Page 222

STATE VARIABLES, THE FREDHOLM THEORY AND OPTIMAL COMMUNICATIONS

by

ARTHUR BERNARD BAGGEROER

B. S. E. E., Purdue University
(1963)

S. M., Massachusetts Institute of Technology
(1965)

E. E., Massachusetts Institute of Technology
(1965)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
January, 1968

Signature of Author-_____
Department of Electrical Engineering, January 15, 1968

Certified by_____
Thesis Supervisor

Accepted by_____
Chairman Departmental Committee on Graduate Studies

STATE VARIABLES, THE FREDHOLM THEORY AND OPTIMAL COMMUNICATIONS

by

ARTHUR BERNARD BAGGEROER

Submitted to the Department of Electrical Engineering on January 15, 1968 in partial fulfillment of the requirements for the degree of Doctor of Science.

ABSTRACT

This thesis is concerned with the use of state variable techniques for solving Fredholm integral equations, and the application of the resulting theory to several optimal communications problems. The material may be divided into the following areas;

(i) the solution of homogeneous and nonhomogeneous Fredholm integral equations;
(ii) optimal signal design for additive colored noise channels;
(iii) optimal smoothing and filtering with delay;
(iv) smoothing and filtering for nonlinear modulation systems;
(v) estimation theory for a distributed environment.

A method for solving Fredholm integral equations of the second kind by state variable techniques is derived. The principal advantage of this method is that it leads to effective computer algorithms for calculating numerical solutions. The only assumptions that are made are: (a) the kernel of the integral equation is the covariance function of a random process; (b) this random process is the output of a linear system having a white noise input; (c) this linear system has a finite dimensional state-variable description of its input-output relationship.

Both the homogeneous and nonhomogeneous integral equations are reduced to two linear first-vector differential equations plus an associated set of boundary conditions. The coefficients of these differential equations follow directly from the matrices that describe the linear system. In the case of the homogeneous integral equation, the eigenvalues are found to be the solutions to the transcendental equation. The eigenfunctions also follow directly.

In addition, the Fredholm determinant function is related to the transcendental equation for the eigenvalues. For the nonhomogeneous equation, the vector differential equations are identical to those that have appeared in the literature for

optimal smoothing. The methods for solving these equations are discussed with particular consideration given to numerical procedures. In both types of equations, several analytical and numerical examples are presented.

The results for the nonhomogeneous equation are then applied to the problem of signal design for additive colored noise channels. Pontryagin's Principle is used to derive a set of necessary conditions for the optimal signal when both its energy and bandwith are constrained. These conditions are then used to devise a computer algorithm to effect the design. Two numerical examples of the technique are presented.

The nonhomogeneous Fredholm results are applied to deriving a structured approach to the optimal smoothing problem. By starting with the finite time Wiener-Hopf equation we are able to find the estimator structure. The smoother results are then used to find the filter realizable with the delay. The performance of both of these estimators are extensively discussed and illustrated.

The methods for deriving the Fredholm theory results are extended so as to be able to treat nonlinear modulation systems. The smoother equations and an approximate realization of the realizable filter are derived for these systems.

Finally, an approach to estimation on a distributed medium is introduced. The smoother structure when pure delay enters the observation method is derived, and a simple case is illustrated.

Thesis Supervisor: Harry L. Van Trees
Title Associate Professor of Electrical Engineering

ACKNOWLEDGEMENT

I wish to acknowledge and express my sincerest thanks to Professor Harry L. Van Trees for his patient supervision of my doctoral studies. His constant encouragement and interest provided a continual source of inspiration for which I cannot be too grateful.

I also wish to thank Professors W. B. Davenport and W. M. Siebert for their interest as the readers of my thesis committee.

Several people have made valuable contributions to this thesis. I have profited greatly from my association with my fellow student Lewis D. Collins and Professor Donald L. Snyder. Discussions with Theodore J. Cruise have clarified several issues of the material.

Earlier drafts of much of the material were typed by Miss Camille Tortorici, while the final manuscript was done by Mrs. Vera Conwicke. Their efforts are sincerely appreciated.

The work was supported in part by the National Science Foundation, a U. S. Navy Purchasing Contract and by the Joint Service Electronics Program. I am grateful for this support. The computations were performed at the M. I. T. Computation Center.

Finally, I thank my wife Carol for both her encouragement and her patience in the pursuit of my graduate education.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

One of the more powerful techniques used in the analysis of communication problems is the Fredholm integral equation theory.[1, 2, 3, 4] Often, however, this theory is difficult to use because solution methods are either too tedious for analytic procedures or too awkward for convenient implementation on a digital computer.

In recent years, state variable techniques have become increasingly useful, especially in optimal control theory.[5, 6, 7] This is primarily due to their adaptability to computational approaches. In this thesis we shall develop a state variable theory for solving Fredholm integral equations. We shall then apply this theory to several problems in optimal communications.

State variable methods have already provided solutions to several important problems in communication theory. Undoubtably the most significant of these is the original work of Kalman and Bucy on linear filtering theory.[8] Starting with this work, many people have used these techniques for the detection and estimation of random processes.[9, 10]

The major advantage that these techniques offer is that they lead to solutions which are readily implemented on a digital computer. Their essential aspect is that the systems or random processes which are involved are represented in terms of differential equations rather than by impulse responses and covariance functions. Since the digital computer is ideally suited for integrating differential equations, we can see how this type of formulation leads to convenient computational methods of solutions.

There is a second important advantage of state variable techniques. Historically, the concept of state found its first application in optimal control theory. Over the years, control theorists have developed a vast literature pertaining to state variable methods. As a result, in using a state variable approach to our problems, we can expropriate many of the methods that have been developed in this area.

The application of the Fredholm integral equation theory to communications is certainly well known. The homogeneous integral equation, with its eigenfunctions and eigenvalues, is probably most familiar in the context of a Karhunen-Loeve expansion of a random process.[11] Since this expansion theory is often the starting point in the analysis of a particular problem, it is easy to see why we are interested in being able to determine the eigenfunctions and eigenvalues for this equation. Similarly, the

nonhomogeneous equation is often encountered. Its solution specifies the optimal receiver and its performance for detecting a known signal in additive colored noise. [3,4] Other applications of it include the solution to the finite time Wiener-Hopf equation and accuracy bounds for parameter estimation.

The major difficulty in using the Fredholm theory is in obtaining solutions to these equations. With the exception of a limited number of cases, finding analytic solutions is difficult at best, while current numerical methods tend to use a large amount of computer time. This is where major contribution of this thesis lies. For a wide class of Fredholm equations of interest we shall apply state variable methods to the problem of finding solutions to these equations. Because of the inherent computational advantages that these techniques offer, we shall be able to devise a solution algorithm that is both well suited and efficient for implementing on a digital computer.

We shall find, however, that our solutions are of more general interest than for just solving these integral equations. We shall apply our results to several problems in communications.

By coupling our method for solving the nonhomogeneous equation with the Minimal Principle of optimal control theory, we shall be able to formulate and solve a signal design problem for additive colored noise channels where bandwidth and energy constraints are imposed. [13,6]

We can also use the nonhomogeneous equation to derive the state variable form of the optimal smoother.[14, 15, 16] With the smoother equations we shall be able to find a structure for the filter realizable with a delay. We shall also be able to analyze the performance of both structures.

We shall also find that the technique we used in solving the Fredholm equations may be used to solve new problems. By extending these techniques, we shall be able to find the smoother equations for nonlinear modulation systems. Again borrowing results from control theory, we shall also derive an approximation to the realizable filter from the smoother structure for this problem.[17]

The final topic that we shall treat is apparently quite divorced from the Fredholm theory. This is due only to our approach. In processing array data, delay factors often enter the signals. In this topic we shall use a variational approach to derive the smoother structure when pure delay enters our observation process. We use a variational approach for simplicity. It is possible to develop and use a Fredholm theory approach; however, this leads to a significantly longer derivation.

Before proceeding let us outline the sequence of material in the thesis. In Chapter 2 we shall introduce the concepts for describing random processes by state variable methods. We shall also derive an important result that is used throughout

the thesis. In Chapter 3 we shall consider the solution of the homogeneous Fredholm integral equation. We shall derive a transcendental equation for the eigenvalues, and then determine the eigenfunctions. We shall also show how to calculate the Fredholm determinant. [2, 4] In Chapter 4, we shall consider the nonhomogeneous equation solution. We shall derive a set of differential equations that specify its solution. Then we shall present several solution methods which exist for solving this particular set of equations.

In the remainder of the thesis we shall apply the results of Chapters 2-4. In Chapter 5 we consider optimal signal design for detection in additive colored noise channels. In Chapter 6, we present an extensive discussion of linear smoothing and filtering with delay. In Chapter 7, we extend our results to treat smoothing and filtering for nonlinear modulation systems, while in Chapter 8 we present an approach to estimation theory when delays occur in the observation method.

We shall present many examples. We do this for two reasons. We shall work a number of analytic examples to illustrate the use of the methods we derive. We shall also present a number of examples analyzed by numerical methods. In the course of the thesis we shall emphasize the numerical aspects of our methods. We feel this is where the major application of much of the material lies. Most problems are too complex to be

analyzed analytically, so finding effective numerical procedures is a very relevant problem.

We also want to indicate our notational conventions. Generally, scalars are lower case symbols which are not underscored; vectors are lower case symbols which are underscored; and matrices are upper case symbols which are not underscored.

CHAPTER II

STATE VARIABLE RANDOM PROCESSES

In this chapter we shall introduce some of the concepts and properties of state variable random processes that we shall need. First, we shall review the ideas of the description and generation of random processes using state variable methods. Then we shall develop some of the properties of the second order moments of these processes. We shall also introduce two processes which we use in many of our examples. Finally, we shall present a derivation which is common to many of the problems that we shall analyze.

A. Generation of State Variable Random Processes

In this section we shall briefly review some of the concepts associated with the description and generation of random processes using state variable methods. This is done principally to establish our notation conventions and terminology. For a more detailed discussion we refer to references 3 and 8.

In the application of the state variable methods to communication theory problems, the random processes of interest are usually characterized as being generated by a dynamical system that is excited by a white noise process. Consequently, the relevant information, which must somehow be provided, is the

equations describing the operation of the dynamical system and a description of the white noise excitation rather than the probability density(ies) or moments of the processes. It is this point of view that we shall assume regarding the description of our random processes.

This is not a very restrictive assumption, as we can generate a large class of processes of interest. In particular, we can generate the important class of stationary processes with rational spectra quite conveniently using constant parameter, linear dynamical systems.

The majority of the processes that we shall discuss are generated by a system whose dynamics may be described in terms of a finite dimensional, linear, ordinary differential equation, termed the state equation,

$$\frac{d\underline{x}(t)}{dt} = F(t)\underline{x}(t) + G(t)\underline{u}(t), \quad \text{(linear state equation)}, \tag{2.1}$$

where

$\underline{x}(t)$ is the state variable vector $(n \times 1)$,

$\underline{u}(t)$ is the white excitation process $(m \times 1)$,

$F(t)$ $(n \times n)$ and $G(t)$ $(n \times m)$ are matrices that determine the system dynamics.

(For notational simplicity, we shall work with continuous time processes. In the study of processes generated by a nonlinear dynamical system, this introduces some attendant mathematical difficulties; however, we shall not discuss them here.[10]) In general,

we shall assume that $\underline{u}(t)$ is white, i.e., it may be interpreted as the derivative of an independent increment process. Consequently, we have (assuming zero mean)

$$E[\underline{u}(t)\underline{u}^T(\tau)] = Q\delta(t-\tau). \tag{2.2}$$

In order to describe a state variable equation completely, the initial state of the system must be considered. We are concerned with representing a random process over the time interval $T_o \leq t \leq T_f$. We shall assume that the initial state $\underline{x}(T_o)$ is a zero mean random vector with a covariance matrix given by

$$E[\underline{x}(T_o)\underline{x}^T(T_o)] = P_o. \tag{2.3}$$

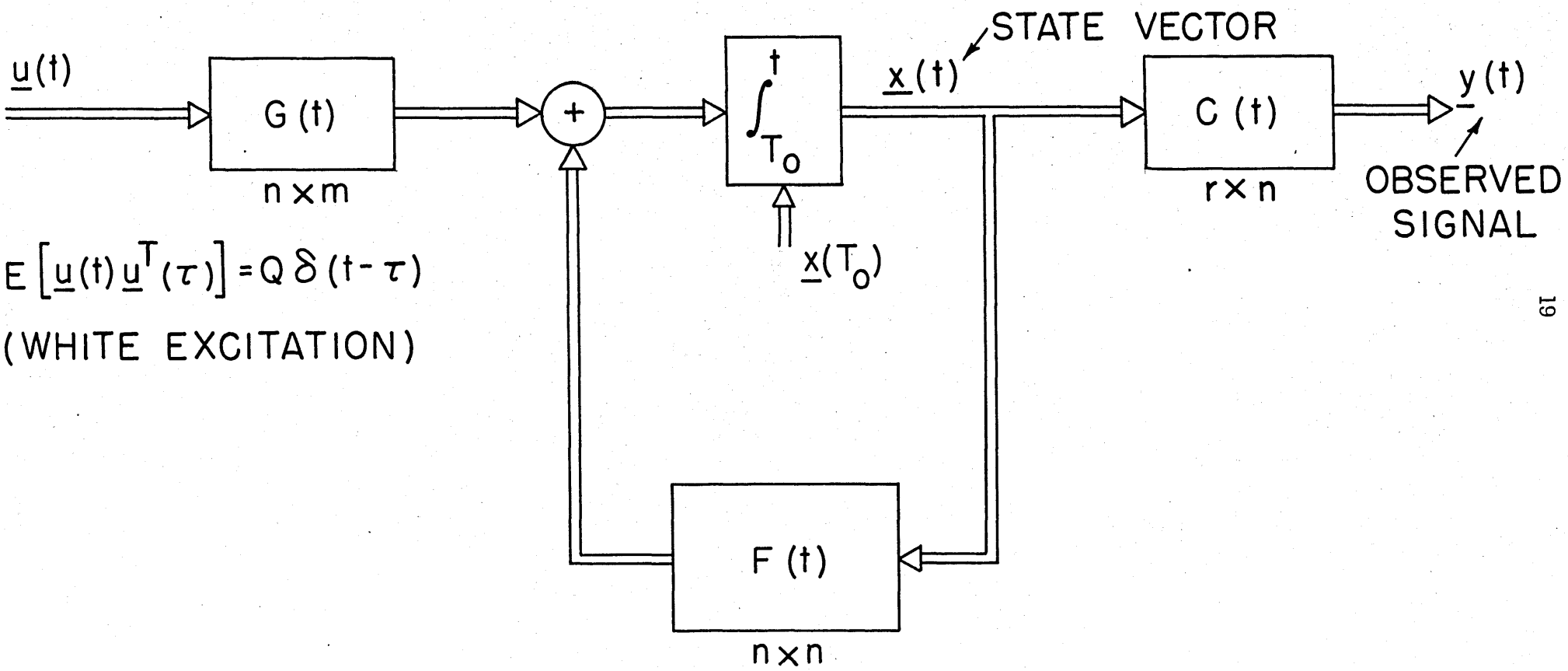
In the case of a deterministic input signal and deterministic initial conditions, knowledge of $\underline{x}(T_o)$ and $u(\tau)$ for $T_o \leq \tau \leq t$ is sufficient to determine $\underline{x}(t)$ for all t. With a random input and/or random initial conditions we can determine the covariance matrix of the state vector,

$$K_{\underline{x}}(t,\tau) = E[\underline{x}(t)\underline{x}^T(\tau)], \tag{2.4}$$

for all t and τ greater than T_o. We should note that this is true only when the state representation is linear as we have assumed in Eq. 2.1.

STATE EQUATION MODEL FOR GENERATING $K_{\underline{y}}(t,\tau)$

$$K_{\underline{y}}(t,\tau) = E\left\{\underline{y}(t)\,\underline{y}^{T}(\tau)\right\}$$



$$E\left[\underline{u}(t)\,\underline{u}^{T}(\tau)\right] = Q\,\delta(t-\tau)$$

(WHITE EXCITATION)

$$\frac{d\,\underline{x}(t)}{dt} = F(t)\,\underline{x}(t) + G(t)\,\underline{u}(t)$$

STATE EQUATION

$$\underline{y}(t) = C(t)\,\underline{x}(t)$$

OBSERVATION EQUATION

Fig. 2.1

Generally, one does not observe the entire state vector at the output of a dynamical system, e.g., in many cases only the first component of the vector is observed. Consequently, we must specify the relationship between the observed random process and the state vector, of the dynamic system. For the majority of the random processes that we shall consider, we shall assume that the observation relationship is a linear, possibly time-varying, no memory transformation, i.e., the observed random process $\underline{y}(t)$ is given by

$$\underline{y}(t) = C(t)\underline{x}(t) \quad \text{(observation equation)}. \tag{2.5}$$

(If the observation is a linear transformation which involves memory, and this transformation is representable in terms of a system of state variables, i.e., there is an ordinary differential equation describing the operation, we can reduce it to the previous case by simply augmenting the state vector and then redefining the matrix C(t).) In Fig. 2.1 we have illustrated a block diagram of the dynamic system that generates the random processes of interest.

Finally, in a communications context, additive white noise is often present in the actual observation. Consequently, we shall consider signals of the form

$$\underline{r}(t) = \underline{y}(t) + \underline{w}(t), \tag{2.6}$$

where

$\underline{y}(t)$ is a process generated as previously discussed,

$\underline{w}(t)$ is a white noise process.

We assume that $\underline{w}(t)$ has zero mean and a covariance matrix given by by

$$E[\underline{w}(t)\underline{w}^T(\tau) = R(t)\delta(t-\tau), \tag{2.7}$$

where R(t) is a positive definite matrix.

Before proceeding several comments are in order.

It is often convenient to describe the random processes in terms of the system that generates them. On occasion we shall do this, e.g. constant parameter systems refer to the processes that may be generated by a dynamical system with a constant state description, e.g. stationary processes, or the Wiener process.

We have avoided introducing the assumption of Gaussian statistics for $\underline{x}(T_o)$, $\underline{u}(t)$ and $\underline{w}(t)$. We shall indicate whenever it is necessary to introduce this assumption; however, for many of our derivations it is unnecessary since we use a structured linear approach rather than an unstructured Gaussian approach.

In two of the chapters we shall analyze problems which involve either processes which are generated by nonlinear dynamical systems or those which are observed linearly through a distributed medium, i.e., the observation method cannot be described in terms of a finite dimensional state equation. Since the notation required for the description of the generation of these processes is peculiar to the individual chapter we shall defer introducing it until then.

Finally, we shall usually work with low pass waveforms. In Appendix B we have introduced the concept of a complex state variable. This concept allows us to analyze bandpass waveforms of

interest with very little modification to the low pass theory that we shall develop.

B. Covariance Functions for State Variable Processes

Several of our derivations concern the covariance matrix $K_{\underline{y}}(t, \tau)$ of a random process $\underline{y}(t)$ which is generated by the methods described in the previous section. In this section we shall briefly review how this covariance can be related to the matrices which describe the system for generating the random processes.

The covariance matrix of $\underline{y}(t)$ is defined to be

$$E[\underline{y}(t)\underline{y}^T(\tau)] \triangleq K_{\underline{y}}(t, \tau). \tag{2.8}$$

By using Eq. 2.5, $K_{\underline{y}}(t, \tau)$ is easily related to the covariance matrix of the state vector $\underline{x}(t)$,

$$K_{\underline{y}}(t, \tau) = C(t)K_{\underline{x}}(t, \tau)C^T(\tau). \tag{2.9}$$

In Appendix A, the following result is shown:

$$K_{\underline{x}}(t, \tau) = \begin{cases} \theta(t, \tau)K_{\underline{x}}(\tau, \tau) & \text{for } t \geq \tau, \\ K_{\underline{x}}(t, t)\theta^T(\tau, t) & \text{for } \tau \geq t, \end{cases} \tag{2.10}$$

where $\theta(t, \tau)$ is the transition matrix for the system defined by the matrix $F(t)$, i.e.,

$$\frac{d}{dt}\theta(t,\tau) = F(t)\theta(t,\tau), \tag{2.11a}$$

$$\theta(\tau,\tau) = I. \tag{2.11b}$$

Furthermore, in Appendix A, it is shown that the matrix $K_{\underline{x}}(t,t)$ satisfies the following differential equation.

$$\frac{d}{dt}K_x(t,t) = F(t)K_x(t,t) + K_x(t,t)F^T(t) + G(t)QG^T(t), \quad t > T_o, \tag{2.12a}$$

with the initial condition

$$K_{\underline{x}}(T_o,T_o) = P_o. \tag{2.12b}$$

Because many of our examples concern stationary processes and constant parameter systems we shall make some comments regarding their generation. Let us assume that the matrices describing the generation of $\underline{y}(t)$ are constant. Consequently, the transition matrix $\theta(t,\tau)$ is given by the matrix exponential

$$\theta(t,\tau) = e^{F(t-\tau)}. \tag{2.13}$$

Furthermore, let us assume that P_o is chosen to be the steady

state solution P_∞ to Eq. 2.12a.* Therefore, $K_{\underline{y}}(t, t+\Delta t)$ is a function only of Δt. We have

$$K_{\underline{x}}(t, t+\Delta t) = \begin{cases} e^{-F\Delta t} P_\infty, & \Delta t \leq 0, \\ P_\infty e^{F^T \Delta t}, & \Delta t \geq 0. \end{cases} \tag{2.14}$$

We shall use two particular stationary processes in several of our examples. The first is the first order Butterworth, or one-pole, process. The covariance function for this process is

$$K_y(t,\tau) = Se^{-k|t-\tau|}. \tag{2.15}$$

The state equations which describe the generation of this process are

---

*One can evaluate P_∞ using transform techniques. It is easily shown that

$$P_\infty = \int_0^\infty e^{Ft} GQG^T e^{Ft^T} dt =$$

$$\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} [Is-F]^{-1} GQG^T [-Is-F^T]^{-1} ds$$

$$\frac{dx(t)}{dt} = -kx(t) + u(t), \qquad t > T_o, \tag{2.16a}$$

$$y(t) = x(t), \tag{2.16b}$$

$$E[u(t)u(\tau)] = 2kS\,\delta(t-\tau), \tag{2.16c}$$

$$E[x^2(T_o)] = P_o \tag{2.16d}$$

The matrices involved are

$$F = -k \qquad C = 1$$
$$G = 1 \qquad P_o = S$$
$$Q = 2kS \tag{2.17 a-e}$$

We shall use this process to illustrate analytically many of the techniques that we shall develop.

The second process we shall use in examples to illustrate the numerical aspects of an analysis using our techniques. This process is generated by a two dimensional state equation where the matrices describing the process generation are

$$F = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix} \qquad C = [1 \quad 0]$$

$$G = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad P_o = \begin{bmatrix} 4 & 0 \\ 0 & 40 \end{bmatrix} \tag{2.18a-e}$$

$$Q = 160$$

The process $\underline{y}(t)$ is a stationary process whose covariance function is

$$K_y(t, t+\Delta t) = \frac{4}{3} e^{-|\Delta t|}(3\cos(3\Delta t) + \sin(3|\Delta t|)), \qquad (2.19a)$$

and whose spectrum is

$$S_y(\omega) = \frac{40}{\omega^2 - 16\omega^2 + 100} . \qquad (2.19b)$$

We have illustrated these functions in Figs. 2.2a and 2.2b. We have included this process principally to illustrate some of the computational aspects of our techniques. We have also chosen to state matrices such that the spectrum of y(t) has a peak in it away from the origin. This will introduce some interesting aspects the some of our examples. We should also note that in all our examples, any analysis involving this process would require a prohibitive amount of time.

C. The Derivation of the Differential Equations for

$$\underline{\zeta}(\tau) = \int_{T_o}^{T_f} K_{\underline{y}}(t, \tau)\underline{f}(\tau)d\tau$$

Many of the problems in communication theory that we shall analyze involve the integral operation

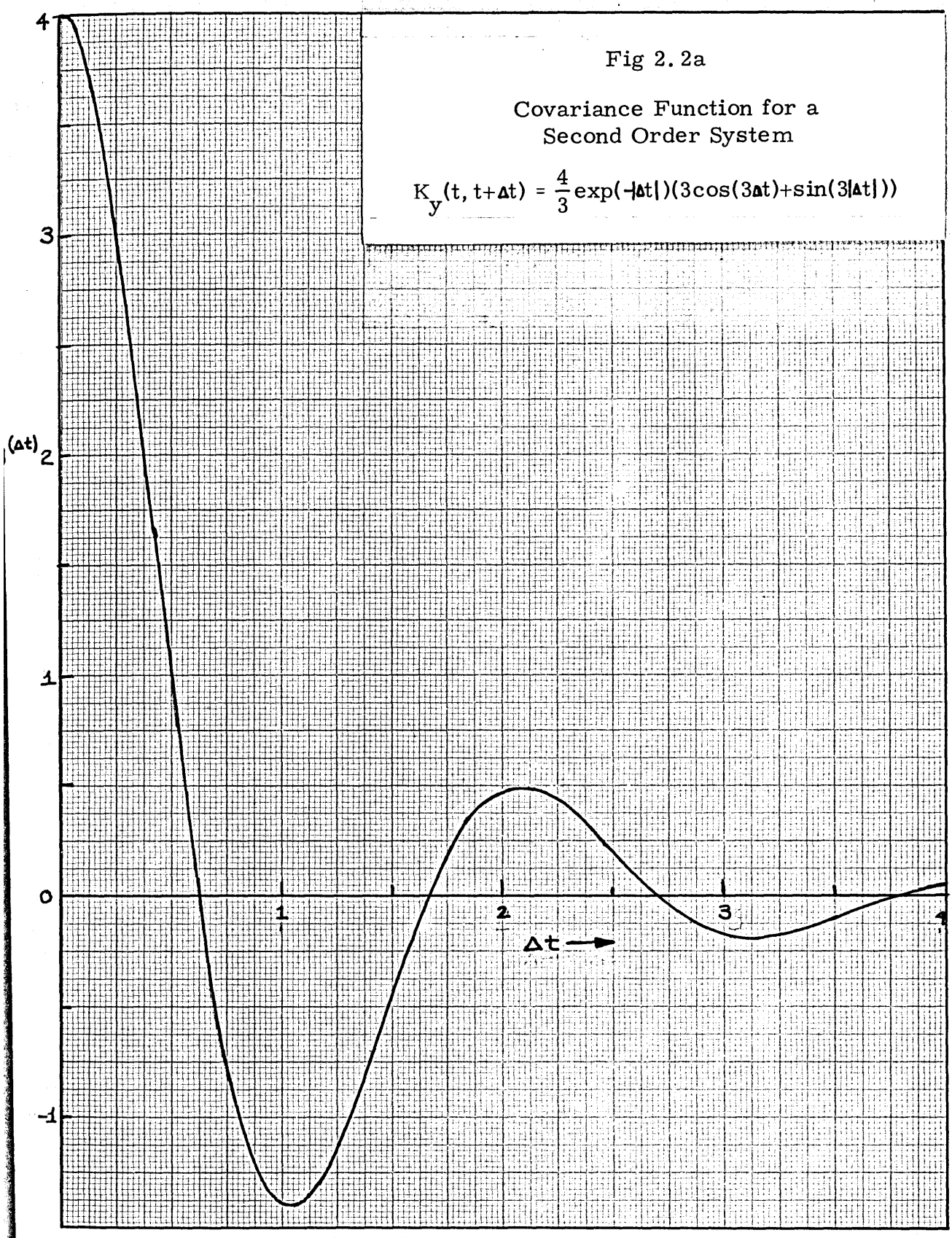


Fig 2.2a

Covariance Function for a
Second Order System

$$K_y(t, t+\Delta t) = \frac{4}{3}\exp(-|\Delta t|)(3\cos(3\Delta t)+\sin(3|\Delta t|))$$

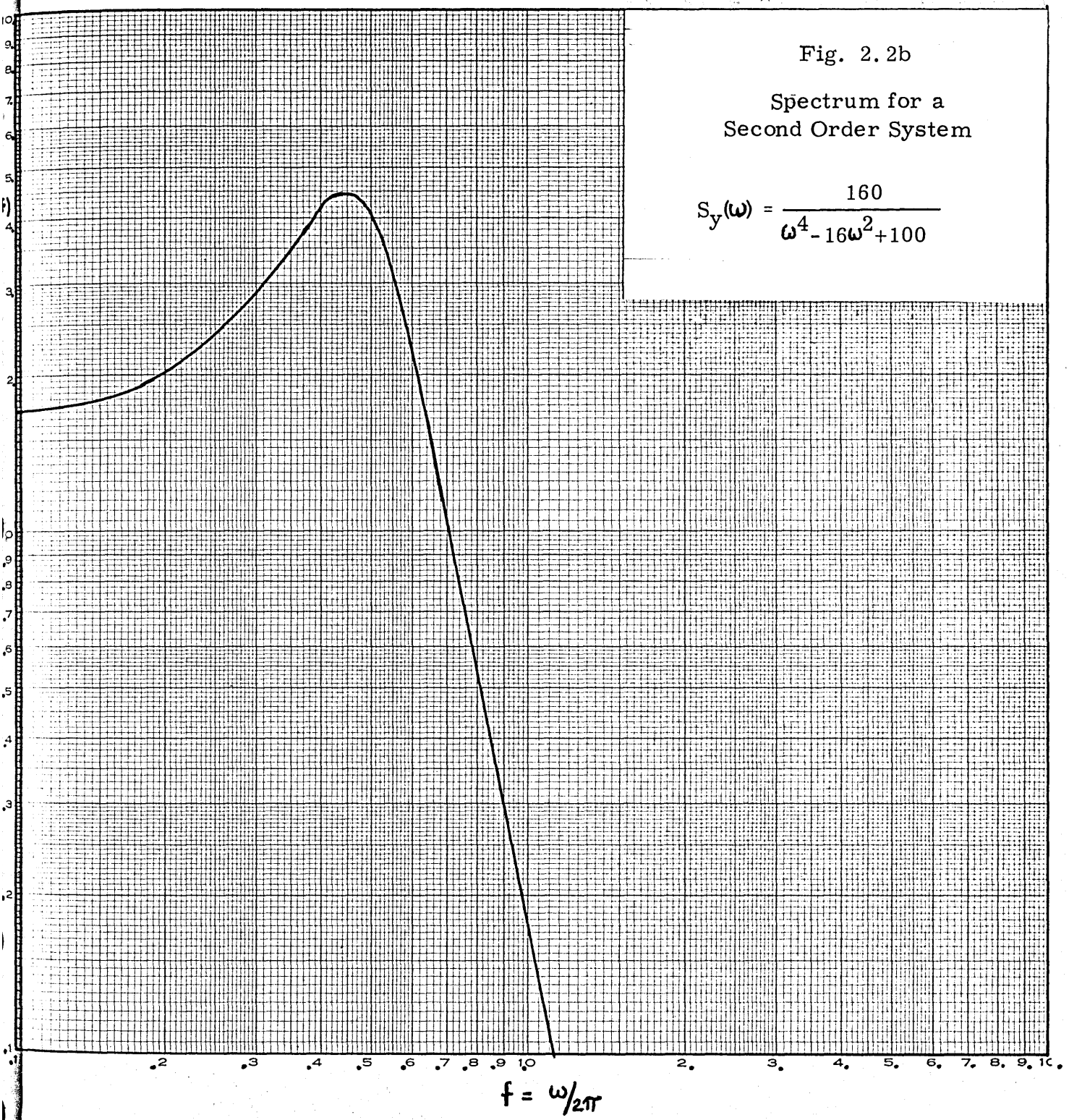


Fig. 2.2b

Spectrum for a
Second Order System

$$S_y(\omega) = \frac{160}{\omega^4 - 16\omega^2 + 100}$$

$$\underline{\zeta}(t) = \int_{T_o}^{T_f} K_{\underline{y}}(t,\tau)\underline{f}(\tau)d\tau, \qquad T_o \leq t \leq T_f, \tag{2.20}$$

In the study of Fredholm integral equations $\underline{f}(t)$ is either related to the eigenfunction $\underline{\phi}(t)$ in the homogeneous case, or it is the solution $\underline{g}(t)$ in the nonhomogeneous case. In linear estimation theory, this is the integral operation specified by the Wiener-Hopf equation. In this section we shall derive a set of differential equations for this integral operation. Solving these differential equations is equivalent to performing the integral operation specified by Eq. 2.20. In many of our derivations, we shall use these differential equations to convert an integral operator to a set of differential equations.

Let us now proceed with our derivation. By using Eq. 2.9, we may write Eq. 2.20 as

$$\underline{\zeta}(t) = C(t)\underline{\xi}(t), \qquad T_o \leq t \leq T_f, \tag{2.21}$$

where

$$\underline{\xi}(t) \triangleq \int_{T_o}^{T_f} K_{\underline{x}}(t,\tau)C^T(\tau)\underline{f}(\tau)d\tau, \quad T_o \leq t \leq T_f. \tag{2.22}$$

We shall now determine a set of differential equations in terms of the function $\underline{\xi}(t)$. Substituting Eq. 2.10 in Eq. 2.22, we have

$$\underline{\xi}(t) = \int_{T_o}^{t} \theta(t,\tau) K_{\underline{x}}(\tau,\tau) C^T(\tau) \underline{f}(\tau) d\tau$$

$$+ K_x(t,t) \int_{t}^{T_f} \theta^T(\tau,t) C^T(\tau) f(\tau) d\tau, \quad T_o \leq t \leq T_f. \qquad (2.23)$$

If we differentiate Eq. 2.23 with respect to t, we obtain

$$\frac{d\underline{\xi}(t)}{dt} = \int_{T_o}^{t} \frac{\partial\theta(t,\tau)}{\partial t} K_{\underline{x}}(\tau,\tau) C^T(\tau) \underline{f}(\tau) d\tau$$

$$+ \frac{dK_{\underline{x}}(t,t)}{dt} \int_{t}^{T_f} \theta^T(\tau,t) C^T(\tau) f(\tau) d\tau$$

$$+ K_{\underline{x}}(t,t) \int_{t}^{T_f} \frac{\partial\theta^T(\tau,t)}{\partial t} C^T(\tau) \underline{f}(\tau) d\tau, \quad T_o \leq t \leq T_f. \qquad (2.24)$$

We have used Eq. 2.11b and cancelled two equal terms. In the first term of the right-hand side of Eq. 2.24 we substitute Eq. 2.11a, and in the last term we use the fact that $\theta^T(\tau,t)$ is the transition matrix for the adjoint equation of the matrix F(t).[6] That is,

$$\frac{\partial}{\partial t} \theta^T(\tau,t) = -F^T(t)\, \theta^T(\tau,t). \qquad (2.25)$$

When we make these two substitutions in Eq. 2.24, we obtain

$$\frac{d\underline{\xi}(t)}{dt} = F(t) \int_{T_o}^{t} \theta(t,\tau) K_{\underline{x}}(\tau,\tau) C^T(\tau) \underline{f}(\tau) d\tau$$

$$+ \left[\frac{dK_{\underline{x}}(t,t)}{dt} - K_{\underline{x}}(t,t) F^T(t) \right] \int_{t}^{T_f} \theta^T(\tau,t) C^T(\tau) f(\tau) d\tau,$$

$$T_o \leq t \leq T_f. \qquad (2.26)$$

By applying Eq. 2.12a, we obtain

$$\frac{d\underline{\xi}(t)}{dt} = F(t) \int_{T_o}^{t} \theta(t,\tau) K_{\underline{x}}(\tau,\tau) C^T(\tau) \underline{f}(\tau) d\tau$$

$$+ [F(t) K_x(t,t) + G(t) Q G^T(t)] \int_{t}^{T_f} \theta^T(\tau,t) C^T(\tau) f(\tau) d\tau,$$

$$T_o \leq t \leq T_f. \qquad (2.27)$$

After rearranging terms and using Eq. 2.23, we finally have

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\xi(t) + G(t) Q G^T(t) \int_{t}^{T_f} \theta^T(\tau,t) C^T(\tau) \underline{f}(\tau) d\tau,$$

$$T_o \leq t \leq T_f. \qquad (2.28)$$

At this point we have derived a differential equation for $\underline{\xi}(t)$; however, we see that an integral operation still remains. Let us simply define this integral operation as a second linear functional of $\underline{f}(t)$:

$$\underline{\eta}(t) = \int_t^{T_f} \theta^T(\tau, t) C^T(\tau) \underline{f}(\tau) d\tau, \quad T_o \leq t \leq T_f. \tag{2.29}$$

Therefore, we have

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\underline{\xi}(t) + G(t)Q\, G^T(t)\underline{\eta}(t), \; T_o \leq t \leq T_f. \tag{2.30}$$

It is now a simple matter to derive a second differential equation which $\underline{\eta}(t)$ satisfies. Differentiating Eq. 2.29 gives us

$$\frac{d\underline{\eta}(t)}{dt} = -C^T(t)f(t) - F^T(t) \int_t^{T_f} \theta^T(\tau, t) C^T(\tau) \underline{f}(\tau) d\tau,$$

$$T_o \leq t \leq T_f, \tag{2.31}$$

where we have again used the adjoint relationship given by Eq. 2.25. After substituting Eq. 2.27, we have

$$\frac{d\underline{\eta}(t)}{dt} = -C^T(t)\underline{f}(t) - F^T(t)\underline{\eta}(t), \quad T_o \leq t \leq T_f. \tag{2.32}$$

We now want to derive two sets of boundary conditions which Eqs. 2.30 and 2.32 satisfy. In all the applications that we shall consider, the function $f(\tau)$ is bounded at the end points, $t = T_o$ and $t = T_f$. Consequently, by setting $t = T_f$ in Eq. 2.29, we obtain

$$\underline{\eta}(T_f) = \underline{0}. \tag{2.33}$$

The second boundary condition follows directly from Eq. 2.23. If we set $t = T_o$, the first term is zero, while the second term may be written

$$\xi(T_o) = K_{\underline{x}}(T_o, T_o) \int_{T_o}^{T_f} \theta^T(\tau, t) C(\tau) \underline{f}(\tau) d\tau, \tag{2.34}$$

or

$$\underline{\xi}(T_o) = K_{\underline{x}}(T_o, T_o) \underline{\eta}(T_o) = P_o \underline{\eta}(T_o). \tag{2.35}$$

It is easy to see that the two boundary conditions given by Eqs. 2.33 and 2.35 are independent.

We may now summarize the results of our derivation. We have derived two differential equations:

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\underline{\xi}(t) + G(t)Q\,G^T(t)\underline{\eta}(t), \quad T_o \leq t \leq T_f \tag{2.30}$$

(repeated)

$$\frac{d\underline{\eta}(t)}{dt} = -C(t)\underline{f}(t) - F^{T}(t)\underline{\eta}(t), \qquad T_o \leq t \leq T_f. \tag{2.32 (repeated)}$$

In addition, we have the boundary conditions

$$P_o\underline{\eta}(T_o) = \underline{\xi}(T_o), \tag{2.35 (repeated)}$$

$$\underline{\eta}(T_f) = \underline{0}. \tag{2.33 (repeated)}$$

The relation to the original integral operation is given by

$$\underline{\zeta}(t) = C(t)\underline{\xi}(t) = \int_{T_o}^{T_f} K_{\underline{y}}(t,\tau)f(\tau)d\tau. \quad T_o \leq t \leq T_f. \tag{2.36}$$

Notice that the only property of $\underline{f}(t)$ which we required was its boundedness at the endpoints of the interval. (This excludes considering equations of the first kind where singularity functions may appear there.) Equations 2.23 and 2.35 each imply n linearly independent boundary conditions. Since the differential equations are linear, any solution that satisfies the boundary conditions is unique. Finally, the derivation of these equations can be reversed in order to obtain the functional defined by Eq. 2.22; that is, we can integrate the differential equations rather than differen- the integral equation. Consequently, the solution $\underline{\xi}(t)$ to the differential equations must be identical to the result of the functional operation of Eq. 2.22. This implies that the existence of a solution $\underline{\xi}(t)$ that satisfies the boundary conditions is both

necessary and sufficient for the existence of the solution to the operation defined by Eq. 2.22.

In this chapter we have developed the concepts that we shall need. In addition, we have presented a derivation which we shall utilize in several subsequent chapters. We shall now use this material to develop a theory for the solution of Fredholm integral equations. We then shall apply this theory to several problems in optimal communications.

CHAPTER III

HOMOGENEOUS FREDHOLM INTEGRAL EQUATIONS

Homogeneous Fredholm integral equations play an important role in theoretical communications. As a theoretical tool, their most important use arises in the theory of Karhunen-Loève expansions of a random processes. One of the more difficult aspects of this theory is that, except in a limited number of cases, it is very difficult to find solutions to these equations. We shall now apply a state variable method to find a solution technique that is both analytically efficient and is especially well suited for determining solutions by computational methods. We shall then work a number of examples to illustrate the technique. Finally, we shall show how our results can be used to find the Fredholm determinant function.

A. The State Variable Solution to Homogeneous Fredholm Integral Equations

The homogeneous Fredholm integral equation is usually written

$$\int_{T_o}^{T_f} K_{\underline{y}}(t,\tau)\underline{\phi}(\tau)d\tau = \lambda\underline{\phi}(\tau), \qquad T_o \leq t \leq T_f, \tag{3.1}$$

where the kernel $\underline{K}_{\underline{y}}(t,\tau)$ is the covariance matrix of a vector random process $\underline{y}(t)$ which is generated by the methods described in the previous chapter, $\underline{\phi}(t)$ is an eigenfunction solution, and λ is the associated eigenvalue. (We should note that we are using a vector eigenfunction-scalar eigenvalue expansion.[18,3] In this expansion we have

$$\underline{y}(t) = \sum_{i=1}^{\infty} y_i \underline{\phi}_i(t), \qquad T_o \leq t \leq T_f, \tag{3.1a}$$

where the generalized Fourier coefficient is a scalar given by

$$y_i = \int_{T_o}^{T_f} \underline{y}^T(\tau)\underline{\phi}_i(\tau)d\tau \tag{3.1b}$$

This type of expansion is necessary if the components of y(t) are correlated.)

The solution to this equation is an eigenvalue problem. There are an at most countable number of values of $\lambda > 0$ for which solutions exist to Eq. 3.1 and there are no solutions for $\lambda < 0$. If $\underline{K}_{\underline{y}}(t,\tau)$ is positive definite, then the solutions to Eq. 3.1 have positive eigenvalues and form a complete orthonormal set. However, if $\underline{K}_{\underline{y}}(t,\tau)$ is only non-negative definite, then there exists solutions $\phi_o(t)$ with zero eigenvalues, i.e., they are orthogonal to the kernel

$$\int_{T_o}^{T_f} \underline{K}_y(t,\tau)\underline{\phi}_o(\tau)d\tau = \underline{0}, \quad T_o \le t \le T_f. \tag{3.2}$$

We shall consider finding only those solutions with positive eigenvalues.

If we index the eigenvalues and their associated eigenfunction by the subscript i, the integral Eq. 3.1 becomes

$$\int_{T_o}^{T_f} K_y(t,\tau)\phi_i(\tau)d\tau = \lambda_i\phi_i(\tau), \quad T_o \le t \le T_f. \tag{3.3}$$

When we employ Eq. 2.9 we may write Eq. 3.3 as

$$C(t) \int_{T_o}^{T_f} \underline{K}_x(t,\tau)C^T(\tau)\underline{\phi}_i(\tau)d\tau = \lambda_i\underline{\phi}_i(t), \; T_o \le t \le T_f. \tag{3.4}$$

Let us now put Eq. 3.4 into such a form that we can employ the results of Section II-C. If in Eq. 3.4 we set

$$\underline{\phi}_i(t) = \underline{f}(t), \tag{3.5}$$

the result is that the integral enclosed by parentheses is the function $\underline{\xi}(t)$ as defined in Eq. 2.22. Consequently, let us define $\underline{\xi}_i(t)$ to be

$$\underline{\xi}_i(t) = \int_{T_o}^{T_f} K_{\underline{x}}(t,\tau) C^T(\tau) \underline{\phi}_i(\tau) d\tau, \quad T_o \leq t \leq T_f; \tag{3.6}$$

so that Eq. 3.4 becomes

$$C(t)\underline{\xi}_i(t) = \lambda_i \underline{\phi}_i(t), \quad T_o \leq t \leq T_f. \tag{3.7}$$

If we assume that λ_i is positive, which is guaranteed if $K_{\underline{y}}(t,\tau)$ is positive definite, we can solve for the eigenfunction in terms of $\underline{\xi}_i(t)$. This gives us

$$\underline{\phi}_i(t) = \frac{1}{\lambda_i} C(t)\underline{\xi}_i(t), \quad T_o \leq t \leq T_f. \tag{3.8}$$

If we examine Eq. 3.6, we see that the integral operation which is defined is of the same form as the operation considered in Section II-C. Consequently, we can reduce it to a set of differential equations with a two point boundary condition. Let us identify $\underline{\phi}_i(t)$ in Eq. 3.6 with $\underline{f}(t)$ in Eq. 2.20. Then, if we substitute

$$\underline{f}(t) = \underline{\phi}_i(t) = \frac{1}{\lambda_i} C(t)\underline{\xi}_i(t), \quad T_o \leq t \leq T_f, \tag{3.9}$$

in Eqs. 2.20 and 2.32, we find that the differential Eqs. 2.30 and 2.32 become

$$\frac{d}{dt}\underline{\xi}_i(t) = F(t)\underline{\xi}_i(t) + G(t)Q\,G^T(t)\underline{\eta}_i(t), \quad T_o \le t \le T_f, \tag{3.10}$$

$$\frac{d}{dt}\underline{\eta}_i(t) = \frac{C^T(t)C(t)}{\lambda_i}\underline{\xi}_i(t) - F^T(t)\underline{\eta}_i(t), \quad T_o \le t \le T_f. \tag{3.11}$$

From Eqs. 2.33 and 2.35, the boundary conditions are

$$\underline{\eta}_i(T_f) = \underline{0}, \tag{3.13a}$$

$$\underline{\xi}_i(T_o) = P_o\underline{\eta}_i(T_o). \tag{3.13b}$$

The desired eigenfunction is related to the solution $\underline{\xi}_i(t)$ by Eq. 3.8, or

$$\underline{\phi}_i(t) = \frac{1}{\lambda_i} C(t)\underline{\xi}_i(t), \qquad T_o \le t \le T_f. \tag{3.8 (repeated)}$$

The net result of Eqs. 3.4 - 3.13 is that we have transformed the homogeneous Fredholm integral equation into a set of differential equations whose coefficients are directly related to the state equations and covariance matrices that are used to generate the random process $\underline{y}(t)$.

Before we use the above results to determine the eigenvalues and eigenfunctions, let us make two observations.

Notice that we have a set of 2n differential equations to solve. This is consistent with previous methods for treating

stationary processes. In these methods, one has a 2n-order differential equation to solve, where 2n is the degree of the denominator polynomial of the spectrum.

Equation 3.8 implies that all of the solutions to Eq. 3.1 with positive λ are contained in the range space defined by C(t). We should note that if C(t) is not onto for a set of t with nonzero measure, then $K_{\underline{y}}(t,\tau)$ is not positive definite. In this situation there may be solutions with λ equal to zero which are not contained in this range space.

We shall now specify a general solution technique for solving these differential equations, which in turn specifies the eigenvalues and eigenfunctions. With this technique we shall first find a transcendental equation that specifies the eigenvalues. Given the eigenvalues, the eigenfunctions follow directly.

Let us define the (2n x 2n) matrix $W(t:\lambda)$ as

$$W(t:\lambda) = \left[\begin{array}{c|c} F(t) & G(t)Q\,G^{T}(t) \\ \hline \dfrac{-C^{T}(t)C(t)}{\lambda} & -F^{T}(t) \end{array}\right], \qquad (3.14)$$

so that in vector form Eqs. 3.10 and 3.11 become

$$\frac{d}{dt}\begin{bmatrix} \underline{\xi}_i(t) \\ \hline \underline{\eta}_i(t) \end{bmatrix} = W(t:\lambda_i)\begin{bmatrix} \underline{\xi}_i(t) \\ \hline \underline{\eta}_i(t) \end{bmatrix}, \quad T_o \le t \le T_f. \tag{3.15}$$

Furthermore, let us define the transition matrix associated with $W(t:\lambda)$ by

$$\frac{\partial}{\partial t}\Psi(t, T_o:\lambda) = W(t:\lambda)\Psi(t, T_o:\lambda), \tag{3.16}$$

$$\Psi(T_o, T_o:\lambda) = I. \tag{3.17}$$

(We have emphasized the λ dependence of $W(t:\lambda)$ and $\Psi(t, T_o:\lambda)$ by including λ as an argument.)

In terms of this transition matrix, the most general solution to Eq. 3.14 is

$$\begin{bmatrix} \underline{\xi}_i(t) \\ \hline \underline{\eta}_i(t) \end{bmatrix} = \Psi(t, T_o:\lambda_i)\begin{bmatrix} \underline{\xi}_i(T_o) \\ \hline \underline{\eta}_i(T_o) \end{bmatrix}, \quad T_o \le t \le T_f. \tag{3.18}$$

After employing the boundary condition specified by Eq. 3.13, we have

$$\begin{bmatrix} \underline{\xi}_i(t) \\ \hline \underline{\eta}_i(t) \end{bmatrix} = \Psi(t, T_o:\lambda_i)\begin{bmatrix} P_o \\ \hline I \end{bmatrix}\underline{\eta}_i(T_o), \quad T_o \le t \le T_f. \tag{3.19}$$

Let us now partition $\Psi(t, T_o:\lambda)$ into four n by n matrices such that

$$\Psi(t, T_o:\lambda) = \left[\begin{array}{c|c} \Psi_{\underline{\xi}\underline{\xi}}(t, T_o:\lambda) & \Psi_{\underline{\xi}\underline{\eta}}(t, T_o:\lambda) \\ \hline \Psi_{\underline{\eta}\underline{\xi}}(t, T_o:\lambda) & \Psi_{\underline{\eta}\underline{\eta}}(t, T_o:\lambda) \end{array}\right] \tag{3.20}$$

Rewriting Eq. 3.19 in terms of these partitions, we have

$$\left[\begin{array}{c} \underline{\xi}_i(t) \\ \hline \underline{\eta}_i(t) \end{array}\right] = \left[\begin{array}{c} \Psi_{\underline{\xi}\underline{\xi}}(t, T_o:\lambda_i)P_o + \Psi_{\underline{\xi}\underline{\eta}}(t, T_o:\lambda_i) \\ \hline \Psi_{\underline{\eta}\underline{\xi}}(t, T_o:\lambda_i)P_o + \Psi_{\underline{\eta}\underline{\eta}}(t, T_o:\lambda_i) \end{array}\right] \underline{\eta}_i(T_o) \tag{3.21}$$

$$T_o \le t \le T_f.$$

Taking the lower partition, the boundary condition given by Eq. 3.12 requires

$$\underline{\eta}_i(T_f) = [\Psi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda_i)P_o + \Psi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda_i)]\underline{\eta}_i(T_o) = 0 \tag{3.22}$$

This implies one of two consequencies. Either

$$\underline{\eta}_i(T_o) = \underline{0}. \tag{3.23}$$

which implies a trivial zero solution; or,

$$\det[\Psi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda_i)P_o + \Psi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda_i)] = 0. \tag{3.24}$$

If the latter is true, Eq. 3.15 has a nontrivial solution which satisfies the requisite boundary conditions. Because of the functional equivalence of these differential equations and the original integral equation, this nontrivial solution to Eq. 3.15 implies that λ_i is an eigenvalue. That is, the eigenvalues of Eq. 3.1 are simply those values of λ_i that satisfy the transcendental equation specified by Eq. 3.24.

Now that we have found an expression for the eigenvalues, we can show how the eigenfunctions follow.

For convenience, define $A(\lambda)$ as

$$A(\lambda) = \Psi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda)P_o + \Psi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda). \tag{3.25}$$

When λ is equal to an eigenvalue, λ_i, $A(\lambda_i)$ has a vanishing determinant. Consequently, the characteristic polynomial of $A(\lambda_i)$ has a root equal to zero and $\underline{\eta}_i(T_o)$ is the characteristic vector associated with this root. (We have used the adjective "characteristic" in order to avoid confusing the eigenvalue properties of the <u>matrix</u> $A(\lambda_i)$ with those of the integral equation, Eq. 3.1.)

Therefore, to determine $\underline{\eta}_i(T_o)$ to a multiplicative factor we need to solve the linear homogeneous equation

$$A(\lambda_i)\underline{\eta}_i(T_o) = 0. \tag{3.26}$$

Given $\underline{\eta}_i(T_o)$ we can find the eigenfunctions by using the upper partition of Eq. 3.21 and Eq. 3.8. This gives us

$$\underline{\phi}_i(t) = \frac{C(t)}{\lambda_i} [\Psi_{\underline{\xi}\,\underline{\xi}}(t, T_o:\lambda_i)P_o + \Psi_{\underline{\xi}\,\underline{\eta}}(t, T_o:\lambda_i)]\underline{\eta}_i(T_o), \qquad (3.27)$$

$$T_o \le t \le T_f.$$

which is the desired result.

Before proceeding we should comment about multiple-order roots of Eq. 3.24. In general, the function det $A(\lambda)$ vanishes with nonzero slope, that is, near an eigenvalue λ_i,

$$\det A(\lambda) = c_1(\lambda-\lambda_i) + c_2(\lambda-\lambda_i)^2 + \ldots, \qquad (3.28)$$

where c_1 is nonzero. In the case of multiple-order eigenvalues, the function det $(A(\lambda))$ vanishes tangentially; that is, near an eigenvalue λ_i of order ℓ

$$\det A(\lambda) = c_\ell(\lambda-\lambda_i)^\ell + c_{\ell+1}(\lambda-\lambda_i)^{\ell+1} + \ldots \ . \qquad (3.29)$$

This implies that there will be ℓ linearly independent vectors $\underline{\eta}_i(T_o)$ satisfying

$$A(\lambda_i)\underline{\eta}_i(T_o) = \underline{0}, \qquad (3.30)$$

<u>i.e.</u> $A(\lambda_i)$ has rank $n - \ell$.

If we examine our solution we see that the only function that we need to determine is the transition matrix of $W(t:\lambda)$, $\Psi(t:T_o:\lambda)$. For the important case of kernels that are covariances of the output of a constant parameter system, we can find an

analytic expression for this transition matrix in terms of the matrix exponential,

$$\Psi(t, T_o:\lambda) = e^{W(\lambda)(t-T_o)}. \tag{3.31}$$

This matrix exponential may be conveniently computed by Laplace transform methods. We have

$$e^{W(\lambda)t} = \mathcal{L}^{-1}\{[I_s - W(\lambda)]^{-1}\}, \tag{3.32}$$

where $\mathcal{L}^{-1}$ is the inverse Laplace operator. (In the inversion the contour must be taken to the right of all pole locations of $[Is-W(\lambda)]^{-1}$.) If one desires a numerical evaluation of this matrix exponential, as is the case for systems of order greater than one, a possible method of calculation is to truncate the series expansion

$$e^{W(\lambda)t} = \sum_{j=0}^{\infty} \frac{[W(\lambda)]^j t^j}{j!}. \tag{3.33}$$

In the case of time varying systems, there is no general analytic method for determining this transition matrix. However, we can still use our technique by evaluating this transition matrix by numerical methods, e.g. by integrating the differential equation defining it.

We can now summarize our results for homogeneous

Fredholm equations. The eigenvalues λ_i are specified by the roots of the transcendental equation

$$\det A(\lambda_i) = 0 \qquad (3.24)$$
(repeated)

where $A(\lambda)$ is given by Eq. 3.25.

$$A(\lambda) = \Psi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda)P_o + \Psi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda) \qquad (3.25)$$
(repeated)

The eigenfunctions are given by Eq. 3.27

$$\underline{\phi}_i(t) = \frac{C(t)}{\lambda_i}\left[\Psi_{\underline{\xi}\underline{\xi}}(t, T_o:\lambda_i)P_o + \Psi_{\underline{\xi}\underline{\eta}}(T_f, T_o:\lambda_i)\right]\underline{\eta}_i(T_o) \qquad (3.27)$$
(repeated)

where $\underline{\eta}_i(T_o)$ satisfies the orthogonality relationship

$$A(\lambda_i)\underline{\eta}_i(T_o) = \underline{0}. \qquad (3.26)$$
(repeated)

(The multiplicative factor may be determined by applying the normality requirement.) The matrices

$$\Psi_{\underline{\xi}\underline{\xi}}(t, T_o:\lambda), \Psi_{\underline{\xi}\underline{\eta}}(t, T_o:\lambda), \Psi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda), \text{ and } \Psi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda)$$

are partitions of the matrix $\Psi(t, T_o:\lambda)$ which is the transition matrix associated with the matrix

$$W(t:\lambda) = \left[\begin{array}{c|c} F(t) & G(t)Q\,G^{T}(t) \\ \hline -\frac{C^{T}(t)C(t)}{\lambda} & -F^{T}(t) \end{array}\right] \qquad (3.14)\ \text{(repeated)}$$

These equations specify the eigenvalues and eigenfunctions for a kernel $K_{\underline{y}}(t,\tau)$ which is the covariance matrix for the random process $\underline{y}(t)$. This random process is generated at the output of a linear system that has a state variable description of its dynamics and a white noise excitation.

To conclude this section we point out some advantages that this technique offers.

1. We can solve Eq. 3.1 in the vector case. For those techniques which rely upon spectral factorization methods the vector case could cause some difficulty. (In some respects we have defined this problem away by our method of characterizing the random processes of interest. It should be pointed out that depending upon the problem this method of characterization may be just as fundamental as the covariance method.)

2. Once the state matrices are chosen to generate $\underline{y}(t)$, the differential equations that must be solved follow directly.

3. One does not have to substitute any functions back into the original integral equation in order to determine the transcendental equation that determines the eigenvalues.

4. We can solve for each eigenvalue and eigenfunction independently of the others, which is significant in actually obtaining accurate solutions.

5. We can study a certain class of time varying kernels.

6. Finally, the most important advantage is that the technique is very well suited to numerical methods. This allows one to determine numerical solutions easily for problems in which an analytic calculation is either difficult or not feasible.

The major disadvantage is that the class of kernels that we may study is limited to those that fit into our state variable model. However, we emphasize that most of the processes of interest in communications do fit within our model.

B. Examples of Eigenvalue and Eigenfunction Determination

In this section we shall illustrate the method developed in the previous section. To do this we shall consider several examples. First, we shall do three examples analytically. We do this principally to illustrate the use of the formulae in the previous section. The processes in these examples are generated by a first order system. In general these are the only systems for which this type of analysis can be done in a reasonable amount of time. We shall then present an example of the numerical analysis of a second-order system. It is this type of problem for which the technique is most useful. It allows one to obtain numerical solutions very quickly with a digital computer.

Example 1 - Eigenvalues and Eigenfunctions for the Wiener Process

The covariance matrix of a Wiener process which starts at $t = 0$ is

$$K_y(t,\tau) = \mu^2 \min(t,\tau), \quad 0 \leq t,\tau. \tag{3.34}$$

A state-variable representation of a system which generates y(t) is

$$\frac{dx(t)}{dt} = u(t), \quad \text{(state equation)}, \tag{3.35a}$$

$$y(t) = \mu x(t), \quad \text{(observation equation)}, \tag{3.35b}$$

where

$$E[u(t)u(\tau)] = \delta(t-\tau), \tag{3.35c}$$

$$E[x^2(0)] = 0. \tag{3.35d}$$

(The Wiener process starts with a known initial state by definition.)

For convenience let us identify the state matrices as indicated in Fig. 2.1

$$F = 0 \qquad C = \mu$$
$$G = 1 \qquad P_o = 0 \tag{3.36a-e}$$
$$Q = 1$$

Let us find the solution to Eq. 3.1 when we choose $T_f = T$ and $T_o = 0$. First, we need the matrix $W(\lambda)$ (the system is constant parameter) specified by Eq. 3.14. After performing the required substitutions of Eq. 3.36 in Eq. 3.14 we have

$$W(\lambda) = \left[\begin{array}{c|c} 0 & 1 \\ \hline -\frac{\mu^2}{\lambda} & 0 \end{array}\right] \tag{3.37}$$

To find the eigenvalues and eigenfunctions we need to find the transition matrix of $W(\lambda)$. If we apply Eq. 3.32 we find that the transition matrix $\Psi(t,0:\lambda)$ is

$$\Psi(t,0:\lambda) = \left[\begin{array}{c|c} \cos(\frac{\mu}{\sqrt{\lambda}}t) & \frac{\lambda}{\mu}\sin(\frac{\mu}{\sqrt{\lambda}}t) \\ \hline -\frac{\mu}{\lambda}\sin(\frac{\mu}{\sqrt{\lambda}}t) & \cos(\frac{\mu}{\sqrt{\lambda}}t) \end{array}\right] \tag{3.38}$$

We now simply apply the results as summarized at the end of the previous section.

First we substitute Eqs. 3.36e and 3.38 evaluated at $t = T$ into Eq. 3.25 to find $A(\lambda)$. In order for an eigenvalue to exist Eq. 3.24 implies

$$\det(A(\lambda_i)) = \cos(\frac{\mu}{\sqrt{\lambda_i}}T) = 0. \tag{3.39}$$

The distinct solutions to Eq. 3.39 are given by

$$\lambda_i = \left[\frac{2\mu T}{(2i + T)\pi} \right]^2 \tag{3.40}$$

$$i = \quad 0, 1, 2, \ldots .$$

The eigenfunctions follow by substituting Eq. 3.38 in Eq. 3.27. After determining the appropriate normalization factor, we have

$$\phi_i(t) = \sqrt{\frac{2}{T}} \sin \left[\left(\frac{2i + 1}{2} \right) \frac{\pi}{T} t \right] \quad 0 \leq t \leq T. \tag{3.41}$$

Example 2 - Eigenvalues and Eigenfunctions for a One Pole, Stationary Process

Let us now consider the kernel of Eq. 3.1 to be

$$K_y(t, \tau) = Se^{-k|t-\tau|}. \tag{3.42}$$

This is the covariance of the output of a first order system with a pole at -k and P_o chosen such that the process is stationary. The state equations that generate this process are given by Eqs. 2.16 and 2.17.

Since the kernel is stationary only the difference between the upper and lower limits of the integral are important. Consequently, we again set $T_f = T$ and $T = 0$.

Proceeding as before, the matrix $W(\lambda)$ follows by substituting Eqs. 3.44 in 3.14.

$$W(\lambda) = \left[\begin{array}{c|c} -k & 2kS \\ \hline -\frac{1}{\lambda} & k \end{array}\right] . \qquad (3.43)$$

The transition matrix for $W(\lambda)$ is

$$\Psi(t,0:\lambda) = \left[\begin{array}{c|c} \cos(kbt) - \frac{\sin(kbt)}{b} & \frac{2P}{b}\sin(kbt) \\ \hline -\frac{1}{\lambda k}\frac{\sin(kbt)}{b} & \cos(kbt) + \frac{\sin(kbt)}{b} \end{array}\right] , \qquad (3.44)$$

where

$$b \triangleq \sqrt{\frac{2S}{k\lambda} - 1} \quad . \qquad (3.45)$$

By substituting Eq. 3.46 in 3.24 and 3.25, we obtain an equation which determines our eigenvalues

$$\det(A(\lambda_i)) = \frac{1}{b_i}\left[1 - \frac{S}{\lambda_i k}\right]\sin(kb_iT) + \cos(kb_iT) = 0 \qquad (3.46)$$

In order to compute the roots by hand, Eq. 3.46 can be put in a more convenient form. This form is

$$\tan(kb_iT) = \frac{2b_i}{b_i^2 - 1} . \qquad (3.47)$$

Solving Eq. 3.46 for λ gives us the expression for the eigenvalues, λ_i, in terms of the b_i,

$$\lambda_i = \frac{2S}{k} [\, 1 + b_i^2 \,]^{-1}. \tag{3.48}$$

Applying Eq. 3.27 gives us the eigenfunctions. They are of the form

$$\phi_i(t) = \gamma_i [\cos(kb_i t) + \frac{1}{b_i} \sin(kb_i t)], \quad 0 \leq t \leq T, \tag{3.49}$$

where γ_i is a normalizing factor.

