THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SYNTHESIS OF OPTIMUM LINEAR MULTIVARIABLE SYSTEMS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

SAYYED MOHAMMAD MOFEEZ

Norman, Oklahoma

1967

SYNTHESIS OF OPTIMUM LINEAR MULTIVARIABLE SYSTEMS

APPROVED BY

C. M. Sliepcevich

[illegible]

C. R. Haden

J. A. Payne

Arthur Bernhart

DISSERTATION COMMITTEE

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to the many individuals from whom assistance and encouragement have been obtained for the preparation of this work. In particular, gratitude is extended to Dr. C. M. Sliepcevich, Dr. M. L. McGuire, Dr. A. F. Bernhart, and Dr. J. A. Payne for the advice and the invaluable help and comments.

The assistance and interest of the personnel of the Process Control Laboratory, Dr. R. H. Luecke, R. A. Sims, and R. L. Loeffler are also greatly appreciated.

Thanks are also extended to the National Science Foundation and the University of Oklahoma Research Institute for the financial assistance during this work.

Many thanks are due to Mrs. Shirley Ray for her patience and aid in organizing and typing of this dissertation.

Last but not least I must thank my wife, Leili, for her patience, understanding, and encouragement throughout the entire work.

ABSTRACT

Analytical synthesis methods for the design of optimum, linear, constant coefficient, multivariable systems are developed. Matrix methods are used to extend well-known concepts in single variable systems to the treatment of systems with multiple inputs and outputs. The performance criterion used is a minimized weighted sum of the mean-square errors between the set of actual outputs and a set of ideal or desired outputs. Standard variational procedures are used to show that the optimization of linear, multivariable systems leads to a set of simultaneous Wiener-Hopf equations. The derivation of these equations has been achieved in two different ways. In one method the derivation is carried out strictly in the time domain. A proposed alternative method is used to show the advantage of working only in the frequency domain. The resulting vector Wiener-Hopf equation is then properly modified and is solved by a combination of the "function method" and the "method of undetermined coefficients." Two non-trivial, comprehensive examples have been worked out in detail for illustrative purposes. Analog computer simulation was used in this phase of the work advantageously.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SYNTHESIS OF OPTIMUM LINEAR MULTIVARIABLE SYSTEMS

CHAPTER I

INTRODUCTION

In recent years much has been done in the area of multivariable control system synthesis and design. The philosophy behind a design-by-synthesis problem is that of cancellation compensation, i.e., the design of a controller to cancel the undesired dynamics of the fixed element or plant and to substitute the desired dynamics. If the desired dynamics are properly chosen, this concept can be applied and a satisfactory design achieved. However, if the desired dynamics are not properly chosen, the resulting design may yield a completely unsatisfactory system. Therefore the crucial step in this type of design is the choice of the desired dynamics.

The purpose of this work is to find the elements of the desired dynamic matrix using the minimized mean-square sum of the weighted errors as the performance index.

General Optimization Problem

In an optimum synthesis problem, one, given an input or a set of inputs and a desired output or a set of desired outputs, is required to determine the mathematical description of the system which minimizes or maximizes a given performance index. The objective of using a performance index is to encompass in a single number a quality measure for the performance of the system. Typical performance indices used are integral - square error for deterministic signals and mean-square error for random or stochastic signals.

If analytical procedures are used in synthesis and design, no explicit statement concerning the degree of stability of the system is required. All the solutions for the optimum system include the twin requirements that the system be stable and realizable in the sense of having an impulse response that is zero for time less than zero. The degree of stability is found as a part of the solution of the problem and is a function of other specifications.

At this point the classification of the problem must be determined, that is, whether it is characterized by a free configuration, a semi-free configuration, or a fixed configuration. A problem is classified as a free configuration if there is no specification for the configuration of the system [N.2]. It is called a semi-free configuration problem if part of the system (fixed elements) are specified but there is no restriction on the configuration of the compensating network. And finally a fixed configuration problem is one for which the design freedom is limited to the adjustments of one or two parameters.

Depending on either of these classifications, the performance index is expressed in terms of the free parameters of the system. Then, using standard variational techniques, the newly defined functional is minimized or maximized which in turn leads to the values of free parameters which minimize or maximize the given performance index.

Statement of Problem

Multivariable systems are those systems in which one or more input variables control one or more output variables. In other words, multivariable systems are characterized by coupling or interaction.

Consider a linear, time-invariant, multivariable system such as shown in Figure 1-1 with inputs $r_1(t)$, ... $r_p(t)$ and outputs $c_1(t)$, ... $c_q(t)$. The number of inputs, p, and the number of outputs, q, are not necessarily equal. Furthermore assume all signals to be stationary, stochastic, continuous time functions with power spectra expressible as a rational function of frequency. Let $w_{ij}(t)$ be the impulse response of the system measured at output terminal i due to a unit impulse applied at jth input-terminal. Then:

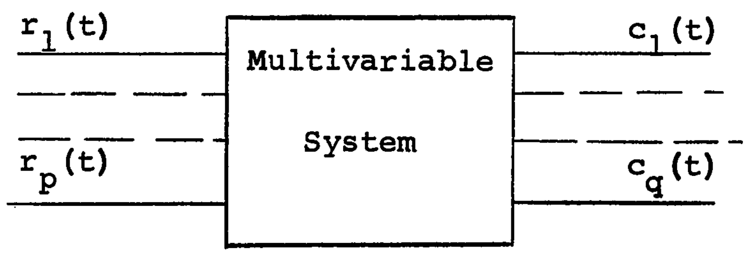

Figure 1-1. A Multivariable Time-Invariant Linear System

$$c_{ij}(t) = \int_{-\infty}^{\infty} w_{ij}(\tau)\; r_j(t-\tau)d\tau \qquad (1\text{-}1)$$

The total output at the ith terminal, when all the inputs $r_1(t)$ through $r_p(t)$ act simultaneously, is

$$c_i(t) = \sum_{j=1}^{p} \int_{-\infty}^{\infty} w_{ij}(\tau)\; r_j(t-\tau)d\tau \qquad (1\text{-}2)$$

If $w_{ij}(\tau)$ is zero for some pair, then there is simply no connection between that input output pair [P.2].

In a synthesis-type problem one assumes knowledge only of the input and the required output. The superposition integral (1-1) provides the mathematical expression of the manner in which the output $c_{ij}(t)$, is explicitly related to the input, $r_j(t-\tau)$, and the impulse response, $w_{ij}(\tau)$, for a linear time-invariant system. Contrary to the design-by-analysis problem, for which this relationship is explicit, the same expression, when employed for design-by-synthesis problem, is implicit. Given the inputs and outputs the solution to the synthesis problem is a $w_{ij}(\tau)$ satisfying (1-1). Solving (1-1) for c_i, given, w_{ij} and r_j is clearly an easier problem than solving for w_{ij}, given r_j and c_i [P.2]. Indeed, there may be even questions regarding whether or not any w_{ij} function exists, satisfying (1-1), for arbitrarily specified functions for r_j and c_i. The difficulties in solving equations such as (1-1) constitute a formidable problem in the field of mathematics covered by the subject of integral equations [P.2].

Suppose that $d_i(t)$ is the desired total output at the terminal i. The problem is to determine the elements of the system matrix, $w_{ij}(t)$, such that the sum of the weighted mean-square errors be minimized.

Before proceeding any further a few words with regard to the choice of the mean-square error criterion is in order. An ideal performance index should have the following properties

i) reliability

ii) ready applicability

iii) selectivity

While in the engineering literature non-mean-square error criteria for predictors are often presented as physically significant and then shunted aside because of mathematical unmanageability, it has been shown [S.1] that, in the case of Gaussian processes, all such criteria yield the same predictor as the linear mean-square predictor of Wiener. Therefore, despite the fact that there has been in the past and there continues to exist a certain amount of controversy regarding the selection of the mean-square error criterion, it is to be used because it has the merit of lending itself to analytical treatment and of being applicable to many real problems.

In returning to the original problem one notes that the problem, as specified, is of the free configuration type since no restriction is imposed on the forms of $w_{ij}(t)$

(other than physical realizability and stability). Note also that when the input signals are the sum of messages and noises, where the message part is the desired or information signal and the noise is the background or unwanted disturbance, and it is desired to recover the desired signals from the total inputs in a mean-square sense, the problem becomes one of determining an optimum linear filter. In solving this type of optimization problem, three basic steps are usually involved. 1) Expressing the performance index in terms of the free parameters of the problem. 2) Deriving the necessary set of equations whose simultaneous solutions are, if there exist such solutions, the required weighting functions. 3) Solving these equations for the solutions.

These three steps are followed closely in the chapter dealing with the detailed presentation of the problem and the method of approach to its solution.

Review of Previous Work

Historically, attempts to develop methods of synthesis began with Wiener's theory for the optimum constant coefficient system. The system obtained by Wiener's approach is optimum in the sense that the time-average mean-square error is minimum. Wiener's early approach embodies the somewhat restrictive condition that the statistical properties of the signals, both messages and noise, be time-invariant.

In 1951, Bootan [B.3] developed a generalization of the Wiener theory which provides minimum mean-square error at any time for linear time-varying systems with a set of nonstationary inputs. Bootan's method resulted in an integral equation which gives the solution as the impulse response of the optimum system. The resulting integral equation did not lend itself to a general solution and it was not until 1956 that Shinbrot [S.2] suggested an approximate method, which from an engineering point of view, "is extremely useful in a wide variety of practical problems" [P.2].

In 1961, Peterson [P.2] used the standard variational technique to develop Bootan's integral equation. This integral equation is also referred to as a modified Wiener-Hopf linear integral. Peterson then used Shinbrot's method to solve the modified Wiener-Hopf equation for both single-input single-output and multiple-input multiple-output linear time-varying systems. Shinbrot's method had the disadvantage that it required that the correlation functions be approximated in a way that permits the integral equation to be solved. It also required that the weighting functions be separable with respect to the time of observation, t, and the time of application of unit impulse, τ.

Through the introduction of the state variable concept, Kalman [K.1] approached the problem of linear filtering and prediction in such a way to avoid attacking the vector Wiener-Hopf integral equation directly. A nonlinear differential equation of the Riccati type is derived for the covariance matrix of the optimum filtering error. Through the duality principle* the optimal control problem with generalized quadratic error criterion is also solved. However, the completion of the synthesis was also dependent on the solution of Riccati equation. It is known that this equation can only be solved analytically in some simple and special cases but in general numerical computation has to be used to obtain the final answer.

With the application of matrix relationships to describe the behavior of multidimensional linear systems, Matyas [M.1], Amara [A.2] and Freeman [F.1] noted that the optimum synthesis problem of free configuration or semi-free configuration type leads to a vector Wiener-Hopf equation. The main problem in performing any design in this area is therefore the solution of this type of equation. The method of undetermined coefficients was initially suggested by Wiener [W.1]. This method was later developed by Amara [A.2] and Hsieh and Leondes [B.4]. Until recently it was the only practical method of solving problems of this type. Its chief disadvantage is that it failed to give an explicit solution so that auxiliary information, such as

---

*Two systems are said to be dual if the equations which characterize one of them have the same mathematical form as the equations that characterize the other.

the mean-square error, is difficult to obtain and in addition it is poorly adapted to either hand or machine calculations.

In 1958 Wiener and Massani [B.4] presented an alternative method for solving these general equations. In a rigorous mathematical style a method of spectral factorization of matrices was developed which enabled them to obtain an explicit solution to the discrete filtering problem. Unfortunately their result is obtained in terms of an infinite series of matrices and, consequently, it is inconvenient for many purposes [B.4]. In 1959 Amara [A.2] extended Wiener's theory of optimum filter design to linear, multivariable, constant coefficient systems. He presented an implicit method of solving the vector Wiener-Hopf equation.

Recently, additional advances have been made in the area of spectral factorization of rational matrices. Youla [Y.1] developed an algebraic technique for spectral factorization which has certain advantages over the Wiener-Massani algorithm. Although this procedure always results in a closed form solution, it is quite complex and even small systems require a great deal of computational labor. A short time after Youla's work, Wong and Thomas [W.2] presented an excellent discussion of the Wiener-Masoni work and developed a third method of spectral factorization of rational matrices. Unfortunately their results are given in terms of a set of simultaneous matrix equations which are difficult to solve by any means [B.4].

In 1962 Brockett and Mesarovic [B.4] developed the so-called "functional method" which completely circumvents the problem of spectral factorization of matrices in the process of solving a vector Wiener-Hopf equation. Unlike previous techniques this method gives an explicit solution without requiring the spectral factorization of matrices. Contrary to the method of undetermined coefficient suggested by Wiener and extended by Amara [A.2] the functional method determines one complete column of the optimum weighting function matrix each time it cycles through the suggested steps.

In 1963 Hsieh [H.1] used the method of steepest-descent in the Hilbert space to solve the optimum synthesis of multivariable control systems with a generalized quadratic error criterion. Apart from the conventional minimization techniques, his approach worked directly on the error criterion functional itself so that the optimal solution can be determined through successive approximations with the aid of a digital computer. Obviously the method has computational advantages. However, as noted by Dorato [D.1] the disadvantage of such techniques, at least for certain problems, is the open loop nature of the resulting control law. Also as has been pointed out by the author himself the main feature of this method is to synthesize an optimal control when the system weighting function matrix is obtained experimentally.

In this work a method is suggested which is a combination of the "functional method" by Brockett and Mesarovi [B.4] and the "method of undetermined coefficients" developed by Amara [A.2]. The details of this method are covered in Chapter III. However, there are some mathematical concepts and definitions which are the backbones of the presentation of Chapter III and are to be treated in the next chapter.

CHAPTER II

FUNDAMENTAL MATHEMATICAL BACKGROUND

There are certain fundamental mathematical concepts and definitions which are used extensively in the next chapter, in particular, and throughout the rest of this work in general. These fundamental theories are reviewed in this chapter mostly without proof. In each case an appropriate reference will be given for interested readers.

The Simplest Problems in the Calculus of Variations

Definition: The problems concerned with the determination of extreme values of integrals whose integrands contain unknown functions belong to the calculus of variations.

The simplest of such problems involves the determination of the unknown function $y = y(x)$ for which the integral

$$I = \int_{x_0}^{x_1} F(x,y,y')\,dx \tag{2-1}$$

between two fixed points P_0 (x_0,y_0) and P_1 (x_1,y_1) is a minimum. Here, it is assumed that F as a function of its arguments, x, y, y', is known and has partial derivatives of the second order. It is also assumed that there exists a curve $y = y(x)$ with continuously turning tangent that minimizes the integral.

Let $h(x)$ be any function with continuous second derivatives and

$$h(x_0) = h(x_1) = 0 \tag{2-2}$$

Then the function

$$\bar{y}(x) = y(x) + \alpha h(x) \tag{2-3}$$

for α, a small parameter, represents a family of curves passing through P_0 and P_1, since $y = y(x)$ passes through these points and $h(x_0) = h(x_1) = 0$. This situation is illustrated in Figure 2-1.

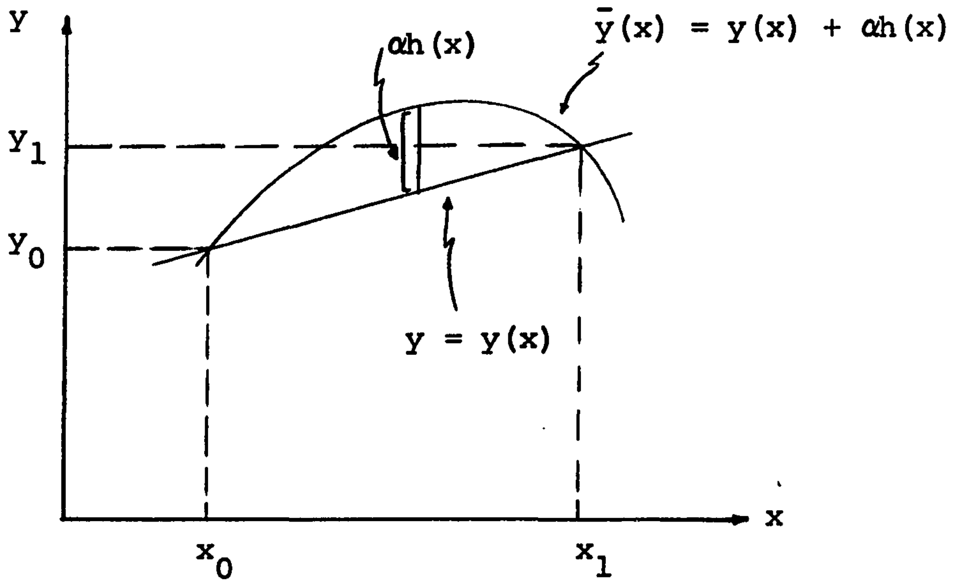

Figure 2-1. An Illustration of the Simplest Problems in the Calculus of Variations

Here the deviation from the minimizing curve, $y = y(x)$, is indicated by $\alpha h(x)$.

Upon substituting (2-3) into (2-1) there results

$$I(\alpha) = \int_{x_0}^{x_1} F[x, y(x) + \alpha h(x), y'(x) + \alpha h'(x)]dx \qquad (2\text{-}4)$$

For $\alpha = 0$, (2-3) gives $\bar{y}(x) = y(x)$ and since $y(x)$ minimizes the integral, one concludes that $I(\alpha)$ must have minimum for $\alpha = 0$. A necessary condition is therefore

$$\left.\frac{dI}{d\alpha}\right|_{\alpha=0} = 0 \qquad (2\text{-}5)$$

This same argument will be used in the presentation of Chapter III.

The Fundamental Lemma of the Calculus of Variations

If a function $\varphi(x)$ is continuous in an interval (x_0, x_1) and if

$$\int_{x_0}^{x_1} \varphi(x)\ h(x)dx = 0 \qquad (2\text{-}6)$$

for an arbitrary function $h(x)$ (subject to some condition of general character only), then $\varphi(x) \equiv 0$ throughout the interval $x_0 \leq x \leq x_1$. The general conditions that $h(x)$ must satisfy are, for instance, that $h(x)$ should have first or higher order derivatives, should vanish at the end points (x_0, x_1), and $|h(x)| < \epsilon$ or both $|h(x)| < \epsilon$ and $|h'(x)| < \epsilon$, where ϵ is an arbitrary positive number [S.4].

Some Fundamental Theorems and Definitions Related to the Functions of a Complex Variable

Definition: Let $f(z)$ be analytic throughout the domain D except for an isolated singularity at a certain finite point z_0 of D, then the integral

$$\frac{1}{2\pi j} \int_C f(z)dz \tag{2-7}$$

will have the same value on all curves C on D which enclose z_0 and no other singularity of f. This value is known as the residue of $f(z)$ at z_0, and is denoted by Res $[f(z_0), z_0]$.

Cauchy's Residue Theorem

If $f(z)$ is analytic in a domain D and C is a simple closed curve in D within which $f(z)$ is analytic except for isolated singularities at $z_1, \ldots z_k$, then

$$\frac{1}{2\pi j} \int_C f(z)dz = \text{Res}\,[f(z), z_1] + \ldots\ldots + \text{Res}\,[f(z), z_k] \tag{2-8}$$

For a proof of this theorem see reference [K.2].

Theorem. The residue of $f(z)$ at a finite point z_0 is given by the equation

$$\text{Res}\,[f(z), z_0] = a_{-1}, \tag{2-9}$$

where

$$f(z) = \ldots + \frac{a_{-N}}{(z-z_0)^N} + \ldots\ldots + \frac{a_{-1}}{(z-z_0)} + a_0 + a_1(z-z_0) + \ldots \tag{2-10}$$

is the Laurent expansion of $f(z)$ at z_0. For proof see reference [K.2].

The following rules are useful in cases where one is interested in the terms of $(z - z_0)^{-1}$ form in the expansion.

