

March 1986

LIDS-P-1546

H^∞ Weighted Sensitivity Minimization for
Systems with Commensurate Input Lags

by

Gilead Tadmor¹

Abstract

Optimal stabilizing feedback which minimizes the H^∞ weighted sensitivity is derived for systems with multiple input lags. The solution of an associated operator interpolation problem in H^2 , occupies most of this note. Reformulation of the problem as a maximal eigenvalue/eigenfunction problem in the time domain, is a key step. The main result is a characterization of these eigenvalue and eigenfunction.

Key Words. Weighted Sensitivity, H^p spaces, Interpolating functions, Maximal eigenvalues, Time domain analysis.

¹Dr. Gilead Tadmor, Massachusetts Institute of Technology, Laboratory for Information and Decision Systems, Room 35-312, 77 Massachusetts Ave., Cambridge, MA 02139 (617)253-6173.

1. Introduction

The H^∞ approach transforms a variety of control problems into problems of interpolation of operators on H^2 . (For overview of that approach and for extensive lists of references, see Francis and Dyle [8] and Helton [11]. The operator interpolation problem and the more general problem of Hankel extensions are discussed in Adamjan, Arov and Krein [1] and in Sarason [12].) The presence of input lags renders the associated interpolation problems infinite dimensional, hence remarkably more difficult than in the ordinary case. This happens already in the relatively simple case of a single pure delay, as discussed in the recent works of Flamm and Mitter [3,4,5] and Foias, Tannenbaum and Zames [6,7].

Inspired by those articles, our aim here is to develop the theory to suit also multiple (commensurate) input lags. In fact, the scope of our technique goes beyond that class, and in a following note [13], it is applied in handling certain distributed delays.

Let us start right away with a description of our system, the problem, and an outline of the solution. Consider a system governed by a scalar transfer function $P(s) = P_0(s)B(e^{-s})$. Here $P_0(s)$ is a proper rational function and $B(z)$, a polynomial. A proper rational function $w(s)$ is the weight function. It is assumed that both $P_0(s)$ and $w(s)$ are stable and minimum phase², and that $P(s)$, hence $B(e^{-s})$, has no zeros on the imaginary axis³.

²Ways for handling instabilities and right half plane zeros in either $P_0(s)$ or $w(s)$ were developed in [4,5] and in [7]. Although these techniques adapt to our case, we prefer to concentrate here on the effect of the countable set of right half plane zeros of $B(e^{-s})$.

³This is a standard technical assumption (c.f. [3,4,5,6,7,14,15]).

0. Notations.

H^2 and H^∞ are the spaces of L_2 and L_∞ functions on the imaginary axis, having analytic continuations in the right half plane. A good source on H^p spaces is [9]. By abuse of notations, we shall **not** distinguish between a function in H^2 and its inverse Laplace transform, which is a function in $L_2[0,\infty]$. (In particular, our use of the symbol " \wedge " will be for a completely different purpose.) Hopefully, this will simplify, rather than obscure the discussion. A star will denote the adjoint of an operator or a matrix (e.g., T^*), and a bar, the complex adjoint of a scalar (e.g., \bar{z}). The kernel of a linear operator is denoted by "ker" (e.g. $\ker(T)$).

Given an internally stabilizing feedback compensator, $F(s)$, in the system described in Figure 1, the system's weighted sensitivity is the function $w(s)(1+P(s)F(s))^{-1}$. The problem is to

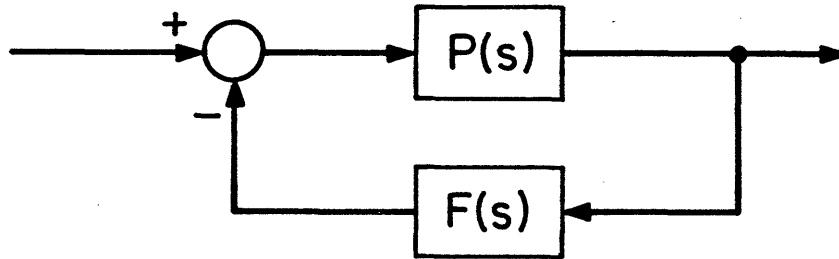


Figure 1

find that stabilizing $F(s)$ which minimizes the supremum of the weighted sensitivity over all frequencies; that is, the H^∞ norm

$$\|w(s)(1+P(s)F(s))^{-1}\|_\infty. \quad (1.1)$$

Following Zames [14], the equivalent interpolation problem is set in two steps. (For details see [14] or any of the following articles, e.g. [15].) First let $\phi(s)$ be the inner part of $P(s)$. Then there exists $h(s) \in H^\infty$ which minimizes this next sup-norm

$$\|w(s) - \phi(s)h(s)\|_\infty$$

Furthermore, having the minimizing function $h(s)$, the optimal feedback compensator is given by

$$F(s) = \frac{\phi(s)h(s)}{P(s)(w(s)-\phi(s)h(s))} \quad (1.3)$$

and the minimum value of the two problems is equal.

The second step follows from an important result of Sarason [12, Proposition 2.1]: Consider the compression of multiplication by $w(s)$ to the space $K =: H^2 \ominus \phi(s)H^2$; that is, the operator

$$Tf = \pi_K(w(s)f(s)) \quad \text{for } f \in K$$

(π_K being the orthogonal projection onto K along H^2). Then the minimum in (1.2) is equal to the operator norm of T . In other words: The function $\phi(s) = w(s) - \phi(s)h(s)$ interpolate T when $h(s)$ minimizes (1.2).

