SIAM J. CONTROL AND OPTIMIZATION
Vol. 25, No. 3, May 1987

011

ON THE H^∞-OPTIMAL SENSITIVITY PROBLEM FOR SYSTEMS WITH DELAYS*

CIPRIAN FOIAS†, ALLEN TANNENBAUM‡ AND GEORGE ZAMES§

Abstract. In this paper we extend some of the results of [IEEE Trans. Automat. Control, AC-31 (1986), pp. 763-766] to more general delay systems. In particular, we analyze the effect of the interaction of delays and nonminimum phase zeros on the H^∞-optimal weighted sensitivity.

Key words. sensitivity minimization, delay, contraction, defect operator, distributed system

AMS(MOS) subject classifications. 93B35, 93C05

Notation and Terminology.

$D =$ open unit disc
$\bar{D} =$ closed unit disc
$\partial D =$ unit circle
$H =$ open right half plane
$\bar{H} =$ closed right half plane
$\tilde{\bar{H}} = \bar{H} \cup \{\infty\}$
$H^p(X) =$ the standard Hardy p-space ($1 \leqq p \leqq \infty$) on X where $X = D$ or H. See Duren [6] or Rudin [19] for details. We will also use some elementary facts about L^p-spaces and Hilbert spaces. Again see [6] or [19] for details.
$H^2(X) \ominus uH^2(X) =$ orthogonal complement of $uH^2(X)$ in $H^2(X)$ where $u \in H^\infty(X)$ is an inner function.

Let S denote an arbitrary Hilbert space with inner product $\langle\ ,\ \rangle$. Then for $x, y \in S$, $x \otimes y$ denotes the operator defined by $(x \otimes y)w := \langle w, y\rangle x$ for $w \in S$.

On the unit circle ∂D we identify $\bar{z}$ and $1/z$ in the usual way.

Finally we use all the standard notation from Hilbert space theory. See, e.g., [6], [19], [24].

Introduction. This paper is the sequel to [9]. We recall that in [9], the authors solved the weighted H^∞-minimization problem for a plant consisting of a pure delay and arbitrary stable (with stable inverse) real rational proper weighting function. We saw that in contrast to the unweighted problem, which reduces to a simple classical Nevanlinna-Pick interpolation problem for a large class of distributed systems [7], [16], even for the simplest weighting function ($W(s) = 1/(as+1)$, $a > 0$), the weighted problem reflects the distributed nature of systems with delays.

In this paper, we give a general procedure for computing the optimal weighted sensitivity for an arbitrary real rational stable (with stable inverse) weight, and for plants of the form $e^{-hs}P_0(s)$ where $P_0(s)$ is a proper real rational function with no poles or zeros on the $j\omega$-axis.

In point of fact, we give a general procedure for solving the following kind of problem: Let $P(s)$ be a plant (perhaps distributed) and suppose that we have a

* Received by the editors December 20, 1985; accepted for publication (in revised form) March 24, 1986.
† Department of Mathematics, Indiana University, Bloomington, Indiana 47405.
‡ Department of Electrical Engineering, McGill University, Montréal, Québec, Canada H3A 2A7 and Department of Mathematics, Ben-Gurion University, Beer Sheva, Israel. Present address, Department of Electrical Engineering, University of Minnesota, Minneapolis, Minnesota 55455.
§ Department of Electrical Engineering, McGill University, Montréal, Québec, Canada H3A 2A7.

factorization $P(s) = P_1(s)P_2(s)$. Then for given weight, we can write down an expression for the H^∞-optimal sensitivity of $P(s)$ in terms of data determined by $P_1(s)$ and $P_2(s)$. Moreover in this expression (see (3.2) below for a more precise statement) the data given by $P_1(s)$ is *decoupled* from the data given by $P_2(s)$. So for example when $P(s) = e^{-hs}P_0(s)$ as above, we can apply our procedure to $P_1(s) = e^{-hs}$ and $P_2(s) = P_0(s)$, a proper real rational function for which the optimal sensitivity problem is easy to solve.

Our methods are in a certain sense a generalization in the rational weighting case of the one-step extension technique of Adamjan, Arov and Krein [1], [2] and actually give new proofs to certain of their results (see Theorem 3.2, § 3.4 and Theorem 3.9 below for details). Basically what we have solved is an "n-step" or even an "∞-step" extension problem Theorem 3.2. Thus our techniques even give a new viewpoint to certain problems in Nevanlinna-Pick interpolation theory [14].

As in [9], our methods are heavily influenced by the results of Sarason [20] and Sz. Nagy and Foias [23], [24]. Consequently, we will be working in $H^2(D)$ where D is the unit disc. Moreover, the techniques we use have a strong complex-analytic flavor.

Finally in § 4, we will apply our procedure to the case

$$P(s) = e^{-hs}\left(\frac{s-b}{s+b}\right), \qquad W(s) = \frac{1}{as+1},$$

a, b, $h > 0$. This will allow us to understand the coupling and effect of the three fundamental parameters a (the inverse of the bandwidth), b (the nonminimum phase zero), and h (the delay) on the optimal sensitivity. As expected for $b \to \infty$, our formula approaches that of [9] (see also § 1), and so our method here actually gives an alternative route to some of the results of [9].

1. Preliminaries. In this section we would like to briefly review some of the material from our paper [9], and set up some of the notation connected with the weighted sensitivity H^∞-minimization problem posed by Zames [26]. We should note that independently David Flamm in his thesis [8] (done while at M.I.T.) has derived some results very similar to the ones that we will describe in this section. Israel Gohberg more recently discussed with the authors an approach to derive (1) below, similar to that of Flamm's using the Hankel operator.

We begin by recalling the general weighted sensitivity H^∞-minimization problem for SISO, LTI plants (see [11] for an excellent survey on all of this). We are given a SISO, LTI plant $P(s)$, and a stable (with stable inverse) proper real rational weight $W(s)$. Let $C(s)$ denote an internally stabilizing LTI controller for $P(s)$ in the feedback system of Fig. 1.

Then following [26], we define the *weighted sensitivity*:

$$S_W(s) := W(s)(1 + P(s)C(s))^{-1}.$$

The problem in which we are interested is in determining the existence of and computing

$$\inf\{\|S_W(s)\|_\infty : C \text{ stabilizing}\}$$

where $\|\ \|_\infty$ denotes the H^∞-norm in the right half plane H.

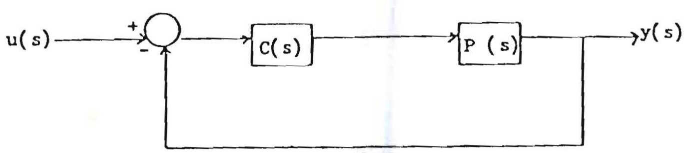

FIG. 1. *Standard feedback configuration.*

In the finite-dimensional case, this problem is discussed and solved in [13], [12], [15]. In our previous paper [9], we considered the case in which $P(s)=e^{-hs}$, and $W(s)$ a stable strictly proper real rational weighting function with stable inverse. Basically we showed that the problem of the computation of the optimal sensitivity could be reduced to computing the eigenvalues of a certain linear ordinary differential operator with constant coefficients of order $2n$ subject to $2n$ boundary conditions, where $n=$ number of poles of $W(s)$. From the associated Wronskian determinant of the problem, we could then find the required minimal sensitivity (actually all of the singular values of the associated Hankel operator).

To see how this goes, let us briefly sketch the argument from [9]. (See [9] for all the rigorous details.) First of all using the results of [26], one can show that the computation of the optimal sensitivity amounts to finding:

$$\mu := \inf_{q\in H^\infty} \| W(s)-e^{-hs}q(s)\|_\infty .$$

(Throughout this section $H^2 := H^2(H)$, $H^\infty := H^\infty(H)$.) Let $\Pi : H^2 \to H^2 \ominus e^{-hs}H^2$ denote orthogonal projection. Moreover, we denote by M_W the operator $H^2\to H^2$ defined by multiplication by W. Then by [1], [20], [23],

$$\mu = \|\Pi M_W|_{H^2\ominus e^{-hs}H^2}\|.$$

Computing this norm is not difficult. Indeed we can show via the Fourier or Laplace transform (see [20]) that there exists an isometric isomorphism

$$\phi : H^2\ominus e^{-hs}H^2 \xrightarrow{\sim} L^2[0,h].$$

Setting

$$\Gamma := \phi \circ (\Pi M_W|_{H^2\ominus e^{-hs}H^2}) \circ \phi^{-1}$$

we are reduced to computing $\|\Gamma\|$. (Notice $\Gamma : L^2[0,h] \to L^2[0,h]$.) But again from [20] it follows that we can identify the operator "$1/s$" on $H^2\ominus e^{-hs}H^2$ with the Volterra operator

$$V : L^2[0,h] \to L^2[0,h]$$

$Vf(x) := \int_0^x f(t)\,dt$ via ϕ. The inverse operator (of course unbounded) of V is the derivative operator $Df=f'$ with domain consisting of

$$\{f\in L^2[0,h] : f'\in L^2[0,h], f(0)=0\}$$

(i.e. the operator D corresponds to "s").

Now to compute $\|\Gamma\|$, we need to compute the largest eigenvalue of $\Gamma^*\Gamma$ (since Γ is compact), or equivalently the smallest positive eigenvalue of $(\Gamma^*\Gamma)^{-1}$. To do this we clearly only need identify the adjoint D^* of D. But it is easy to compute (using integration by parts) that $D^* = -D$ with domain

$$\{f\in L^2[0,h] : f'\in L^2[0,h], f(h)=0\}.$$

With these remarks one can derive the eigenvalue problem alluded to above [9].

