)e^{H\tau}\begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0. \tag{41}$$

We now assume that H is diagonalizable, for simplicity, and write $H = \sum_i w_i v_i^* \mu_i$. (In the proof of Theorem 1, we show that the result also holds without the assumption.) Thus, the term in parentheses in (41) can be written as:

$$\sum_i \left((sI-A)^{-1}e^{-s\tau}(I,\ \ 0) - \partial_{\mu_i}(e^{-s\tau})(I,\ \ 0)\right) w_i v_i^*.$$

From (34) we see that

$$\partial_{\mu_i}\left((sI-A)^{-1}e^{-s\tau}\right) = (\mu_i I - A)^{-1}\left(\partial_{\mu_i}(e^{-s\tau})I - (sI-A)^{-1}e^{-s\tau}\right)$$

which gives

$$\lambda\hat{x} = -C\left(\sum_i (\mu_i I - A)\partial_{\mu_i}((sI-A)^{-1}e^{-s\tau}X_i)\right)$$

where

$$X_i \ \ := \ \ (I,\ \ 0)w_i v_i^* e^{H\tau}\begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0.$$

We will now calculate $\hat{u} = \mathbf{\Pi}_+ F^* \lambda\hat{x}$. Then

$$\begin{aligned} \hat{u}_1 \ &= \ -\mathbf{\Pi}_+ e^{s\tau}B^*(sI+A^*)^{-1}C^*C\sum_i(\mu_i I - A)\partial_{\mu_i}\left((sI-A)^{-1}e^{-s\tau}X_i\right) \\ &= \ -\sum_i \partial_{\mu_i}\left(\mathbf{\Pi}_+ B^*(sI+A^*)^{-1}C^*C(sI-A)^{-1}(\mu_i I - A)X_i\right) \end{aligned}$$

by Proposition 2. (Note that, with the aid of a partial fraction expansion, the proposition can be applied elementwise.) Using (15) we have

$$(sI + A^*)^{-1}C^*C(sI - A)^{-1} = -S(sI - A)^{-1} + (sI + A^*)^{-1}S.$$

Thus

$$\begin{aligned}\hat{u}_1 &= \sum_i \partial_{\mu_i}(B^*S(sI - A)^{-1}(\mu_i I - A)X_i) \\ &= -\sum_i B^*S(sI - A)^{-1}X_i \\ &= -B^*S(sI - A)^{-1}v(\tau). \end{aligned} \tag{42}$$

Now define

$$g_i(s) := C(\mu_i I - A)(sI - A)^{-1}e^{-s\tau}X_i.$$

Then

$$\begin{aligned}-\hat{u}_2 &= \mathbf{\Pi}_+(1 - E^*(sI + A^*)^{-1}C^*)\sum_i \partial_{\mu_i}(g_i(s)) \\ &= \sum_i \partial_{\mu_i}\left(g_i(s) - E^*\mathbf{\Pi}_+(sI + A^*)^{-1}C^*g_i(s)\right) \\ &= \sum_i \partial_{\mu_i}\left(g_i(s) - E^*\partial_{-A^*}C^*g_i(s)\right) \\ &= \sum_i \left(\partial_{\mu_i}g_i(s) - E^*(\mu_i I + A^*)^{-1}(\partial_{\mu_i} - \partial_{-A^*})(C^*g_i(s))\right) \\ &= \sum_i \left((1 - E^*(\mu_i I + A^*)^{-1}C^*)\partial_{\mu_i}(g_i(s))\right) \\ &\quad + E^*\partial_{-A^*}\left(\sum_i(\mu_i I + A^*)^{-1}C^*g_i(s)\right). \end{aligned} \tag{43}$$

Proposition 4. The following identity holds:

$$\sum_i(\mu_i I + A^*)^{-1}C^*g_i(s) = 0.$$

In the proof of the proposition we will make use of the following fact.

Lemma 2. Let $X, Y \in \mathbb{C}^{n\times n}$ be such that

$$\operatorname{rank}(X) + \operatorname{rank}(Y) = n$$

and $X + Y$ be non-singular. Then

$$Y(X + Y)^{-1}X = 0.$$

Proof. We assume that $0 < r := \operatorname{rank}(X) < n$, otherwise the statement is trivial. We now write $X = X_1X_2$ and $Y = Y_1Y_2$, where $X_1 \in \mathbb{C}^{n\times r}$, $X_2 \in \mathbb{C}^{r\times n}$, $Y_1 \in \mathbb{C}^{n\times(n-r)}$ and $Y_2 \in \mathbb{C}^{(n-r)\times n}$. Thus

$$\begin{aligned}
Y(X+Y)^{-1}X &= Y_1Y_2\left((X_1,\ Y_1)\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}\right)^{-1} X_1X_2 \\
&= Y_1Y_2\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}^{-1}(X_1,\ Y_1)^{-1}X_1X_2 \\
&= Y_1\left(0_{(n-r)\times r},\ I_{(n-r)\times(n-r)}\right)\begin{pmatrix} I_{r\times r} \\ 0_{(n-r)\times r} \end{pmatrix} X_2 \\
&= 0.\square
\end{aligned}$$