Sarason showed also [12, Proposition 5.1] that if the operator T has a maximal function, say $f(s)$, in K , then the unique interpolating function is $\phi(s) =: Tf(s)/f(s)$. Moreover, that $\phi(s)$ is a constant multiple (by $\|T\|$) of an inner function.

Our main effort will be to find a maximal function for T , or equivalently a maximal eigenvalue and eigenfunction for T^*T (this follows the philosophy of both [3,4,5] and [6]). The main result will be the construction of a parametrized family of matrices, $\Omega(\lambda^2)$, directly from the system's ingredients, such that λ^2 is an eigenvalue of T^*T precisely when $\Omega(\lambda^2)$ is singular. The corresponding eigenfunctions are computed from the right annihilating vectors of $\Omega(\lambda^2)$.

In some cases, however, a maximal eigenfunction does not exist. We treat this possibility too.

A key step in the analysis is the conversion of the problem to a time domain setup, which proved advantageous already in the single pure delay case. Yet the computations of π_K , T and T^* are significantly more complicated in the present context. The following section 2 is dedicated to these computations. The interpolation is then discussed in section 3.

2. The Time Domain Setup

The polynomial $B(z)$ can be written in the form

$$B(z) = az^m \prod_{i=1}^N (1 - e^{\alpha_i} z),$$

where non of the α_i 's is imaginary (as follows from the assumption $B(e^{-j\omega}) \neq 0$ for all $\omega \in \mathbb{R}$). Let $\alpha_1, \dots, \alpha_n$ ($n \leq N$) be those which lie in the open right half plane, and set

$$B_0(z) =: z^m \prod_{i=1}^n (1 - e^{\alpha_i} z).$$

Observation 2.1. The inner parts of $P(s)$ (namely $\phi(s)$) and of $B_0(e^{-s})$ are equal. Moreover, $B_0(e^{-s})H^p = \phi(s)H^p$ for all $p \geq 1$.

Proof. By assumption and the construction above, $P_0(s)B(e^{-s})/B_0(e^{-s})$ has no right half plane zeros, it is continuous on the imaginary axis and it is not of the form $e^{-\lambda s}h(s)$ for some $\lambda > 0$ and $h(s) \in H^\infty$. Hence it is outer and the first statement follows.

The inner part of $B_0(e^{-s})$ is the product of e^{-ms} and the blaschke product whose zeros are at $\alpha_i + 2k\pi j$ for $i=1, \dots, n$ and $k=0, \pm 1, \pm 2, \dots$. A simple computation shows that $\phi(s)/B_0(e^{-s})$ is uniformly bounded in the vicinity of these zeros, so the outer part of $B_0(e^{-s})$, has an H^∞ inverse. Hence the second part of the observation. Q.E.D.

In view of its time domain meaning as a delay operator, it will be convenient to work with $B_0(e^{-s})$, rather than with $\phi(s)$. From now on we

shall identify H^2 with $L_2[0, \infty)$ and K , with a subspace of the latter. All functions, unless stated otherwise, will be of the time variable.

The Subspace K .

We shall use the following notations: Given a scalar function $f(t)$, defined on $[0, \infty)$, $\tilde{f}(t)$ and $\hat{f}(t)$ will be these next n and $m+n$ -vector valued function

$$\tilde{f}(t) =: \begin{bmatrix} f(t) \\ f(t+1) \\ \vdots \\ f(t+n-1) \end{bmatrix} \quad \text{and} \quad \hat{f}(t) =: \begin{bmatrix} f(t) \\ f(t+1) \\ \vdots \\ f(t+m+n-1) \end{bmatrix}.$$

Here m and n are as in the definition of $B_0(z)$. It is assumed that $n \geq 1$, the case $n=0$ being virtually covered in [2-7].

Let us rewrite $B_0(z)$ in the form

$$B_0(z) = z^m(a_0 + a_1 z + \dots + a_n z^n)$$

Computation of the a_i is left out. It is important to note that both a_0 and a_n are non zero. Thus one can define an $n \times n$ matrix E , as follows

$$E =: - \begin{bmatrix} \bar{a}_0 & \bar{a}_1 & \dots & \bar{a}_{n-1} \\ 0 & \bar{a}_0 & \bar{a}_1 & \dots & \bar{a}_{n-2} \\ 0 & \dots & 0 & \bar{a}_0 \end{bmatrix}^{-1} \begin{bmatrix} \bar{a}_m & 0 & \dots & 0 \\ \bar{a}_{n-1} & \bar{a}_m & 0 & \dots & 0 \\ \bar{a}_1 & \bar{a}_2 & \dots & \bar{a}_n \end{bmatrix}$$

The matrix E plays a major role in our developments:

Theorem 2.2. The subspace K is determined by the matrix E as follows

$$K = : \{f: [0, \infty) \rightarrow \mathbb{C} \mid f|_{[0, m+n]} \in L_2[0, m+n];$$

$$\tilde{f}(t+in) = E^i \tilde{f}(t) \text{ for } t \geq m \text{ and } i=0,1,2,\dots\}$$

In particular, the geometric series $\sum (E^*E)^i$ converges.

