SRC TR 88-27

Digital Controller Design for Multivariable Systems with Structural Closed-Loop Performance Specifications

by

Evanglelos Zafiriou
and Manfred Morari

INT. J. CONTROL, 1987, VOL. 46, NO. 6, 2087–2111

Digital controller design for multivariable systems with structural closed-loop performance specifications

EVANGHELOS ZAFIRIOU†‡ and MANFRED MORARI†

The problem of the direct design of the closed-loop transfer function matrix is addressed for multivariable discrete systems. The limitations imposed by unstable zeros, time delays and the structure associated with these are quantified. A design procedure is formulated that provides the designer with quantitative measures for evaluating the tradeoffs between different closed-loop interaction structures and durations. The problem of intersample rippling is also considered. The procedure requires only linear-algebra operations, includes the eventual construction of the feedback controller in state space, and is presented in a way that allows its straightforward computer implementation.

1. Introduction

One can find in the control literature numerous different types of criteria for synthesizing or evaluating a control system. In most cases a number of performance considerations is lumped together into some objective function, which is then optimized with respect to the control system. Such approaches have been proven satisfactory in many cases. However there are situations in which one cannot simply optimize a single scalar objective function. In process control, such a case is that of setpoint tracking for multivariable systems. Quite often, it is necessary to look at the closed-loop transfer function matrix relating the setpoints to the process outputs and require that certain elements of the matrix are equal to zero, so that setpoint changes in some outputs do not upset other important ones. Also, one may sometimes wish to allow such closed-loop interactions in order to improve setpoint tracking for the important outputs at the expense of upsetting less valuable ones. The same arguments carry over to certain cases of disturbance rejection. The paper treats setpoint tracking and disturbance rejection in a uniform way.

2. Achievable input/output mappings

The discretized plant is described by the transfer matrix $P(z)$, which is obtained by adding a zero-order hold in front of the continuous plant and then taking the z-transform. $P(z)$ is assumed to be square.

Let $H_{oi}(z)$ denote the transfer matrix between output o and input i. We can define the following relations with respect to Fig. 1.

$$H_{ur} = C(I + PC)^{-1} \tag{1}$$

$$H_{ud} = -H_{ur} \tag{2}$$

Received 21 January 1987.

† Chemical Engineering Department, California Institute of Technology, 206-41 Pasadena, CA 91125, U.S.A.

‡ Current affiliation: Chemical and Nuclear Engineering Department and Systems Research Center, University of Maryland, College Park, MD 20742, U.S.A.

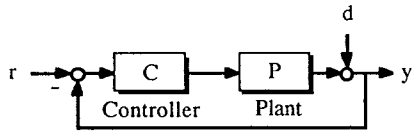


Figure 1. Feedback control structure.

$$H_{yr} = PC(I + PC)^{-1} = PH_{ur} \tag{3}$$

$$H_{yd} = (I + PC)^{-1} = I - H_{yr} = I - PH_{ur} \tag{4}$$

From (4) it follows that if the control system provides good setpoint tracking ($H_{yr}r \approx r$) then one also has good disturbance rejection ($H_{yd}d \approx 0$) provided that the disturbance d is of a type similar to the setpoint r. If this is not the case, then one has to design a two-degree-of-freedom controller (Vidyasagar 1985) by separately designing two different controllers C, one for setpoint tracking and one for disturbance rejection, and then appropriately combining them into one unified block structure (see, for example, Zafiriou and Morari 1987). Hence it is sufficient to cover here only the design of C (Fig. 1) for good setpoint tracking or disturbance rejection.

From (1) we can obtain

$$C = H_{ur}(I - PH_{ur})^{-1} \tag{5}$$

and so designing C is equivalent to designing H_{ur}, which is the controller of the internal model control structure (Garcia and Morari 1982) or the parameter of the Q-parametrization (Zames 1981). It can be shown (e.g. Callier and Desoer 1982) that necessary and sufficient conditions for the internal stability of the system in Fig. 1 are as follows.

Condition C.1

(i) H_{ur} stable

(ii) PH_{ur} stable

(iii) $H_{ur}P$ stable

(iv) $(I - PH_{ur})P$ stable

C.1 (ii), (iii), (iv) are implied by C.1 (i) if P is stable. Hence the following assumption, which will be made throughout this paper allows us to consider only C.1 (i).

Assumption A.1

P is stable.

It should be pointed out however, that for setpoint tracking, the above assumption need not be made. In that case, the use of the two-degree-of-freedom structure makes it sufficient to consider C.1 (i) only, even when P is unstable. The problem is then reduced to the one discussed in this paper in which A.1 holds (Vidyasagar 1985, Zafiriou and Morari 1987).

The controller $C(z)$ has to be causal, since future measurements of the plant output are not known. It follows from (5) that an equivalent condition is as follows.

Condition C.2

H_{ur} causal.

One can see from the above discussion that the control objective can be reduced to finding an $H_{yr}(z)$ with the desired structure and properties, which can be produced through (3) by an $H_{ur}(z)$ that satisfies C.1 (i) and C.2. However, looking only at $H_{yr}(z)$ for checking the performance of the control system may be insufficient because of the phenomenon of intersample rippling. This phenomenon is present when $H_{ur}(z)$ has poles near $(-1, 0)$ which are cancelled by the zeros of $P(z)$ in (3). Hence, in order to make it sufficient to judge performance by looking at $H_{yr}(z)$ only, $H_{ur}(z)$ must also satisfy the following condition.

Condition C.3

H_{ur} cancels no zeros of P that are 'near' $(-1, 0)$.

One can use a number of different regions on the z-plane to define 'near' $(-1, 0)$ (Åström and Wittenmark 1984). A simple and satisfactory way in practice is to include all zeros with negative real parts (Zafiriou and Morari 1985).

3. Characterization of all permissible $H_{yr}(z)$

From (3) it follows that

$$H_{ur} = P^{-1} H_{yr} \tag{6}$$

Hence the conditions of § 2 on H_{ur} can be translated into the following condition on H_{yr}.

Condition C.4

H_{yr} is a stable, causal transfer matrix that makes $P^{-1}H_{yr}$ causal and cancels the poles of P^{-1} (zeros of P) that are outside the unit circle or near $(-1, 0)$.

The time delays in $P(z)$, which make P^{-1} non-causal, appear as zeros at infinity. We shall now exploit this fact to make the treatment of time delays and undesirable zeros of P uniform. The transformation $\lambda = z^{-1}$ will be used. Define

$$\hat{P}(\lambda) \stackrel{\text{def}}{=} P(\lambda^{-1}) \leftrightarrow P(z) \tag{7}$$

$$\hat{H}_{yr}(\lambda) \stackrel{\text{def}}{=} H_{yr}(\lambda^{-1}) \leftrightarrow H_{yr}(z) \tag{8}$$

Let $a_1, \ldots, a_f$ be the zeros of $P(z)$, which, according to C.4, we do not wish to appear as poles of $P(z)^{-1}H_{yr}(z)$. These will appear in $\hat{P}(\lambda)^{-1}$ as poles at $b_1, \ldots, b_f$ where

$$b_i = \frac{1}{a_i}, \quad i = 1, \ldots, f \tag{9}$$

The time delays in $P(z)$ will give rise to zeros at 0 in $\hat{P}(\lambda)$ and consequently the non-causal terms in $P(z)^{-1}$ will produce poles at 0 in $\hat{P}(\lambda)^{-1}$. Hence C.4 is equivalent to C.5 as follows.

Condition C.5

(i) $H_{yr}(z)$ is a stable, causal transfer matrix

(ii) $\hat{P}(\lambda)^{-1}\hat{H}_{yr}(\lambda)$ has no poles at $b_0, b_1, \ldots, b_f$.

In the above, the following notation has been used:

$$b_0 = 0 \tag{10}$$

Some additional notation and definitions are now needed. $P(z)$ (and $\hat{P}(\lambda)$) is assumed to have dimension $r \times r$ and to be of normal rank r. In the following, it will be assumed that $\hat{P}(\lambda)$ has no poles at $b_0, ..., b_f$. This is certainly the case for b_0, since all elements of $P(z)$ are proper, but, in general, $P(z)$ may have poles at $a_1, ..., a_f$ resulting in poles at $b_1, ..., b_f$ in $\hat{P}(\lambda)$. The existence of poles and zeros at the same location is a clearly multivariable characteristic (Kailath 1980). The assumption that this is not the case for $P(z)$ serves considerably to simplify the notation and it is not restrictive since such a phenomenon is caused by exact cancellations in det $[P(z)]$ which will not happen if a slight perturbation in the terms of $P(z)$ is introduced. Let $\{n_0, n_1, ..., n_f\}$ be a set of integers greater than or equal to zero such that

$$\hat{P}^{(k)}(b_i) = 0, \quad ((k = 0, ..., n_i - 1), i = 0, ..., f) \tag{11}$$

$$\text{rank } [\hat{P}^{(n_i)}(b_i)] \neq 0, \quad i = 0, ..., f \tag{12}$$

where $\hat{P}^{(k)}(\lambda)$is the kth derivative of $\hat{P}(\lambda)$. Also, let m_i, $i = 0, ..., f$, be the order of the zero b_i of $\hat{P}(\lambda)$, as this order is defined from the Smith–McMillan form of $\hat{P}(\lambda)$ (Desoer and Schulman 1974). The computation of m_i without going through the Smith–McMillan form is briefly discussed in § 4.2. From (11), (12) and the definition of the order of a zero, it follows that

$$m_i \geqslant n_i, \quad i = 0, ..., f \tag{13}$$

The following theorem quantifies C.5 (ii).

