2-1-1961

Limits on the Identification Time for Linear Systems

G. R. Cooper
Purdue University

J. C. Lindenbaud
Purdue University

TR-EE 61-3

PR 7 2358

TECHNICAL REPORT NO. 2
CONTRACT AF 33(616)-6890
PRF 2358

FINAL REPORT VOL. II
PROJECT 8225
TASK 82181

PURDUE UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

Limits on the Identification Time for Linear Systems

G. R. Cooper, Principal Investigator

J. C. Lindenlaub

February 1, 1961
Lafayette, Indiana

FOR
U. S. AIR FORCE
WRIGHT AIR DEVELOPMENT DIVISION
WRIGHT-PATTERSON AIR FORCE BASE
DAYTON, OHIO

TECHNICAL REPORT NO. 2

FINAL REPORT

VOLUME II

CONTRACT AF 33(616) - 6890

Project 8225

Task 82181

PRF 2358

LIMITS ON THE IDENTIFICATION TIME FOR LINEAR SYSTEMS

for

U. S. AIR FORCE

WRIGHT AIR DEVELOPMENT DIVISION

WRIGHT-PATTERSON AIR FORCE BASE

DAYTON, OHIO

by

G. R. Cooper, Principal Investigator

J. C. Lindenlaub

School of Electrical Engineering

Purdue University

Lafayette, Indiana

February 1, 1961

PREFACE

This report is the second volume of a two-volume final report prepared by the School of Electrical Engineering, Purdue University, under USAF Contract No. AF 33(616)-6890, Project No. 8225, Task No. 82181. The contract is administered under the direction of the Flight Control Laboratory, Wright Air Development Division, Wright-Patterson Air Force Base, Dayton, Ohio, by Lt. P. C. Gregory, the initiator of the study.

The first volume presented the development and analysis of a particular class of adaptive controls under the assumption of the availability of identification information. This second volume deals with the limits on the identification time for linear systems for a number of identification techniques.

For the past year Purdue University has had partial support by the Air Force in a rather broad study of adaptive control systems. The study was initiated some two and one half years ago and is still continuing. During this general research effort a number of critical areas in the theory of adaptive control have been uncovered. In several of these areas specific research objectives were set and results obtained, while in other areas work remains to be done.

One of these critical areas, covered in Volume I of this final report by Gibson and Meditch, is the unnecessary restriction of the adjustment procedure to incremental or continuous adjustment of physical parameters. This is the <u>parameter adjustment</u> solution to the <u>control signal modification</u> problem. The more general procedure, discussed in Volume I, lies in <u>control signal synthesis</u>, in which a new signal is generated with which to drive the plant so as to achieve optimum response.

A second critical area that has been under investigation is the identification problem. This study, which is reported in this volume of the final report, has established the minimum time required to identify the impulse response of any linear system in the presence of random disturbances and in the absence of *a priori* knowledge. This result has been obtained for several different practical identification techniques as well as for the ideal identifier.

Independent of Air Force support, Schiewe has reported on his analysis of multi-dimensional adaptive systems which measure not the impulse response of the plant but only certain important aspects of that response and Eveleigh has compared incremental vs. sinusoidal perturbation in multi-dimensional adaptive systems for speed of response and hunting loss. Tou and his co-workers, Joseph and Lewis, have been actively studying the digital adaptive problem and achieved very encouraging results.

Work is continuing now on new, fast identification schemes and theoretical analyses of identification with *a priori* information, as well as in the newer and relatively unexplored area of systems which exhibit learning. These require memory capacity and extended logic in the adaptive loop and the capacity for modifying the control law in accord with generalized performance criteria.

CONTENTS

	Page
LIST OF FIGURES	vi
LIST OF SYMBOLS	viii
ABSTRACT	xiii
THE IDENTIFICATION PROBLEM	1
1.1 The Identification Problem and its Relation to Adaptive Systems	1
1.2 Statement of the Problem and Basic Assumptions	4
1.3 Characteristics and Representation of Impulse Responses	8
IDENTIFICATION TIME REQUIRED BY AN IDEAL IDENTIFIER	14
2.1 The Ideal Identifier	14
2.2 The Joint Estimation of P Unknown Parameters	16
2.3 A P-Parameter Representation of an Impulse Response	20
2.4 Identification Time Required by an Ideal Identifier	21
IDENTIFICATION TIME REQUIRED BY CROSSCORRELATION IDENTIFICATION	27
3.1 Theoretical Basis for Crosscorrelation Identification	27
3.2 Nature of the Errors Associated with Crosscorrelation Identification	30
3.3 Analysis of Output Noise	33
3.4 A Practical Periodic Test Signal	37
3.5 Identification Time Required by Crosscorrelation	37
3.6 Equivalence of Crosscorrelation Identification and the Ideal Identifier	43
SYSTEM IDENTIFICATION USING SAMPLING TECHNIQUES	44
4.1 Introduction	44
4.2 Least Squares Estimates of the Impulse Response	47
4.3 Optimum Test Signals	49
4.4 Identification Time Using Sampling Techniques	53
4.5 Equivalence to the Ideal Identifier	55
IDENTIFICATION TIME REQUIRED BY MATCHED FILTER IDENTIFICATION	56
5.1 Description of the Matched Filter Identification Technique	56
5.2 Variance of the Impulse Response Estimate	60
5.3 Reduction of Variance by Periodic Excitation	63
5.4 Identification Time Required by the Matched Filter Technique	68
5.5 Comparison with the Ideal Identifier	71
5.6 Some Practical Considerations	71

CONTENTS (continued)

Page

EXAMPLES 73
6.1 Introduction 74
6.2 T_I for a Minimum Phase Second Order System - White Output Noise 77
6.3 T_I for a Nonminimum Phase Second Order System- White Output Noise 85
6.4 White Noise Originating Within the Feedback Loop 87
6.5 Consideration of a Narrowband Noise Process 94

SUMMARY OF RESULTS, CONCLUSIONS, AND RELATED PROBLEMS 97
7.1 Significance of the Ideal Identifier 97
7.2 Equivalence of Crosscorrelation, Sampling, and Matched Filter Identification 99
7.3 Operational Similarities and Relative Advantages of the Various Identification Methods 101
7.4 Significance of Identification Time Results for Adaptive Systems 104
7.5 Related Problems 105

BIBLIOGRAPHY 112

LIST OF FIGURES

Figure		Page
1-1	The Basic Identification Problem	6
1-2	Impulse Response of a Typical Time-Invariant System	10
1-3	Impulse Response of a Typical Time-Varying System	10
1-4	Sample Point Approximation of an Impulse Response	12
2-1	An Impulse Response and Sample Values	22
3-1	Crosscorrelation Identification	28
3-2	Convolution of $\phi_x(\tau)$ and $g(\lambda)$	32
3-3	Two Practical Test Signal Autocorrelation Functions	32
4-1	Identification by Sampling Techniques	45
5-1	Matched Filter Identification	58
5-2	Generation of Output Signal when x(t) is Periodic	66
5-3	Matched Filter Spectrum for a Periodic Test Signal	69
5-4	Signal and Noise Frequency Characteristics	69
5-5	Ideal Comb Filter Frequency Characteristics	69
6-1	White Noise Within the Feedback Loop and Equivalent Output Noise	76
6-2	Second Order Pole-Zero Configuration	78
6-3	Nonminimum Phase Second Order Pole-Zero Plot	78
6-4	Impulse Response - Second Order System $K_G = \omega_o = 1$, $\zeta = 1/2$	79
6-5	Normalized Identification Time vs. Variance	82

LIST OF FIGURES (continued)

Figure Page

6-6 Comparison of T_I for Gaussian and Periodic Test Signals 84

6-7 Impulse Response - Nonminimum Phase System $K_G = \omega_0 = 1$, $\zeta = 1/2$, $\alpha = 2$ 86

6-8 Equivalent Output Noise Power Density Spectrum 89

6-9 Signal and Noise Spectra at the Output of the Matched Filter 89

6-10 Convolution of $\Phi_n(\omega)$ and $\Phi_x(\omega)$ 93

6-11 Elimination of Narrowband Noise by Filtering 95

7-1 Identification by Means of a Model 110

LIST OF SYMBOLS

		Page
$x_1, x_2, \ldots, x_j$	input signals	1
$y_1, y_2, \ldots, y_k$	output signals	1
t	time	6
$x(t)$	input test signal	6
$w(t)$	output signal	6
$n(t)$	external noise	6
$y(t)$	observed signal	6
λ	time	6
$g(\lambda)$	unknown impulse response	6
$a_0, a_1, \ldots, a_n$	a set of coefficients	8
$b_0, b_1, \ldots, b_m$	a set of coefficients	8
p	the operator $\frac{d}{dt}$	8
$\delta(t)$	unit delta function	8
$g(\lambda, t)$	impulse response function	8
k	a constant	11
ω_0	undamped natural frequency	11
ζ	relative damping ratio	11
e	base of Naperian Logarithms	11
i	index of summation	13
α_i	set of coefficients	13
ψ_i	set of orthonormal functions	13
P	number of unknown parameters	14
$p[g(\lambda)\|y(t), x(t)]$	*a posteriori* probability density function	14

LIST OF SYMBOLS (continued)

		Page
$p[g(\lambda)\|y(t)]$	a posteriori probability density function	15
$p[g(\lambda)]$	prior probability density function	15
$p[y(t)\|g(\lambda)]$	likelihood function	15
$p[y(t)]$	probability density function	15
K	normalizing constant	15
L	likelihood function	16
$q_1, q_2, \ldots, q_p$	set of unknown parameters	16
$\{q_i\}$	set of unknown parameters	16
$w_a(t)$	actual output signal	16
q_{i_a}	actual value of i-th parameter	16
Φ_n	noise power spectral density	16
T_I	identification time	16
exp	exponential operator	16
$p[y(t)\| \{q_i\}]$	likelihood function	16
$\{q_{i_a}\}$	actual value of all parameters of the set q_i	16
Δq_i	deviation of q_i from actual value	17
j	index of summation	17
(_)	a matrix	17
b_{ij}	i, j element of (B)	18
$P[\{q\} \| y(t)]$	a posteriori probability density function	18
Σ_2	covariance matrix	19
σ_{ij}	covariance of q_i and q_j	19

LIST OF SYMBOLS (continued)

		Page
$\lvert B\rvert$	determinant of (B)	19
$\lvert B_{ij}\rvert$	cofactor of b_{ij}	19
g_p	sample values of $g(\lambda)$	21
$\Delta\lambda$	sampling interval	21
W_M	maximum system bandwidth in cps	21
Λ	a constant	21
$\overline{x^2}$	mean square value of x(t)	24
E_x	test signal energy	24
β	a constant	24
(I)	identity matrix	24
$\sigma_{\hat{g}}^2$	variance of a sample point estimate	25
τ	delay in seconds	29
z(t)	multiplier output signal	29
$\phi_n(\tau)$	external noise autocorrelation function	29
ω	radian frequency	29
$\Phi_n(\omega)$	external noise power spectral density	29
$\phi_{xy}(\tau)$	crosscorrelation function of x(t) and y(t)	29
E	expectation operation	29
$\phi_x(\tau)$	test signal autocorrelation function	30
Φ_x	area under $\phi_x(\tau)$ function	30
$\overline{z(t)}$	average value of z(t)	30
X	a constant	31
a	a constant	31

LIST OF SYMBOLS (continued)

		Page
t_1	minimum interval between changes of state	31
$\phi_z(\tau)$	autocorrelation function of z(t)	34
T_x	fundamental period of x(t)	36
$\phi_{x_1}(\tau)$	one period of a periodic correlation function	36
$\overline{n_o^2}$	mean square value of noise	37
$F(\omega)$	averaging filter transfer function	38
$\sim$	ensemble average	38
$\Phi_{no}(\omega)$	output noise power spectral density	38
W_F	equivalent noise bandwidth of averaging filter	38
W_G	equivalent noise bandwidth of system	41
K_G	low frequency power gain	41
W_x	equivalent noise bandwidth of test signal	41
γ	signal-to-noise ratio	41
$\overline{n^2}_{eff}$	effective mean square value of noise	41
t_a	sampling interval	44
$x(m)$	sample values of x(t)	44
Φ_n	noise correlation matrix	44
$(\hat{g})$	estimate of (g)	48
$\tilde{\phi}_x(r)$	empirical autocorrelation function	48
$\tilde{\phi}_{xy}(r)$	empirical crosscorrelation function	48
Φ_g	covariance matrix of (g)	49
$h(\lambda)$	matched filter impulse response	57

LIST OF SYMBOLS (continued)

		Page
$x_1(t)$	a test signal which is zero outside the interval 0, T_x	57
Δ	a delay	57
k	a constant	59
$\hat{g}_1(t)$	output of matched filter	59
$X_1(\omega)$	Fourier transform of $x_1(t)$	59
$H(\omega)$	matched filter transfer function	61
M	a constant	64
$h'(\lambda)$	an estimating filter	65
$f(\lambda)$	envelope function	65
$F(\omega)$	Fourier transfer of $f(\lambda)$	67
$f_1(t)$, $f_2(t)$, $f_3(t)$	time functions	67
$F_1(\omega)$, $F_2(\omega)$, $F_3(\omega)$	Fourier transforms of $f_1(t)$, $f_2(t)$, $f_3(t)$	67
ω_a	arbitrary value of radian frequency	68
E_W	energy of observed signal	75
$G_1(\omega)$	a transfer function	76
$W_G\vert_{max}$	maximum system bandwidth	81
c	a constant	81
α	right-half plane zero location	85
N	white noise power spectral density	88
Ω	variable of integration	91
β	a constant	92
s	Laplace transform variable	92
P(s), Q(s)	polynominals in s	92
$h_i(\lambda)$	impulse response of a filter	105

ABSTRACT

The problem of estimating the impulse response of a linear system arises in adaptive control problems and elsewhere. Often it is necessary to make the system identification in the presence of external noise disturbances. This work considers the problem of determining the time that is necessary to estimate the impulse response of a linear system with a specified variance. It is assumed that essentially no a priori knowledge about the unknown system is available, and that the output signal of the system is corrupted by an additive stationary noise signal.

An ideal identifier is defined as a device that yields, for a given identification time, minimum variance estimates of samples of the unknown impulse response function. Statistical parameter estimation techniques are used to determine the identification time required by an ideal identifier. The results show that, when the external disturbance is Gaussian and white, and the output signal energy is large compared to the power spectral density of the noise, the identification time is proportional to the power spectral density of the noise, and inversely proportional to the variance of the estimate and the mean square value of the input test signal. The identification time is independent of the impulse response being estimated.

The identification times required by several practical identification schemes are calculated and compared to the identification time of the ideal identifier. It is established that, when the input test signal is optimized and the noise is white, the methods of crosscorrelation, sampling input-output data, and matched filter identification are all equivalent to the ideal identifer.

Depending upon the size of the variance in the impulse response estimate that is required it is concluded that, in the absence of *a priori* knowledge about the system, and when the rms response of the system to the input test signal is of the same order of magnitude as the variance of the external noise, the time required to identify an unknown system is an order of magnitude or more greater than the significant length of the impulse response. It is also concluded that, when the noise is white and the test signal is optimized, no measurement technique will yield a smaller identification time than that of the ideal identifier. It is pointed out that further reduction in identification time could probably be achieved by identification schemes making maximum use of all available *a priori* knowledge about the system.

CHAPTER 1

THE IDENTIFICATION PROBLEM

The problem of system identification and its relation to the currently active area of adaptive systems is discussed in this introductory chapter. The framework, into which the primary problem considered in this research fits, is set by the identification requirements of an adaptive system and the classification of identification techniques. The specific problem is stated, and the basic assumptions, upon which the analysis is built, are given. Finally, to provide a starting point, a review of some of the properties of impulse responses is presented.

1.1 The Identification Problem and its Relation to Adaptive Systems

The general identification problem consists of determining a complete description of the relationships between the input and output signals of an unknown system having input signals $x_1, x_2, \ldots, x_j$ and output signals $y_1, y_2, \ldots, y_k$. In general the unknown system may be non-linear and time-varying and the number of input signals, j, need not equal the number of output signals, k. The behavior of the unknown system is to be determined by making suitable tests among the various inputs and outputs. This problem has been discussed by Zadeh [33], Lee [15], Woodrow [30], Moore [19], and others. Current interest in the identification problem has been stimulated by recent work in the area of adaptive control systems.

There is not, as yet, a generally accepted definition of an adaptive control system, but one which has been widely used is the following: "An adaptive system is one which is provided with a means of continuously monitoring its own performance in relation to a given index of performance or optimum condition and a means of modifying its own parameters by closed loop action so as to approach this optimum." [5, Cooper, Gibson, et. al.] This definition implies that an adaptive system must be capable of performing the following functions: provide continuous information about the present state of the system or identify the process; compare present system performance to the desired or optimum performance and make a decision to adapt the system so as to achieve optimum performance; and finally, initiate a proper modification so as to drive the control system to the optimum. These three principles, identification, decision, and modification are inherent in any adaptive system. This functional breakdown of an adaptive system is similar to that proposed by Aseltine et. al. [2]. Furthermore, this breakdown is a useful concept for the design of an adaptive system as it clearly places the adaptive nature in evidence.

An identification technique to be useful in adaptive control systems must meet two conditions: first, the identification must be made in the presence of normal operating signals, and any tests performed upon the system must not unduly disturb the normal operation of the control system; second, the identification must be made relatively quickly if the information is to be useful for the decision-making and modification phases of the adaptive process. In order to measure the characteristics of an unknown process it is necessary to supply energy to the system. The former requirement makes it necessary to use low-level test signals or normal operating signals to furnish the energy necessary for system identification. As a result, the response of the system is small and the effects of noise disturbances become important. The influence of noise upon the observation of the system's response determines the length of time that is required to identify the process, and hence, is directly related to the latter requirement of an identification technique.

Although this research has been motivated by the particular requirements placed upon identification techniques by adaptive systems the results are not restricted to this particular application. In view of this, details of the theoretical work have not been specifically related to the adaptive problem; however, the examples which are considered are discussed from the viewpoint of adaptive systems, and a section of the last chapter is devoted to the discussion of the significance of the results of this research for adaptive systems.

In any application requiring the identification of an unknown system it is necessary to specify how the process is to be described, what prior knowledge is available, and how the system is to be excited. The unknown system may be described in a complete or partial manner. Examples of complete system specification include such items as the values of all independent parameters, and the time response or transfer function relationships between the various inputs and outputs. The latter two methods of description apply only to linear systems. An unknown process can be described in a partial manner by specifying such quantities as gain, rise time, and overshoot, resonant frequency and relative damping ratio, or the describing function.

Identification techniques can also be classified in terms of whether or not they require some _a priori_ knowledge about the systems characteristics. The availability of _a priori_ knowledge about the system to be identified can range from a complete lack of any prior knowledge at one extreme to complete knowledge of the system behavior at the other. In most engineering situations some _a priori_ knowledge is available; in some instances the order of the system is known, while in others the ranges and/or rates of change of the system parameters may be known.

The source of energy used to excite the system offers a third useful method of classifying identification techniques. Identification can be made from observations of the output signals due to the systems normal operating signals. Alternatively, a test signal, designed solely for the purpose of identification, can be applied to the input terminals, and the system response observed. Advantages and disadvantages of each method have been discussed by Cooper and Gibson, et. al. [5].

1.2 Statement of the Problem and Basic Assumptions

The aim of this research is to determine fundamental limits on the time that is required to estimate, with a specified accuracy, the impulse response of a linear system when the measurement technique is corrupted by external noise signals. Thus, the problem is one of making a complete identification as opposed to a partial identification. The analysis, for simplicity, is restricted to systems with a single input and a single output although the results are applicable, with a suitable modification of notation, to multidimensional systems. Of primary interest is the determination of a conservative limit, a greatest lower bound, on the identification time. For this reason only identification techniques that do not require any a priori knowledge of the system are considered, because prior knowledge, if properly used, can only serve to reduce the identification time. As an example, consider the limiting case where the system is known exactly. Then it is not even necessary to make a measurement to identify the system . . . . identification can be achieved in zero time. Rather than tie the identification process to the properties of normal operating signals, which vary depending upon the particular application, only identification techniques using test signals are treated. In summary, the problem is the investigation of the identification time requirements of the class of identification techniques which, completely identify a linear system in terms of its impulse response, do not require any a priori knowledge about the system, and receive their energy from special test signals.

