SELF-ADAPTIVE OPTIMAL FILTER DESIGN FOR STATISTICALLY-DESCRIBABLE SIGNALS

by

Robert Eugene Nabours

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1965

THE UNIVERSITY OF ARIZONA

GRADUATE COLLEGE

I hereby recommend that this dissertation prepared under my direction by Robert Eugene Nabours

entitled Self-Adaptive Optimal Filter Design for Statistically Describable Signals

be accepted as fulfilling the dissertation requirement of the degree of Doctor of Philosophy

D. G. Schultz — April 14, 1965

Dissertation Director — Date

After inspection of the dissertation, the following members of the Final Examination Committee concur in its approval and recommend its acceptance:*

E R Hansenbauer	14 April 1965
Frederick C. Hill	14 April 1965
Thomas Regan	14 April 1965
Robert M. Wright	15 April 1965

*This approval and acceptance is contingent on the candidate's adequate performance and defense of this dissertation at the final oral examination. The inclusion of this sheet bound into the library copy of the dissertation is evidence of satisfactory performance at the final examination.

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Robert E. Mahoney

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Dr. Donald G. Schultz and Dr. Aladdin N. Perkins for their guidance and inspiration throughout this study. The assistance and constructive criticism of the author's colleagues is also gratefully acknowledged.

This research was carried out while the author was employed at the Tucson Research Laboratory of Textron's Bell Aerosystems Company. The author would like to express his appreciation to the Management of the Tucson Research Laboratory for their financial support in the pursuance of the goals set by this work.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS--Continued

ABSTRACT

A theory is presented for the design of time-varying filters which yield minimal mean square estimates of the incoming signal for the case of nonstationary, additive, uncorrelated signal plus noise inputs. It is shown how certain aspects of Wiener theory apply to the more general problem and under what conditions the nonstationary design is optimum.

An engineering solution to the problem is obtained based upon an assumption of "quasi-stationarity" of the short-time power spectral densities of signal and noise. A recent control system approach is used to simplify the analytical solution of a set of equivalent Wiener problems, along with a linearization of the input power spectral density function by the asymptotic Bode technique which yields a measurement procedure for determining the nonstationary channel statistics. An analysis is performed to place in evidence the channel identification time requirements for optimum adaptive action of the time-varying filter.

Two examples are simulated on an analog computer to both verify the theory and illustrate the hardware complexity that arises in practical realizations of adaptive filters. It is demonstrated by theory and experiment that significant gains in filter output signal-to-noise ratio can be achieved through the application of the adaptive concepts to time-varying communication channels.

A practical example of an empirical nature is given to illustrate the design of an optimum adaptive phase-locked loop detector for the Echo II, FM passive satellite, communication system. A theoretical improvement in information transfer of approximately 27 times is demonstrated for the satellite system when a closed-loop combination of adaptive receiver-variable rate transmitter is employed.

Chapter 1

INTRODUCTION

1.1 PROLOGUE

In the most general terms, a filter is a device for separating or reducing certain entities in a system of heterogeneous collection. The heterogeneous collection which is of interest in electrical communications is composed of signal and noise, where the two real time processes are distinguished as the desired or information bearing electrical fluctuations (signal) and the undesired or disturbing electrical fluctuations (noise).

Information transfer is accomplished in a communications system when the signal consists of random electrical fluctuations and when these time variations are capable of being separated, to some extent, from the noise.

Figure 1.1-1 illustrates a general open-loop adaptive communications system in block diagram form.

This work deals with the receiver although it is impossible to ignore the effects of the other elements of the system; in particular the behavior of the physical communication media and to a somewhat lesser degree the form of the encoder, decoder, and information source and sink are contributing factors to the receiver design.

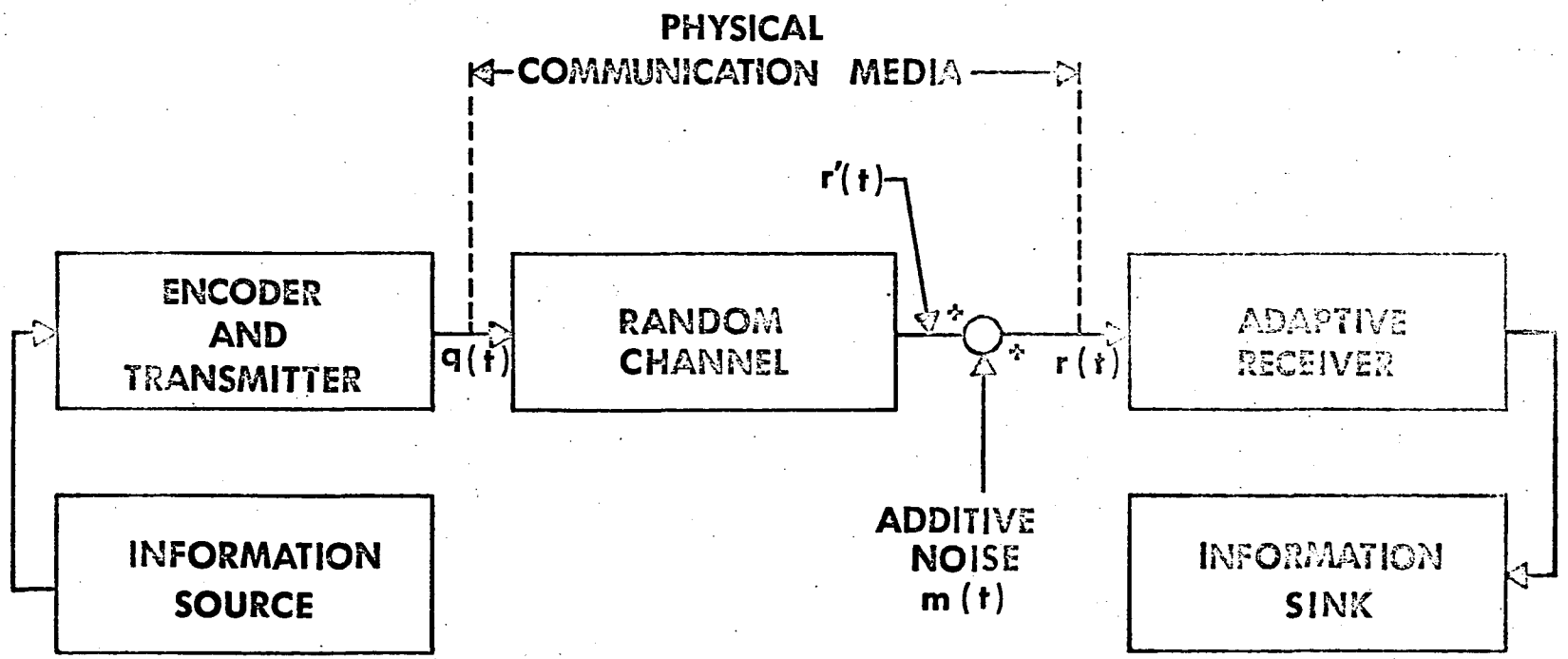


Figure 1.1-1. General Open-Loop Adaptive Communication System

Fundamentally, it is the function of the receiver to transform the electromagnetic energy incident at its input into another energy form or state. The output from the receiver must be in a form which is acceptable to the information sink, such as an audio or visual presentation if the sink is a human, or an acceptable range of voltage variations with time if the sink is an electronic computer. In any event, the receiver is faced with the problem of filtering (general sense) the incoming real time random process in order to separate the signal from the noise. The filtering operation may take place directly in the domain of the received energy form, i.e., a radio frequency (rf) filter, or it may occur after decoding (demodulation) of the high frequency signal into a low frequency or base-band variation. In practice both techniques are used, since they yield additive improvements in signal separation from noise.

Independently and almost simultaneously in the early 1940's, Norbert Wiener [1] and A. N. Kolmogorov [2] developed a theory for the solution to the problem of optimum (in a minimal mean square error sense) filter design for the separation of signal from noise when both are stationary, additive, stochastic processes.

However, when either or both the signal and noise are nonstationary, then the optimal filtering problem cannot be solved directly by the Wiener-Kolmogorov theory.

Unfortunately, it is usually the case for physical communications media, that neither signal nor noise possess stationary statistical descriptors, such as power spectral densities, auto-correlation functions, or even means. If one adopts the viewpoint that all continuous but nonstationary stochastic processes may be represented as discrete random functions of time, then it follows that a filter structure which is optimum in the Wiener sense at one instant of time will not, in general, be of optimum structure at some later time displacement. Thus, if an optimum filter exists for nonstationary stochastic processes, it must of necessity have a time-varying structure. Only if such a structure can vary deterministically as a function of the time-varying statistical descriptors of the stochastic processes will such a filter be, in the Wiener sense, optimum.

The objective of this work is to present a theory for the design of minimal mean square error filters for nonstationary, additive, uncorrelated, signal and noise. It is shown that certain aspects of the Wiener-Kolmogorov theory carry over to the more general problem and that, under certain qualifying conditions, the nonstationary design of such filters follows directly the Wiener techniques. Of more than cursory interest to communications engineers should be the results of Chapter 4, wherein the simulation of two particular cases of time-varying adaptive filters is presented along with the empirical performance of these filters.

Obviously, as is the case in most engineering problems, certain approximations and assumptions are required in order to effect a useful solution. Wherever these occur the author has called them to the attention of the reader, and note is made of any important consequences which are inherent in such departures from the exactness of the mathematics.

The fundamental contribution of this work is the reduction of theory to practice wherein the important factor of engineering judgement is exercised in order to simplify the design procedures identified by the mathematics.

1.2 CONTENTS

The material to follow has been arranged in the following sequence: (1) a survey of the existing theory and approaches to the problem, (2) a statement of the approach to be used in this work, (3) the application of this approach to theoretical and empirical examples, and (4) analog computer realizations to experimentally verify the adaptive filter theory presented.

The history of the problem is treated in Chapter 2 by presenting the general mathematical framework in which the problem lies and by discussing briefly the approaches to a solution as used by previous investigators. The notation used is presented in the early part of Chapter 2.

The basis for the approach used to effect an engineering solution to the problem is set down in Chapter 3. A necessity for certain assumptions and approximations is demonstrated, and these assumptions and approximations are treated in detail to establish their merit and effectiveness in the context of this investigation. The proposed design method is outlined in the last section of Chapter 3.

Three examples of adaptive filter design and an analog computer simulation of the first two of these examples are presented in Chapter 4. A comparison is made between the theoretical and experimental performances of the simulated examples. A filter performance measure, which indicates the improvement in filter action of the adaptive structure in relation to the fixed optimal filter, is defined and the experimental behavior of this measure with nonstationary inputs is presented. The design of an adaptive phase-locked loop FM demodulator is considered in the last section of Chapter 4 in order to give a more practical example of the use of the adaptive theory.

Practical applications of the method are discussed and conclusions drawn to the work in Chapter 5.

Chapter 2

HISTORY OF THE PROBLEM

2.1 INTRODUCTION AND ORGANIZATION OF THE CHAPTER

The background material pertaining to the problem of adaptive optimal filtering of stochastic processes is presented in this chapter. A mathematical formulation of the problem, which dates back to the early 1940's, is presented, and the contributions of investigators in this area since that time are discussed, approximately in their chronological order. Section 2.2 is devoted to a discussion of the notation to be used in the paper.

2.2 NOTATION

The nonstationary and time-varying aspects of the problem to be considered in this work make it necessary to use certain notational forms which, although a part of the modern work in communications and random process theory, are not standardized and consequently require definition here.

"Adaptive Filter." An adaptive filter (or self-adaptive filter, which is slightly redundant) is defined as a passive or active device which automatically adjusts its parameters in accordance with some algorithm dependent upon changes in the input signal statistics.

Lower case letters such as s(t), n(t), and i(t) denote time dependent or stochastic random variables. The stochastic variables considered in this work are all assumed to be real-valued, bounded, continuous or discrete functions of time. This is in agreement with the physical world that generates such entities. In general, the stochastic variables treated are not stationary; i.e., their statistics may be time dependent, and consequently they will not in general satisfy the ergodic hypothesis.[1]

Since the stochastic processes to be treated are non-stationary, the notation used to denote mathematical expectation must differ from that used to denote time average. To this end, the following definitions are employed:

Mathematical Expectation

$$E\{f[s(t)]\} = \int_S f(s)dP(s), \qquad (2\text{-}1)$$

where E is the expectation operator, P(s) is the probability measure over the space S of s(t), and the integral is taken in the Lebesgue-Stieltjes sense.[2] E {f} will be referred to as the expectation of the function f.

---

1. For a discussion of the ergodic hypothesis, see Davenport and Root [3], pp. 66-68.

2. See Korn and Korn [4], p. 113.

Time Average

$$A\{f[s(t)]\} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f[s(t)]\, dt, \qquad (2\text{-}2)$$

where A is the average operator and the integral is taken in the Riemann sense. $A\{f\}$ will be referred to as the average of the function f.

Certain special expectations and averages arise in the study of random processes and these are defined below.

Autocorrelation Function

$$\Phi_{xx}(t_1,t_2) = E\{x(t_1)x(t_2)\} = \overline{x(t_1)x(t_2)}. \qquad (2\text{-}3)$$

Crosscorrelation Function

$$\Phi_{xy}(t_1,t_2) = E\{x(t_1)y(t_2)\} = \overline{x(t_1)y(t_2)}. \qquad (2\text{-}4)$$

Time Autocorrelation Function

$$\varphi_{xx}(t_1,t_2) = A\{x(t_1)x(t_2)\} = \widetilde{x(t_1)x(t_2)}. \qquad (2\text{-}5)$$

Time Crosscorrelation Function

$$\varphi_{xy}(t_1,t_2) = A\{x(t_1)y(t_2)\} = \widetilde{x(t_1)y(t_2)}. \qquad (2\text{-}6)$$

For stationary stochastic processes an extremely valuable relationship exists between the power spectral density associated with the random variable x(t) and its

time autocorrelation function.[1] This relationship, known as the Wiener-Khintchine relation, states that

$$G_{xx}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi_{xx}(\tau)\bar{e}^{j\omega\tau}\, d\tau, \qquad (2\text{-}7)$$

where $G_{xx}(\omega)$ is the power spectral density function associated with the stationary random variable $x(t)$. Physically $|G_{xx}(\omega)|$ represents the power that would be dissipated in a one-ohm resistor if a random voltage or current $x(t)$ were passed through an "ideal" zero bandwidth filter centered at frequency ω.

In order to extend the useful ideas of autocorrelation and power spectral density to nonstationary random variables, the short-time autocorrelation functions and corresponding power density spectra as first used by Fano [5], will be defined as follows:

Short-Time Autocorrelation Function

$$\varphi_{xx}^{t}(\tau,\alpha) = 2\alpha\int_{-\infty}^{t} x(\eta)x(\eta-\tau)\bar{e}^{2\alpha(t-\eta)}d\eta. \qquad (2\text{-}8)$$

Short-Time Power Spectral Density

$$G_{xx}^{t}(\alpha,\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varphi_{xx}^{t}(\tau,\alpha)\bar{e}^{\alpha|\tau|-j\omega\tau}d\tau, \qquad (2\text{-}9)$$

---

1. Note that the time autocorrelation function is a function of only one variable $t = |t_2 - t_1|$, since, for a stationary process, only the relative time displacement determines φ_{xx}, not the initial or final time values t_1 and t_2.

where α is the time constant of a low-pass filter whose impulse response or weighting function is given by

$$h(t) = \begin{cases} 2\alpha\ e^{-2\alpha t}, & t > 0 \\ 0, & t < 0. \end{cases} \tag{2-10}$$

The two-sided Laplace transform proves to be a useful mathematical concept in certain aspects of optimization theory and is defined by the reciprocal relations

$$F(s) = \int_{-\infty}^{\infty} f(t) e^{-st} dt, \tag{2-11}$$

$$f(t) = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} F(s)\ e^{st} ds, \tag{2-12}$$

where $s = \sigma + j\omega$ is the so-called complex frequency. The absence of an abscissa of absolute convergence in the inverse transform (2-12), restricts the transform to functions which satisfy the convergence relations

$$\int_0^{\infty} |f(t)|^2\ dt < \infty,$$
$$-\int_0^{-\infty} |f(t)|^2\ dt < \infty. \tag{2-13}$$

2.3 MATHEMATICAL FORMULATION

With reference to figure 2.3-1, the following definitions are made for the purpose of mathematical modeling:

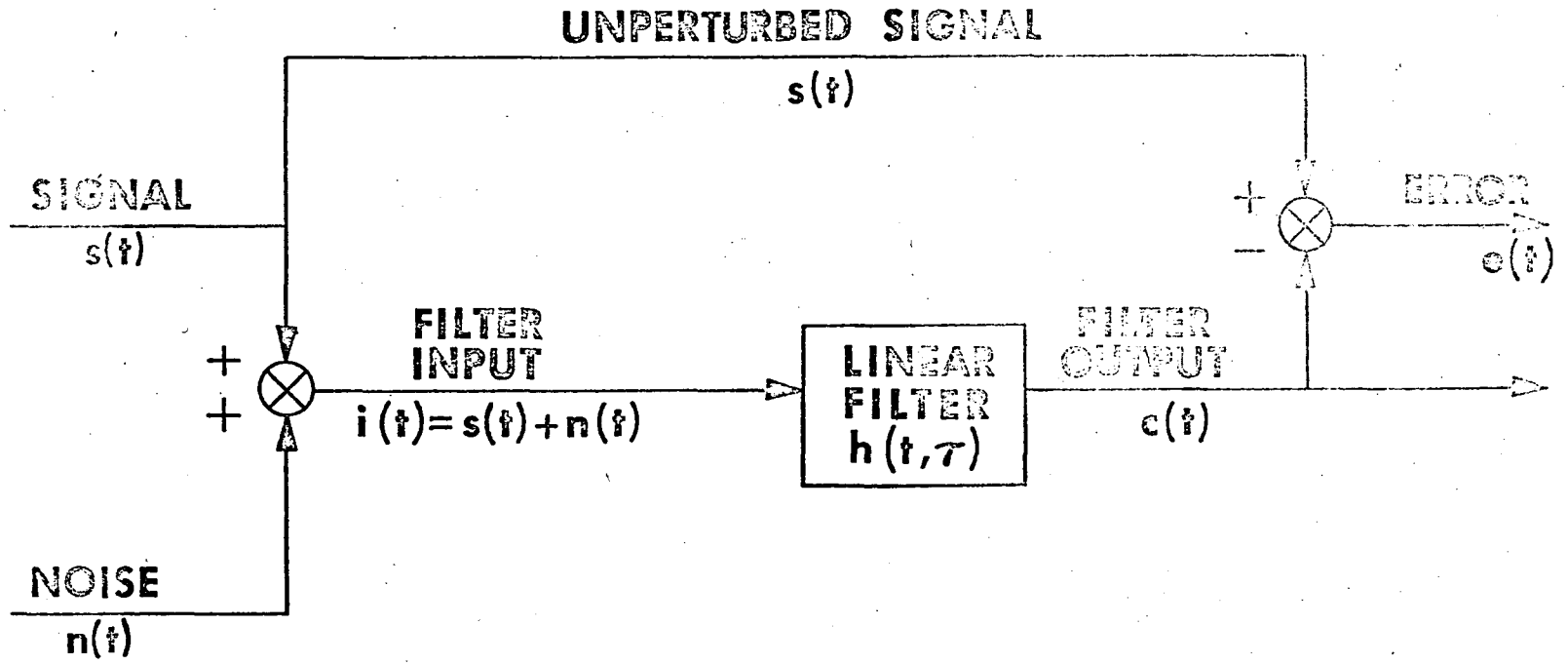


Figure 2.3-1. Block Diagram Defining Symbols

The signal s(t) and noise n(t) are assumed to be stochastically independent, continuous or discrete, random variables possessing certain time dependent statistical descriptors.

The impulse-response function or weighting function of the linear filter, $h(t,\tau)$ is characterized by two variables, t and τ, for the general time-varying linear filter, where $h(t,\tau)$ represents the output of the filter at time t due to an impulse input at time τ.

The symbol e(t) is the filter error, i.e., the difference between actual output c(t) and the desired output s(t).

The output c(t) may be related to the filter input i(t) by the superposition integral,

$$c(t) = \int_{-\infty}^{\infty} h(t,\tau)\, i(t-\tau)\, d\tau = \int_{0}^{t} h(t,\tau)[s(t-\tau) + n(t-\tau)]\, d\tau, \tag{2-14}$$

where the upper limit ∞ in the first integral may be replaced by t in the second integral, since, for physical realizability, $h(t,\tau) \equiv 0$ for $\tau > t$. The lower limit $-\infty$ in the first integral may be replaced by 0 in the second integral if we define $h(t,\tau) = 0$ for $\tau < 0$.

As a consequence of the definition of filter error, figure 2.3-1 and equation (2-14), the error may be written as a function of the input random process as follows:

$$e(t) = s(t) - \int_0^t h(t,\tau)[s(t-\tau) + n(t-\tau)]d\tau. \qquad (2\text{-}15)$$

In most communication signals it is not the instantaneous error $e(t)$ which is of interest, but rather the mean square of $e(t)$, which is the criterion used here. This is denoted by $\overline{e^2(t)} = mse$, where the bar indicates that averaging is to be performed with respect to the ensemble of input functions. A discussion of other error criteria such as $|e(t)|$ and $\overline{e(t)}$ is not presented or considered here, but rather note is made of the correspondence between minimal mse and maximum output signal-to-noise power ratio. The reader is referred to Knapp [6] for a thorough discussion of error criteria.

The squared instantaneous error yields

$$e^2(t) = s^2(t) - 2\int_0^t h(t,\tau)s(t)[s(t-\tau) + n(t-\tau)]d\tau + \int_0^t\int_0^t h(t,\tau_1)h(t,\tau_2)[s(t-\tau_1) + n(t-\tau_1)][s(t-\tau_2) + n(t-\tau_2)]d\tau_1 d\tau_2. \qquad (2\text{-}16)$$

The auto- and cross-correlation functions for the various stochastic processes are defined as follows:

$$\Phi_{ii}(t_1,t_2) = \overline{i(t_1)i(t_2)} = \overline{[s(t_1) + n(t_1)][s(t_2) + n(t_2)]}$$

$$\Phi_{is}(t_1,t_2) = \overline{i(t_1)s(t_2)} = \overline{[s(t_1) + n(t_1)]s(t_2)}$$

$$\Phi_{ss}(t_1,t_2) = \overline{s(t_1)s(t_2)} \qquad (2\text{-}17)$$

$$\Phi_{nn}(t_1,t_2) = \overline{n(t_1)n(t_2)}$$

$$\Phi_{sn}(t_1,t_2) = \overline{s(t_1)n(t_2)}$$

where the bar indicates that ensemble averaging, (2-3) and (2-4), is to be applied. The mse may now be expressed in terms of certain correlation functions as

$$\overline{e^2(t)} = \Phi_{ss}(0) - 2\int_0^t h(t,\tau)\Phi_{si}(t,t-\tau)d\tau + \int_0^t\int_0^t h(t,\tau_1)h(t,\tau_2)\Phi_{ii}(t-\tau_1,t-\tau_2)d\tau_1 d\tau_2, \qquad (2\text{-}18)$$

where the operation of ensemble averaging is assumed to commute with that of integration.

As was demonstrated by Booton [7], using the calculus of variations technique, the mse is minimized with respect to all possible choices of the system weighting function $h(t,\tau)$, if and only if, $h(t,\tau) \equiv h_o(t,\tau)$, where $h_o(t,\tau)$ is related to the system correlation functions as follows:

$$\Phi_{si}(t-\tau, t) = \int_0^t h_o(t,\eta)\Phi_{ii}(t-\tau, t-\eta)\,d\eta, \qquad (2\text{-}19)$$

for $0 < \tau < \infty$.

2.4 CONTRIBUTIONS TO THE THEORY

Unfortunately (2-19) cannot be solved in general for $h_o(t,\tau)$ explicitly in terms of Φ_{si} and Φ_{ii} since (2-19) is the general form of the integral equation of the first kind (see references [7] and [8]). Shinbrot [8] was able to obtain a solution for $h_o(t,\tau)$ explicitly in terms of Φ_{si} and Φ_{ii} for the special case of white noise which is stochastically independent of the signal. Even under this special case the solution was obtained in the form of a finite summation of products of functions of t and functions of τ as an approximation to the general infinite summation required for an exact solution. For non-white noise the only tractable technique for effecting a solution to (2-19) appears to be numerical and, in all but the simplest of cases, would require machine computation.

Beginning in 1954, Bendat [9] treated the problem of optimum-time variable filtering for nonstationary random processes. His approach was to consider those system weighting functions $h(t,\tau)$ which are separable, i.e., $h(t,\tau) = f(t)g(\tau)$, where $f(t)$, the time-varying factor, is a function of t alone, while $g(\tau)$, the constant parameter factor, is a

function of τ alone. In addition, more complicated weighting functions of the form

$$h(t,\tau) = \sum_{n=1}^{M} f_n(t)g_n(\tau) + U(t)\,\delta(\tau) + V(t)\delta(t-\tau), \quad (2\text{-}20)$$

where M is an integer and δ is the Dirac delta function, were considered by Bendat. Included in Bendat's work is a discussion of asymptotic filter design and synthesis, where $h_a(t,\tau)$ is defined as an asymptotic weighting function to an optimum weighting function $h_o(t,\tau)$ if

$$h_a(t,\tau)/h_o(t,\tau) \longrightarrow 1 \text{ as } t \longrightarrow \infty \ (t \longrightarrow 0), \quad (2\text{-}21)$$

thus $h_a(t,\tau)$ can be either the long time ($t \longrightarrow \infty$) or short time ($t \longrightarrow 0$) asymptote to the optimum weighting function $h_o(t,\tau)$. The asymptotic design ideas presented by Bendat offer a solution to the synthesis of time-variable filters when exact optimal synthesis is difficult and when asymptotic operating regions are involved. The major drawback to Bendat's approach lies in the fact that nothing can be inferred in general about the asymptotic filter's performance during the non-asymptotic period. The question arises also, as to what restrictive conditions exist in order for the optimal weighting function to be of separable form.

Several ad hoc engineering approaches to the design and synthesis of received signal controlled time-varying filters have appeared in the literature. Keiser [10] proposed a

filter whose passband is controlled by a signal derived from a measurement of certain parameters in the input signal plus noise autocorrelation function. Using an experimental filter whose bandwidth and passband characteristics were fixed, but whose center frequency was a controlled variable, Keiser was able to demonstrate a significant improvement in signal-to-noise ratio for the variable-pass network in comparison to a fixed network filter even at fractional signal-to-noise ratios. Although his approach was essentially the first example of a forward loop adaptive filter, the experimental example treated was of such a simplified nature that the significant design problems associated with the general adaptive filter concept were rendered trivial.