Given any function $f(t)$ in $L_2[0, \infty)$, here is its projection $f_0(t) =: \pi_K f(t)$

$$f_0(t) = f(t) \quad \text{for } t \in [0, m)$$

$$\tilde{f}_0(t) = (I - E^*E) \sum_{i=1}^m E^{*i} \tilde{f}(t+in) \quad \text{for } t \in [m, m+1)$$

$$\tilde{f}_0(t+in) = E^i \tilde{f}_0(t) \quad \text{for } t \in [m, m+1) \text{ and } i=1,2,\dots$$

Corollary 2.3. The subspace K is isometric to $L_2([0,1], \mathbb{C}^{m+n})$, where the latter is endowed with the inner product

$$\langle \hat{f}, \hat{g} \rangle_K = \int_0^1 \langle \hat{f}(t), \hat{g}(t) \rangle dt$$

with

$$Q =: \begin{bmatrix} I_{m \times m} & 0_{m \times n} \\ 0_{n \times m} & (I - E^* E)^{-1} \end{bmatrix}$$

Proof of the Corollary. Theorem 2.2 establishes a continuous 1:1 correspondence between K and $L_2([0, m+n], \mathbb{C})$, or equivalently (using the " \wedge " symbol) - between K and $L_2([0, 1], \mathbb{C}^{m+n})$. The inner product is obtained by direct computation: Suppose $f_0(t)$ and $g_0(t)$ were in K . Then

$$\begin{aligned} \langle f_0, g_0 \rangle &= \int_0^\infty \langle f_0(t), g_0(t) \rangle dt \\ &= \int_0^m \langle f_0(t), g_0(t) \rangle dt + \sum_{i=0}^\infty \int_m^{m+1} \langle \tilde{f}_0(t+in), \tilde{g}_0(t+in) \rangle dt \\ &= \int_0^m \langle f_0(t), g_0(t) \rangle dt + \sum_{i=0}^\infty \int_m^{m+1} \langle E^i \tilde{f}_0(t), E^i \tilde{g}_0(t) \rangle dt \\ &= \int_0^1 \langle \hat{f}_0(t), \hat{g}_0(t) \rangle dt. \end{aligned}$$

Q.E.D.

Proof of Theorem 2.2. The space $K^\perp (= B_0(e^{-s})H^2)$ consists of the functions $f_1(t)$ of the form

$$f_1(t) = \sum_{i=0}^n a_i g(t-i-m) \quad \text{for } t \in [0, \infty),$$

allowing $g(t)$ to vary in $L_2[0, \infty)$ and taking $g(t)=0$ for negative t . Thus, $f_0(t)$ belongs to K if and only if for all $g(t)$ in $L_2[0, \infty)$, it satisfies

$$\begin{aligned} 0 &= \int_0^{\infty} \langle f_0(t), \sum_{i=0}^n a_i g(t-i-m) \rangle dt \\ &= \int_m^{\infty} \langle \sum_{i=0}^n \bar{a}_i f_0(t+i), g(t-m) \rangle dt. \end{aligned}$$

Our subspace K is thereby characterized by the difference equation

$$\sum_{i=0}^n \bar{a}_i f_0(t+i) = 0 \quad \text{for } t \geq m. \quad (2.1)$$

Observe that, in vector form, Eq. (2.1) becomes

$$\tilde{f}_0(t+n) = E\tilde{f}_0(t) \quad \text{for } t \geq m. \quad (2.2)$$

In order to complete the proof of the first part in the theorem, it remains to demonstrate the decay of E^i in operator norm. Indeed, notice that Eq. (2.1) gives rise to a group of operators on $L_2[0, n]$, denoted $S(t)$, which shifts along solutions. Precisely: given $\xi(\tau) \in L_2[0, n]$ and the

unique solution, $f(t)$, of Eq. (2.1) which satisfies $f(\tau) = \xi(\tau)$, $\tau \in [0, n]$, set $[S(t)f](\tau) =: f(t+\tau)$. The (point) spectrum of the generator of that group is the set of zeros of the equation

$$0 = \sum_{i=0}^n \bar{a}_i e^{is} = \prod_{i=1}^n (1 - e^{\bar{a}_i + s}),$$

and is therefore uniformly contained within the left half plane. In particular, the sequence $\|S(in)\|$ decays exponentially. Yet, the equivalence between Eqs. (2.1) and (2.2) implies that $\|E^i\| = \|S(in)\|$ which proves our claim (namely, the convergence of $\sum (E^*E)^i$).

As for the second part, one can easily check by direct computation that our definition of π_K is that of an orthogonal projection on K . For completeness, let us derive the stated formulas: Choose any function $f_1(t)$ in K^\perp . Obviously, $f_1(t)$ vanishes for $t < m$. Given any $f_0(t)$ in K we have

$$\begin{aligned} 0 &= \int_m^\infty \langle f_0(t), f_1(t) \rangle dt = \sum_{i=0}^\infty \int_m^{m+1} \langle E^i \tilde{f}_0(t), \tilde{f}_1(t+in) \rangle dt \\ &= \int_m^{m+1} \langle \tilde{f}_0(t), \sum_{i=0}^\infty E^{*i} \tilde{f}_1(t+in) \rangle dt. \end{aligned}$$

Since $\tilde{f}_0|_{[m, m+1]}$ could be any L_2 function, the following holds

$$\sum_{i=0}^{\infty} E^{*i} \tilde{f}_1(t+in) = 0 \quad \text{for } t \in [m, m+1].$$

Thereby, if $f(t) \in L_2[0, \infty)$ and $f_0(t) = \pi_K f(t)$, then

$$\begin{aligned} 0 &= \sum_{i=0}^{\infty} E^{*i} (\tilde{f}(t+in) - E^i \tilde{f}_0(t)) \\ &= \sum_{i=0}^{\infty} E^{*i} \tilde{f}(t+in) - (I - E^* E)^{-1} \tilde{f}_0(t) \quad \text{for } t \in [m, m+1]. \end{aligned}$$

Q.E.D.