The basic identification problem is illustrated in Fig. 1-1 along with the notation used for the input test signal, output signal, external noise, and observed signal. The input test signal is assumed, in most instances, to be a known deterministic quantity. Desireable properties of a test signal include: a small mean square value, and a small peak power, so that the normal operation of the system is not seriously disturbed; and a wide bandwidth, so that the high frequency characteristics of the system can be measured. Practical systems are not truly band-limited (zero transmission above some cutoff frequency) so that it would be necessary, in theory, to use an infinite bandwidth test signal in order to obtain an exact representation of the systems impulse response. In practice, however, if the equivalent noise bandwidth [18, Middleton, p. 684] of the test signal is wide compared to the equivalent noise bandwidth of the system, the errors in the estimate of the impulse response due to finite bandwidth test signals are small. Since the main interest of this research is the errors in the impulse response estimate due to the external disturbances it is assumed that any errors due to the practical limitations of the test signal are much smaller than those caused by the noise. Equivalently, the test signal bandwidth is assumed to be large compared to the bandwidth of the known system.

The output signal is assumed to be unmeasureable thereby requiring the identification to be based upon measurements of the observed signal. A stationary ergodic random process, with zero mean, is assumed for the external noise disturbance. Chapter 6 illustrates how noise signals originating at the input, or within the system, can be represented by

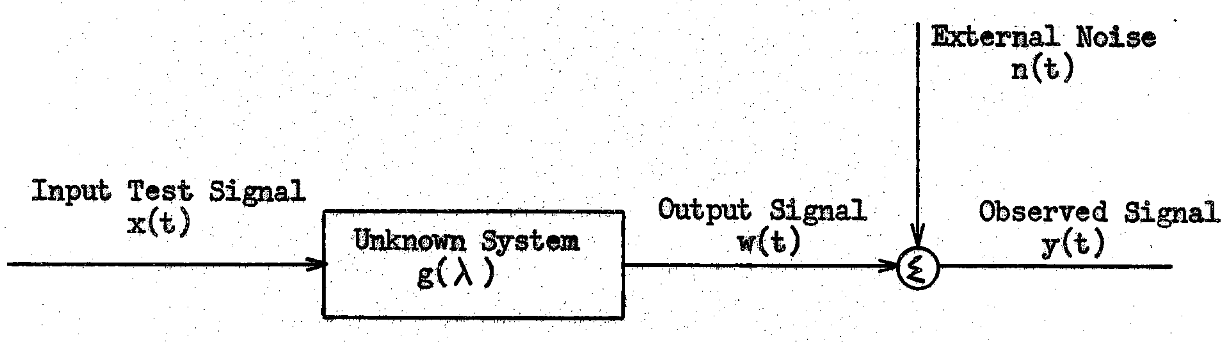


Fig. 1-1

The Basic Identification Problem

equivalent noise signals introduced at the output as shown in Fig. 1-1.

Chapter 2 introduces the concept of an ideal identifier and determines the identification time required by an ideal identification technique. The results are independent of the particular system that is being identified, and do not depend upon any particular measurement technique. Thus the identification time of the ideal identifier serves as a useful basis to which practical identification schemes can be compared.

The effects of external noise upon the identification time of three practical identification techniques are analyzed in Chapters 3, 4, and 5. No initial claim is made with regard to whether or not these techniques are optimum or not, and it is gratifying that each of the three methods, crosscorrelation, sampling, and matched filter, turn out to be equivalent to the ideal identifier.

Examples are considered in Chapter 6 in order that the theoretical results of the preceding chapters can be tied down to some practical problems. The importance of the identification problem in adaptive systems justifies discussion of the examples from the adaptive viewpoint. However, in keeping with the objective of this work, detailed analysis of the effects of normal operating signals upon the various identification techniques is not considered. The operating signals, from the viewpoint of the identification problem, are unwanted or noise signals which tend to increase the identification time. The particular requirements of an identification technique for an adaptive system make it convenient to express the identification time in terms of the systems gain-bandwidth product and the signal-to-noise ratio found at the output of the system under test. Since the exact nature of the system is generally unknown, it is only possible to consider average values of the gain and bandwidth. The final chapter summarizes the work, comments on the equivalence of the various identification techniques, discusses the significance of this work with respect to adaptive systems, and presents some related problems.

1.3 Characteristics and Representation of Impulse Responses

A review of some of the properties of impulse responses, along with several analytical and graphical techniques of representing impulse responses, is given in this section. This summary of facts will serve as a starting point for consideration of the identification problem outlined above.

A linear system is one whose input-output characteristics are described by a linear differential equation of the form

$$(a_n p^n + a_{n-1} p^{n-1} + \ldots + a_1 p + a_o)\, x(t) = (b_m p^m + b_{m-1} p^{m-1} + \ldots + b_1 p + b_o)\, w(t) \qquad m \leq n \tag{1-1}$$

where $x(t)$ is the input signal, $w(t)$ is the output signal, and p is the operator $\frac{d}{dt}$. The condition $m \leq n$ is necessary for the physical realizability of the system. In general the coefficients a_i and b_i are functions of time but are independent of x. The behavior of the system is completely determined if all the a_i and b_i are known as functions of time. A useful description of a linear system is the unit impulse response which is the solution of Eq. (1-1) for $w(t)$ when the input signal is a unit impulse, i.e., $x(t) = \delta(t)$, where $\delta(t)$ denotes the unit delta function. A knowledge of the impulse response of a linear system gives a complete description of the system. Assuming there is no initial stored energy, it is possible, by means of the integral equation

$$w(\lambda) = \int_{-\infty}^{\lambda} g(\lambda, t)\, x(t)\, dt \tag{1-2}$$

to predict the behavior of the system to any input $x(t)$ if the behavior is known when $x(t) = \delta(t)$. The impulse response is denoted by $g(\lambda, t)$, and is interpreted as the value of the output at time λ when a unit impulse is applied at time t. When the system is time-invariant $g(\lambda, t)$ becomes $g(\lambda - t)$ and the impulse response may be represented graphically as in Fig. 1-2. For the time-varying case $g(\lambda, t)$ may be represented as the height of a surface above the λ, t plane as shown in Fig. 1-3. It is a property of physical systems that $g(\lambda, t) = 0$ for $\lambda < t$. This restriction is due to the fact that the system cannot respond before the excitation is applied. A second property of physically realizable systems is

$$\int_0^{\infty} |g(\lambda, t)|^2 \, d\lambda < \infty \qquad (1\text{-}3)$$

In most practical cases expression (1-3) implies that the impulse response approaches zero as $(\lambda - t)$ becomes large. Thus in both Fig. 1-2 and Fig. 1-3, $g(\lambda, t)$ is zero for $\lambda < t$, and the impulse response function is essentially zero for large values of $(\lambda - t)$.

In the absence of external noise the time required to measure the impulse response of an unknown system is equal to the significant duration of the impulse response. The identification could be achieved by applying an impulse to the system's input and observing the output. Information about the form of $g(\lambda, t)$ cannot be obtained any faster than the inherent delay of the system allows; thus, the significant duration of $g(\lambda, t)$ represents a lower limit on the identification time in the absence of noise. The effects of external noise will increase the identification time, and a limitation on the identification time under these conditions is the main result of this work.

The techniques for system identification considered here require that

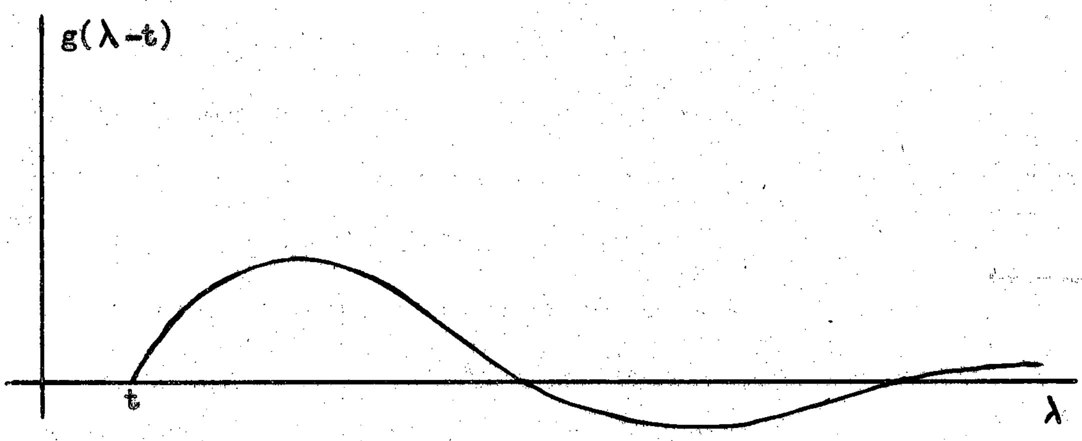


Fig. 1-2

Impulse Response of a Typical Time-Invariant System

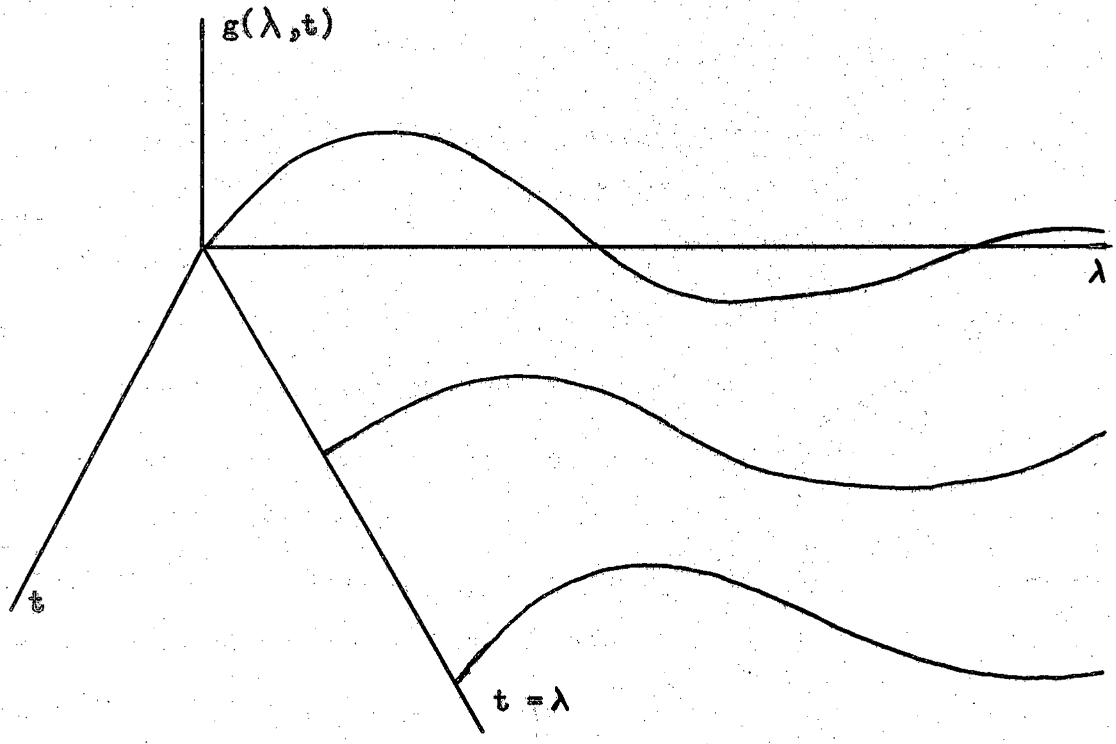


Fig. 1-3
Impulse Response of a Typical Time-Varying System

the system be, at most, slowly time-varying. In terms of Fig. (1-3) this means that variations in the height of the $g(\lambda, t)$ surface along lines parallel to the $\lambda = t$ line must be slow compared to the significant length of the impulse response.

The impulse response of a system may be given as a mathematical function of time such as

$$g(\lambda) = k\, e^{-\zeta\omega_o\lambda} \sin \omega_o \sqrt{1-\zeta^2}\, \lambda \qquad (1\text{-}4)$$

for a simple second-order system. Another common way of representing an impulse response is by means of a graph such as the one in Fig. 1-2 (or Fig. 1-3 for the time-varying case). Sometimes instead of a complete graph only sample points of the impulse response curve are given. (Fig. 1-4) In practice some error is introduced by the sampling process, but in most engineering applications this error approaches zero as the number of sampling points approaches infinity. The relation between the sampling rate and the test signal bandwidth is pointed out at the end of Chapter 3.

Another method of representing an impulse response is by a Taylor's series expansion.

$$g(\lambda) = g(\lambda_o) + (\lambda-\lambda_o)\, g'(\lambda_o) + \frac{(\lambda-\lambda_o)^2}{2!} g''(\lambda_o) + \dots + \frac{(\lambda-\lambda_o)^n}{n!} g^{(n)}(\lambda_o) + \dots \qquad (1\text{-}5)$$

The nature of impulse responses of practical systems indicates that, in general, a large number of terms will be required in the Taylor's series expansion to achieve a good approximation to the actual impulse response.

An identification technique based upon a Taylor's series expansion of the impulse response has been suggested by Braun [4] . In his paper Braun shows that a Taylor's series expansion of the impulse response about a point t_o can be computed by applying an abrupt change in the input signal, $\Delta x(t)$,

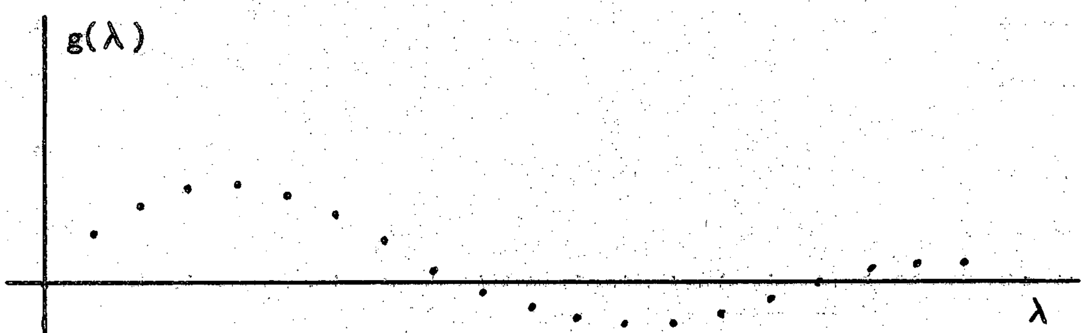


Fig. 1-4

Sample Point Approximation of an Impulse Response at t_0 and measuring the derivatives of the output signal just prior to, and just after $t = t_0$. It is felt that the necessity of measuring the derivatives of the output signal, especially in the presence of noise, imposes a serious practical limitation on the method. For this reason no further consideration is given to this identification scheme.

A different kind of series representation of the impulse response, also useful in the identification problem, is a series of orthonormal functions. $g(\lambda)$ may be expressed as

$$g(\lambda) = \sum_{i=1}^{\infty} \alpha_i \Psi_i(\lambda) \qquad (1\text{-}6)$$

where the Ψ_i are a set of orthogonal functions satisfying the conditions

$$\int_0^{\infty} \Psi_i(\lambda)\Psi_j(\lambda)\, d\lambda = 0 \qquad i \neq j \quad i,j = 1,2,3\ldots \qquad (1\text{-}7)$$

and $$\int_0^{\infty} \Psi_i^2(\lambda)\, d\lambda = 1 \qquad i = 1,2,3,\ldots \qquad (1\text{-}8)$$

and the constants, α_i, are given by

$$\alpha_i = \int_0^{\infty} g(\lambda)\, \Psi_i(\lambda)\, d\lambda \qquad (1\text{-}9)$$

Methods of measuring the coefficients of an orthonormal series expansion of an impulse response are considered briefly in Chapter 7.

CHAPTER 2

IDENTIFICATION TIME REQUIRED BY AN IDEAL IDENTIFIER

The concept of an ideal identifier is introduced in this chapter, and the identification time required by an ideal identifier is determined. Ideal identification is based on statistical parameter estimation, and, therefore, the results do not depend upon any particular data processing technique. Following the definition of the ideal identifier a discussion on the estimation of P unknown parameters is given. These results are applied to the measurement of an unknown impulse response, and an expression for the corresponding identification time is derived. The results obtained serve as a basis to which the practical identification techniques discussed in succeeding chapters can be compared.

2.1 The Ideal Identifier

An ideal identification scheme is one which, for a given observation period called the identification time, has as its output signal a minimum variance estimate of the unknown impulse response. It is well known [3, Bello] that the mean value of the *a posteriori* probability density function of an unknown parameter is a minimum variance estimator. In the context of the identification problem this *a posteriori* probability density function is

$$p\left[g(\lambda) \mid y(t), x(t)\right] \tag{2-1}$$

This is the conditional density function associated with the event $g(\lambda)$ being present given the conditions that a test signal $x(t)$ was applied to the input of the system and a signal (plus noise) $y(t)$ was observed at the output. For any given situation the test signal $x(t)$, is not a random quantity as it is assumed to be known exactly. Thus expression (2-1) may be replaced by

$$p\,[g(\lambda) \mid y(t)] \tag{2-2}$$

The *a posteriori* probability density function of $g(\lambda)$ may be expressed as

$$\begin{aligned} p[g(\lambda) \mid y(t)] &= \frac{p\,[g(\lambda)]\; p\,[y(t) \mid g(\lambda)]}{p\,[y(t)]} \\ &= K\;\; p\,[g(\lambda)]\; p\,[y(t) \mid g(\lambda)] \end{aligned} \tag{2-3}$$

where $p[g(\lambda)]$ is the prior probability density function of $g(\lambda)$ and $p[y(t) \mid g(\lambda)]$ is the likelihood function.* In Eq. (2-3) and in what follows K is a constant chosen so that the area under the associated density function is normalized to unity.

Under certain rather general conditions, maximum likelihood estimates are very nearly equal to the minimum variance estimates obtained from the mean of the *a posteriori* probability density function. These conditions are; first, the prior probability density function must not be sharply peaked, or it must at least be slowly varying compared to the *a posteriori* probability density function; second, the likelihood function must have a center of symmetry at which its maximum is located. The first condition is satisfied here because a minimum of *a priori* knowledge is assumed about $g(\lambda)$ which is equivalent to stating that all functions $g(\lambda)$ are equally likely *a priori*, i. e., the prior probability density function is a constant. It will be shown below that in the case where the noise is additive and Gaussian the second condition is also satisfied.

---

*The term likelihood function is preferred here because $p[y\,(t) \mid g(\lambda)]$ is considered to be a function of $g(\lambda)$ and, as such, is not interpreted as a density function.

In summary, an ideal identifier is one which, for a specified identification time, gives minimum variance estimates of the unknown impulse response. In many cases maximum likelihood estimates are equivalent to minimum variance estimates.

2.2 The Joint Estimation of P Unknown Parameters

Consider a received signal $w(t)$ which is a function of P unknown parameters $q_1, q_2, \ldots, q_P$. The observed signal, $y(t)$, is equal to the actual signal, $w_a(t)$ (obtained by letting $q_i = q_{i_a}$ $i = 1,2,\ldots, P$, where q_{i_a} is the actual or true value of q_i), plus white Gaussian noise having a power spectral density $\Phi_n(0)$ watts per cps. Woodward has shown [32, p. 66] that the likelihood function for this situation may be expressed as

$$L = p\left[y(t) \mid \{q_i\}\right] = K \exp\left\{-\frac{1}{2\Phi_n(0)}\int_{T_I}\left[y(t)-w(t)\right]^2 dt\right\} \tag{2-4}$$

The integration is to be carried out over the identification period T_I.

Under the symmetry assumption made above the likelihood function is a maximum when all parameters take on their true value. Consequently,

$$\left.\frac{\partial L}{\partial q_i}\right|_{\{q_{i_a}\}} = 0 \quad i = 1, 2, \ldots P \tag{2-5}$$

where the subscript $\{q_{i_a}\}$ indicates that the partial derivative is to be evaluated at the point where all parameters of the set $\{q_i\}$ take on their actual value. Substitution of Eq. (2-4) into Eq. (2-5) results in the conditions

$$\int_{T_I}\left[y(t) - w_a(t)\right] \left.\frac{\partial w}{\partial q_i}\right|_{\{q_{i_a}\}} dt = 0 \quad i = 1,2,\ldots P \tag{2-6}$$

It shall be assumed that the received signal energy is sufficiently large compared to the power spectral density of the noise that near the true values of the $\{q_i\}$ the received signal may be adequately represented by

$$w(t) = w_a(t) + \sum_{i=1}^{P} \left.\frac{\partial w}{\partial q_i}\right|_{\{q_{i_a}\}} \Delta q_i \tag{2-7}$$

where $\Delta q_i = q_{i_a} - q_i$ is the deviation of the i^{th} parameter from its true value. With this substitution and making use of the conditions of Eq. (2-6) the likelihood function, near its maximum value becomes

$$L = K \exp\left\{\frac{-1}{2\Phi_n(0)}\int_{T_I}\left[[y(t) - w_a(t)]^2 + \sum_{i=1}^{P}\sum_{j=1}^{P}\Delta q_i \Delta q_j \left.\frac{\partial w}{\partial q_i}\right|_{\{q_{i_a}\}}\left.\frac{\partial w}{\partial q_j}\right|_{\{q_{i_a}\}}\right]dt\right\} \tag{2-8}$$

The term

$$\exp\left\{\frac{-1}{2\Phi_n(0)}\int_{T_I}[y(t) - w_a(t)]^2 \, dt\right\} \tag{2-9}$$

is not a function of the $\{q_i\}$ and can be absorbed into the normalizing constant K.

