Sequential design of linear multivariable systems

Prof. D.Q. Mayne, M.Sc. (Eng), Ph.D., D.I.C., C.Eng., M.I.E.E.

Indexing terms: *Control-system c.a.d., Multivariable control systems*

Abstract

The paper describes and critically assesses, in the light of recent advances, the sequential return difference method for the computer-aided design of linear multivariable control systems. In this method, a sequence of single-loop designs, using classical procedures, yields a multivariable design satisfying various criteria such as stability, disturbance attenuation, low interaction and integrity.

1 Introduction

Rosenbrock, in a pioneering paper,[1] redirected attention to extending the immensely successful classical design procedures to multivariable problems. This paper proposed a powerful design technique, based on the inverse Nyquist array, for designing linear multivariable systems, but also provided a useful theoretical basis and a substantial impetus for further studies in this area. The subsequent developments include the important work by MacFarlane and his colleagues[2] on the characteristic locus method as well as the sequential return defference method[3-5] discussed here.

The design objectives considered in these papers include the following:

(*a*) stability
(*b*) insensitivity to parameter variations
(*c*) insensitivity to external disturbances (and possibly complete rejection of specific types of disturbances)
(*d*) low interaction
(*e*) loop by loop implementability
(*f*) integrity (maintenance of stability in the face of component change or failure)
(*g*) controller simplicity.

Different applications will assign different priorities to these objectives and hence make different trade offs. It is desirable to have a design method which is flexible enough to concentrate on some objectives and ignore others. For example, integrity in the face of sensor failure may well require a reduction in performance (e.g. an increase in sensitivity to external disturbances); if the latter is judged to be more important, the designer should be free to disregard the former.

2 Design objectives

It is desirable to specify the design objectives quantitatively. To do this we must specify the system being considered; this is shown in Figs. 1 and 2. Thus we consider the regulator problem (the problem of regulating available outputs), and assume initially (for simplicity) that the plant transfer function $G(s)$ and controller transfer function $G_c(s)$ are square $(m \times m)$. $R(s) \triangleq G(s)G_c(s)$ is the loop transfer-function matrix and $T(s) \triangleq I + R(s)$ is the return difference. Let r denote the desired output, d the disturbance, u the input to the plant, y the output of the plant and $e \triangleq r - y$ the error. All these quantities are vectors of dimension m. Let the complete open-loop system (plant and controller) have the following state vector description:

$$\dot{x} = Ax + Be \quad (1)$$

$$y = Cx + d \quad (2)$$

We can express the transfer function $R(s)$ (from e to y) in terms of (A, B, C) as follows:

$$R(s) = C(sI - A)^{-1}B \quad (3)$$

The open-loop characteristic polynomial is

$$\rho_0(s) = \det(sI - A) \quad (4)$$

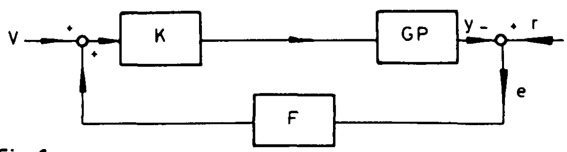

Fig. 1
Closed-loop system

The closed-loop system satisfies eqns. 1 and 2 with e replaced by $r - y$, and hence has the state vector description

$$\dot{x} = (A - BC)x + Br - Bd \quad (5)$$

$$y = Cx + d \quad (6)$$

The closed-loop characteristic polynomial is

$$\rho_c(s) = \det(sI - A + BC) \quad (7)$$

The closed-loop transfer from r to y is

$$H(s) = T^{-1}(s)R(s) = R(s)T^{-1}(s) \quad (8)$$

where

$$T(s) = I + R(s) \quad (9)$$

The closed-loop transfer function from d to y is

$$H_d(s) = T^{-1}(s) \quad (10)$$

The system described above has two limitations. Not all control problems require regulation of the measured outputs; an example of this arises in the control of turbofans.[6] A second limitation arises from the fact that there may be more outputs than inputs; how these extra outputs may be employed is discussed later.

We are now in a position to express the design criteria quantitatively.

(*a*) *Stability*
The closed-loop system is (asymptotically) stable if the zeros of $\rho_c(s)$ lie in C_l (the open left half of the complex plane C). It follows from eqns. 4 and 7 (proofs of all results stated in this paper are given in References 3, 4 and 5) that

$$\det[T(s)] = \rho_c(s)/\rho_0(s) \quad (11)$$

Since the zeros of $\rho_c(s)$ are the zeros of $\det[T(s)]$, stability of the closed-loop system can be assessed in the usual way from the Nyquist locus of $\det[T(j\omega)]$. Thus if the open-loop system is stable, all the zeros of $\rho_0(s)$ lie in C_l], then the closed-loop system is stable if the locus of $\det[T(j\omega)]$ does not encircle the origin. More accurately, let the contour D consist of the imaginary axis, indented to avoid any zeros of $\rho_0(s)$ on this axis, and a semicircle in the right halfplane, large enough to include all zeros of $\rho_0(s)$; then the image of D under $\det[T(s)]$ encircles the origin clockwise $N_c - N_0$ times, where N_c is the number of zeros of ρ_c in D and N_0 the number of zeros of ρ_0 in D.