The Operator T.

In order to formulate T we need some specific representation of the weight function $w(s)$. For simplicity we assume that its poles are all of multiplicity 1. Then it can be written as

$$w(s) = \eta + \sum_{i=1}^p \frac{\gamma_i}{\beta_i + s}.$$

By assumption ($w(s)$ is stable) all β_i lie in the open right half plane. In the time domain (i.e., by inverse Laplace transform) $w(s)$ becomes

$$w(t) = \eta\delta(t) + \sum_{i=1}^p \gamma_i e^{-\beta_i t}$$

The operator T (on K) is defined by convolution with $w(t)$, followed by projection onto K . That is

$$Tf(t) = \pi_K(\eta f(t) + \sum_{i=1}^p \gamma_i \int_0^t e^{\beta_i(\tau-t)} f(\tau) d\tau) \quad \text{for } f(t) \in K.$$

Since $\eta f(t)$ is in K , it is invariant under π_K . Also, by Theorem 2.2, the projection does not affect the convolution for $t \leq m$. In order to compute the projection for $t \geq m$, let us isolate one of the elements (suppressing for the moment the index i) $g(t) =: \int_0^t e^{\beta(\tau-t)} f(\tau) d\tau$. For fixed $t \in [m, m+1)$ and an integer i , straightforward computation yields

$$\begin{aligned} \tilde{g}(t+in) &= e^{-\beta t} [e^{-in\beta} v \int_0^m e^{\beta\tau} f(\tau) d\tau \\ &+ (F_0 \sum_{q=0}^{i-1} e^{(q-i)n\beta} E^q + F_1 E^i) \int_m^{m+1} e^{\beta\tau} \tilde{f}(\tau) d\tau \\ &+ E^i \int_m^t e^{\beta\tau} \tilde{f}(\tau) d\tau], \end{aligned} \quad (2.3)$$

where the vector v and the matrices F_0 and F_1 are as follows

$$v =: \begin{bmatrix} 1 \\ e^{-\beta} \\ \vdots \\ e^{-(n-1)\beta} \end{bmatrix}, \quad F_0 =: \begin{bmatrix} 1 & e^\beta & \dots & e^{(n-1)\beta} \\ e^{-\beta} & 1 & \dots & e^{(n-2)\beta} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-(n-1)\beta} & e^{-(n-2)\beta} & \dots & 1 \end{bmatrix}$$

$$\text{and } F_1 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ e^{-\beta} & 0 & \dots & 0 \\ \vdots & e^{-\beta} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ e^{-(n-1)\beta} & e^{-(n-2)\beta} & \dots & e^{-\beta} \end{bmatrix}.$$

For simplicity, we assume that $e^{-n\beta}$ is **not** an eigenvalue of E , and obtain a simpler formula for the second summand on the right hand side of Eq. (2.3), namely

$$\dots + (F_0 (1 - e^{n\beta} E)^{-1} (e^{in\beta} - E^i) + F_1 E^i) \int_m^{m+1} e^{\beta\tau} \tilde{f}(\tau) d\tau \dots$$

Next we have to compute $(I - E^* E) \sum E^{*i} g(t + in)$. Doing so term by term in Eq. (2.3), the last term becomes simply

$$\int_m^t e^{\beta(\tau-t)} \tilde{f}(\tau) d\tau.$$

The first term sums up over i , to

$$e^{-\beta t} (I - E^* E) (I - e^{-n\beta} E^*)^{-1} \int_0^m e^{\beta \tau} f(\tau) d\tau$$

The more complicated one is the middle one, which yields

$$e^{-\beta t} G \int_m^{m+1} e^{\beta \tau} \tilde{f}(\tau) d\tau$$

where

$$G = (I - E^* E) [(I - e^{-n\beta} E^*)^{-1} F_0 (I - e^{n\beta} E)^{-1} \\ + \sum_{i=0}^{\infty} E^{*i} (F_1 - F_0 (I - e^{n\beta} E)^{-1}) E^i].$$

It should be pointed out, however, that the infinite sum in the last expression converges to the solution of the Lyapunov type matrix equation

$$X - E^* X E = F_1 - F_0 (I - e^{n\beta} E)^{-1}$$

which is solvable via a finite procedure (for details see Djaferis and Mitter [2]).

We shall now collect all these details and describe the action of T as an operator on $L_2([0,1], C^{m+n})$ (which is possible, in view of Corollary 2.3). In doing so, however, we still need a few more notations: We

substitute G_i and v_i for G and v , which were constructed above, to indicate that β_i substitutes for β in their definitions. Then we build $(m+n) \times (m+n)$ matrices H_i , $i=1, \dots, p$, from the following blocks

$$(H_i)_{11} =: \begin{bmatrix} 0 & \dots & & & 0 \\ e^{-\beta} & & 0 & \dots & 0 \\ e^{-2\beta} & & e^{-\beta} & 0 & \dots & 0 \\ \vdots & & & & & \\ e^{-(m-1)\beta} & \dots & & e^{-\beta} & & 0 \end{bmatrix},$$

$$(H_i)_{21} =: (I - E^* E) (1 - e^{-n\beta} E^*)^{-1} [e^{-m\beta} v, \dots, e^{-\beta} v]$$

$$(H_i)_{12} =: 0_{m \times n} \text{ and } (H_i)_{22} =: G_i.$$

Observation 2.4. The operator T acts via

$$\begin{aligned} T\hat{f}(t) &= \eta\hat{f}(t) + \sum_{i=1}^p \gamma_i \int_0^t e^{\beta_i(\tau-t)} \hat{f}(\tau) d\tau \\ &\quad + \sum_{i=1}^p \gamma_i H_i \int_0^1 e^{\beta_i(\tau-t)} \hat{f}(\tau) d\tau \end{aligned} \tag{2.4}$$

for $\hat{f}(t)$ in $L_2([0,1], \mathbb{C}^{m+n})$.