By defining the column matrix or vector

$$(Q) = \begin{bmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ \cdot \\ q_P \end{bmatrix} \tag{2-10}$$

and the square matrix

$$(B) = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1P} \\ b_{21} & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \\ b_{P1} & \cdot & \cdot & b_{PP} \end{bmatrix} \tag{2-11}$$

with elements

$$b_{ij} = \int_{T_I} \left.\frac{\partial w}{\partial q_i}\right|_{\{q_{i_a}\}} \left.\frac{\partial w}{\partial q_j}\right|_{\{q_{j_a}\}} dt \qquad (2\text{-}12)$$

the likelihood function can be put into the form

$$L = K \exp\left\{-\frac{1}{2}(Q)' \frac{1}{\Phi_n(0)}(B) \quad (Q)\right\} \qquad (2\text{-}13)$$

where (Q)' indicates the transpose of the matrix (Q).

The assumption that the received signal energy is large compared to the power spectral density of the noise results in a likelihood function which is a multivariate Gaussian distribution near its maximum. The likelihood function is specified by the elements b_{ij} defined in Eq. (2-12), the actual values of the parameters, and the power spectral density of the noise.

From a single observation of finite duration only an estimate of the likelihood function can be constructed. Thus the location of the maximum of L is itself a random variable. Since $\{q_i\}$ is being estimated by the coordinates of the maximum value of the likelihood function the variances and covariances associated with the estimates of the $\{q_i\}$ are given by the covariance matrix describing the location of the maximum of the likelihood function. Under the assumption that the prior probability density function of $\{q_i\}$ is constant in the neighborhood of the true values of the $\{q_i\}$, the likelihood function is proportional to the <u>a posteriori</u> probability density function, $p[\{q_i\}|y(t)]$.

Now $p[\{q_i\} \mid y(t)] \Delta q_1 \Delta q_2 \ldots \Delta q_p$ is the probability that the true parameters lie in the interval $\{q_i\}$, $\{q_i + \Delta q_i\}$, but since the estimates of the true values of the parameters are given by the set of q_i's that makes L a maximum this is also equal to the probability that the maximum value of L lies in the interval $\{q_i\}$, $\{q_i + \Delta q_i\}$. The likelihood function and the location of the estimate of its maximum are described by the same density function. Therefore, the covariance matrix associated with the maximum value of L is equal to the covariance matrix of the likelihood function.

From Eq. (2-13) and the above discussion it is clear that the covariance matrix associated with estimating the set of parameters $\{q_i\}$ is

$$\Sigma = \Phi_n(0)\,(B)^{-1} \tag{2-14}$$

where the i, j element of this matrix is the covariance of q_i and q_j and is given by

$$\sigma_{ij}^2 = \Phi_n(0)\frac{|B_{ij}|}{|B|} \tag{2-15}$$

where $|B|$ is the determinant of the matrix (B) and $|B_{ij}|$ is the cofactor of the element b_{ij}.

Eq. (2-15) can be used to determine the variance associated with the estimation of any of the parameters, q_i, or the covariance associated with any pair of parameters q_i, q_j. The evaluation of the determinants in Eq. (2-15) may be difficult if the number of parameters under consideration is large. When the covariances are zero, however, or equivalently where a single parameter is estimated under the assumption that all other parameters are known the variance can be expressed as

$$\sigma_{ii}^2 = \frac{\Phi_n(0)}{b_{ii}} \tag{2-16}$$

since $|B| = b_{ii}\,|B_{ii}|$ when all of the elements not on the major diagonal are zero.

There is reason to believe that in any event when the ratio of received signal energy to noise power spectral density is large Eq. (2-16) represents the theoretically minimum variance because correlation between errors can only increase the uncertainty of the estimate.

A physical interpretation of the minimum variance obtained in Eq. (2-16) can be obtained by substituting the definition of b_{ij} into the equation. Thus the minimum variance associated with the estimation of a single parameter is

$$\sigma_{ii}^2 = \frac{\Phi_n(0)}{\int_{T_I} \left[\frac{\partial w}{\partial q_i} \bigg|_{\{q_{i_a}\}} \right]^2 dt} \qquad (2\text{-}17)$$

The quantity

$$\int_{T_I} \left[\frac{\partial w}{\partial q_i} \bigg|_{\{q_{i_a}\}} \right]^2 dt \qquad (2\text{-}18)$$

may be interpreted as the received signal sensitivity with respect to the parameter q_i. The larger this sensitivity the smaller the variance. A parameter will be estimated with a small variance if a small change in this parameter causes a large change in the received signal. The role of $\Phi_n(0)$ in Eq. (2-17) is clear, a large noise power spectral density results in a large minimum variance and vice versa.

2.3 A P Parameter Representation of an Impulse Response

In order to apply the results of the previous section to the estimation of an impulse response function it is first necessary to approximate the continuous function $g(\lambda)$ so that it is describable by a finite set of parameters.

A convenient and common way of approximating an impulse response function is to represent it by a set of numbers obtained by sampling the function at intervals of $\Delta\lambda$ seconds. The sample values, g_p, are taken at the instants $\lambda_p = p\Delta\lambda$, $p = 0, 1, 2, 3, \ldots$. A typical impulse response and its sample values are shown in Fig. 2-1. By taking $\Delta\lambda$ sufficiently small it is possible to approximate $g(\lambda)$ to any degree of accuracy desired. If $g(\lambda)$ is bandlimited with maximum frequency W_M cycles per second the function may be represented exactly by sample values spaced at intervals of $\frac{1}{2 W_M}$ seconds [21, Shannon]. It is assumed that any errors due to approximating $g(\lambda)$ in this manner are much smaller than the estimation errors due to the external noise. This assumption is further justified by the fact that approximation errors are fundamentally deterministic in nature and can be compensated for if necessary, whereas the errors due to the presence of external noise signals are random in nature, and hence cannot be predicted or compensated for.

Physical realizability requires that $g(\lambda) = 0$ for $\lambda < 0$ and in most practical cases impulse responses are essentially zero for λ greater than some Λ. Therefore, $g(\lambda)$ may be represented by the set of parameters

$$g_0, g_1, \ldots, g_P \tag{2-19}$$

where P is an integer greater than $\Lambda/\Delta\lambda$.

2.4 Identification Time Required by an Ideal Identifier

In order to estimate an unknown impulse response it is necessary to excite the input of the system with a test signal, $x(t)$, and then observe the output signal (plus noise), $y(t) = w(t) + n(t)$. The procedure and nomenclature is shown in Fig. 1-1.

The output signal is related to the input signal by the convolution integral

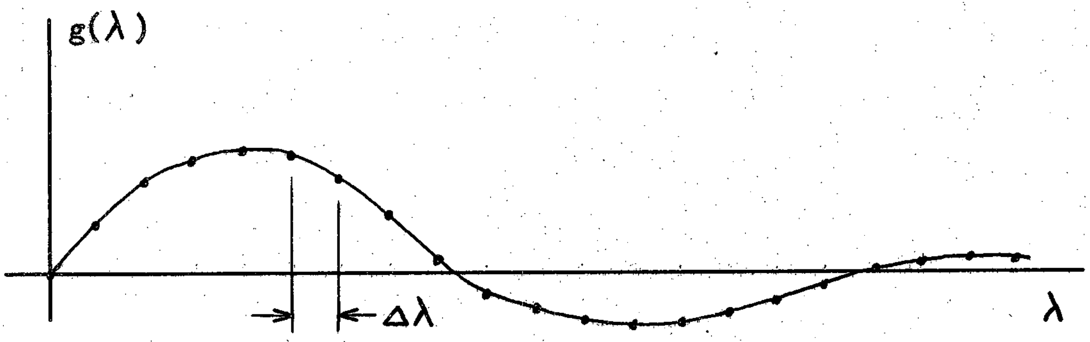


Fig. 2-1

An Impulse Response and Sample Values

$$w(t) = \int_0^{\infty} x(t-\lambda)\, g(\lambda)\, d\lambda \qquad (2\text{-}20)$$

or using the sampling technique introduced in the last section

$$w(t) = \Delta\lambda \sum_{p=0}^{P} x(t-\lambda_p)\, g_p \qquad (2\text{-}21)$$

In order to apply the results of Eq. (2-15) and Eq. (2-16) it is necessary to calculate the partial derivatives $\left.\frac{\partial w(t)}{\partial g_i}\right|_{\{g_{i_a}\}}$. From the definition of the partial derivative

$$\frac{\partial w}{\partial g_i} = \lim_{\Delta g_i \to 0} \frac{\Delta\lambda \left\{ \sum_{p=0}^{P} x(t-\lambda_p)\,(g_p + \Delta g_i) - \sum_{p=0}^{P} x(t-\lambda_p)\, g_p \right\}}{\Delta g_i}$$

$$= \Delta\lambda\, x(t-\lambda_i) \qquad (2\text{-}22)$$

Thus, the elements of the matrix in Eq. (2-15) are

$$b_{ij} = (\Delta\lambda)^2 \int_{T_I} x(t-\lambda_i)\, x(t-\lambda_j)\, dt \qquad (2\text{-}23)$$

The problem of choosing a test signal in some optimum way so as to minimize the variances and covariances associated with estimating the parameters $g_0, g_1, \ldots, g_P$ is now considered. It is apparent from Eq. (2-15) and Eq. (2-23) that the variances become smaller as the amplitude of the test signal increases. In a practical situation, however, the amplitudes of the test signal x(t), are restricted. It is, therefore, of interest to determine the form of x(t) which minimizes the variances subject to some amplitude constraint. The constraint that the signal energy

$$\int_{T_I} x^2(t)\, dt = E_x = T_I \overline{x^2} \tag{2-24}$$

remain fixed will be imposed. If, in addition to this, the test signal is periodic with the identification time, T_I, equal to the fundamental period of $x(t)$, then all elements along the major diagonal of the matrix (B) are equal. Note also that (B) is symmetric and positive definite.

It can be shown [16, Levin] that if a symmetric positive definite matrix (M) has for each element along its principal diagonal the value β and arbitrary values elsewhere, then the elements along the principal diagonal of $(M)^{-1}$ will all reach their minimum value of $1/\beta$ if and only if $(M) = \beta\,(I)$, where (I) is the identity matrix.

In order to satisfy these conditions the test signal must satisfy the following set of conditions

$$\int_{T_I} x(t - \lambda_i)\, x(t - \lambda_j)\, dt = \begin{cases} 0 & i \neq j \\ E_x & i = j \end{cases} \tag{2-25}$$

The integral of Eq. (2-25) is proportional to the autocorrelation function of the test signal, and from the well known properties of autocorrelation functions the requirement of Eq. (2-25) is equivalent to requiring that the test signal be white.

This result is not too surprising since if the identification scheme is to reproduce the fine structure of $g(\lambda)$ it is necessary that $x(t)$ contain high frequency components. Also, since the external noise is white it seems entirely reasonable to spread the test signal energy equally over all frequencies.

For the optimum test signal then

$$b_{ij} = \begin{cases} 0 & i \neq j \\ E_x(\Delta\lambda)^2 = T_I\,\overline{x^2}\,(\Delta\lambda)^2 & i = j \end{cases} \tag{2-26}$$

and from Eq. (2-17), since the variances of each g_p are equal, the variance associated with the estimation of the impulse response $g(\lambda)$ is

$$\sigma_{\hat{g}}^2 = \frac{\Phi_n(0)}{T_I\,\overline{x^2}\,(\Delta\lambda)^2} \tag{2-27}$$

The corresponding identification time is

$$T_I = \frac{\Phi_n(0)}{\overline{x^2}\,\sigma_{\hat{g}}^2\,(\Delta\lambda)^2} \tag{2-28}$$

It is important to note that the identification time required by the ideal identifier depends only upon the power spectral density of the external noise, the mean square value of the test signal, and the sampling interval, $\Delta\lambda$. The identification time is independent of the impulse response being estimated, as long as the assumptions implied by Eq. (2-7) are valid; that is, as long as the received signal energy is large compared to the noise power density spectrum. It will be shown in Chapter 6 that this condition is generally satisfied in practical situations. One might suppose that it would be possible to reduce the settling time to any desired degree by simply increasing the mean square value of the test signal, or by increasing the sampling interval. In practice the mean square value of the test signal is limited by such considerations as the effects of normal control signals, power limitations, and possible large signal nonlinear effects. The constraint upon $\overline{x^2}$ is determined by the particular application. $\Delta\lambda$ cannot be increased arbitrarily either, because then the errors due to sampling would become as large or larger than the errors resulting from the external noise. This condition would be contrary to the assumption made in Section 2-3.

CHAPTER 3

IDENTIFICATION TIME REQUIRED BY CROSSCORREOATION IDENTIFICATION

The application of crosscorrelation techniques to system identification is analyzed in this chapter. The sources of errors associated with crosscorrelation identification are discussed at some length, and the noise terms of the output signal are analyzed. When a random test signal is used the output signal has noise terms arising from two sources, the external noise, and the test signal itself. It is pointed out, that by using a periodic test signal and an ideal finite-memory integrator for an averaging filter, that the latter noise term can be eliminated entirely. A considerable saving in identification time results. The identification time requirement of the corsscorrelation technique is compared to the requirement of an ideal identifier. Crosscorrelation is found to be equivalent to the ideal identifier.

3.1 Theoretical Basis for Crosscorrelation Identification

The use of crosscorrelation techniques for the identification of linear systems is not new. It was probably first introduced by Lee [15] and the method has been applied to the identification problem of adaptive systems by Anderson, Buland, and Cooper [1] . Measurement of the impulse response of a time-invariant or slowly time-varying linear system by means of crosscorrelation is based upon the following theoretical development. A test signal, which for the present will be assumed to be a sample of a stationary ergodic random process, is applied to the input of the system under test. The output signal of the system is then crosscorrelated with the input test signal. The details of the method are illustrated in Fig. 3-1.

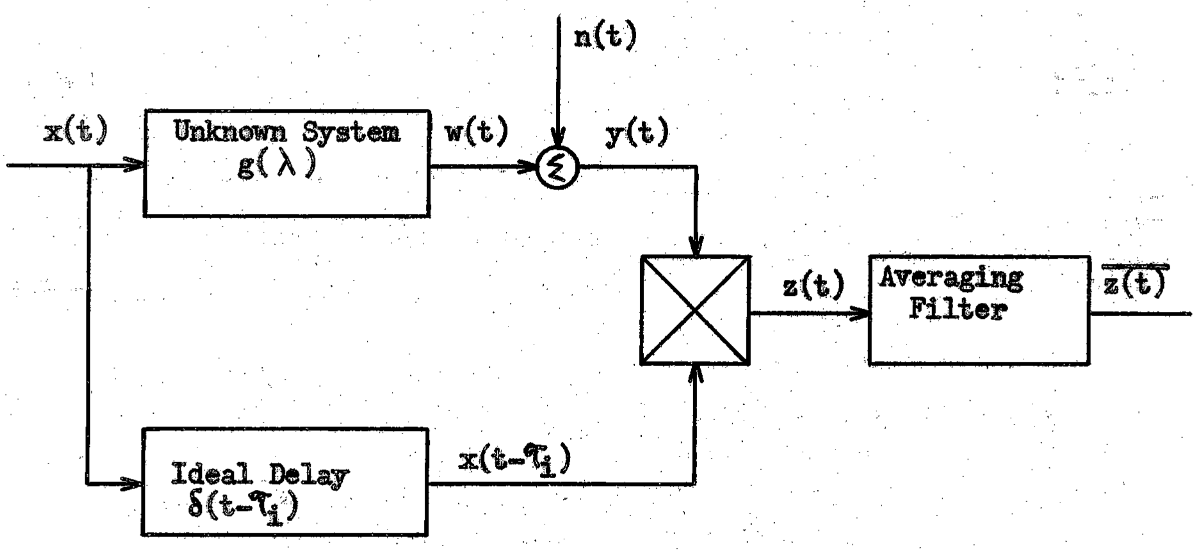


Fig. 3-1

Crosscorrelation Identification

The output signal w(t) is given by the relation

$$w(t) = \int_0^\infty x(t - \lambda_1)\, g(\lambda_1)\, d\lambda_1 \tag{3-1}$$

It is assumed that w(t) is unmeasurable and that only y(t) = w(t) + n(t) is available for observation. n(t) is an external noise signal assumed to be from a zero mean stationary random process described by a noise autocorrelation function $\phi_n(\tau)$, and corresponding power spectral density, $\Phi_n(\omega)$. The crosscorrelation of x(t) and y(t) is achieved by the multiplier and averaging filter. The signal at the output of the multiplier is

$$z(t) = \int_0^\infty x(t - \tau_i)\, x(t - \lambda_1)\, g(\lambda_1)\, d\lambda_1 + n(t)\, x(t - \tau_i) \tag{3-2}$$

The crosscorrelation function between the input and output signal, $\phi_{xy}(\tau)$, of the system under test is the mathematical expectation of Eq.(3-2). Under appropriate conditions the interchange of the integration and expectation operations is justified [9, Doob, Theorem 2.7]. Thus

$$\phi_{xy}(\tau) = \int_0^\infty E[x(t - \tau_i)\, x(t - \lambda_1)]\, g(\lambda_1)\, d\lambda_1 + E[n(t)\, x(t - \tau_i)] \tag{3-3}$$

Since the input test signal and the external noise are assumed to be statistically independent the last term in Eq. (3-3) is zero. The expected value of the product in the first term of Eq. (3-3) is recognized as the test signal autocorrelation function, $\phi_x(\tau)$, so that

$$\phi_{xy}(\tau_i) = \int_0^{\infty} \phi_x(\tau_i - \lambda_1)\, g(\lambda_1)\, d\lambda_1 \tag{3-4}$$

A solution of this integral equation for $g(\lambda)$ is difficult in general, but when the test signal is wideband compared to the bandwidth of the system $\phi_x(\tau_i - \lambda_1)$ can be approximated as

$$\phi_x(\tau) \approx \Phi_x\, \delta(\tau) \tag{3-5}$$

where Φ_x is the area under the $\phi_x(\tau)$ function and $\delta(\tau)$ is the unit delta function. With this approximation, Eq. (3-4) becomes

$$\phi_{xy}(\tau_i) = \Phi_x\, g(\tau_i) \tag{3-6}$$

It is evident from Eq. (3-6) that the crosscorrelation technique shown in Fig. 3-1 can be used to measure a particular sample point of the unknown impulse response. Complete identification is achieved by using a number of such correlation channels in parallel.

3.2 Nature of the Errors Associated with Crosscorrelation Identification

From the definition of the autocorrelation function of an ergodic random process [14, Laning and Battin, p. 113]

$$\phi_x(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} x(t)\, x(t + \tau)\, dt \tag{3-7}$$

it is evident that the output of the averaging filter in Fig. 3-1, $\overline{z(t)}$, will equal $\phi_{xy}(\tau_i)$ only if the filter integration time is infinite. When the integration time is finite, as it will be for any physically realizable filter, there are error or noise terms in $\overline{z(t)}$ as well as the signal term, $\Phi_x g(\tau_i)$. The errors arise from three sources: first, there is an error due to the randomness of the test signal; second, there is an error component resulting from the presence of the external noise; and third, there is an error introduced by the assumption implied by Eq. (3-5).

It will be shown below that the error due to the randomness of x(t) can be eliminated by the use of periodic test signal, and an ideal finite-memory integrator as an averaging filter. The external noise error cannot be eliminated completely.

The approximation made in Eq.(3-5) is equivalent to assuming that the test signal is white noise. It is not possible, in practice, to generate truly white noise, but if the bandwidth of x(t) is much wider, say 100 times wider, than the bandwidth of $g(\lambda)$, the two functions whose product is to be integrated will resemble those shown in Fig. 3-2. The exact nature of the statistical properties of x(t) is unimportant. If, however, the time duration of $\phi_x(\tau)$ is much smaller than the time duration of $g(\lambda)$, as it is in the case shown in Fig. 3-2, then, from an engineering viewpoint, the integral of the product of the two functions is adequately given by Eq.(3-6).