Proof of Proposition 4. It is sufficient to show that

$$\sum_i U_iTV_i = 0 \tag{44}$$

for an arbitrary constant matrix T, where

$$U_i = (\mu_i I + A^*)^{-1},$$

$$V_i = (\mu_i I - A)(I,\ \ 0)w_iv_i^*e^{H\tau}\begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0.$$

In fact it is sufficient to prove (44) for an arbitrary T of rank one, since the general fact then follows by linear superposition. We first take $T = z_jy^*$, where z_j is an eigenvector of A^*, i.e., $A^*z_j = \bar{\gamma}_jz_j$, and y^* is an arbitrary row vector. Observe that

$$e^{H\tau}\begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0 = \begin{pmatrix} v(\tau) \\ w(\tau) \end{pmatrix} = \begin{pmatrix} I \\ \lambda^{-1}S \end{pmatrix} v(\tau) \tag{45}$$

from (12)–(14). Thus

$$\begin{aligned}
\sum_i U_iTV_i &= z_jy^*\left(\sum_i \frac{1}{\mu_i+\bar{\gamma}_j}(\mu_iI - A)(I,\ \ 0)w_iv_i^*\right)\begin{pmatrix} I \\ \lambda^{-1}S \end{pmatrix} v(\tau) \\
&= z_jy^*\left((I,\ \ 0)\sum_i \frac{\mu_i}{\mu_i+\bar{\gamma}_j}w_iv_i^* - (A,\ \ 0)\sum_i \frac{1}{\mu_i+\bar{\gamma}_j}w_iv_i^*\right)\begin{pmatrix} I \\ \lambda^{-1}S \end{pmatrix} v(\tau) \\
&= z_jy^*\left((I,\ \ 0)H(H+\bar{\gamma}_jI)^{-1} - (A,\ \ 0)(H+\bar{\gamma}_jI)^{-1}\right)\begin{pmatrix} I \\ \lambda^{-1}S \end{pmatrix} v(\tau) \\
&= z_jy^*\left((0,\ \ \lambda^{-1}BB^*)(H+\bar{\gamma}_jI)^{-1}\right)\begin{pmatrix} I \\ \lambda^{-1}S \end{pmatrix} v(\tau).
\end{aligned} \tag{46}$$

Since, by assumption, $-\bar{\gamma}_j$ is distinct from the eigenvalues of H, we have

$$
\begin{aligned}
(H+\bar{\gamma}_j I)^{-1} &= \begin{pmatrix} A+\bar{\gamma}_j I & \lambda^{-1}BB^* \\ -\lambda^{-1}C^*C & -A^*+\bar{\gamma}_j I \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \star & \star \\ \Delta^{-1}\lambda^{-1}C^*C(A+\bar{\gamma}_j I)^{-1} & \Delta^{-1} \end{pmatrix},
\end{aligned}
$$

where $\Delta = (-A^*+\bar{\gamma}_j I) + \lambda^{-2}C^*C(A+\bar{\gamma}_j I)^{-1}BB^*$. Note that Δ is non-singular since $\det(H+\bar{\gamma}_j I) \neq 0$. Thus

$$
\begin{aligned}
&(0,\ \ \lambda^{-1}BB^*)(H+\bar{\gamma}_j I)^{-1}\begin{pmatrix} I \\ \lambda^{-1}S \end{pmatrix} \\
&= \lambda^{-2}BB^*\Delta^{-1}\left(C^*C + S(A+\bar{\gamma}_j I)\right)(A+\bar{\gamma}_j I)^{-1} \\
&= \lambda^{-2}BB^*\Delta^{-1}(-A^*+\bar{\gamma}_j I)S(A+\bar{\gamma}_j I)^{-1}.
\end{aligned}
$$

We now claim that $B^*\Delta^{-1}(-A^*+\bar{\gamma}_j I)=0$. It suffices to show that

$$
\lambda^{-2}C^*C(A+\bar{\gamma}_j I)^{-1}BB^*\Delta^{-1}(-A^*+\bar{\gamma}_j I) = 0, \tag{47}
$$

since $C(A+\overline{\gamma}_j I)^{-1}B \neq 0$ by Lemma 1 and C^* is a non-zero column vector. To show (47) let $Y = \lambda^{-2}C^*C(A+\bar{\gamma}_j I)^{-1}BB^*$ and $X = (-A^*+\bar{\gamma}_j I)$. Observe that $\Delta = X+Y$ is non-singular and $\mathrm{rank}(X)+\mathrm{rank}(Y) = \mathrm{rank}(\Delta)$. This follows since X is rank deficient and Y has rank at most one (B, C are vectors). Thus the conditions of Lemma 2 are satisfied and so $Y(X+Y)^{-1}X = 0$. This shows that (47) and hence (46) is zero. Since y^* is arbitrary in (46) and z_j is any eigenvector of A^* this proves that (44) holds for any $n\times n$ matrix T having range the span of the eigenvectors of A^*. This completes the proof if A is diagonalizable.