Example 3 - Eigenvalues and Eigenfunctions for a One Pole, Non-Stationary Process

The output process of a constant parameter system is not necessarily stationary e.g. the Wiener process. A second example of this can be generated from the previous example. Instead of setting P_o equal to the mean-square power of the stationary process, assume that we know the state at t = 0 exactly; that is,

$$P_o = 0. \tag{3.50}$$

In this case the covariance function becomes

$$K_y(t,\tau) = \begin{cases} S\, e^{-kt}(e^{k\tau} - e^{-k\tau}) & \text{for } t > \tau, \\ S\, e^{-k\tau}(e^{kt} - e^{-kt}) & \text{for } \tau > t. \end{cases} \tag{3.51}$$

By substituting Eq. 3.44 in Eqs. 3.24 and 3.25 and setting P_o equal to zero, the equation for the eigenvalues follows directly:

$$\det(A(\lambda_i)) = \cos(kb_iT) + \frac{\sin(kb_iT)}{b_i} = 0, \tag{3.52}$$

where, as before,

$$b = \sqrt{\frac{2S}{k\lambda} - 1} \tag{3.45}$$

or equivalently,

$$\tan(kb_iT) = -b_i. \tag{3.53}$$

From Eqs. 3.27 and 3.44, the eigenfunctions have the form

$$\phi_i(t) = \gamma_i \sin(kb_it) \qquad 0 \le t \le T, \tag{3.54}$$

where γ_i is again a normalizing factor.

Example 4 - Eigenvalues for a Two Pole, Stationary Process

In this example we want to consider the analysis when the kernel is the covariance of the output of a second-order system.

In contrast with the previous examples, however, we shall consider a particular system and analyze it by using numerical methods. Obtaining analytic results for systems whose dimension is greater than one is straightforward, but extremely tedious. Let us assume that the kernel $K_{\underline{y}}(t, \tau)$ of Eq. 3.1 is the covariance function given by Eq. 2.19 and illustrated by Fig. 2.2a. The state equations for generating y(t) are specified by Eq. 2.18. In addition, let us set $T_f = 2$ and $T = 0$.

First, we need the matrix $W(\lambda)$. By substituting Eq. 2.18 into Eq. 3.14 we obtain

$$W(\lambda) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -10 & -2 & 0 & 160 \\ -\frac{1}{\lambda} & 0 & 0 & 10 \\ 0 & 0 & -1 & 2 \end{bmatrix} \tag{3.55}$$

In order to determine $\det[A(\lambda)]$, we need to find the transition matrix of $W(\lambda)$ evaluated at $T = 2$, i.e. we need to exponentiate the matrix $W(\lambda) \cdot T$. To do this, we used a straightforward approach by applying Eq. 3.33 and taking the first 30 terms in a nested fashion. (For further discussion regarding our numerical procedures, see Reference 32.

Once we find this transition matrix, we performed the operations indicated by Eq. 3.25 to find $\det[A(\lambda)]$. By varying the parameter λ and repeating this procedure, we can plot this function versus λ. The resulting curve is indicated in Fig. 3.1.

In this figure the zero crossings are the desired eigenvalues. This type of behavior for the function det(A(λ)) is typical of those that we have observed. In the region of the larger eigenvalues, corresponding to those with significant energy, the function is well-behaved and oscillating (nonperiodically). As we approach the less significant eigenvalues, however, the amplitude of the oscillation rapidly increases. Eventually, the eigenvalues become so small that it becomes difficult to compute A(λ) accurately. In this region the eigenvalues are approaching their asymptotic behavior, as discussed by Capon.[19] This behavior is governed by the "tail" of $S_y(\omega)$. (In Fig. 3.1 we have used arrowheads to indicate the location of the eigenvalues as indicated by Capon's formulae. As the eigenvalues become small, the comparison is quite good.)

Since this state-variable technique is well suited for finding the significant eigenvalues, one could combine this method with an asymptotic method in order to find all of the eigenvalues conveniently. In all cases that we have studied, we could account for at least 95 percent (often as much as 99 percent) of the total energy

$$\int_0^{T_f} \mathrm{Tr}[K_y(t,t)]\,dt = E = \sum_{i=1}^{\infty} \lambda_i, \tag{3.56}$$

by our method. The asymptotic method, or a comparable one for non-stationary kernels, could be used to calculate the eigenvalues

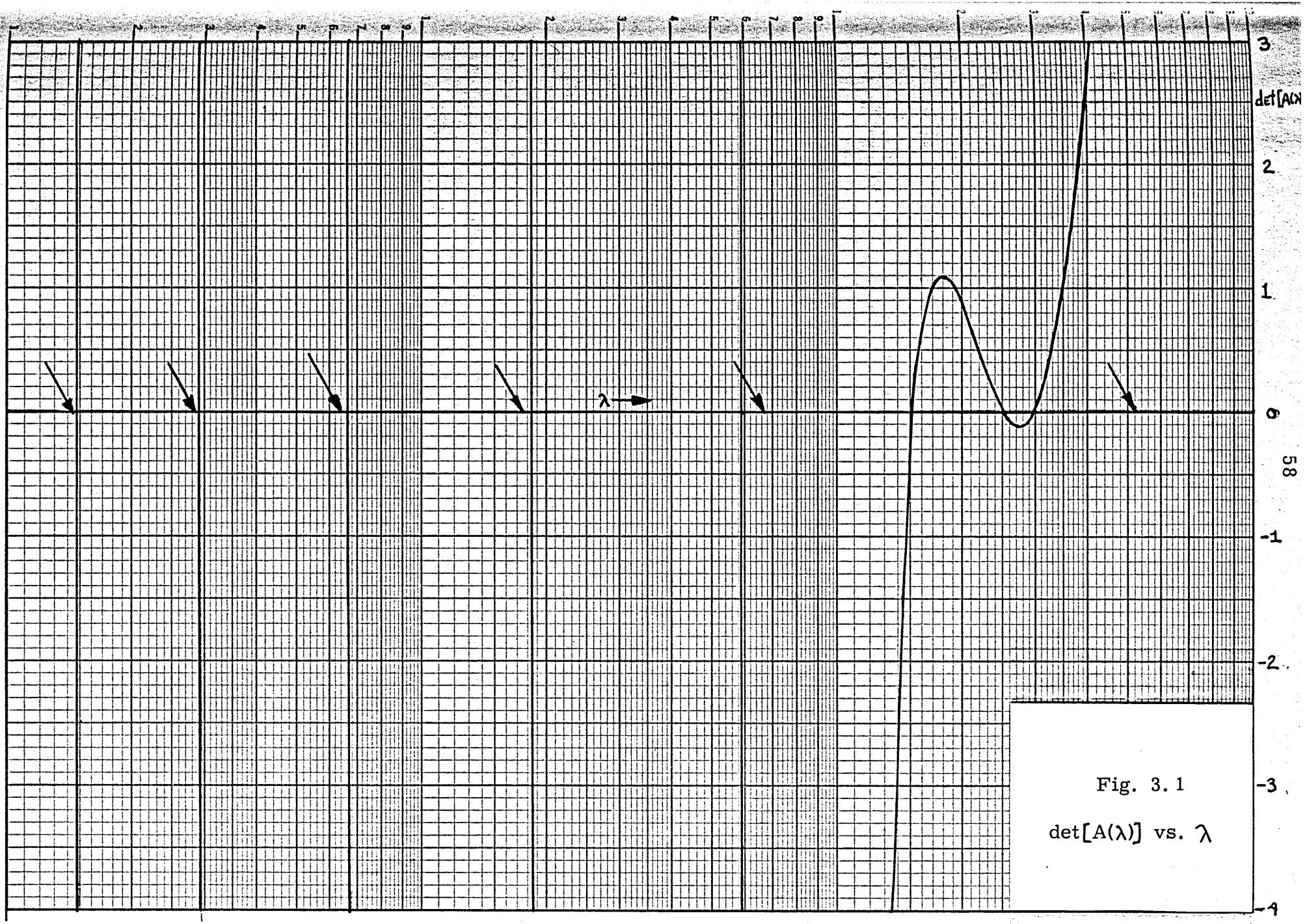


Fig. 3.1

det[A(λ)] vs. λ

corresponding the residual energy not accounted for.

We can summarize the behavior of the eigenvalues for this example by plotting in Fig. 3.2 the first six eigenvalues against T, the length of the interval ($T = T_f - T_o$). We see that the curves satisfy the monotonicity requirement,[20,3]

$$\frac{\partial \lambda_i(T)}{\partial T} = \lambda_1(T) \underline{\phi}_i^T(T) \underline{\phi}_i(T). \tag{3.57}$$

In addition, the number of significant eigenvalues increases with T reflecting the increase in the "2WT product".

We shall conclude this example by discussing the computer time required to find the eigenvalues for this kernel. We used the Fortran language on the IBM 7094 computer (the method has recently been reprogrammed for the IBM 360 as a general purpose routine.) As indicated earlier, we did not employ any sophisticated algorithms for computing $A(\lambda)$. The time required to compute the data for Fig. 3.2 in order to find the first eight eigenvalues (99.8 percent of the energy) is approximately 20 seconds; in addition, the eigenfunctions may be found with very little additional computer time.

C. The Fredholm Determinant Function

In the application of the Fredholm theory to communication theory problems, the Fredholm determinant function is often used, e.g. in the design of receivers and the calculation of their performance for detecting Gaussian signals in Gaussian noise. This function is defined as

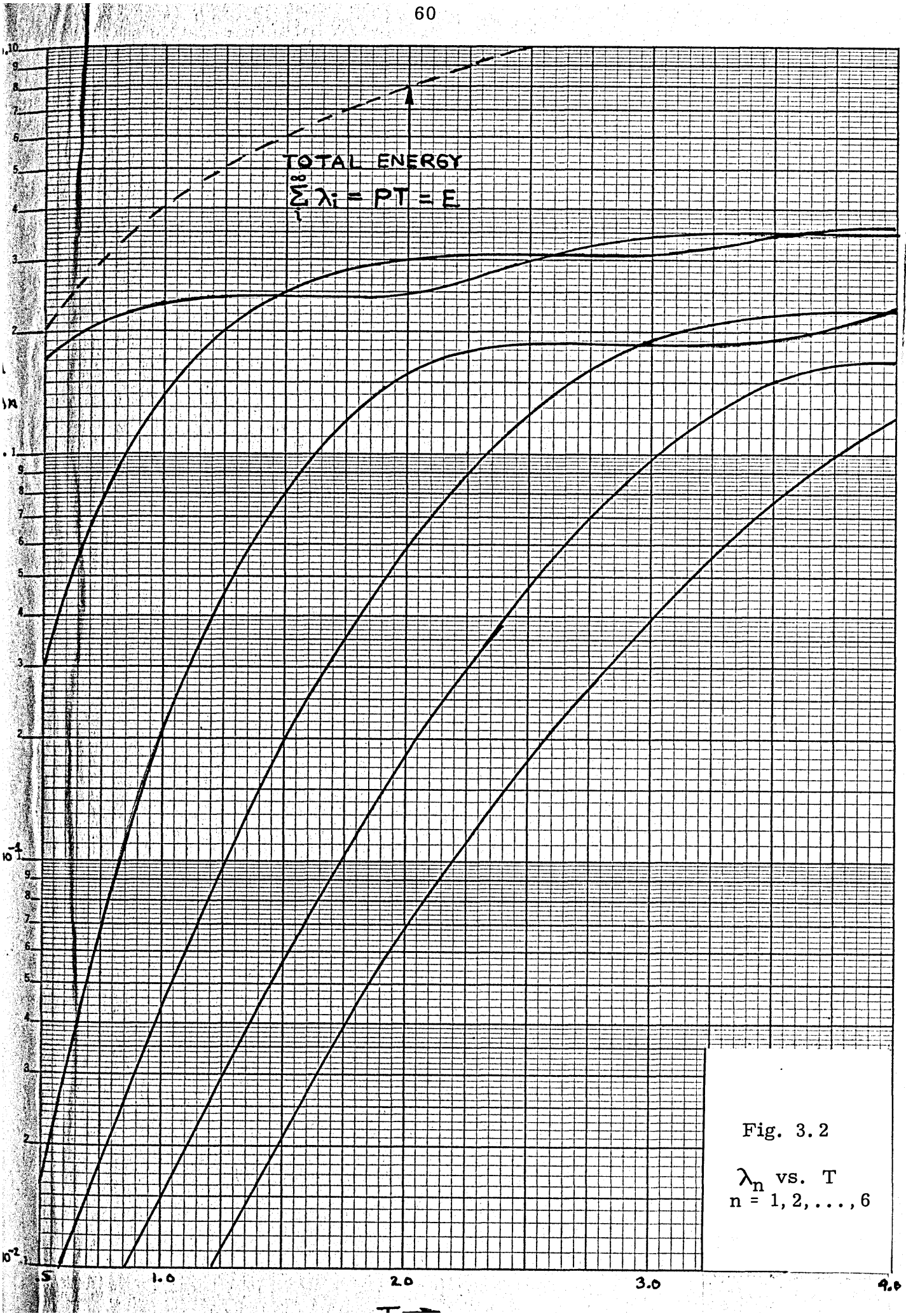


Fig. 3.2

λ_n vs. T
n = 1, 2, . . . , 6

$$D_{\mathcal{F}}(z) = \prod_{i=1}^{\infty} (1+z\lambda_i) \tag{3.57}$$

where the λ_i's are the eigenvalues of Eq. 3.1. In this section, we shall show how the theory developed in Section 3.1 of this chapter can be used to find a closed form expression for this function. This expression can be determined either analytically, it can be readily calculated by the same numerical procedures employed in the previous section.

To do this we make some observations regarding the function $\det[A(\frac{1}{z})]$ where z may be a complex variable. (1.) It is easy to argue that $\det[A(\frac{1}{z})]$ is analytic in the finite plane. (2.) Because our test for the eigenvalues is necessary and sufficient, $\det[A(\frac{1}{z})]$ has zeroes only at $\frac{1}{\lambda_i}$. (3.) The sum of the eigenvalues converges to E(Eq. 3.56). Given these observations, one can show that $\det[A(\frac{1}{z})]$ has the infinite product form[21]

$$\det[A(\frac{1}{z})] = A_o \prod_{i=1}^{\infty} (1-\lambda_i z) \tag{3.58}$$

where

$$A_o = \lim_{z \to 0} A(\frac{1}{z}) \tag{3.59}$$

Comparing this with Eq. 3.57, we find that the Fredholm determinant function is given by

$$D_{\mathcal{F}}(z) = \frac{1}{A_o} \det[A(-\frac{1}{z})] , \tag{3.60}$$

i.e. we can evaluate D(z) by using the same function that we employed for determining the eigenvalues of Eq. 3.1 except we use a negative argument.

The only issue that remains is a convenient method for evaluating the constant A_o. In a direct proof of Eq. 3.60 Collins has done this.[22] For completeness, we shall include this part of his derivation.

If we let $z \to 0$ in Eq. 3.14, $W(t:\frac{1}{z})$ becomes

$$\lim_{z \to 0} W(t:\frac{1}{z}) = \begin{bmatrix} F(t) & G(t)\, Q\, G^T(t) \\ 0 & -F^T(t) \end{bmatrix} \tag{3.61}$$

Let us examine the differential equation, Eq. 3.16 defining $\Psi(t, T_o:\frac{1}{z})$ when we substitute Eq. 3.61. We see that

$$\frac{d}{dt} \lim_{z \to 0} \Psi_{\eta\xi}(t, T_o:\frac{1}{z}) = -F^T(t) \lim_{z \to 0} \Psi_{\eta\xi}(t, T_o:\frac{1}{z}) \tag{3.62}$$

Since the initial condition for this homogeneous equation is the zero matrix, its solution is the zero matrix, i.e.

$$\lim_{z \to 0} \Psi_{\eta\xi}(t, T_o:\frac{1}{z}) = 0 \tag{3.63}$$

Similarly we see that

$$\frac{d}{dt} \lim_{z \to 0} \Psi_{\eta\eta}(t, T_o : \frac{1}{z}) = -F^T(t) \lim_{z \to 0} \Psi_{\eta\eta}(t, T_o : \frac{1}{z}) \qquad (3.64)$$

However, its initial condition is the identity matrix. Therefore, from the adjoint relationship we have

$$\lim_{z \to 0} \Psi_{\eta\eta}(T_f, T_o : \frac{1}{z}) = \theta^{-1^T}(T_f, T_o) \qquad (3.65)$$

where $\theta(T_f, T_o)$ is the transition matrix of F(t). Consequently, from Eq. 3.25 we obtain

$$A_o = \lim_{z \to 0} \det[A(\frac{1}{z})] = \det[\theta^{-1^T}(T_f, T_o)] = [\det[\theta(T_f, T_o)]]^{-1} \qquad (3.66)$$

Since a transition matrix is nonsingular, its determinant cannot change sign, i.e., it must always be positive. This implies that A_o is positive, and we can take its logarithm. This yields[23]

$$\ell n(A_o) = -\ell n(\det[\theta(T_f, T_o)]) =$$

$$\int_{T_o}^{T_f} \frac{d}{dt} \ell n(\det[\theta(t, T_o)])dt =$$

$$\int_{T_o}^{T_f} Tr[\theta^{-1^T}(t, T_o)(\frac{d\theta(t, T_o)}{dt})^T]dt =$$

$$- \int_{T_o}^{T_f} \mathrm{Tr}[\, F(t)]\, dt \tag{3.67}$$

Therefore, we finally have

$$A_o = e^{- \int_{T_o}^{T_f} \mathrm{Tr}[\, F(t)]\, dt} \tag{3.68}$$

and

$$D_{\mathcal{F}}(z) = e^{\int_{T_o}^{T_f} \mathrm{Tr}[F(t)]\, dt} \det[\, A(-\tfrac{1}{z})] \tag{3.69}$$

For many of our problems, $D_{\mathcal{F}}(z)$ is not the most convenient function to use; instead, we shall use the function

$$d(\lambda) \triangleq \frac{\det[\, A(-\lambda)]}{A_o} = \prod_{i=1}^{\infty} (1 + \frac{\lambda_i}{\lambda}), \tag{3.70}$$

from which $D_{\mathcal{F}}(z)$ can be quickly determined. The function $d(\lambda)$ approaches one as $\lambda \to +\infty$ or $-\infty$. For positive values of λ it is a monotonically decreasing function of λ, while for negative values of λ it has the same behavior as we have earlier discussed regarding the behavior of $\det[\, A(\lambda)]$.

Before proceeding, we shall pause to mention an

asymptotic expression for $d(\lambda)$ when $\underline{y}(t)$ is a stationary process and the time interval, or 2WT product is large. Under these assumptions it is easy to show that

$$\ell n\, d(\lambda) \simeq T \int_{-\infty}^{\infty} \ell n\left(1 + \frac{S_y(\omega)}{\lambda}\right) d\omega \tag{3.71}$$

This formula allows use to determine the asymptotic behavior of $d(\lambda)$ in our calculations.

Let us briefly determine $d(\lambda)$ for three of the examples in the previous section. As before, we shall do two examples analytically and the third by numerical methods.

Example 5 - $d(\lambda)$ for a Wiener Process

Since $F = 0$, we have $A_o = 1$. From Eq. 3.38 we have,

$$d(\lambda) = \det A(-\lambda) = \cosh\left(\frac{\mu}{\sqrt{\lambda}}\, T\right). \tag{3.72}$$

Example 6 - $d(\lambda)$ for a One Pole, Stationary Process

Using Eq. 2.17 in Eq. 3.63 we find

$$A_o = e^{+kT} \tag{3.73}$$

To determine $\det[A(-\lambda)]$, we use Eq. 3.44. Let us define

$$\beta = \left[1 + \frac{2S}{k\lambda}\right]^{\frac{1}{2}}. \tag{3.74}$$

After some routine algebra, we obtain

$$d(\lambda) = e^{-kT}\left[\left(\frac{\beta^2+1}{2}\right)\frac{\sinh(k\beta T)}{} + \cosh(k\beta T)\right] =$$

$$\frac{e^{-kT}}{2\beta}\left[(\beta+1)^2 e^{k\beta T} - (\beta-1)^2 e^{-k\beta T}\right] \qquad (3.75)$$

Example 7 - d(λ) for a Two Pole, Stationary Process

Let us find d(λ) for the same stationary kernel considered in Example 4. As in that example we shall use a numerical analysis. From Eq. 3.68 we find

$$A_o = e^4 \qquad (3.76)$$

Next, we use the same computational method that we used to determine the eigenvalues; however, we must use a negative argument. In Fig. 3.3a we have plotted the resulting d(λ) with solid lines for time interval lengths of 1, 2 and 3. As a means of comparing our results with those derived by assuming the time interval is large, we have plotted the results indicated by Eq. 3.71 with dotted lines. We can see that for $T \geq 2$, we are very close to the asymptotic results indicated by a stationary process, large time interval analysis.

Finally, the function log (d(λ)) is often used in communications, e.g., as it related to the realizable fitter error. In Fig. 3.3b we have plotted log(d(x)) vs. λ as found by our exact analysis, the solid lines, and the approximate results found using Eq. 3.71. Again we have the close comparison for $T \geq 2$.

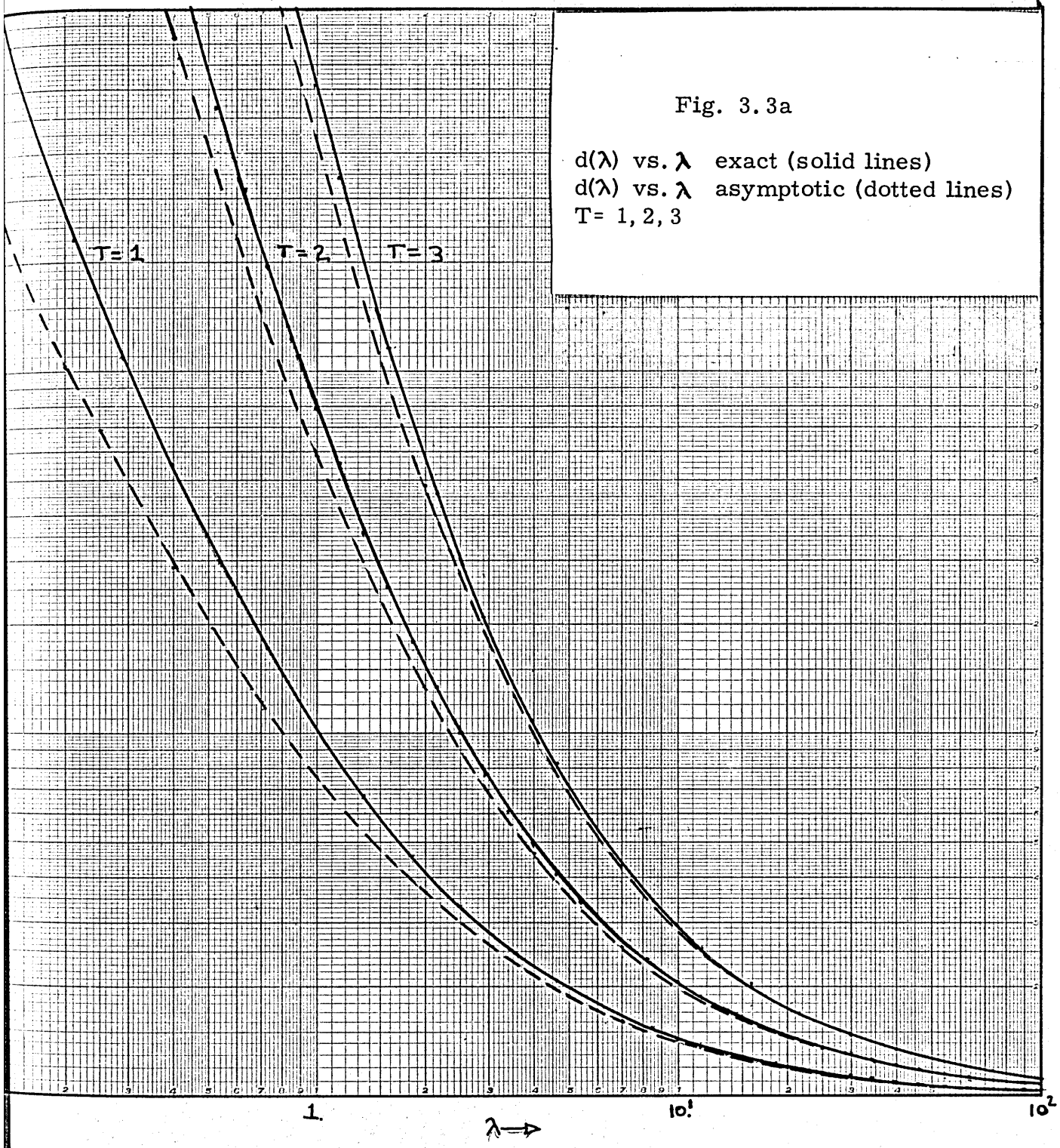


Fig. 3.3a

$d(\lambda)$ vs. λ exact (solid lines)
$d(\lambda)$ vs. λ asymptotic (dotted lines)
T= 1, 2, 3

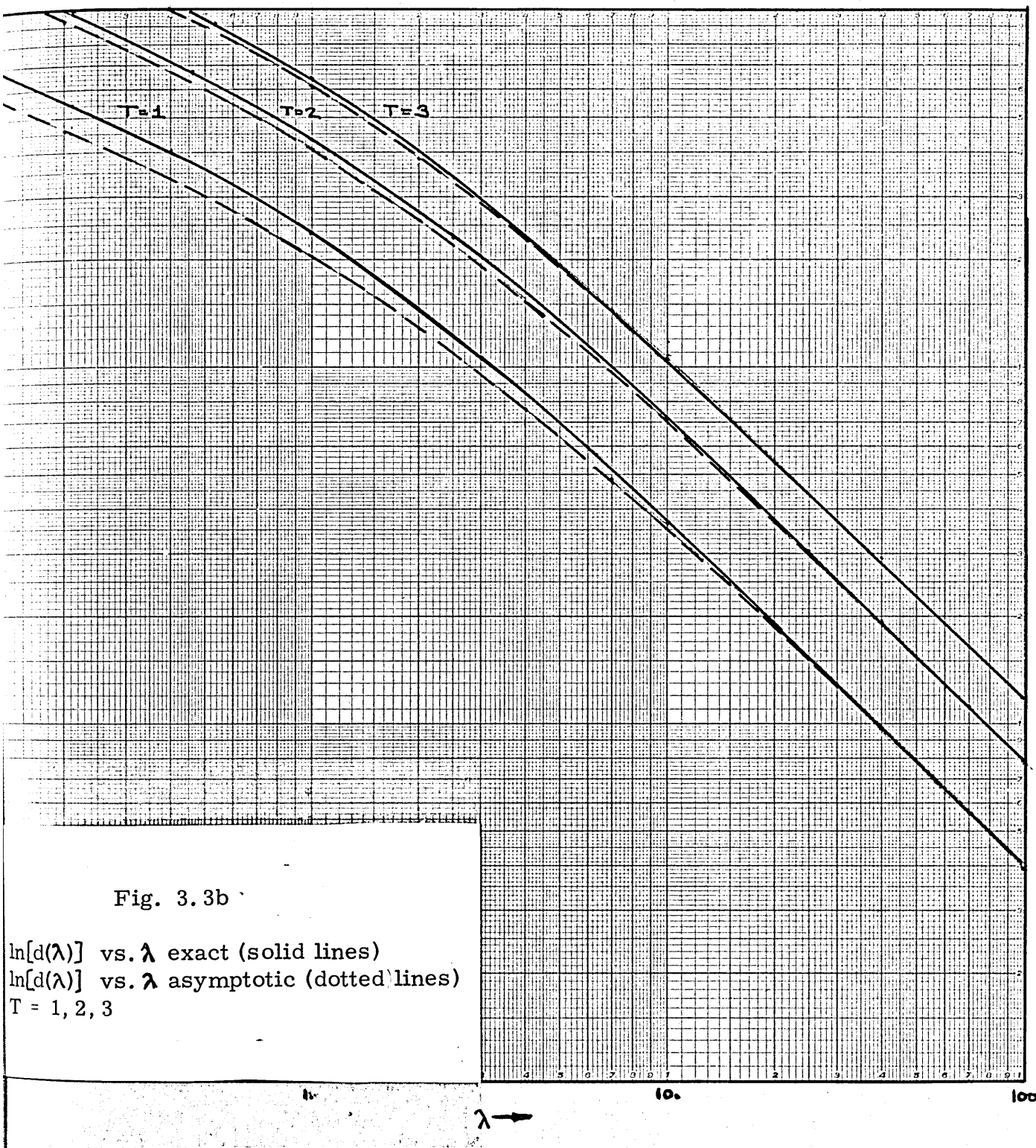


Fig. 3.3b

$\ln[d(\lambda)]$ vs. λ exact (solid lines)
$\ln[d(\lambda)]$ vs. λ asymptotic (dotted lines)
$T = 1, 2, 3$

(This plot also indicates the behavior for large λ, whereas Fig. 3.3a did not.)

D. Discussion of Results for Homogeneous Fredholm Integral Equations

In this chapter we have formulated a state variable approach for solving homogeneous Fredholm integral equations. As we indicated earlier, the technique has several advantages particularly from a computational viewpoint. Consequently, in problems where we need to evaluate the eigenvalues and/or eigenfunctions directly, we have a very general method available which allows us to make just such an evaluation with a minimum of effort.

Quite often we do not need these solutions directly, but we require an expression involving them. In many cases, we should be able to determine such expressions in a closed form by using the theory that we have developed, for example, as we did with the Fredholm determinant function.

Proceeding with our discussion, several comments are in order regarding the desirability of having a convenient method of evaluating this Fredholm determinant function.

1. In the problem of detecting Gaussian signals in Gaussian noise, it enters in two ways. First, it appears as the bias in calculating the threshold of the likelihood receiver. More importantly, it is intimately involved in the calculation of performance bounds for this problem.[12]

2. In an estimation theory context if we consider the problem of estimating the parameters of a Gaussian process

(or the system identification problem), we can show that the Cramér Rao bound can be determined using this function.

Finally, the solution method we have developed is important in itself. As we shall see in Chapter V, the concept of requiring a determinant to vanish for the existence of a solution is important. In this chapter, we shall find a similar set of homogeneous differential equations and boundary conditions that specify the solution to the optimal signal design problem for additive colored noise problems. The method of solving these equations is exactly analogous to the eigenvalue problem in that it requires the vanishing of a determinant for a solution to exist.

Now that we have studied the solution of homogeneous Fredholm equations, let us examine the nonhomogeneous equation.

CHAPTER IV

NONHOMOGENEOUS FREDHOLM INTEGRAL EQUATIONS

Nonhomogeneous Fredholm integral equations are of considerable interest in communication theory. In this chapter we shall use the results of Chapter II to develop a state variable method of solving these equations. As before, there are some significant advantages to the methods introduced.

One of the more important applications of this integral equation is the problem of determining the optimal receiver for the detection of a known signal in the presence of additive colored noise. Since we shall study this problem in detail both in this chapter and in the next, we shall first pause briefly to review the communication model for this problem. This is not meant, however, to imply that this problem is the only place where we can apply our methods. Other applications include the solution of finite time Wiener-Hopf equations in order to find optimal estimators, as we shall do in Chapter VI, and the calculation of the Cramer-Rao bound for the estimation of signal parameters.

After this brief review we shall present our derivation. The results come quickly since we have derived the required formulae in Chapter II. The result of our derivation is a pair of vector differential equations and boundary conditions. Since these equations have appeared in the literature in another context, namely

the linear smoother, we have several solution methods available.[14,15,16] We shall devote a section to introducing these methods and comments upon their applicability.

Finally, we shall consider three examples. Two of these will be worked analytically while we shall resort to numerical procedures for the third.

A. Communication Model for the Detection of a Known Signal in Colored Noise

Let us briefly introduce the communication model for the problem of detecting a known vector signal in additive colored noise. We have illustrated the model in Fig. 4.1. We have a transmitter which on hypothesis 1 transmits $\underline{s}_1(t)$, while on hypothesis 0, it transmits $\underline{s}_0(t)$ over the time interval $T_0 \leq t \leq T_f$. For discussion purposes let us assume that $\underline{s}_1(t)$ is $\underline{s}(t)$ while $\underline{s}_0(t)$ is $-\underline{s}(t)$. The channel adds a vector colored Gaussian noise process to the signal. We assume that this colored noise consists of two independent components. The first component is a random process $\underline{y}(t)$ that is generated according to the methods we discussed in Chapter II. The second component is a white Gaussian process $\underline{w}(t)$ that has a covariance as specified by Eq. 2.7. Consequently, we have the following detection problem

$$\text{on } H_1 \quad \underline{r}(t) = \underline{s}(t) + \underline{y}(t) + \underline{w}(t), \qquad T_0 \leq t \leq T_f,$$

$$\text{on } H_0 \quad \underline{r}(t) = -\underline{s}(t) + \underline{y}(t) + \underline{w}(t), \qquad T_0 \leq t \leq T_f. \tag{4.1}$$

Under these assumptions it is straightforward to show that the optimal receiver can be realized as indicated in Fig. 4.1 This realization is a correlation receiver. We multiply (dot product) the received signal $\underline{r}(t)$ with a function $\underline{g}(t)$, and then integrate over the observation interval, i.e., the sufficient statistic for the decision device is

$$\ell(\underline{r}) = \int_{T_o}^{T_f} \underline{r}^T(\tau)\underline{g}(\tau)d\tau \tag{4.2}$$

The correlating signal is the solution to a nonhomogeneous Fredholm integral equation.

This integral equation has the form

$$\int_{T_o}^{T_f} K_{\underline{y}}(t,\tau)\underline{g}(\tau)d\tau + R(t)\underline{g}(t) = \underline{s}(t), \quad T_o \leq t \leq T_f; \tag{4.3}$$

where $K_{\underline{y}}(t,\tau)$, the kernel of the equation, is the covariance of the random process $\underline{y}(t)$, which we assumed is generated according to the methods discussed in Chapter II;

$\underline{s}(t)$ is a known vector function, the transmitted signal;

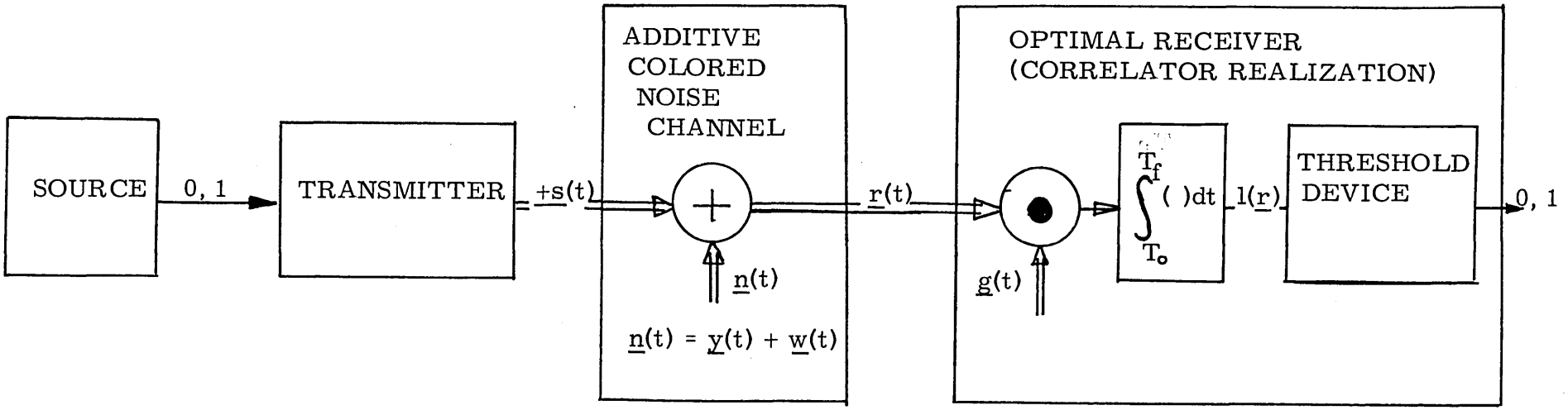


Fig. 4.1

System Model for the Detection
of a Known Signal in Colored Noise

R(t) is the covariance weighting matrix of $\underline{w}(t)$, which is assumed to be positive definite;

and

$\underline{g}(t)$ is the desired solution, the correlating signal.

One can also find the performance measure for this system. It is again straightforward to show that this measure, usually termed d^2, is given by

$$d^2 = \int_{T_o}^{T_f} \underline{s}^T(\tau)\underline{g}(\tau)d\tau \tag{4.4}$$

where s(t) and $\underline{g}(t)$ are defined above. Error probabilities, detection and false alarm probabilities can all be determined in terms of this measure.

As we metnioned earlier we shall study the nonhomogeneous integral Eq. 4.3 in the context of this detection problem. We emphasize, however, that the techniques developed are general in that they do not need to be considered in this particular context. Let us now proceed to develop our solution method.

B. The State Variable Approach to Nonhomogeneous Fredholm Integral Equations

Let us make two remarks regarding the solution. First, in contrast to the homogeneous equation that has an at most countable number of solutions, this equation has a unique solution

when R(t) is positive definite. Second, we may find a series solution in terms of the eigenvalues and eigenfunctions of the homogeneous equation; for example, when there is equal white noise level σ in each component channel, we have

$$\underline{g}(t) = \sigma^{-1}\left[\, s(t) - \sum_{i}^{\infty} \frac{\lambda_i s_i}{\sigma + \lambda_i} \underline{\phi}_i(t)\right], \quad T_o \le t \le T_f, \tag{4.5}$$

(the general vector case requires a simple, but straightforward modification).

Let us now proceed with the derivation. We rewrite Eq. 4.3. (The first part of the derivation is due to a suggestion by L. D. Collins.[24]) By using Eq. 2.9, we have

$$\underline{g}(t) = R^{-1}(t)\left[\, \underline{s}(t) - C(t) \int_{T_o}^{T_f} K_{\underline{x}}(t, \tau) C^T(\tau) \underline{g}(\tau) d\tau\right], \quad T_o \le t \le T_f. \tag{4.6}$$

If in Eq. 4.6 we set

$$\underline{g}(t) = \underline{f}(t), \tag{4.7}$$

we have the result that the integral enclosed in parantheses is the function $\underline{\xi}(t)$ as defined in Eq. 2.22. Consequently, we define $\underline{\xi}(t)$ to be

$$\underline{\xi}(t) = \int_{T_o}^{T_f} K_{\underline{x}}(t,\tau)C^T(\tau)\underline{g}(\tau)d\tau, \quad T_o \leq t \leq T_f, \tag{4.8}$$

so that Eq. 4.3 becomes

$$\underline{g}(t) = R^{-1}(t)[\underline{s}(t) - C(t)\underline{\xi}(t)], \qquad T_o \leq t \leq T_f. \tag{4.9}$$

For the class of $K_{\underline{y}}(t,\tau)$ that we are considering, we have shown in Chapter II that the functional defined by Eq. 4.8 may be represented as the solution to the following differential equations:

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\underline{\xi}(t) + G(t)Q\,G^T(t)\underline{\eta}(t), \quad T_o \leq t \leq T_f \tag{4.10}$$

$$\frac{d\underline{\eta}(t)}{dt} = -C^T(t)\underline{g}(t) - F^T(t)\underline{\eta}(t), \qquad T_o \leq t \leq T_f \tag{4.11}$$

plus a set of boundary conditions. If we substitute Eq. 4.9 in Eq. 4.11, we obtain

$$\frac{d\underline{\eta}(t)}{dt} = C^T(t)R^{-1}(t)C(t)\underline{\xi}(t) - F^T(t)\underline{\eta}(t) \tag{4.12}$$

$$- C^T(t)R^{-1}(t)\underline{s}(t) \qquad T_o \leq t \leq T_f$$

Consequently, we have shown that the nonhomogeneous Fredholm equation can be reduced to the following set of differential equations and associated boundary conditions:

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\xi(t) + G(t)QG^T(t)\underline{\eta}(t), \quad T_o \le t \le T_f, \tag{4.10 (repeated)}$$

$$\frac{d\underline{\eta}(t)}{dt} = C^T(t)R^{-1}(t)C(t)\underline{\xi}(t) - F^T(t)\underline{\eta}(t) - \frac{C^T(t)}{\sigma}\underline{s}(t) \quad T_o \le t \le T_f \tag{4.12 (repeated)}$$

$$\underline{\xi}(T_o) = P_o\underline{\eta}(T_o), \tag{4.13}$$

$$\underline{\eta}(T_f) = 0. \tag{4.14}$$

The desired solution is given by Eq. 4.3 to be

$$\underline{g}(t) = R^{-1}(t)[\underline{s}(t) - C(t)\underline{\xi}(t)]; \quad T_o \le t \le T_f. \tag{4.9 (repeated)}$$

Quite often it will be convenient to write Eq. 4.10 and Eq. 4.12 as one differential equation in the form

$$\frac{d}{dt}\begin{bmatrix}\underline{\xi}(t)\\ \hline \underline{\eta}(t)\end{bmatrix} = W(t)\begin{bmatrix}\underline{\xi}(t)\\ \hline \underline{\eta}(t)\end{bmatrix} - \begin{bmatrix}0\\ \hline C^T(t)R^{-1}(t)\underline{s}(t)\end{bmatrix} \tag{4.15}$$

$$T_o \le t \le T_f,$$

where we define w(t) to be

$$w(t) \triangleq \left[\begin{array}{c|c} F(t) & G(t)QG^T(t) \\ \hline C^T(t)R^{-1}(t)C(t) & -F^T(t) \end{array}\right] \tag{4.16}$$

We note that the coefficient matrix of Eq. 4.16, W(t) is similar to that which appeared in the homogeneous equation. The major difference is that positive definite matrix R(t) appears where $-\lambda$ did. We also note from our discussion of the Fredholm determinant that one wants to find the transition matrix of W(t) in order to compute $D_{\mathcal{F}}(\lambda)$.

The solution indicated by Eq. 4.9 has an appealing interpretation when one calculates d^2 as given by Eq. 4.4. For simplicity, let us again assume $R(t) = \sigma I$, σ a scalar. Substituting Eq. 4.9 into Eq. 4.4 gives us

$$d^2 = \int_{T_o}^{T_f} \frac{\underline{s}^T(t)\underline{s}(t)}{\sigma}\, dt - \int_{T_o}^{T_f} \frac{\underline{s}^T(t)C(t)\underline{\xi}(t)}{\sigma}\, dt. \tag{4.17}$$

The first term is simply the pure white noise performance, d_w^2. The second term represents the degradation, d_g^2, caused by the presence of colored noise in the observation. Therefore, we have

$$d^2 = d_w^2 - d_g^2 \tag{4.18a}$$

$$d_w^2 = \int_{T_o}^{T_f} \left(\frac{\underline{s}^T(t)\underline{s}(t)}{\sigma} \right) dt \triangleq \frac{E}{\sigma} \tag{4.18b}$$

$$d_g^2 = \int_{T_o}^{T_f} \left(\frac{\underline{s}^T(t)C(t)\underline{\xi}(t)}{\sigma} \right) dt. \tag{4.18c}$$

In the next chapter we shall consider the problem of maximizing the performance by choosing s(t) when it is subject to energy and band width constraints.

C. Methods of Solving the Differential Equations for the Nonhomogeneous Fredholm Equation

In the last section we derived a pair of vector differential equations that implicitly specified the solution of Eq. 4.3 As we shall see in Chapter VI, these equations appear in the optimal smoother, or the state variable formulation of the unrealizable filter*. In this section we shall exploit the methods that have been developed in the literature for solving the smoothing equations in order to solve the differential equations for the nonhomogeneous equation. Since these methods have evolved from the smoothing theory literature, the material in this section draws heavily upon References 14 and 15.

Let us outline our approach. We shall develop three methods, each of which has a particular application. The first method is useful for obtaining analytic solutions. The second is intermediate result in the development of the third method. It is useful because it introduces some important concepts and results.

The third method is applicable to finding numerical solutions since the first two methods have some undesireable aspects. We should also note that the methods build upon one

---

*This is consistent with the estimator-subtractor realization of the optimal receiver for detecting $\underline{s}(t)$.[3]

another. Consequently, one needs to read the entire section to understand the development of the last method.

Before proceeding, let us summarize the results from the previous section that we need and introduce some notation that we shall require. We want to solve the differential equations

$$\frac{d}{dt}\begin{bmatrix}\underline{\xi}(t)\\ \hline \underline{\eta}(t)\end{bmatrix} = W(t)\begin{bmatrix}\underline{\xi}(t)\\ \hline \underline{\eta}(t)\end{bmatrix} - \begin{bmatrix}0\\ \hline C^T(t)R^{-1}(t)\underline{s}(t)\end{bmatrix}, \quad T_o \le t \le T_f, \tag{4.15 (repeated)}$$

where we have defined

$$W(t) = \begin{bmatrix} F(t) & G(t)QG^T(t) \\ C^T(t)R^{-1}(t)C(t) & -F^T(t)\end{bmatrix}, \tag{4.16 (repeated)}$$

and we have imposed the boundary conditions

$$\underline{\xi}(T_o) = P_o\underline{\eta}(T_o), \tag{4.13 (repeated)}$$

$$\underline{\eta}(T_f) = \underline{0}. \tag{4.14 (repeated)}$$

Furthermore, let us introduce the following notation. We define the transition matrix of W(t) to be $\Psi(t,\tau)$, i.e.,

$$\frac{d}{dt}\Psi(t,\tau) = W(t)\Psi(t,\tau), \tag{4.19a}$$

$$\Psi(t,\tau) = I \,. \tag{4.19b}$$

In addition, let us partition this transition matrix into four n x n submatrices in the form

$$\Psi(t,\tau) = \left[\begin{array}{c|c} \Psi_{\underline{\xi}\,\underline{\xi}}(t,\tau) & \Psi_{\underline{\xi}\,\underline{\eta}}(t,\tau) \\ \hline \Psi_{\underline{\eta}\,\underline{\xi}}(t,\tau) & \Psi_{\underline{\eta}\,\underline{\eta}}(t,\tau) \end{array}\right] . \tag{4.20}$$

Method 1

The basic approach of the first method is to use the superposition of a particular and a homogeneous solution. First, we generate a convenient particular solution in order to incorporate the forcing term dependence. Then we add a homogeneous solution so as to satisfy the boundary conditions. In order to find a particular solution, let $\underline{\xi}_p(t)$ and $\underline{\eta}_p(t)$ be the solution to Eq. 4.15 with the initial conditions

$$\underline{\xi}_p(T_o) = \underline{\eta}_p(T_o) = \underline{0} \,. \tag{4.21}$$

Since we have specified a complete set of initial conditions we can uniquely solve the equation

$$\frac{d}{dt}\left[\begin{array}{c} \underline{\xi}_p(t) \\ \hline \underline{\eta}_p(t) \end{array}\right] = W(t)\left[\begin{array}{c} \underline{\xi}_p(t) \\ \hline \underline{\eta}_p(t) \end{array}\right] - \left[\begin{array}{c} \\ \hline C^T(t)R^{-1}(t)\underline{s}(t) \end{array}\right], \; T_o \leq t. \tag{4.22}$$

In order to match the boundary conditions, let us add to this particular solution a homogeneous solution of the form

$$\begin{bmatrix} \underline{\xi}_h(t) \\ \hline \underline{\eta}_h(t) \end{bmatrix} = \Psi(t, T_o) \begin{bmatrix} P_o \\ \hline I \end{bmatrix} \underline{\eta}_h(T_o), \tag{4.23}$$

where $\underline{\eta}_h(T_o)$ is to be chosen. Notice that the sum of these two solutions satisfies the initial boundary condition (Eq. 4.13) independent of $\underline{\eta}_h(T_o)$. Therefore, we want to chose $\underline{\eta}_h(T_o)$ such that we satisfy the final boundary condition (Eq. 4.14). To do this, let us rewrite Eq. 4.23 in the form

$$\begin{bmatrix} \underline{\xi}_h(t) \\ \hline \underline{\eta}_h(t) \end{bmatrix} = \begin{bmatrix} \Phi_{\underline{\xi}}(t, T_o) \\ \hline \Phi_{\underline{\eta}}(t, T_o) \end{bmatrix} \underline{\eta}_h(T_o), \tag{4.24}$$

where we define the matrices

$$\Phi_{\underline{\xi}}(t, T_o) \triangleq \Psi_{\underline{\xi}\,\underline{\xi}}(t, T_o) P_o + \Psi_{\underline{\xi}\,\underline{\eta}}(t, T_o), \tag{4.25}$$

$$\Phi_{\underline{\eta}}(t, T_o) \triangleq \Psi_{\underline{\eta}\,\underline{\xi}}(t, T_o) P_o + \Psi_{\underline{\eta}\,\underline{\eta}}(t, T_o). \tag{4.26}$$

Consequently, we have

$$\begin{bmatrix} \underline{\xi}(t) \\ \underline{\eta}(t) \end{bmatrix} = \begin{bmatrix} \underline{\xi}_p(t) + \Phi_{\underline{\eta}}(t, T_o)\underline{\eta}_h(T_o) \\ \hline \underline{\eta}_p(t) + \Phi_{\underline{\eta}}(t, T_o)\underline{\eta}_h(T_o) \end{bmatrix} \qquad (4.27)$$

$$T_o \leq t \leq T_f.$$

Applying the final boundary condition, requires that

$$\underline{\eta}_h(T_o) = -\Phi_{\underline{\eta}}^{-1}(T_f, T_o)\underline{\eta}_p(T_f) \qquad (4.28)$$

Substituting this in Eq. 4.27 gives us the final result for this method,

$$\underline{\xi}(t) = \underline{\xi}_p(t) - \Phi_{\underline{\xi}}(t, T_o)\Phi_{\underline{\eta}}^{-1}(T_f, T_o)\underline{\eta}_p(T_f), \quad T_o \leq t \leq T_f \qquad (4.29)$$

$$\underline{\eta}(t) = \underline{\eta}_p(t) - \Phi_{\underline{\eta}}(t, T_o)\Phi_{\underline{\eta}}^{-1}(T_f, T_o)\underline{\eta}_p(T_f). \quad T_o \leq t \leq T_f \qquad (4.30)$$

(The matrix $\Phi_{\underline{\eta}}(t, T_o)$ can be shown to be nonsingular for all t.[8])

Let us briefly summarize the method. First, we need to determine $\Phi_{\underline{\xi}}(t, T_o)$ and $\Phi_{\underline{\eta}}(t, T_o)$ as defined by Eqs. 4.19, 4.25 and 4.26. (This can be done independent of the signal, $\underline{s}(t)$.) We then find the particular solutions $\underline{\xi}_p(t)$ and $\underline{\eta}_p(t)$ by solving Eq. 4.15 with the initial conditions specified by Eq. 4.21. Finally, we substitute these functions into Eqs. 4.27 and 4.28 to find $\underline{\xi}(t)$ and $\underline{\eta}(t)$.

Two comments are in order. For a large class of problems the differential equations that one needs to solve using this method have constant coefficients. Consequently, the method is well suited for finding $\underline{\xi}(t)$ and $\underline{\eta}(t)$ analytically. We shall illustrate the use of this method in the next section with two examples.

We also observe that the differential equations we need to solve are unstable. e.g. if the system parameters are constant W has eigenvalues with the positive real parts. Consequently, one can (and does) encounter difficulty in numerically solving these equations when the time interval $[T_o, T_f]$ is long. This leads us to the problem of finding an effective numerical procedure for solving one equation.

In order to solve this problem we shall introduce two more methods. The first of these will develop some important concepts and results. The second shall use these concepts to develop the solution method which has the desired numerical properties.

Method 2

The most difficult aspect of solving Eq. 4.15 is satisfying the two point, or mixed, boundary conditions. The essential aspect of the second method is to introduce a third function from which we can determine $\underline{\xi}(T_f)$. Since $\underline{\eta}(T_f)$ is always identically zero, this allows us to specify a complete set of boundary conditions at $t = T_f$. With these conditions we then integrate Eq. 4.15 backwards over the interval.

From Method 1, let us define

$$\Sigma(t/t) \stackrel{\Delta}{=} \Phi_{\underline{\xi}}(t, T_o)\Phi_{\underline{\eta}}^{-1}(t, T_o)^{*}. \tag{4.31}$$

*The notation $\Sigma(t/t)$ is consistent with Chapter VI.

We shall find a matrix differential equation that $\Sigma(t/t)$ satisfies. We have

$$\frac{d\Phi_{\underline{\xi}}(t, T_o)}{dt} = \frac{d\Sigma(t/t)}{dt} \Phi_{\underline{\eta}}(t, T_o) + \Sigma(t/t) \frac{d\Phi_{\underline{\eta}}(t, T_o)}{dt} \qquad (4.32)$$

Substituting from Eqs. 4.25, 4.26, and 4.19, we find

$$F(t)\Phi_{\underline{\xi}}(t, T_o) + G(t)Q G^T(t)\Phi_{\underline{\eta}}(t, T_o) =$$

$$\frac{d\Sigma(t/t)}{dt} \Phi_{\underline{\eta}}(t, T_o) + \Sigma(t/t)(C^T(t)R^{-1}(t)C(t)\Phi_{\underline{\xi}}(t, T_o) - F^T(t)\Phi_{\underline{\eta}}(t, T_o)) \qquad (4.33)$$

Multiplying by $\Phi^{-1}_{\underline{\eta}}(t, T_o)$ and using Eq. 4.31 yields

$$\frac{d\Sigma(t/t)}{dt} = F(t)\Sigma(t/t) + \Sigma(t/t)F^T(t) - \Sigma(t/t)C^T(t)R^{-1}(t)C(t)\Sigma(t/t) + G(t)Q G^T(t) \qquad (4.34)$$

The initial condition follows from Eqs. 4.25 and 4.26

$$\Sigma(T_o/T_o) = P_o \qquad (4.35)$$

Consequently, we have the expected result that $\Sigma(t/t)$ is identical to the realizable filter covariance matrix since it satisfies the same matrix Ricatti differential equation and has the same initial condition.

Let us define a function $\underline{\xi}_r(t)$

$$\underline{\xi}_r(t) = \underline{\xi}_p(t) - \Phi_{\underline{\xi}}(t, T_o)\Phi_{\eta}^{-1}(t, T_o)\underline{\eta}_p(t) = \underline{\xi}_p(t) - \Sigma(t/t)\underline{\eta}_p(t) \tag{4.36}$$

We note that

$$\underline{\xi}(T_f) = \underline{\xi}_r(T_f) \tag{4.37}$$

Now we shall find a differential equation for $\underline{\xi}_r(t)$. Differentiating Eq. 4.36 and using Eq. 4.34, yields

$$\frac{d\underline{\xi}_r(t)}{dt} = F(t)\underline{\xi}_p(t) + G(t)QG^T(t)\underline{\eta}_p(t) -$$

$$\left(F(t)\Sigma(t/t) + \Sigma(t/t)F^T(t) + G(t)QG^T(t) - \Sigma(t/t)C^T(t)R^{-1}(t)C(t)\Sigma(t/t)\right)\eta_p(t)$$

$$- \Sigma(t/t)\left(C^T(t)R^{-1}(t)C(t)\underline{\xi}_p(t) - F^T(t)\underline{\eta}_p(t)\right.$$

$$\left. - C^T(t)R^{-1}(t)\underline{s}(t)\right) \tag{4.38}$$

After cancelling and combining terms by using Eq. 4.36 we have

$$\frac{d\underline{\xi}_r(t)}{dt} = F(t)\underline{\xi}_r(t) + \Sigma(t/t)G^T(t)R^{-1}(t)(\underline{s}(t) - G(t)\underline{\xi}_r(t)) \tag{4.39}$$

The initial condition follows from Eqs. 4.36, 4.31 and 4.22.

$$\underline{\xi}_r(T_o) = \underline{0} \tag{4.40}$$

From Eq. 4.38 we have the expected result that $\underline{\xi}_r(t)$ satisfies a differential equation of the same form as realizable filter estimator equation. We should note that in this particular application of solving the nonhomogeneous integral equation, the input $\underline{s}(t)$ is deterministic rather than a random process, *e.g.*, some received signal $\underline{r}(t)$.

Equations 4.34 and 4,38 are the key to the second method. We simply integrate them forward in time to $t = T_f$, apply Eq. 4.37 to find $\underline{\xi}(T_f)$, and then integrate Eq. 4.15 backwards in time using the complete set of boundary conditions at $t = T_f$. Expressed in terms of an integral operation we have

$$\begin{bmatrix} \underline{\xi}(t) \\ \hline \underline{\eta}(t) \end{bmatrix} = \Psi(t, T_f) \begin{bmatrix} \underline{\xi}_r(T_f) \\ \hline \underline{0} \end{bmatrix} + \int_t^{T_f} \Psi(t, t') \begin{bmatrix} \underline{0} \\ \hline C^T(t')R^{-1}(t')\underline{s}(t') \end{bmatrix} dt',$$

$$T_o \leq t \leq T_f. \qquad (4.41)$$

Let us examine this method for a moment. The basic approach was to convert a two point boundary problem into an initial, or final, value problem. Since the $\underline{\xi}_r(t)$ function that we developed for this conversion is the output of a realizable filter structure, it has many desirable properties. In particular, a lot is known regarding effective procedures for calculating $\underline{\xi}_r(t)$ numerically.

However, we will still have difficulty integrating

Eq. 4.41 backwards in time since Eq. 4.16 is also unstable when integrated backwards. For example, with constant parameter system W has eigenvalues in the left half plane. These produce growing exponentials as we integrate Eq. 4.16 backwards from the endpoint. With our third method we shall eliminate this undesirable feature.

Method 3

In this method we shall derive a result which allows us to uncouple the equations for $\underline{\xi}(t)$ and $\underline{\eta}(t)$. After we do this we shall observe that the resulting differential equations for $\underline{\xi}(t)$ and $\underline{\eta}(t)$ have some desirable features from a computational viewpoint. First, we need to derive one key result.

Let us consider the difference of $\underline{\xi}(t)$ and $\underline{\xi}_r(t)$. Substituting Eqs. 4.39 and 4.36

$$\underline{\xi}(t) - \underline{\xi}_r(t) = \Phi_{\underline{\xi}}(t, T_o)\{-\Phi_{\underline{\eta}}^{-1}(T_f, T_o)\underline{\eta}_p(T_f) + \Phi_{\underline{\eta}}^{-1}(t, T_o)\underline{\eta}_p(t)\} =$$

$$\Phi_{\underline{\xi}}(t, T_o)\Phi_{\underline{\eta}}^{-1}(t, T_o)\{\underline{\eta}_p(t) - \Phi_{\underline{\eta}}(t, T_o)\Phi_{\underline{\eta}}^{-1}(T_f, T_o)\underline{\eta}_p(T_f)\} =$$

$$\Sigma(t/t)\underline{\eta}(t) \qquad (4.42)$$

Consequently, we have the result

$$\underline{\xi}(t) - \underline{\xi}_r(t) = \Sigma(t/t)\underline{\eta}(t), \qquad (4.43)$$

or

$$\underline{\eta}(t) = \Sigma^{-1}(t/t)(\underline{\xi}(t) - \underline{\xi}_r(t)). \qquad (4.44)$$

If we substitute this into Eq. 4.10 we can obtain separate the differential equations for $\underline{\xi}(t)$ and $\underline{\eta}(t)$. We find substituting for $\underline{\eta}(t)$

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\underline{\xi}(t) + G(t)Q\,G^T(t)\Sigma^{-1}(t/t)(\underline{\xi}(t) - \underline{\xi}_r(t)) =$$

$$(F(t) + G(t)Q\,G^T(t)\,\Sigma^{-1}(t/t))\underline{\xi}(t) - G(t)Q\,G^T(t)\Sigma^{-1}(t/t)\underline{\xi}_r(t),$$

$$T_o \le t \le T_f. \qquad (4.45)$$

Similarly, substituting for $\underline{\xi}(t)$ yields

$$\frac{d\underline{\eta}(t)}{dt} = C^T(t)R^{-1}(t\,C(t)(\Sigma(t/t)\underline{\eta}(t) + \underline{\xi}_r(t))$$

$$-F^T(t)\underline{\eta}(t) - C^T(t)R^{-1}(t)\underline{s}(t) =$$

$$-(F(t)-\Sigma(t/t)G^T(t)R^{-1}(t)G(t)^T\underline{\eta}(t) - {}^T(t)R^{-1}(t)(s(t)-C(t)\underline{\xi}_r(t)),$$

$$T_o \le t \le T_f. \qquad (4.46)$$

We now note that by finding $\underline{\xi}_r(t)$ we can solve either Eq. 4.45 or Eq. 4.46 for $\underline{\xi}(t)$ or $\underline{\eta}(t)$ respectively. The initial (or final) condition for Eq. 4.45 is given by Eq. 4.37 while for Eq. 4.46 is given by Eq. 4.14. Either function can be obtained from the other by using Eq. 4.43.

Now let us examine the stability aspects of these equations when integrated backwards. Our discussion is essentially qualitative. First, we need to examine the coefficient matrices of

Eq. 4.45 and Eq. 4.46. In Eq. 4.46, this matrix is $-(F(t) - \Sigma(t/t)C^T(t)R^{-1}(t)C(t))^T$, which is the negative transpose of the coefficient matrix of the realizable filter. Consequently, if it is stable, so is Eq. 4.46 when integrated backwards.

For constant parameter systems, we can see heuristically that Eq. 4.45 is also stable when integrated backwards over large time intervals. If the interval is "long",

$$\frac{d\Sigma(t/t)}{dt} \cong 0 \tag{4.47}$$

over most of the interval. If we assume equality, i.e., $\Sigma(t/t) = \Sigma_\infty$, we have

$$F + G Q G^T \Sigma_\infty = \Sigma_\infty [-(F - \Sigma_\infty C^T R^{-1} C)^T] \Sigma_\infty^{-1}; \tag{4.48}$$

or, $F + G Q G \Sigma_\infty^{-1}$ is similar to $-(F-\Sigma_\infty C^T R^{-1} C)^T$. This implies that both matrices have the same eigenvalues; therefore, Eq.4.45 is also stable when integrated backwards over the interval. Consequently, one can numerically solve either Eq. 4.45 or Eq. 4.46 and obtain stable solutions. However, if $\overset{\vee}{\underset{\wedge}{\Sigma}}(t/t)$ is also available we should point out that Eq. 4.45 requires its inversion whereas Eq. 4.46 does not.

Summary of Methods

In this section we have developed in considerable detail methods which exist for solving the differential equations we derived

for the nonhomogeneous Fredholm integral equation. Let us now summarize these methods by discussing their applicability.

If we have a constant parameter system and want to find an analytical solution, method 1 is probably most useful since the differential equations have constant coefficients for a large class of problems. However, if we want to obtain a numerical solution, especially over a long time interval, method 3 is probably the best, since methods 1 and 2 can create some difficulties when integrated numerically. We really never use method 2. The essential reason for introducing it is that it was an intermediate result in our derivation of method 3.