The autocorrelation functions of two practical noise signals which would be suitable as test signals are illustrated in Fig. 3-3. Both are binary noise signals with states $\pm$ X. The first signal has Poisson distributed zero crossings and takes the values $\pm$ X with equal probability. The autocorrelation function for this signal is

$$\phi(\tau) = X^2 e^{-2a|\tau|} \qquad (3\text{-}8)$$

where a is the average number of zero crossings per unit time. The second signal is called discrete-interval binary noise [1, Anderson, et.al.]. The signal changes state only at the specific times ℓt_1, where ℓ is an integer and t_1 is the minimum interval between changes in

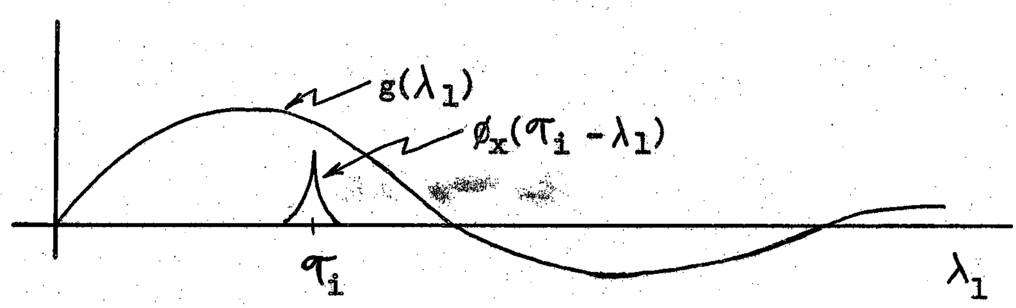
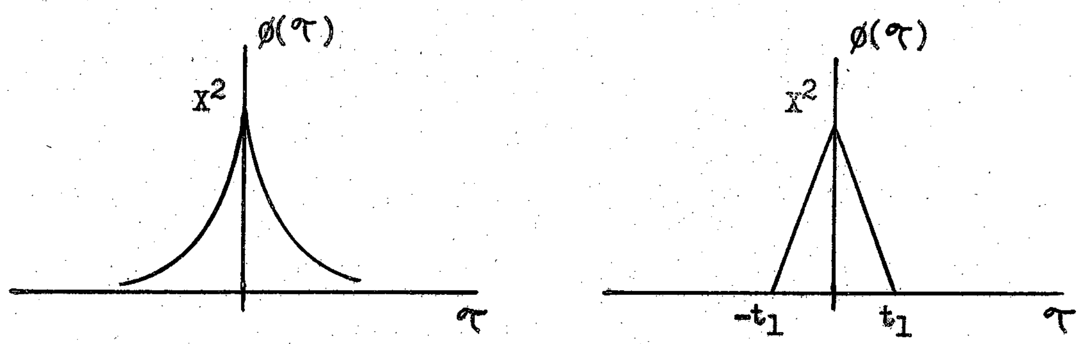


Fig. 3-2

Convolution of $\phi_x(\tau)$ and $g(\lambda)$



a) Binary noise with Poisson distributed zero crossings.

b) Discrete-interval binary noise.

Fig. 3-3

Two Practical Test Signal Autocorrelation Functions state; the value in successive intervals are independent, and both states are equally probable.

The autocorrelation function for discrete-interval binary noise is [22, Truxal, p. 433]

$$\phi(\tau) = \begin{cases} X^2\left(1 - \frac{|\tau|}{t_1}\right) & |\tau| \le t_1 \\ 0 & |\tau| > t_1 \end{cases} \tag{3-9}$$

It is evident from Fig. 3-3 that the autocorrelation functions of these signals can be made as narrow as desired by a suitable choice of a or t_1 thereby making the error associated with the assumption of Eq. (3-5) as small as desired.

Since the size of the error due to the fact that the test signal bandwidth is not infinite is to some extent controllable, and since the nature of the error allows its exact calculation from a knowledge of $\phi_x(\tau)$, the error in Eq. (3-6) due to assuming that x(t) is white noise will be assumed to be negligible compared to the error introduced by the finite integration time of the averaging filter.

3.3 Analysis of Output Noise

The noise components of z(t), the output signal of the multiplier, can be studied by obtaining the autocorrelation function of z(t). From Eq. (3-2)

$$z(t)\, z(t+\tau) = n(t)\, n(t+\tau)\, x(t-\tau_i)\, x(t+\tau-\tau_i)$$

$$+ \int_0^\infty n(t)\, x(t-\tau_i)\, x(t+\tau-\tau_i)\, x(t+\tau-\lambda_2)\, g(\lambda_2)\, d\lambda_2$$

$$+\int_0^\infty n(t+\tau)\, x(t+\tau-\tau_i)\, x(t-\tau_i)\, x(t-\lambda_1)\, g(\lambda_1)\, d\lambda_1$$

$$+\int_0^\infty \int_0^\infty x(t-\tau_i)\, x(t+\tau-\tau_i)\, x(t-\lambda_1)\, x(t+\tau-\lambda_2) g(\lambda_1) g(\lambda_2) d\lambda_1 d\lambda_2 \tag{3-10}$$

The autocorrelation function $\phi_z(\tau)$ is the ensemble average of Eq. (3-10). Since n(t) and x(t) have been assumed to be statistically independent the fourth product moments which result from the first three terms in Eq. (3-10) can be factored. The expected value of the second and third terms are zero since $E[n(t)] = 0$. Thus $\phi_z(\tau)$ becomes

$$\phi_z(\tau) = \phi_x(\tau)\, \phi_n(\tau)$$

$$+\int_0^\infty \int_0^\infty E[x(t-\tau_i)\, x(t+\tau-\tau_i) x(t-\lambda_1) x(t+\tau-\lambda_2)]\, g(\lambda_1) g(\lambda_2)\, d\lambda_1 d\lambda_2 \tag{3-11}$$

In order to simplify the rest of the analysis x(t) will be assumed to be a Gaussian process. While this assumption probably is not valid in general, consideration of this simplified problem will permit the study of the general behavior of the various noise terms in z(t) without unnecessarily complicating the mathematics. This action is further justified by the fact that it will be shown later that errors caused by noise components resulting from the second term of Eq. (3-11) can be completely eliminated by a judicious choice of test signal and averaging filter.

For a Gaussian process the fourth product moment appearing in Eq. (3-11) can be factored into [14, Laning and Battin, p. 161]

$$E[x(t-\tau_i)\,x(t+\tau-\tau_i)]\;E[x(t-\lambda_1)\,x(t+\tau-\lambda_2)]$$
$$+\;E[x(t-\tau_i)\,x(t-\lambda_1)]\;E[x(t+\tau-\tau_i)\,x(t+\tau-\lambda_2)]$$
$$+\;E[x(t-\tau_i)\,x(t+\tau-\lambda_2)]\;E[x(t+\tau-\tau_i)\,x(t-\lambda_1)] \qquad (3\text{-}12)$$

Each factor in the above expression can be identified with an autocorrelation function. Using Eq. (3-12) and Eq. (3-5) the expression for $\phi_z(\tau)$ becomes

$$\phi_z(\tau) = \phi_n(\tau)\,\phi_x(\tau)$$

$$+\;\Phi_x^2 \int_0^\infty \int_0^\infty \Big\{ \delta(\tau)\,\delta(\tau+\lambda_1-\lambda_2) + \delta(\tau_i-\lambda_1)\,\delta(\tau_i-\lambda_2)$$

$$+\;\delta(\tau+\tau_i-\lambda_2)\,\delta(\tau+\lambda_1-\tau_i)\Big\}\, g(\lambda_1)\, g(\lambda_2)\; d\lambda_1\; d\lambda_2 \qquad (3\text{-}13)$$

Carrying out the integrations over the delta functions,

$$\phi_z(\tau) = \phi_x(\tau)\phi_n(\tau) + \Phi_x^2\, g^2(\tau_i) + \Phi_x^2\,\delta(\tau)\int_0^\infty g(\lambda_1)\, g(\lambda_1+\tau)\; d\lambda_1$$

$$+\;\Phi_x^2\, g(\tau_i+\tau)\; g(\tau_i-\tau) \qquad (3\text{-}14)$$

The second term in Eq. (3-14) is recognized as stemming from the signal component of z(t); the remaining terms represent noise.

The $\phi_x(\tau)\ \phi_n(\tau)$ term in the $\phi_z(\tau)$ expression results from the external noise, n(t), and this noise component in z(t) cannot, in general, be entirely eliminated by the use of an averaging filter with a finite integration time. The other two noise terms are a result of the random properties of the test signal alone. If instead of a continuous noise sample, x(t) is generated by taking a noise sample of length T_x seconds and repeating it periodically the autocorrelation function $\phi_x(\tau)$ as well as the noise components in z(t) due to x(t) will also be periodic with period T_x. Hence, the average of these noise terms over one period will be equal to the average over all time. If the test signal has a zero mean value this average will also be zero. The optimum averaging filter for a periodic test signal is an ideal finite-memory integrator with memory time ℓT_x, where ℓ is a positive integer. By using a periodic test signal and an ideal finite-memory integrator all noise terms except those resulting from the external noise, can be eliminated.

When x(t) is periodic $\phi_x(\tau)$ takes the form

$$\phi_x(\tau) = \sum_{k=-\infty}^{+\infty} \phi_{x_1}(\tau - k\,T_x) \qquad (3\text{-}15)$$

where T_x is the period, and $\phi_{x_1}(\tau)$ is zero outside the interval $-T_x/2 < \tau < +T_x/2$. In order to preserve the quality of the signal component of z(t) the periodic noise sample must be chosen so that $\phi_{x_1}(\tau)$ is narrow compared to the time duration of the impulse response, and in addition the period, T_x, must be large compared to the significant length of $g(\lambda)$ so that only the k = 0 term in the output signal expression

$$\overline{z(t)} = \Phi_x \sum_{k=-\infty}^{+\infty} g(\tau_i - k\,T_x) \qquad (3\text{-}16)$$

is important. (When x(t) is periodic, Φ_x is the area under $\phi_{x_1}(\tau)$.)

3.4 A Practical Periodic Test Signal

The analysis above clearly points out the noise reduction advantages that can be gained by using a periodic test signal. Other advantages of a periodic test signal are that, once the test signal has been chosen, the problems of noise generation and delay are greatly simplified. This is particularly true if discrete interval binary noise is used. The discrete nature of the noise permits the use of standard digital storage devices to obtain the ideal delay required for crosscorrelation, and, once chosen, the fundamental period of x(t) can be permanently stored in the delay mechanism thereby simplifying the noise generation problem. The binary property of x(t) also simplifies the multiplication operation; multiplication can now be achieved by a simple gating circuit.

The problem of choosing a noise sample that will possess the desired $\phi_{x_1}(\tau)$ is not a simple one. A sample chosen at random may have statistics that differ widely from those of an ideal sample, and the resulting $\phi_{x_1}(\tau)$ may be entirely unsatisfactory. Here again, discrete interval binary noise offers a considerable advantage over other types of excitation. It is possible, because of the binary nature of the noise, to synthesize a nearly ideal sample by computation [27, WADD Technical Report 60-201, Appendix A].

3.5 Identification Time Required by Crosscorrelation

The identification time of the crosscorrelation identification technique is closely related to the mean square value of the noise at the output of the averaging filter $\overline{n_o^2}$. $\overline{n_o^2}$ can be obtained from a knowledge of the power spectral density of the noise components in z(t) and the frequency characteristics of the averaging filter. Thus,

$$\overline{n_o^2} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widetilde{\Phi}_{no}(\omega)\, F(\omega)\, F(-\omega)\, d\omega \tag{3-17}$$

where $F(\omega)$ is the transfer function of the averaging filter and $\widetilde{\Phi}_{no}(\omega)$ is the average output noise power spectral density. It is evident from Eq. (3-14) that in the general case the properties of the noise depends upon $g(\lambda)$ so that it becomes necessary to use an average noise power spectral density in Eq. (3-17). The $\sim$ notation is used to emphasize the fact that the average is with respect to the ensemble of all possible impulse responses. In such applications as adaptive control systems an approximate value of $g(\lambda)$ is known if the adaptive loop is functioning properly. If statistical data is not available the desired or optimum $g(\lambda)$ can be considered to be equal to the average value. When the test signal is periodic the output noise can be made independent of $g(\lambda)$ and the averaging operation becomes trivial.

Since the bandwidth of the averaging filter will be small compared to the noise terms, the variance at the filter output is approximately equal to

$$\overline{n_o^2} = \frac{1}{2\pi}\int_{-\frac{\pi}{T_I}}^{+\frac{\pi}{T_I}} \widetilde{\Phi}_{no}(\omega)\, d\omega \tag{3-18}$$

where T_I is the identification time defined in terms of the filters equivalent noise bandwidth, W_F, by the relation

$$T_I = \frac{1}{2\, W_F} \tag{3-19}$$

In obtaining Eq. (3-18) from Eq. (3-17) it was assumed that the low frequency gain of the filter is unity. No loss in generality results from this assumption because both signal and noise terms are multiplied by the same gain factor.

The identification time required to obtain a specified variance, $\sigma_{\hat{g}}^2$, in the estimation of $g(\lambda)$ can be computed from Eq. (3-18). Recalling from Eq. (3-16) that the signal term is $\Phi_x\, g(\tau_i)$ it follows that

$$\sigma_{\hat{g}}^2 = \frac{\overline{n_o^2(t)}}{\Phi_x^2} = \frac{1}{2\pi \Phi_x^2} \int_{-\frac{\pi}{T_I}}^{+\frac{\pi}{T_I}} \widetilde{\Phi}_{no}(\omega)\, d\omega \qquad (3\text{-}20)$$

In the general case, it is not possible to obtain an analytical expression for T_I, and it is necessary to compute T_I by numerical techniques. However, when $\widetilde{\Phi}_{no}(\omega)$ is constant and equal to $\widetilde{\Phi}_{no}(0)$ over the frequency range $-W_F$ to $+W_F$ cps Eq. (3-20) can be simplified so that the identification time becomes

$$T_I = \frac{\widetilde{\Phi}_{no}(0)}{\Phi_x^2\, \sigma_{\hat{g}}^2} \qquad (3\text{-}21)$$

The above condition on $\widetilde{\Phi}_{no}(\omega)$ is nearly always satisfied in practical cases if a wideband periodic test signal and optimum averaging filter are used. (See discussion in Section 6.5.)

A comparison of the identification times required by white Gaussian and periodic test signals will now be made. From the noise terms of Eq. (3-14) $\widetilde{\Phi}_{no}(0)$ for the Gaussian case is

$$\widetilde{\Phi_{no}(0)} = \int_{-\infty}^{+\infty} \phi_n(\tau)\, \phi_x(\tau)\, d\tau + \Phi_x^2 \int_0^{\infty} [\widetilde{g(\lambda)}]^2\, d\lambda$$

$$+ \Phi_x^2 \int_{-\infty}^{+\infty} \widetilde{g(\tau_i + \tau)}\, \widetilde{g(\tau_i - \tau)}\, d\tau \qquad (3\text{-}22)$$

An upper bound on the third integral may be obtained by the use of the Schwarz inequality [6, Courant, p. 131].

$$\int_{-\infty}^{+\infty} \widetilde{g(\tau_i + \tau)}\, \widetilde{g(\tau_i - \tau)}\, d\tau$$

$$\leq \sqrt{\int_{-\infty}^{+\infty} [\widetilde{g(\tau_i + \tau_1)}]^2\, d\tau_1 \int_{-\infty}^{+\infty} [\widetilde{g(\tau_i - \tau_2)}]^2\, d\tau_2} \qquad (3\text{-}23)$$

and assuming

$$\int_0^{\infty} [\widetilde{g(\lambda)}]^2\, d\lambda$$

is finite it is easy to see that $\widetilde{\Phi_{no}(0)}$ is bounded by

$$\widetilde{\Phi_{no}(0)} \leq \int_{-\infty}^{+\infty} \phi_n(\tau)\, \phi_x(\tau)\, d\tau + 2\Phi_x^2 \int_0^{\infty} [\widetilde{g(\lambda)}]^2\, d\lambda \qquad (3\text{-}24)$$

It is convenient at this point to introduce the concept of the signal-to-noise ratio appearing at the output of $g(\lambda)$. As discussed in Chapter 1 this is an especially important concept in adaptive system applications as it allows the mean square value of the test signal to be specified in terms of its effect at the output of the system being identified. Since the mean square value of the output signal, $\overline{w^2}$, depends upon the transmission characteristics of $g(\lambda)$ as well as the test signal properties it is necessary to consider ensemble averages once again. In terms of the average equivalent noise bandwidth of the system, $\widetilde{W}_G$, and the average low frequency power gain, $\widetilde{K}_G$, the mean square value of w(t) is

$$\overline{w^2} = \frac{\overline{x^2}\,\widetilde{W}_G\,\widetilde{K}_G}{W_x} \tag{3-25}$$

where $\overline{x^2}$ is the mean square value, and W_x is the equivalent noise bandwidth of the test signal.

The output signal-to-noise ratio, γ, is defined as the ratio of $\overline{w^2}$ to the effective mean square value of the external noise, $\overline{n^2}_{eff}$. $\overline{n^2}_{eff}$ is, in turn, defined as

$$\overline{n^2}_{eff} = \frac{1}{2\pi}\int_{-2\pi W_x}^{+2\pi W_x} \Phi_n(\omega)\, d\omega \tag{3-26}$$

This definition is prompted by the fact that any practical test signal will have a finite bandwidth; hence the system's output signal will be bandlimited, and the signal component of y(t) will not be affected if the observed signal is filtered so as to eliminate all components above the frequency W_x.

When the external noise is white $\overline{n^2}_{eff}$ equals $2\,\Phi_n(0)\,W_x$; and when the external noise bandwidth is small compared to the test signal bandwidth $\overline{n^2}_{eff}$ equals $\overline{n^2}$, the mean square value of the external disturbance. For either of these conditions and a Gaussian test signal the identification time becomes

$$T_I \leq \frac{2\ \widetilde{W}_G\ \widetilde{K}_G}{\sigma^2_{\hat{g}}\ \gamma}\ (1 + 2\gamma) \qquad (3\text{-}27)$$

This expression was obtained by substituting Eq. (3-24) into Eq. (3-21), introducing the output signal-to-noise ratio constraint, and using the fact that Φ_x can be expressed as $\overline{x^2}/2W_x$.

When the test signal is periodic and an ideal finite-memory integrator is used as an averaging filter the only noise term remaining is due to the external disturbance. For the two special cases cited above the zero frequency power spectral density of this term is

$$\Phi_{no}(0) = \frac{\overline{n^2}_{eff}\ \overline{x^2}}{2\ W_x} \qquad (3\text{-}28)$$

and the corresponding identification time is

$$T_I = \frac{\overline{n^2}_{eff}\ 2\ W_x}{\sigma^2_{\hat{g}}\ \overline{x^2}} = \frac{2\ \widetilde{W}_G\ \widetilde{K}_G}{\sigma^2_{\hat{g}}\ \gamma} \qquad (3\text{-}29)$$

A comparison of Eq. (3-29) and Eq. (3-27) shows that the identification time required by crosscorrelation techniques using periodic test signals can be as small as $1/(1 + 2\gamma)$ times the required identification time using an arbitrary random test signal. Even if γ is only of the order of unity, as it probably would be in most adaptive control applications, there is a threefold saving in identification time. For applications which allow γ to be larger the saving is even greater.

3.6 Equivalence of Crosscorrelation Identification and the Ideal Identifier

It is extremely interesting and important to note that when the external disturbance, n(t), is white the identification time required by crosscorrelation using a periodic test signal and an optimum averaging filter is equal to the identification time required by an ideal identifier. This can be shown by first recalling that when the noise is white $\overline{n^2}_{eff}$ is equal to $\Phi_n(0) 2 W_x$ and Eq. (3-29) can be put into the form

$$T_I = \frac{\Phi_n(0) \quad (2 \quad W_x)^2}{\overline{x^2} \quad \sigma_{\hat{g}}^2} \tag{3-30}$$

All that remains to establish the equivalence is to relate the test signal bandwidth to the impulse response sampling interval, and this is easily done by means of the sampling theorem [21, Shannon]. In order to be able to independently specify individual sample points, spaced at intervals of $\Delta\lambda$ seconds, it is necessary that the observed signal contain frequency components at least as great as $1/2\Delta\lambda$. Thus the minimum allowable test signal bandwidth is $1/2\Delta\lambda$, and with this substitution Eq. (3-30) equals the identification time of the ideal identifier as given in Chapter 2, Eq. (2-28).