The same conclusion holds in the non-diagonalizable case as well. To see this consider a sequence of matrices A_k $(k=1,2,\ldots)$ tending to A, each of which is diagonalizable, and let $P_k(s)$ $(k=1,2,\ldots)$ denote the corresponding perturbed transfer functions. It can be easily seen that $P_k(s)\to P(s)$ in the gap metric (this follows because the graph symbol F depends continuously on A; see [16] for the definition and expressions for the gap metric). Further, due to the metric property of the gap, see e.g. [16], the optimal robustness radius $b_{\mathrm{opt}}(P)$ depends continuously on $P(s)$. Since the largest root of $\det(L(\lambda)=0$ is $\lambda=\sqrt{1-b_{\mathrm{opt}}^2(P)}$, λ depends continuously on A as well. Clearly R and S depend continuously on A, hence so does $L(\lambda)$. Note that $L(\lambda)$ remains singular as A varies. Now suppose that $w_{0,k}$ are unit vectors satisfying $L_k(\lambda_k)w_{0,k}=0$, $k=1,2,\ldots$ (where L_k is defined via (16) for the case of each $P_k(s)$ respectively). Without loss of generality we can take $w_{0,k}$ to be a convergent sequence (otherwise we select an appropriate subsequence), and let the limit point be w_0. Then $L(\lambda)w_0=0$, so this limit point can be taken as the definition of w_0. Since H is assumed diagonalizable and since it depends continuously on A, we can assume that, in each case, H_k is also diagonalizable (possibly for k sufficiently large). Further, we can select eigenvectors for H_k tending to corresponding ones for H. Compiling the above, it is seen that $\sum_i U_i T V_i$ depends continuously on the perturbations of A. By the earlier part of the proof this expression is identically zero in the case of each A_k. Thus in the limit it is also zero for the given $P(s)$ and A. This completes the proof. $\square$

Proposition 4 and equation (43) show that

$$\begin{aligned} \hat{u}_2 &= -\sum_i \left(1 - E^*(\mu_i I + A^*)^{-1} C^*\right) \partial_{\mu_i}(g_i(s)) && (48) \\ &= -\sum_i M^*(\mu_i) \partial_{\mu_i}(g_i(s)). && (49) \end{aligned}$$

We now expand $\partial_{\mu_i}(g_i(s))$ using (35) so that the ∂-operator acts only on the infinite-dimensional element $e^{-s\tau}$:

$$\begin{aligned} \partial_{\mu_i}(g_i(s)) &= C(\mu_i I - A)\partial_{\mu_i}\left((sI - A)^{-1} e^{-s\tau} X_i\right) \\ &= C(\mu_i I - A)(sI - A)^{-1}\left(\partial_{\mu_i}(e^{-s\tau})I - (\mu_i I - A)^{-1} e^{-\mu_i \tau}\right) X_i \\ &= C(sI - A)^{-1}(\mu_i I - A) X_i \partial_{\mu_i}(e^{-s\tau}) - C(sI - A)^{-1} X_i e^{-\mu_i \tau}. \end{aligned}$$

Using the following equalities:

$$\begin{aligned} \sum_i \mu_i M^*(\mu_i)\partial_{\mu_i}(e^{-s\tau}) X_i &= (I,\ \ 0) H M^*(H) \partial_H(e^{-s\tau}) e^{H\tau} \begin{pmatrix} \lambda^{-1} R \\ I \end{pmatrix} w_0, \\ \sum_i A M^*(\mu_i)\partial_{\mu_i}(e^{-s\tau}) X_i &= (A,\ \ 0) M^*(H) \partial_H(e^{-s\tau}) e^{H\tau} \begin{pmatrix} \lambda^{-1} R \\ I \end{pmatrix} w_0, \\ \sum_i M^*(\mu_i) X_i e^{-\mu_i \tau} &= (I,\ \ 0) M^*(H) \begin{pmatrix} \lambda^{-1} R \\ I \end{pmatrix} w_0 \end{aligned}$$

and (49) we obtain

$$\begin{aligned} \hat{u}_2 &= C(sI - A)^{-1}(I,\ \ 0) M^*(H) \begin{pmatrix} \lambda^{-1} R \\ I \end{pmatrix} w_0 \\ &\quad - C(sI - A)^{-1} B (0,\ \ \lambda^{-1} B^*) M^*(H) \partial_H(e^{-s\tau}) \begin{pmatrix} I \\ \lambda^{-1} S \end{pmatrix} v(\tau). \qquad (50) \end{aligned}$$

We are now ready to state the form of the optimal compensator.

Theorem 1. Let $P(s) = e^{-s\tau} P_0(s)$ be a single-input/single-output plant with $P_0(s) = C(sI - A_0)^{-1} B$ a minimal realization and $\tau > 0$. Suppose that $P_0(s)$ satisfies condition (a) of Lemma 1. Then the optimal robustness radius for gap ball uncertainty is given by the formula

$$b_{\mathrm{opt}}(P) = (1 - \lambda_{\max}^2)^{1/2},$$

where $\lambda_{\max}$ is the largest solution of the transcendental equation $\det L(\lambda) = 0$ on $[0, 1]$ (which is defined by (16)). The transfer function of the optimal controller is given by:

$$K_{\mathrm{opt}}(s) = \frac{B^* S (sI - A)^{-1} B_1}{C(sI - A)^{-1} B_2 - C(sI - A)^{-1} B h(\tau, s)}, \qquad (51)$$

where $L(\lambda)w_0 = 0$,

$$
\begin{aligned}
w_1 &= \begin{pmatrix} \lambda^{-1}R \\ I \end{pmatrix} w_0, \\
B_1 &= (I,\ 0)e^{H\tau}w_1, \\
B_2 &= (I,\ 0)M^*(H)w_1, \\
h(\tau, s) &= (0,\ \lambda^{-1}B^*)M^*(H)\partial_H(e^{-s\tau})e^{H\tau}w_1,
\end{aligned}
$$

with R, S, H defined as in (5), (15), (11), and $\lambda = \lambda_{\max}$ throughout.