Theorem 1

Condition C.5 (ii) holds if and only if both (*a*) and (*b*) hold, as follows:

(*a*) $$\hat{H}_{yr}(\lambda) = (\lambda - b_i)^{n_i} \hat{H}_i(\lambda), \quad i = 0, ..., f$$

where $\hat{H}_i(\lambda)$ is a rational $r \times r$ matrix in λ, with no poles at b_i;

(*b*) for any $i = 0, ..., f$ such that $m_i > n_i$, the columns of

$$\left[\hat{H}_i^{(0)}(b_i)^T \ ... \ \frac{1}{(m_i - n_i - 1)!} \hat{H}_i^{(m_i - n_i - 1)}(b_i)^T \right]^T$$

are in the column space of

$$M_i \overset{\text{def}}{=} \begin{bmatrix} \frac{1}{n_i!} \hat{P}^{(n_i)}(b_i) & 0 & ... & 0 \\ \frac{1}{(n_i+1)!} \hat{P}^{(n_i+1)}(b_i) & \frac{1}{n_i!} \hat{P}^{(n_i)}(b_i) & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(m_i-1)!} \hat{P}^{(m_i-1)}(b_i) & \frac{1}{(m_i-2)!} \hat{P}^{(m_i-2)}(b_i) & ... & \frac{1}{n_i!} \hat{P}^{(n_i)}(b_i) \end{bmatrix} \tag{14}$$

where the superscript (k) indicates the kth derivative and T indicates matrix transposition.

Proof

See Appendix A.

The value of Theorem 1 lies in the fact that it provides a characterization of all acceptable $\hat{H}_{yr}(\lambda)$ without requiring the inversion of $\hat{P}(\lambda)$. The theorem applies to the general case. However in practice one is usually faced with a situation where the order of the zeros $a_1, ..., a_f$ of the model $P(z)$ is equal to 1. Hence of the zeros $b_0, b_1, ..., b_f$ of $\hat{P}(\lambda)$ only b_0 has an order larger than 1. The fact that b_0 is equal to zero (see (10)) can then be used to obtain a simpler form for Theorem 1. The following two corollaries describe these situations.

Corollary 1

Let the order of the zero a_i of $P(z)$ be equal to one. Then $\hat{P}(\lambda)^{-1}\hat{H}_{yr}(\lambda)$ has no poles at b_i if and only if the columns of $\hat{H}_{yr}(b_i)$ are in the column space of $\hat{P}(b_i)$ $(= P(a_i))$.

Proof

The proof follows directly from Theorem 1 for $m_i = 1$.

Corollary 2

Let $P(z)$ have the impulse response coefficient description

$$P(z) = z^{-N}(A_0 + A_1 z^{-1} + A_2 z^{-2} + ...) \tag{15}$$

where

$$\text{rank } [A_0] \neq 0 \tag{16}$$

$$N \geqslant 0 \tag{17}$$

Then

$$n_0 = N \tag{18}$$

$$M_0 = \begin{bmatrix} A_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ A_{m_0-1-N} & \cdots & A_0 \end{bmatrix} \tag{19}$$

and $\hat{P}(\lambda)^{-1}\hat{H}_{yr}(\lambda)$ has no poles at $b_0 = 0$ if and only if both (a) and (b) hold, as follows:

(a) $\hat{H}_{yr}(\lambda) = \lambda^N \hat{H}_0(\lambda)$ where $\hat{H}_0(\lambda)$ is a rational matrix in λ with no poles at $b_0 = 0$;

(b) if $m_0 > N$, the columns of

$$\left[\hat{H}_0^{(0)}(0)^T \cdots \frac{1}{(m_0 - N - 1)!}\hat{H}_0^{(m_0-N-1)}(0)^T\right]^T$$

are in the column space of M_0.

Proof

From (7), we obtain $\hat{P}(\lambda) = P(\lambda^{-1}) = \lambda^N(A_0 + A_1\lambda + A_2\lambda^2 + ...)$. Equations (18) and (19) can now be obtained by repeated differentiation and evaluation at $\lambda = 0$. The rest follows as a restatement of Theorem 1 for this special case.

whose dimension can grow very large, might be difficult and time consuming. Van Dooren *et al.* (1979) have exploited the Toeplitz matrix form of $M_{i,k}$ to develop a fast recursive algorithm that performs the rank search in a numerically stable way. In each step, the rank of $M_{i,k}$ is computed for some k by obtaining the SVD of an $r \times r$ matrix. At the same time, $M_{i,k}$ is reduced to a form with linearly independent columns. Hence, to obtain m_i and an orthonormal basis for the column space of M_i one has to obtain the SVD of only $(m_i - n_i + 1)$ matrices of dimension $r \times r$.

4.3. *Design of a column of* $H_{yr}(z)$

The requirements of Theorem 1 apply to each column of H_{yr} separately, and so each column can be designed independently. Let us write

$$H_{yr}(z) = [h_1(z) \quad \dots \quad h_r(z)] \tag{24}$$

where $h_j(z)$ has the dimension $r \times 1$, $j = 1, \dots, r$. Also let

$$\hat{h}_j(\lambda) = h_j(\lambda^{-1}) \leftrightarrow h_j(z), \quad j = 1, \dots, r \tag{25}$$

We shall now proceed with the design of $\hat{h}_j(\lambda)$ for some j. Let U_i be a matrix whose columns form an orthonormal basis for the column space of M_i given in (14). U_i can be obtained from M_i by the procedure of Van Dooren *et al.* (1979), briefly discussed in § 4.2. Also let

$$\rho_i = \text{rank } [M_i] = \text{rank } [U_i] \tag{26}$$

According to Theorem 1 we must have

$$\hat{h}_j(\lambda) = (\lambda - b_i)^{n_i} \hat{h}_{j,i}(\lambda), \quad i = 0, \dots, f \tag{27}$$

where

$$\eta_{j,i} \overset{\text{def}}{=} \left[\hat{h}_{j,i}^{(0)}(b_i)^{\mathrm{T}} \dots \frac{1}{(m_i - n_i - 1)!} \hat{h}_{j,i}^{(m_i - n_i - 1)}(b_i)^{\mathrm{T}} \right]^{\mathrm{T}} \tag{28}$$

is a linear combination of the columns of U_i, i.e.

$$\eta_{j,i} = U_i \chi_i^1 \tag{29}$$

where χ_i^1 is any vector of dimension ρ_i. The freedom allowed in the choice of χ_i^1 will now be gradually reduced by requiring certain properties for $\hat{h}_j(\lambda)$, according to the designer's specifications and decisions. First, the limitations imposed by the desired structure of h_j will be quantified. Then, the undesirable zeros and time delays that have to be present in the diagonal element will be determined and the design of this element will be reduced to that of an SISO system. Finally, the non-diagonal elements will be designed so that the closed-loop interactions are minimized. It should be pointed out that if for some i we have $m_i = n_i$, then part (*b*) of Theorem 1 and therefore (29) do not apply for that i and so all equations in this section corresponding to that particular i should be ignored.

4.3.1. *Structure of* h_j. Let the design specification be that the $l_1, l_2, \dots, l_g$ elements of $h_j(z)$ be identically equal to zero, where

$$g \leqslant r - 1 \tag{30}$$

$$l_k \neq j, \quad k = 1, \dots, g \tag{31}$$

We shall use l to denote the set

$$l \stackrel{\text{def}}{=} \{l_1, ..., l_g\} \tag{32}$$

Define

$$\bar{l} \stackrel{\text{def}}{=} \{1, 2, ..., r\} - \{j\} - l \tag{33}$$

Let

$$e_k \stackrel{\text{def}}{=} [0 \quad ... \quad 0 \quad 1 \quad 0 \quad ... \quad 0]^T \tag{34}$$

where the 1 is the kth element and e_k has dimension $r \times 1$.