CHAPTER 4

SYSTEM IDENTIFICATION USING SAMPLING TECHNIQUES

This chapter deals with the identification technique which estimates the impulse response of the unknown system from data obtained by sampling the system input and output signals. Least squares estimates of the impulse response sample points are obtained. The identification time is derived for the case of a white external noise and an optimum test signal. The sampling method of system identification is shown to have an identification time which is equal to that of an ideal identifier.

4.1 Introduction

The computation of estimates of points of an unknown impulse response function from data obtained by sampling the input and output signals of the unknown system will now be considered. The continuous signals x(t), w(t), and y(t) are sampled every t_a seconds and the values of these signals at the instants mt_a, where m is an integer, will be denoted by x(m), w(m), and y(m) respectively. The sequence of random variables n(m), obtained by sampling the zero mean stationary random process n(t), are described statistically by the covariance matrix Σ_n whose i, j element is given by

$$\sigma_{ij}^2 = E[n(i)\ n(j)] \tag{4-1}$$

The sampling procedure and notation is summarized schematically in Fig. 4-1.

The sampled data representation of the unknown systems impulse response will be denoted by g(p). In order to be able to reproduce the fine structure of $g(\lambda)$ it is necessary that the sampling rate be large

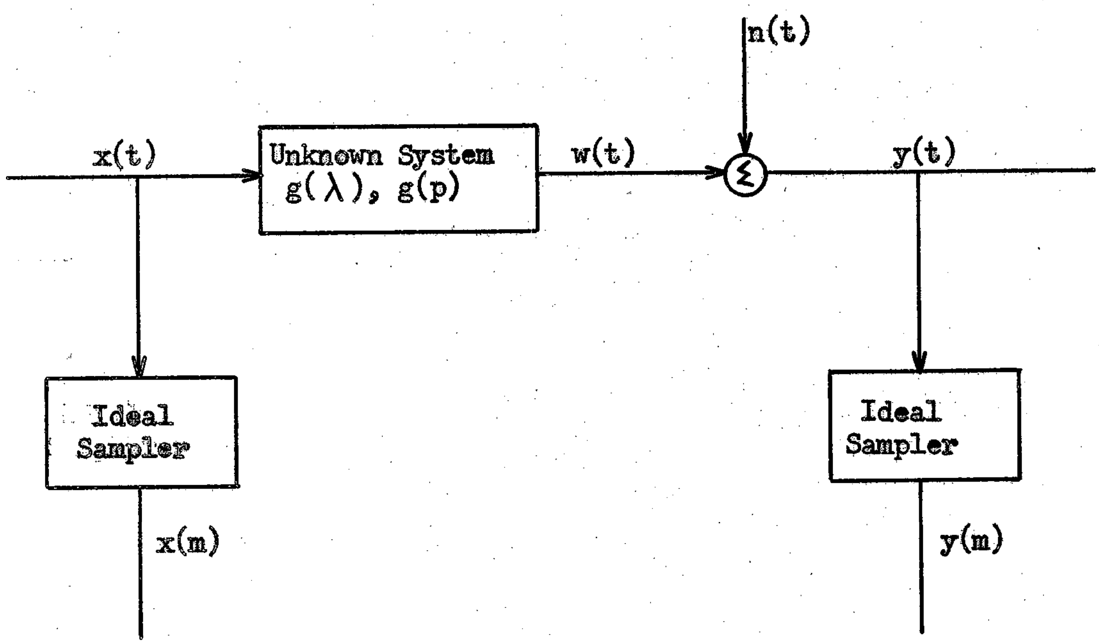


Fig. 4-1

Identification by Sampling Techniques

compared to the systems bandwidth. Some error in the estimation of $g(\lambda)$ will, of course, be introduced by the sampling procedure, but by using a sufficiently large sampling rate this error can be made small compared to the errors introduced by the external disturbance. The fact that impulse responses of practical systems approach zero as λ approaches infinity permits $g(p)$ to be represented, to a good approximation, by the vector

$$(g) = \begin{bmatrix} g(0) \\ g(1) \\ \circ \\ \circ \\ \circ \\ g(P) \end{bmatrix} \tag{4-2}$$

where P is chosen so that $P \cdot t_a$ is greater than the significant duration of $g(\lambda)$.

An estimate of (g) is computed from data obtained by recording a sequence of M samples of the signal $y(t)$ and a corresponding sequence of samples derived from the test signal $x(t)$. The delay caused by the system being identified makes it necessary to record M + P sample values of $x(t)$. The number of data points, M, required to achieve the specified variances of the $g(p)$ estimates will determine the identification time. It is convenient to introduce the following notation

$$(w) = \begin{bmatrix} w(P) \\ w(P+1) \\ \circ \\ \circ \\ \circ \\ w(P+M) \end{bmatrix} \tag{4-3a}$$

$$(n) = \begin{bmatrix} n(P) \\ n(P+1) \\ \circ \\ \circ \\ \circ \\ n(P+M) \end{bmatrix} \tag{4-3b}$$

$$(y) = (w) + (n) = \begin{bmatrix} y(P) \\ y(P+1) \\ \cdot \\ \cdot \\ \cdot \\ y(P+M) \end{bmatrix} \tag{4-3c}$$

and the matrix

$$(x) = \begin{bmatrix} x(P) & x(P-1) & \cdots & x(0) \\ x(P+1) & x(P) & \cdots & x(1) \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ x(P+M) & x(P+M-1) & \cdots & x(M) \end{bmatrix} \tag{4-3d}$$

With this notation the sample data version of the convolution integral

$$w(m) = \sum_{p=0}^{P} x(m-p) \; g(p) \; t_a \; ; m=P, P+1, \ldots, P+M \tag{4-4}$$

can be replaced by the matrix equation

$$(w) = t_a \; (x) \; (g) \tag{4-5}$$

The external disturbance prohibits the direct measurement of (w) and it is necessary to estimate (g) from the equation

$$(y) = t_a \; (x) \; (g) \tag{4-6}$$

4.2 Least Squares Estimates of the Impulse Response

A criterion for chosing estimates of the g(p) that is mathematically reasonable, and, in addition, leads to expressions which, from a computational standpoint, are easy to implement is to choose those values of g(p) which minimize the sum of squared deviations

$$\sum_{m=0}^{M} \left[y(m) - \sum_{p=0}^{P} x(m-p) \; g(p) \; t_a\right]^2 = \left[(w) - t_a(x)\;(g)\right]' \; \left[(w)-t_a(x)(g)\right] \tag{4-7}$$

Levin [16] shows that these estimates are given by the set of so called normal equations

$$t_a (x)'(x)\ (\hat{g}) = (x)'(y) \qquad (4\text{-}8)$$

where $(x)'$ is the transpose of the matrix (x) and $(\hat{g})$ is the estimate of (g).

The set of normal equations can be put into a more familiar form by defining empirical correlation functions

$$\tilde{\phi}_x(r) = \frac{1}{M+1} \sum_{m=0}^{M} x(m)\ x(m+r) \qquad (4\text{-}9a)$$

and

$$\tilde{\phi}_{xy}(r) = \frac{1}{M+1} \sum_{m=0}^{M} x(m)\ y(m+r) \qquad (4\text{-}9b)$$

Then the $\hat{g}(p)$ can be obtained from the set of linear simultaneous equations

$$\sum_{p=0}^{P} \tilde{\phi}_x(p-i)\ \hat{g}(p)\ t_a = \tilde{\phi}_{xy}(i) \qquad i = 0,1,2,\ldots,P \qquad (4\text{-}10)$$

Eq. (4-10) appears to be similar to the sample data Wiener-Hopf equation, but in this case the quantities $\tilde{\phi}_x(r)$ and $\tilde{\phi}_{xy}(r)$ are not correlation functions in the usual sense, $\tilde{\phi}_x(r)$ and $\tilde{\phi}_{xy}(r)$ are empirical correlation functions calculated from the finite sequences of $x(m)$ and $y(m)$. If $x(t)$ is from a stationary random process then

$$\lim_{M\to\infty} \tilde{\phi}_x(r) = \phi_x(r) \qquad (4\text{-}11a)$$

and

$$\lim_{M \to \infty} \tilde{\phi}_{xy}(r) = \phi_{xy}(r) \tag{4-11b}$$

where $\phi_x(r)$ and $\phi_{xy}(r)$ are obtained by sampling $\phi_x(\tau)$ and $\phi_{xy}(\tau)$, the correlation functions defined in the usual manner [14, Laning and Battin, p. 113].

The covariance matrix of $(\hat{g})$ which by definition is equal to

$$\not{\xi}_{\hat{g}} = E[(\hat{g})\ (\hat{g})'] - E[(\hat{g})]\ E[(\hat{g})'] \tag{4-12}$$

can be shown to be equal to

$$\not{\xi}_{\hat{g}} = [(x)'(x)]^{-1}\ (x)'\ \not{\xi}_n\ (x)\ [(x)'(x)]^{-1} \frac{1}{t_a^2} \tag{4-13}$$

where $[(x)'(x)]^{-1}$ is the inverse of the matrix $[(x)'(x)]$ and $\not{\xi}_n$ is the covariance matrix of the external noise. When the noise samples n(m) are uncorrelated

$$\not{\xi}_n = \overline{n^2}_{eff}(I) \tag{4-14}$$

and

$$\not{\xi}_{\hat{g}} = \overline{n^2}_{eff}\ [(x)'(x)]^{-1} \frac{1}{t_a^2} \tag{4-15}$$

4.3 Optimum Test Signals

Up to this point nothing has been said about the test signal to be used for the identification of $g(\lambda)$, and indeed whether or not a solution to the set of equations Eq. (4-8) or Eq. (4-10) even exists. Premultiplying both sides of Eq. (4-8) by $[(x)'(x)]^{-1}$ gives

$$t_a(\hat{g}) = [(x)'(x)]^{-1}\ (x)'(y) \tag{4-16}$$

From this expression it is clear that in order for a unique solution to these equations to exist $(x)'(x)$ must not be singular, i.e., $[(x)'(x)]^{-1}$ must exist. Since $(x)'(x)$ is proportional to a correlation matrix it is positive definite [7, Cramer, p. 295] and its inverse always exists [7, Cramer, p. 115]. In order that the solution be non-trivial the column matrix $(x)'(y)$ must not be identically zero [11, Guillemin, p. 18 and 105]. To insure this condition it is necessary and sufficient that (y) not be identically zero, and that the row vectors of (x) be linearly independent.

In Section 4.1 it was pointed out that the sampling rate must be large compared to the bandwidth of the system under test. This condition is necessary if the fine structure of $g(\lambda)$ is to be adequately reproduced, but it is not sufficient. In addition, the test signal must contain components at these higher frequencies. The most logical bandwidth to choose for the test signal is W_x equal to $1/2t_a$. Because of the sampling process, components in $x(t)$ at frequencies higher than $1/2t_a$ will not contribute to the estimate of $g(\lambda)$, and a test signal bandwidth less than $1/2t_a$ will not make full use of the possibilities offered by the sampling rate $1/t_a$. In what follows W_x will be assumed to be equal to $1/2t_a$.

The expression for the covariance matrix of $(\hat{g})$, Eq. (4-13), shows that the mean square errors in the measurement of $(\hat{g})$ can be reduced by increasing the amplitude of the test signal. In physical applications the amplitude of $x(t)$ will be limited by practical considerations. If the mean square value of $w(t)$, the system's response to the test signal, is to be equal to γ times the effective mean square value of the external noise, as defined by Eq. (3-26), the following condition must be met

$$\overline{x^2}\, t_a\, 2\, \widetilde{\widetilde{W}}_G\, \widetilde{\widetilde{K}}_G = \gamma\, \overline{n^2}_{eff} \qquad (4\text{-}17)$$

where, as in the previous chapter, $\widetilde{\widetilde{W}}_G$ is the average equivalent noise bandwidth, and $\widetilde{\widetilde{K}}_G$ is the average low frequency power gain of the system being identified. $2t_a$ is recognized as being equal to the reciprocal of the test signal bandwidth so that Eq. (4-17) is equivalent to the output signal-to-noise ratio constraint imposed in the preceding chapter. The constraint upon the test signal can be expressed as a constraint upon the mean square value of x(t),

$$\overline{x^2} = \frac{\gamma\, \overline{n^2}_{eff}}{2\, t_a\, \widetilde{\widetilde{W}}_G\, \widetilde{\widetilde{K}}_G} \qquad (4\text{-}18)$$

The conditions that the optimum test signal must satisfy can be determined by expressing the covariance matrix as

$$\not{\Sigma}_{\hat{g}} = \Big[[(x)'(x)]\ [(x)'\not{\Sigma}_n\,(x)]^{-1}\ [(x)'(x)]\Big]^{-1} \frac{1}{t_a{}^2} \qquad (4\text{-}19)$$

Since $[(x)'(x)]\ [(x)'\not{\Sigma}_n\,(x)]^{-1}\ [(x)'(x)]$ is symmetric and positive definite it follows from the discussion in Section 2.4 that the variances associated with the $\hat{g}(p)$ will be minimized if

$$[(x)'(x)]\ [(x)'\not{\Sigma}_n(x)]^{-1}\ [(x)'(x)] = k(I) \qquad (4\text{-}20)$$

where k is a constant that is determined by the constraint Eq. (4-18). The important property of the optimum test signal is that it is a signal which, when acting in combination with the external noise, causes the estimates of the impulse response to be uncorrelated with one another. This property is expressed mathematically in Eq. (4-20).

When the external disturbance is white Eq. (4-20) is simplified considerably, and becomes

$$\overline{n^2_{eff}} \quad [(x)'(x)] \quad = k(I) \tag{4-21}$$

The i, j element of $(x)'(x)$ is equal to

$$(M+1) \quad \tilde{\phi}_x(i-j) \tag{4-22}$$

so, for the case of white noise, the optimum test signal must satisfy the conditions

$$\tilde{\phi}_x(0) = \overline{x^2} \neq 0 \tag{4-23a}$$

$$\tilde{\phi}_x(r) = 0 \qquad 0 < r \leq P \tag{4-23b}$$

that is, the x(m) values must be white over a range of P samples.

If x(m) is a stationary random process with zero mean, then a reasonable choice would be to choose x(m) so that

$$E\,[\tilde{\phi}_x(0)] = \phi_x(0) \neq 0 \tag{4-24a}$$

$$E\,[\tilde{\phi}_x(r)] = \phi_x(r) = 0 \qquad 0 < r \leq P \tag{4-24b}$$

Additional errors are introduced in the estimates of g(p) when x(m) is an arbitrary random process because even if $\phi_x(r)$ satisfies the optimum conditions the empirical correlation function, $\tilde{\phi}_x(r)$, obtained for a particular observation of M sample points, may not be optimum and hence the variances of $\hat{g}(p)$ will not be minimized. It is, therefore, concluded that a periodic test signal, chosen in such a manner so as to satisfy the optimum conditions on $\tilde{\phi}_x(r)$, is better than an arbitrary test signal whose empirical autocorrelation functions is only optimum on the average.

Some of the practical advantages of a periodic test signal have been mentioned above (Section 3.4). A periodic test signal can also be used in the sampling technique for system identification provided the period of the test signal is large compared to the significant length of the impulse response. The synthesis of an optimum test signal is greatly simplified if the test signal is periodic and the number of sample points observed, M, is chosen so that the identification time is an integral number of periods of the test signal. When the external noise is white, periodic discrete interval binary noise can be used as a test signal. Another class of zero-correlation codes, useful when the noise is white, has been discussed by Tompkins [24].

4.4 Identification Time Using Sampling Techniques

Computation of the identification time will be illustrated for the case when the external noise is white and the test signal is optimized. The insight gained from the results of this simple case will then be used to establish, in a heuristic manner, that a similar result holds in general when the test signal is optimized.

When the noise is white and the test signal is optimized all of the off-diagonal terms of the correlation matrix, Eq. (4-15), are zero. In addition, the variances associated with the estimation of each g(p), p = 0, 1, 2 ... P, are equal. Substituting from expression (4-22) and Eq. (4-23) into Eq. (4-15) gives

$$\sigma_{\hat{g}}^{2} = \frac{\overline{n_{eff}^{2}}}{t_a^{2}\,(M+1)\,\overline{x^{2}}} \qquad (4\text{-}25)$$

where $\sigma_{\hat{g}}^2$ is the variance associated with the estimate of a single sample point, g(p), p = 0, 1, 2 ... P. The corresponding identification time is

$$T_I = (M + 1) t_a = \frac{\overline{n^2}_{eff}}{t_a \overline{x^2} \sigma_{\hat{g}}^2} \qquad (4\text{-}26)$$

Using the constraint Eq. (4-18) the identification time may also be expressed as

$$T_I = \frac{2 \tilde{W}_G \tilde{K}_G}{\gamma \sigma_{\hat{g}}^2} \qquad (4\text{-}27)$$

In arriving at these expressions for T_I the computation time required to solve the set of equations, Eq. (4-8) or Eq. (4-10), has been neglected. Thus, in this respect, T_I as given by Eq. (4-26) or Eq. (4-27) is a lower bound.

In the general case optimization of the test signal has the effect of putting the covariance matrix in the form

$$\Sigma_{\hat{g}} = \sigma_{\hat{g}}^2 (I) \qquad (4\text{-}28)$$

that is, the estimates of sample points of $g(\lambda)$ are uncorrelated with each other, and the variances of each estimate are equal. A similar result is obtained for the special case of white external noise. Since, in each case, signal optimization leads to estimates of $g(\lambda)$ which are uncorrelated it is reasonable to expect that the form of Eq. (4-25) and Eq. (4-26) is valid in the general case also.

Very often the exact nature of the external noise correlation function is not known. If this is the case, it is convenient to pick a test signal satisfying the condition of Eq. (4-23) because then only the $p = i$ term in Eq. (4-10) is important, and the computation of the impulse response estimates is greatly simplified. Choosing a test signal in this manner is equivalent to using an optimum test signal under the assumption that the external disturbance is white.

4.5 Equivalence to the Ideal Identifier

The expression for the identification time using sampling techniques, Eq. (4-26) or Eq. (4-27) is equivalent to the identification time expression obtained for crosscorrelation identification. It was shown previously that crosscorrelation identification time is equal to the time required by an ideal identifier to estimate the impulse response. Hence it is concluded that identification by sampling techniques is also equivalent to an ideal identifier.

CHAPTER 5

IDENTIFICATION TIME REQUIRED BY MATCHED FILTER IDENTIFICATION

A third identification technique is described and analyzed in this chapter. The variance in the impulse response estimate resulting from matched filter identification is determined, and a means of reducing the variance by the use of a periodic test signal and a comb filter is developed. The matched filter identification time is determined, and upon comparison with the results of Chapter 2, it is established that matched filter identification is also ideal. The chapter ends with a brief discussion on the synthesis of suitable test signals, matched filters, and comb filters.

5.1 Description of the Matched Filter Identification Technique

The output signal resulting from the application of a unit impulse to an unknown system would be the impulse response of the system. A distinct advantage of this identification scheme would be that the resulting output signal is a continuous representation of the impulse response, not just sample points of this function. An adaptive system using this type of identification scheme has been discussed by Aseltine, et al. [2]. From a practical standpoint this apparently simple method of identification has serious drawbacks; some systems that can be represented adequately by a linear model under normal operating conditions may exhibit non-linear characteristics for large input signals, such as impulses. Also, in situations where it is necessary to make the identification in the presence of normal operating signals, such as in adaptive control applications, large amplitude test signals may not be tolerable.

A technique, suggested by Turin [25], for obtaining a continuous real time estimation of an unknown impulse response that does not require the use of impulse type test signals is illustrated in Fig. 5-1. The input test signal is deterministic in nature, and for the moment, will be assumed to be zero outside the interval $0 \leq t \leq T_x$. This signal will be denoted by $x_1(t)$ to distinguish it from a test signal which exists for all values of t. The observed signal, y(t), which is equal to the system output signal, w(t), plus a zero-mean stationary noise signal, n(t), is passed through an estimating filter, $h(\lambda)$. The extimating filter is designed so that its output signal is an estimate of the unknown system.

It is established below that if $h(\lambda)$ is proportional to $x_1(\Delta - \lambda)$, and $x_1(t)$ has a bandwidth which is wide compared to the bandwidth of the system being identified, then the signal component at the output of the estimating filter is proportional to $g(t - \Delta)$. Such a filter is called a matched filter [26, Turin]. The delay, Δ, must be greater than or equal to T_x in order to guarantee the physical realizability of the estimating filter.