In the particular case in which

$$W(s)=\frac{1}{as+1}, \qquad a>0,$$

we get $\Gamma \cong (aD+1)^{-1}$ and one derives the eigenvalue problem of finding the largest positive ρ (it is straightforward to check $\rho < 1$) such that

$$(-a^2D^2+1)f = \frac{1}{\rho^2}f, \quad f(h)=0, \quad -af'(0)+f(0)=0.$$

From the associated Wronskian, one is reduced to finding the largest $\rho \in (0, 1)$, say ρ_1, that satisfies

$$\left(\sqrt{\frac{1}{\rho^2}-1}\right)+\tan\left(\frac{h\sqrt{(1/\rho^2)-1}}{a}\right)=0. \tag{1}$$

Then ρ_1 is the required norm (and the first singular value of the associated Hankel operator).

Note that if $\rho_2 \in (0, 1)$, $\rho_2 < \rho_1$, is the next largest root of (1), then ρ_2 will be the second singular value of the associated Hankel, and so on. In other words we have an explicit procedure for computing all of the singular values of the associated Hankel from the Wronskian of a certain elementary eigenvalue problem. Moreover we can clearly even write down the Schmidt vectors using this procedure. (See [18] for the relevant definitions.)

In §§ 3 and 4 below, we will offer another procedure for computing the optimal sensitivity applicable to more general delay systems. Our new method only makes use of elementary properties of $H^2(D)$ and $H^\infty(D)$ and reduces the optimal sensitivity problem to an algebraic one. We will generalize (1) in § 4 to the case of a plant with a delay and a nonminimum phase zero.

2. Triangular operators. In this section we collect some standard facts about certain types of lower block triangular operators. Our basic references are [22], [23].

Let H_1, H_2 denote (complex) Hilbert spaces, and set $H := H_1 \oplus H_2$. Let $S: H \to H$ be a bounded linear operator such that H_2 is S-invariant subspace of H, i.e., $S|H_2: H_2 \to H_2$. Then clearly we can write

$$S = \begin{bmatrix} S_1 & 0 \\ Y & S_2 \end{bmatrix}$$

where $S_1 := (S^*|H_1)^*$, $S_2 := S|H_2$, and $Y: H_1 \to H_2$ is the *coupling operator.*

Next let $A: H \to H$ be an arbitrary contraction, i.e., $\|A\| \leqq 1$. Then in the usual way [24] we can define the associated *defect operators* and *defect spaces*:

$$D_A := (I - A^*A)^{1/2}, \qquad D_{A^*} := (I - AA^*)^{1/2},$$

$$\mathcal{D}_A := \overline{D_A H}, \qquad \mathcal{D}_{A^*} := \overline{D_{A^*} H}.$$

We can now state one of the key results of [22].

THEOREM 2.1. *With the above notation,* $\|S\| \leqq \rho$ *if and only if* $\|S_i\| \leqq \rho$ $(i=1, 2)$ *and* $Y = D_{S_2^*/\rho} L D_{S_1/\rho}$ *for some* $L: \mathcal{D}_{S_1/\rho} \to \mathcal{D}_{S_2^*/\rho}$ *such that* $\|L\| \leqq \rho$. *Moreover, if we set* $\theta := \max\{\|S_1\|, \|S_2\|\}$, *and assume* $\rho > \theta$, *then* $\|S\| = \rho$ *if and only if* $\|L\| = \rho$.

Proof. The first statement is Theorem 1 of [22]. The second statement is standard, but since we do not know a convenient reference, we will include the proof. By scaling we can assume $\rho = 1$. Therefore under the hypothesis that $1 > \theta$, we want to show $\|S\| = 1$ if and only if $\|L\| = 1$.

Suppose first $\|S\|=1$. Then following [22, pp. 205-207], one can define an isometry $\sigma:\mathscr{D}_S\to\mathscr{D}_L\oplus\mathscr{D}_{S_2}$ such that

$$\sigma D_S h=\begin{bmatrix} D_L D_{S_1} & 0 \\ -S_2^* L D_{S_1} & D_{S_2}\end{bmatrix}\begin{bmatrix} h_1 \\ h_2\end{bmatrix}.$$

Now $\|S\|=1$ if and only if there exists a sequence

$$h^{(n)}=\begin{bmatrix} h_1^{(n)} \\ h_2^{(n)}\end{bmatrix},$$

$\|h^{(n)}\|=1$ such that $D_S h^{(n)}\to 0$, which in turn is equivalent to

$$(*)\qquad \|D_{S_2}h_2^{(n)}-S_2^* L D_{S_1}h_1^{(n)}\|^2+\|D_L D_{S_1}h_1^{(n)}\|^2\to 0.$$

We claim now that

$$\limsup_{n\to\infty}\|D_{S_1}h_1^{(n)}\|=:q>0.$$

Indeed, suppose not. Then $\|h_1^{(n)}\|\to 0$ since D_{S_1} is invertible (S_1 by hypothesis is a strict contraction), and therefore by $(*)$ $\|D_{S_2}h_2^{(n)}\|\to 0$, and so $\|h_2^{(n)}\|\to 0$ since D_{S_2} is invertible (S_2 is a strict contraction). But this contradicts our hypothesis that $\|h^{(n)}\|=1$.

Choose a subsequence $\{h_1^{(n')}\}$ such that $\|D_{S_1}h_1^{(n')})\|>0$, and $\|D_{S_1}h_1^{(n')}\|\to q$. Since by $(*)$ $\|D_L D_{S_1}h_1^{(n)}\|\to 0$, we get that $\|D_L(D_{S_1}h_1^{(n')}/\|D_{S_1}h_1^{(n')}\|)\|\to 0$ which implies $\|L\|=1$.

Conversely suppose $\|L\|=1$ (and $1>\theta$). By hypothesis $D_{S_1}^{-1}$ exists. Then we can choose a sequence $\{h_1^{(n)}\}$ such that $\|D_{S_1}h_1^{(n)}\|=1$, and $\|D_L D_{S_1}h_1^{(n)}\|\to 0$. Set $h_2^{(n)}:=D_{S_2}^{-1}S_2^* L D_{S_1}h_1^{(n)}$ (note $D_{S_2}^{-1}$ exists). Then clearly

$$\|D_{S_2}h_2^{(n)}-S_2^* L D_{S_1}h_1^{(n)}\|^2+\|D_L D_{S_1}h_1^{(n)}\|^2\to 0$$

and hence $D_S h^{(n)}\to 0$ where

$$h^{(n)}=\begin{bmatrix} h_1^{(n)} \\ h_2^{(n)}\end{bmatrix}.$$

To complete the proof therefore we need only show $\|h^{(n)}\|\geqq M>0$ for fixed positive constant M for all n. But clearly

$$\|h^{(n)}\|\geqq\|h_1^{(n)}\|\geqq\frac{1}{\|D_{S_1}\|}. \qquad \Box$$

Remark 2.2. For results related to (2.1) see [5] in which arbitrary block 2×2 matrices are considered.

So far we have been considering results about general contractions. In point of fact however, for our purposes the contractions we will need have a special form.

More precisely, let $m_1, m_2\in H^\infty(D)$ be inner functions. Let $H_i:=H^2\ominus m_i H^2$ $i=1$, 2 and set $H:=H^2\ominus m_1 m_2 H^2$ (where throughout this section $H^2:=H^2(D)$). We denote by T the *compression* (i.e. projection) of the unilateral shift on H^2 (defined by multiplication by z) to H. (Recall $T:=\Pi M_z|_H$ where $M_z: H^2\to H^2$ denotes multiplication by z and $\Pi: H^2\to H$ is the orthogonal projection.)

Next we have that

$$H^2\ominus m_1 m_2 H^2=(H^2\ominus m_1 H^2)\oplus(m_1 H^2\ominus m_1 m_2 H^2)$$
$$\cong H_1\oplus H_2.$$

Note that by abuse of notation, the direct sum symbol in $H_1 \oplus m_1 H_2$ stands for "orthogonal direct sum," while the direct sum symbol in $H_1 \oplus H_2$ stands for "external direct sum." (See [19] for the relevant definitions.)

Moreover, we have the following.

LEMMA 2.3. *$m_1 H^2 \ominus m_1 m_2 H^2$ is an invariant subspace of $H^2 \ominus m_1 m_2 H^2$ with respect to T.*

Proof. Let $v \in m_1 H^2 \ominus m_1 m_2 H^2$. Set $m = m_1 m_2$. Clearly $\bar{m} v \perp H^2$. Let $v_{-1} = \langle \bar{m} v, \bar{z} \rangle$. Then it is easy to compute that

$$Tv = zv - v_{-1} m.$$

But then Tv is divisible by m_1, i.e., $Tv \in m_1 H^2 \ominus m_1 m_2 H^2$. □

Lemma 2.3 means that if we identify $H_2 \cong m_1 H^2 \ominus m_1 m_2 H^2$, then we can regard H_2 as an invariant subspace of H with respect to T. Thus with these identifications, we can write

$$T = \begin{bmatrix} T_1 & 0 \\ X & T_2 \end{bmatrix}$$

where the T_i are defined as above. Clearly T_1, T_2, T are contractions.