(*b*) *Insensitivity to parameter variation*
Owing to variation in parameters, let $R(s)$ vary by $\delta R(s)$, and let $\delta H(s)$ denote the corresponding 1st-order variation in $H(s)$. Then it can be shown that

$$\|\delta H(s)H^{-1}(s)\| \leqslant \|T^{-1}(s)\| \quad (12)$$

Hence the effect of parameter variations on $H(j\omega)$ in the frequency band $[0, \omega_1]$ is small if

$$\|T^{-1}(j\omega)\| \ll 1, \quad \text{for all} \quad \omega \in [0, \omega_1] \quad (13)$$

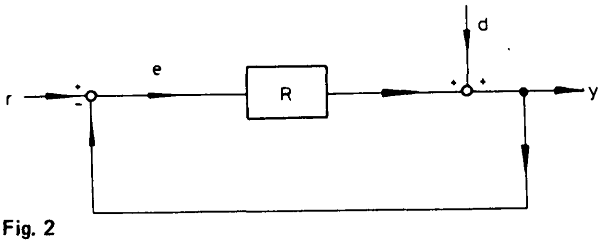

Fig. 2
Equivalent closed-loop system

Paper 8313C, first received 12th December 1978 and in revised form 19th February 1979

Prof. Mayne is with the department of Computing and Control, Imperial College of Science & Technology, London SW7 2BZ, England

0020–3270/79/060568 + 05 $01·50/0

(c) Insensitivity to external disturbances
It follows from eqn. 10 that the effect of the disturbance on the output is very much reduced in the frequency band $[0, \omega_1]$ if eqn. 13 is satisfied.

Sometimes complete rejection (at all outputs) of a class of disturbances [in the sense that $e(t) \to 0$ as $t \to \infty$] is required. A common example is the suppression of a constant disturbance which can be achieved in the single-variable case by employing integral action. In the multivariable case, a (scalar) disturbance is completely rejected if the closed-loop system is stable and $\beta(s)$ is a factor of the numerator polynomial of $T^{-1}(s)$, i.e. if

$$T^{-1}(s) = [1/\rho_c(s)]\,[\beta(s)N(s)] \tag{14}$$

where $N(s)$ is a polynomial matrix, and $\beta(s)$ is a polynomial whose zeros are those poles of $\bar{d}(s)$ (the Laplace transform of d) which lie in C_r, the *closed* right half of the complex plane C. Thus a constant disturbance is rejected in the multivariable case if the return difference has the form

$$T^{-1}(s) = [1/\rho_c(s)]\,[sN(s)] \tag{15}$$

(d) Low interaction
The condition given by eqn. 13 is satisfied if the loop transfer function $R(s)$ is large (i.e. feedback is tight) in the sense that

$$\|R^{-1}(j\omega)\| \ll 1, \quad \text{for all} \quad \omega \in [0, \omega_1] \tag{16}$$

If eqn. 16 is satisfied then

$$H(j\omega) \doteq I \quad \text{for all} \quad \omega \in [0, \omega_1] \tag{17}$$

so that tight feedback automatically ensures low interaction at low frequencies.

As $\omega \to \infty$, $R(j\omega) \to 0$ (our system is strictly proper) so that $T(j\omega) \to I$ and $H(j\omega) \to R(j\omega)$. Therefore, low interaction at high frequencies requires that G_c be chosen so that $R(j\omega)$ is approximately diagonal at these frequencies.

(e) Loop by loop implementability
If the design procedure is such that the system is stable with loops $1, \dots, i$ closed (and the remaining loops open), for $i = 1, 2, \dots, m$, then it is possible to implement the controller by closing, in sequence, loops $1, 2, \dots, m$ rather than closing all the loops simultaneously.

(f) Integrity
Let $T^\alpha(s)$ denote the return difference under fault α, and let A denote the set of possible faults. Then the closed-loop system is stable under all fault conditions if the Nyquist loci of det $[T(j\omega)]$ and det $[T^\alpha(j\omega)]$ for all $\alpha \in A$, satisfy the usual encirclement criterion.