Referring to Fig. 5-1 it can be seen that

$$w(\lambda_2) = \int_{-\infty}^{+\infty} x_1(\lambda_2 - \lambda_1)\, g(\lambda_1)\, d\lambda_1 \qquad (5\text{-}1)$$

and $\hat{g}_1(t)$, the output of the estimating filter due to the test signal $x_1(t)$, is

$$\hat{g}_1(t) = \int_{-\infty}^{+\infty} y(t - \lambda_2)\, h(\lambda_2)\, d\lambda_2 \qquad (5\text{-}2)$$

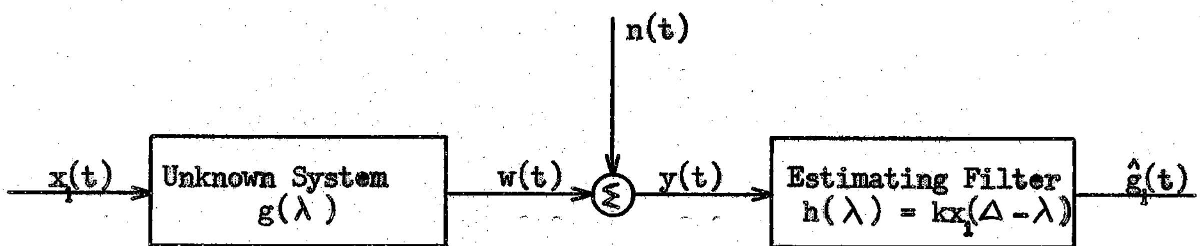


Fig. 5-1

Matched Filter Identification

By combining Eq. (5-1) and Eq. (5-2) and using the fact that

$$y(t) = w(t) + n(t) \tag{5-3}$$

and

$$h(\lambda) = k\, x_1(\Delta - \lambda) \tag{5-4}$$

the output of the estimating filter can be expressed as

$$\hat{g}_1(t) = k \int_{-\infty}^{+\infty} g(\lambda_1)\, d\lambda_1 \left\{ \int_{-\infty}^{+\infty} x_1(t - \lambda_2 - \lambda_1)\, x_1(\Delta - \lambda_2)\, d\lambda_2 \right\}$$

$$+ k \int_{-\infty}^{+\infty} n(t - \lambda_2)\, x_1(\Delta - \lambda_2)\, d\lambda_2 \tag{5-5}$$

In this expression k is an arbitrary constant of proportionality.

Now, for a moment, focus attention on the integral within the curly brackets of Eq. (5-5) and introduce the change of variable, $\lambda_3 = t - \lambda_2 - \lambda_1$. This integral then becomes

$$\int_{-\infty}^{+\infty} x_1(\lambda_3)\, x_1(\lambda_3 + \Delta + \lambda_1 - t)\, d\lambda_3 \tag{5-6}$$

Using Parseval's Theorem expression (5-6) can also be put into the form

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} X_1^*(\omega)\, X_1(\omega)\, e^{j(\lambda_1 - t + \Delta)}\, d\omega \tag{5-7}$$

where $X_1(\omega)$ is the Fourier transform of $x_1(t)$ and the asterisk denotes the complex conjugate. If $x_1(t)$ is a wide band signal and $X_1(\omega)$ is nearly constant and equal to $X_1(0)$ expression (5-7) becomes

$$\frac{1}{2\pi} \quad X_1^{\,2}(0) \int_{-\infty}^{+\infty} e^{j(\lambda_1 - t + \Delta)} \; d\omega \tag{5-8}$$

which, from the properties of Fourier transforms, is recognized as being equal to

$$X_1^{\,2}(0) \;\; \delta(\lambda_1 - t + \Delta) \tag{5-9}$$

The same result can be obtained by noting that expression (5-6) is the autotranslation function [20, Newton, Gould, and Kaiser, p. 51] of the test signal, and recalling that the autotranslation function of a wide-band signal is approximately equal to a delta function. The derivation is similar to the one given in Section 3.1 of the chapter on cross-correlation identification.

Replacing the term in curly brackets in Eq. (5-5) by its equivalent, expression (5-9), and integrating with respect to λ_1 yields

$$\hat{g}_1(t) = k\, X_1^{\,2}(0)\; g(t - \Delta) \;\; + k \int_{-\infty}^{+\infty} n(t - \lambda_2)\; x_1(\Delta - \lambda_2) \;\; d\lambda_2 \tag{5-10}$$

The first term in Eq. (5-10) is the signal component of $\hat{g}_1(t)$, and as was indicated previously it is proportional to $g(t - \Delta)$. For convenience throughout the rest of the chapter k will be set equal to $1/X_1^{\,2}(0)$.

5.2 Variance of the Impulse Response Estimate

The crucial assumption that was made in deriving Eq. (5-10) was that the test signal spectrum is white. In reality signals with a perfectly flat spectrum cannot be generated, and this fact leads to an error in the signal term of $\hat{g}_1(t)$. This error is, however, deterministic in nature and with a knowledge of the exact test signal spectrum it can be compensated for. Furthermore, since there is usually a considerable amount of freedom in the choice of $x_1(t)$, by selecting a test signal with a wide spectrum the effects of this error can be minimized. Primary consideration will be given to the effects of the random noise errors in $\hat{g}_1(t)$ by assuming that deterministic errors, due to a finite bandwidth test signal, are small compared to the random errors.

The second term in Eq. (5-10) represents the error in the estimate of $g(\lambda)$ due to the external noise, $n(t)$. The contribution of this term to the variance of $\hat{g}_1(\lambda)$ can be computed from a knowledge of the noise power spectral density and the frequency characteristics of the estimating filter. Thus,

$$\sigma_{\hat{g}_1}^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi_n(\omega) \left|H(\omega)\right|^2 d\omega$$

$$= \frac{1}{2\pi X_1^4(0)} \int_{-\infty}^{+\infty} \Phi_n(\omega) \left|X_1(\omega)\right|^2 d\omega \qquad (5\text{-}11)$$

The variance can be put into a form that is more nearly related to the expression of previous chapters by defining $\overline{n_o^2}$, the mean square value of the noise at the output of a unity-gain matched filter.

$$\overline{n_o^2} = \frac{1}{2\pi X_1^2(0)} \int_{-\infty}^{+\infty} \Phi_n(\omega) \left|X_1(\omega)\right|^2 d\omega \qquad (5\text{-}12)$$

With this new notation $\sigma_{\hat{g}_1}^2$ becomes

$$\sigma_{\hat{g}_1}^2 = \frac{\overline{n_o^2}}{X_1^2(0)} \tag{5-13}$$

There are two special cases of interest. First, if the test signal is wide band with respect to n(t), as well as being wide band with respect to $g(\lambda)$, $X_1(\omega)$ in Eq. (5-12) is constant and

$$\overline{n_o^2} = \overline{n^2} \tag{5-14}$$

In this case the estimating filter does a good job in producing a signal term in $\hat{g}_1(t)$ that is proportional to the unknown impulse response, but it is of no value in reducing the effects of the external noise, because all of the frequency components of the external noise are transmitted through $h(\lambda)$ without distortion. The gain factor $1/X_1^2(0)$ has no effect upon the signal-to-noise ratio because both the signal and noise components of y(t) are multiplied by the same gain. For this situation the signal-to-noise ratio at the output of the estimating filter is equal to the signal-to-noise ratio at the input. This ratio, γ, is determined by the restriction placed on the test signal.

The second important special case occurs when the bandwidth of the external noise is wide compared to the bandwidth of the test signal. Of course, $x_1(t)$ is still wide band compared to $g(\lambda)$. Then $\Phi_n(\omega)$ in Eq. (5-13) is approximately constant and equal to $\Phi_n(0)$, and the normalized output noise is

$$\overline{n_o^2} = \Phi_n(0)\, 2W_x \tag{5-15}$$

where W_x is the equivalent noise bandwidth of the test signal. In this type of situation the estimating filter reduces the effective bandwidth of the random noise as well as producing a signal term proportional to $g(\lambda)$. Now only those noise frequency components within the pass-band of the estimating filter appear in $\hat{g}_1(t)$, and since $h(\lambda) = k\, x(\Delta - \lambda)$, it is easy to see that the effective bandwidth of the random noise term in $\hat{g}_1(t)$ is equal to W_x.

The reader may ask if it is not possible to design the estimating filter to reduce the total error in $\hat{g}_1(t)$, the "smear" error, due to the fact that $x_1(t)$ does not have an infinite bandwidth, as well as the error due to $n(t)$. This problem has been considered in detail by Turin [25]. In his paper Turin shows that when the test signal is optimized and the external noise is white the optimum estimating filter is a matched filter. When the noise is not white the expression for the optimum filter is somewhat more complicated. The present work is concerned with finding limits upon the identification time due to external noise, and, therefore has not considered in detail the errors introduced by the practical limitations of the test signal. The optimum estimating filter for the problem considered here is a matched filter regardless of the shape of the noise spectrum.

5.3 Reduction of Variance by Periodic Excitation

The identification time required to obtain the impulse response estimate variance given by Eq. (5-13), is equal to Δ, the delay necessary to make $h(\lambda)$ physically realizable. Δ is of the order of magnitude of T_x, the duration of the test signal. Often the variance given by this equation will be much larger than can be tolerated. This will be true especially if the signal-to-noise ratio at the output of the unknown system is required to be small. In order to achieve a smaller variance it becomes necessary to average over several observations of the impulse response estimate. If an average is taken over M observations the general expression for the variance becomes

$$\sigma_{\hat{g}}^2 = \frac{\overline{n_o^2}}{M \; X_1^2(0)} \tag{5-16}$$

Note the subscript on $\hat{g}$ has been dropped to indicate that the variance no longer corresponds to a single observation.

One of the nice features of the matched filter identification technique is that a continuous real time estimate of $g(\lambda)$ is presented at the output of the matched filter. A continuous train of impulse response estimates can be generated at the output of the estimating filter by using a suitable periodic test signal, x(t), where

$$x(t) = \sum_{i=-\infty}^{+\infty} x_1(t - i\,T_x) \tag{5-17}$$

Then averaging over M observations is simply equivalent to averaging over M periods of the output signal. It is shown below that one method of mechanizing the averaging operation is to place a comb filter in cascade with the estimating filter.

An arbitrary periodic test signal will not give a satisfactory estimate of $g(\lambda)$. x(t) must be rich in harmonic content and in addition the period T_x must be larger than the significant length of $g(\lambda)$. The generation of a train of impulse response estimates can be visualized with the aid of Fig. 5-2. Since x(t) is periodic the autocorrelation function of x(t) will also be periodic with the properties

$$\phi_x(\tau) = \sum_{i=-\infty}^{+\infty} \phi_{x_1}(\tau - iT_x) \qquad (5\text{-}18a)$$

and

$$\phi_{x_1}(\tau) = 0 \qquad |\tau| > \frac{T_x}{2} \qquad (5\text{-}18b)$$

As t increases the pulse labeled $\phi_{x_1}(\lambda - t)$ in Fig. 5-2 sweeps over the impulse response $g(\lambda)$. The output signal is proportional to the integral of the product of $\phi_x(t - \lambda)$ and $g(\lambda)$ so that if $\phi_{x_1}(\tau)$ is narrow compared to the significant length of $g(\lambda)$ and T_x is large enough so that only a single pulse of $\phi_x(\tau)$ overlaps $g(\lambda)$ at any single instant of time the output signal is a good estimate of $g(\lambda)$. As time progresses the labeled pulse passes beyond the overlap region, and the next pulse in the train begins to contribute to the output. The mechanization of the matched filter identification scheme using a periodic test signal is identical to that used when $x(t)$ is aperiodic. The block diagram is shown in Fig. 5-1.

Averaging of the impulse response estimate over several periods of $x(t)$ could be accomplished by replacing the estimating filter that is matched to a single period of $x(t)$ by a filter that is matched to M periods of the test signal, i.e. use a new estimating filter $h'(\lambda)$ with an impulse response

$$\begin{aligned} h'(\lambda) &= kf(\lambda)x(-\lambda) \qquad (5\text{-}19) \\ &= kf(\lambda) \sum_{i=-\infty}^{+\infty} x_1(-\lambda + iT_x) \end{aligned}$$

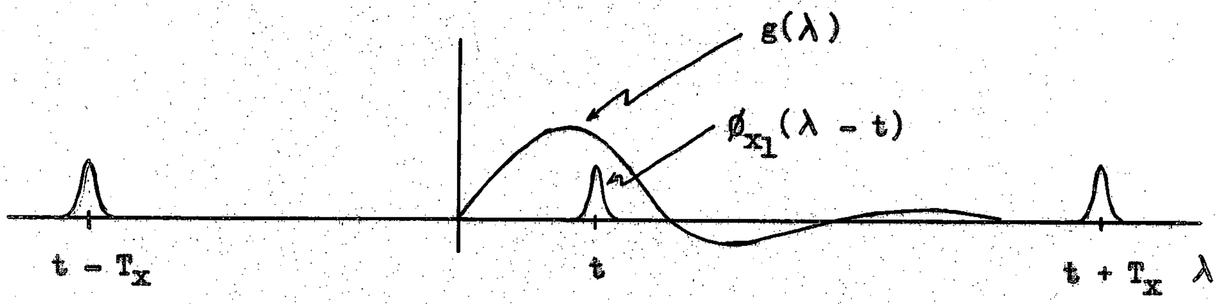


Fig. 5-2

Generation of Output Signal when x(t) is Periodic where

$$f(\lambda) = \begin{cases} 1 & 0 \le \lambda \le MT_x \\ 0 & \lambda < 0 \; ; \; \lambda > MT_x \end{cases} \tag{5-20}$$

The construction of such a filter is easier to view in the frequency domain than in the time domain. If a periodic function, with fundamental period T_1 is multiplied by an envelope function, $f_2(t)$, then the product, $f_3(t)$, can be represented in the frequency domain by [29, Reference Data for Radio Engineers, p. 1018]

$$F_3(\omega) = \frac{2\pi}{T_1} F_1(\omega) \sum_{i=-\infty}^{+\infty} F_2\left(\omega - \frac{i\,2\pi}{T_1}\right) \tag{5-21}$$

where $F_1(\omega)$ is the Fourier transform of a single period of the periodic function. Thus, by analogy with Eq. (5-21), it can be seen that $h'(\lambda)$ can be considered as being made up of two filters in cascade. One filter is matched to a single period of the test signal and has a transform $X_1(-\omega)$; the other filter is a comb filter with pass-bands centered around the harmonic frequencies of $x(t)$. The frequency response of the pass-bands are determined by $F(\omega)$. The frequency response of the two filters in cascade is shown in Fig. 5-3.

In the limiting case as M approaches infinity $f(\lambda)$ is equal to unity, $F(\omega)$ becomes a delta function, and the frequency response reduces to a line spectra proportional to the test signal spectra. The addition of a comb filter to Fig. 5-1 is all that is required to account for the process of averaging over M periods. The output of the comb filter, $\hat{g}(t)$, is a train of estimates of $g(\lambda)$, each impulse response estimate being the average over the past M periods of $\hat{g}_1(t)$. Both signal and noise terms pass through the comb filter so the gain has no effect upon the signal-to-noise ratio. Hence, without loss of generality, the gain may be taken as unity.

5.4 Identification Time Required by the Matched Filter Technique

The noise reduction properties of the comb filter become evident by studying Fig. 5-4 and Fig. 5-5. The signal term at the output of $h(\lambda)$ is a periodic display of impulse response estimates, the fundamental period being determined by T_x, the period of the test signal. The frequency domain representation of the signal is, as a result of its periodic nature, a line spectra with components at the harmonic frequencies of the test signal. The spectrum of the external noise is, in general, continuous so that the frequency characteristics of the noise term present at the output of $h(\lambda)$ is also continuous and is given by $\Phi_n(\omega)|H(\omega)|^2$.

The comb filter may be represented ideally by rectangular passbands with unity gain and width 1/2 MT_x cps., the equivalent noise bandwidth of $F(\omega)$. The product of Fig. 5-4 and Fig. 5-5 is the frequency domain representation of the signal and noise terms of $\hat{g}(t)$. The signal term spectra of $\hat{g}(t)$ is identical to the signal spectra at the input of the comb filter if $g(\lambda)$ is not changing, but the noise spectrum is altered considerably. Assuming that $\Phi_n(\omega)|H(\omega)|^2$ is nearly constant over any frequency range $\omega_a < \omega < \omega_a + \pi/MT_x$ where ω_a is arbitrary, the mean square value of the noise at the output of the comb filter is

$$\frac{1}{M\,T_x} \cdot T_x = \frac{1}{M} \tag{5-22}$$

times the mean square value of the noise into the filter. The settling

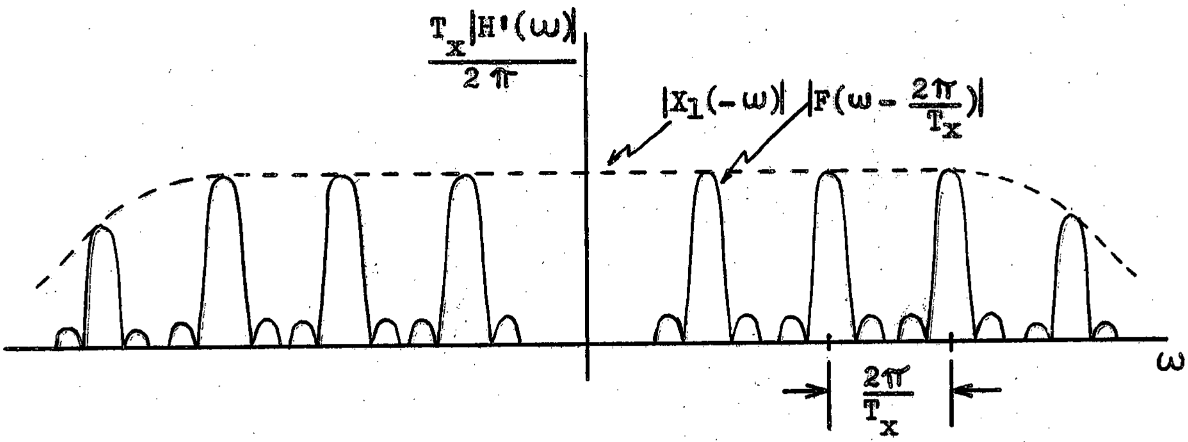


Fig. 5-3

Matched Filter Spectrum for a Periodic Test Signal

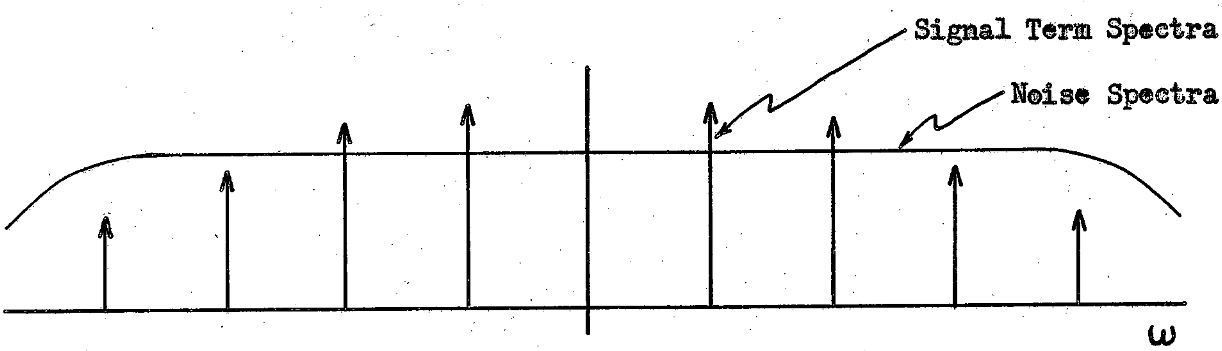


Fig. 5-4

Signal and Noise Frequency Characteristics

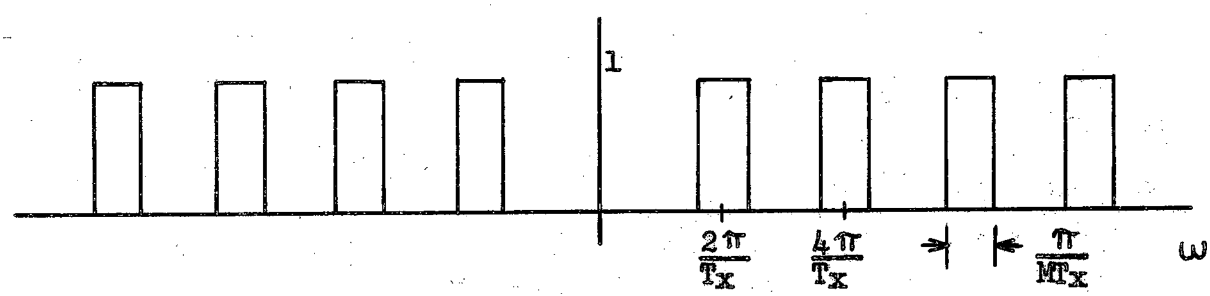


Fig. 5-5

Ideal Comb Filter Frequency Characteristics time associated with the comb filter is equal to one-half times the reciprocal of the bandwidth or MT_x seconds.