The results of a General Electric Company investigation are described by Klass [11] wherein a filter was developed which adapts to a condition of pass or no-pass as determined by the presence or absence of a desired pulse-type signal at the input. The adaptation time for this filter was found to be dependent both upon input signal-to-noise ratio and upon pulse repetition rate, being approximately inverse to their product. In general, several seconds were required for the filter to adapt or in a rudimentary sense to learn about its environment.

Glaser [12], in 1961, reported the development of an adaptive filter for the detection of a pulse signal whose waveform is fixed, but unknown to the receiver. The algorithm

used was that of the maximum likelihood ratio, and a digitally simulated system exhibited "satisfactory" performance for signal-to-noise ratios as low as 12 db.

In 1962, Bucy and Follin [13] published the results of their investigation of an adaptive filter which was simulated on an analog computer to measure the spectral densities of the input signal and noise processes and adjust its bandpass characteristics to yield optimal filtering in the Wiener sense. The experiment treated two unknown parameters and the adaptation time required was slow, being of the order of several seconds.

Many worthwhile contributions have been made in less closely related areas such as adaptive or self-optimizing control systems and Stromer [14] gives an excellent bibliography of 47 references on pre-1959 work in this field. The texts by Chang [15], Newton et al [16], and Mishkin and Braun [17], are among the books which treat the general area of optimal design theory.

Chapter 3

ANALYSIS AND SYNTHESIS OF ADAPTIVE FILTERS

3.1 INTRODUCTION AND ORGANIZATION OF THE CHAPTER

A general form for the channel adaptive filter is suggested by the time-varying aspects of the problem, and this form is described, first as an over-all system, and then in detail for each functional block. An analysis is made in order to place the design factors in evidence and to examine their interrelationships for performance trade-offs and limitations. The proposed design method is outlined in the last section of this chapter.

3.2 THE GENERAL ADAPTIVE FILTER

The random process input to the adaptive filter is taken to be a nonstationary stochastic variable composed of signal s(t) and noise n(t), where the two are assumed to be stochastically independent and additive. In order for the time-varying linear filter to adapt to or follow the non-stationary behavior of this input in an optimum sense, it is necessary that the following operations be performed by the adaptive structure:

(1) The nonstationary behavior of the input process must be observed,

(2) the requisite filter parameter variations must be determined from a combination of both the physical measurements on the input process and the optimization algorithm for the filter, and

(3) the physical realization of the filter parameter variations necessary for optimal performance must be accomplished.

Based upon the foregoing operational description, a block diagram for the general adaptive filter structure is postulated as shown in figure 3.2-1.

The output of the filter, shown in figure 3.2-1 as $\hat{s}(t-\lambda)$, is an estimate of the input signal waveform $s(t)$, where λ is a time delay which arises due to (a) the inherent response time of a physical system to an excitation at $t=0$, and (b) the time required for the operations of channel identification, parameter computation, and implementation of the computed parameter values within the filter structure. Ideally, one would seek to cause $\lambda \longrightarrow 0$ and $\hat{s}(t) \equiv s(t)$, however the physical limitations of the system never allow such an ideal realization, and, in most communication systems, it is unnecessary that $\lambda \equiv 0$ or that $\hat{s}(t) \equiv s(t)$. An example in point is the situation in acceptable delayed voice communications where not only the individual speaker and listener determine the extent to which time delay and distortion can enter the system, but also the material being spoken.

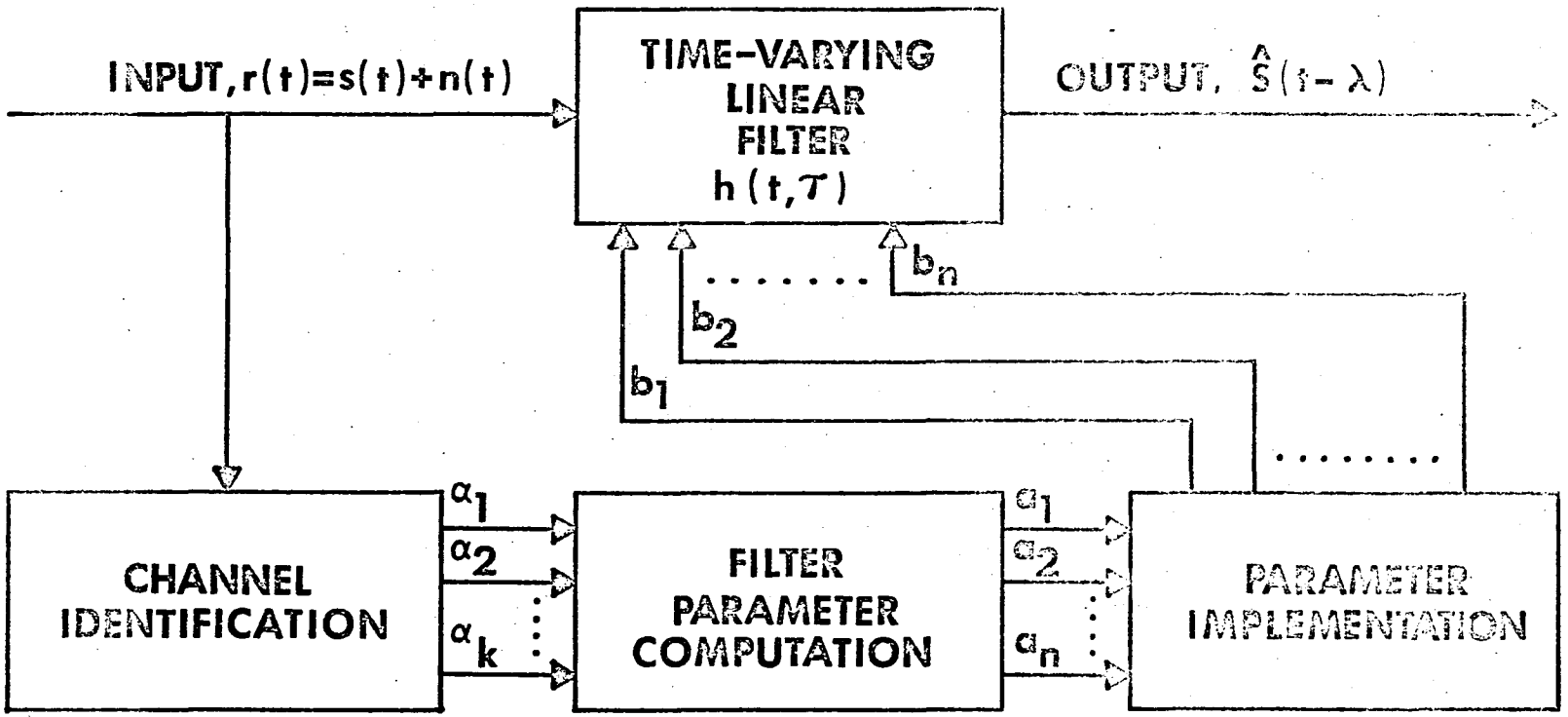


Figure 3.2-1. Block Diagram of General Adaptive Filter

While average output signal-to-noise ratio is not the only criterion for performance of a filter, in many applications it proves to be the most significant measure, and often no other is readily useable or tractable from an analysis point of view. Thus, the criterion for adaptive filter performance is chosen as minimal average square error, i.e.,

$$A\{e^2(t)\} \triangleq A\{[s(t) - \hat{s}(t-\lambda)]^2\} \qquad (3\text{-}1)$$

is to be a minimum. In addition, the system delay λ should be small, where the qualitative "small" is necessary until a specific application is determined or a quantitative restriction is imposed on this delay.

Of the two sources of time delay ((a) and (b) listed above) in the adaptive filter structure, the delay due to the adaptive operations of identification, computation, and implementation is the determining factor with regard to the allowable time rate of variation in the input nonstationarity parameters. This delay must be short with respect to the rate of change in the random process parameters. The effect of changing the filter's structure may otherwise be such that the filter response $h(t,\tau)$ does not correspond to the optimal response function $h_o(t,\tau)$ for the instantaneous input which exists at the time the adaptation process has been completed. Stated another way, the nonstationary

time dependence of the input signal must be slowly varying with respect to the time required for the over-all adaptive operation.

The following assumption, hereafter referred to as quasi-stationarity, is made regarding the random channel and additive noise portions of the communication system model depicted earlier in figure 1.1-1:

> Both the gain and phase-delay characteristics of the random channel and the nonstationary parameters of the additive noise are varying slowly with respect to transmitted signal message changes and receiver filter adaptation time.

The quasi-stationarity assumption regarding channel parameters, and hence received signal nonstationarities, can be justified for high information rate systems. Consider an FM/FM telemetry system employing 16 quantization levels and information source bandwidths of 30-50 kc. The average information rate of transmission over the channel will be approximately 29-48 kilobits/sec.[1] For a representative fading rate of 20 db in 100 milliseconds and a relatively fine quantization of 100 levels, the bit rate for the fading nonstationarity will be 66.5 bits/sec. Thus,

---

1. See Nichols and Rauch [18], p. 184, and Harman [19], pp. 109-111.

the average signaling rate is approximately three orders of magnitude higher than the fading rate.[1]

With respect to the additive noise source, the quasi-stationarity assumption is less easily justified, and for certain types of man-made interference it appears definitely incorrect. Nevertheless, such an assumption is inherent in the adaptive filter concept as herein proposed, and one must content oneself with the restrictions of the assumption or modify the philosophy of adaptation suggested here.

Despite the restrictions on the use of the proposed adaptive filter concept, there exist many physical applications for such systems. Ionospheric fading, multi-path interference, time-varying white background noise, and all types of slowly changing atmospheric conditions which corrupt the signal transmission are examples of the nonstationary problems for which adaptive filtering will be shown to be a useful and feasible receiver technique.

3.3 TIME-VARYING FILTER STRUCTURE

When one adopts the quasi-stationarity assumption, as stated in section 3.2, certain aspects of the Wiener-Kolmogorov theory become applicable to the problem. In

---

1. The fading rate used here is representative of the measured maximum short-term fading in ionospheric propagation tests. See the National Bureau of Standards [20] for empirical data on fading and noise in ionospheric propagation.

particular, it is useful to consider the stochastic input to the adaptive filter as a "piecewise" stationary and ergodic random variable, or equivalently, the problem may be stated as follows:

> The input stochastic process is composed of a sequence $\{i_a(t)\}$ of random variables, for which there exists a finite or infinite number of subsequences, each of which is defined to be stationary and ergodic within its particular subspace of the entire input space. The ergodic hypothesis, however, does not apply to the ensemble members of the different subsequences.

The optimal Wiener filter solution to this problem is thus a finite or infinite number of solutions to each subspace problem if one has the a priori knowledge that the ensemble member, which is the input to the system for any given period of time, belongs to a certain stationary and ergodic class of functions. It is the function of the channel identification block, figure 3.2-1, to provide the adaptive structure with this a priori knowledge.

In order to determine the specific information required for a solution to each of the equivalent Wiener problems which compose the quasi-stationary problem to be treated, a discussion of the, by now, classic Wiener problem is given in the appendix. There it is shown that the optimum

system function $H_o(s)$ which minimizes the averaged squared error for the Wiener filtering problem is

$$H_o(s) = \frac{\left\{ \dfrac{\Phi_{ss}(s)}{[\Phi_{ss}(s) + \Phi_{nn}(s)]_-} \right\}_{P.R.}}{[\Phi_{ss}(s) + \Phi_{nn}(s)]_+} \tag{3-2}$$

where the symbols $[\ ]_+$ and $\{\ \}_{P.R.}$ are as defined in the appendix.

An examination of (3-2) discloses that the a priori information required for a solution to the piecewise equivalent Wiener problems is the complex power spectral densities of signal and signal plus noise at the input to the filter. The two-sided transforms $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$ used in (3-2) are related to the signal and noise power spectral densities by the relation

$$\Phi_{xx}(s)]_{s=j\omega} = 2\pi\, G_{xx}(\omega). \tag{3-3}$$

If one considers the piecewise application of (3-2) to the nonstationary problem, it becomes necessary, in general, to first determine the power spectral densities (PSD's) of input signal and noise, and then to implement the solution of (3-2) for the optimum filter function $H_o(s)$. Moreover, these operations must be performed in real time such that the quasi-stationarity assumption is valid. Aside from the computational difficulties of implementing a machine (digital or analog) computation according to (3-2),

the time required for an accurate determination of $G_{ss}(\omega)$ and $G_{nn}(\omega)$ is ordinarily long if a serial procedure is used, or extremely complex and expensive, from a hardware viewpoint, if a parallel technique is employed.

The single most effective method for achieving a practical realization to the indicated solution lies in the use of all available a priori knowledge concerning the communication system. The form of modulation, physical medium of communication, information source statistics, and expected noise and interference process descriptions are all likely to be available to the system designer. Thus, it is probable that the channel identification block of the adaptive filter structure need not determine the continuous PSD of signal and noise for all frequencies, or even for a large number of discrete frequencies, in order for an accurate description of $G_{ss}(\omega)$ and $G_{nn}(\omega)$ to be obtained.

By using the a priori system knowledge available, it is possible to pre-design the time-varying filter structure; incorporating within this design the provision for adjusting certain elements in the structure, and hence parameters of the filter's response function. Three examples of this procedure are treated in Chapter 4.

There remain however, additional analytical complications associated with the computation and physical realization of the optimal filter form (3-2), even when

one is given the spectral densities of signal and noise. The operations of spectrum factorization (A-25) and partial fractions expansion (A-27) necessary to the physically realizable solution (3-2) make it necessary that $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$ be expressible as the quotient of rational, finite algebraic polynomials in s.

Such a representation cannot always be made exact for the signals and noise of interest in communications. An example[1] is the power spectrum of the pulse-width modulated waveform depicted in figure 3.3-1. Here it is assumed that the pulse-width ξ of an individual pulse is equally likely to be any value from zero to T. The pulse repetition rate is 2T. For this waveform the power spectrum will contain two components, one due to the random behavior of the pulse widths and the other component due to the periodic behavior of the pulse repetitions.

The two-sided Fourier transform of the time autocorrelation function for the random component may be shown to be

$$\Phi_{rr}(s) = K^2 \left[\frac{T^2S^2 - 4 \sinh^2 (TS/2)}{-T^4S^4} \right] , \qquad (3\text{-}4)$$

where $K = K(E,T)$. Thus the power spectrum is not a quotient of rational, finite algebraic polynomials in s.

---

1. This example is contained in Lee [21], pp. 225-235.

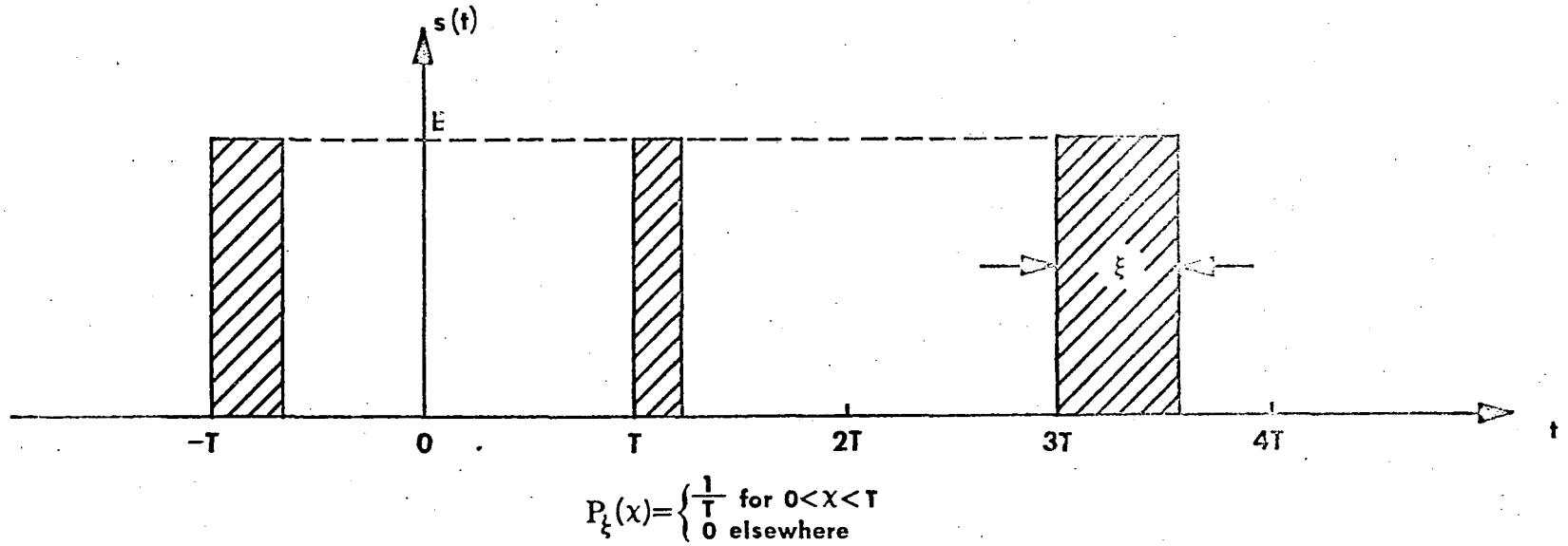


Figure 3.3-1. A Pulse-Width Modulated Wave

Various methods are available, e.g., the Butterworth approximation,[1] for representing such power spectra by a finite algebraic polynomial quotient. Consequently, it is assumed that the spectra of interest in this work are either of this exact nature or can be approximated by such a representation.

The second aspect of difficulty to the application of (3-2) is the algebraic complexity which is present. This complexity effectively limits the use of Wiener theory to simple problems. In all but the simplest of cases, the algebraic complexity of finding $H_o(s)$, which results from an exact solution to (3-2), is such that the coefficients of the s-terms in numerator and denominator are high order functions of the spectrum parameters. Such solutions, even when the algebraic difficulties have been overcome, do not disclose to the synthesist a procedure for adjustment of the filter elements which will cause $H_o(s)$ to correspond to the quasi-stationary forms of $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$.

R. B. Streets, Jr. [23] has developed a technique for the determination of an approximation to the Wiener filter which greatly simplifies the solution of (3-2) and in addition presents $H_o(s)$ in a concise algebraic form. Streets' method is founded upon a high degree of intuition resulting from his examination of many particular cases of exact Wiener design.

---

1. See Tuttle [22] pp. 741-831.

The method proposed by Streets is to normalize the optimum unrealizable filter transfer function (A-22) with respect to $\Phi_{nn}(s)$. Thus one begins with (A-22), repeated below

$$H_o(s) = \frac{\Phi_{ss}(s)}{\Phi_{ss}(s) + \Phi_{nn}(s)} ,$$

and then factors the normalized $\Phi_{ss}(s)$ function into a LHP product Φ_s^+ and a RHP product Φ_s^-, each containing the square root of the gain.

$$H_o(s) = \frac{\Phi_s^+ \Phi_s^-}{1 + \Phi_s^+ \Phi_s^-} . \qquad (3-5)$$

He then naïvely[1] assumes that all that is necessary in order to make $H_o(s)$ physically realizable is to retain only the LHP factor Φ_s^+ of the normalized $\Phi_{ss}(s)$, thus the physically realizable optimal response function is

$$H_o(s) = \frac{\Phi_s^+}{1 + \Phi_s^+} . \qquad (3-6)$$

Streets then identifies Φ_s^+ with the open loop transfer function of a unity feedback system as follows

$$G_o(s) = \mathcal{O}(1/s), \text{ for } \Phi_s^+ < 1,$$
$$G_o(s) \approx \Phi_s^+, \text{ for } \Phi_s^+ > 1, \qquad (3-7)$$

---

1. This is Streets' term for this assumption. See [23] p. 34.

where the notation $\mathcal{O}(1/s)$ is used to indicate that $G_o(j\omega)$ has a Bode plot for $\omega > \omega_c/\sqrt{2}$ which has a -1 slope (-20 db per decade), and where ω_c is the radian frequency at which $|G_o(j\omega)| = 1$.

The explanation for the $\sqrt{2}$ factor by which the break frequency ω_c is reduced lies in the closed loop "Butterworth" system approach that Streets assumes. An example of the method follows.

Example:

$$\Phi_{ss}(s) = \frac{R^2}{1 - S^2/V^2}, \quad \Phi_{nn}(s) = N^2.$$

Then the normalized Φ_{ss} is

$$\Phi_{ss} = \frac{R^2/N^2}{1 - S^2/V^2}$$

and

$$\Phi_s^+ = \frac{R/N}{1 + S/V}, \quad \Phi_s^- = \frac{R/N}{1 - S/V}.$$

Thus, we let

$$G_o(s) = \frac{R/N}{1 + S/V},$$

which is valid for all ω since in this particular case $G_o(j\omega) = \Phi_s^+(j\omega)$ for all $\Phi_s^+(j\omega)$.

The closed loop approximate filter form will then be given by

$$H_o(s) \approx \frac{G_o(s)}{1 + G_o(s)} = \frac{R/N}{1 + R/N + S/V}. \qquad (3\text{-}8)$$

The exact Wiener filter for this example is given by

$$H_o(s) = \frac{\sqrt{1 + (R/N)^2} - 1}{\sqrt{1 + (R/N)^2} + \frac{S}{V}} . \qquad (3\text{-}9)$$

Thus one may observe that (3-8) and (3-9) agree "very nearly" when $R/N > 10$. However, the true measure of the approximation lies in the differences in performance between the approximate (3-8) and the exact system (3-9). Streets has compared the RMS value of the error for this approximate filter (and others) as a percentage increase over the exact filter RMS error. His analysis shows that even for a signal-to-noise ratio as low as 2:1 (6 db) the approximation error is only about 10% greater than the exact filter RMS error, and for a signal-to-noise ratio of 10:1 (20 db), the approximation error is less than 0.5% greater than the exact filter RMS error.

It is of interest to note that even for this simple example the approximate filter function offers a considerable simplification in its "computational" form over the exact filter. That is to say that the machine computations and measurements necessary in order to perform the system operations as outlined in section 3.2 are greatly simplified by the approximate filter form.

3.4 CHANNEL IDENTIFICATION

As has been shown in the preceding section, the function of the channel identification block in figure 3.2-1 is to provide the adaptive structure with enough a priori knowledge of the input processes for a determination and adjustment of the variable elements in the filter. Basically this is an estimation problem in the sense of mathematical statistics, and one is led to consider certain aspects of the theory of statistical inference.

Parzen [26,27,28,29,30,31], Blackman and Tukey [32], Grenander and Rosenblatt [33], and others [34,35,36,37] have treated the general problems of estimating the spectral densities of stochastic processes. Any attempt to develop a criterion for the "goodness" of a spectral density estimate is inseparably connected with the purposes for which the estimate will be used. However, there are certain fundamental properties of statistical estimates which act as quantitative measures of the estimate's perfection. A basic requirement for an estimate is consistency in quadratic mean.[1] If a stationary time series x(t), is observed over the sample period $0 \leq t \leq T$, and an estimate k_T of an unknown parameter k of the series is formed on the basis of the sample, then k_T is consistent in quadratic mean if

$$\lim_{T \to \infty} \{E[|k_T - k|^2]\} = 0,$$

---

1. Parzen [27], p. 331.

where the expected value is taken under the assumption that k is the true parameter value. If an estimate is consistent, it is asymptotically unbiased, i.e.,

$$\lim_{T\to\infty} \{E[k_T]\} = k.$$

Parzen [31] points out that empirical spectral analysis of a stationary time series is not an "easy problem" due to the fact that the obvious estimate of the spectral density function; namely, $f_T(\omega)$, defined by

$$f_T(\omega) = \frac{1}{2\pi}\int_{-T}^{T} e^{j\tau\omega} R_T(\tau)\,d\tau,$$

where

$$R_T(\tau) = \frac{1}{T}\int_0^{T-|\tau|} x(t)x(t+|\tau|)\,dt, \qquad |\tau| < T,$$

is <u>not</u> a consistent estimate of the true spectral density function $f(\omega)$.

Many ways have been suggested in the literature for forming consistent estimates of the value $f(\omega_o)$ of a spectral density function at a given frequency ω_o. Generally these methods belong to one or the other of two classes, (i) the use of a continuous time record of finite length, and (ii) the use of discrete data taken at equally spaced time values. Of these two classes of methods (i) offers the advantage of empirical calculation in real time such that it should be possible to continuously, or at lease piecewise continuously,

estimate the spectral densities of interest at certain discrete frequencies.

Analog computer techniques for estimating the power spectral density of a random process at a discrete frequency have been suggested by various authors, and Rideout [38] gives the square-law realization shown in figure 3.4-1.

Certain significant practical difficulties arise when the technique of power spectral density estimation illustrated by figure 3.4-1 is employed. An ideal flat frequency response function $H_f(j\omega)$ centered at f_i and having bandwidth Δf cannot be physically realized, the operation of squaring is difficult to implement over a large dynamic range of inputs, and the integration must be cyclic with the estimate readout occurring at the end of each integration period T before the integrator is reset to a zero initial condition.

Blackman and Tukey [32] have treated the problem of non-ideal filtering for the continuous analysis case by a consideration of several realizable spectral window forms $H_f(j\omega)$ for which an equivalent ideal filter bandwidth is defined. In addition, they introduced the concept of the equivalent number of degrees of freedom of the square-law estimate, where

$$k = \frac{2}{[\Delta\Phi(\omega)]^2} , \qquad (3\text{-}10)$$

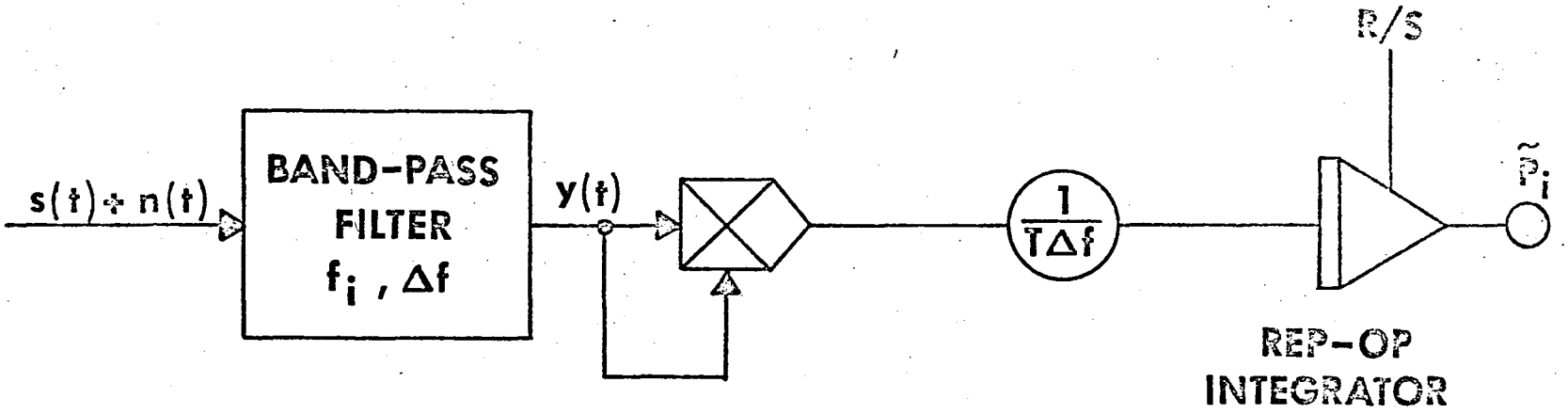


Figure 3.4-1. Square-Law Power Spectral Density Analyzer

is the equivalent number of degrees of freedom; $\Delta\Phi(\omega)$ is termed the coefficient of variation and is given by

$$\Delta\Phi(\omega) = \frac{\sigma_1}{\langle a^2 \rangle} , \tag{3-11}$$

where σ_1 is the standard deviation of the estimate and $\langle a^2 \rangle$ is the mean of the estimate.