<u>Rule I.</u> At a simple pole z_0 (i.e., a pole of first order),

$$\text{Res}\,[f(z),\, z_0] = \lim_{z \to z_0} (z-z_0) f(z) \tag{2-11}$$

<u>Rule II.</u> At a pole z_0 of order N, (N = 2, 3 ...)

$$\text{Res}\,[f(z),\, z_0] = \lim_{z-z_0} \frac{g^{(N-1)}(z)}{(N-1)!} \tag{2-12}$$

where $g(z) = (z - z_0)^N f(z)$.

<u>Theorem.</u> Let $f(z)$ be analytic in a domain D which includes the real axis and all of the half-plane $y > 0$ except for a finite number of points. If

$$\lim_{R\to\infty} \int_0^{\pi} f(Re^{j\theta})\, Re^{j\theta} d\theta = 0$$

and

$$\int_{-\infty}^{\infty} f(x)\,dx$$

exists, then

$$\int_{-\infty}^{\infty} f(x)\,dx = 2\pi j\{\text{Sum of residues of } f(z) \text{ in the upper half-plane}\} \qquad (2\text{-}13)$$

Note that $z = x + iy = e^{i\theta}$.

Parseval's Theorem

If $x_1(t)$ and $x_2(t)$ are two Fourier transformable functions and it is desired to find the integral of the product of these two functions in the time domain, from $-\infty$ to $+\infty$ directly in terms of this same integral in the frequency domain, it can be shown [N.2]

$$\int_{-\infty}^{\infty} x_1(t)x_2(t)\,dt = \frac{1}{2\pi j}\int_{-j\infty}^{+j\infty} x_1(-s)x_2(s)\,ds \qquad (2\text{-}14)$$

In particular if $x_1(t) \equiv x_2(t)$ then

$$\int_{-\infty}^{\infty} x^2(t)\,dt = \frac{1}{2\pi j}\int_{-j\infty}^{+j\infty} x(-s)x(s)\,ds \qquad (2\text{-}15)$$

This result is known as Parseval's Theorem. This theorem represents a very convenient way of expressing the integral-square value of a time function in terms of its transform.

Some Well-Known Relations in Concise Forms

Superposition Integral

Let $c_{ij}(s)$, $w_{ij}(s)$, and $r_j(s)$ be the Fourier transforms of $c_{ij}(t)$, $w_{ij}(t)$ and $r_j(t)$ respectively, then Equation (1-1) in frequency domain may be written as

$$c_{ij}(s) = w_{ij}(s)r_j(s) \tag{2-16}$$

and similarly Equation (1-2) may be written

$$c_i(s) = \sum_{j=1}^{p} w_{ij}(s)r_j(s) \tag{2-17}$$

In matrix notation this may be written as

$$C(s) = W(s)\ R(s) \tag{2-18}$$

where C(s) is a q-dimensional column vector with $c_i(s)$ as its ith component. W(s) is a q x p matrix with general components $w_{ij}(s)$ and R(s) is a p-dimensional column vector with $r_j(s)$ as its jth component.

<u>Autocorrelation and Cross-Correlation</u>

Let $r_i(t)$ be a representative member of a stochastic ensemble. The autocorrelation function is defined as the ensemble average of the product of $r_1(t)$ and $r_1(t+\tau)$. Symbolically this is expressed as

$$\varphi_{rr}^{ii}(t,\tau) \triangleq \widetilde{r_i(t)r_i(t+\tau)} \tag{2-19}$$

where the symbols are used in the same sense as in reference [N.2].

When the stochastic signal is stationary; i.e., when the statistical properties of the signal are time independent then the ensemble average may be interchanged with the time average. That is

$$\varphi_{rr}^{ii}(\tau) = \overline{r_i(t)r_i(t+\tau)} \tag{2-20}$$

or equivalently,

$$\varphi_{rr}^{ii}(\tau) = \int_{-\infty}^{\infty} r_i(t)r_i(t+\tau)dt \tag{2-21}$$

In this same sense cross-correlation functions are defined as

$$\varphi_{rd}^{ij}(\tau) = \int_{-\infty}^{\infty} r_i(t)d_j(t+\tau)dt \tag{2-22}$$

Taking the Fourier transform of Equation (2-21) one gets

$$\Phi_{rr}^{ii}(s) = \int_{-\infty}^{\infty} \varphi_{rr}^{ii}(\tau)\ e^{-s\tau}\ d\tau$$

$$= \int_{-\infty}^{\infty} dt \int_{-\infty}^{\infty} r_i(t)\ r_i(t+\tau)\ e^{-s\tau}\ d\tau$$

$$= \int_{-\infty}^{\infty} dt\ r_i(t)\ e^{st} \int_{-\infty}^{\infty} r_i(t+\tau)\ e^{-s(t+\tau)}\ d\tau$$

$$= r_i(-s)\ r_i(s)$$

Therefore:

$$\Phi_{rr}^{ii}(s) = r_i(-s)\ r_i(s) \tag{2-23}$$

In matrix notation this may be written as

$$\Phi_{rr}(s) = R(-s)R^T(s) \tag{2-24}$$

since R is a p-dimensional column vector or a p x 1 matrix.

Following the same argument the Fourier transform of Equation (2-22), in matrix form, may be written as

$$\Phi_{rd}(s) = R(-s)D^T(s) \tag{2-25}$$

where, as usual, $D^T(s)$ is the transpose of D(s).

Note that the autocorrelation is an even function

$$\Phi_{rr}(s) = \Phi_{rr}(-s) \tag{2-26}$$

but

$$\Phi_{rd}(s) = \Phi_{dr}^T(-s) \tag{2-27}$$

since cross-correlation is an odd function.

In concluding this chapter the following facts concerning the fourier transform are worthy of mentioning. The direct Fourier transform of a time function which is zero for t < 0 will have all its poles in the left half plane. On the other hand a time function which is zero for $t > 0$ will have all its poles in the right half plane. A time function which is nonzero for all values of time will have a transfer function with poles on both sides of the imaginary axis. For more detail see Appendix A of reference [N.2].

CHAPTER III

PRESENTATION OF THE METHOD OF OPTIMUM DESIGN-BY SYNTHESIS

With the ideas of the preceding chapters in mind it is now possible to give a more precise statement of the problem and the proposed method of approach to its solution.

Assume that a set of random stationary inputs to a system and a set of desired outputs for this system are given. It is required to synthesize this system so that the sum of the weighted mean-square of the differences between the set of actual outputs and the desired outputs is minimized. There is no limit on the choice of the elements of this system other than physical realizability. Therefore this optimization problem is of "free configuration" type. The system considered here is shown in Figure 3-1. Solid lines are used to indicate the multiplicity of inputs and outputs. The input, $r(t)$, is a p-dimensional column vector with its components, $r_i(t)$, being members of stationary stochastic processes. The output, $c(t)$, is a q-dimensional column vector and $w(t)$ is a $q \times p$ weighting function matrix. As shown in Figure 3-1 the error is the difference between the desired output vector, $d(t)$, and

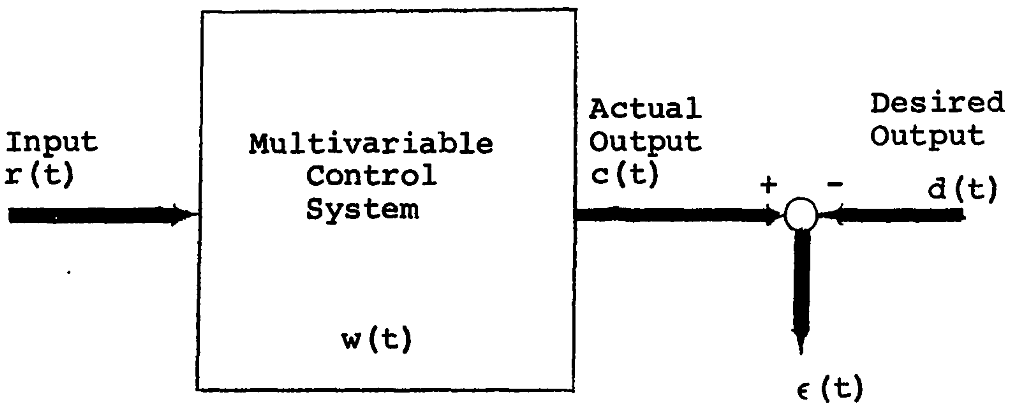

Figure 3-1. Multivariable Control System, Free Configuration

the actual output vector c(t). The objective is to minimize the sum of the weighted mean-square value of the error by properly choosing the elements of the weighting function matrix.

The proposed method of solution of this problem has two basic parts. First, an implicit solution is derived in the form of the vector Wiener-Hopf equation. This equation is equivalent to the Euler's equation in the calculus of variations. Therefore the solution to this equation is the explicit description of the minimizing system in analogy with the solution to Euler's equation which is the minimizing curve. The derivation of the vector Wiener-Hopf equation has been achieved in two different ways. In one method the derivation is performed strictly in the time domain. This method is essentially the extension of the ideas developed in reference [N.2] concerning the synthesis of optimum linear constant parameter systems. The alternative method is used to show the advantage of working strictly in the frequency domain. This approach is new and has not appeared in the literature yet.

The second part of the problem deals with the methods of solution of a vector Wiener-Hopf equation. The method developed in this work is a combination of the "functional method" presented in reference [B.4] and the "method of undetermined coefficients" presented in reference [A.2].

Derivation of the Wiener-Hopf Equation in Vector Form

In this section the vector Wiener-Hopf equation is derived using two different methods of analysis. The solutions to the vector equation are the elements of the optimum weighting matrix. The steps followed in deriving this equation are as follows: (1) the performance functional is defined; (2) this functional is expressed in terms of the weighting function matrix, the inputs, and the ideal outputs; (3) standard variational procedure is employed and the condition for minimization of the performance functional is developed into the vector Wiener-Hopf equation.

The Performance Functional

Referring to Figure 3-1, the error is by definition

$$\epsilon(t) \triangleq c(t) - d(t) \tag{3-1}$$

The general error matrix is given by the autocorrelation matrix, $\varphi_{\epsilon\epsilon}(\tau)$ [N.1]. For reasons given previously it is usual to consider a sum of the weighted individual errors. This implies that the main terms of interest in the matrix $\varphi_{\epsilon\epsilon}(\tau)$ are the diagonal terms. Therefore the performance functional P is defined as

$$P \triangleq \text{tr. } [L\, \varphi_{\epsilon\epsilon}(\tau)] \tag{3-2}$$

where L is a diagonal matrix which weights the diagonal elements of $\varphi_{\epsilon\epsilon}(\tau)$. For a particular case when the performance criterion is the sum of the weighted mean-square of the individual errors, P may be written as

$$P = \text{tr. } [L\, \varphi_{\epsilon\epsilon}(0)] \tag{3-3}$$

since the autocorrelation evaluated at $\tau = 0$ is precisely the mean-square error. Equation (3-3) is essentially equivalent to

$$P = \overline{\epsilon^T(t)\; L\, \epsilon(t)} \tag{3-4}$$

where $\epsilon(t)$ is q x 1 matrix and L is a q x q diagonal matrix. Note that P is a scalar and $P = P^T$ since L is diagonal.

<u>Minimization Method I</u>

Using the definition of the $\epsilon(t)$ given by Equation (3-1) in Equation (3-4) yields

$$P = \overline{c^T(t)\ L\ c(t)} - [\overline{d^T(t)\ L\ c(t)} + \overline{c^T(t)\ L\ d(t)}]$$

$$+ \overline{d^T(t)\ L\ d(t)} \qquad (3\text{-}5)$$

But note that the input-output relation, as expressed in Equation (1-2), may be written in matrix notation as follows

$$c(t) = w(\tau) * r(t-\tau) \qquad (3\text{-}6)$$

where * means the convolution of appropriate components of w(t) matrix and r(t) matrix. Knowing that

$$c^T(t) = r^T(t-\tau) * w^T(\tau) \qquad (3\text{-}7)$$

the first term on the right of Equation (3-5) may be written as

$$\overline{c^T(t)\ L\ c(t)} = \overline{r^T(t-\tau) * w^T(\tau)\ L\ w(\sigma) * r(t-\sigma)} \qquad (3\text{-}8)$$

Similarly

$$\overline{c^T(t)\ L\ d(t)} = \overline{r^T(t-\tau) * w^T(\tau)\ L\ d(t)} \qquad (3\text{-}9)$$

Note also that

$$\overline{c^T(t)\ L\ d(t)} \equiv \overline{d^T(t)\ L\ c(t)} \qquad (3\text{-}10)$$

since L is diagonal and each element on the right of Equation (3-5) is a scalar.

Upon substituting Equations (3-8), (3-9), and (3-10) into Equation (3-5) P becomes

$$P = \overline{r^T(t-\tau) * w^T(\tau)\ L\ w(\sigma) * r(t-\sigma)}$$
$$- 2\ \overline{r^T(t-\tau) * w^T(\tau)\ L\ d(t)} + \overline{d^T(t)\ L\ d(t)} \qquad (3\text{-}11)$$

This equation given us the sum of the weighted mean-square of the individual errors in terms of the weighting functions, inputs, and ideal outputs.

The next step is to determine the system functions that minimize this functional. This is the same type of problem as the simplest problem in the calculus of variations which was treated in the early part of Chapter II. Therefore, the same method of solution is followed. To do this, it is assumed that such solutions do exist and is denoted by a q x p matrix $w_m(t)$. Corresponding to Equation (2-3) we construct

$$w(t) = w_m(t) + \alpha h(t) \qquad (3\text{-}12)$$

Here $w_m(t)$ is the assumed solution, h(t) is any physically realizable weighting function matrix and α is a parameter which may vary to test whether $w_m(t)$ is the solution. Substitution of Equation (3-12) into the right member of Equation (3-11) makes the performance functional P a function of the parameter α. By setting the derivative of P with respect to α evaluated at $\alpha = 0$ equal to zero; i.e.

$$\frac{d}{d\alpha} P(\alpha) \Big|_{\alpha=0} = 0 \quad , \qquad (3\text{-}13)$$

the necessary condition for $w_m(t)$ which must be satisfied, in order for it to be the solution, is arrived.

Substituting Equation (3-12) into the right hand side of Equation (3-11) and performing the operation suggested by Equation (3-13) gives

$$\overline{r^T(t-\tau) * w_m^T(\tau) \; L \; h(\sigma) * r(t-\sigma)}$$

$$+ \overline{r^T(t-\tau) * h^T(\tau) \; L \; w_m(\sigma) * r(t-\sigma)}$$

$$- 2 \; \overline{r^T(t-\tau) * h^T(\tau) \; L \; d(t)} = 0 \qquad (3\text{-}14)$$

In tensor notation Equation (3-14) may be written as

$$\overline{r_{jl}(t-\tau) * w_{kj}(\tau) \; l_{kk} \; h_{ki}(\sigma) * r_{il}(t-\sigma)}$$

$$+ \overline{r_{il}(t-\tau) * h_{ki}(\tau) \; l_{kk} \; w_{kj}(\sigma) \; r_{jl}(t-\sigma)}$$

$$- 2 \; \overline{r_{il}(t-\tau) * h_{ki}(\tau) \; l_{kk} \; d_{kl}(t)} = 0 \qquad (3\text{-}15)$$

Here it is understood that the repeated indices mean summation. Also the index m on w(t) was dropped to avoid confusion with dummy indices.*

---

*For further detail see Appendix A.

Recalling the definitions of autocorrelation and the cross-correlation functions as given by Equation (2-20), (2-21), and (2-22), Equation (3-15) may be written as

$$l_{kk}\left[\int_{-\infty}^{\infty} d\tau w_{kj}(\tau) \int_{-\infty}^{\infty} h_{ki}(\sigma)\ \varphi_{rr}^{ji}(\tau-\sigma)\, d\sigma\right.$$

$$+ \int_{-\infty}^{\infty} d\tau\ h_{ki}(\tau) \int_{-\infty}^{\infty} w_{kj}(\sigma)\ \varphi_{rr}^{ji}(\tau-\sigma)\, d\sigma$$

$$\left. - 2 \int_{-\infty}^{\infty} h_{ki}(\tau)\ \varphi_{rd}^{ik}(\tau)\, d\tau\right] = 0 \qquad (3\text{-}16)$$

Because of the even property of autocorrelation function of stationary signals, we know that

$$\varphi_{rr}(\tau-\sigma) = \varphi_{rr}(\sigma-\tau) \qquad (3\text{-}17)$$

This equation implies that

$$\int_{-\infty}^{\infty} d\tau\ w_{kj}(\tau) \int_{-\infty}^{\infty} h_{ki}(\sigma)\ \varphi_{rr}^{ji}(\tau-\sigma)\, d\sigma$$

$$= \int_{-\infty}^{\infty} d\tau\ h_{ki}(\tau) \int_{-\infty}^{\infty} w_{kj}(\sigma)\ \varphi_{rr}^{ji}(\tau-\sigma)\, d\sigma \qquad (3\text{-}18)$$

since we can change the order of integration and interchange the variables of integration with the result that the only difference between the two sides will be in the sign of the autocorrelation variable. Using Equation (3-18) in Equation (3-16) yields

$$2\ l_{kk} \left[\int_{-\infty}^{\infty} d\tau\ h_{ki}(\tau)\ \{w_{kj}(\sigma)\ *\ \varphi_{rr}^{ji}(\tau-\sigma)\ -\ \varphi_{rd}^{ik}(\tau)\} \right] = 0 \tag{3-19}$$

Now the elements of the h(t) matrix are physically realizable weighting functions but other than that are perfectly arbitrary. Therefore the only way that Equation (3-19) can be satisfied for values of τ equal or greater than zero is for the factor inside the inner bracket to be equal to zero. Thus we are led to the condition

$$w_{kj}(\sigma)\ *\ \varphi_{rr}^{ji}(t-\sigma)\ -\ \varphi_{rd}^{ik}(t)\ =\ 0 \tag{3-20}$$

for

$$t \geq 0$$

$$\text{all} \quad \begin{matrix} k = 1,\ 2,\ \ldots\ldots\ q \\ i = 1,\ 2,\ \ldots\ldots\ p \end{matrix}$$

In matrix notation, this equation may be written as

$$\varphi_{rr}(t-\sigma)\ *\ w_m^T(\sigma)\ -\ \varphi_{rd}(t)\ =\ 0 \quad \text{for } t \geq 0 \tag{3-21}$$

The integral equation of the form of Equation (3-21) is known as the vector form of the Wiener-Hopf equation [B.4]. The system weighting function matrix, $w_m(t)$, that minimizes the sum of the weighted mean-square of the individual errors must satisfy this condition.

Strictly speaking, the condition expressed by the Equation (3-21) is a necessary but not sufficient to ensure a minimum of P. To show that the solution obtained from Equation (3-21) yields a minimum of P, the second derivative of P with respect to α is examined. Substitute Equation (3-12) into Equation (3-11) to obtain

$$P = \overline{r^T(t-\tau) * [w_m^T(\tau) + \alpha h^T(\tau)]\ L\ [w_m(\sigma) + \alpha h(\sigma)] * r(t-\sigma)}$$

$$- 2\ \overline{r^T(t-\tau) * [w_m^T(\tau) + \alpha h^T(\tau)]\ L\ d(t)} + \overline{d^T(t)\ L\ d(t)}$$

(3-22)

The first derivative of Equation (3-22) with respect to α is

$$\frac{dP}{d\alpha} = \overline{r^T(t-\tau) * h^T(\tau)\ L\ [w_m(\sigma) + \alpha h(\sigma)] * r(t-\sigma)}$$

$$+ \overline{r^T(t-\tau) + [w_m^T(\tau) + \alpha h^T(\tau)]\ L\ h(\sigma) * r(t-\sigma)}$$

$$- 2\ \overline{r^T(t-\tau) * h^T(\tau)} \tag{3-23}$$

The second derivative of P with respect to α is given by

$$\frac{d^2P}{d\alpha^2} = 2\ \overline{r^T(t-\tau) * h^T(\tau)\ L\ h(\sigma) * r(t-\sigma)} \tag{3-24}$$

But according to Equation (3-8) the right hand side of Equation (3-24) is proportional to $\overline{C^T(t)\ L\ C(t)}$. Therefore

$$\frac{d^2P}{d\alpha^2} = 2\ \overline{C^T(t)\ L\ C(t)} \tag{3-25}$$

where C(t) is the output of the system shown in Figure 3-1 with w(t) being replaced by h(t). The physical interpretation of Equation (3-25) is that the second derivative of P is nothing more than a sum of the weighted mean-square values of the set of filtered input functions obtained by passing this input set through a system having a weighting functions matrix h(t). Since this summation is always greater than zero, it is shown that the second derivative of P with respect to α is always positive. However, in strict mathematical terms, the necessary and sufficient conditions for a functional of the form*

$$J[y] = \int_a^b F(x, y, y')dx, \qquad y(a) = A, \quad y(b) = B \tag{3-26}$$

to have a strong extremum for the curve $y = y(x)$ are

1) y(x) should satisfy Euler's equation
2) the strengthened Legendre condition
3) the strengthened Jacobi condition
4) the Weierstrass condition

But on the other hand, P, the functional in the above argument, is not of the type expressed by Equation (3-26)

---

*See reference [G.1] for more details.

