KYBERNETIKA — VOLUME 11 (1974), NUMBER 2

Algebraic Approach to Discrete Stochastic Control

VLADIMÍR KUČERA

The paper presents a unified formulation and solution of various problems of stochastic control, viz. random disturbance compensation, random signal following, etc. for multivariable linear discrete systems. The algebraic method developed originally for the problems of deterministic control is succesfully applied. The synthesis procedure is reduced to solving a linear Diophantine equation in polynomial matrices. Due to the algebraic approach, the classical results are extended to unstable systems with possibly different number of inputs and outputs and not necessarily of full rank, and to a class of nonstationary random sequences with possibly singular correlation matrix.

INTRODUCTION

The design of optimum systems with random inputs is one of the most significant problems in optimal control. There are many related problems with many modifications to be found in the literature. We shall summarize some typical problems below.

(1) Following a random signal.

Given a system $\mathscr{S}$, find a controller $\mathscr{R}$ such that the system output $\boldsymbol{Y}$ follows a given random signal $\boldsymbol{C}$ in an optimal way, see Fig. 1.

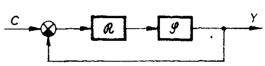

Fig. 1. Following a random signal.

(2) Compensating a random disturbance.

Given a system $\mathscr{S} = \mathscr{S}_1\mathscr{S}_2$ and a random disturbance $\boldsymbol{V}$ passing through the part $\mathscr{S}_1$ of $\mathscr{S}$, find a controller $\mathscr{R}$ that minimizes the effect of $\boldsymbol{V}$ on the system output $\boldsymbol{Y}$ in some specified sense, see Fig. 2.

(3) Following a random signal contaminated by noise.

Given a system $\mathscr{S}$ and a random signal $\boldsymbol{C}$ contaminated by additive noise $\boldsymbol{V}$, find a controller $\mathscr{R}$ such that the system output $\boldsymbol{Y}$ follows the $\boldsymbol{C}$ in an optimal manner, see Fig. 3.

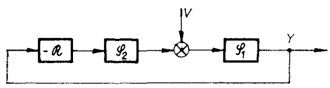

Fig. 2. Compensating a random disturbance.

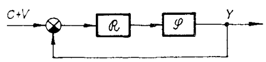

Fig. 3. Following a random signal contaminated by noise.

(4) Following a random signal in the presence of disturbance.

Given a system $\mathscr{S} = \mathscr{S}_1 \mathscr{S}_2$ and a random disturbance $\boldsymbol{V}$ passing through the part $\mathscr{S}_1$ of $\mathscr{S}$, find a controller $\mathscr{R}$ such that the system output $\boldsymbol{Y}$ follows a given random signal $\boldsymbol{C}$ in a prespecified sense, see Fig. 4.

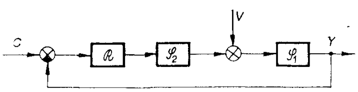

Fig. 4. Following a random signal in the presence of disturbance.

In all cases the closed-loop system is required to be stable and so is the system input. Otherwise the results would be of limited engineering relevance.

There are other problems of stochastic control which are combinations of the above and, therefore, will not be explicitely mentioned.

A recognized optimality criterion is the minimum of the sum of steady-state variances of certain stochastic components, i.e. the components of the follow-up error in problems (1), (3), (4) and the components of the system output in problem (2).

The above problems have been considered by many authors who applied different approaches. There are essentially two major directions – the complex-domain and the time-domain formulations. The complex-domain approach rests on the solution of a Wiener-Hopf-like equation by spectral factorization and can be found in [2; 7; 24; 30; 31]. The solution is restricted to stable systems with nonsingular impulse response matrix and to stationary random inputs with nonsingular correlation matrix. In [27] a modified approach has been applied to obtain the solution for the special case of single-input single-output possibly unstable systems. On the other hand, the time-domain approach is based on the solution of a matrix algebraic equation derived by Kalman [10; 11] and a comprehensive treatment can be found in [1; 6; 9; 26; 29]. The solution can easily by generalized to nonconstant or nonlinear systems but it requires the knowledge of the system state, which is rarely accessible in a real system. Moreover, solving matrix algebraic equations is not a simple task due to nonuniqueness of solutions [19; 20; 21; 22].

A similar status quo was also in the field of deterministic control. Recently, the author has developed a new algebraic theory of discrete deterministic optimal control [12; 13; 14; 15; 16; 17; 18] in an attempt to obtain a general solution well-adapted to machine processing. In this paper the algebraic approach is applied to the problems of discrete stochastic optimal control. The mathematical machinery needed to solve these problems is relatively very simple and is based on polynomial algebra. The synthesis procedure reduces to solving a linear Diophantine equation in polynomials or polynomial matrices, and can be effectively algorithmized. The method is general enough to accommodate unstable systems with different number of inputs and outputs and random signals with singular correlation matrix and possibly unbounded covariance matrix.

Problems (1) through (4) and other related problems, though different in nature, can be cast into a common scheme defined in the following sections. This will make it possible to present a unified general treatment of all cases.

PRELIMINARIES

Referring for details to [4; 18; 23; 33] we first summarize some preliminary results.

Let $\mathfrak{R}$ be a commutative ring. If an element $e \in \mathfrak{R}$ has a multiplicative inverse in $\mathfrak{R}$, we call e a unit of $\mathfrak{R}$. If $a, b \in \mathfrak{R}$, $b \neq 0$ we write $b \mid a$ to denote that b divides a. A greatest common divisor of $a, b \in \mathfrak{R}$ will be denoted as (a, b). Note that (a, b) is determined by a and b to within units in $\mathfrak{R}$.

Given the field $\mathfrak{F}$ of reals, let $\mathfrak{F}[z^{-1}]$ denote the ring of polynomials over $\mathfrak{F}$ in the indeterminate z^{-1}. If $a \in \mathfrak{F}[z^{-1}]$ then ∂a denotes the degree of a. By convention, $\partial 0 = -\infty$. The units of $\mathfrak{F}[z^{-1}]$ are polynomials of zero degree.

Let $\mathfrak{F}(z^{-1})$ denote the quotient field of $\mathfrak{F}[z^{-1}]$, i.e. the field of rational functions

$$a = \frac{q}{p} \tag{5}$$

with $p \neq 0$, $q \in \mathfrak{F}[z^{-1}]$. Then we denote $\mathfrak{F}\{z^{-1}\}$ the ring of elements (5) such that $(p, z^{-1}) = 1$, i.e. the ring of realizable rational functions. They can be written as

$$a = \alpha_0 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots, \quad \alpha_k \in \mathfrak{F}. \tag{6}$$

The elements (6) for which the sequence $\{\alpha_0, \alpha_1, \alpha_2, \dots\}$ converges to zero form the ring of stable realizable rational functions, denoted by $\mathfrak{F}^+\{z^{-1}\}$.

Now denote

$$\begin{aligned}
&\mathfrak{F}_{l,m} &&= \text{set of} \quad l \times m \quad \text{matrices over} \quad \mathfrak{F}\\
&\mathfrak{F}_{l,m}[z^{-1}] &&= \text{set of} \quad l \times m \quad \text{matrices over} \quad \mathfrak{F}[z^{-1}]\\
&\mathfrak{F}_{l,m}(z^{-1}) &&= \text{set of} \quad l \times m \quad \text{matrices over} \quad \mathfrak{F}(z^{-1})\\
&\mathfrak{F}_{l,m}\{z^{-1}\} &&= \text{set of} \quad l \times m \quad \text{matrices over} \quad \mathfrak{F}\{z^{-1}\}\\
&\mathfrak{F}^{+}_{l,m}\{z^{-1}\} &&= \text{set of} \quad l \times m \quad \text{matrices over} \quad \mathfrak{F}^{+}\{z^{-1}\}\,.
\end{aligned}$$

These sets are noncommutative rings when $l = m > 1$. The $\mathfrak{F}_{1,1}$ is viewed as isomorphic with $\mathfrak{F}$. We shall write I_l for the $l \times l$ identity matrix over $\mathfrak{F}$.

By the classical invariant-factor theorem [4; 9; 18] any polynomial matrix $A \in \mathfrak{F}_{l,m}[z^{-1}]$ can be written in the form

$$A = E_1 \operatorname{diag}\{a_1, a_2, \ldots, a_r, 0, \ldots, 0\}\, E_2\,, \tag{7}$$

where $E_1 \in \mathfrak{F}_{l,l}[z^{-1}]$ and $E_2 \in \mathfrak{F}_{m,m}[z^{-1}]$ are matrices such that $\det E_1$ and $\det E_2$ are units of $\mathfrak{F}[z^{-1}]$, and where $\operatorname{diag}\{\cdot\}$ is a matrix in $\mathfrak{F}_{l,m}[z^{-1}]$ all of whose elements are zero except those on the main diagonal, which are $a_1, a_2, \ldots, a_r$, possibly followed by zeros. The polynomials a_k are the invariant polynomials of A; they are uniquely determined by A up to units of $\mathfrak{F}[z^{-1}]$ and satisfy $a_k \mid a_{k+1}$, $k = 1, 2, \ldots, r-1$. The integer r is the rank of A. We shall call (7) the canonical decomposition of A.

The polynomial matrices of $\mathfrak{F}_{l,m}[z^{-1}]$ can also be written [18] as matrix polynomials over $\mathfrak{F}_{l,m}$,

$$A = A_0 + A_1 z^{-1} + \ldots + A_n z^{-n}\,, \quad A_k \in \mathfrak{F}_{l,m}\,.$$

If $A_n \neq 0$ then n is the degree of A, denoted by ∂A. We define $\partial 0 = -\infty$.

Let $a \in \mathfrak{F}[z^{-1}]$ and $B \in \mathfrak{F}_{l,m}[z^{-1}]$ with elements b_{ij}. Then we write (a, B) to denote $(a, (b_{11}, b_{12}, \ldots, b_{lm}))$.

A polynomial $p \in \mathfrak{F}[z^{-1}]$ is said to be stable if $1/p \in \mathfrak{F}^{+}\{z^{-1}\}$. Then any nonzero polynomial $a \in \mathfrak{F}[z^{-1}]$ can be factorized as

$$a = a^{+} a^{-}\,,$$

where a^{+} is the stable factor of a having highest degree and belonging to $\mathfrak{F}[z^{-1}]$. Given a nonzero polynomial matrix $A \in \mathfrak{F}_{l,m}[z^{-1}]$ and its canonical decomposition (7), we define the factorizations

$$A = A_1^{+} A_2^{-} = A_1^{-} A_2^{+}\,, \tag{8}$$

where

$$\begin{aligned}
A_1^{+} &= E_1 \operatorname{diag}\{a_1^{+}, a_2^{+}, \ldots, a_r^{+}, 1, \ldots, 1\} &&\in \mathfrak{F}_{l,l}[z^{-1}]\,,\\
A_2^{-} &= \operatorname{diag}\{a_1^{-}, a_2^{-}, \ldots, a_r^{-}, 0, \ldots, 0\}\, E_2 &&\in \mathfrak{F}_{l,m}[z^{-1}]\,,\\
A_1^{-} &= E_1 \operatorname{diag}\{a_1^{-}, a_2^{-}, \ldots, a_r^{-}, 0, \ldots, 0\} &&\in \mathfrak{F}_{l,m}[z^{-1}]\,,\\
A_2^{+} &= \operatorname{diag}\{a_1^{+}, a_2^{+}, \ldots, a_r^{+}, 1, \ldots, 1\}\, E_2 &&\in \mathfrak{F}_{m,m}[z^{-1}]\,.
\end{aligned}$$

 Observe that A_1^+ and A_2^+ are nonsingular matrices [18].