Summarising, the overriding objective of stability is satisfied if, and only if, det $[T(j\omega)]$ satisfies the appropriate encirclement criterion. The performance objectives *(b)*, *(c)* and *(d)* are met at low frequencies ($\omega \in [0, \omega_1]$) if

$$\|T^{-1}(j\omega)\| \ll 1, \quad \text{for all} \quad \omega \in [0, \omega_1]$$

3 Sequential return difference method

It is obviously desirable to restrict the poles of the controller $G_c(s)$ to lie in C_l. It can be shown that, as the feedback becomes tight [more precisely as $\alpha \to \infty$, where $R(s) = \alpha\tilde{R}(s)$, and $\tilde{R}(s)$ is nonsingular], the finite zeros of $\rho_c(s)$ tend to the zeros of det $R(s)$ {or det $[\tilde{R}(s)]$}. Hence it is also desirable to restrict the zeros of det $[G_c(s)]$ to lie in C_l. Rosenbrock[1] has shown that any nonsingular rational $G_c(s)$ satisfying these pole-zero conditions has the representation

$$G_c(s) = PK(s)F(s) \tag{18}$$

where P is a column permutation matrix (equivalent to relabelling of the inputs), $K(s)$ is a unimodular matrix (representing a sequence of elementary column operations), and $F(s)$ a nonsingular diagonal rational matrix $\{F(s) = \text{diag}\,[f_1(s), \dots, f_m(s)]\}$. Because (modulo a sign change) det$(P) = 1$ and det $[K(s)] = 1$, we have

$$\det[G_c(s)] = \det[F(s)] \tag{19}$$

so that only F affects the poles and zeros of det $[G_c(s)]$. Hence (bearing in mind the limitations of our system description), the designer must choose suitable values for P, K and F.

Most of the design methods separate the design into two phases; first, the choice of P and K to satisfy various criteria, e.g. diagonal dominance in the inverse Nyquist array (i.n.a.) method, 'equalisation' of single-loop transfer functions in the sequential return difference (s.r.d.) method, or 'high-frequency alignment' in the characteristic locus (c.l.) method, although the c.l. method has a prior stage of compensation employing feedback of extra outputs and does not restrict F to be diagonal. Let us assume initially that P and K have been chosen and examine the design of F. We can then return to the design of P and K.

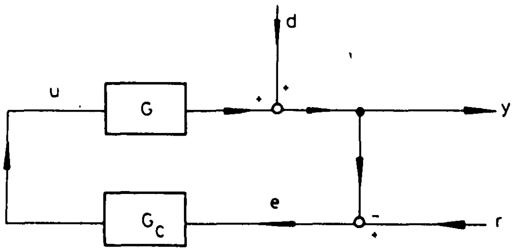

Fig. 3
Structure of controller

3.1 Design of F

Let $Q(s) \triangleq G(s)PK(s)$ denote the forward-path transfer function. Clearly, $R(s) = Q(s)F(s)$ (see Fig. 3). The sequential return difference method proceeds as follows:

(i) $f_1(s)$ is chosen (using classical single-loop design techniques) so that the scalar return difference $t_1(s)$ defined by

$$t_1(s) = 1 + f_1(s)q_{11}(s) \tag{20}$$

is satisfactory [$t_1(j\omega)$ has a sufficiently large magnitude in the range $[0, \omega_1]$ to satisfy the performance requirement, and the Nyquist locus of $t_1(j\omega)$ satisfies the Nyquist criteria to ensure stability].

(ii) Loop 1 is then closed, yielding the system S_1, and the *modified* forward-path transfer function $Q^1(s)$ is calculated. By virtue of of eqn. 1, the system S_1 (with loop 1 closed) is stable.

(iii) $f_2(s)$ is then chosen so that the scalar return difference $t_2(s)$ (seen on breaking loop 2) defined by

$$t_2(s) = 1 + f_2(s)q^1_{22}(s) \tag{21}$$

is satisfactory (since S_1 is stable, stability of S_2, the system with loops 1 and 2 closed, is ensured if the Nyquist locus of $t_1(j\omega)$ does not encircle the origin).

(iv) Loop 2 is then closed, yielding the (stable) system S_2, and the modified transfer function $Q^2(s)$ calculated etc.

In general, suppose loops 1 to i − 1 have been closed and the forward-path transfer function $Q^{i-1}(s)$ of S_{i-1} has been obtained. Then $f_i(s)$ is chosen so that:

$$t_i(s) = 1 + f_i(s)q^{i-1}_{ii}(s) \tag{22}$$

is satisfactory. $Q^i(s)$ is then calculated using

$$q^i_{r,s}(s) = q^{i-1}_{r,s}(s)/t_i(s), \quad \text{if } r = i \quad \text{or} \quad s = i$$

$$q^i_{r,s}(s) = q^{i-1}_{r,s}(s) - f_i(s)q^{i-1}_{r,i}(s)q^{i-1}_{i,s}(s)/t_i(s) \quad \text{otherwise} \tag{23}$$

In a practical computer-aided procedure, $Q^i(s)$ would be calculated at a set of frequencies ($s = j\omega_1, j\omega_2 \dots j\omega_N$) using eqn. 23 or directly from the specification of S_{i-1}. However, algebraic formulas for updating Q^i are given in References 4 and 5. Let $F_i(s) \triangleq \text{diag}\,[f_1(s), \dots, f_i(s), 0, \dots, 0]$ and $T_i(s) \triangleq I + Q(s)F_i(s)$. Then as shown in References 3 and 5, $Q^i(s) = T_i^{-1}(s)Q(s)$ and

$$\det[T_i(s)] = \det[T_{i-1}(s)]\,t_i(s) \tag{24}$$

so that

$$\det[T_i(s)] = \prod_{j=1}^{i} t_j(s), i = 1, \dots, m \tag{25}$$

Hence det $[T(s)] = \prod_{j=1}^{m} t_j(s)$, so that the number of encirclements of the origin by the Nyquist locus of det $[T(j\omega)]$ is equal to the sum of the encirclements by the Nyquist loci of $t_1(j\omega), t_2(j\omega), \dots, t_m(j\omega)$.