Let us now apply the results of this section to analyze some examples of solving nonhomogeneous Fredholm equations with our method developed in Section B of this chapter.

D. Examples of Solutions to the Nonhomogeneous Fredholm Equation

In this section we shall consider three examples to illustrate the results of the last two sections. Again, we shall work two examples analytically while we shall use numerical procedures for the third. In those examples that we work analytically, we shall use method 1 as discussed in the last section. The computer program used in the numerical example integrated the second differential equation of method 3.

We shall present the examples in the context of detection in colored noise. After determining the solution $g(t)$, we shall compute the d^2 (and d_g^2) performance measures discussed at the end

of the previous section. Finally, in all three examples we shall set $T_o = 0$ and $T_f = T$ and assume that the forcing function, or transmitted signal, $s(t)$ is a pulsed sine wave with unit energy over this interval, i.e.,

$$s(t) = \sqrt{\frac{2}{T}} \sin\left(\frac{n\pi t}{T}\right), \quad 0 \le t \le T. \tag{4.49}$$

We also assume that $R(t)$, or the white noise level, is a scalar constant

$$R(t) = \sigma > 0 \tag{4.50}$$

Example 1 - g(t) for a Wiener Process

Let us consider the problem of finding $g(t)$ when the kernel is the covariance of a Wiener process. Equations 3.35 and 3.36 describe a system for generating this process. First, we substitute the parameters of these equations into Eq. 4.15 in order to find the equation that we need to solve. We obtain

$$\frac{d}{dt}\begin{bmatrix}\xi(t)\\ \hline \eta(t)\end{bmatrix} = \left[\begin{array}{c|c} 0 & 1 \\ \hline \frac{\mu^2}{\sigma} & 0 \end{array}\right]\begin{bmatrix}\xi(t)\\ \hline \eta(t)\end{bmatrix} - \left[\begin{array}{c} 0 \\ \hline \frac{\mu}{\sigma}\sqrt{\frac{2}{T}}\sin\left(\frac{n\pi t}{T}\right) \end{array}\right], \quad 0 \le t \le T \tag{4.51}$$

From Eq. 4.13 and Eq. 4.14 the boundary conditions are

$$\xi(0) = 0, \tag{4.52}$$

$$\eta(T) = 0. \tag{4.53}$$

Referring to the last section, method 1, we find particular solutions $\xi_p(t)$ and $\eta_p(t)$. These are the solutions to Eq. 4.15 with $\xi_p(0) = \eta_p(0) = 0$. Doing this we obtain

$$\begin{bmatrix} \xi_p(t) \\ \hline \eta_p(t) \end{bmatrix} = \frac{\mu}{\sigma\gamma^2}\sqrt{\frac{2}{T}} \begin{bmatrix} \sin(\frac{n\pi t}{T}) \\ \hline \frac{n\pi}{T}\cos(\frac{n\pi t}{T}) \end{bmatrix}, \tag{4.54}$$

where, for this problem, we define γ^2 to be

$$\gamma^2 = (\frac{n\pi}{T})^2 + \frac{\mu^2}{\sigma}. \tag{4.55}$$

Next, we need to find the transition matrix associated with Eq. 4.15. After some straightforward calculation we find

$$\Psi(t,0) = \left[\begin{array}{c|c} \cosh([\frac{\mu^2}{\sigma}]^{\frac{1}{2}} t) & [\frac{\mu^2}{\sigma}]^{-\frac{1}{2}} \sinh([\frac{\mu^2}{\sigma}]^{\frac{1}{2}} t) \\ \hline [\frac{\mu^2}{\sigma}]^{\frac{1}{2}} \sin h([\frac{\mu^2}{\sigma}]^{\frac{1}{2}} t) & \cosh([\frac{\mu^2}{\sigma}]^{\frac{1}{2}} t) \end{array}\right]. \tag{4.56}$$

We need to add a homogeneous equation of the form Eq. 4.23

$$\begin{bmatrix} \xi_h(t) \\ \eta_h(t) \end{bmatrix} = \begin{bmatrix} \left[\frac{\mu^2}{\sigma}\right]^{-\frac{1}{2}} \sinh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} t\right) \\ \cosh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} t\right) \end{bmatrix} \eta_h(0) \qquad (4.57)$$

From Eqs. 4.57 and 4.54 we find that we must choose $\eta_h(0)$ to be Eq. 4.28

$$\eta_h(0) = -\frac{\mu}{\sigma\gamma^2}\sqrt{\frac{2}{T}}\left(\frac{n\pi}{T}\right)(-1)^n \Big/ \cosh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} T\right) . \qquad (4.58)$$

Consequently, we have

$$\underline{\xi}(t) = \frac{\mu}{\sigma\gamma^2}\sqrt{\frac{2}{T}}\left\{\sin\left(\frac{n\pi t}{T}\right) - \left(\frac{n\pi}{T}\right)\left[\frac{\mu^2}{\sigma}\right]^{-\frac{1}{2}}(-1)^n\left[\frac{\sinh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} t\right)}{\cosh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} T\right)}\right]\right\} ,$$

$$0 \le t \le T \qquad (4.59)$$

The solution g(t) is found by substituting Eq. 4.59 into Eq. 4.9

$$g(t) = \sqrt{\frac{2}{T}} \left(\frac{n\pi}{T\gamma}\right)^2 \left\{ \sin\left(\frac{n\pi t}{T}\right) + \left(\frac{n\pi}{T}\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}}\right)^{-1} (-1)^n \cdot \left[\frac{\sinh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} t\right)}{\cosh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} T\right)}\right] \right\}, \quad 0 \le t \le T. \tag{4.60}$$

After some straightforward but tedious manipulation we can also calculate d_g^2 as given by Eq. 4.18c

$$d_g^2 = \frac{\mu^2}{\sigma\gamma^2} \left\{ 1 + \frac{2}{\gamma^2} \left(\frac{n\pi}{T}\right)^2 \frac{\tanh\left(\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} T\right)}{\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} T} \right\}. \tag{4.61}$$

Example 2 - g(t) for a One Pole Stationary Spectrum

Let us consider the kernel to be the covariance function for a one pole stationary process as described by Eqs. 2.16 and 2.17. From these equations and Eq. 4.15, the differential equations that we need to solve are

$$\frac{d}{dt}\begin{bmatrix} \xi(t) \\ \hline \eta(t) \end{bmatrix} = \left[\begin{array}{c|c} -k & 2kS \\ \hline 1/\sigma & k \end{array}\right] \begin{bmatrix} \xi(t) \\ \eta(t) \end{bmatrix} - \begin{bmatrix} 0 \\ \hline \frac{1}{\sigma}\sqrt{\frac{2}{T}} \sin\left(\frac{n\pi t}{T}\right) \end{bmatrix}, \quad 0 \le t \le T, \tag{4.62}$$

subject to the boundary conditions specified by Eqs. 4.13 and 4.14. We shall now use the first solution method which we discussed. After some manipulation, we find

$$\xi_p(t) = \sqrt{\frac{2}{T}}\,\frac{1}{\sigma\gamma^2}\left(2kS\sin\left(\frac{n\pi t}{T}\right) - \frac{S}{\lambda}\,\frac{n\pi}{T}\left(e^{k\lambda t} - e^{-k\lambda t}\right)\right), \quad (4.63)$$

$$\eta_p(t) = \sqrt{\frac{2}{T}}\,\frac{1}{\sigma\gamma^2}\left(k\sin\left(\frac{n\pi t}{T}\right) + \frac{n\pi}{T}\cos\left(\frac{n\pi t}{T}\right)\right.$$

$$\left. - \frac{n\pi}{2T}\left(1 - \frac{1}{\lambda}\right)e^{k\lambda t} + \frac{n\pi}{2T}\left(1 + \frac{1}{\lambda}\right)e^{-k\lambda t}\right), \quad (4.64)$$

where

$$\lambda = \left[1 + \frac{2S}{k\sigma}\right]^{\frac{1}{2}}, \quad (4.65)$$

$$\gamma^2 = \left(\frac{n\pi}{T}^2 + k^2\lambda^2\right). \quad (4.66)$$

In this example, the transition matrix, $\Psi(t, 0)$ that we need is

$$\Psi(t,0) = \left[\begin{array}{c|c} \frac{1}{2}\left(1-\frac{1}{\lambda}\right)e^{k\lambda t} + \frac{1}{2}\left(1+\frac{1}{\lambda}\right)e^{-k\lambda t} & \frac{S}{\lambda}\,e^{k\lambda t} - \frac{S}{\lambda}\,e^{-k\lambda t} \\ \hline \frac{1}{2k\lambda\sigma}\,e^{k\lambda t} - \frac{1}{2k\lambda\sigma}\,e^{-k\lambda t} & \frac{1}{2}\left(1+\frac{1}{\lambda}\right)e^{k\lambda t} + \frac{1}{2}\left(1-\frac{1}{\lambda}\right)e^{-k\lambda t} \end{array}\right] \quad (4.67)$$

Therefore, according to Eqs. 4.29 and 4.30

$$\xi(t) = \sqrt{\frac{2}{T}} \frac{1}{\sigma\gamma^2} \left(2kS \sin\left(\frac{n\pi t}{T}\right) - \frac{S}{\lambda} \frac{n\pi}{T} \left(e^{k\lambda t} - e^{-k\lambda t} \right) \right)$$

$$- \left(\frac{S}{2} \left(1 + \frac{1}{\lambda} \right) e^{k\lambda t} + \frac{S}{2} \left(1 - \frac{1}{\lambda} \right) e^{-k\lambda t} \right) n_h(0), \quad 0 \le t \le T \quad (4.68)$$

$$\eta(t) = \sqrt{\frac{2}{T}} \frac{1}{\sigma\gamma^2} \left(k \sin\left(\frac{n\pi t}{T}\right) + \frac{n\pi}{T} \cos\left(\frac{n\pi t}{T}\right) - \frac{n\pi}{2T} \left(1 - \frac{1}{\lambda} \right) e^{k\lambda t} - \frac{n\pi}{2T} \left(1 + \frac{1}{\lambda} \right) e^{-k\lambda t} \right)$$

$$- \left(\frac{(\lambda+1)^2}{4\lambda} e^{k\lambda t} - \frac{(\lambda-1)^2}{4\lambda} e^{-k\lambda t} \right) n_h(0), \quad 0 \le t \le T \quad (6.69)$$

where

$$\eta_h(0) = \left(\frac{(\lambda+1)^2}{4\lambda} e^{k\lambda T} - \frac{(\lambda-1)^2}{4\lambda} e^{-k\lambda T} \right)^{-1}$$

$$\sqrt{\frac{2}{T}} \frac{1}{\sigma\gamma^2} \frac{n\pi}{T} \left((-1)^n - \frac{1}{2} \left(1 + \frac{1}{\lambda} \right) e^{k\lambda T} - \frac{1}{2} \left(1 - \frac{1}{\lambda} \right) e^{-k\lambda T} \right).$$

(4.70)

Finally, using Eq. 4.9 we find the solution g(t)

$$g(t) = \sqrt{\frac{2}{T}} \frac{1}{\sigma\gamma^2} \left\{ \left[\left(\frac{n\pi}{T} \right)^2 + k^2 \right] \sin\left(\frac{n\pi t}{T}\right) - \right.$$

$$\frac{S}{\sigma} \frac{n\pi}{\lambda T} \left(\frac{(\lambda+1)^2}{2\lambda} e^{k\lambda T} - \frac{(\lambda-1)^2}{2\lambda} e^{-k\lambda T} \right)^{-1}$$

(continued)

$$x\left\{\left((\lambda+1)\ (-1)^n e^{k\lambda t}-e^{-k\lambda(T-t)}\ +(\lambda-1)\ (-1)^n e^{-k\lambda t}-e^{-k\lambda(T-t)}\right)\right\}$$

$$0 \le t \le T \qquad (4.71)$$

One can continue and evaluate the performance by computing d^2 and d_g^2 according to Eq. 4.18; however, the result is rather complex and is not too illustrative. Instead of presenting an analytic formula, we shall plot d^2 and d_g^2 against n for a particular choice of parameters. The results are presented in Fig. 4.2 when $k = 1$, $\sigma = 1$, $S = 1$, and $T = 2$. For the case $n = 1$, we see that the presence of the colored noise degrades our performance approximately 50 percent from that of the white noise. For $n = 8$, however, our performance is within 2 percent of the performance for the white noise only performance.

We can easily see that this is what we would intuitively expect. For $n = 1$, the bandwidth of the signal is approximately $\pi/T = 1.57$. Consequently, most of the signal energy appears in the frequencies where the spectrum of the colored component of the noise is significant compared to white noise level. For $n = 8$ we have a bandwidth of $8\pi/T = 12.56$. Therefore, most of its energy appears where the white noise $\underline{w}(t)$ is the dominant component of channel observation noise.

It should be apparent from the complexity of these two simple examples that an analytic solution for $g(t)$ (and the performance) is indeed a very difficult task when the kernel is the covariance of a process generated by higher order system. Consequently, it is desirable to have an efficient numerical method.

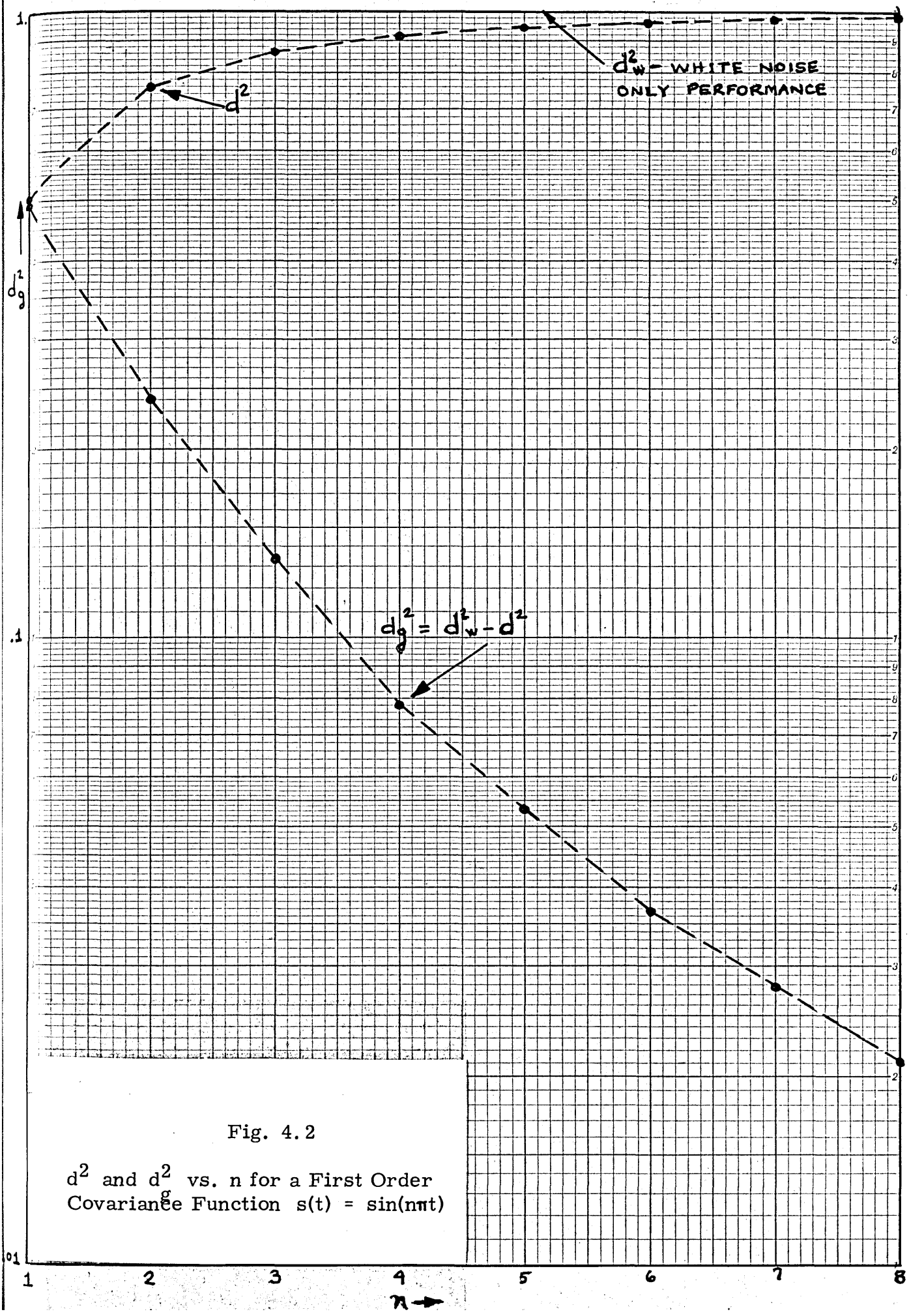


Fig. 4.2

d^2 and d_g^2 vs. n for a First Order Covariance Function $s(t) = \sin(n\pi t)$

Example 3 - g(t) for a Two Pole Stationary Spectrum

For the second example, we shall consider a numerical approach to the analysis of a second-order system. In particular, we choose $K_y(t,\tau)$ to be the covariance given by Eqs. 2.18 and 2.19 $\sigma = 1$ and $T = 2$. To find g(t) we determined $\xi(t)$ by using method 3 in the last section.

In Figs. 4.3 and 4.4 we have drawn the signal s(t) and the corresponding solution g(t) for n = 2 and n = 8. For the low-frequency (n = 2) case, we find that functionally s(t) and g(t) differ significantly only near the end points of the interval, while for the high-frequency (n = 8) case we find that s(t) and g(t) are nearly identical. Here, we are approaching the white noise, or matched-filter, solution. We have summarized the results for this example in Fig. 4.5 by plotting the d^2 vs n behavior. We see that for $n > 8$ we are within 4 percent of the white noise only performance.

Again we can see the effect of the colored noise upon the detection performance. For n = 2, most of the signal energy appears at the peak of the spectrum of the colored component of the noise. Consequently, the performance is degraded the most. For n = 8, the signal energy is centered around f = 12.56 where the white noise is dominant.

The computer time that we required to find a solution g(t) and to calculate its performance is approximately 5-10 sec of IBM 7094 computer time, using the Fortran language.

The method has recently been rewritten as a general purpose routine for the IBM 360.

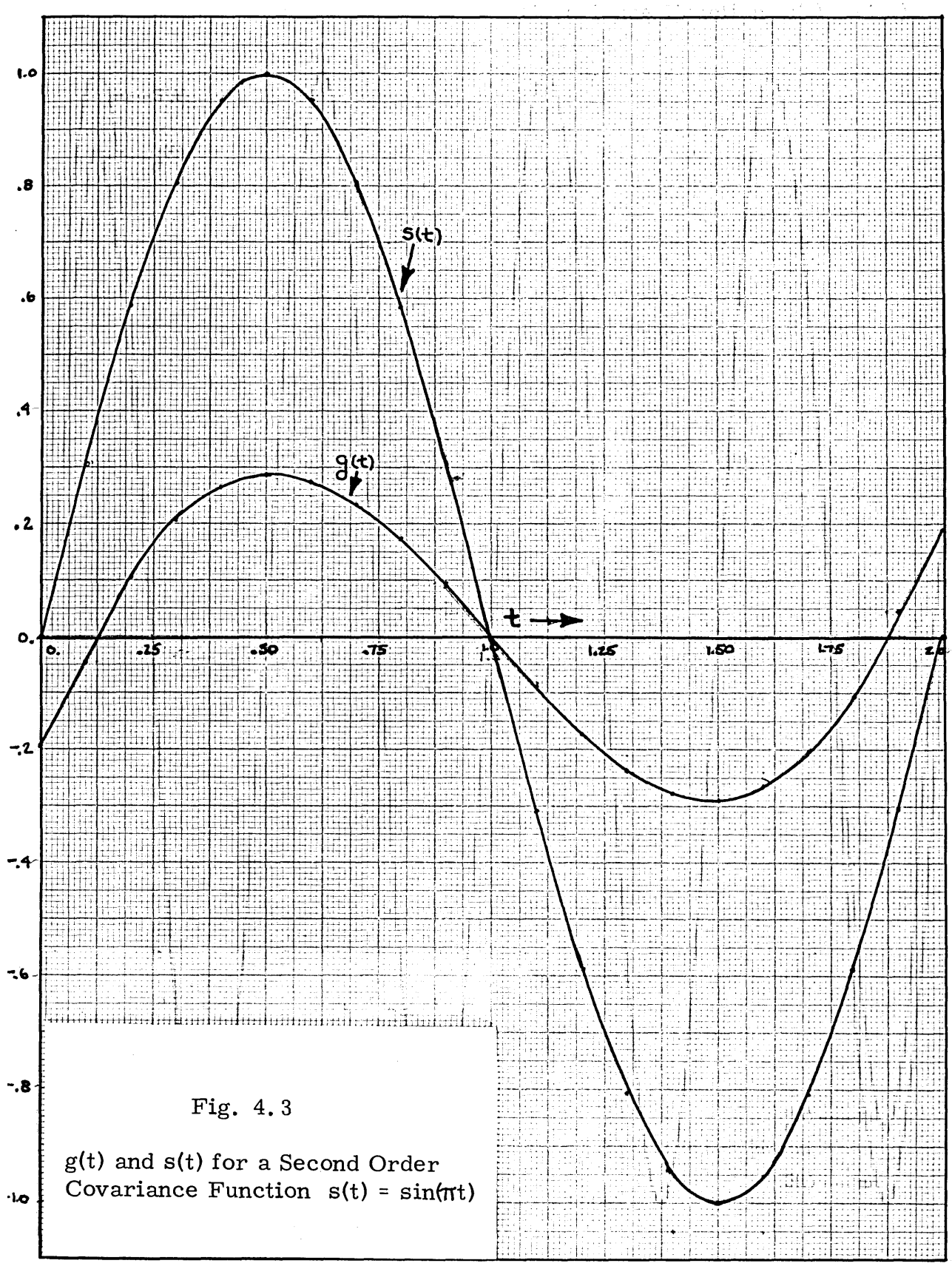
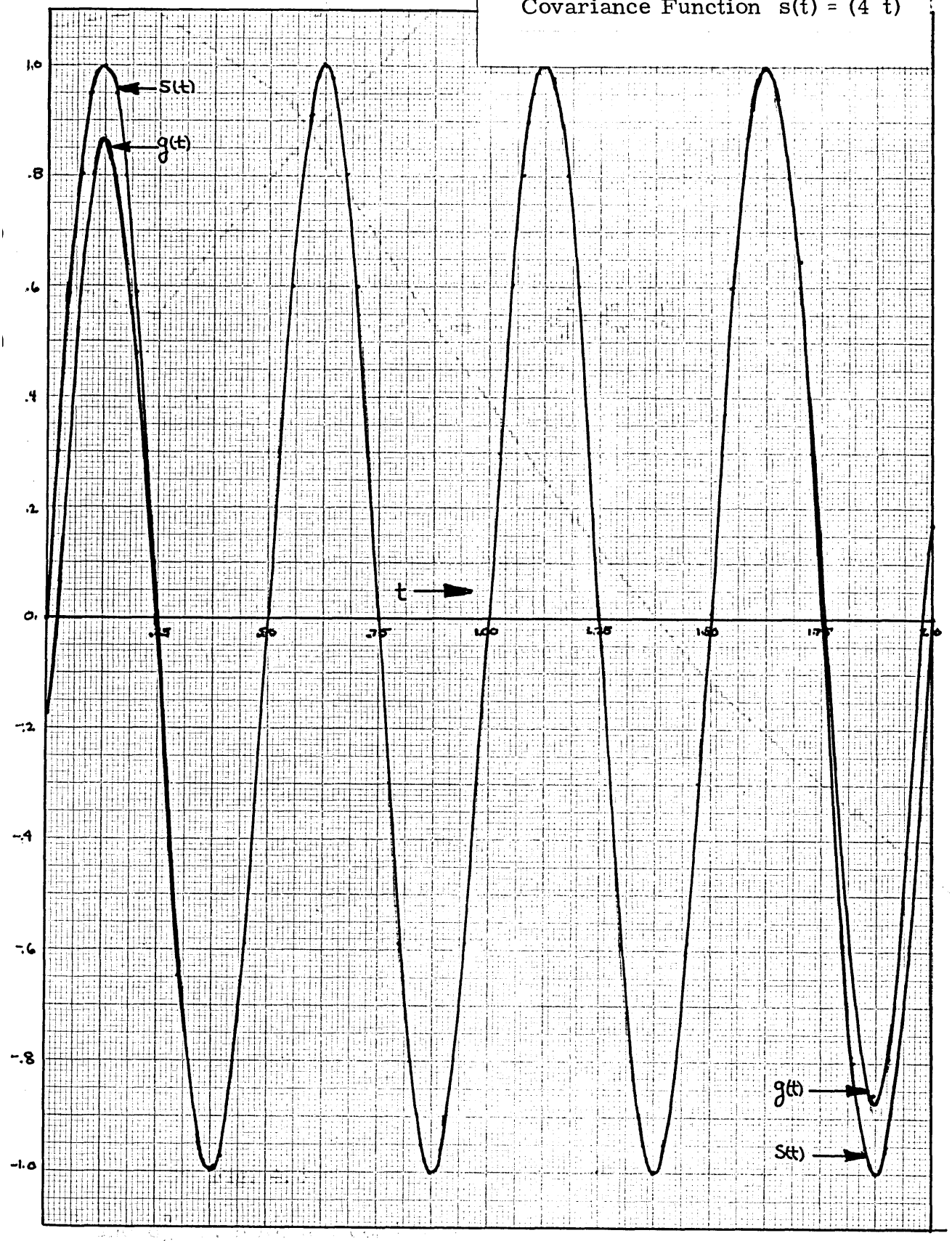


Fig. 4.3

g(t) and s(t) for a Second Order Covariance Function $s(t) = \sin(\pi t)$

Fig. 4.4

g(t) and s(t) for a Second Order Covariance Function s(t) = (4 t)



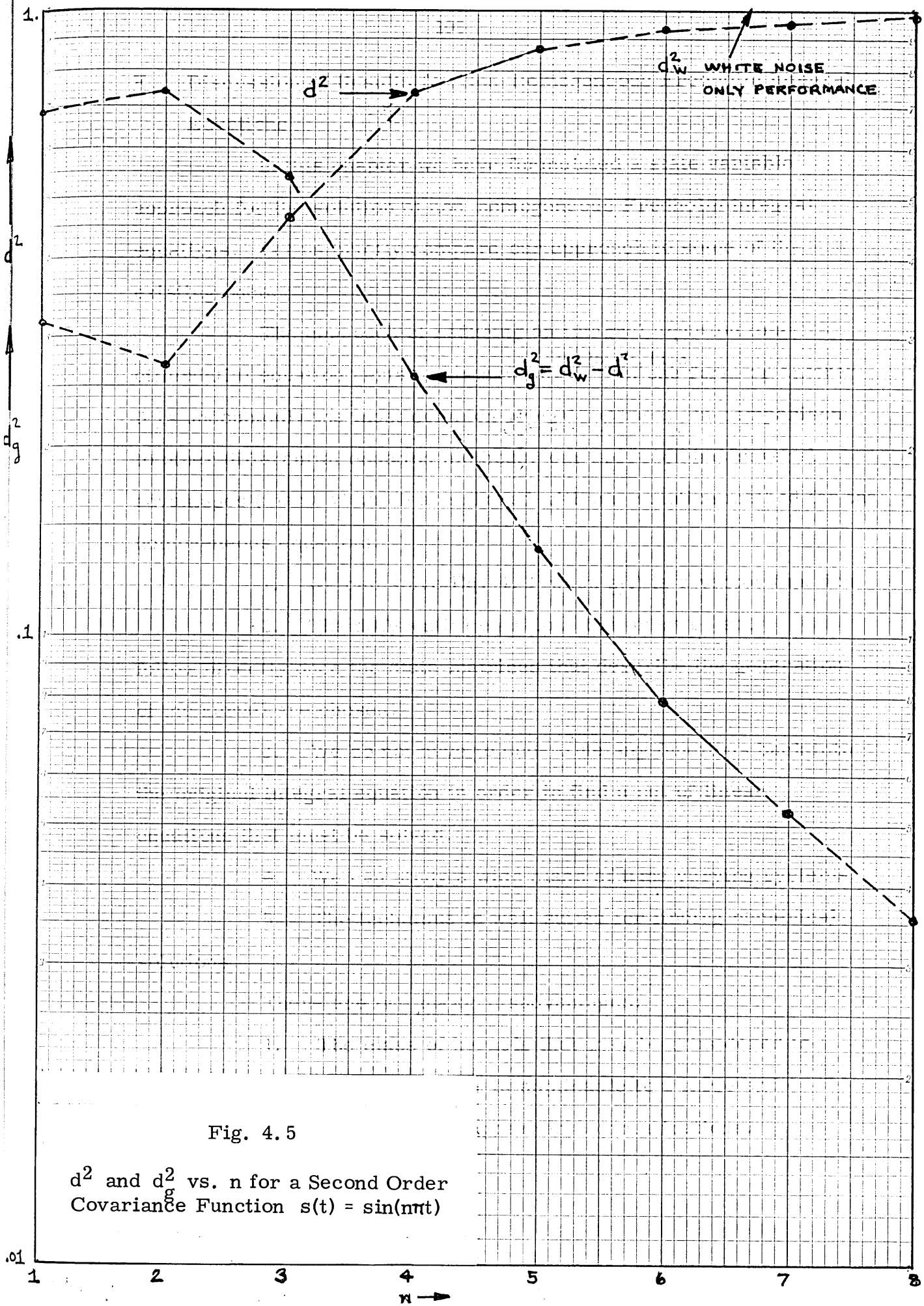


Fig. 4.5

d^2 and d_g^2 vs. n for a Second Order Covariance Function $s(t) = \sin(n\pi t)$

E. Discussion of Results for Non-Homogeneous Fredholm Integral Equations

In this chapter we have formulated a state variable approach for the solution of non-homogeneous Fredholm integral equations. Let us briefly compare our approach to some of the existing ones.

The approach of reducing an integral equation to a differential equation certainly is now new.[25,26,27,11,3] In one form or another it is undoubtedly the most common procedure used. In comparison to other differential equation methods, our approach has several advantages (many of which are shared with our solution method for the homogeneous equation).

1. We can solve Eq. 4.3 when $\underline{s}(t)$ is a vector function.

2. The differential equations that must be solved follow directly once the state matrices that describe the generation of the kernel are chosen.

3. We do not have to substitute any functions back into the original integral equation in order to find a set of linear equations that must be solved.

4. We can study a wide class of time varying systems.

5. The technique is well suited for finding numerical solutions.

There are two major disadvantages.

1. The class of kernels that may be considered is limited. However, the technique is applicable to a large and important class of kernels that appear in communications.

2. We cannot handle integral equations of the first kind, e.g. when the white component of the noise is identically zero.[3,27] For these equations singularity functions appear at the interval endpoints. We precluded these in our derivation. We should note that we have observed the limiting behavior of our solution approaching these singularity functions when the white noise is small.

A second method that is in contrast to the differential equation approaches is to find the inverse kernel of the integral equation. The inverse kernel $Q(\tau, u)$ is defined so that it satisfies a second integral equation

$$\int_{T_o}^{T_f} \{R(t)\delta(t-\tau) + K_{\underline{y}}(t, \tau)\} Q(\tau, u) d\tau = I\delta(t-u)$$

$$T_o < t,\ u < T_f \qquad (4.72)$$

In terms of the inverse kernel the solution $\underline{g}(t)$ is found to be

$$\underline{g}(t) = \int_{T_o}^{T_f} Q(t, u)\underline{s}(u) du, \quad T_o < t < T_f \qquad (4.73)$$

One common numerical method that uses this approach is to approximate the integral operations 4.72 and 4.73 by matrices

$$[K]\,[Q] = I \qquad (4.74)$$

$$\underline{g} = [Q]\underline{s} \qquad (4.75)$$

Let us briefly compare the computation required using this approach to that of our approach. If we assume that we sample the interval at NI points, Equations 4.73 and 4.74 are NI dimensional. One can show that the number of computations required to find [Q] as given by 4.73 goes as (NI) cubed. If we assume that we find Q by specifying the coefficients of our differential equations, the computations we require increases only linearly with NI. The computation required to implement Eq. 4.3 is proportional to NI squared, whereas the computations required to solve our equations again is linearly proportional to NI. The conclusion is that for large NI, which is required for high accuracy, the differential equation approach is superior.

Before we leave the topic of the inverse kernel, we shall point out an important concept that we shall use in a later chapter. We can consider that the non-homogeneous integral equation specifies a linear operation. In an explicit integral representation, this linear operation is given by Eq. 4.72. It is completely equivalent, however, to specify g(t) implicitly as the solution to our differential equations.

In the two following chapters we shall apply the results of this chapter. In the next we shall apply them to the problem of designing optimal signals for detection in additive colored noise channels. Our basic approach is to regard the differential equations we have developed as a dynamic system with initial and final boundary conditions When the problem is expressed in this form we can apply the Maximal Principle of Pontryagin for the optimization.[13.6]

In the subsequent chapter we shall present a new approach to solving Wiener-Hopf equations by using the results of this chapter. We shall then proceed to develop a unified theory of linear smoothing and filtering with delay.

CHAPTER V

OPTIMAL SIGNALS DESIGN FOR ADDITIVE COLORED NOISE CHANNELS VIA STATE VARIABLES

In the problem of detecting a known signal in the presence of additive colored noise, the signal waveform affects the performance of the receiver. For a given energy level, certain signals result in lower probabilities of error than do others. Consequently, by choosing the signal waveform in some optimal manner, we may maximize the performance of the system.

If one does this optimization, however, the signals that result tend to have large bandwidths. For example, when the noise is stationary, it places the signal energy in a frequency band that is on the tail of the colored noise spectrum. Often the available bandwidth is restricted; therefore, in this case one must perform this optimization with some form of constraint upon the bandwidth of the signal waveform. This is the problem which we want to consider in this chapter. For a given energy level, we want to find the signal waveform that optimizes the detection performance when the bandwidth (defined later) is constrained.

A. Problem Statement

We introduced the problem of detecting a known signal in the presence of additive colored noise in Chapter IV. Let us

briefly review some of its aspects for the optimization problem that we want to consider. The model of the system that we want to study is illustrated in Fig. 4. 1. Depending upon the hypothesis, a known signal, s(t) or -s(t) is transmitted over an additive scalar colored noise channel. In general, we shall assume that this colored noise is a zero mean Gaussian process that consists of a white component w(t) plus an independent component y(t) with finite power. It is easy to show that the optimal receiver computes the log-likelihood ratio by correlating the received signal with a second known function g(t).

This correlating signal may be determined by solving a Fredholm integral equation, Eq. 5. 1, of the second kind,

$$\int_{T_o}^{T_f} K_y(t,\tau)g(\tau)d\tau + \frac{N_o}{2} g(t) = s(t), \quad T_o \leq t \leq T_f; \tag{5.1}$$

where

s(t)	is the transmitted signal;
g(t)	is the optimal correlating signal;
$N_o/2$	is the power per unit bandwidth level of the white noise (identified as σ in the previous chapter);

and

$K_y(t, \tau)$	is the covariance function of the colored (finite power) component of the additive noise.
$[T_o, T_f]$	is the observation interval.

(We shall study only scalar channels here. However, all the results can be extended to vector channels with little or no difficulty.)

For this problem the appropriate performance measure is given by

$$d^2 = \frac{2}{N_o} \int_{T_o}^{T_f} s(t)\, g(t)\, dt. \tag{5.2}$$

Under the Gaussian assumption that we made, we can determine probabilities of error, false alarm and detection. If we relax this assumption, we can still interpret d^2 as the receiver output signal to noise ratio.

The choice of the signal s(t) is not completely free. We impose two constraints upon it. The first is an energy constraint, or

$$\int_{T_o}^{T_f} s^2(\tau) d\tau = E. \tag{5.3}$$

Since Eq. 5.1 is linear, it is easy to see that d^2 is linearly dependent upon E. Secondly, one can define bandwidth in a multitude of ways. Initially, we shall require that our signal satisfy the constraint

$$\int_{-\infty}^{\infty} \omega^2 |S_s(\omega)|^2 d\omega = \int_{T_o}^{T_f} \left(\frac{ds(\tau)}{d\tau}\right)^2 d\tau \leq EB^2. \qquad (5.4)$$

i.e., we have used a mean square constraint upon the derivative of the signal. (We do not need a normalization factor because of Eq. 5.3.)

We shall set up and solve the optimization problem using the constraint of Eq. 5.4. After we have done this, we shall show how we can extend our results to other types of constraints, e.g.,

$$\left| \frac{ds(t)}{dt} \right| \leq E^{\frac{1}{2}} B,$$

We can now state the optimization problem in terms of Eqs. 5.1 through 5.4. We want to find a signal s(t) that maximizes the performance measure d^2 given by Eq. 5.2 where g(t) is related to s(t) by Eq. 5.1, and yet satisfies the energy and bandwidth constraints imposed by Eqs. 5.3 and 5.4.

The first approach which one may want to consider is to formulate the optimization problem in terms of the eigenfunctions and eigenvalues of the homogeneous equation which may be associated with Eq. 5.1. If one does this, he finds that the optimal signal is the eigenfunction with the smallest eigenvalue which satisfies both Eqs. 5.3 and 5.4. This approach neglects two important issues. Unless Eq. 5.4 is satisfied with equality, we can find better signals. In addition, it neglects discontinuity effects caused by turning the signal on and off at T_o and T_f respectively.

A second approach as proposed by Van Trees is to apply the calculus of variations while introducing Lagrange multipliers to incorporate the constraints.[28] The resulting integral equation can then be converted to a set of differential equations by using results we derived in Chapter IV. For the particular form of constraints upon the signal that we have initially used, i.e., Eqs. 5.3 and 5.4, this is undoubtedly the most direct method of the minimization. However, the approach that we shall use is more general. Many of the results that we shall develop can be extended to constraints which cannot be readily handled with the classical calculus of variations. We assume that the colored component of the noise y(t) is a random process that is generated as we described in Chapter II. Making this assumption we recognize that we can represent the linear integral Eq. 5.1 as a set of differential equations as discussed in the previous chapter. Next we consider that this set of differential equations can be viewed as a dynamic system with boundary conditions and an input s(t); consequently, the Minimum Principle of Pontryagin can be used to perform the optimization.[13] By using this approach we shall first find a general solution to the problem, then we shall consider two specific examples in order to illustrate the techniques involved.

B. The Application of the Minimum Principle

In this section we shall develop a state variable formulation for the problem. Using this formulation we shall apply the minimum principle to find the necessary conditions for the existence of an optimal signal. We shall then exploit these conditions

to find an algorithm for determining the optimal signal.

Since there are several important issues that arise in the course of our derivation, we shall divide this section into subsections as listed below:

1. The State Variable Formulation of the Problem
2. The Minimum Principle and the Necessary Conditions
3. The Reduction of the Order of the Equations by 2n
4. The Transcendental Equation for λ_E and λ_B and the Selection of the Optimal Signal Candidates
5. The Asymptotic Solutions

1. The State Variable Formulation of the Problem

In order to apply the Minimum Principle we need to formulate the problem in terms of differential equations, boundary conditions, cost functionals, and a control. First, we need to find a set of differential equations and boundary conditions which relate g(t), the solution of the Fredholm integral equation expressed by the solution of Eq. 5.1 to the signal s(t).

We can do this by using the results derived in the previous chapter. Reviewing these results we have shown that g(t), the solution to Eq. 5.1 is given by Eq. 4.9

$$g(t) = \frac{2}{N_o} (s(t) - C(t)\,\underline{\xi}(t)), \qquad T_o \leq t \leq T_f \tag{5.5}$$

The vector function, $\underline{\xi}(t)$, satisfies the differential equations, Eqs. 4.10 and 4.12

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\,\underline{\xi}(t) + G(t)\,Q\,G^{T}(t)\,\underline{\eta}(t), \quad T_o \le t \le T_f, \tag{5.6}$$

$$\frac{d\underline{\eta}(t)}{dt} = C^{T}(t)\,\frac{2}{N_o}\,C(t)\,\underline{\xi}(t) - F^{T}(t)\,\underline{\eta}(t) - C^{T}(t)\,\frac{2}{N_o}\,s(t), \quad T_o \le t \le T_f \tag{5.7}$$

The boundary conditions which specify the solution uniquely are Eqs. 4.13 and 4.14.

$$\underline{\xi}(T_o) = P_o\,\underline{\eta}(T_o) \tag{5.8}$$

$$\underline{\eta}(T_f) = \underline{0}. \tag{5.9}$$

Consequently, we have the desired result that we can relate g(t) to s(t) by solving two vector differential equations where we have a two point boundary value condition imposed upon them.

Let us now develop the cost functional for the problem. The performance measure of our system is given by Eq. 5.2. If we substitute Eq. 5.5 in Eq. 5.2 and use Eq. 5.3, we find that

$$d^2 = \int_{T_o}^{T_f} s(\tau)\,\frac{2}{N_o}\,(s(\tau) - C(\tau)\,\underline{\xi}(\tau))\,d\tau$$

$$= \frac{2E}{N_o} - \frac{2}{N_o}\int_{T_o}^{T_f} s(\tau)C(\tau)\underline{\xi}(\tau)\,d\tau. \tag{5.10}$$

The first term in Eq. 5.10 is the performance when there is just white noise present. The second term represents the degradation in performance caused by the presence of the colored component of the noise. As in Chapter IV, let us define d_g^2 to be

$$d_g^2 = \frac{2}{N_o} \int_{T_o}^{T_f} s(\tau) C(\tau) \underline{\xi}(\tau) d\tau , \tag{5.11}$$

and the function $L(\xi(t), s(t))$ to be

$$L(\underline{\xi}(t), s(t)) = \frac{2}{N_o} s(t) C(t) \underline{\xi}(t), \qquad T_o \leq t \leq T_f, \tag{5.12a}$$

$$d_g^2 = \int_{T_o}^{T_f} L(\xi(\tau), s(\tau) d\tau \tag{5.12b}$$

Since the energy E and the white noise level $N_o/2$ are constants, it is obvious that we can maximize d^2 by minimizing d_g^2.

The state variable formulation requires that the system variables be related by derivative rather than integral operations. Since we are constraining both the signal and its derivative, we cannot use $s(t)$ as the control. Instead, let us define the control function, $v(t)$, to be the derivative of the signal.

$$v(t) = \frac{ds(t)}{dt} \tag{5.13}$$

Furthermore, we require

$$s(T_o) = s(T_f) = 0 \tag{5.14}$$

Eq. 5.14 is a logical requirement. Since we are constraining the derivative of the signal, it is reasonable to require that there be no jump discontinuities (implying singularities in v(t)) at the endpoints of the interval.

We now have all the state equations and boundary conditions that describe the dynamics of the system. The state equations are given by Eqs. 5.6, 5.7, and 5.13.

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\underline{\xi}(t) + G(t)\,Q\,G^T(t)\underline{\eta}(t), \qquad T_o \le t \le T_f \tag{5.6 (repeated)}$$

$$\frac{d\underline{\eta}(t)}{dt} = C^T(t)\,\frac{2}{N_o}\,C(t)\underline{\xi}(t) - F^T(t)\underline{\eta}(t) - C^T(t)\frac{2}{N_o}\,s(t), \qquad T_o \le t \le T_f \tag{5.7 (repeated)}$$

$$\frac{ds(t)}{dt} = v(t), \qquad T_o \le t \le T_f \tag{5.13 (repeated)}$$

We have (2n + 1) individual equations. The boundary conditions are given by Eqs. 5.8, 5.9, and 5.14.

$$\underline{\xi}(T_o) = P_o\underline{\eta}(T_o) \tag{5.8 (repeated)}$$

$$\underline{\eta}(T_f) = \underline{0} \tag{5.9 (repeated)}$$

$$s(T_o) = s(T_f) = 0 \tag{5.14 (repeated)}$$

Notice that there are $(2n+2)$ individual boundary conditions. Consequently, these conditions cannot be satisfied for an arbitrary $v(t)$.

In order to introduce the energy and bandwidth constraints, we need to augment the state equations artificially by adding the two equations

$$\frac{dx_E(t)}{dt} = \frac{s^2(t)}{2}, \qquad T_o \le t \le T_f, \tag{5.15}$$

$$\frac{dx_B(t)}{dt} = \frac{v^2(t)}{2}, \qquad T_o \le t \le T_f. \tag{5.16}$$

(We have introduced the factor of $1/2$ for a later convenience.) The boundary conditions are

$$x_E(T_o) = x_B(T_o) = 0, \tag{5.17}$$

$$x_E(T_f) = E/2 \tag{5.18}$$

$$x_B(T_f) = EB^2/2 \tag{5.19}$$

It is easy to see that these differential equations and boundary conditions represent the constraints described by Eqs. 5.3 and 5.4.

With these last results, we have formulated the problem in a form where we can apply the Minimum Principle.

2. The Minimum Principle and the Necessary Conditions

In this section we shall use Pontryagin's Minimum Principle to derive the necessary conditions for optimality. Before proceeding, two comments are in order. First, the control function is v(t) not s(t), which is one of the components of the state vector for the system. Secondly, we shall not develop much background material on the Principle itself. For further information we refer to References 6 and 13.

The Hamiltonian for this system is

$$H(\underline{\xi}, \underline{\eta}, s, x_E, x_B, p_o, p_{\underline{\xi}}, p_{\underline{\eta}}, p_s, \lambda_E, \lambda_B, v, t) \overset{*}{=}$$

$$p_o L(\xi(t), s(t)) + p_{\underline{\xi}}^T(t)\, \dot{\underline{\xi}}(t) + p_{\underline{\eta}}^T(t)\, \dot{\underline{\eta}}(t) + p_s(t)\, \dot{s}(t)$$

$$+ \lambda_E(t)\, \dot{x}_E(t) + \lambda_B(t)\, \dot{x}_B(t) =$$

$$p_o \frac{2}{N_o} \underline{s}(t)\, C(t)\, \underline{\xi}(t)$$

$$+ p_{\underline{\xi}}^T(t)(F(t)\, \underline{\xi}(t) + G(t)\, Q\, G^T(t)\, \underline{\eta}(t))$$

$$p_{\underline{\eta}}^T(t)(C^T(t) \frac{2}{N_o} C(t)\, \underline{\xi}(t) - F^T(t)\, \underline{\eta}(t) - C^T(t) \frac{2}{N_o} s(t))$$

$$p_s(t)\, v(t) + \lambda_E(t) \frac{s^2(t)}{2} + \lambda_B(t) \frac{v^2(t)}{2}, \quad T_o \le t \le T_f \tag{5.20}$$

---

*We shall drop the arguments when there is no specific need for them.

We have denoted the costate vector of the state equation describing the dynamics of the system (Eqs. 5.6, 5.7 and 5.13) by the variable $\underline{p}(t)$. The subscript indicates the corresponding state variable. The costates of the constraint equations are denoted by $\lambda_E(t)$ and $\lambda_B(t)$. The system that we want to optimize may be explicitly time-dependent (non-autonomous), has a fixed time interval and has boundary conditions at both ends of the time interval.

Let $\underline{\xi}(t)$, $\underline{\eta}(t)$, and s(t) be the functions that satisfy the differential equations expressed by Eqs. 5.6, 5.7 and 5.13; the boundary conditions given by Eqs. 5.8, 5.9 and 5.14; and the constraints of Eqs. 5.3 and 5.4, when the control function is v(t). The Minimum Principle states: In order that $\hat{v}(t)$ by optimum, it is necessary that there exist a constant p_o and functions $\hat{\underline{p}}_{\underline{\xi}}(t)$, $\hat{\underline{p}}_{\underline{\eta}}(t)$, $\hat{p}_s(t)$, $\hat{\lambda}_E(t)$, and $\hat{\lambda}_B(t)$ (not all identically zero) such that the following four assertions hold:

a. $$\hat{\underline{p}}_{\underline{\xi}}(t) = \nabla_{\underline{\xi}} \hat{H},^* \tag{5.21}$$

$$\hat{\underline{p}}_{\underline{\eta}}(t) = \nabla_{\underline{\eta}} \hat{H}, \tag{5.22}$$

$$\hat{p}_s(t) = -\partial \hat{H}/\partial s, \tag{5.23}$$

$$\hat{\lambda}_E(t) = -\partial \hat{H}/\partial E \tag{5.24}$$

---

*$H = H(\hat{\underline{\xi}}, \hat{\underline{\eta}}, \hat{s}, \hat{x}_E, \hat{x}_B, \hat{p}_o, \hat{\underline{p}}_{\underline{\xi}}, \hat{\underline{p}}_{\underline{\eta}}, \hat{p}_s, \hat{\lambda}_E, \hat{\lambda}_B, \hat{v}, t)$

As expected the energy and bandwidth constraint costates are constants (therefore, we shall drop the time dependence notation.)

Since we have no boundary upon the control region, we can minimize the Hamiltonian vs. the variable v(t) by equating the derivative to zero.

$$\frac{\partial H}{\partial v}\Big|_{v = \hat{v}(t)} = 0 = p_s(t) + \lambda_B\, v(t), \quad T_o \le t \le T_f. \tag{5.31}$$

Furthermore, for this to be a minimum, we require that

$$\frac{\partial^2 H}{\partial v^2}\Big|_{v = \hat{v}(t)} \ge 0 \qquad \text{equivalently, from 5.31} \tag{5.32}$$

$$\lambda_B \ge 0.$$

In general, we can show that $\lambda_B > 0$. Then we can solve Eq. 5.31 for v(t). This yields

$$v(t) = -p_s(t)/\lambda_B \quad , \qquad T_o \le t \le T_f. \tag{5.33}$$

Substitute Eq. 5.33 in Eq. 5.13. Now we shall write the canonical equations expressed by Eqs. 5.6, 5.7, 5.13, 5.26, 5.27 and 5.28 in an augmented vector form. This yields a homogeneous set of 4n + 2 equations,

$$\hat{\lambda}_B(t) = \partial\hat{H}/\partial x_B; \tag{5.25}$$

b. for all t in the interval $[T_o, T_f]$ the function $H(\hat{\underline{\xi}}, \hat{\underline{\eta}}, \hat{\underline{s}}, \hat{x}_E, \hat{x}_B, \hat{\underline{p}}_{\underline{\xi}}, \hat{\underline{p}}_{\underline{\eta}}, \hat{p}_s, \hat{\lambda}_E, \hat{\lambda}_B, v, t)$ is minimized as a function of the variable v;

c. p_o is a constant with $p_o \geq 0$;

d. the costate vector is perpendicular to the manifold defined by boundary conditions at each end of the interval.

Let us now examine what each of these assertions implies. If we perform the derivative operations indicated by Eqs. 5.21 to 5.25, we find

$$\hat{\underline{p}}_{\underline{\xi}}(t) = -p_o \frac{2}{N_o} C^T(t)\hat{\underline{s}}(t) - F^T(t)\hat{\underline{p}}_{\underline{\xi}}(t) - C^T(t)\frac{2}{N_o}C(t)\hat{\underline{p}}_{\underline{\eta}}(t), \quad T_o \leq t \leq T_f; \tag{5.26}$$

$$\hat{\underline{p}}_{\underline{\eta}}(t) = -G(t)\,Q\,G^T(t)\,\hat{\underline{p}}_{\underline{\xi}}(t) + F(t)\,\hat{\underline{p}}_{\underline{\eta}}(t), \quad T_o \leq t \leq T_f; \tag{5.27}$$

$$\hat{p}_s(t) = -p_o \frac{2}{N_o} C^T(t)\hat{\underline{\xi}}(t) + C^T(t)\frac{2}{N_o}\hat{\underline{p}}_{\underline{\eta}}(t) - \lambda_E(t)\hat{\underline{s}}(t), \quad T_o \leq t \leq T_f; \tag{5.28}$$

$$\lambda_E(t) = 0, \tag{5.29}$$

$$\lambda_B(t) = 0. \tag{5.30}$$

$$\frac{d}{dt}\begin{bmatrix} \hat{\underline{\xi}}(t) \\ \hat{\underline{\eta}}(t) \\ \hat{s}(t) \\ \hat{\underline{p}}_{\underline{\xi}}(t) \\ \hat{\underline{p}}_{\underline{\eta}}(t) \\ \hat{p}_s(t) \end{bmatrix} = \begin{bmatrix} F(t) & G(t)QG^T(t) & 0 & 0 & 0 & 0 \\ C^T(t)\frac{2}{N_o}C(t) & -F^T(t) & C^T(t)\frac{2}{N_o} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1/\lambda_B \\ 0 & 0 & -p_o C^T(t)\frac{2}{N_o} & -F^T(t) & -C^T(t)\frac{2}{N_o}C(t) & 0 \\ 0 & 0 & 0 & -G^T(t)QG^T(t) & F^T(t) & 0 \\ -p_o\frac{2}{N_o}C(t) & 0 & -\lambda_E & 0 & \frac{2}{N_o}C(t) & 0 \end{bmatrix} \begin{bmatrix} \hat{\underline{\xi}}(t) \\ \hat{\underline{\eta}}(t) \\ \hat{s}(t) \\ \hat{\underline{p}}_{\underline{\xi}}(t) \\ \hat{\underline{p}}_{\underline{\eta}}(t) \\ \hat{p}_s(t) \end{bmatrix},$$

$$T_o \le t \le T_f. \qquad (5.34)$$

(From now on, we shall drop the ^ notation and assume that we refer to the optimal solution.)

If, in assertion (c) of the Minimum Principle, the constant p_o is identically zero, then we have what we shall call an asymptotic case. We shall return to this case later; however, let us for the interim set p_o equal to unity. Since the costate equations are linear, this entails no loss of generality.

Let us consider the boundary, or transversality, conditions implied by assertion (d). At the initial time, these conditions imply

$$\underline{p}_{\underline{\eta}}(T_o) = -P_o \underline{p}_{\underline{\xi}}(T_o), \tag{5.35}$$

and $p_s(T_o)$ is unspecified. At the final, or endpoint, time we have

$$\underline{p}_{\underline{\xi}}(T_f) = \underline{0}, \tag{5.36}$$

and $\underline{p}_{\underline{\eta}}(T_f)$ and $p_s(T_f)$ are unspecified. We also have that λ_E and λ_B are unspecified constants ($\lambda_B \geq 0$). We also have the boundary conditions given by Eqs. 5.8, 5.9 and 5.14. Therefore, we have a total of 4n + 2 boundary conditions. In addition, we should notice that we do not have to find the control v(t) in order to find s(t), although we may easily deduce it from Eq. 5.33.

3. The Reduction of the Order of the Equations by 2n

We are now in a position to show how the assertions of the Minimum Principle may be used to find the candidates for the optimal signal. However, before proceeding we shall derive a

result that significantly simplifies the solution method. We shall prove that, in general,

$$\underline{\xi}(t) = -\underline{p}_{\underline{\eta}}(t), \qquad T_o \le t \le T_f, \tag{5.37}$$

$$\underline{\eta}(t) = \underline{p}_{\underline{\xi}}(t), \qquad T_o \le t \le T_f. \tag{5.38}$$

This reduction was suggested by the variational approach of Van Trees. We point out though that our derivation is independent of the type of constraint imposed upon the signal; i.e., it only depends upon the differential equations for $\underline{\xi}(t)$ and $\underline{\eta}(t)$.

Let us define two vectors $\underline{\epsilon}_1(t)$ and $\underline{\epsilon}_2(t)$ as

$$\underline{\epsilon}_1(t) = \underline{\epsilon}(t) + \underline{p}_{\underline{\eta}}(t), \quad T_o \le t \le T_f, \tag{5.39}$$

$$\underline{\epsilon}_2(t) = \underline{\eta}(t) - \underline{p}_{\underline{\xi}}(t), \quad T_o \le t \le T_f. \tag{5.40}$$

If we differentiate these two equations and substitute Eqs. 5.6, 5.7, 5.26 and 5.27 (with p_o equal to 1), we find

$$\begin{aligned}\underline{\dot{\epsilon}}_1(t) &= F(t)\,\underline{\xi}(t) + G(t)\,Q\,G^T(t)\,\underline{\eta}(t) - G(t)\,Q\,G^T(t)\,\underline{p}_{\underline{\xi}}(t) + F(t)\,\underline{p}_{\underline{\eta}}(t) \\ &= F(t)\,\underline{\epsilon}_1(t) + G(t)\,Q\,G^T(t)\,\underline{\epsilon}_2(t), \quad T_o \le t \le T_f,\end{aligned} \tag{5.41}$$

$$\begin{aligned}\underline{\dot{\epsilon}}_2(t) &= C^T(t)\,\frac{2}{N_o}\,C(t)\,\underline{\xi}(t) - F^T(t)\,\underline{\eta}(t) - C^T(t)\,\frac{2}{N_o}\,s(t) \\ &\quad + C^T(t)\,\frac{2}{N_o}\,C(t)\,\underline{p}_{\underline{\eta}}(t) + F^T(t)\,\underline{p}_{\underline{\xi}}(t) + C^T(t)\,\frac{2}{N_o}\,s(t) \\ &= C^T(t)\,\frac{2}{N_o}\,C(t)\,\underline{\epsilon}_1(t) - F^T(t)\,\underline{\epsilon}_2(t), \quad T_o \le t \le T_f.\end{aligned} \tag{5.42}$$

The boundary conditions that the solution to these differential equations satisfy may be found by using Eqs. 5.8, 5.9, 5.35 and 5.36. They are

$$\underline{\epsilon}_1(T_o) = P_o\,\underline{\eta}(T_o) - P_o\,\underline{p}_{\xi}(T_o) = P_o\,\underline{\epsilon}_2(T_o) \tag{5.43}$$

$$\underline{\epsilon}_2(T_f) = \underline{\eta}(T_f) - \underline{p}_{\xi}(T_f) = \underline{0}. \tag{5.44}$$

Consequently, Eqs. 5.41 to 5.44 specify two vector linear differential equation with a two point boundary value condition. However, these equations are just those that specify the eigenvalues and eigenfunctions for the homogeneous Fredholm integral equation as shown in Chapter III, Eqs. 3.10 to 3.13 . We have shown that in order to have a nontrivial solution to this problem, we require that

$$-\frac{N_o}{2} = \lambda_i > 0. \tag{5.45}$$

where λ_i is an eigenvalue of a Karhunen-Loève expansion of the colored noise. Clearly, this is impossible since $\lambda_i \geq 0$. Consequently, the only solution is the trivial one, i.e., $\underline{\epsilon}_1(t) = \underline{\epsilon}_2(t) = \underline{0}$, which proves the assertion of Eqs. 5.37 and 5.38.

4. The Transcendental Equation for λ_E and λ_B and the Selection of the Optimal Signal Candidates

In this subsection we shall use the necessary conditions to derive a transcendental equation that must be satisfied for an

optimal solution to exist. The method and result are very similar to that which we used in Chapter III to find the eigenvalues of the homogeneous Fredholm equation. The most important distinction is that this equation is in terms of two parameters, λ_E and λ_B, whereas we had but one before. Once we satisfy this equation, we can generate a signal which is a candidate for the optimum solution.

Because of the linear dependencies derived in the previous subsection we can reduce the 4n + 2 equation specified in Eq. 5.34 to a set of 2n + 2 equations. We have

$$\frac{d}{dt}\begin{bmatrix} \underline{\xi}(t) \\ \underline{\eta}(t) \\ s(t) \\ p_s(t) \end{bmatrix} = \begin{bmatrix} F(t) & G(t)QG^T(t) & 0 & 0 \\ C^T(t)\frac{2}{N_o}C(t) & -F^T(t) & -C^T(t)\frac{2}{N_o} & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda_B} \\ -\frac{4}{N_o}C(t) & 0 & -\lambda_E & 0 \end{bmatrix} \begin{bmatrix} \underline{\xi}(t) \\ \underline{\eta}(t) \\ s(t) \\ p_s(t) \end{bmatrix},$$

$$T_o \leq t \leq T_f. \qquad (5.46)$$

The boundary conditions are given by Eqs. 5.8, 5.9, and 5.14. These conditions specify 2n + 2 boundary conditions that must be satisfied for an optimum to exist.

Since Eq. 5.46 is a homogeneous linear equation, we may not, in general, have a nontrivial solution. In order to find where we may obtain a nontrivial solution, let us define the transition matrix associated with Eq. 5.46 to be $\chi(t, T_o: \lambda_E, \lambda_B)$. We emphasize the

dependence of χ upon λ_E and λ_B by including them as arguments.

Since Eq. 5.46 is linear, we can determine any solution to Eq. 5.46 in terms of this transition matrix. If we use the boundary conditions specified by Eqs. 5.8 and 5.14, we find that any solution that satisfies the initial conditions may be written in the form

$$\begin{bmatrix} \underline{\xi}(t) \\ \underline{\eta}(t) \\ s(t) \\ p_s(t) \end{bmatrix} = \chi(t, T_o : \lambda_E, \lambda_B) \begin{bmatrix} P_o\, \underline{\eta}(T_o) \\ \underline{\eta}(T_o) \\ 0 \\ p_s(T_o) \end{bmatrix}, \quad (5.47)$$

$$T_o \leq t \leq T_f.$$

The final boundary condition requires that $\underline{\eta}(T_f)$ and $s(T_f)$ both be zero. In order to see what this implies, let us partition this transition matrix as follows (we drop the arguments temporarily):

$$\chi = \left[\begin{array}{c|c|c|c} \chi_{\underline{\xi}\underline{\xi}} & \chi_{\underline{\xi}\underline{\eta}} & \chi_{\underline{\xi}s} & \chi_{\underline{\xi}p_s} \\ \hline \chi_{\underline{\eta}\underline{\xi}} & \chi_{\underline{\eta}\underline{\eta}} & \chi_{\underline{\eta}s} & \chi_{\underline{\eta}p_s} \\ \hline \chi_{s\underline{\xi}} & \chi_{s\underline{\eta}} & \chi_{ss} & \chi_{sp_s} \\ \hline \chi_{p_s\underline{\xi}} & \chi_{p_s\underline{\eta}} & \chi_{p_s} & \chi_{p_sp_s} \end{array}\right] \quad (5.48)$$

By substituting Eq. 5.38 in Eq. 5.37, we find that in terms of these partitions, the requirement that $\underline{\eta}(T_f)$ vanish, implies

$$\underline{0} = [\chi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda_E, \lambda_B)P_o + \chi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda_E, \lambda_B)]\underline{\eta}(T_o)$$

$$+ \chi_{\underline{\eta}\, p_s}(T_f, T_o:\lambda_E, \lambda_B)]\, p_s(T_o) \qquad (5.49)$$

Similarly, we find that $s(T_f)$ being zero requires

$$0 = [\chi_{s\underline{\xi}}(T_f, T_o:\lambda_E, \lambda_B)P_o + \chi_{s\underline{\eta}}(T_f, T_o:\lambda_E, \lambda_B)]\underline{\eta}(T_o)$$

$$+ \chi_{s\, p_s}(T_f, T_o:\lambda_E, \lambda_B)p_s(T_o) \qquad (5.50)$$

We can write Eqs. 5.49 and 5.50 more concisely in matrix-vector form

$$0 = \left[\begin{array}{c|c} \chi_{\underline{\eta}\underline{\xi}}(T_f, T_o:\lambda_E, \lambda_B)P_o + \chi_{\underline{\eta}\underline{\eta}}(T_f, T_o:\lambda_E, \lambda_B) & \chi_{\underline{\eta}\, p_s}(T_f, T_o:\lambda_E, \lambda_B) \\ \hline \chi_{s\,\underline{\xi}}(T_f, T_o:\lambda_E, \lambda_B)P_o + \chi_{s\, p_s}(T_f, T_o:\lambda_E, \lambda_B) & \chi_{s\, p_s}(T_f, T_o:\lambda_E, \lambda_B) \end{array}\right] \left[\begin{array}{c} \underline{\eta}(T_o) \\ \hline p_s(T_o) \end{array}\right] \qquad (5.51)$$

or by defining the matrix in Eq. 5.51 to be $D(\lambda_E, \lambda_B)$, we have

$$\underline{0} = D(\lambda_E, \lambda_B) \begin{bmatrix} \underline{\eta}(T_o) \\ \hline p_s(T_o) \end{bmatrix} \tag{5.52}$$

Equation 5.51 specifies a set of n + 1 linear homogeneous algebraic equations. The only way that this set of equations can have a nontrivial solution is for the determinant of the matrix $D(\lambda_E, \lambda_B)$ to be identically zero. Consequently, the test for candidates for the optimal signal is to find those values of λ_E and λ_B (> 0) such that

$$\det[D(\lambda_E, \lambda_B)] = 0. \tag{5.53}$$

Once Eq. 5.53 is satisfied, we can find a non-zero solution to Eq. 5.52 up to a multiplicative constant. Knowing $\underline{\eta}(T_o)$ and $p_s(T_o)$, allows us to determine the candidate signal(s), $s_{\lambda_E, \lambda_B}(t)$, for the particular values of λ_E and λ_B that satisfy Eq. 5.53. The multiplicative constant may be determined by applying the energy constraint of Eq. 5.3; i.e.,

$$\int_{T_o}^{T_f} s^2_{\lambda_E, \lambda_B}(\tau)d\tau = E \tag{5.54}$$

By using Eq. 5.33 we can determine the bandwidth of the signal. We have that

$$B^2 = \frac{1}{E}\int_{T_o}^{T_f}\left(\frac{d\, s_{\lambda_E,\lambda_B}(\tau)}{d\tau}\right)^2 d\tau = \frac{1}{E\lambda_B^2}\int_{T_o}^{T_f} p^2_{s\,\lambda_E,\lambda_B}(\tau)d\tau \tag{5.55}$$

In order to satisfy Eq. 5.53, we require that only the rank of $D(\lambda_E,\lambda_B)$ be less than or equal to n the dimension of $\underline{\xi}(t)$ and $\underline{\eta}(t)$. The case when this rank is less than n presents an important aspect of this optimization. For convenience, let us define

$$n_D = n + 1 - \text{Rank}[D(\lambda_E,\lambda_B)]. \tag{5.56}$$

n_D specifies the number of linearly independent solutions to Eq. 5.52 that we may obtain for the given values of λ_E and λ_B. These solutions in turn specify n_D functions, v(t), that satisfy the necessary conditions for optimality given by the Minimum Principle.