Morrow [39] has shown that

$$\frac{\sigma_1}{\langle a^2 \rangle} = \frac{1}{\sqrt{2(BW)\ T}} , \tag{3-12}$$

where BW is the "effective" bandwidth of the filter in cps and T is the integration time in seconds. When the coefficient of variation is less than 0.2, the error in the PSD estimate has an approximate normal distribution and a confidence level can be established for the estimate.

If a permissible error of 10 per cent in the PSD estimate is assumed, the coefficient of variation becomes 0.1 and the required bandwidth-integration time product must be 50. In this case the equivalent number of degrees of freedom, k, is equal to 200, and as is well known, the chi-square distribution may be approximated by the normal distribution as $k \longrightarrow \infty$.[1]

From a table of the normal distribution one determines that for k = 200, 90 per cent of the PSD estimates

---

1. Burington and May [40] state that if $k > 30$, in practical calculations it is often sufficient to approximate the χ^2 distribution by a normal distribution with mean k and variance 2k.

lie within the range of 0.836 to 1.165 of the true PSD, i.e., a spread of 1.443 db. Thus, one has an *a priori* probability of 0.9 (90% confidence) that a single PSD estimate will lie within an interval of width 1.44 db about the true PSD.

To illustrate the magnitude of integration time required for the estimation of PSD in a communication channel, consider the problem of a large index FM wave having a 1-mcps bandwidth after modulation. If one selects five distinct, equally spaced frequencies at which PSD is to be estimated, narrow-band filters having "effective" bandwidths of approximately 50 kcps could be used.

Filter Q's on the order of 2,000 would be required at perhaps 100 mcps, hence the use of crystal filters would be indicated. The required integration time, in order that all five PSD estimates be made at a 90% confidence level of 1.44 db will be

$$T = \frac{k}{4 \cdot BW}$$

$$T = \frac{200}{4(10^4)5} = 1 \text{ millisecond.}$$

Hence, short integration times are required when compared to the representative maximum fading rate of 20 db in 100 milliseconds.

It has been suggested by Cohn [41] that the square-law device shown in figure 3.4-1 can be replaced in practice by a linear detector, which is much simpler to realize

accurately. Cobb shows that both types of detectors will give the correct rms amplitude for an input which has a Gaussian amplitude distribution. Since the input to the square-law device has been narrow band-pass filtered, it follows[1] that the amplitude distribution of the random variable y(t) in figure 3.4-1 is approximately Gaussian.

The analog realization shown in figure 3.4-2 is suggested as a practical power spectral density analyzer for estimating the mean-square amplitude of the random variable x(t) at the discrete frequency f_i.

When the random variable y(t), in either figure 3.4-1 or figure 3.4-2, has a Gaussian amplitude distribution with standard deviation σ and zero mean, as has been postulated, the expectation of the output of the square-law analyzer is

$$E\{P_y\} = \int_{-\infty}^{\infty} y^2 \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-y^2/2\sigma^2\right] dy = \sigma^2, \quad (3\text{-}13)$$

and that of the linear analyzer is

$$E\{P_y\} = \int_{0}^{\infty} y \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-y^2/2\sigma^2\right] dy = \sigma\left(\frac{2}{\pi}\right)^{1/2}. \quad (3\text{-}14)$$

Thus, for ergodic random inputs, the square-law analyzer gives an asymptotically unbiased estimate of the variance

---

1. M. Rosenblatt [42] has examined the rigorous mathematical conditions necessary for this ascertion and has shown that a fairly broad domain of inputs exists within which the assertion holds.

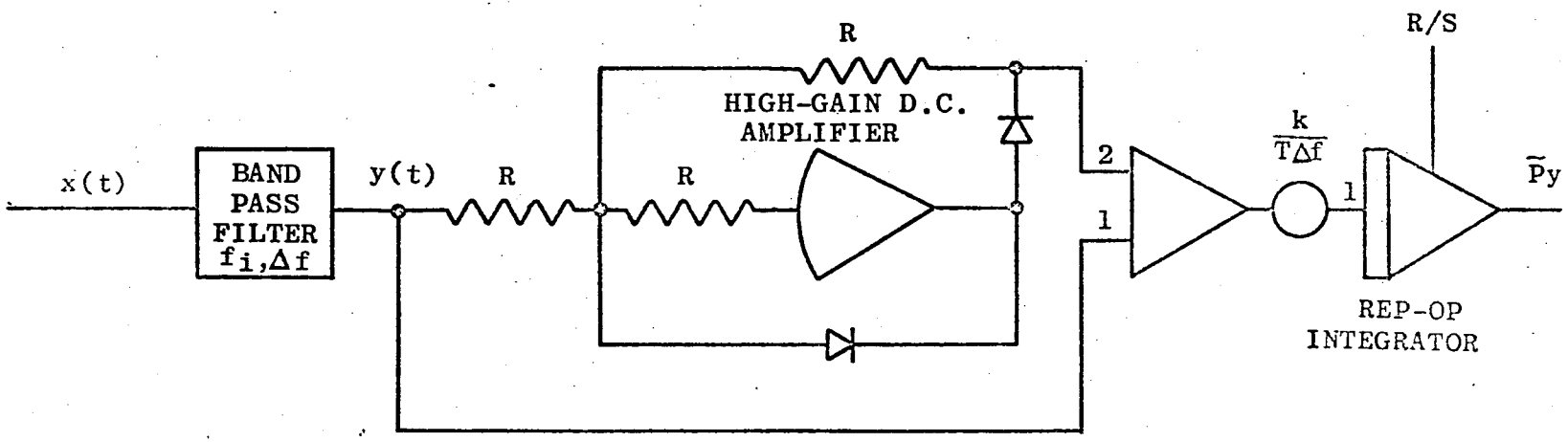


Figure 3.4-2. Linear Detector Power Spectral Density Analyzer

of y(t), while the linear analyzer does not. The asymptotic bias of the linear analyzer estimate is, however, removable and this estimate offers certain other advantages in computational simplicity which will be discussed in Chapter 4.

Since it is the final values of the integrator outputs (figures 3.4-1 and 3.4-2) which are the PSD estimates, only those output values occurring at integer multiples of the repetitive integration time T represent the correct estimates $\{\alpha_k\}$ of PSD for use as inputs to the parameter computation block. Consequently, a sample-hold device must be used following the integrator shown in either figure 3.4-1 or 3.4-2 in order to form the piecewise-constant estimates of PSD. This sample-hold circuit must operate in time synchronism with the repetitive operating integrators such that the samples of $P_y(t)$ are taken slightly in time advance of the integrator's reset operation.

The realization of such a sample-hold device is discussed in Chapter 4 along with other aspects of the analog simulation of adaptive filters.

3.5 FILTER PARAMETER COMPUTATION

Referring to the general adaptive filter diagram, figure 3.2-1, the input to the parameter computation block consists of the $\{\alpha_i\}$ sample-held PSD estimates from the channel identifier, as discussed in the preceding section. Since it has been assumed that enough a priori knowledge is

available to pre-design the form of the time-varying filter structure, it is then the function of the parameter computation block to determine the requisite values of the adjustable parameters within the filter structure.

Although the parameter computations can be made either digitally or by analog techniques, only analog techniques are discussed here since this technique operates in a parallel fashion in real time and has sufficient accuracy for the intended utilization.

Parameter computation is accomplished as follows. The k point-frequency PSD estimates are obtained from the channel identification block at times T, 2T, 3T, ... , and these estimates act as the constant (for a period T) inputs to the parameter computer. From these inputs, utilizing a straight line or asymptotic Bode approximation to the input (signal plus noise) PSD curve, the unknown filter parameters $\{a_m\}$ can be determined.

An example to illustrate the method is given as follows. Let the signal and noise have complex spectra of the forms

$$\Phi_{ss}(s) = \frac{R^2\left(1 - \frac{s^2}{r^2}\right)}{\left(1 - \frac{s^2}{q^2}\right)\left(1 - \frac{s^2}{u^2}\right)}, \quad q < r < u,$$

and

$$\Phi_{nn}(s) = N^2,$$

respectively.

The power spectral densities for this example, along with the asymptotic Bode approximations, are illustrated in figure 3.5-1.

It is assumed that R^2, N^2, q, r, and u are random variables satisfying the quasi-stationarity assumption and such that the ranges of q, r, and u do not overlap. The frequencies at which PSD estimates are taken are identified as ω_1, ω_2, ω_3, ω_4, and ω_5. Since the random variables comprising signal and noise are taken to be uncorrelated, the complex spectra input to the adaptive structure is

$$\Phi_{ii}(s) = \Phi_{ss}(s) + \Phi_{nn}(s).$$

For this example, the PSD estimates at $\omega_1, \ldots, \omega_5$ are identified as $P_1, \ldots, P_5$ and may be written, using the asymptotic Bode approximation, as

$$\begin{aligned} P_1 &= R^2 + N^2 \\ P_2 &= N^2 + \left(\frac{q}{\omega_2}\right)^2 R^2 \\ P_3 &= N^2 + \left(\frac{q}{r}\right)^2 R^2 \\ P_4 &= N^2 + \left(\frac{q}{r}\,\frac{u}{\omega_4}\right)^2 R^2 \\ P_5 &= N^2 + \left(\frac{q}{r}\,\frac{u}{\omega_5}\right)^2 R^2 \end{aligned} \qquad (3\text{-}15)$$

Equations (3-15) comprise a set of five independent linear equations in the five unknowns R^2, N^2, q^2R^2, $\left(\frac{1}{r}\right)^2 (q)^2R^2$, $(u)^2\left(\frac{q}{r}\right)^2R^2$. Since the system determinant

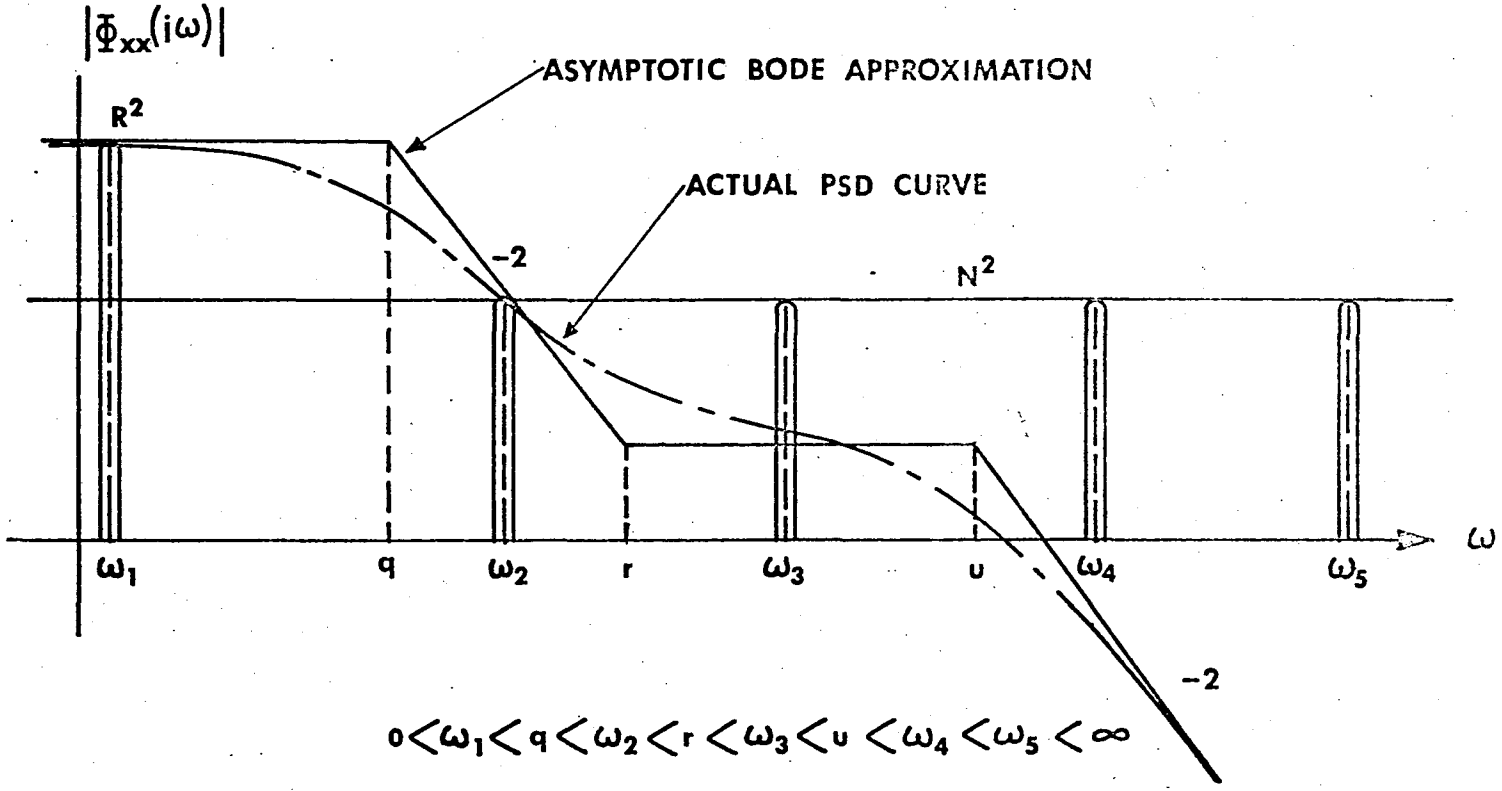


Figure 3.5-1. Example Signal and Noise Power Spectral Density Curves

$$\Delta = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & (1/\omega_2)^2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & (1/\omega_4)^2 \\ 0 & 1 & 0 & 0 & (1/\omega_5)^2 \end{vmatrix}$$

is non-zero for $\omega_4 \neq \omega_5$. unique solutions exist and are given by

$$R^2 = \frac{P_1(\omega_5^2 - \omega_4^2) - \omega_5^2 P_5 + \omega_4^2 P_4}{\omega_5^2 - \omega_4^2} ,$$

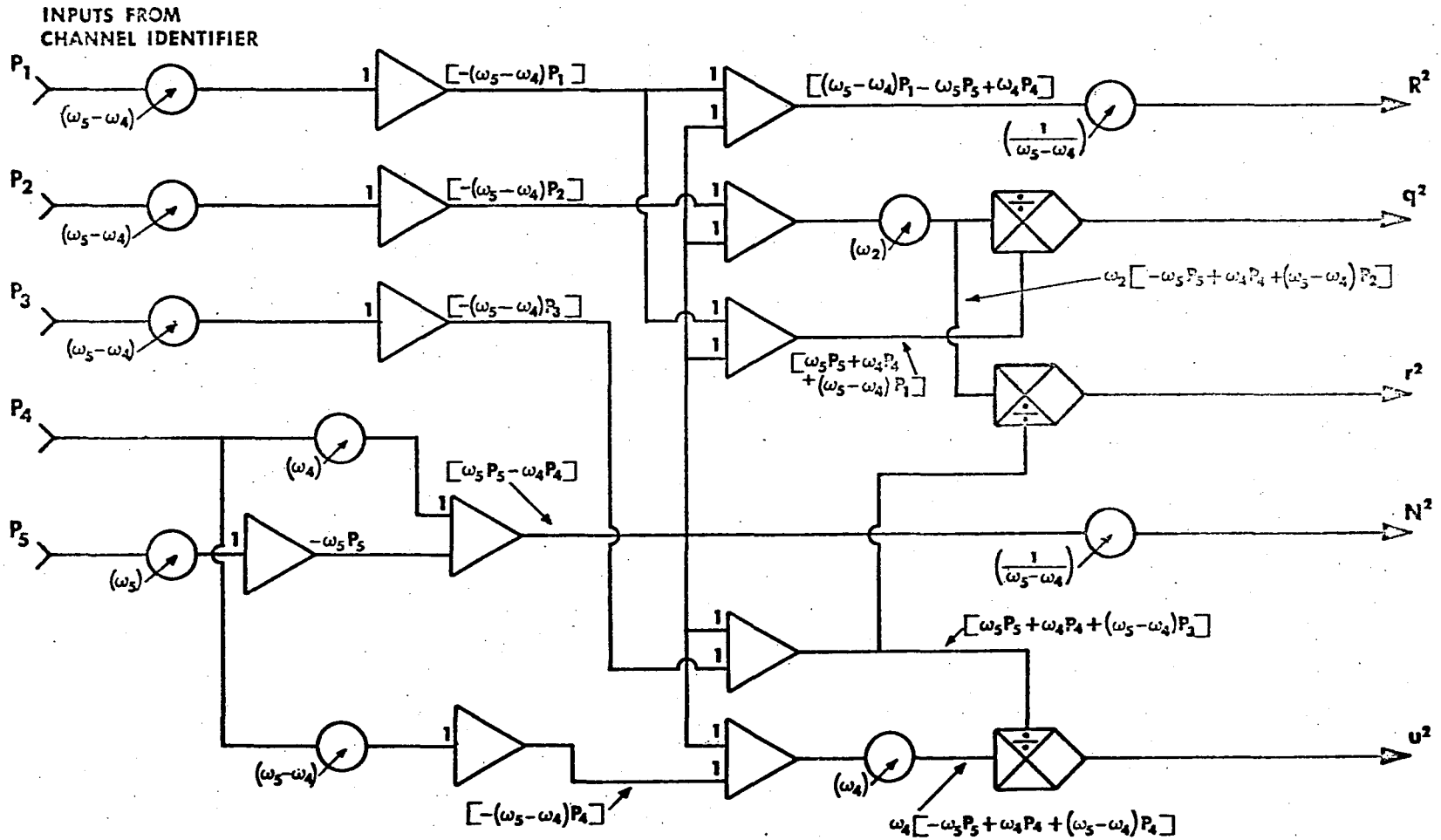
$$N^2 = \frac{\omega_5^2 P_5 - \omega_4^2 P_4}{\omega_5^2 - \omega_4^2} ,$$

$$q^2 = \omega_2^2 \frac{\omega_5^2 P_5 - \omega_4^2 P_4 - (\omega_5^2 - \omega_4^2)P_2}{\omega_5^2 P_5 - \omega_4^2 P_4 - (\omega_5^2 - \omega_4^2)P_1} , \qquad (3\text{-}16)$$

$$r^2 = \omega_2^2 \frac{\omega_5^2 P_5 - \omega_4^2 P_4 - (\omega_5^2 - \omega_4^2)P_2}{\omega_5^2 P_5 - \omega_4^2 P_4 - (\omega_5^2 - \omega_4^2)P_3} ,$$

$$u^2 = \omega_4^2 \frac{\omega_5^2 P_5 - \omega_4^2 P_4 - (\omega_5^2 - \omega_4^2)P_4}{\omega_5^2 P_5 - \omega_4^2 P_4 - (\omega_5^2 - \omega_4^2)P_3} .$$

An analog realization for the determination of the unknown parameters in this example is shown in figure 3.5-2. Higher ordered power spectral density representations than those used in the example will require a more complicated computation block, however, the types of computing elements will not change.



NOTE: FOR NOTATION SIMPLICITY THE QUANTITIES ω_1, ω_2, ω_3, ω_4, ω_5 SHOWN IN THE DIAGRAM ARE THE NUMERICAL VALUES OF THE CORRESPONDING ω_i VALUES SQUARED.

Figure 3.5-2. Example Parameter Computation Diagram

3.6 PARAMETER IMPLEMENTATION

The computed parameter estimates $\{a_m\}$ from the filter parameter computation block act as the inputs to the parameter implementation block, whose function it is to effect the adjustment of the variable components of the time-varying filter structure. Both the given inputs and required outputs determine the structure of the implementation block.

In general the inputs to the parameter implementation block are piecewise d.c. random voltages produced by the algebraic transformations of the computation block on the sampled and held PSD outputs from the channel identifier. The required outputs from the implementation block are determined both in quantity and functional form by the physical realization chosen for the time-varying filter.

Since the time-varying filter used is one whose transfer function is specified in form but whose pole and zero positions may be required to vary from sampling interval to sampling interval, both dynamic response and static accuracy are implied requirements of the implementing block. Response speed to step changes in the inputs $\{a_m\}$ is related to the integration time T used for PSD estimation. A rule of thumb would be to require the 10% to 90% response time of the implementing system to be less than $\frac{T}{10}$, such that the time-varying filter parameters will be correctly realized, within 10%, for approximately 90% of the time.

Only one method of varying the parameters of the filter form is described in this work, although other techniques have been considered. By using an analog quarter-square multiplier within the loop of a differential analyzer form it is possible to control accurately ($\pm$0.5% of saturation) and rapidly (10% to 90% response time $\sim$50 μsec.) the positions of poles and zeros in a linear transfer function.

As an example of this technique for time-varying filter realization, consider the following transfer function

$$H(s) = \frac{A/B}{1 + A/B + S/B} ,$$

where A and B are to be controlled parameters. An analog realization of this transfer function is shown in figure 3.6-1, wherein two multipliers are required to realize the transfer function dependence upon parameters A and B.

In general, for a transfer function of the form

$$H(s) = \frac{K(1 + \frac{s}{a_1})(1 + \frac{s}{a_2}) \cdots (1 + \frac{s}{a_m})}{(1 + \frac{s}{b_1})(1 + \frac{s}{b_2}) \cdots (1 + \frac{s}{b_n})} , \qquad (3\text{-}17)$$

where $n > m$, the differential-analyzer realization with multiplier controlled parameters will require $m + n + 1$ multipliers for the simulation.

The control signals produced by the parameter implementation block act as the inputs to the various multipliers used in the analog simulation of the filter transfer function. The examples given in Chapter 4 further illustrate the type

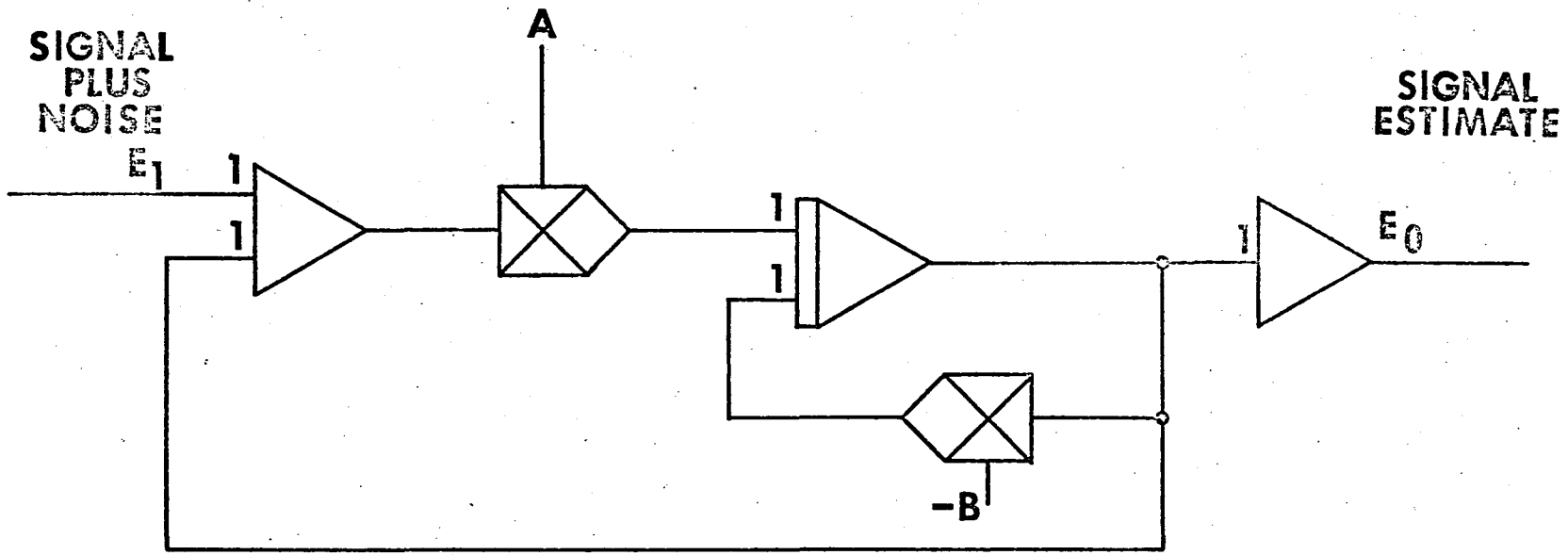


Figure 3.6-1. Analog Realization of the Transfer Function $H_o(s) = \frac{A/B}{1 + A/B + s/B}$

of operations required and hardware realizations necessary for the operation of parameter implementation.

3.7 SUMMARY OF THE PROPOSED DESIGN METHOD

An a priori knowledge of the communication system is assumed as follows: $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$ are of known forms; these forms contain certain quasi-stationary random parameters. The ranges over which the random parameters vary is known or at worst a confidence statement can be made in regard to the parameter ranges.

The form of the time-varying optimal filter is determined, based upon the available a priori knowledge of the system descriptors $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$. In practical cases, the use of Streets' technique to determine the approximate form of $H_o(s)$ will greatly simplify the analytical work necessary for this step. Once the realization of $H_o(s)$ has been found, the requisite quasi-stationary process parameters necessary to the time-varying form of $H_o(s)$ are placed in evidence. The parameter estimate outputs from the channel identifier are then known and channel identification circuit design is accomplished utilizing the $\Phi_{ss}(s) + \Phi_{nn}(s)$ form, which is the input, and the required parameter estimates, which compose the output of the identification block.

Special purpose analog computing circuits are then designed in order to compute the required filter parameter

variations necessary for optimal performance. In certain cases, the hardware complexity necessary to the parameter computer realization can be simplified by using linear rather than square-law detectors for PSD estimation by the identifier.

The parameter implementation block, needed to realize the time-varying filter parameters, is dependent upon the method of parameter adjustment chosen within the filter structure and upon the inputs from the computation block. Consequently, the implementation circuit design is the final step in the procedure.

Examples of the application of the design method are presented in the next chapter.

Chapter 4

ANALOG SIMULATION OF ADAPTIVE FILTERS

4.1 INTRODUCTION AND ORGANIZATION OF THE CHAPTER

Three examples of adaptive filter design are presented and the analog simulations of the adaptive filters to be designed and analyzed as examples in sections 4.2 and 4.4 are described in this chapter. The experimental work presented is intended to both justify the proposed theory and to illustrate the hardware complexities which arise in the realization of adaptive filters.