The identification time may be expressed in terms of the desired variance and the properties of the test signal and external disturbance by using Eq. (5-16) and relating $X_1^2(0)$ to the mean square value and the equivalent noise bandwidth of x(t). The result is

$$T_I = \frac{\overline{n_o^2} \; 2 \, W_x}{\sigma_{\hat{g}}^2 \; \overline{x^2}} \tag{5-23}$$

With the introduction of the output signal-to-noise ratio constraint T_I can also be written in terms of γ and the average power gain and bandwidth of the system being identified. Thus,

$$T_I = \frac{\overline{n_o^2} \; 2 \, \tilde{W}_G \; \tilde{K}_G}{\sigma_{\hat{g}}^2 \; \gamma \; \overline{n^2}_{eff}} \tag{5-24}$$

When the bandwidth of the external disturbance is smaller than that of the test signal $\overline{n_o^2} = \overline{n^2}_{eff} = \overline{n^2}$ and Eq. (5-24) reduces to

$$T_I = \frac{2 \, \tilde{W}_G \; \tilde{K}_G}{\sigma_{\hat{g}}^2 \; \gamma} \tag{5-25}$$

Similarly, if the noise is white with respect to the test signal

$\overline{n_o^2} = \overline{n^2}_{eff} = \Phi_n(0) \, 2W_x$ and T_I is again given by Eq. (5-25).

5.5 Comparison with the Ideal Identifier

Matched filter identification can also be shown to be equivalent to the ideal identifier. This fact is established by replacing $\overline{n_e^2}$ in Eq. (5-23) with its equivalent expression for the white noise case, $\Phi_n(0)\ 2W_x$, and comparing the resulting expression for the identification time to the results obtained for the ideal identifier in Chapter 2.

5.6 Some Practical Considerations

The subject of synthesizing a test signal, the matched filter, and the comb filter has been carefully avoided up to this point with the result that it has been possible to establish the fact that matched filter identification is theoretically possible, and is indeed equivalent to the ideal identifier when the external disturbance is white. Synthesis of a suitable test signal for identifying $g(\lambda)$ is no more difficult in this case than it is for identification by crosscorrelation or sampling techniques. This topic has been discussed in the previous chapters.

It might appear that the construction of a suitable matched filter would impose the primary practical limitation upon this identification technique, because, in general, the synthesis of a matched filter is difficult. The problem, however, may be partially avoided in this application by building the wide band estimating filter <u>first</u> and then synthesizing the test signal by simply applying an impulse to the filter. The filter's impulse response, reversed in time, would be the corresponding test signal. For use in adaptive systems it would be desirable to chose a filter whose impulse response did not have large peaks, otherwise the test signal might disturb the systems normal operation.

The identification problem does not suggest any short cuts for constructing the comb filter. As is generally the case, the comb filter would be a rather complex device to build, particularly with regard to the amount of hardware that would be required.

CHAPTER 6

EXAMPLES

The previous chapters have developed and presented expressions for the identification time required by an ideal identifier and the identification techniques employing crosscorrelation, sampling, and matched filters. T_I was expressed in terms of the test signal and noise parameters, and also in terms of the unknown system's gain-bandwidth product and the signal-to-noise ratio at the output of the system. Several examples are presented in this chapter in order that some insight may be gained as to the order of magnitude of T_I for practical situations. The results are presented in such a manner that the identification time can be obtained from either a knowledge of the noise and test signal parameters, or from a specification of the gain-bandwidth product and output signal-to-noise ratio. In the absence of any knowledge about the system under test the identification time can only be specified in terms of the systems measurement environment, i.e., in terms of the test signal and noise parameters. On the other hand, there are situations when it is meaningful to discuss such things as the average impulse response, average gain, or average bandwidth. This would be true particularly in adaptive control problems where the design or optimum impulse response may be taken as the average response. The importance of the identification problem in the adaptive control field warrants slanting the discussion of the examples presented here towards that area.

6.1 Introduction

The results for the identification time that have been obtained do not depend upon the particular system that is being identified. For simplicity two second order examples will be considered, one being a nonminimum phase system. While seemingly restrictive, second order systems have impulse responses that are typical of a large class of higher order systems characterized by a single dominant pole pair. Thus, the results of these examples may be used as a guide to what may be expected from higher order systems. When suitably placed, the right-half-plane zero of the nonminimum phase system has the effect of widening the system's bandwidth while retaining the oscillatory properties of the impulse response. The identification time is highly dependent upon the bandwidth so that an interesting comparison of the identification times of the minimum phase and nonminimum phase systems can be made.

White noise introduced at the output of the unknown system is an important example to consider because it represents the worst possible case insofar as the identification time is concerned. If the noise autocorrelation function has a non-zero width, and the nature of the correlation is known, this additional knowledge can be used to reduce the identification time. An example of this is discussed for the special case of a narrowband noise process in Section 6-5. White noise introduced within the feedback loop, as illustrated in Fig. 6-1, is more realistic in some cases than white noise at the output of the system; therefore, examples of this type are also treated. An approximation to this kind of situation occurs when the actuators of a control system are subjected to broadband disturbances, such as wind gusts acting upon the control surfaces of an aircraft traveling at high speeds. The equivalent output noise can be determined by considering the power transmission between the point where the noise is applied and the output of the system.

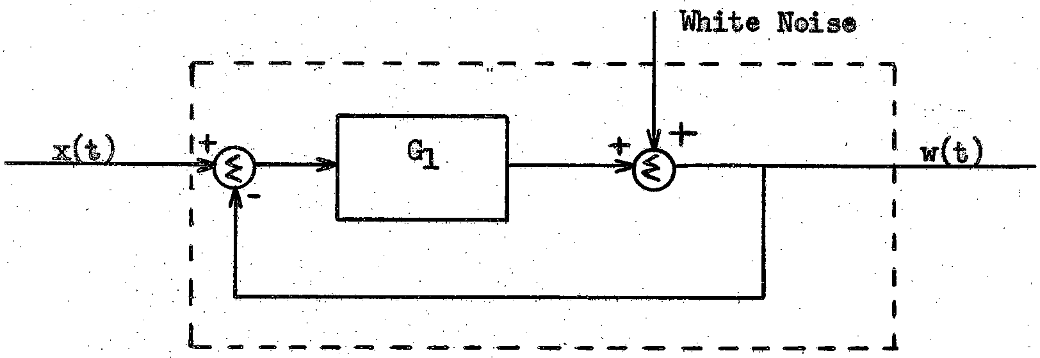
The results obtained in Chapter 2 for the ideal identifier are valid only if the signal energy at the output of the system being identified is large compared to the power spectral density of the external noise. Before considering specific examples it will be established that this condition does exist for a large class of problems. The energy of the observed signal may be expressed as

$$E_W = \frac{T_I \, \overline{x^2} \, W_G}{W_x} \qquad (6\text{-}1)$$

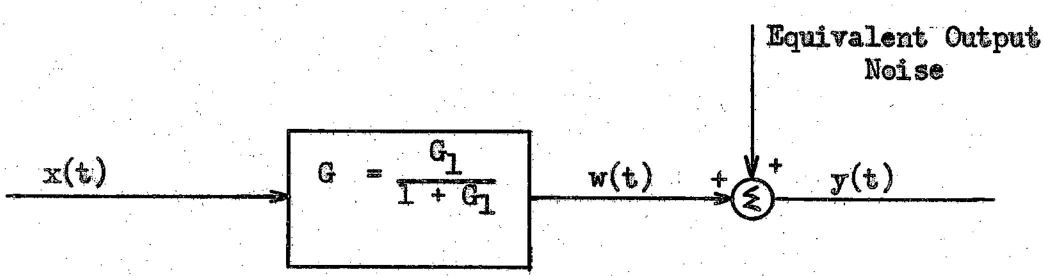
and the ratio $E_W / \Phi_n(0)$ can, by using the output signal-to-noise ratio constraint, be expressed as

$$\frac{E_W}{\Phi_n(0)} = T_I \, W_x \, \gamma \qquad (6\text{-}2)$$

It was pointed out in Chapter 1 that T_I is at least of the order of magnitude of the significant duration of the impulse response. By relating the significant duration of $g(\lambda)$ to the system bandwidth and using the fact that W_x is much larger than W_G it can be shown that Eq. (6-2) is, in most practical cases, a large number if $\gamma > 1$. If, for instance, T_I is considered to be at least as large as 2 system time constants and $W_x = 100\, W_G$ the ratio is greater than or equal to $100\,\gamma$. These conditions will be satisfied in all the examples presented in this chapter so that the results may be considered to be equivalent to those that would be obtained by an ideal identifier.



(a)



(b)

Fig. 6-1

White Noise Within the Feedback Loop and Equivalent Output Noise

6.2 T_I for a Minimum Phase Second Order System - White Output Noise

As an example of the type of problem that might be encountered in the design of an adaptive control system a second order system with a pole-zero configuration as shown in Fig. 6-2 will be considered. The impulse response has the form

$$g(\lambda) = \frac{\sqrt{K_G}\ \omega_0\, e^{-\zeta\omega_0\lambda}}{\sqrt{1-\zeta^2}} \sin\omega_0\sqrt{1-\zeta^2}\ \lambda \qquad (6\text{-}3)$$

where K_G is the zero frequency power gain, ω_0 is the undamped natural frequency, and ζ is the relative damping ratio. The average or nominal values of K_G and ω_0 will be considered to be unity. Frequency and/or magnitude scaling may be applied to the final results if the identification time is desired for the more general case of arbitrary values for K_G and ω_0. The figure $\zeta = 1/2$ is a convenient value to choose for the average damping ratio; a ζ of 1/2 results in a step response with a moderate amount of overshoot (16.36%). The average impulse response is plotted in Fig. 6-4, the vertical bars have been added to illustrate the size of the standard deviations associated with three different values of the variance of the impulse response estimate. Note that, for engineering purposes, the impulse response may be considered essentially zero for values of λ greater than ten seconds.

The equivalent noise bandwidth of the average system must be calculated if T_I is to be expressed in terms of the output signal-to-noise ratio. By definition [18, Middleton, p. 684] the equivalent noise bandwidth is equal to

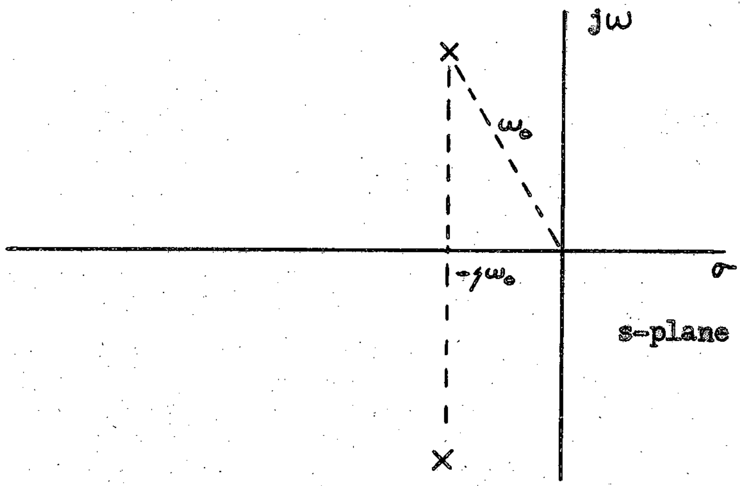


Fig. 6-2

Second Order Pole-Zero Configuration

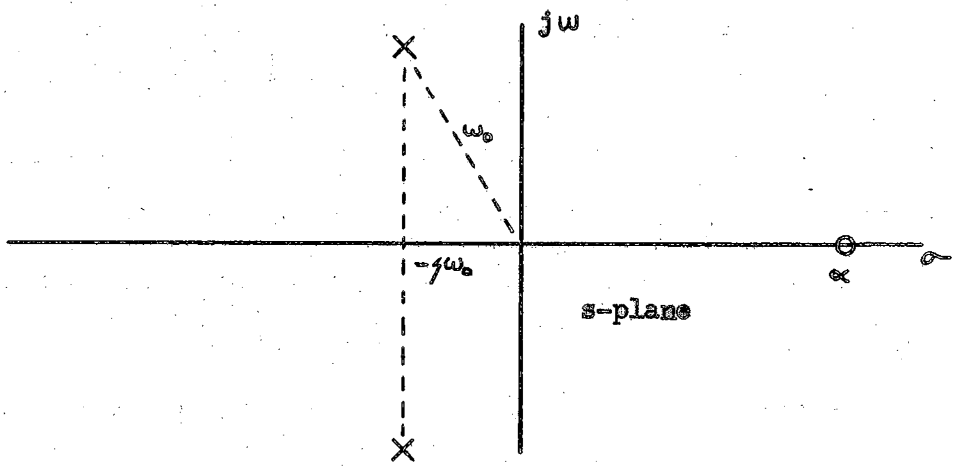


Fig. 6-3

Nonminimum Phase Second Order Pole-Zero Plot

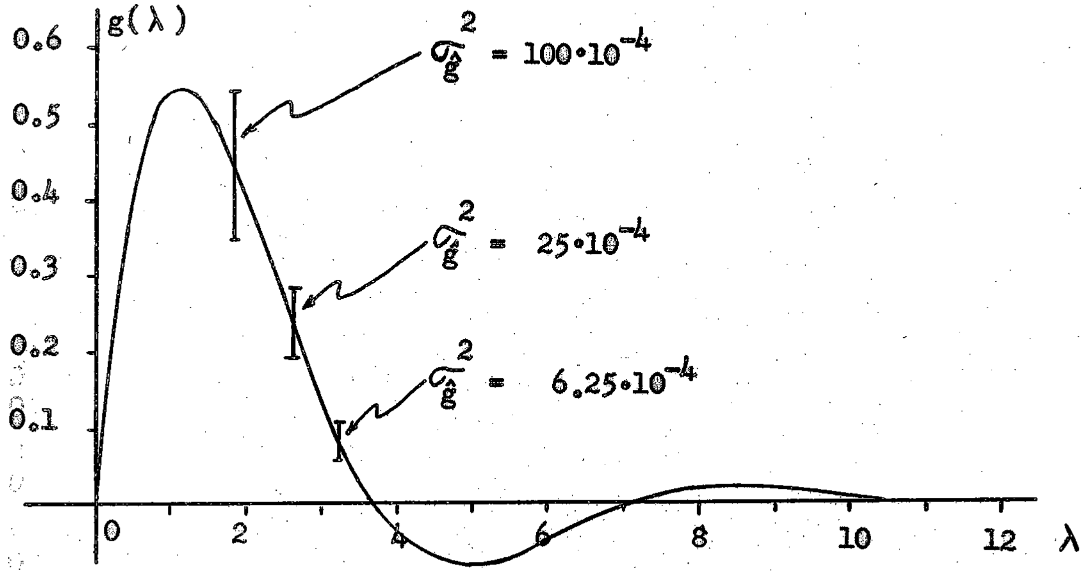


Fig. 6-4

Impulse Response - Second Order System - $K_G = \omega_o = 1$, $\zeta = 1/2$

$$\tilde{W}_G = \frac{1}{4\pi \tilde{K}_G} \int_{-\infty}^{+\infty} \widetilde{G(\omega)} \;\; \widetilde{G(-\omega)} \, d\omega \tag{6-4}$$

or by using Parseval's Theorem

$$\tilde{W}_G = \frac{1}{2 \tilde{K}_G} \int_{0}^{\infty} [\widetilde{g(\lambda)}]^2 \; d\lambda \tag{6-5}$$

For this example application of Eq. (6-4) or Eq. (6-5) yields

$$\tilde{W}_G = \frac{\omega_0}{8\zeta} = 0.25 \text{ cps} \tag{6-6}$$

The bandwidth of the test signal must be large compared to the bandwidth of the system being identified if the errors in the estimate due to the limitations of the test signal are to be small compared to the errors from the external noise. Therefore, before the bandwidth of the test signal can be specified with any degree of assurance it is necessary to consider the range of parameter variations and the effects of these variations upon the bandwidth of the system. Suppose the resonant frequency is allowed to vary between the values 0.25 and 4, and that the relative damping ratio varies from 0.25 to 0.75. These limits on the parameters allow a 16:1 change in ω_0 and impulse responses ranging from highly oscillatory to near critical damping. A completely arbitrary but reasonable starting point for the specification of the test signal bandwidth would be to set W_x equal to 100 times the average system bandwidth. In this example the maximum system bandwidth occurs when ω_0 is at its largest value, and ζ takes on its minimum value. For $\omega_0 = 4$ and $\zeta = 0.25$

$W_G\Big|_{max}$ is equal to 2 cps. In this case

$$\frac{W_G\Big|_{max}}{\widetilde{\widetilde{W}}_G} = 8 \qquad (6\text{-}7)$$

so that if W_x equals 100 $\widetilde{\widetilde{W}}_G$, the test signal bandwidth is at least 12.5 times the bandwidth of the system being identified, and this choice of W_x is justified.

Identification techniques using a periodic test signal require that the test signal period be large compared to the significant duration of the impulse response. An indication of the test signal period that would be required for this example may be obtained by defining the significant duration of the impulse response to be equal to five times the reciprocal of the damping factor (here equal to $\zeta\omega_o$) and computing the maximum and average value of this figure. The maximum significant duration of $g(\lambda)$ is 80 seconds, while the significant duration of the average system is 10 seconds. If, for instance, T_x is taken as two times the maximum significant duration of $g(\lambda)$, T_x would equal 160 seconds.

The identification time is

$$T_I = \frac{2\,\widetilde{W}_G\ \widetilde{\widetilde{K}}_G}{\gamma\,\sigma_{\hat{g}}^2} = \frac{\Phi_n(0)\ \ (2W_x)^2}{\overline{x^2}\ \ \sigma_{\hat{g}}^2} \qquad (6\text{-}8)$$

This expression is valid for any of the identification techniques discussed in Chapters 3, 4, and 5, when the external noise is white, and also represents the identification time of an ideal identifier. The identification time, multiplied by a constant c, is shown plotted vs the variance of the impulse response estimate in Fig. 6-5. c may be expressed either in terms of the gain-bandwidth product and output signal-

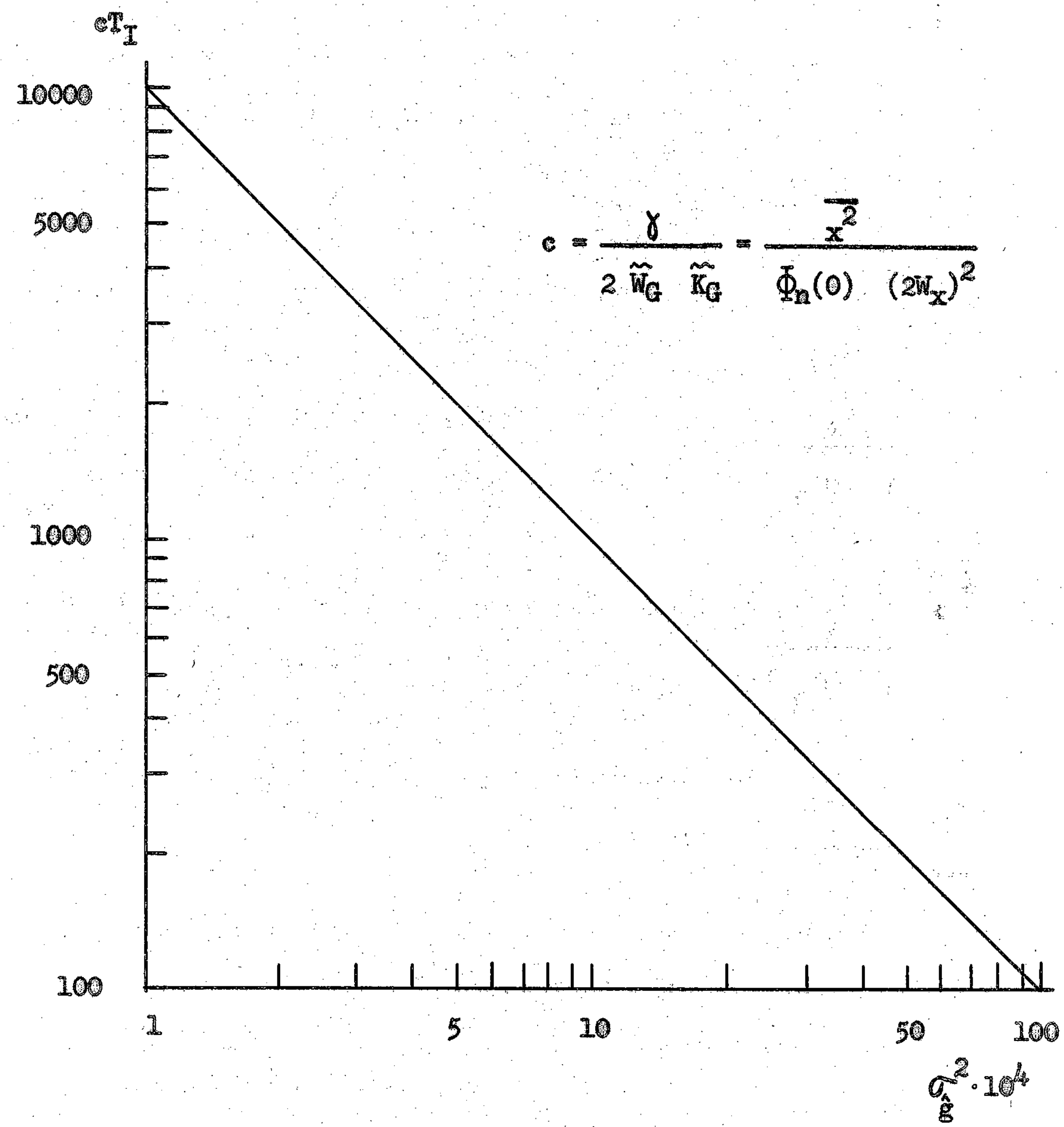


Fig. 6-5

Normalized Identification Time vs Variance to-noise ratio, or in terms of the noise and test signal parameters $\Phi_n(0)$, $\overline{x^2}$, and W_x.