If the above steps can be carried out successfully, then the resultant system S_m has the following properties:

(a) It is stable since the Nyquist locus of det $[T(j\omega)]$ satisfies the required encirclement criterion.

(b) It attenuates the effect of parameter variations and external disturbances in $[0, \omega_1]$, since for $\omega \in [0, \omega_1]$

$$\|T^{-1}(j\omega)\| \doteq \|[Q(j\omega)F(j\omega)]^{-1}\|$$

$$\leqslant \|Q^{-1}(j\omega)\| \Big/ \sum_{i=1}^{m} |f_i(j\omega)| \tag{26}$$

(*c*) It is loop by loop implementable, since $S_1^1, S_2, \ldots, S_m$ are all stable.

(*d*) It is stable for the following fault condititions: simultaneous opening of loops $2-m$, or loops $3-m$, or ..., or loop m.

A disturbance d is completely rejected, at all outputs, if the denominator of each $f_i(s)$ has $\beta(s)$ as a factor [where $\beta(s)$ is a polynomial whose zeros are the unstable poles of the Laplace transform of d]. Thus a constant disturbance ($Ld = c/s$) is completely rejected if each $f_i(s)$ has s in the denominator (integral action).

To achieve integrity against *other* fault conditions (parameterised by α), it is necessary to choose f_i, for each i, so that not only is $t_i(s)$ *satisfactory* (in the sense defined above), but also the return difference $t_i^\alpha(s)$ satisfies the Nyquist criterion for stability, for all fault conditions α in A.

3.2 Design of P and K

The procedure for designing F is simple, and, apart from the initial choice of P and K, one often used in practice. However, the method will only succeed if the transfer functions $q_{11}(s), q_{22}^1(s), \ldots, q_{mm}^{m-1}(s)$ 'seen' in designing loops $1, 2, \ldots, m$ are satisfactory (e.g. have no zeros in C_r). Conditions under which there exists a P and K such that $q_{ii}^{i-1}(s), i = , \ldots, m$, are satisfactory have been established by Rosenbrock. Let the plant transfer function $G(s)$ be nonsingular and rational, and let all the poles of $G(s)$ and zeros of det $[G(s)]$ lie in C_l. Then there exists a permutation matrix P and a compensation matrix $K(s)$ {where det $[K(s)] = 1$} such that $Q(s) = G(s)PK(s)$ is nonsingular and diagonal, and whose elements have all their poles and zeros in C_l. Since, in this case, feedback around the ith loop does not affect any other loop, it follows that $q_{ii}^{i-1}(s) = q_{ii}(s), i = 1, \ldots, m$, and hence that $q_{ii}^{i-1}(s), i = 1, \ldots, m$, are all satisfactory. In practice, of course, one hopes to achieve satisfactory transfer functions with a PK simpler than one which makes Q diagonal.

Another useful result is the following. If feedback is tight in the band $[0, \omega_1]$, then {since det $[Q(s)]$ = det $[Q(s)]$ det $[PK(s)]$ = det $[G(s)]$ } we have

$$\det [G(j\omega)] = \det [Q(j\omega)] \doteqdot \prod_{i=1}^{m} q_{ii}^{i-1}(j\omega) \tag{27}$$

for all $\omega \in [0, \omega_1]$. Thus, if det $[G(s)$ has zeros in C_r relatively close to the origin, so will some of the transfer functions $q_{ii}^{i-1}(s)$; in this case it is not possible to have m tight loops. However, if G satisfies Rosenbrock's conditions, satisfactory transfer functions are achieved if P and K are chosen suitably; the task of K is to distribute det $[G(j\omega)]$ suitably between the loops.

As an example consider

$$G(s) = [1/(s+1)^2] \begin{bmatrix} 1-s & (1/3-s) \\ 2-s & 1-s \end{bmatrix} \tag{28}$$

If $PK = I$, then $q_{11}(s) = (1-s)/(s+1)^2$ and it is not possible to choose f_1 so that the first loop is tight [because $q_{11}(s)$ has a zero in C_r]. However

$$\det [G(s)] = (1/2)(s+1)/(s+1)^4 \tag{29}$$

so that in principle there exists a PK such that q_{11} and q_{22}^1 are satisfactory. In fact, with

$$PK = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \tag{30}$$

then

$$Q(s) = [1/(s+1)^2] \begin{bmatrix} (1/3)+s & (1/3)-s \\ s & 1-s \end{bmatrix} \tag{31}$$

Clearly $q_{11}(s)$ is satisfactory. Making use of eqns. 27 and 29 we obtain

$$\begin{aligned} q_{22}^1(s) &\doteqdot \det [G(s)]/q_{11}(s) \\ &= 1/[(s+1)(3s+1)] \end{aligned} \tag{32}$$

which is also satisfactory. This approximation is valid at low frequencies.