We see that because of the linearity and quadratic constraint of any linear combination of these functions that have the same values of λ_E and λ_B also satisfy the necessary conditions given by the Minimum Principle. Consequently, any time we find n_D is greater than 1, we must consider these linear combinations when checking to see which candidate is indeed optimum. Of course,

these candidates are subject to the same constraints as any other; i.e., the energy and bandwidth constraints given by Eqs. 5.54 and 5.55.

5. The Asymptotic Solutions

An issue which we deferred was the question of the asymptotic case when p_o equals zero. Since these solutions provide an useful in the analysis of a particular problem, it is worthwhile to examine them before proceeding with the discussion of the algorithm of our design procedure. We shall call the solutions that satisfy the necessary conditions of the Minimum Principle when p_o is zero the asymptotic solutions. (They are often called pathological solutions.) In order to test for their existence, we set p_o equal to zero in Eq. 5.34 and examine the differential equations for $\underline{p}_{\underline{\xi}}(t)$ and $\underline{p}_{\underline{\eta}}(t)$. If we write these equations in augmented vector form, we obtain the following homogeneous equation

$$\frac{d}{dt}\begin{bmatrix}\underline{p}_{\underline{\xi}}(t)\\ \hline \underline{p}_{\underline{\eta}}(t)\end{bmatrix} = \left[\begin{array}{c|c} -F^T(t) & -C^T(t)\frac{2}{N_o}C(t) \\ \hline -G(t)QG^T(t) & F(t)\end{array}\right]\begin{bmatrix}\underline{p}_{\underline{\xi}}(t)\\ \hline \underline{p}_{\underline{\eta}}(t)\end{bmatrix}$$

$$T_o \leq t \leq T_f. \qquad (5.57)$$

The appropriate boundary conditions are specified by Eqs. 5.35 and 5.36.

From Chapter III, Eq. 3.13,

$$\begin{bmatrix} -F^T(t) & -C^T(t)\frac{2}{N_o}C(t) \\ -G(t)QG^T(t) & F(t) \end{bmatrix} = -W^T(t:-N_o/2) \qquad (5.58)$$

Let us define the transition matrix associated with Eq. 5.57 to be $\Phi(t,\tau:-N_o/2)$. We note that $\Phi(t,\tau:-N_o/2)$ is related to the transition matrix $\Psi(t,\tau:-N_o/2)$ associated with $W(t:-N_o/2)$

$$\Phi(t,\tau:-N_o/2) = \Psi^T(\tau,t:-N_o/2) \qquad (5.59)$$

Any solution to Eq. 5.57 may be found in terms of $\Phi(t,\tau:-N_o/2)$. In order to find the solution to Eq. 5.57 that satisfies the boundary conditions, we partition $\Phi(t,\tau:-N_o/2)$ into four n x n submatrices,

$$\Phi(t:\tau:-\frac{N_o}{2}) = \begin{bmatrix} \Psi^T_{\xi\xi}(\tau,t:-\frac{N_o}{2}) & \Psi^T_{\eta\xi}(\tau,t:-\frac{N_o}{2}) \\ \Psi^T_{\xi\eta}(\tau,t:-\frac{N_o}{2}) & \Psi^T_{\eta\eta}(\tau,t:-\frac{N_o}{2}) \end{bmatrix} \qquad (5.60)$$

If we incorporate the boundary condition specified by Eq. 5.36 that the solution to Eq. 5.57 is

$$\begin{bmatrix} \underline{p}_\xi(t) \\ \underline{p}_\eta(t) \end{bmatrix} = \begin{bmatrix} \Psi^T_{\eta\xi}(T_f,t:-\frac{N_o}{2}) \\ \Psi^T_{\eta\eta}(T_f,t:-\frac{N_o}{2}) \end{bmatrix} \underline{p}_\eta(T_f). \qquad (5.61)$$

The initial condition specified by Eq. 5.35 requires

$$\underline{p}_{\eta}(T_o) = \Psi^T_{\eta\eta}(T_f, T_o{:-}\frac{N_o}{2})\underline{p}_{\eta}(T_f) \tag{5.62}$$

$$= -P_o\underline{p}_{\xi}(T_o) = -P_o\Psi^T_{\eta\xi}(T_f, T_o{:-}N_o 2)\underline{p}_{\eta}(T_f),$$

or

$$\underline{0} = \left[\Psi^T_{\eta\eta}(T_f, T_o{:-}\frac{N_o}{2}) + P_o\,\Psi^T_{\eta\xi}(T_f, T_o{:-}\frac{N_o}{2})\right]\underline{p}_{\eta}(T_f) \tag{5.63}$$

The only way that Eq. 5.36 can have a nontrivial solution is for the determinant of the matrix enclosed by brackets in Eq. 5.63 to vanish.

$$\det\left[\Psi^T_{\eta\eta}(T_f, T_o{:-}\frac{N_o}{2}) + P_o\,\Psi^T_{\eta\xi}(T_f, T_o{:-}\frac{N_o}{2})\right] = 0. \tag{5.64}$$

If we transpose the matrix in Eq. 5.64 (this does not change the determinant value), we find that we have the test for an eigenvalue that we developed in Chapter III. There we showed that the only way for this determinant to vanish is for

$$-\frac{N_o}{2} = \lambda_i > 0$$

where λ_i is an eigenvalue associated with the Karhunen Loeve expansion colored noise process. Clearly, this is impossible. Consequently, the only solution is the trivial one, i.e.,

C. The Signal Design Algorithm

In the previous section we have derived the results necessary for solving the signal design problem. In this section we shall use these results (and a fair amount of experience) to devise an algorithm which when implemented on a digital computer will find the optimal signals and their performance.

Ideally, we want to be able to find the optimum signal for given values of E and B. Although this is certainly possible, it is far more efficient to solve a particular problem where we let B be a parameter and then select the specific value in which we are interested (the energy is normalized to unity). The reason for this approach will become apparent when we consider some specific examples.

The Minimum Principle has provided us with a set of necessary condition from which we found a test (Eq. 5.53) for an optimum signal. The result of this test is that we essentially have an eigenvalue problem in two dimensions. The most difficult aspect of the problem becomes finding the particular values of λ_E and λ_B that both satisfy this test and correspond to a signal the desired bandwidth. The algorithm that we suggest here is simply a systematic method of approaching this aspect of the problem.

The algorithm has several steps. First, we shall outline it. Then in the next section we shall discuss it in the context of two examples.

a. Find the loci points in the λ_E, λ_B plane that satisfy Eq. 5.53. This requires that we have an effective procedure for calculating transition matrices. In general, one can simply

$$\underline{p}_{\underline{\xi}}(t) = \underline{p}_{\underline{\eta}}(t) = \underline{0}, \qquad T_o \leq t \leq T_f. \tag{5.65}$$

If we substitute Eq. 5.65 in Eq. 5.34, we find that the differential equations for s(t) and $p_s(t)$ are

$$s(t) = -\frac{1}{\lambda_B} p_s(t), \qquad T_o \leq t \leq T_f, \tag{5.66}$$

$$p_s(t) = -\lambda_E s(t), \qquad T_o \leq t \leq T_f. \tag{5.67}$$

The only solution to these equations that satisfies the boundary conditions specified by Eq. 5.14 is

$$s(t) = \sqrt{\frac{2E}{T}} \sin\left(n\pi \frac{t - T_o}{T_f - T_o}\right), \quad T_o \leq t \leq T_f, \tag{5.68}$$

with

$$-\frac{\lambda_E}{\lambda_B} = \left(\frac{n\pi}{T_f - T_o}\right)^2, \quad \lambda_E < 0. \tag{5.69}$$

Several comments are in order. (1.) The bandwidth of these signals is easily shown to be $n\pi/T$. (2.) We have not violated assertion (a) in our application of the Minimum Principle since $p_s(t)$ is non zero. (3.) We did not require our system to be time invariant, nor did we specify the dimension of the system. Consequently, these solutions, which are probably the most practical to transmit, exist for all types of colored noise that fit within our model.

numerically integrate the differential equations that specify the transition matrix. However, if the matrices that describe the generation of the channel noise (F, G, Q, C) are constants, then we can use the matrix exponential, i.e.,

$$\chi(t, T_o : \lambda_E, \lambda_B) = e^{Z(\lambda_E, \lambda_B)(t-T_o)}, \tag{5.70}$$

where $Z(\lambda_E, \lambda_B)$ is the coefficient matrix of Eq. 5.45.

b. For a particular point on these loci, solve Eq. 5.52 for $\underline{\eta}(T_o)$ and $p_s(T_o)$. Then use Eq. 5.47 to determine the signals $s_{\lambda_E, \lambda_B}(t)$, $p_{s\,\lambda_E, \lambda_B}(t)$ and $\underline{\xi}_{\lambda_E, \lambda_B}(t)$.

c. Since the performance is linearly related to the energy, normalize these signals such that $s_{\lambda E, \lambda B}(t)$ has unit energy.

d. Calculate the bandwidth and performance of the normalized signals as specified by Eqs. 5.55 and 5.11, respectively.

e. Repeat parts b, c, and d at appropriate intervals along these loci in the λ_E, λ_B plane. (The interval should be small enough so that the bandwidth and performance as calculated in part d vary in a reasonably continuous manner.) As we move along a particular locus in the λ_E, λ_B plane, plot the degradation vs. bandwidth in a second plane, a d_g^2, B^2 plane.

f. As mentioned earlier, we need to pay special attention to the case when Eq. 5.53 has more than one solution. This situation corresponds to a crossing of two or more loci in the λ_E, λ_B plane. In this case find the solutions and plot the locus produced in the d_g^2, B^2 plane by linearly combining the different

signals. Probably the most convenient means of doing this is to use the Fredholm integral equation technique discussed in Chapter IV.

g. For a given value of bandwidth constraint the optimal signal is the one that corresponds to the lowest value of d_g^2. This signal is the absolute minimum, while the others correspond to relative minima.

h. We recommend that one use his engineering judgment in determining which loci on the λ_E, λ_B plane are on interest and in deciding when the white noise performance is being approached.

The problem of determining which loci in the λ_E, λ_B plane generate the optimal solution is time consuming. As we shall see, these loci can be rather complex. In the two examples that we have studied, we have observed a phenomenon which, if it were generally true, would considerably simplify this aspect of the algorithm. If one chooses a value of λ_E and finds the solution of Eq. 5.53 with the largest value of λ_B, the signal that corresponds to this point is globally optimum for the value of B that the signal has. (Similarly, one can fix λ_B and find the solution of Eq. 5.53 with the largest λ_E in absolute value.) With this conjecture the bandwidth then becomes a monotonic function of either λ_E or λ_B. The only possible difficulty that would be in finding the points corresponding to loci crossing.

We have not been able to prove that this phenomenon is true in general. However, it seems quite promising in view of the evidence that we have and the analogy to the smallest eigenvalue optimization procedure when there is no bandwidth constraint imposed.[3]

D. Examples of Optimal Signal Design

In this section we shall illustrate our technique. We shall consider the colored noise to have the one and two pole spectra described in Chapter II.

Example 1 - Signal Design with a One Pole Spectrum

The equations describing the generation of the colored component of the observation noise are given by Eqs. 2.15 and 2.16. Since the process is stationary, let us also set $T_o = 0$ and $T_f = T$.

In order to set up the test for the optimum signal, we need to find the coefficient matrix in Eq. 5.46. If we substitute the coefficients of the Eqs. 2.15 and 2.16 into Eq. 5.46 we obtain

$$\frac{d}{dt}\begin{bmatrix} \xi(t) \\ \eta(t) \\ s(t) \\ p_s(t) \end{bmatrix} = \begin{bmatrix} -k & 2kS & 0 & 0 \\ \frac{2}{N_o} & k & -\frac{2}{N_o} & 0 \\ 0 & 0 & 0 & -\frac{1}{\lambda_B} \\ -\frac{4}{N_o} & 0 & -\lambda_E & 0 \end{bmatrix} \begin{bmatrix} \xi(t) \\ \eta(t) \\ s(t) \\ p_s(t) \end{bmatrix} \tag{5.71}$$

From Eqs. 5.8, 5.9 and 5.14 the boundary conditions are

$$\xi(0) = S\eta(0), \tag{5.72a}$$

$$\eta(T) = 0, \tag{5.72b}$$

$$s(0) = s(T) = 0 \tag{5.72c}$$

Step (a) in our algorithm requires us to find the loci in λ_E, λ_B plane that satisfy Eq. 5.53. To do this, we need to calculate the transition matrix (4 x 4) of Eq. 5.71 $\chi(T, 0:\lambda_E, \lambda_B)$. From this matrix we can compute $D(\lambda_E, \lambda_B)$ (2 x 2) by using Eq. 5.51.

To illustrate this, let us choose specific values and use a computer to perform the required calculations for the parameters k, $N_o/2$, S, and T. Let us set

$$\begin{aligned} k &= 1 \\ \frac{N_o}{2} &= 1 \\ S &= 1 \\ T &= 2 \end{aligned} \qquad (5.73\ a\text{-}d)$$

The loci of points in the λ_E, λ_B plane that satisfy Eq. 5.53 are illustrated in Fig. 5.1.

In this figure only six loci have been plotted. Other loci with lower values of λ_B for a given λ_E exist but are not significant in the final solution. (They correspond to signals with a large bandwidth and performance very close to the "white noise only" limit. We found these loci by first fixing λ_E and then locating the zeroes of Eq. 5.53 as a function of λ_B.

According to Eq. 5.32, we need to look for solutions only in region of the plane where $\lambda_B > 0$. In this particular example we did not observe any loci in the first quadrant of the plane. Consequently, the only region of interest is where $\lambda_B > 0$ and $\lambda_E < 0$.

The reason for considering the asymptotic solutions should now be apparent. We have sketched their loci as given by Eq. 5.69 with dotted lines. We see that the loci asymptotically approach those of the asymptotic solutions. Therefore, for large values of λ_E a convenient place to start searching for the loci is near those specified by Eq. 5.69. We should point out that the loci of the asymptotic solutions do not cause the determinant specified by Eq. 5.53 to vanish. This is because in determining their existence we used the equations directly derived from the Minimum Principle rather than the reduced set of equations that were used to derive Eq. 5.53.

Now, if we take the solutions specified by these loci and determine the bandwidth and performance of the corresponding signal according to steps (b) through (e), we produce a second set of loci in a d_g^2-B^2 plane. These are illustrated in Fig. 5.2. We can identify the corresponding loci by the numbers 1 through 6). In addition we have indicated the bandwidth and performance of the asymptotic solutions by a large dot near the identifying number of the loci aht approaches it. (We can indicate it by a single dot because the entire loci for the asymptotic solution corresponds to just one signal.)

In the λ_E - λ_B plane of Fig. 5.1 as we move from left to right on the solid lines, those that cause the determinant of Eq. 5.53 to vanish, we generate the loci in the d_g^2 -B^2 plane with the solid lines in the direction indicated by the arrowheads. We see that they evolve from the dots corresponding to the asymptotic solutions. As can also be seen, the loci corresponding to numbers

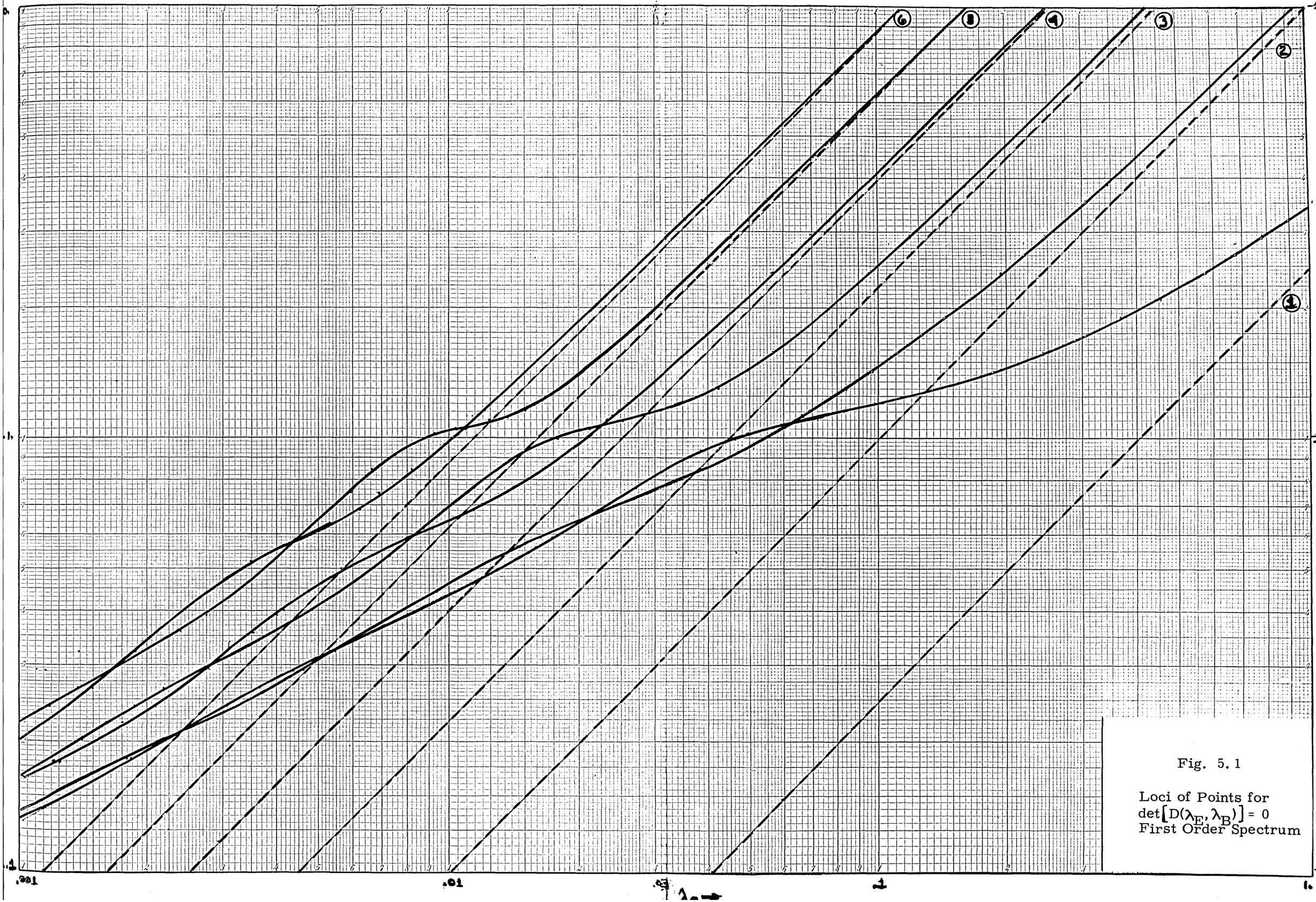


Fig. 5.1

Loci of Points for
$\det[D(\lambda_E, \lambda_B)] = 0$
First Order Spectrum

one and two are well behaved, while those corresponding to the remaining numbers have a rather erratic behavior.

As indicated by step (f) the other aspect of the optimization that needs to be considered is the crossings of the loci that occur in the λ_E - λ_B plane. At these crossings, there are two signals with the same values of λ_E and λ_B that satisfy the necessary conditions. If we find the degradation and bandwidth of all possible linear combinations of these signals (normalized to unit energy), we produce the loci in the d_g^2 - B^2 plane indicated by the dotted lines. (There is no relation between the dotted lines in each plane.) Since it is evident that only some crossings are significant, i.e., are candidates for the absolute minimum, we have not plotted the loci produced by all the crossings. Once we have generated these loci over the bandwidth constraint region of interest, we can find the optimal signal corresponding for any particular bandwidth constraint value. We merely select the signal which produces the absolute minimum degradation, or the lowest point in the d_g^2 - B^2 plane for the specified value of B.

At this point we pause to discuss the phenomenon we mentioned in the last section. To do this, let us trace the loci in the two planes. Let us start at the top right of the graph with λ_E large in absolute value. As we start to increase λ_E by moving downward, the locus in the d_g^2 - B^2 plane eminates from point one and proceeds across to the right. The point on this locus where the dotted line starts corresponds to the first loci crossing in the λ_E - λ_B plane. All the points on this dotted line correspond to this crossing point. As we again move downward increasing λ_E

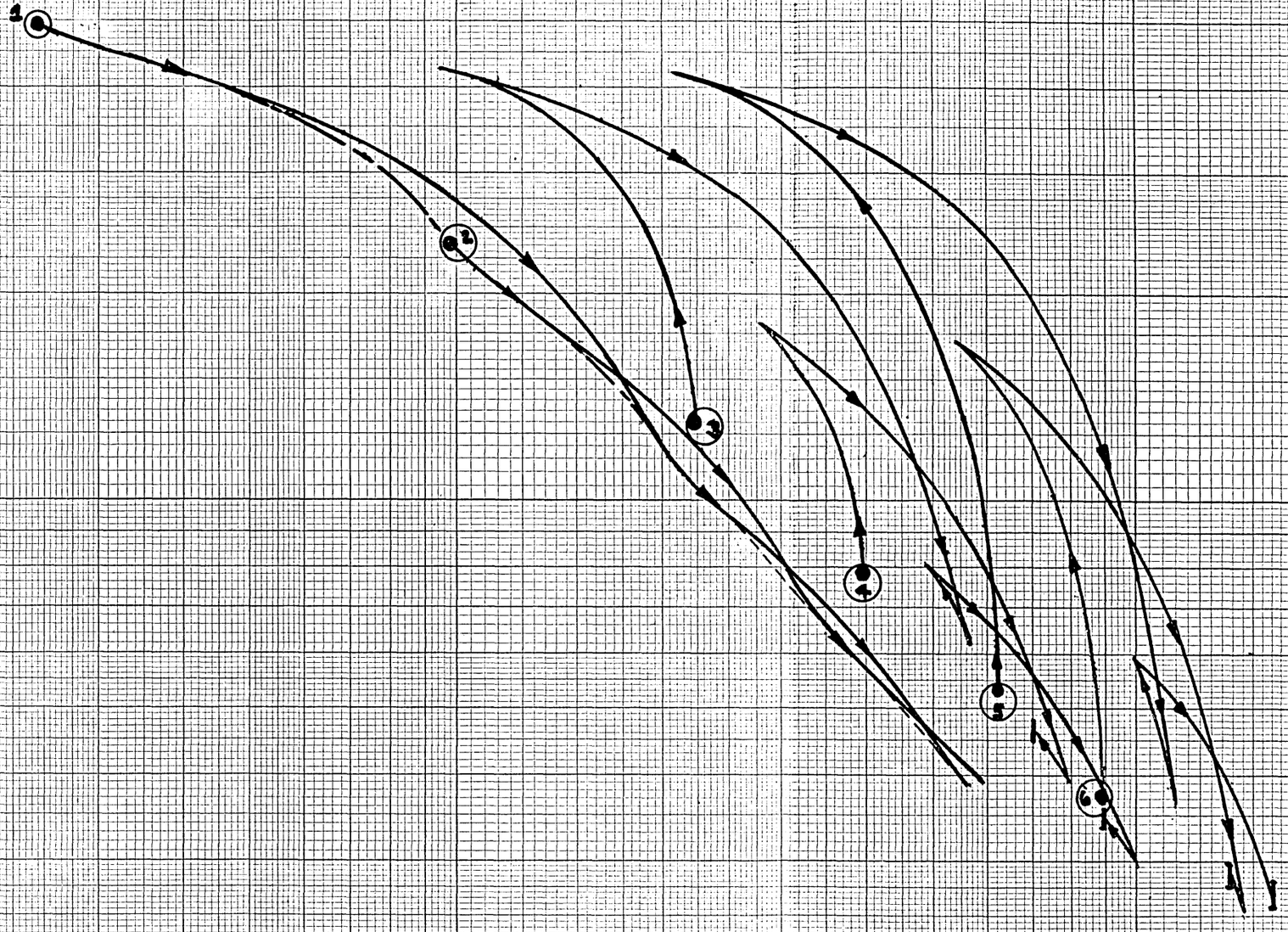


Fig. 5.2

Performance vs. Bandwidth
for the Optimal Signal Candidates
First Order Spectrum

and take the largest λ_B solution, we generate the solid line. This continues until we reach another dotted line which again corresponds to a loci crossing. We can continue in this manner as far as we wish. The important point is that we could ignore the complicated patterns corresponding to remaining values of λ_B that satisfied our test.

Let us illustrate the actual form of some of the optimal signals for two different values of the bandwidth constraint, B. In the first illustration we have B^2 equal to 9. If we examine the $d_g^2 - B^2$ of Fig. 5.3, we see that the optimal signal is a linear combination of the two solutions that are produced by the first crossing of loci one and two in the $\lambda_E - \lambda_B$ plane. The optimal signal for this constraint is illustrated in Fig. 5.3. We see that we can achieve a degradation of 23.6 percent less than the white noise only performance.

In addition, we have drawn g(t), the correlation signal for the optimal receiver. The signal(s) exhibit no particular symmetry for this constraint value of B equal to 3 as it is composed principally of the signals $\sin(\pi t/T)$ and $\sin(2\pi t/T)$. We should note that because of the possibility of two different sign reversals there are actually four signals that are optimal. All of these, however, basically have the same waveshape.

In Fig. 5.4 we show the optimal signal s(t) and its correlating signal g(t) when we allow twice as much bandwidth, i.e., $B^2 = 36$. In this case, the optimal signal does not correspond to a crossing in the $\lambda_E - \lambda_B$ plane, as it did previously. The bandwidth is sufficiently large enough so that we can attain a performance

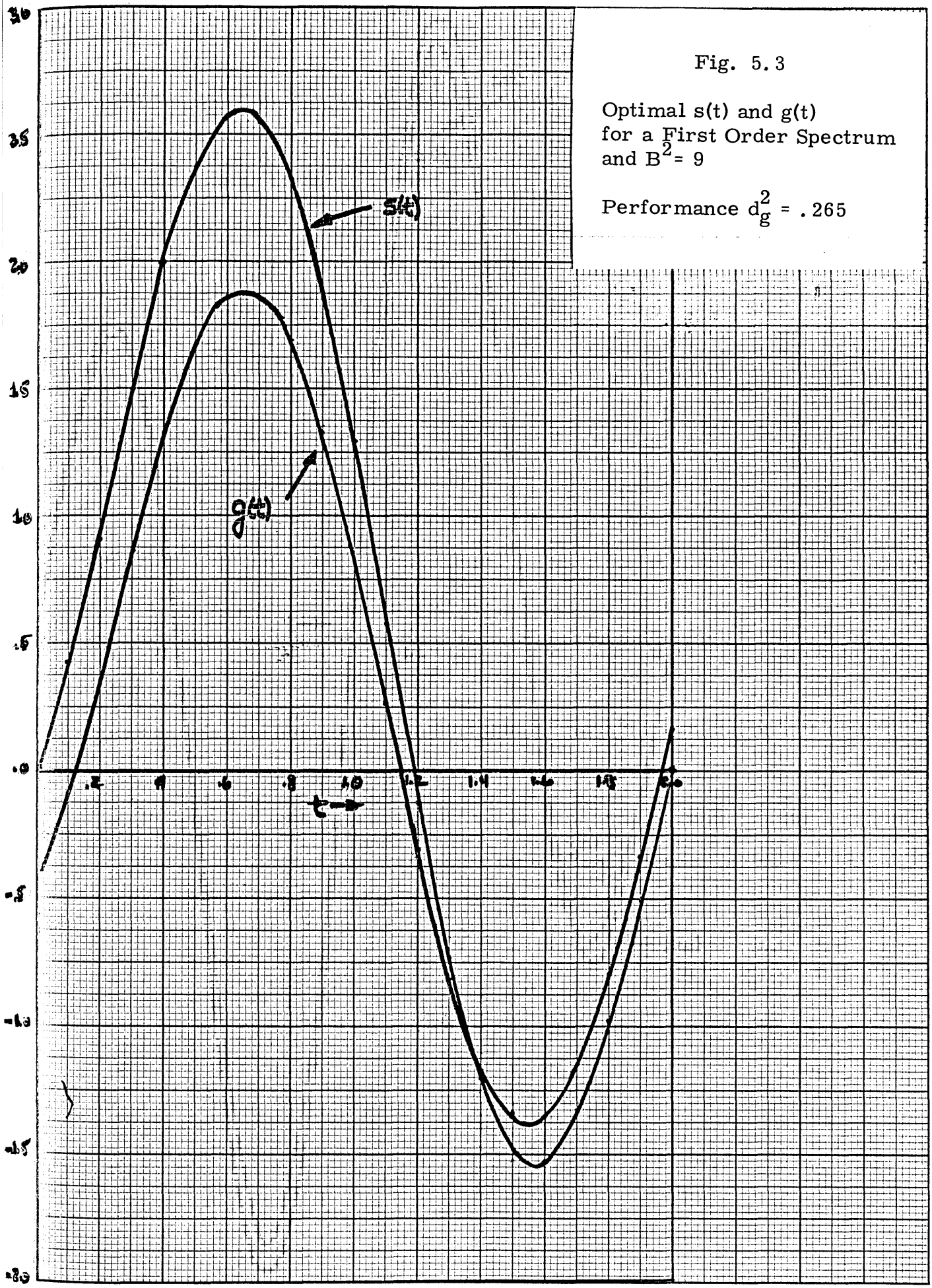


Fig. 5.3

Optimal s(t) and g(t)
for a First Order Spectrum
and $B^2 = 9$

Performance $d_g^2 = .265$

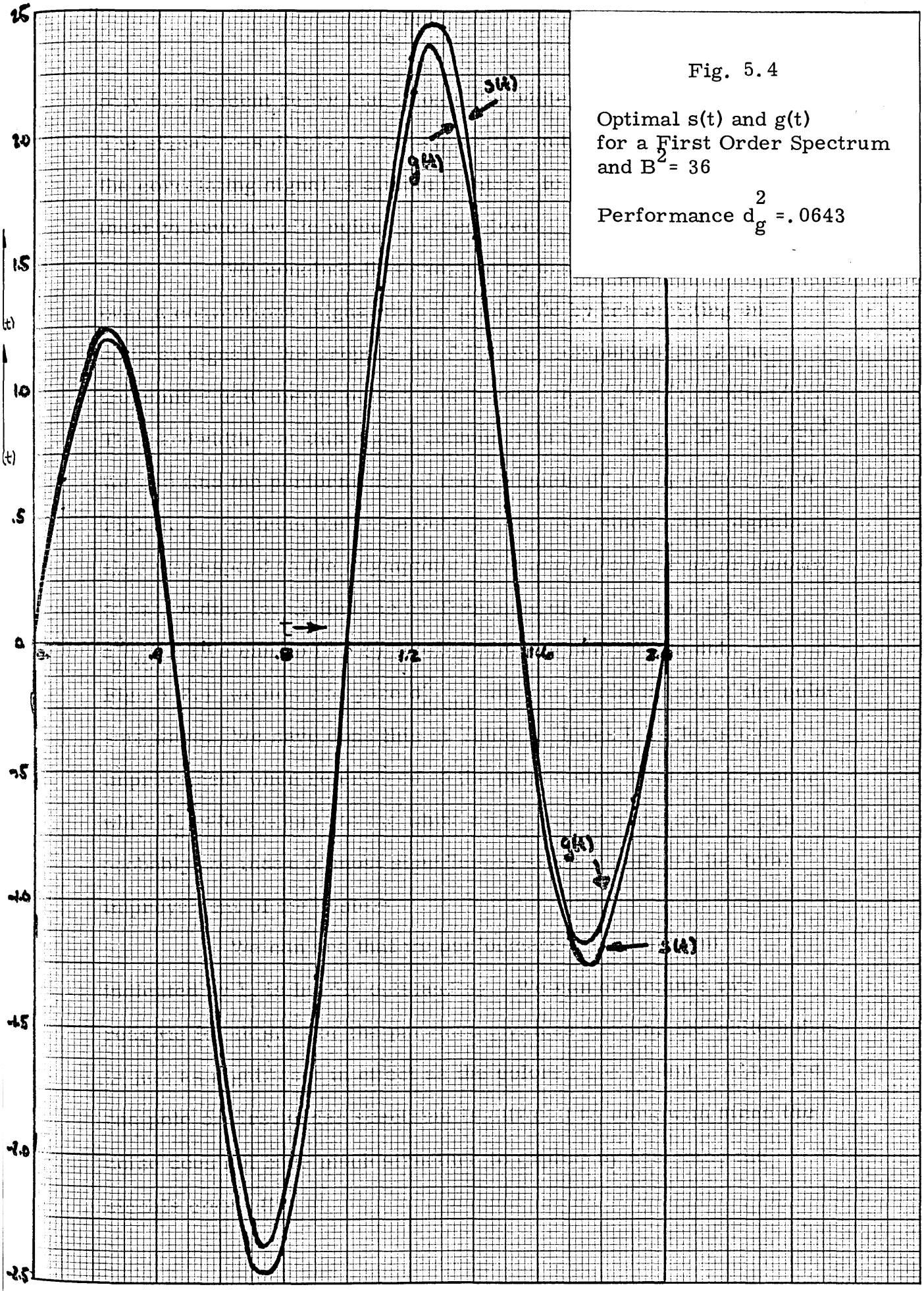


Fig. 5.4

Optimal s(t) and g(t)
for a First Order Spectrum
and $B^2 = 36$

Performance $d_g^2 = .0643$

only 6.4 percent below that of only the white noise being present. The signal does display some symmetry in that $s(t) = -s(T-t)$; and the correlating signal is almost identical to $s(t)$ indicating that we (nearly) have a white noise type receiver.

We conclude this example by discussing the improvement over a conventional signal that is attainable by transmitting an optimal signal. We propose the following comparison. Let us transmit a unit energy pulsed sine wave, i.e., $\sqrt{2E/T}\ \sin(n\pi t/T)$. The bandwidth that this signal consumes is $n\pi/T$, i.e., $B^2 = (n\pi/T)^2$. We can find the performance of these signals $d^2_{g\,sin}$ by referring to Fig. 4.2 or from the large dots on Fig. 5.2 since these signals also correspond to the asymptotic solutions. For this same amount of bandwidth, let us find the performance of the optimal signal $d^2_{g\ opt}$. (We can do this by looking directly beneath the dots in Fig. 5.2.) we now compare $d^2_{g\ opt}$ with the performance of the pulsed sine wave. We have done this in Table I. Index I_1 reflects the improvement when referenced to the pulsed sine wave performance while I_2 references it to the white noise only performance. I_2 is probably the index of most significance since it reflects the total improvement of the system.

We can see that the performance improvement indicated by I_2 is hardly outstanding. We can conjecture two reasons for this. For the parameter values chosen, the white noise is just as significant in effect as the colored component. Also, the one pole or first order spectrum is difficult to work with since the noise does not have much structure to exploit by designing the signals correctly.

TABLE I

Comparison of Performance for Optimum

and Pulsed Sine Wave Signals

n	$d^2_{g_{sin}}$	$d^2_{g_{opt}}$	$I_1 = \frac{d^2_{g_{sin}} - d^2_{g_{opt}}}{d^2_{g_{sin}}}$ 100%	$I_2 = \frac{d^2_{sin} - d^2_{opt}}{d^2_w} \times 100\%$
1	0.490	0.490	0	0
2	0.239	0.235	1.7	.4
3	0.130	0.105	19	2.5
4	0.078	0.057	27	2.3
5	0.053	0.037	30	1.6

Example 2 - Signal Design With a Two Pole Spectrum

In this example we shall consider the signal design problem when the colored component of the observation noise has a two pole spectrum. We do this primarily for several reasons. First, we want to develop a better understanding of the problems involved in implementing our algorithm. Secondly, we want to verify the phenomenom that we observed with the one pole spectrum. Thirdly, the one pole spectrum considered in the previous example can give deceiving results at times. Finally, we want to demonstrate that our optimization procedure can produce more significant improvements when working with a process that has more "structure" to it.

We assume that the colored component of the observation noise is generated by Eq. 2.18. In addition, we set the white noise spectral level to unity and assume that the interval length is two, i.e., $T_2 = 0$, $T_f = T = 2$. Hence, the colored component of the observation noise dominates for all frequencies less than say .8 Hz.

This particular colored noise spectrum is particularly interesting because of its shape. Suppose we constrain the bandwidth to be less than 3, or $B^2 = 9$. This corresponds to allowing frequencies which are lower than that frequency where the spectrum has its peak. We certainly do not want to put all the signal energy in frequencies near the peak. Yet, if the bandwidth is available we should be able to use it to our advantage. Although one can conjecture from either our previous example, or his engineering judgment that a linear combination of pulsed sine waves should be close to optimum, it is not apparent that this is so.

Let us proceed with the steps in our algorithm. The first six loci in the λ_E - λ_B plane are illustrated and numbered in Fig. 5.5. Again, we observe the asymptotic behavior for large negative values of λ_E. In two respects the loci exhibit a somewhat more complicated behavior than those for the one pole spectrum. First, they do not cross simply pairwise. Apparently, they cross with the adjacent locus; i.e., the second crosses the first and third, the third crosses the second and fourth, etc. Secondly, in some regions they are extremely close together.

In Fig. 5.6 we have plotted the corresponding loci in the d_g^2-B^2 plane. Their behavior is very similar to the one pole case. We see that the first and second loci are well behaved, while the others exhibit the same type of erratic behavior. As indicated by the dotted lines, there are regions where we must consider the linear combinations of signals corresponding to a loci crossing in the λ_E - λ_B plane. Most importantly, the maximum λ_B phenomenom still occurs.

Let us now examine the shape of an optimal signal when we constrain the bandwidth such that the colored noise is dominant over the allowed frequency range. If we chose $B^2 = 13$, the optimum signal s(t) and its correlating signal g(t) are illustrated in Fig. 5.7. The signal is principally composed of functions of the form $\sin(\pi t)$ and $\sin(3\pi t)$. The degradation realized by this signal is .53 which is approximately 15 percent below what one might expect using a non-optimum approach of selecting the pulsed sine wave with the best performance that still satisfies the bandwidth constraint.

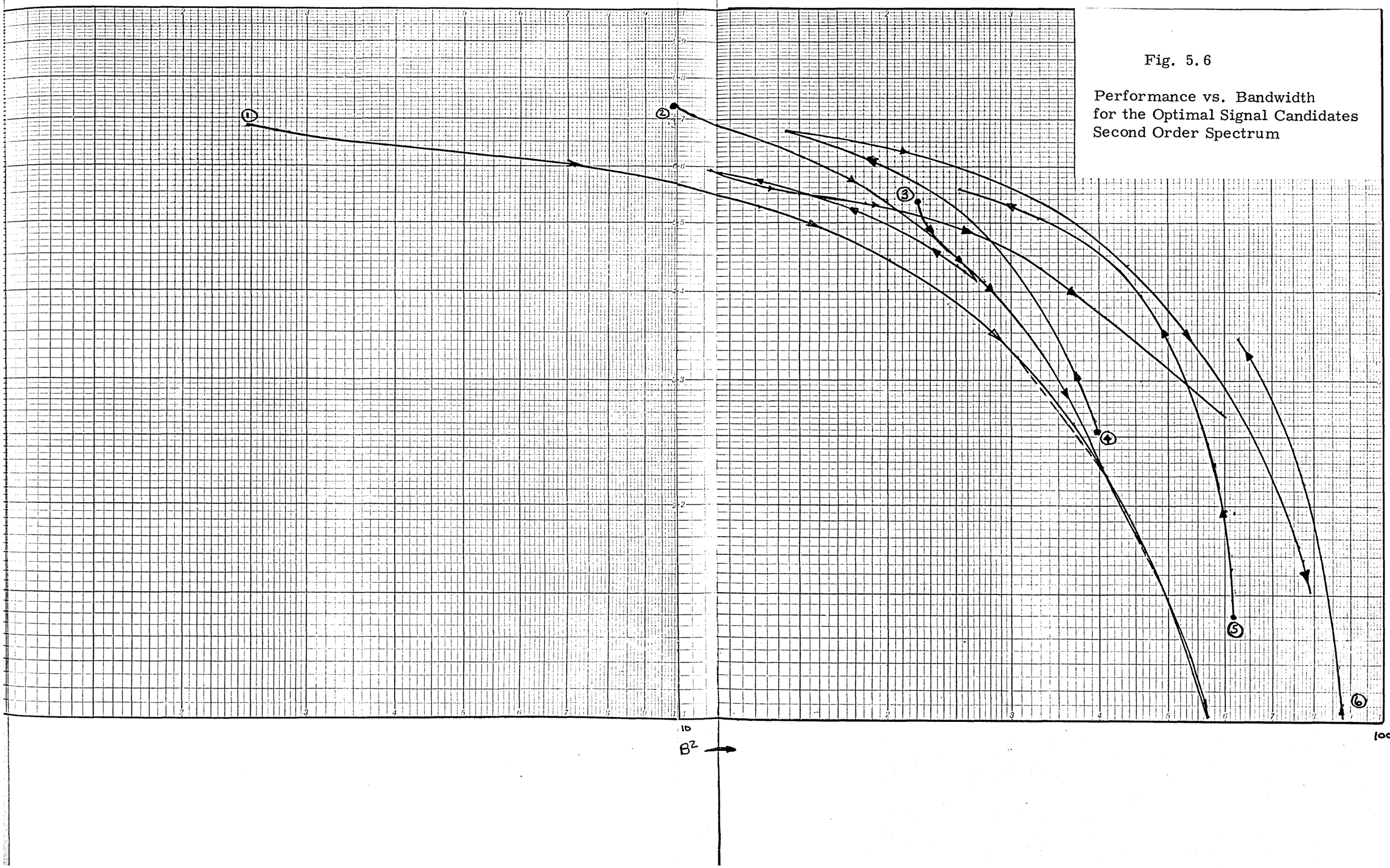


Fig. 5.6

Performance vs. Bandwidth
for the Optimal Signal Candidates
Second Order Spectrum

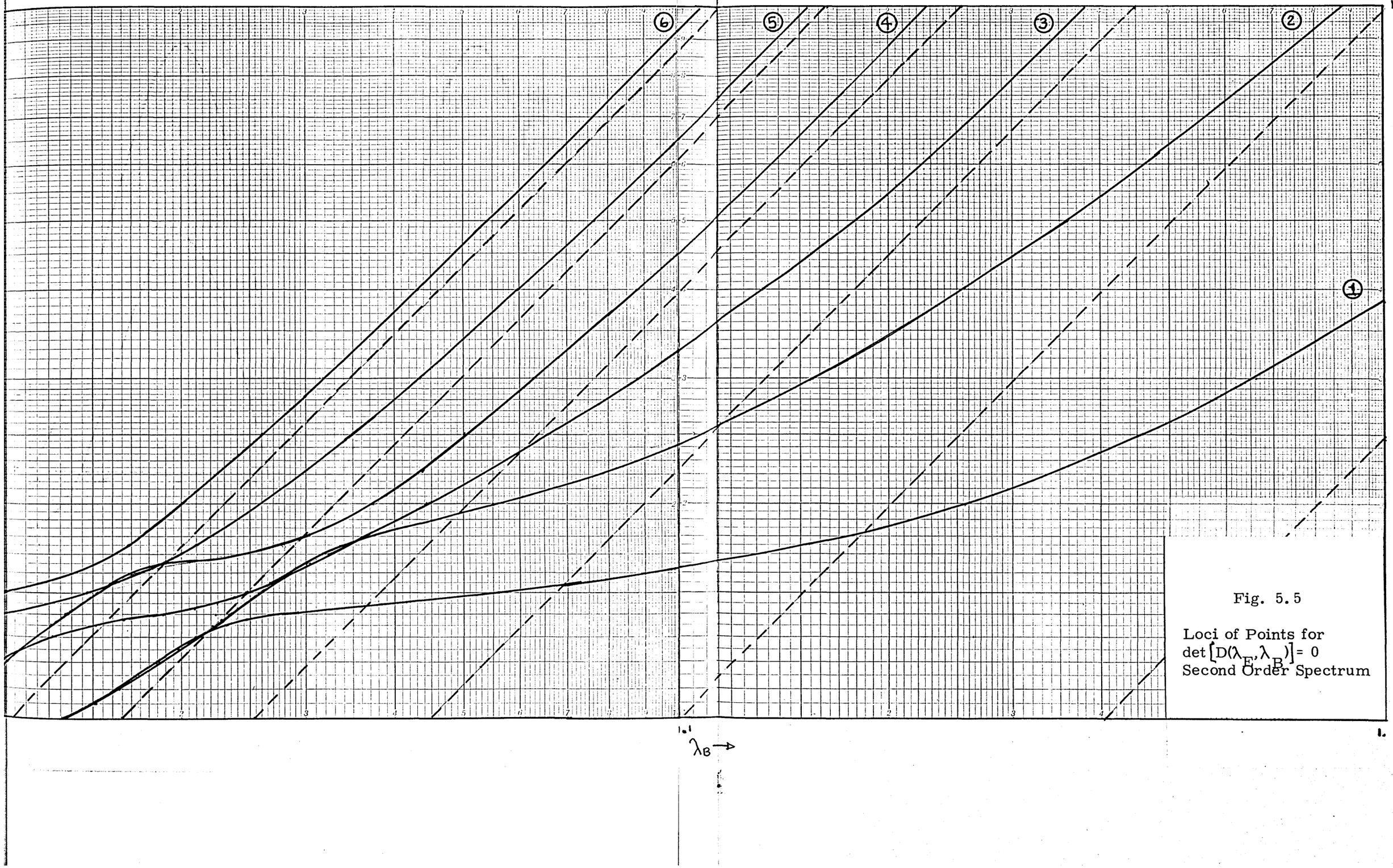


Fig. 5.5

Loci of Points for
$\det[D(\lambda_E, \lambda_B)] = 0$
Second Order Spectrum

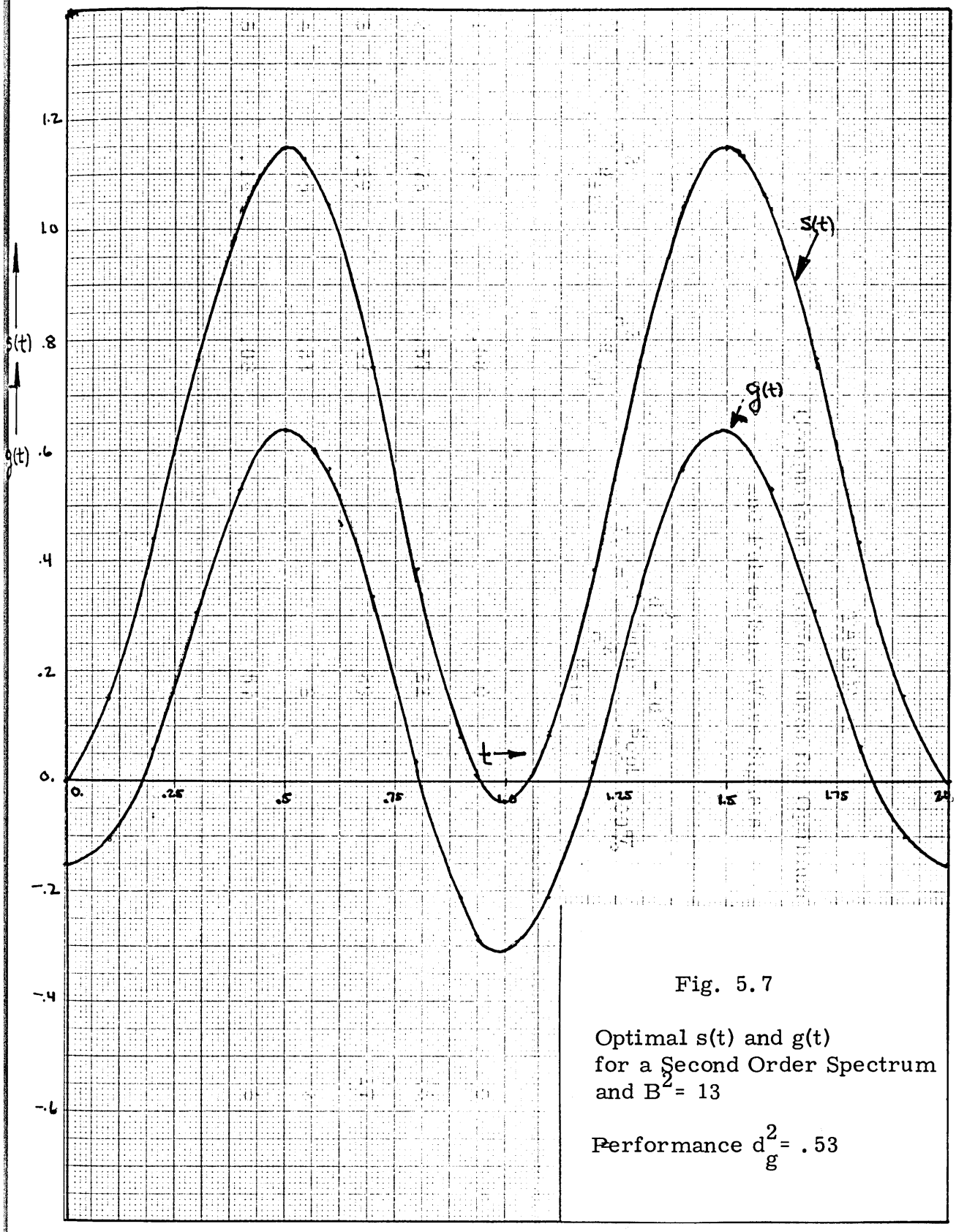


Fig. 5.7

Optimal s(t) and g(t)
for a Second Order Spectrum
and $B^2 = 13$

Performance $d_g^2 = .53$

TABLE II

Comparison of Performance for Optimum

and Pulsed Sine Wave Signals

n	$d^2_{g_{sin}}$	$d^2_{g_{opt}}$	$I_1 = \frac{d^2_{g_{sin}} - d^2_{g_{opt}}}{d^2_{g_{sin}}} 100\%$	$I_2 = \frac{d^2_{sin} - d^2_{opt}}{d^2_w} \times 100\%$
1	.68	.68	0	0
2	.73	.57	22	16
3	.59	.42	29	17
4	.25	.23	8	2
5	.14	.08	57	6

In Table II we have summarized the performance in the same way that we did in the first example of this section. In this example the optimization has improved our performance more significantly. This is probably due to the colored noise being more dominant in this example.

E. Signal Design With a Hard Bandwidth Constraint

In this section we shall derive the differential equations that specify the necessary condition for optimality when we constrain the absolute value of the derivative of the signal, i.e., we require

$$|v(t)| = \left| \frac{ds(t)}{dt} \right| \leq E^{\frac{1}{2}} B, \qquad T_o \leq t \leq T_f \tag{5.74}$$

As before, we constrain the signal energy by

$$\int_{T_o}^{T_f} s^2(\tau)d\tau \leq E \tag{5.75}$$

We can formulate the problem in a manner very similar to that used previously. The resulting Hamiltonian is

$$H(\xi, \eta, s, x_E, \underline{p}_{\xi}, \underline{p}_{\eta}, p_s, \lambda_E, v) =$$

$$p_o \frac{2}{N_o} s(t)C(t)\underline{\xi}(t) + \underline{p}_{\xi}^T(t)(F(t)\underline{\xi}(t) + G(t)QG^T(t)\underline{\eta}(t))$$

(continued)

$$\underline{p}_{\underline{\eta}}^{T}(t)(C^{T}(t)\frac{2}{N_o}C(t)\underline{\xi}(t) - F^{T}(t)\underline{\eta}(t) - C^{T}(t)\frac{2}{N_o}s(t))$$

$$+ p_s(t)v(t) + \frac{\lambda_E(t)}{2}s^2(t) \tag{5.76}$$

Taking the required derivatives, we find

$$\frac{\partial H}{\partial \underline{\xi}} = -\dot{\underline{p}}_{\underline{\xi}}(t) = p_o C^{T}(t)\frac{2}{N_o}s(t) + F^{T}(t)\underline{p}_{\underline{\xi}}(t)$$

$$+ C^{T}(t)\frac{2}{N_o}G(t)\underline{p}_{\underline{\eta}}(t) \tag{5.77}$$

$$\frac{\partial H}{\partial \underline{\eta}} = -\dot{\underline{p}}_{\underline{\eta}}(t) = G(t)Q\,G^{T}(t)\underline{p}_{\underline{\xi}}(t) - F(t)\underline{p}_{\underline{\eta}}(t) \tag{5.78}$$

$$\frac{\partial H}{\partial s} = -\dot{p}_s(t) = p_o\frac{2}{N_o}C(t)\underline{\xi}(t) - \frac{2}{N_o}C(t)\underline{p}_{\underline{\eta}}(t)$$

$$+ \lambda_E(t)s(t) \tag{5.79}$$

$$\frac{\partial H}{\partial E} = -\dot{\lambda}_E(t) = 0 \tag{5.80}$$

The transversatility conditions imply

$$P_o\underline{p}_{\underline{\xi}}(T_o) = -\underline{p}_{\underline{\eta}}(T_o), \tag{5.81}$$

$$\underline{p}_{\underline{\xi}}(T_f) = 0. \tag{5.82}$$

$p_s(T_f)$ and $p_s(T_o)$ are free.

We again have that λ_E is a constant. Furthermore, we note that Eqs. 5.77, 5.75, 5.81, 5.82 are identical to Eqs. 5.26, 5.27, 5.35, 5.34 respectively. Consequently, we can again show by using the results of Section B-3 that

$$\underline{\xi}(t) = -\underline{p}_{\underline{\eta}}(t) \tag{5.83a}$$

$$\underline{\eta}(t) = \underline{p}_{\underline{\xi}}(t) \tag{5.83b}$$

Therefore, Eq. 5.79 becomes (assuming p_o equals unity)

$$p_s(t) = -\frac{4}{N_o} C(t)\underline{\xi}(t) - \lambda_E s(t) \tag{5.84}$$

The major difference between the application of the Minimum Principle to this problem and the one in the text comes in the minimization of the Hamiltonian as a function of the control v(t). If $p_s(t)$ is non-zero, this minimization implies

$$v(t) = -E^{\frac{1}{2}}B \operatorname{sgn}(p_s(t)) \tag{5.85}$$

This implies that the optimal signal has a constant linear slope of $\pm E^{1/2}B$ when $p_s(t)$ is non-zero.

The differential equations, now non-linear, that specify the necessary conditions are

$$\frac{d\underline{\xi}(t)}{dt} = F(t)\underline{\xi}(t) + G(t)Q\,G^T(t)\underline{\eta}(t) \tag{5.86}$$

$$\frac{d\underline{\eta}(t)}{dt} = C^T(t)\frac{2}{N_o}C(t)\underline{\xi}(t) - F^T(t)\underline{\eta}(t) - C^T(t)\frac{2}{N_o}s(t) \tag{5.87}$$

$$\frac{ds(t)}{dt} = v(t) = -E^{\frac{1}{2}}B\,\mathrm{sgn}(p_s(t)) \tag{5.88}$$

$$\frac{dp_s(t)}{dt} = -\frac{4}{N_o}C(t)\underline{\xi}(t) - \lambda_E s(t) \tag{5.89}$$

The boundary conditions are

$$\underline{\xi}(T_o) = P_o\underline{\eta}(T_o) \tag{5.90}$$

$$\underline{\eta}(T_f) = 0 \tag{5.91}$$

$$s(T_o) = s(T_f) = 0 \tag{5.92}$$

If $p_s(t)$ is zero over any region, we cannot determine v(t) by Eq. 5.85. In such region

$$p_s(t) = 0 \tag{5.93}$$

implies

$$s(t) = -\frac{4}{\lambda_E N_o}C(t)\underline{\xi}(t) \tag{5.94}$$

If we combine this equation with Eq. 5.86, we find that the two equations have the same form as those that specify the eigenfunctions associated with the colored noise.

This leads us to the following conjecture: the optimal signal consists of regions of a constant slope of $\pm E^{1/2}B$ and regions where the signal has the same functional form as the eigenfunction of the colored noise. Finding a solution technique for solving these differential equations will be part of future research.

F. Summary and Discussion

We have presented a state variable method for designing optimal signals for detection in colored noise when there are energy and bandwidth constraints. The performance measure was given by d_g^2, which specified the loss of receiver performance due to the colored noise being present. We used the differential equations and their associated boundary condition that specified the optimal receiver and performance measure as if they described a dynamic system. We then applied Pontryagin's Minimal Principle to derive the necessary conditions that the optimal signal must satisfy. These conditions specified a characteristic value problem. We could determine the optimal signal by solving this characteristic value problem.

We suggested a computer algorithm for doing this. By using this algorithm we were able to analyze two examples of colored noise spectra. In addition to finding the optimal signal and its associated performance, the algorithm displayed several interesting features.

One may argue that we did not need to use the Minimum Principle to solve this problem since we could proceed directly to the 2n + 2 differential equations and boundary conditions via the usual calculus of variations and the estimator-subtractor realization of the receiver. The advantage of this formulation becomes apparent when we change the type of constraints upon the signal. In Section E we derived the differential equations that specify the optimal signal when we impose an energy constraint and a hard (bandwidth) constraint, $|ds(t)|dt| \leq B$. This problem is readily solved using a Minimal Principle, whereas the calculus of variations would require much more effort.

We intend to continue studying the issue of hard constraints by examining the design problem when we impose peak constraints on the signal itself and/or its derivative. We expect that the important aspect will be to find an efficient computer algorithm to solve the resulting differential equations.

There are several other important issues.

1. The most difficult aspect of our method is to select the signal that is globally optimum from all those that satisfy the necessary conditions optimum. As we have discussed in the text, we have observed a phenomenom which eliminates this problem. We feel that there is sufficient empirical evidence to spend some time trying to verify this theoretically.

2. The equivalence of $\underline{\xi}(t)$ and $-\underline{p}_{\eta}(t)$, and $\underline{\eta}(t)$ and $\underline{p}_{\xi}(t)$ is a very general result. Because of this generality it seems that there should be a means of formulating these problems with a stochastic minimum principle which would allow us to obtain the

final differential equations directly. However, this means has not been apparent so far.[29]

3. One can alsopose the optimization problem when we use a suboptimum receiver. For example, if we constrain the receiver to be a matched filter, one wants to minimize the functional

$$d_g^{2\prime} = \int_{T_o}^{T_f}\int_{T_o}^{T_f} s(t)K_y(t,\tau)s(\tau)d\tau = \int_{T_o}^{T_f} s(t)C(t)\underline{\xi}(t)dt \tag{5.95}$$

where we have defined

$$\underline{\xi}(t) = \int_{T_o}^{T_f} K_{\underline{x}}(t,\tau)C(\tau)s(\tau)d\tau \qquad T_o \leq t \leq T_f \tag{5.96}$$

The results of Chapter II are now directly applicable, and we can proceed in a manner analogous to the way we did in this chapter. Similar results should exist for other types of suboptimal receivers, e.g., those operating in a reverberation environment.

4. The algorithm that we used requires an accurate method for calculating transition matrices. All the problems that we have examined to date have involved constant parameter systems. Consequently, we could use the matrix exponential. By being reasonably careful we have not encountered any difficulty when we used a straightforward series sum approximation. However, this is certainly no guarantee that this series approach will always

work, for example as the dimension of the system increases. We think that it is worthwhile to investigate other methods to compare their accuracy and speed to the present method.

5. Finally, the results in this section can be extended to bandpass signal design by using the material in Appendix B. The possibility of a noise spectrum that is not symmetrical about the carrier introduces some interesting questions on how to best take advantage of this situation.

CHAPTER VI

LINEAR SMOOTHING AND FILTERING WITH DELAY

In this chapter we shall use the results derived in Chapter IV for solving nonhomogeneous Fredholm integral equations to develop a unified approach to linear smoothing and filtering with delay. In the smoothing problem we receive a signal over a fixed time interval $[T_o, T_f]$. We then want to find $\hat{\underline{x}}(t)$, $T_o \leq t \leq T_f$, the estimate of the state vector that generated this signal over the same fixed interval. In a filtering problem we receive a signal continuously, i.e. the endpoint time of the interval, T_f, is constantly increasing. We then want to produce an estimate which evolves in time as a function of this endpoint. For the realizable filter we want to find $\hat{\underline{x}}(T_f)$ vs. T_f, the estimate of the state vector right at the endpoint time, T_f. For the filter with delay, we are allowed a delay before we make our estimate. We want to find $\hat{\underline{x}}(T_f - \Delta)$ vs. T_f, the estimate of the state vector Δ units prior to the endpoint of the interval.

Our approach to these problems is straightforward. We start with the Wiener-Hopf equation that specifies the impulse response for the optimal linear estimator.[3] We then show how we can find a set of differential equations that specify the optimal estimate implicitly as part of their solution. From these equations we can derive matrix differential equations which determine the covariance

of error.

Our approach to the problem of filtering with delay is also straightforward. In the solution to the optimal smoother, we simply allow the endpoint time of the observation interval to be a variable. We then derive a set of differential equations which are a function of this variable endpoint time rather than the time within a fixed observation interval as for the smoother. The performance is also derived in an analogous manner.