An EAI-TR-48 general purpose analog computer having 32 operational amplifiers, 8 repetitive-operation integrators, 3 quarter-square multipliers, 2-X^2 diode function generators, and 20 coefficient setting potentiometers, was used for the simulations. In addition, several pieces of special purpose analog equipment were constructed using Burr-Brown #1512 operational amplifiers.

The over-all experimental complex of equipment was state-of-the-art representative for the intended use, and all instruments used were calibrated and determined to be operating within the manufacturer's published specifications.

4.2 ADAPTIVE FILTER EXAMPLE NO. 1 - DESIGN AND ANALYSIS

The general class of problems to which this example relates is that of low-pass signals corrupted by white noise.

The signal complex spectral density function is given by

$$\Phi_{ss}(s) = \frac{R^2}{1 - S^2/v^2} \tag{4-1}$$

and the noise complex spectral density function by

$$\Phi_{nn}(s) = N^2. \tag{4-2}$$

It may be noted that the signal spectral density is representative of a random square wave having Poisson distributed event points where 2/v is the expected number of event points per second.

The parameter v is taken as time-invariant for this example, while the parameters R^2 and N^2 are taken as non-stationary and are postulated to obey the quasi-stationarity assumption of section 3.2.

The exact stationary Wiener filter for this example has a transfer function which is given by

$$H_o(s) = \frac{\sqrt{1 + (R/N)^2} - 1}{\sqrt{1 + (R/N)^2} + S/v}, \tag{4-3}$$

while the Street's approximation yields

$$H_a(s) = \frac{R/N}{1 + R/N + S/v} \qquad (4\text{-}4)$$

One may observe that (4-3) and (4-4) agree "very nearly" when $R/N > 10$, and Streets [23] has shown that when $R/N = 2$ (6 db) the approximate filter has only about a 10% greater RMS error than the exact filter.

It is interesting to note that even for this "simple" example the approximate filter form offers a considerable simplification in computational requirements over the exact filter. In addition, since (4-4) contains R/N directly as a parameter rather than $(R/N)^2$ as in (4-3), it is obvious that a properly designed linear detector realization of the channel identification block will greatly simplify the adaptive structure.

A consideration of the Bode approximation to the PSD curves shown in figure 4.2-1, shows that the two discrete frequency PSD estimates using a linear detector at ω_1 and ω_3, denoted by F_1 and F_3, will yield respectively

$$F_1 = \sqrt{R^2 + N^2} \quad \text{and} \quad F_3 = \sqrt{N^2 + \left(\frac{v}{\omega_3}\right)^2 R^2}. \qquad (4\text{-}5)$$

By choosing $\omega_3 >> v$, and forming the quotient of F_1/F_2 from (4-5) one has

$$\frac{F_1}{F_3} = \sqrt{\frac{R^2 + N^2}{N^2 + \left(\frac{v}{\omega_3}\right)^2 R^2}} \approx \sqrt{\frac{R^2}{N^2} + 1}, \qquad (4\text{-}6)$$

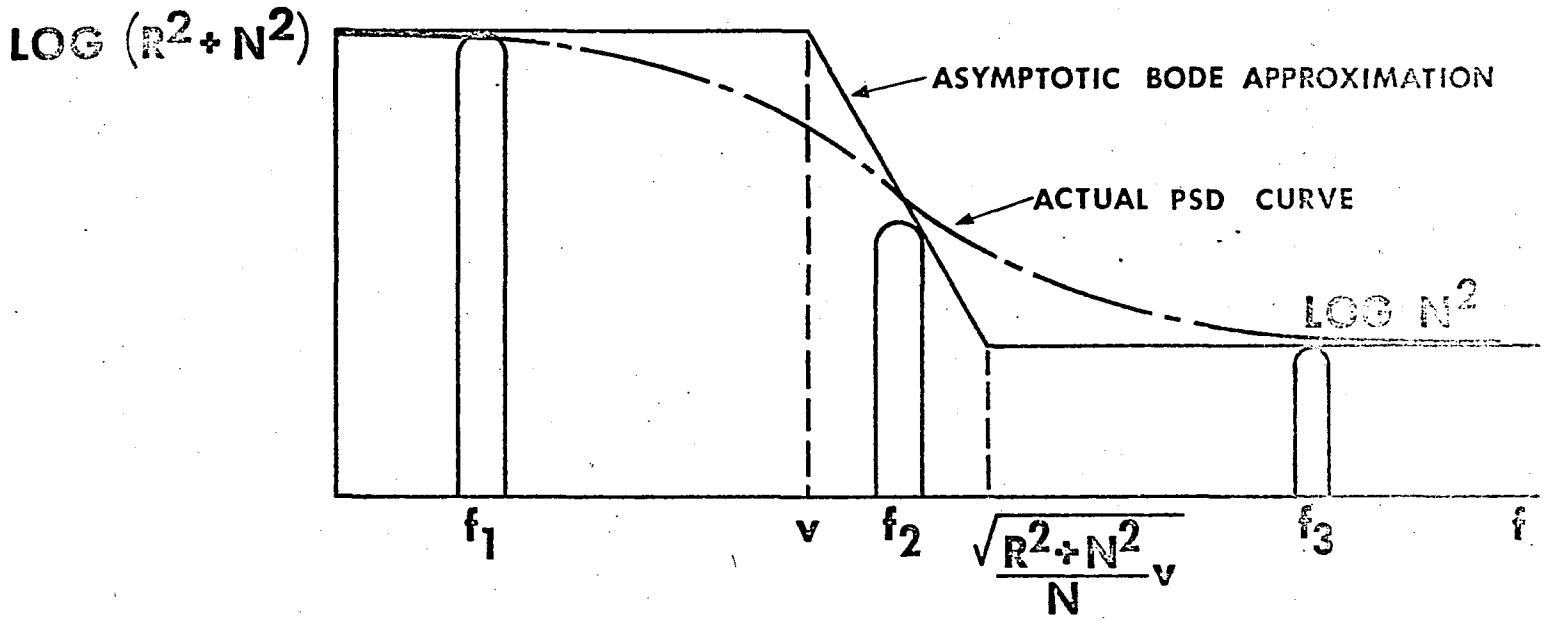


Figure 4.2-1. Bode Plot of $\Phi_{ss}(s) + \Phi_{nn}(s)$ Where $\Phi_{ss}(s) = \frac{R^2}{1-s^2/v^2}$; $\Phi_{nn}(s) = N^2$

and for the case where $\frac{R}{N} > 1$, (4-6) may be further reduced to

$$\frac{F_1}{F_2} \approx \frac{R}{N} \tag{4-7}$$

Thus, by employing the linear detection method for PSD estimation, an estimate of the requisite R/N parameter variation in (4-4) can be formed directly from the quotient of the two PSD estimates at f_1 and f_3 under the assumptions stated above.

To predict the theoretical improvement that an "ideal" adaptive filter[1] realizes, in increased output signal-to-noise ratio, over a fixed optimum filter, one makes the following analysis.

Let γ denote the design value of R/N chosen for the fixed optimum filter of equation (4-3). Since v is not changing with time in this example, it is convenient to normalize the complex variable s in (4-3) and in (4-4) such that

$$H_o(s) = \frac{\sqrt{1 + \gamma^2} - 1}{\sqrt{1 + \gamma^2} + s} \tag{4-8}$$

and

$$H_a(s) = \frac{R/N}{1 + R/N + s} .$$

---

1. The term "ideal", as used here, denotes an adaptive filter that makes zero error in identifying the signal and noise PSD parameters, and requires no time for the identification, computation, and implementation operations.

The error spectrum for the output of the linear ideal filter may be shown to be[1]

$$\Phi_{ee}(s) = \Phi_{ss}(s)[1-H(s)][1-H(-s)] + \Phi_{nn}(s)H(s)H(-s) \quad (4\text{-}9)$$

and when the inverse transformation is invoked, the mean-square filter error is given by

$$\overline{e^2(t)} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \Phi_{ee}(s)ds. \quad (4\text{-}10)$$

Upon substituting the two filter forms (4-8) into (4-9), one obtains the error spectral representations for the adaptive and fixed filters as follows:

$$\Phi_{ee_a}(s) = \frac{2R^2}{(1+R/N)^2 - s^2},$$

and (4-11)

$$\Phi_{ee_o}(s) = \frac{R^2 + N^2[\sqrt{1+\gamma^2} - 1]^2}{1+\gamma^2 - s^2}.$$

The mean-square filter error may now be obtained by substituting the two spectral forms in (4-11) into (4-10) and using a contour integration around the left-half plane. One obtains

$$\overline{e_a^2} = \frac{R^2}{1 + R/N} \quad (4\text{-}12)$$

and

$$\overline{e_o^2} = \frac{R^2 + N^2[\sqrt{1+\gamma^2} - 1]^2}{2\sqrt{1+\gamma^2}}. \quad (4\text{-}13)$$

---

1. See Newton, Gould, and Kaiser [16], pp. 123-125.

It is convenient to normalize (4-12) and (4-13) with respect to the signal "strength parameter" R^2 and to define the adaptive filter gain F_G (expressed in db) as

$$F_G \triangleq 10 \log_{10}\left[\frac{\overline{e_o^2}}{\overline{e_a^2}}\right], \tag{4-14}$$

which is equivalent to the ratio of the output signal-to-noise ratios of the filters. Thus, if $F_G > 0$ db, the output signal-to-noise ratio of the adaptive filter is greater, by the value of F_G, than that of the fixed optimum filter.

For an input condition on the quasi-stationary factor R/N of $R/N = \gamma$, the value of F_G is ≤ 0 db, since the exact optimum filter $H_o(s)$ in (4-8) will produce the minimum possible mean-square error and $H_a(s)$, from (4-8), will produce an error greater than or equal to that of $H_o(s)$ due to the Streets approximation.

However, for a sufficiently large variation in R/N from the fixed design value γ, the "ideal" adaptive filter will produce a lower mean-square error than the fixed optimum filter. Figure 4.2-2 is a graph of equations (4-12) and (4-13), normalized with respect to R^2, vs the quasi-stationary value of R/N.

Either analytically, using (4-12) and (4-13), or from the graph of mean-square errors for the two filters figure 4.2-2, one can determine the theoretical filter gain F_G of the adaptive structure for a given probability

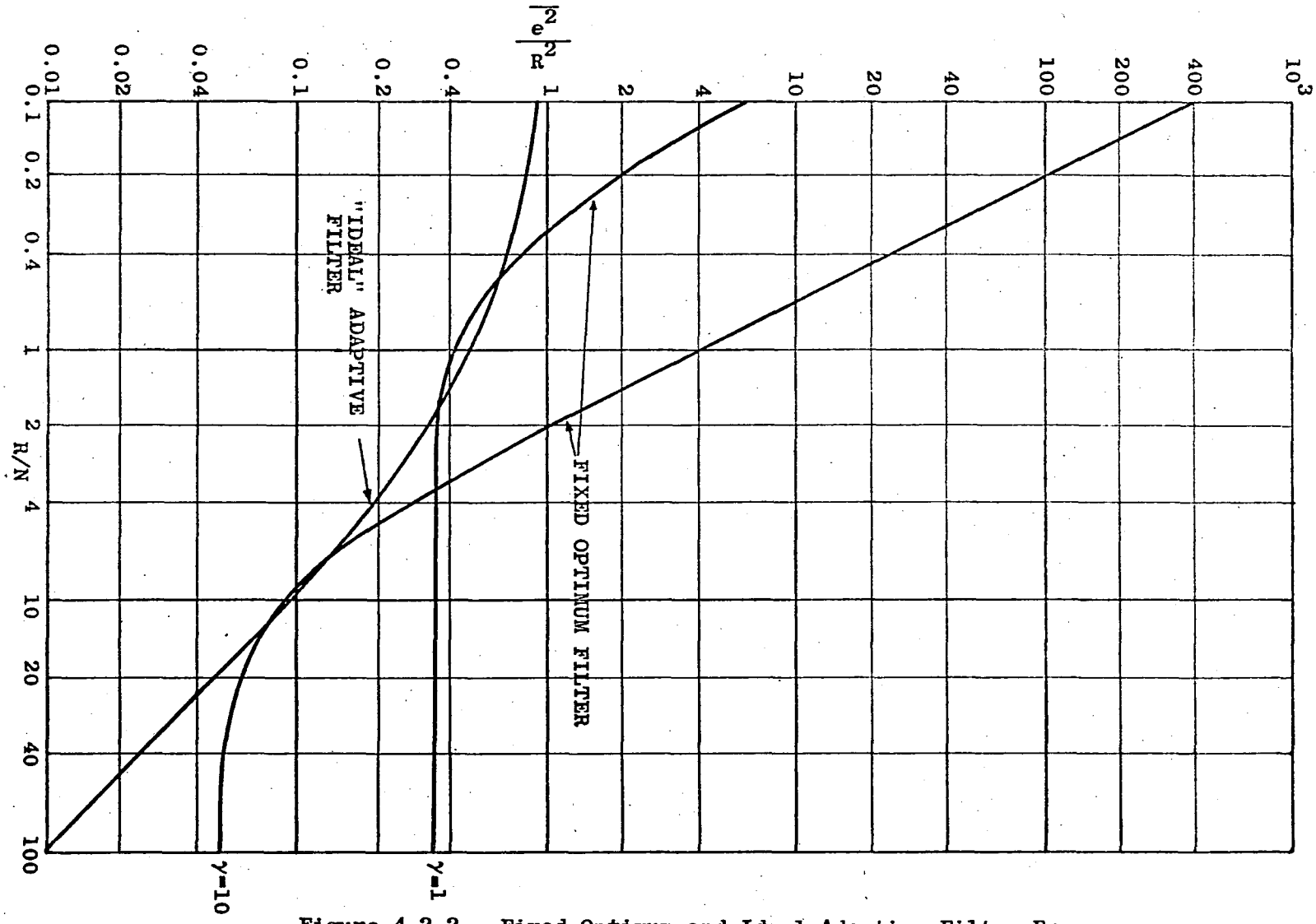


Figure 4.2-2. Fixed Optimum and Ideal Adaptive Filter Error Vs. Parameter R/N for Example No. 1

distribution of the random parameter variable R/N. For example if R/N is known to be distributed "equally likely" on the range $1 \leq R/N \leq 100$ and the fixed filter has been designed for the logrithmic median of this range, i.e., $\gamma = 10$, then the ideal adaptive filter will give an expected F_G of

$$E\{F_G\} = 10 \log_{10} \left\{ \frac{1}{99} \int_1^{100} \frac{[1+x][1+x^{-2}(\sqrt{1+\gamma^2} - 1)^2]}{2\sqrt{1+\gamma^2}} \, dx \right\} \quad (4\text{-}15)$$

$$E\{F_G\} = 4.6 \text{ db.}$$

Thus, for this example, if the signal were fading such that the input signal-to-noise ratio has a dynamic range from 0 db to 40 db, the adaptive structure will theoretically give an average improvement in output signal-to-noise ratio of 4.6 db. At the extremes of the signal strength, i.e., for R/N = 1 and R/N = 100, the ideal adaptive filter gain is 9.15 db and 9.6 db, respectively. At the minimum input signal-to-noise value, a 9 db increase in signal-to-noise ratio could significantly improve the over-all threshold of a receiver.

The experimental verification of the performance of the adaptive filter for this example is presented in the next section.

4.3 SIMULATION OF EXAMPLE NO. 1 - WHITE NOISE CORRUPTED LOW-PASS SIGNAL

The stochastic processes which have been identified as signal and noise in example No. 1 were generated by two, stochastically independent, GR-1390-B random noise generators. Each generator produced a random signal having a Gaussian amplitude distribution and a flat ($\pm$1 db) power spectral density from 20 cps to 20 kcps. The generator output used to create the signal process was low-pass filtered by a first order filter having an upper cutoff frequency of 398 cps (2500 rps). For all practical purposes the noise, which was the output from the second generator, was flat over and far beyond the signal bandwidth of approximately 400 cps.

The input to the adaptive structure was formed by combining the two stochastic processes in an analog summing amplifier, the output of which has the spectral representation of (4-16).

$$\Phi_{ii}(s) = \Phi_{ss}(s) + \Phi_{nn}(s) = \frac{R^2}{1 - s^2/(2500)^2} + N^2. \qquad (4\text{-}16)$$

For experimental purposes N^2 was held constant and R^2 was varied by varying the attenuator on the output of the random noise generator used to form the signal process.

As has already been discussed in section 4.2, the adaptive filter transfer function for this example is

$$H_a(s) = \frac{R/N}{1 + R/N + s^2/(2.5 \times 10^3)^2} \qquad (4\text{-}17)$$

A differential analyzer realization of the form shown in figure 3.6-1, with parameter B fixed, was considered for the simulation. However, since the numerical value of B required for the experiment was >>1, and consequently considerable gain was required both to realize B = v and $A = \frac{R}{N}\, v$, another hardware form was chosen to realize $H_a(s)$, as shown in figure 4.3-1.

Two point-PSD estimates were used to identify the parameters R and N for this example. The low-frequency filter used to estimate $R^2 + N^2$ was arbitrarily selected as a flat band-pass filter having a 3 db bandwidth of 190 cps with a center frequency of 160 cps. The high-frequency filter used to estimate N^2 was a Gaussian band-pass filter centered at 10 kcps and having a 3 db bandwidth of 3.4 kcps.

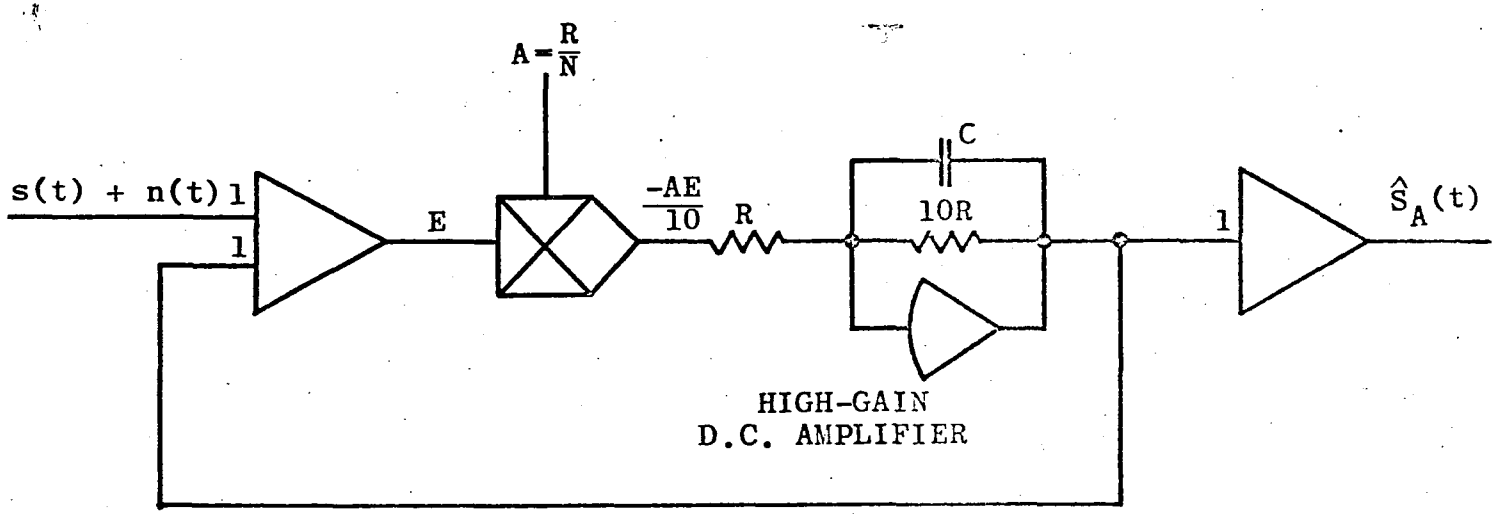
Utilizing the maximum integration time available from the TR-48 equipment in repetitive-operation, which is 500 milliseconds, the coefficients of variation for the two estimates are, from (3-11),

$$\Delta\Phi(\omega_1) = \frac{1}{\sqrt{2(190)0.5}} = 0.0725$$

and

$$\Delta\Phi(\omega_2) = \frac{1}{\sqrt{2(3.4 \times 10^3)0.5}} = 0.0172.$$

In terms of confidence intervals (C.I.) for the two estimates, the estimate of $R^2 + N^2$ had a 96% C.I. of 1.79 db about the true value and the estimate of N^2 had a 96% C.I.



$$RC = \frac{1}{10v}$$

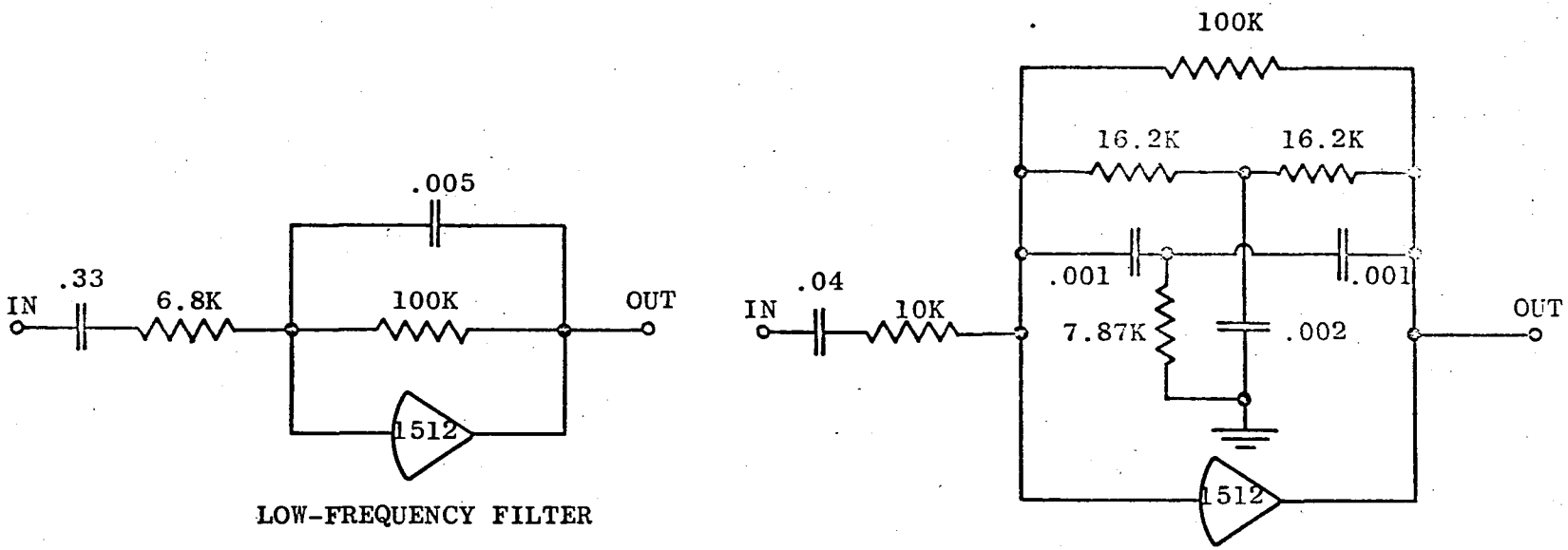
Figure 4.3-1. Realization of $H_A(s) = \frac{A}{1 + A + S/v}$, for v Fixed

of 0.43 db about the true value. The operational amplifier configurations constructed to realize the filters used for PSD estimation at f_1 = 160 cps and at f_2 = 10 kcps are shown in figure 4.3-2.

The linear detector circuits which follow the band-pass filters, were constructed in accordance with figure 3.4-2 using transistorized Burr-Brown 1512 operational amplifiers and GE-IN3873 computer diodes.

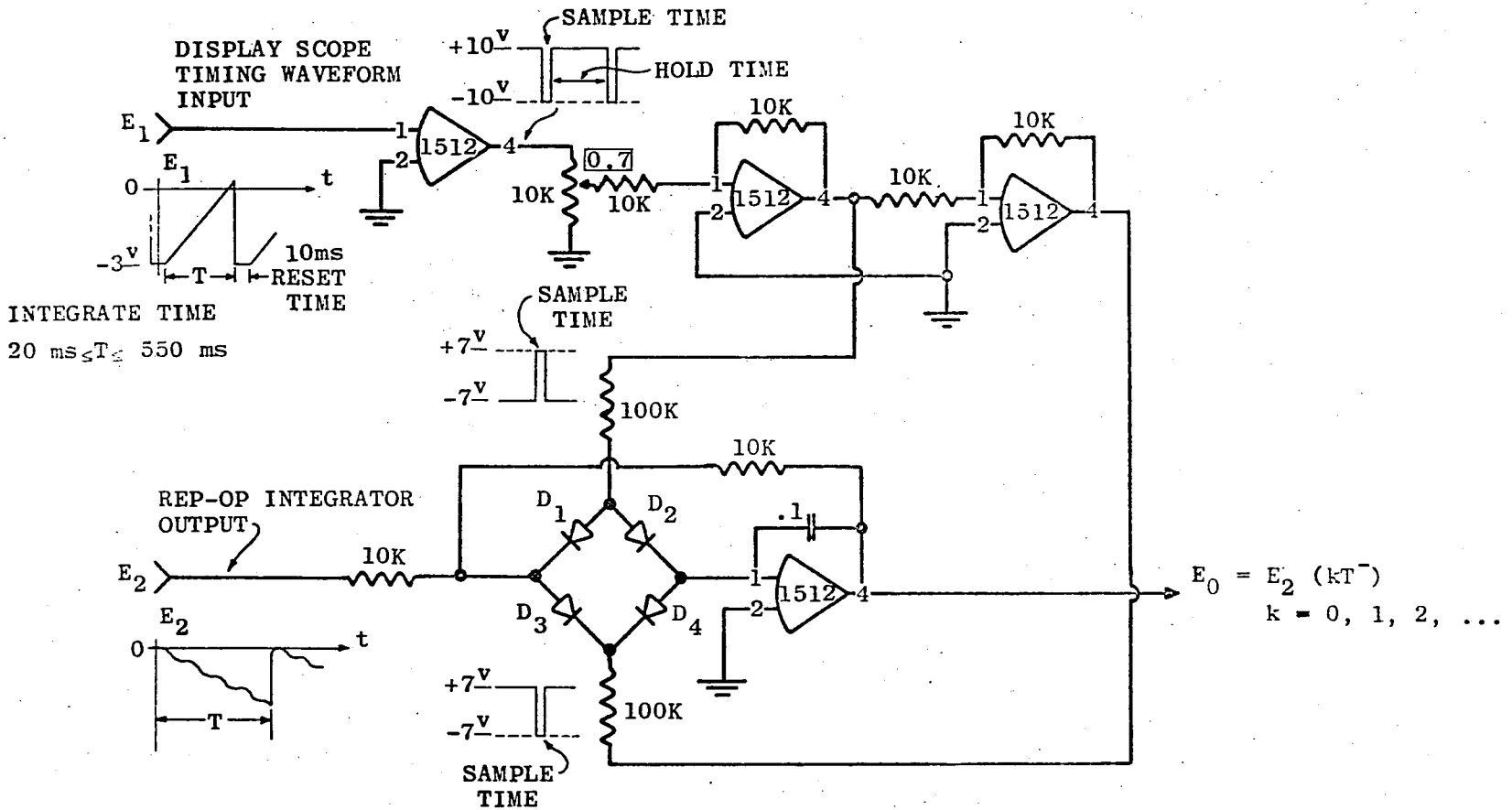
In order to provide the parameter computation section of the adaptive structure with the final sample-held values of the outputs from the channel identifier it was necessary to construct sample-hold circuits and a timing generator. The function of the timing generator was to cause the sample taken from the REP-OP integrators to be slightly in time advance of the integrators' reset operations. Since an internally generated sawtooth timing waveform is available at the patch panel of the TR-48, this waveform being intended for the synchronous operation of a display scope, the timing waveform was used to generate a time-advanced pulse sequence necessary for the operation of the sample-hold circuits. Figure 4.3-3 shows the circuit diagram of the timing generator and one sample-hold circuit. A separate sample-hold circuit is required for each REP-OP integrator output from the channel identification block.