Before proving the theorem we will establish the following result.

Lemma 3. Consider any (non-zero) strictly proper rational function $P_0(s)$. Then $b_{\text{opt}}(e^{-s\tau}P_0(s))$ is a *strictly* decreasing, continuous function of τ.

Proof. Let $P_0 = N_0/M_0$ be a normalized coprime factorization over $\mathcal{H}_\infty$. Then $(e^{-s\tau}N,\ M)$ varies continuously with τ in the $\mathcal{H}_\infty$ norm. Thus $\left\| \mathbf{\Pi}_- \begin{pmatrix} -e^{s\tau}N_0^* \\ M_0^* \end{pmatrix} \Big|_{\mathcal{H}_2} \right\|$ varies continuously with τ, and the same is true for $b_{\text{opt}}(e^{-s\tau}P_0(s))$ using equation (2). Thus, the statement of the lemma requires proof to the following fact:

$$
\left\| \mathbf{\Pi}_- \begin{pmatrix} -e^{s\tau_2}N_0^* \\ M_0^* \end{pmatrix} \Big|_{\mathcal{H}_2} \right\| > \left\| \mathbf{\Pi}_- \begin{pmatrix} -e^{s\tau_1}N_0^* \\ M_0^* \end{pmatrix} \Big|_{\mathcal{H}_2} \right\| \tag{52}
$$

for $\tau_2 > \tau_1$ (cf. equation (2)). We will prove a slightly more general fact than (52). Let $F \in \mathcal{H}_\infty^{1\times 2}$ be continuous (and non-constant) on the compactified right half plane, co-inner and right invertible over $\mathcal{H}_\infty$. Let $\Phi(s) \in \mathcal{H}_\infty^{2\times 2}$ be unitary ($\Phi^*\Phi = \Phi\Phi^* = I$) with $\det\Phi(s)$ not a constant. We will show that $\|\mathbf{\Pi}_-\Phi^*F^*|_{\mathcal{H}_2}\| > \|\mathbf{\Pi}_-F^*|_{\mathcal{H}_2}\|$. First note that, for any $\hat{x} \in \mathcal{H}_2$,

$$
\begin{aligned}
\|\mathbf{\Pi}_-\Phi^*F^*\hat{x}\|_2^2 &= \|\mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_-F^*\hat{x} + \mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_+F^*\hat{x}\|_2^2 \\
&= \|\mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_-F^*\hat{x}\|_2^2 + \|\mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_+F^*\hat{x}\|_2^2 \\
&= \|\mathbf{\Pi}_-F^*\hat{x}\|_2^2 + \|\mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_+F^*\hat{x}\|_2^2.
\end{aligned} \tag{53}
$$

The second equality follows since the inner product

$$\langle \mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_-F^*\hat{x}, \mathbf{\Pi}_-\Phi^*\mathbf{\Pi}_+F^*\hat{x}\rangle = 0.$$

We now take $\hat{x}$ to be the maximal vector for $\mathbf{\Pi}_-F^*|_{\mathcal{H}_2}$ and write $\hat{v} = \mathbf{\Pi}_+F^*\hat{x}$. Let us suppose that $\mathbf{\Pi}_-\Phi^*\hat{v} = 0$, i.e., the second term in (53) vanishes. This means that $\hat{v} = \Phi\hat{w}$ for some $\hat{w} \in \mathcal{H}_2$, so $\hat{v}$ has a non-trivial inner factor. But this contradicts the fact that $\hat{x}$ is outer and $\hat{v}/\hat{x}$ is left invertible over $\mathcal{H}_\infty$ for the assumed conditions on F [18]. This establishes the required result. $\square$.

Proof of Theorem 1. Under the assumption that $H = H_\lambda$ is diagonalizable the result follows from Proposition 1, equation (2) and steps P1.–P3. It follows that $K_{\text{opt}}(s) = -\hat{u}_1/\hat{u}_2$, where $\hat{u}_1$ and $\hat{u}_2$ are given in (42) and (50). To show that the formula is also valid when H is not diagonalizable we use a perturbation argument. We break the reasoning into two steps.

STEP 1. We claim that H_λ has distinct eigenvalues for almost all λ if the conditions of Lemma 1 hold. To see this we write $C(sI - A)^{-1}B = p(s)/r(s)$ where $p(s)$ and $r(s)$ are coprime and $r(s)$ has degree n (this is possible since (A, B) is controllable). We note that p^*p and r^*r are coprime (this follows as in the proof of $(a_1) \Leftrightarrow (b_1)$ in Lemma 1). Thus $\lambda^2 r^*r - p^*p$ has distinct roots for almost all λ (this follows by taking derivatives with respect to s, eliminating λ^2 to give a non-zero polynomial in s, and verifying that these root locations can only be achieved by finitely many λ). The claim now follows by noting that the eigenvalues of H_λ coincide with the zeros of the expression

$$1 - \lambda^{-2}C(sI - A)^{-1}BB^*(-sI - A^*)^{-1}C^*$$

which are the same as the zeros of $\lambda^2 r^*r - p^*p$.