If

$$A = A_n z^{-n} + A_{n+1} z^{-(n+1)} + \ldots \in \mathfrak{F}_{l,m}(z^{-1}),$$

we can denote

$$A'_k = \text{transpose of } A_k,$$
$$\operatorname{tr} A = \text{trace of } A,$$
$$\langle A \rangle = A_0, \quad \text{the term of } A \text{ at } z^0,$$
$$A^{=} = A_n z^n + A_{n+1} z^{n+1} + \ldots.$$

Then the set $\mathfrak{F}^+_{l,m}\{z^{-1}\}$, viewed as a vector space over $\mathfrak{F}$, can be normed by introducing the quadratic norm $\|\cdot\|$ as follows

$$\|A\|^2 = \operatorname{tr} \langle A^{='} A \rangle. \tag{9}$$

In particular, consider

$$A = A_0 + A_1 z^{-1} + \ldots + A_n z^{-n} \in \mathfrak{F}_{l,m}[z^{-1}]$$

with $\partial A = n \geqq 0$. Then we define

$$A^{\sim} = z^{-n} A^{=} = A_0 z^{-n} + A_1 z^{-(n-1)} + \ldots + A_n \in \mathfrak{F}_{l,m}[z^{-1}]. \tag{10}$$

For any nonzero polynomial $a \in \mathfrak{F}[z^{-1}]$ we define the polynomial

$$a^* = a^+ a^{-\sim},$$

belonging again to $\mathfrak{F}[z^{-1}]$ and satisfying [18]

$$a^{=} a = a^{*=} a^*. \tag{11}$$

Given a nonzero polynomial matrix $A \in \mathfrak{F}_{l,m}[z^{-1}]$ and let

$$A^{='} A = E_1^{='} \operatorname{diag}\{p_1^{=} p_1, \ldots, p_s^{=} p_s, 0, \ldots, 0\} E_1,$$
$$A A^{='} = E_2 \operatorname{diag}\{q_1 q_1^{=}, \ldots, q_s q_s^{=}, 0, \ldots, 0\} E_2^{='}$$

be the canonical decompositions of $A^{='}A$ and $AA^{='}$. Then we define the matrix $A_1^* \in \mathfrak{F}_{s,m}[z^{-1}]$ by

$$\operatorname{diag}\{p_1^*, \ldots, p_s^*, 0, \ldots, 0\} E_1 = \begin{bmatrix} A_1^* \\ 0 \end{bmatrix} \tag{12}$$

and the matrix $A_2^* \in \mathfrak{F}_{l,s}[z^{-1}]$ by

$$E_2 \operatorname{diag}\{q_1^*, \ldots, q_s^*, 0, \ldots, 0\} = [A_2^* \; 0]. \tag{13}$$

It is clear [18; 32] that the A_1 and A_2 satisfy the relations

$$A^{=\prime}A = A_1^{*=\prime}A_1^*, \quad AA^{=\prime} = A_2^*A_2^{*=\prime},$$
$$\operatorname{rank} A_1^* = s, \qquad \operatorname{rank} A_2^* = s.$$

MATRIX DIOPHANTINE EQUATIONS

When dealing with single-input single-output systems we have to solve linear Diophantine equations of the form

$$ax + by = c, \tag{14}$$

where a, b, c are given polynomials of $\mathfrak{F}[z^{-1}]$ and x, y are unknown polynomials. It is shown in [12; 23] that equation (16) has a solution if and only if $(a, b) \mid c$. When x_0, y_0 is a particular solution of (14) then all solutions can be written as

$$x = x_0 + \frac{b}{(a, b)} t, \tag{15}$$

$$y = y_0 - \frac{a}{(a, b)} t,$$

where t is an arbitrary polynomial of $\mathfrak{F}[z^{-1}]$. If equation (14) is viewed over $\mathfrak{F}^+\{z^{-1}\}$, then t in (15) is an arbitrary element of $\mathfrak{F}^+\{z^{-1}\}$. An effective algorithm to find x_0, y_0 is presented in [12].

In applications, we often seek for a particular solution x^0, y^0 such that $\partial y^0 < \partial a$. To find the solution we apply the division algorithm

$$y_0 = \frac{a}{(a, b)} q + r, \quad \partial r < \partial \frac{a}{(a, b)}$$

and, in view of (15),

$$x^0 = x_0 + \frac{b}{(a, b)} (t_0 + q),$$

$$y^0 = r - \frac{a}{(a, b)} t_0,$$

where t_0 is an arbitrary polynomial of $\mathfrak{F}[z^{-1}]$ with

$$\partial t_0 < \partial(a, b).$$

In case $(a, b) = 1$ the solution x^0, y^0 is uniquely determined by setting $t_0 = 0$ and has the property that $\partial y^0 < \partial y$ for any y satisfying (15).

In multivariable control problems we encounter linear Diophantine equations of the form

$$AX + YB = C\,, \tag{16}$$

where $A \in \mathfrak{F}_{l,p}[z^{-1}]$, $B \in \mathfrak{F}_{q,m}[z^{-1}]$, $C \in \mathfrak{F}_{l,m}[z^{-1}]$ are given polynomial matrices and $X \in \mathfrak{F}_{p,m}[z^{-1}]$, $Y \in \mathfrak{F}_{l,q}[z^{-1}]$ are unknown matrices. It is shown in [18] that equation (16) has a solution if and only if the matrices

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

have the same invariant polynomials.

Let

$$A = E_{1A} \operatorname{diag}\{a_1, a_2, \ldots, a_r, 0, \ldots, 0\}\, E_{2A}\,,$$

$$B = E_{1B} \operatorname{diag}\{b_1, b_2, \ldots, b_s, 0, \ldots, 0\}\, E_{2B}$$

be the canonical decompositions of A and B and write $\bar{x}_{ij}$ for the elements of $\bar{X} = E_{2A} X E_{2B}^{-1}$, $\bar{y}_{ij}$ for the elements of $\bar{Y} = E_{1A}^{-1} Y E_{1B}$, and $\bar{c}_{ij}$ for the elements of $\bar{C} = E_{1A}^{-1} C E_{2B}^{-1}$. Then any solvable equation (16) is equivalent to the following sets of polynomial equations

$$a_i \bar{x}_{ij} + \bar{y}_{ij} b_j = \bar{c}_{ij}\,, \quad i = 1, 2, \ldots, r \quad \text{and} \quad j = 1, 2, \ldots, s\,, \tag{17}$$

$$a_i \bar{x}_{ij} = \bar{c}_{ij}\,, \quad i = 1, 2, \ldots, r \quad \text{and} \quad j = s + 1, \ldots, m\,, \tag{18}$$

$$\bar{y}_{ij} b_j = \bar{c}_{ij}\,, \quad i = r + 1, \ldots, l \quad \text{and} \quad j = 1, 2, \ldots, s\,, \tag{19}$$

$$0 = \bar{c}_{ij}\,, \quad i = r + 1, \ldots, l \quad \text{and} \quad j = s + 1, \ldots, m\,. \tag{20}$$

The remaining elements $\bar{x}_{ij}$ and $\bar{y}_{ij}$ can be chosen arbitrarily within $\mathfrak{F}[z^{-1}]$.

As a consequence [18], a particular solution of equation (16) can be written as

$$X_0 = E_{2A}^{-1} \begin{bmatrix} \bar{X}_{0,11} & \bar{X}_{0,12} \\ 0 & 0 \end{bmatrix} E_{2B} \in \mathfrak{F}_{p,m}[z^{-1}]$$

$$Y_0 = E_{1A} \begin{bmatrix} \bar{Y}_{0,11} & 0 \\ \bar{Y}_{0,21} & 0 \end{bmatrix} E_{1B}^{-1} \in \mathfrak{F}_{l,q}[z^{-1}]\,,$$

where the elements $\bar{x}_{0,ij}$ of $\bar{X}_{0,11} \in \mathfrak{F}_{r,s}[z^{-1}]$ and the elements $\bar{y}_{0,ij}$ of $\bar{Y}_{0,11} \in \mathfrak{F}_{r,s}[z^{-1}]$ are particular solutions of (17), the elements $\bar{x}_{0,ij}$ of $\bar{X}_{0,12} \in \mathfrak{F}_{r,m-s}[z^{-1}]$ are particular solutions of (18), and the elements $\bar{y}_{0,ij}$ of $\bar{Y}_{0,21} \in \mathfrak{F}_{l-r,s}[z^{-1}]$ are particular solutions of (19).

Then it is proved in [18] that the general solution of equation (16) becomes

$$X = X_0 + E_{2A}^{-1} T E_{2B}, \tag{21}$$

$$Y = Y_0 - E_{1A} S E_{1B}^{-1},$$

$$T = \begin{bmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{bmatrix}, \quad S = \begin{bmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{bmatrix}.$$

The elements of $T_{11} \in \mathfrak{F}_{r,s}[z^{-1}]$ are $t_{ij} b_j / (a_i, b_j)$ and the elements of $S_{11} \in \mathfrak{F}_{r,s}[z^{-1}]$ are $a_i t_{ij} / (a_i, b_j)$, where t_{ij} are arbitrary polynomials of $\mathfrak{F}[z^{-1}]$. The matrices $T_{21} \in \mathfrak{F}_{p-r,s}[z^{-1}]$, $T_{22} \in \mathfrak{F}_{p-r,m-s}[z^{-1}]$ and $S_{12} \in \mathfrak{F}_{r,q-s}[z^{-1}]$, $S_{22} \in \mathfrak{F}_{l-r,q-s}[z^{-1}]$ are arbitrary polynomial matrices.

It is to be noted that a particular solution X^0, Y^0 such that $\partial Y^0 < \partial A$ cannot be, in general, found by application of the division algorithm but, instead, by analysis of the general solution Y, see [18].

When the X and Y are allowed to be matrices over $\mathfrak{F}^+\{z^{-1}\}$ then the t_{ij} in (21) are arbitrary elements of $\mathfrak{F}^+\{z^{-1}\}$ and so are the elements of T_{21}, T_{22} and S_{12}, S_{22}.

SYSTEM DESCRIPTION

Throughout the paper we shall consider finite-dimensional discrete linear constants m-input l-output systems defined over the field $\mathfrak{F}$. They are described by the equations

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \tag{22}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k,$$

where k ranges over integers, $\mathbf{u} \in \mathfrak{F}^m$ is the m-vector input, $\mathbf{y} \in \mathfrak{F}^l$ is the l-vector output, $\mathbf{x} \in \mathfrak{F}^n$ is the n-dimensional state vector, and $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are matrices over $\mathfrak{F}$ of appropriate dimensions [9].

The matrix sequence

$$\boldsymbol{S} = \mathbf{C} z^{-1} (\mathbf{I}_n - z^{-1}\mathbf{A})^{-1} \mathbf{B} + \mathbf{D} \in \mathfrak{F}_{l,m}\{z^{-1}\} \tag{23}$$

is called the impulse response matrix of the system. Conversely, any quadruple $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ satisfying (23) is a realization of $\boldsymbol{S}$; if $\mathbf{A}$ is of least possible size the realization is minimal [9; 15; 18].

The $\boldsymbol{S}$ can be written as the ratio of a polynomial matrix and a polynomial, viz.

$$\boldsymbol{S} = \frac{B}{a}, \tag{24}$$

where $a \in \mathfrak{F}[z^{-1}]$, $B \in \mathfrak{F}_{l,m}[z^{-1}]$ and

$$(a, B) = 1, \quad (a, z^{-1}) = 1.$$

If $l = m = 1$ (single-variable system) we obtain

$$S = \frac{b}{a} \in \mathfrak{F}\{z^{-1}\} \tag{25}$$

where both a and b belong to $\mathfrak{F}[z^{-1}]$.