The complete adaptive filter simulation for example No. 1 is shown schematically in figure 4.3-4.



ALL CAPACITOR VALUES IN MICROFARADS
ALL OPERATIONAL AMPLIFIERS ARE BURR-BROWN
TRANSISTORIZED UNITS

Figure 4.3-2. Selective Amplifier Realizations for the PSD Estimates of Example No. 1



ALL OPERATIONAL AMPLIFIERS ARE BURR-BROWN
ALL C'S ARE IN MICRO-FARADS
ALL VOLTAGES SHOWN ARE REFERENCED TO EQUIP. GROUND
D_1, D_2, D_3, D_4 ARE IN4009

Figure 4.3-3. Timing Generator and Sample-Hold Circuit for PSD Estimator

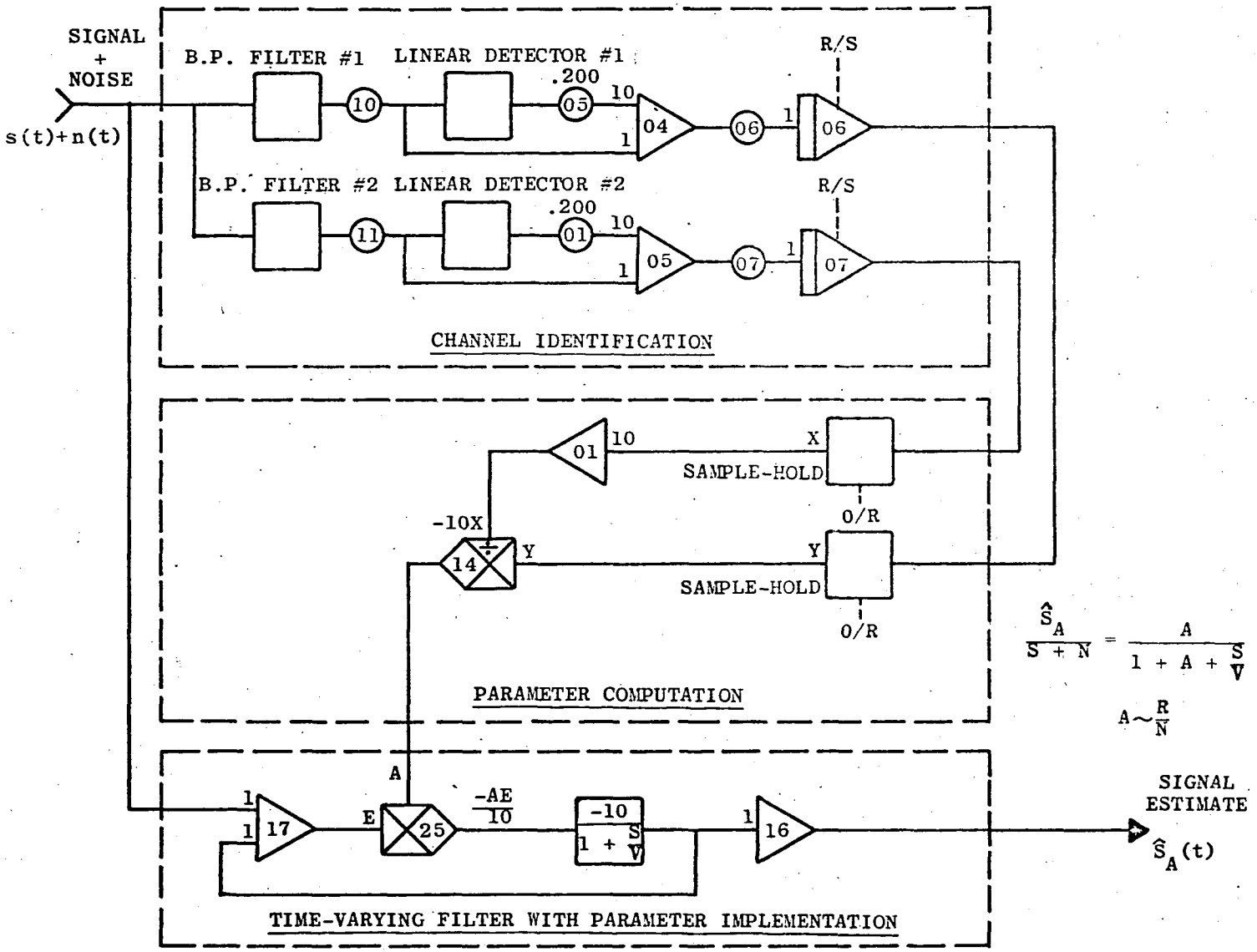


Figure 4.3-4. Adaptive Filter Simulation - Example No. 1

In order to experimentally determine the adaptive filter performance, the error signal

$$\epsilon_a(t) = s(t) - \hat{s}_a(t) \tag{4-18}$$

was formed using a differencing amplifier and then squared using an X^2 diode function generator, which yields

$$\epsilon_a^2(t) = [s(t) - \hat{s}_a(t)]^2. \tag{4-19}$$

The average squared-error was experimentally determined by a weighted integration of $\epsilon_a^2(t)$, yielding

$$\overline{\epsilon_a^2(t)} = \frac{1}{T_o}\int_o^{T_o} [s(t) - \hat{s}_a(t)]^2 dt. \tag{4-20}$$

The integration T_o used in (4-20) was made long with respect to the REP-OP integration time T. In all evaluations $T_o \geq 30$ seconds, and T = 0.5 second.

A similar computational scheme was used to obtain the average squared-error of the filtered output from a computer simulated fixed optimum filter. The analog computer realization for comparing the mean-square error of the fixed optimum and adaptive filters is shown in figure 4.3-5.

Two cases of fixed optimum filter design were compared with the adaptive filter for this example. Figure 4.3-6 shows the experimental data and theoretical curve for F_G in decibels for the fixed filter designed at an R/N ratio of 1.22 (1.73 db). The experimental data may be seen to

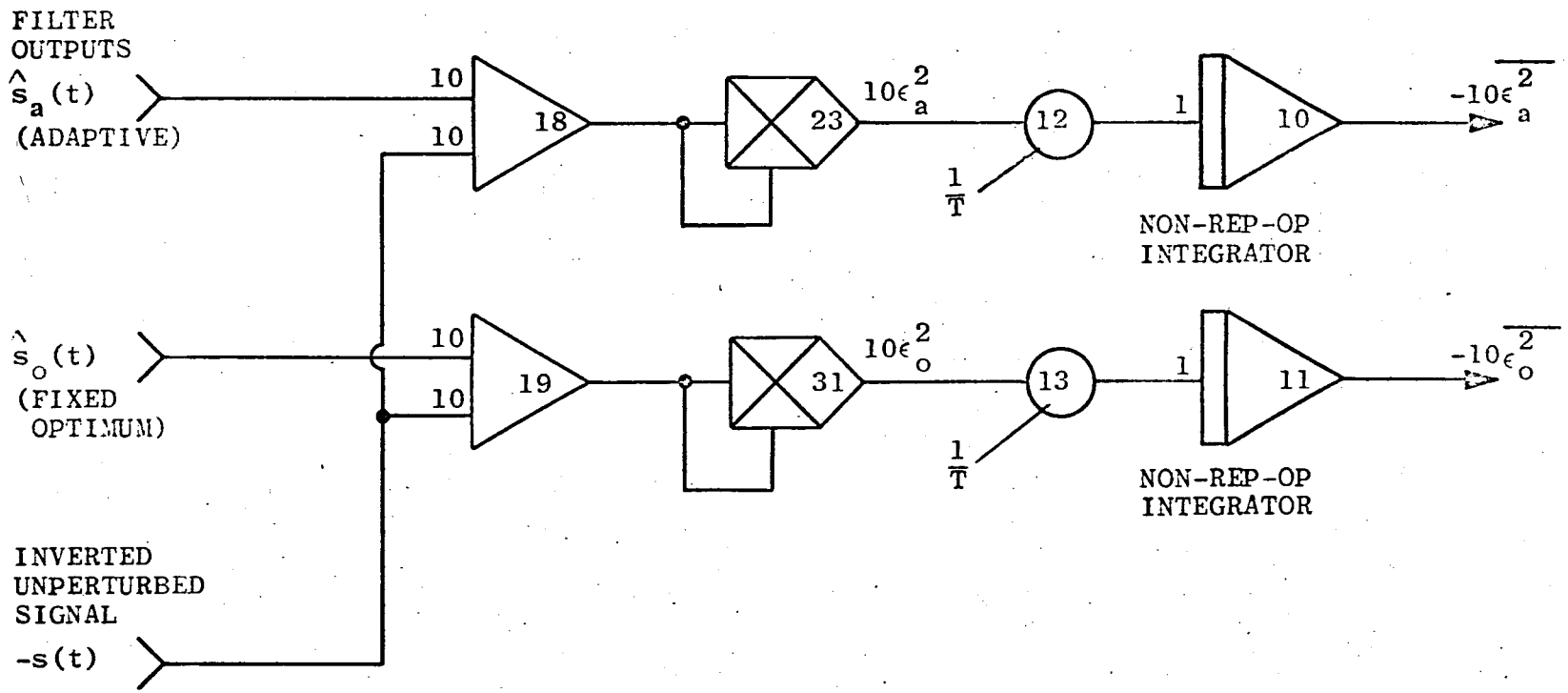


Figure 4.3-5. Analog Method of Computation of Average Squared Filtering Errors

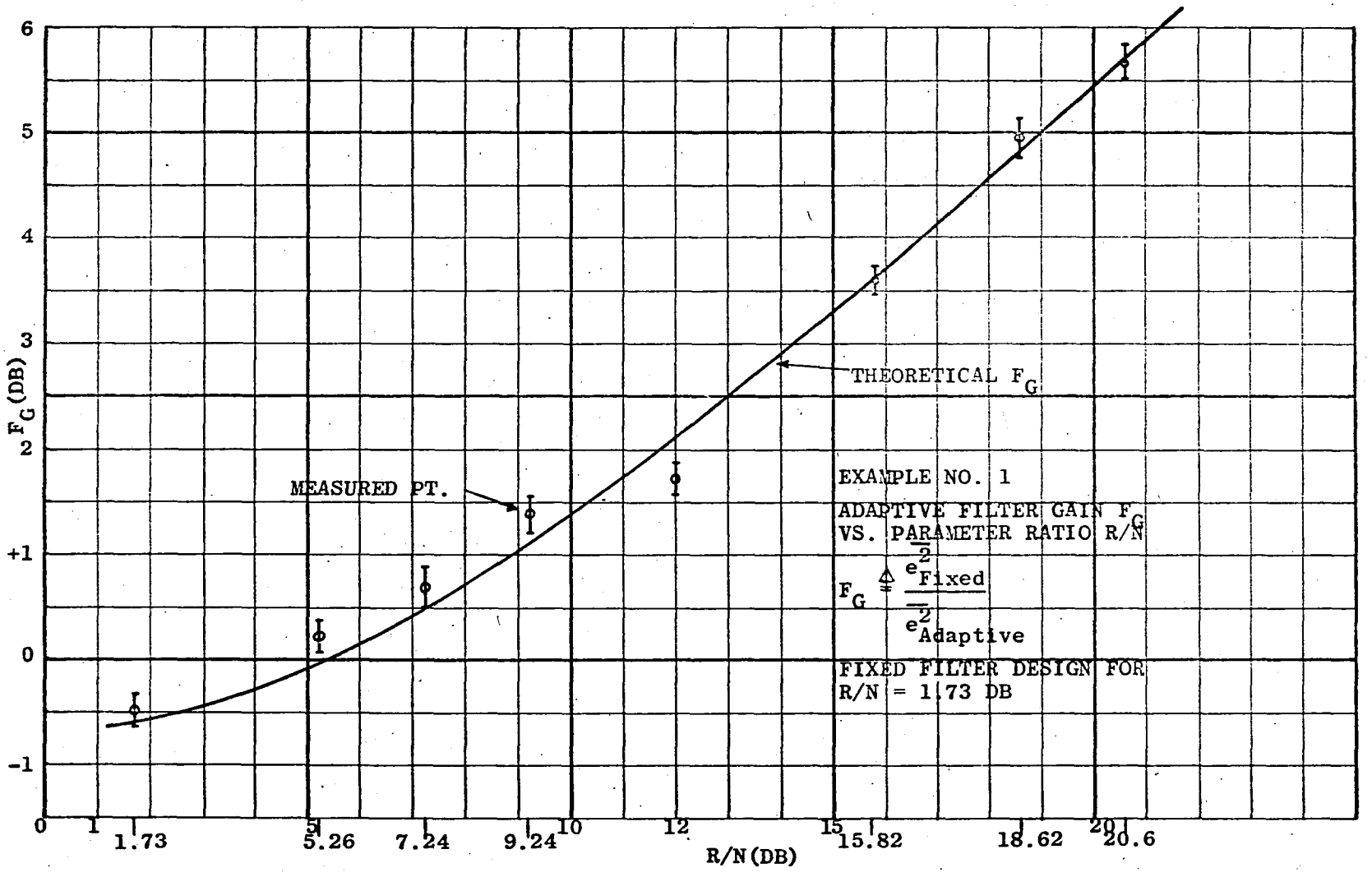


Figure 4.3-6. Experimental and Theoretical F_G Vs. R/N for Adaptive Filter Example No. 1

agree with the theoretical to within $\pm$0.5 db as R/N varies over a 20 db range.

Figure 4.3-7 gives the results of the measured adaptive filter gain vs the theoretical for the fixed optimum filter designed at an R/N ratio of 10.7 (20.6 db). The measured data lie within a spread of $\pm$0.5 db of the theoretical for R/N from 5.26 db to 18.62 db. The experimental F_G was 1.5 db low at the lower extreme value of 1.73 db and 0.95 db low at the upper extreme value of 20.6 db.

The data presented in figures 4.3-6 and 4.3-7 were taken for a static condition on R/N; that is, several seconds were allowed after a change was made in R/N before the mean squared-errors of fixed and adaptive filters were computed.

In order to assess the effect of time-variations in R/N on the adaptive filter's performance, the step function response of the channel identification and parameter computation portion of the adaptive structure was determined. Figure 4.3-8 shows the response of the parameter control signal A(t) (see figure 4.3-4) to a step change in the random process R/N ratio from 1.22 to 10.7, i.e., from 1.73 to 20.6 db. This result was obtained from a recording using a Moseley Model 2D X-Y recorder. The measured 10-90% rise time of the Moseley recorder was 40 milliseconds.

Since the measured rise-time of the control circuitry-recorder cascade as shown in figure 4.2-8 was 95 milliseconds, the actual rise time of the control circuitry may be estimated

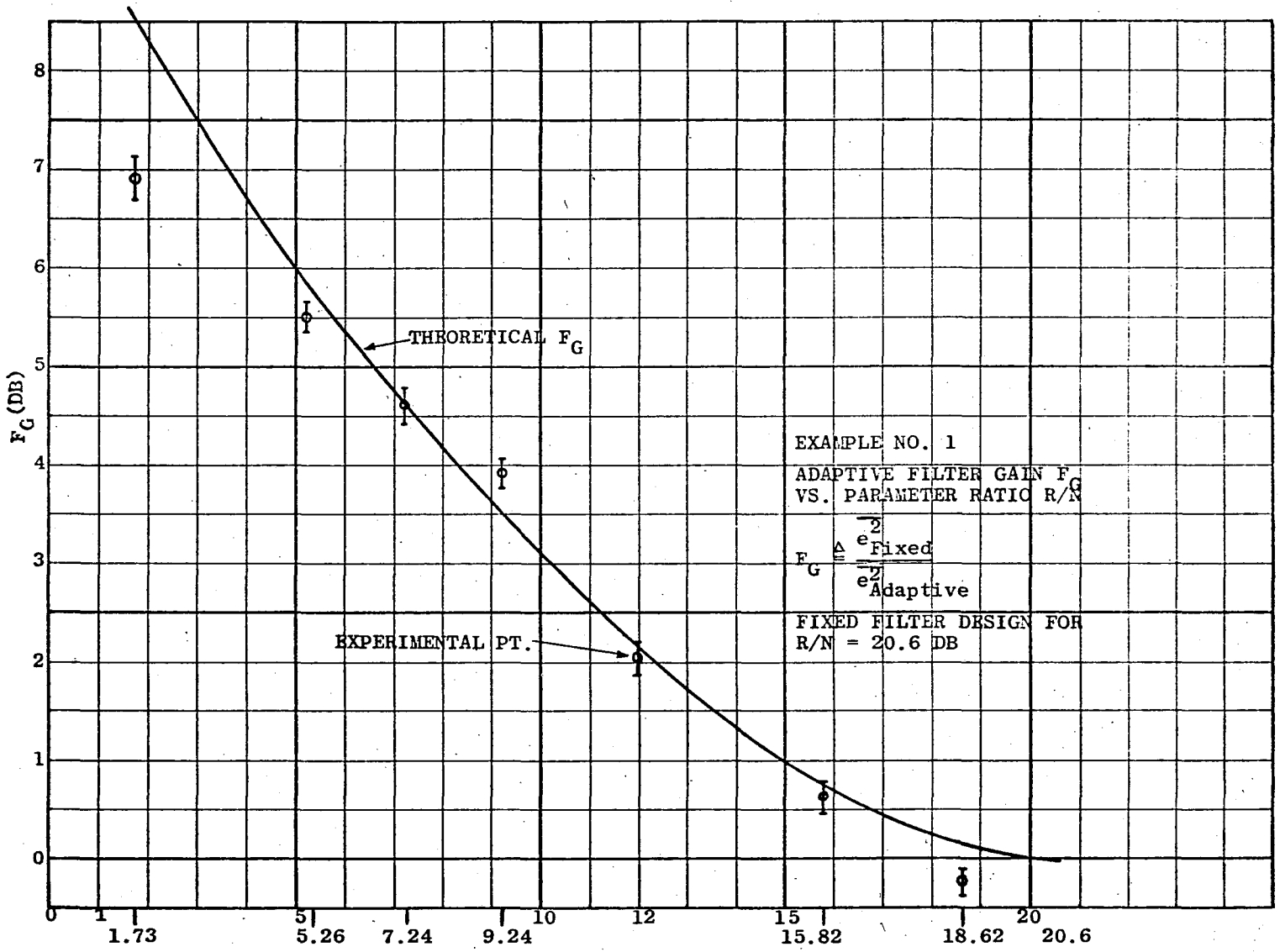


Figure 4.3-7. Experimental and Theoretical F_G Vs. R/N for Adaptive Filter Example No. 1

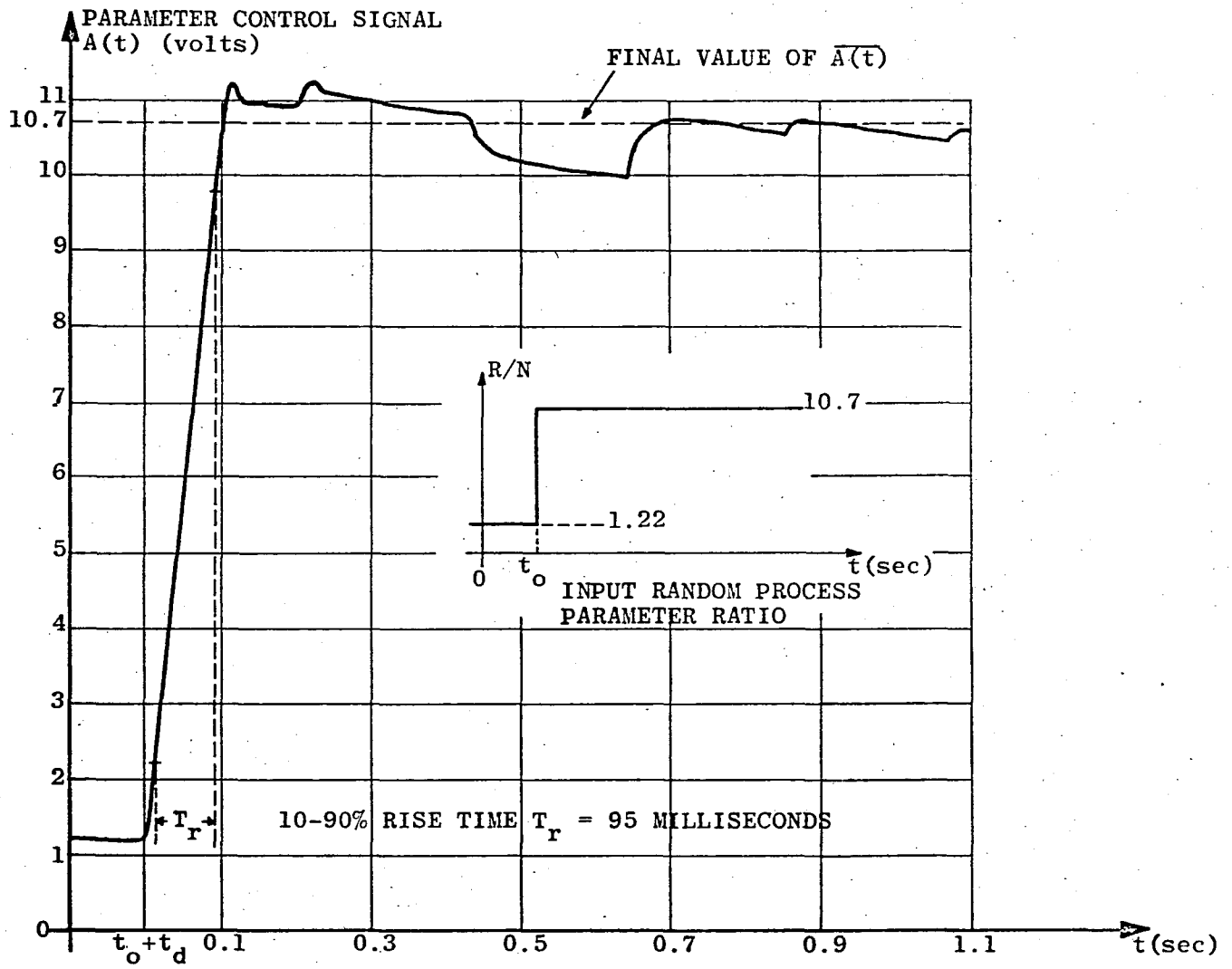


Figure 4.3-8. Step Function Response of Parameter Control Signal - Example No. 1

from the theory of rise time composition in cascade circuitry.[1] The control circuit rise time is approximately given by

$$T_r \approx \sqrt{(95 \times 10^{-3})^2 - (40 \times 10^{-3})^2}$$

$$T_r \approx 86 \text{ milliseconds.}$$

Thus, the break frequency for the channel identification parameter computation portion of the adaptive structure is approximately

$$f = \frac{1}{2.2(2\pi)T_r} = 1.85 \text{ cps.} \qquad (4\text{-}21)$$

The delay time, noted in figure 4.3-8 as t_d, is a random variable, the value of which for any particular performance of the step response experiment is determined by the time separation between application of the step change in R/N and the sample-time of the PSD estimating circuitry. Thus, once a PSD estimate of the value of R/N has been made for $t > t_o$, the sample-hold and parameter computation circuitry will respond with an 86 millisecond 10-90% rise time.

Since a REP-OP integration time of 500 milliseconds was used, this being primarily determined by a desire to reduce the variance in the PSD estimates, the 86 millisecond response time of the remaining circuitry appears negligible by comparison.

---

1. See any modern text on amplifier theory, e.g., Martin [43], pp. 222-224.

Thus, in this example, the time required for PSD estimation limits the ability of the adaptive structure to follow the nonstationarities in the input process.

4.4 ADAPTIVE FILTER EXAMPLE NO. 2 - DESIGN AND ANALYSIS

The case of two band-pass signals having mutually interferring spectra is considered in this example.

The complex spectral density of signal a(t), the signal to be recovered, is given by

$$\Phi_{aa}(s) = \frac{-A^2(s^2/u^2)}{(1-s^2/u^2)(1-s^2/v^2)} \qquad (4\text{-}22)$$

and the complex spectral density of signal b(t), the interferring signal, by

$$\Phi_{bb}(s) = \frac{-B^2(s^2/w^2)}{(1-s^2/w^2)(1-s^2/q^2)} , \qquad (4\text{-}23)$$

where the frequencies u, v, w, q are fixed and related by $u < v \leq w < q$. For convenience let $v = w$ and let the signals have equal bandwidths, i.e., $v-u = q-w$.

The signal "strength parameters" A^2 and B^2 are stochastic variables, assumed to satisfy the quasi-stationarity condition. In order to simplify both the analysis and the hardware simulation to follow, signal a(t) will be taken as a low-pass signal, i.e., set $u = 0$ in (4-22).