Several comments are in order before proceeding. First, all the derivations in this chapter are original. Many of the results, however, have appeared in the literature. They are referenced where appropriate. The most important point to be made concerning our methods is the approach. The entire theory can be developed concisely and directly starting from a few basic results.

Secondly, we shall employ a structured approach to this topic. We shall require the estimator structure to be linear, regardless of the statistics of the processes involved. Existing approaches to this problem are unstructured; it is assumed that the processes involved are Gaussian and then the estimator structure is derived. It is well known, however, that both approaches yield the same estimator.

Thirdly, we shall assume that the reader is familiar with the well known results for realizable filtering by using state variable techniques, i.e., the Kalman-Bucy filter.[8,3]

Finally, although these topics have been extensively studied in the literature, several incorrect derivations and results have appeared. We shall point out these errors and give the

correct results.

A. The Optimal Linear Smoother

In this section, we shall derive a state variable realization of the optimum linear smoother, the state variable equivalent of the unrealizable filter. First, we shall establish our model.

Let us assume that we generate a random process $\underline{y}(t)$ by the methods described in Chapter II. Let us also assume that we observe this process in the presence of an additive white noise $\underline{w}(t)$ over the interval $T_o \leq t \leq T_f$. That is, we receive the signal

$$\underline{r}(t) = \underline{y}(t) + \underline{w}(t) = C(t)\,\underline{x}(t) + \underline{w}(t), \qquad T_o \leq t \leq T_f, \tag{6.1}$$

where $\underline{x}(t)$ is the state vector of the system that generates $y(t)$, $\underline{w}(t)$ is an additive white observation noise that has a covariance $R(t)\,\delta(t-\tau)$ as given by Eq. 2.7.

In the optimal smoothing problem we want to find a state-variable description of the linear system that minimizes the mean-square error in estimating each component of the state vector $\underline{x}(t)$. We shall find that this description consists of two first-order vector differential equations having a two-point boundary restrictions.

Let us now proceed with the derivation of these equations.

First, we define the matrix impulse response of the optimal linear smoother to be $\underline{h}_o(t,\tau)$, or

$$\hat{\underline{x}}(t) = \int_{T_o}^{T_f} \underline{h}_o(t,\tau)\underline{r}(\tau)d\tau, \qquad T_o \le t \le T_f.^{*} \tag{6.2}$$

This operator produces an estimate of the state vector, $\hat{\underline{x}}(t)$ at time t by observing $\underline{r}(\tau)$ over the entire interval $[T_o, T_f]$.

It is well known and can easily be shown by classical methods that this impulse response satisfies the following Wiener-Hopf integral equation.

$$K_{\underline{dr}}(t,\tau) = \int_{T_o}^{T_f} \underline{h}(t,\nu)K_{\underline{r}}(\nu,\tau)d\nu, \qquad T_o \le t,\ \tau \le T_f, \tag{6.3}$$

where $K_{\underline{dr}}(t,\tau)$ is the cross covariance of the desired signal and the received signal, $K_{\underline{r}}(t,\tau)$ is the covariance of the received signal.

For our application, the desired signal is the state vector, $\underline{x}(t)$. Therefore, we have

$$K_{\underline{x}}(t,\tau)C^{T}(\tau) = \int_{T_o}^{T_f} \underline{h}(t,\nu)K_{\underline{r}}(\nu,\tau)d\nu, \qquad T_o \le t,\ \tau \le T_f,^{\dagger} \tag{6.4}$$

---

*Although $\underline{h}_o(t,\tau)$ is a matrix and should be denoted by a capital letter in our notation convention, we shall defer to the conventional notation.

†We have assumed zero means for $\underline{x}(T_o)$, $\underline{\nu}(t)$ and $\underline{w}(t)$. If the means were non zero, we would need to add a bias term to Eq. 6.2.

with

$$\underline{K}_r(t,\tau) = \underline{K}_y(t,\tau) + R(t)\delta(t-\tau). \tag{6.5}$$

The first step in our derivation is to solve this equation for $\underline{h}_o(t,\tau)$. In order to do this, we need to introduce the inverse kernel $Q_r(t,\tau)$ of $K_r(t,\tau)$, the covariance of the received signal.

Let us now introduce some material from Chapter IV on the nonhomogeneous that we need. From Eq. 4.3 we can write the nonhomogeneous Fredholm integral equation as

$$\int_{T_o}^{T_f} \underline{K}_r(t,\tau)\underline{g}(\tau) = \underline{s}(t), \qquad T_o \le t \le T_f. \tag{6.6}$$

As we discussed in Chapter IV, we can consider that this integral equation specifies a linear operator upon $\underline{s}(t)$, with the solution $\underline{g}(t)$ being the result of this linear operation. We define the integral representation of this operation to be (Eq. 4.73).

$$\underline{g}(t) = \int_{T_o}^{T_f} \underline{Q}_r(t,\tau)\underline{s}(\tau)d\tau, \quad T_o \le t \le T_f. \tag{6.7}$$

Operating upon $\underline{s}(t)$ with $\underline{Q}_r(t,\tau)$ to find $\underline{g}(t)$ is equivalent to solving the integral equation by means of our differential equation approach. It is easy to show that the inverse kernel satisfies the following

integral equation in two variables

$$\int_{T_o}^{T_f} \underline{K}_r(t,\nu)\underline{Q}_r(\nu,\tau)d\nu = I\delta(t-\tau), \quad T_o \le t,\ \tau \le T_f,^{*} \tag{6.8}$$

Let us multiply both sides of Eq. 6.4 by $\underline{Q}_r(\tau,z)$ and then integrate with respect to τ. This yields

$$\int_{T_o}^{T_f} \underline{K}_x(t,\tau)C^T(\tau)\underline{Q}_r(\tau,z)d\tau = \int_{T_o}^{T_f} \underline{h}(t,\nu) \int_{T_o}^{T_f} \underline{K}_r(\nu,\tau)\underline{Q}_r(\tau,z)d\tau d\nu$$

$$= \int_{T_o}^{T_f} \underline{h}_o(t,\nu)I\delta(\nu-z)d\nu = \underline{h}_o(t,z), \quad T_o \le t,\ z \le T_f. \tag{6.9}$$

We are not directly interested in the impulse response of the optimum estimator. What we really want to find is the estimate $\underline{\hat{x}}(t)$, which is the output of the estimator. We can obtain this by substituting Eq. 6.9 into Eq. 6.2,

$$\underline{\hat{x}}(t) = \int_{T_o}^{T_f} \underline{K}_x(t,\tau)C^T(\tau) \left(\int_{T_o}^{T_f} \underline{Q}_r(\tau,z)\underline{r}(z)dz \right) d\tau, \quad T_o \le t \le T_f. \tag{6.10}$$

---

*We should observe that the inverse kernel can be shown to be symmetric, i.e. $Q_r(t,\tau) = Q_r^T(\tau,t)$; therefore, we can define it as a pre- or post-multiplier operator.

Thus, the optimum estimate is the result of two integral operations. We now want to show how we can reduce Eq. 6.10 to two differential equations with an associated set of boundary conditions. The estimate $\hat{\underline{x}}(t)$ is specified implicitly by their solution.

Let us define the term in parenthesis in Eq. 6.10 as $\underline{g}_{\underline{r}}(\tau)$, so that we have

$$\underline{g}_{\underline{r}}(\tau) = \int_{T_o}^{T_f} Q_{\underline{r}}(\tau, z)\underline{r}(z)dz, \qquad T_o \leq \tau \leq T_f. \tag{6.11}$$

Substituting this into Eq. 6.9 gives us

$$\hat{\underline{x}}(t) = \int_{T_o}^{T_f} K_{\underline{x}}(t, \tau)C^T(\tau)\underline{g}_{\underline{r}}(\tau)d\tau, \qquad T_o \leq t \leq T_f. \tag{6.12}$$

Observe that Eqs. 6.10 and 6.11 are integral operations of the type encountered in Chapters IV and II, respectively. Consequently, we can convert each into two vector differential equations with an associated set of boundary conditions.

From our previous discussion, $\underline{g}_{\underline{r}}(\tau)$ is the solution to the nonhomogeneous Fredholm integral equation when the signal $\underline{s}(t)$ is replaced by $\underline{r}(t)$. From Chapter IV, we have

$$\underline{g}_{\underline{r}}(\tau) = R(\tau)[\underline{r}(\tau) - C(\tau)\underline{\xi}(\tau)], \tag{6.13}$$

where $\underline{\xi}(\tau)$ is the solution to the differential equations, (Eqs. 4.10 and 4.12

$$\frac{d\underline{\xi}(\tau)}{d\tau} = F(\tau)\underline{\xi}(t) + G(\tau)Q\,G^{T}(\tau)\underline{\eta}_1(\tau), \quad T_o \le \tau \le T_f \tag{6.14}$$

$$\frac{d\underline{\eta}_1(\tau)}{d\tau} = C^{T}(\tau)R^{-1}(\tau)C(\tau)\underline{\xi}(\tau) - F^{T}(\tau)\underline{\eta}_1(\tau) - C^{T}(\tau)R^{-1}(\tau)\underline{r}(\tau)$$

$$T_o \le \tau \le T_f \tag{6.15}$$

The boundary conditions are (Eqs. 4.13 and 4.14)

$$\underline{\xi}(T_o) = P_o\underline{\eta}_1(T_o), \tag{6.16}$$

$$\underline{\eta}_1(T_f) = \underline{0}. \tag{6.17}$$

In Chapter II we found that the integral operation given by Eq. 6.12 also has a differential equation representation. If in Eq. 2.22 we set

$$\underline{f}(t) = \underline{g}_{\underline{r}}(t), \qquad T_o \le t \le T_f, \tag{6.18}$$

and then substitute Eq. 6.13 for $\underline{g}_r(t)$, we find that $\underline{x}(t)$ can be found by solving the differential equations

$$\frac{d\hat{\underline{x}}(t)}{dt} = F(t)\hat{\underline{x}}(t) + G(t)Q\,G^{T}(t)\underline{\eta}_2(t), \quad T_o \le t \le T_f \tag{6.19}$$

$$\frac{d\underline{\eta}_2(t)}{dt} = -C^T(t)\underline{g}_{\underline{r}}(t) - F^T(t)\underline{\eta}_2(t)$$

$$= C^T(t)R^{-1}(t)C(t)\underline{\xi}(t) - F^T(t)\underline{\eta}_2(t) - C^T(t)R^{-1}(t)\underline{r}(t).$$

$$T_o \leq t \leq T_f \qquad (6.20)$$

The boundary conditions for these equations (Eqs. 2.33 and 2.35)

$$\hat{\underline{x}}(T_o) = P_o\underline{\eta}_2(T_o) \qquad (6.21)$$

$$\underline{\eta}_2(T_f) = 0 \qquad (6.22)$$

Upon a first inspection it appears that we need to solve four coupled vector differential equations with their associated boundary conditions. However, if we examine Eqs. 6.15 and 6.20, we find that $\underline{\eta}_1(t)$ and $\underline{\eta}_2(t)$ satisfy the same differential equation. Since both equations have the same boundary condition at $t = T_f$ (Eqs. 6.17 and 6.22), they must have identical solutions. Therefore, we have

$$\underline{\eta}_1(t) = \underline{\eta}_2(t) \triangleq \underline{p}(t), \qquad T_o \leq t \leq T_f \qquad (6.23)$$

By replacing $\underline{\eta}_1(t)$ and $\underline{\eta}_2(t)$ by $\underline{p}(t)$ in Eqs. 6.13, 6.15, 6.18 and 6.20, we see that $\underline{\xi}(t)$ and $\hat{\underline{x}}(t)$ satisfy the same differential equations (Eqs. 6.14 and 6.19) and have the same boundary conditions (Eqs. 6.16 and 6.21).

Therefore, we must also have

$$\underline{\hat{x}}(t) = \underline{\xi}(t), \qquad T_o \leq t \leq T_f . \tag{6.24}$$

Consequently, we have shown that two of the four differential equations are redundant.

We finally obtain the state variable representation of the optimum linear smoother. The optimum estimate $\hat{x}(t)$ satisfies the differential equations

$$\frac{d\underline{\hat{x}}(t)}{dt} = F(t)\underline{\hat{x}}(t) + G(t)Q\,G^T(t)\underline{p}(t), \qquad T_o \leq t \leq T_f, \tag{6.25}$$

$$\frac{d\underline{p}(t)}{dt} = C^T(t)R^{-1}(t)C(t)\underline{\hat{x}}(t) - F^T(t)\underline{p}(t) - C^T(t)R^{-1}(t)\underline{r}(t), \qquad T_o \leq t \leq T_f \tag{6.26}$$

where we impose the boundary conditions

$$\underline{\hat{x}}(T_o) = P_o\,\underline{p}(T_o), \tag{6.27}$$

$$\underline{p}(T_f) = \underline{0}. \tag{6.28}$$

The smoother realization specified by Eqs. 6.25 to 6.28 is well known. It was first derived by Bryson and Frazier in reference 14 be assuming Gaussian statistics and then using a variational approach to maximize the a posteriori probability of the state vector.

When we compare these equations to those in Chapter IV that specified our solution to the nonhomogeneous Fredholm integral

equation, we observe that they are identical in form. The major difference is that our input is now a random process $\underline{r}(t)$, whereas before we had a known signal $\underline{s}(t)$. The result of this observation is that the solution methods developed in Chapter IV, Section C are also applicable to solving the above estimation equations for the smoother. (In fact, the methods presented there were originally developed in the literature for solving these estimator equations. We have exploited the above identity of form to solve the equations we derived in Chapter IV. In order to make use of these methods, it is obvious that one must identify $\underline{\hat{x}}(t)$ with $\underline{\xi}(t)$ and $\underline{p}(t)$ with $\underline{\eta}(t)$.

Since we shall need the results in the next section, it is useful to relate the smoothing structure derived above to the realizable filter structure. To do this we shall review some of the relationships that we derived in Chapter IV. If we make the identity suggested above, the variable $\underline{\xi}_r(t)$ is easily seen to correspond to the realizable filter estimate, $\underline{\hat{x}}_r(t)$. It is well known, or it can be seen from Eq. 4.39 , that $\underline{\hat{x}}_r(t)$ satisfies the equation

$$\frac{d}{dt}\underline{\hat{x}}_r(t) = F(t)\underline{\hat{x}}_r(t) + \Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)(\underline{r}(t)-C(t)\underline{\hat{x}}_r(t)),^*$$

$$T_o < t, \qquad (6.29)$$

where $\Sigma(t/t)$ satisfies the variance equation (Eq. 4.34).

$$\frac{d\Sigma(\frac{t}{t})}{dt} = F(t)\Sigma(\tfrac{t}{t}) + F^T(t)\Sigma(\tfrac{t}{t}) - \Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)C(t)\Sigma(\tfrac{t}{t})+G(t)QG^T(t) \qquad (6.30)$$

---

*We shall use the notation $\Sigma(t/T)$ and $\Sigma(\frac{t}{T})$ interchangeably. Both symbols denote the covariance of error at time t ($\geq T_o$) when we have observed the signal $\underline{r}(\tau)$ over the interval $[T_o, T]$.

We have not demonstrated directly that $\Sigma(t/t)$ is indeed the covariance matrix of the realizable filter. It is straightforward to do so, therefore, we omit it.

The smoother estimate $\hat{\underline{x}}(t)$ is related to the realizable estimate $\hat{\underline{x}}_r(t)$ in two ways. Quite obviously, the estimate correspond at the end of the interval as stated by Eq. 4.37

$$\hat{x}(t)\Big|_{t = T_f} = \hat{\underline{x}}_r(T_f). \tag{6.31}$$

In addition we have the important relationship throughout the interval as expressed by Eq. 4.43

$$\Sigma(\tfrac{t}{t})\underline{p}(t) = \hat{\underline{x}}(t) - \hat{\underline{x}}_r(t), \qquad T_o \leq t \leq T_f \tag{6.32}$$

We shall often exploit this relationship in the remainder of the chapter.

There is one important contrast between the two structures. In the realizable filter, the covariance of error, $\Sigma(t/t)$, was implicit in the filter structure. In the smoother, the corresponding covariance $\Sigma(t/T_f)$ is not. Deriving an equation for this covariance is the topic of our next section.

B. Covariance of Error for the Optimum Smoother

In this section we shall derive four matrix differential equations, each of which specifies the performance of the optimal smoother. Since this is a rather long section, we shall pause briefly to outline our development.

First, we shall derive differential equations for both the smoother error and the realizable filter error processes. We shall evaluate three expectations which involve these error processes and the noise processes $\underline{u}(t)$ and $\underline{w}(t)$. Using these expectations we shall then derive the four differential equations that specify the covariance of error, $\Sigma(t/T_f)$, of the optimal smoother. We shall point out and correct some errors that have appeared in the literature on this topic.[14,15] Finally, we shall suggest an algorithm for finding the steady state covariance of error of the realizable filter, Σ_∞, by using some of the results that we have derived for the smoother performance. Let us now proceed.

The starting point for our analysis is finding the differential equations for $\underline{\epsilon}(t)$ and $\underline{\epsilon}_r(t)$, the error process for the optimal smoother and the realizable filter respectively. First, we consider the smoother error.

We note that the process $\underline{x}(t)$ as generated by Eq. 2.1 satisfies the differential equation

$$\frac{d\underline{x}(t)}{dt} = F(t)\underline{x}(t) + G(t)\underline{u}(t), \qquad T_o \leq t. \tag{2.1 (repeated)}$$

We define the smoother estimation error to be

$$\underline{\epsilon}(t) \triangleq \hat{\underline{x}}(t) - \underline{x}(t), \qquad T_o \leq t \leq T_f. \tag{6.33}$$

If we subtract Eq. 2.1 from Eq. 6.25, we find that the error satisfies the differential equation

$$\frac{d\underline{\epsilon}(t)}{dt} = F(t)\underline{\epsilon}(t) + G(t)Q\,G^T(t)\underline{p}(t) - G(t)\underline{u}(t),\ T_o \leq t \leq T_f. \qquad (6.34)$$

We can also write Eq. 6.26 in terms of the error, $\underline{\epsilon}(t)$. By substituting Eq. 6.1 into Eq. 6.26, we obtain

$$\begin{aligned}\frac{d\underline{p}(t)}{dt} &= C^T(t)R^{-1}(t)C(t)\hat{\underline{x}}(t) - F^T(t)\underline{p}(t) - C^T(t)R^{-1}(t)(C(t)\underline{x}(t) + \underline{w}(t)) \\ &= C^T(t)R^{-1}(t)C(t)\underline{\epsilon}(t) - F^T(t)\underline{p}(t) - C^T(t)R^{-1}(t)\underline{w}(t), \\ &\qquad T_o \leq t \leq T_f. \qquad (6.35)\end{aligned}$$

We can also find the boundary conditions for Eqs. 6.34 and 6.35. We still have at $t = T_f$

$$\underline{p}(T_f) = \underline{0} \qquad (6.28)$$
(repeated)

To find the initial conditions, let us consider Eq. 6.27. We have

$$\hat{\underline{x}}(T_o) = \hat{\underline{x}}(T_o) - \underline{x}(T_o) - (-\underline{x}(T_o)) = P_o\underline{p}(T_o);\ \text{or}$$

$$\underline{\epsilon}(T_o) - (-\underline{x}(T_o)) = P_o\underline{p}(T_o). \qquad (6.36)$$

However, $-\underline{x}(T_o)$ is the a priori error, $\underline{\epsilon}_o$, in the estimate of the initial state. From our assumptions in Chapter II, this a priori

error has zero mean and a covariance of P_o and is independent of $\underline{u}(t)$ and $\underline{w}(t)$. Consequently, we may write Eq. 6.36 as

$$\underline{\epsilon}(T_o) - \underline{\epsilon}_o = P_o \underline{p}(T_o). \tag{6.37}$$

Let us now find a differential equation for the realizable filter error. We define this error to be

$$\underline{\epsilon}_r(t) = \hat{\underline{x}}_r(t) - \underline{x}(t), \qquad T_o \leq t. \tag{6.38}$$

If we substitute Eq. 6.1 into Eq. 6.29 and then subtract 2.1, we find

$$\frac{d}{dt}\underline{\epsilon}_r(t) = (F(t) - \Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)C(t))\underline{\epsilon}_r(t) - G(t)\underline{u}(t)$$

$$- \Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)\underline{w}(t), \quad T_o \leq t. \tag{6.39}$$

We need to specify the initial condition of Eq. 6.39. Since we have assumed zero a priori mean for $\underline{x}(T_o)$, we have

$$\hat{\underline{x}}_r(T_o) = \underline{0}. \tag{6.40}$$

Therefore, from Eq. 6.38 the realizable filter error at $t = T_o$ equals the a priori error

$$\underline{\epsilon}_r(T_o) = -\underline{x}(T_o) = \underline{\epsilon}_o \tag{6.41}$$

Finally, we can relate the smoother error to the realizable filter error. By using Eq. 6.32, we can write

$$\underline{\hat{x}}(t) - \underline{\hat{x}}_r(t) = (\underline{\hat{x}}(t) - \underline{x}(t)) - (\underline{\hat{x}}_r(t) - \underline{x}(t)) =$$

$$\underline{\epsilon}(t) - \underline{\epsilon}_r(t) = \Sigma(\tfrac{t}{t})\,\underline{p}(t), \qquad T_o \le t \le T_f. \tag{6.42}$$

(Since $\underline{\epsilon}_r(T_o) = \underline{\epsilon}_o$, Eq. 6.37 is obviously a special case of Eq. 6.42 evaluated at $t = T_o$.)

Let us now summarize the important equations that we shall use in our analysis.

1. Writing the smoothing equations 6.34 and 6.35 in augmented vector form, we have

$$\frac{d}{dt}\begin{bmatrix}\underline{\epsilon}(t)\\ \underline{p}(t)\end{bmatrix} = \begin{bmatrix}F(t) & G(t)QG^T(t)\\ C^T(t)R^{-1}(t)C(t) & -F^T(t)\end{bmatrix}\begin{bmatrix}\underline{\epsilon}(t)\\ \underline{p}(t)\end{bmatrix}$$

$$-\begin{bmatrix}G(t)\underline{u}(t)\\ C^T(t)R^{-1}(t)\underline{w}(t)\end{bmatrix} =$$

$$W(t)\begin{bmatrix}\underline{\epsilon}(t)\\ \underline{p}(t)\end{bmatrix}\begin{bmatrix}G(t)\underline{u}(t)\\ C^T(t)R^{-1}(t)\underline{w}(t)\end{bmatrix}, \quad T_o \le t \le T_f. \tag{6.43}$$

where we can identify the coefficient matrix as it has been defined previously by Eq. 4. 16 . The boundary conditions for this equation are

$$\underline{\epsilon}(T_o) - \underline{\epsilon}_o = P_o \underline{p}(T_o) \qquad (6.37)$$
(repeated)

$$\underline{p}(T_f) = \underline{0} \qquad (6.28)$$
(repeated)

where $\underline{\epsilon}_o$ is the a priori error of the estimate of the initial state.

2. The realizable filter error is the solution to Eq. 6.39

$$\frac{d}{dt}\,\underline{\epsilon}_r(t) = (F(t) - \Sigma(t)t)C^T(t)R^{-1}(t)C(t))\underline{\epsilon}_r(t) - G(t)\underline{u}(t)$$
$$-\Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)\underline{w}(t), \quad T_o < t, \qquad (6.39)$$
(repeated)

where the initial condition is

$$\underline{\epsilon}_r(T_o) = \underline{\epsilon}_o, \qquad (6.41)$$
(repeated)

and $\Sigma(\frac{t}{t})$ is the covariance of $\underline{\epsilon}_r(t)$.

3. The smoothing error and the realizable filter error are related by

$$\underline{\epsilon}(t) - \underline{\epsilon}_r(t) = \Sigma(\tfrac{t}{t})\underline{p}(t) \qquad (6.42)$$
(repeated)

In order to derive the differential equations for the performance of the smoother, we shall need to evaluate three

expectations. We shall simply state the results of the evaluation of these expectations. The interested reader can follow the derivations which follow these statements.

Expectation 1

$$E[\underline{\epsilon}(t)\underline{p}^T(\tau)] = 0, \qquad T_o \leq t, \ \tau \leq T_f. \tag{6.44}$$

Proof: By classical arguments, the received signal $\underline{r}(\tau)$ can be shown to be uncorrelated with the error $\underline{\epsilon}(t)$ for all $t, \tau \epsilon [T_o, T_f]$. As given by Eqs. 6.25 through 6.28, the random process $\underline{p}(\tau)$ is the result of a linear operation upon the received signal; consequently, it is also uncorrelated with $\underline{\epsilon}(t)$ for all $t, \tau \epsilon [T_o, T_f]$.

Expectation 2

$$A(t) = E\left[\begin{Bmatrix}\underline{\epsilon}(t)\\ \hline \underline{p}(t)\end{Bmatrix} \underline{\epsilon}_r^T(t)\right] = \Psi(t, T_f)\begin{bmatrix}\phi(T_f, t)\Sigma(\frac{t}{t})\\ \hline 0\end{bmatrix}, \quad T_o \leq t \leq T_f, \tag{6.45}$$

where $\Psi(t, T_f)$ is the transition matrix for $W(t)$ and $\phi(T_f, t)$ is the realizable filter transition matrix.

Proof: In order to evaluate this expectation, we need to write the solution of Eq. 6.43 as an integral operation. From Eq. 6.31, we have

$$\underline{\epsilon}(t)\Big|_{t = T_f} = \underline{\epsilon}_r(T_f) \tag{6.46}$$

With Eq. 6.28 this provides a complete set of boundary condition for Eq. 6.43 at the endpoint of the interval. Therefore, we can write

$$\begin{bmatrix} \underline{\epsilon}(t) \\ \hline \underline{p}(t) \end{bmatrix} = \Psi(t, T_f) \begin{bmatrix} \underline{\epsilon}_r(T_f) \\ \hline \underline{0} \end{bmatrix} + \int_t^{T_f} \Psi(t,\tau) \begin{bmatrix} G(\tau)\underline{u}(\tau) \\ \hline C^T(\tau)R^{-1}(\tau)\underline{w}(\tau) \end{bmatrix} d\tau,$$

$$T_o \leq t \leq T_f \qquad (6.47)$$

Let us now substitute this equation into the above expectation. Since the realizable error $\underline{\epsilon}_r(t)$ is independent of $\underline{u}(\tau)$ and $\underline{w}(\tau)$ for $\tau > t$, one of the terms is identically zero. Therefore, we have

$$A(t) = E\left[\left\{ \begin{matrix} \underline{\epsilon}(t) \\ \hline \underline{p}(t) \end{matrix} \right\} \underline{\epsilon}_r^T(t) \right] = \Psi(t, T_f) \quad E[\underline{\epsilon}_r(T_f)\underline{\epsilon}_r^T(t)],$$

$$T_o \leq t \leq T_f. \qquad (6.48)$$

The expectation in this equation can be evaluated by writing $\underline{\epsilon}_r(T_f)$ as an integral operation,

$$\underline{\epsilon}_r(T_f) = \phi(T_f, t)\underline{\epsilon}_r(t) - \int_t^{T_f} \phi(T_f, \tau)[G(\tau)\underline{u}(\tau) + \Sigma(\tfrac{\tau}{\tau})C^T(\tau)R^{-1}(\tau)\underline{w}(\tau)]\, d\tau,$$

$$T_o \leq t \leq T_f. \qquad (6.49)$$

When we substitute Eq. 6.49 in Eq. 6.48, we find that the second term vanishes since $\underline{\epsilon}_r(t)$ is again independent of $\underline{u}(\tau)$ and $\underline{w}(\tau)$ for

$\tau \geq t$. Therefore, we finally obtain the desired result.

$$A(t) = E\left\{\begin{bmatrix}\underline{\epsilon}(t)\\ \hline \underline{p}(t)\end{bmatrix}\underline{\epsilon}_r^T(t)\right\} = \Psi(t, T_f)\begin{bmatrix}\phi(T_f, t)\Sigma(\frac{t}{t})\\ \hline 0\end{bmatrix}, \quad T_o \leq t \leq T_f. \tag{6.45}$$

(repeated)

Expectation 3

$$B(t) = E\left\{\begin{bmatrix}\underline{\epsilon}(t)\\ \hline \underline{p}(t)\end{bmatrix}\begin{bmatrix}G(t)\underline{u}(t)\\ \hline C^T(t)R^{-1}(t)\underline{w}(t)\end{bmatrix}^T\right\} = \frac{1}{2}\left[\begin{array}{c|c} G(t)QG^T(t) & \\ \hline 0 & C^T(t)R^{-1}(t)C(t)\end{array}\right]$$

$$+ \Psi(t, T_f)\left[\begin{array}{c|c} \phi(T_f, t)G(t)QG^T(t) & \phi(T_f, t)\Sigma(\frac{t}{t})C^T(t)R^{-1}(t)C(t) \\ \hline 0 & 0\end{array}\right], \quad T_o \leq t \leq T_f. \tag{6.50}$$

Let us substitute Eq. 6.44 into the above. This yields

$$B(t) = E\left\{\left(\Psi(t, T_f)\begin{bmatrix}\underline{\epsilon}_r(T_f)\\ \\ \underline{0}\end{bmatrix} + \int_t^{T_f}\Psi(t, \tau)\begin{bmatrix}G(\tau)\underline{u}(\tau)\\ \hline C^T(\tau)R^{-1}(\tau)\underline{w}(\tau)\end{bmatrix}d\tau\right)\right.$$

$$\left.\times\begin{bmatrix}G(t)\underline{u}(t)\\ \hline C(t)R^{-1}(t)\underline{w}(t)\end{bmatrix}^T\right\} \quad T_o \leq t \leq T_f. \tag{6.51}$$

We can relate $\underline{\epsilon}_r(T_f)$ to $\underline{u}(t)$ and $\underline{w}(t)$, and we can also evaluate the second expectation. We find

$$B(t)=\Psi(t,T_f)E\left\{\left[\begin{array}{c}\phi(T_f,T_o)\underline{\epsilon}_r(T_o)\\ \hline +\int_{T_o}^{T_f}\phi(T_f,\tau)[G(\tau)\underline{u}(\tau)+\Sigma(\frac{\tau}{\tau})C^T(\tau)R^{-1}(\tau)\underline{w}(\tau)]d\tau\\ \hline \underline{0}\end{array}\right]\right.$$

$$\left.\times\left[\begin{array}{c}G(t)\underline{u}^T(t)\\ \hline C^T(t)R^{-1}(t)w(t)\end{array}\right]^T\right\}$$

$$+\int_t^{T_f}\Psi(t,\tau)\left[\begin{array}{c|c}G(\tau)QG^T(t) & 0\\ \hline 0 & C^T(t)R^{-1}(t)C(\tau)\end{array}\right]\delta(t-\tau)d\tau,$$

$$T_o\le t\le T_f. \qquad (6.52)$$

In Eq. 6.52 we note that $\underline{\epsilon}_r(T_o)$ is independent of $\underline{u}(t)$ and $\underline{w}(t)$; consequently, we can easily evaluate the first term. In order to evaluate the second term we need to split the impulse since it is evaluated at $\tau = t$, the endpoint of the integration region. This yields the desired result of Eq. 6.50.

With these expectations we can now derive the differential equations for the covariance of error of the optimal smoother. Let us proceed by differentiating the definition of the covariance. We

want to evaluate

$$\frac{d}{dt}\Sigma(\frac{t}{T_f}) = \frac{d}{dt}E[\underline{\epsilon}(t)\underline{\epsilon}^T(t)] = E[\dot{\underline{\epsilon}}(t)\underline{\epsilon}^T(t)+\underline{\epsilon}(t)\dot{\underline{\epsilon}}^T(t)]^*, \quad T_o \leq t \leq T_f. \tag{6.53}$$

Since the terms in the above are transposes of each other, we shall consider only one of them. Substituting Eq. 6.34, we have

$$E[\dot{\underline{\epsilon}}(t)\underline{\epsilon}^T(t)] = F(t)E[\underline{\epsilon}(t)\underline{\epsilon}^T(t)] + G(t)QG^T(t)E[\underline{p}(t)\underline{\epsilon}^T(t)]$$

$$G(t)E[\underline{u}(t)\underline{\epsilon}^T(t)] = F(t)\Sigma(\frac{t}{T}) - G(t)QG^T(t)\phi^T(T_f,t)\Psi^T(t,T_f) -$$

$$\frac{1}{2}G(t)QG^T(t), \quad T_o \leq t \leq T_f \tag{6.54}$$

where we have used Expectations 1 and 3. Adding the transpose term we have the first differential equation for the smoother covariance.

Differential Equation 1 for $\Sigma(t/T_f)$

$$\frac{d}{dt}\Sigma(\frac{t}{T_f}) = F(t)\Sigma(\frac{t}{T_f}) + \Sigma(\frac{t}{T_f})F^T(t) - G(t)QG^T(t)$$

---

*We use the notation $\Sigma(t/T_f)$ to indicate the covariance of error at a time $t(\geq T_o)$ where we have observed the signal $r(\tau)$ over the interval $T_o \leq \tau \leq T_f$.

$$- [\Psi(t\ T_f)\phi(T_f, t)G(t)Q\,G^T(t)] - [G(t)Q\,G^T(t)\phi^T(T_f, t)\Psi^T(t, T_f)],$$

$$T_o \leq t \leq T_f,^* \qquad (6.55)$$

There are two difficulties with this expression. First, the forcing term is difficult to evaluate. Secondly when integrated backwards over long time interval it becomes unstable.

We shall now derive a second differential equation form for $\Sigma(t/T_f)$ which eliminates these difficulties.

If we solve Eq. 6.42 for $\underline{p}(t)$ we obtain

$$\underline{p}(t) = \Sigma^{-1}(\tfrac{t}{t})(\underline{\epsilon}(t) - \underline{\epsilon}_r(t)). \qquad (6.56)$$

Now we substitute for $\underline{p}(t)$ in Eq. 6.34 (this is similar to the approach of Method 3 in Chapter IV) we find

$$\frac{d\underline{\epsilon}(t)}{dt} = F(t)\epsilon(t) + G(t)Q\,G^T(t)\Sigma^{-1}(\tfrac{t}{t})(\underline{\epsilon}(t)-\underline{\epsilon}_r(t)) - G(t)\underline{v}(t) =$$

$$(F(t) + G(t)Q\,G^T(t)\Sigma^{-1}(\tfrac{t}{t}))\underline{\epsilon}(t) + G(t)Q\,G^T(t)\Sigma^{-1}(\tfrac{t}{t})\underline{\epsilon}_r(t)$$

$$-G(t)\underline{v}(t), \qquad T_o \leq t \leq T_f. \qquad (6.57)$$

Let us now substitute Eq. 6.57 in Eq. 6.53. If we use Expectations 2 and 3, we have a cancellation of terms, and we obtain the desired result

---

*A relationship of this form was first derived by Baggeroer in the 1966 WESCON Proceedings.

Differential Equation 2 for $\Sigma(t/T_f)$

$$\frac{d\Sigma(\frac{t}{T_f})}{dT} = (F(t) + G(t)Q G^T(t)\Sigma^{-1}(\frac{t}{t})\Sigma(\frac{t}{T_f}) +$$

$$\Sigma(\frac{t}{T_f})(F(t) + G(t)Q G^T(t)\Sigma^{-1}(\frac{t}{t}))^T - G(t)Q G^T(t),\ T_o \leq t \leq T_f. \tag{6.58}$$

Several comments are in order here. This formula was first derived by Rauch, Tung and Striebel using a discrete time analysis and then passing to a continuous limit.[15] The covariance of error for the realizable estimate, $\Sigma(t/t)$ enters Eq. 6.58 in two ways. Obviously, its inverse appears as part of the coefficient terms. In addition, it supplies the required boundary condition at $t = T_f$, for

$$\Sigma(\frac{t}{T_f})\Big|_{t = T_f} = \Sigma(\frac{t}{t})\Big|_{t = T_f} \tag{6.59}$$

Consequently, we can solve Eq. 6.58 backwards in time from $t = T_f$. Finally, we discussed the stability of equations with this coefficient in Chapter IV. We can reach the same conclusions here as we did there, i.e., when integrated backwards from $t = T_f$, the solution to Eq. 6.58 will in general be stable. Consequently, obtaining solutions for long time intervals does not cause any numerical difficulties when done this way.

We can derive a third differential equation for finding $\Sigma(t/T_f)$ by using the covariance of $\underline{p}(t)$, which we denote by $\Pi(t/T_f)$. First, we relate this covariance to $\Sigma(t/T_f)$. We can rewrite Eq. 6.42 in the form

$$\underline{\epsilon}(t) - \Sigma(\tfrac{t}{t})\underline{p}(t) = \underline{\epsilon}_r(t), \qquad T_o \le t \le T_f. \tag{6.60}$$

Post multiplying Eq. 6.60 by its transpose, taking the expected value of the result and using Expectation 1 gives us

$$\Sigma(\tfrac{t}{T_f}) + \Sigma(\tfrac{t}{t})\Pi(\tfrac{t}{T_f})\Sigma(\tfrac{t}{t}) = \Sigma(\tfrac{t}{t}), \quad T_o \le t \le T_f \tag{6.61}$$

By using an analysis similar to that used in deriving form 2 for $\Sigma(t/T_f)$, we can find a differential equation for $\Pi(t/T_f)$. This gives us our third form.

Differential Equation 3 for $\Sigma(t/T_f)$

$$\frac{d\Pi(\frac{t}{T_f})}{dt} = -(F(t) + \Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)C(t))^T\Pi(\tfrac{t}{T_f})$$

$$-\Pi(\tfrac{t}{T_f})(F(t)+\Sigma(\tfrac{t}{t})C^T(t)R^{-1}(t)C(t)) - C^T(t)R^{-1}(t)C(t),$$

$$T_o \le t \le T_f. \tag{6.62}$$

where $\Pi(t/T_f)$ is related to $\Sigma(t/T_f)$ by Eq. 6.61. Since $\underline{p}(T_f)$ is identically zero, the boundary condition at $t = T_f$ is obviously

$$\Pi\left(\frac{T_f}{T_f}\right) = 0. \tag{6.63}$$

Again, some comments are in order. The same issues of integrating backwards and the stability of the solution still apply as discussed earlier for form 2 still apply. In addition, if $\Sigma(t/t)$ is already available, it may be better to use this form rather than form 2 since taking an inverse matrix at all points in the interval is not required.

The first three differential equations that we have derived are not well suited for finding analytic solutions for $\Sigma(t/T_f)$. Even in the case of constant parameter systems, the matrix differential equations have time varying coefficients. We can eliminate this difficulty by considering a 2n x 2n matrix differential equation rather than an n x n one as we have considered before. To do this we need to combine forms 2 and 3 in an appropriate manner.

One can show by straightforward manipulation of Eqs. 6.58, 6.61 and 6.62 that the following matrix differential equation is satisfied.

Differential Equation 4 for $\Sigma(t/T_f)$

$$\frac{d}{dt} P\left(\frac{t}{T_f}\right) = W(t)P\left(\frac{t}{t}\right) + P\left(\frac{t}{t}\right)W^T(t) + \left[\begin{array}{c|c} G(t)Q\,G^T(t) & 0 \\ \hline 0 & (C^T(t)R^{-1}(t)C(t) \end{array}\right],$$

$$T_o \le t \le T_f, \tag{6.64a}$$

where W(t) is defined by Eq. 4.15

$$\text{and } P(\tfrac{t}{T_f}) \triangleq \left[\begin{array}{c|c} \Sigma(\frac{t}{T_f}) & -\Pi(\frac{t}{T_f})\Sigma(\frac{t}{t}) \\ \hline -\Sigma(\frac{t}{t})\,\Pi(\frac{t}{T_f}) & -\Pi(\frac{t}{T_f}) \end{array}\right] \tag{6.64b}$$

We can specify the boundary condition at $t = T_f$ by using Eqs. 6.59) and 6.63)

$$P(\tfrac{T_f}{T_f}) = \left[\begin{array}{c|c} \Sigma(\frac{T_f}{T_f}) & 0 \\ \hline 0 & 0 \end{array}\right] \tag{6.65}$$

where $\Sigma(T_f/T_f)$ is the realizable filter error at T_f.

Consequently we can solve Eq. 6.64a backwards over the interval using this boundary condition. This is analogous to method 2 that we developed in Chapter IV.

If we extend the concept developed there, it is easy to find an integral representation for the solution to Eq. 6.64
We have,[30]

$$P\left(\frac{t}{T_f}\right) = \Psi(t, T_f) \begin{bmatrix} \Sigma\left(\frac{T_f}{T_f}\right) & 0 \\ \hline 0 & 0 \end{bmatrix} \Psi^T(t, T_f)$$

$$- \int_t^{T_f} \Psi(t, \tau) \begin{bmatrix} G(\tau) Q G^T(\tau) & 0 \\ \hline 0 & C^T(\tau) R^{-1}(\tau) C(\tau) \end{bmatrix} \Psi^T(t, \tau)\, d\tau,$$

$$T_o \leq t \leq T_f. \qquad (6.66)$$

We should point out that solving for $\Sigma(t/T_f)$ using this form does have certain advantages when the system parameters are constant. In this case, the coefficient matrix W is a constant matrix, which allows us to see the matrix exponential. This is certainly an analytic advantage. However, we should remember that we have a larger dimensional set of equations with which to deal. Finally, we observe that this form is not well suited for finding numerical solutions of $\Sigma(t/T_f)$. Since it involves W(t) as a coefficient matrix, it introduces virtually the same stability problems that we had for method 2 in Chapter IV.

This completes our discussion of the differential equation forms for $\Sigma(t/T_f)$. Before proceeding, we shall discuss the merits of each. Form 1 was the starting point of our derivation. Although it was relatively simple in appearance, the forcing term for it was difficult to evaluate. In addition, it introduced stability problems. Forms 2 and 3 were well suited for numerical procedures, since they

were stable when integrated backwards of the interval. However, they were not well suited for analytic procedures because they involved time varying coefficients. Form 4 was well suited for analytic methods, since the coefficient matrix was a constant for systems with constant parameters. It, however, also introduced stability problems when used numerically over long time intervals.

Several errors have been made in the literature regarding the covariance of $\underline{p}(t)$, $\Pi(t/T_f)$.[14, 15] If we apply the results of these papers, we can quickly show that they imply

$$\Pi\left(\frac{T_f - dt}{T_f}\right) = -C^T(T_f)R^{-1}(T_f)C(T_f)dt \qquad (6.67)$$

which is clearly impossible for a covariance. We shall now correct these errors.

The differential equation of 6.64, form 4, for $\Sigma(t/T_f)$ was first derived by Bryson and Frazier.[14] However, they erroneously interpreted the partitions of P(t). They state that the lower right partition should be $\Pi(t/T_f)$; we have seen that it is $-\Pi(t, T_f)$ which agrees with what Eq. 6.67 would indicate. They also interpret the off diagonal partitions as $E[\underline{\epsilon}(t)\underline{p}^T(t)]$. From expectation 1, we know that this is identically zero. These partitions are actually $\Pi(t/T_f)\Sigma(t/t)$ as we have indicated. Their upper left partition for $P(t/T_f)$ however is $\Sigma(t/T_f)$. Consequently, if we naively apply their results to find $\Sigma(t/T_f)$, we would obtain correct results. However, their incorrect assertions regarding the covariance of $\underline{p}(t)$ leaves their derivation rather suspect.

Rauch, Tung and Striebel state that the error $\underline{\epsilon}(t)$ is correlated with $\underline{p}(t)$ which is incorrect.[15] They also have not related $\Sigma(t/T_f)$ and $\Pi(t/T_f)$ correctly, as indicated by Eq. 6.63. Their equation for $\Sigma(t/T_f)$ is correct as we have discussed previously form 2.

Equation 6.58 also provides an interesting relationship between the ralizable and unrealizable errors when calculated using the classical Wiener approach. This corresponds to the case when the system parameters are constant and the time interval $[T_o, T_f]$ is large. If we are somewhere in the "middle" of the interval

$$\frac{d\Sigma(\frac{t}{T_f})}{dt} \simeq 0, \qquad T_o \ll t \ll T_f. \tag{6.68}$$

Denoting the realizable Wiener error by Σ_∞ and the unrealizable error by $\Sigma(t/\infty)$, we obtain

$$(F+GQG^T\Sigma_\infty^{-1})\Sigma(\tfrac{t}{\infty}) + \Sigma(\tfrac{t}{\infty})(F+GQG^T\Sigma_\infty^{-1})^T - GQG^T = 0 \tag{6.69}$$

This equation has an interesting implication. $\Sigma(t/\infty)$ is easy to determine since it is the unrealizable covariance; however, Σ_∞^{-1} is usually difficult to determine. Eq. 6.69) provides a <u>linear</u> matrix equation for finding Σ_∞ (really Σ_∞^{-1}). Since there are several ways of solving this linear equation, Σ_∞ follows directly. Notice the linearity of the equation eliminates the ambiguity of solutions which arises by setting the Ricatti equation to steady state.

This completes our discussion of the performance of

the optimal smoother. Let us now illustrate it with several examples.

C. Examples of the Optimal Smoother Performance

In this last section we analyzed the performance, or covariance of error, of the optimal smoother. In this section we shall illustrate this performance by considering several examples. First, we shall work two examples for first order systems analytically. We shall apply Eqs. 6.64 and 6.65 to do this. Then we shall consider the analysis of a second order system by numerically integrating Eq. 6.62 and then applying Eq. 6.61.

Example 6.1 - Covariance of Error for a Wiener Process

In this first example we shall find the covariance of error for a Wiener Process that is received in the presence of additive white noise. The system for generating this process is described in Eqs. 3.35 and 3.36. We repeat the system matrices here for convenience

$$F = 0 \qquad C = \mu$$
$$G = 1 \qquad P_o = 0 \qquad (3.36)$$
$$Q = 1 \qquad \text{(repeated)}$$

In addition we assume that the spectral height of the additive white noise is σ and the observation interval is $[0, T]$, i.e., $T_o = 0$ and $T_f = T$. Fortunately, we have available many of the required results from previous examples which concern the process. As a result, we can find the solutions rather quickly.

The first step is to find the matrix W(t) and its associated transition matrix $\Psi(t,\tau)$. We did this in Chapter IV and the results are indicated by Eqs. 4.51 and 4.56 . (We need to substitute $t-\tau$ for t as the system has constant parameters.) Next we need to find $\Sigma(T/T)$. We do this by using the partitions of $\Psi(t,\tau)$ as indicated by Eq. 4.31 . This yields

$$\Sigma\left(\frac{T}{T}\right) = \left[\frac{\mu^2}{\sigma}\right]^{-\frac{1}{2}} \tanh\left(\left[\frac{\mu^2}{\sigma}\right]^{-\frac{1}{2}} T\right). \tag{6.70}$$

Consequently, Eq. 6.66 becomes

$$P\left(\frac{t}{T}\right) = \Psi(t,T)\left[\begin{array}{c|c} \left[\frac{\mu^2}{\sigma}\right]^{-\frac{1}{2}} \tanh\left(\left[\frac{\mu^2}{\sigma}\right]^{-\frac{1}{2}} T\right) & 0 \\ \hline 0 & 0 \end{array}\right]\Psi^T(t,T)$$

$$\int_t^T \Psi(t,\tau)\left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \frac{\mu^2}{\sigma} \end{array}\right]\Psi^T(t,\tau)d\tau, \quad 0<t<T. \tag{6.71}$$

where $\Psi(t,\tau)$ is as defined by Eq. 4.56 as indicated above. Separating our the upper, left partition for $\Sigma(t/T)$, we find after some straightforward manipulation

$$\Sigma(t/T) = \left[\frac{\mu^2}{\sigma}\right]^{\frac{-1}{2}} \frac{\cosh\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}}(T-t) \cdot \sinh\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} t}{\cosh\left[\frac{\mu^2}{\sigma}\right]^{\frac{1}{2}} T}, \quad 0 \le t \le T. \tag{6.72}$$

Example 6.2 - Covariance of Error for a One Pole Process

Let us now consider the performance for a random process generated by a system with a pole at -k. The generation of this process is described by Eqs. 2.16 and 2.17. The state matrices are repeated here for convenience

$$F = -k \qquad C = 1$$
$$G = 1 \qquad P_o = S \tag{2.17 (repeated)}$$
$$Q = 2kS$$

Let us consider a slight variation of this representation. Instead of choosing $P_o = P$ such that the process generated is stationary, let it remain a free variable. In addition, we chose the level of the observation noise to be σ and the observation interval to be $[0, T]$. Again, we have many results available from previous examples.

The matrix W(t) and its associated transition matrix $\Psi(t, \tau)$ are given by Eqs. 4.62 and 4.68, respectively. (Again, we need to substitute $(t-\tau)$ for t.) Next, we need to find $\Sigma(T/T)$. By using Eq. 4.31 the realizable filter covariance of error $\Sigma(T/T)$ for an arbitrary covariance of the initial state, P_o, is given by

$$\Sigma\left(\frac{T}{T}\right) = S\,\frac{\left[\frac{P_o}{S}+\frac{1}{\lambda}\left(2-\frac{P_o}{S}\right)\right]e^{k\lambda T} + \left[\frac{P_o}{S}-\frac{1}{\lambda}\left(2-\frac{P_o}{S}\right)\right]e^{-k\lambda T}}{\left[1+\frac{1}{\lambda}+\frac{P_o}{S}\,\frac{\lambda^2-1}{2}\right]e^{k\lambda T} + \left[1-\frac{1}{\lambda}-\frac{P_o}{S}\,\frac{\lambda^2-1}{2}\right]e^{-k\lambda T}} \tag{6.73 a}$$

where

$$\lambda = \left(1+\frac{4S}{kN_o}\right)^{\frac{1}{2}} \tag{6.73b}$$

For this system, Eq. 6.66 becomes

$$P\left(\frac{t}{T}\right) = \Psi(t,T)\begin{bmatrix}\Sigma\left(\frac{T}{T}\right) & 0\\ 0 & 0\end{bmatrix}\Psi^T(t,T) -$$

$$\int_t^T \Psi(t,\tau)\begin{bmatrix}2kS & 0\\ 0 & \frac{1}{\sigma}\end{bmatrix}\Psi^T(t,\tau)\,d\tau, \qquad 0 \le t \le T. \tag{6.74}$$

where $\Sigma(T/T)$ is defined above by Eq. 6.73a. Taking the upper left portion for $\Sigma(t/T)$ we find after some straight forward manipulation

$$\Sigma\left(\frac{t}{T}\right) = \left[\cosh\left[k\lambda(T-t)\right] + \frac{\sinh\left[k\lambda(T-t)\right]}{\lambda}\right] \times$$

$$\left[\Sigma\left(\frac{T}{T}\right)\cosh\left[k\lambda(T-t)\right] + \left(\Sigma\left(\frac{T}{T}\right)-2S\right)\frac{\sinh\left[k\lambda(T-t)\right]}{\lambda}\right], \quad 0 \le t \le T. \tag{6.75}$$

In order to illustrate this result let us consider three choices of P_o.

Case a - $P_o = S$ (Stationary Process)

Here we have chosen P_o such that we are estimating a stationary process. In this case

$$\Sigma\left(\frac{T}{T}\right) = 2S \left\{ \frac{(\lambda+1)e^{k\lambda T} + (\lambda-1)e^{-k\lambda T}}{(\lambda+1)^2 e^{k\lambda T} - (\lambda-1)^2 e^{-k\lambda T}} \right\}. \qquad (6.76)$$

If we substitute this equation into Eq. 6.75 , we can show that

$$\Sigma\left(\frac{t}{T}\right) = \Sigma\left(\frac{T}{T}\right)\left\{ \frac{(\cosh[k\lambda(T-t)] + \frac{1}{\lambda}\sinh[k\lambda(T-t)])(\cosh[k\lambda t] + \frac{1}{\lambda}\sinh[k\lambda t])}{\cosh[k\lambda T] + \frac{1}{\lambda}\sinh[k\lambda T]} \right\},$$

$$0 < t < T. \qquad (6.77)$$

Therefore, we see that the covariance of error is symmetric with respect to the midpoint of the interval. We can easily show that they approach the large time results, i.e. the Wiener filtering results, quite easily. In Fig. 6.1 we have plotted Eq. 6.77 where we have chosen k = 1, σ = 1, S = 1 and T = 2.

Case b - $P_o = 2S/\lambda+1$ (Steady State Realizable Filtering Error)

In this case we have chosen the initial covariance such that we do not gain any improvement by realizable filtering, i.e.

$$\Sigma\left(\frac{T}{T}\right) = P_o = \frac{2S}{\lambda + 1} \qquad (6.78)$$

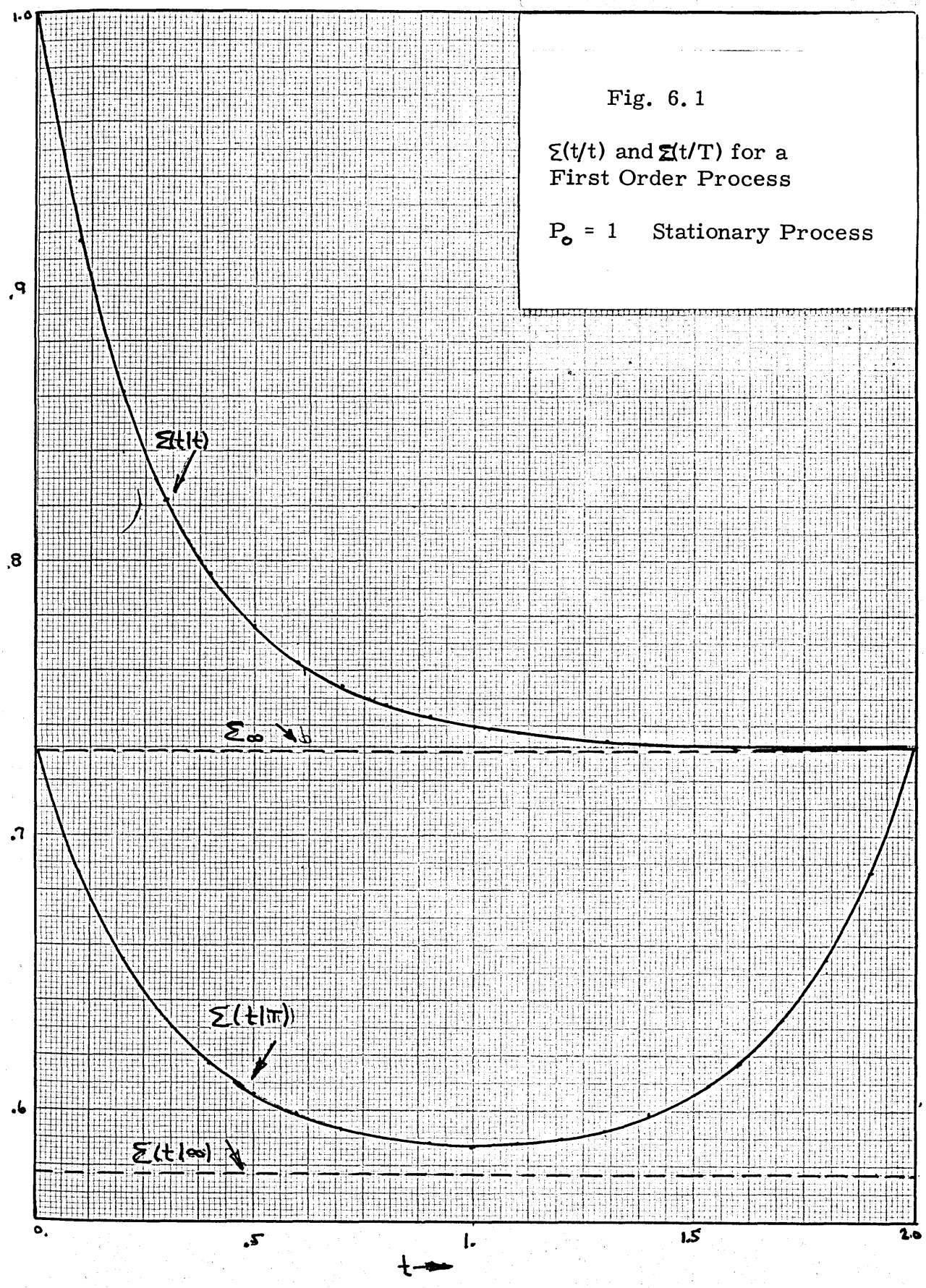


Fig. 6.1

$\Sigma(t/t)$ and $\Sigma(t/T)$ for a First Order Process

$P_o = 1$ Stationary Process

for all T. Substituting this equation into Eq. 6.7 we find

$$\Sigma(\frac{t}{T}) = \frac{S}{\lambda} + S\left(\frac{\lambda - 1}{\lambda + 1}\right) e^{-2k(T-t)} =$$

$$\frac{2S}{\lambda + 1}\left\{\frac{1}{2}(1+\frac{1}{\lambda}) + \frac{1}{2}(1-\frac{1}{\lambda})e^{-2k(T-t)}\right\}, \quad 0 \leq t \leq T. \tag{6.79}$$

Therefore, we see that the performance approaches a constant as we move backwards from the endpoint. This constant is the unrealizable filter error as we could calculate from the classical theory. The behavior near the endpoint, i.e., near T, reflects the gain in performance which the realizable filter could attain by allowing some delay before making its estimate. (We shall develop this concept in detail in the next section.) In Fig. 6.2 we have plotted Eq. 6.79 for the choice of parameters in the previous figure.

Case c - $\underline{P}_o = \underline{0}$ (Known Initial State)

Let us consider the case when we know the initial state exactly, i.e., $P_o = 0$. This case reflects the extent of correlation time between estimates of the state of the system at different times. For this choice of P_o we have

$$\Sigma(\frac{T}{T}) = \frac{2S \sinh[k\lambda T]}{\lambda \cosh[k\lambda T] + \sinh[k\lambda T]} \tag{6.80}$$

When we substitute this into Eq. 6.75 we find

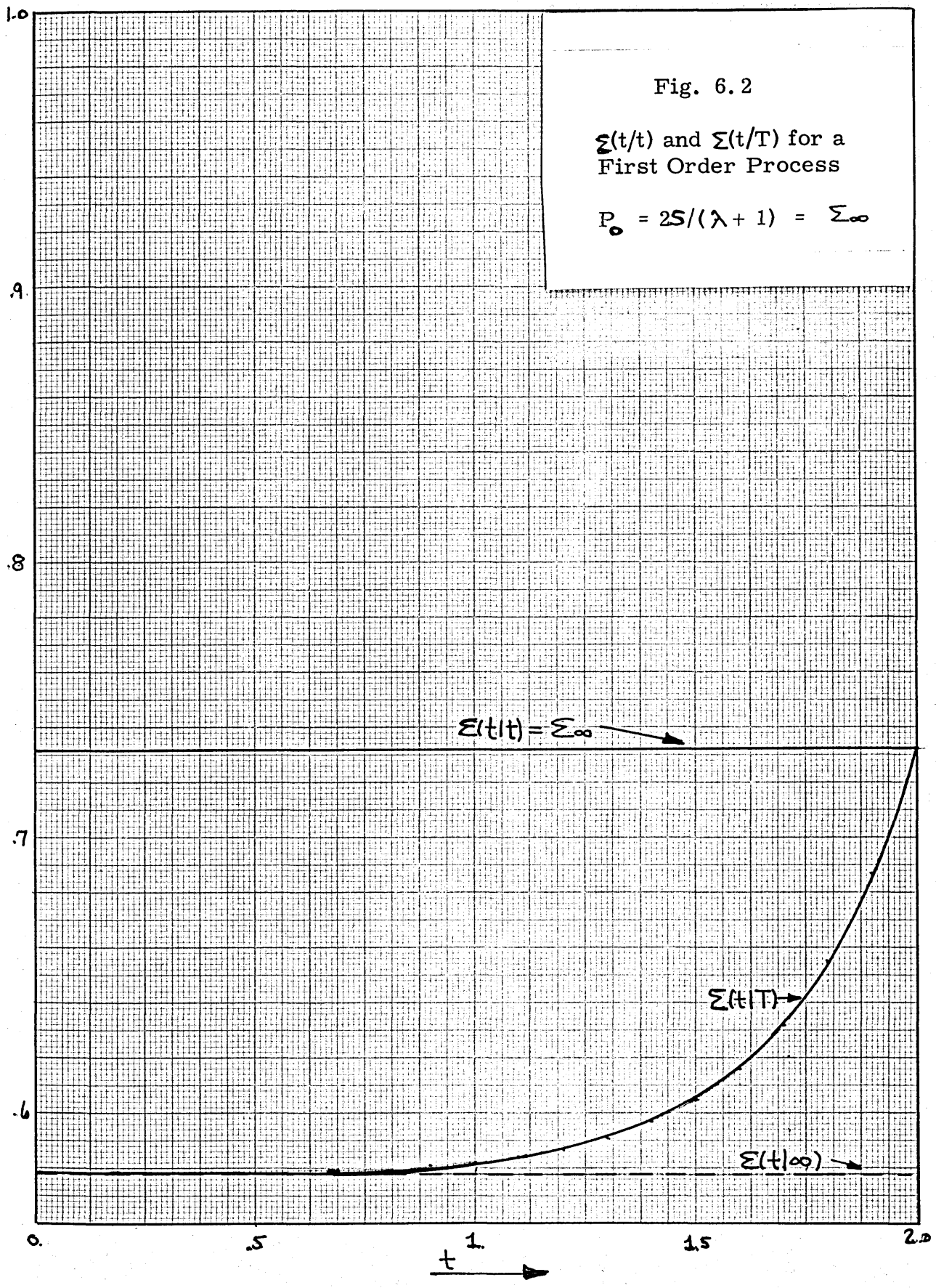


Fig. 6.2

$\Sigma(t/t)$ and $\Sigma(t/T)$ for a First Order Process

$P_o = 2S/(\lambda + 1) = \Sigma_\infty$

$$\Sigma(\frac{t}{T}) = \frac{2S}{\lambda}\left(\frac{\sinh[k\lambda t](\cosh[k\lambda(T-t)] + \frac{1}{\lambda}\sinh[k\lambda(T-t)])}{\cosh(k\lambda T) + \frac{1}{\lambda}\sinh[k\lambda T]}\right), \quad 0 < t < T. \tag{6.81}$$

In Fig. 5.3 we have plotted $\Sigma(t/T)$ for the same choice of parameters. We can see that after .6 secs. where it approaches $\Sigma(t|\infty)$, the information regarding the initial state is virtually useless in making our estimate. (We should point out that we did not take the observation interval quite long enough for this case. There is a plateau in middle of the interval if the interval is long enough.)

Example 3 - Covariance of Error for a Two Pole Process

We have analyzed about all the systems that one can do analytically in a reasonable amount of time. In order to study higher dimensioned systems, we use numerical methods. To do this we shall numerically integrate Eq. 6.62 to determine $\Pi(t/T_f)$. Given this function we find $\Sigma(t/T_f)$ by using 6.61. First though, let us consider the system that we wish to study.

We shall assume that we want to estimate the stationary process y(t) which is described by Eqs. 2.18 and 2.19. It has a covariance function and a spectrum as illustrated in Figs. 2.19a and 2.19b, respectively.

Instead of performing the type of analysis done in the previous example, let us consider a variate of a pre-emphasis probelm.