(for example P involves double integrations and, contrary to F in Equation (3-26), does not have derivative of the minimizing curve as the argument of its integrand). Therefore it is quite plausible to consider conditions expressed by Equations (3-21) and (3-25) as necessary and sufficient conditions that the solutions obtained from Equation (3-21) do yield a minimum.

With Equation (3-21) satisfied one can proceed to find an expression for the minimum P. To do this Equation (3-11) may be written in tensor notation; i.e.

$$P = 1_{ii} \int_{-\infty}^{\infty} d\tau \; w_{ji}(\tau) \; \{w_{ij}(\sigma) * \varphi_{rr}^{kj}(\tau-\sigma)\}$$

$$- 2 \int_{-\infty}^{\infty} \varphi_{rd}^{ji}(\tau) \; w_{ij}(\tau) d\tau + \varphi_{dd}^{ii}(0)$$

Substituting the result of Equation (3-21) into this equation yields

$$P_{min.} = 1_{ii} \; [\varphi_{dd}^{ii}(0) - \int_{-\infty}^{\infty} \varphi_{rd}^{ji} \; (\tau) \; w_{ij}(\tau) d\tau] \qquad (3\text{-}27)$$

In matrix notation Equation (3-27) may be written as

$$P_{min.} = \text{tr.} \; [L \; \varphi_{dd}(0) - \int_{-\infty}^{\infty} L \; w_m(\tau) \; \varphi_{rd}(\tau) d\tau] \qquad (3\text{-}28)$$

where tr. [...] denotes the trace of matrix.

The derivation of the vector Wiener-Hopf equation is therefore complete. As was pointed out earlier in this chapter, the development of this method was based on the extension of the procedures presented by reference [N.2] for the case of single-input single-output systems. The derivation was performed strictly in the time domain. In the next section it is shown that transformation to the frequency domain at the early stages of the derivation helps to simplify the operations involved considerably.

Minimization Method II

The general error matrix, $\varphi_{\epsilon\epsilon}(\tau)$, is by definition

$$\varphi_{\epsilon\epsilon}(\tau) \triangleq \overline{\epsilon(t)\ \epsilon^{T}(t+\tau)} \qquad (3\text{-}29)$$

but

$$\epsilon(t) = c(t) - d(t)$$

Therefore

$$\varphi_{\epsilon\epsilon}(\tau) = \varphi_{dd}(\tau) - \varphi_{dc}(\tau) - \varphi_{cd}(\tau) + \varphi_{cc}(\tau) \qquad (3\text{-}30)$$

Denote G(s) with proper subscripts as the spectral density of the function represented by the subscript. For example let $G_{\epsilon\epsilon}(s)$ be the spectral density of $\epsilon(t)$ and $G_{rr}(s)$ the spectral density for r(t), etc. Thus corresponding to Equation (3-30) one has

$$G_{\epsilon\epsilon}(s) = G_{dd}(s) - G_{dc}(s) - G_{cd}(s) + G_{cc}(s) \qquad (3\text{-}31)$$

By definition

$$G_{\epsilon\epsilon}(s) = \int_{-\infty}^{\infty} \varphi_{\epsilon\epsilon}(\tau) e^{-s\tau} d\tau \qquad (3\text{-}32)$$

Similarly

$$G_{dc}(s) = \int_{-\infty}^{\infty} \varphi_{dc}(\tau) e^{-s\tau} d\tau \qquad (3\text{-}33)$$

But

$$\varphi_{dc}(\tau) = \overline{d(t) c^T(t+\tau)} = \int_{-\infty}^{\infty} w(\tau_1)\ \varphi_{dr}^T(\tau-\tau_1) d\tau_1 \qquad (3\text{-}34)$$

since

$$c(t) = \int_{-\infty}^{\infty} w(\tau_1)\ r(t-\tau_1) d\tau_1 \qquad (3\text{-}35)$$

Here no physical realizability condition is imposed upon $w(t)$. Upon taking the Fourier transform of Equation (3-34), using Equation (3-33) there results

$$G_{dc}(s) = \int_{-\infty}^{\infty} d\tau\ e^{-s\tau} \int_{-\infty}^{\infty} w(\tau_1)\ \varphi_{dr}^T(\tau-\tau_1) d\tau_1$$

$$= \int_{-\infty}^{\infty} d\tau_1\ w(\tau_1) \int_{-\infty}^{\infty} \varphi_{dr}^T(\tau-\tau_1)\ e^{-s(\tau-\tau_1)}\ e^{-s\tau_1}\ d\tau$$

$$\int_{-\infty}^{\infty} w(\tau_1)\ e^{-s\tau_1}\ d\tau_1 \int_{-\infty}^{\infty} \varphi_{dr}^{T}(t)\ e^{-st}\ dt$$

$$= W(s)\ G_{dr}^{T}(s) \tag{3-36}$$

Following the same procedure it can be shown that

$$G_{cd}(s) = \bar{W}(s)\ G_{rd}(s) \tag{3-37}$$

and

$$G_{cc}(s) = \bar{W}(s)\ G_{rr}(s)\ W^{T}(s) \tag{3-38}$$

where $\bar{W}(s) = W(-s)$.

Substituting Equations (3-36) through (3-38) into Equation (3-31) yields

$$G_{\epsilon\epsilon}(s) = G_{dd}(s) - W(s)\ G_{dr}^{T}(s) - \bar{W}(s)\ G_{rd}(s)$$

$$+ \bar{W}(s)\ G_{rr}(s)\ W^{T}(s) \tag{3-39}$$

Using the result of Parseval's Theorem it can be shown that

$$P = \text{tr.}\ [L\ \varphi_{\epsilon\epsilon}(0)] = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \text{tr.}\ [L\ G_{\epsilon\epsilon}(s)]ds \tag{3-40}$$

Now at $\tau = 0$

$$\text{tr.}\ [W(s)\ G_{dr}^{T}(s)] = \text{tr.}\ [\bar{W}(s)\ G_{rd}(s)] \tag{3-41}$$

since $W(s)\ G_{dr}^{T}(s) = G_{dc}(s)$ and $\bar{W}(s)\ G_{rd}(s) = G_{cd}(s)$ are the spectral densities of

$$\varphi_{dc}(\tau) = \int_{-\infty}^{\infty} d(t)\ c(t+\tau)dt$$

and

$$\varphi_{cd}(\tau) = \int_{-\infty}^{\infty} c(t)\ d(t+\tau)dt$$

and at $\tau = 0$ one notes that $\varphi_{dc}(0) \equiv \varphi_{cd}(0)$. With the introduction of Equations (3-31) and (3-41) into Equation (3-40) one may write

$$P = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr. } [L\ G_{dd}(s) - 2L\ \bar{W}(s)\ G_{rd}(s)$$

$$+ L\ \bar{W}(s)\ G_{rr}(s)\ W^{T}(s)]ds \tag{3-42}$$

The next step is to determine the system transfer matrix, $W(s)$, which minimizes P. Here again standard variational techniques, presented in Chapter II and used in the preceding section of this chapter, are followed. Construct a transfer matrix in the variational form

$$W(s) = W_{m}(s) + \alpha H(s) \tag{3-43}$$

Substituting Equation (3-43) in Equation (3-42), differentiating with respect to α, and setting the result, evaluated at $\alpha = 0$, equal to zero yields*

$$\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr.} \; [L \; \bar{H}(s) \; G_{rr}(s) \; W_m^T(s)$$

$$+ L \; \bar{W}_m(s) \; G_{rr}(s) \; H^T(s) - 2L \; \bar{H}(s) \; G_{rd}(s)]ds = 0 \tag{3-44}$$

Note that the trace of each of the matrix terms in the bracket under the integral is a scalar. Interchanging $W_m(s)$ with $H(s)$ in the second term in the bracket corresponds to interchanging the order of integration and the labeling of the integration variables in time domain with the net result being the sign change of the autocorrelation variable and corresponding change of $G_{rr}(s)$ to $G_{rr}(-s)$. It implies that Equation (3-38) may be written as

$$\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr.} \; [2L \; \bar{H}(s) \; \{G_{rr}(s) \; W_m^T(s) - G_{rd}(s)\}]ds = 0 \tag{3-45}$$

since $\bar{G}_{rr}(s) = G_{rr}(s)$. Now $H(s)$ is an arbitrary matrix and L a diagonal matrix with non-zero diagonal elements proving the only way in which Equation (3-45) can be satisfied is for the factor in the inner bracket to be equal to zero. Thus

$$G_{rr}(s) \; W_m^T(s) - G_{rd}(s) = 0 \tag{3-46}$$

---

*See Appendix B for details.

Up to this point no physical realizability condition has been imposed on H(s). If it is required that H(s) to satisfy this condition then Equation (3-46) becomes

$$\left[G_{rr}(s)\ W_m^T(s)\right]_+ = \left[G_{rd}(s)\right]_+ \qquad (3\text{-}47)$$

where $[.....]_+$ is used in the same sense as suggested in reference [N.2]. That is, $[G(s)]_+$ is defined as the components of G(s) which has all its poles in the left half-plane (LHP) and stems from the fact that the direct transform of a time function which is zero for negative time will have all its poles in the LHP. Note that Equation (3-47) is independent of matrix L and hence the resulting $W_m(t)$ is independent of the diagonal weighting matrix L.

Substituting Equation (3-46) into Equation (3-42) gives

$$P = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr.}\ [L\ G_{dd}(s) - L\ \bar{W}(s)\ G_{rd}(s)]ds \qquad (3\text{-}48)$$

where L, as defined before, is a diagonal weighting matrix. Note that Equations (3-47) and (3-48) are the counterparts of the Equations (3-21) and (3-28) in the frequency domain.

A Method of Solution of the Vector Wiener-Hopf Equation

As was pointed out early in this chapter the second part of the problem discussed herein deals with the solution of the vector Wiener-Hopf equation. The method presented in this section is a modified, combined form of the two recent methods which appeared in the literature, namely the method of "undetermined coefficients" presented by reference [A.2] and the "functional method" developed in reference [B.4]. However, before proceeding further along this line a change of notation is in order to simplify and ease manipulation of Equation (3-46). Let

$$\begin{aligned} A &\equiv G_{rr}(s) && \text{be a } p \times p \text{ matrix} \\ X &\equiv W^T(s) && \text{be a } p \times q \text{ matrix} \qquad (3\text{-}49) \\ C &\equiv G_{rd}(s) && \text{be a } p \times q \text{ matrix} \end{aligned}$$

In terms of these new notations Equation (3-41) may be written as

$$[AX]_+ = [C]_+ \tag{3-50}$$

Note that for each column of matrix C, say C_K, the following matrix relation holds.

$$[AX_K]_+ = [C_K]_+ \qquad k = 1, \ldots q \tag{3-51}$$

where X_K is the kth column of matrix X.

The objective is to solve for the components of the column vector X_K with the column vector C_K and matrix A given. Once this is done the whole procedure is repeated q times until all of the components of matrix X are found.

The first step in the solution of Equation (3-51) is to factor matrix A into two matrices P' and P such that $A = P'\ P$ and P is stable. This is always possible since A, being a spectral density matrix, is a rational, Hermitian matrix; i.e.,

$$A = \bar{A}^T$$

Different authors have developed methods of factorization of this type of matrix. The most recent ones are that of Kavanagh [K.32] and Youla [Y.1]. The method developed by Youla was further modified by Brockett and Mesarovic [B.4]. Their work has been summarized in Appendix C for easy reference and has been used to modify Equation (3-51) into the following form

$$[P'PX_K]_+ = [C_K]_+ \qquad K = 1, \ldots q \tag{3-52}$$

where

$$P' \equiv \bar{P}^T$$

Theorem 2 of reference [B.4] states that if P' contains no stable poles then

$$[P'D]_+ = [P'[D]_+]_+ \tag{3-53}$$

Now the result of this theorem is used to define

$$D_K \triangleq [P'^{-1}\ C_K]_+ \tag{3-54}$$

Inside the bracket Equation (3-52) is premultiplied by $P' \, P'^{-1}$ and then the definition for D_K is used so that

$$[P' \, PX_K]_+ = [P' \, D_K]_+ \qquad K = 1, \ldots q \tag{3-55}$$

Using Theorem 2 of reference [B.4] again, it can be shown that Equation (3-55) is equivalent to

$$[P' \, [P \, X_K]_+]_+ = [P' \, [D_K]_+]_+ \tag{3-56}$$

since P, X_K, and D_K are assumed to have stable poles only. On the other hand, P' has only unstable poles. Therefore if the operation represented by $[\quad]_+$ is peformed, P' can only modify the residues of the stable poles on each side of Equation (3-56). This implies that as far as stable poles are concerned one can write

$$[PX_K] = [D_K]_+ \tag{3-57}$$

or

$$[X_K]_+ = [P^{-1}[D_K]_+]_+$$

$$= [P^{-1} \, D_K]_+ \tag{3-58}$$

As a result of performing the operation suggested by Equation (3-58) all of the probable stable poles of the system are determined. The next step is to modify the residues of these poles by taking into account the effect of P' on each pole. At this point one can use the method of undetermined coefficients, as worked out in reference [A.2], to determine the modified partial fraction expansion coefficient of each stable pole.

The significance of the modified method presented in this section is that, contrary to the method of undetermined coefficient presented in reference [A.2], it is not necessary to find the "natural" poles by solving the determinant of $G_{rr}(s)$ and then add to it the "forced" poles if there are any. This is automatically taken care of in the process of determining the unknown coefficients of the probable poles. Vanishing of any of these coefficients means the cancellation of the corresponding poles and hence the modification of probable poles into actual poles of the system matrix. The method differs from the functional method, presented in reference [B.4], in the sense that it results in stable system directly and does not require that the poles of each member of the system matrix be the same as those of all other elements as required by the method of undetermined coefficients presented in reference [A.2]. These points will be clarified in the next chapter.

Summary of the Derivation

The procedure for solving an equation of the form represented by Equation (3-50) is summarized as follows:

Step 1. Factorize A into P' and P using the steps suggested in Appendix C.

Step 2. Use Equation (3-54) to find column vectors D_K.

Step 3. Perform the operation suggested by Equation (3-58) and determine the system's probable stable poles. It is not necessary to find the coefficients of the expansion at this point.

Step 4. To each pole assign an unknown coefficient a_{ij} where $i = 1, \ldots, p$, and $j = 1, \ldots, m$ (m being the number of stable poles). Now substitute the resulting solution for X_K into Equation (3-51) and determine the a_{ij} coefficients.

Step 5. Repeat steps 3 and 4 q times.

Step 6. Use Equation (3-48) to find the numerical value for $P_{min.}$.

A formalized block diagram of these steps is shown in Figure 3-2.

A Method of Generating Minimum P

In cases where elements of $w_m(t)$ are physically realizable functions, Equation (3-23) may be written as

$$P_{min.} = \sum_{i=1}^{q} l_{ii}\, \varphi_{dd}^{ii}(0) - \sum_{j=1}^{p} l_{ii} \int_0^t w_{ij}(t)\, \varphi_{rd}^{ji}(t)\,dt \tag{3-59}$$

Equation (3-59) is shown in block diagram form in Figure 3-3.

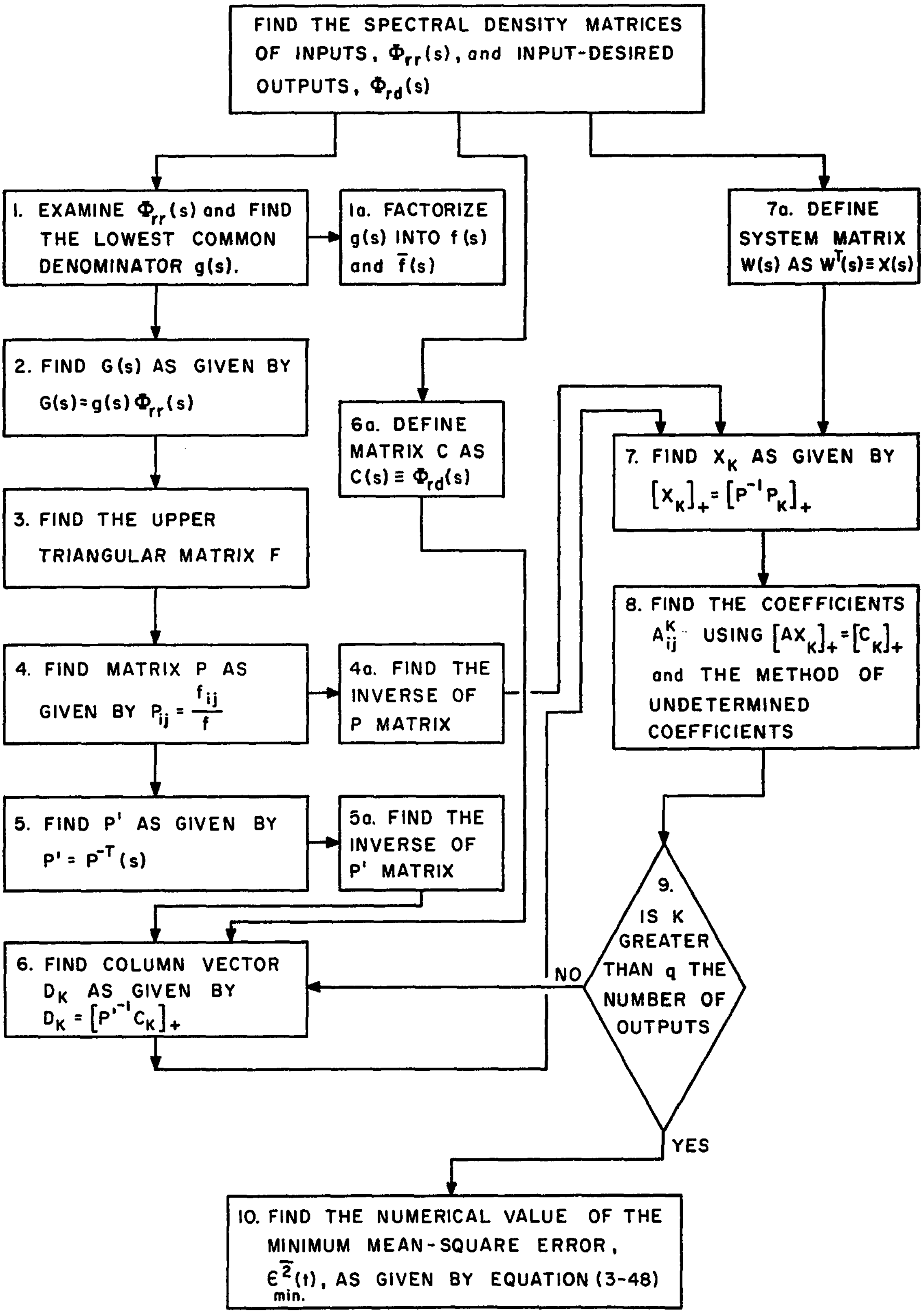

Figure 3-2. Formalized Block Diagram of the Proposed Method of Solution of Vector Wiener-Hopf Equation.