The stationary optimum filter for this example has a transfer function given by

$$H_o(s) = \frac{K_o(1 + s/q)}{(1 + s/a)} , \qquad (4\text{-}24)$$

where

$$K_o = \frac{1 + v/q}{1 + v/a} ,$$

and

$$\frac{1}{a} = \left[\frac{1}{q^2} + \frac{1}{(\frac{A}{B})^2 v^2} \right]^{1/2} .$$

Using Streets' procedure, the approximate optimum filter is

$$H_a(s) = \frac{K_a(1 + s/q)}{(1 + s/b)} , \qquad (4\text{-}25)$$

where

$$K_a = \frac{A/B}{1 + A/B} ,$$

and

$$\frac{1}{b} = \frac{\frac{A}{B}\frac{1}{q} + \frac{1}{v}}{1 + A/B} .$$

The plot of the power spectral densities of signals a(t) and b(t) is shown in figure 4.4-1, along with the location of two narrow band-pass filters for PSD estimation at $\omega_1 < v$ and at ω_2, where $v < \omega_2 < q$. The use of square-law detectors to estimate the power at ω_1 and ω_2 yields, respectively

$$P_1 = A^2 + (\frac{\omega_1}{v})^2 B^2$$

and (4-26)

$$P_2 = B^2 + (\frac{v}{\omega_2})^2 A^2 .$$

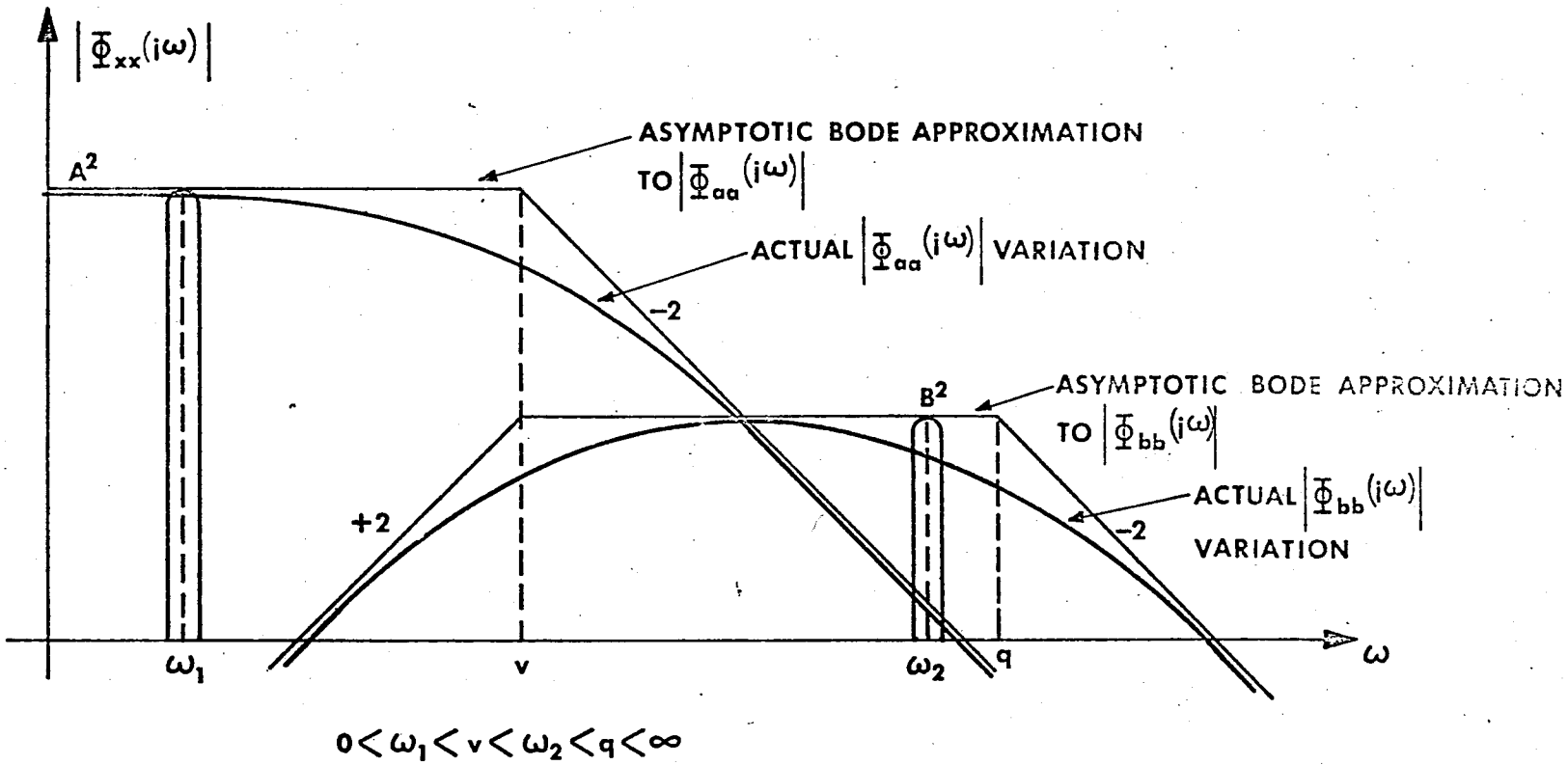


Figure 4.4-1. Plot of $\Phi_{aa}(j\omega)$ and $\Phi_{bb}(j\omega)$ for Adaptive Filter Example No. 2

The time-varying parameter, required for the adaptive filter realization of (4-25), is A/B, which from (4-26) is given by

$$\frac{A}{B} = \sqrt{\frac{P_1 - \left(\frac{\omega_1}{v}\right)^2 P_2}{P_2 - \left(\frac{v}{\omega_2}\right)^2 P_1}} \quad . \tag{4-27}$$

The theoretical improvement of the adaptive filter over the fixed optimum filter is again computed by utilizing relations (4-9) and (4-10). The evaluation of the equivalent contour integral (4-10) is greatly simplified by the prepared tables found in Newton, Gould, and Kaiser [16], pp. 366-381. The calculated mean-square error expressions using (4-24) and (4-25) are of the same form and hence only the result for $\overline{e_o^2}$ is given below.

$$\overline{e_o^2} = \frac{\left[A^2\left(\frac{q-K_o a}{aq}\right)^2 + K_o^2 B^2 \left(\frac{1}{v}\right)^2\right](av)^2 + B^2(1-K_o)^2 av}{2(a+v)} \quad . \tag{4-28}$$

The parameters K_o and a are constants in the fixed optimum filter design and are determined by the A/B ratio selected for the design of this filter. For the adaptive structure, the parameters K_a and b are time-varying in correspondence to the nonstationary behavior of A/B as estimated by (4-27).

Because of the numerical complexity of (4-28), the theoretical F_G curves vs A/B are not shown until q, v, and the range of A/B have been selected for simulation in

section 4.5. The theoretical and measured F_G curves for this example are shown in figure 4.5-4.

Both the theoretical F_G and the measured values shown for this example indicate that an improvement in filter output signal-to-noise ratio of approximately 9 db is realized for a 20 db departure in the A/B ratio from the fixed filter design value.

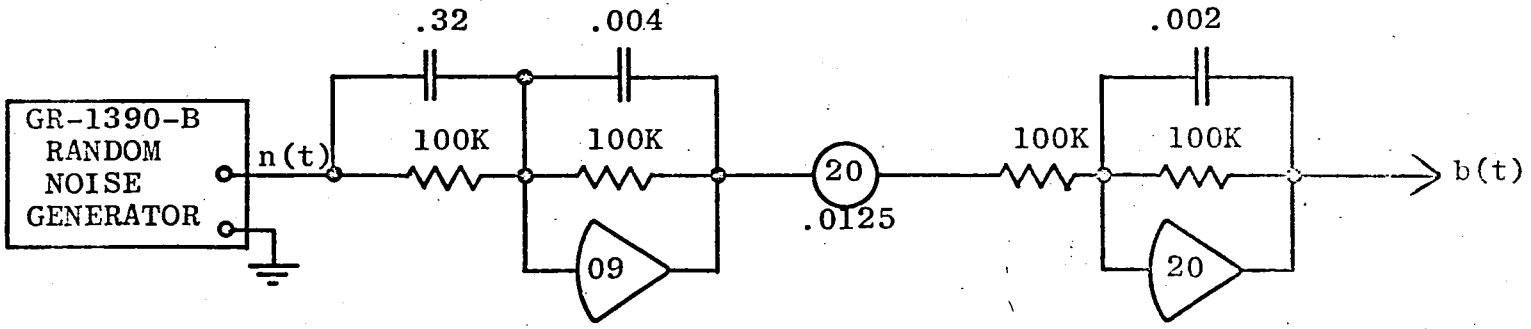
4.5 SIMULATION OF EXAMPLE NO. 2 - MUTUALLY INTERFERRING BAND-PASS SIGNALS

The stochastic process a(t), which was identified in section 4.4 as the lower frequency band-pass signal in example no. 2, was produced by shaping the output of a GR-1390-B random noise generator using a first order low-pass filter having an upper break frequency of 2500 rps. Since the lower break frequency of the noise generator was 31.4 rps, the bandwidth of signal a(t) was essentially 2500 rps.

The stochastically independent interferring random signal b(t) was produced by shaping the spectrum of a second GR-1390-B noise generator using the analog circuitry shown in figure 4.5-1.

The fixed optimum filter structure described by (4-24) and the adaptive filter (4-25) are of the same H(s) form, that is

$$H(s) = \frac{K(1 + s/z)}{(1 + s/p)} . \qquad (4\text{-}29)$$



$$\frac{B(s)}{N(s)} = \frac{1}{80} \frac{(1 + s/31.4)}{(1 + s/2500)(1 + s/5000)}$$

Figure 4.5-1. Analog Circuitry for Shaping the PSD of White Noise n(t) to Form b(t)

For both filters, the zero location s = -z was held fixed at z = 5000 rps. The design value selected for the nonstationary parameter ratio A/B determines K and p in the fixed filter, while the time-varying estimate of A/B controls K and p in the adaptive filter.

By the analysis of section 4.4 it was determined that the fixed optimum filter parameters are

$$p = \frac{A/B}{\left(\frac{A^2}{B^2}\frac{1}{q^2} + \frac{1}{r^2}\right)^{1/2}} \qquad (4\text{-}30)$$

and

$$K = \frac{1 + v/q}{1 + v/p}\ , \qquad (4\text{-}31)$$

where

v = 2500 rps, and q = 5000 rps.

Experimentally it was again convenient, as in the simulation of example no. 1, to fix B and vary A by using the variable attenuator on the output of the noise generator used to form the process a(t). The range of variation used for A/B was from 1.22 to 10.7.

When the lower extreme value of A/B is selected as the design point for the fixed optimum filter, the fixed filter parameters are determined by (4-30) and (4-31) to be

p = 2600 rps and K = 0.765.

Referring once again to the analysis of section 4.4, it was found that the adaptive filter parameters are

functions of the nonstationary parameter ratio A/B as follows:

$$K = \frac{A/B}{1 + A/B} \tag{4-32}$$

and

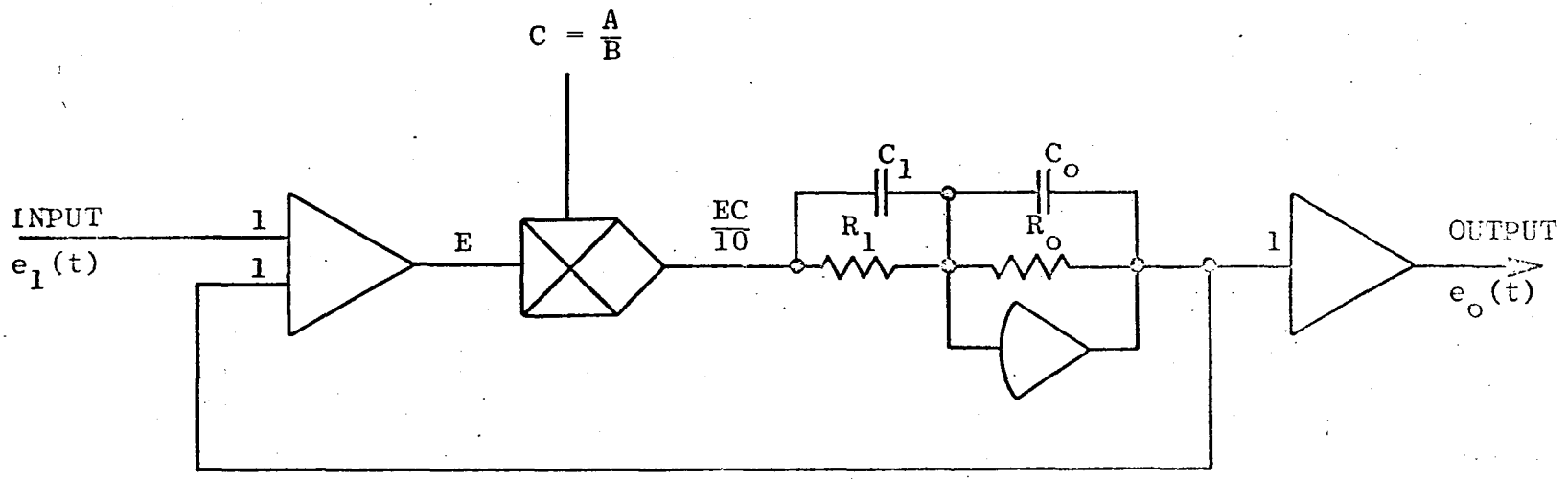
$$p = \frac{1 + A/B}{\frac{A}{B}\frac{1}{q} + \frac{1}{v}} \tag{4-33}$$

The analog realization, using a single multiplier that was used to simulate the transfer function form (4-29) with parameters (4-32) and 4-33), is shown in figure 4.5-2.

In order to form the parameter control signal c(t), it was necessary to estimate the time-varying parameter A/B. Since the analog equipment needed to estimate A/B by relationship (4-27), which is the required method when both A and B vary, was not available, it was convenient to estimate A alone. Because B was held fixed in the experimental simulation, an estimate of A was sufficient to determine A/B.

The complete simulation used for the adaptive filter is shown in figure 4.5-3.

To compare experimental and theoretical performance of the adaptive filter for this example, the technique illustrated in figure 4.3-5 was employed. The result of the measured vs theoretical F_G for the fixed optimum filter designed at A/B = 1.22 is shown in figure 4.5-4. The measured data lie within a spread of +1.4 to -0.6 db of the theoretical curve which was calculated using (4-28) with the numerical values of K_o, a, K_a and b substituted.



$$\frac{R_o}{R_1} = 10, \qquad K = \frac{A/B}{1 + A/B},$$

$$R_1C_1 = \frac{1}{z}, \qquad p = \frac{1 + A/B}{\frac{A}{B}\frac{1}{q} + \frac{1}{v}}.$$

$$R_oC_o = \frac{1}{v}$$

Figure 4.5-2. Analog Simulation of the Transfer Function $H(s) = \frac{K(1 + s/z)}{(1 + s/p)}$

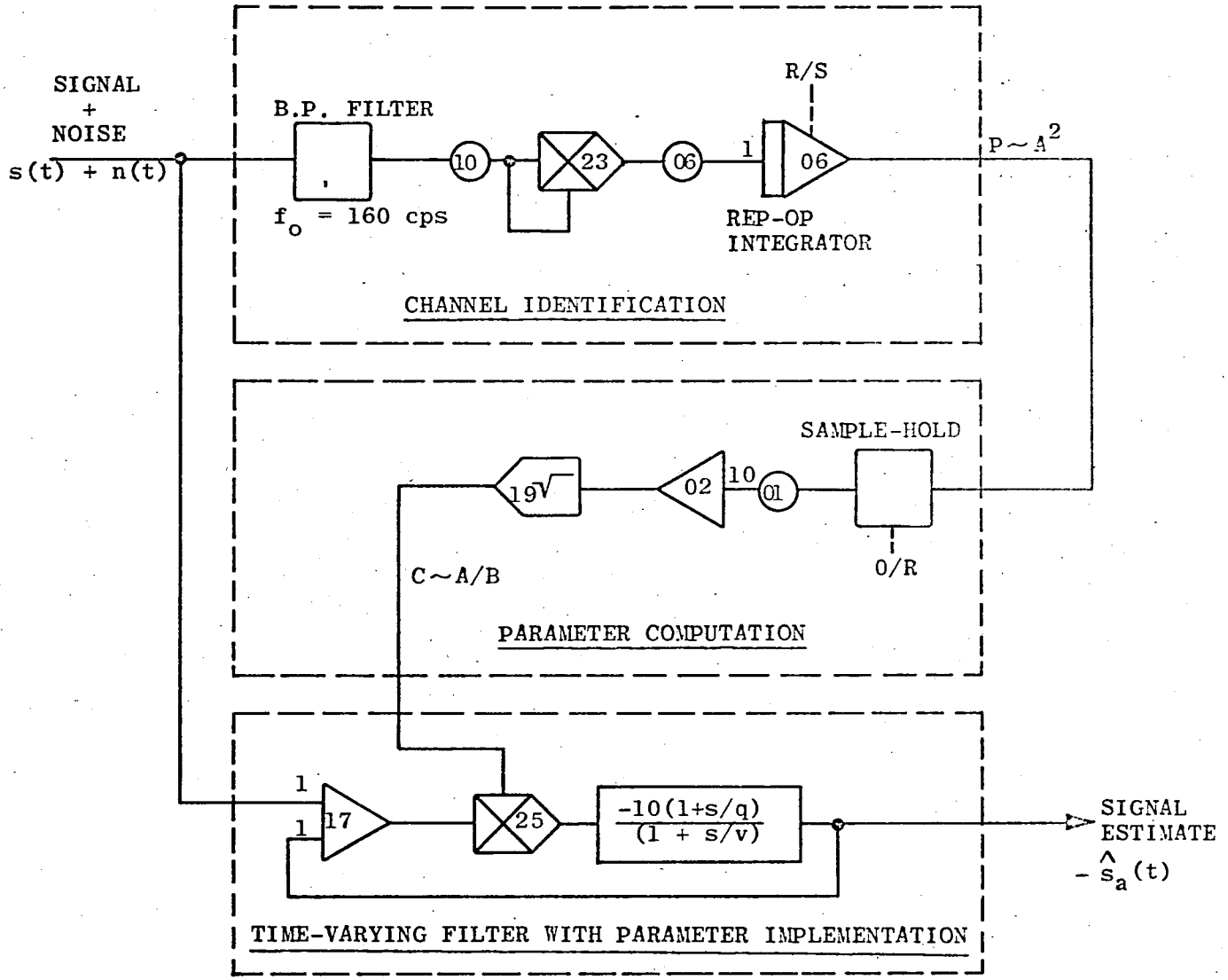


Figure 4.5-3. Adaptive Filter Simulation-Example No. 2

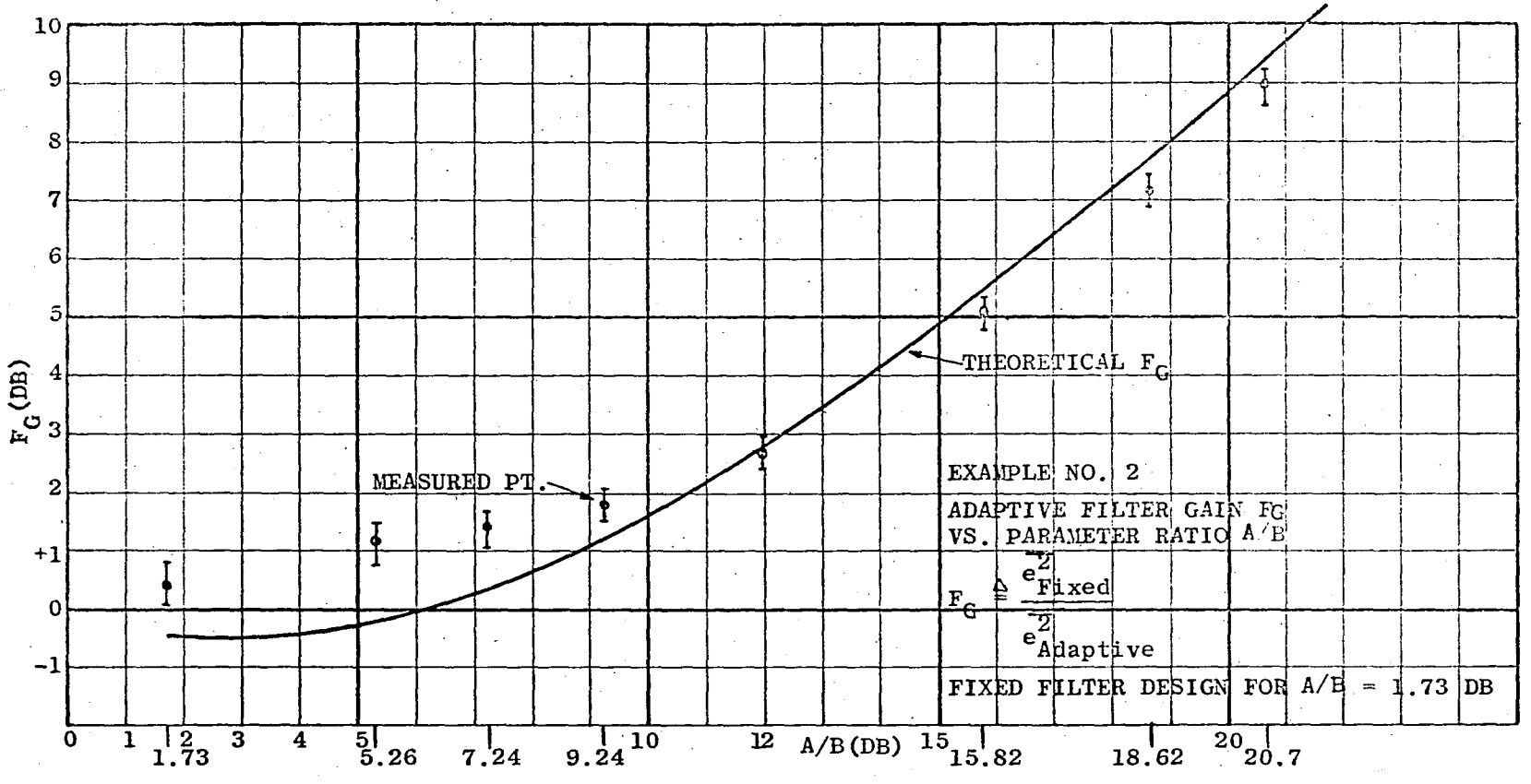


Figure 4.5-4. Experimental and Theoretical F_G Vs. A/B for Adaptive Filter Example No. 2

4.6 AN ADAPTIVE PHASE-LOCKED LOOP FM DEMODULATOR

A practical example of an empirical nature is now given to illustrate the use of the proposed adaptive filter theory for the design of an improved FM demodulator structure. The phase-locked loop (PLL) demodulator was chosen for this example because of its proven capability in weak signal detection and because of the observed time-varying nature of the received signal-to-noise ratio that such a detector experiences in a modern communication satellite system.

The system used for illustration is the Echo-II passive satellite S- and X-band transportable terminal. The gross received signal strength variation pattern is shown in figure 4.6-1 as a function of earth central angle and time from zenith for an antenna-tracked overhead pass of the satellite [44]. A close approximation to the function shown in figure 4.6-1 on a linear scale, is given by the relationship

$$S(t) = S(0)\, \bar{e}^{|t|/\tau}, \qquad (4\text{-}34)$$

for each orbit in the total communications interval T.[1]

Additionally, since the receiver antenna tracks the overhead pass of the satellite, the gross received noise power variation varies with antenna elevation angle.

---

1. The parameters used in equation (4-34) are τ = 113 sec, T/2 - 530 sec., S(0)/S(T/2) = 100.

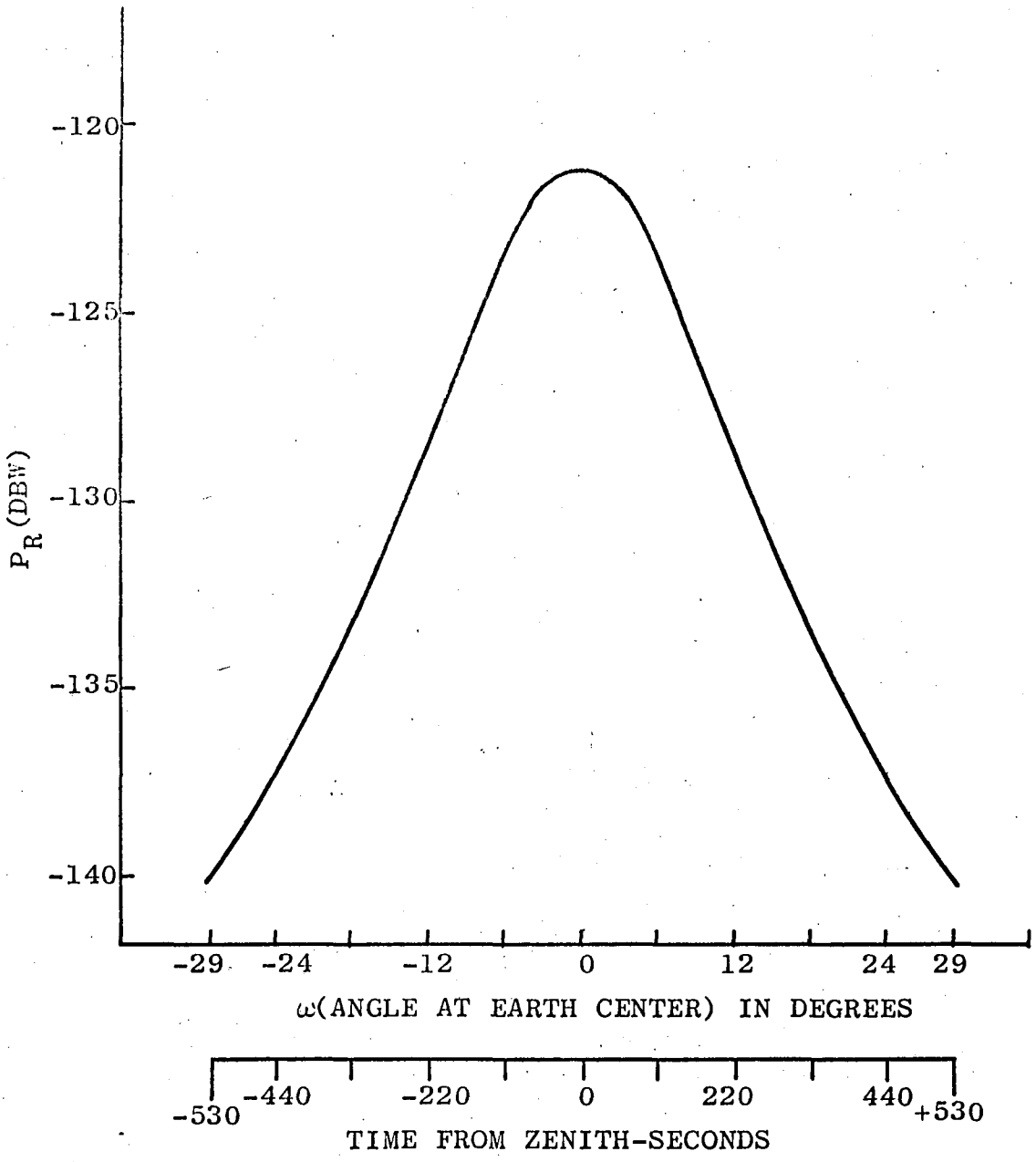


Figure 4.6-1. Signal Strength for Overhead Pass

Pritchard [45] indicates that antenna noise temperature varies with frequency and elevation angle as shown in figure 4.6-2. Using 8 kmc as the FM carrier frequency, one has that the received noise power varies by a factor of 7:1 as antenna elevation angle θ varies from 5^{o} to 90^{o} with respect to the horizontal.

Upon normalizing the received carrier-to-noise ratio (CNR), such that for an elevation angle $\theta = 5^{o}$ the CNR will be 10 db,[1] one has that the CNR as a function of time and angle at earth's center is as plotted in figure 4.6-3. The assumption made here is that the receiver front-end noise is extremely low ($T < 10^{o}K$). Thus a cryogenic maser type of supersensitive receiver front-end is implied.