Corollary 2.5. When restricted to $L_2([0,1], \mathbb{C}^{m+n})$, the adjoint operator, T^* , is given by

$$T^* \dot{f}(t) = \bar{\eta} \dot{f}(t) + \sum_{i=1}^p \bar{\gamma}_i \int_t^1 e^{\bar{\beta}_i(t-\tau)} \dot{f}(\tau) d\tau$$

(2.5)

$$+ \sum_{i=1}^p \bar{\gamma}_i Q^{-1} H_i^* Q \int_0^1 e^{\bar{\beta}_i(t-\tau)} \dot{f}(\tau) d\tau .$$

The observation is obvious. The corollary follows by direction computation, taking into account the inner product induced by the isometry to K on $L_2([0,1], C^{m+n})$, as described in Corollary 2.3.

3. The Main Steps in the Interpolation.

The operator T^*T is of the form $|\eta|^2 I +$ (a compact operator). By Weyl's Theorem (see e.g. [10, pp. 92 & 295]), its spectrum consists of countably many eigenvalues (each of a finite multiplicity) and their only cluster point, $|\eta|^2$, which, as we shall later show, is not an eigenvalue. In choosing an interpolation strategy, we distinguish between two possibilities: Should there be a maximal eigenvalue for T^*T , then the corresponding eigenfunction, say f , is maximal for T . This solves the interpolation problem, as explained in the introduction. Otherwise (i.e., when $\eta \neq 0$ and all eigenvalues are smaller than $|\eta|^2$) we check whether $w(s)$ itself interpolates T .

Here are sufficient conditions for these two possibilities:

Proposition 3.1. If there exists $\omega_0 > 0$ such that $|w(j\omega)| > |\eta|$ for all $\omega \in \mathbb{R}$, $|\omega| > \omega_0$, then infinitely many eigenvalues of T^*T are larger than $|\eta|^2$. In particular, a maximal eigenvalue does exist.

Proposition 3.2. If $|w(j\omega)| \leq |\eta|$ for all $\omega \in \mathbb{R}$, then $w(s)$ interpolates T and the zero feedback $F(s)$ is optimal.

We defer the proofs to this section's end. Meanwhile we assume that a maximal eigenvalue exists, and try to find it. The following is a nice and very useful observation of Flamm [4,5]:

Proposition 3.3. Suppose λ^2 is an eigenvalue of T^*T and let s_1, \dots, s_{2p} be the roots of the equation

$$w(s)\overline{w(-\bar{s})} = \lambda^2, \quad (3.1)$$

repeated according to multiplicity. Let $\xi_1(t), \dots, \xi_{2p}(t)$ be scalar functions defined as follows: If $s_i = s_{i+1} = \dots = s_{i+q}$ is a root of multiplicity $q+1$, then

$$\xi_i(t) =: e^{s_i t}, \quad \xi_{i+1}(t) =: t e^{s_i t}, \dots, \quad \xi_{i+q}(t) =: t^q e^{s_i t}.$$

Then, any associated eigenfunction lies in the subspace $U =: \text{sp}\{\xi_1(t), \dots, \xi_{2p}(t)\}$. (Note: Since eigenfunction take values in C^{m+n} and $\xi_i(t)$ are scalars, the coefficients in the relevant linear combinations are $(m+n)$ -vectors.)

Proof. Consider this next set of differential equations

$$\begin{aligned} \frac{d}{dt} \hat{x}_i(t) &= -\beta_i \hat{x}_i(t) + \gamma_i \hat{f}(t) \\ \frac{d}{dt} \hat{x}_{i+p}(t) &= \bar{\beta}_i \hat{x}_{i+p}(t) - \bar{\gamma}_i \hat{g}(t), \end{aligned} \quad (3.2)$$

for $i=1, \dots, p$, and assume

$$\hat{g}(t) = \eta \hat{f}(t) + \sum_{i=1}^p \hat{x}_i(t)$$

and

$$\dot{\mathbf{h}}(t) = \bar{\eta} \dot{\mathbf{g}}(t) + \sum_{i=1}^p \dot{\mathbf{x}}_{i+p}(t). \quad (3.3)$$

It is easy to see that for properly chosen initial data (defined by $\dot{\mathbf{f}}(t)$) we have $\dot{\mathbf{g}}(t) = T\dot{\mathbf{f}}(t)$ and $\dot{\mathbf{h}}(t) = T^*\dot{\mathbf{g}}(t)$.