At high frequencies, $Q^1(s), Q^2(s), \ldots, Q^m(s)$ all approach $Q(s)$ so that, for our example, $q_{11}(s) = (1/3+s)/(s+1)^2$ has an asymptotic phase lag of $\pi/2$, and $q_{22}^1(s) \doteqdot (1-s)/(s+1)^2$ has an asymptotic phase lag or $3\pi/2$. Without compensation, $Q(s) = G(s)$, so $q_{11}(s) = (1-s)/(s+1)^2$ and $q_{22}^1(s) = q_{11}(s)$ both have asymptotic phase lag of $3\pi/2$.

If $G(s)$ satisfies Rosenbrock's conditions, and $m = 2$, then only P and $K(s)$ need to be chosen to make $q_{11}(s)$ satisfactory. Tight feedback in the first loop will then automatically make $q_{22}^1(s)$ satisfactory, since at low frequencies $q_{22}^1(s) \doteqdot \det [G(s)]/q_{11}(s)$. An extension to this procedure for $m > 2$ is described in Reference 5.

It must be admitted, however, that this procedure for choosing $K(s)$ is not particularly simple. Hence it is appropriate to ask whether precompensation techniques employed in alternative design procedures (which must implicitly satisfy similar requirements) can be used in s.r.d. Recent work[7] on multivariable root loci appears relevant. $G(s)$ can be expanded in inverse powers of s as follows:

$$G(s) = (1/s)G_1 + (1/s^2)G_2 + , \ldots, \tag{33}$$

Similarly

$$Q(s) = (1/s)Q_1 + (1/s^2)Q_2 + , \ldots, \tag{34}$$

If Q_1 is nonsingular and $F(s) = \alpha I$, then as $\alpha \to \infty$, the finite closed-loop poles of H tend to the zeros of det $[Q(s)]$ = det $[G(s)]$, and the remaining m closed-loop poles tend to the zeros of det $(sI - \alpha Q_1)$. Hence, if G_1 is nonsingular and $PK(s) = G_1^{-1}$ so that $Q_1 = I$, then the m unbounded closed-loop poles tend to $-\alpha$, a 1st-order pattern. If $Q_1 = 0$ but $Q_2 = I$, then $F(s) = \alpha I$ together with compensator $PK(s) = G_2^{-1}$ results in a system with $2m$ unbounded closed-loop poles tending to $\beta_i \pm j\sqrt{\alpha}$ (where the offset $\beta_i, i = 1, \ldots, m$, is easily calculated) as $\alpha \to \infty$, i.e. a 2nd-order pattern. Of course more complicated patterns may occur. Thus, if G_1 is singular and G_2 is nonsingular, the asymptotic pattern will be a combination of 1st- and higher-order patterns.

The situation can be re-examined in the frequency domain. Returning to our example, it can be seen that the *sum* of the asymptotic phase lags seen in designing each loop is 3π if no compensation is employed, and 2π if the compensator defined by eqn. 30 is used. But if the compensator were such that $Q(s)$ is approximately diagonal or triangular at high frequencies, then [since $\prod_{i=1}^{m} q_{ii}(s) = \det Q(s)$ if $Q(s)$ is diagonal or triangular] it follows that the sum of the asymptotic phase lags would equal the asymptotic phase lag of det $[Q(j\omega)]$. In our example, the asymptotic phase lag of det $[Q(j\omega)]$ is $3\pi/2$ (which is *lower* than both the previous values). Hence an obvious objective of the precompensator would be to ensure that the sum of the asymptotic phase lags of $q_{ii}(j\omega), i = 1, \ldots, m$, is equal to the asymptotic phase lag of det $[Q(j\omega)]$, confirming the desirability of MacFarlane's high frequency alignment.