The limiting rate at which information can be transferred in an ideal, strictly band-limited channel is given by Shannon's channel capacity formula

$$C = W \log_2 (1 + S/N), \text{ bits per second.} \qquad (4\text{-}35)$$

Since the receiver output signal-to-noise ratio varies in proportion to the received CNR according to

$$\left(\frac{S}{N}\right)_o \doteq 3M^2 \, C/N_i, \qquad (4\text{-}36)$$

where $(S/N)_o$ is the output signal-to-noise ratio of the PLL detector, C/N_i is the received CNR, and M is the FM modulation

---

1. This corresponds to an FM threshold of approximately 25 db for the output SNR of the PLL detector.

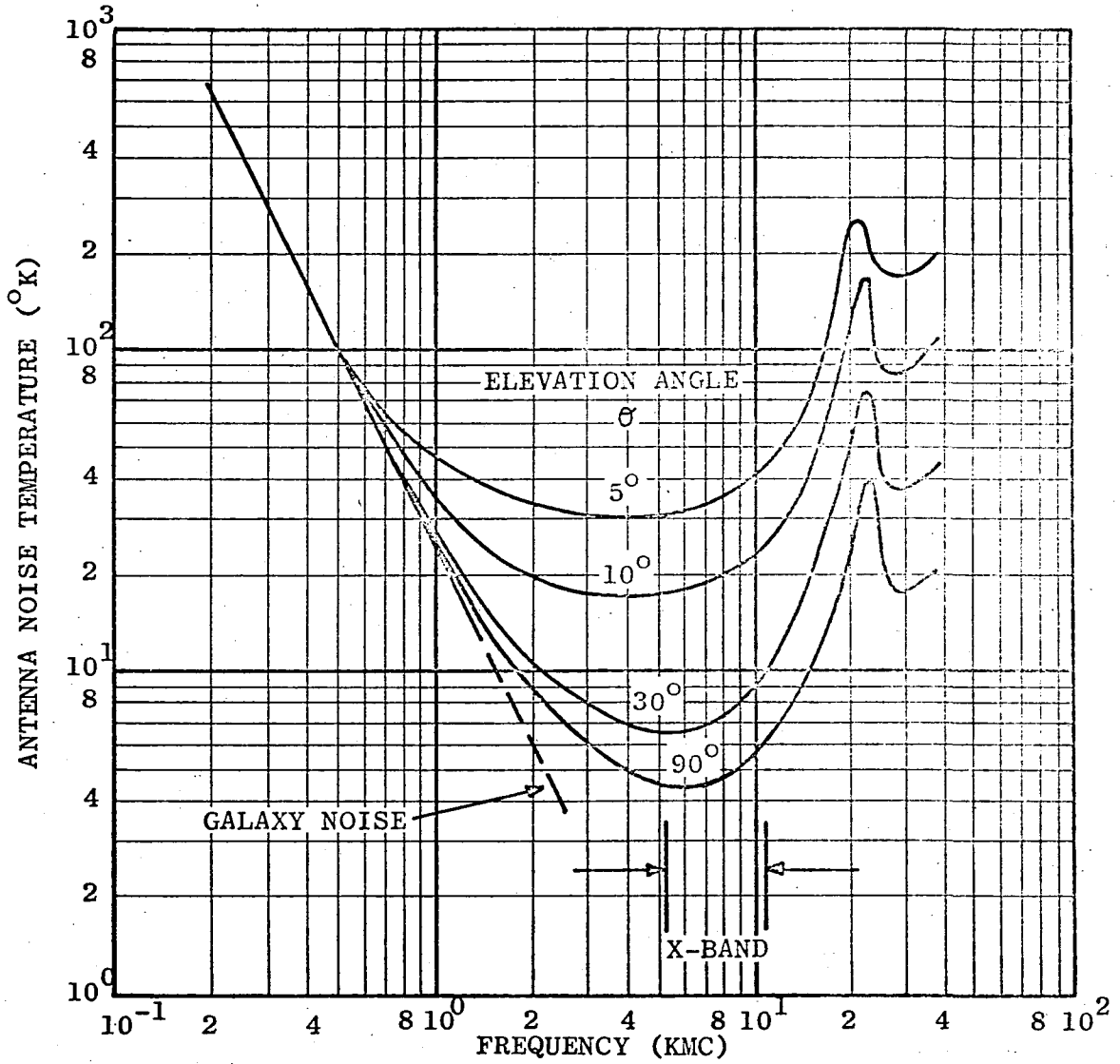


Figure 4.6-2. Antenna Noise Temperature Vs. Frequency With Antenna Elevation Angle as a Parameter

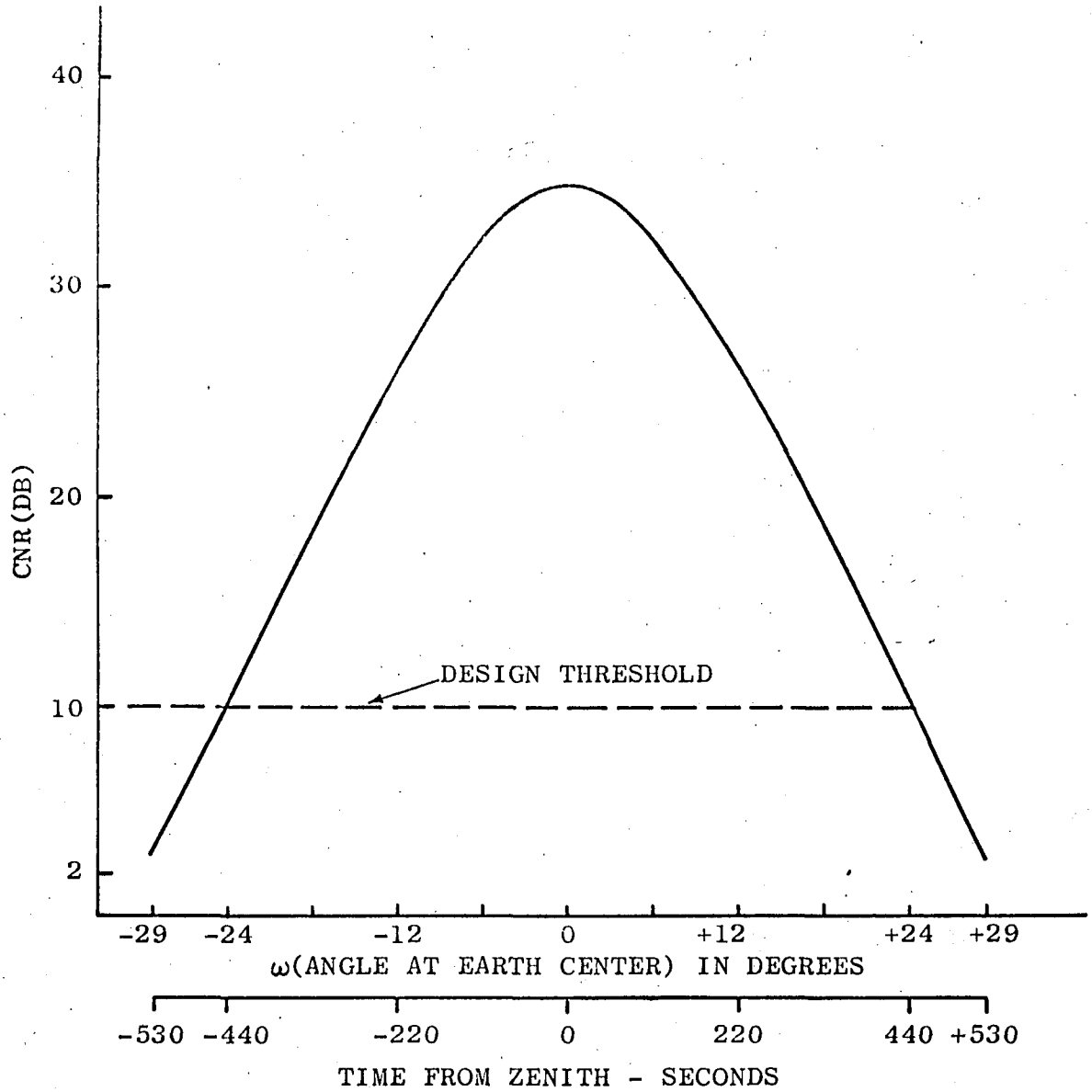


Figure 4.6-3. Carrier-to-Noise Ratio, Received for an Overhead Pass

index; it follows that the information capacity of the satellite communication system varies with time.

Thus, if the PLL detector were to operate within a constant margin of the threshold condition, i.e., at a constant output SNR, the transmitter could transmit with a time-varying information rate. This would allow a larger total amount of information to be transferred during each satellite pass of T seconds than could be transferred in a fixed capacity system.

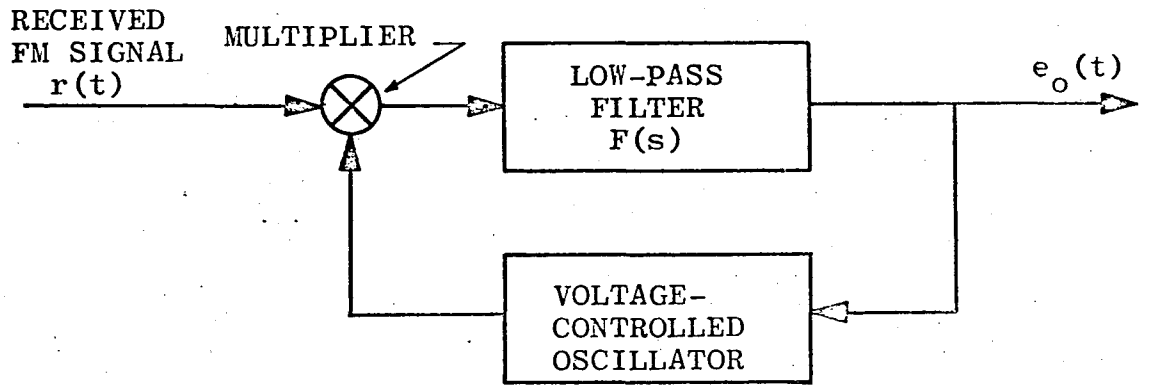
A block diagram of the PLL detector is shown in figure 4.6-4, along with the equivalent linear form.

Analytically, the problem of optimum design may be stated as follows [46]:

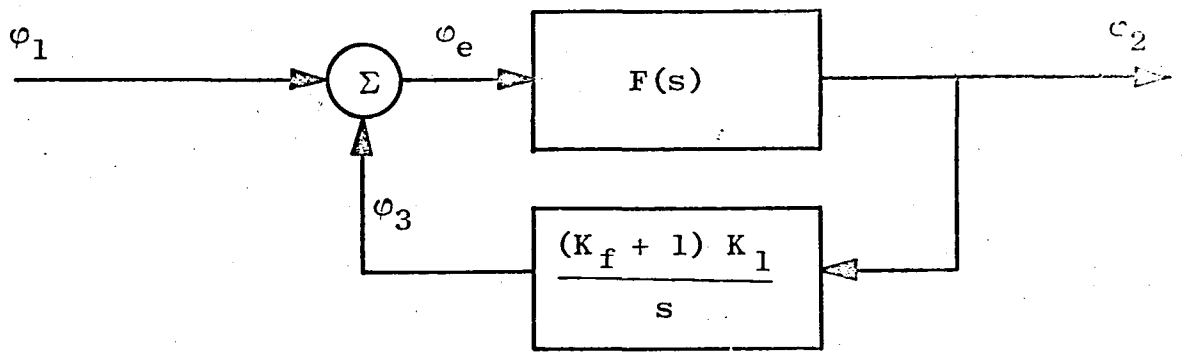
$$\sigma_N^2 + \lambda^2 E_T^2 = \Psi^2 = \text{minimum}, \qquad (4\text{-}37)$$

where σ_N^2 is the mean-square phase or frequency jitter due to noise interference and E_T^2 is the transient error measure due to non-ideal tracking of the signal phase or frequency variations. Lambda is a Lagrangian multiplier as used in the calculus of variations technique.

The minimization of detector error (4-37) is to be accomplished subject to a semi-fixed configuration constraint as depicted in figure 4.6-4(b). Under these conditions, the optimum selection of F(s), K_f, and K_1 becomes a function of the input PSD, i.e., $\Phi_{ss}(s) + \Phi_{nn}(s)$.



(a) PHASE-LOCKED LOOP



(b) EQUIVALENT LINEAR FORM

Figure 4.6-4. Phase-Locked Loop Diagrams

When one has signal and noise PSD's which satisfy the quasi-stationarity assumption, the solution of (4-37) yields a time-varying system. The time variations in $F(s)$, K_f, and K_1 are then determined by the nonstationary behavior of $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$.

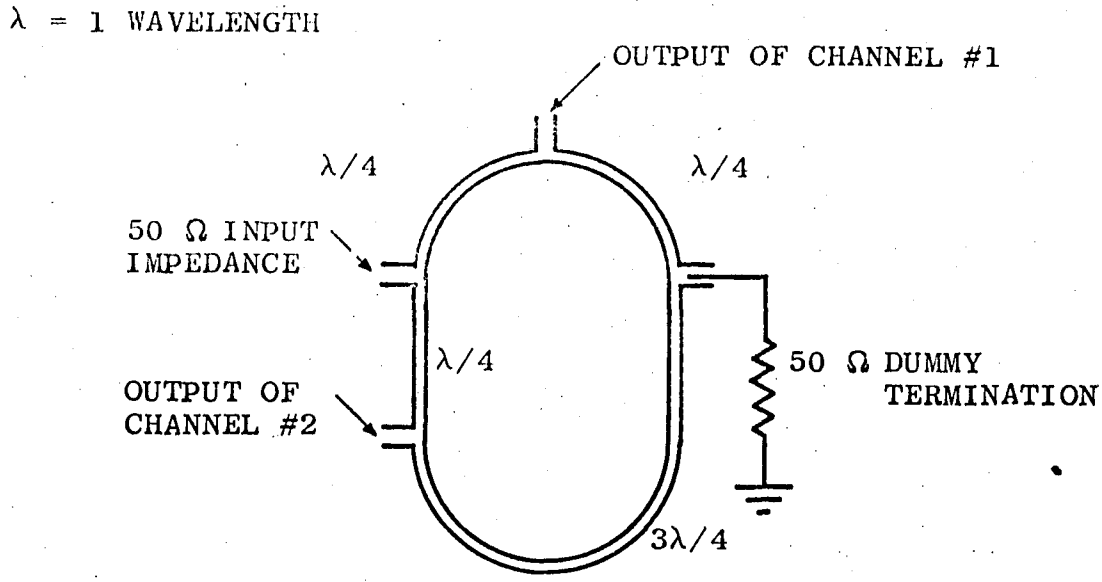
Since the maximum rate of change in CNR (figure 4.6-3) is approximately

$$\left(\frac{\Delta CNR}{\Delta T}\right)_{MAX} = \frac{17}{220} = 0.0773 \text{ db/sec},$$

the application of the channel identification techniques proposed in this work is realistic.

The channel identification is required at RF with the result that measuring devices must be inserted without significantly altering the signal which is of concern to the PLL demodulator. One approach to this problem is to utilize a three-port hybrid to split the RF signal into two channels which have interchannel isolation with respect to interconnected impedances. Such a hybrid device can be constructed from transmission line sections as shown in figure 4.6-5. On the order of 20-db isolation can be achieved with bandwidths of approximately 35 per cent.

In the channel identifier output of the hybrid, a straightforward approach is to place a series of η narrow-band filters in parallel to obtain the inputs for the point power spectral density estimates. This approach,



Z_o = 70.7 Ω = TRANSMISSION LINE CHARACTERISTIC IMPEDANCE

Figure 4.6-5. Transmission Line Hybrid

however, is quite complicated from a hardware standpoint and consequently another more practical method is presented.

The second approach is shown in block diagram form in figure 4.6-6. The approach used here is to feed the channel identifier output of the hybrid into a wide-band mixer with center frequency corresponding to the center frequency of the received FM signal band. The IF frequency is at dc so that the entire spectrum has been translated to the low-pass case. A highly stable, crystal controlled, temperature stabilized oscillator is utilized in the mixing process. A series of narrow crystal filters at center frequencies uniformly distributed over the frequency band from dc to several megacycles are utilized as inputs for the PSD estimators. Standard operational amplifiers and integrators operating over the bandwidth and integration times required are state-of-the-art and should present no problem in the realization.

Care must be taken to design the mixer to keep spurious output terms low since these spurious outputs could cause error in the spectrum translation. With sufficient care a 60-db dynamic range can be achieved with spurious response low enough to meet the required objectives.

The parameter computation and implementation portions of the adaptive PLL structure could be realized by the same types of special-purpose analog computing elements already described in sections 4.2 and 4.4. Since the low-pass

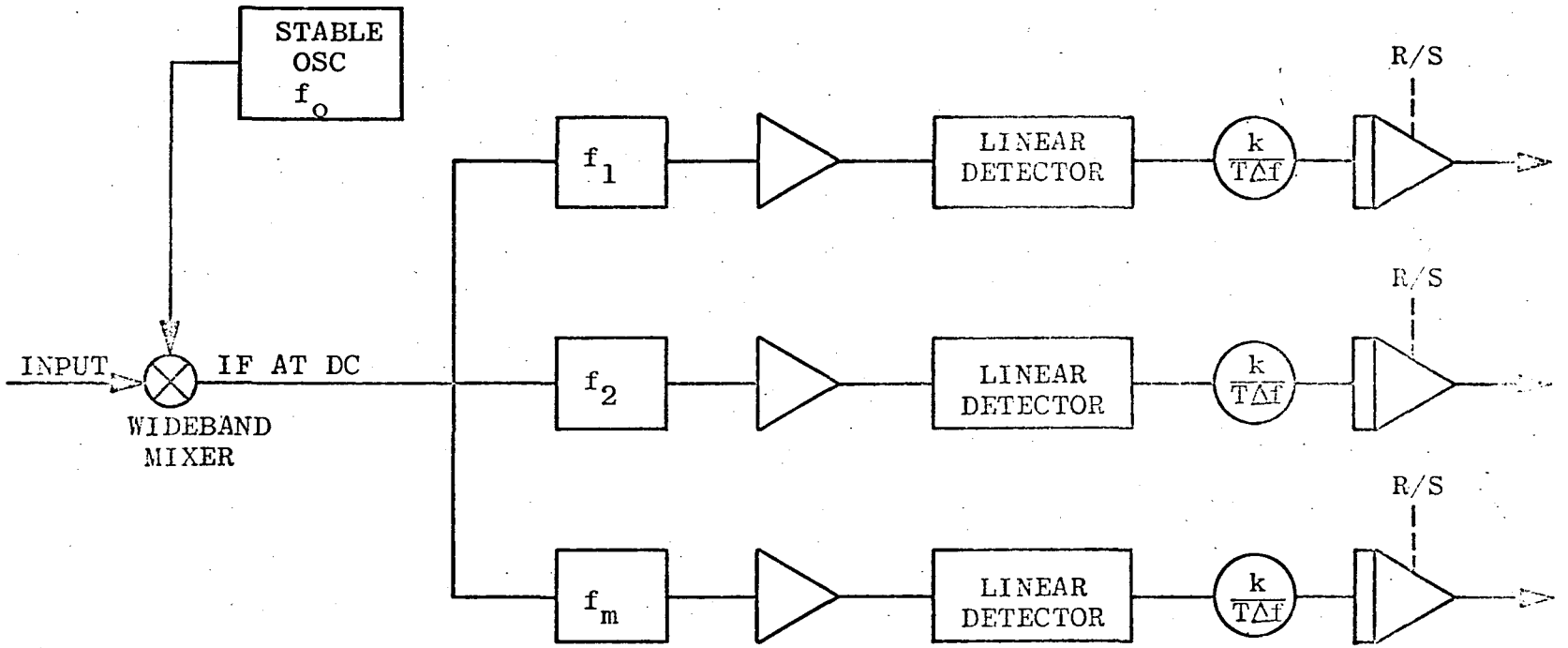


Figure 4.6-6. Practical Channel Identifier

filter operating within the phase-locked loop is required to have maximum bandwidths on the order of 100 kcps for large index FM systems. state-of-the-art multipliers could be used to effect the parameter variations.

Develet [47] has derived an expression for the output signal-to-noise ratio of the PLL receiver having an optimum fixed filter function F(s). Under the conditions of modulation with a band-limited white Gaussian signal, the output signal-to-noise ratio of the PLL detector when operating above threshold is given by

$$\left(\frac{S}{N}\right)_o = \frac{2D^2 (C/N)_i}{\exp(\sigma^2)} , \qquad (4\text{-}38)$$

where σ^2 is the minimum following error of the PLL.

Utilizing Develet's work and the CNR data of figure 4.6-3, the time-varying rate at which information could be transferred in the Echo II system, employing an optimum adaptive receiver, is shown in figure 4.6-7.

In order to realize the advantage in variable data rate offered by the channel it would be necessary to close the communication loop. Thus a return path from receiver to transmitter would be required, through which the received CNR information could be transferred. This channel would require an extremely low information rate, hence bandwidth requirements would be slight. Using this return information, the transmitter would time-vary its data rate

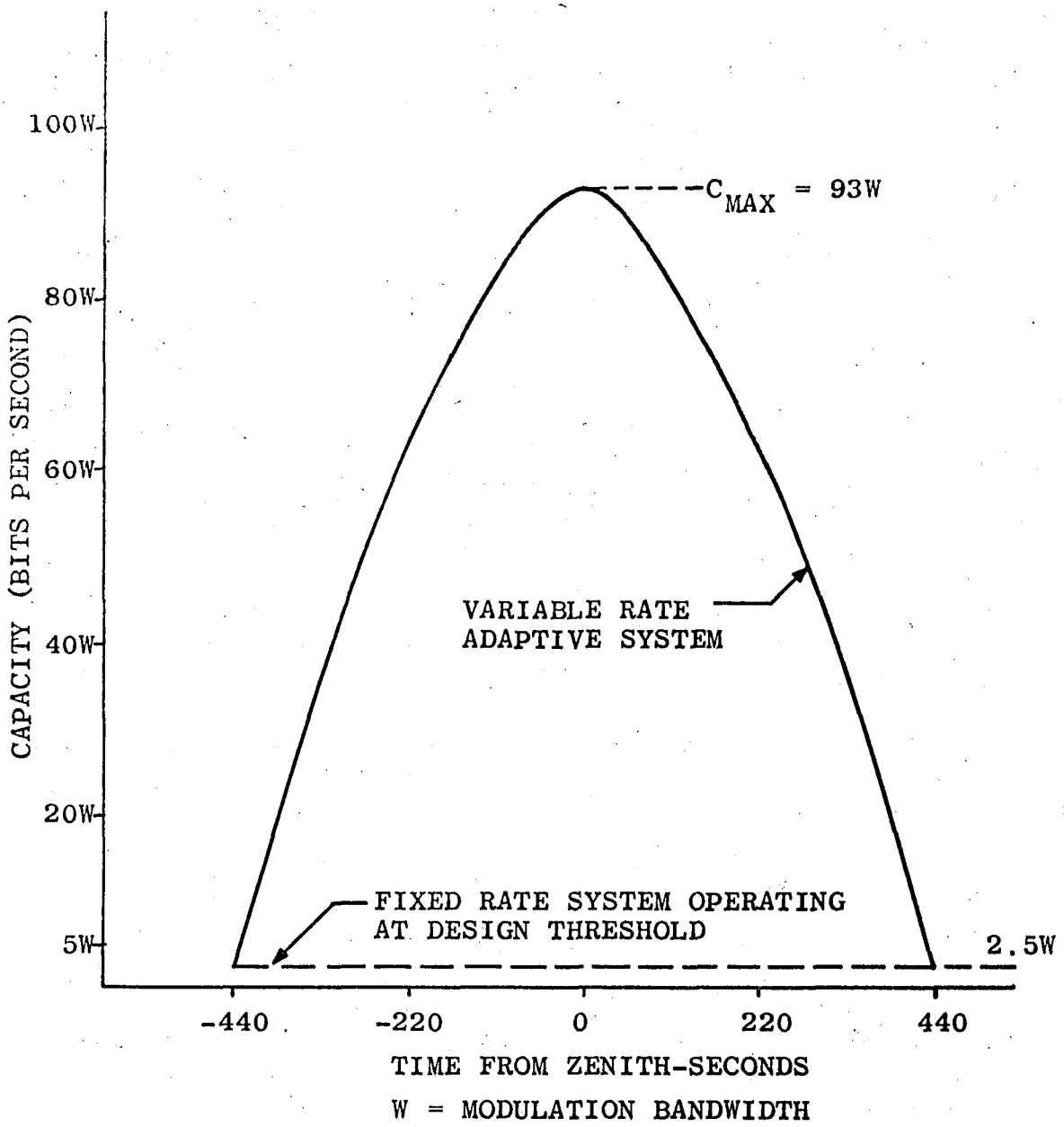


Figure 4.6-7. Available Information Rate of Optimum PLL Receiver for the Echo II System

in correspondence to the receiver CNR and hence the link information capacity.

The area under the two curves in figure 4.6-7 represents the total information transfer possible during a pass of the satellite. A graphical integration yields the result that the time-varying adaptive system could transfer 27 times the information of a fixed rate system during each overhead pass of Echo II.

In summary, this example has shown that the theory presented for the design of adaptive filters has practical applications in modern communication systems. Additionally, the point is made that by using adaptive time-varying transmitters and receivers, more efficient use can be taken of the available communication channels.

Chapter 5

DISCUSSION AND SUMMARY

5.1 DISCUSSION OF WORK PRESENTED

A theory has been proposed for the design of near-optimal time-varying filters which offer a significant improvement over fixed parameter Wiener filters when the input is a nonstationary stochastic process, having slowly varying nonstationary parameters, composed of additive, uncorrelated signal and noise.

Examples of the design method have been given and two experimental simulations have been presented to verify the proposed theory. The experimental results obtained adequately substantiate the theory and indicate the hardware complexities which arise when adaptive filtering is applied to practical problems. Both the theoretical work and the practical realizations are believed to be of use in the area of modern communications.

It has been demonstrated by the two experimental analog simulations described in sections 4.3 and 4.5, that in practice adaptive filter structures yield performances considerably better than the best fixed filter. Gains in output signal-to-noise ratios of 6 to 9 db are representative of the adaptive structures considered when the filters

are subject to some 20 db of variability in input signal-to-noise ratio.

A theoretical example of a modern satellite communication system was given in section 4.6 to illustrate both the feasibility of the design methods presented and the order of improvement that may be expected when adaptive concepts are applied to time-varying channels. It was demonstrated in section 4.6 that some 27 times as much information could theoretically be transferred in an adaptive satellite communication system as can be transferred using a fixed, minimum threshold design for the receiver.