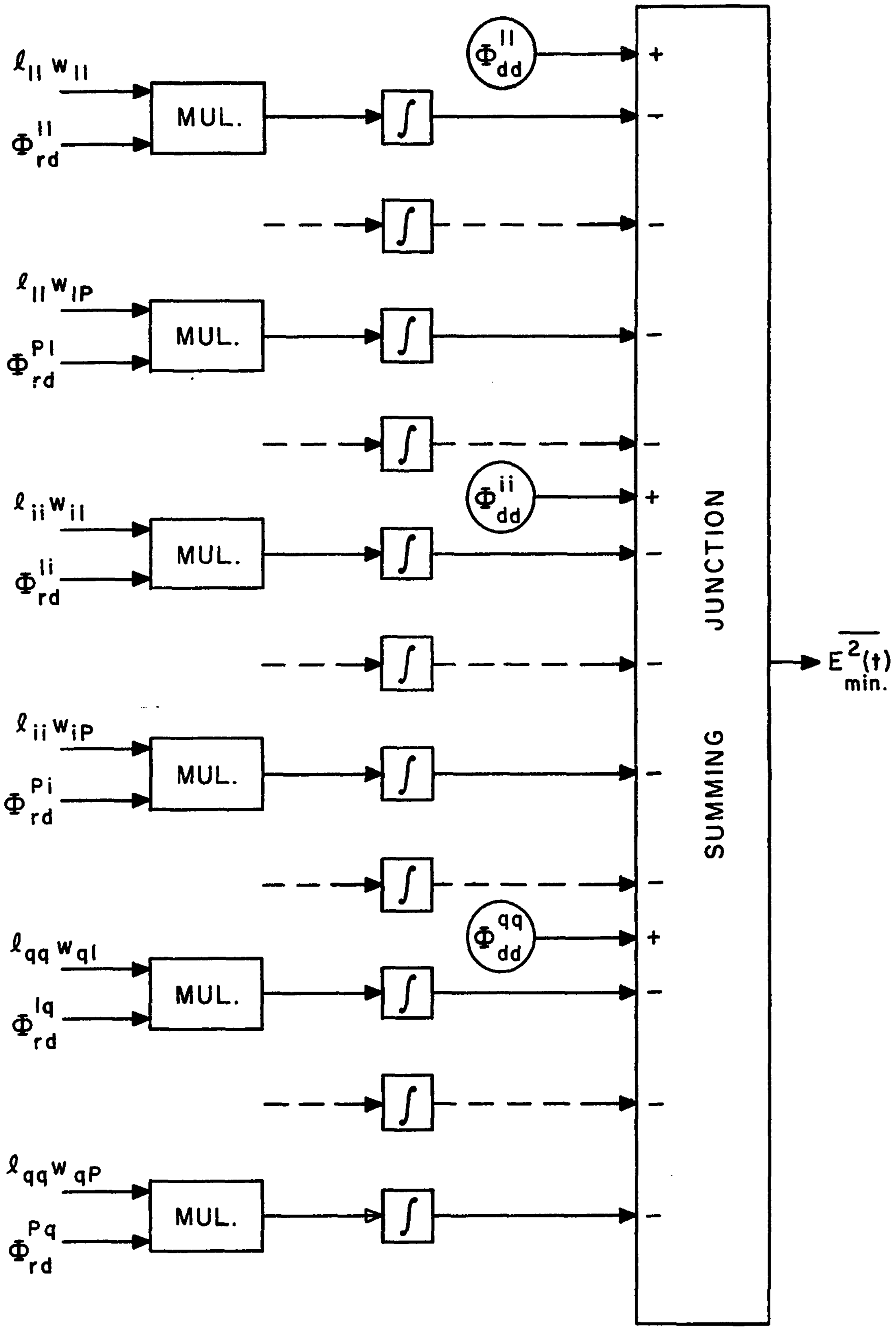

Figure 3-3. Block Diagram Representation of Equation (3-59).

CHAPTER IV

APPLICATIONS

In this chapter the results of the previous chapter are applied to two illustrative examples. Each example is solved by two different methods. First the functional method is used. A summary of the steps involved in this method is repeated in Appendix D for easy reference. Then the method suggested in Chapter III is applied. The first example solved is the one cited by Amara [A.2]. This solution is carried out to provide a base for comparison of the various methods. The second example is more comprehensive and is being carried out in detail for illustrative purposes. In each case an analog computer is used to simulate the resulting optimum system and generate the minimized mean-square error. Finally the results are tabulated for easy comparison.

Example 1 -

Given: The system under consideration is diagrammatically shown in Figure 4-1. Each of the inputs, r_i, is composed of message, m_i, corrupted with white noise, n_i. The noises are uncorrelated with each other and with each message, while the message at the second input is related to the message at the first input.

Required: It is desired to extract the best message m_1 (in a mean-square sense) by operating on the inputs r_1 and r_2 simultaneously.

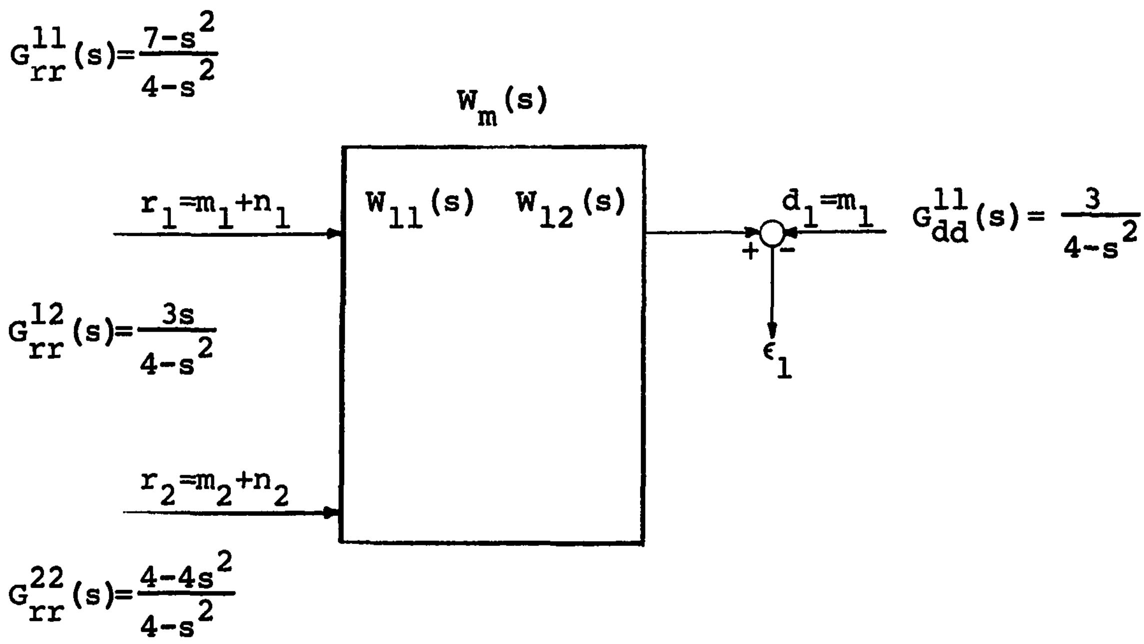

Figure 4-1. Diagrammatic Representation of Example 1

Solution: From the information given in the problem and shown on Figure 4-1, one obtains the following:

$$A = \begin{vmatrix} \frac{7-s^2}{4-s^2} & \frac{3s}{4-s^2} \\ \\ \frac{-3s}{4-s^2} & \frac{4-4s^2}{4-s^2} \end{vmatrix}$$

$$X_1 = \begin{vmatrix} x_{11} \\ \\ x_{21} \end{vmatrix} , \quad C_1 = \begin{vmatrix} \frac{3}{4-s^2} \\ \\ \frac{-3s}{4-s^2} \end{vmatrix}$$

The weighting matrix L is assumed to be the identity matrix. Corresponding to Equation (3-51),

$$[AX_1]_+ = [C_1]_+ \quad , \tag{3-51}$$

one has

$$\left[\begin{vmatrix} \frac{7-s^2}{4-s^2} & \frac{3s}{4-s^2} \\ \\ \frac{-3s}{4-s^2} & \frac{4-4s^2}{4-s^2} \end{vmatrix} \begin{vmatrix} x_{11} \\ \\ x_{21} \end{vmatrix}\right]_+ = \begin{bmatrix} \frac{3}{4-s^2} \\ \\ \frac{-3s}{4-s^2} \end{bmatrix}_+$$

<u>Modification Steps</u>

The first step in the solution of this problem is the factorization of matrix A into two matrices P and P' such that P is stable. To perform this operation the steps presented in Appendix C are followed.

Step 1. The lowest common denominator of matrix A is

$$g(s) = 4-s^2$$

Step 2. The newly defined matrix G then becomes

$$G = gA$$

$$= \begin{vmatrix} 7-s^2 & 3s \\ -3s & 4-4s^2 \end{vmatrix}$$

Step 3. Function g(s) may now be factored as

$$g(s) = (2-s)(2+s) = \bar{f}(s)f(s)$$

Step 4. The elements of the upper triangular matrix F then becomes

$$f_{11}(s) = (7-s^2)^- = \sqrt{7}-s$$

$$f_{12}(s) = \frac{3s}{(7-s^2)^+} = \frac{3s}{\sqrt{7}+s} \; ; \; f_{21}(s) = 0$$

$$f_{22}(s) = \left((4-4s^2) - \frac{3s}{(\sqrt{7}+s)} \cdot \frac{-3s}{(\sqrt{7}-s)}\right)^+$$

$$= \frac{(2+s)(\sqrt{7}+2s)}{(\sqrt{7}+s)}$$

Step 5. The matrix P(s) then becomes

$$P(s) = \frac{1}{f(s)} F(s)$$

$$= \begin{vmatrix} \dfrac{\sqrt{7}-s}{2+s} & \dfrac{3s}{(\sqrt{\ }+s)(2+s)} \\ \\ 0 & \dfrac{\sqrt{7}+2s}{\sqrt{7}+s} \end{vmatrix}$$

and the matrix $P'(s)$ is

$$P'(s) = \bar{P}^T(s)$$

$$= \begin{vmatrix} \dfrac{\sqrt{7}+s}{2-s} & 0 \\ \\ \dfrac{-3s}{(\sqrt{7}-s)(2-s)} & \dfrac{\sqrt{7}-2s}{\sqrt{7}-s} \end{vmatrix}$$

The inverses of P and P' are calculated to be respectively

$$P^{-1}(s) = \begin{vmatrix} \dfrac{2+s}{\sqrt{7}-s} & \dfrac{-3s}{(\sqrt{7}-s)(\sqrt{7}+2s)} \\ \\ 0 & \dfrac{\sqrt{7}+s}{\sqrt{7}+2s} \end{vmatrix}$$

$$P'^{-1}(s) = \begin{vmatrix} \dfrac{2-s}{\sqrt{7}+s} & 0 \\ & \\ \dfrac{3s}{(\sqrt{7}+s)(\sqrt{7}-2s)} & \dfrac{\sqrt{7}-s}{\sqrt{7}-2s} \end{vmatrix}$$

Using Equation (3-54) the column vector D_1 is found to be

$$D_1 = [P'^{-1}C_1]_+ \tag{3-54}$$

$$= \left[\begin{vmatrix} \dfrac{2-s}{\sqrt{7}+s} & 0 \\ & \\ \dfrac{3s}{(\sqrt{7}+s)(\sqrt{7}-2s)} & \dfrac{\sqrt{7}-s}{\sqrt{7}-2s} \end{vmatrix} \begin{vmatrix} \dfrac{3}{4-s^2} \\ \\ \dfrac{-3s}{4-s^2} \end{vmatrix} \right]_+$$

$$= \begin{vmatrix} \dfrac{3}{(\sqrt{}+s)(2+s)} \\ \\ \dfrac{1}{\sqrt{7}+s} \end{vmatrix}$$

At this point the modification steps are concluded. In the next section the functional method presented in reference [B.4] is used to solve the modified form of the vector Wiener-Hopf equation for this example.

Functional Method Steps

The steps involved in the functional method are summarized in Appendix D. Originally only six steps were suggested in reference [B.4]. Step 6 of the original reference has to be modified to guarantee the stability of the elements of the resulting system matrix. Step 7 has been added to include calculation of $\overline{\epsilon^2_{min.}(t)}$ in each case. At some points corrections have been made to take care of the misprints and/or mistakes contained in the original paper [B.4].

Step 1. $|P(s)| = 0$ has only one nonminimum phase zero in this case. Therefore

$$m = 1,\ z_1 = \sqrt{7},\ Z = (z_1 + z_1)^{-1} = \frac{1}{2\sqrt{7}}$$

Step 2. The adjoint of P(s) is found to be

$$A_{(p)}(s) = \begin{vmatrix} \frac{\sqrt{7}+2s}{\sqrt{7}+s} & \frac{-3s}{(\sqrt{7}+s)(2+s)} \\ 0 & \frac{\sqrt{7}-s}{2+s} \end{vmatrix}$$

The first row of the above matrix is chosen to construct the B matrix

$$B(z_1) = \begin{vmatrix} \frac{\sqrt{7}+2z_1}{\sqrt{7}+z_1} & \frac{-3z_1}{(\sqrt{7}+z_1)(2+z_1)} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{3}{2} & \frac{-3}{2(2+\sqrt{7})} \end{vmatrix}$$

which is a 1 x 2 matrix in contrast to a 2 x 1 matrix as suggested in reference [B.4].

Step 3. The matrix F is defined as $F = BB^T$ and is computed to be

$$F = BB^T = \begin{vmatrix} \frac{3}{2} & \frac{-3}{2(2+\sqrt{7})} \end{vmatrix} \begin{vmatrix} \frac{3}{2} \\ \\ \frac{-3}{2(2+\sqrt{7})} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{9}{4}\left[\frac{(2+\sqrt{7})^2+1}{(2+\sqrt{7})^2}\right] \end{vmatrix}$$

The matrix H is defined as the Schur product [B.1] of F and Z; i.e.,

$$h_{ij} = f_{ij}z_{ij}$$

Therefore

$$H = \begin{vmatrix} \frac{9}{4}\left(\frac{(2+\sqrt{7})^2+1}{(2+\sqrt{7})^2}\right) \frac{1}{2\sqrt{7}} \end{vmatrix}$$

and hence

$$H^{-1} = \begin{vmatrix} \frac{4}{9}\left(\frac{(2+\sqrt{7})^2}{(2+\sqrt{7})^2+1}\right) \; 2\sqrt{7} \end{vmatrix}$$

Step 4. The components of column vector V_1 denoted by v_{k1}, are computed from

$$v_{k1} = B^K D_K(z_K)$$

where, of course, B^K means Kth row of matrix B and D_K denotes the Kth column of the matrix D. Then

$$v_{11} = \begin{vmatrix} \frac{3}{2} & \frac{-3}{2(2+\sqrt{7})} \end{vmatrix} \begin{vmatrix} \frac{3}{2\sqrt{7}(2+\sqrt{7})} \\ \\ \frac{1}{2\sqrt{7}} \end{vmatrix}$$

$$= \frac{3}{2} \; \frac{1}{2\sqrt{7}+\sqrt{7}} = 0.122$$

In this particular case

$$V_1 = v_{11} = 0.122$$

Step 5. Column vector U_1 is defined as $H^{-1}V_1$ and matrix Γ_1 is a diagonal matrix whose elements γ_{ii} are the ith components of column vector U_1. Therefore

$$U_1 = \begin{vmatrix} \frac{9}{4}\left[\frac{(2+\sqrt{7})^2}{(2+\sqrt{7})^2+1}\right] 2\sqrt{7} \end{vmatrix} \begin{vmatrix} \frac{3}{2(2\sqrt{7}+\sqrt{7})} \end{vmatrix}$$

$$= 0.274$$

and

$$\Gamma_1 = U_1$$

Matrix R_1 is defined as

$$R_1 = B^T\Gamma_1 = \begin{vmatrix} 0.411 \\ \\ -0.088 \end{vmatrix}$$

Step 6. The column vector y_K is defined as

$$Y_K = D_K - R_K S$$

where S is a vector whose ith component is given by $(s+z_i)^{-1}$. Therefore

$$S = \frac{1}{\sqrt{7}+s}$$

and

$$Y_1 = \begin{vmatrix} \dfrac{3}{(\sqrt{7}+s)(2+s)} \\ \\ \dfrac{1}{\sqrt{7}+s} \end{vmatrix} - \begin{vmatrix} \dfrac{0.411}{\sqrt{7}+s} \\ \\ \dfrac{-0.088}{\sqrt{7}+s} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{3-0.411(2+s)}{\sqrt{7}+s)(2+s)} \\ \\ \dfrac{1.088}{\sqrt{7}+s} \end{vmatrix}$$

Now the column vector X_1 is calculated from

$$[X_1]_+ = [P^{-1}\, Y_1]_+$$

or

$$X_1 = \left[\begin{bmatrix} \dfrac{2+s}{\sqrt{7}-s} & \dfrac{-3s}{(\sqrt{7}-s)(\sqrt{7}+2s)} \\ \\ 0 & \dfrac{\sqrt{7}+s}{\sqrt{7}+2s} \end{bmatrix} \begin{vmatrix} \dfrac{3-0.411(2+s)}{(\sqrt{\ }+s)(2+s)} \\ \\ \dfrac{1.088}{\sqrt{7}+s} \end{vmatrix} \right]_+$$

$$= \begin{vmatrix} \dfrac{0.411}{s+1.32} \\ \\ \dfrac{0.544}{s+1.32} \end{vmatrix}$$

which is in complete agreement with the corresponding results obtained by Amara [A.2].

Step 7. In this step the numerical value of the minimum mean-square error, $\overline{\epsilon^2_{min.}(t)}$ for the example cited here, is found. From the derivation of the results of the functional method, as presented in reference [B.4], one obtains

$$\overline{\epsilon^2_{min.}(t)} = q_{10} + q_1$$

where

$$q_{10} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \{tr.[G_{dd}(s)] - \bar{D}_1^T(s)\ D_1(s)\}ds$$

and

$$q_1 = U_1^T V_1$$

Therefore

$$q_{10} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \left[\frac{3}{4-s^2} - \frac{13-s^2}{(7-s^2)(4-s^2)} \right] ds$$

$$= \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \frac{8-2s^2}{(s^2-7)(s^2-4)}\ ds$$

Performing this integration using the residue theorem* it can be shown that

$$q_{10} = \frac{1}{\sqrt{7}} = 0.378$$

Now q_1 is given by

$$q_1 = U_1^T V_1$$

$$= (0.274)(0.122) = 0.0334$$

Hence

$$\overline{\epsilon^2_{min.}(t)} = 0.378 + 0.0334 = 0.411$$

---

*See Chapter II or reference [K.2] for more detail.

which again agrees with the result calculated by Amara [A.2]. This completes the solution of this problem using the functional method. In the next section the proposed alternative method of solution of the vector Wiener-Hopf equation is applied to the same example cited above.

Proposed Alternative Method Steps

A summary of the steps involved in this method was given in Chapter III. A reference to this summary shows that the first two steps are modifiers in nature therefore the same modification steps taken early in the present chapter still are valid.