To illustrate this consider our example. Expanding in inverse powers of s yields

$$G(s) = -(1/s)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + (1/s^2)\begin{bmatrix} 3 & 7/3 \\ 4 & 3 \end{bmatrix} + , \ldots, \tag{35}$$

Clearly G_1 is singular, so we cannot set $PK(s)$ equal to the inverse of G_1. However, with

$$PK(s) = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix} \tag{36}$$

{so that det $[PK(s)] = +1$} we obtain

$$\begin{aligned} Q(s) &= G(s)PK(s) \\ &= (1/s)\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + (1/s^2)\begin{bmatrix} -3 & 2/3 \\ -4 & 1 \end{bmatrix} + , \ldots, \end{aligned} \tag{37}$$

so that the asymptotic phase lag of $q_{11}(s)$ is $\pi/2$ and of $q_{22}(s)$ is π, yielding a total lag of $3\pi/2$. Hence, asymptotically this compensator superior to the preceding two. With this compensator

$$Q(s) = [1/(s+1)^2] \begin{bmatrix} s-1 & 2/3 \\ s-2 & 1 \end{bmatrix} \tag{38}$$

Because $q_{11}(s)$ has a zero in C_r, s.r.d. requires that loop 2 (which has an asymptotic phase lag of π) be designed first. Suppose (for illustrative purposes) we choose $f_2(s) = \alpha(s+\alpha)$, then (for large α) $t_2(s) = 1 + q_{22}(s)f_2(s) \doteqdot (s^2 + \alpha s + \alpha^2)/(s+1)^2$, which is satisfactory. Closing the second loops yields $q_{11}^2(s)$ (the transfer function seen in designing loop 1) where

$$q_{11}^2(s) = q_{11}(s) - f_2(s)q_{21}(s)q_{12}(s)/t_2(s) \doteq \frac{s^3 + (\alpha/3)s^2 + (\alpha^2/3)s + \alpha^2/3}{(s+1)^2(s^2+\alpha s+\alpha^2)} \tag{39}$$

has an asymptotic phase lag of $\pi/2$ and is relatively simple to handle.

3.3 Additional outputs

Suppose the plant has an output (vector) z in addition to the output y which is to be regulated. These outputs can be combined with y to give an effective output y', where

$$\bar{y}'(s) = \bar{y}(s) + L(s)\bar{z}(s) \tag{40}$$

(the overbar denotes Laplace transform) in such a way that the resultant (square) forward-path transfer function $G'(s)$ is superior to $G(s)$. [If the transfer function from u to z is $G_z(s)$, then the transfer function from u to y' is $G'(s) = G(s) + L(s)G_z(s)$.] By superior, we mean that the Nyquist locus of det $[G'(s)]$ is better (e.g. has less phase lag) than that of det $[G(s)]$. To evaluate the effect of incorporating a *single* extra observation, let z be a scalar, $G_z(s) = g_z^T(s)$ a row vector (dimension m) of rational functions, and restrict $L(s)$ to be a real column vector l. Then

$$\begin{aligned}\det[D'(s)] &= \det[G(s) + lg_z^T(s)]\\ &= \det[G(s)]\det[I + G^{-1}(s)\,lg_z^T(s)]\\ &= \det[G(s)]\,[1 + g_z^T(s)G^{-1}(s)l]\end{aligned} \tag{41}$$

Thus, in our example, if

$$g_z^T(s) = [s/(s+1)^2, 0] \tag{42}$$

and

$$\begin{aligned}1 + g_z^T(s)G^{-1}(s)l &= 1 + (3s/(s+1))[l_1(1-s) + l_2(s-1/3)]\\ &= [3(l_2-l_1)s^2 + (3l_1 - l_2 + 1)s + 1]/(s+1)\end{aligned} \tag{43}$$

which becomes $(s+1)$ with $l_1 = 2/3$ and $l_2 = 1$, providing useful extra phase shift. Indeed, $G'(s)$ becomes

$$G'(s) = [1/(s+1)^2]\begin{bmatrix}1-s/3 & 1/3-s\\ 2 & 1-s\end{bmatrix} \tag{44}$$

and

$$\det[G'(s)] = 1/[3(s+1)^2] \tag{45}$$

which has an asymptotic phase lag of only π (compared with $3\pi/2$) which can be shared among the two loops with proper high-frequency compensation.

Expanding in inverse powers of s yields

$$G'(s) = (1/s)\begin{bmatrix}-1/3 & -1\\ 0 & -1\end{bmatrix} + ,\ldots, \tag{46}$$

which is nonsingular so that high-frequency compensation is simplified. Indeed improvement of high-frequency compensation provides a useful guide to the choice of $L(s)$.

If z is a vector rather than a scalar, each component can be treated in turn as above. Effectively, L is chosen to add suitable zeros to det $[G(s)]$; in our example, det $[G'(s)] = (s+1)$ det $[G(s)]$.

3.4 Critique of s.r.d.

Any method can be criticised if it forces the design to achieve more objectives than are required for the particular application, since satisfaction of the additional objectives is achieved at the expense of the main objectives. Both the s.r.d. and the inverse Nyquist array methods can be criticised from this point of view if integrity is not required; the former because it requires stability of the subsystems $S_1, S_2, \ldots, S_m$, and the latter because it requires integrity against any sensor failure.

A second, important criticism of the s.r.d. method is that design is an iterative process, and it is unlikely that the first choice of compensator $PK(s)$ and feedback controller $F(s)$ will be satisfactory.