A consideration of the departure from optimum, which results for a Wiener filter that has been designed for a stationary input, when the input is in fact nonstationary, indicates that worst case fixed filter design is not, in general, an efficient practice. Certainly a more sophisticated approach in the single-parameter nonstationary problem would be to minimize the analytical performance index

$$E\{F_G\} = \int F_G(\xi)P(\xi)d\xi,$$

where ξ is the nonstationary parameter, with respect to the function F_G. This would be accomplished by selecting the design value $\xi = \lambda$ for $H_o(s)$, such that the probability weighted area under the $F_G(\xi)$ curve is a minimum. For a uniform density, $P(\xi)$, this corresponds to minimizing the integral

$$\int_{\xi_{min}}^{\xi_{max}} F_G'(\xi)d\xi.$$

Thus, it is suggested as a corollary to the adaptive filter theory presented, that fixed parameter Wiener filters should be designed by using the a priori probability concerning the likelihood of the input stochastic processes departing from stationarity.

5.2 SUGGESTIONS FOR FUTURE STUDY

There remain several areas of interest to which additional attention should be given in regard to the adaptive filter theory presented.

The selection of frequencies at which power spectral density estimates are to be made has been largely arbitrary in this work. It seems likely that adaptive performance would be improved if this selection were made in some optimum fashion, perhaps with a consideration for minimizing the PSD estimator time required for a specified variance in the estimate.

The problem of channel identification time has not been entirely resolved. It may be possible to describe in a more satisfying analytical manner the class of input PSD's to which the proposed adaptive theory applies.

The matter of parameter implementation is another area in which further study is needed. The dynamic ranges

of state-of-the-art multipliers restrict the range of continuous parameter control to approximately 10-20 to 1. Other devices, such as the field-effect transistor, may prove of use for this purpose.

5.3 SUMMARY

In conclusion, the contributions of this work have been to further the knowledge in the area of optimum filter theory and to present a design procedure for realizing near-optimum time-varying filters for solving the realistic and practical information transfer problems presented by non-stationary communication channels.

Appendix

DERIVATION OF THE WIENER OPTIMAL FILTER

For an input stochastic process composed of stationary and ergodic signal and noise which are uncorrelated, the instantaneous filter error is given by

$$e(t) = s(t) - c(t) = s(t) - \int_0^t h(\tau)[s(t-\tau) + n(t-\tau)]d\tau, \quad \text{(A-1)}$$

where $h(t,\tau)$, the time-varying filter impulse response in (2-15), has been replaced by $h(\tau)$, the impulse response function of a time-invariant or constant parameter filter.

Since a physically realizable system can only respond to past inputs, it is necessary that $h(\tau) = 0$ for $\tau > t$. In addition one is free to define the impulse response $h(\tau)$, to be zero for $\tau < 0$. Thus, the lower and upper limits of integration in (A-1) may be replaced by $-\infty$ and ∞, respectively. Making this substitution for the limits of integration, one has

$$e(t) = s(t) - \int_{-\infty}^{\infty} h(\tau)[s(t-\tau) + n(t-\tau)]d\tau. \quad \text{(A-2)}$$

[1]For the correlation function of a stationary stochastic process

---

1. This analysis of correlation functions for nonstationary random variables is due to Laning and Battin [24], pp. 108-109.

$$\varphi_{xx}(\tau_1+t,\ \tau_2+t) = \varphi_{xx}(\tau_1,\ \tau_2) \tag{A-3}$$

for every t. Choosing $t = -\tau_1$, one has

$$\varphi_{xx}(\tau_1,\ \tau_2) = \varphi_{xx}(0,\ \tau_2-\tau_1). \tag{A-4}$$

Thus, the correlation function is dependent only upon the interval $\tau_2 - \tau_1$, and not only τ_1 and τ_2 individually. The notation, in the stationary case, may just as well be

$$\varphi_{xx}(\tau_1,\ \tau_2) = \varphi_{xx}(\tau_2-\tau_1) = \varphi_{xx}(\tau_1-\tau_2), \tag{A-5}$$

the latter relationship existing because of the symmetry of the autocorrelation function.

Now forming the ensemble autocorrelation of the error random variable,

$$\begin{aligned}\Phi_{ee}(\lambda) &= E\{e(t)e(t+\lambda)\} \\ &= E\{[s(t)-\int_{-\infty}^{\infty} h(\tau_1)[s(t-\tau_1)+n(t-\tau_1)]d\tau_1]\cdot \\ &\quad [s(t+\lambda)-\int_{-\infty}^{\infty} h(\tau_2)[s(t+\lambda-\tau_2)+n(t+\lambda-\tau_2)]d\tau_2]\},\end{aligned} \tag{A-6}$$

and interchanging the operations of ensemble averaging and integration, yields

$$\Phi_{ee}(\lambda) = E\{s(t)s(t+\lambda)\} - \int_{-\infty}^{\infty} h(\tau_2)E\{s(t)[s(t+\lambda-\tau_2) + n(t+\lambda-\tau_2)]\}d\tau_2$$

$$- \int_{-\infty}^{\infty} h(\tau_1)E\{s(t+\lambda)[s(t-\tau_1)+n(t-\tau_1)]\}d\tau_1 \qquad \text{(A-7)}$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)E\{[s(t-\tau_1)+n(t-\tau_1)][s(t+\lambda-\tau_2) + n(t+\lambda-\tau_2)]\}d\tau_1 d\tau_2.$$

From the definitions of autocorrelation and cross-correlation, equations (2-3) and (2-4) in section 2.4, certain simplifications in the notation of (A-7) can be made, thus

$$\Phi_{ee}(\lambda) = \Phi_{ss}(\lambda) - \int_{-\infty}^{\infty} h(\tau_2)[\Phi_{ss}(\lambda-\tau_2) + \Phi_{sn}(\lambda-\tau_2)]d\tau_2$$

$$- \int_{-\infty}^{\infty} h(\tau_1)[\Phi_{ss}(\lambda-\tau_1) + \Phi_{sn}(\lambda-\tau_1)]d\tau_1 \qquad \text{(A-8)}$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)[\Phi_{ss}(\lambda+\tau_1-\tau_2) + \Phi_{sn}(\lambda+\tau_1-\tau_2) + \Phi_{ns}(\lambda+\tau_1-\tau_2) + \Phi_{nn}(\lambda+\tau_1-\tau_2)]d\tau_1 d\tau_2$$

By the assumption, in the statement of the problem, that signal and noise are uncorrelated, all crosscorrelation terms in (A-8) are zero, thus (A-8) becomes

$$\Phi_{ee}(\lambda) = \Phi_{ss}(\lambda) - \int_{-\infty}^{\infty} h(\tau)\Phi_{ss}(\lambda-\tau)d\tau - \int_{-\infty}^{\infty} h(\tau_1)\Phi_{ss}(-\lambda-\tau_1)d\tau_1 \qquad \text{(A-9)}$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)[\Phi_{ss}(\lambda+\tau_1-\tau_2) + \Phi_{nn}(\lambda+\tau_1-\tau_2)]d\tau_1 d\tau_2.$$

Taking the two-sided Laplace transform of (A-9) gives

$$\Phi_{ee}(s) \triangleq \int_{-\infty}^{\infty} \Phi_{ee}(\lambda) e^{-s\lambda} d\lambda = \int_{-\infty}^{\infty} \Phi_{ss}(\lambda) \bar{e}^{s\lambda} d\lambda$$

$$- \int_{-\infty}^{\infty} h(\tau) \int_{-\infty}^{\infty} \Phi_{ss}(\lambda-\tau) \bar{e}^{s\lambda} d\lambda d\tau$$

$$- \int_{-\infty}^{\infty} h(\tau_1) \int_{-\infty}^{\infty} \Phi_{ss}(-\lambda-\tau_1) \bar{e}^{s\lambda} d\lambda d\tau_1 \qquad \text{(A-10)}$$

$$+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\tau_1) h(\tau_2) \int_{-\infty}^{\infty} [\Phi_{ss}(\lambda+\tau_1-\tau_2)$$

$$+ \Phi_{nn}(\lambda+\tau_1-\tau_2)] \bar{e}^{s\lambda} d\lambda d\tau_1 d\tau_2,$$

where $\Phi_{xx}(s)$ is the symbol for the two-sided Laplace transform of the function $\Phi_{xx}(\lambda)$.

Multiplying certain terms in (A-10) by unity, where unity may be written as

$$\bar{e}^{s\tau_1} \cdot e^{s\tau_1} \equiv 1, \qquad \text{(A-11)}$$

and rewriting the integrals in a slightly different form, yields the following:

$$\Phi_{ee}(s) = \Phi_{ss}(s) - \int_{-\infty}^{\infty} h(\tau) \bar{e}^{s\tau} d\tau \int_{-\infty}^{\infty} \Phi_{ss}(\lambda-\tau) \bar{e}^{s(\lambda-\tau)} d\lambda \qquad \text{(A-12)}$$

$$- \int_{-\infty}^{\infty} h(\tau) e^{s\tau} d\tau \int_{-\infty}^{\infty} \Phi_{ss}(-\lambda-\tau) \bar{e}^{s(\lambda+\tau)} d\lambda + \int_{-\infty}^{\infty} h(\tau_1) e^{s\tau_1} d\tau_1 \cdot$$

$$\int_{-\infty}^{\infty} h(\tau_2) \bar{e}^{s\tau_2} d\tau_2 \int_{-\infty}^{\infty} [\Phi_{ss}(\lambda+\tau_1-\tau_2)$$

$$+ \Phi_{nn}(\lambda+\tau_1-\tau_2)] \bar{e}^{s(\lambda+\tau_1-\tau_2)} d\lambda.$$

One may now identify the various integral forms shown with the corresponding two-sided Laplace transforms and re-write (A-12) as

$$\Phi_{ee}(s) = \Phi_{ss}(s) - \Phi_{ss}(s)[H(s)+H(-s)] + H(s)H(-s)[\Phi_{ss}(s)+\Phi_{nn}(s)]. \tag{A-13}$$

Since the stochastic signal and noise processes s(t) and n(t), with respect to which the Laplace transform of the error autocorrelation function (A-13) is expressed, are stationary and ergodic random variables, one has that

$$\Phi_{ee}(\lambda) = \varphi_{ee}(\lambda), \tag{A-14}$$

with probability one. Consequently, the following relation (known as Parseval's relation) exists between $\Phi_{ee}(s)$ and the average squared error,

$$\overline{e^2(t)} = \widetilde{e^2(t)} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \Phi_{ee}(s)\, ds. \tag{A-15}$$

In order to minimize the average squared error, one is led through the foregoing analysis to consider the problem of choosing H(s) such that

$$A\{e^2(t)\} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \{\Phi_{ss}(s) - [H(s)+H(-s)]\Phi_{ss}(s) + H(s)H(-s)[\Phi_{ss}(s)+\Phi_{nn}(s)]\}\, ds \tag{A-16}$$

will be minimized.

The minimization problem as thus stated, is an example of the class of mini-max problems treated by the calculus of variations technique. One proceeds to a solution by setting the general linear system response function H(s) equal to the optimal response function $H_o(s)$ plus a perturbation term, thus let

$$H(s) = H_o(s) + \epsilon\eta(s). \tag{A-17}$$

where $H_o(s)$ is that response function which causes the integral form (A-16) to be a minimum, or in general, a minimum or maximum. Physical reasoning is usually sufficient to show whether the solution function $H_o(s)$ that is obtained produces a minimum or a maximum.

A necessary, but not sufficient, condition that H(s) produce a minimum value of the integral (A-16) is that

$$\left.\frac{\partial A\{e^2(t)\}}{\partial \epsilon}\right]_{\epsilon=0} = 0, \tag{A-18}$$

independent of $\eta(s)$.

Substituting (A-17) for H(s) in (A-16) gives

$$A\{e^2(t)\} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \{\Phi_{ss}(s) - [H_o(s)+\epsilon\eta(s)+H_o(-s)+\epsilon\eta(-s)]\Phi_{ss}(s) + [H_o(s)+\epsilon\eta(s)][H_o(-s)+\epsilon\eta(-s)][\Phi_{ss}(s)+\Phi_{nn}(s)]\}ds. \tag{A-19}$$

Thus,

$$\left.\frac{\partial A\{e^2(t)\}}{\partial \epsilon}\right]_{\epsilon=0} = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \{-[\eta(s)+\eta(-s)]\Phi_{ss}(s)+\eta(s)H_o(-s)[\Phi_{ss}(s) + \Phi_{nn}(s)]+\eta(-s)H_o(s)[\Phi_{ss}(s)+\Phi_{nn}(s)]\}ds$$

$$= \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \eta(s)\{H_o(-s)[\Phi_{ss}(s)+\Phi_{nn}(s)]-\Phi_{ss}(s)\}ds + \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \eta(-s)\{H_o(s)[\Phi_{ss}(s)+\Phi_{nn}(s)] - \Phi_{ss}(s)\}ds, \qquad \text{(A-20)}$$

must be zero, independent of $\eta(s)$. The two integral forms in (A-20) can be made contour integrals in the s-plane by closing the line integral from $-j\infty$ to $j\infty$ with an infinite radius semi-circle which closes either through a left half plane (LHP) path or through a right half plane (RHP) path. Since e(t) is zero for $t < 0$, i.e., the filter is not responding until the input begins at $t = 0$, one concludes that an LHP contour is to be used in order to evaluate the physically realizable average squared error, and correspondingly its partial derivative with respect to ϵ.

It will be shown, in the discussion that follows, that such an LHP closure of the integrals in (A-20) results in an equivalence between the setting of either one to zero, independent of $\eta(s)$. That is to say, if the kernel of the second integral form in (A-20),

$$H_o(s)[\Phi_{ss}(s)+\Phi_{nn}(s)] - \Phi_{ss}(s), \qquad \text{(A-21)}$$

is set to zero the first integral form will necessarily be zero also.

Thus, as a formal solution to the Wiener problem, one obtains from equating (A-21) to zero,

$$H_o(s) = \frac{\Phi_{ss}(s)}{\Phi_{ss}(s) + \Phi_{nn}(s)} . \qquad \text{(A-22)}$$

This solution for $H_o(s)$ is, however, not physically realizable, i.e., the substitution of the two-sided Laplace transforms $\Phi_{ss}(s)$ and $\Phi_{nn}(s)$ for real time-autocorrelation functions $\varphi_{ss}(\lambda)$ and $\varphi_{nn}(\lambda)$ in (A-22) will require $H_o(s)$ to have right half plane poles.

Since H(s) must be physically realizable, i.e., regular in the right half of the s-plane (RRHP), and $H_o(s)$ must be RRHP, one concludes that $\eta(s)$ must be RRHP. Now if one considers the second integral form in (A-20),

$$\frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \eta(-s)\{H_o(s)[\Phi_{ss}(s)+\Phi_{nn}(s)]-\Phi_{ss}(s)\}ds, \qquad \text{(A-23)}$$

and noting that $\eta(-s)$ must be RLHP, by Morera's theorem[1] the integral expression in (A-23) will be zero if

---

1. See any text on complex variable theory, e.g., Churchill [25], pp. 94-95.

$$H_o(s)[\Phi_{ss}(s)+\Phi_{nn}(s)]-\Phi_{ss}(s) \qquad \text{(A-24)}$$

is regular in the left half plane (RLHP).

Since $\Phi_{ss}(s)$ and $\Phi_{ss}(s) + \Phi_{nn}(s)$ are the two-sided Laplace transforms of real time-autocorrelation functions, these transforms are symmetrical about both the real and imaginary axes in the s-plane. Consequently for (A-26) to be RLHP, the LHP poles of $H_o(s)\ [\Phi_{ss}(s) + \Phi_{nn}(s)]$ must be identical to the LHP poles of $\Phi_{ss}(s)$.

Let $\Phi_{ss}(s) + \Phi_{nn}(s)$ be written in factored form as

$$\Phi_{ss}(s) + \Phi_{nn}(s) = \Delta(s)\Delta(-s), \qquad \text{(A-25)}$$

where $\Delta(s)$ contains the poles and zeros of $[\Phi_{ss}(s) + \Phi_{nn}(s)]$ which lie in the LHP and $\Delta(-s)$ contains the (necessarily symmetrical about the imaginary axis) poles and zeros of $[\Phi_{ss}(s) + \Phi_{nn}(s)]$ which lie in the RHP.

The integral form (A-23) may thus be written as

$$\frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \eta(-s)\Delta(-s)\left\{\Delta(s)H_o(s) - \frac{\Phi_{ss}(s)}{\Delta(-s)}\right\}ds. \qquad \text{(A-26)}$$

By a partial fractions expansion, one may write

$$\frac{\Phi_{ss}(s)}{\Delta(-s)} = \Gamma_1(s) + \Gamma_s(-s), \qquad \text{(A-27)}$$

where $\Gamma_1(s)$ is RHHP and $\Gamma_2(-s)$ is RLHP. Then (A-26) becomes

$$\frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \eta(-s)\Delta(-s)\{\Delta(s)H_o(s) - \Gamma_1(s)\}ds$$
$$- \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \eta(-s)\Delta(-s)\Gamma_2(-s)ds. \qquad \text{(A-28)}$$

The integral forms shown in (A-28) are of necessity contour integrals and an inspection of the original integral forms (A-20), from which (A-28) was derived, shows that it is the LHP contour by which (A-28) is to be evaluated. Thus, using a LHP contour for the integrals in (A-28) and applying Morera's theorem, one has that the second integral form in (A-28) is zero since $\eta(-s)\Delta(-s)\Gamma_2(-s)$ is regular in the LHP.

Setting the remaining, in general non-zero, integral in (A-28) equal to zero, independent of $\eta(s)$, gives

$$\Delta(s)H_o(s) - \Gamma_1(s) = 0, \qquad \text{(A-29)}$$

and hence

$$H_o(s) = \frac{\Gamma_1(s)}{\Delta(s)}. \qquad \text{(A-30)}$$

Equation (A-30) is the physically realizable solution for the optimum Wiener filter. Frequently this is written in a slightly different form obtained by setting

$$[\Phi_{ss}(s) + \Phi_{nn}(s)]_- = \Delta(-s), \qquad \text{(A-31)}$$

$$[\Phi_{ss}(s) + \Phi_{nn}(s)]_+ = \Delta(s), \qquad \text{(A-32)}$$

and

$$\left\{\frac{\Phi_{ss}(s)}{\Delta(-s)}\right\}_{P.R.} = \Gamma_1(s). \qquad \text{(A-33)}$$

In (A-33) the notation P.R., which stands for physically realizable, indicates the operation of (A-27), whereby $\Phi_{ss}(s)/\Delta(-s)$ is expanded by partial fractions, and then only that portion of the expansion, $\Gamma_1(s)$, which is RRHP is retained.

Using (A-31), (A-32), and (A-33), the optimum system function $H_o(s)$, which minimizes the average squared error for the Wiener problem, is expressed as

$$H_o(s) = \frac{\left\{ \dfrac{\Phi_{ss}(s)}{[\Phi_{ss}(s) + \Phi_{nn}(s)]_-} \right\}_{P.R.}}{[\Phi_{ss}(s) + \Phi_{nn}(s)]_+} . \qquad \text{(A-34)}$$

REFERENCES

[1] Wiener, N. Extrapolation, Interpolation, and Smoothing of Stationary Time Series, published jointly by the Technology Press of the Massachusetts Institute of Technology and John Wiley and Sons, Inc., New York, 4th printing, 1960.

[2] Kolmogorov, A.N., "Interpolation and Extrapolation von stationären zufälligen Folgen, Bulletin de l'académie des sciences de U.R.S.S., Ser. Math. 5, 1941, pp. 3-14.

[3] Davenport, W.B., Jr. and W.L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill Book Company, Inc., New York, 1958.

[4] Korn, G.A. and T.M. Korn, Mathematical Handbook for Scientists and Engineers, McGraw-Hill Book Company, Inc., 1961.

[5] Fano, R.M., "Short-Time Autocorrelation Functions and Power Spectra," Journal of the Acoustical Society of America, Vol. 22, No.5, September 1950, pp. 546-550.

[6] Knapp, C.H., "Synthesis of Fixed and Variable Networks for Filtering Non-Stationary Random Signals," IEEE Transactions on Information Theory, Vol. IT-9, No.4, October 1963, pp. 276-280.

[7] Booton, R.C., Jr., "An Optimization Theory for Time-Varying Linear Systems with Non-Stationary Statistical Inputs," IRE Proceedings, Vol. 40, August 1952, pp. 977-981.

[8] Shinbrot, M., "Optimization of Time-Varying Linear Systems with Non-Stationary Inputs," Transactions ASME, Vol. 80, February 1958.

[9] Bendat, J.S., Principles and Applications of Random Noise Theory, John Wiley and Sons, Inc., New York, 1958.

[10] Keiser, B.E., "The Linear Input-Controlled Variable-Pass Network," IRE Transactions on Information Theory, Vol IT-1, March 1955, pp. 34-39.

[11] Klass, P.J., "Filter Detects Signals Buried in Noise," Aviation Week September 26, 1960, pp. 83-95.

[12] Glaser, E.M., "Signal Detection by Adaptive Filters," IRE Transactions on Information Theory, Vol. IT-7, April 1961, pp. 87-98.

[13] Bucy, R.S. and J.W. Follin, Jr., "Adaptive Finite Time Filtering," IRE Transactions on Automatic Control, July 1962, pp. 10-19.

[14] Stromer, P.R., "Adaptive or Self-Optimizing Control Systems - A Bibliography," IRE Transactions on Automatic Control Vol. AC-4, May 1959.

[15] Chang, S.S.L., Synthesis of Optimum Control Systems, McGraw-Hill Book Company, Inc., New York, 1961.

[16] Newton, G.C., L.A. Gould, and J.F. Kaiser, Analytical Design of Linear Feedback Controls, John Wiley & Sons, Inc., New York, 1957.

[17] Mishkin, E., and L.Braun, Jr., Adaptive Control Systems, McGraw-Hill Book Company, Inc., New York, 1961.

[18] Nichols, R.H. and L.L. Rauch, Radio Telemetry, Second Edition, John Wiley and Sons, Inc., 1956.

[19] Harman, W.W., Principles of the Statistical Theory of Communication, McGraw-Hill Book Company, Inc.,1963.

[20] Radio Science, Journal of Research of the National Bureau of Standards, Formerly Radio Propagation. See particularly Vol. 67D, No.5, pp. 459-480; Vol. 67D, No.6, pp. 753-762; and Vol. 68D, No. 6, pp. 723-736.

[21] Lee, Y.W., Statistical Theory of Communication, John Wiley and Sons, Inc., 1960.

[22] Tuttle, D.F., Jr., Network Synthesis, Vol. 1, John Wiley and Sons, Inc., 1958.

[23] Streets, R.B., Jr., "Integration of Analytical and Conventional Design Techniques for Optimum Control Systems," PhD Dissertation, University of Arizona, 1963.

[24] Laning, J.H., Jr. and R.H. Battin, Random Processes in Automatic Control, McGraw-Hill Book Company, Inc., 1948.

[25] Churchill, R.V., Introduction to Complex Variables and Applications, McGraw-Hill Book Company, Inc., 1948.

[26] Parzen, E. "On Consistent Estimates of the Spectral Density of a Stationary Time Series." Proceedings of the National Academy of Science, Vol. 42, 1956, pp. 154-157.

[27] Parzen, E., "On Consistent Estimates of the Spectrum of a Stationary Time Series," Annuals of Mathematical Statistics, Vol. 28, 1957, pp. 329-348.

[28] Parzen, E., "On Choosing an Estimate of the Spectral Density Function of a Stationary Time Series," Annuals of Mathematical Statistics, Vol. 28, 1957, pp. 921-932.

[29] Parzen, E., "On Asymptotically Efficient Consistent Estimates of the Spectral Density Function of a Stationary Time Series," Journal of the Royal Statistical Society, Section B, Vol. 20, 1958, pp. 303-322.

[30] Parzen, E., "An Approach to Time Series Analysis," Annuals of Mathematical Statistics, Vol. 32, No. 4, 1961, pp. 951-989.

[31] Parzen, E., "Mathematical Considerations in the Estimation of Spectra," Technometrics, Vol. 3, No.2, 1961, pp. 167-190.

[32] Blackman, R.B. and J.W. Tukey, The Measurement of Power Spectra From the Point of View of Communications Engineering, Dover Publications, Inc., New York, 1958.

[33] Grenander, U. and M. Rosenblatt, Statistical Analysis of Stationary Time Series, John Wiley and Sons, Inc., 1957.

[34] Priestley, M.B., "Estimation of the Spectral Density Function in the Presence of Harmonic Components," Journal of the Royal Statistical Society, Vol. 26, Part 1, 1964, pp. 123-132.

[35] Cooper, G.R. and H. Berndt, "Estimation of Time-Varying Correlation Functions," Purdue University Publication TR-EE-63-1, March 1963.

[36] Steiglitz, K., "Power Spectrum Identification for Adaptive Systems," IEEE Transactions on Applications and Industry, Vol. 83, May 1964, pp. 195-197.

[37] Kharkevich, A.A., Spectra and Analysis, Translated from Russian, Consultants Bureau, New York, 1960.

[38] Rideout, V.C. "Part 9: Random-Process Studies," Computer Handbook, ed. by H.D. Huskey and G.A. Korn, McGraw-Hill Book Company, Inc., 1962, pp. 5-94 to 5-110.

[39] Morrow, Charles T., "Averaging Time and Data Reduction Time for Random Vibration Spectra," Journal of the Acoustical Society of America, Vol. 30, Nos. 5 and 6, June 1958.

[40] Burington, R.S. and D. C. May, Jr., Handbook of Probability and Statistics With Tables, Handbook Publishers, Inc., Sandusky, Ohio, 1958.

[41] Cohn, C.E., "Linear Detectors for Noise Measurements," Review of Scientific Instruments, Vol. 35, No. 5, May 1964, pp. 638-639.

[42] Rosenblatt, M., "Some Comments on Narrow Band-Pass Filters," Quarterly of Applied Mathematics, Vol. XVIII, No. 4, 1961, pp. 387-393.

[43] Martin, T.L., Electronic Circuits, Prentice-Hall, Inc., 1955.

[44] Research on Devices for the Detection of Weak Signals, Tech. Documentary Report No. AL TDR 64-224, published by the Air Force Avionics Laboratory, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, September 1964.

[45] Pritchard, W.L., "Microwave Aspects of Space Communications," Microwave Journal, Vol. VI, No. 8, August 1963, pp. 52-64.

[46] Jaffe, R. and E. Rechtin, "Design and Performance of Phase-Lock Circuits Capable of Near-Optimum Performance Over a Wide Range of Input Signal and Noise Levels," IRE Trans. on Information Theory, Vol. IT-1, March 1955, pp. 66-76.

[47] Develet, J.A., Jr., "A Threshold Criterion for Phase-Lock Demodulation," Proc. of IEEE, Vol. 51, No. 2, February 1963, pp. 349-356.