STEP 2. Suppose for some τ and $\lambda = \lambda_{\max}(\tau)$ that H_λ has repeated roots. Take a sequence τ_i $(i = 1, 2, \ldots)$ with $\tau_i < \tau$ and $\tau_i \to \tau$. By Lemma 3, $\lambda_i = \lambda_{\max}(\tau_i) < \lambda$ and $\lambda_i \to \lambda$. Using STEP 1 we can assume, without loss of generality, that H_{λ_i} has distinct eigenvalues. Now suppose that w_i are unit vectors satisfying $L(\lambda_i)w_i$. Without loss of generality we can take w_i to be a convergent sequence (otherwise we select an appropriate subsequence), and let the limit point be w_0. Then $L(\lambda)w_0 = 0$. Let $\hat{x}$ be defined as in equation (41) with $H = H_\lambda$, and let $\hat{x}_i$ be defined by the same equation but with w_0 replace by w_i, τ replace by τ_i and H by H_{λ_i}. We note that $\hat{x}_i \to \hat{x}$. Define $\hat{u} = \mathbf{\Pi}_+ F^*\lambda\hat{x}$ and let $(\hat{u})_i$ be defined similarly but with $\hat{x}$ replaced by $\hat{x}_i$ and τ replace by τ_i in F. Then $(\hat{u})_i \to \hat{u}$. The proof is now completed by noting that, with the first and second components of $\hat{u}$ defined through equations (42) and (50), we also have $(\hat{u})_i \to \hat{u}$ by virtue of the fact that $M^*(H)$ and $\partial_H(e^{-s\tau})$ are continuous functions of H. Thus (42) and (50) are correct expressions even when H is not diagonalizable. $\square$

Remark 1. The form of $K_{\mathrm{opt}}(s)$ given in (51) is the ratio of two $\mathcal{H}_2$ functions. Consequently, it is not a coprime factorization over $\mathcal{H}_\infty$ since there will be a common zero at infinity. In fact it was proved in [18] that $(s+1)\hat{x}$ is a unit in $\mathcal{H}_\infty$ and that $\hat{u}/\hat{x}$ belongs to $\mathcal{H}_\infty$ and is left invertible. Thus (51) becomes a coprime factorization over $\mathcal{H}_\infty$ on multiplication of the numerator and denominator by $(s+1)$.

Remark 2. (On the failure of conditions of Lemma 1.) If the poles and zeros of $P_0(s)$ do not satisfy the genericity assumption (a) of Lemma 1 then $M^*(H)$ may not be defined. Examples where such a situation arises are: $P_0(s) = (s-1)/s(s+1)$ and $P_0(s) = (s+1)/s(s-1)$. However, it seems to be the case in such situations that $(I,\ 0)M^*(H)w_1$, as well as the second term in (50), is bounded in the limit as $P_0(s)$ is approached. If a well-defined limit can be proved to exist then the controller formula of Theorem 1 would also be valid in this sense when Lemma 1 fails, by virtue of known results about the continuity of $\hat{u}/\hat{x}$ (see [24]).

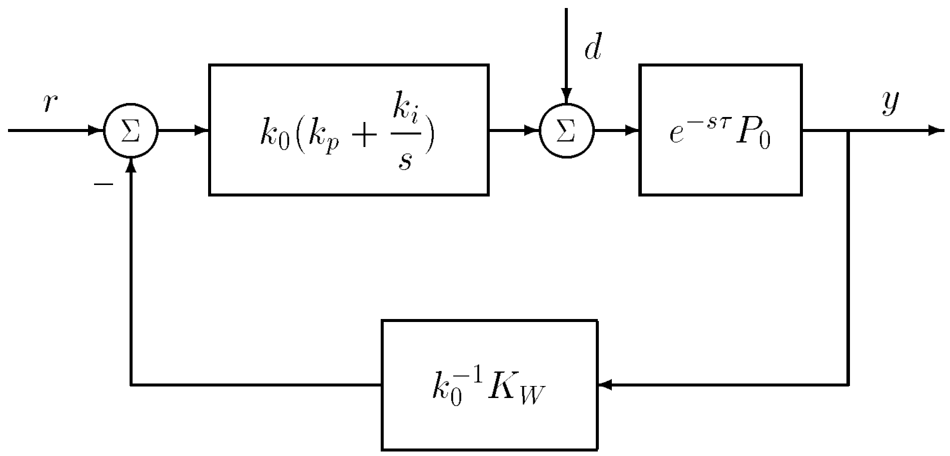

Figure 2: Scheme for optimal robustness controller.

Remark 3. It should be observed that the infinite dimensionality of $K_{\rm opt}(s)$ is confined to the term $h(\tau, s)$, and more specifically to $\partial_H(e^{-s\tau})$. It is possible to approximate $\partial_H(e^{-s\tau})$ uniformly in the $\mathcal{H}_\infty$ norm, which then implies that the corresponding finite dimensional controller converges to $K_{\rm opt}(s)$ in the gap metric. (This follows since the corresponding coprime factors then converge in the $\mathcal{H}_\infty$ norm, cf. Remark 1.) To approximate $\partial_H(e^{-s\tau})$ uniformly, it suffices to find a sequence of functions $f_k(s)$, with uniformly bounded $\mathcal{H}_\infty$ norm, such that $e^{-s\tau}$ is approximated uniformly on any finite interval of the imaginary axis. However, if H has an eigenvalue of multiplicity r on the imaginary axis then it is also necessary to have the first r derivatives of $f_k(s)$ tending to those of $e^{-s\tau}$ at that point. Possible choices satisfying these conditions are: $f_k(s) = 1/(1+s\tau/k)^k$, $f_k(s) = (1-0.5s\tau/k)^k/(1+0.5s\tau/k)^k$ or general Padé approximants. An alternative approach to finite dimensional controller design for the $\mathcal{H}_\infty$ design problem of this paper has been given in [18] which explores approximation of the time-domain vector $x(t)$.