Define

$$e^l \stackrel{\text{def}}{=} \begin{bmatrix} e_{l_1}^T \\ \vdots \\ e_{l_g}^T \end{bmatrix} \tag{35}$$

$$\Lambda_i^l \stackrel{\text{def}}{=} \text{diag}\, [\underbrace{e^l, ..., e^l}_{(m_i - n_i)}], \quad i = 1, ..., f \tag{36}$$

In order for the specified elements of h_j to be zero, the vector χ_i^1 must solve:

$$\Lambda_i^l \eta_{j,i} = 0$$

or

$$\Lambda_i^l U_i \chi_i^1 = 0 \tag{37}$$

Let

$$\rho_i^l = \text{rank}\, [\Lambda_i^l M_i] = \text{rank}\, [\Lambda_i^l U_i] \tag{38}$$

Then $\rho_i \geqslant \rho_i^l$. Hence the null space of $\Lambda_i^l U_i$ has dimension

$$\xi_i^l = \rho_i - \rho_i^l \tag{39}$$

Let V_i^l be a matrix whose columns form an orthonormal basis for the null space of $\Lambda_i^l U_i$. Both V_i^l and ρ_i^l can be obtained from an SVD of $\Lambda_i^l U_i$. Then the solutions to (37) are:

$$\chi_i^1 = V_i^l \chi_i^2 \tag{40}$$

where χ_i^2 can be any vector of dimension ξ_i^l when $\xi_i^l \neq 0$. If $\xi_i^l = 0$ then of course $V_i^l = 0$ and $\chi_i^1 = 0$.

Hence we must have

$$\eta_{j,i} = U_i V_i^l \chi_i^2 \tag{41}$$

where $\eta_{j,i}$ was defined in (28). Note that (41) includes the case $\xi_i^l = 0$, where $V_i^l = 0$ yields $\eta_{j,i} = 0$.

4.3.2. *Diagonal element of* h_j. We shall now proceed with the determination of the jth element of h_j. Up to this point, no assumption has been made on the order of the zeros $b_0, b_1, ..., b_f$. However, if more than one zero has order larger than 1, then the number of possible choices to be examined at this point could grow enormously.

On the other hand, in practice one is usually faced with a situation where the order and degree (as defined from the Smith–McMillan form—see Desoer and Schulman 1974 and Lemma A.1 in Appendix A) of the zeros $a_1, ..., a_f$ of $P(z)$ and therefore of the zeros $b_1, ..., b_f$ of $\hat{P}(\lambda)$, is equal to 1. The following assumption will be made here to simplify the procedure.

Assumption A.2

The degree of the zeros $a_1, ..., a_f$ of $P(z)$ is equal to 1.

No assumption is made however about the zero $b_0 = 0$ of $\hat{P}(\lambda)$ corresponding to time delays in $P(z)$. We shall examine the two cases separately.

Case (a) $b_i, i = 1, ..., f$

It follows from Assumption A.2 that for the order of the zeros m_i we have $m_i = 1$. Also since $r \geqslant 2$, Assumption A.2 implies that $n_i = 0$. Then from (27)–(29) it follows that since 0 is a linear combination of the columns of U_i, the highest power of $(\lambda - b_i)$ that is sufficient to be included in the elements of $\hat{h}_j(\lambda)$ is $(\lambda - b_i)^1$. However, according to Rule 1 we wish to have the smallest possible power in the jth element, and that is $(\lambda - b_i)^0 = 1$. In order for this to be possible, the following equation must have a solution

$$e_j^T \eta_{j,i} = 1$$

or

$$e_j^T U_i V_i^l \chi_i^2 = 1 \tag{42}$$

where e_j is defined in (34). Equation (42) will have no solution only if the matrix $e_j^T U_i V_i^l$ is identically zero. If this happens for some i, then the factor $(\lambda - b_i)$ must be included in the jth element of $\hat{h}_j(\lambda)$. Let us assume that this matrix is zero only for $i = 1, ..., \phi_1$. Also let the zeros $b_1, ..., b_{\phi_2}$ $(\phi_2 \leqslant \phi_1)$ have positive real part and the zeros $b_{\phi_2+1}, ..., b_{\phi_1}$ have negative real part. Then, according to Rule 2, the factor

$$\zeta_0(\lambda) = \prod_{i=1}^{\phi_2} \frac{(1 - b_i^{-1})(\lambda - b_i)}{(1 - b_i)(\lambda - b_i^{-1})} \prod_{i=\phi_2+1}^{\phi_1} \frac{(\lambda - b_i)}{(1 - b_i)} \tag{43}$$

should be included in the jth element.

Note that one does not always have to follow Rule 1. One may wish to include the factor $(\lambda - b_i)$ for some i in the jth element even when one does not have to do it, if that will result in significantly smaller interactions (non-diagonal elements) and if the jth output is not so important. The procedure for determining the magnitude of the interactions will then be exactly the same (see § 4.3.3) and at the end the designer can decide whether the inclusion of $(\lambda - b_i)$ pays off. A simple qualitative way to figure out *a priori* whether it may pay off, without going through the whole design procedure, is the following. For $m_i = 1$ and $n_i = 0$ we have rank $[U_i] = r - 1$. U_i consists of the first $(r - 1)$ columns of the left singular vector matrix in an SVD of M_i. The rth column u_i is orthogonal to all the columns of U_i. If the jth element of u_i is large compared with the kth elements, where k belongs to the set $\bar{I}$ defined in (33), then it is likely that the inclusion of $(\lambda - b_i)$ in the jth element will result in significantly smaller interactions in the non-zero non-diagonal elements of h_j.

Case (*b*) $b_0(=0)$ (Time delays)

Define

$$\zeta_\tau(\lambda) = \lambda^\tau \zeta_0(\lambda) \tag{44}$$

where τ is an integer and $\zeta_0(\lambda)$ is given by (43). Then, according to Rule 1 we need to find the smallest τ such that $\zeta_\tau(\lambda)$ is possible as the jth element of $\hat{h}_{j,0}(\lambda)$. From (28), (29) it follows that in order for a τ to be possible, the following equation must have a solution

$$\varepsilon_j \eta_{j,0} = Z_\tau$$

or

$$\varepsilon_j U_0 V_0^l \chi_0^2 = Z_\tau \tag{45}$$

where

$$\varepsilon_j = \text{diag}\,[\underbrace{e_j^T, \ldots, e_j^T}_{(m_0 - n_0)}] \tag{46}$$

and

$$Z_\tau = \left[\zeta_\tau^{(0)}(0) \ldots \frac{1}{(m_0 - n_0 - 1)!}\zeta_\tau^{(m_0 - n_0 - 1)}(0)\right]^T \tag{47}$$

Hence one can obtain the smallest possible τ as

$$\tau_0 = \min\{\tau \in N_0 \,|\, \text{rank}\,[\varepsilon_j U_0 V_0^l \,|\, Z_\tau] = \text{rank}\,[\varepsilon_j U_0 V_0^l]\} \tag{48}$$

where N_0 is the set of positive integers, including zero.

Still, contrary to Rule 1 one may wish to choose a τ larger than τ_0 if that results in smaller interactions for a given set l. Equation (45) should, of course, have a solution for this τ, i.e. the rank condition in (48) should hold. In the following paragraph τ_0 is used, but any other possible τ can be used instead without affecting the procedure for determining the magnitude of the interactions in § 4.3.3.

The jth element of $\hat{h}_j(\lambda)$ has been completely determined at this point as

$$e_j^T \hat{h}_j(\lambda) = \lambda^N \zeta_{\tau_0}(\lambda) \tag{49}$$

Let us now quantify the limitations that the selection of this diagonal element imposes on χ_i^2, $i = 0, 1, \ldots, f$. The following equations have to be satisfied.

$$e_j^T U_i V_i^l \chi_i^2 = b_i^{n_0} \zeta_{\tau_0}(b_i), \quad i = 1, \ldots, f \tag{50}$$

$$\varepsilon_j U_0 V_0^l \chi_0^2 = Z_{\tau_0} \tag{51}$$

Let χ_i^0, $i = 0, \ldots, f$, be a particular solution for each corresponding equation, obtained with some method for solving systems of linear equations. Also let W_i^l, $i = 1, \ldots, f$ be a matrix whose columns form an orthonormal basis for the null space of $e_j^T U_i V_i^l$, and W_0^l the corresponding matrix for $\varepsilon_j U_0 V_0^l$. These matrices and their ranks w_i^l, can be obtained from an SVD. Then the χ_i^2 that solve the set of equations (50), (51) are:

$$\chi_i^2 = \chi_i^0 + W_i^l \chi_i^3, \quad i = 0, 1, \ldots, f \tag{52}$$

where χ_i^3 is any vector of dimension w_i^l, when $w_i^l \neq 0$. If $w_i^l = 0$ then $W_i^l = 0$ and $\chi_i^2 = \chi_i^0$.