Step 3. To determine the system's probable stable poles Equation (3-58),

$$[X]_+ = [P^{-1}D]_+ \quad , \tag{3-58}$$

is utilized. There results

$$[X_1]_+ = \left[\begin{vmatrix} \dfrac{2+s}{\sqrt{7}-s} & \dfrac{-3s}{(\sqrt{7}-s)(\sqrt{7+2s})} \\ & \\ 0 & \dfrac{\sqrt{7}+s}{\sqrt{7+2s}} \end{vmatrix} \begin{vmatrix} \dfrac{3}{(\sqrt{7}+s)(2+s)} \\ \\ \dfrac{1}{\sqrt{7}+s} \end{vmatrix}\right]_+$$

$$= \begin{vmatrix} \dfrac{a_{11}}{\sqrt{7+2s}} \\ \\ \dfrac{a_{21}}{\sqrt{7+2s}} \end{vmatrix}$$

Step 4. With X_1 as defined above, Equation (3-51)

$$[AX_1]_+ = [C_1]_+ \ , \tag{3-51}$$

becomes

$$\left[\begin{bmatrix} \dfrac{7-s^2}{4-s^2} & \dfrac{3s}{4-s^2} \\ & \\ \dfrac{-3s}{4-s^2} & \dfrac{4-4s^2}{4-s^2} \end{bmatrix} \begin{bmatrix} \dfrac{a_{11}}{\sqrt{7}+2s} \\ \\ \dfrac{a_{21}}{\sqrt{7}+2s} \end{bmatrix}\right]_+ = \begin{bmatrix} \dfrac{3}{4-s^2} \\ \\ \\ \end{bmatrix}_+$$

Performing the matrix multiplication and equating the corresponding elements on each side results

$$\frac{7-s^2}{4-s^2} \cdot \frac{a_{11}}{\sqrt{7}+2s} + \frac{3s}{4-s^2} \cdot \frac{a_{21}}{\sqrt{7}+2s} = \frac{3}{4-s^2}$$

The operation $[\ldots\ldots]_+$ implies that the resides of similar stable poles on each side of the above equation must be equal. In other words the following two equations must hold.

$$\frac{(7-4)a_{11} - 6a_{21}}{4(\sqrt{7}-4)} = \frac{3}{4}$$

$$\frac{(7-\frac{7}{4})a_{11} - \frac{3\sqrt{7}}{2}a_{21}}{(4-\frac{7}{4})} = 0$$

from which one obtains

$$a_{11} = 0.822$$

$$a_{21} = 1.09$$

which are the same results as before.

Step 5. This step is not needed for this particular problem.

Step 6. The numerical value for the minimum mean-square error is given by Equation (3-48); i.e.,

$$\overline{\epsilon^2_{min.}(t)} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \text{tr. } [G_{dd}(s) - \bar{W}(s)G_{rd}(s)]ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} [G^{11}_{dd}(s) - \{\bar{W}_{11}(s)G^{11}_{rd}(s)$$

$$+ \bar{W}_{12}(s)G^{21}_{rd}(s)\}]ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \left[\frac{3}{4-s^2}\left(1 - \frac{0.411}{1.32-s} + \frac{0.545s}{1.32-s}\right)\right] ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \frac{2.73-1.365s}{(4-s^2)(1.32-s)} ds$$

$$= 0.411$$

which again agrees with the previous result. This completes the solution.

A comparison of the proposed alternative method of solution of the vector Wiener-Hopf equation, used in this section, with the functional method utilized in the previous section shows that the method presented in this section is more direct and fewer computations are necessary.

In the next section a technique for generating $\overline{\epsilon^2_{min.}(t)}$ on an analog computer is applied to the present example.

Generating $\overline{\epsilon^2_{min.}(t)}$ on an Analog Computer

The minimum sum of the weighted mean-square error, $\overline{\epsilon^2_{min.}(t)}$, in the time domain is given by

$$\overline{\epsilon^2_{min.}(t)} = \text{tr. } [L\,\varphi_{dd}(0) - \int_{-\infty}^{-\infty} L\, w_m(t)\ \varphi_{rd}(t)dt] \quad (3\text{-}28)$$

For this particular example the above equation is expanded to give

$$\overline{\epsilon^2_{min.}(t)} = \varphi^{11}_{dd}(0) - \int_0^t [w_{11}(\tau)\ \varphi^{11}_{rd}(\tau)d\tau$$

$$- \int_0^t w_{12}(\tau)\ \varphi^{21}_{rd}(\tau)d\tau$$

since the elements of the diagonal matrix L are all equal to unity. It can be shown that in this example

$$\varphi^{11}_{dd}(0) = \frac{3}{4}$$

$$w_{11}(t) = 0.411\, e^{-1.32t} \quad ; \quad w_{12}(t) = 0.544\, e^{-1.32t}$$

$$\varphi_{rd}^{11}(t) = \frac{3}{4}\, e^{-2t}, \qquad \varphi_{rd}^{21}(t) = \frac{3}{2}\, e^{-2t}$$

Although in this particular case the solution is simple, it is generated for illustrative purposes. The analog computer program for this problem is shown in Figure 4-2.

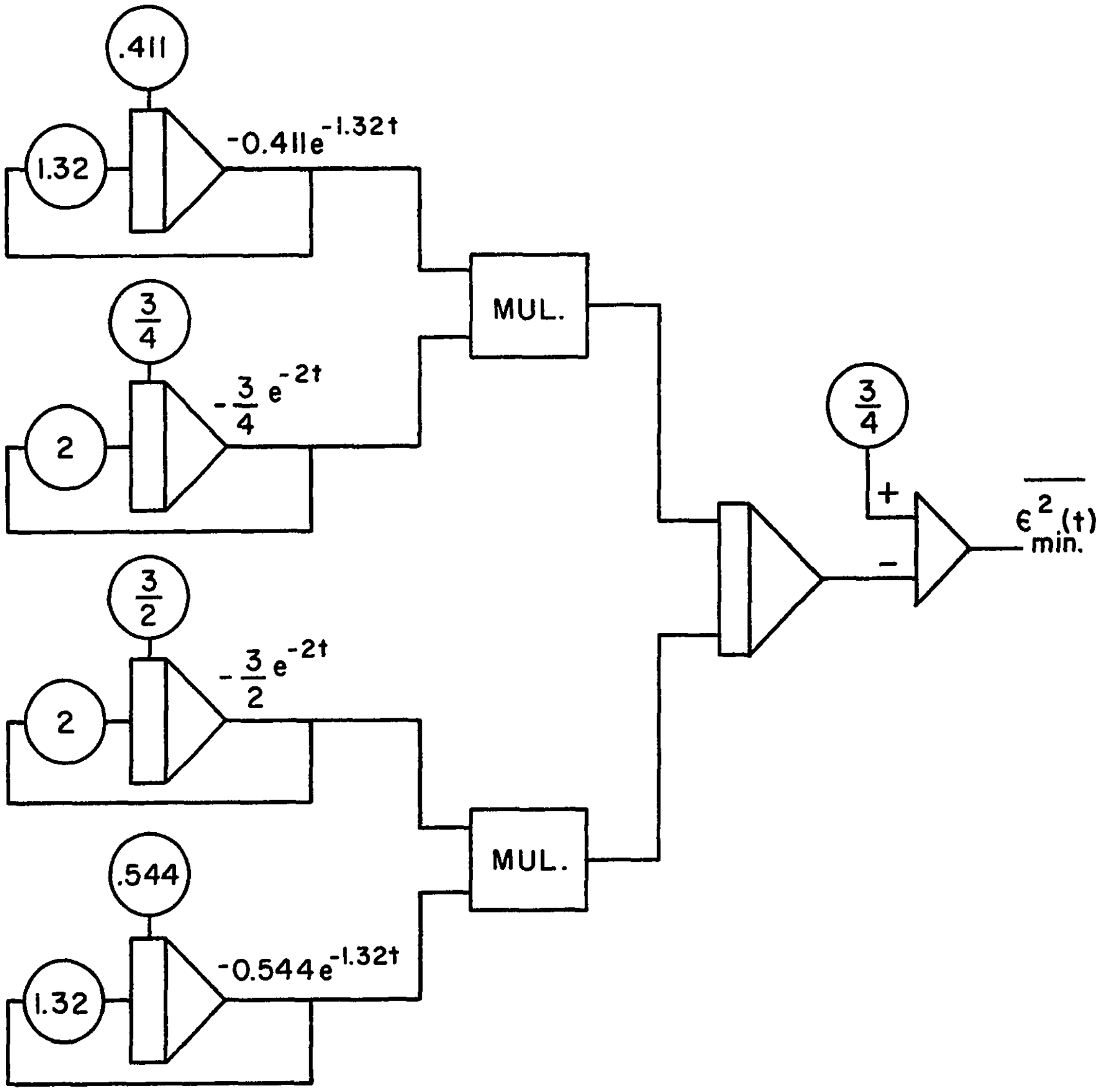

Figure 4-2. Analog Computer Program for Generating $\overline{\epsilon^2_{min.}(t)}$ for Example 1.

Example 2 -

Given: The system under consideration in this case is shown in Figure 4-3. Again each of the inputs, r_i, is composed of message m_i corrupted with white noise n_i. The noises are uncorrelated with each other and with each message. The two messages are statistically independent random variables.

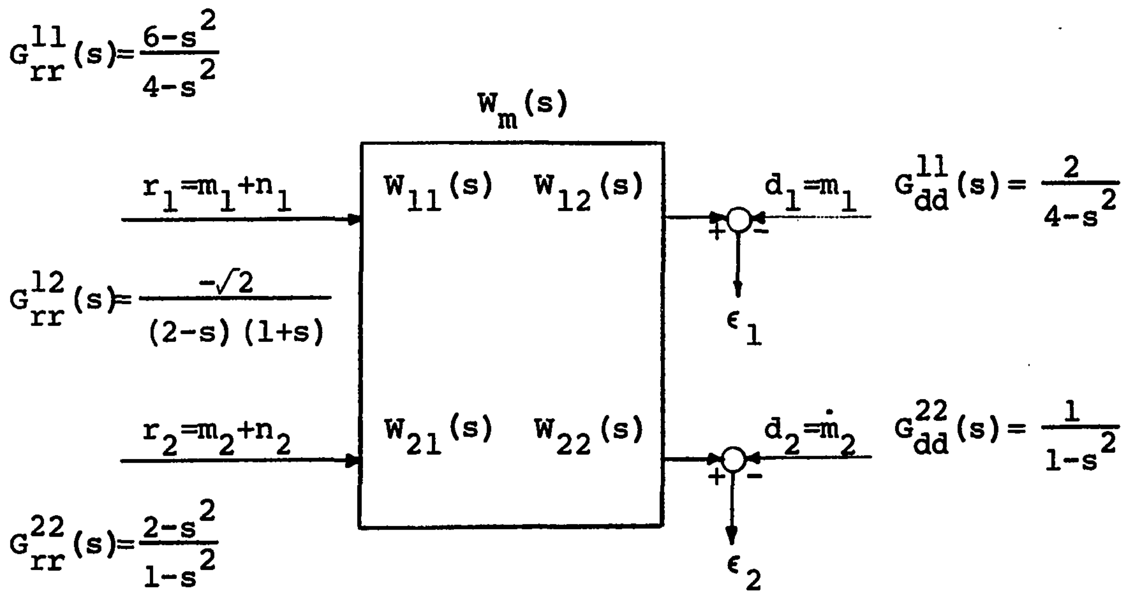

Figure 4-3. Diagrammatic Representation of Example 2

Required: It is desired to extract the best m_1 and the best derivative of m_2 (in a mean-square sense) by operating on the inputs r_1 and r_2 simultaneously.

Solution: From the information given in the problem and shown in Figure 4-3 one obtains the following

$$A = \begin{vmatrix} \dfrac{6-s^2}{4-s^2} & \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ \\ \dfrac{-\sqrt{2}}{(2+s)(1-s)} & \dfrac{2-s^2}{1-s^2} \end{vmatrix}$$

$$X = \begin{vmatrix} X_{11}(s) & X_{12}(s) \\ \\ X_{21}(s) & X_{22}(s) \end{vmatrix}$$

$$G_{dd}(s) = \begin{vmatrix} \dfrac{2}{4-s^2} & \dfrac{1}{(2-s)(1+s)} \\ \\ \dfrac{1}{(2+s)(1-s)} & \dfrac{1}{1-s^2} \end{vmatrix}$$

$$C = \begin{vmatrix} \dfrac{2}{4-s^2} & \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ \\ \dfrac{-\sqrt{2}}{(2+s)(1-s)} & \dfrac{-1}{1-s^2} \end{vmatrix}$$

The weighting matrix L is again assumed to be an identity matrix. Corresponding to Equation (3-51),

$$[AX]_+ = [C]_+$$

$$\left[\begin{vmatrix} \dfrac{6-s^2}{4-s^2} & \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ \\ \dfrac{-\sqrt{2}}{(2+s)(1-s)} & \dfrac{2-s^2}{1-s^2} \end{vmatrix} \begin{vmatrix} X_{11}(s) & X_{12}(s) \\ \\ X_{21}(s) & X_{22}(s) \end{vmatrix}\right]_+$$

$$= \begin{bmatrix} \dfrac{2}{4-s^2} & \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ \\ \dfrac{-\sqrt{2}}{(2+s)(1-s)} & \dfrac{-1}{1-s^2} \end{bmatrix}$$

<u>Modification Steps</u>

Step 1. The lowest common denominator of matrix A is

$$g(s) = (4-s^2)(1-s^2)$$

Step 2. The newly defined matrix G then becomes

$$G = gA$$

$$= \begin{vmatrix} (1-s^2)(6-s^2) & -\sqrt{2}(1-s)(2+s) \\ \\ -\sqrt{2}(1+s)(2-s) & (2-s^2)(4-s^2) \end{vmatrix}$$

Step 3. Function g(s) may now be factored as

$$g(s) = \bar{f}(s)\ f(s)$$

where

$$\bar{f}(s) = (2-s)(1-s)$$

$$f(s) = (2+s)(1+s)$$

Step 4. The elements of the upper triangular matrix F then becomes

$$f_{11}(s) = (1-s)(\sqrt{6}-s)$$

$$f_{12}(s) = \frac{-\sqrt{2}(2+s)(1-s)}{(1+s)(\sqrt{6}+s)} \quad ; \quad f_{21}(s) = 0$$

$$f_{22}(s) = \left|(2-s^2)(4-s^2) - \frac{2(4-s^2)}{6-s^2}\right|^+$$

$$= \frac{(2+s)(a+s)(b+s)}{(\sqrt{6}+s)}$$

where $a^2 = 4+\sqrt{6}$ and $b^2 = 4-\sqrt{6}$.

Step 5. The matrices P and P' then become

$$P(s) = \frac{1}{f(s)} F(s)$$

$$= \begin{vmatrix} \dfrac{(1-s)(\sqrt{6}-s)}{(2+s)(1+s)} & \dfrac{-\sqrt{2}(1-s)}{(1+s)^2(\sqrt{6}+s)} \\ & \\ 0 & \dfrac{(a+s)(b+s)}{(1+s)(\sqrt{6}+s)} \end{vmatrix}$$

$$P'(s) = \bar{P}^T(s)$$

$$= \begin{vmatrix} \dfrac{(1+s)(\sqrt{6}+s)}{(2-s)(1-s)} & 0 \\ \\ \dfrac{-\sqrt{2}(1+s)}{(1-s)^2(\sqrt{6}-s)} & \dfrac{(a-s)(b-s)}{(1-s)(\sqrt{6}-s)} \end{vmatrix}$$

The inverses of P and P' are calculated to be respectively

$$P^{-1}(s) = \begin{vmatrix} \dfrac{(2+s)(1+s)}{(1-s)(\sqrt{6}-s)} & \dfrac{\sqrt{2}(2+s)}{(a+s)(b+s)(\sqrt{6}-s)} \\ \\ 0 & \dfrac{(1+s)(\sqrt{6}+s)}{(a+s)(b+s)} \end{vmatrix}$$

$$P'^{-1}(s) = \begin{vmatrix} \dfrac{(1-s)(2-s)}{(1+s)(\sqrt{6}+s)} & 0 \\ \\ \dfrac{\sqrt{2}(2-s)}{(a-s)(b-s)(\sqrt{6}+s)} & \dfrac{(1-s)(\sqrt{6}-s)}{(a-s)(b-s)} \end{vmatrix}$$

Using Equation (3-54) the columns of matrix D are found to be

$$D_1(s) = [P'^{-1}C_1]_+ \qquad (3\text{-}54)$$

$$= \left[\begin{vmatrix} \dfrac{(1-s)(2-s)}{(1+s)(\sqrt{6}+s)} & 0 \\ \\ \dfrac{\sqrt{2}(2-s)}{(a-s)(b-s)(\sqrt{6}+s)} & \dfrac{(1-s)(\sqrt{6}-s)}{(a-s)(b-s)} \end{vmatrix} \begin{vmatrix} \dfrac{2}{4-s^2} \\ \\ \dfrac{-\sqrt{2}}{(1-s)(2+s)} \end{vmatrix} \right]_+$$

$$= \begin{vmatrix} \dfrac{2(1-s)}{(1+s)(\sqrt{6}+s)(2+s)} \\ \\ \dfrac{-0.3408}{\sqrt{6}+s} \end{vmatrix}$$

and

$$D_2(s) = [P'^{-1} C_2]$$

$$= \left[\begin{vmatrix} \dfrac{(2-s)(1-s)}{(1+s)(\sqrt{6}+s)} & 0 \\ \\ \dfrac{\sqrt{2}(2-s)}{(a-s)(b-s)(\sqrt{6}+s)} & \dfrac{(1-s)(\sqrt{6}-s)}{(a-s)(b-s)} \end{vmatrix} \begin{vmatrix} \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ \\ \dfrac{-1}{1-s^2} \end{vmatrix} \right]_+$$

$$= \begin{vmatrix} \dfrac{-\sqrt{2}(1-s)}{(1+s)^2(\sqrt{6}+s)} \\ \\ \dfrac{0.0747}{\sqrt{6}+s} - \dfrac{0.6076}{(1+s)} \end{vmatrix}$$

<u>Functional Method Steps</u>

Step 1. $|P(s)| = 0$ has two nonminimum phase zeroes in this case. Therefore

$$m = 2, \; z_1 = 1, \; z_2 = \sqrt{6}$$

The matrix Z then is

$$Z = \begin{vmatrix} \frac{1}{2} & \frac{1}{1+\sqrt{6}} \\ \\ \frac{1}{1+\sqrt{6}} & \frac{1}{2\sqrt{6}} \end{vmatrix}$$

Step 2. The adjoint of P(s) is found to be

$$A_{(p)}(s) = \begin{vmatrix} \frac{(a+s)(b+s)}{(1+s)(\sqrt{6}+s)} & \frac{\sqrt{2}(1-s)}{(1+s)^2(\sqrt{6}+s)} \\ \\ 0 & \frac{(1-s)(\sqrt{6}-s)}{(2+s)(1+s)} \end{vmatrix}$$

The first row of $A_{(p)}(s)$ is chosen to construct the B matrix; i.e.,

$$B^K(z_K) = \begin{vmatrix} \frac{(a+z_K)(b+z_K)}{(1+z_K)(\sqrt{6}+z_K)} & \frac{\sqrt{2}(1-z_K)}{(1+z_K)^2(\sqrt{6}+z_K)} \end{vmatrix}$$

Hence

$$B^1(1) = \begin{vmatrix} \frac{(a+1)(b+1)}{(1+1)(\sqrt{6}+1)} & 0 \end{vmatrix}$$

$$B^2(\sqrt{6}) = \begin{vmatrix} \frac{(a+\sqrt{6})(b+\sqrt{6})}{(1+\sqrt{6})(\sqrt{6}+\sqrt{6})} & \frac{\sqrt{6}(1-\sqrt{6})}{(1+\sqrt{6})^2(\sqrt{6}+\sqrt{6})} \end{vmatrix}$$