Now in this case it is well known [24] that D_{T_1}, $D_{T_2^*}$ are of rank 1. Indeed we can compute that

$$I - T_1^* T_1 = \mu_1 \otimes \mu_1, \qquad I - T_2 T_2^* = \mu_{2*} \otimes \mu_{2*}$$

where

$$\mu_1 := \bar{z}(m_1(z) - m_1(0)); \qquad \mu_{2*} := 1 - m_2(z)\overline{m_2(0)}$$

and where $(x \otimes y) w := \langle w, y \rangle x$.

For such T_1 and T_2, it is easy to compute X.

PROPOSITION 2.4. $X = \mu_{2*} \otimes \mu_1$.

Proof. Since T is a contraction, by (2.1) we can write $X = D_{T_2^*} L D_{T_1}$ where $L: \mathscr{D}_{T_1} \to \mathscr{D}_{T_2^*}$ is a contraction between the corresponding defect spaces. Set

$$\hat{\mu}_1 := \frac{\mu_1}{\|\mu_1\|} \quad \text{and} \quad \hat{\mu}_{2*} := \frac{\mu_{2*}}{\|\mu_{2*}\|}$$

(note that $\|\mu_1\|^2 = 1 - |m_1(0)|^2$ and $\|\mu_{2*}\|^2 = 1 - |m_2(0)|^2$). Then $L: \mathscr{D}_{T_1} \to \mathscr{D}_{T_2^*}$ is such that $L\hat{\mu}_1 = \lambda \hat{\mu}_{2*}$ for some constant λ (since the defect spaces are one-dimensional). Hence, using the facts that $D_{T_1} = \|\mu_1\|(\hat{\mu}_1 \otimes \hat{\mu}_1)$, $D_{T_2^*} = \|\mu_{2*}\|(\hat{\mu}_{2*} \otimes \hat{\mu}_{2*})$, and $X = D_{T_2^*} L D_{T_1}$, we get that $X = \lambda \mu_{2*} \otimes \mu_1$. Note that since T is a contraction $|\lambda| \leqq 1$. We have still not used the fact that T is the compression to $H^2 \ominus m_1 m_2 H^2$ of the shift. We do this now.

Indeed we can apply T to μ_1. It is easy to compute that

$$\begin{aligned} T\mu_1 &= z\mu_1 - \overline{m_2(0)}(1 - |m_1(0)|^2) m_1 m_2 \\ &= z\mu_1 - (1 - |m_1(0)|^2) m_1 \\ &\quad + (1 - |m_1(0)|^2) m_1 + \overline{m_2(0)}(1 - |m_1(0)|^2) m_1 m_2. \end{aligned}$$

Under our isomorphisms,

$$\begin{aligned} H^2 \ominus m_1 m_2 H^2 &= (H^2 \ominus m_1 H^2) \oplus (m_1 H^2 \ominus m_1 m_2 H^2) \\ &\cong (H^2 \ominus m_1 H^2) \oplus (H^2 \ominus m_2 H^2), \end{aligned}$$

we can write

$$T\mu_1 = \begin{bmatrix} z\mu_1 - (1-|m_1(0)|^2)m_1(z) \\ (1-|m_1(0)|^2)(1-\overline{m_2(0)}m_2(z)) \end{bmatrix}.$$

Finally it is easy to compute that

$$\begin{bmatrix} T_1 & 0 \\ \lambda\mu_{2*}\otimes\mu_1 & T_2 \end{bmatrix}\begin{bmatrix} \mu_1 \\ 0 \end{bmatrix} = \begin{bmatrix} z\mu_1 - (1-|m_1(0)|^2)m_1(z) \\ \lambda(1-|m_1(0)|^2)(1-\overline{m_2(0)}m_2(z)) \end{bmatrix}.$$

Thus $\lambda = 1$ as required. □

3. Weighted sensitivity minimization. In this section we explicitly solve the weighted sensitivity minimization problem for L^∞-plants of the form $e^{-hs}P_0(s)$, where $P_0(s)$ is a real rational proper function with no poles or zeros on the $j\omega$-axis. Actually our procedure does much more. Basically for given weight $W(s)$ (with the hypotheses discussed in § 1), we give a technique for solving the weighted H^∞-minimization problem for a plant $P(s) = P_1(s)P_2(s)$ in terms of data determined independently by $P_1(s)$, and independently by $P_2(s)$. Our method only depends on one knowing the maximum of the optimal sensitivities of $P_1(s)$ and $P_2(s)$, and from this one can find the optimal sensitivity for $P(s)$.

As in [9], for simplicity we initially will take a weight of the form

$$W(s) = \frac{qs+r}{ms+n}$$

stable with stable inverse, and such that $\|W(s)\|_\infty \leqq 1$. In § 3.8 we will explain how our method immediately applies to general real rational weights. Moreover we will assume that $P(s)$ is proper and stable with no zeros on the $j\omega$-axis. Again in Remarks 3.10 we show how to extend our method to unstable plants. The example to keep in mind is $P(s) = e^{-hs}P_0(s)$ where $P_0(s)$ is a stable proper plant with no zeros on the $j\omega$-axis. However, the technique we give applies much more generally.

Let $\phi: H \to D$ be a fixed conformal equivalence. Set

$$\hat{W}(z) = W(\phi^{-1}(z)), \qquad \hat{P}(z) = P(\phi^{-1}(z)).$$

Let $\hat{P}_i(z)$ be the inner part of $\hat{P}(z)$. Then we assume $\hat{P}_i(z) = m_1(z)m_2(z)$ where the $m_i(z)$ are inner functions. As in § 2, set $(H^2 := H^2(D))$:

$$H := H^2 \ominus m_1 m_2 H^2,$$

$$H_i := H^2 \ominus m_i H^2, \qquad i = 1, 2,$$

$$T := \text{compression of the unilateral shift on } H^2(D) \text{ to } H.$$

Then if we make the identifications

$$\begin{aligned} H &= H^2 \ominus m_1 m_1 H^2 \\ &= (H^2 \ominus m_1 H^2) \oplus (m_1 H^2 \ominus m_1 m_2 H^2) \\ &\cong H_1 \oplus H_2 \end{aligned}$$

we can regard H_2 as an invariant subspace of H with respect to T.

When $P(s) = e^{-hs}P_0(s)$ as above, we can take $m_1(z)$ to be the Blaschke product in D whose zeros consist of the images under ϕ of the nonminimum phase zeros of $P_0(s)$, and $m_2(z) = e^{-h\phi^{-1}(z)}$.

Then following the notation of § 1 and the constructions of [26], [13], [12] the problem of computing the optimal sensitivity

$$\inf_{C \text{ stabilizing}} \|W(1+PC)^{-1}\|_\infty$$

can be reduced to computing

$$\mu := \inf_{q\in H^\infty(D)} \|\hat{W}(z) - m_1(z)m_2(z)q(z)\|_\infty.$$

Remark 3.1. We should note that the existence of a $q(z)$, achieving the infimum μ for the given $m(z) = m_1(z)m_2(z)$ inner as above, only depends on the hypothesis that $\hat{W}(z) \in H^\infty(D)$. See [14], [20].

Now as in § 2, we have

$$T = \begin{bmatrix} T_1 & 0 \\ X & T_2 \end{bmatrix}$$

relative to the decomposition $H \cong H_1 \oplus H_2$. If we write

$$\hat{W}(z) = \frac{\alpha z + \beta}{\gamma z + \delta}$$

where $\Delta := \alpha\delta - \beta\gamma \neq 0$, then $\|\hat{W}(z)\|_\infty \leqq 1$ since we assumed $\|W(s)\|_\infty \leqq 1$. Moreover without loss of generality we may clearly assume $\|\hat{W}(z)\|_\infty = 1$. (Indeed, if necessary, we can always replace $\hat{W}(z)$ by $\hat{W}(z)/\|\hat{W}(z)\|_\infty$.) Thus

$$\|\hat{W}(T)\| \leqq 1, \quad \|\hat{W}(T_1)\| \leqq 1, \quad \|\hat{W}(T_2)\| \leqq 1.$$

Moreover it is easy to compute that

$$\hat{W}(T) = \begin{bmatrix} \hat{W}(T_1) & 0 \\ \Delta(\gamma T_2 + \delta)^{-1} X (\gamma T_1 + \delta)^{-1} & \hat{W}(T_2) \end{bmatrix}.$$

Now it is well known (see [20], [23], [24]) that the infimum

$$\inf_{q\in H^\infty(D)} \|\hat{W}(z) - m_1(z)m_2(z)q(z)\|_\infty = \|\hat{W}(T)\|,$$

and what we will do now is give an explicit procedure for computing the latter norm in terms of data determined separately by the $\hat{W}(T_1)$ and $\hat{W}(T_2)$ parts of $\hat{W}(T)$. In effect we will decouple these in order to compute $\|\hat{W}(T)\|$. First note, however, that

$$\|\hat{W}(T)\| \leqq \|\hat{W}(z)\|_\infty = 1,$$

$$\|\hat{W}(T)\| \geqq \theta := \max\{\|\hat{W}(T_1)\|, \|\hat{W}(T_2)\|\}$$

and so $\theta = 1$ implies that $\|\hat{W}(T)\| = 1$. Therefore we can clearly assume $\theta < 1$.