From (41) and (52) we obtain

$$\eta_{j,i} = U_i V_i^l \chi_i^0 + U_i V_i^l W_i^l \chi_i^3, \quad i = 0, 1, \ldots, f \tag{53}$$

4.3.3. *Non-diagonal elements of h_j.* The part of the procedure that was developed in § 4.3.1 makes sure that the elements of h_j corresponding to the set l defined in (32) are identically equal to zero. We shall now proceed to compute the terms in the non-zero non-diagonal elements of h_j, i.e. the elements corresponding to the set $\bar{l}$, defined in (33). To do so, the freedom allowed in the choice of χ_i^3 will be used.

Let $\bar{l}_1, \bar{l}_2, \ldots, \bar{l}_q$ be the elements of the set $\bar{l}$. According to Rule 3, the $\bar{l}_k$th element of $\hat{h}_j(\lambda)$ should be of the form

$$e_{\bar{l}_k}^{\mathrm{T}} \hat{h}_j(\lambda) = B_k(\lambda) \stackrel{\mathrm{def}}{=} \lambda^{\delta_k}(\beta_{k,0} + \beta_{k,1}\lambda + \ldots + \beta_{k,\nu}\lambda^{\nu})(1-\lambda), \quad k = 1, \ldots, q \tag{54}$$

From Corollary 2 (a) it follows that

$$\delta_k = n_0 = N, \quad k = 1, \ldots, q \tag{55}$$

The values of $\beta_{k,0}, \ldots, \beta_{k,\nu}$, $k = 1, \ldots, q$ will be computed from (28) and (53). Note that any of the β can be zero, including the first ones, β_0, β_1, etc.

Define

$$e^{\bar{l}} \stackrel{\mathrm{def}}{=} \begin{bmatrix} e_{\bar{l}_1}^{\mathrm{T}} \\ \vdots \\ e_{\bar{l}_q}^{\mathrm{T}} \end{bmatrix} \tag{56}$$

$$\varepsilon^{\bar{l}} \stackrel{\mathrm{def}}{=} \begin{bmatrix} \varepsilon_{\bar{l}_1} \\ \vdots \\ \varepsilon_{\bar{l}_q} \end{bmatrix} \tag{57}$$

where $\varepsilon_{\bar{l}_k}$ is defined as in (46) for $\bar{l}_k$ instead of j. As explained in Case 1 of § 4.3.2 we have $n_i = 0$ for $i = 1, \ldots, f$. From (27), (28), (54), (55) it follows that

$$e^{\bar{l}}\eta_{j,i} = \begin{bmatrix} B_1(b_i) \\ \vdots \\ B_q(b_i) \end{bmatrix} = \operatorname{diag} \underbrace{[\gamma_{i,\nu}, \ldots, \gamma_{i,\nu}]}_{q} \begin{bmatrix} \theta_{1,\nu} \\ \vdots \\ \theta_{q,\nu} \end{bmatrix}, \quad i = 1, \ldots, f \tag{58}$$

where

$$\gamma_{i,\nu} \stackrel{\mathrm{def}}{=} [(1-b_i)b_i^N \quad (1-b_i)b_i^{(N+1)} \quad \ldots \quad (1-b_i)b_i^{(N+\nu)}], \quad i = 1, \ldots, f \tag{59}$$

$$\theta_{k,\nu} \stackrel{\mathrm{def}}{=} [\beta_{k,0} \quad \beta_{k,1} \quad \ldots \quad \beta_{k,\nu}]^{\mathrm{T}}, \quad k = 1, \ldots, q \tag{60}$$

It also follows that

$$\varepsilon^{\bar{l}}\eta_{j,0} = \begin{bmatrix} B_1^0 \\ \vdots \\ B_q^0 \end{bmatrix} = \operatorname{diag} \underbrace{[\Gamma_\nu, \ldots, \Gamma_\nu]}_{q} \begin{bmatrix} \theta_{1,\nu} \\ \vdots \\ \theta_{q,\nu} \end{bmatrix} \tag{61}$$

where

$$B_k^0 = [\beta_{k,0} \quad (\beta_{k,1} - \beta_{k,0}) \quad \ldots \quad (\beta_{k,m_0-N-1} - \beta_{k,m_0-N-2})]^{\mathrm{T}} \quad k = 1, \ldots, q \tag{62}$$

$$\beta_{k,\mu} = 0 \quad \text{for} \quad \mu > \nu, \quad k = 1, \ldots, q \tag{63}$$

and Γ_v is a matrix containing the first $v+1$ columns of $[\Gamma \quad 0 \quad 0 \quad 0 \quad \ldots]$ with

$$\Gamma = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ -1 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & -1 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & -1 & 1 \end{bmatrix}_{(m_0-N)\times(m_0-N)} \tag{64}$$

Then the use of (53) in (58), (61) and combination of the resulting equations yields:

$$K_v\theta_v = T_1 + T_2 X \tag{65}$$

where

$$\theta_v \overset{\text{def}}{=} \begin{bmatrix} \theta_{1,v} \\ \vdots \\ \theta_{q,v} \end{bmatrix} \tag{66}$$

$$K_v \overset{\text{def}}{=} \begin{bmatrix} \text{diag}\,[\Gamma_v, \ldots, \Gamma_v] \\ \text{diag}\,[\gamma_{1,v}, \ldots, \gamma_{1,v}] \\ \vdots \\ \text{diag}\,[\underbrace{\gamma_{f,v}, \ldots, \gamma_{f,v}}_{q}] \end{bmatrix} \tag{67}$$

$$X \overset{\text{def}}{=} [(\chi_0^3)^{\mathrm{T}} \quad (\chi_1^3)^{\mathrm{T}} \quad \ldots \quad (\chi_f^3)^{\mathrm{T}}]^{\mathrm{T}} \tag{68}$$

$$T_1 \overset{\text{def}}{=} \begin{bmatrix} \varepsilon^I U_0 V_0^l \chi_0^0 \\ e^I U_1 V_1^l \chi_1^0 \\ \vdots \\ e^I U_f V_f^l \chi_f^0 \end{bmatrix} \tag{69}$$

$$T_2 \overset{\text{def}}{=} \text{diag}\,[\varepsilon^I U_0 V_0^l W_0^l, e^I U_1 V_1^l W_1^l, \ldots, e^I U_f V_f^l W_f^l] \tag{70}$$

Equation (65) can also be written as

$$[K_v | -T_2] \begin{bmatrix} \theta_v \\ X \end{bmatrix} = T_1 \tag{71}$$

Hence the smallest possible v for h_j and for the particular choice of the set l can be obtained as the smallest v for which (71) has a solution, i.e.

$$v_{\min} = \min\,\{v \in N_0 \,|\, \text{rank}\,[K_v | -T_2] = \text{rank}\,[K_v | -T_2 | T_1]\} \tag{72}$$

However instead of trying to minimize v, a better alternative to use a larger v and use the extra degrees of freedom to minimize the sum of the squared errors for the step response of the $\bar{I}_1, \ldots, \bar{I}_q$ system outputs to the jth external input (jth element of r or d). This means minimizing

$$J_v \overset{\text{def}}{=} \sum_{k=1}^{q} \left(\phi_k \sum_{\mu=0}^{v} \beta_{k,\mu}^2 \right) = \theta_v^{\mathrm{T}} \Phi^{\mathrm{T}} \Phi \theta_v \tag{73}$$

where ϕ_k, $k = 1, ..., q$ are optional weights (positive real numbers) and

$$\Phi = \text{diag}\,[\phi_1^{1/2} I_{v+1}, ..., \phi_q^{1/2} I_{v+1}] \tag{74}$$

where I_{v+1} is the $(v+1) \times (v+1)$ identity matrix.

Equation (65) can be written as

$$K_v \Phi^{-1}(\Phi\theta_v) = T_1 + T_2 X \tag{75}$$

Hence the $\Phi\theta_v$ that minimizes J_v can be obtained as the minimum norm solution to (75). For v large enough, K_v is full rank, i.e. rank $[K_v] = q(f + m_0 - N)$, and for a given X, the solution is

$$\theta_v(X) = \Phi^{-1} F_v^*(F_v F_v^*)^{-1}(T_1 + T_2 X) \tag{76}$$

where

$$F_v \overset{\text{def}}{=} K_v \Phi^{-1} \tag{77}$$

and the superscript * indicates complex conjugate transpose. Note that although the matrices involved may be complex, the solution θ_v will be real because any complex zeros of $P(z)$ come in complex conjugate pairs. However the form in which the solution is given in (76) may cause numerical problems in some cases. One can avoid them by computing the pseudo-inverse $F_v^\dagger = F_v^*(F_v F_v^*)^{-1}$ from an SVD of F_v (Stewart 1973, p. 324).