Thence

$$B = \begin{vmatrix} 1.152 & 0 \\ \\ 1.092 & -0.0352 \end{vmatrix}$$

Step 3. The matrix F is defined as

$$F = BB^T$$

$$= \begin{vmatrix} 1.152 & 0 \\ 1.092 & -0.0352 \end{vmatrix} \begin{vmatrix} 1.152 & 1.092 \\ 0 & -0.0352 \end{vmatrix}$$

$$= \begin{vmatrix} 1.33 & 1.26 \\ 1.26 & 1.194 \end{vmatrix}$$

The matrix H, being a Schur product [B.1] of F and Z, becomes

$$H = \begin{vmatrix} 0.665 & 0.365 \\ 0.365 & 0.244 \end{vmatrix}$$

and H^{-1} becomes

$$H^{-1} = \begin{vmatrix} 8.41 & -12.59 \\ -12.59 & 22.93 \end{vmatrix}$$

Step 4-1. The components of column vector V_1, denoted by v_{K1} are computed from

$$v_{K1} = B^K D_1(z_K)$$

Then:

$$v_{11} = B^1 D_1(1)$$

$$= \begin{vmatrix} 1.152 & 0 \end{vmatrix} \begin{vmatrix} 0 \\ \\ \dfrac{-0.34}{\sqrt{6}+1} \end{vmatrix} = 0$$

and

$$v_{21} = B^2 D_1(\sqrt{6})$$

$$= \begin{vmatrix} 1.092 & -0.0352 \end{vmatrix} \begin{vmatrix} \dfrac{2(1-\sqrt{6})}{(1+\sqrt{6})(2\sqrt{6})(2+\sqrt{6})} \\ \\ \dfrac{-0.34}{2\sqrt{6}} \end{vmatrix}$$

$$= -0.0445$$

Hence

$$V_1 = \begin{vmatrix} 0 \\ \\ -0.0445 \end{vmatrix}$$

Step 5-1. Column vector U_1 is defined as $H^{-1}V_1$ and matrix Γ_1 is a diagonal matrix whose elements γ_{ii} are the ith components of column vector U_1. Therefore

$$U_1 = \begin{vmatrix} 8.41 & -12.59 \\ \\ -12.59 & 22.93 \end{vmatrix} \begin{vmatrix} 0 \\ \\ -0.0445 \end{vmatrix}$$

$$= \begin{vmatrix} 0.56 \\ \\ -1.02 \end{vmatrix}$$

$$R_1 = B^T\Gamma_1$$

$$= \begin{vmatrix} 1.152 & 1.092 \\ 0 & -0.0352 \end{vmatrix} \begin{vmatrix} 0.56 & 0 \\ 0 & -1.02 \end{vmatrix}$$

$$= \begin{vmatrix} 0.645 & -1.114 \\ 0 & 0.036 \end{vmatrix}$$

Step 6-1. The column vector Y_K is defined as

$$Y_K = D_K - R_K S$$

where

$$S = \begin{vmatrix} \dfrac{1}{1+s} \\ \dfrac{1}{\sqrt{6}+s} \end{vmatrix}$$

Then:

$$R_1 S = \begin{vmatrix} 0.645 & -1.114 \\ 0 & 0.036 \end{vmatrix} \begin{vmatrix} \dfrac{1}{1+s} \\ \dfrac{1}{\sqrt{6}+s} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{0.645}{1+s} - \dfrac{1.114}{\sqrt{6}+s} \\ \dfrac{0.036}{\sqrt{6}+s} \end{vmatrix}$$

and

$$Y_1 = D_1 - R_1 S$$

$$= \begin{vmatrix} \dfrac{2(1-s)}{(1+s)(\sqrt{6}+s)(2+s)} \\ \\ \dfrac{-0.34}{\sqrt{6}+s} \end{vmatrix} - \begin{vmatrix} \dfrac{0.645}{1+s} - \dfrac{1.114}{\sqrt{6}+s} \\ \\ \dfrac{0.036}{\sqrt{6}+s} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{0.469(2.25-s)(1.014-s)}{(1+s)(\sqrt{6}+s)(2+s)} \\ \\ \dfrac{-0.3768}{\sqrt{6}+s} \end{vmatrix}$$

At this point reference [B.4] suggests that the column vector X_1 be calculated from $X_1 = P^{-1}Y_1$. However, if one requires that the components of X matrix have stable poles only then X_1 should be calculated from the following equation

$$[X_1]_+ = [P^{-1}Y_1]_+$$

since the fact that P has stable poles does not guarantee that its inverse be stable. Therefore

$$X_1 = [P^{-1}Y_1]_+$$

$$
= \left[\begin{vmatrix} \dfrac{(2+s)(1+s)}{(1-s)(\sqrt{6}-s)} & \dfrac{\sqrt{2}(2+s)}{(a+s)(b+s)(\sqrt{6}-s)} \\ \\ 0 & \dfrac{(1+s)(\sqrt{6}+s)}{(a+s)(b+s)} \end{vmatrix} \begin{vmatrix} \dfrac{0.469(2.25-s)(1.014-s)}{(1+s)(\sqrt{6}+s)(2+s)} \\ \\ \dfrac{-0.3768}{\sqrt{6}+s} \end{vmatrix} \right]_+
$$

$$
= \begin{vmatrix} \dfrac{0.494}{a+s} - \dfrac{0.07}{b+s} \\ \\ \dfrac{-0.448}{a+s} + \dfrac{0.0713}{b+s} \end{vmatrix}
$$

But X_1 is only the first column of matrix X. To determine the second column; X_2, of matrix X steps 4 through 6 are repeated with column vector D_1 being replaced by D_2.

Step 4-2. The components of column vector V_2, denoted by v_{K2}, are computed from

$$v_{K2} = B^K D_2(z_K)$$

Then:

$$
\begin{aligned}
v_{12} &= B^1 D_2(1) \\
&= \begin{vmatrix} 1.152 & 0 \end{vmatrix} \begin{vmatrix} 0 \\ \\ \dfrac{0.0747}{\sqrt{6}+1} - \dfrac{0.6076}{1+1} \end{vmatrix} \\
&= 0
\end{aligned}
$$

and

$$v_{22} = B^2 D_2 (\sqrt{6})$$

$$= \begin{vmatrix} 1.096 & -0.0352 \end{vmatrix} \begin{vmatrix} 0.035 \\ -0.1608 \end{vmatrix}$$

$$= 0.0326$$

Hence

$$V_2 = \begin{vmatrix} 0 \\ 0.0326 \end{vmatrix}$$

Step 5-2. Column vector U_2 is defined as $H^{-1}V_2$ and matrix Γ_2 is a diagonal matrix whose elements γ_{ii} are the ith components of column vector U_2. Therefore

$$U_2 = \begin{vmatrix} 8.41 & -12.59 \\ -12.59 & 22.93 \end{vmatrix} \begin{vmatrix} 0 \\ 0.0326 \end{vmatrix}$$

$$= \begin{vmatrix} -0.410 \\ 0.748 \end{vmatrix}$$

and

$$\Gamma_2 = \begin{vmatrix} -0.410 & 0 \\ 0 & 0.748 \end{vmatrix}$$

Matrix R_2 is defined as

$$R_2 = B^T\Gamma_2$$

$$= \begin{vmatrix} 1.152 & 1.092 \\ 0 & -0.0352 \end{vmatrix} \begin{vmatrix} -0.410 & 0 \\ 0 & 0.748 \end{vmatrix}$$

$$= \begin{vmatrix} -0.472 & 0.817 \\ 0 & -0.026 \end{vmatrix}$$

Step 6-2. The column vector Y_K is defined as

$$Y_K = D_K - R_K S$$

where again

$$S = \begin{vmatrix} \dfrac{1}{1+s} \\ \\ \dfrac{1}{\sqrt{6}+s} \end{vmatrix}$$

Then:

$$R_2 S = \begin{vmatrix} -0.472 & 0.817 \\ 0 & -0.026 \end{vmatrix} \begin{vmatrix} \dfrac{1}{1+s} \\ \\ \dfrac{1}{\sqrt{6}+s} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{0.817}{\sqrt{6}+s} - \dfrac{0.472}{1+s} \\ \\ \dfrac{-0.026}{\sqrt{6}+s} \end{vmatrix}$$

and

$$Y_2 = D_2 - R_2S$$

$$= \begin{vmatrix} \dfrac{-\sqrt{2}(1-s)}{(1+s)^2(\sqrt{6}+s)} \\ \\ \dfrac{0.0747}{\sqrt{6}+s} - \dfrac{0.0607}{1+s} \end{vmatrix} - \begin{vmatrix} \dfrac{0.817}{\sqrt{6}+s} - \dfrac{0.472}{1+s} \\ \\ \dfrac{-0.026}{\sqrt{6}+s} \end{vmatrix}$$

$$= \begin{vmatrix} \dfrac{-0.345(3.11-4.08s+s^2}{(1+s)^2(\sqrt{\ }+s)} \\ \\ \dfrac{-0.5076(2.74+s)}{(1+s)(\sqrt{6}+s)} \end{vmatrix}$$

Now the column vector X_2 is calculated from

$$X_2 = [P^{-1}Y_2]_+$$

$$= \left[\begin{vmatrix} \dfrac{(2+s)(1+s)}{(1-s)(\sqrt{6}-s)} & \dfrac{\sqrt{2}(2+s)}{(a+s)(b+s)(\sqrt{6}-s)} \\ \\ 0 & \dfrac{(1+s)(\sqrt{6}+s)}{(a+s)(b+s)} \end{vmatrix} \begin{vmatrix} \dfrac{-0.345(3.11-4.08s+s^2)}{(1+s)^2(\sqrt{6}+s)} \\ \\ \dfrac{-0.5076(2.74+s)}{(1+s)(\sqrt{6}+s)} \end{vmatrix} \right]_+$$

$$= \begin{vmatrix} \dfrac{-0.0876}{a+s} + \dfrac{0.574}{b+s} - \dfrac{0.945}{1+s} \\ \\ \dfrac{0.0784}{a+s} - \dfrac{0.586}{b+s} \end{vmatrix}$$

At this point the determination of matrix X is complete.

Step 7. The numerical value of the minimum mean-square error at the first output is given by

$$\overline{\epsilon^2_{min.}(t)} = q_{10} + q_1$$

where

$$q_{10} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}[G^{11}_{dd}(s) - D_1^{-T}(s)D_1(s)]ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\left[\frac{2}{4-s^2} - \left(\frac{4(1-s^2)}{(1-s^2)(6-s^2)(4-s^2)} + \frac{(0.34)^2}{6-s^2}\right)\right] ds$$

$$= 0.3845$$

and q_1 is given by

$$q_1 = U_1^T V_1$$

$$= \begin{vmatrix} 0.56 & -1.02 \end{vmatrix} \begin{vmatrix} 0 \\ -0.0445 \end{vmatrix}$$

$$= 0.0454$$

Hence the numerical value of the minimum mean-square error at the first output is

$$\overline{\epsilon^2_{min.}(t)} = 0.345 + 0.0454$$

$$= 0.4299$$

Similarly at the second output $\overline{\epsilon^2_{min.}(t)} = q_{20} + q_2$ where

$$q_{20} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} [G_{dd}^{22}(s) - D_2^{-T}(s)\, D_2(s)]ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \left[\frac{1}{1-s^2} - \left(\frac{2}{6-s^2} + \left(\frac{-0.6076}{1-s} + \frac{0.0747}{\sqrt{6}-s}\right)\right.\right.$$

$$\left.\left.\left(\frac{-0.6076}{1+s} + \frac{0.0747}{\sqrt{6}+s}\right)\right)\right] ds$$

Applying the result of residue theorem the value of q_{20} is found to be

$$q_{20} = 0.2216$$

Now q_2 is given by

$$q_2 = U_2^T\, V_2$$

$$= \begin{vmatrix} -0.41 & 0.748 \end{vmatrix} \begin{vmatrix} 0 \\ 0.0326 \end{vmatrix}$$

$$= 0.0244$$

Hence the numerical value of $\overline{\epsilon^2_{min.}(t)}$ at the second output is

$$\overline{\epsilon^2_{min.}(t)} = 0.2216 + 0.0244$$

$$= 0.246$$

This completes the solution of the second example using the functional method. In the next section the proposed alternative method is applied to Example 2.

Proposed Alternative Method Steps

Steps 1 and 2 of this method have already been taken in the early part of the solution of the second example. Therefore the results of the Modification Steps are used in the following:

Step 3. To determine probable stable poles of the system Equation (3-58)

$$[X_K] = [P^{-1}D_K]_+ \quad , \tag{3-58}$$

is used. There results

$$X_1' = \left[\overset{P^{-1}}{\begin{vmatrix} \dfrac{(2+s)(1+s)}{(1-s)(\sqrt{6}-s)} & \dfrac{\sqrt{2}(2+s)}{(a+s)(b+s)(\sqrt{6}-s)} \\ & \\ 0 & \dfrac{(1+s)(\sqrt{6}+s)}{(a+s)(b+s)} \end{vmatrix}} \overset{D_1}{\begin{vmatrix} \dfrac{2(1-s)}{(1+s)(\sqrt{6}+s)(2+s)} \\ \\ \dfrac{-0.34}{\sqrt{6}+s} \end{vmatrix}} \right]_+$$

$$= \begin{vmatrix} \dfrac{A_{11}'}{a+s} + \dfrac{A_{12}'}{b+s} + \dfrac{A_{13}'}{\sqrt{6}+s} \\ \\ \dfrac{A_{21}'}{a+s} + \dfrac{A_{22}'}{b+s} \end{vmatrix}$$

and

$$X_2' = \left[\begin{bmatrix} \dfrac{(2+s)(1+s)}{(1-s)(\sqrt{6}-s)} & \dfrac{\sqrt{2}(2+s)}{(a+s)(b+s)(\sqrt{6}-s)} \\ \\ 0 & \dfrac{(1+s)(\sqrt{6}+s)}{(a+s)(b+s)} \end{bmatrix} \begin{bmatrix} \dfrac{-\sqrt{2}(1-s)}{(1+s)^2(\sqrt{6}+s)} \\ \\ \dfrac{0.0747}{\sqrt{6}+s} - \dfrac{0.6076}{(1+s)} \end{bmatrix}\right]_+$$

$$= \begin{bmatrix} \dfrac{B_{11}'}{a+s} + \dfrac{B_{12}'}{b+s} + \dfrac{B_{13}'}{1+s} + \dfrac{B_{14}'}{\sqrt{6}+s} \\ \\ \dfrac{B_{21}'}{a+s} + \dfrac{B_{22}'}{b+s} \end{bmatrix}$$

Now the coefficients A_{ij}' and B_{ij}' must be modified by considering the effect of matrix P'. Let the modified coefficients be A_{ij} and B_{ij} respectively.

Step 4 and 5. With matrix X defined, Equation (3-51)

$$[AX]_+ = [C]_+ \tag{3-51}$$

becomes

$$\begin{array}{ccc} A(s) & X_1(s) & C_1(s) \\ \left[\begin{bmatrix} \dfrac{6-s^2}{4-s^2} & \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ \\ \dfrac{-\sqrt{2}}{(2+s)(1-s)} & \dfrac{2-s^2}{1-s^2} \end{bmatrix} \begin{bmatrix} \dfrac{A_{11}}{a+s} + \dfrac{A_{12}}{b+s} + \dfrac{A_{13}}{\sqrt{6}+s} \\ \\ \dfrac{A_{21}}{a+s} + \dfrac{A_{22}}{b+s} \end{bmatrix}\right]_+ & = & \begin{bmatrix} \dfrac{2}{4-s^2} \\ \\ \dfrac{-\sqrt{2}}{(1-s)(2+s)} \end{bmatrix}_+ \end{array}$$

and

$$\begin{array}{ccc} A(s) & X_2(s) & C_2(s) \\ \left[\left[\begin{matrix} \dfrac{6-s^2}{4-s^2} & \dfrac{-\sqrt{2}}{(2-s)(1+s)} \\ & \\ \dfrac{-\sqrt{2}}{(2+s)(1-s)} & \dfrac{2-s^2}{1-s^2} \end{matrix}\right] \left[\begin{matrix} \dfrac{B_{11}}{a+s}+\dfrac{B_{12}}{b+s}+\dfrac{B_{13}}{1+s}+\dfrac{B_{14}}{\sqrt{6}+s} \\ \\ \dfrac{B_{21}}{a+s}+\dfrac{B_{22}}{b+s} \end{matrix}\right]\right]_+ = & & \left[\begin{matrix} \dfrac{-\sqrt{2}}{(1+s)(2-s)} \\ \\ \dfrac{-1}{1-s^2} \end{matrix}\right]_+ \end{array}$$

Since the residue of similar stable poles on each side of the above two matrix equations must be equal then a set of simultaneous equations may be written in terms of the unknown coefficients A_{ij} and B_{ij}. From the simultaneous solution of these equations the following numerical values for A_{ij}'s and B_{ij}'s are determined

$$A_{11} = 0.49,\ A_{12} = -0.0694,\ A_{13} = 0$$

$$A_{21} = -0.448,\ A_{22} = 0.071$$

$$B_{11} = -0.0871,\ B_{12} = 0.5746,\ B_{13} = -0.9427,\ B_{14} = 0$$

$$B_{21} = 0.079,\ B_{22} = -0.586$$

These coefficients are almost the same as the corresponding coefficients found using the functional method. Note also how the vanishing of coefficients A_{13} and B_{14} resulted in the cancellation of the corresponding poles in $X_1(s)$ and $X_2(s)$ such that the optimal system matrix for this particular example becomes

$W(s) \equiv X^T(s)$

$$= \begin{vmatrix} \dfrac{0.49}{a+s} \; \dfrac{0.069}{b+s} & \dfrac{-0.448}{a+s} + \dfrac{0.071}{b+s} \\ & \\ \dfrac{-0.087}{a+s} + \dfrac{0.575}{b+s} - \dfrac{0.943}{1+s} & \dfrac{0.079}{1+s} - \dfrac{0.588}{b+s} \end{vmatrix}$$

where

$$a = 2.54 \quad ; \quad b = 1.245$$

Step 6. The minimum mean-square error at the first output, using the coefficient values found by the functional method, is given by

$$\overline{\epsilon^2_{\min.}(t)} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\left[\frac{2}{4-s^2} - \left(\frac{0.494(b-s)-0.0698(a-s)}{(a-s)(b-s)} \cdot \frac{2}{4-s^2}\right)\right.$$

$$\left. - \left(\frac{-0.448(b-s)+0.0.0713(a-s)}{(a-s)(b-s)} \cdot \frac{-\sqrt{2}}{(1-s)(2+s)}\right)\right] ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\left[\frac{2(a-s)(b-s)(1-s)-2(1-s)[0.494(b-s)-0.0698(a-s)]}{(4-s^2)(a-s)(b-s)(1-s)}\right.$$

$$\left. + \frac{\sqrt{2}(2-s)[-0.448(b-s)+0.0713(a-s)]}{(4-s^2)(a-s)(b-s)(1-s)}\right] ds$$

$$= 0.4254$$

Similarly the minimum mean-square error at the second output is found to be

$$\overline{\epsilon^2_{min.}(t)} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\left[\frac{1}{1-s^2} + \left(\frac{0.0876(b-s)(1-s)}{(a-s)(b-s)(1-s)}\right)\frac{-\sqrt{2}}{(1+s)(2-s)}\right.$$

$$+\left(\frac{0.574(a-s)(1-s)-0.945(a-s)(b-s)}{(a-s)(b-s)(1-s)}\right)\frac{\sqrt{2}}{(1+s)(2-s)}$$

$$\left.-\left(\frac{0.0784(b-s)-0.586(a-s)}{(b-s)(a-s)}\right)\frac{1}{1-s^2}\right] ds$$

$$= 0.296$$

Generating the Total Sum of $\overline{\epsilon^2_{min.}(t)}$ on an Analog Computer

The simulation equation in this case is given by

$$\overline{\epsilon^2_{(T)min.}(t)} = \sum_{i=1}^{2}\varphi^{ii}_{dd}(0) - \sum_{j=1}^{2}\int_0^t w_{ij}(t)\,\varphi^{ji}_{rd}(t)dt$$

where

$$\varphi^{11}_{dd}(t) = \sqrt{2}\,e^{-2t}\ ;\quad \varphi^{22}_{dd}(t) = e^{-t}$$

$$\varphi^{11}_{rd}(t) = \sqrt{2}\,e^{-2t}\ ;\quad \varphi^{12}_{rd}(t) = -\frac{\sqrt{2}}{3}\,e^{-t}$$

$$\varphi^{21}_{rd}(t) = -\frac{\sqrt{2}}{3}\,e^{-2t}\ ;\quad \varphi^{22}_{rd}(t) = -\frac{1}{2}\,e^{-t}$$

$$w_{11}(t) = (0.49\,e^{-2.54t} - 0.0694\,e^{-1.245t})$$

$$w_{12}(t) = (-0.448\ e^{-2.54t} + 0.071\ e^{-1.245t})$$

$$w_{21}(t) = (-0.0871\ e^{-2.54t} + 0.575\ e^{-1.245t} - 0.943\ e^{-t})$$

$$w_{22}(t) = (0.079\ e^{-2.54t} - 0.588\ e^{-1.245t})$$

The analog program for this problem is shown in Figure 4-4. A summary of the results of this chapter is tabulated in Table 1.