Using the defect operator notation of § 2 (as well as the functions μ_1 and μ_{2*}), define for $j = 1, 2$ and $\rho \in (0, 1]$ such that $\rho > \theta$

$$\mu_1^{(j)} = D^{-j}_{\hat{W}(T_1)/\rho}(\bar{\gamma}T_1^* + \bar{\delta})^{-1}\mu_1,$$

$$\mu_{2*}^{(j)} = D^{-j}_{\hat{W}(T_2)^*/\rho}(\gamma T_2 + \delta)^{-1}\mu_{2*}.$$

We can now state (finally!) the following key result.

THEOREM 3.2. *With the above notation* $\|\hat{W}(T)\| \leqq \rho$ *if and only if*

(2) $$\langle(\bar{\gamma}T_2^* + \bar{\delta})^{-1}\mu_{2*}^{(2)}, \mu_{2*}\rangle \cdot \langle(\gamma T_1 + \delta)^{-1}\mu_1^{(2)}, \mu_1\rangle \leqq \rho^2\Delta^{-2}.$$

Moreover $\|\hat{W}(T)\| = \rho$ *if and only if* (2) *is an equality.* (*Note we are assuming* $\rho \in (0, 1]$ *is such that* $\rho > \theta$.)

Proof by Theorem 2.1. $\|\hat{W}(T)\| \leqq \rho$ if and only if

$$\frac{1}{\rho}\Delta(\gamma T_2 + \delta)^{-1} X(\gamma T_1 + \delta)^{-1} = D_{\hat{W}(T_2)^*/\rho} L_\rho D_{\hat{W}(T_1)/\rho}$$

where

$$L_\rho : \mathcal{D}_{\hat{W}(T_1)/\rho} \to \mathcal{D}_{\hat{W}(T_2)^*/\rho}$$

defines a contraction of the corresponding defect spaces. But then it is easy to compute that

$$L_\rho = \frac{1}{\rho}\Delta\mu_{2*}^{(1)} \otimes \mu_1^{(1)}.$$

Indeed this follows immediately from the definition of the $\mu_{2*}^{(1)}$ and $\mu_1^{(1)}$ once we show that

$$\mu_{2*}^{(1)} \otimes \mu_1(\gamma T_1 + \delta)^{-1} D_{\hat{W}(T_1)/\rho}^{-1} = \mu_{2*}^{(1)} \otimes D_{\hat{W}(T_1)/\rho}^{-1}(\bar{\gamma} T_1^* + \bar{\delta})^{-1}\mu_1.$$

But to see this just apply the first operator to an element ψ. We get

$$\begin{aligned}\mu_{2*}^{(1)} \otimes \mu_1(\gamma T_1 + \delta)^{-1} D_{\hat{W}(T_1)/\rho}^{-1}\psi &= \langle(\gamma T_1 + \delta)^{-1} D_{\hat{W}(T_1)/\rho}^{-1}\psi, \mu_1\rangle\mu_{2*}^{(1)} \\ &= \langle\psi, D_{\hat{W}(T_1)/\rho}^{-1}(\bar{\gamma} T_1^* + \bar{\delta})^{-1}\mu_1\rangle\mu_{2*}^{(1)}\end{aligned}$$

since $D_{\hat{W}(T_1)/\rho}^{-1}$ is self-adjoint.

Therefore

$$\begin{aligned}\|L_\rho\| \leqq 1 &\Leftrightarrow \|\mu_{2*}^{(1)}\|\,\|\mu_1^{(1)}\| \leqq \rho\Delta^{-1} \\ &\Leftrightarrow \text{the inequality (2) holds.}\end{aligned}$$

Finally, under the assumption that $\rho > \theta$, by (2.1) $\|\hat{W}(T)\| = \rho$ if and only if $\|L_\rho\| = 1$ if and only if (2) is an equality. □

Remarks 3.3. (i) In case $m_1(z) = (z - a)/(1 - \bar{a}z)$, $|a| < 1$, (2) is equivalent to certain inequalities derived by Adamjan, Arov and Krein [1], [2] in connection with the one-step extension problem. Hence what we have derived here is an expression for the norm of an "n-step extension" (in case m_1 is a finite Blaschke product), or even an "∞-step extension" (e.g., when m_1 is an infinite Blaschke product).

(ii) We will assume from now on that $\rho \geqq \|\hat{W}(T)\|$, and $\rho > \theta$. Note that in our procedure below, we can compute $\|\hat{W}(T)\|$ explicitly once we know θ. Thus if we can find the optimal sensitivity for plants $P_1(s)$, $P_2(s)$ we can find it for $P(s) = P_1(s)P_2(s)$.

We now come to the crucial question of how to compute the inner products of (2). Again we can give an explicit procedure.

3.4. *Computation of inner products.* We will start with the computation of

$$\langle(\bar{\gamma} T_2^* + \bar{\delta})^{-1}\mu_{2*}^{(2)}, \mu_{2*}\rangle.$$

Set $\nu_* := (\bar{\gamma} T_2^* + \bar{\delta})^{-1}\mu_{2*}^{(2)}$. Since $\mu_{2*} = 1 - m_2(z)\overline{m_2(0)}$, and since $\nu_* \in H^2 \ominus m_2 H^2$, we have that $\langle\nu_*, \mu_{2*}\rangle = \nu_*(0)$. Thus we must show how to find $\nu_*(0)$. We give a simple algebraic procedure for doing this.

First note that

$$(\gamma T_2 + \delta)^{-1}\mu_{2*} = \left(1 - \frac{1}{\rho^2}\hat{W}(T_2)\hat{W}(T_2)^*\right)\mu_{2*}^{(2)}.$$

Therefore

$$\begin{aligned}\mu_{2*} &= \left[(\gamma T_2+\delta)(\bar{\gamma}T_2^*+\bar{\delta})-\frac{1}{\rho^2}(\alpha T_2+\beta)(\bar{\alpha}T_2^*+\bar{\beta})\right]\nu_* \\ &= (A+BT_2+\bar{B}T_2^*+CT_2T_2^*)\nu_*\end{aligned} \tag{3}$$

where

$$A:=|\delta|^2-\left(\frac{1}{\rho^2}\right)|\beta|^2,$$

$$B:=\left(\gamma\bar{\delta}-\left(\frac{1}{\rho^2}\right)\alpha\bar{\beta}\right),$$

$$C:=\left(|\gamma|^2-\left(\frac{1}{\rho^2}\right)|\alpha|^2\right).$$

Now $(1-T_2T_2^*)\nu_*=\langle\nu_*,\mu_{2*}\rangle\mu_{2*}=\nu_*(0)\mu_{2*}$. Therefore, from (3) we see that

$$(1+C\nu_*(0))\mu_{2*}=(F+BT_2+\bar{B}T_2^*)\nu_* \tag{4}$$

where $F:=A+C$. But $\bar{m}_2\nu_*\perp H^2$ so we can write $\bar{m}_2\nu_*=\nu_{-1}\bar{z}+\nu_{-2}\bar{z}^2+\cdots$. Then

$$T_2^*\nu_*=\bar{z}(\nu_*-\nu_*(0)),\qquad T_2\nu_*=z\nu_*-m_2\nu_{-1}.$$

Consequently, from (4) we see that

$$(1+C\nu_*(0))\mu_{2*}+\bar{B}\bar{z}\nu_*(0)+Bm_2\nu_{-1}=(F+Bz+\bar{B}\bar{z})\nu_*. \tag{5}$$

Finally, multiplying both sides of (4) by z and rearranging terms, we derive the following key relationship:

$$(C\mu_{2*}z+\bar{B})\nu_*(0)+Bm_2z\nu_{-1}=(Bz^2+Fz+\bar{B})\nu_*-\mu_{2*}z. \tag{6}$$

(Note that even though this relationship has been derived on the boundary on D, since all the functions are in $H^2(D)$, they can be analytically continued to D.)

We are almost done! Indeed it is easy to see that the roots z_1, z_2 of $Bz^2+Fz+\bar{B}$ are such that $|z_1z_2|=1$. If $B=\bar{B}$ is real (which always occurs in cases of interest in engineering) $z_1z_2=1$. We can always assume $|z_1|\leqq 1$. We have three cases.

CASE (i). $|F|>2|B|$. Then $z_1\in D$, $z_2=1/\bar{z}_1$. Now multiply (5) by $\bar{m}_2$ to get

$$(C\bar{m}_2\mu_{2*}z+\bar{B}\bar{m}_2)\nu_*(0)+Bz\nu_{-1}=(Bz^2+Fz+\bar{B})\bar{m}_2\nu_*-\bar{m}_2\mu_{2*}z. \tag{7}$$

Note that $\bar{m}_2\mu_{2*}$ and $\bar{m}_2\nu_*$ can be continued analytically in the complement of the unit disc and are 0 at ∞. (On the boundary of D we identify $\bar{z}$ and $1/z$.)

Then plugging z_1 into (6) and z_2 into (7) we get

$$(C\mu_{2*}(z_1)z_1+\bar{B})\nu_*(0)+Bm_2(z_1)z_1\nu_{-1}=-\mu_{2*}(z_1)z_1, \tag{8}$$

$$(C(\bar{m}_2\mu_{2*})(z_2)\cdot z_2+\bar{B}m_2(z_2))\nu_*(0)+Bz_2\nu_{-1}=-(\bar{m}_2\mu_{2*})(z_2)\cdot z_2. \tag{9}$$

Using the fact that $z_2=1/\bar{z}_1$, one can solve these equations for $\nu_*(0)$ and show

$$\frac{1}{|\nu_*(0)|}=\left|\frac{A-C}{2}+\frac{1}{2}\sqrt{F^2-4|B|^2}\frac{1+|m_2(z_1)|^2}{1-|m_2(z_1)|^2}\right|. \tag{10}$$

Note that the case in which $B=0$ is a limiting case of Case (i) in which $z_1=0$, $z_2=\infty$. When this occurs one can compute

$$\frac{1}{|\nu_*(0)|}=\left|\frac{A+C|m_2(0)|^2}{1-|m_2(0)|^2}\right|. \tag{11}$$

Before stating Cases (ii) and (iii) we will need the following lemma.