One can now compute X by minimizing $J_v(X)$ for the solution $\theta_v(X)$ of (76). From (76) we obtain

$$J_v(X) = (T_1 + T_2 X)^*(F_v F_v^*)^{-1}(T_1 + T_2 X) = (T_1 + T_2 X)^*(F_v^\dagger)^* F_v^\dagger (T_1 + T_2 X) \tag{78}$$

By setting the gradient of $J_v(X)$ equal to zero we get

$$T_2^*(F_v^\dagger)^* F_v^\dagger T_2 X = -T_2^*(F_v^\dagger)^* F_v^\dagger T_1 \tag{79}$$

from which a solution X which minimizes $J_v(X)$ can be obtained. The optimum θ_v can then be computed from (76).

It is clear that by increasing v, the value of the obtained minimum of J_v will either be reduced or it will remain the same. Hence the designer has the option to choose interactions with smaller magnitude in exchange for a longer duration of the interactions. The knowledge of the value of this minimum at the limit as $v \to \infty$ would be quite helpful in making this decision. From (78) we see that we need to compute $F_v F_v^*$ as $v \to \infty$. The fact that the elements of $\gamma_{i,v}$, $i = 1, ..., f$, are terms in a geometric progression allows us to do so easily when $b_1, ..., b_f$ are inside the unit circle. We cannot do so however if some of them are outside the unit circle, i.e. when some of the undesirable zeros of $P(z)$ are inside the unit circle. This is actually a situation where for numerical reasons it would be strongly recommended to compute $F_v^\dagger$ from an SVD of F_v, as mentioned above.

5. Construction of $H_{ur}(z)$ and $C(z)$

After the desired $H_{yr}(z)$ has been designed, $H_{ur}(z)$ can be obtained from (80):

$$H_{ur}(z) = P(z)^{-1} H_{yr}(z) \tag{80}$$

Substitution of (80) into (5) yields:

$$C(z) = H_{ur}(z)[I - H_{yr}(z)]^{-1} \tag{81}$$

If one attempts to construct $H_{ur}(z)$ and $C(z)$ from (6), (81) by doing the computations in terms of transfer function matrices, the procedure would be extremely tedious. Instead, the computations can be made quite simply by working in the state space. One can obtain realizations of $P(z)$, $H_{yr}(z)$ to get the state space descriptions

$$P(z) = C(zI - A)^{-1}B + D \tag{82}$$

$$H_{yr}(z) = C_0(zI - A_0)^{-1}B_0 + D_0 \tag{83}$$

Here, $P(z)$ represents a physical system and so it can be assumed to be strictly proper, i.e. $D = 0$. Then, from Corollary 2 it follows that $D_0 = 0$. Construction of $H_{yr}(z)$, $C(z)$ involves the following steps.

Step 1. Inversion of $P(z)$.

Silverman (1969) developed a computationally simple algorithm for the inversion of a linear multivariable system whose state space description is known. The result of the inversion will be

$$P(z)^{-1} = (C_1(zI - A_1)^{-1}B_1 + D_1)(K_0 + K_1 z + \ldots K_{m_0} z^{m_0})z^N \tag{84}$$

where

$$A_1 = A - B\bar{D}^{-1}\bar{C} \tag{85}$$

$$B_1 = B\bar{D}^{-1} \tag{86}$$

$$C_1 = -\bar{D}^{-1}\bar{C} \tag{87}$$

$$D_1 = \bar{D}^{-1} \tag{88}$$

m_0 is the order of the zero b_0 of $\hat{P}(\lambda)$ obtained from Corollary 2 and N is defined in (15). The matrices $\bar{C}, \bar{D}, K_0, \ldots, K_{m_0}$ are determined with Silverman's (1969) procedure.

Step 2. Computation of $H_{ur}(z)$.

The following Theorem will be used.

Theorem 3

Let

$$G(z) = C(zI - A)^{-1}B + D \tag{89}$$

then

$$G(z)z^k = C(zI - A)^{-1}A^k B + \sum_{l=1}^{k} CA^{l-1}Bz^{k-l} + Dz^k, \quad \forall k \geqslant 1 \tag{90}$$

Proof

See Appendix B.

We can now apply Theorem 3 to $P(z)^{-1}$, to obtain

$$P(z)^{-1} = C_1(zI - A_1)^{-1}\left(\sum_{i=0}^{m_0} A_1^{i+N}B_1K_i\right)$$

$$+\sum_{i=0}^{m_0}\sum_{l=1}^{i+N} C_1 A_1^{l-1} B_1 K_i z^{i+N-l} + \sum_{i=0}^{m_0} D_1 K_i z^{i+N}$$
$$= C_2(zI - A_2)^{-1} B_2 + D_{2,0} + D_{2,1} z + \ldots + D_{2,m_0+N} z^{m_0+N} \tag{91}$$

where

$$A_2 = A_1 \tag{92}$$

$$B_2 = \sum_{i=0}^{m_0} A_1^{i+N} B_1 K_i \tag{93}$$

$$C_2 = C_1 \tag{94}$$

$$D_{2,k} = \sum_{i=0}^{m_0} C_1 A^{N-1-k+i} B_1 K_i, \quad k = 0, \ldots, N-1 \tag{95}$$

$$D_{2,k} = D_1 K_{k-N} + \sum_{i=k-N+1}^{m_0} C_1 A^{N-1-k+i} B_1 K_i, \quad k = N, \ldots, N+m_0 \tag{96}$$

Then from (6), (83), (91) we obtain

$$H_{ur}(z) = (C_2(zI - A_2)^{-1} B_2 C_0 + D_{2,0} C_0)(zI - A_0)^{-1} B_0$$
$$+ \left(\sum_{i=1}^{m_0+N} D_{2,i} z^i\right) C_0 (zI - A_0)^{-1} B_0 \tag{97}$$

Application of Theorem 3 on the second term of the right-hand side yields the term

$$\left(\sum_{i=1}^{m_0+N} D_{2,i} C_0 A_0^i\right)(zI - A_0)^{-1} B_0 + \sum_{i=1}^{m_0+N} D_{2,i} C_0 A_0^{i-1} B_0 + \sum_{i=1}^{m_0+N-1} \Psi_i z^i$$

where the fact $(zI - A)^{-1} A^k = A^k (zI - A)^{-1}$ was used. However, by construction, $P(z)^{-1} H_{yr}(z)$ is proper. Therefore $\Psi_i = 0$ for all $i = 1, \ldots, m_0 + N - 1$. Hence

$$H_{ur}(z) = \left(C_2(zI - A_2)^{-1} B_2 C_0 + D_{2,0} C_0 + \sum_{i=1}^{m_0+N} D_{2,i} C_0 A_0^i\right)$$
$$(zI - A_0)^{-1} B_0 + \sum_{i=1}^{m_0+N} D_{2,i} C_0 A_0^{i-1} B_0 \tag{98}$$

All that is necessary now is to compute the product of two proper transfer function matrices, whose state space descriptions are known. The following Theorem takes care of that.

Theorem 4 (Doyle 1984)

Let

$$G_1(z) = C_1(zI - A_1)^{-1} B_1 + D_1 \tag{99}$$

$$G_2(z) = C_2(zI - A_2)^{-1} B_2 + D_2 \tag{100}$$

then

$$G_1(z) G_2(z) = C(zI - A)^{-1} B + D \tag{101}$$

where

$$A=\begin{bmatrix} A_1 & B_1C_2 \\ 0 & A_2 \end{bmatrix},\quad B=\begin{bmatrix} B_1D_2 \\ B_2 \end{bmatrix},\quad C=[C_1|D_1C_2],\quad D=[D_1D_2]$$

Application of Theorem 4 to (98) yields a state-space description for H_{ur}.

Step 3. Computation of $C(z)$

All that is needed is to compute a state-space description of $(I-H_{yr}(z))^{-1}$. After that, application of Theorem 4 to (81) will give a state-space description for $C(z)$. From (83) we obtain

$$I-H_{yr}(z)=-C_0(zI-A_0)^{-1}B_0+I \tag{102}$$

and a state-space description of $(I-H_{yr}(z))^{-1}$ can easily be computed as

$$(I-H_{yr}(z))^{-1}=C_0(zI-(A_0+B_0C_0))^{-1}B_0+I \tag{103}$$

The result of the procedure described is the state-space descriptions of $H_{ur}(z)$ and $C(z)$. One can always obtain a matrix transfer function form, but since the control law can easily be implemented with a state-space description, it would be advisable to avoid further computations by implementing it as such. It is important to point out, however, that the realizations obtained for $H_{ur}(z)$ and $C(z)$ are not minimal. It is essential to obtain minimal realizations of them before the implementation so that the undesirable zeros $a_1,\ldots,a_f$ of $P(z)$ do not appear as poles of $H_{ur}(z)$.