Since we have a two state system, let us see if we can improve our performance by including the second state, or the

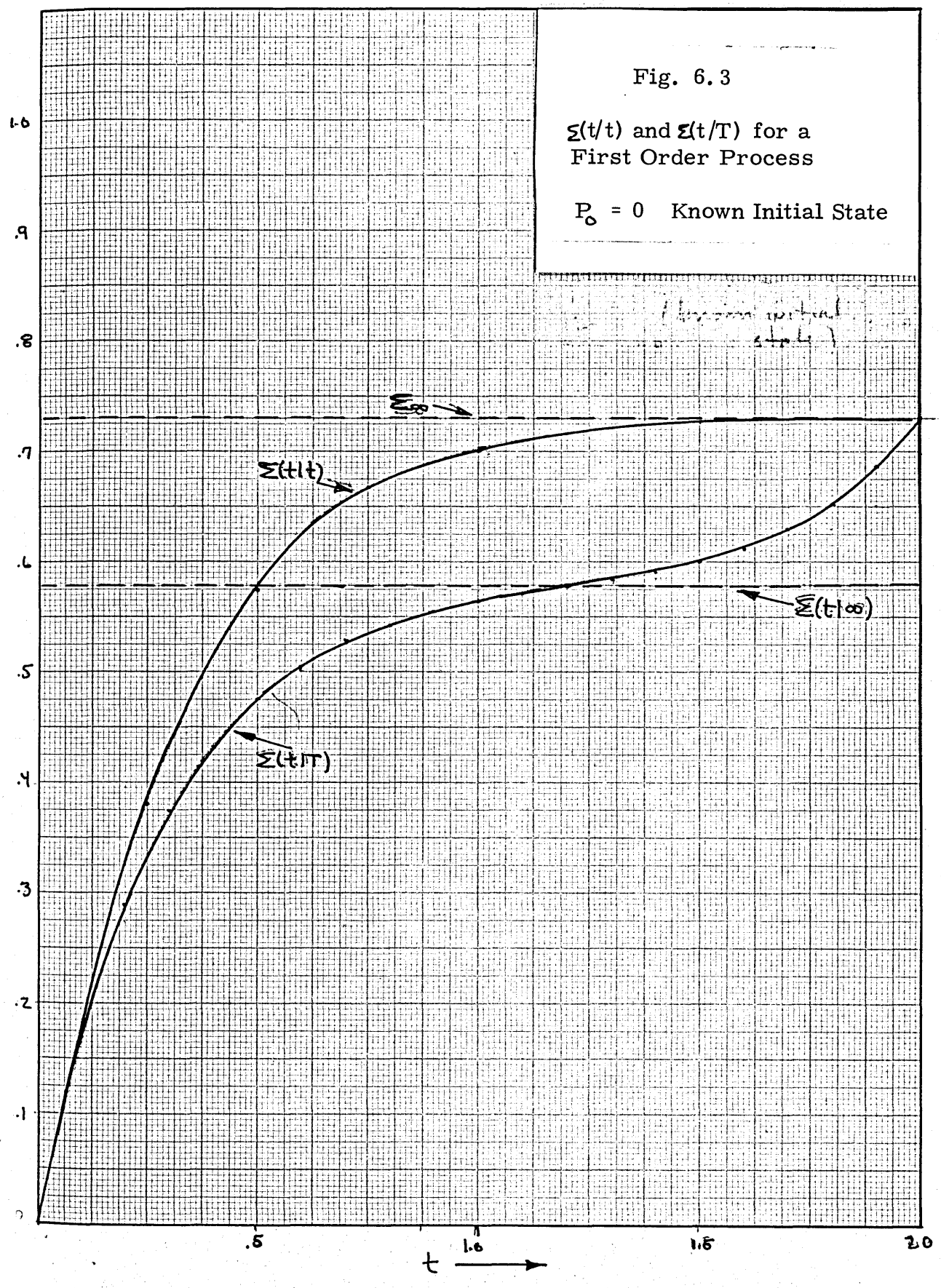


Fig. 6.3
Σ(t/t) and Σ(t/T) for a First Order Process
P₀ = 0 Known Initial State
Σ∞
Σ(t|t)
Σ(t|∞)
Σ(t|T)
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
.5
1.0
1.5
2.0
t

derivative of the message, in our transmitted signal. In this context, we then have that the output of the system is

$$y(t) = \alpha_1 x_1(t) + \alpha_2 x_2(t) = \alpha_1 x_1(t) + \alpha_2 \frac{dx_1(t)}{dt} \tag{6.82}$$

where we desire to estimate $x_1(t)$ as our message. It would not be a fair comparison to simply add the second state, since this would increase the transmitted power. Let us, therefore, constrain the power to be fixed to its original level of 4. We have

$$E[y^2(t)] = \alpha_1^2 E[x_1^2(t)] + \alpha_2^2 E[x_2^2(t)] =$$

$$4\alpha_1^2 + 40\alpha_2^2 = 4 \tag{6.83}$$

$(E[x_1(t) \frac{dx_1(t)}{dt}] = 0$ for a stationary differentiable random process). Therefore, we shall vary α_1 and α_2 within the constraint of Eq. 6.83 to see if it can improve our performance. We shall also assume that the additive white noise level is 1 and the interval length is 2.

In summary, the state matrices for our system are

$$F = \begin{bmatrix} 0 & 1 \\ -10 & -2 \end{bmatrix} \qquad C = [\alpha_1 \; ! \; \alpha_2]$$

$$G = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad R = 1$$

$$Q = 160 \qquad P_o = \begin{bmatrix} 4 & 0 \\ 0 & 40 \end{bmatrix} \qquad \text{(6.84 a-e)}$$

with α_1 and α_2 constrained to be on the ellipse

$$\alpha_1^2 + 10\,\alpha_2^2 = 1 \qquad \text{(6.83) (repeated)}$$

In addition we choose the interval length to be 2, or $T_o = 0$ and $T_f = T = 2$.

If we proceed with our analysis by numerically integrating Eq. 6.63 to find $\Sigma(t/T)$ for various value α_1 and α_2 that satisfy Eq. 6.83 , we generate the curves of Fig. 6.4. Here we have plotted $\Sigma_{11}(t/T)$ (normalized), the covariance of the message.

We can make several observations regarding these curves. Although we are estimating a stationary process, the curves are not symmetrical about the midpoint of the interval as in Fig 6.1 unless the observed signal contains only one of the states, i.e., $\alpha_1 = 0$, or $\alpha_2 = 0$. Secondly, we have plotted the curves only for positive values of α_2; if α_2 is negative with respect to α_1, we generate curves which are exactly inverted in time. At the present time, we do not have a good physical interpretation for this observation.

Let us also consider the improvement in performance. We have summarized this in Fig. 6.5. In this figure we have plotted $\Sigma_{11}(T \mid T)$(normalized), the realizable estimate at the endpoint of the interval, and $\Sigma_{11}(T/2 \mid T)$, the smoother performance of the midpoint of the interval. These points are very close to their asymptotic

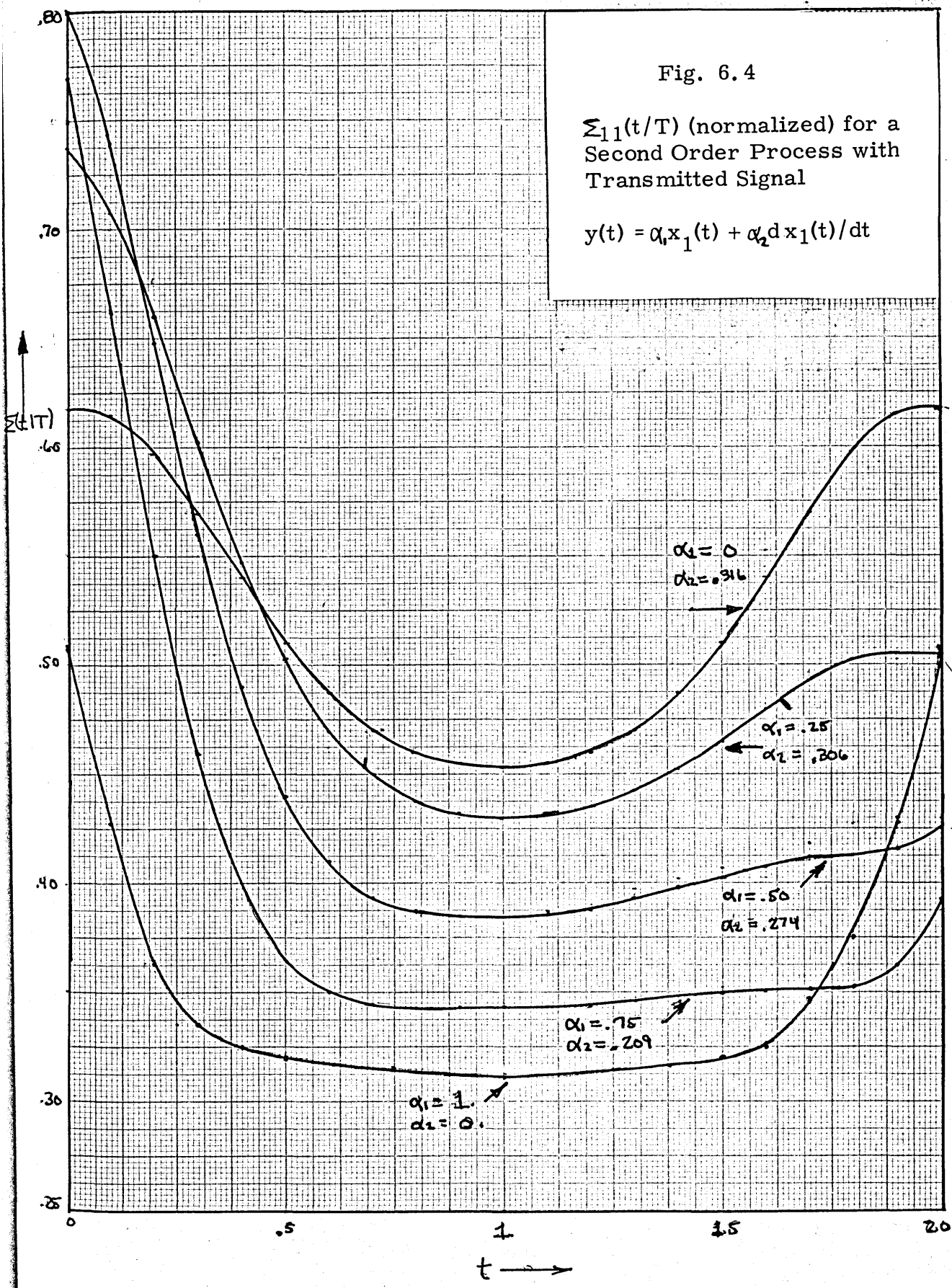


Fig. 6.4
$\Sigma_{11}(t/T)$ (normalized) for a Second Order Process with Transmitted Signal
$y(t) = \alpha_1 x_1(t) + \alpha_2 d x_1(t)/dt$
$\Sigma(t/T)$
.80
.70
.60
.50
.40
.30
.25
0
.5
1
1.5
2.0
t
$\alpha_1 = 0$
$\alpha_2 = .316$
$\alpha_1 = .25$
$\alpha_2 = .306$
$\alpha_1 = .50$
$\alpha_2 = .274$
$\alpha_1 = .75$
$\alpha_2 = .209$
$\alpha_1 = 1.$
$\alpha_2 = 0.$

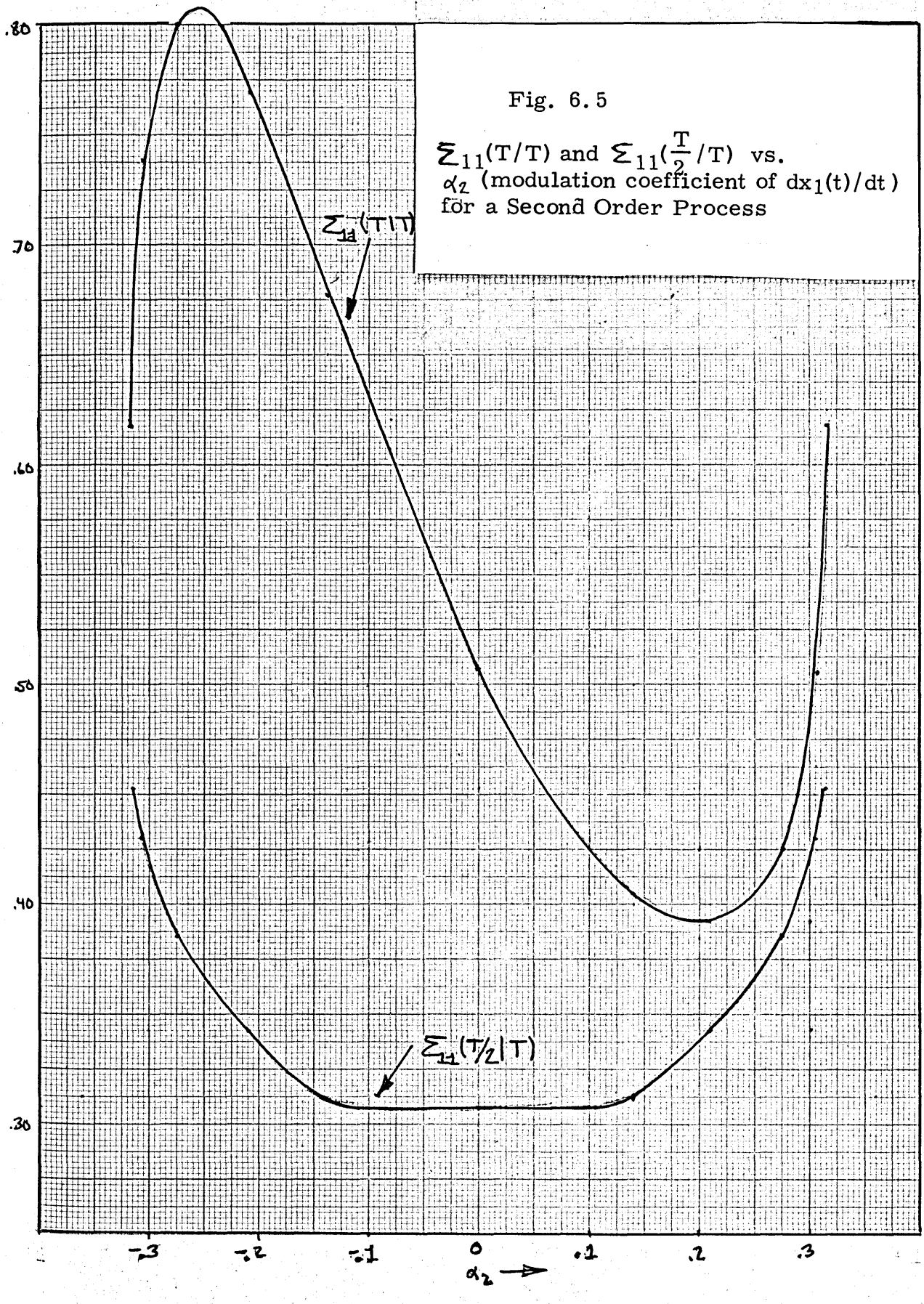


Fig. 6.5

$\Sigma_{11}(T/T)$ and $\Sigma_{11}(\frac{T}{2}/T)$ vs. α_2 (modulation coefficient of $dx_1(t)/dt$) for a Second Order Process

limits, $\Sigma_{11\,\infty}$ and $\Sigma_{11}(t/\infty)$. First, we see that transmitting some of the derivative in the signal does not improve the smoother performance. However, it can either degrade or improve the realizable estimate significantly. Choosing $\alpha_2 = .2$ improves our performance (over $\alpha_2=0$) by approximately 25 percent, whereas choosing $\alpha_2 = -.25$ degrades it by 50 percent. This would indicate from this particular problem that this type of preemphasis is not useful when doing smoothing (or filtering with delay), while it can yield significant improvement for realizable filtering.

D. Filtering with Delay

The optimal smoother uses all the available data, both that in the past and in the future, in making its estimate at a particular point. However, one of the disadvantages of the optimal smoother structure is that it operates over a fixed time interval $[T_o, T_f]$. Consequently, as more data is accumulated, i.e., T_f increases, we must resolve the smoothing equations if we are to use this added data.

In contrast to this, the realizable filter produces an estimate that evolves as the data is accumulated. It, however, uses only past data in making its estimate, whereas, the smoother makes use of both past and future data for its estimate.

The filter realizable with delay combines the advantages of both the realizable filter and the smoother. By allowing a fixed delay before we are required to make our estimate, we can find a filter whose output evolves in time as the data is accumulated yet it is able to take advantage of a limited amount of future data in making

its estimate.

Let us discuss the filter structure in more detail. We assume that we have received the signal $\underline{r}(t)$ over the interval $[T_o, T_f]$. We want to estimate the state vector at the point $t = T_f - \Delta$, i.e., find $\hat{\underline{x}}(T_f - \Delta)(T_f - \Delta > T_o)$, where the independent variable in our filter structure is T_f, the endpoint of observation interval. We note that like the realizable filter, the filter with delay is a point estimator whereas the smoother estimates the signal over an entire interval.

Our approach to finding the filter structure is straightforward.* We use the integral representation specified by solution method 2 of Chapter IV to find $\hat{\underline{x}}(t)$ for the optimal smoother. We then set $t = T_f - \Delta$ in this integral representation. This gives us $\hat{x}(T_f - \Delta)$ in terms of the received data and the realizable filter. We then differentiate this integral to find a set of differential equations for the desired estimate $\hat{\underline{x}}(T_f - \Delta)$. We note that the independent variable for these equations is T_f rather than t, some internal point in a fixed interval.

Let us proceed with our derivation. First we write the smoothing equations Eqs. 6.25 and 6.26 in an augmented matrix form,

---

*This approach to the problem was first used by Baggeroer in the 1966 WESCON Proceedings.[19]

$$\frac{d}{dt}\begin{bmatrix}\hat{\underline{x}}(t)\\ \hline \underline{p}(t)\end{bmatrix} = W(t)\begin{bmatrix}\hat{\underline{x}}(t)\\ \hline \underline{p}(t)\end{bmatrix} - \begin{bmatrix}\underline{0}\\ \hline C^T(t)R^{-1}(t)\underline{r}(t)\end{bmatrix}, \qquad (6.85)$$

$$T_o \leq t \leq T_f.$$

From Eq. 4.41 the solution to these equations in an integral form is

$$\begin{bmatrix}\hat{\underline{x}}(t)\\ \hline \underline{p}(t)\end{bmatrix} = \Psi(t, T_f)\begin{bmatrix}\hat{\underline{x}}_r(T_f)\\ \hline \underline{0}\end{bmatrix} + \int_t^{T_f} \Psi(t,\tau)\begin{bmatrix}\underline{0}\\ \hline C^T(\tau)R^{-1}(\tau)\underline{r}(\tau)\end{bmatrix} d\tau,$$

$$T_o \leq t \leq T_f, \qquad (6.86)$$

where $\hat{\underline{x}}_r(T_f)$ is the realizable filter output at the endpoint of the interval T_f. Let us evaluate Eq. 6.86 at $t = T_f - \Delta$, with $T_f - \Delta > T_o$,

$$\begin{bmatrix}\hat{\underline{x}}(T_f-\Delta)\\ \hline \underline{p}(T_f-\Delta)\end{bmatrix} = \Psi(T_f-\Delta, T_f)\begin{bmatrix}\hat{\underline{x}}_r(T_f)\\ \hline 0\end{bmatrix} + \int_{T_f-\Delta}^{T_f} \Psi(T_f-\Delta, \tau)\begin{bmatrix}\underline{0}\\ \hline \underline{C}^T(\tau)R^{-1}(\tau)\underline{r}(\tau)\end{bmatrix} d\tau$$

(6.87)

This is the desired integral representation. We note that the only time variable involved is T_f. Let us now determine a differential equation that Eq. 6.87 satisfies, where the independent variable is T_f the endpoint of an increasing interval rather than t, some internal time in a fixed interval. Differentiating Eq. 6.87 we obtain

$$\frac{d}{dT_f}\begin{bmatrix}\hat{\underline{x}}(T_f-\Delta)\\ \hline \underline{p}(T_f-\Delta)\end{bmatrix} = \frac{d}{dT_f}\left(\Psi(T_f-\Delta, T_f)\right)\begin{bmatrix}\hat{\underline{x}}_r(T_f)\\ \hline \underline{0}\end{bmatrix} + \Psi(T_f-\Delta, T_f)\begin{bmatrix}\frac{d\hat{\underline{x}}_r(T_f)}{dT_f}\\ \hline \underline{0}\end{bmatrix} +$$

$$\Psi(T_f-\Delta, T_f)\begin{bmatrix}\underline{0}\\ \hline C^T(T_f)R^{-1}(T_f)\underline{r}(T_f)\end{bmatrix} - \begin{bmatrix}\underline{0}\\ \hline C^T(T_f-\Delta)R^{-1}(T_f-\Delta)\underline{r}(T_f-\Delta)\end{bmatrix}$$

$$+\int_{T_f-\Delta}^{T_f} \frac{d}{dT_f}(\Psi(T_f-\Delta, \tau))\begin{bmatrix}0\\ \hline C^T(\tau)R^{-1}(\tau)\underline{r}(\tau)\end{bmatrix} d\tau \qquad (6.88)$$

It is a straightforward exercise to show

$$\frac{d}{dT_f}\Psi(T_f-\Delta, T_f) = W(T_f-\Delta)\Psi(T_f-\Delta, T_f) - \Psi(T_f-\Delta, T_f)W(T_f). \qquad (6.89)$$

We also have

$$\frac{d}{dT_f}\Psi(T_f-\Delta, \tau) = W(T_f-\Delta)\Psi(T_f-\Delta, \tau). \qquad (6.90)$$

Let us substitute these two equations plus the expression for $d\hat{\underline{x}}_r(T_f)/dT_f$ from Eq. 6.29 . We obtain

$$\frac{d}{dT_f}\begin{bmatrix}\hat{\underline{x}}(T_f-\Delta)\\ \hline \underline{p}(T_f-\Delta)\end{bmatrix} = W(T_f-\Delta)\Bigg\{\Psi(T_f-\Delta, T_f)\begin{bmatrix}\hat{\underline{x}}_r(T_f)\\ \hline 0\end{bmatrix} + \int_{T_f-\Delta}^{T_f}\Psi(T_f-\Delta,\tau)\begin{bmatrix}\underline{0}\\ \hline \underline{C}^T(\tau)R^{-1}(\tau)\underline{r}(\tau)\end{bmatrix}d\tau$$

$$+\Psi(T_f-\Delta, T_f)\Bigg\{-W(T_f)\begin{bmatrix}\hat{\underline{x}}_r(T_f)\\ \hline \underline{0}\end{bmatrix} + \begin{bmatrix}\underline{0}\\ \hline C^T(T_f)R^{-1}(T_f)\underline{r}(T_f)\end{bmatrix} +$$

$$\begin{bmatrix}F(T_f)\hat{\underline{x}}_r(T_f) + \Sigma(\frac{T_f}{T_f})C^T(T_f)R^{-1}(T_f)(\underline{r}(T_f)-C(T_f)\hat{\underline{x}}_r(T_f))\\ \hline \underline{0}\end{bmatrix}\Bigg\}$$

$$-\begin{bmatrix}\underline{0}\\ \hline C^T(T_f-\Delta)R^{-1}(T_f-\Delta)\underline{r}(T_f-\Delta)\end{bmatrix} \qquad (6.91)$$

First, we identify the term in the first bracket as

$$\begin{bmatrix}\hat{\underline{x}}(T_f-\Delta)\\ \hline \underline{p}(T_f-\Delta)\end{bmatrix}$$

from Eq. 6.85. If we write out the term in the second bracket, we find that it can be written as

$$\begin{bmatrix} \Sigma\left(\frac{T_f}{T_f}\right) \\ \hline \\ I \end{bmatrix} \left(C^T(T_f)R^{-1}(T_f)(\underline{r}(T_f) - C(T_f)\hat{\underline{x}}_r(T_f)\right)$$

Therefore, we finally obtain the desired differential equation

$$\frac{d}{dT_f}\begin{bmatrix} \hat{\underline{x}}(T_f-\Delta) \\ \hline \underline{p}(T_f-\Delta) \end{bmatrix} = W(T_f-\Delta)\begin{bmatrix} \hat{\underline{x}}(T_f-\Delta) \\ \hline \underline{p}(T_f-\Delta) \end{bmatrix} - \begin{bmatrix} \underline{0} \\ \hline C^T(T_f-\Delta)R^{-1}(T_f-\Delta)\underline{r}(T_f-\Delta) \end{bmatrix}$$

$$+\Psi(T_f-\Delta, T_f)\begin{bmatrix} \Sigma\left(\frac{T_f}{T_f}\right) \\ \hline \\ I \end{bmatrix} C^T(T_f)R^{-1}(T_f)(\underline{r}(T_f) - C(T_f)\hat{\underline{x}}_r(T_f))^* \tag{6.92}$$

The only issue that remains are the initial conditions. In order to specify $\hat{\underline{x}}(T_o)$ and $\underline{p}(T_o)$ we must solve the smoothing equations over the interval $[T_o, T_o+\Delta]$.

Let us now see if we can simplify our structure by eliminating $\underline{p}(T_f-\Delta)$. From Eq. 6.42 we have

$$\underline{p}(T_f-\Delta) = \Sigma\left(\frac{T_f-\Delta}{T_f-\Delta}\right)(\hat{\underline{x}}(T_f-\Delta) - \hat{\underline{x}}_r(T_f-\Delta)) \tag{6.93}$$

---

*This equation was first obtained by Baggeroer in the 1966 WESCON Proceedings.[16]

Substituting this into Eq. 6.92 we find

$$\frac{d\underline{x}(T_f-\Delta)}{dT_f} = \left(F(T_f-\Delta)+G(T_f-\Delta)QG^T(T_f-\Delta)G^T(T_f-\Delta)\Sigma^{-1}\left(\frac{T_f-\Delta}{T_f-\Delta}\right)\right)\hat{\underline{x}}(T_f-\Delta)$$

$$G(T_f-\Delta)QG^T(T_f-\Delta)\Sigma^{-1}\left(\frac{T_f-\Delta}{T_f-\Delta}\right)\hat{\underline{x}}_r(T_f-\Delta)$$

$$\left(\Psi_{\underline{\xi}\underline{\xi}}(T_f-\Delta, T_f)\Sigma\left(\frac{T_f}{T_f}\right)+\Psi_{\underline{\xi}\eta}(T_f-\Delta, T_f)\right)C^T(T_f)R^{-1}(T_f) \times$$

$$\left(\underline{r}(T_f) - C(T_f)\hat{\underline{x}}_r(T_f)\right) \qquad (6.94)$$

One of the difficulties in implementing Eq. 6.94 is the calculation of the coefficient matrix $\Psi_{\underline{\xi}\underline{\xi}}(T_f-\Delta, T_f)\Sigma(T_f/T_f) + \Psi_{\underline{\xi}\eta}(T_f-\Delta, T_f)$. For constant parameter system $\Psi_{\underline{\xi}\underline{\xi}}(T_f-\Delta, T_f)$ and $\Psi_{\underline{\xi}\eta}(T_f-\Delta, T_f)$ can be computed independent of T_f. Therefore, one need only evaluate $\Sigma(T_f/T_f)$ which is already available from the realizable filter structure. For time varying systems it may be more efficient to compute this coefficient matrix by solving a differential equation for it. It is a straightforward exercise using Eq. 6.89 to show that the coefficient matrix satisfies the equation

$$\frac{d}{dT_f}\left[\Psi_{\underline{\xi\xi}}(T_f-\Delta, T_f)\Sigma\left(\frac{T_f}{T_f}\right) + \Psi_{\underline{\xi\eta}}(T_f-\Delta, T_f)\right] =$$

$$\left[F(T_f-\Delta) + G(T_f-\Delta)Q\,G^T(T_f-\Delta)\Sigma^{-1}\left(\frac{T_f-\Delta}{T_f-\Delta}\right)\right] x$$

$$\left[\Psi_{\xi\xi}(T_f-\Delta, T_f)\,\Sigma\left(\frac{T_f}{T_f}\right) + \Psi_{\xi\eta}(T_f-\Delta, T_f)\right]$$

$$-\left[\Psi_{\underline{\xi\xi}}(T_f-\Delta, T_f)\Sigma\left(\frac{T_f}{T_f}\right) + \Psi_{\underline{\xi\eta}}(T_f-\Delta, T_f)\right] x$$

$$\left[F(T_f) + G(T_f)QG^T(T_f)\Sigma^{-1}\left(\frac{T_f}{T_f}\right)\right]. \qquad (6.95)$$

The initial condition follows by setting $T_f = T_o + \Delta$. Eqs. 6.94 together with 6.95 are the same as those derived by Meditch.[31] using a discrete time approach. We simply point out that depending upon the system, we may be able to compute the coefficient matrix more conveniently then by solving a matrix differential equation. The most important aspect of implementing Eq. 6.94 is that it appears to be unstable. Indeed, if implemented directly it would be. To illustrate where the difficulty lies let us pause a moment in our discussion.

Let us consider a differential equation representation for the linear time invariant system with an impulse response

$$h(t) = \begin{cases} e^{\beta t} & 0 < t < \Delta \qquad \beta > 0 \\ 0 & \text{elsewhere} \end{cases} \qquad (6.96)$$

Certainly, this system is stable under any realistic criterion. The output of this system is given by

$$y(t) = \int_{t-\Delta}^{t} x(\tau) e^{\beta(t-\tau)} d\tau. \qquad (6.97)$$

One can easily show that y(t) satisfies the following differential equation

$$\frac{dy(t)}{dt} = \beta y(t) + x(t) - e^{\beta\Delta} x(t-\Delta) \qquad (6.98)$$

Since β is positive, this would indicate an unstable system which contradicts what we had above. The difficulty lies in that we are trying to subtract two responses which are in general unstable to yield a stable net response. This is not very feasible to do in a practical sense.

The consequence of our discussion is that Eq. 6.94 which has the same form as Eq. 6.98 is not suitable for implementation. We should manipulate into an integral form like Eq. 6. which has inputs of $\underline{x}_r(T_f - \Delta)$, $\underline{x}(T_f)$ and $\underline{r}(T_f)$. We should then realize this representation with a topped delay line. Note that our delay line will have a finite length of Δ; consequently, we can always realize it as closely as desired by decreasing the top spacing.

This ends our discussion regarding the structure of the filter with delay. Before proceeding with our discussion of its

performance two comments are in order. First, it is straightforward to find a differential equation for $\underline{p}(T_f-\Delta)$ as well as $\hat{\underline{x}}(T_f-\Delta)$ from Eqs. 6.92 and 6.93 . Second, we emphasize how quickly the filter equations have been derived from the smoother structure by using our technique.

E. Performance of the Filter of Delay

In this section we shall employ the techniques that we have developed to derive a matrix differential equation for $\Sigma(T_f-\Delta/T_f)$, the covariance of error for the filter with delay. This equation is important in two respects. First, it tells us how much we gain in performance by allowing the delay and using the more complicated filter structure. Second, we can find how long the transient effects are before we can attain steady state performance. Since we have already derived many of the required results our derivation is short.

There are two ways in which we can proceed. We can work with Eq. 6.58 and derive $\Sigma(T_f-\Delta/T_f)$ from it. However, to be consistent with our approach to deriving the filter structure we shall use Eq. 6.64 and separate out the partition for $\Sigma(T_f-\Delta/T_f)$. This also has an advantage in identifying some terms in a form which is easily to compute.

To proceed let us work with Eq. 6.64, the integral representation for the solution of Eq. 6.66 . Setting $t = T_f - \Delta$ we have for $T_f - \Delta > T_o$

$$P\left(\frac{T_f-\Delta}{T_f}\right)=\left[\begin{array}{c|c} \Sigma\left(\frac{T_f-\Delta}{T_f}\right) & -\Pi\left(\frac{T_f-\Delta}{T_f}\right)\ \Sigma\left(\frac{T_f}{T_f}\right) \\ \hline -\Sigma\left(\frac{T_f}{T_f}\right)\Pi\left(\frac{T_f-\Delta}{T_f}\right) & -\Pi\left(\frac{T_f-\Delta}{T_f}\right) \end{array}\right]=$$

$$\Psi(T_f-\Delta,T_f)\left[\begin{array}{c|c} \Sigma\left(\frac{T_f}{T_f}\right) & 0 \\ \hline 0 & 0 \end{array}\right]\Psi^T(T_f-\Delta,T_f)-$$

$$\int_{T_f-\Delta}^{T_f}\Psi(T_f-\Delta,\tau)\left[\begin{array}{c|c} G(\tau)QG^T(\tau) & \\ \hline 0 & C^T(\tau)R^{-1}(\tau)C(\tau) \end{array}\right]\Psi^T(T_f-\Delta,\tau)d\tau \tag{6.99}$$

Let us now differentiate this expression with respect to T_f. We need to use Eqs. 6.89 , 6.90 and 6.29 . Doing this we obtain

$$\frac{d}{T_f}P\left(\frac{T_f-\Delta}{T_f}\right)=\left[W(T_f-\Delta)\Psi(T_f-\Delta,T_f)-\Psi(T_f-\Delta,T_f)W(T_f)\right]\left[\begin{array}{c|c} \Sigma\left(\frac{T_f}{T_f}\right) & 0 \\ \hline 0 & 0 \end{array}\right]\Psi^T(T_f-\Delta,T_f)$$

$$+\Psi(T_f-\Delta,T_f)\left[\begin{array}{c|c} F(T_f)\Sigma\left(\frac{T_f}{T_f}\right)+\Sigma\left(\frac{T_f}{T_f}\right)F^T(T_f)+G(T_f)QG^T(T_f) & 0 \\ -\Sigma\left(\frac{T_f}{T_f}\right)G^T(T_f)R^{-1}(T_f)C(T_f)\Sigma\left(\frac{T_f}{T_f}\right) & \\ \hline 0 & 0 \end{array}\right]\Psi^T(T_f-\Delta,T_f)$$

$$+\Psi(T_f-\Delta, T_f)\left[\begin{array}{c|c} \Sigma\left(\frac{T_f}{T_f}\right) & 0 \\ \hline 0 & 0 \end{array}\right]\left[\Psi^T(T_f-\Delta, T_f)W^T(T_f-\Delta)-W^T(T_f)\Psi^T(T_f-\Delta, T_f)\right]$$

$$-\Psi(T_f-\Delta,T_f)\left[\begin{array}{c|c} G(T_f)QG^T(T_f) & \\ \hline 0 & C^T(T_f)R^{-1}(T_fC(T_f) \end{array}\right]\Psi^T(T_f-\Delta, T_f)$$

$$+\left[\begin{array}{c|c} G(T_f-\Delta)QG^T(T_f-\Delta) & 0 \\ \hline 0 & C^T(T_f-\Delta)R^{-1}(T_f-\Delta)C(T_f-\Delta) \end{array}\right]$$

$$+W(T_f-\Delta)\left(\int_{T_f-\Delta}^{T_f}\Psi(T_f-\Delta, \tau)\left[\begin{array}{c|c} G(\tau)QG^T(\tau) & 0 \\ \hline 0 & C^T(\tau)R^{-1}(\tau)C(\tau) \end{array}\right]\Psi^T(T_f-\Delta, \tau)d\tau\right)$$

$$\left(\int_{T_f-\Delta}^{T_f}\Psi(T_f-\Delta, \tau)\left[\begin{array}{c|c} G(\tau)QG^T(\tau) & 0 \\ \hline 0 & C^T(\tau)R^{-1}(\tau)C(\tau) \end{array}\right]\Psi^T(T_f-\Delta, \tau)d\tau\right)W^T(T_f-\Delta)$$

(6.100)

To reduce this equation, we first note that

$$W(T_f)\begin{bmatrix} \Sigma(\frac{T_f}{T_f}) & 0 \\ \hline 0 & 0 \end{bmatrix} = \left[\begin{array}{c|c} F(T_f)\Sigma(\frac{T_f}{T_f}) & 0 \\ \hline C^T(T_f)R^{-1}(T_f)C(T_f)\Sigma(\frac{T_f}{T_f}) & 0 \end{array}\right] \tag{6.101}$$

After we combine terms with common factors and use Eqs. 6.99 and 6.101 , we have

$$\frac{dP(\frac{T_f - \Delta}{T_f})}{dT_f} = W(T_f-\Delta)P(\frac{T_f - \Delta}{T_f}) + P(\frac{T_f - \Delta}{T_f})W^T(T_f-\Delta) -$$

$$\Psi(T_f-\Delta,T_f)\left[\begin{array}{c|c} \Sigma(\frac{T_f}{T_f})C^T(T_f)R^{-1}(T_f)C(T_f)\Sigma(\frac{T_f}{T_f}) & \Sigma(\frac{T_f}{T_f})C^T(T_f)R^{-1}(T_f)C(T_f) \\ \hline C^T(T_f)R^{-1}(T_f)C(T_f)\Sigma(\frac{T_f}{T_f}) & C^T(T_f)R^{-1}(T_f)C(T_f) \end{array}\right] x$$

$$\Psi^T(T_f-\Delta, T_f)$$

$$+\left[\begin{array}{c|c} G(T_f-\Delta)Q\,G^T(T_f-\Delta) & 0 \\ \hline 0 & C^T(T_f-\Delta)R^{-1}(T_f-\Delta)C(T_f-\Delta) \end{array}\right] \tag{6.102}$$

Simplifying the second term on the right, we have finally

$$\frac{dP(\frac{T_f - \Delta}{T_f})}{dT_f} W(T_f-\Delta)P(\frac{T_f - \Delta}{T_f}) + P(\frac{T_f - \Delta}{T_f})W(T_f-\Delta)$$

$$+ \left[\begin{array}{c|c} G(T_f-\Delta)Q\,G^T(T_f-\Delta) & 0 \\ \hline 0 & C^T(T_f-\Delta)R^{-1}(T_f-\Delta)C(T_f-\Delta) \end{array}\right] -$$

$$\Psi(T_f-\Delta,T_f)\left[\begin{array}{c} \Sigma(\frac{T_f}{T_f}) \\ \hline I \end{array}\right] C^T(T_f)R^{-1}(T_f)C(T_f)[\,\Sigma(\frac{T_f}{T_f})\,|\,I]\Psi^T(T_f-\Delta,T_f) \tag{6.103}$$

We see that with the exception of the added term the final equation is very similar to the corresponding one for the smoother.

To find the initial condition $P(T_o/T_o+\Delta)$ we need to solve one of the form(s) of the smoother performance equations for $\Sigma(T_o/T_o+\Delta)$ and $\Pi(T_o/T_o+\Delta)$. Doing this we find

$$P(\frac{T_o}{T_o+\Delta}) = \left[\begin{array}{c|c} \Sigma(\frac{T_o}{T_o+\Delta}) & -\Pi(\frac{T_o}{T_o+\Delta})P_o \\ \hline P_o\Pi(\frac{T_o}{T_o+\Delta}) & -\Pi(\frac{T_o}{T_o+\Delta}) \end{array}\right] \tag{6.104}$$

PAGES (S) MISSING FROM ORIGINAL

Let us now consider the upper left partition for $\Sigma(T_f - \Delta/T_f)$. Using Eqs. 6.99 and 6.61 with $t = T_f - \Delta$, we find

$$\frac{d\Sigma(\frac{T_f - \Delta}{T_f})}{dT_f} = [F(T_f - \Delta) + G(T_f - \Delta)QG^T(T_f - \Delta)\Sigma^{-1}(\frac{T_f - \Delta}{T_f - \Delta})]\Sigma(\frac{T_f - \Delta}{T_f})$$

$$+\Sigma(\frac{T_f - \Delta}{T_f}[F(T_f - \Delta) + G(T_f - \Delta)QG^T(T_f - \Delta)\Sigma^{-1}(\frac{T_f - \Delta}{T_f - \Delta})]^T$$

$$+G^T(T_f - \Delta)QG^T(T_f - \Delta)$$

$$-[\Psi_{\xi\xi}(T_f, T_f - \Delta)\Sigma(\frac{T_f}{T_f}) + \Psi_{\xi\eta}(T_f, T_f - \Delta)]C^T(T_f)R^{-1}(T_f)C(T_f)$$

$$[\Psi_{\xi\xi}(T_f, T_f - \Delta)\Sigma(\frac{T_f}{T_f}) + \Psi_{\xi\eta}(T_f, T_f - \Delta)]^T \qquad (6.105)$$

To find the initial condition $\Sigma(T_o/T_o + \Delta)$ we again need to solve the smoother equations over the interval $[T_o, T_o + \Delta]$. We can again identify the coefficient matrix $\Psi_{\xi\xi}(T_f - \Delta/T_f)\Sigma(T_f/T_f) + \Psi_{\xi\eta}(T_f - \Delta/T_f)$ that we had in the filter structure. As before we can evaluate this by one of two ways. We finally note that Eq. 6.105 together with Eq. 6.95 was first derived by Meditch in Reference 31.

One interesting aspect of Eq. 6.105 is that it is a linear equation unless $\Delta = 0$ whereupon it becomes the matrix Riccati equation for $\Sigma(T_f/T_f)$ as would be expected. The second aspect is

that it is unstable when integrated forward. We, therefore, suggest a backwards integration as we have done in previous problems. To do this one will need to solve the smoother variance equation over the entire interval of interest .

F. Example of Performance of the Filter With Delay

Let us illustrate the results of the previous section with an example. To do this we numerically integrated Eq. 6. The coefficient matrix was evaluated by using the matrix exponential.

The example that we shall study is a one pole process with the initial covariance matrix chosen such that a stationary process is generated. The equations that describe the generation of the process are Eq. 2.16 and Eq. 2.17 . In Fig. 6.6 we have plotted the $\Sigma(T_f - \Delta/T_f)$ that results from our numerical integration vs. $T_f - \Delta$ for various values of Δ. The top curve is for $\Delta = 0$, which is the realizable filter covariance as calculated by solving the Ricatti equation; therefore, we can quickly see how much we gain by allowing a delay. We have also indicated the lower limit on the covariance of error, as calculated from the classical Wiener theory.

Several observations should be made. The transient behavior has about the same duration as that of the realizable filter. The curves approach the asymptotic limits very closely over the interval length considered. Finally, we integrated Eq. 6.92 forward in time. We have previously indicated that this equation is unstable when integrated in this direction. If we plotted these curves over several more time constants, we could see this instability entering. The curves no longer remain constant as they start to grow exponentially.

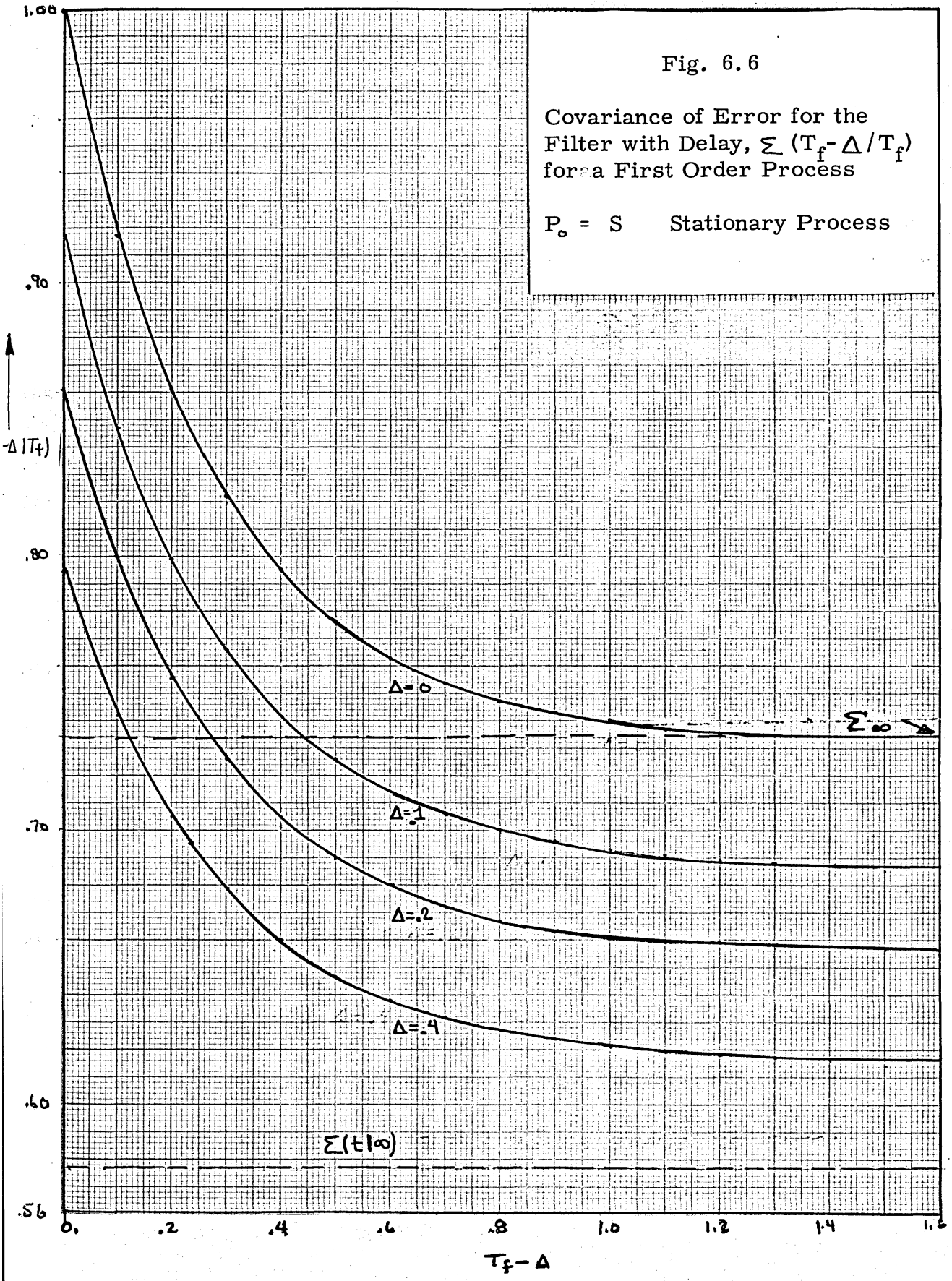


Fig. 6.6

Covariance of Error for the Filter with Delay, $\Sigma (T_f - \Delta / T_f)$ for a First Order Process

P_o = S Stationary Process

This oscillation is even more pronounced in the second order systems that we have studied. This is indicative of the type of behavior we could expect if we integrated the differential equation, Eq. 6.92 describing the estimation structure directly.

G. Discussion and Conclusions

In this chapter we have extensively discussed linear smoothing and filtering with delay. As we stated in the introduction, we feel that our approach is a unified one in that everything follows from the differential equations for the optimal smoother.

The starting point for our development was the finite time Wiener-Hopf equation. We presented a method for deriving the smoothing differential equations from this integral equation by using our results for the Fredholm theory developed in Chapters II and IV. We then derived several different forms for the differential equations specifying the covariance of error for the smoother.

After working several examples of the smoother performance, we discussed the filter with delay. Both its structure and performance could be derived directly from the smoother results. We also indicated possible instability difficulty in implementing these results. We suggested a way to avoid this difficulty; however, we did not develop the suggestion extensively. We concluded the discussion of filter with delay be presenting an example of its performance for various amounts of delay allowed.

The smoother and/or filter with delay are not specifically limited to the area of estimation theory. For example, quite frequently in problems in radar/sonar the receiver has the optimal

smoother as one of its components.[3] Consequently, we can apply our results to realize this part of the receiver. Certainly this is not the only application, for further discussion we refer to Reference 12.

CHAPTER VII

SMOOTHING AND FILTERING FOR NONLINEAR MODULATION SYSTEMS

In the previous chapters we have assumed that the generation of the random processes of interest could be described by a linear system with a finite dimensional linear state representation. In this chapter we shall extend our techniques so as to treat the problems of smoothing and filtering when the observation method is a non-linear function of the state of the system.

With this extension we can represent many modulation schemes and channels of current interest in communications. We should point out though that this is not the most general problem that can be incorporated in the state variable framework. We shall still require that our state equation be linear.

Let us outline our procedure. We shall briefly review our model in order to introduce the notation required for the non-linearity. We shall then show how we can use our techniques to convert an integral equation for the optimal smoothed estimate to a pair of differential equations for it. We shall solve the filtering problem by use of an approximation technique for converting the smoothing equations described over a fixed time interval to a pair of equations which described the realizable filter with a moving end-point.

A. State Variable Modelling of Non-Linear Modulation Systems

In Chapter II we introduced the notation required for the generation of random processes by systems with a linear state representation. Let us modify our generation method to allow the observed output, $\underline{y}(t)$, to be a non-linear function of the state vector, $\underline{x}(t)$. As before, we shall require that the internal dynamics of the generation be described by a linear state equation with a random excitation process and random initial conditions. However, here we need to assume that the state vector generated is a Gaussian random process. Consequently, we have

$$\frac{d\underline{x}(t)}{dt} = F(t)\underline{x}(t) + G(t)u(t), \text{ (linear state equation)}, \tag{7.1}$$

$$E[u(t)u^{T}(\tau)] = Q\delta(t-\tau), \tag{7.2}$$

$$E[\underline{x}(T_o)\underline{x}^{T}(T_o)] = P_o. \tag{7.3}$$

where (t) is a white Gaussian process, with a "spectral height" of Q, and $\underline{x}(T_o)$ is a Gaussian random vector.

The non-linear aspect of this problem enters in how we observe the state vector. We shall assume that the output, or observed process, is a continuous, non-linear, no memory function of the state vector,

$$\underline{y}(t) = \underline{s}(\underline{x}(t), t), \qquad T_o \leq t. \tag{7.4}$$

Let us also introduce for the gradient of $\underline{s}$ with respect to the state vector $\underline{x}(t)$,

$$C(\underline{x}(t),t) \triangleq \frac{\partial\, \underline{s}\,(\underline{x}(t,t)}{\partial(\underline{x}(t))} \triangleq \begin{bmatrix} \frac{\partial s_1(\underline{x}(t),t)}{\partial(x_1(t))} & \frac{\partial s_1(\underline{x}(t),t)}{\partial(x_2(t))} & \cdots & \frac{\partial s_1(x(t),t)}{\partial(x_n(t))} \\ \frac{\partial s_2(\underline{x}(t),t)}{\partial(x_1(t))} & & & \\ \vdots & & & \\ \frac{\partial s_m(\underline{x}(t),t)}{\partial(x_1(t))} & & & \frac{\partial s_m(x(t),t)}{\partial(x_n(t))} \end{bmatrix}, \tag{7.5}$$

$$T_o \leq t$$

Therefore, in the special case of linear systems $C(\underline{x}(t),t)$ is independent of $\underline{x}(t)$ and may be identified as the $C(t)$ which we have been using previously.

We can incorporate certain types of non-linear memory operation in our structure. If we can interpret the modulation, or observation, operation as the cascade of linear system (with memory) which has a state representation for its dynamics then followed by a non-linear no memory operation, we can simply augment the state vector to incorporate the memory operation and then redefine the modulation operation. Probably the most important application of this is FM modulation, where we interpret it as the cascade of an

integrator followed by a phase modulator.

The final aspect of our model is the received signal $\underline{r}(t)$. Again, we shall assume that the signal which the receiver observes over the interval $[T_o, T_f]$ is corrupted by added white noise,

$$\underline{r}(t) = \underline{s}(\underline{x}(t), t) + \underline{w}(t), \qquad T_o \leq t \leq T_f \tag{7.6}$$

where

$$E[\underline{w}(t)\underline{w}^T(\tau)] = R(t)\,\delta(t-\tau) \tag{7.7}$$

with R(t) positive definite. Here we must assume that $\underline{w}(t)$ is also Gaussian.

Let us summarize the differences between what we have assumed here and in Chapter II. First, we allowed the observed signal to be a nonlinear function of the state vector. Secondly, we have assumed that $\underline{v}(t)$ and $\underline{w}(t)$ are Gaussian random processes and $\underline{x}(T_o)$ is a Gaussian random vector. In Chapter II, we made assumptions regarding only their second order statistics.

can
We/consider the problems of smoothing and filtering when the state equation is also non-linear. However, we need to use a different approach. In this approach, we maximize the a posteriori density directly where we incorporate the constraint of the state equation by using a time varying Lagrangian multiplier. This was first done in Ref. and later in Ref.

Finally, we have termed the observation operation as a modulation. This is simply a convenience. One can model a large class of communication systems and channels using this formulation. One simply has a large augmented state vector to incorporate the dynamics of all the systems involved.[10]

B. Smoothing For Non-Linear Modulation Systems

In this section we shall derive the differential equations and their boundary conditions which implicitly generate the optimal smoothed estimate $\hat{\underline{x}}(t)$ when the received signal is generated according to the methods of the previous section. The starting point for our derivation is an integral equation for $\hat{\underline{x}}(t)$ as derived by Van Trees in Ref. 3 using an eigenfunction expansion approach. From Eq. 5.160 in this reference, it is necessary that the optimal estimate x(t) satisfy the following integral equation

$$\hat{\underline{x}}(t) = \int_{T_o}^{T_f} K_{\underline{x}}(t,\tau) C^T(\hat{\underline{x}}(\tau),\tau) R^{-1}(\tau)(\underline{r}(\tau) - \underline{s}(\hat{\underline{x}}(\tau),\tau)) d\tau, \quad T_o \le t \le T_f \tag{7.8}$$

We now observe that we have the same type of kernel for the integral operation as we discussed in Chapter II, Section C. The only difference is that before the kernel operated upon $C^T(\tau)\underline{f}(\tau)$ whereas now we have $C^T(\hat{\underline{x}}(\tau),\tau))R^{-1}(\tau)(\underline{r}(\tau) - \underline{s}(\hat{\underline{x}}(\tau),\tau))$. Consequently, if we want to use the results that we derived in that section, we must examine how this difference of terms affects the derivation. When we do this, we see that the derivation is unaffected. Therefore,

in Eq. 2.22 we can replace $C^T(\tau)\underline{f}(\tau)$ by $C^T(\underline{\hat{x}}(\tau),\tau)R^{-1}(\tau)(\underline{r}(\tau)-\underline{s}(\underline{\hat{x}}(\tau),\tau)$ and reduce Eq. 7.8 to two differential equations with a set of boundary conditions. The differential equations that describe the optimal smoother for this problem are

$$\frac{d\underline{\hat{x}}(t)}{dt} = F(t)\underline{\hat{x}}(t) + G(t)QG^T(t)\underline{p}(t), \quad T_o \le t \le T_f \tag{7.9}$$

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) - C^T(\underline{\hat{x}}(t),t)R^{-1}(t)(\underline{r}(t)-\underline{s}(\underline{\hat{x}}(t),t)), \quad T_o \le t \le T_f. \tag{7.10}$$

The boundary conditions that are imposed are the same. We have

$$\underline{\hat{x}}(T_o) = P_o\underline{p}(T_o), \tag{7.11}$$

$$\underline{p}(T_f) = \underline{0}. \tag{7.12}$$

We shall make two comments regarding Eqs. 7.9 through 7.12. First, our derivation of these differential equations and boundary conditions from Eq. 7.8 is exact. There are no approximations involved. Second, we emphasize that the integral Eq. 7.8 is only a necessary condition. It is usually not sufficient; consequently, its solution need not be unique.

Contrasting Eqs. 7.9 through 7.12 to those derived for the corresponding linear problem in Chapter VI, Section A; we see that the equation for $\underline{\hat{x}}(t)$ is the same, while the nonlinear aspects of the

estimator are introduced in the equation for $\underline{p}(t)$.

Now that we have derived the equations for $\underline{x}(t)$ we must consider methods of solving them. First, let us run through the methods discussed for solving the parallel set of linear equations developed in section C of Chapter IV. Method 1 employed the superposition principle; therefore, it is not applicable. In method 2 we found a complete set of boundary conditions at $t = T_f$ by introducing a function that corresponded to realizable filter output. If we could parallel this we could also solve Eq. 7.9 and 7.10 backwards from the endpoint. In the next section we shall derive an approximate solution for the realizable filter, so this technique is certainly feasible. However, we shall still have instability problems with this method. The key to the third method was a linear relation between the functions $\hat{\underline{x}}(t)$, $\hat{\underline{x}}_r(t)$ and $\underline{p}(t)$ as given by Eq. 4.42. Unfortunately, we do not have such a relationship at the current time. Therefore, only method 2 seems at all promising.

Certainly, there do exist other methods of solving nonlinear two point boundary value problems. One technique is the method of quasi-linearization. With this technique, the estimation equations are linearized around some a priori estimate of the solution. Then these linear equations are solved exactly, by use of the transition matrix associated with this system. This new solution provides the next estimate around which the nonlinear equations are linearized. This technique is repeated until a satisfactory convergence criterion has been satisfied. One of the problems, however, is generating the required a priori estimate of $\hat{\underline{x}}(t)$ and $\underline{p}(t)$. A reasonable choice might be to use the estimates

found by applying method 2, in effect combining this method with a quasilinearization approach.

We have seen how the realizable filter is important in solving the smoothing equations. This filter is certainly of much more interest than its use for this application. Let us now discuss how we can use the smoothing equations to find an approximate realization of the realizable filter.

C. Realizable Filtering for Non-linear Modulation Systems

In this section we shall present a derivation for an approximate realization of the optimum realizable filter. Our derivation is a modified version of that presented by Detchmendy and Shridar.[17,33] In particular, the modifications that we shall make eliminate some of the troublesome issues in their results.

The fundamental difference between the interval estimator and the realizable filter is the time variable involved. In the smoother the important time variable is the time within the fixed observation interval, whereas in the realizable filter the important time variable is the end-point time of the observation interval, which is not fixed but increases continually as the data are accumulated. For the realizable filter we want the estimate at the end point of the observation interval as a function of the length of the interval.

In order to make the transition between the two time variables, we shall use the concept of "invariant imbedding". However, before we do this, let us motivate its use for this particular problem.

We consider a sample function of $\underline{p}(t)$ near the end point

of the observation interval or near $t = T_f$. From Eq. 7.12 we have that it vanishes at $t = T_f$. However, at $t = T_f - \Delta T$, we have from Eq. 7.10

$$\underline{p}(T_f-\Delta T) = -\frac{d\underline{p}(t)}{dt}\Big|_{t = T_f} \Delta T =$$

$$C^T(\hat{\underline{x}}(T_f), T_f)R^{-1}(T_f)(\underline{r}(T_f)-\underline{s}(\hat{\underline{x}}(T_f), T_f))\Delta T \triangleq -\Delta\underline{\eta} \qquad (7.13)$$

Now, we shall examine the same problem with the same sample functions from a slightly different viewpoint. Let us consider the trajectories for $\underline{x}(t)$ and $\underline{p}(t)$ over the interval $[T_o, T_f-\Delta T]$. We can way that these trajectories solve a second problem defined over this shortened interval. For this problem the initial conditions are the same. However, the endpoint time is now $T_f - \Delta T$, and $\underline{p}(T_f-\Delta T)$ is equal to $\Delta\underline{\eta}$ as defined in Eq. 7.13 instead of being identically zero. We can produce the same trajectory by considering an appropriately chosen non-zero boundary condition for $\underline{p}(T_f-\Delta T)$.

This leads us to the following hypothetical question. If we imbed our smoothing problem in a larger class of problems with the boundary condition

$$\underline{p}(T_f) = \underline{\eta} \qquad (7.14)$$

for Eq. 7.10, how does the solution of Eq. 7.9 at T_f depend upon changes in T_f and $\underline{\eta}$? This question is answered by the invariant

imbedding equation, which is a partial differential equation for the solution of Eq. 7.9 at T_f as a function of T_f and $\underline{\eta}$.

First, we shall sketch its derivation.[17]

Let us consider the solutions to Eqs. 7.9 and 7.10 when we impose the final bundary condition specified by Eq. 7.14. We shall denote these solutions by $\underline{x}(t, T_f, \underline{\eta})$ and $\underline{p}(t, T_f, \underline{\eta})$. We have introduced arguments T_f and $\underline{\eta}$ to emphasize that these solutions are dependent upon these parameters. We also point out that we are assuming $\underline{\eta}$ to be an independent variable. We have

$$\underline{x}(t, T_f, \underline{0}) = \hat{\underline{x}}(t), \qquad T_o \leq t \leq T_f, \tag{7.15}$$

$$\underline{p}(T_f, T_f, \underline{\eta}) = \underline{\eta}. \tag{7.16}$$

We define

$$\underline{\xi}(T_f, \underline{\eta}) = \underline{x}(T_f, T_f, \eta). \tag{7.17}$$

We note

$$\underline{\xi}(T_f, 0) = \hat{\underline{x}}(T_f) \tag{7.18}$$

For future convenience let us also define the function Γ by

$$\Gamma(\underline{x}(T_f, T_f, \underline{\eta}), \underline{\eta}, T_f) \triangleq \left.\frac{d\underline{x}(t, T_f, \underline{\eta})}{dt}\right|_{t = T_f} =$$

$$F(T_f)\underline{\xi}(T_f, \underline{\eta}) + G(T_f)QG^T(T_f)\underline{\eta} \tag{7.19}$$

and the function Δ by

$$\Delta(\underline{x}(T_f, T_f, \eta), \underline{\eta}, T_f) \triangleq \left.\frac{d\underline{p}(t, T_f)\underline{\eta})}{dt}\right|_{t = T_f} =$$

$$-F^T(T_f)\underline{\eta} - C^T(\underline{\xi}(T_f, \underline{\eta}), T_f)R^{-1}(T_f)(\underline{r}(T_f) - \underline{s}(\underline{\xi}(T_f, \eta), T_f) \tag{7.20}$$

which is the desired realizable filter estimate at T_f. We shall now determine a partial differential equation for $\underline{\xi}(T_f, \underline{\eta})$ in terms of the variables T_f and $\underline{\eta}$.