While (25) completely describes a single-variable system, the ratio (24) tells very little about a multivariable system. We have to refine it as follows. Let

$$B = E_1 \operatorname{diag}\{g_1, g_2, \ldots, g_r, 0, \ldots, 0\}\, E_2$$

be a canonical decomposition of B and let

$$\frac{g_i}{a} = \frac{b_i}{a_i}, \quad i = 1, 2, \ldots, r,$$

after cancelling common factors. Then

$$\frac{B}{a} = E_1 \operatorname{diag}\left\{\frac{b_1}{a_1}, \frac{b_2}{a_2}, \ldots, \frac{b_r}{a_r}, 0, \ldots, 0\right\} E_2$$

and, defining the matrices

$$\begin{aligned}
B_1 &= E_1 \operatorname{diag}\{b_1, b_2, \ldots, b_r, 0, \ldots, 0\} &&\in \mathfrak{F}_{l,m}[z^{-1}],\\
A_2 &= E_2^{-1} \operatorname{diag}\{a_1, a_2, \ldots, a_r, 1, \ldots, 1\} &&\in \mathfrak{F}_{m,m}[z^{-1}],\\
A_1 &= \operatorname{diag}\{a_1, a_2, \ldots, a_r, 1, \ldots, 1\}\, E_1^{-1} &&\in \mathfrak{F}_{l,l}[z^{-1}],\\
B_2 &= \operatorname{diag}\{b_1, b_2, \ldots, b_r, 0, \ldots, 0\}\, E_2 &&\in \mathfrak{F}_{l,m}[z^{-1}],
\end{aligned} \tag{26}$$

we can write

$$S = B_1 A_2^{-1} = A_1^{-1} B_2\,. \tag{27}$$

The above decomposition of S into the product of a polynomial matrix and the inverse of another polynomial matrix is fundamental and plays the role similar to (25).

RANDOM SEQUENCES

For convenience, we shall review some elementary facts about random sequences. For details consult [3; 5; 8; 24; 25; 28; 30].

An l-vector random variable over $\mathfrak{F}$ is a vector function whose values belong to $\mathfrak{F}_{l,1}$ and depend on the outcome of a chance event. The (ensemble) expectation of a random variable Δ will be denoted by $\mathsf{E}\Delta$.

A sequence of l-vector random variables

$$A = \{\ldots, A_{-1}, A_0, A_1, \ldots\}$$

is called an l-vector random sequence over $\mathfrak{F}$. The k function with values $\mathsf{E}A_k$ is called the mean-value vector of $\boldsymbol{A}$. The s, t function whose values are $\mathsf{E}A_sA'_t$ is the correlation matrix of $\boldsymbol{A}$. If

$$\boldsymbol{B} = \{\ldots, B_{-1}, B_0, B_1, \ldots\}$$

is another vector random sequence, the s, t function with values $\mathsf{E}A_sB'_t$ is the cross-correlation matrix of $\boldsymbol{A}$ and $\boldsymbol{B}$ (in this-order).

A vector random sequence is said to be (weakly) stationary if its mean-value vector is independent of k and its correlation matrix depends only on $s - t$ and is bounded.

A stationary l-vector random sequence is called white if

$$\begin{aligned} \mathsf{E}(A_s - \mathsf{E}A_s)(A_t - \mathsf{E}A_t)' &= \Omega\,, \quad s = t\,, \\ &= 0\,, \quad s \neq t\,, \end{aligned}$$

where $\Omega \in \mathfrak{F}_{l,l}$ is a symmetric nonnegative definite matrix.

An l-vector random sequence $\boldsymbol{A}$ over $\mathfrak{F}$ can be thought of as the output of an q-input l-output system $\mathscr{F}_A$ over $\mathfrak{F}$ excited by a white q-vector random sequence $\boldsymbol{D}$, see Fig. 5. The $\mathscr{F}_A$ is usually called the shaping filter of $\boldsymbol{A}$. This representation of $\boldsymbol{A}$ is essential for obtaining the main results of the paper.

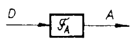

Fig. 5. Random sequence model.

In all that follows we shall confine ourselves to vector random sequences whose shaping filters are systems governed by equations (22). Such a sequence $\boldsymbol{A}$ is stationary if and only if the impulse response matrix $\boldsymbol{F}_A$ of $\mathscr{F}_A$ belongs to $\mathfrak{F}^+_{l,q}\{z^{-1}\}$. Then the sequence

$$\ldots + \Phi_{-1}z + \Phi_0 + \Phi_1 z^{-1} + \ldots,$$

where

$$\Phi_k = \mathsf{E}(A_{k+s} - \mathsf{E}A_{k+s})(A_s - \mathsf{E}A_s)' = \Phi'_{-k}\,,$$

is the correlation matrix of $\boldsymbol{A}$. The Φ_0 is called the covariance matrix of $\boldsymbol{A}$. If $\boldsymbol{B}$ is another stationary vector random sequence, the cross-correlation matrix of $\boldsymbol{A}$ and $\boldsymbol{B}$ can be written as

$$\ldots + \Psi_{-1}z + \Psi_0 + \Psi_1 z^{-1} + \ldots,$$

where

$$\Psi_k = \mathsf{E}(A_{k+s} - \mathsf{E}A_{k+s})^{=}(B_s - \mathsf{E}B_s)'\,.$$

It can be shown that

$$\ldots + \Phi_{-1}z + \Phi_0 + \Phi_1 z^{-1} + \ldots = F_A^{=} F'_A$$

and hence

$$\operatorname{tr} \Phi_0 = \operatorname{tr} \langle F_A^{=} F'_A \rangle = \operatorname{tr} \langle F_A^{=\prime} F_A \rangle = \|F_A\|^2 \tag{28}$$

by (9). In words, the trace of the covariance matrix (= sum of the variances of individual components) of a stationary vector random sequence A can be interpreted as the squared quadratic norm of its shaping filter impulse response matrix F_A. In case of a scalar random sequence A the Φ_0 itself is the variance of A and hence the squared quadratic norm of F_A.

CLOSED-LOOP STABILITY

Consider the closed-loop system configuration shown in Fig. 6, where $\mathscr{S}$ is the system to be controlled and $\mathscr{R}$ is the controller. The most important condition imposed on closed-loop control systems is that of stability. An extensive discussion of the closed-loop stability problem is given in [15] and [18].

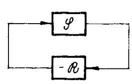

Fig. 6. Closed-loop system configuration.

We shall first summarize the fundamental results for single-variable systems [15] and then proceed to *multivariable systems* [18].

Let $\mathscr{S}$ be a minimal realization of

$$S = \frac{b}{a} \in \mathfrak{F}\{z^{-1}\},$$

$\mathscr{R}$ be a minimal realization of some $R \in \mathfrak{F}\{z^{-1}\}$ and denote

$$K = \frac{SR}{1 + SR}. \tag{29}$$

Then the closed-loop system is stable if and only if

$$K = bM, \quad 1 - K = aN, \tag{30}$$

where M and N are elements of $\mathfrak{F}^+\{z^{-1}\}$ such that

$$bM + aN = 1. \tag{31}$$

This is a linear Diophantine equation over $\mathfrak{F}^+\{z^{-1}\}$ which has infinitely many solutions M, N. The freedom in choosing M and N can be exploited for optimization.

Now let $\mathscr{S}$ be a minimal realization of

$$S = B_1 A_2^{-1} = A_1^{-1} B_2 \in \mathfrak{F}_{l,m}\{z^{-1}\},$$

$\mathscr{R}$ be a minimal realization of some $R \in \mathfrak{F}_{m,l}\{z^{-1}\}$ and denote

(32) $$K_1 = SR(I_l + SR)^{-1}, \quad K_2 = RS(I_m + RS)^{-1}.$$

Then the closed-loop system is stable if and only if

(33) $$K_1 = B_1 M_1, \quad I_l - K_1 = N_1 A_1,$$
$$K_2 = M_2 B_2, \quad I_m - K_2 = A_2 N_2$$

where $M_1 \in \mathfrak{F}^+_{m,l}\{z^{-1}\}$, $N_1 \in \mathfrak{F}^+_{l,l}\{z^{-1}\}$ and $M_2 \in \mathfrak{F}^+_{m,l}\{z^{-1}\}$, $N_2 \in \mathfrak{F}^+_{m,m}\{z^{-1}\}$ obey the linear Diophantine equations

(34) $$B_1 M_1 + N_1 A_1 = I_l,$$

(35) $$A_2 N_2 + M_2 B_2 = I_m.$$

It is shown in [18] that the M_1, N_1 and M_2, N_2 satisfy the mutual relations

(36) $$A_2 M_1 = M_2 A_1,$$
$$N_1 B_2 = B_1 N_2$$

by virtue of (32) and (33).

If $l = m = 1$ we have

$$A_1 = A_2 = a, \quad B_1 = B_2 = b, \quad K_1 = K_2 = K$$

and equations (34) and (35) reduce to equation (31).

As shown in [15, 18] the closed-loop system need not be a minimal realization of K_1 and K_2 even if the $\mathscr{S}$ and $\mathscr{R}$ are minimal realizations of S and R. Then the above result demands that, in addition to stability of the minimal realization of K_1 and K_2, the remaining part of the closed-loop system, which has no relation to K_1 and K_2, should also be stable. This part appears due to the mode cancellations in the cascades $\mathscr{S}\mathscr{R}$ and $\mathscr{R}\mathscr{S}$.

STOCHASTIC CONTROL

In this section we shall transform problems (1) through (4) into a common framework and give the formal definition and complete solution of the general stochastic control problem.

Consider the closed-loop configuration shown in Fig. 7, where

$\mathscr{S}$ = system to be controlled ,
$\mathscr{R}$ = controller ,
$\mathscr{F}_W$ = shaping filter of $\boldsymbol{W}$,
$\boldsymbol{W}$ = random input sequence ,
$\boldsymbol{D}$ = white random sequence ,
$\boldsymbol{U}$ = system input sequence ,
$\boldsymbol{Y}$ = system output sequence ,
$\boldsymbol{E}$ = error sequence .

Further denote respectively $\boldsymbol{F}_U$ and $\boldsymbol{F}_E$ the impulse response matrices of the shaping filters $\mathscr{F}_U$ and $\mathscr{F}_E$ that generate the random sequences $\boldsymbol{U}$ and $\boldsymbol{E}$, i.e.

$$\boldsymbol{U} = \boldsymbol{F}_U \boldsymbol{D} , \quad \boldsymbol{E} = \boldsymbol{F}_E \boldsymbol{D} .$$

The optimality criterion to be minimized will be chosen as $\|\boldsymbol{F}_E\|^2$, which can be interpreted as the sum of steady-state variances of the error sequence components

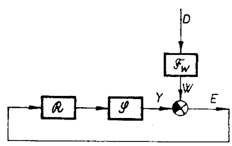

Fig. 7. The general stochastic control problem.

The mean value of the error sequence is immaterial since it has no effect upon the steady-state variances. The optimality criterion simply disregards the mean values. If the error is a zero-mean random sequence then this criterion coincides with the root-mean-square error criterion [2; 30; 31].

For the moment, we introduce the following notation:

$$\boldsymbol{C} = \boldsymbol{F}_C \boldsymbol{D} , \quad \boldsymbol{V} = \boldsymbol{F}_V \boldsymbol{D} ,$$

$\boldsymbol{S}_1$ = impulse response matrix of $\mathscr{S}_1$,
Φ_{CC} = correlation matrix of $\boldsymbol{C}$,
Φ_{VV} = correlation matrix of $\boldsymbol{V}$,
Φ_{CV} = cross-correlation matrix of $\boldsymbol{C}$ and $\boldsymbol{V}$.

Then it is clear that problems (1) through (4) can be expressed in terms of Fig. 7 when identifying

sub 1 $$\boldsymbol{F}_W = \boldsymbol{F}_C ,$$

sub 2 $$F_W = S_1 F_V,$$

sub 3 $$F_W^{=} F_W' = \Phi_{CC} + \Phi_{CV} + \Phi_{CV}^{='} + \Phi_{VV},$$

sub 4 $$F_W^{=} F_W' = \Phi_{CC} - \Phi_{CV} S_1' - S_1^{=} \Phi_{CV}^{='} + \Phi_{VV}.$$

Now we can give an exact formulation of the general stochastic control problem. It is instructive and certainly worthwhile to begin with the special case of single-variable systems and then generalize.