LEMMA 3.5. *Let z_1 be such that $Bz_1^2+Fz_1+\bar{B}=0$, and such that $|z_1|=1$. Then $m_2(z)$ admits an analytic extension to a neighborhood of z_1, and $|m_2(\zeta)|=1$ for all ζ in an arc neighborhood of z_1 on the unit circle.*

Proof. First we claim $z_1\notin\sigma(T_2)$ (where $\sigma(T_2)$ denotes the spectrum of the contraction T_2). Indeed if to the contrary $z_1\in\sigma(T_2)$, then $\hat{W}(z_1)\in\sigma(\hat{W}(T_2))$. But by definition, since $Bz_1^2+Fz_1+\bar{B}=0$, we have that $(1-(|\hat{W}(z_1)|^2/\rho^2))=0$, that is $|\hat{W}(z_1)|=\rho$. But this would imply that $\|\hat{W}(T_2)\|\geqq\rho$, which contradicts our assumption in Remark 3.3(ii) that $\|\hat{W}(T_2)\|<\rho$.

But since $z_1\notin\sigma(T_2)$ we get the required result from [24, Chap. III, Thm. (5.1)]. □

Remark 3.6. With the notation of (3.5), note that since $m_2(z)$ is analytic in a neighborhood of z_1, μ_{2*} must be analytic in this neighborhood of z_1, and ν_* can have at most a pole at z_1. But since $\nu_*\in H^2(D)$, in point of fact ν_* must be analytic at z_1 as well. Moreover the derivatives of these functions will also be analytic in a neighborhood of z_1, since the derivative of an analytic function is itself analytic.

We can now state Cases (ii) and (iii) (z_1 and z_2 are the roots of $Bz^2+Fz+\bar{B}$).

CASE (ii). $|F|<2|B|$ i.e. $|z_1|=|z_2|=1$, $z_1\neq z_2$. In this case plug the z_i $i=1,2$ into (6) to get two linear equations (one of which will be (7), and the other (7) with z_2 substituted for z_1) in the two unknowns $\nu_*(0)$, ν_{-1} and solve for $\nu_*(0)$. By (3.5) and (3.6) this is valid since the functions m_2, μ_{2*}, ν_* are analytic in neighborhoods of z_1 and z_2.

We can then compute that

$$\frac{1}{|\nu_*(0)|}=\left|\frac{A-C}{2}+\frac{j\sqrt{4|B|^2-F^2}}{2}\cdot\frac{1+m_2(z_1)\overline{m_2(z_2)}}{1-m_2(z_1)\overline{m_2(z_2)}}\right|. \tag{12}$$

(When $m_2(z_1)=m_2(z_2)$, $z_1\neq z_2$, it is easy to show that $\nu_*(0)=0$.)

CASE (iii). $|F|=2|B|$, i.e. $z_1=z_2$. Then plug z_1 into (6), and z_2 into the derivative of (6). Once more by (3.5) and (3.6) this makes sense, and we can solve the two resulting equations in the two unknowns $\nu_*(0)$, ν_{-1} for $\nu_*(0)$.

Making the computation, we get that

$$\frac{1}{|\nu_*(0)|}=\left|-C+\varepsilon|B|-B\frac{m_2(z_1)}{m_2'(z_1)}\right| \tag{13}$$

where

$$\varepsilon=\begin{cases}1 & \text{if } F>0,\\ -1 & \text{if } F<0,\end{cases}$$

for $m_2'(z_1)\neq 0$. When $m_2'(z_1)=0$, it is easy to show that $\nu_*(0)=0$.

In short from (6), using simple linear algebra, we can find $\nu_*(0)$, the value of the first inner product. Notice that Cases (i) and (ii) are generic, while Case (iii) is the nongeneric case in this situation.

Next we come to the computation of the second inner product of (2), namely

$$\langle(\gamma T_1+\delta)^{-1}\mu_1^{(2)},\mu_1\rangle.$$

We will propose two methods for doing this. The first works for any inner function $m_1(z)$, and the second for a finite Blaschke product.

The first method is simply to imitate the procedure that we used previously in evaluating $\langle \nu_*, \mu_{2*}\rangle$. Indeed, set

$$\nu := (\gamma T_1 + \delta)^{-1}\mu_1^{(2)}.$$

Then we want to evaluate $\langle \nu, \mu_1\rangle$. But

$$\begin{aligned}\langle \nu, \mu_1\rangle &= \langle \nu, \bar{z}(m_1(z) - m_1(0))\rangle\\ &= \langle \bar{m}_1\nu, \bar{z}\rangle.\end{aligned}$$

Since $\bar{m}_1\nu \perp H^2$, we may write $\bar{m}_1\nu = \hat{\nu}_{-1}\bar{z} + (\text{higher order terms in } \bar{z})$ and so $\langle \nu, \mu_1\rangle = \hat{\nu}_{-1}$. Playing the same game as above we end up with the following analogue of equation (6):

$$(C\mu_1 z + Bm_1 z)\hat{\nu}_{-1} + \bar{B}\nu(0) = (Bz^2 + Fz + \bar{B})\nu - \mu_1 z. \tag{14}$$

We again divide the analysis of (14) into the identical Cases (i), (ii) and (iii) depending upon the roots of $Bz^2 + Fz + \bar{B}$ from which we derive analogous formulae for $\hat{\nu}_{-1}$ (the required value of $\langle \nu, \mu_1\rangle$) to those we found above for $\nu_*(0)$.

We should note that a deeper explanation of the analogy between (6) and (14) can be given via a beautiful result from [21]. In point of fact using this result it is possible to write down (14) immediately from (6) and the analogous formulae to those of (10)-(13) for $\hat{\nu}_{-1}$ just by inspection. However since these formulae may be derived by elementary linear algebra as above, we will leave it to the interested reader to consult [21].

The second method for finding $\hat{\nu}_{-1}$ works when $m_1(z)$ is a finite Blaschke product. In this case, $H^2 \ominus m_1 H^2$ is finite dimensional and it is easy to compute a basis for this space (see e.g. [21], [17]). Therefore the computation of the second inner product of (2) amounts to finite matrix operations once a suitable basis is chosen. For example, if

$$m_1(z) = \prod_{i=1}^{n}\left(\frac{z - a_i}{1 - \bar{a}_i z}\right)$$

with $a_i \neq a_j$ for $i \neq j$, then the elements

$$v_k := \frac{(1 - |a_k|^2)^{1/2}}{1 - \bar{a}_k z}\prod_{i=1}^{k-1}\left(\frac{z - a_i}{1 - \bar{a}_i z}\right)$$

for $k = 1, \cdots, n$ form a unitary basis for $H^2 \ominus m_1 H^2$ relative to which all the relevant linear operators may be given a finite matrix form. For Blaschke products (in the unit disc) which have roots with multiplicities, it is again easy and standard to write down a similar unitary basis (see [20], [10], [17]).

We thus have an explicit procedure for computing the inner products (3.2). We now give an explicit algorithm for the computation of the optimal sensitivity.

3.7. *Computation of optimal sensitivity.* We will use the notation of (3.4). Note moreover that the computation of $\nu_*(0)$ and $\hat{\nu}_{-1}$ as functions of ρ divide into the identical Cases (i), (ii) and (iii) depending on the roots $Bz^2 + Fz + \bar{B}$.

To make the dependence of ρ explicit, let us set

$$\nu_1(\rho) := \nu_*(0), \qquad \nu_2(\rho) := \hat{\nu}_{-1}.$$

Then (3.2) reads

$$\nu_1(\rho)\nu_2(\rho) \leqq \rho^2\Delta^{-2}. \tag{15}$$

Let us now recall some of our assumptions:

(a) $\hat{W}(z)=(\alpha z+\beta)/(\gamma z+\delta)\in H^{\infty}(D)$, $\hat{W}^{-1}(z)$ has no poles in D, the open unit disc. (This follows since we assumed $W(s)\in H^{\infty}(H)$ with stable inverse.) We should note that our methods immediately go through without the hypothesis that $\hat{W}^{-1}(z)$ has no poles in D, but we retain it since it will be easier to explain our algorithm this way.

(b) $\hat{W}(z)$ is normalized so that $\|\hat{W}(z)\|_{\infty}=1$. (Again this can be done without loss of generality, by replacing if necessary $\hat{W}(z)$ by $\hat{W}(z)/\|\hat{W}(z)\|_{\infty}$. We make this normalization since it will be a bit easier to state our algorithm this way. Of course, one can easily write down a similar algorithm without such a normalization.)

(c) $\rho\in(0,1]$ is such that $\rho\geqq\|\hat{W}(T)\|$, and $\rho>\theta:=\max\{\|\hat{W}(T_1)\|, \|\hat{W}(T_2)\|\}$. Note that for the algorithm to work we must know θ.