Suppose now that $\dot{\mathbf{f}}(t)$ is an eigenfunction associated with the eigenvalue λ^2 . It then follows (combining Eqs. (3.3) and (3.2), and substituting terms in $\dot{\mathbf{f}}(t)$ and $\dot{\mathbf{x}}_i(t)$ for $\dot{\mathbf{g}}(t)$ and $\dot{\mathbf{h}}(t)$) that both $\dot{\mathbf{f}}(t)$ and $\dot{\mathbf{x}}_i(t)$, $i=1, \dots, 2p$, are analytic. Taking the Laplace transforms of their analytic continuations to the whole half line $[0, \infty)$, one obtains

$$(\lambda^2 - |\eta|^2) \dot{\mathbf{f}}(s) = [\bar{\eta}B, B](sI - \begin{bmatrix} -D & 0 \\ -CB^* & D^* \end{bmatrix})^{-1} \left(\begin{bmatrix} C \\ -\bar{C}\eta \end{bmatrix} \dot{\mathbf{f}}(s) + x(0) \right), \quad (3.4)$$

where $x(0)$ is the $2p(m+n)$ -vector of initial data for $\dot{\mathbf{x}}_1(t), \dots, \dot{\mathbf{x}}_{2p}(t)$,

$$B =: [I_{m+n}, \dots, I_{m+n}] \quad \text{p times} \quad C =: \begin{bmatrix} \gamma_1 I_{m+n} \\ \vdots \\ \gamma_p I_{m+n} \end{bmatrix},$$

\bar{C} is the like of C with $\bar{\gamma}_i$ substituting for γ_i , and finally, D is the block diagonal matrix

$$D =: \text{diag}\{\beta_1 I_{m+n}, \dots, \beta_p I_{m+n}\}.$$

Direct computation shows that the coefficient of $\dot{\mathbf{f}}(s)$ on the right hand

side of Eq. (3.4) is $\overline{w(s)w(-\bar{s}) - |\eta|^2}$. Eq. (3.4) thus becomes

$$\hat{f}(s) = \frac{1}{\lambda^2 \overline{w(s)w(-\bar{s})}} [\bar{\eta}B, BI - \begin{bmatrix} -D & 0 \\ -CB^* & D^* \end{bmatrix}]^{-1} x(0). \quad (3.5)$$

Now, the poles of the inverse matrix in (3.5) are exactly those of $w(s)w(-\bar{s})$ (i.e., $-\beta_1, \dots, -\beta_p, \bar{\beta}_1, \dots, \bar{\beta}_p$), so they cancel each other. It remains that the poles of $\hat{f}(s)$ are zeros of Eq. (3.1), as stated. Q.E.D.

Having in mind one (candidate for being an) eigenvalue, λ^2 , we can now consider the restriction of T to the corresponding subspace $U =: \text{sp}\{\xi_1(t), \dots, \xi_{2p}(t)\}$. The range of T is then $V =: U + \text{sp}\{e^{-\beta_1 t}, \dots, e^{-\beta_p t}\}$. Restricted to V , the adjoint operator T^* takes values in $W =: V + \text{sp}\{e^{\bar{\beta}_1 t}, \dots, e^{\bar{\beta}_p t}\}$.

Thus restricted, T and T^* can be described by $3p(m+n) \times 2p(m+n)$ and $4p(m+n) \times 3p(m+n)$ matrices, which we denote (block-wise)

$$\begin{matrix} 2p \\ p \end{matrix} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} \quad \text{and} \quad \begin{matrix} 2p & p \\ p & \\ p & \end{matrix} \begin{bmatrix} T_{11}^* & \vdots & 0 \\ \hline 0 & \vdots & T_{22}^* \\ \hline T_{31}^* & \vdots & T_{32}^* \end{bmatrix}$$

The corresponding vectors describe the coefficients of $\xi_1(t), \dots, \xi_{2p}(t)$, $e^{-\beta_1 t}, \dots, e^{-\beta_p t}$ and $e^{\bar{\beta}_1 t}, \dots, e^{\bar{\beta}_p t}$, according to this order. (Recall that each of the coefficients is an $(m+n)$ -vector.) We leave out the detail of these matrices, which are read directly from Formulas (2.4) and (2.5).

Now set

$$\Omega(\lambda^2) =: \left[\begin{array}{c} T_2 \\ \hline T_{31}^* T_1 + T_{31}^* T_2 \end{array} \right].$$

(This $2p(m+n) \times 2p(m+n)$ matrix depends on λ^2 , since our whole construction starts with Eq. (3.1).) Here is our main result.

Theorem 3.2. A positive number λ^2 is an eigenvalue of T^*T if and only if the corresponding matrix $\Omega(\lambda^2)$ is singular (i.e., $\det \Omega(\lambda^2) = 0$). If this is the case and $z \in C^{2p(m+n)}$ satisfies $\Omega(\lambda^2)z = 0$, then an associated eigenfunction is

$$\hat{f}(t) = \sum_{i=1}^{2p} z_i \xi_i(t), \quad (3.6)$$

Proof. We leave it to the reader to check these next two facts: (i) $T_{11}^* T_1 = \lambda^2 I_{2p(m+n)}$. (Hint: Check first the case where s_1, \dots, s_{2p} are distinct values and $\xi_i(t) = e^{s_i t}$.) (ii) T_{22}^* is invertible. (Explicitly: $T_{22}^* = \text{diag}\{w(\beta_1), \dots, w(\beta_p)\}$, and since $w(s)$ is minimum phase, T_{22}^* is invertible.)

Now if $\hat{f}(t)$ is an eigenfunction associated with λ^2 , it belongs to U (by Flamm's result). Hence so does $T^*T\hat{f}(t)$. Letting z be the vector of coefficients, as described in Formula (3.6), we should have $T_{22}^* T_1 z = 0$ and $(T_{31}^* T_1 + T_{32}^* T_2)z = 0$. In view of fact (ii), above, it means that $\Omega(\lambda^2)z = 0$. The converse direction is by now obvious. Q.E.D.

Remark 3.3. A characterization of eigenvalues of T^*T could also be

made in terms of certain boundary constraints that Eqs. (3.2) and (3.3) must satisfy in order to be a true model for T and T^* . It seems that solution of the associated two sided boundary value problem will be computationally far more complicated than our criterion.