8 Design example

In this section we will use the formula in Theorem 1 to design a controller to achieve disturbance attenuation and robust set-point following for a delay system. We will take a plant with two lightly damped poles and a non-trivial delay. This type of system is a suitable nominal model in a number of process control applications. We will show that our design has significantly better performance than a Smith predictor.

The plant we will consider is:

$$P(s) = e^{-1.5s}\frac{1}{s^2 + 0.2s + 1}.$$

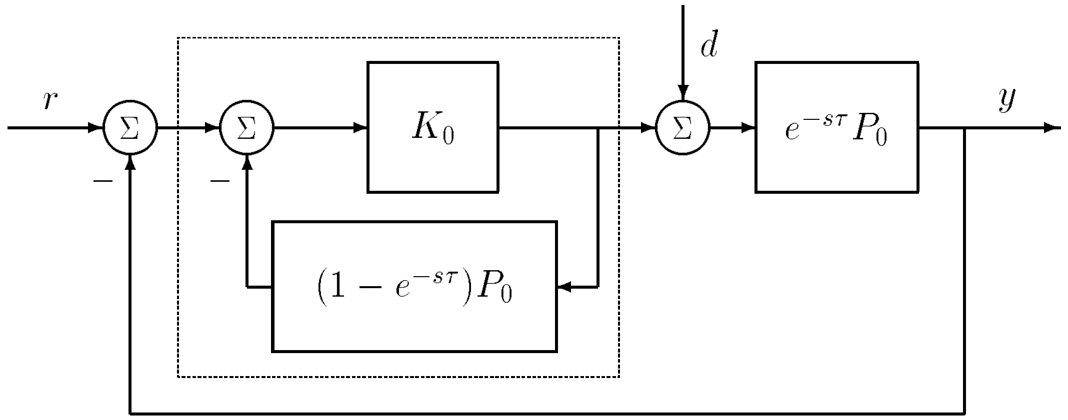

Figure 3: Smith predictor control system.

We will demand increased damping of the oscillations in response to disturbances at the plant input, together with robust asymptotic tracking of constant reference inputs. To achieve these goals, we follow the H_∞ loop shaping (optimal robustness) design procedure of Glover-McFarlane [20]. We select an appropriate weighting function $W(s)$ and compute the optimal controller $K_W(s)$ for $W(s)P(s)$. The final controller for $P(s)$ is then given by $K(s) = K_W(s)W(s)$.

Since an integral term is required in the controller for robust tracking, a proportional-plus-integral weighting function $W(s) = k_p + k_i/s$ was selected. The values $k_p = 1.4$ and $k_i = 0.4$ were found to give an acceptable loop shape and a sufficiently large robustness margin of $b_{\rm opt} = 0.3296$. (Increasing k_p and k_i gives larger loop gain but causes $b_{\rm opt}$ to decrease.) The final $K(s)$ obtained is stable (except for the pole at $s = 0$) and is non-minimum phase as evidenced by the lagging phase characteristic between 0.3 and 1.0 rad/sec (see Fig. 4). A suitable implementation of the controller is shown in Fig. 2, where $k_0 = K_W(0)$. (Note that the specifications are also met with $K(s)$ in the forward path and unity negative feedback. However, the arrangement of Fig. 2 ensures that the response from r to y is not adversely affected by the non-minimum phaseness of $K_W(s)$. Moreover, any admissible command response—a stable system with at least second order roll-off in series with a time delay of τ seconds—can be achieved in this set-up using an additional stable filter outside the loop. In particular, the command response could be made the same as the Smith predictor below.)

The idea of the Smith predictor [26] is that a pre-compensator K_0 is designed for P_0 to give a desired command response from r to y in the delay-free situation. The pre-compensator

$$K_s = \frac{K_0}{1 + (1 - e^{-s\tau})P_0 K_0}$$

stabilizes the plant $e^{-s\tau}P_0$, if P_0 is stable, and gives the same command response but with a delay of τ seconds (see Fig. 3). For this example, K_0 was designed using the H_∞

loop shaping (optimal robustness) procedure using the same weighting function as above: $W(s) = 1.4 + 0.4/s$. This gives:

$$K_0(s) = \frac{(2.302s^2 + 0.470s + 0.612)(s + 0.286)}{s(s^2 + 2.693s + 0.719)}.$$

The Bode plots for both compensators K and K_s are shown for comparison in Fig. 4. The Nyquist plots of the return ratio transfer functions PK and PK_s are shown in Fig. 5. Although there are some broad similarities in the forms of these plots, it should be pointed out that a basic property of the Smith predictor is that open loop poles of P_0 are cancelled by the zeros of the compensator K_s. Thus, the open loop lightly damped poles in P are not shifted by feedback in Fig. 3, in contrast to the situation in Fig. 2. The adverse effects of this cancellation become evident if the response at the plant output y is compared to step disturbances d at the plant input (see Fig. 6). We remark that the bad properties of the Smith predictor in case the plant has poles near the imaginary axis, have already been noted in the literature, and this has led to modifications of the scheme being proposed [30].