(37) Given a system $\mathscr{S}$ which is a minimal realization of

$$S = \frac{b}{a} \in \mathfrak{F}\{z^{-1}\}, \quad b \neq 0$$

and a random sequence W by its shaping filter F_W which is a (not necessarily minimal) realization of

$$F_W = \frac{q}{p} \in \mathfrak{F}\{z^{-1}\}, \quad q \neq 0.$$

Find a controller $\mathscr{R}$ which is a minimal realization of some $R \in \mathfrak{F}\{z^{-1}\}$ such that the closed-loop system is stable, the F_U is stable, and the $\|F_E\|^2$ is minimized.

For further reference define

(38) $$\frac{a}{p} = \frac{a_0}{p_0}$$

after cancelling common factors. Then we have the following result, which is a generalization of a similar result in [27].

Theorem 1. *Problem (37) has a solution if and only if the linear Diophantine equation*

(39) $$b^- x + p a_0^- y = b^{-\sim} q^* a_0^{-\sim}$$

has a solution x^0, y^0 such that $\partial y^0 < \partial b^-$ and

$$M = \frac{x^0}{b^* q^* a_0^{-\sim}}, \quad N = \frac{p_0 y^0}{b^{-\sim} q^* a_0^*},$$

$$F_U = aMF_W, \quad F_E = aNF_W$$

belong to $\mathfrak{F}^+\{z^{-1}\}$.

The optimal controller is unique and it is given as a minimal realization of

(40) $$R = \frac{M}{N}.$$

 Moreover,

$$\|\boldsymbol{F}_E\|^2_{\min} = \left\langle \left(\frac{y^0}{b^-}\right)^{=} \left(\frac{y^0}{b^-}\right) \right\rangle. \tag{41}$$

Proof. In order to minimize the variance $\|\boldsymbol{F}_E\|^2$ of $\boldsymbol{E}$ we shall assume that $\boldsymbol{F}_E$ is stable, whereby the $\boldsymbol{E}$ is a stationary random sequence and

$$\|\boldsymbol{F}_E\|^2 = \langle \boldsymbol{F}_E^{=} \boldsymbol{F}_E \rangle \tag{42}$$

in view of (28). Then we will manipulate the expression $\langle \boldsymbol{F}_E^{=} \boldsymbol{F}_E \rangle$ so as to make the minimizing choice of $\boldsymbol{R}$ obvious.

Write

$$\boldsymbol{F}_E = (1 - \boldsymbol{K}) \boldsymbol{F}_W .$$

Denoting

$$\boldsymbol{F}_E^* = (1 - \boldsymbol{K}) \frac{q^*}{p}, \tag{43}$$

it is clear that

$$\boldsymbol{F}_E^{=} \boldsymbol{F}_E = \boldsymbol{F}_E^{=*} \boldsymbol{F}_E^* \tag{44}$$

by virtue of (11) and

$$\boldsymbol{F}_E = \boldsymbol{F}_E^* \frac{q^-}{q^{-\sim}}. \tag{45}$$

To guarantee a stable closed-loop system we have to set $\boldsymbol{K} = b\boldsymbol{M}$ for some $\boldsymbol{M} \in \mathfrak{F}^+\{z^{-1}\}$, see (30). Then

$$\boldsymbol{F}_E^* = \frac{q^*}{p} - b\boldsymbol{M} \frac{q^*}{p} \tag{46}$$

and

$$\boldsymbol{F}_E^{*=} \boldsymbol{F}_E^* = \frac{q^{*=} q^*}{p^{=} p} - \frac{q^{*=}}{p} b\boldsymbol{M} \frac{q^*}{p} - \frac{q^{*=}}{p^{=}} \boldsymbol{M}^{=} b^{=} \frac{q^*}{p} +$$

$$+ \frac{q^{*=}}{p^{=}} \boldsymbol{M}^{=} b^{=} b\boldsymbol{M} \frac{q^*}{p} = \left(\frac{b^{=} q^*}{b^{*=} p} - \frac{b^* q^*}{p} \boldsymbol{M}\right)^{=} \left(\frac{b^{=} q^*}{b^{*=} p} - \frac{b^* q^*}{p} \boldsymbol{M}\right)$$

after rearranging. Since

$$\frac{b^{=}}{b^{*=}} = \frac{b^{-=}}{b^-} = \frac{b^{-\sim}}{b^-}$$

by (10) and the definition of b^*, and since

$$\left(\frac{a_0^{-\sim}}{a_0^-}\right)^{=} \left(\frac{a_0^{-\sim}}{a_0^-}\right) = 1,$$

we can write

(47)
$$F_E^{*=}F_E^* = F_{E0}^{=}F_{E0}\,,$$

where

(48)
$$F_{E0} = \frac{b^{-\sim}q^*a_0^{-\sim}}{b^-pa_0^-} - \frac{b^*q^*a_0^{-\sim}}{pa_0^-}M\,.$$

Now take the partial fraction expansion

(49)
$$\frac{b^{-\sim}q^*a_0^{-\sim}}{b^-pa_0^-} = \frac{y}{b^-} + \frac{x}{pa_0^-}\,.$$

It follows that the polynomials x and y are governed by equation (39).
In view of (48) and (49) we can write

(50)
$$F_{E0} = \frac{y}{b^-} + Z\,,$$

where

(51)
$$Z = \frac{x}{pa_0^-} - \frac{b^*q^*a_0^{-\sim}}{pa_0^-}M\,.$$

Then, by (44), (47), and (50),

(52)
$$\langle F_E^{=}F_E\rangle = \left\langle\left(\frac{y}{b^-}\right)^{=}\left(\frac{y}{b^-}\right)\right\rangle + \left\langle\left(\frac{y}{b^-}\right)^{=}Z\right\rangle + \left\langle Z^{=}\left(\frac{y}{b^-}\right)\right\rangle + \langle Z^{=}Z\rangle\,.$$

Any solution of equation (39) can be written as

(53)
$$x = x^0 + \frac{pa_0^-}{(b^-, pa_0^-)}t\,,$$

(54)
$$y = y^0 - \frac{b^-}{(b^-, pa_0^-)}t\,,$$

where $t \in \mathfrak{F}[z^{-1}]$ arbitrary and

(55)
$$\partial y^0 < \partial b^-\,.$$

The key observation is that

$$\left(\frac{y^0}{b^-}\right)^{=} = \frac{y^{0\sim}}{b^{-\sim}}z^{-(\partial b^- - \partial y^0)}$$

is divisible by z^{-1} due to (55). Therefore

$$\left\langle\left(\frac{y^0}{b^-}\right)^{=}Z\right\rangle = 0\,, \quad \left\langle\left(\frac{y^0}{b^-}\right)^{=}\frac{t}{(b^-, pa_0^-)}\right\rangle = 0$$

and after substituting (54) into (52) we have

$$(56) \quad \langle \boldsymbol{F}_E^{=} \boldsymbol{F}_E \rangle = \left\langle \left(\frac{y^0}{b^-}\right)^{=} \left(\frac{y^0}{b^-}\right) \right\rangle + \left\langle \left(\boldsymbol{Z} - \frac{t}{(b^-, pa_0^-)}\right)^{=} \left(\boldsymbol{Z} - \frac{t}{(b^-, pa_0^-)}\right) \right\rangle .$$

The first term on the right-hand side of (56) cannot be affected by any choice of $\boldsymbol{M}$ (and hence $\boldsymbol{R}$). The best we can do to minimize (56) is to set

$$(57) \qquad \boldsymbol{Z} - \frac{t}{(b^-, pa_0^-)} = 0 ,$$

i.e.

$$\frac{x}{pa_0^-} - \frac{b^* q^* a_0^{-\sim}}{pa_0^-} \boldsymbol{M} - \frac{t}{(b^-, pa_0^-)} = 0$$

by (51). But

$$\frac{x}{pa_0^-} - \frac{t}{(b^-, pa_0^-)} = \frac{x^0}{pa_0^-} ,$$

see (53). Hence (56) is minimized by setting

$$(58) \qquad \boldsymbol{M} = \frac{x^0}{b^* q^* a_0^{-\sim}} .$$

Substituting (58) into (46) we obtain

$$(59) \qquad \boldsymbol{F}_E^* = \frac{q^*}{p} - \frac{b^- x^0}{b^{-\sim} pa_0^{-\sim}} = \frac{b^{-\sim} q^* a_0^{-\sim} - b^- x^0}{b^{-\sim} pa_0^{-\sim}} =$$

$$= \frac{pa_0^- y^0}{b^{-\sim} pa_0^{-\sim}} = \frac{a_0^- y^0}{a_0^{-\sim} b^{-\sim}}$$

on using equation (39). Now consider (43) and set $1 - \boldsymbol{K} = a\boldsymbol{N}$ for some $\boldsymbol{N} \in \mathfrak{F}^+\{z^{-1}\}$ to guarantee a stable closed-loop system, see (30). Then

$$(60) \qquad \boldsymbol{F}_E^* = a\boldsymbol{N} \frac{q^*}{p} = a_0 \boldsymbol{N} \frac{q^*}{p_0}$$

and the comparison of (59) and (60) yields

$$(61) \qquad \boldsymbol{N} = \frac{p_0 y^0}{b^{-\sim} q^* a_0^*} .$$

Observe that the M and N satisfy the Diophantine equation (31). Therefore, the closed-loop system will be stable if and only if both M and N are stable. Moreover, we have to require that

$$F_U = aMF_W = \frac{a_0^- q^-}{a_0^{-\sim} q^{-\sim}} \frac{a_0^+ x^0}{p_0 b^*}$$

be stable according to the problem statement, and that

$$F_E = aNF_W = \frac{a_0^- q^-}{a_0^{-\sim} q^{-\sim}} \frac{y^0}{b^{-\sim}}$$

be stable to satisfy hypothesis (42). It can be seen that the F_E is always stable, see equation (39), and hence the minimized steady-state variance of E is given by (41) on applying (42), (56) and (57).

The optimal controller is given as a minimal realization of

$$R = \frac{1}{S} \frac{K}{1 - K} = \frac{M}{N} = \frac{a_0^+ x^0}{p_0 b^* y^0}$$

by virtue of (29), (30) and (58), (61).

In order that the F_U may be stable it is necessary that p_0 be stable, i.e. $p^- \mid a$ by (42). Then $(b^-, pa_0^-) = 1$ and the solution x^0, y^0 satisfying $\partial y^0 < \partial b^-$ exists and has also the property that y^0 is of least possible degree among all solutions of (39). As such a solution is unique, the optimal controller $\mathscr{R}$ is also unique. □

It is to be noted that the input random sequence W was not assumed stationary in the sense that $\mathscr{F}_W$ does not have to be stable. For unstable $\mathscr{F}_W$, however, the problem (41) can have a solution only if $p^- \mid a$. This can hardly be exactly satisfied in practice due to the parameter fluctuations unless the unstable part of $\mathscr{F}_W$ is actually a part of the system $\mathscr{S}$ through which the random sequence passes on its way to system output. As typical examples in this line serve problems (2) and (4) when a stationary disturbance V passes through an unstable part $\mathscr{S}_1$ of $\mathscr{S}$, see Example 1.

Example 1. Consider problem (2) for the system $\mathscr{S} = \mathscr{S}_1 \mathscr{S}_2$, where

$$S_1 = \frac{z^{-1}}{(1 - z^{-1})(1 - 1{\cdot}5z^{-1})}, \quad S_2 = 1,$$

and the disturbance V which is a zero-mean white random sequence over $\mathfrak{F}$ with shaping filter

$$F_V = 1 .$$

Note that the $\mathscr{S}$ is not stable. It can be thought of as a two-phase servomotor with small rotor resistance, operating at low speeds. At this mode of operation the shaft torque increases with increasing speed.