4 Improved s.r.d. design procedure

The first criticism of the s.r.d. method (requiring stability of the subsystems $S_1, S_2, \ldots, S_m$) can be removed by assessing stability from the locus of det $[T(j\omega)]$ [or, equivalently, from the sum of the encirclement of the origin by the loci of $t_1(j\omega), t_2(j\omega), \ldots, t_m(j\omega)$]. The second criticism (that it provides no satisfactory means for refining the first design) arises from limitations on the initially proposed computational procedure, wherein each scalar return difference $t_i(s)$ $(i = 1, \ldots, m)$ was examined in turn. This was done because the original program[4,5] was algebraic in nature, and employed algebraic formulas in Reference 5 to calculate the forward-path transfer function $Q^i(s)$ of S^i, $i = 1, \ldots, m$, sequentially [the formulas show how to calculate $Q^i(s)$ given $Q^{i-1}(s)$ and $f_i(s)$]. A bad choice for, say, $f_3(s)$ might result in an unsatisfactory $q_{55}^4(s)$, and no method for correcting this was given.

Modern c.a.d. methods store a (state space) model of the current system (plant plus controller); using the model, the loci of $t_i(j\omega)$, $i = 1, \ldots, m$, and of det $[T(j\omega)]$ can be *simultaneously* displayed to the designer, and the effect of *any* parameter changes [in P, $K(s)$ or $F(s)$] on *all* of these loci can be easily observed. With this facility, the s.r.d. method can be improved as follows. We assume that additional outputs have already been incorporated as described in Section 3.3.

Relabel inputs (equivalent to choosing P) and outputs, and let $G(s)$ denote the resultant transfer function. Choose the real matrix K_h (high-frequency compensator) so that the asymptotic zeros of det $[I + \alpha G(s)K_h]$ as $\alpha \to \infty$ (i.e. the nonfinite closed loop poles with $F = \alpha I$) are satisfactory. [Recall that the finite closed-loop poles tend, as $\alpha \to 0$, to the zeros of $G(s)$.] If $G(s)$ is uniform so that $G(s) = (1/s^k)G_k + (1/s^{k+1})G_{k+1} + ,\ldots,$ where G_k is nonsingular then $K_h = G_k^{-1}$ will suffice; if not, more *ad hoc* methods, such as MacFarlane's alignment procedure, may be employed. (The crux of the problem is that if $G(s)$ is not uniform, a polynomial, rather than a scalar matrix, is needed to diagonalise or triangularise $G(s)$, and a typical compromise between simplicity and effectiveness is involved). Our *initial* controller has $K(s) = K_h$ and $F(s) = \alpha I$, where a suitable magnitude for α has to be guessed. The *initial* value of $Q(s)$ is $G(s)K(s) = G(s)K_h$.

The loci of $t_i(j\omega)$ [or $q_{ii}^{i-1}(j\omega)f_i(j\omega)$], $i = 1, \ldots, m$, and det $[T(j\omega)]$ corresponding to the current design are displayed. The design objectives are simply to ensure that $t_i(j\omega)$ [or $q_{ii}^{i-1}(j\omega)f_i(j\omega)$], $i = 1, \ldots, m$, have sufficiently large magnitudes in the band $[0, \omega_1]$, *and* that the locus of det $[T(j\omega)]$ obeys the appropriate Nyquist criterion {since det $[T(s)] = \prod_{i=1}^{m} t_i(s)$, the effect of parameter changes in $F(s)$ on det $[T(s)]$ can be easily predicted}. A typical design proceeds by setting $f_i(s)$ [$f_i(s) = \alpha$ in the initial design] equal to a conventional single-loop controller, such as $\alpha_i(1 + s\beta_i)/s$ or $\alpha_i(1 + s\beta_i)/(1 + s\gamma_i)$, $i = 1, \ldots, m$, and choosing the coefficients $(\alpha_i, \beta_i, \gamma_i)$ so that $t_i(j\omega)$, $i = 1, \ldots, m$, and det $[T(j\omega)]$ satisfy the above criteria. [Since $F(s)$ is diagonal, the asymptotic properties of the nonfinite closed-loop poles are preserved; e.g. if G_1 is invertible, $K_h = G_1^{-1}$, and $\alpha_i = (1 + s\beta_i)/s$, then the nonfinite closed-loop poles tend to $-\alpha_1\beta_1, -\alpha_2\beta_2, \ldots, -\alpha_m\beta_m$, i.e. to the roots of det $[sI - \mathrm{diag}(\alpha_1\beta_1, \alpha_2\beta_2, \ldots, \alpha_m\beta_m)]$ as $\alpha_i \to \infty$, $i = 1, \ldots, m$.)