6. Illustrations

The first example in this section is used to illustrate the tradeoff between the time duration of the closed-loop interactions and the magnitude of the sum-of-squared errors that they cause. This simple example is also used to demonstrate the procedure step by step. The second example examines different structures for H_{yr} and illustrates how the structure associated with a zero outside the UC can produce large or small closed-loop interactions, depending on the structure chosen for H_{yr}.

6.1. *Example* 1

Consider the system

$$P(z)=\begin{bmatrix} \dfrac{0{\cdot}6}{z-0{\cdot}4} & \dfrac{0{\cdot}5}{z-0{\cdot}5} \\ 1{\cdot}2\dfrac{0{\cdot}5}{z-0{\cdot}5} & \dfrac{0{\cdot}6}{z-0{\cdot}4} \end{bmatrix} \tag{104}$$

Computation of the roots of det $[P(z)]$ shows that the system has one zero outside the UC, at $z=1{\cdot}547$. Garcia and Morari (1985) pointed out that an acceptable lower triangular H_{yr} is

$$H_{yr,1}(z)=\begin{bmatrix} z^{-1} & 0 \\ 3{\cdot}095(1-z^{-1})z^{-1} & \dfrac{-z+1{\cdot}547}{1{\cdot}547z-1}z^{-1} \end{bmatrix} \tag{105}$$

Clearly the interactions in output 2 for a setpoint change in output 1 are very large in magnitude (over 300% of the setpoint change) although of short duration.

We shall now use the procedure of § 4.3 to design a lower triangular $H_{yr}(z)$. For the time delays ($b_0 = 0$) we have $n_0 = m_0 = 1$. Hence part (*b*) of Theorem 1 (or Corollary 2) does not apply for $i = 0$ and therefore (29) does not apply for $i = 0$. Thus none of the equations of § 4.3 that correspond to $i = 0$ should be considered. For the zero $a_1 = 1{\cdot}547$ ($b_1 = a_1^{-1}$), we have $n_1 = 0$, $m_1 = 1$. Also

$$U_1 = \begin{bmatrix} 0{\cdot}675 \\ 0{\cdot}739 \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{106}$$

and $\rho_1 = 1$.

6.1.1. *Design of* $h_1(z)$. In this case, l is the empty set and $\bar{I} = \{2\}$. Hence $\rho_1^l = 0$ and therefore $\xi_1^l = 1$, $V_1^l = 1$. Thus (42) has a solution for $i = 1$ and as a result b_1 should not be included in $\zeta_0(\lambda)$. Thus from (43) it follows that $\zeta_0(\lambda) = 1$. The case of b_0 should not be considered as mentioned above. Hence $\zeta_{\tau_0}(\lambda) = \zeta_0(\lambda) = 1$ and from (49) it follows that the first element of the column (diagonal element) is equal to $\lambda\,(= z^{-1})$. Then (50) needs to be satisfied for $i = 1$. Since the null space of $e_1^T U_1 V_1^l$ is the empty set, it follows that $W_1^l = 0$ and

$$\chi_1^2 = \chi_1^0 = \frac{b_1}{u_1} \tag{107}$$

We shall now proceed with the design of the second element of the column. In (54), (55) we have $q = 1$, $\bar{I}_1 = 2$, $\delta_1 = 1$. In (67), (69) the part corresponding to b_0 is omitted and so from (59), (67), (69), (107) it follows that

$$K_v = [(1 - b_1)b_1 \quad (1 - b_1)b_1^2 \quad \ldots \quad (1 - b_1)b_1^{v+1}] \tag{108}$$

$$T_1 = \frac{b_1 u_2}{u_1} \tag{109}$$

From $W_1^l = 0$ and (70) it follows that $T_2 = 0$. In this case $\Phi = I$ and from (76) we obtain

$$\theta_v = K_v^*(K_v K_v^*)^{-1} T_1 \tag{110}$$

Equations (59), (73), (108)–(110) yield

$$\beta_{i,j} = \frac{(1 + b_1)b_1^j u_2}{(1 - b_1^{2v+2})u_1}, \quad j = 0, \ldots, v \tag{111}$$

$$J_v = \frac{(1 + b_1)u_2^2}{(1 - b_1)(1 - b_1^{2v+2})u_1^2} \tag{112}$$

For $v = 0$ we get $\beta_{1,0} = 3{\cdot}095$, i.e. the design in (105). However the error caused by the interactions in this case is $J_0 = 9{\cdot}58$, which is quite large. Increasing the duration v of the interactions reduces J_v as (112) indicates. Since $b_1 < 1$ we can compute the limit:

$$\lim_{v \to \infty} J_v = \frac{(1 + b_1)u_2^2}{(1 - b_1)u_1^2} = 5{\cdot}58 \tag{113}$$

The designer can of course select a relatively small v, for which J_v is sufficiently close to the limit given by (113). A plot of J_v as a function of v is given in Fig. 2. One can see that a selection of $v = 4$, is satisfactory. It yields $J_v = 5{\cdot}65$. For $v = 4$, the second

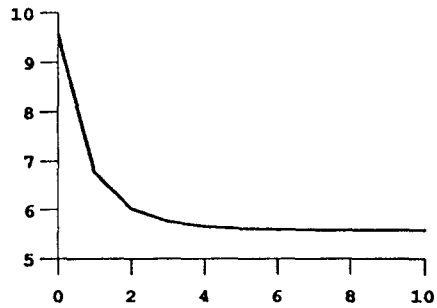


Figure 2. Example 1; J_v for column 1.

element of $h_1(z)$ becomes equal to $(1{\cdot}82 + 1{\cdot}18z^{-1} + 0{\cdot}76z^{-2} + 0{\cdot}49z^{-3} + 0{\cdot}32z^{-4})$ $(1 - z^{-1})z^{-1}$.

6.1.2. *Design of* $h_2(z)$. Since we require the first element of the column to be zero, we have $l = \{1\}$ and therefore $\bar{l}$ is the empty set. Hence $\rho_1^l = 1$, $\xi_1^l = 0$, $V_1^l = 0$. Then (42) does not have a solution for $i = 1$. As a result the second element (diagonal element) has to have a zero at $\lambda = b_1$. From (43) we obtain

$$\zeta_0(\lambda) = \frac{(1 - b_1^{-1})(\lambda - b_1)}{(1 - b_1)(\lambda - b_1^{-1})} \tag{114}$$

Then, since $\zeta_{\tau_0}(\lambda) = \zeta_0(\lambda)$, (49) implies that the diagonal element is $\lambda\zeta_0(\lambda)$. Substitution of z^{-1} for λ and a_1 $(= 1{\cdot}547)$ for b_1^{-1} produces the expression in (105).

6.2. *Example* 2

Consider the system:

$$P(z) = \begin{bmatrix} \dfrac{0{\cdot}90}{z - 0{\cdot}35} & \dfrac{0{\cdot}50}{z - 0{\cdot}35}z^{-1} & \dfrac{1{\cdot}00}{z - 0{\cdot}35} \\ \dfrac{2{\cdot}70}{z - 0{\cdot}60}z^{-1} & \dfrac{5{\cdot}80}{z - 0{\cdot}60}z^{-1} & \dfrac{0{\cdot}60}{z - 0{\cdot}60}z^{-1} \\ \dfrac{0{\cdot}40}{z - 0{\cdot}50} & \dfrac{-0{\cdot}45}{z - 0{\cdot}50} & \dfrac{1{\cdot}00}{z - 0{\cdot}50}z^{-1} \end{bmatrix} \tag{115}$$

The computation of the roots of $\det[P(z)]$ yields one zero outside the UC at $z = a_1 = 1{\cdot}3088$. We shall limit ourselves to the design of the first column of H_{yr}. Two different structures will be examined:

$$h_1 = \begin{bmatrix} x \\ x \\ 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} x \\ 0 \\ x \end{bmatrix} \tag{116}$$

The SVD of $P(a_1)$ yields the following left singular vector matrix:

$$U = \begin{bmatrix} 0{\cdot}125 & -0{\cdot}700 & -0{\cdot}703 \\ 0{\cdot}992 & 0{\cdot}0689 & 0{\cdot}107 \\ -0{\cdot}0267 & -0{\cdot}711 & 0{\cdot}703 \end{bmatrix} \tag{117}$$

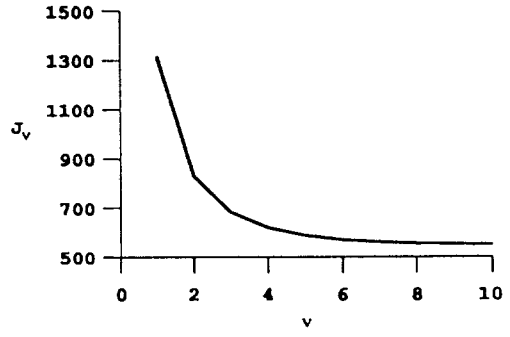
The two first columns of U form U_1. The third, u, is orthogonal to U_1. Then from Corollary 1 it follows that $u^* H_{yr}(a_1) = 0$ for all acceptable H_{yr}. Equation (117) suggests that if the first structure of (116) is selected, the value of the non-diagonal element at $z = a_1$ will have to be larger than the one for the second structure, because of the smaller corresponding element in u.