Transforming the problem into the general configuration shown in Fig. 7 we obtain

$$S = \frac{z^{-1}}{(1 - z^{-1})(1 - 1{\cdot}5z^{-1})},$$

$$F_W = \frac{z^{-1}}{(1 - z^{-1})(1 - 1{\cdot}5z^{-1})}.$$

Applying Theorem 1 we compute

$$b^+ = 1, \qquad b^- = z^{-1},$$
$$a_0^+ = 1, \qquad a_0^- = 1,$$
$$q^* = 1, \qquad p_0 = 1,$$

and solve the equation

$$z^{-1}x + (1 - z^{-1})(1 - 1{\cdot}5z^{-1})\, y = 1. \tag{62}$$

Using the algorithm described in [12] we find the solution x^0, y^0 satisfying $\partial y^0 < 1$ to be

$$x^0 = 2{\cdot}5 - 1{\cdot}5z^{-1},$$
$$y^0 = 1.$$

Then

$$M = 2{\cdot}5 - 1{\cdot}5z^{-1}, \qquad N = 1,$$
$$F_U = 2{\cdot}5z^{-1} - 1{\cdot}5z^{-2}, \quad F_E = z^{-1}$$

all belong to $\mathfrak{F}^+\{z^{-1}\}$ and hence the optimal controller exists and is given as a minimal realization of

$$R = 2{\cdot}5 - 1{\cdot}5z^{-1}$$

by (40).

Since $F_E^= F_E = 1$, the random sequence E (the output of $\mathscr{S}$ in Fig. 2) approaches a white random sequence with variance $\|F_E\|^2_{\min} = 1$, even though the random sequence W is not stationary.

Example 2. Consider problem (1) for the system $\mathscr{S}$ over $\mathfrak{F}$ which is a minimal realization of

$$S = \frac{z^{-1}}{(1 - z^{-1})(1 - 0{\cdot}4z^{-1})}$$

and the zero-mean random sequence C by its shaping filter

$$F_C = \frac{1}{1 - 0{\cdot}5z^{-1}}.$$

This is a typical positioning system. Recasting the problem in terms of Fig. 7 we have

$$F_W = \frac{1}{1 - 0{\cdot}5z^{-1}}.$$

Hence

$$b^+ = 1, \qquad b^- = z^{-1},$$
$$a_0^+ = 1 - 0{\cdot}4z^{-1}, \quad a_0^- = 1 - z^{-1},$$
$$q^* = 1, \qquad p_0 = 1 - 0{\cdot}5z^{-1},$$

and equation (39) reads

$$z^{-1}x + (1 - 0{\cdot}5z^{-1})(1 - z^{-1})\, y = z^{-1} - 1.$$

The solution x^0, y^0 with $\partial y^0 < 1$ is

$$x^0 = -0{\cdot}5(1 - z^{-1}),$$
$$y^0 = -1$$

and it yields

(63) $$M = 0{\cdot}5, \quad N = \frac{1 - 0{\cdot}5z^{-1}}{(1 - z^{-1})(1 - 0{\cdot}4z^{-1})},$$
$$F_U = 0{\cdot}5\,\frac{(1 - z^{-1})(1 - 0{\cdot}4z^{-1})}{1 - 0{\cdot}5z^{-1}}, \quad F_E = 1.$$

Since the N is not stable the closed-loop system would not be stable, either, and hence the error variance never reaches its steady state. We conclude that the problem has no solution. As a rule, it is impossible to design a stable closed-loop positioning system that would follow a zero-mean stationary random signal in the minimum variance sense. To avoid this impass we usually introduce a nonzero mean value and take it into computations in several ways, see [2; 30].

Example 3. Consider problem (2) for the system $\mathcal{S} = \mathcal{S}_1\mathcal{S}_2$ over $\mathfrak{F}$ which is a minimal realization of

$$S_1 = 1, \quad S_2 = \frac{z^{-1}(1 - z^{-1})}{z^{-1} - 2}$$

and for the random disturbance V given by

$$F_V = \frac{4 - 3z^{-1} + z^{-2}}{z^{-1} - 2}.$$

Clearly, we have

$$S = \frac{z^{-1}(1 - z^{-1})}{z^{-1} - 2}, \quad F_W = \frac{4 - 3z^{-1} + z^{-2}}{z^{-1} - 2}$$

in Fig. 7 and

$$b^+ = 1, \qquad b^- = z^{-1}(1 - z^{-1}),$$

$$a_0^+ = 1\,, \qquad a_0^- = 1\,,$$
$$q^* = 4 - 3z^{-1} + z^{-2}\,, \quad p_0 = 1\,.$$

Thus equation (39) becomes

$$z^{-1}(1 - z^{-1})\,x + (z^{-1} - 2)\,y = (z^{-1} - 1)(4 - 3z^{-1} + z^{-2})$$

and the solution x^0, y^0 with $\partial y^0 < 2$ is

$$x^0 = 1 - z^{-1}\,,$$
$$y_0 = 2(1 - z^{-1})\,.$$

We obtain

$$\boldsymbol{M} = -\frac{1}{4 - 3z^{-1} + z^{-2}}\,, \quad \boldsymbol{N} = -\frac{2}{(4 - 3z^{-1} + z^{-2})(z^{-1} - 2)}\,,$$

$$\boldsymbol{F}_U = -1\,, \quad \boldsymbol{F}_E = \frac{2}{2 - z^{-1}}$$

which all belong to $\mathfrak{F}^+\{z^{-1}\}$, and hence the optimal controller is a minimal realization of

$$\boldsymbol{R} = \frac{z^{-1} - 2}{2}\,.$$

It should be noted that the optimal and stable solution exists (the steady-state system output is even a white random sequence) though the system $\mathscr{S}$ does not enjoy the minimum-phase property. This result serves as a counterexample to the common fallacy that a system with zeros at the stability boundary cannot be stably controlled in the minimum variance sense [27].

Now we proceed to the multivariable case.

(64) Given a system $\mathscr{S}$ which is a minimal realization of

$$\boldsymbol{S} = \frac{B}{a} \in \mathfrak{F}_{l,m}\{z^{-1}\}\,, \quad B \neq 0\,,$$

and a vector random sequence $\boldsymbol{W}$ by its shaping filter $\boldsymbol{F}_W$ which is a (not necessarily minimal) realization of

$$\boldsymbol{F}_W = \frac{Q}{p} \in \mathfrak{F}_{l,q}\{z^{-1}\}\,, \quad Q \neq 0\,.$$

Find a controller $\mathscr{R}$ which is a minimal realization of some $\boldsymbol{R} \in \mathfrak{F}_{m,l}\{z^{-1}\}$ such that the closed-loop system is stable, the $\boldsymbol{F}_U$ is stable, and the $\|\boldsymbol{F}_E\|^2$ is minimized.

For further reference denote rank $B = r$ and write

$$\boldsymbol{S} = B_1 A_2^{-1} = A_1^{-1} B_2\,.$$

By (26) and (8) the B_1 can be written as $B_1 = B_1^- B_2^+$ and

$$B_1^- = [B_{11}^- \; 0],$$

where $B_{11}^- \in \mathfrak{F}_{l,r}[z^{-1}]$, $0 \in \mathfrak{F}_{l,m-r}[z^{-1}]$ and rank $B_{11}^- = r$. Then, using (12),

$$B_{11}^{-=\prime} B_{11}^- = (B_{11}^-)^{*=\prime} (B_{11}^-)^{*=},$$

where $(B_{11}^-)^* \in \mathfrak{F}_{r,r}[z^{-1}]$ and rank $(B_{11}^-)^* = r$. For convenience, denote $H = (B_{11}^-)^*$ and

$$d = \partial B_{11}^- - \partial H. \tag{65}$$

Further let rank $Q = s$ and employing (8) write $Q = Q_1^+ Q_2^-$, where by definition

$$Q_2^- = \begin{bmatrix} Q_{21}^- \\ 0 \end{bmatrix}$$

with $Q_{21}^- \in \mathfrak{F}_{s,q}[z^{-1}]$, $0 \in \mathfrak{F}_{l-s,q}[z^{-1}]$ and rank $Q_{21}^- = s$. Then, using (13),

$$Q_{21}^{-=} Q_{21}^{-\prime} = (Q_{21}^-)^{*=} (Q_{21}^-)^{*\prime},$$

where $(Q_{21}^-)^* \in \mathfrak{F}_{s,s}[z^{-1}]$ and rank $(Q_{21}^-)^* = s$. For convenience, denote $L = (Q_{21}^-)^*$ and

$$Q^* = Q_1^+ \begin{bmatrix} L \\ 0 \end{bmatrix} \in \mathfrak{F}_{l,s}[z^{-1}]. \tag{66}$$

We shall also use the notation

$$A_1 \frac{Q^*}{p} = \frac{F}{p_0},$$

where $(p_0, F) = 1$, and write $F = F_1^+ F_2^-$. In view of (8) the F_2^- can be written in the form

$$F_2^- = \begin{bmatrix} F_{21}^- \\ 0 \end{bmatrix}$$

with $F_{21}^- \in \mathfrak{F}_{s,s}[z^{-1}]$, $0 \in \mathfrak{F}_{l-s,s}[z^{-1}]$ and rank $F_{21}^- = s$. Then, using (13),

$$F_{21}^- F_{21}^{-=\prime} = (F_{21}^-)^* (F_{21}^-)^{*=\prime},$$

where $(F_{21}^-)^* \in \mathfrak{F}_{s,s}[z^{-1}]$ and rank $(F_{21}^-)^* = s$. For convenience, denote $G = (F_{21}^-)^*$. It can be shown [18] that

$$\partial F_{21}^- - \partial G = 0. \tag{67}$$

We have the following fundamental result.

Theorem 2. *Problem* (64) *has a solution if and only if the linear Diophantine equation*

$$z^{-d}H^{\sim\prime}X + YG^{\sim\prime}p = B_{11}^{-\sim\prime}Q^*F_{21}^{-\sim\prime} \tag{68}$$

has a solution $X^0 \in \mathfrak{F}_{r,s}[z^{-1}]$, $Y^0 \in \mathfrak{F}_{r,s}[z^{-1}]$ *such that* $\partial Y^0 < \partial z^{-d}H^{\sim\prime}$ *and the linear Diophantine equations*

$$B_1 M_1 + N_1 A_1 = I_l\,, \tag{69}$$

$$A_2 N_2 + M_2 B_2 = I_m \tag{70}$$

and

$$\begin{aligned} A_2 M_1 &= M_2 A_1\,, \\ B_1 N_2 &= N_1 B_2 \end{aligned} \tag{71}$$

have solutions $M_1 \in \mathfrak{F}^+_{m,l}\{z^{-1}\}$, $N_1 \in \mathfrak{F}^+_{l,l}\{z^{-1}\}$ *and* $M_2 \in \mathfrak{F}^+_{m,l}\{z^{-1}\}$, $N_2 \in \mathfrak{F}^+_{m,m}\{z^{-1}\}$ *satisfying*

$$HM_{11}F_{21}^{-\sim\prime} = X^0\,, \qquad B_2^+ M_1 Q^* = \begin{bmatrix} M_{11} \\ M_{21} \end{bmatrix}, \tag{72}$$

$$B_{11}^{-\sim\prime}N_{11}G = Y^0 p_0\,, \quad N_1 F_1^+ = [N_{11}\ N_{12}]\,, \tag{73}$$

and

$$F_U = A_2 M_1 F_W \quad \textit{belongs to} \quad \mathfrak{F}^+_{m,q}\{z^{-1}\}\,,$$

$$F_E = N_1 A_1 F_W \quad \textit{belongs to} \quad \mathfrak{F}^+_{l,q}\{z^{-1}\}\,.$$

The optimal controller is not unique, in general, and all optimal controllers are given as minimal realizations of

$$R = M_2 N_1^{-1} = N_2^{-1} M_1\,. \tag{74}$$

Moreover,

$$\begin{aligned} \|F_E\|^2_{\min} = {} & \operatorname{tr}\langle((H^{\sim\prime})^{-1}Y^0)^{=\prime}((H^{\sim\prime})^{-1}Y^0)\rangle + \\ & + \operatorname{tr}\langle F_W^{=\prime}F_W\rangle - \operatorname{tr}\langle F_W^{=\prime}B_{11}^{-}H^{-1}(H^{=\prime})^{-1}B_{11}^{-=\prime}F_W\rangle\,. \end{aligned} \tag{75}$$

Proof. In order to minimize the sum of variances $\|F_E\|^2$ we shall assume that F_E is stable, whereby the E is a stationary vector random sequence and

$$\|F_E\|^2 = \operatorname{tr}\langle F_E^{=}F_E'\rangle = \operatorname{tr}\langle F_E^{=\prime}F_E\rangle \tag{76}$$

in view of (28). Then we will manipulate the expression $\operatorname{tr}\langle F_E^{=\prime}F_E\rangle$ so as to make the minimizing choice of R obvious.