The closed-loop system can now be examined by inspecting the frequency or step response of the closed-loop system, whose transfer function is $H(s) = Q^m(s)F(s)$. At this stage, it may be noted that interaction in the closed-loop system (which is low at low frequencies due to tight feedback, and at high frequencies due to K_h) may be too high at intermediate frequencies. Further compensation $K_m(s)$ [setting $K(s) = K_hK_m(s)$] such that det $[K_m(s)] = 1$ (at least asymptotically as $s \to \infty$) may be introduced to reduce interaction (such a $K_m(s)$ may be obtained as a sequence of elementary column operations).

To illustrate the procedure consider an unstable chemical reactor[2] described by

$$A = \begin{bmatrix}1{\cdot}38 & -0{\cdot}2077 & 6{\cdot}715 & -5{\cdot}676\\ -0{\cdot}5814 & -4{\cdot}29 & 0 & 0{\cdot}675\\ 1{\cdot}067 & 4{\cdot}273 & -6{\cdot}654 & 5{\cdot}893\\ 0{\cdot}048 & 4{\cdot}273 & 1{\cdot}343 & -2{\cdot}104\end{bmatrix}$$

$$B = \begin{bmatrix}0 & 0\\ 5{\cdot}679 & 0\\ 1{\cdot}136 & -3{\cdot}146\\ 1{\cdot}136 & 0\end{bmatrix} \quad C = \begin{bmatrix}1 & 0 & 1 & -1\\ 0 & 1 & 0 & 0\end{bmatrix}$$

The eigenvalues of A are 0·06318, 1·991, −5·067, −8·66, and the zeros of $G(s)$ are −1·192, −5·039.

Also

$$G_1 = CB = \begin{bmatrix}0 & -3{\cdot}146\\ 5{\cdot}679 & 0\end{bmatrix}$$

so the system is uniform, of 1st-order. For our initial design, set

$$K(s) = K_h = G_1^{-1}$$

and

$$F(s) = 10I$$

The Nyquist criteria for det $[T(j\omega)]$ is satisfied with this design and the closed-loop eigenvalues are $-1{\cdot}53 \pm j{\cdot}44$ and $-14{\cdot}3 \pm j{\cdot}19$. However, the loop gains are not tight, with $q_{11}(0) = 1{\cdot}96$ and $q_{22}^1(0) = 0{\cdot}43$ resulting in a closed-loop transfer function at zero frequency given by

$$H(0) = \begin{bmatrix} 1{\cdot}35 & 0{\cdot}9 \\ -0{\cdot}09 & 0{\cdot}79 \end{bmatrix}$$

which shows considerable offset and interaction. Since the asymptotic behaviour is 1st-order, no difficulty in increasing loop gain is expected, so $F(s)$ is changed to

$$F(s) = (1/s)\begin{bmatrix} 120(s+1) & 0 \\ 0 & 60(s+1) \end{bmatrix}$$

higher gain being introduced in loop 1 because of the larger offset in loop 1 (35% compared with 21% in loop 1), and the larger interaction indicated in $h_{12}(s)$ compared with $h_{21}(s)$. The resultant system is satisfactory, the step responses settling within 1 s; interaction is also low, the peak interaction being less than 2·5% on output 1 and less than 0·75% on output 2. The complete controller has a transfer function

$$G_c(s) = G_h F(s)$$

$$= [(s+1)/s]\begin{bmatrix} 0 & 10{\cdot}57 \\ 38{\cdot}24 & 0 \end{bmatrix}$$

Because of the widely differing phase angles in the elements of $G(s)$, it was not profitable to attempt to reduce interaction using a scalar matrix K_m.

5 Conclusion

We have outlined the sequential difference method, but have also taken advantage of recent developments in the design of multivariable control, in particular the work of MacFarlane and his colleagues, to critically assess this method and suggest various improvements. The amended version naturally now bears certain similarities to the characteristic locus method, the major difference being that attention is focused on the (real) transfer functions $q_{ii}^{i-1}(s)f_i(s), i = 1, \ldots, m$, rather than on the eigenvalues of $Q(s)$.

6 References

1 ROSENBROCK, H.H.: 'Design of multivariable control systems using the inverse Nyquist array', *Proc. IEE*, 1969, 116, (11), pp. 1929–1936
2 MACFARLANE, A.G.J., and KOUVARITAKIS, B.: 'A design technique for linear multivariable systems', *Int. J. Control*, 1977, 25, pp. 837–874
3 MAYNE, D.Q.: 'The design of linear multivariable systems', *Automatica*, 1973, 9, pp. 201–207
4 MAYNE, D.Q.: 'The effect of feedback in linear multivariable systems', *ibid.*, 1974, 10, pp. 405–412
5 MAYNE, D.Q.: 'The sequential return difference design method'. Proceedings of the SRC vacation school for postgraduate control engineers, 1976 (to be published)
6 SAIN, M.K., PECZKOWSKI, J.L., and MELSA, J.L.: 'Alternatives for linear multivariable control'. National Engineering Consortium, Chicago, 1978
7 OWENS, D.H.: 'Dynamic transformations and the calculation of multivariable root-loci, *Int. J. Control*, 1978, 28, pp. 345–352