Finally, a comparison of the command response to step inputs at r is shown in Fig. 7. Note that in Figs. 4-7 the solid line denotes the optimal robustness compensator and the broken line the Smith predictor.

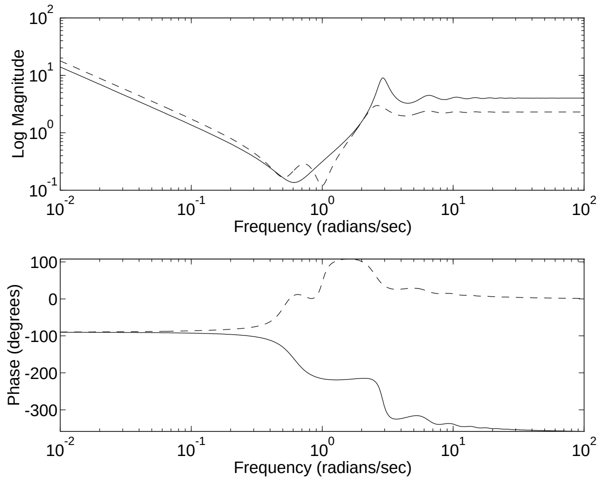

Figure 4: Bode plots of optimal robustness and Smith predictor compensators: K and K_s.

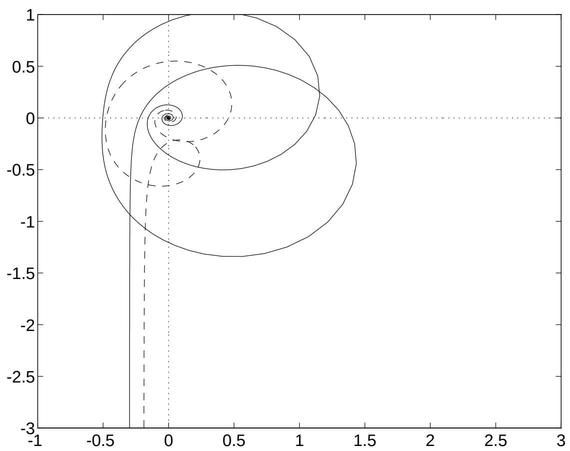

Figure 5: Nyquist diagrams of return ratio transfer functions $P(s)K(s)$ and $P(s)K_s(s)$.

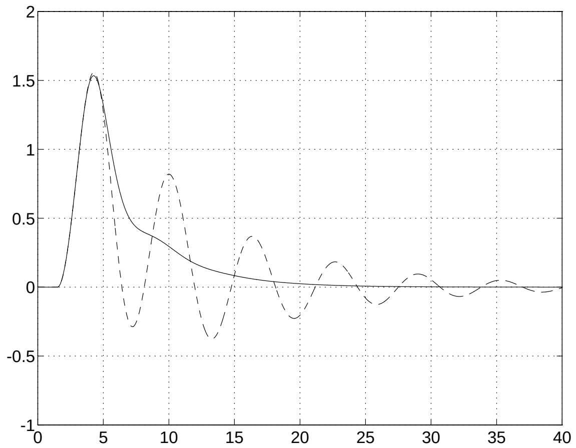

Figure 6: Response of y to a unit step at d with optimal robustness and Smith predictor compensators.

References

[1] H. Bercovici, C. Foias and A. Tannenbaum, "On skew Toeplitz operators," *Operator Theory: Advances and Applications*, **32**, 21-43, 1988.

[2] L. de Branges, "Some Hilbert spaces of analytic functions I," *Trans. Amer. Math. Soc.*, **106**, 445–468, 1963.

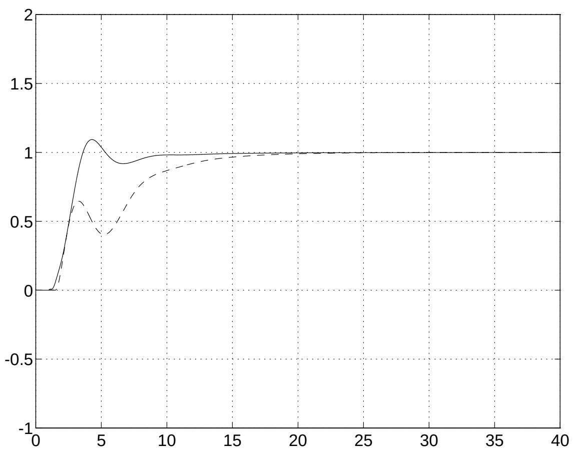

Figure 7: Response of y to a unit step at r with optimal robustness and Smith predictor compensators.

[3] L. de Branges and J. Rovnyak, "Canonical models in quantum scattering theory," in *Perturbation Theory and its Applications in Quantum Mechanics*, edited by C.H. Wilcox, John Wiley, 1966.

[4] S. Buddie, T.T. Georgiou, Ü. Özgüner and M.C. Smith, "Flexible Structure Experiments at JPL and WPAFB", *International Journal of Control*, **58(1)**, 1-19, 1993.