Write

$$\boldsymbol{F}_E = (I_l - \boldsymbol{K}_1)\, \boldsymbol{F}_W .$$

Denoting

$$F_E^* = (I_l - \boldsymbol{K}_1) \frac{Q^*}{p}, \tag{77}$$

it is clear that

$$\boldsymbol{F}_E^{=} \boldsymbol{F}_E' = \boldsymbol{F}_E^{*=} \boldsymbol{F}_E^{*\prime} \tag{78}$$

by virtue of (15) and (43), and

$$\boldsymbol{F}_E = \boldsymbol{F}_E^* L^{-1} Q_{21}^- . \tag{79}$$

To guarantee a stable closed-loop system we have to set $\boldsymbol{K}_1 = B_1 \boldsymbol{M}_1$ for some $\boldsymbol{M}_1 \in \mathfrak{F}_{m,l}^+\{z^{-1}\}$, see (33). Then

$$F_E^* = \frac{Q^*}{p} - B_2 \boldsymbol{M}_1 \frac{Q^*}{p} = \frac{Q^*}{p} - [B_{11}^- \; 0]\, B_1^+ \boldsymbol{M}_1 \frac{Q^*}{p} = \frac{Q^*}{p} - \frac{B_{11}^- \boldsymbol{M}_{11}}{p}, \tag{80}$$

where

$$B_2^+ \boldsymbol{M}_1 Q^* = \begin{bmatrix} \boldsymbol{M}_{11} \\ \boldsymbol{M}_{21} \end{bmatrix}$$

and $\boldsymbol{M}_{11} \in \mathfrak{F}_{r,s}^+\{z^{-1}\}$, $\boldsymbol{M}_{21} \in \mathfrak{F}_{m-r,s}^+\{z^{-1}\}$. Substituting into (78) we obtain

$$\begin{aligned} \boldsymbol{F}_E^{*=\prime} \boldsymbol{F}_E^* &= \frac{Q^{*=\prime} Q^*}{p^= p} - \frac{Q^{*=\prime}}{p^= p} B_{11}^- \boldsymbol{M}_{11} - \boldsymbol{M}_{11}^{=\prime} B_{11}^{-=\prime} \frac{Q^*}{p^= p} + \\ &\quad + \boldsymbol{M}_{11}^{=\prime} B_{11}^{-=\prime} \frac{1}{p^= p} B_{11}^- \boldsymbol{M}_{11} = \\ &= \left(\frac{(H^{=\prime})^{-1} B_{11}^{-=\prime} Q^*}{p} - \frac{H\boldsymbol{M}_{11}}{p} \right)^{=\prime} \left(\frac{(H^{=\prime})^{-1} B_{11}^{-=\prime} Q^*}{p} - \frac{H\boldsymbol{M}_{11}}{p} \right) + \\ &\quad + \frac{Q^{*=\prime} Q^*}{p^= p} - \frac{Q^{*=\prime}}{p^=} B_{11}^- H^{-1} (H^{=\prime})^{-1} B_{11}^{-=\prime} \frac{Q^*}{p} . \end{aligned}$$

Now, by definition,

$$G^{-1} F_{21}^- F_{21}^{-=\prime} (G^{=\prime})^{-1} = I_s$$

 and using the well-known property of the trace of a matrix, we obtain

$$(81)\quad \operatorname{tr} \boldsymbol{F}_E^{*=\prime}\boldsymbol{F}_E =$$

$$= \operatorname{tr}\,(F_{21}^{-=\prime}(G^{=\prime})^{-1})^{=\prime}\left(\frac{(H^{=\prime})^{-1}\,B_{11}^{-=\prime}Q^*}{p} - \frac{HM_{11}}{p}\right)^{=\prime}\cdot$$

$$\cdot\left(\frac{(H^{=\prime})^{-1}\,B_{11}^{-=\prime}Q^*}{p} - \frac{HM_{11}}{p}\right)(F_{21}^{-=\prime}(G^{=\prime})^{-1}) +$$

$$+ \operatorname{tr}\frac{Q^{*=\prime}Q^*}{p^{=}p} - \operatorname{tr}\frac{Q^{*=\prime}}{p^{=}}\,B_{11}^{-}H^{-1}(H^{=\prime})^{-1}\,B_{11}^{-=\prime}\,\frac{Q^*}{p}.$$

Since the last two terms in (81) are independent of $\boldsymbol{M}_{11}$ (and hence $\boldsymbol{M}_1$ and, in turn, $\boldsymbol{R}$) the expression $\operatorname{tr}\langle \boldsymbol{F}_E^{=\prime}\boldsymbol{F}_E\rangle$ attains its minimum for the same controller $\boldsymbol{R}$ as the expression $\operatorname{tr}\langle \boldsymbol{F}_{E0}^{=\prime}\boldsymbol{F}_{E0}\rangle$ does, where

$$(82)\qquad \boldsymbol{F}_{E0} = \frac{(H^{=\prime})^{-1}\,B_{11}^{-=\prime}Q^*F_{21}^{-=\prime}(G^{=\prime})^{-1}}{p} - \frac{HM_{11}F_{21}^{-=\prime}(G^{=\prime})^{-1}}{p} =$$

$$= \frac{(H^{\sim\prime})^{-1}\,B_{11}^{-\sim\prime}Q^*F_{21}^{-\sim\prime}(G^{\sim\prime})^{-1}}{z^{-d}p} - \frac{HM_{11}F_{21}^{-\sim\prime}(G^{\sim\prime})^{-1}}{p}.$$

Here we have used the substitutions

$$(H^{=\prime})^{-1}\,B_{11}^{-=\prime} = \frac{(H^{\sim\prime})^{-1}\,B_{11}^{-\sim\prime}}{z^{-d}},\quad F_{21}^{-=\prime}(G^{=\prime})^{-1} = F_{21}^{-=\prime}(G^{\sim\prime})^{-1}$$

obtained from (10) and (65), (67).

Now take the partial fraction expansion

$$(83)\qquad \frac{(H^{\sim\prime})^{-1}\,B_{11}^{-\sim\prime}Q^*F_{21}^{-\sim\prime}(G^{\sim\prime})^{-1}}{z^{-d}p} = \frac{(H^{\sim\prime})^{-1}\,Y}{z^{-d}} + \frac{X(G^{\sim\prime})^{-1}}{p}.$$

It follows that the matrices X and Y are governed by the Diophantine equation (68).

In view of (82) and (83) we can write

$$(84)\qquad \boldsymbol{F}_{E0} = \frac{(H^{\sim\prime})^{-1}\,Y}{z^{-d}} + \boldsymbol{Z},$$

where

$$(85)\qquad \boldsymbol{Z} = \frac{X(G^{\sim\prime})^{-1}}{p} - \frac{HM_{11}F_{21}^{-\sim\prime}(G^{\sim\prime})^{-1}}{p}.$$

Then (84) implies

$$\text{(86)}\qquad \operatorname{tr}\langle F_{E0}^{=\prime}F_{E0}\rangle = \operatorname{tr}\left\langle\left(\frac{(H^{\sim\prime})^{-1}Y}{z^{-d}}\right)^{=\prime}\left(\frac{(H^{\sim\prime})^{-1}Y}{z^{-d}}\right)\right\rangle +$$

$$+\operatorname{tr}\left\langle\left(\frac{(H^{\sim\prime})^{-1}Y}{z^{-d}}\right)^{=\prime}Z\right\rangle + \operatorname{tr}\left\langle Z^{=\prime}\left(\frac{(H^{\sim\prime})^{-1}Y}{z^{-d}}\right)\right\rangle + \operatorname{tr}\langle Z^{=\prime}Z\rangle\,.$$

Now let

$$\text{(87)}\qquad \begin{aligned} z^{-d}H^{\sim\prime} &= E_{1A}\operatorname{diag}\{a_1, a_2, \ldots, a_r\}\,E_{2A}\,,\\ G^{\sim\prime}p &= E_{1B}\operatorname{diag}\{b_1, b_2, \ldots, b_s\}\,E_{2B} \end{aligned}$$

be the respective canonical decompositions. Then, using (21), any solution of equation (68) can be written in the form

$$\text{(88)}\qquad X = X^0 + E_{2A}^{-1}TE_{2B}\,,$$

$$\text{(89)}\qquad Y = Y^0 - E_{1A}SE_{1B}^{-1}\,,$$

where

$$\text{(90)}\qquad \partial Y^0 < \partial z^{-d}H^{\sim\prime}\,,$$

the matrix $T\in\mathfrak{F}_{r,s}[z^{-1}]$ has elements $t_{ij}b_j/(a_i, b_j)$, the matrix $S\in\mathfrak{F}_{r,s}[z^{-1}]$ has elements $a_it_{ij}/(a_i, b_j)$, and where t_{ij} are arbitrary polynomials of $\mathfrak{F}[z^{-1}]$.

The key observation is that

$$\left(\frac{(H^{\sim\prime})^{-1}Y^0}{z^{-d}}\right)^{=} = z^{-(\partial z^{-d}H^{\sim\prime}-\partial Y^0)}\,(H^{\prime})^{-1}\,Y^{0\sim}$$

is divisible by z^{-1} due to (90). Therefore

$$\left\langle\left(\frac{(H^{\sim\prime})^{-1}Y^0}{z^{-d}}\right)^{=\prime}\left(\frac{(H^{\sim\prime})^{-1}}{z^{-d}}\,E_{1A}SE_{1B}^{-1}\right)\right\rangle = 0\,,$$

$$\left\langle\left(\frac{(H^{\sim\prime})^{-1}Y^0}{z^{-d}}\right)^{=\prime}Z\right\rangle = 0$$

and after substituting (89) into (86) we have

$$\text{(91)}\qquad \operatorname{tr}\langle F_{E0}^{=\prime}F_{E0}\rangle = \operatorname{tr}\left\langle\left(\frac{(H^{\sim\prime})^{-1}Y^0}{z^{-d}}\right)^{=\prime}\left(\frac{(H^{\sim\prime})^{-1}Y^0}{z^{-d}}\right)\right\rangle +$$

$$+\operatorname{tr}\left\langle\left(Z - \frac{(H^{\sim\prime})^{-1}}{z^{-d}}\,E_{1A}SE_{1B}^{-1}\right)\left(Z - \frac{(H^{\sim\prime})^{-1}}{z^{-d}}\,E_{1A}SE_{1B}^{-1}\right)\right\rangle.$$

The first term on the right-hand side of (91) cannot be affected by any choice of $\boldsymbol{M}_{11}$ (and hence $\boldsymbol{R}$). The best we can do to minimize (92) is to set

$$Z - \frac{(H^{\sim\prime})^{-1}}{z^{-d}} E_{1A} S E_{1B}^{-1} = 0\,, \tag{92}$$

i.e.

$$\frac{X(G^{\sim\prime})^{-1}}{p} - \frac{H\boldsymbol{M}_{11} F_{21}^{-\sim\prime}(G^{\sim\prime})^{-1}}{p} - \frac{(H^{\sim\prime})^{-1}}{z^{-d}} E_{1A} S E_{1B}^{-1} = 0$$

by (85). But

$$\frac{X(G^{\sim\prime})^{-1}}{p} - \frac{(H^{\sim\prime})^{-1}}{z^{-d}} E_{1A} S E_{1B}^{-1} = \frac{X^0(G^{\sim\prime})^{-1}}{p}$$

because

$$X - \frac{(H^{\sim\prime})^{-1}}{z^{-d}} E_{1A} S E_{1B}^{-1} G^{\sim\prime} p =$$

$$= X - E_{2A}^{-1} \operatorname{diag}\left\{\frac{1}{a_1}, \frac{1}{a_2}, \ldots, \frac{1}{a_r}\right\} S \operatorname{diag}\{b_1, b_2, \ldots, b_s\} E_{2B} =$$

$$= X - E_{2A}^{-1} T E_{2B} = X^0$$

on successive application of (87), the definition of T and S, and (88).