The consideration of the time delays (b_0) makes the situation even more favourable for the second structure. We have $n_0 = 1$, $m_0 = 2$ and

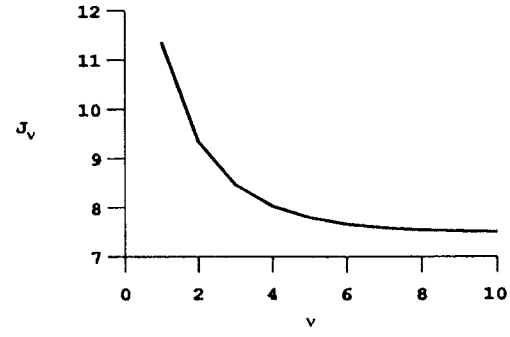
$$U_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{118}$$

The fact that the second row of U_0 is zero allows in the case of the second structure satisfaction of (37) for $i = 0$ without using any of the available degrees of freedom. This results in a non-zero w_0^l and T_2, and the additional freedom in choosing χ_0^3 through (79) reduces J_v even more.

The above qualitative observations are confirmed from the quantitative results of the design procedure. The corresponding plots of J_v versus v shown in Fig. 3 for both structures of (116) show a huge difference in the closed-loop interactions of the two structures.



(a)



(b)

Figure 3. Example 2; J_v for column 1: (a) $h_1 = [x \quad x \quad 0]^T$; (b) $h_1 = [x \quad 0 \quad x]^T$.

7. Conclusions

The results in this paper quantify the effects of the undesirable zeros and time delays of a multivariable discrete system on its closed-loop performance in a way that can be used for the direct design of the closed-loop transfer function matrix. The designer is provided with quantitative criteria for comparing different designs and evaluating the tradeoffs. The entire procedure is based on linear-algebra operations and its implementation on the computer is straightforward.

The design is based on the knowledge of a system model. Hence it may not be robust to model-plant mismatch. However it can be used in the first step of the controller design for the standard two-step internal model control design procedure, in which robustness properties are incorporated in the second step with the design of a low pass filter. Details on the filter design can be found in the literature (Zafiriou and Morari 1986 b, Zafiriou 1987).

Appendix A

Proof of Theorem 1

The following lemma will be used in the proof.

Lemma A.1 (Van Dooren *et al.* 1979, Vandewalle and Dewilde 1974)

Let a rational matrix $A(\lambda)$ of normal rank r have the following Laurent expansion at α:

$$A(\lambda) = \sum_{k=-l}^{\infty} (\lambda - \alpha)^k A_k(\alpha) \tag{A 1}$$

Define

$$T_k(\alpha) = \begin{bmatrix} A_{-l}(\alpha) & A_{-l+1}(\alpha) & \dots & A_k(\alpha) \\ 0 & A_{-l}(\alpha) & \dots & A_{k-1}(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{-l}(\alpha) \end{bmatrix} \tag{A 2}$$

$$\rho_k(\alpha) \overset{\text{det}}{=} \text{rank } [T_k(\alpha)] - \text{rank } [T_{k-1}(\alpha)] \tag{A 3}$$

Let p and z be a pole and a zero respectively of $A(\lambda)$ of orders ω_p, ω_z and degrees δ_p, δ_z, as these are defined from the Smith–McMillan form of $A(\lambda)$ (Van Dooren *et al.* 1979, Desoer and Schulman 1974).

The following hold:

(i) $\omega_p = -\min \{k \mid \rho_k \neq 0\}$

(ii) $\omega_z = \min \{k \mid \rho_k = r\}$

(iii) $\delta_p = \sum_{k=-\omega_p}^{-1} \rho_k$

(iv) $\delta_z = \sum_{k=0}^{\omega_z} (r - \rho_k)$

Proof of the Theorem

$\hat{P}(\lambda)^{-1}$ has as its poles exactly the zeros of $\hat{P}(\lambda)$ with the same order and degree

(Desoer and Schulman 1974). Hence, since b_i is a zero of $\hat{P}(\lambda)$ of order m_i, it is also a pole of $\hat{P}(\lambda)^{-1}$ and we can write

$$\hat{P}(\lambda)^{-1} = \sum_{k=1}^{m_i} (\lambda - b_i)^{-k} R_{i,k} + G_i(\lambda), \quad i = 0, 1, ..., f \tag{A 4}$$

where

$$\text{rank } [R_{i,m_i}] \neq 0, \quad i = 0, 1, ..., f \tag{A 5}$$

and $G_i(\lambda)$ has no poles at b_i, $i = 0, 1, ..., f$.

Postmultiplication of (A 4) with $\hat{H}_{yr}(\lambda)$ yields

$$\hat{P}(\lambda)^{-1}\hat{H}_{yr}(\lambda) = \sum_{k=1}^{m_i} (\lambda - b_i)^{-k} R_{i,k}\hat{H}_{yr}(\lambda) + G_i(\lambda)\hat{H}_{yr}(\lambda) \tag{A 6}$$

Now take a partial fraction expansion for each term in the sum of the right-hand side of (A 6) to obtain:

$$\begin{aligned}\hat{P}(\lambda)^{-1}\hat{H}_{yr}(\lambda) = \sum_{k=1}^{m_i} \left[\sum_{k=0}^{k-1} (\lambda - b_i)^{-k+h} R_{i,k} \frac{1}{h!} \hat{H}_{yr}^{(h)}(b_i) + R_{i,k} G_{k,i}(\lambda)\right] \\ + G_i(\lambda)\hat{H}_{yr}(\lambda) = \sum_{k=1}^{m_i} \left((\lambda - b_i)^{-k} \sum_{h=k}^{m_i} R_{i,h} \frac{1}{(h-k)!} \hat{H}_{yr}^{(h-k)}(b_i)\right) \\ + \sum_{k=1}^{m_i} R_{i,k} G_{k,i}(\lambda) + G_i(\lambda)\hat{H}_{yr}(\lambda)\end{aligned} \tag{A 7}$$

where $G_{k,i}(\lambda)$ has no poles at b_i. Also recall that $G_i(\lambda)$ has no poles at b_i either. Hence in order for Condition C.5 (ii) to hold, we must have for all $i = 0, ..., f$:

$$\sum_{h=k}^{m_i} R_{i,h} \frac{1}{(h-k)!} \hat{H}_{yr}^{(h-k)}(b_i) = 0, \quad k = 1, ..., m_i \tag{A 8}$$

Satisfaction of (A 8) is equivalent to requiring that the columns of

$$\left[\hat{H}_{yr}^{(0)}(b_i)^T \ ... \ \frac{1}{(m_i - 1)!} \hat{H}_{yr}^{(m_i - 1)}(b_i)^T\right]^T$$

are in the null space of N_i, where:

$$N_i \overset{\text{def}}{=} \begin{bmatrix} R_{i,m_i} & 0 & ... & 0 \\ R_{i,m_i - 1} & R_{i,m_i} & ... & 0 \\ \vdots & \vdots & & \vdots \\ R_{i,1} & R_{i,2} & ... & R_{i,m_i} \end{bmatrix} \tag{A 9}$$

We shall now proceed to determine the null space of N_i.

Postmultiply both sides of (A 4) with $\hat{P}(\lambda)$ to obtain:

$$I = \sum_{k=1}^{m_i} (\lambda - b_i)^{-k} R_{i,k}\hat{P}(\lambda) + G_i(\lambda)\hat{P}(\lambda) \tag{A 10}$$

Since I has no poles at $b_0, ..., b_f$, taking a partial fraction expansion leads to a

condition similar to (A 8), in exactly the same manner. Hence (A 10) yields

$$\sum_{h=k}^{m_i} R_{i,h} \frac{1}{(h-k)!} \hat{P}^{(h-k)}(b_i) = 0, \quad k = 1, \ldots, m_i \tag{A 11}$$

The equations implied by (A 11) for $k = l, \ldots, m_i$ can be put together in matrix form:

$$N_i \Big[\underbrace{0 \quad \ldots \quad 0}_{l-1} \quad \hat{P}^{(0)}(b_i)^{\mathrm{T}} \quad \ldots \quad \frac{1}{(m_i-l)!} \hat{P}^{(m_i-l)}(b_i)^{\mathrm{T}} \Big]^{\mathrm{T}} = 0, \quad l = 1, \ldots, m_i \tag{A 12}$$

The equations obtained from (A 12) for $l = 1, \ldots, m_i$ can be written together as:

$$N_i L_i = 0, \quad i = 0, \ldots, f \tag{A 13}$$

where

$$L_i \overset{\text{def}}{=} \begin{bmatrix} \hat{P}^{(0)}(b_i) & 0 & \ldots & 0 \\ \hat{P}^{(1)}(b_i) & \hat{P}^{(0)}(b_i) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ \dfrac{1}{(m_i-1)!} \hat{P}^{(m_i-1)}(b_i) & \dfrac{1}{(m_i-2)!} \hat{P}^{(m_i-2)}(b_i) & \ldots & \hat{P}^{(0)}(b_i) \end{bmatrix} \tag{A 14}$$

Hence the column space of L_i is a subspace of the null space of N_i. It will now be shown that it is exactly the null space of N_i.