Let us examine the solutions $\underline{x}(t, T_f, \eta)$ and $\underline{p}(t, T_f, \eta)$ as illustrated in Fig. 7.1. We have from Eq. 7.10

$$\underline{p}(T_f - \Delta T, T_f, \eta) = \underline{\eta} - \left.\frac{d\underline{p}(t, T_f, \underline{\eta})}{dt}\right|_{t = T_f} \Delta T =$$

$$\underline{\eta} - \Delta(\underline{x}(T_f, T_f, \underline{\eta}), \underline{\eta}, T_f)\Delta T =$$

$$\eta - \Delta(\underline{\xi}(T_f, \underline{\eta}), \underline{\eta}, T_f)\Delta T \triangleq \eta - \Delta\eta \tag{7.21}$$

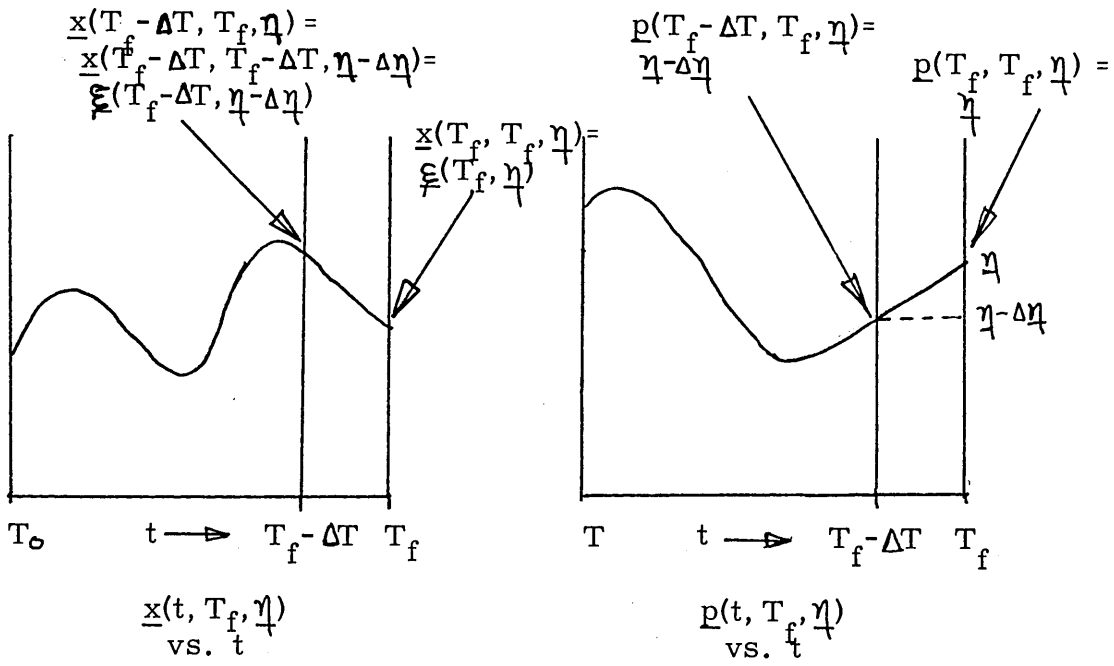


Fig. 7.1

Diagram for Invariant Imbedding Argument

Now, we can interpret $\underline{x}(T_f - \Delta T, T_f, \underline{\eta})$ as being the solution to a second problem producing the same trajectory over the interval $[T_o, T_f - \Delta T]$ with the boundary condition $\underline{\eta} - \Delta\underline{\eta}$, i.e.,

$$\underline{x}(T_f - \Delta T, T_f, \eta) = \underline{x}(T_f - \Delta T, T_f - \Delta T, \underline{\eta} - \Delta\underline{\eta}) =$$

$$\underline{\xi}(T_f - \Delta T, \underline{\eta} - \Delta\underline{\eta}). \tag{7.22}$$

We also have

$$\underline{x}(T_f - \Delta T, T_f, \eta) = \underline{x}(T_f, T_f, \underline{\eta}) - \left.\frac{d\underline{x}(t, T_f, \eta)}{dt}\right|_{t = T_f} \Delta T =$$

$$\underline{\xi}(T_f, \underline{\eta}) - \Gamma(\underline{\xi}(T_f, \underline{\eta}), \underline{\eta}, T_f)\, \Delta T \tag{7.23}$$

Combining Eqs. 7.22 and 7.23, we find

$$\underline{\xi}(T_f - \Delta T, \underline{\eta} - \Delta\underline{\eta}) = \underline{\xi}(T_f, \underline{\eta}) - \Gamma(\underline{\xi}(T_f, \underline{\eta}), \underline{\eta}, T_f)\Delta T \tag{7.24}$$

We also can expand $\xi(T_f, \eta)$ in a two dimensional Taylor series. Doing this, we obtain

$$\underline{\xi}(T_f - \Delta T, \eta - \Delta\underline{\eta}) = \underline{\xi}(T_f, \underline{\eta}) - \frac{\partial\underline{\xi}(T_f, \eta)}{\partial T_f}\Delta T - \frac{\partial\underline{\xi}(T_f, \underline{\eta})}{\partial\underline{\eta}}\Delta\underline{\eta} \tag{7.25}$$

---

*We interpret $\partial\underline{\xi}(T_f, \underline{\eta})/\partial\underline{\eta}$ as in Eq. 7.5.

However, as given by Eq. 7.21, we have constrained $\Delta\underline{\eta}$ to be

$$\Delta\underline{\eta} = \Delta(\underline{\xi}(T_f, \underline{\eta}), \underline{\eta}, T_f)\Delta T \tag{7.26}$$

Substituting Eq. 7.26 in Eq. 7.25, equating the result to Eq. 7.24 and dividing by ΔT we obtain the desired invariant imbedding equation

$$\frac{\partial \underline{\xi}(T_f, \underline{\eta})}{\partial T_f} + \frac{\partial \underline{\xi}(T_f, \underline{\eta})}{\partial \underline{\eta}} \Delta(\underline{\xi}(T_f, \eta), \eta, T_f) = \Gamma(\underline{\xi}(T_f, \eta), \underline{\eta}, T_f) \tag{7.27}$$

This equation relates the value of the solution to Eq. 7.9 at $t = T_f$ to changes in T_f and η, the end point time of the interval and the boundary condition for Eq. 7.10 at $t = T_f$.

We now want to solve Eq. 7.27 and evaluate its solution at $\underline{\eta} = 0$ as prescribed by Eq. 7.18 to find the realizable estimate $x(T_f)$ vs T_f.*

Let us see if we can find a set of ordinary differential equation that will generate the solution to Eq. 7.27. This equation is a partial differential equation; therefore, we would expect that its solution would require an infinite set of equations. In general, this is true; however, let us try a finite order approximation. Since we are interested in the solution at $\underline{\eta} = 0$, let us use a power series

---

*We point out that we have made an expansion in terms of ΔT. Since there is currently some controversy regarding the significance of the terms in the expansion, we are certainly involved in this issue with our approach.[10]

expansion in η. We shall try a solution of the form

$$\underline{\xi}(T_f, \underline{\eta}) = \hat{\underline{x}}(T_f) + P_1(T_f)\underline{\eta} + \text{terms of } O(\tag{7.28}$$

Eq. 7.28 implies that we shall consider explicitly only terms linear in $\underline{\eta}$.

Let us substitute our trial solution into Eq. 7.27. Using Eqs. 7.17 and 7.18 and expanding terms to first order we obtain

$$\frac{d\underline{x}(T_f)}{dT_f} + \frac{dP_1(T_f)}{dT_f}\underline{\eta} +$$

$$P_1(T_f)\left\{ -F^T(T_f)\underline{\eta} - C^T(\hat{\underline{x}}(T_f), T_f)R^{-1}(T_f)(\underline{r}(T_f) - \underline{s}(\hat{\underline{x}}(T_f), T_f) \right.$$

$$\left. - \frac{\partial}{\partial \underline{x}}(C^T(\hat{\underline{x}}, T_f)R^{-1}(T_f)(\underline{r}(T_f) - \underline{s}(\hat{x}, T_f))\Big|_{\hat{\underline{x}} = \hat{\underline{x}}(T_f)} P_1(T_f)\underline{\eta} \right\} +$$

terms of $O(|\eta|^2)$

$$F(T_f)[\hat{\underline{x}}(T_f) + P_1'(T_f)\underline{\eta}] + G(T_f)Q\,G^T(T_f) + \text{terms of } O(|\eta|^2)$$

(7.29)

Now we combine terms of the same order in η. We find

$$\left(\frac{d\hat{\underline{x}}(T_f)}{dT_f} - F(T_f)\hat{\underline{x}}(T_f) - P_1(T_f)C^T(\hat{\underline{x}}(T_f), T_f)R^{-1}(T_f)(\underline{r}(T_f) - \underline{s}(\hat{\underline{x}}(T_f), T_f) \right)\eta^0 +$$

$$\left(\frac{dP_1(T_f)}{dT_f} - P_1(T_f)F^T(T_f) - P_1(T_f)\frac{\partial}{\partial \underline{x}}(C^T(\hat{\underline{x}}, T_f)R^{-1}(T_f)(\underline{r}(T_f) - s(\hat{\underline{x}}, T_f)\Big|_{\hat{\underline{x}} = \hat{\underline{x}}(T_f)}^{P_1(T_f)}\right.$$

$$-F(T_f)P_1(T_f) - G(T_f)Q\,G^T(T_f)\Big)\ \underline{\eta} + \text{terms of } O(|\underline{\eta}|^2) = 0 \tag{7.30}$$

Demanding a solution for arbitrary $\underline{\eta}$ gives us a first order approximation to the realizable filter. For arbitrary $\underline{\eta}$ the coefficients of each power of $\underline{\eta}$ must vanish. We find

$$\frac{d\hat{\underline{x}}(T_f)}{dT_f} = F(T_f)\hat{\underline{x}}(T_f) + P_1(T_f)C^T(\hat{\underline{x}}(T_f), T_f)R^{-1}(T_f)\left(\underline{r}(T_f) - \underline{s}(\hat{\underline{x}}(T_f), T_f)\right)\ , \tag{7.31}$$

$$\frac{dP_1(T_f)}{dT_f} = F(T_f)P_1(T_f) + P_1(T_f)F\,(T_f) + G(T_f)QG^T(T_f)$$

$$-P_1(T_f)\,\frac{\partial}{\partial \underline{x}}\,(C^T(\hat{\underline{x}}, T_f)R^{-1}(T_f)(\underline{r}(T_f) - \underline{s}(\hat{\underline{x}}, T_f))\Big|_{\hat{\underline{x}} = \hat{\underline{x}}(T_f)}\ P_1(T_f) \tag{7.32}$$

To complete the solution we need to specify some initial conditions for Eqs. 7.32 and 7.23. To do this we set $T_f = T_o$ in Eq. 7.28. We have

$$\underline{\xi}(T_o, \underline{\eta})\Big|_{\underline{\eta} = \underline{0}} = \hat{\underline{x}}(T_o) = 0\ , \tag{7.33}$$

since we have assumed zero a priori means. We also have that Eq. 7.28 must satisfy the condition of Eq. 7.11. This implies

$$P_1(T_o) = P_o \tag{7.34}$$

(This also requires that the initial condition for the coefficients in any higher order expansion must be zero.)

Several comments are in order.

1. Although we have not made an issue about terms of order ΔT in our expansion, we have derived the same approximation as found by Snyder who approximated the solution to an a posteriori Fakker-Planck equation.[10]

2. If the observation method is linear, Equations 7.31 through 7.34 are identical to those describing the Kalman-Bucy filter. This is certainly to be expected. For the linear case, it is easy to show that a first order expansion yields an exact solution to the invariant imbedding equation.

3. $P_1(T_f)$ is conditioned upon the received signal; therefore, it cannot be computed a priori as in the linear case. We also point out that we have no reason from this method to equate $P_1(T_f)$ to a conditional covariance matrix. However, in Snyder's approach one can make this identification.

4. Finally, if we want to consider higher order approximations, we should observe how $P_1(T_f)$ couples to the estimate. In general, the higher order terms will couple both ways also, i.e., with the estimate and with the other terms. In addition, the number of elements in the higher order approximations are going to be large, e.g. on the order of $(NF)^n$ where NF is the state vector dimension and n is the approximation order.

D. Summary

In this chapter we have briefly outlined an approach to smoothing and realizable filtering for non-linear modulation systems. We started with an integral equation that specified a necessary condition for the optimal smoothed estimate. We then demonstrated how some of the techniques which we developed earlier in Chapter II could be extended to reduce this integral equation to a pair of non-linear differential equations for the optimal smoothed estimate. This reduction was an exact procedure; therefore, solving the differential equations is equivalent to solving the integral equations.

These differential equations specified the smoother structure for our problem; however, we were still faced with the issue of solving them. One of the methods suggested employed the realizable filter.

With this motivation in addition to the general desirability of solution for the realizable filter, we introduced the concept of invariant imbedding. This concept used the smoother structure to derive a partial differential equation for the realizable filter estimate. This equation was difficult to implement; therefore, we introduced an approach which allowed us to find an approximate solution which could be implemented conveniently. Using this approach, the filter structure followed directly.

CHAPTER VIII

STATE VARIABLE ESTIMATION IN A DISTRIBUTED ENVIRONMENT

In the application of state variable techniques to communication problems, the issue of delay arises quite naturally. For example, since these methods are well adapted to vector processes, they appear to be well suited for problems in array data processing. In these problems, the delay enters because of the finite propagation time of the signals between elements in the array.

When we try to incorporate the issue of delay into our model, several difficulties arise. Although delay is certainly a linear operation, there does not exist a way of representing it with the methods that we introduced in Chapter II. If we examine the issue of delay more closely, we can see why this is true.

A delay operation inherently involves spatial as well as temporal aspects, whether it is caused by a wavefront propagating across an array, or a signal passing through a delay, or transmission. We must recognize that this type of operation is created by a mechanism that is distributed across a spatial environment. Since the model introduced in Chapter II involved only a single time variable, we should not expect that they would be able to handle this situation where there is both a time and spatial variable.

In this chapter we shall extend our state variable techniques so that we can handle a certain class of problems that involve a distributed environment. This class of problems arises quite often in array data processing. We shall discuss our approach in the context of incorporating a delay operation; however, the methods are directly extendable to other distributed environments where the medium may have several spatial coordinates, be nonhomogeneous and/or lossy. As we shall see, even the simple delay mechanism causes a fair amount of difficulty.

Our approach here is also going to be different. Previously, we worked with integral equations and used our methods to reduce them to differential equations. In this chapter, we are going to derive the differential equations directly by assuming that the processes involved are Gaussian and then maximizing the a posteriori density by a variational approach. This does not mean that we cannot extend our previous methods, e.g., the Fredholm integral equation theory to this problem. We have done this; however, it is not of enough general interest to develop it here for this single problem.

A. A State Variable Model for Observation in a Distributed Environment

In this section we shall extend the concepts for the generation of random processes of Chapter II so that we can incorporate the distributed aspects of the delay operation. Our goal is to find a set of state and observation equations to represent signals of the form

$$\underline{y}(t) = \alpha_o \underline{s}(t) + \sum_{l=1}^{m} \alpha_l \underline{s}(t-\tau_i), \quad 0 \leq \tau_1 \leq \tau_2 \leq \ldots \leq \tau_m \tag{8.1}$$

where $\underline{s}(t)$ is a signal process generated by the methods we have previously discussed. (We shall allow vector observations; this feature does not require any additional modifications to the theory.)

Let us assume that we have generated a signal $\underline{s}(t)$ according to these methods. This implies that we have a state and observation equation describing its generation,

$$\frac{d\underline{x}(t)}{dt} = F(t)\underline{x}(t) + G(t)\underline{u}(t), \qquad T_o < t \text{ (state equation)}, \tag{8.2}$$

$$\underline{s}(t) = C(t)\underline{x}(t), \qquad T_o < t \text{ (observation equation)}. \tag{8.3}$$

This also implies that we have made the following assumptions regarding the statistics of the initial state $\underline{x}(T_o)$ and the driving noise process $\underline{v}(t)$,

$$E[\underline{x}(T_o)\underline{x}^T(T_o)] = P_o \tag{8.4}$$

$$E[\underline{u}(t)\underline{u}^T(\tau)] = Q\,\delta(t-\tau) \tag{8.5}$$

For the purposes of the derivation in the next section, we shall also assume that these are a Gaussian random vector and a white Gaussian random process respectively. We emphasize that the

result is not dependent upon this assumption, since an alternate, but more complex,derivation can be made without it by structuring the filter to be linear.

As can easily be seen, the generation of $\underline{s}(t)$ is described by a set of ordinary differential equations, or equations with only one independent variable. However, if we want to use state variable techniques, the functional description of the delay operation as expressed in Eq. 8.1 is not applicable. We need to introduce an equation which describes the dynamics of the system which produces the delay operation. This operation cannot be described in terms of a finite dimensional ordinary differential equation; it is a partial differential equation with a spatial variable z as well as a temporal variable z,

$$\frac{\partial \underline{\Psi}(t,z)}{\partial t} + \frac{\partial \underline{\Psi}(t,z)}{\partial z} = \underline{0} \tag{8.6}$$

For this problem with pure delay, we have assumed a unity velocity of propagation with no loss of generality.

It is easy to see how this equation describes the dynamics of the delay operation. Let us show how we can represent $\underline{s}(t-\tau_i)$ in terms of this equation and a boundary condition at $z = 0$. The general solution to Eq. 8.6 is

$$\underline{\Psi}(t,z) = \underline{\Psi}_o(t-z) \tag{8.7}$$

---

*We should not confuse the vector $\underline{\Psi}(t,z)$ with the earlier transition matrices used.

where Ψ_o is an arbitrary function to be determined. If we impose the boundary condition

$$\underline{\Psi}(t,z)\Big|_{z=0} = \underline{s}(t), \tag{8.8}$$

we find

$$\underline{\Psi}_o(t) = \underline{\Psi}(t,z)\Big|_{z=0} = \underline{s}(t). \tag{8.9}$$

(We could consider that Eq. 8.8 describes the input to a delay line.) If we evaluate $\underline{\Psi}(t,z)$ at $z = \tau_i$, or at the output of the delay line, we obtain

$$\underline{\Psi}(t,z)\Big|_{z=\tau_i} = \underline{\Psi}_o(t-\tau_i) = \underline{s}(t-\tau_i)\,. \tag{8.10}$$

Consequently, we can rewrite Eq. 8.1 as

$$\underline{y}(t) = \alpha_o s(t) + \sum_{i=1}^{m} \alpha_i \Psi(t,\tau_i), \quad T_o < t\,, \tag{8.11}$$

where $\underline{\Psi}(t,z)$ is the output of the dynamic system whose operation is described by Eq. 8.6 and whose input is described by the boundary condition of Eq. 8.8. Thus, we can eliminate the functional description of the delay that appears in Eq. 8.1.

We are now able to write a set of state and observation equations to represent the effect of a delay in our observation. We have the state equations

$$\frac{d\underline{x}(t)}{dt} = F(t)\underline{x}(t) + G(t)\underline{v}(t), \qquad T_o \leq t, \tag{8.2 repeated}$$

$$\frac{\partial \underline{\Psi}(t,z)}{\partial t} = - \frac{\partial \underline{\Psi}(t,z)}{\partial z}, \qquad T_o \leq t, \quad 0 \leq z \tag{8.6 repeated}$$

We impose a boundary condition upon $\underline{\Psi}(t,z)$ at $z = 0$,

$$\underline{\Psi}(t,0) = \underline{s}(t) = C(t)\underline{x}(t) \tag{8.8 repeated}$$

upon $\underline{\Psi}(t,z)$. The observation equation becomes

$$\underline{y}(t) = \alpha_o C(t)\underline{x}(t) + \sum_{i=1}^{m} \alpha_i C(t-\tau_i)\underline{x}(t-\tau_i) =$$

$$\alpha_o s\ (t) + \sum_{i=1}^{m} \alpha_i \Psi(t,\tau_i) \tag{8.11}$$

As we can see, the equation describing the dynamics of the delay operation enters as a state equation. However, there is an important point to be made regarding the difference between Eqs. 8.2 and 8.6. The state of Eq. 8.2 at a particular time can be

described by a set of numbers the state vector. To describe the state of Eq. 8.6, we must specify a set of functions of the variable z. One cannot determine the state of this equation at a future time without knowledge of these functions. In dealing with distributed media, the extension of the state vector to a state function is intrinsic to the idea of a system state.

The concept of the state of the delay operation leads us to the problem of the assumptions that must be made regarding its initial state $\underline{\Psi}(T_o, z)$, $0 \leq z \leq \tau_m$. For the present we shall assume that it is a zero mean Gaussian random process as a function of z with a covariance

$$E[\underline{\Psi}(T_o, z)\underline{\Psi}(T_o, \zeta)] = K_o(z, \zeta) \qquad 0 \leq z, \zeta \leq \tau_m \tag{8.12a}$$

To complete our assumptions regarding our model, we shall assume that we observe $\underline{y}(t)$ over the interval $T_o \leq t \leq T_f$ in the presence of an additive white Gaussian noise. Therefore, our received signal $\underline{r}(t)$ is

$$\underline{r}(t) = \underline{y}(t) + \underline{w}(t), \qquad T_o \leq t \leq T_f, \tag{8.13}$$

where $\underline{y}(t)$ is given by Eq. 8.9 and the covariance of $\underline{w}(t)$ is

$$E[\underline{w}(t)\,\underline{w}^T(\tau)] = R(t)\,\delta(t-\tau) \tag{8.14}$$

Before proceeding to the next section where we shall derive the receiver structure for estimating $\underline{x}(t)$, let us briefly discuss how we could generalize our model for other types of propagation media. The key to this is recognizing that the equation we introduced to describe the delay operation is actually a state equation which describes the dynamics of the distributed media of interest. It is, therefore, appropriate to consider modifying equation 8.6 as a generalization of our model. For example, the dynamics of the medium may be described by the equation

$$A_t(t,z)\frac{\partial \underline{\Psi}(t,z)}{\partial t} + A_z(t,z)\frac{\partial \underline{\Psi}(t,z)}{\partial z} + B(t,z)\underline{\Psi}(t,z) = \underline{0},\ T_o < t,\ 0 < z \tag{8.15}$$

where we impose the boundary condition of Eq. 8.9. Eq. 8.15 is obviously a generalization of the delay mechanism equation. By choosing the coefficient matrices appropriately, one can model non-homogeneous and/or lossy media.

We can also generalize Eq. 8.5 to include time varying gains at each element. In this case Eq. 8.11 becomes

$$y(t) = \alpha_o(t)\underline{s}(t) + \sum_{i=1}^{m} \alpha_i(t)\underline{\Psi}(t,\tau_i),\quad T_o < t \tag{8.16}$$

where $\underline{\Psi}(t,z)$ is the solution to Eq. 8.15.

We have illustrated our model in Fig. 8.1. In this model, $\underline{s}(t)$ is generated by a system as described in Chapter II. It

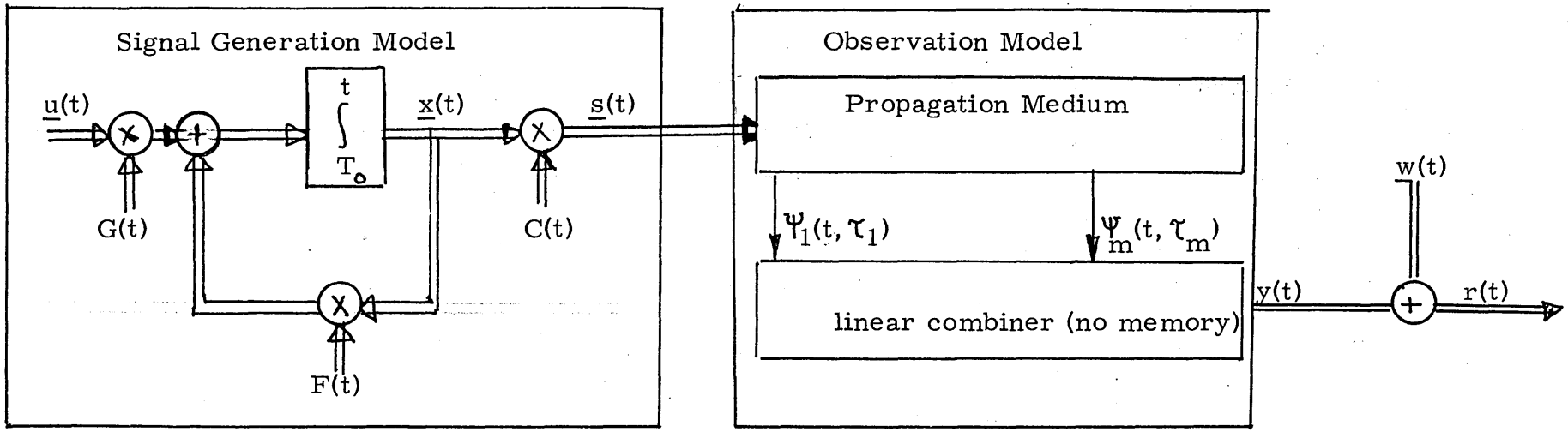


Fig. 8.1a Distributed System Observation Model

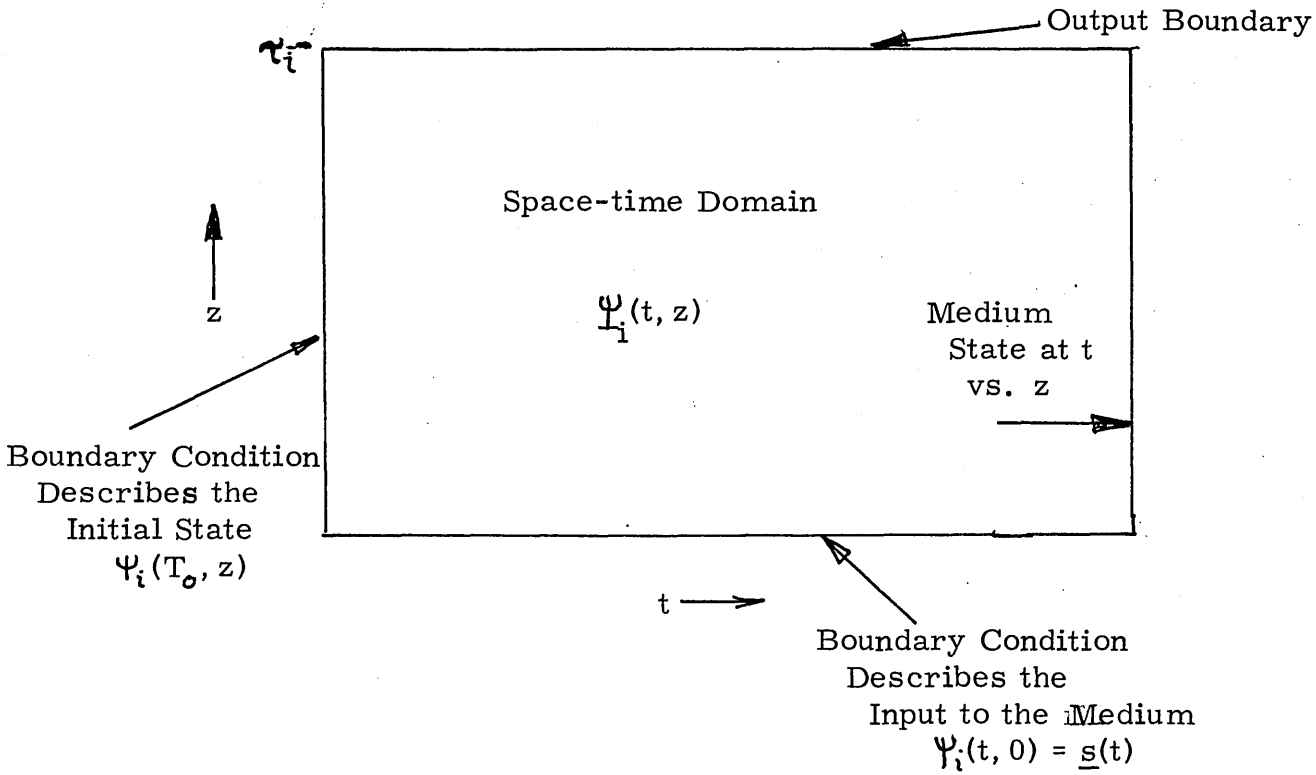


Fig. 8. 1b Space-time Domain for the Propagation Medium

passes through a propagation, a distributed, medium whose dynamics may be described by Eq. 8.15. The signal is then spatially sampled at various points in the medium and linearly combined according to Eq. 8.16. The resulting signal then has white noise added to it before being observed at the receiver. We can generalize the model even further by allowing different propagation paths. In this case, we would have a separate medium description for each path.

The model assumed allows a large degree of flexibility. In some respects too much, since it requires us to make detailed assumptions regarding the medium, which in turn makes the receiver quite detailed and difficult to implement. However, if one is faced with a problem which does have this type of problem entering in its observation process, one must somehow incorporate the spatial aspects of the problem. We feel that our approach is a reasonable attempt at doing this.

B. State Variable Estimation in the Presence of Pure Delay

In this section we shall derive the equations that specify the optimal smoother when pure delay enters into the observation method. We shall find that these equations are a set of differential-difference equations that specify the desired estimate implicitly as part of their solution.

We shall assume that the generation and observation of the received signal may be described by Eqs. 8.1 - 8.14 of the previous section. Our approach to deriving the smoother structure is different than that used in Chapter VI. We shall use a variational approach to maximizing an a posteriori density. In maximizing this

density we shall need to introduce Lagrangian multipliers to incorporate the constraints among the different variables. We should also note that since the delay operation is introduced as a dynamic system with a state equation, we shall be required to estimate its state also. Let us proceed with our derivation.

Since the processes involved were assumed to be Gaussian, it is straightforward to show that the joint a posteriori density for $\underline{x}(t)$ and $\underline{\Psi}(t,z)$ has the form

$$p_{\underline{x},\underline{\Psi}|\underline{r}}(\underline{x}(t),\underline{\Psi}(t,z)|\underline{r}(t) = k \exp(-J(\underline{x}(T_o),\underline{u}(t),\underline{\Psi}(T_o,z)) \quad (8.17)$$

where k is a constant and the functional J is given by

$$J(\underline{x}(T_o),\underline{u}(t),\underline{\Psi}(T_o,z)) = \frac{1}{2}||\underline{x}(T_o)||_{P_o^{-1}} + \frac{1}{2}\int_{T_o}^{T_f}||\underline{u}(t)||_{Q^{-1}}dt$$

$$+\frac{1}{2}\int_{T_o}^{T_f}||\underline{r}(t)-\underline{y}(t)||_{R^{-1}(t)}dt + \frac{1}{2}\int_0^{\tau_m}\int_0^{\tau_m}\underline{\Psi}^T(T_o,z)Q_o(z,\zeta)\underline{\Psi}(T_o,\zeta)dzd\zeta^* \quad (8.18)$$

---

* $||x||_A \triangleq \underline{x}A\underline{x}$

We note that $\underline{x}(t)$ is related to $\underline{x}(T_o)$ and $\underline{u}(t)$ by Eqs. 8.2, $\underline{\Psi}(t,z)$ is specified by Eqs. 8.6 and 8.8, and $\underline{y}(t)$ is given by Eq. 8.11. The arguments for J are $\underline{x}(T_o)$, $\underline{v}(t)$ and $\underline{\Psi}(T_o,z)$ since these variables uniquely determine $\underline{x}(t)$ and $\underline{\Psi}(t,z)$.

It is easy to see that to maximize this density, we could just as well minimize J as a function of the variables $\underline{x}(T_o)$, $\underline{u}(t)$ and $\underline{\Psi}(T_o,z)$. However, to perform this minimization we need to incorporate the constraints between the different variables. First, we need to relate $\underline{x}(t)$ to $\underline{u}(t)$ and $\underline{x}(T_o)$. We can do this by using a time-varying Lagrangian multiplier $\underline{p}(t)$ for Eq. 8.2. We shall add to the functional J an identically zero term L_o of the form

$$L_o = \int_{T_o}^{T_f} \underline{p}^T(t) \left(\frac{d\underline{x}(t)}{dt} - F(t)\underline{x}(t) - G(t)\underline{u}(t) \right) dt \qquad (8.19)$$

It will be useful to integrate the first term by parts. Doing this, we have

$$L_o = \underline{p}^T(T_f)\underline{x}(T_f) - \underline{p}^T(T_o)\underline{x}(T_o) -$$

$$\int_{T_o}^{T_f} \left(\frac{d\underline{p}^T(t)}{dt} \underline{x}(t) + \underline{p}^T(t)F(t)\underline{x}(t) + \underline{p}^T(t)G(t)\underline{u}(t) \right) dt \qquad (8.20)$$

To incorporate each of the delay operations $\underline{s}(t-\tau_i)$ need to introduce a Lagrangian multiplier $\underline{\mu}_i(t,z)$ which is a function of

both space and time. The state equation describing the delay operation is given by Eq. 8.6. Consequently, we add to J in terms L_i which are all identically zero

$$L_i = \int_{T_o}^{T_f} \int_{0}^{\tau_i} \underline{\mu}_i^T(t,z) \left(\frac{\partial \underline{\Psi}(t,z)}{\partial t} + \frac{\partial \underline{\Psi}(t,z)}{\partial z} \right) dt\, dz, \quad i = 1, m. \tag{8.21}$$

We shall also need to have this term integrated by parts. Doing this with respect to both variables t and z, we obtain

$$L_i = \int_{T_o}^{T_f} \left(\underline{\mu}^T(t,\tau_i)\underline{\Psi}(t,\tau_i) - \underline{\mu}^T(t,0)C(t)\underline{x}(t) \right) dt$$

$$+ \int_{T_o}^{\tau_i} \left(\underline{\mu}^T(T_f,z)\underline{\Psi}(T_f,z) - \underline{\mu}^T(T_o,z)\underline{\Psi}(T_o,z) \right) dz$$

$$- \int_{T_o}^{T_f} \int_{0}^{\tau_i} \left(\frac{\partial \underline{\mu}^T(t,z)}{\partial t} + \frac{\partial \underline{\mu}^T(t,z)}{\partial z} \right) \underline{\Psi}(t,z)\, dt\, dz, \quad i = 1, m, \tag{8.22}$$

where we have substituted $C(t)\underline{x}(t)$ for $\underline{\Psi}(t,0)$ according to Eq. 8.8.

We add L_o and L_i as given by Eq. 8.20 and Eq. 8.22, and then we substitute the expressions for $\underline{y}(t)$ and $\underline{s}(t)$ as given by Eq. 8.9 and Eq. 8.3. Doing this, we find,

$$J(\underline{x}(T_o),\underline{u}(t),\Psi(T_o,z)) = \frac{1}{2}\left|\left|\underline{x}(T_o)\right|\right|_{P_o^{-1}} + \frac{1}{2}\int_{T_0}^{T_f}\left|\left|u(t)\right|\right|_{Q^{-1}} dt +$$

$$\frac{1}{2}\int_{T_o}^{T_f}\left|\left|r(t) - \alpha_o\, C(t)\underline{x}(t) - \sum_{i=1}^{m}\alpha_i\,\underline{\Psi}(t,\tau_i\,\right|\right|_{R^{-1}(t)} dt +$$

$$\frac{1}{2}\int_0^{\tau_m}\int_0^{\tau_m}\underline{\Psi}^T(T_o,z)Q_o(z,\zeta)\underline{\Psi}(T_o,\zeta)d\zeta\, dz$$

$$+\,\underline{p}^T(T_f)\underline{x}(T_f) - \underline{p}^T(T_o)\underline{x}(T_o) - \int_{T_o}^{T_f}\left(\frac{d\underline{p}^T(t)}{dt}x(t)+\underline{p}^T(t)F(t)\underline{x}(t)+\underline{p}^T(t)G(t)\underline{u}(t)\right)dt$$

$$\sum_{i=1}^{m}\left(\int_{T_o}^{T_f}\left(\underline{\mu}_i^T(t,\tau_i)\underline{\Psi}(t,\tau_i) - \underline{\mu}_i^T(t,0)\,C(t)\underline{x}(t)\right)dt\right.$$

$$\left.\int_0^{\tau_i}\left(\underline{\mu}_i^T(T_f,z)\underline{\Psi}(T_f,z) - \underline{\mu}_i^T(T_o,z)\underline{\Psi}(T_o,z)\right)dz\right)$$

$$\int_{T_o}^{T_f}\int_0^{\tau_i}\left(\frac{\partial\underline{\mu}_i^T(t,z)}{\partial t} + \frac{\partial\underline{\mu}_i^T(t,z)}{\partial z}\right)\underline{\Psi}(t,z)dt\, dz \qquad (8.23)$$

Let us now proceed with our optimization procedure to minimize J. First, we define $\underline{\hat{x}}(T_o)$, $\underline{\hat{u}}(t)$ and $\underline{\hat{\Psi}}(T_o, z)$ to be the optimum choice for minimizing J. Now we expand J about these optimal estimates, i.e., we let

$$\underline{x}(T_o) = \underline{\hat{x}}(T_o) + \epsilon\, \delta\, \underline{x}(T_o) \tag{8.24}$$

$$\underline{u}(t) = \underline{\hat{u}}(t) + \epsilon\, \delta\, \underline{u}(t) \tag{8.25}$$

$$\underline{\Psi}(T_o, z) = \underline{\hat{\Psi}}(T_o, z) + \epsilon \delta\, \underline{\Psi}(T_o, z) \tag{8.26}$$

Substituting Eqs. 8.24 through 8.26 into Eq. 8.23, we find

$$J(x(T_o), \underline{u}(t), \underline{\Psi}(T_o, z)) = J(\underline{\hat{x}}(T_o), \underline{u}(t), \underline{\Psi}(T_o, z))$$

$$\epsilon \left\{ \underline{\hat{x}}^T(T_o) P_o^{-1}\, \delta \underline{x}(T_o) + \int_{T_o}^{T_f} \underline{u}^T(t) Q^{-1} \delta \underline{u}(t) + \right.$$

$$\int_{T_o}^{T_f} (\underline{r}(t) - \alpha_o C(t) \underline{\hat{x}}(t) - \sum_{i=1}^{m} \alpha_i \underline{\hat{\Psi}}(t, \tau_i))^T \chi$$

$$\left. R^{-1}(t)(-\alpha_o C(t) \delta \underline{x}(t) - \sum_{i=1}^{m} \alpha_i \delta \underline{\Psi}(t, \tau_i)) dt \right.$$

$$+\int_0^{\tau_m}\left(\int_0^{\tau_m}\hat{\underline{\Psi}}^T(T_o,\zeta)Q_o(\zeta,z)d\zeta\right)\delta\underline{\Psi}(T_o,z)dz$$

$$+\underline{p}^T(T_f)\delta\underline{x}(T_f)-\underline{p}^T(T_o)\delta\underline{x}(T_o)-\int_{T_o}^{T_f}\left(\frac{d\underline{p}^T(t)}{dt}\delta\underline{x}(t)+\underline{p}^T(t)F(t)\delta\underline{x}(t)+\underline{p}^T(t)G(t)\delta\underline{u}(t)\right)dt$$

$$\sum_{i=1}^{m}\left[\int_{T_o}^{T_f}\left(\underline{\mu}_i^T(t,\tau_i)\delta\underline{\Psi}(t,\tau_i)-\underline{\mu}_i^T(t,0)C(t)\delta\underline{x}(t)\right)dt\right.$$

$$\int_0^{\tau_i}\left(\underline{\mu}_i^T(T_f,z)\delta\underline{\Psi}(T_f,z)-\underline{\mu}_i^T(T_o,z)\delta\underline{\Psi}(T_o,z)\right)dz$$

$$\left.\int_{T_o}^{T_f}\int_0^{\tau_i}\left(\frac{\partial\underline{\mu}_i^T(t,z)}{\partial t}+\frac{\partial\underline{\mu}_i^T(t,z)}{\partial z}\right)\delta\underline{\Psi}(t,z)\,dt\,dz\right]+\text{terms of } O(\epsilon)^2 \tag{8.27}$$

Now let us combine the common variations. We obtain

$$J(\underline{x}(T_o),\underline{u}(t),\underline{\Psi}(T_o,z))=J(\hat{\underline{x}}(T_o),\hat{\underline{u}}(t),\hat{\underline{\Psi}}(T_o,z))+$$

$$\underline{p}(T_f)\,\delta\underline{x}(T_f)+(\hat{\underline{x}}^T(T_o)P_o^{-1}-\underline{p}^T(T_o))\delta\underline{x}(T_o)+\int_{T_o}^{T_f}\left[(\hat{\underline{u}}^T(t)Q^{-1}-\underline{p}^T(t)G(t))\delta\underline{u}(t)\right.$$

$$+\left(\underline{d}^T(t)\alpha_o C(t) - \frac{d\underline{p}^T(t)}{dt} - \underline{p}^T(t)F(t) - \sum_{i=1}^{m}\underline{\mu}_i^T(t,0)C(t)\right)\delta x(t)$$

$$\sum_{i=1}^{m}\left(-\underline{d}^T(t)\alpha_i + \underline{\mu}_i^T(t,\tau_i)\right)\delta\underline{\Psi}(t,\tau_i)\Bigg]dt +$$

$$\int_0^{\tau_m}\left[\int_0^{\tau_m}\hat{\underline{\Psi}}^T(T_o,\zeta)Q_o(\zeta,z)d\zeta - \sum_{i=1}^{m}\underline{\mu}_i^T(T_o,z)u_{-1}(\tau_i - z)\right]\delta\underline{\Psi}(T_o,z)dz$$

$$\sum_{i=1}^{m}\int_0^{\tau_i}\underline{\mu}_i^T(T_f,z)\delta\underline{\Psi}(T_f,z)dz - \sum_{i=1}^{m}\int_{T_o}^{T_f}\int_0^{\tau_i}\left(\frac{\partial\underline{\mu}_i^T(t,z)}{\partial t} + \frac{\partial\underline{\mu}_i^T(t,z)}{\partial z}\right)\delta\underline{\Psi}(t,z)dt\,dz$$

$$+ \text{terms of } O(\epsilon^2), \tag{8.28}$$

where

$$\underline{d}(t) \triangleq R^{-1}(t)\left(\underline{r}(t) - \alpha_o C(t)\hat{\underline{x}}(t) - \sum_{i=1}^{m}\alpha_i\hat{\underline{\Psi}}(t,\tau_i)\right) \tag{8.29}$$

We shall now make a series of arguments to cause the ϵ variation of the functional to vanish. First, we shall require that the Lagrangian multiplier functions satisfy equations such that the coefficient of some of the variations vanishes identically. We shall then argue that the coefficients of the remaining variations must

vanish because of the optimality of $\hat{\underline{x}}(T_o)$, $\hat{\underline{u}}(t)$, and $\underline{\Psi}(T_o, z)$.

First, we shall impose restrictions upon $\underline{p}(t)$, the Lagrangian multiplier for the state Eq. 8.2. We shall require that it satisfy the differential equation

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) - C^T(t)\left(\alpha_o \underline{d}(t) + \sum_{i=1}^{m} \underline{\mu}_i(t, 0)\right), \quad T_o \leq t \leq T_f. \tag{8.30}$$

In addition, we shall impose the boundary condition

$$\underline{p}(T_f) = \underline{0} \tag{8.31}$$

Next, we shall restrict the Lagrangian multipliers for each of the delays τ_i. We shall require

$$\frac{\partial \underline{\mu}_i(t, z)}{\partial t} + \frac{\partial \underline{\mu}_i(t, z)}{\partial z} = \underline{0}, \quad T_o \leq t \leq T_f, \quad 0 < z < \tau_i, \quad i = 1, m \tag{8.32}$$

We note that these equations can be solved functionally,

$$\underline{\mu}_i(t, z) = \underline{\mu}_{o_i}(t-z), \quad T_o \leq t \leq T_f, \quad 0 < z < \tau_i, \quad i = 1, m \tag{8.33}$$

where $\underline{\mu}_{o_i}(t)$ is a function yet to be determined. In addition, we shall impose both temporal and spatial boundary conditions. We shall require

$$\text{a.}\quad \underline{\mu}_i(T_f, z) = \underline{0}, \qquad 0 \le z \le \tau_i, \tag{8.34}$$

or from Eq. 8.33

$$\underline{\mu}_{o_i}(T_f - z) = \underline{0}, \qquad 0 \le z \le \tau_i; \tag{8.35}$$

$$\text{b.}\quad \underline{\mu}_i(t, \tau_i) = \alpha_i \underline{d}(t), \qquad T_o \le t \le T_f, \tag{8.36}$$

or from Eq. 8.33

$$\underline{\mu}_{o_i}(t - \tau_1) = \alpha_i \underline{d}(t), \quad T_o \le t \le T_f. \tag{8.37}$$

We can interpret each of these conditions in terms of results we developed earlier. Eq. 8.34 is the parallel to Eq. 8.31 in that it specifies a terminal state. Eq. 8.36 is a spatial boundary condition parallel to the one imposed upon $\Psi(t, 0)$ as specified by Eq. 8.8.

If we substitute Eqs. 8.30, 8.31, 8.32, 8.34 and 8.36 in Eq. 8.28

$$J(\underline{x}(T_o), \underline{u}(t), \underline{\Psi}(T_o, z)) = J(\hat{\underline{x}}(T_o), \hat{\underline{u}}(t), \hat{\underline{\Psi}}(T_o, z) +$$

$$\epsilon \left(\left(\hat{\underline{x}}^T(T_o) P_o^{-1} - \underline{p}^T(T_o) \right) \delta \underline{x}(T_o) + \int_{T_o}^{T_f} \left(\hat{\underline{u}}^T(t) Q^{-1} - p^T(t) G(t) \right) \delta \underline{u}(t) \right.$$

(continued)

$$\int_0^{\tau_m} \left[\int_0^{\tau_m} \hat{\underline{\Psi}}^T(T_o,\zeta) Q_o(\zeta, z) d\zeta - \sum_{i=1}^{m} \underline{\mu}_i^T(T_o, z) u_{-1}(\tau_i - z) \right] \delta\underline{\Psi}(T_o, z).$$

(8.38)

If we examine the variations in the above we see that they are the control parameters. Consequently, we can argue that in order for $\hat{\underline{x}}(T_o)$, $\hat{\underline{u}}(t)$ and $\hat{\underline{\Psi}}(T_o, z)$ to be optimum, their coefficients must vanish. We have then

$$\hat{\underline{x}}(T_o) = P_o \underline{p}(T_o) \tag{8.39}$$

$$\hat{\underline{u}}(t) = Q G^T(t)\underline{p}(t), \qquad T_o \le t \le T_f, \tag{8.40}$$

$$\sum_{i=1}^{m} \underline{\mu}_i(T_o, z) u_{-1}(\tau_i - z) = \int_0^{\tau_m} Q_o(z,\zeta)\hat{\underline{\Psi}}(T_o, \zeta) d\zeta, \quad 0 \le z \le \tau_m$$

(8.41)

Eq. 8.40 allows us to eliminate $\hat{\underline{u}}(t)$ in the differential equation for $\hat{\underline{x}}(t)$. We obtain

$$\frac{d\hat{\underline{x}}(t)}{dt} = F(t)\hat{\underline{x}}(t) + G(t)Q\, G^T(t)\underline{p}(t), \qquad T_o \le t \le T_f. \tag{8.42}$$

Eqs. 8.39 and 8.41 specify initial conditions upon the state of the system. Eq. 8.41 presents a problem since it is in an integral equation form. Quite often we shall be able to reduce this integral equation to a differential equation by techniques similar to those we have previously discussed. It may be realistic to assume that the initial state estimate is spatially white. This is certainly a worst case situation. In this case we have

$$K_o(z,\zeta) = K_o(z)\,\delta(z-\zeta), \qquad 0 \le z,\ \zeta \le \tau_m \tag{8.43}$$

for which Eq. 8.41 becomes

$$\sum_{i=1}^{m} \underline{\mu}_i(T_o, z)\, u_{-1}(\tau_i - z) = K_o^{-1}(z)\hat{\underline{\Psi}}(T_o, z), \quad 0 \le z \le \tau_m \tag{8.44}$$

This completes our derivation of the estimator structure by minimizing the functional J. Let us now summarize our results for this structure. We have for $T_o \le t \le T_f$ (all the equations are repeated).

$$\frac{d\hat{\underline{x}}(t)}{dt} = F(t)\hat{\underline{x}}(t) + G(t)Q\,G^T(t)\underline{p}(t) \tag{8.42}$$

$$\frac{\partial\hat{\underline{\Psi}}(t,z)}{\partial t} = -\frac{\partial\hat{\underline{\Psi}}(t,z)}{\partial z}, \qquad 0 \le z \le \tau_m \tag{8.6}$$

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) - C^T(t)\left[\alpha_o\underline{d}(t) + \sum_{i=1}^{m}\underline{\mu}_i(t,0)\right] \tag{8.30}$$

$$\frac{\partial\underline{\mu}_i(t,z)}{\partial t} = -\frac{\partial\underline{\mu}_i(t,z)}{\partial z}, \quad 0 \le z \le \tau_i, \ i = 1, \ m, \tag{8.32}$$

where

$$\underline{d}(t) = R^{-1}(t)\left(\underline{r}(t) - \alpha_o C_1(t)\underline{x}(t) - \sum_{i=1}^{m}\alpha_i\underline{\Psi}_i(t,\tau_i)\right) \tag{8.29}$$

The temporal boundary condition are

$$P_o\underline{p}(T_o) = \hat{\underline{x}}(T_o), \tag{8.39}$$

$$\sum_{i=1}^{m}\underline{\mu}_i(T_o,z)u_{-1}(\tau_i - z) = K_o^{-1}(z)\hat{\underline{\Psi}}(T_o,z), \quad 0 \le z \le \tau_m, \tag{8.44}$$

$$\underline{p}(T_f) = 0, \tag{8.31}$$

$$\underline{\mu}_i(T_f,z) = 0, \qquad 0 < z < \tau_i, \ i = 1, \ m. \tag{8.34}$$

The spatial boundary conditions are

$$\hat{\underline{\Psi}}(t,0) = C(t)\hat{\underline{x}}(t), \qquad T_o \leq t \leq T_f, \tag{8.8}$$

$$\underline{\mu}_i(t,\tau_i) = \alpha_i \underline{d}(t), \qquad T_o \leq t \leq T_f, \; i = 1, \; m. \tag{8.36}$$

If we compare these equations to those for the case of ordinary state variable equations, we can see that these equations are a logical extension to them.

In the above representation, we used the partial differential equations to specify $\hat{\underline{\Psi}}(t,z)$ and the $\underline{\mu}_i(t,z)$. When we are concerned with a pure delay operations, we can solve these partial differential equations and convert the above representation to a set of differential-difference equations. In the general case of arbitrary delay spacing, this representation can be rather complex due to the various time intervals involved. However, in the case of equally spaced delays we can obtain a structure which is easier with which to work. Let us illustrate this with an example.

Example 1 - Estimator Structure for Two Equal Delays

We shall consider the case when m = 2, and

$$T_o = 0$$

$$T_f = T > 4\Delta T$$

$$\tau_1 = \Delta T$$

$$\tau_2 = 2\Delta T \tag{8.37 a-d}$$

This model would correspond to a three element receiver. We shall allow the system for the generation of $\underline{s}(t)$ to have an arbitrary state description.

To convert our receiver to a differential-difference equation form we need to solve the equations for $\hat{\underline{\Psi}}(t, z)$ and $\underline{\mu}_i(t, z)$. We have done this in Eqs. 8.10 and 8.33. First, we want to determine $\underline{d}(t)$ over the interval $[0, T]$. To do this, we need $\hat{\underline{\Psi}}(t, z)$. We have

$$\hat{\underline{\Psi}}(t, z) = \begin{cases} \underline{\Psi}_o(t-z) = C(t-z)\hat{\underline{x}}(t-z) & 0 \leq t - z < T \\ \hat{\underline{\Psi}}(0, z-t) & -2\Delta T < t - z < 0 \end{cases} \tag{8.38}$$

Therefore, we obtain

$$\underline{d}(t) = \begin{cases} R^{-1}(t)[\underline{r}(t) - \alpha_o C(t)\hat{\underline{x}}(t) - \alpha_1 \hat{\underline{\Psi}}(0, \Delta T - t) - \alpha_2 \hat{\underline{\Psi}}(0, 2\Delta T - t)], & 0 \leq t \leq \Delta T, \\ R^{-1}(t)[\underline{r}(t) - \alpha_o C(t)\hat{\underline{x}}(t) - \alpha_1 C(t-\Delta T)\hat{\underline{x}}(t-\Delta T) - \alpha_2 \hat{\underline{\Psi}}(0, 2\Delta T - t)], & \Delta T \leq t \leq 2\Delta T, \\ R^{-1}(t)[\underline{r}(t) - \alpha_o C(t)\hat{\underline{x}}(t) - \alpha_1 C(t-\Delta T)\hat{\underline{x}}(t-\Delta T) - \alpha_2 C(t-2\Delta T)\hat{\underline{x}}(t-2\Delta T)] & 2\Delta T \leq t \leq T. \end{cases} \tag{8.39}$$

Now we need to find the functions $\underline{\mu}_{o_i}(t)$. From Eqs. 8.35 and 8.37 we have

$$\underline{\mu}_{0_1}(t) = \begin{cases} \underline{0}, & T - \Delta T \le t \le T, \\ \alpha_1 \underline{d}(t+\Delta t), & -\Delta T \le t \le T - \Delta T; \end{cases} \tag{8.40a}$$

$$\underline{\mu}_{o_2}(t) = \begin{cases} \underline{0}, & T - 2\Delta T \le t \le T, \\ \alpha_2 \underline{d}(t+2\Delta T), & -2\Delta T \le t \quad T - 2\Delta T. \end{cases} \tag{8.40b}$$

We are now able to write the estimation equations solely in terms of $\hat{\underline{x}}(t)$, $\underline{p}(t)$ and $\hat{\underline{\Psi}}(0, z)$. The equation for $\hat{\underline{x}}(t)$ is the same throughout the interval

$$\frac{d\hat{\underline{x}}(t)}{dt} = F(t)\hat{\underline{x}}(t) + G(t)Q\,G^T(t)\underline{p}(t), \quad 0 \le t \le T \tag{8.41}$$

By substituting Eqs. 8.39 and 8.40 for $\underline{d}(t)$ and $\underline{\mu}_i(t, 0)$ respectively, we find for the different time intervals

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) -$$

$$C^T(t) \Big(\alpha_o R^{-1}(t)[\underline{r}(t) - \alpha_o C(t)\hat{\underline{x}}(t) - \alpha_1 \hat{\underline{\Psi}}(0, \Delta T - t) - \alpha_2 \hat{\underline{\Psi}}(0, 2\Delta T - t)]$$

$$+ \alpha_1 R^{-1}(t+\Delta T)[\underline{r}(t+\Delta T) - \alpha_o C(t+\Delta)\hat{\underline{x}}(t+\Delta T) - \alpha_1 C(t)\hat{\underline{x}}(t) - \alpha_2 \hat{\underline{\Psi}}(0, \Delta T - t)]$$

$$+ \alpha_2 R^{-1}(t+2\Delta)[\underline{r}(t+2\Delta T) - \alpha_o C(t+2\Delta T)\hat{\underline{x}}(t+2\Delta T) - \alpha_1 C(t+\Delta T)\hat{\underline{x}}(t+\Delta T) - \alpha_2 C(t)\hat{\underline{x}}(t) \Big)$$

$$0 \le t \le \Delta t$$

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) =$$

$$C^T(t)\Bigg(\alpha_o R^{-1}(t)[\underline{r}(t)-\alpha_o C(t)\hat{\underline{x}}(t)-\alpha_1 C(t-\Delta T)\hat{\underline{x}}(t-\Delta T)-\alpha_2\hat{\underline{\Psi}}(0,2\Delta T-t)] +$$

$$+\alpha_1 R^{-1}(t+\Delta t)[\underline{r}(t+\Delta T)-\alpha_o C(t+\Delta T)\hat{\underline{x}}(t+\Delta T)-\alpha_1 C(t)\hat{\underline{x}}(t)-\alpha_2 C(t-\Delta T)\hat{\underline{x}}(t-\Delta T)]$$

$$\alpha_2 R^{-1}(t+2\Delta T)[\underline{r}(t+2\Delta T)-\alpha_o C(t+2\Delta T)\hat{\underline{x}}(t+2\Delta T)-\alpha_1 C(t+\Delta T)\hat{\underline{x}}(t+\Delta T)-\alpha_2 C(t)\hat{\underline{x}}(t)]\Bigg),$$

$$\Delta T < t < 2\Delta T, \quad (8.42c)$$

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) -$$

$$C^T(t)\Bigg(\alpha_o R^{-1}(t)[\underline{r}(t)-\alpha_o C(t)\hat{\underline{x}}(t)-\alpha_1 C(t-\Delta T)\underline{x}(t-\Delta T)-\alpha_2 C(t-2\Delta T)\underline{x}(t-2\Delta T)$$

$$\alpha_1 R^{-1}(t+\Delta T)[\underline{r}(t+\Delta T)-\alpha_o C(t+\Delta T)\hat{\underline{x}}(t+\Delta T)-\alpha_1 C(t)\hat{\underline{x}}(t)-\alpha_2 C(t)\Delta T)\hat{\underline{x}}(t-\Delta T)]$$

$$+\alpha_2 R^{-1}(t+2\Delta T)[\underline{r}(t+2\Delta T)-\alpha_o C(t+2\Delta T)\hat{\underline{x}}(t+2\Delta T)-\alpha_1 C(t+\Delta T)\hat{\underline{x}}(t+\Delta T)-\alpha_2 C(t)\hat{\underline{x}}(t)]\Bigg)$$

$$2\Delta T < t < T - 2\Delta T$$

(8.42d)

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) -$$

(continued)

$$C^T(t)\left\{\alpha_o R^{-1}(t)[\underline{r}(t)-\alpha_o C(t)\hat{\underline{x}}(t)-\alpha_1 C(t-\Delta T)\hat{\underline{x}}(T-\Delta T)-\alpha_2 C(t-2\Delta T)\hat{\underline{x}}(t-2\Delta T)]\right.$$

$$\left.\alpha_1 R^{-1}(t+\Delta T)[\underline{r}(t)-\alpha_o C(t+\Delta T)\hat{\underline{x}}(t+\Delta T)-\alpha_1 C(t)\underline{x}(t)-\alpha_2 C(t-\Delta T)\hat{\underline{x}}(t-\Delta T)]\right\}$$

$$T - 2\Delta T < t < T - \Delta T. \qquad (8.42d)$$

$$\frac{d\underline{p}(t)}{dt} = -F^T(t)\underline{p}(t) -$$

$$C^T(t)\left\{\alpha_o R^{-1}(t)[\underline{r}(t)-\alpha_o C(t)\hat{\underline{x}}(t)-\alpha_1 C(t-\Delta T)\hat{\underline{x}}(t-\Delta T)-\alpha_2 C(t-2\Delta T)\underline{x}(t-2\Delta T)]\right\}$$

$$T - \Delta T < t < T. \qquad (8.42e)$$

The boundary conditions are

$$P_o\underline{p}(0) = \hat{\underline{x}}(0) \qquad (8.43)$$

$$\underline{p}(T) = \underline{0} \qquad (8.44)$$

$$K_o^{-1}(t)\underline{\Psi}(0,t) =$$

$$\alpha_1 R^{-1}(\Delta T-t)[\underline{r}(\Delta T-t)-\alpha_o C(\Delta T-t)\hat{\underline{x}}(\Delta T-t)-\alpha_1\hat{\underline{\Psi}}(0,t)-\alpha_2\hat{\underline{\Psi}}(0,\Delta T+t)]$$

$$\alpha_2 R^{-1}(2\Delta T-t)[\underline{r}(2\Delta T-t)-\alpha_2 C(2\Delta T-t)\hat{\underline{x}}(2\Delta T-t)-\alpha_1 C(\Delta T-t)\hat{\underline{x}}(\Delta T-t)-\alpha_2\underline{\Psi}(0,t)],$$

$$0 \le t \le \Delta T, \qquad (8.45a)$$

$$K_o^{-1}(t)\underline{\Psi}(0,t) = [\underline{r}(2\Delta T-t) - C(2\Delta T-t)\hat{\underline{x}}(2\Delta T-t) - \alpha_1\hat{\underline{\Psi}}(0,\Delta T+t) - \alpha_2\hat{\underline{\Psi}}(0,t)],$$

$$\Delta T \leq t \leq 2\Delta T. \qquad (8.45b)$$

Eqs. 8.41 through 8.45 specify our receiver structure. It would be an understatement to say it is simply complex.

The major difficulty encountered is that $\hat{\underline{x}}(t)$ enters the equations both delayed and retarted, i.e., $\hat{\underline{x}}(t)$, $\hat{\underline{x}}(t-\Delta T)$, and $\hat{\underline{x}}(t+\Delta T)$ all can enter the same equation. Unfortunately the mathematical theory for solving this type of equation has not been developed very extensively, if at all. Therefore, we shall suggest two possible approaches.

First, one can approximate the delay operation by some finite Padé approximation. The order would of course be dependent on how large the delay is compared to the correlation time of the process involved. We should point out that we are approximating the receiver, not the environment and then designing the receiver.

A second approach is to augment the state vector to include $\hat{\underline{x}}(t-2\Delta T)$, $\hat{\underline{x}}(t-\Delta T)$, $\hat{\underline{x}}(t+\Delta T)$ and $\hat{\underline{x}}(t+2\Delta T)$, and the corresponding function $\hat{\underline{p}}(t-2\Delta T)$, $\hat{\underline{p}}(t-\Delta T)$, $\hat{\underline{p}}(t+\Delta T)$ and $\hat{\underline{p}}(t+\Delta T)$. Obviously, this increases the dimension of the system involved quite quickly which will impose a severe computational demand. Finding efficient solution algorithms is one of the issues for our future research.

C. Discussion of Estimation for Distributed Systems

To conclude this chapter we shall make some general comments on the status of estimation theory for distributed systems.

As we have done in many parts of this thesis we have borrowed heavily from optimal control theory methods. The research concerning the optimal control for distributed systems if far more advanced than it currently is for estimation theory. Because of the relationship between the two we shall briefly mention some of the more pertinent work that has been done in this area.

The initial research in the area of optimal control theory for distributed systems was done in a series of articles by Butkovskii and Lerner.[35-38] In these articles they developed a maximal (or minimal) principle for such systems. They have also studied approximation procedures by the use of truncating an orthonormal expansion series.

From the aspect of applications to a particular system, all their studies have been concerned with the heat, or diffusion, equation. This particular equation has been investigated in almost all the studies of distributed systems which have appeared in the control literature. This is perhaps unfortunate since this equation, which is of the parabolic type, is not representative of those which commonly describe a dynamic system. These equations are usually of the hyperbolic type, and they are, in particular, related to propagation phenomena.[34, 39]

Since Butkovskii's work several people have made studies in the area. Most notable is Wang.[40, 41] He has employed dynamic programming principles to the control problem. This gives a

functional equation which the optimal process satisfies. He has also formulated the concepts of observability and controllability for distributed systems. In his examples (the heat equation) he resorts to expansion techniques in order to determine the solution. Wang's studies are probably the most concise formulation of the concepts which arise in distributed systems which has appeared to date.

Sakawa[42] has also studied the optimal control of a system described by the heat equation. He exploits the fact that the input-output relationship for such a system can be analytically determined. Again expansion techniques are used to actually find a solution. Other investigators in this area include Goodson, Pierre, and Murray.[43-45]

Recently, the realizable filter structure for the problem that we studied in the last section was obtained by extending the Kalman-Bucy approach.[46] Many of the same issues enter, in particular, delayed and advanced differential-difference equations. In addition the variance equation the results is a function of two essentially spatial variables which further complicates the issue. We could derive the realizable filter from the smoother in Chapter VI. Whether we can do this for the problem considered here remains to be seen. In array processing the realizable filter is not usually employed; however, we have seen that it is very useful in conjunction with the solution of the interval estimation equations.

A current disadvantage of the state variable approach is that a method to determine the mean square error performance has not been determined, whereas one does exist for the classical theory. In the classical theory, however, the system is limited to be

stationary and of infinite time. In addition, numerical integration techniques must be used because of the non-rationality of the spectra involved. A state variable approach to this problem may possibly be computationally easier and at the same time allow one to study the transient or time varying aspects of the problem. This would allow us to determine how much we lose by using an asymptotic analysis.

We mentioned in the introduction that we could derive a Fredholm theory for this problem. We can employ these results to find a second, or modal, approach to this problem. If we actually compute the eigenvalues and eigenfunctions involved we can expand our receiver in terms of them. By truncating this expansion we can obtain an approximate our receiver structure. (This is similar in concept to the approach of Butkovskii and Lerner.[35-38]) A bank of correlations may be an appreciable simpler structure than that derived in last section. Furthermore, we can calculate the performance using this method.

CHAPTER IX

SUMMARY

We have presented an extensive discussion ot the use of state variable techniques for solving Fredholm integral equations and the application of the resulting theory to problems in optimal communications. The material in Chapters 2-4 developed the solution techniques for the integral equations, while the remaining Chapters 5-8 exploited various aspects of the theory to solve various communication theory problems.

In Chapter 2 we introduced the concepts of generating random processes with systems described by state variables. We were particularly interested in the properties of the covariance of the state vector $K_{\underline{x}}(t,\tau)$. By using these properties we were able to reduce the linear operator specified by this covariance to a pair of differential equations with an associated set of boundary conditions. These equations were the key to many of our derivations.