Therefore (91), and, in view of the above discussion and (80), also the tr $\langle \boldsymbol{F}_E' \boldsymbol{F}_E \rangle$ is minimized by setting

$$H\boldsymbol{M}_{11} F_{21}^{-\sim\prime} = X^0\,. \tag{93}$$

Substituting (93) into (80) we obtain

$$B_{11}^{-\sim\prime} \boldsymbol{F}_E^* F_{21}^{-\sim\prime} = \frac{B_{11}^{-\sim\prime} Q^* F_{21}^{-\sim\prime} - z^{-d} H^{\sim\prime} X^0}{p} = Y^0 G^{\sim\prime} \tag{94}$$

on making use of equation (68). Now consider (77) and set $I_l - \boldsymbol{K}_1 = N_1 A_1$ for some $N_1 \in \mathfrak{F}_{l,l}^{+}\{z^{-1}\}$ to guarantee a stable closed-loop system, see (33). Then

$$B_{11}^{-\sim\prime} \boldsymbol{F}_E^* F_{21}^{-\sim\prime} = B_{11}^{-\sim\prime} N_1 A_1 \frac{Q^*}{p} F_{21}^{-\sim\prime} = \tag{95}$$

$$= B_{11}^{-\sim\prime} N_1 \frac{F}{p_0} F_{21}^{-\sim\prime} = B_{11}^{-\sim\prime} N_1 \frac{F^+}{p_0} \begin{bmatrix} F_{21}^{-} \\ 0 \end{bmatrix} F_{21}^{-\sim\prime} = \frac{B_{11}^{-\sim\prime} N_{11} G G^{\sim\prime}}{p_0}\,,$$

where

$$N_1 F^+ = [N_{11}\ N_{12}]$$

and $N_{11} \in \mathfrak{F}^+_{l,s}\{z^{-1}\}$, $N_{12} \in \mathfrak{F}^+_{l,l-s}\{z^{-1}\}$. The comparison of (94) and (95) yields

(96) $$B_{11}^{-\sim\prime} N_{11} G = Y^0 p_0 \,.$$

To summarize, we have to first solve the stability equations (69) and (70) for stable $\boldsymbol{M}_1$, N_1 and $\boldsymbol{M}_2$, N_2 and then restrict the solutions so as to satisfy the mutual relations (71). Then the closed-loop system will be stable. Further we solve the optimality equation (68) for X^0, Y^0 such that $\partial Y^0 < \partial z^{-d} H^{\sim\prime}$ and, in order to obtain an optimal as well as stable closed-loop system, we have to satisfy relations (72), (73). Moreover, it is necessary to require that $\boldsymbol{F}_U$, which is given by

$$\boldsymbol{S}\boldsymbol{F}_U = \boldsymbol{K}_1 \boldsymbol{F}_W \,, \quad \boldsymbol{S} = B_1 A_2^{-1} \,, \quad \boldsymbol{K}_1 = B_1 \boldsymbol{M}_1 \,,$$

as

$$\boldsymbol{F}_U = A_2 \boldsymbol{M}_1 \boldsymbol{F}_W \,,$$

be stable according to the problem statement and that

$$\boldsymbol{F}_E = (I_l - \boldsymbol{K}_1)\, \boldsymbol{F}_W = N_1 A_1 \boldsymbol{F}_W$$

be stable to satisfy hypothesis (76).

Then all optimal controllers are given as minimal realizations of (74) because

$$\boldsymbol{S}\boldsymbol{R} = \boldsymbol{K}_1 (I_l - \boldsymbol{K}_1)^{-1} \,,$$

$$\boldsymbol{S} = B_1 A_2^{-1} \,, \quad \boldsymbol{K}_1 = B_1 \boldsymbol{M}_1 \,, \quad I_l - \boldsymbol{K}_1 = N_1 A_1$$

yields

$$\boldsymbol{R} = A_2 \boldsymbol{M}_1 A_1^{-1} N_1^{-1} = \boldsymbol{M}_2 A_1 A_1^{-1} N_1^{-1} = \boldsymbol{M}_2 N_1^{-1}$$

on applying (71), and

$$\boldsymbol{R}\boldsymbol{S} = (I_m - \boldsymbol{K}_2)^{-1}\, \boldsymbol{K}_2 \,,$$

$$\boldsymbol{S} = A_1^{-1} B_2 \,, \quad \boldsymbol{K}_2 = \boldsymbol{M}_2 B_2 \,, \quad I_m - \boldsymbol{K}_2 = A_2 N_2$$

yields

$$\boldsymbol{R} = N_2^{-1} A_2^{-1} \boldsymbol{M}_2 A_1 = N_2^{-1} A_2^{-1} A_2 \boldsymbol{M}_1 = N_2^{-1} \boldsymbol{M}_1$$

again applying (71).

The minimized sum of steady-state variances of the individual components of $\boldsymbol{E}$ is given by (75) when (76), (78), (81), (82), (91) and relations

$$z^{-d=} z^{-d} = 1$$

$$\operatorname{tr} Q^{*=\prime} Q^* = \operatorname{tr} Q^{*=} Q^{*\prime} = \operatorname{tr} Q^{=} Q' = \operatorname{tr} Q^{=\prime} Q$$

are taken into account. Note that when $r = l$ the B_{11}^- is invertible and, by definition,

$$B_{11}^- H^{-1} (H^{=\prime})^{-1}\, (B_{11}^{-=\prime})^{-1} = I_l \,.$$

Then (75) simplifies to

$$\|F_E\|^2_{\min} = \operatorname{tr} \langle ((H^{\sim\prime})^{-1} Y^0)^{=\prime} ((H^{\sim\prime})^{-1} Y^0) \rangle . \qquad \square$$

Example 4. Consider problem (2) for the system $\mathscr{S} = \mathscr{S}_1 \mathscr{S}_2$ over $\mathfrak{F}$ which is a minimal realization of

$$S_1 = \frac{\begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix}}{(1 - z^{-1})(1 - 0{\cdot}5z^{-1})}, \quad S_2 = 1$$

and the disturbance given by

$$F_V = 1 .$$

The system is depicted in Fig. 8 and can be interpretted as a two-dimensional single-variable system in which the auxiliary output Y_2 is used to improve the control.

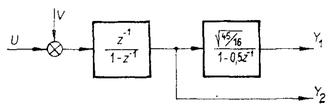

Fig. 8. The system in Example 4.

The problem can be recast in terms of Fig. 7 by setting

$$S = \frac{\begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix}}{(1 - z^{-1})(1 - 0{\cdot}5z^{-1})}, \quad F_W = \frac{\begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix}}{(1 - z^{-1})(1 - 0{\cdot}5z^{-1})} .$$

Computing the canonical decomposition

$$\begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix} = \begin{bmatrix} \sqrt{45/16} & 0 \\ 1 - 0{\cdot}5z^{-1} & 1 \end{bmatrix} \begin{bmatrix} z^{-1} \\ 0 \end{bmatrix},$$

we obtain

$$B_1 = \begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix}, \quad A_2 = (1 - z^{-1})(1 - 0{\cdot}5z^{-1}),$$

$$A_1 = \begin{bmatrix} \sqrt{16/45}\,(1 - z^{-1})(1 - 0{\cdot}5z^{-1}) & 0 \\ -\sqrt{16/45}\,(1 - 0{\cdot}5z^{-1}) & 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} z^{-1} \\ 0 \end{bmatrix}$$

by (26) and also

$$Q = \begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix}, \quad p = (1 - z^{-1})(1 - 0{\cdot}5z^{-1}) .$$

It is seen that

$$l = 2 , \quad m = q = 1 , \quad r = s = 1 .$$

We shall first guarantee the closed-loop stability by solving equations (69), (70) and (71).
Equation (69) becomes

$$\begin{bmatrix} \sqrt{45/16}\, z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix} M_1 + N_1 \begin{bmatrix} \sqrt{16/45}\,(1 - z^{-1})(1 - 0{\cdot}5z^{-1}) & 0 \\ -\sqrt{16/45}\,(1 - 0{\cdot}5z^{-1}) & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and it is equivalent to the polynomial equations

$$z^{-1}\hat{m}_{11}^1 + \hat{n}_{11}^1(1 - z^{-1})(1 - 0{\cdot}5z^{-1}) = 1\,, \quad z^{-1}\hat{m}_{12}^1 + \hat{n}_{12}^1 = 0\,,$$

$$\hat{n}_{21}^1(1 - z^{-1})(1 - 0{\cdot}5z^{-1}) = 0\,, \qquad \hat{n}_{22}^1 = 1$$

where

$$M_1 = [\hat{m}_{11}^1\ \hat{m}_{12}^1] \begin{bmatrix} \sqrt{16/45} & 0 \\ -\sqrt{16/45}\,(1 - 0{\cdot}5z^{-1}) & 1 \end{bmatrix},$$

$$N_1 = \begin{bmatrix} \sqrt{45/16} & 0 \\ 1 - 0{\cdot}5z^{-1} & 1 \end{bmatrix} \begin{bmatrix} \hat{n}_{11}^1 & \hat{n}_{12}^1 \\ \hat{n}_{21}^1 & \hat{n}_{22}^1 \end{bmatrix}.$$

We obtain

$$\hat{m}_{11}^1 = 1{\cdot}5 - 0{\cdot}5z^{-1} + (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, t_{11}\,, \quad \hat{m}_{12}^1 = t_{12}\,,$$

$$\hat{n}_{11}^1 = 1 - z^{-1}t_{11}\,, \quad \hat{n}_{12}^1 = -z^{-1}t_{12}\,,$$

$$\hat{n}_{21}^1 = 0\,, \qquad \hat{n}_{22}^1 = 1\,,$$

for arbitrary $t_{11}, t_{12} \in \mathfrak{F}^+\{z^{-1}\}$.

Equation (70) becomes

$$(1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, N_2 + M_2 \begin{bmatrix} z^{-1} \\ 0 \end{bmatrix} = 1$$

and it is equivalent to the polynomial equation

$$(1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, \hat{n}_{11}^2 + \hat{m}_{11}^2 z^{-1} = 1$$

where

$$N_2 = [\hat{n}_{11}^2]\,, \quad M_2 = [\hat{m}_{11}^2\ v_{12}]\,.$$

We obtain

$$\hat{n}_{11}^2 = 1 - z^{-1}v_{11}\,,$$

$$\hat{m}_{11}^2 = 1{\cdot}5 - 0{\cdot}5z^{-1} + (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, v_{11}$$

for arbitrary $v_{11}, v_{12} \in \mathfrak{F}^+\{z^{-1}\}$.