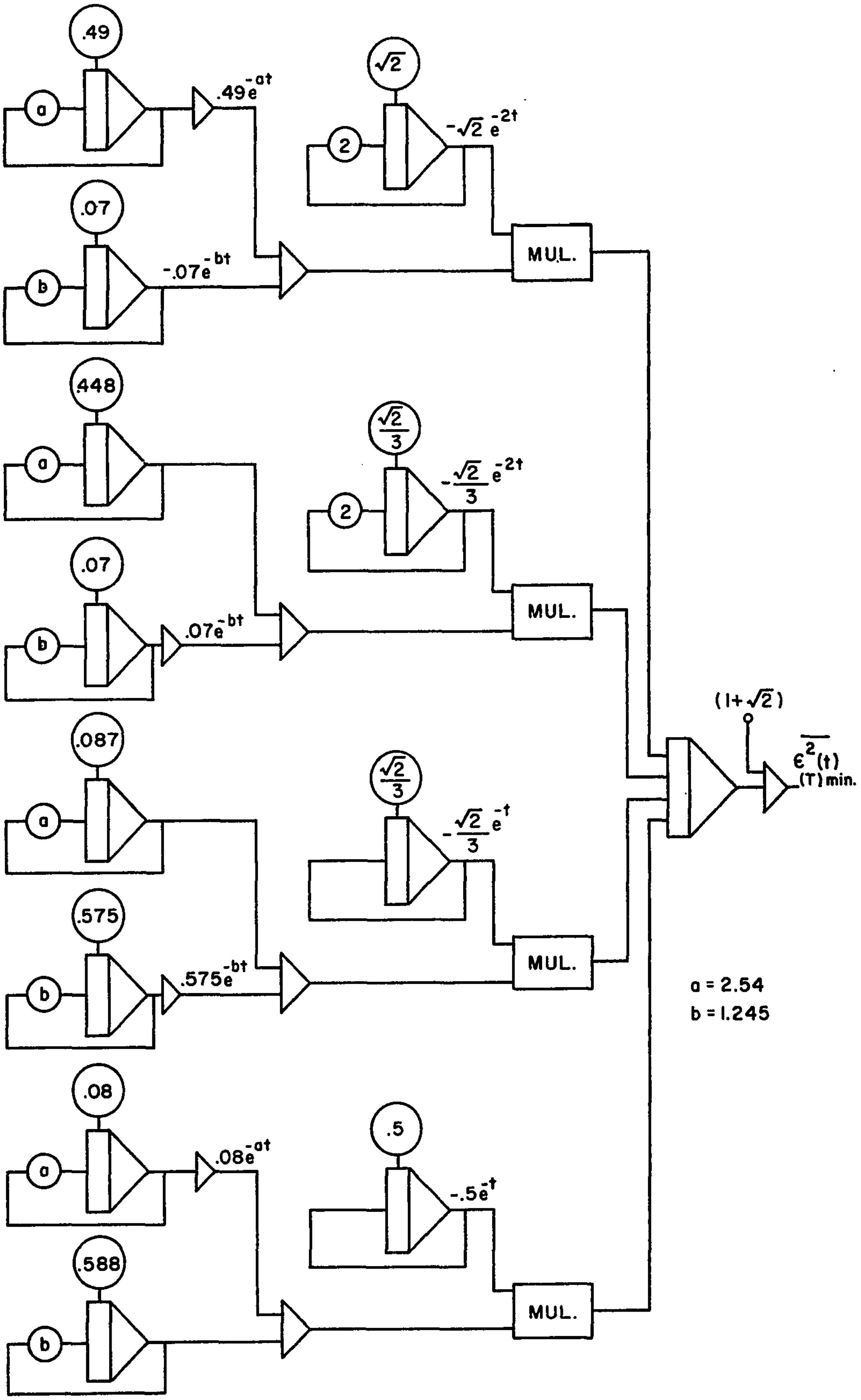

Figure 4-4. Generating $\overline{\epsilon^2_{(T)\min.}(t)}$ on Analog Computer.

TABLE 1

A SUMMARY OF THE RESULTS

			METHOD OF UNDETERMINED COEFFICIENTS	FUNCTIONAL METHOD	PROPOSED ALTERNATIVE METHOD
EXAMPLE NO. 1	ELEMENTS OF SYSTEM MATRIX	$W_{11}(s)$	$\frac{0.411}{s+1.32}$	$\frac{0.411}{s+1.32}$	$\frac{0.411}{s+1.32}$
		$W_{12}(s)$	$\frac{0.545}{s+1.32}$	$\frac{0.544}{s+1.32}$	$\frac{0.545}{s+1.32}$
	$\overline{\epsilon^2_{min.}(t)}$		0.41	0.411	0.411
EXAMPLE NO. 2	ELEMENTS OF SYSTEM MATRIX	$W_{11}(s)$		$\frac{0.494}{s+2.54} - \frac{0.07}{s+1.245}$	$\frac{0.49}{s+2.54} - \frac{0.069}{s+1.245}$
		$W_{12}(s)$		$\frac{-0.448}{s+2.54} + \frac{0.0713}{s+1.245}$	$\frac{-0.448}{s+2.54} + \frac{0.071}{s+1.245}$
		$W_{21}(s)$		$\frac{-0.0876}{s+2.54} + \frac{0.574}{s+1.245} - \frac{0.945}{s+1}$	$\frac{-0.087}{s+2.54} + \frac{0.575}{2+1.245} - \frac{0.943}{s+1}$
		$W_{22}(s)$		$\frac{0.0784}{s+2.54} - \frac{0.586}{s+1.245}$	$\frac{0.079}{s+2.54} - \frac{0.588}{s+1.245}$
	$\overline{\epsilon^2_{min.}(t)}$			0.676	0.7214

CHAPTER V

SUMMARY AND CONCLUSIONS

Through the use of matrix relationships the well-known concepts in linear, time-invariant, single input-single output system synthesis can be extended to solve the design-by-synthesis problems for systems with a multiplicity of inputs and outputs. Because of their practical significance, systems with random inputs are considered. The proposed method of approach, as presented in Chapter III, has basically two parts.

First, the derivation of an implicit solution to the synthesis problem in the form of vector Wiener-Hopf equation. The derivation of this type of equation has been achieved in two different ways. In one method the derivation has been performed strictly in the time domain. This method is an extension of the ideas developed in reference [N.2] concerning the synthesis of optimum linear constant parameter systems with a single input and output. In the alternative method the derivation has been performed strictly in the frequency domain. The frequency domain approach, as presented in Chapter III, is new and has not been cited in the existing literature. A comparison of these two methods shows that transformation from time to frequency domain at the early stages of optimization procedure facilitates algebraically the derivation of the vector Wiener-Hopf equation.

Although a sum of the weighted mean-square errors is used as a performance functional, it has been shown that the implicit solution to the minimization problem is independent of the weighting matrix L.*

The main difficulty in the optimization problem is to find the explicit solution to the vector Wiener-Hopf equation. The second part of the presentation of Chapter III deals with the methods of solution of a vector Wiener-Hopf equation. The method presented in that chapter is a combination of the "functional method" presented in reference [B.4] and the "method of undetermined coefficients" presented in reference [A.2].

Two hypothetical, illustrative examples are treated in Chapter IV. Each problem is solved first by functional method and then by the proposed alternative method developed in this work and presented in Chapter III. It has been shown that the functional method as presented in reference [B.4] may result in system matrix with some unstable poles. Therefore an operation is presented that should be performed on the final system as obtained by the functional method to guarantee the stability of the system poles.** The proposed

---

*See Equation (3-21) or equivalently Equation (3-46).

**See solution to Example 2 for more detail.

method of this work, on the other hand, is developed in such a way that it always results (with less computational effort) in systems with stable poles. The advantage of this method over the one suggested by reference [A.2] is that here it is not necessary to find the "natural" poles by solving the determinant $G_{rr}(s)$ and then add the "forced" poles if there are any. In this technique this is automatically taken care of in the process of determining the unknown coefficients of the probable poles. Vanishing of any of these coefficients means the cancellation of the corresponding poles and hence modification of probable poles into actual poles of the system matrix.* The method results in stable system directly and does not require that the poles of each member of the system matrix be the same as those of all other elements.

None of the existing methods of solution of the vector Wiener-Hopf equation including the one presented in this work are fully adaptable to digital computation unless a specialized program is written that can handle the functional operation properly. The method of generating the minimum mean-square error on an analog computer enables the designer to find the value of the performance functional, once the system matrix is determined.

A summary of the results obtained in Chapter IV are tabulated in Table I. Comparison of these results

---

*See solution to Example 2 for more detail.

reveals some interesting findings. However before proceeding to the discussion of these findings a few words are in order regarding the selection of the particular types of input messages used as representative of random input functions.

Many writers such as Lanning and Battin [L.1] have noted the presence of damped exponential-cosine or exponential processes. Bendat [B.2] studied different mathematical-physical sources and assumptions under which exponential autocorrelation functions occur. The net result of his studies is a strong appreciation for the exponential cosine and exponential autocorrelation function as a means of describing much of the random noise phenomena of interest. Calculation of these correlation functions or their associated spectral density functions are carried out in detail in references [B.2], [L.1], and [S.5]. A summary of the results of these derivations is given in Appendix E. These results are used as a justification for the type of input signals used particularly in the second example.

In going back to the discussion of the results of the examples, the first interesting point is that when the inputs to the system are somehow related to each other, i.e., when one is, for instance, the derivative of the other, etc., the solution matrix will have all its elements with identical, stable-only poles. And furthermore, the numerical value of the minimum mean-square error at the first output is the same as the partial fraction expansion coefficient of the pole of the first element in the system matrix. In this case each element of the system matrix has only one pole; i.e., the interconnecting sub-systems are first order systems.* The functional method without modification will produce the same resulting system matrix as the method developed in this work namely the proposed alternative method. These points are shown in Example I.

However, when the inputs are not related such as in Example II, the elements of system matrix may not have identical poles and the system resulting from the application of functional method must be modified to obtain the desired stable system matrix.** The resulting sub-systems (coupling systems) are of 2nd and 3rd order. As is the case in free configuration type of synthesis, the choice of the system is dictated by the nature of the inputs and the desired outputs subject to the physical realizability condition.

In conclusion it may be added that although the optimization procedure is carried out for free configuration case, the semi-free configuration type problem can be handled by modifying the system matrix W such that it includes the fixed plant matrix by using proper feedback loops.

---

*See Table I.

**See the solution to Example II.

Recommendation for Future Work

In Chapter III the necessary and sufficient conditions for the existence of a strong extremum for a functional were enumerated. It is worthwhile to study and investigate just how close a functional of the form of Equation (3-11) will satisfy the necessary and sufficient conditions mentioned previously. The first step in this direction should be the modification of these conditions since Equation (3-11) is not quite of the same form as Equation (3-26).

In the optimization method presented in this work constant parameter linear systems are used. Time-varying systems do occur in practical problems so often that an attempt to optimize such systems, using perhaps an extension of the present work, should prove worthwhile. There has been some work done along this line. Reference [P.2] gives a good account by using the method which was developed by Shinbrot [S.2] to solve the modified vector Wiener-Hopf equation. However, this method of solution is based on the assumptions that the system impulse response is separable and that the random input functions have a particular type of autocorrelation function. A more general and less restrictive method of solution of the modified vector Wiener-Hopf equation is desirable.

The technique of generating mean-square errors at the output of linear time-varying systems which was developed in reference [M.4] may be used to generate minimum mean-square error when dealing with time-varying systems.

BIBLIOGRAPHY

A.1 Alexandro, F. J., Jr. "Compensation of Plant Variations in Optimal Control Systems," IEEE Trans. on Automatic Control, Vol. AC-10, No. 3, pp. 275-280, July 1965.

A.2 Amara, R. C. "The Linear Least Square Synthesis of Multi-Variable Control Systems," AIEE Trans. on Applications and Industry, pt. II, Vol. 78, pp. 115-120, May 1959.

A.3 Athanas, M. and Flab, P. L. Optimal Control. New York: McGraw-Hill Book Co., 1966.

B.1 Bellman, R. E. Introduction to Matrix Analysis. New York: McGraw-Hill Book Co., 1960.

B.2 Bendat, J. S. Principles and Applications of Random Noise Theory. New York: John Wiley & Sons, Inc., 1958.

B.3 Bootan, R. C., Jr. "An Optimization Theory for Time-Varying Linear Systems with Non-Stationary Statistical Inputs," IRE Proc., Vol. 40, pp. 977-981, August 1952.

B.4 Brockett, R. W. and Mesarovic, M. D. "Synthesis of Multi-Variable Systems," AIEE Trans. on Applications and Industry, pp. 216-221, Sept. 1963.

C.1 Chang, S. S. L. Synthesis of Optimum Control Systems. New York: McGraw-Hill Book Co., 1961.

C.2 Curzy, J. B., Jr. and Perkins, W. R. "A New Approach to the Sensitivity Problem in Multivariable Feedback System Design," IEEE Trans. on Automatic Control, Vol. AC-9, pp. 216-223, July 1964.

D.1 Dorato, P. Discussion of Reference H.1.

E.1 Elsgolts, L. S. Calculus of Variations. Reading, Mass.: Addison-Wesley Publishing Co., 1962.

F.1 Freeman, H. "A Synthesis Method of Multipole Control Systems," AIEE Trans. on Applications and Industry, Vol. 76, pt. II, pp. 28-31, March 1957.

G.1 Gelfand, I. M. and Fomin, S. V. Calculus of Variations. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

G.2 Graham, D. and Lathrop, R. C. "The Synthesis of Optimum Transient Response: Criteria and Standard Forms," AIEE Trans. on Applications and Industry, Vol. 72, pt. II, pp. 273-288, November 1953.

G.21 Graham, D. and Lathrop, R. C. "The Performance of Servomechanisms with Derivative and Integral Control," AIEE Trans. on Applications and Industry, Vol. 73, pt. II, pp. 10-17, March 1954.

G.3 Graham, D. "The Influence of Time Scale and Gain on Criteria for Servomechanisms Performance," AIEE Trans. on Applications and Industry, Vol. 73, pt. II, pp. 153-158, July 1954.

H.1 Hsieh, H. C. "Synthesis of Optimum Multivariable Control Systems by the Method of Steepest Descent," IEEE Trans. on Applications and Industry, Vol. 82, pp. 125-130, May 1963.

H.2 Horowitz, I. M. Synthesis of Feedback Systems. New York: Academic Press, 1963.

J.1 James, H. M., Nicholas, N. B. and Phillips, R. S. Theory of Servomechanism. New York: McGraw-Hill Book Co., 1947.

K.1 Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problem," ASME Trans., Series D, Journal of Basic Engineering, pp. 35-45, March 1960.

K.2 Kaplan, W. Advanced Calculus. Reading, Mass.: Addison-Wesley Publishing Co., Inc., 1952.

K.31 Kavanagh, R. J. "Multivariable Control Systems Synthesis," AIEE Trans. on Applications and Industry, Vol. 77, pp. 425-429, November 1958.

K.32 Kavanagh, R. J. "A Note on Optimum Linear Multivariable Filters," Proc. Institute of Electrical Engineers, London, England, Vol. 108, pt. C, pp. 412-417, September 1961.

K.4 Kipiniak, W. Dynamic Optimization and Control. New York: M.I.T. and John Wiley & Sons, Inc., 1961.

L.1 Lanning, J. H., Jr., and Battin, R. H. Random Process in Automatic Control. New York: McGraw-Hill Book Co., 1956.

L.2 Liu, S. H. and Thaler, G. J. "An Analytical Method for Analysis and Design of Control Systems," IEEE Trans. on Automatic Control, Vol. AC-10, No. 1, pp. 49-59, January 1965.

M.1 Matyas, J. and Silhanek, Y. "The Description of Multidimensional Linear Systems in Matrix Form," Automation and Remote Control, Vol. 22, No. 7, pp. 768-777, July 1961.

M.21 Mesarovic, M. D. The Control of Multivariable Systems. New York: The Technology Press of M.I.T. and John Wiley & Sons, Inc., 1960.

M.22 Mesarovic, M. D. Foundation for a Control Theory of a Multivariable System. New York: John Wiley & Sons, Inc., 1960.

M.3 McBride, L. E., Jr., and Narendra, K. S. "Optimization of Time-Varying Systems," IEEE Trans. on Automatic Control, Vol. AC-10, No. 3, pp. 289-294, July 1965.

M.4 Mofeez, S. M. "Measurement of the Mean-Square Error at the Output of a Certain Class of Linear Systems with Noise Input Using Analog Computer Techniques," M. S. Thesis, The University of Oklahoma, 1966.

M.5 Morgan, B. S. "Multivariable Systems," IEEE International Convention Record, Vol. 13, pt. 6, pp. 87-95, 1965.

N.1 Narendra, K. S. and Goldwyn, R. M. "Application of Matrix Methods to the Optimum Synthesis of Multivariable System Subject to Constraints," AIEE Trans. on Applications and Industry, Vol. 81, pt. II, pp. 151-157, July 1962.

N.2 Newton, G. C., Jr., Gould, L. A. and Kaiser, J. K. Analytical Design of Linear Feedback Controls. New York: John Wiley & Sons, Inc., 1957.

P.1 Papoulis, A. The Fourier Integral and Its Applications. New York: McGraw-Hill Book Co., 1962.

P.2 Peterson, E. L. Statistical Analysis of Optimization of Systems. New York: John Wiley and Sons, Inc., 1961.

P.3 Povejsil, D. J. and Fucls, A. M. "A Method for Preliminary Synthesis of a Complex Multiple Control System," AIEE Trans. on Applications and Industry, Vol. 74, pt. II, pp. 129-134, July 1955.

S.1 Sherman, S. "Non-Mean Square Error Criteria," IRE Trans. on Information Theory, Vol. 1T-4, No. 3, pp. 125-126, September 1959.

S.2 Shinbrot, M. "A Generalization of a Method for the Solution of Integral Equations Arising in Optimization of Time-Varying Linear Systems with Nonstationary Inputs," IRE Trans. on Information Theory, Vol. 1T-3, No. 4, pp. 220-224, December 1957.

S.3 Shinners, S. M. Control System Design. New York: John Wiley & Sons, Inc., 1964.

S.4 Sokolnikoff, I. S. and Reaheffer, R. M. Mathematics of Physics and Modern Engineering. New York: McGraw-Hill Book Co., 1958.

S.5 Solodnikov, V. V. Introduction to the Statistical Dynamics of Automatic Control Systems. New York: Dover Publications, Inc., 1960.

T.1 Thiede, E. C. and Stubberud, A. R. "Control System Synthesis Using Linear Transformation," IEEE Trans. on Automatic Control, Vol. AC-10, No. 2, pp. 172-177, April 1965.

T.2 Tou, J. Modern Control Theory. New York: McGraw-Hill Book Co., 1964.

T.3 Tyler, J. S., Jr., and Tuteur, F. B. "The Use of a Quadratic Performance Index to Design Multivariable Control System," IEEE Trans. on Automatic Control, Vol. AC-11, No. 1, pp. 84-92, 1966.

W.1 Wiener, N. The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. New York: John Wiley & Sons, Inc., 1949.

W.2 Wong, E. and Thomas, J. B. "On the Multidimensional Prediction and Filtering Problem and the Factorization of Spectral Matrices," Journal, The Franklin Institute, Vol. B, pp. 87-99, August 1961.

Y.1 Youla, D. C. "The Factorization of Rational Matrices," IRE Trans. on Information Theory, Vol. IT-7, No. 3, pp. 172-189, July 1961.