Here then is our algorithmic procedure for the computation of $\|\hat{W}(T)\|$ and hence the optimal sensitivity. We consider two cases.

(A) $|\beta/\alpha|=1$. Then the algorithm is as follows:

(i) We first consider Case (iii) of (3.4), i.e. $|F|^2=4|B|^2$. Regarding this as an equation in $\rho\in(0,1]$, it is easy to see that the unique solution will be $\rho=1$. (Just consider the locus

$$\{z: \rho^2-|W(z)|^2=0\}$$

and notice that there exists $z_0\in\partial D$ such that $\hat{W}(z_0)=0$. See Fig. 2.)

We now check if $\rho=1$ gives equality for (15) using the Case (iii) formulae of (3.4) ((13) and the analogous formula for $\nu_2(\rho)$). If we do get equality, then by Theorems 2.1 and 3.2, $\|\hat{W}(T)\|=1$, and the algorithm terminates. If not, i.e. if we get strict inequality, we go to step (ii).

(ii) If $\rho<1$, then it is easy to check we are in Case (ii) of (3.4). (See Fig. 2.) Using the formulae we derived for Case (ii) ((12) and the analogous formula for $\nu_2(\rho)$), we check if there exists $\rho\in(0,1)$ with $\rho>\theta$ which gives equality in (15). If there exists such a solution, say ρ_1, then by Theorems 2.1 and 3.2 it is unique and $\|\hat{W}(T)\|=\rho_1$, i.e. the algorithm terminates. If not, i.e. if we get strict inequality for all $\rho\in(0,1)$ with $\rho>\theta$, we go to step (iii).

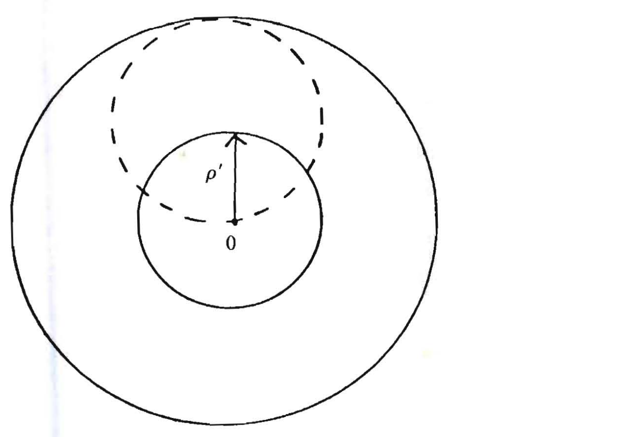

FIG. 2. *Representation of the case $|\beta/\alpha|=1$. Both solid circles are centered at 0, the larger being ∂D, the unit circle. The dashed circle represents the locus $\hat{W}(\partial D)$. Since $\|\hat{W}(z)\|_{\infty}=1$, $\hat{W}^{-1}(z)$ has no poles in D, and $|\beta/\alpha|=1$, $\hat{W}(\partial D)$ passes through the origin 0 and is tangent to ∂D. Note that the circle of radius ρ' intersects $\hat{W}(\partial D)$ in two points, i.e. for any $0<\rho'<1$ we are in Case* (ii). *When $\rho'=1$, we are in Case* (iii).

(iii) If steps (ii) and (iii) fail to find the norm, then from our hypotheses and Theorems 2.1 and 3.2, we have

$$\| \hat{W}(T) \| = \theta = \max \{ \| \hat{W}(T_1) \|, \| \hat{W}(T_2) \| \}$$

and once more we are done.

This completes the analysis of case (A).

(B) $|\beta/\alpha| \neq 1$. Then the algorithmic procedure for finding $\| \hat{W}(T) \|$ is as follows:

(i) As in (A), we first consider Case (iii) of (3.4), that is $|F|^2 = 4|B|^2$. Regarding this as an equation in $\rho \in (0, 1]$ and with the above hypotheses (a), (b), (c) one can easily show that we get precisely two solutions, namely $\rho = 1$, and a unique $0 < \rho_0 < 1$. (Again to see this, just consider the locus $\{\rho^2 - |\hat{W}(z)|^2 = 0\}$ and examine the cases $|\beta/\alpha| < 1$, $|\beta/\alpha| > 1$. See Figs. 3 and 4 below.)

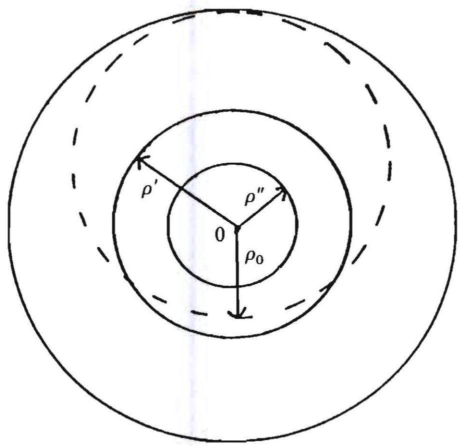

FIG. 3. *Representation of the case $|\beta/\alpha| < 1$. All three solid circles are centered at 0, the largest being ∂D, the unit circle. The dashed circle represents the locus $\hat{W}(\partial D)$. Since $\| \hat{W}(z) \|_{\infty} = 1$, and $\hat{W}^{-1}(z)$ has no poles in D, $\hat{W}(\partial D)$ is tangent to ∂D. ρ_0 is the distance of 0 to the closest point on $\hat{W}(\partial D)$. Note that the circle of radius ρ' intersects $\hat{W}(\partial D)$ in two points, i.e. for $\rho_0 < \rho' < 1$ we are in Case* (ii). *For $0 < \rho'' < \rho_0$ we are in Case* (i). *For $\rho' = 1$, or $\rho' = \rho_0$, we are in Case* (iii).

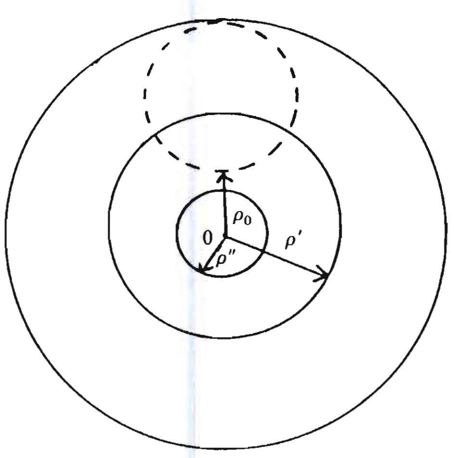

FIG. 4. *Representation of the case $|\beta/\alpha| > 1$. Same explanation as for Fig. 3, except here the origin 0 lies to the exterior of $\hat{W}(\partial D)$ (which is represented by the dashed circle).*

We now check if $\rho = 1$ gives equality for (15) using the Case (iii) formulae of (3.4) and if it does, then as before $\|\hat{W}(T)\| = 1$. If we get strictly inequality, we consider ρ_0. If $\rho_0 > \theta$, and if it gives equality when substituted in (15) (using the Case (iii) formulae), then by Theorems 2.1 and 3.2, $\|\hat{W}(T)\| = \rho_0$. If not, we go to step (ii).

(ii) If from step (i) we have failed to find the required norm, we consider now ρ such that $\rho > \theta$ and $\rho_0 < \rho < 1$. Then it is easy to check that we will be in Case (ii) of (3.4) (see Figs. 3 and 4). If we can find such a ρ, say ρ_2, which gives equality in (15) (using the Case (ii) formulae of (3.4)), then by Theorems 2.1 and 3.2 ρ_2 will be unique and $\|\hat{W}(T)\| = \rho_2$, and so the algorithm terminates. If all such ρ with $\rho > \theta$ and $\rho_0 < \rho < 1$ give strict inequality, we go to step (iii).

(iii) We consider ρ such that $\rho > \theta$ and $0 < \rho < \rho_0$. (Of course we need $\rho_0 > \theta$ in this step. If not, just go to step (iv).) Then one can easily check that we will be in Case (i) of (3.4) (see Figs. 3 and 4). If we can find such a ρ, say ρ_3, which gives equality in (15) (using the Case (i) formulae of (3.4)), then by Theorems 2.1 and 3.2 ρ_3 will be unique and $\|\hat{W}(T)\| = \rho_3$, i.e., we are done. If all such ρ with $\rho > \theta$ and $0 < \rho < \rho_0$ give strict inequality, we go to step (iv).

(iv) If in all three steps above we have failed to find the norm, then by Theorems 2.1 and 3.2 and the above hypotheses

$$\|\hat{W}(T)\| = \theta = \max\{\|\hat{W}(T_1)\|, \|\hat{W}(T_2)\|\}$$

and once again the algorithm terminates.

In short, (3.7) gives an easily computable algorithm for finding $\|\hat{W}(T)\|$ once we know θ. Thus we have a technique for computing the H^∞-optimal sensitivity for distributed systems like $e^{-hs}P_0(s)$, $P_0(s)$ rational stable, since we know the optimal sensitivities for e^{-hs} and $P_0(s)$ already. We now will discuss what occurs for more general weights.

3.8. *General weights and one-step extensions.* The above analysis was made for linear weights. Still keeping our assumptions on $P(s)$ (i.e. $P(s)$ is stable, proper, with no zeros on the $j\omega$-axis), we would like to explicitly show how our methods carry over for a general real rational weight $W(s)$, $W(s) \in H^\infty$ with stable inverse, $\|W(s)\|_\infty \leqq 1$. Using the above conformal equivalence $\phi: H \to D$ we set as before

$$\hat{W}(z) := W(\phi^{-1}(z))$$

and we write $\hat{W}(z) = p(z)/q(z)$ a ratio of relatively prime polynomials in z.