Since we look for a maximal eigenvalue, our suggested scheme is simply: Compute $\det \Omega(\lambda^2)$ for a decreasing sequence of numbers λ^2 and find the first λ^2 which yields the value zero. This next simple observation is of great help.

Observation 3.4. The maximal eigenvalue λ^2 lies in the interval $(|\eta|^2, \|w(s)\|_\infty^2]$.

Proof. Since $\|w(s)\|_\infty$ is the norm of the uncompressed multiplication by $w(s)$, we have $\lambda^2 = \|T^*T\|^2 \leq \|w(s)\|_\infty^2$. Since $|\eta|^2$ is a cluster point of eigenvalues of T^*T , λ^2 cannot be strictly smaller than $|\eta|^2$. Finally, one easily finds out that if $|\eta|^2$ is substituted for λ^2 in Eq. (3.5), above, then the right hand side of that equation is not strictly proper, hence f could not be in K . Q.E.D.

Assume that a maximal eigenfunction, $f(s)$, for T^*T (i.e., a maximal function for T) is at hand. Let us recall now the further steps described in the introduction towards the solution of the interpolation, and of the original sensitivity minimization, problems: The interpolating function is $\varphi(s) = Tf(s)/f(s)$. Using the notations of section 1, $\varphi(s)$ is of the form

$$\varphi(s) = w(s) - \phi(s)h(s),$$

for some $h(s) \in H^\infty$, and the corresponding optimal feedback compensator is

$$F(s) = \frac{\phi(s)h(s)}{P(s)(w(s)-\phi(s)h(s))} .$$

Substituting $w(s)-TF(s)/f(s)$ for $\phi(s)h(s)$, we get

$$\begin{aligned} F(s) &= \frac{w(s)f(s)-TF(s)}{P(s)TF(s)} \\ &= \frac{(I-\pi_K)(w(s)f(s))}{P(s)TF(s)} \end{aligned}$$

which is the solution to the original problem.

Before bringing two more proofs that we owe, let us make one more comment: The formula we have just displayed is of the optimal compensator, when optimality is measured in terms of minimal weighted sensitivity. This compensator is likely to have, however, serious drawbacks. As in simpler cases, it might be unstable, and have delays. This has already been observed by Flamm in [5]. So in order to obtain a more practical compenstor, one has to look for an approximating, stable and rational suboptimal $F(s)$. This search goes beyond the scope of the present discussion.

Here are the proofs of Propositions 3.1 and 3.2.

Proof of Proposition 3.1. A key fact is that the operator $f: \rightarrow Rf =: (I-\pi_K)(w*f): K \rightarrow K$ is finite dimensional. One can easily see that (when restricted to $f(t)*K$) the terms in $w*f(t)$ which contribute to the projection onto K are the first and second summands on the right hand side of Eq. (2.3) (with β_k , $k=1, \dots, p$, substituting for β). Each of these terms

defines a finite dimensional operator; in particular, so are their projections onto K^\perp .

Henceforth the proof goes mainly in the frequency domain; i.e., members of K are regarded as functions of the imaginary argument $j\omega$.

Claim. The restrictions of functions in K to any (fixed) finite interval $[-j\omega_1, j\omega_1]$, are dense in the associated space $L_2[-j\omega_1, j\omega_1]$.

Proof. Suppose a function $g(j\omega) \in L_2[-j\omega_1, j\omega_1]$ is orthogonal to the restrictions of all functions in K to $[-j\omega_1, j\omega_1]$. Each member in K is of the form

$$f(j\omega) = \{f(j\omega)\} \int_0^1 e^{-j\omega t} f(t) dt,$$

where $\{f(j\omega)\}$ is this next row vector

$$\{f(j\omega)\} =: [1, e^{-j\omega}, \dots, e^{-(m+n-1)j\omega}] \left[\begin{array}{c|c} I_{m \times m} & 0_{m \times n} \\ \hline 0_{n \times m} & (I_{n \times n} - Ee^{-j\omega n})^{-1} \end{array} \right].$$

(This follows from Theorem 2.2.)

Hence $\langle g, f \rangle = 0$ for all $f \in K$, means

$$\begin{aligned}
0 &= \int_{-j\omega_1}^{j\omega_1} \langle g(j\omega), \zeta(j\omega) \int_0^1 e^{-j\omega t} \dot{\hat{f}}(t) dt \rangle d\omega \\
&= \int_0^1 \langle \int_{-j\omega_1}^{j\omega_1} e^{j\omega t} \zeta(j\omega)^* g(j\omega) d\omega, \dot{\hat{f}}(t) \rangle dt
\end{aligned}$$

In effect, $\zeta^*(j\omega)g(j\omega)$ should vanish. Since $\zeta(j\omega) \neq 0$, the observation $g(j\omega) = 0$ follows.