In Chapter 3 we applied these concepts and results to solving homogeneous Fredholm integral equations. We first reduced the integral equation to a homogeneous set of differential equations with imposed boundary conditions. The coefficients for

these equations and the boundary conditions were determined directly in terms of the state matrices that describe the generation of the kernel of the integral equation. We then used the transition matrix associated with these equations to find a transcendental function whose roots specified the eigenvalues. Given these eigenvalues, the eigenfunctions follow directly from the same transition matrix. Finally, we used the same transition matrix that specified the eigenvalues to find the Fredholm determinant. As a result, we had that the only function that needed to be calculated in order to find the eigenvalues, eigenfunctions and the Fredholm determinant function was the transition matrix.

In Chapter 4 we again used the results of Chapter 2 to reduce the nonhomogeneous Fredholm integral equation to a set of nonhomogeneous differential equations with a set of boundary conditions. The coefficients of the differential equations and boundary conditions were again directly related to the state matrices which describe the generation of the kernel. We noted that the differential equations and boundary conditions derived were the same as those that specify the optimal smoother structure. Then we exploited the methods that have been developed in the literature for solving the smoother equations so as to solve the nonhomogeneous integral equation. We were also careful to note the applicability of each method that we introduced.

Chapter 5 considered a problem in optimal signal

design. One of the more important applications of the nonhomogeneous integral equation occurs in the communication problem of detecting a known signal in additive colored noise. We viewed the results of Chapter 4 as describing a dynamic system which related the correlating signal in the optimal receiver to the transmitted signal. We then used Pontryagin's Principle to derive a set of necessary conditions for the signal that optimizes the system performance when both the signal energy and its bandwidth are constrained. By using the necessary conditions we devised an algorithm to design optimal signals and their resulting performance when the channel noise had first and second order spectrums. In the course of doing these examples, the algorithm displayed several interesting features.

Chapter 6 was an extensive presenation on a unified approach to optimal smoothing and filtering realizable with a delay. Our starting point was the finite time Wiener-Hopf equation. We used the nonhomogeneous integral equation results from Chapter 4 and those from Chapter 2 to derive the state variable structure for the smoother. We then found the differential equations for the realizable with delay filter directly from the smoothing equations. We presented several different methods for calculating the covariance of error of the smoother and the filter realizable with delay.

In Chapter 7 we extended the results derived in Chapter 2 so as to treat nonlinear modulation systems. Using these results we reduced an integral equation that specifies a necessary

condition for the smooth estimate to a set of differential equations and boundary condition for it. Our derivation was exact, so the results were equivalent to the original integral equation. We then introduced the concept of invariant imbedding in order to derive an approximate realization of the realizable filter from the nonlinear smoothing equations.

In Chapter 8 we recognized that when pure delay enters our observation, we cannot describe it with a finite dimensional state equation. Consequently, we extended our concept of state to include function states in a space-time domain. With this concept we were able to derive the smoother structure for delayed observations. The methods we used were extendable to other types of distributed media; however, the structure even in the case of pure delay was rather complex.

This completes the general summary of the results. In the course of our discussion we worked many examples to illustrate the methods derived. Although the methods were certainly analytically efficient when used properly, we emphasized the numerical aspects of our methods since this is where we think their major application lies.

We also indicated that the techniques that we used were quite often powerful enough to be either extended or applied to more complex problems. These problems suggest topics for further

research. We shall list them by chapters:

2-4. extending the results on solving the Fredholm integral equations when the kernel is generated by a nonlinear or dispersive system;

5 signal design for spread channels; solution algorithms for problems when hard constraints are imposed;

6 effective implementation of filters realizable with delay;

7 solving the smoothing equations for nonlinear modulation systems;extension of invariant imbedding to treat filters realizable with delay; finding the relationship between the covariance of error and the invariant imbedding terms;

8 evaluation of the smoother performance; effective solution procedures for the smoother; the use of a modal or eigenfunction expansion for realizing the filter structure.

In addition there are many direct applications of the theory we have already developed. Many of these are mentioned at the end of each of the respective chapters.

APPENDIX A

DERIVATION OF THE PROPERTIES OF COVARIANCES FOR STATE VARIABLE RANDOM PROCESSES

In this Appendix we shall sketch the derivation of the properties stated in Chapter II-B. We assume that $\underline{x}(t)$ is generated as discussed in Chapter II-A.

First, we derive Eq. 2.10. Consider the case when $t > \tau$. The state at time t is related to the state at time τ and the input $u(t')$ over the interval $t > t' > \tau$ by

$$\underline{x}(t) = \theta(t,\tau)\underline{x}(\tau) + \int_{\tau}^{t} \theta(t,t')G(t')\underline{u}(t')dt' \qquad (A-1)$$

where $\theta(t,t')$ is defined by Eq. 2.11. If we post-multiply A-1 by $\underline{x}^T(\tau)$, and take expectations we obtain

$$E[\underline{x}(t)\underline{x}^T(\tau)] = \theta(t,\tau)K_{\underline{x}}(\tau,\tau) + \int_{\tau}^{t} \theta(t,t')G(t')E[\underline{u}(t')\underline{x}^T(\tau)]dt' \qquad (A-2)$$

However, because of the Markov nature of the state vector, $\underline{u}(t')$ and $x(\tau)$ are independent over the range of integration. Consequently, the second term of A-2 is zero, and we have

$$K_{\underline{x}}(t,\tau) = \theta(t,\tau)K_{\underline{x}}(\tau,\tau), \qquad t \geq \tau \tag{A-3}$$

which is the first part of the desired result. The derivation of the second part of Eq. 2.10 is identical; therefore, we omit it.

We now derive Eq. 2.12. We proceed by differentiating the definition of $K_{\underline{x}}(t,t)$

$$\frac{d}{dt} K_{\underline{x}}(t,t) = E\left[\frac{d\underline{x}(t)}{dt} \underline{x}^T(t) + \underline{x}(t) \frac{d\underline{x}^T(t)}{dt} \right] \tag{A-4}$$

By substituting the state equation, we obtain

$$\frac{d}{dt} K_{\underline{x}}(t,t) = F(t)K_{\underline{x}}(t,t) + K_{\underline{x}}(t,t)F^T(t)$$

$$+ G(t)\, E\left[\underline{u}(t)\underline{x}^T(t)\right] + E\left[\underline{x}(t)\underline{u}^T(t)\right]G^T(t) \tag{A-5}$$

Since the last two terms are transposes of each other, we consider only the second term. The state at time t in terms of the initial state $\underline{x}(T_o)$ and the input $u(t')$ for $T_o < t' < t$ is given by

$$\underline{x}(t) = \theta(t,T_o)\,\underline{x}\,(T_o) + \int_{T_o}^{t} \theta(t,t')G(t')\underline{u}(t')dt' \tag{A-6}$$

Therefore, we have

$$E[\underline{x}(t)\underline{u}^T(t)]G^T(t) = \{\theta(t, T_o)E[\underline{x}(T_o)\underline{u}^T(t)] + \int_{T_o}^{t} \theta(t, t')G(t')E[\underline{u}(t')\underline{u}^T(t)]\,dt'\}G^T(t) \quad (A-7)$$

We assume that the first term is zero for $t > T_o$. The second term becomes upon performing the expectation

$$E[\underline{x}(t)\underline{u}^T(t)]G^T(t) = \int_{T_o}^{t} \theta(t, t')G(t')Q\,\delta(t'-t)dt'G^T(t) \quad (A-8)$$

The integral is non-zero at only the endpoint of the integration interval. We must assume that the limiting form of the delta function is symmetrical; therefore, only one-half the "area" is included in the integration region. Integration Eq. A-8 thus yields

$$E[\underline{x}(t)\underline{u}^T(t)]G^T(t) = \frac{1}{2}\,G(t)\,Q\,G^T(t) \quad (A-9)$$

Substituting this term plus its transpose into Eq. A-5 gives the desired result

$$\frac{dK_{\underline{x}}(t,t)}{dt} = F(t)K_{\underline{x}}(t,t) + K_{\underline{x}}(t,t)F^T(t) + G(t)Q\,G^T(t), \quad t > T_o \quad (A-9)$$

The initial condition $K_{\underline{x}}(T_o, T_o)$ must be specified.

APPENDIX B

COMPLEX RANDOM PROCESS GENERATION

All of the waveforms considered in the text were low pass signals. In many applications, e.g. the signal design problem that we considered, it is useful to be able to extend these concepts to the case of bandpass waveforms. In this section we shall show how we can do this by using the concept of a complex state variable.

The use of complex notation for representing narrow band processes and functions is well known. For example, if y(t) is a random process that has a spectrum which is narrow band about a carrier ω_c we can represent it in the form

$$y(t) = y_c(t)\cos(\omega_c t) + y_s(t)\sin(\omega_c t) = \mathrm{Re}[\tilde{y}(t)e^{j\omega_c t}] \qquad \text{(B-1a)}$$

where

$$\tilde{y}(t) = y_c(t) - j\, y_s(t) \qquad \text{(B-1b)}$$

$y_c(t)$ and $y_s(t)$ are low pass processes. Under the narrow band and stationarity assumptions, one can show

$$K_{y_c y_c}(\Delta t) = K_{y_s y_s}(\Delta t) \qquad \text{(B-2a)}$$

and

$$K_{y_c y_s}(\Delta t) = -K_{y_c y_s}(-\Delta t) \qquad \text{(B-2b)}$$

Both components have the same covariance, and the cross covariance is an odd function. There are two important points to be made here. We can represent y(t) as sinusoidal modulation of a low pass complex process y(t), termed the complex envelope.The real and imaginary parts of y(t), $y_c(t)$ and $y_s(t)$ respectively, are random processes that have identical auto covariances, and a very particular form for their cross covariance. This provides a key to our analysis.

The use of complex state variables is purely a notational convenience. It is obvious that all the problems which one wants to consider may be solved by expanding the terms into their real and imaginary components. However, for the problem where the complex notation is applicable, this expansion is too general a formulation and it leads to a needlessly cumbersome description of the processes involved.

If we are to describe random processes by complex notation, we want to have a convenient form for representing the various covariances of the components. Obviously, if we have to enumerate them all individually, we have not gained anything over using the higher dimensioned representation. The complex notation for describing random processes is applicable when there are two processes; e.g., the quadrature components $y_c(t)$ and $y_s(t)$ of a narrow band process, which have the same auto

covariances and the particular form for the cross covariance between them. We now want to show that we can find a state representation which generates a complex random process which satisfies Eq. B-2. In addition we want to generalize our concepts so as to include the non-stationary vector process case.

Random Process Generation with Complex Notation

In this section we shall develop the theory needed to describe the generation of complex random processes. Let us assume that we have the following state variable description of a linear system

$$\frac{d\underline{\tilde{x}}(t)}{dt} = \tilde{F}(t)\underline{\tilde{x}}(t) + \tilde{G}(t)\underline{\tilde{u}}(t) \quad \text{(linear state equation)}, \tag{B-3a}$$

$$\underline{\tilde{y}}(t) = \tilde{C}(t)\underline{\tilde{x}}(t), \quad \text{(linear observation equation)} \tag{B-3b}$$

where all the coefficient matrices may be complex. In order to describe the generation we shall make two assumptions regarding the statistics of the driving noise $\underline{\tilde{u}}(t)$ and the initial state vector $\underline{\tilde{x}}(T_o)$. With these two assumptions we shall develop the entire theory. Finally, we shall demonstrate that such representations can indeed be used as a convenience to describe the complex envelope of narrow band processes by showing that they yield results consistent with Eqs. B-1 and B-2 in the stationary case.

First, let us consider the white noise driving function $\underline{\tilde{u}}(t)$. The complex covariance function for the process assuming zero

mean is

$$\tilde{K}_{\underline{\tilde{u}}}(t,\tau) = E[\underline{\tilde{u}}(t)\underline{\tilde{u}}^{+}(\tau)] = \tilde{Q}\,\delta(t-\tau), \tag{B-4}$$

where we have used the notation

$$\underline{\tilde{u}}^{+}(t) = [\underline{\tilde{u}}(t)^{*}]^{T}, \tag{B-5}$$

i.e., the conjugate transpose. Let us expand this complex covariance in terms of the quadrature components.

$$\tilde{K}_{\underline{\tilde{u}}}(t,\tau) = E[(\underline{u}_c(t) - j\underline{u}_s(t))(\underline{u}_c{}^{T}(\tau) + j\underline{u}_s(\tau))] =$$

$$K_{\underline{u}_c\underline{u}_c}(t,\tau) + K_{\underline{u}_s\underline{u}_s}(t,\tau) + jK_{\underline{u}_c\underline{u}_s}(t,\tau) - jK_{\underline{u}_s\underline{u}_c}(t,\tau) =$$

$$\tilde{Q}\,\delta(t-\tau) \tag{B-6}$$

In order that the covariance matrix be a convenient method of representing the covariances and cross covariances of the components of $\underline{\tilde{u}}(t)$, we shall require that

$$K_{\underline{u}_c\underline{u}_c}(t,\tau) = K_{\underline{u}_s\underline{u}_s}(t,\tau) = \frac{1}{2}\,\mathrm{Re}[\tilde{Q}]\,\delta(t-\tau), \tag{B-7a}$$

$$K_{\underline{u}_c\underline{u}_s}(t,\tau) = -K_{\underline{u}_s\underline{u}_c}(t,\tau) = \frac{1}{2}\,\mathrm{Im}[\tilde{Q}]\,\delta(t-\tau). \tag{B-7b}$$

The covariance matrices for the two components are identical non-negative definite matrices, and the cross covariance matrix is a skew symmetric matrix. This implies that $\tilde{Q}(t)$ is a Hermitian matrix with a non-negative definite real part.

We also note that the conjugate operation in the definition of the complex covariance matrix, for under the above assumption we have

$$E[\underline{\tilde{u}}(t)\underline{\tilde{u}}^{T}(\tau)] = Q^{1}\,\delta(t-\tau) = 0, \tag{B-8}$$

Quite often, one does not have correlation between the components of $\underline{u}(t)$ (i.e., $E[\underline{u}_c(t)\,\underline{u}_s{}^{T}(t)] = \underline{0}$) since any correlation between the components of the state vector may be represented in the coefficient matrices $\underline{F}(t)$ and $\underline{G}(t)$. In this case, $\underline{Q}(t)$ is a real non-negative definite symmetric matrix. Also, note that under the assumptions we made, u_{c_i} is uncorrelated with u_{s_i} for all i.

The next issue which we want to consider is the initial conditions. In order that we be consistent with the concept of state, whatever assumptions which we make regarding the state vector at the initial time T_o, should be satisfied at an arbitrary time t ($t \geq T_o$).

First, we shall assume that $\underline{x}(T_i)$ is a complex random vector (we assume zero mean for simplicity). The complex covariance matrix for this random vector is

$$\tilde{P}_o = \tilde{K}_{\underline{\tilde{x}}}(T_o, T_o) = E[\underline{\tilde{x}}(T_o)\underline{\tilde{x}}^+(T_o)] =$$

$$K_{\underline{x}_c\underline{x}_c}(T_o, T_o) + K_{\underline{x}_s\underline{x}_s}(T_o, T_o) +$$

$$jK_{\underline{x}_c\underline{x}_s}(T_o, T_o) - jK_{\underline{x}_s\underline{x}_c}(T_o, T_o) . \qquad \text{(B-9)}$$

The assumptions which we shall make about the covariance of this random vector are

$$K_{\underline{x}_c\underline{x}_c}(T_o, T_o) = K_{\underline{x}_s\underline{x}_s}(T_o, T_o) = \frac{1}{2} \mathrm{Re}[\tilde{P}_o] , \qquad \text{(B-10a)}$$

$$K_{\underline{x}_c\underline{x}_s}(T_o, T_o) = K_{\underline{x}_c\underline{x}_s}(T_i, T_i) = \frac{1}{2} I_m[\tilde{P}_o] . \qquad \text{(B-10b)}$$

Consequently, the complex covariance matrix of initial condition is a Hermitian matrix with a non-negative definite real part. We also note that under the above assumptions

$$E[\underline{\tilde{x}}(T_o)\underline{\tilde{x}}^T(T_o)] = 0 \qquad \text{(B-11)}$$

Let us now consider what these assumptions imply about the covariance of the state vector $\underline{x}(t)$ and the observed signal $\underline{y}(t)$. Since we can relate the covariance of $\underline{y}(t)$ directly to that of the state vector, we shall consider $K_{\underline{x}}(t,\tau)$ first.

In the study of real state variable random processes one can determine $\tilde{K}_{\underline{\underline{x}}}(t, \tau)$ in terms of the state equation matrices, the function Q associated with the covariance of the excitation noise $\underline{\tilde{u}}(t)$, and the covariance $\tilde{K}_{\underline{\underline{x}}}(T_o, T_o)$ of the initial state vector, $\underline{\tilde{x}}(T_o)$. The results for complex state variables are exactly parallel. The only change is that the transpose operation is replaced by a conjugate transpose operation. The methods for determining $\tilde{K}_{\underline{\underline{x}}}(t, \tau)$ are first to find $\tilde{K}_{\underline{\underline{x}}}(t, t)$ as the solution of a linear differential equation and then use the transition matrix associated with the matrix F(t) to relate $\tilde{K}_{\underline{\underline{x}}}(t, \tau)$ to $\tilde{K}_{\underline{\underline{x}}}(t, t)$. $\tilde{K}_{\underline{\underline{x}}}(t, t)$ satisfies the linear matrix differential equation

$$\frac{d\tilde{K}_{\underline{\underline{x}}}(t, t)}{dt} = \tilde{F}(t)\tilde{K}_{\underline{\underline{x}}}(t, t) + \tilde{K}_{\underline{\underline{x}}}(t, t)\tilde{F}^{+}(t) + \tilde{G}(t)\tilde{Q}\tilde{G}^{+}(t) \qquad \text{(B-12)}$$

where the initial condition $\tilde{K}_{\underline{\underline{x}}}(T_o, T_o)$ is given as part of the system description, $\tilde{K}_{\underline{\underline{x}}}(t, \tau)$ is given by

$$\tilde{K}_{\underline{\underline{x}}}(t, \tau) = \begin{cases} \tilde{\theta}(t, \tau)\tilde{K}_{\underline{\underline{x}}}(\tau, \tau), & t > \tau, \\ \tilde{K}_{\underline{\underline{x}}}(t, t)\tilde{\theta}^{+}(\tau, t), & \tau > t. \end{cases} \qquad \text{(B-13a)}$$

where $\theta(t, \tau)$ is the complex transition matrix associated with F(t). We also note for future reference that

$$\tilde{\tilde{K}}_{\underset{\sim}{x}}(t,\tau) = \tilde{\tilde{K}}_{\underset{\sim}{x}}^{+}(\tau,t) \tag{B-13b}$$

We can readily show that $K_{\underset{\sim}{x}}(t,t)$ is Hermitian for all t. In order to do this, we simply perform the conjugate transpose operation upon Eq. B-12. This yields

$$\frac{d\tilde{\tilde{K}}_{\underset{\sim}{x}}^{+}(t,t)}{dt} = \tilde{\tilde{K}}_{\underset{\sim}{x}}^{+}(t,t)\tilde{\tilde{F}}^{+}(t) + \tilde{\tilde{F}}(t)\tilde{\tilde{K}}_{\underset{\sim}{x}}^{+}(t,t) + \tilde{G}(t)\tilde{Q}^{+}\tilde{G}^{+}(t) \tag{B-14a}$$

Q is Hermitian; therefore, $\tilde{K}_{\underset{\sim}{x}}(t,t)$ and $\tilde{K}_{\underset{\sim}{x}}^{+}(t,t)$ satisfy the same linear differential equation. Since $\tilde{K}_{\underset{\sim}{x}}(t,t)$ and $\tilde{K}_{\underset{\sim}{x}}^{+}(t,t)$ have the same initial conditions ($\tilde{K}_{\underset{\sim}{x}}(t,t)$ and $\tilde{K}_{\underset{\sim}{x}}(T_o, T_o)$ is Hermitian by assumption), they must be identical. Consequently, complex covariance matrix $\tilde{K}_{\underset{\sim}{x}}(t,t)$ is Hermitian for all t. We can also show that

$$E[\tilde{\underset{\sim}{x}}(t)\tilde{\underset{\sim}{x}}^{T}(t)] = 0 \tag{B-14b}$$

for all t. In order to do this we note that this expectation satisfies the linear differential equation

$$\frac{d}{dt} E[\tilde{\underset{\sim}{x}}(t)\tilde{\underset{\sim}{x}}^{T}(t)] = \tilde{\tilde{F}}(t)E[\tilde{\underset{\sim}{x}}(t)\tilde{\underset{\sim}{x}}^{T}(t)] + E[\tilde{\underset{\sim}{x}}(t)\tilde{\underset{\sim}{x}}^{T}(t)]\tilde{\tilde{F}}^{T}(t) +$$

$$\tilde{G}(t)Q^{1}\tilde{\tilde{G}}^{T}(t) \tag{15}$$

Since Q^1 equals zero (Eq. B-8), the forcing term in this equation is zero. In addition, the homogeneous solution is zero for all t

(Eq. B-11). This proves the assertion.

By using the above, we may prove that $E[\tilde{\underline{x}}(t)\tilde{\underline{x}}^T(\tau)]$ equals zero for all t and τ. To do this we note that

$$E[\tilde{\underline{x}}(t)\tilde{\underline{x}}^T(\tau)] = \begin{cases} \tilde{\theta}(t,\tau)E[\tilde{\underline{x}}(\tau)\tilde{\underline{x}}^T(\tau)], & t > \tau, \\ E[\tilde{\underline{x}}(t)\tilde{\underline{x}}^T(t)]\tilde{\theta}^T(t,\tau), & \tau > t. \end{cases} \qquad \text{(B-16)}$$

Since the expectations on the right side of the above equation are zero, the expectation on the left equals zero for all t and τ.

We note here that the assumptions which we have made on the covariance of the initial state vector $\tilde{\underline{x}}(T_o)$ are satisfied by the covariance of the state vector $\tilde{\underline{x}}(t)$ for all $t \geq T_o$.

Usually, we are not concerned directly with the state vector of a system. The vector interest is the observed signal, $\underline{y}(t)$, which is related to the state vector by Eq. B-3b. We can simply state the properties of the covariance $\tilde{K}_{\underline{y}}(t,\tau)$ since it is related directly to the covariance of the state vector. This relationship is given by

$$\tilde{K}_{\underline{y}}(t,\tau) = \tilde{C}(t)\tilde{K}_{\underline{x}}(t,\tau)\tilde{C}^+(\tau) \qquad \text{(B-17)}$$

Consequently, it is clear that $\tilde{K}_{\underline{y}}(t,t)$ is Hermitian and that $E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau)]$ is zero.

We are now in a position to derive some of the properties regarding the individual components of the observed

signal. First, let us prove that the covariances of $\underline{y}_c(t)$ and $\underline{y}_s(t)$ are identical. We have

$$E[\underline{y}_c(t)\underline{y}_c^T(\tau)] = E\left\{\left(\frac{\tilde{\underline{y}}(t)+\tilde{\underline{y}}^*(t)}{2}\right)\left(\frac{\tilde{\underline{y}}^T(\tau)+\tilde{\underline{y}}^{*T}(\tau)}{2}\right)\right\}$$

$$= \frac{1}{4}\left(E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau)] + E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^{*T}(\tau)] + E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^{*T}(\tau)] + E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau)]^*\right)$$

$$= \frac{1}{2}\,\mathrm{Re}[\tilde{K}_{\underline{y}}(t,\tau)] \qquad \text{(B-18a)}$$

$$E[\underline{y}_s(t)\underline{y}_s(\tau)] = E\left\{ j\left(\frac{\tilde{\underline{y}}(t)-\tilde{\underline{y}}^*(t)}{2}\right) j\left(\frac{\tilde{\underline{y}}^T(\tau)-\tilde{\underline{y}}^{*T}(\tau)}{2}\right)\right\}$$

$$= \frac{1}{4}\left(-E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau)] + E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^{*T}(\tau)] + E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^{*T}(\tau)]^* - E[\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau)]^*\right)$$

$$\frac{1}{2}\,\mathrm{Re}[\tilde{K}_{\underline{y}}(t,\tau)] \qquad \text{(B-18b)}$$

Consequently, the covariances of both components are equal to one-half the real part of the complex covariance matrix.

The cross covariance between components may be found in a similar fashion. We have

$$E[\underline{y}_c(t)\underline{y}_s(\tau)] = E\left\{\left(\frac{\tilde{\underline{y}}(t)+\tilde{\underline{y}}^*(t)}{2}\right) j\left(\frac{\tilde{\underline{y}}^T(\tau)-\tilde{\underline{y}}^{*T}(\tau)}{2}\right)\right\}$$

$$= \frac{1}{4}(E(\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau)) - E(\tilde{\underline{y}}(t)\tilde{\underline{y}}^{*T}(\tau)) +$$

$$E(\tilde{\underline{y}}(t)\tilde{\underline{y}}^{*T}(\tau))^* - E(\tilde{\underline{y}}(t)\tilde{\underline{y}}^T(\tau))^*)$$

$$= -\frac{1}{2}\,\mathrm{Im}[\tilde{K}_{\underline{y}}(t,\tau)] \qquad \text{(B-19)}$$

By using Eqs. B-18 and B-19, we have a convenient method for finding the auto- and cross-covariance of the real and imaginary components of the complex signal $\underline{y}(t)$ in terms of the complex covariance function $K_{\underline{y}}(t,\tau)$. This provides the notational convenience of working with just one covariance matrix, yet it allows us to determine the covariances of the individual components. This is a major advantage of our complex notation.

Stationary Random Processes

We now want to show that the assumptions which we made lead to results which are consistent with those which have been developed for stationary scalar random processes. By choosing $\tilde{\underline{P}}_o$ to be the steady state solution to Eq. B-12, i.e.,

$$\tilde{\underline{P}}_o = \lim_{t\to\infty} \tilde{K}_{\underline{x}}(t,t), \qquad \text{(B-20)}$$

We can show that the covariance matrix $\tilde{\underline{K}}_{\underline{y}}(t,\tau)$ is stationary. Equivalently, we could say that under this assumption for the initial state vector $\tilde{\underline{x}}(T_o)$, $\tilde{\underline{y}}(t)$ is a segment of stationary random processes. For stationary random processes we shall use the notation

$$\tilde{\underline{K}}_{\underline{y}}(\Delta t) = \tilde{\underline{K}}_{\underline{y}}(t,\ t+\Delta t), \tag{B-21}$$

i.e., we shall use only one argument.

Since we have already proven that the real and imaginary components have the same covariance (Eq. B-18), then Eq. B-2a is certainly satisfied. Therefore, we need only to prove that the cross covariance is an odd function. In general, we have (Eq. B-14),

$$\mathrm{Re}[\tilde{\underline{K}}_{\underline{y}}(\Delta t)] + j\,\mathrm{Im}[\tilde{\underline{K}}_{\underline{y}}(\Delta t)]$$

$$= \mathrm{Re}[\tilde{\underline{K}}_{\underline{y}}(\Delta t)]^T - j\,\mathrm{Im}[\tilde{\underline{K}}_{\underline{y}}(\Delta t)]^T. \tag{B-22a}$$

By equating the imaginary parts and substituting Eq. B-19, we obtain

$$\underline{K}_{\underline{y}_c\underline{y}_s}(\Delta t) = -\underline{K}^T_{\underline{y}_c\underline{y}_s}(\Delta t) = -\underline{K}_{\underline{y}_s\underline{y}_c}(-\Delta t), \tag{B-22b}$$

which for the scalar case is consistent with Eq. B-2b. These conditions are sufficient to show that $\underline{y}(t)$ has a real, positive spectrum in the scalar case.

Summary

In this section we introduced the idea of generating a complex random process by exciting a linear system having a complex state variable description with a complex white noise. We then showed how we could describe the second order statistics of this process in terms of a complex covariance function and then we discussed how we could determine this function from the state variable description of the system. The only assumptions which we were required to make were on $\underset{\sim}{\tilde{u}}(t)$ and $\underset{\sim}{\tilde{x}}(T_o)$. Our results were independent of the form of the coefficient matrices F(t), G(t), and C(t). Our methods were exactly parallel to those for real state variables, and our results were similar to those which have been developed for describing narrow band processes by complex notation. We shall now consider two examples to illustrate the type of random processes which we may generate.

Example 1

The first example which we consider is the first order (scalar) case. The equations which describe this system are

$$\frac{d\tilde{x}(t)}{dt} = -k\tilde{x}(t) + \tilde{u}(t) \qquad \text{(state equation)}, \qquad \text{(B-23)}$$

and

$$\tilde{y}(t) = \tilde{x}(t) \qquad \text{(observation equation)}. \qquad \text{(B-24)}$$

The assumptions regarding $\tilde{u}(t)$ and $\tilde{x}(T_o)$ are

$$E[\tilde{u}(t)\tilde{u}^*(\tau)] = 2\mathrm{Re}[\tilde{k}]P\delta(t-\tau) \tag{B-25}$$

and

$$E[|\tilde{x}(T_o)|^2] = P_o \tag{B-26}$$

Since we have a scalar process, both P and P_o must be real. In addition, we have again assumed zero means.

First, we shall find $\tilde{K}_x(t,t)$. The differential equation, Eq. B-12, which it satisfies is

$$\frac{d\tilde{K}_x(t,t)}{dt} = -\tilde{k}\tilde{K}_x(t,t) - \tilde{k}^*\tilde{K}_x(t,t) + 2\mathrm{Re}[\tilde{k}]P$$

$$= -2\mathrm{Re}[\tilde{k}]\tilde{K}_x(t,t) + 2\mathrm{Re}[\tilde{k}]P \tag{B-27}$$

The solution to this equation is

$$\tilde{K}_x(t,t) = P + (P-P_o)e^{-2\mathrm{Re}[\tilde{k}](t-T_o)}, \quad t > T_o \tag{B-28}$$

In order to find $\tilde{K}_x(t,\tau)$, we need to find $\theta(t,\tau)$, the transition matrix for this system. This is easily found to be

$$\tilde{\theta}(t,\tau) = e^{-\tilde{k}(t-\tau)} \tag{B-29}$$

By substituting Eqs. B-28 and B-29 into Eq. B-13, we find $\tilde{K}_x(t,\tau)$ which is also $\tilde{K}_y(t,\tau)$ for this particular example. Furthermore, we

find the auto- and cross-covariance of the individual components by applyting Eqs. B-18 and B-19 respectively.

Let us now consider the stationary case in more detail. In this case

$$P_o = P. \tag{B-30}$$

If we perform the indicated substitutions, we obtain

$$\tilde{K}_{\tilde{x}}(\Delta t) = \begin{cases} Pe^{-\tilde{k}\Delta t} & \Delta t \geq 0 \\ Pe^{-\tilde{k}^*\Delta t} & \Delta t \leq 0 \end{cases} \tag{B-31}$$

This may be written as

$$\tilde{K}_{\tilde{x}}(\Delta t) = Pe^{-\mathrm{Re}[\tilde{k}]\,|\Delta t|}\, e^{-j\mathrm{Im}[\tilde{k}]\Delta t} \tag{B-32}$$

By applying Eqs. B-18 and B-19 we find

$$K_{x_c x_c}(\Delta t) = K_{x_s x_s}(\Delta t) = \frac{P}{2} e^{-\mathrm{Re}[\tilde{k}]\,|\Delta t|} \cos(\mathrm{Im}[\tilde{k}]\Delta t) \tag{B-32a}$$

$$K_{x_c x_s}(\Delta t) = \frac{P}{2} e^{-\mathrm{Re}[\tilde{k}]\,|\Delta t|} \sin(\mathrm{Im}[k]\Delta t) \tag{B-33b}$$

The spectrum of the complex process follows easily as

$$S_{\tilde{x}}(\omega) = \frac{2\mathrm{Re}[k]P}{(\omega-\mathrm{Im}[k])^2 + \mathrm{Re}[k]^2} \qquad \text{(B-34)}$$

The spectra and cross spectra of $x_c(t)$ and $x_s(t)$ may easily be found in terms of the even and odd parts of $S_{\tilde{x}}(\omega)$.

From Eq. B-34, we see that in the stationary case, the net effect of allowing a complex gain is that the spectrum has a frequency shift equal to the imaginary part of the gain. In a narrow band interpretation this would correspond to a mean Doppler shift about the carrier. In general, we would not expect such a simple interpretation of the effect of a complex state representation. This suggests that we should consider a second example where we have a second order system and two feedback gains.

Example 2

In this example we want to analyze a second order system. In the steady state, it corresponds to a system with two poles. Note that the pole locations $-k_1$ and $-k_2$ need not be complex conjugates of one another. Again, we shall analyze the stationary case, the analysis for the non-stationary case is straightforward, but is not especially informative The state and observation equations for this system are

$$\frac{d}{dt}\begin{bmatrix}\tilde{x}_1(t)\\ \tilde{x}_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\tilde{k}_1\tilde{k}_2 & -(\tilde{k}_1+\tilde{k}_2)\end{bmatrix}\begin{bmatrix}\tilde{x}_1(t)\\ \tilde{x}_2(t)\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}\tilde{u}(t) \qquad \text{(B-35)}$$

and

$$\tilde{y}(t) = [\,1:0\,] \begin{bmatrix} \tilde{x}_1(t) \\ \tilde{x}_2(t) \end{bmatrix} \tag{B-36}$$

The covariance of the driving noise is

$$E[\,\tilde{u}(t)\tilde{u}^*(\tau)] = 2P\Big(\mathrm{Re}[\tilde{k}_1\tilde{k}_2]\,\mathrm{Re}[\tilde{k}_1\tilde{k}_2] + \frac{\mathrm{Im}[\tilde{k}_1\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1\tilde{k}_2]}\,[\,\mathrm{Im}(\tilde{k}_1\tilde{k}_2)\mathrm{Re}(\tilde{k}_1+\tilde{k}_2) - \mathrm{Im}(\tilde{k}_1+\tilde{k}_2)]\Big)\delta(t-\tau) \tag{B-37}$$

From the steady state solution to Eq. B-12, we have

$$\tilde{P}_o = P \left[\begin{array}{c|c} 1 & \dfrac{j\mathrm{Im}[\tilde{k}_1\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]} \\ \hline \dfrac{-j\mathrm{Im}[\tilde{k}_1\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]} & \dfrac{(\mathrm{Re}(\tilde{k}_1\tilde{k}_2)\mathrm{Re}[\tilde{k}_1+\tilde{k}_2] + \mathrm{Im}(\tilde{k}_1\tilde{k}_2)\mathrm{Im}(\tilde{k}_1+\tilde{k}_2))}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]} \end{array}\right] \tag{B-38}$$

Therefore, we have a stationary process $\tilde{y}(t)$ with power P.

In order to find the covariance matrix we need to find the transition matrix for the system. We do this by using matrix Laplace transform techniques. We find.

$$\tilde{\theta}(t,\tau) = \frac{1}{\tilde{k}_2 - \tilde{k}_1} \left[\begin{array}{c|c} [\tilde{k}_2 e^{-\tilde{k}_1(t-\tau)} - k_1 e^{-\tilde{k}_2(t-\tau)}] & e^{-\tilde{k}_1(t-\tau)} - e^{-\tilde{k}_2(t-\tau)} \\ \hline -(\tilde{k}_1\tilde{k}_2)[e^{-\tilde{k}_1(t-\tau)} - e^{-\tilde{k}_2(t-\tau)}] & -[\tilde{k}_1 e^{-\tilde{k}_1(t-\tau)} - \tilde{k}_2 e^{-\tilde{k}_2(t-\tau)}] \end{array}\right]$$

(B-39)

If we substitute Eqs. B-38 and B-39 in Eq. B-13 and then use Eq. B-17, we obtain

$$\tilde{K}_y(\Delta t) = \begin{cases} \frac{P}{(\tilde{k}_2 - \tilde{k}_1)^*}(\tilde{k}_2^* e^{\tilde{k}_1^*\Delta_1} - \tilde{k}_1^* e^{\tilde{k}_2^*\Delta t}) + \frac{\mathrm{Im}[\tilde{k}_1\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]}(e^{k_1^*\Delta t} - e^{k_2^*\Delta t}) & \Delta t < 0 \\ \frac{P}{(\tilde{k}_2 - \tilde{k}_1)}(\tilde{k}_2 e^{-\tilde{k}_1\Delta_t} - \tilde{k}_1 e^{-\tilde{k}_2\Delta t}) - \frac{j\mathrm{Im}[\tilde{k}_1\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]}(e^{-\tilde{k}_1\Delta_t} - e^{-\tilde{k}_2\Delta t}) & \Delta > 0 \end{cases}$$

(B-40)

We now want to determine the spectrum $S_y(\omega)$ from Eq. B-40. Let us define two coefficients for convenience:

$$\tilde{A}_1 = \frac{\tilde{k}_2 + \tilde{k}_1^*}{(\tilde{k}_2 - \tilde{k}_1)} \cdot \frac{\mathrm{Re}[\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]} \qquad \text{(B-41)}$$

$$\tilde{A}_2 = \frac{\tilde{k}_2^* + \tilde{k}_1}{(\tilde{k}_2 - \tilde{k}_1)} \cdot \frac{\mathrm{Re}[\tilde{k}_2]}{\mathrm{Re}[\tilde{k}_1+\tilde{k}_2]} \qquad \text{(B-42)}$$

After a fair amount of manipulation, we can compute $S_y(\omega)$.

$$S_y(\omega) = 2P \left(\frac{\mathrm{Re}[\tilde{A}_1]\,\mathrm{Re}[\tilde{k}_1] + \mathrm{Im}[\tilde{A}_1](\omega+\mathrm{Im}[\tilde{k}_1])}{\mathrm{Re}^2[\tilde{k}_1] + (\omega+\mathrm{Im}[\tilde{k}_1])^2} + \right.$$

$$\left. \frac{\mathrm{Re}[\tilde{A}_2]\,\mathrm{Re}[\tilde{k}_2] + \mathrm{Im}[\tilde{A}_2](\omega+\mathrm{Im}[\tilde{k}_2])}{\mathrm{Re}^2[\tilde{k}_2] + (\omega+\mathrm{Im}[\tilde{k}_2])^2} \right) . \quad \text{(B-43)}$$

We have plotted this function for various values of $\tilde{k}_1$ and $\tilde{k}_2$ in Figure B-1 through B-4. The values of $\tilde{k}_1$ and $\tilde{k}_2$ for a particular figure are illustrated on the figures by the pole location they produce, i.e., the system has a pole at $-\tilde{k}_1$ and $-\tilde{k}_2$.

In Figures B-1 and B-2 we illustrate that by simply choosing the poles as complex conjugates we can produce either spectrum which is either very flat near $\omega = 0$ or is peaked with two symmetric lobes. If one wanted to use real state variables to generate this spectrum, a fourth order system would be required. In this case, the complex notation has significantly reduced the computation required.

Figure B-3 illustrates an interesting observation about mean Doppler shifts. Let us draw the pole-zero locations for the complex system. If there exists a line $\omega = \omega_c$ about which the pole-zero pattern is symmetric, then in the stationary case the complex notation effectively produces a spectrum which is symmetric about $\omega = \omega_c$. For example, in Figure B-3, the pole pattern is symmetric about $\omega = 1/2$. We see that the spectrum is symmetric about

$\omega = -1/2$ also. Consequently, we can use the complex notation to introduce mean Doppler shifts of narrow band processes.

Figure B-4 illustrates that we can obtain spectra which are not symmetric about any axis. This is a relatively important case. In dealing with narrow band processes if one can find a frequency about which the spectrum is symmetric, then the component processes $y_c(t)$ and $y_s(t)$ are uncorrelated. However, if there is no axis of symmetry (or if the choice of carrier is not at our disposal) then the components are definitely correlated. This example shows that we can model narrow band processes with non-symmetric spectra very conveniently with our complex state variable notation.

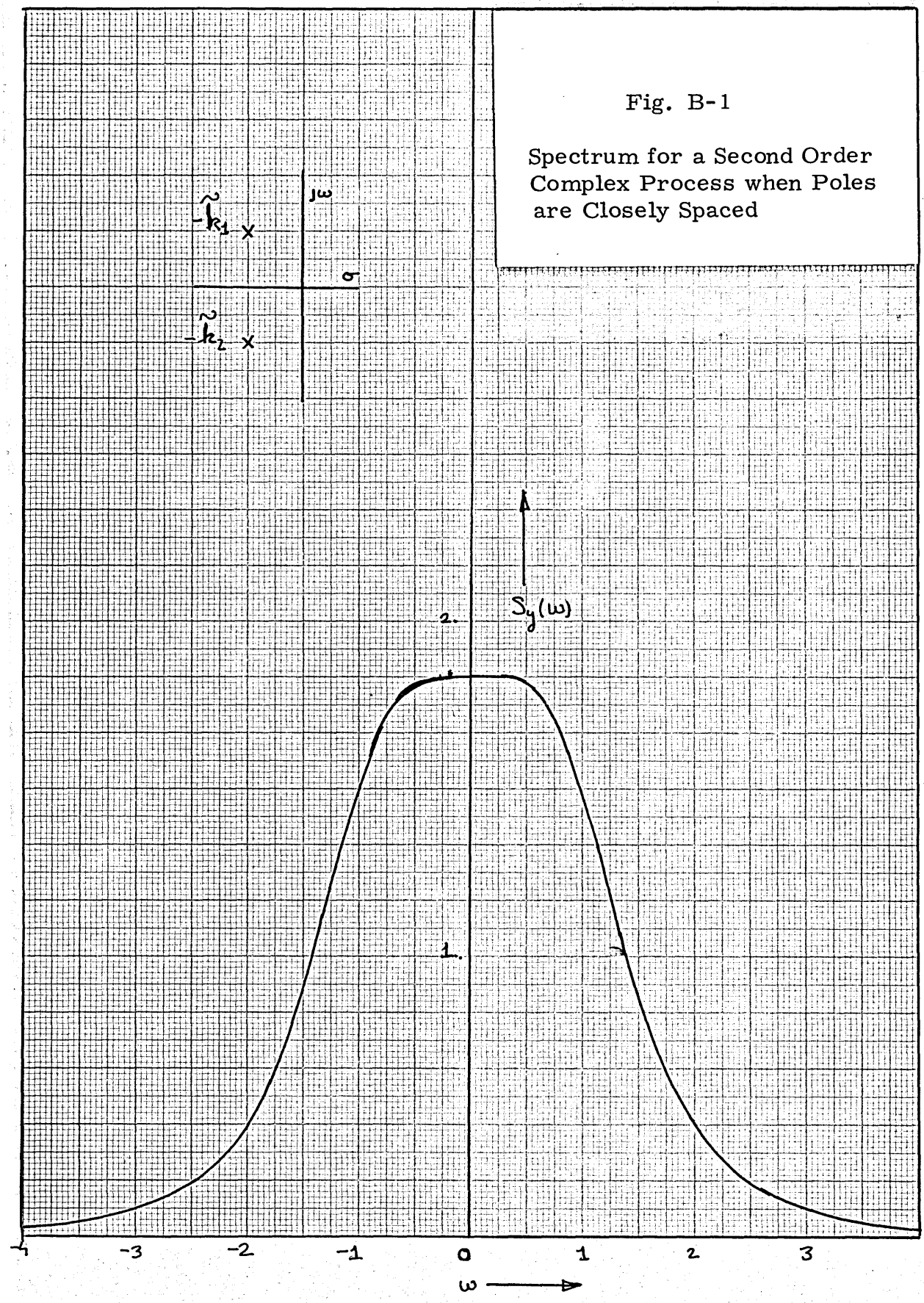


Fig. B-1

Spectrum for a Second Order Complex Process when Poles are Closely Spaced

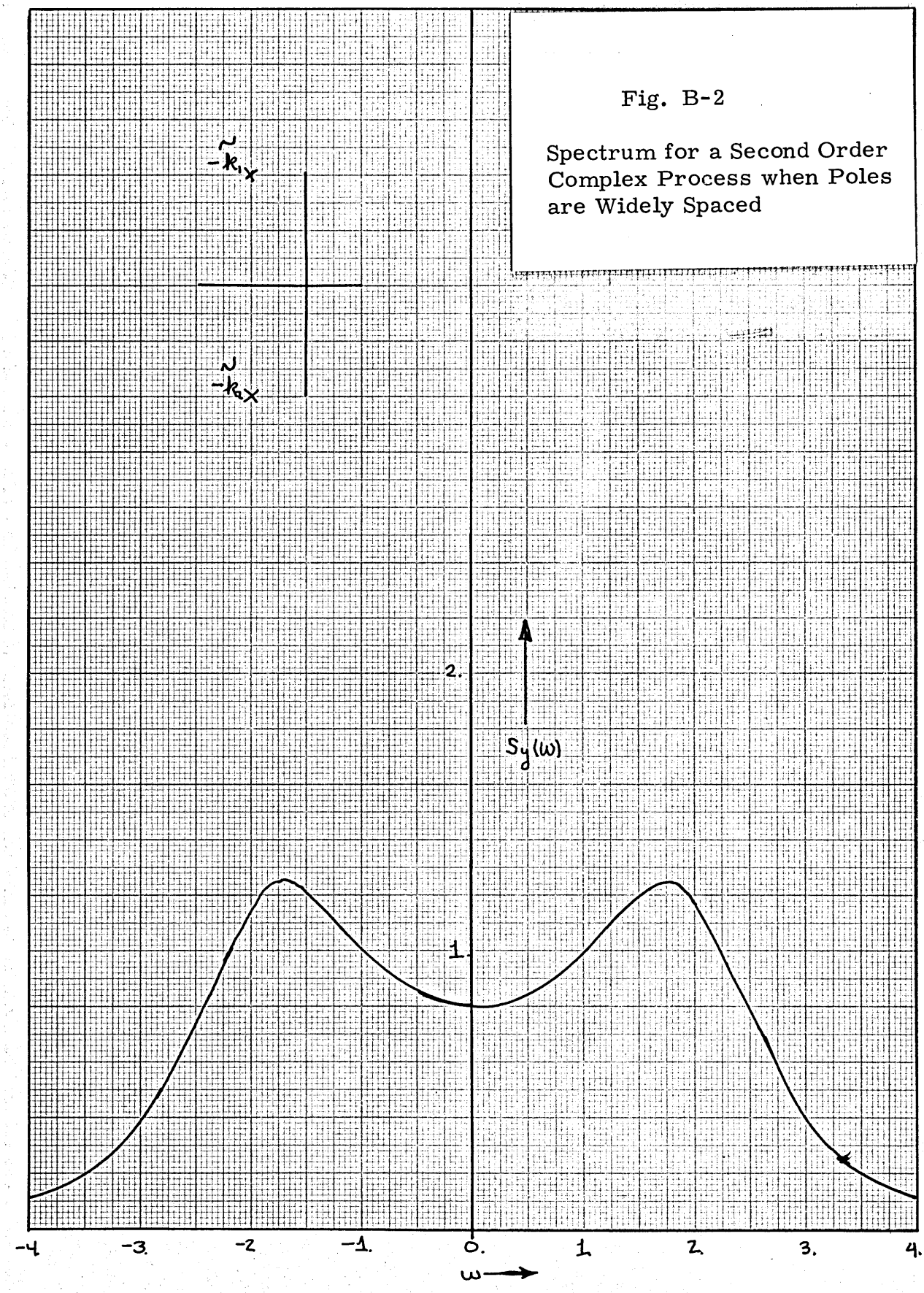


Fig. B-2

Spectrum for a Second Order Complex Process when Poles are Widely Spaced

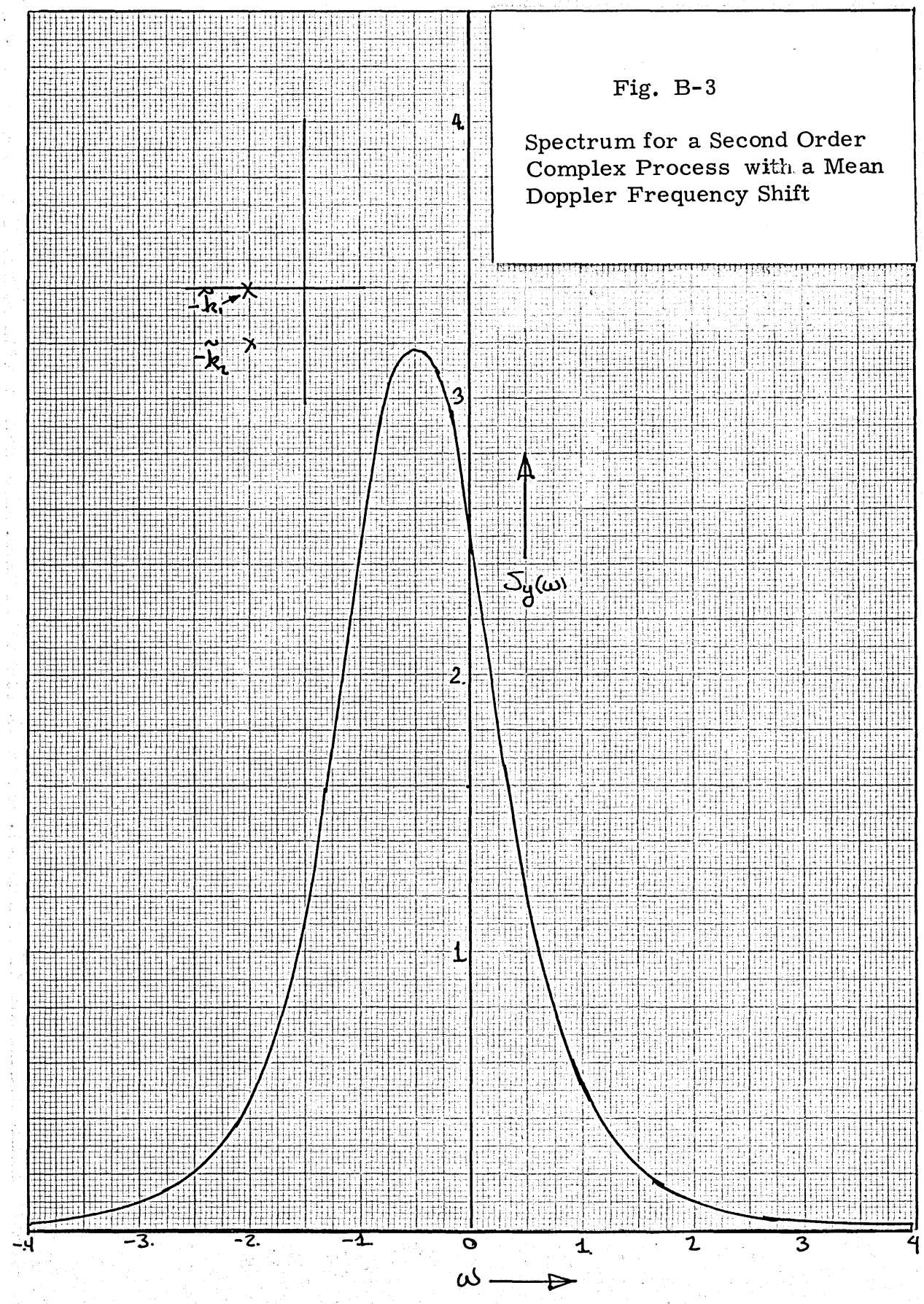


Fig. B-3

Spectrum for a Second Order Complex Process with a Mean Doppler Frequency Shift

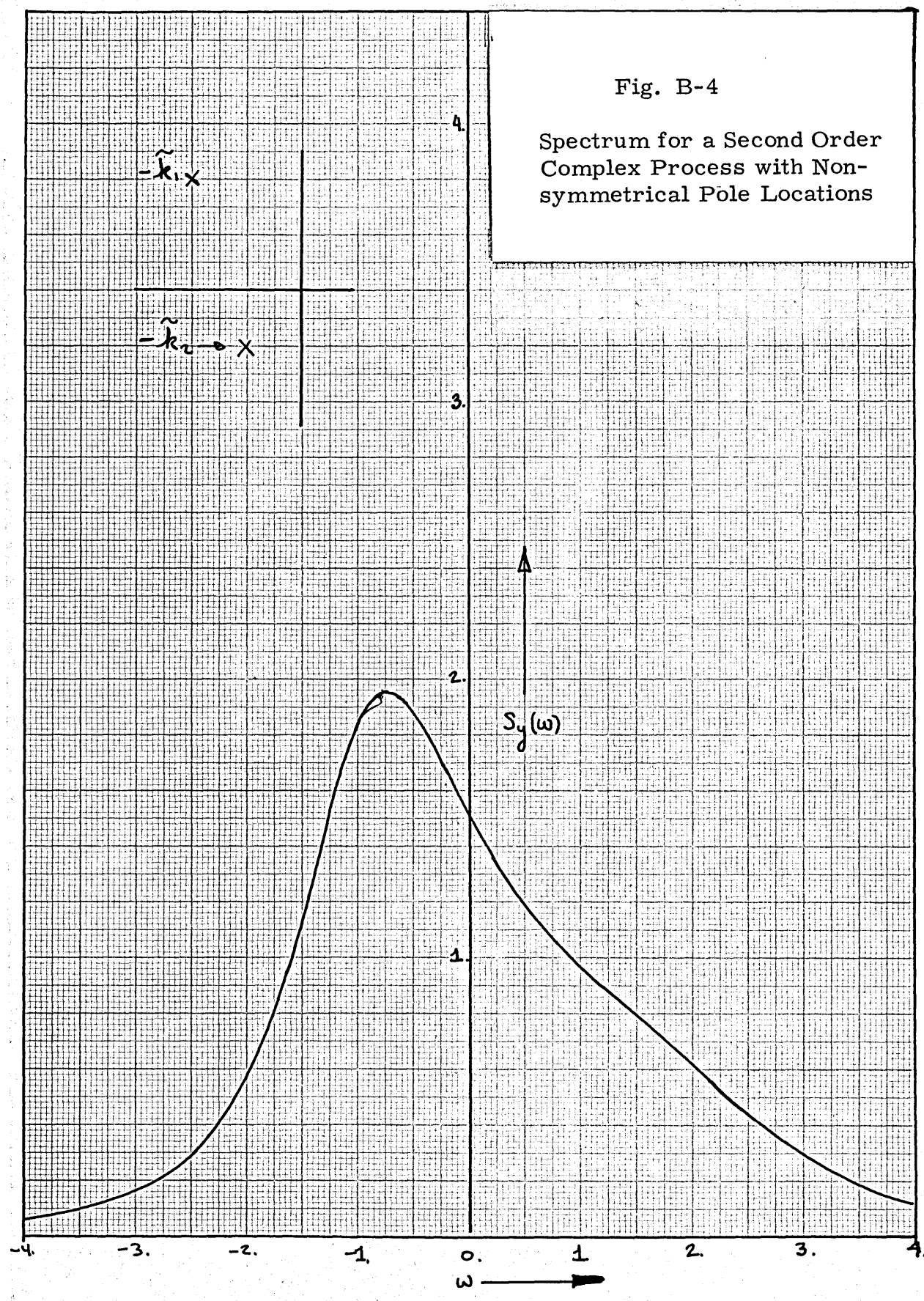


Fig. B-4

Spectrum for a Second Order Complex Process with Non-symmetrical Pole Locations

REFERENCES

1. R. Courant and D. Hilbert, Methods of Mathematical Physics, Vol. I, Interscience Publishers, New York, 1953.

2. W. V. Lovitt, Linear Integral Equations, McGraw-Hill, New York, 1924.

3. H. L. Van Trees, Detection, Estimation and Modulation Theory, Part I, John Wiley and Sons, New York, 1968.

4. C. W. Helstrom, Statistical Theory of Signal Detection, Pergamon, London, 1960.

5. P. M. DeRusso, R. J, Roy, and C. M. Close, State Variables for Engineers, John Wiley and Sons, New York, 1965.

6. M. Athans and P. L. Falb, Optimal Control, McGraw-Hill, New York, 1966.

7. L. Zadeh and C. Dosoer, Linear System Theory, McGraw-Hill, New York, 1963.

8. R. E. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction Theory," A. S. M. E. Journal of Basic Engineering, Vol. 83, March 1961, 95-108.

9. F. Schweppe, "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Transactions on Information Theory, Vol. IT-11, No. 1, January 1965, pp. 61-70.

10. D. Snyder, "The State Variable Approach to Continuous Estimation," M. I. T., Ph. D, Thesis, February 1966.

11. W. Davenport and W. Root, Random Signals and Noise, McGraw-Hill, New York, 1958.

12. H. L. Van Trees, Detection, Estimation, and Modulation Theory, Part II, John Wiley and Sons, (to appear).

13. L. S. Pontryagin, V. G. Boltyanski, R. Gamkrelidze, and E. Mishenko, The Mathematical Theory of Optimal Processes, Interscience Publishers, New York, 1962.

14. A. E. Bryson and M. Frazier, "Smoothing for Linear and Nonlinear Dynamic Systems", Proceedings Optimum Systems Conference, Wright - Patterson Air Force Base, Sept. 1962.

15. H. Rauch, F. Tung, and C. Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems", AIAA Journal, Vol. 3, No 8, August 1966.

16. A. Baggeroer, "Maximum Aposteriori Interval Estimation," 1966 WESCON Convention Record, Paper 7-3, August 1966.

17. D. M. Detchmendy and Sridar, "Sequential Estimation of States and Parameters in Noisy Dynamical Systems," A. S. M. E. Journal of Basic Engineering, June 1966, pp. 362-368.

18. E. J. Kelley and W. L. Root, "A Representation of Vector-Valued Random Processes," Group Report 55-21, (revised), M. I. T. Lincoln Laboratory, April 22, 1960.

19. J. Capon, "Asymptotic Eigenfunctions and Eigenvalues of a Homogeneous Integral Equation," IRE Transactions on Information Theory, Vol. IT-8, No. 1, Jan. 1962, pp. .

20. R. Y. Huang and R. A. Johnson, "Information Transmission with Time Continuous Processes," IEEE Transactions on Information Theory, Vol. IT-9, No. 2, April 1963, pp. 84-95.

21. E. T. Whitaker and G. N. Watson, A Course of Modern Analysis, Cambridge University Press, Cambridge, England, 1962.

22. L. D. Collins, "Evaluations of the Fredholm Determinant for State Variable Covariance Functions," accepted for publication in IEEE Proceedings, Decenber 1967.

23. M. Athans and F. C. Schweppe, "Gradient Matrices and Matrix Calculations," Technical Note 1965-53, M. I. T. Lincoln Laboratory, 1965.

24. L. D. Collins, "A Simplified Derivation of the Differential Equation Solution to Fredholm Integral Equations of the Second Kind," Internal Memo No. IM-LDC-19, Detection and Estimation Theory Group, Resaerch Laboratory of Electronics, M. I. T., Cambridge, Mass., January 1967.

25. L. A. Zadeh and K. S. Miller, "Solution to an Integral Equation Occuring in the Theories of Prediction and Detection," IRE Transactions in Information Theory, Vol. IT-2, No. 3, June 1956, pp. 72-76.

26. D. Youla, "The Solution of a Homogeneous Wiener -Hopf Integral Equation Occuring In the Expansion of Second Order Stationary Random Functions," IRE Transactions on Information Theory, Vol. IT-3, No. 4, Sept. 1957, pp. 187-193.

27. C. W. Helstrom, "Solution of the Detection Integral Equation for Stationary Filtered White Noise," IEEE Transactions on Information Theory, Vol. IT-11, No. 3, July 1965, pp. 335-339.

28. H. L. Van Trees, Private communication.

29. H. J. Kushner and F. C. Schweppe, "A Maximum Principle for Stochastic Control Systems," Journal of Mathematical Analysis and Applications, Vol. 8, pp. 287-302, 1964.

30. M. Athans, "Solution of the Matrix Equation $X(t) = A(t)X(t) + X(t)B(t) + U(t)$," Technical Note 1965-26, M. I. T. Lincoln Laboratory, June 1965

31. J. S. Meditch, "On Optimal Linear Smoothing Theory," Technical Report 67-105, Information Processing and Control System Laboratory, Northwestern University, March 1967.

32. A. Baggeroer, "Fortran Programs for a) Solving Fredholm Integral Equations, b) Optimal Signal Design For Additive Colored Noise Channels, c) Calculating Smoother and Filler with Delay Covariances of Error," Internal Memo, Detection and Estimation Theory Group, Research Laboratory for Electronics, M. I. T. (to appear).

33. R. Bellman, R. Kalaba, and R. Sridhar, "Sensitivity Analysis and Invariant Imbedding," Research Memorandum, No. 4039-PR, March 1964.

34. P. R. Garabedian, Partial Differential Equations, John Wiley and Sons, New York, 1964

35. A. C. Butkovskii and A. Y. Lerner, "The Optimal Control of Systems with Distributed Parameters," Automatika i Telemekhanika, Vol. 21, No. 6, June 1960, pp. 682-692

36. A. G. Butkovskii, "The Maximum Principle for Optimum Systems with Distributed Parameters," Automatika i Telemekhanika,, Vol. 22, No. 9, September 1961, pp. 1288-1301

37. A. Butkovskii, "Optimum Processes in Systems With Distributed Parameters," Automatika i Telemekhanika, Vol. 22, No. 1 pp. 17-25, January, 1961.

38. A. Butkovskii, "Some Approximate Methods For Solving Problems of Optimal Control of Distributed Parameter Systems," Automatika i Telemeknanika, Vol. 22, No. 12, pp. 1565-1575, December, 1961.

39. Courant and Hilbert, Methods of Mathematical Physics, Vol. II, Interscience Publishers, New York, 1962.

40. P. K. C. Wang, "Control of Distributed Parameter Systems,"in Advances in Control Systems, C. T. Leondes (Editor), Academic Press, New York, 1964.

41. P. K. C. Wang and F. Tung, "Optimum Control of Distributed Parameter Systems," ASME Journal of Basic Engineering, pp. 67-79, March 1964.

42. Y. Sakawa, "Solution of an Optimal Control Problem in a Distributed Parameter System," IEEE Transactions on Automatic Control, October 1964, pp. 420-470.

43. R. Goodson, "Optimal Control of Systems with Distributed Parameters," JACC Proceedings, June 1964, pp. 390-397.

44. P. Pierre, "Minimum Mean Square Error Design of Distributed Parameter Control Systems," JACC Proceedings, June 1964, pp. 381-389.

45. M. Lasso, "The Modal Analysis and Synthesis of Linear Distributed Control Systems," Ph. D. Thesis, Department of Electrical Engineering, M. I. T. , Camridge, Massachusetts, June 1965.

46. H. Kwakernaak, "Optimum Filtering in Linear Systems with Time Delays," IEEE Transactions on Automatic Control, Vol. AC-12, No. 2, April 1967, pp. 169-173.

BIOGRAPHY

Arthur B. Baggeroer was born in Weymouth, Massachusetts on June 9, 1942. He graduated from Wakefield Memorial High School, Wakefield, Massachusetts. He received the B. S. E. E. degree from Purdue University in 1963, the S. M. and E. E. degrees from M. I. T. in 1965 and the Sc. D. degree in 1968. While at M. I. T. he was a National Science Foundation Graduate Fellow from 1963 to 1966. Currently, he is an Instructor in Electrical Engineering at M. I. T.

He spent the summers of 1959 to 1965 working for the Mitre Corporation and the summers of 1966 and 1967 at M. I. T.'s Lincoln Laboratory. He has presented papers at the 1966 WESCON and the 1967 International Conference on Communication Technology in the area of the application of state variable techniques to communication problems. He is a member of the Tau Beta Pi, Eta Kappa Nu, Sigma Xi and IEEE.

He is married to the former Carol Coleman , has two children, and currently lives in Watertown, Massachusetts.