The mutual relations (71) then yield

$$v_{11} = t_{11}\,,$$

$$v_{12} = (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, t_{12}$$

 and, finally,

$$M_1 = [\sqrt{16/45}\,(1{\cdot}5 - 0{\cdot}5z^{-1}) + \\ + \sqrt{16/45}\,(1 - z^{-1})(1 - 0{\cdot}5z^{-1})\,t_{11} - \sqrt{16/45}\,(1 - 0{\cdot}5z^{-1})\,t_{12} \quad t_{12}],$$

$$N_1 = \begin{bmatrix} \sqrt{45/16} - \sqrt{45/16}\,z^{-1}t_{11} & -\sqrt{45/16}\,z^{-1}t_{12} \\ 1 - 0{\cdot}5z^{-1} - z^{-1}(1 - 0{\cdot}5z^{-1})\,t_{11} & 1 - z^{-1}(1 - 0{\cdot}5z^{-1})\,t_{12} \end{bmatrix},$$

and

$$M_2 = [1{\cdot}5 - 0{\cdot}5z^{-1} + (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\,t_{11} \quad (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\,t_{12}],$$

$$N_2 = 1 - z^{-1}t_{11}.$$

Further we shall solve equation (68) to minimize the optimality criterion. Computing

$$B_{11}^- = \begin{bmatrix} \sqrt{45/16}\,z^{-1} \\ z^{-1}(1 - 0{\cdot}5z^{-1}) \end{bmatrix}, \quad B_2^+ = 1, \quad B_{11}^{-\sim\prime} = [\sqrt{45/16}\,z^{-1} \quad z^{-1} - 0{\cdot}5],$$

$$H = 2 - 0{\cdot}25z^{-1}, \quad H^{\sim\prime} = 2z^{-1} - 0{\cdot}25, \quad d = 1$$

and

$$Q_1^+ = \begin{bmatrix} \sqrt{45/16} & 0 \\ 1 - 0{\cdot}5z^{-1} & 1 \end{bmatrix}, \quad Q_2^- = \begin{bmatrix} z^{-1} \\ 0 \end{bmatrix}, \quad Q_{21}^- = z^{-1},$$

$$L = 1, \quad Q^* = \begin{bmatrix} \sqrt{45/16} \\ 1 - 0{\cdot}5z^{-1} \end{bmatrix}$$

and

$$F = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad F_1^+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad F_2^- = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad F_{21}^- = 1,$$

$$G = 1, \quad p_0 = 1,$$

equation (68) becomes

$$z^{-1}(2z^{-1} - 0{\cdot}25)\,X + Y(1 - z^{-1})(1 - 0{\cdot}5z^{-1}) = \\ = [\sqrt{45/16}\,z^{-1} \quad z^{-1} - 0{\cdot}5] \begin{bmatrix} \sqrt{45/16} \\ 1 - 0{\cdot}5z^{-1} \end{bmatrix} = -0{\cdot}5 + \tfrac{65}{16}z^{-1} - 0{\cdot}5z^{-2}.$$

Its general solution is

$$X = 2{\cdot}75 - z^{-1} + (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\,t,$$

$$Y = -0{\cdot}5 + 4z^{-1} - z^{-1}(2z^{-1} - 0{\cdot}25)\,t$$

for arbitrary polynomial $t \in \mathfrak{F}[z^{-1}]$, and the solution X^0, Y^0 satisfying $\partial Y^0 < 2$ reads

$$X^0 = 2{\cdot}75 - z^{-1}, \quad Y^0 = -0{\cdot}5 + 4z^{-1}.$$

Now we have to satisfy relations (72) and (73) thereby putting the conditions of stability and optimality together. Computing

$$M_{11} = 1{\cdot}5 - 0{\cdot}5z^{-1} + (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, t_{11},$$

$$N_{11} = \begin{bmatrix} \sqrt{45/16}\backslash - \sqrt{45/16}\backslash\, z^{-1} t_{11} \\ 1 - 0{\cdot}5z^{-1} - z^{-1}(1 - 0{\cdot}5z^{-1})\, t_{11} \end{bmatrix}$$

by (79) and (80), we obtain

$$t_{11} = -\frac{0{\cdot}25}{2 - 0{\cdot}25z^{-1}},$$

$$t_{12} \in \mathfrak{F}^{+}\{z^{-1}\} \quad \text{arbitrary}.$$

Therefore

$$M_1 = \left[\sqrt{16/45}\backslash \frac{2{\cdot}75 - z^{-1}}{2 - 0{\cdot}25z^{-1}} - \sqrt{16/45}\backslash\, (1 - 0{\cdot}5z^{-1})\, t_{12} \quad t_{12}\right],$$

$$N_1 = \begin{bmatrix} \sqrt{45/16}\backslash \dfrac{2}{2 - 0{\cdot}25z^{-1}} & -\sqrt{45/16}\backslash\, z^{-1} t_{12} \\ 2\dfrac{1 - 0{\cdot}5z^{-1}}{2 - 0{\cdot}25z^{-1}} & 1 - z^{-1}(1 - 0{\cdot}5z^{-1})\, t_{12} \end{bmatrix}$$

and

$$M_2 = \left[\frac{2{\cdot}75 - z^{-1}}{2 - 0{\cdot}25z^{-1}} \quad (1 - z^{-1})(1 - 0{\cdot}5z^{-1})\, t_{12}\right],$$

$$N_2 = \frac{2}{2 - 0{\cdot}25z^{-1}}$$

and also

$$F_U = z^{-1}\frac{2{\cdot}75 - z^{-1}}{2 - 0{\cdot}25z^{-1}}, \quad F_E = \begin{bmatrix} 2z^{-1}\dfrac{\sqrt{45/16}}{2 - 0{\cdot}25z^{-1}} \\ 2z^{-1}\dfrac{1 - 0{\cdot}5z^{-1}}{2 - 0{\cdot}25z^{-1}} \end{bmatrix}.$$

As the above matrices are stable, our problem has a solution. The optimal controller is not unique and all optimal controllers are given by (74) as minimal realizations of

$$(97) \qquad R = [\tfrac{1}{2}\sqrt{16/45}\backslash\, (2{\cdot}75 - z^{-1}) - \\ - \tfrac{1}{2}\sqrt{16/45}\backslash\, (1 - 0{\cdot}5z^{-1})(2 - 0{\cdot}25z^{-1})\, t_{12} \quad \tfrac{1}{2}(2 - 0{\cdot}25z^{-1})\, t_{12}].$$

All these controllers give the minimized sum of steady-state output variances equal to

$$\|F_E\|^2_{\min} = \frac{20}{7} + \frac{8}{7} = 4$$

by (75).

The nonuniqueness of optimal controller can be utilized to meet additional requirements. For instance, choosing

$$\boldsymbol{t}_{12} = \frac{2}{2 - 0{\cdot}25z^{-1}}$$

in (97), the resulting controller

$$\boldsymbol{R}_0 = \left[\tfrac{3}{8}\sqrt{16/45} \quad 1\right]$$

has the least possible dimension (equal to zero) among all optimal controllers.

Choosing

$$\boldsymbol{t}_{12} = \frac{1}{1 - 0{\cdot}5z^{-1}} \frac{2{\cdot}75 - z^{-1}}{2 - 0{\cdot}25z^{-1}}$$

we get

$$\boldsymbol{R}_1 = \left[0 \quad \frac{1}{1 - 0{\cdot}5z^{-1}} \frac{2{\cdot}75 - z^{-1}}{2 - 0{\cdot}25z^{-1}}\right], \tag{98}$$

and choosing

$$\boldsymbol{t}_{12} = 0$$

we get

$$\boldsymbol{R}_2 = \left[\tfrac{1}{8}\sqrt{16/45}\,(11 - 4z^{-1}) \quad 0\right]. \tag{99}$$

Therefore, the controller $\mathscr{R}_0$ is the simplest one to realize, while the controllers $\mathscr{R}_1$ and $\mathscr{R}_2$ might be used as emergency controllers in case of breakdown of the first and second control channel respectively. Of course, all these controllers give the same optimal performance.

(Received May 20, 1974.)

REFERENCES

[1] Åström K. J.: Introduction to the Stochastic Control Theory. Academic Press, New York 1970.

[2] Chang S. S. L.: Synthesis of Optimum Control Systems. McGraw-Hill, New York 1961.

[3] Doob J. L.: Stochastic Processes. Wiley, New York 1955.

[4] Гантмахер Ф. Р.: Теория матриц. Гостехиздат, Москва 1953.

[5] Гнеденко Б. В.: Курс теории вероятностей. Физматгиз, Москва 1961.

[6] Gunckel T. L., Franklin G. F.: A general solution for linear, sampled-data control. Trans. ASME, J. Basic Eng. *85D* (1963), 197—203.

[7] Halousková A.: Syntéza mnohaparametrových lineárních diskrétních regulačních obvodů podle kvadratických kritérií. Rept. No. 174, ČSAV ÚTIA, Prague 1966.

[8] Jazwinski A. H.: Stochastic Processes and Filtering Theory. Academic Press, New York 1970.

[9] Kalman R. E., Falb P. L., Arbib M. A.: Topics in Mathematical System Theory. McGraw-Hill, New York 1969.

[10] Kalman R. E., Koepcke R. W.: Optimal synthesis of linear sampling control systems using generalized performance indexes. Trans. ASME, J. Basic Eng. *80D* (1958), 1812—1817.

[11] Kalman R. E.: A new approach to linear filtering and prediction problems. Trans. ASME, J. Basic Eng. *82D* (1960), 35—45.

[12] Kučera V.: Algebraic theory of discrete optimal control for single-variable systems I — Preliminaries. Kybernetika *9* (1973), 1, 94—107.

[13] Kučera V.: Algebraic theory of discrete optimal control for single-variable systems II — Open-loop control. Kybernetika *9* (1973), 3, 206—221.
[14] Kučera V.: Algebraic theory of discrete optimal control for single-variable systems III — Closed-loop control. Kybernetika *9* (1973), 4, 291—312.
[15] Kučera V.: Closed-loop stability of discrete linear single-variable systems. Kybernetika *10* (1974), 2, 162—184.
[16] Kučera V.: Constrained optimal control — The algebraic approach. Kybernetika *10* (1974), 4, 317—349.
[17] Kučera V.: Algebraic approach to discrete linear control. IEEE Trans. Automatic Control. *AC-20* (1975), 1, 116—120.
[18] Kučera V.: Algebraic theory of discrete optimal control for multivariable systems. Supplement of Kybernetika, Academia, Prague 1974—1975.
[19] Kučera V.: The discrete Riccati equation of optimal control. Kybernetika *8* (1972), 5, 430—447.
[20] Kučera V.: A contribution to matrix quadratic equations. IEEE Trans. Automatic Control *AC-17* (1972), 3, 344—347.
[21] Kučera V.: On nonnegative definitive solutions to matrix quadratic equations. Proc. 5th IFAC Congress, Paris 1972.
[22] Kučera V.: On nonnegative definite solutions to matrix quadratic equations. Automatica *8* (1972), 4, 413—423.
[23] Lang S.: Algebra. Addision-Wesley, Reading, Mass. 1961.
[24] Laning J. H., Battin R. H.: Random Processes in Automatic Control. McGraw-Hill, New York 1956.
[25] Levin B. R.: Teorie náhodných procesů a její aplikace v radiotechnice. SNTL, Prague 1965.
[26] Meditch J. S.: Stochastic Optimal Linear Estimation and Control. McGraw-Hill, New York 1969.
[27] Peterka V.: On steady-state minimum variance control strategy. Kybernetika *8* (1972), 3, 219—232.
[28] Пугачев Б. С.: Теория случайных функций. Физматгиз, Москва 1960.
[29] Sage A. P.: Optimum Systems Control. Prentice-Hall, Englewood Cliffs, N. J., 1968.
[30] Strejc V. et al.: Syntéza regulačních obvodů s číslicovým počítačem. NČSAV, Prague 1965.
[31] Wiener N.: The Extrapolation, Interpolation and Smoothing of Stationary Time Series. Wiley, New York 1949.
[32] Youla D. C.: On the factorization of rational matrices. IRE Trans. Information Theory *IT-7* (1961), 3, 172—189.
[33] Zariski O., Samuel P.: Commutative Algebra, vols. I and II. Van Nostrand, Princeton, N.J. 1958.

Ing. Vladimír Kučera, CSc.; Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 180 76 Praha 8. Czechoslovakia.