By hypothesis $|w(j\omega)| > |\eta|$ for $\omega \in \mathbb{R}$, $|\omega| > \omega_0$. Choose positive ω_1 , σ and ε , such that $\omega_1 > \omega_0 + \varepsilon$ and $|w(j\omega)| > |\eta| + \varepsilon$ when $\omega_0 + \sigma \leq |\omega| \leq \omega_1$. Then denote by X the infinite dimensional subspace of $L_2[-j\omega_1, j\omega_1]$, which contains those functions that vanish along $[-j(\omega_0 + \sigma), j(\omega_0 + \sigma)]$. If $g(j\omega)$ is in X , we have

$$\|w(j\omega)g(j\omega)\|_{L_2[-j\omega_1, j\omega_1]} > (|\eta| + \varepsilon) \|g(j\omega)\|_{L_2[-j\omega_1, j\omega_1]}. \quad (3.7)$$

Now, recall R , the operator defined at the beginning of the proof, and let Y be the closure in $L_2[-j\omega_1, j\omega_1]$ of the (restrictions to $[-j\omega_1, j\omega_1]$ of functions in) $\ker(R)$. The subspace Y is of finite codimension. It thus easily follows from the definition of X that $Y \cap X$ is infinite dimensional. Let $g(j\omega)$ belong to that intersection, and let $f(j\omega) \in \ker(R)$ approximate $g(j\omega)$ well enough, which is possible by the definition of Y . Since $g(j\omega)$ satisfies inequality (3.7), the function $f(j\omega)$ satisfies

$$\|w(j\omega)f(j\omega)\|_{L_2[-j\omega_1, j\omega_1]}^2 > |\eta|^2 \|f(j\omega)\|_{L_2[j\omega_1, j\omega_1]}^2 \quad (3.8)$$

(In fact, the degree of the approximation of g by f is made so that (3.8) will hold.)

By definition of R , the choice of $f(j\omega)$ in $\ker(R)$ implies that $Tf(j\omega) = w(j\omega)f(j\omega)$. Since for $|\omega| > \omega_0$, the inequality $|w(j\omega)| > |\eta|^2$ holds, we can substitute ∞ for ω_1 in (3.8) without affecting the validity of that inequality.

Summing up: We found a function $f(j\omega) \in K$ for which $\|Tf\| > |\eta| \|f\|$. The completeness of the set of eigenfunctions of T^*T implies the existence of an eigenvalue which is larger than $|\eta|^2$. In order to show that infinitely many such eigenvalues exist, one subtracts from Y , above, the (finite dimensional) eigenspace associated with the eigenvalue we have just found. Then the same argument implies the existence of a second eigenvalue which is larger than $|\eta|^2$, and so on. Q.E.D.

Proof of Proposition 3.1. Since $|\eta| = \lim |w(j\omega)|$, the assumption $|w(j\omega)| \leq |\eta|$ implies $\|w(s)\|_\infty = |\eta|$. As when arguing Observation 3.4, we now obtain

$$|\eta|^2 = \|w(s)\|_\infty^2 \geq \|T\|^2 = \|T^*T\| = \sup\{\lambda^2 : \lambda^2 \text{ is an eigenvalue}$$

$$\text{of } T^*T^2\} \geq |\eta|^2;$$

namely $\|w(s)\|_\infty = \|T\|$. This means that T is interpolated by $w(s)$. (The other requirement in the definition of an interpolating function is

$Tf = \pi_K(w(s)f(s))$, and is met by definition of T .)

Q.E.D.

Acknowledgement

I wish to thank Professor Sanjor K. Mitter and Dr. David S. Flamm for interesting discussions and enlightening remarks concerning this work.

REFERENCES

- [1] V.M. Adamjan, D.Z. Arov, and M.G. Krein, "Infinite Block Hankel Matrices and Related Extension Problems," AMS Translations 111 (1978), pp. 133-156.
- [2] T.E. Djaferis and S.K. Mitter, "Algebraic Methods for the Study of Some Linear Matrix Equations", Linear Algebra and Its Applications, 44(1982), pp.125-142.
- [3] D.S. Flamm and S.K. Mitter, "Progress on H^∞ Optimal Sensitivity for Delay Systems I: Minimum Phase Plant with Input Delay, 1 Pole/zero Weighting Function", LIDS Technical Report P-1513, MIT, November 1985.
- [4] D.S. Flamm and S.K. Mitter, "Progress on H^∞ Optimal Sensitivity for Delay Systems II: Minimum Phase Plant with Input Delay, General Rational Weighting Function", preprint LIDS-MIT, January 1986.
- [5] D.S. Flamm, Ph.D. Thesis, LIDS-MIT, to appear.
- [6] C. Foias, A. Tannenbaum, and G. Zames, "Weighted Sensitivity Minimization for Delay Systmes," Technical Report, Dept. of Electrical Engineering, McGill University, September 1985.
- [7] C. Foias, A Tannenbaum, and G. Zames, "On the H^∞ -Optimal Sensitivity Problem for Systems with Delays," Technical Report, Dept. of Electrical Engineering, McGill University, December 1985.
- [8] B.A. Francis and J. Doyle, "Linear Control Theory with an H^∞ Optimality Criterion," Systems Control Group report #8501, University of Toronto, October 1985.
- [9] J.B. Garnett, Bounded Analytical Functions, Academic Press, New York, 1981.
- [10] P.R. Halmos, A Hilbert Space Problem Book. American Book Co., New York (1967).
- [11] J.W. Helton, "Worst Case Analysis in the Frequency Domain: The H^∞ Approach to Control," IEEE Trans. Auto. Control, AC-30 (1985), pp. 1154-1170.
- [12] D. Sarason, "Generalized Interpolation in H^∞ ," Trans. AMS, 127 (1967), pp. 179-203.
- [13] G. Tadmor, "An Interpolation Problem Associated with H^∞ Optimal Design in Systems with Distributed Input Lags", LIDS Technical

Report P-1545, MIT, March 1986.

- [14] G. Zames, "Feedback and Optimal Sensitivity: Model Reference Transformations, Seminorms and Approximate Inverses," IEEE Trans. Auto. Control, AC-23 (1981), pp. 301-320.
- [15] G. Zames and B.A. Francis, "Feedback Minimax Sensitivity and Optimal Robustness", IEEE Trans. Auto. Control, AC-28 (1983) pp. 585-601.