[5] R.F. Curtain, "$\mathcal{H}_\infty$ control for distributed parameter systems: a survey," Proc. of the 29th IEEE Conference on Decision and Control, Honolulu, Hawaii, December 1990, pp. 22–26.

[6] R.F. Curtain, "A synthesis of time and frequency domain methods for the control of infinite-dimensional systems: a system theoretic approach," *SIAM Frontiers in Applied Mathematics*, to appear.

[7] H. Dym, "J-Contractive matrix functions, reproducing kernel Hilbert spaces and interpolation," Regional conference series in mathematics/Conference Board of the Mathematical Sciences, **71**, 1989.

[8] H. Dym, "Shifts, realizations and interpolation, redux," *Oper. Theory: Adv. Appl.,* to appear.

[9] H. Dym, "On the zeros of some continuous analogues of matrix orthogonal polynomials and a related extention problem with negative squares," *Comm. Pure Appl. Math.*, in press.

[10] D. S. Flamm, "Control of Delay Systems for Minimax Sensitivity," Massachusetts Institute of Technology, report LIDS-TH-1560, 1986.

[11] D. S. Flamm, "Outer factor 'absorption' for $\mathcal{H}_\infty$ control problems," *Int. J. Robust and Nonlinear Control*, to appear.

[12] D.S. Flamm and H. Yang, "$\mathcal{H}_\infty$ optimal mixed sensitivity for general distributed systems," Proc. of the 29th IEEE Conference on Decision and Control, Honolulu, Hawaii, December 1990, pp. 134–139.

[13] C. Foias, A. Tannenbaum and G. Zames, "Weighted sensitivity minimization for delay systems," *IEEE Trans. on Automatic Control*, **AC-31(8)**, 763-766, 1986.

[14] C. Foias, A. Tannenbaum and G. Zames, "Some explicit formulae for the singular values of a certain Hankel operators with factorizable symbol," *SIAM J. Math. Analysis*, **19**, 1081–1091, 1988.

[15] T.T. Georgiou and M.C. Smith, Robust control of feedback systems with combined plant and controller uncertainty, Proceedings of the 1990 American Control Conference, pp. 2009–2013, May 1990.

[16] T.T. Georgiou and M.C. Smith, "Optimal robustness in the gap metric," *IEEE Trans. on Automat. Control*, **35**, 673–686, 1990.

[17] T.T. Georgiou and M.C. Smith, "Robust stabilization in the gap metric: controller design for distributed plants," *IEEE Trans. on Automat. Control*, **37**, 1133–1143, 1992.

[18] T.T. Georgiou and M.C. Smith, "Topological approaches to robustness," in *Lecture Notes in Control and Information Sciences*, Springer-Verlag, **185**, 222-241, 1993.

[19] D.G. Meyer and G.F. Franklin, "A connection between normalized coprime factorizations and linear quadratic regulator theory," *IEEE Trans. Automat. Contr.*, **AC-32**: 227-228, 1987.

[20] D.C. McFarlane and K. Glover, "Robust controller design using normalized coprime factor plant descriptions," *Lecture Notes in Control and Information Sciences*, Springer-Verlag, **138**, 1989.

[21] H. Özbay and A. Tannenbaum, "A skew Toeplitz approach to the $\mathcal{H}_\infty$ optimal control of multivariable distributed systems," *SIAM J. Control and Optimization*, **28**, 653-670, 1990.

[22] H. Özbay, M.C. Smith and A. Tannenbaum, "Mixed-sensitivity optimization for a class of unstable infinite-dimensional systems," *Linear Algebra and its Applications*, **178**, 43–83, 1993.

[23] J.R. Partington and K. Glover, "Robust stabilization of delay systems by approximation of coprime factors," *Systems & Control Letters*, **14**, 325–331, 1990.

[24] V.V. Peller, "Hankel operators and continuity properties of best approximation operators," Steklov Institute of Mathematics, preprint.

[25] M. Schechter, "Principles of functional analysis," Academic Press, 1971.

[26] O.J.M. Smith, "Closer control of loops with dead time," *Chem. Eng. Progress*, **53**, 217–219, 1957.

[27] G. Tadmor, "An interpolation problem associated with $\mathcal{H}_\infty$-optimal design in systems with distributed input lags," *Systems & Control Letters*, **8**, 313-319, 1987.

[28] O. Toker and H. Özbay, "Suboptimal robustness in the gap metric for MIMO delay systems," The Ohio State University, Dept. of Electrical Engineering, submitted to the 1994 ACC, September 1993.

[29] M. Vidyasagar, "Normalized coprime factorizations for non strictly proper systems," *IEEE Trans. on Automatic Contr.*, **33**, 300-301, 1988.

[30] K. Watanabe, Y. Ishiyama and M. Ito, "Modified Smith predictor control for multivariable systems with delays and unmeasurable step disturbances," *Int. J. Control*, **37**, 959–973, 1983.

[31] G. Zames and S.K. Mitter, "A note on essential spectrum and norms of mixed Hankel-Toeplitz operators," *Systems and Control Letters*, **10**, 159–165, 1988.

[32] K. Zhou and P. P. Khargonekar, "On the weighted sensitivity minimization problem for delay systems," *Systems & Control Letters*, **8**, 307-312, 1987.