Then given as above that

$$T = \begin{bmatrix} T_1 & 0 \\ X & T_2 \end{bmatrix}$$

it is easy to compute that

$$\hat{W}(T) = \begin{bmatrix} \hat{W}(T_1) & 0 \\ q^{-1}(T_2)r(X)q^{-1}(T_1) & \hat{W}(T_2) \end{bmatrix}$$

where $r(X)$ has the form

$$r(X) = \sum_{0 \leqq j,k \leqq n-1} a_{jk} T_2^j \mu_{2*} \otimes T_1^{*k} \mu_1 \tag{16}$$

for some constants a_{jk}, and where $n = \max\{\text{degree } p(z), \text{degree } q(z)\}$.

Indeed (16) may be derived as follows: Set for $k \in N$

$$X^{(k)} := \sum_{\substack{0 \leq i, j \leq k-1 \\ i+j=k-1}} T_2^i X T_1^j.$$

Given a polynomial

$$b(z) = \sum_{k=0}^{s} b_k z^k,$$

set

$$\hat{b}(X) = \sum_{k=1}^{s} b_k X^{(k)}.$$

Notice in $\hat{b}(X)$ we have dropped the constant term. Then by direct computation, one gets that

$$r(X) = q(T_2)\hat{p}(X) - p(T_2)\hat{q}(X).$$

Equation (16) now follows from the facts that $X = \mu_{2*} \otimes \mu_1$, and that

$$T_2^j X T_1^k = T_2^j \mu_{2*} \otimes T_1^{*k} \mu_1.$$

In short, $r(X)$ is a *finite rank operator*, and is composed of tensor products of the $T_2^j \mu_{2*}$ and $T_1^{*k}\mu_1$ all of which may be explicitly computed. Hence as in the linear weight case, the computation of $\|\hat{W}(T)\|$ may be reduced to an analogous (but of course messier) algebraic problem using the procedures discussed in Theorem 3.2 and (3.4).

In the most important (from a practical point of view) special case, in which $m_1(z)$ is a finite Blaschke product, we can even get simple closed form formulae as we did above. We will do this now.

Indeed first note that when $m_1(z)$ is a finite Blaschke product, we can in point of fact always reduce ourselves to the case in which $m_1(z) = (z-a)/(1-\bar{a}z)$ for some $a \in D$. To see this let us suppose that

$$m_1(z) = \prod \left(\frac{z - a_i}{1 - \bar{a}_i z} \right).$$

Suppose moreover that we give a procedure for solving the optimal sensitivity problem for $((z-a_1)/(1-\bar{a}_1 z))m_2(z)$ in terms of (decoupled) data determined by $m_2(z)$ and $(z-a_1)/(1-\bar{a}_1 z)$ as we did in Theorem 3.2. Then we can take $\hat{m}_2(z) := ((z-a_1)/(1-\bar{a}_1 z))m_2(z)$ as our new "$m_2(z)$", and $(z-a_2)/(1-\bar{a}_2 z)$ as our new "$m_1(z)$," and solve the resulting problem for $((z-a_2)/(1-a_2 z))\hat{m}_2(z)$ in terms of $\hat{m}_2(z)$ and $(z-a_2)/(1-\bar{a}_2 z)$, and so on. In other words, when $m_1(z)$ is a finite Blaschke product in order to solve the optimal sensitivity problem, it is enough to describe the solution to the problem when we add the zeros of $m_1(z)$ one at a time. This is, of course, the basic idea behind the classical recursive procedure of Nevanlinna-Pick interpolation [14], and the one-step extension procedure of Adamjan, Arov and Krein [1], [2].

Consequently, we will give an explicit solution now of the kind we gave for a linear weight, for a general real rational weight, $\hat{W}(z) = p(z)/q(z)$, $\|\hat{W}(z)\|_\infty \leq 1$, such that $\hat{W}^{-1}$ has no poles in the unit disc, and an inner function $m(z) = m_1(z)m_2(z)$ where $m_1(z) = (z-a)/(1-\bar{a}z)$, $a \in D$. Then with this notation, $\mu_1 = (1-|a|^2)/(1-\bar{a}z)$, and $T_1\mu_1 = a\mu_1$.

From (16), we can write that

$$\begin{aligned} r(X) &= \sum_{0\le j,k\le n-1} a_{jk} T_2^j \mu_{2*} \otimes \bar{a}^k \mu_1 \\ &= \sum_{j=0}^{n-1} b_j T_2^j \mu_{2*} \otimes \mu_1 \end{aligned}$$

(where $n = \max\{\text{degree } p(z), \text{degree } q(z)\}$), for some (explicitly computable) constants b_j.

Set

$$\mu_* = \sum_{j=0}^{n-1} b_j T_2^j \mu_{2*}.$$

Then

$$r(X) = \mu_* \otimes \mu_1.$$

Therefore we have

$$\hat{W}(T) = \begin{bmatrix} \hat{W}(T_1) & 0 \\ q(T_2)^{-1}(\mu_* \otimes \mu_1) q(T_1)^{-1} & \hat{W}(T_2) \end{bmatrix}.$$

We can now play precisely the same game that we did in the linear weight case. Once more without loss of generality we can assume that

$$\theta := \max\{\|\hat{W}(T_1)\|, \|\hat{W}(T_2)\|\} < 1.$$

Let $\rho \in (0, 1]$ and suppose $\rho > \theta$. Then we set for $j = 1, 2$

$$\begin{aligned} \mu_*^{(j)} &:= D^{-j}_{\hat{W}(T_2)^*/\rho}(q(T_2)^{-1})\mu_*, \\ \mu_1^{(j)} &:= D^{-j}_{\hat{W}(a)/\rho}\overline{(q(a))}^{-1}\mu_1 \end{aligned}$$

(since T_1 is multiplication by a).

Then the analogue of Theorem 3.2 in this case is the next theorem.

THEOREM 3.9. *$\|\hat{W}(T)\| \le \rho$ if and only if*

$$\langle \bar{q}(T_2^*)^{-1}\mu_*^{(2)}, \mu_*\rangle \cdot \left\langle \frac{\mu_1^{(2)}}{q(a)}, \mu_1 \right\rangle \le \rho^2. \tag{17}$$

Moreover $\|\hat{W}(T)\| = \rho$ if and only if equality holds in (17). (We are assuming $\rho > \theta$.)

Proof. As in Theorem 3.2, $\|\hat{W}(T)\| \le \rho$ if and only if

$$\frac{1}{\rho} q(T_2)^{-1} r(X) q(T_1)^{-1} = D_{(1/\rho)\hat{W}(T_2)^*} L_\rho D_{(1/\rho)\hat{W}(T_1)}$$

for some contraction L_ρ. But it is easy to compute that

$$L_\rho = \frac{1}{\rho}(\mu_*^{(1)} \otimes \mu_2^{(1)}).$$

Therefore $\|L_\rho\| \le 1$ if and only if we have the inequality (17). The second part of the theorem follows immediately from Theorem 2.1. □

Remarks 3.10. (i) Clearly in this case the second inner product of (17) is trivial to compute. As for the first inner product, it is clear that one can use the same algebraic technique that we discussed in (3.4). Here from the roots of a polynomial of degree $2n$ one gets $2n$ linear equations in $2n$ unknowns from which one can solve for the required value of the inner product, where $n := \max\{p(z), q(z)\}$. Depending upon the multiplicities of the roots and where they lie in relation to D, one can derive a procedure analogous to that of (3.7). We did this by hand for a simple quadratic weight ($W(s) = 1/(as+1)^2$) and admittedly the computation becomes very messy. However, our procedure can certainly be programmed on computer for the kind of rational weights we have considered above.

(ii) Now we finally come to the case in which $P(s) \in L^{\infty}$ is not stable (but has no zeros on the $j\omega$-axis). This poses no problem (at least theoretically). Indeed using the arguments of [13], [27] one may reduce the sensitivity minimization problem to the problem of computing

$$\inf_{q \in H^{\infty}} \|V - Bq\|_{\infty}$$

where $B(s) =$ inner part of $P(s)$, $V \in H^{\infty}$.

Now from (3.1), with these hypotheses, the minimization problem will have a solution. If we then assume that the outer part of $P(s)$ is rational (of course we always consider rational weights that are in H^{∞}), V will be rational, and we can apply our techniques to the solution of the minimization problem. More explicitly, if $P(s) = P_1(s)P_2(s)$ and we could compute the minimal sensitivities of $P_1(s)$, $P_2(s)$, then we could use our preceding procedure in order to solve the problem for $P(s)$. This occurs for example when $P(s) = e^{-hs}P_0(s)$, $P_0(s)$ real rational and proper, $P_0 \in L^{\infty}$ with no zeros on the $j\omega$-axis.

4. An explicit example. Given the general procedures of § 3, an illustrative nontrivial example is certainly called for. We will take

$$W(s) = \frac{1}{as+1}, \qquad a > 0$$

$$P(s) = e^{-hs}\left(\frac{s-b}{s+b}\right), \qquad h, b > 0.$$

The minimum sensitivity in this case will allow us to understand the relationship among the quantities a, b, h.