As explained earlier, the order ω_p of the pole b_i of $\hat{P}(\lambda)^{-1}$ is equal to the order m_i of the zero b_i of $\hat{P}(\lambda)$, i.e.:

$$\omega_p(b_i) = m_i \tag{A 15}$$

Lemma A.1 will now be applied to $A(\lambda) = \hat{P}(\lambda)^{-1}$, for $\alpha = b_i$. From (A 1), (A 4), (A 5) it follows that $l = m_i$ and $A_{-k} = R_{i,k}$ for $k = 1, \ldots, m_i$. By using (A 15) and Lemma A.1 (iii), we obtain

$$\delta_p(b_i) = \sum_{k=-m_i}^{-1} \rho_k(b_i) = \operatorname{rank} [T_{-1}(b_i)] = \operatorname{rank} [N_i] \tag{A 16}$$

since $T_{m_i-1}(b_i) = 0$ and $T_{-1}(b_i)$ can be obtained from N_i simply by permuting its rows and columns.

By definition, the order ω_z of the zero b_i of $\hat{P}(\lambda)$ is equal to m_i:

$$\omega_z(b_i) = m_i \tag{A 17}$$

Lemma A.1 will now be applied on $A(\lambda) = \hat{P}(\lambda)$, for $\alpha = b_i$. In this case, since $\hat{P}(\lambda)$ is assumed to have no poles at b_i, we have $l \leqslant 0$ and $A_k = (1/k!)\hat{P}^{(k)}(b_i)$ for $k = 1, \ldots, m-1$. By using (A 17) and Lemma A.1 (iv), we obtain

$$\begin{aligned} \delta_z(b_i) &= \sum_{k=0}^{m_i} (r - \rho_k(b_i)) = \sum_{k=0}^{m_i-1} (r - \rho_k(b_i)) \\ &= m_i r - \operatorname{rank} [T_{m_i-1}(b_i)] = m_i r - \operatorname{rank} [L_i] \end{aligned} \tag{A 18}$$

since from Lemma A.1 (ii) we have $\rho_{m_i}(b_i) = r$, and we also have $T_{-1}(b_i) = 0$. $T_{m_i-1}(b_i)$ can be obtained from L_i by permuting its rows and columns.

The degree δ_z of the zero b_i of $\hat{P}(\lambda)$ is the same as the degree δ_p of the pole b_i of

$\hat{P}(\lambda)^{-1}$ and so from (A 16), (A 18) we obtain

$$\text{rank } [N_i] + \text{rank } [L_i] = m_i r \tag{A 19}$$

However, N_i and L_i are matrices of dimension $m_i r \times m_i r$. Therefore (A 13) and (A 19) imply that the column space of L_i is exactly the null space of N_i.

Hence from (A 8) it follows that Condition C.5 (ii) is satisfied if and only if the columns of

$$\left[\hat{H}_{yr}^{(0)}(b_i)^T \quad \dots \quad \frac{1}{(m_i-1)!}\hat{H}_{yr}^{(m_i-1)}(b_i)^T\right]^T$$

are in the column space of L_i. From (11) it follows that the first $n_i r$ rows and the last $n_i r$ columns of L_i are identically zero. Hence

$$\hat{H}_{yr}^{(k)}(b_i) = 0, \quad k = 0, \dots, n_i - 1$$

which implies that

$$\hat{H}_{yr}(\lambda) = (\lambda - b_i)^{n_i}\hat{H}_i(\lambda) \tag{A 20}$$

where $\hat{H}_i(\lambda)$ has no poles at b_i. Equation (A 20) completes the proof of part (*a*) of Theorem 1. If $m_i = n_i$, then this is the only requirement since then $L_i = 0$. If, however, $m_i > n_i$ then rank $[L_i] \neq 0$ and additional requirements on $\hat{H}_i(\lambda)$ are necessary. We have for $l = n_i, \dots, m_i - 1$:

$$\frac{1}{l!}\hat{H}_{yr}^{(l)}(b_i) = \frac{1}{l!}\binom{l}{n_i} n_i!\hat{H}_i^{(l-n_i)}(b_i) = \frac{1}{(l-n_i)!}\hat{H}_i^{(l-n_i)}(b_i)$$

and so the requirement on $\hat{H}_i(\lambda)$ is that the columns of

$$\left[\hat{H}_i^{(0)}(b_i)^T \quad \dots \quad \frac{1}{(m_i-n_i-1)!}\hat{H}_i^{(m_i-n_i-1)}(b_i)^T\right]^T$$

are in the column space of M_i, where M_i is defined in (14). □

Appendix B

Proof of Theorem 3

The proof is by induction

$k = 1$:

$$G(z)z = C(zI - A)^{-1}zB + Dz = C(zI - A)^{-1}(A + zI - A)B + Dz$$
$$= C(zI - A)^{-1}AB + CB + Dz$$

$k = n$:

Let

$$G(z)z^n = C(zI - A)^{-1}A^nB + \sum_{l=1}^{n} CA^{l-1}Bz^{n-l} + Dz^n \tag{B 1}$$

hold.

$k = n + 1$:

From (B 1) it follows that

$$G(z)z^{n+1} = C(zI - A)^{-1}A^nzB + \sum_{l=1}^{n} CA^{l-1}Bz^{n+1-l} + Dz^{n+1}$$

and by using the result for $k = 1$ we obtain

$$G(z)z^{n+1} = C(zI - A)^{-1}A^{n+1}B + \sum_{l=1}^{n+1} CA^{l-1}Bz^{n+1-l} + Dz^{n+1} \qquad \square$$

ACKNOWLEDGMENTS

Support from the National Science Foundation and the Department of Energy is gratefully acknowledged.

REFERENCES

ÅSTRÖM, K. J., and WITTENMARK, B., 1984, *Computer Controlled Systems* (Englewood Cliffs, NJ: Prentice-Hall).

CALLIER, F. M., and DESOER, C. A., 1982, *Multivariable Feedback Systems* (New York: Springer-Verlag).

DESOER, C. A., and SCHULMAN, J. D., 1974, *I.E.E.E. Trans. Circuits Syst.*, **21**, 3.

DOYLE, J. C., 1984, Lecture Notes. ONR/Honeywell Workshop on Advances on Multivariable Control.

GARCIA, C. E., and MORARI, M., 1982, *I&EC Process Des. Dev.*, **21**, 308; 1985, *Ibid.*, **24**, 472.

KAILATH, T., 1980, *Linear Systems* (Englewood Cliffs, NJ: Prentice-Hall).

LAUB, A. J., and MOORE, B. C., 1978, *Automatica*, **14**, 557.

SILVERMAN, L. M., 1969, *I.E.E.E. Trans. autom. Control*, **4**, 270.

STEWART, G. C., 1973, *Introduction to Matrix Computations* (New York: Academic Press).

VANDEWALLE, J., and DEWILDE, P., 1974, *I.E.E.E. Trans. autom. Control*, **19**, 608.

VAN DOOREN, P. M., DEWILDE, P., and VANDEWALLE, J., 1979, *I.E.E.E. Trans. Circuits Syst.*, **26**, 180.

VIDYASAGAR, M., 1985, *Control System Synthesis* (Cambridge, Mass: MIT Press).

ZAFIRIOU, E., 1987, A Methodology for the synthesis of robust control systems for multivariable sampled-data processes. Ph.D. thesis, California Institute of Technology.

ZAFIRIOU, E., and MORARI, M., 1985, *Int. J. Control*, **42**, 855; 1986 a, *Ibid.*, **44**, 711; 1986 b, *Proc. American Control Conf.*, Seattle, WA, p. 1; 1987, *Proc. American Control Conf.*, Minneapolis, MN, p. 649.

ZAMES, G., 1981, *I.E.E.E. Trans. autom. Control*, **26**, 301.