Z.1 Zadeh, L. A. and Desoer, C. A. Linear System Theory. New York: McGraw-Hill Book Co., 1963.

APPENDICES

APPENDIX A

EXPANSION OF EQUATIONS (3-14)

Equation (3-14) may be written in the following equivalent form with the averaging operation understood.

$$\underset{r^T(t-\tau)}{[r_{11}\cdots r_{j1}\cdots r_{k1}]} * \underset{w_m^T(\tau)}{\begin{bmatrix} w_{11}\cdots w_{k1}\cdots w_{q1} \\ \\ w_{1j}\cdots w_{kj}\cdots w_{qj} \\ \\ w_{1k}\cdots w_{kp}\cdots w_{qp} \end{bmatrix}} \underset{L}{\begin{bmatrix} l_{11} & & \\ & l_{kk} & \\ & & l_{qq} \end{bmatrix}} \times$$

$$\underset{h(\sigma)}{\begin{bmatrix} h_{11}\cdots h_{1i}\cdots h_{1p} \\ h_{k1}\cdots h_{ki}\cdots h_{kp} \\ h_{q1}\cdots h_{qi}\cdots h_{qp} \end{bmatrix}} * \underset{r(t-\sigma)}{\begin{bmatrix} r_{11} \\ r_{i1} \\ r_{k1} \end{bmatrix}}$$

$$+\underset{r^T(t-\tau)}{[r_{11}\cdots r_{i1}\cdots r_{k1}]} * \underset{h^T(\tau)}{\begin{bmatrix} h_{11}\cdots h_{k1}\cdots h_{q1} \\ \\ h_{1i}\cdots h_{ki}\cdots h_{qi} \\ \\ h_{1p}\cdots h_{kp}\cdots h_{qp} \end{bmatrix}} \underset{L}{\begin{bmatrix} l_{11} & & \\ & l_{kk} & \\ & & l_{qq} \end{bmatrix}} \times$$

$$
\begin{array}{cc}
w_m(\sigma) & r(t-\sigma) \\
\begin{bmatrix} w_{11}\cdots w_{1j}\cdots w_{1p} \\ w_{k1}\cdots w_{kj}\cdots w_{kp} \\ w_{q1}\cdots w_{qj}\cdots w_{qp} \end{bmatrix} * & \begin{bmatrix} r_{11} \\ r_{j1} \\ r_{p1} \end{bmatrix}
\end{array}
$$

$$
\begin{array}{cc}
r^T(t-\tau) & hT(\tau) \\
-2[r_{11}\cdots r_{i1}\cdots r_{p1}]* & \begin{bmatrix} h_{11}\cdots h_{k1}\cdots h_{q1} \\ h_{1i}\cdots h_{ki}\cdots h_{qi} \\ h_{1p}\cdots h_{kp}\cdots h_{qp} \end{bmatrix} \quad x
\end{array}
$$

$$
\begin{array}{cc}
L & d(t) \\
\begin{bmatrix} l_{11} & & \\ & l_{kk} & \\ & & l_{qq} \end{bmatrix} & \begin{bmatrix} d_{11} \\ d_{k1} \\ d_{q1} \end{bmatrix} = 0
\end{array} \qquad \text{(A-1)}
$$

Now since each term on the LHS of Equation (A-1) is a scalar it may be written as

$$\overline{r_{jl}(t-\tau)*w_{kj}(\tau)l_{kk}h_{ki}(\sigma)*r_{il}(t-\sigma)}$$

$$+\ \overline{r_{il}(t-\tau)*h_{ki}(\tau)l_{kk}w_{kj}(\sigma)*r_{jl}(t-\sigma)}$$

$$-\ 2\overline{r_{il}(t-\tau)*h_{ki}(\tau)l_{kk}d_{kl}(t)} = 0 \qquad \text{(A-2)}$$

which is Equation (3-15) with summation on repeated indices understood.

APPENDIX B

DERIVATION OF EQUATION (3-44)

Substitute Equation (3-43) in Equation (3-42) to get

$$P = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr.} \Big[L\, G_{dd}(s) - 2L[\bar{W}_m(s) + \alpha\bar{H}(s)]G_{rd}(s)$$

$$+ L[\bar{W}_m(s) + \alpha\bar{H}(s)]G_{rr}(s)\ [W_m^T(s) + \alpha H^T(s)] \Big]\, ds \qquad \text{(B-1)}$$

Differentiation of Equation (B-1) with respect to α gives

$$\frac{dP}{d\alpha} = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr.} \Big[-2\, L\, \bar{H}(s)\ G_{rd}(s)$$

$$+ L\, \bar{H}(s)\ G_{rr}(s)\ [W_m^T(s) + \alpha H^T(s)] \qquad \text{(B-2)}$$

$$+ L[\bar{W}_m(s) + \alpha\bar{H}(s)]G_{rr}(s)\ H^T(s) \Big]\, ds$$

Set Equation (B-2), evaluated at $\alpha = 0$, equal to zero to get

$$\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \text{tr.} \Big[[L\, \bar{H}(s) G_{rr}(s)\ W_m^T(s) \qquad \text{(B-3)}$$

$$+ L\, \bar{W}_m(s) G_{rr}(s)\ H^T(s) - 2\, L\, \bar{H}(s) G_{rd}(s) \Big]\ d(s) = 0$$

which is the desired result; i.e., Equation (3-44).

APPENDIX C

STEPS INVOLVED IN FACTORIZATION OF A HERMITIAN MATRIX

The problem considered in this appendix can be formulated as follows:

Given a rational, Hermitian matrix A, find P' and P such that A = P' P and P be stable.

This problem has been treated in the literature by several different authors such as Youla [Y.1], Kavanagh [K.32], and Brockett-Mesarovic [B.4]. The latter authors presented a simplified version of the work done by Youla. The summary of their presentation is as follows:

Step 1. Find the lowest common denominator of matrix A and call it g.

Step 2. Define a new matrix G such that

$$G = g\,A$$

with the result that elements of G, $g_{ij}(s)$, all become polynomials of s.

Step 3. Factor g into f' and f such that

$$g = f'f$$

where $f'(s) = f(-s) = \bar{f}(s)$.

Step 4. Construct an upper triangular matrix F, with f_{ij} as elements, such that

$$f_{11}(s) = \left(g_{11}(s)\right)^{-}$$

$$f_{1j}(s) = \frac{g_{1j}(s)}{\left(g_{11}(s)\right)^{+}} \qquad j > 1$$

$$f_{ij}(s) = 0 \qquad i > j$$

$$f_{ii}(s) = \left(g_{ii}(s) - \sum_{j=1}^{i-1} f_{ji}(s)\, \bar{f}_{ji}(s)\right)^{+} \qquad i > 1$$

$$f_{ni}(s) = \frac{g_{ni}(s) - \sum_{j=1}^{n-1} f_{ji}(s)\, \bar{f}_{jn}(s)}{\bar{f}_{nn}(s)} \qquad i > n$$

Where $(.....)^{-}$ denotes the factors of the argument which contains all the right half-plane poles and zeroes and $(.....)^{+}$ means the factors of the argument which contains all the left half-plane poles and zeroes.

Step 5. Define

$$P = \frac{1}{f(s)} F(s)$$

and

$$P' = \bar{P}^T$$

This completes the factorization steps.

APPENDIX D

STEPS INVOLVED IN USING FUNCTIONAL METHOD OF SOLUTION

The procedure for solving an equation of the form $[P'PX]_+ = [P'D]_+$ where P and D are stable is summarized. Originally only six steps were suggested in the literature. Step 7 has been added to include calculation of $\overline{\epsilon^2_{min.}(t)}$ in each case. At some points corrections have been made to take care of the misprints and/or mistakes that appeared in the original paper [B.4].

Steps suggested are as follows:

Step 1. Find the right half-plane (non-minimum phase) zeroes of the determinantial equation $|P(s)| = 0$ and label them z_K. Assume there are m such zeroes. Define z_{jK} as $(z_j + z_K)^{-1}$ and form the m x m matrix Z whose elements are z_{jK}.

Step 2. Compute any nonzero row of the adjoint of $P(z_K)$ and label this row vector B^K. Form the m x n (not n x m as suggested in the original paper) B matrix whose Kth row is the row vector B^K.

Step 3. Compute the m x m matrix BB^T and call it F. Define H as the Schur product [B.1] of matrices F and Z; i.e., let $h_{jK} = (f_{jK} z_{jK})$. Compute H^{-1}.

Step 4. Compute the inner product $B^K D_i(z_K)$ and label it v_{Ki}. The column vector whose components are v_{Ki} is written as V_K.

Step 5. Define U_K as $H^{-1}V_K$ and define Γ_K to be a diagonal matrix whose diagonal terms γ_{ii} are equal to the u_i of the vector U_K and whose off-diagonal terms are zero. Define R_K to be $B^T\Gamma_K$.

Step 6. Define a column vector S whose ith component is $(s+z_i)^{-1}$. If Y is defined as D-RS then X is given by $P^{-1}Y$ and the mean-square error, q, is given by U^TV.

At this point the steps suggested by reference [B.4] are exhausted. However, if one requires that the components of X matrix have stable poles only then X should be calculated from the following equation

$$[X]_+ = [P^{-1}Y]_+$$

since the fact that P has stable poles does not guarantee that its inverse be stable. To determine the individual errors, q_i, one must use Equation (25) of the reference [B.4] instead of Equation (39) which is suggested by the authors of that paper, i.e.

$$q_i = R^i Z (R^i)^T$$

where R^i is the ith row of matrix R.

In discussing the filter design problem the authors of reference [B.4] denoted the autocorrelation of the desired output as c. They later defined q_o, in Equation (20), as

$$q_o = \frac{1}{2\pi i} \int_{-i\infty}^{+i\infty} (c'c - D'D)ds$$

This seems confusing. To avoid this confusion a change of notation was made as follows:

Let d(s) be a column vector corresponding to the Fourier transform of c(t) then define

$$q_o = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} (d'd - D'D)ds$$

$$= \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \{tr. [G_{dd}] - D'D\}ds$$

where $G_{dd}(s)$ is the spectral density matrix corresponding to $\varphi_{cc}(t)$.

Step 7. Define

$$\overline{\varepsilon^2_{min.}(t)} = q_o + q$$

where

$$q_o = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \{tr.[G_{dd}(s)] - D'D\}\, ds$$

and

$$q = U^T V$$

APPENDIX E

SOME SPECTRAL DENSITIES OF WELL-KNOWN RANDOM SIGNALS

Spectral densities of some of the well-known random signals are mentioned here. For details of derivations the readers are referred to the references [B2], [L1], and [S5].

1 - Telegraph Signals

$$G_{rr}(s) = \frac{4a^2\mu}{4\mu^2 - s^2}$$

where:

μ = average number of sign reversal per unit time

a = constant amplitude of the signal

2 - Typical Signal at the Input of a Servomechanism

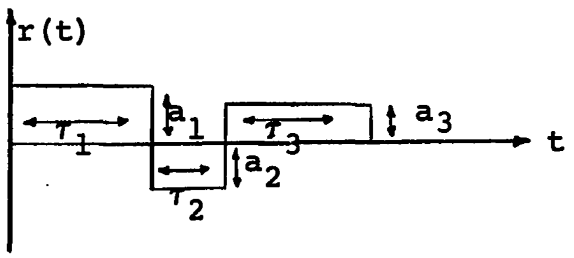

Figure E-1. A Square Wave with Random Width and Amplitude

$$G_{rr}(s) = \overline{a^2} \frac{2\mu}{\mu^2 - S^2}$$

where:

$r(t)$ = a stationary random input

$$\mu = \frac{1}{\tau}$$

$$a_n^2 = \overline{r(t)\, r(t+\tau)}$$

3 - Thermal Noise in Electric Circuits

a) <u>At the Output of an L-R Circuit</u>

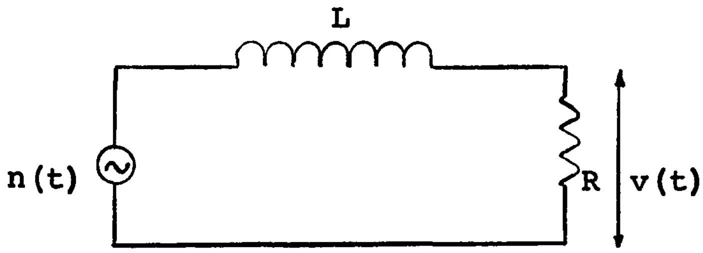

Figure E-2. An R-L Network

$$G_{vv}(s) = \frac{\beta^2 K}{\beta^2 - S^2}$$

where

$$\beta = \frac{R}{L}$$

$$K = G_{nn}(s)$$

b) <u>At the Output of An L-R-C Circuit</u>

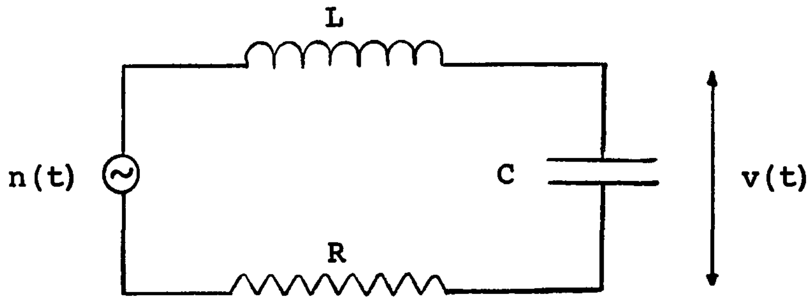

Figure E-3. An L-R-C Network

$$G_{vv}(s) = \frac{Ka^4}{(a^2+s^2)^2-4b^2s^2}$$

where

$$a^2 = \frac{1}{LC}$$

$$2b = R/L$$

$$K = G_{nn}(s)$$

4 - Random Impact of Particles in Brownian Motion

Let x(t) be a random variable with its mean value equal to zero and its variance equal to σ^2 then

$$\sigma^2 = C\lambda$$

where λ is uniform average density of point distribution along time axis and C is a constant. For this type of input

$$G_{xx}(s) = C$$

5 - Atmospheric Turbulence Phenomena

$$G_{rr}(s) = Ak\left[\frac{(k^2+c^2) - s^2}{s^4-2(k^2+c^2)s^2+(k^2+c^2)^2}\right]$$

where A, k, and c are some constant.

APPENDIX F

NOMENCLATURE

A	= A spectral density matrix (40), a constant (32)*
A'_{ij}, A_{ij}	= Constant coefficients (81)
$A_{(p)}(s)$	= Adjoint of matrix P(s) (53)
a	= A constant (32)
a_i, a_{ij}	= Coefficients of expansion (15,59)
B	= A matrix (53), a constant (32)
B^K	= A raw vector (55)
B'_{ij}, B_{ij}	= Constant coefficients (82)
b	= A constant (32)
C	= A closed curve (15) spectral matrix (40) capacitor (114)
C_K	= A column vector of matrix C (40)
$C(s)$	= Laplace transform of c(t) (18)
c	= A constant (115)
$c(t)$	= Total output matrix (22)
$c_g(t)$	= Total output of gth terminal (3)
$c_{ij}(t)$	= Output at terminal j due to ith input (4)
$c_{ij}(s)$	= Laplace transform of $c_{ij}(t)$ (17)

*The numbers in parentheses at the end of each definition refer to the page number where the symbol was defined or first appeared.

D	= A domain (15), a matrix (41)
D_K	= A column vector (42)
$d(t)$	= Desired output column vector (22)
$\frac{d_i(t)}{2}$	= ith components of d(t) (5)
$\epsilon_{min.}(t)$	= Minimum mean-square error (30)
e	= Base of Naperian logarithms, 2.71812 (17)
F	= A matrix (15)
$F(s,y,y')$	= A function of x, y, y' (12)
f_1*f_2	= Convolution of f_1 and f_2 (26)
$f(z)$	= A function of complex variable z (15)
G	= A matrix (45)
$G(s)_{\epsilon\epsilon}$	= An error spectral density matrix (34)
$g(z)$	= A function of complex variable z (16)
H	= A matrix (54)
H^{-1}	= Inverse of H (54)
$h(t)$	= An arbitrary matrix (27)
$h(x)$	= An arbitrary function of x (13)
$h'(x)$	= Derivative of h(x) with respect to x (14)
$h_{ij}(t)$	= Element of h(t) (28)
h_{ij}	= Element of H (54)
i	= An integer used as subscript or superscript (3)
j	= An integer used as subscript or superscript (3) an imaginary number $\sqrt{-1}$ (15)
k	= An integer used as subscript or superscript (28) a constant (115)
L	= A diagonal matrix (25), inductance (113)

LHS = Left hand side (103)

l_{KK} = A diagonal element of L (28)

m = An integer (53)

m_i = Input message (47)

N = An integer (15)

n_i = Noise input (47)

P = A performance functional (25), a matrix with stable poles (41)

P' = Transposed - conjugate of P matrix (41)

P^{-1} = Inverse of P matrix (42)

P_o, P_1 = Fixed points with coordinate x_o, y_o and x_1, y_1 respectively (12)

p = An integer used as subscript (3)

q = An integer used as subscript (3) mean-square error (57)

q_{10} = Mean-square error modifier (57)

R = Radius (16), resistance (113), a matrix (56)

RHS = Right hand side (104)

R(s) = Laplace transform of r(t) (18)

r(t) = A column input vector (22)

$r_i(t)$ = ith component of r(t) (3)

$r_j(s)$ = jth component of R(s) (17)

$\widetilde{r_i(t) r_i(t+\tau)}$ = Ensemble average (18)

$\overline{r_i(t) r_i(t+\tau)}$ = Time average (19)

S = A column vector (56)

s = Laplace transform variable (17)

t = Time (7)

tr. = Trace of a matrix (25)

U_l = A column vector (55)

V_j = A column vector (55)

v_{ij} = ith component of vector V_j (55)

$W(s)$ = Laplace transform of $W(t)$ (18)

$w_{ij}(s)$ = Element of $W(s)$ (17)

$w(t)$ = Matrix of weighting functions (22)

$w_m(t)$ = Optimum weighting function matrix (27)

$w_{ij}(t)$ = Element of $w(t)$ (3)

X = A matrix with unknown elements (40)

X_K = kth column of X (40)

x = An independent variable (12)

$k_{x(t)}$ = A representative of an ensemble of random input (114)

Y_K = A column vector (56)

y = A dependent variable (12)

y' = Derivative of y (13)

$\bar{y}$ = A family of curves (13)

Z = A square matrix (53)

z = A complex variable (15)

z_i = Zeroes of a polynomial (53)

Greek Letters

α = A parameter (13)

β = A constant defined by R/L. (113)

Γ = Lagrange multiplier matrix (55)

$\varepsilon(t)$ = Error function (23)

ε = A small positive number (14)

$\overline{\varepsilon^2(t)}$ = Mean-square error (57)

θ = An angular variable (16)

λ = Uniform average density (114)

μ = Average number of sign reversal per unit time (112)

π = 3.1416 (15)

σ = Time variable (26)

σ^2 = Variance (114)

τ = Time variable (7)

$\varphi(x)$ = A function of x (14)

$\varphi_{\epsilon\epsilon}(\tau)$ = General error matrix (24)

$\varphi_{rr}^{ii}(t,\tau)$ = Autocorrelation of $r_i(t)$ when it is non-stationary random time function (18)

$\varphi_{rr}^{ii}(\tau)$ = Autocorrelation of $r_i(t)$, a stationary random time function (19)

φ_{rr} = Matrix of auto and cross correlation functions (20)

$\Phi_{rr}(s)$ = Laplace transform of $\varphi_{rr}(\tau)$ (20)

$\Phi_{rr}^{ii}(s)$ = Diagonal elements of $\Phi_{rr}(s)$ (19)

$\varphi_{rd}(\tau)$ = Crosscorrelation matrix (20)

$\varphi_{rd}^{ij}(\tau)$ = General element of $\varphi_{rd}(\tau)$ (19)

$\Phi_{rd}(s)$ = Laplace transform of $\varphi_{rd}(\tau)$ (20)