So, let us plug these parameters into our machine and compute. First we choose $\phi: H \to D$ to be

$$z = \phi(s) := \frac{s-b}{s+b}.$$

Then

$$\hat{W}(z) = W(\phi^{-1}(z)) = \frac{1-z}{(ab-1)z + (ab+1)},$$

$$e^{h\phi^{-1}(z)} = e^{hb((z+1)/(z-1))}.$$

We now use the notation of (3.4). Note that $m_1(z) = z$, $m_2(z) = e^{hb((z+1)/(z-1))}$,

$$A = (ab+1)^2 - \frac{1}{\rho^2}, \qquad B = \bar{B} = ((ab)^2 - 1) + \frac{1}{\rho^2},$$

$$C = (ab-1)^2 - \frac{1}{\rho^2}, \qquad F = 2\left((ab)^2 + 1 - \frac{1}{\rho^2}\right).$$

Then in this case, the two roots of the quadratic equation $Bz^2+Fz+\bar{B}$ are

$$z_1=\frac{-((ab)^2+1-(1/\rho^2))+2abj\sqrt{(1/\rho^2)-1}}{((ab)^2-1)+(1/\rho^2)}$$

and $z_2=\bar{z}_1$. It is clear for our plant $P(s)$ that for a, b, $h>0$ and finite, that the optimal sensitivity will always be strictly less than 1. Hence we can immediately remove Case (iii) of (3.4) from our considerations. (Note $|\beta/\alpha|=1$ here. See (3.7) (A) above.)

Therefore since $|z_1|=|z_2|=1$, we are in Case (ii) of the procedure (3.4) and (3.7). Then solving the corresponding linear equations (or using the formulae of (3.4)) we get

$$\nu_*(0)=\frac{\sin\left(\dfrac{h\sqrt{(1/\rho^2)-1}}{a}\right)}{2ab\sin\left(\dfrac{h\sqrt{(1/\rho^2)-1}}{a}\right)+2ab\sqrt{(1/\rho^2)-1}\cos\left(\dfrac{h\sqrt{(1/\rho^2)-1}}{a}\right)}.$$

The second inner product is trivial to compute and turns out to be

$$\hat{\nu}_{-1}:=\frac{\rho^2}{(\rho^2(ab+1)^2-1)}.$$

Next it is trivial to compute that $\Delta=-2ab$, and therefore from (2) we see

$$\nu_*(0)\hat{\nu}_{-1}\leqq\frac{\rho^2}{4a^2b^2}.$$

Hence we get that

$$\frac{\sin\left(\dfrac{h\sqrt{(1/\rho^2)-1}}{a}\right)}{\sin\left(\dfrac{h\sqrt{(1/\rho^2)-1)}}{a}\right)+\sqrt{(1/\rho^2)-1}\cos\left(\dfrac{h\sqrt{(1/\rho^2)-1}}{a}\right)}\leqq\frac{\rho^2(ab+1)^2-1}{ab}. \tag{18}$$

Using our above notation set

$$\theta:=\max\{\|\hat{W}(T_1)\|,\|\hat{W}(T_2)\|\}$$

where T_1 is the compressed shift corresponding to $m_1(z)=z$, and T_2 is the compressed shift corresponding to $m_2(z)=e^{hb((z+1)/(z-1))}$. It is easy to compute that $\|\hat{W}(T_1)\|=1/(ab+1)$, and $\|\hat{W}(T_2)\|=\rho_1$, the largest root of (1) (of § 1), $\rho_1\in(0,1)$.

Then if we algebraically manipulate (18) and invoke Theorems 2.1 and 3.2, (3.7) (A) we see that we are required to find ρ_{opt}, the unique root contained in $(\theta,1)$ of the following equation (it is easy to check ρ_{opt} exists for a, b, $h\in(0,\infty)$):

$$\left(1-\frac{2ab}{\rho^2(ab+1)^2-1}\right)\tan\left(\frac{h\sqrt{(1/\rho^2)-1}}{a}\right)+\sqrt{\frac{1}{\rho^2}-1}=0. \tag{19}$$

By our above theory, $\|\hat{W}(T)\|=\rho_{\mathrm{opt}}$. Equation (19) has a number of interesting properties a few of which we discuss here. For example, as $b\to\infty$, (19) approaches (1) of § 1; this just relates the a and the h. Hence in this sense (19) generalizes (1). As $b\to 0$, it is simple to check $\rho_{\mathrm{opt}}\to 1$. In short, (19) gives the exact relationship among the fundamental parameters a, b, h in optimal sensitivity theory.

5. Conclusions. Once again we have seen the utility of the complex and functional-analytical methods of [21], [24] in dealing with systems with delays. In this paper we have solved (or at least given an implementable procedure to solve) the weighted H^∞-minimization problem for an interesting class of delay systems. From our techniques, we have derived a precise picture of the interaction of a delay, nonminimum phase zero, and given weight in an H^∞-optimal sensitivity problem. Our work in a certain sense gives mathematically rigorous justification to results that one would hope to be true from just purely engineering considerations.

Finally, we have generalized some of the one-step extension results of Adamjan, Arov and Krein [1], [2], and perhaps given a new perspective to certain kinds of (generalized) interpolation problems. It should be interesting to try to push through the techniques we have given here for broader classes of distributed systems, for example those considered in [3], or even in [7].

REFERENCES

[1] V. M. ADAMJAN, D. Z. AROV AND M. G. KREIN, *Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem*, Math. USSR-Sb., 15 (1971), pp. 31-73.

[2] ———, *Infinite block Hankel matrices and related extension problems*, Amer. Math. Soc. Transl., 111 (1978), pp. 133-156.

[3] F. M. CALLIER AND C. A. DESOER, *An algebra of transfer functions for distributed time-invariant systems*, IEEE Trans. Circuits and Systems, CAS-25 (1978), pp. 651-662.

[4] D. N. CLARK, *Concrete model theory for a class of operators*, J. Funct. Anal., 14 (1973), pp. 269-280.

[5] C. DAVIS, W. M. KAHAN AND H. F. WEINBERGER, *Norm-preserving dilations and their applications to optimal error bounds*, SIAM J. Numer. Anal., 19 (1982), pp. 445-469.

[6] P. L. DUREN, *Theory of H^P-Spaces*, Academic Press, New York, 1970.

[7] A. FEINTUCH AND A. TANNENBAUM, *Gain optimization for distributed systems*, Systems Control Lett., 6 (1986), pp. 295-301.

[8] D. FLAMM, M.I.T. thesis proposal, Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 1985.

[9] C. FOIAS, A. TANNENBAUM AND G. ZAMES, *Weighted sensitivity minimization for delay systems*, IEEE Trans. Automat. Control, AC-31 (1986), pp. 763-766.

[10] B. A. FRANCIS, *Notes on H^∞-optimal linear feedback systems*, Lecture Notes, Linköping Univ., 1983.

[11] B. A. FRANCIS AND J. DOYLE, *Linear control theory with an H^∞ optimality criterion*, Systems Control Group report #8501, Univ. Toronto, October 1985.

[12] B. A. FRANCIS, J. W. HELTON AND G. ZAMES, *H^∞-optimal feedback controllers for linear multivariable systems*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 888-900.

[13] B. A. FRANCIS AND G. ZAMES, *On H^∞-optimal sensitivity theory for SISO feedback systems*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 9-16.

[14] J. B. GARNETT, *Bounded Analytical Functions*, Academic Press, New York, 1981.

[15] K. GLOVER, *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds*, Internat. J. Control, 39 (1984), pp. 1115-1193.

[16] P. P. KHARGONEKAR AND K. POOLLA, *Robust stabilization of distributed systems*, Automatica, to appear.

[17] M. G. KREIN AND A. A. NUDELMAN, *The Markov Moment Problem and External Problems*, Translations of Mathematical Monographs, 50, American Mathematical Society, Providence, RI, 1977.

[18] S. C. POWER, *Hankel Operators on Hilbert Space*, Pitman Advanced Publishing Program, Boston, 1982.

[19] W. RUDIN, *Functional Analysis*, McGraw-Hill, New York, 1973.

[20] D. SARASON, *Generalized interpolation in H^∞*, Trans. Amer. Math. Soc. 127 (1967), pp. 179-203.

[21] B. SZ. NAGY, *Unitary dilations of Hilbert space operators and related topics*, CBMS, Regional Conference Series in Mathematics 19, American Mathematical Society, Chap. 9, Par. 2, 1974.

[22] B. SZ. NAGY AND C. FOIAS, *Forme triangulaire d'une contraction et factorisation de la fonction caractéristique*, Acta Sci. Math., 28 (1967), pp. 201-212.

[23] ———, *Dilation des commutants*, C.R. Acad. Sci. Paris Sér. I Math., 266 (1968), pp. 493-495.

[24] ———, *Harmonic Analysis of Operators on Hilbert Space*, North-Holland, Amsterdam, 1970.

[25] A. TANNENBAUM, *Invariance and System Theory: Algebraic and Geometric Aspects*, Lecture Notes in Mathematics 845, Springer-Verlag, Berlin, New York, 1981.

[26] G. ZAMES, *Feedback and optimal sensitivity: model reference transformations, seminorms, and approximate inverses*, IEEE Trans. Automat. Control, AC-23 (1981), pp. 301-320.

[27] G. ZAMES AND B. FRANCIS, *Feedback, minimax sensitivity, and optimal robustness*, IEEE Trans. Automat. Control, AC-28 (1981), pp. 585-601.