The curve is used as follows: suppose design requirements call for the mean square value of the signal at the output of $g(\lambda)$ due to the test signal to be equal to the mean square value of the effective output noise, that is $\gamma = 1$. The average gain for this example is unity, and from Eq. (6-6) the average bandwidth is 0.25 cps. Thus, the constant c is equal to 2. If the desired variance is $25 \cdot 10^{-4}$ (see Fig. 6-4) the identification time, as obtained from Fig. 6-5, is equal to 200 seconds. A smaller variance, such as $6.25 \cdot 10^{-4}$, would require an identification time of 800 seconds whereas a variance of $100 \cdot 10^{-4}$ would require only 50 seconds. These identification times are directly proportional to γ so that if a larger signal-to-noise ratio can be tolerated the identification times can be reduced substantially. For instance if $\gamma = 3$ and $6.25 \cdot 10^{-4}$ is taken as the variance T_I is 267 seconds.

Before leaving this example the identification time required by crosscorrelation identification using Gaussian noise as a test signal will be compared with the identification time required by the same technique using a periodic test signal. Using Eq. (3-27) the upper bound on the identification time for a Gaussian test signal is plotted for $\gamma = 1$, $\gamma = 2$, and $\gamma = \infty$ in Fig. 6-6. Also plotted on the same figure is T_I as determined from Eq. (3-29) (or equivalently from Eq. (6-8)) for $\gamma = 1/2$, $\gamma = 1$, and $\gamma = 2$. For $\gamma = 1$ the technique using the Gaussian signal takes three times longer to identify the system than the method using a periodic test signal; for $\gamma = 2$ it takes five times longer. The smallest identification time obtainable with a

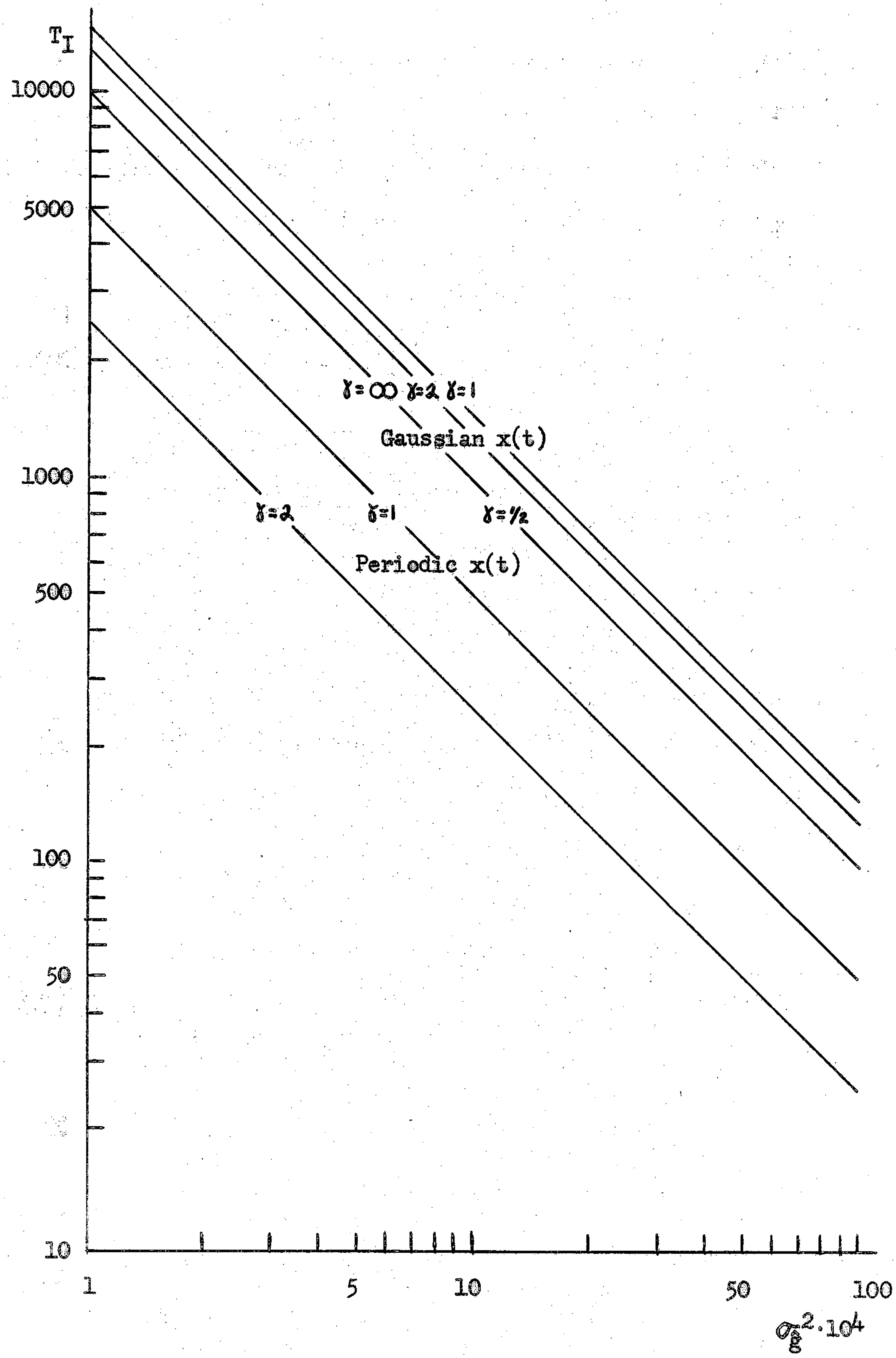


Fig. 6-6

Comparison of T_I for Gaussian and Periodic Test Signals

Gaussian test signal occurs when γ approaches infinity, and this value for T_I equals the identification time obtained by using a periodic test signal and a γ of 1/2. The advantage of periodic test signals over other random test signals is obvious.

6.3 T_I for a Nonminimum Phase Second Order System - White Output Noise

The pole-zero configuration for the nonminimum phase, second order example is shown in Fig. 6-3. The corresponding impulse response is

$$g(\lambda) = \frac{\sqrt{K_G}\,\omega_o e^{-\zeta\omega_o\lambda}}{\alpha}\left[\frac{(\alpha+\zeta\omega_o)}{\sqrt{1-\zeta^2}}\sin\omega_o\sqrt{1-\zeta^2}\,\lambda-\omega_o\cos\omega_o\sqrt{1-\zeta^2}\,\lambda\right] \tag{6-9}$$

The average values for ω_o and K_G will again be taken as unity, and the average relative damping ratio, ζ, will be set equal to 1/2. The location of the right-half plane zero for this example will be $\alpha = 2$. The impulse response, Eq. (6-9), is plotted for this set of parameters in Fig. 6-7, and the vertical bars again indicate the size of the standard deviation for several values of $\sigma_{\hat{g}}^2$. In this example also, $g(\lambda)$ is essentially zero for values of λ greater than ten seconds.

The equivalent noise bandwidth for the nonminimum phase example can be obtained from Eq. (6-4) or Eq. (6-5). The result is

$$\tilde{W}_G = \frac{\omega_o(\alpha^2+\omega_o^2)}{8\zeta\alpha^2} \tag{6-10}$$

If it is assumed that the zero location is fixed and ω_o and ζ are allowed to vary over the same ranges as in the previous example substitution of a few numbers into Eq. (6-10) quickly establishes that the maximum W_G again occurs when $\omega_o = 4$ and $\zeta = 0.25$. $W_G\big|_{max} = 10$ cps

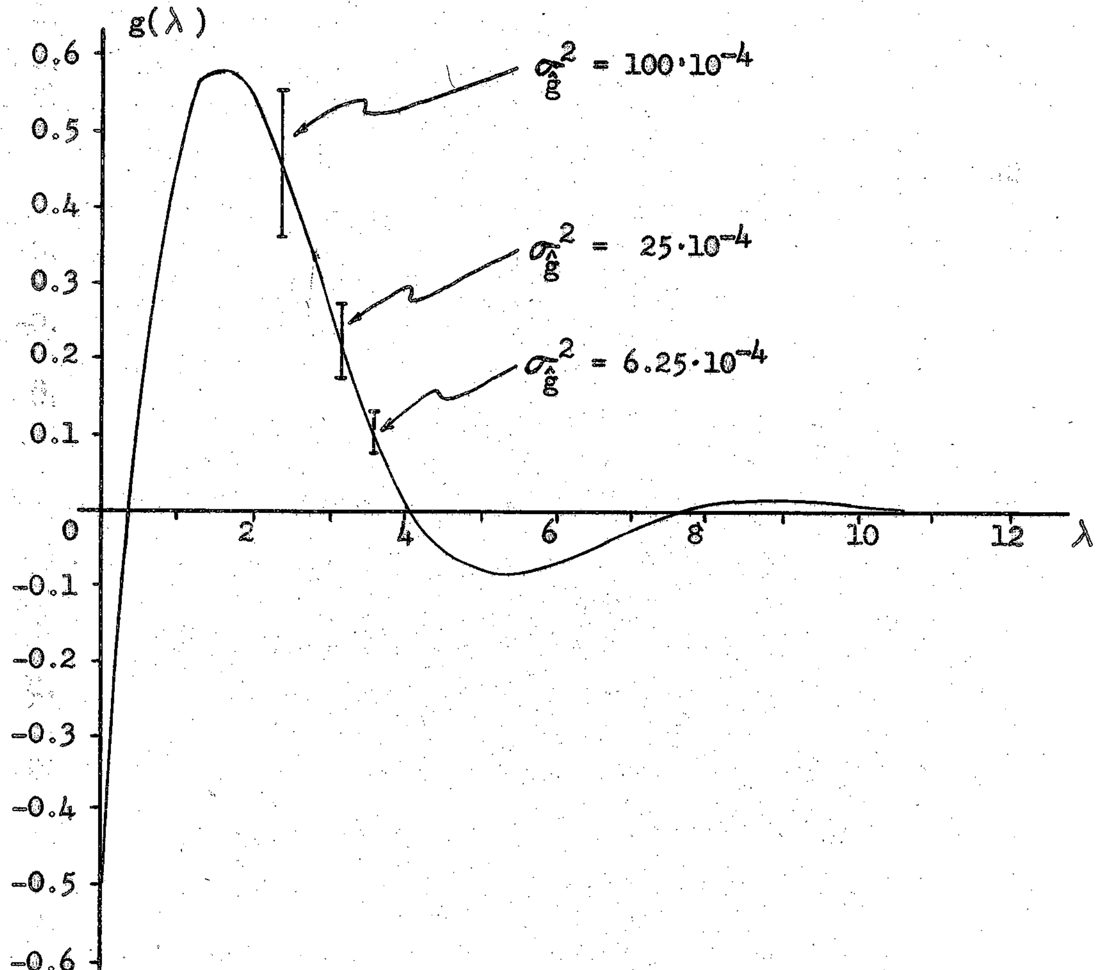


Fig. 6-7

Impulse Response - Nonminimum Phase System - $K_G = \omega_o = 1$, $\zeta = 1/2$, $\alpha = 2$

as compared to $\widetilde{W}_G = 0.3125$ cps. In this case choosing W_x equal to $100\,\widetilde{W}_G$ may not be very satisfactory because the test signal bandwidth would only be about three times the system bandwidth when the parameters approached the values given above. This condition can, of course, be improved by using a test signal with a wider bandwidth.

Consideration of the significant duration of the impulse response for this example yields results that are identical to those for the minimum phase example. The maximum significant duration is 80 seconds, and the period of the test signal must be picked accordingly.

Fig. 6-5 may be used for this example also. Here c = 1.6 so that the identification time is 250 seconds when $\sigma_{\hat{g}}^2 = 25\cdot10^{-4}$, 62.50 seconds when $\sigma_{\hat{g}}^2 = 100\cdot10^{-4}$, and 1000 seconds when $\sigma_{\hat{g}}^2 = 6.25\cdot10^{-4}$ if γ is unity. Comparing these results with those of the previous example shows that a 25% longer identification time is required because of the increase in the average system bandwidth.

The results of these two examples, which are typical of systems of any order with impulse responses resembling those of Fig. 6-4 or Fig. 6-7, show that for output signal-to-noise ratios of the order of unity the identification time ranges from about 10 to 100 times the significant length of the impulse response depending upon the degree of accuracy that is required. With signal-to-noise ratios of 10 to 20 db the identification time could be reduced to the order of magnitude of the significant length of the impulse response.

6.4 White Noise Originating Within the Feedback Loop

This work has considered the problem of identifying an unknown linear system in the presence of an external disturbance introduced at the output of the system. The results that have been obtained may be extended to situations where disturbances are introduced at other points of the system by determining the output disturbance that is equivalent to the actual disturbance [20, Newton, Gould, and Kaiser, p. 37]. This may be achieved by considering the power transfer function from the origin of the noise to the output of the system. The techniques will be illustrated here for the case of white noise introduced within the feedback loop of a control system as illustrated in Fig. 6-1(a). Wideband disturbances acting on the actuators of a control system could be represented in this manner.

The signal transfer function from the origin of the external noise to the output of the system is

$$\frac{1}{1 + G_1(\omega)} \tag{6-11}$$

It is the impulse response relating the signal w(t) to x(t) that is to be determined by the identification technique. By calling the Fourier transform of this impulse response $G(\omega)$ it is easy to show that Eq. (6-11) may be put into the form

$$1 - G(\omega) \tag{6-12}$$

Let the power spectral density of the external noise be N, then the power spectral density of the equivalent output noise is given by the expression

$$\Phi_n(\omega) = N\left|1 - G(\omega)\right|^2 \tag{6-13}$$

and is sketched for a typical unity-gain system in Fig. 6-8.

The effects of this kind of noise upon the identification time for the matched filter technique will now be determined. Consideration of

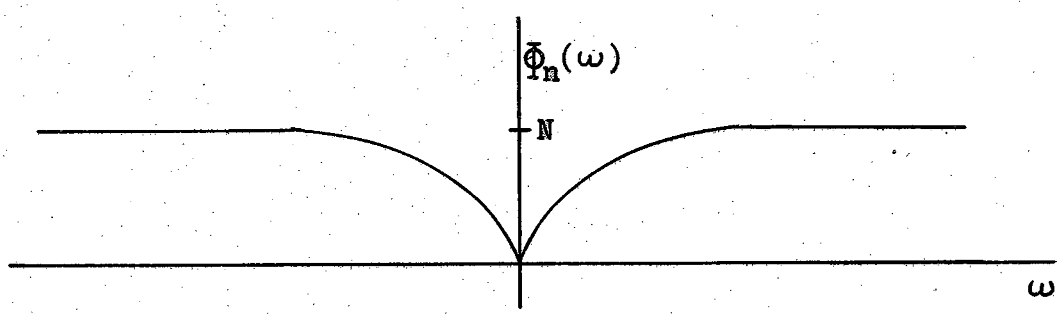


Fig. 6-8

Equivalent Output Noise Power Density Spectrum

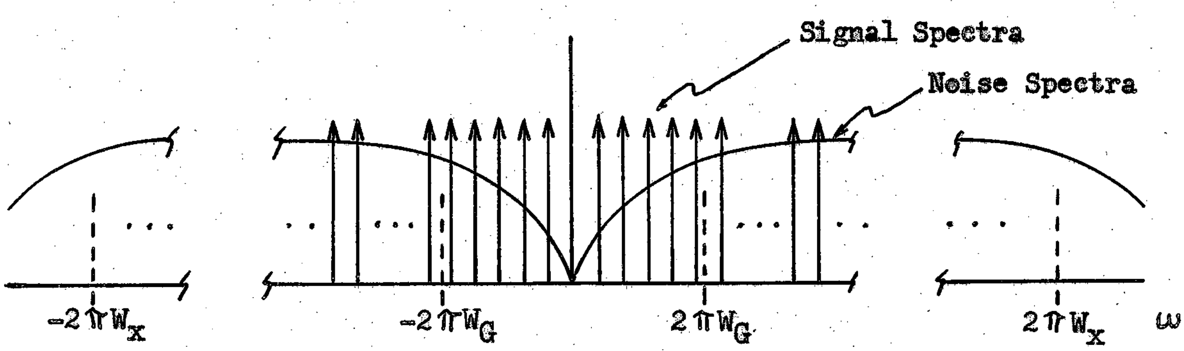


Fig. 6-9

Signal and Noise Spectra at the Output of the Matched Filter the noise and signal spectra at the output of the matched filter will facilitate the analysis. These spectra are plotted in Fig. 6-9. The variance of the impulse response estimate is determined by the total amount of noise energy that is accepted by the comb filter. Since W_x is much larger than W_G the noise out of the comb filter is nearly equal to that which would be obtained if the white noise were introduced at the output of the system instead of within the feedback loop. This is because the noise spectra in Fig. 6-9 differs from that shown in Fig. 5-4 only in the low frequency region.

Actually, for a fixed comb filter bandwidth, the variance is reduced a small amount and this reduction could be reflected in a reduced identification time. An indication of the difference in variance (or identification time) between the case of introducing white noise within the feedback loop to that of a white output noise can be obtained by considering the equivalent output noise spectrum, Eq. (6-13), to be zero within the equivalent noise bandwidth of $g(\lambda)$, and equal to N outside the equivalent noise bandwidth. Then the variance is proportional to $(W_x - W_G)$ when the noise originates within the loop and is proportional to W_x when the noise originates at the output. The ratio of the variances is

$$\frac{\sigma_{\hat{g}}^2\Big|_{\text{noise in feedback loop}}}{\sigma_{\hat{g}}^2\Big|_{\text{noise at output}}} = \frac{W_x - W_G}{W_x} \qquad (6\text{-}14)$$

The saving is negligible if $W_x \gg W_G$.

For crosscorrelation identification the power spectral density of the noise at the output of the multipler, $\Phi_{me}(\omega)$, determines the identification time. The power spectral density of the component of this noise due to the external disturbance can be obtained by convolving the spectrum of the external noise with that of the test signal [10, Gardner and Barnes, p. 275]. This convolution is pictured in Fig. 6-10 for a value of ω which is less than the order of $2\pi W_x$. $\Phi_{ne}(\omega)$ is proportional to the area under the product of these two functions and is given by

$$\Phi_{ne}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi_n(\omega - \Omega)\, \Phi_x(\Omega) \quad d\Omega \qquad (6\text{-}15)$$

It is apparent that $\Phi_{ne}(\omega)$ will be constant from $\omega = 0$ to a value of ω near $2\pi W_x$ if $W_G \ll W_x$. By using the same approximation for $\Phi_n(\omega)$ that was used above, and the fact established in Chapter 3, that T_I is proportional to $\Phi_{ne}(0)$, when $\Phi_{ne}(\omega)$ can be considered to be constant over the passband of the averaging filter, the following relationships can be established: T_I is proportional to $(W_x - W_G)$ if the external noise is introduced within the feedback loop; and T_I is proportional to W_x if the external noise is introduced at the output of the system. The result is equivalent to that obtained for matched filter identification.

Detailed analysis of the sampling identification technique when white noise is introduced within the feedback loop is not as easy because of the relative complexity of the covariance matrix, Eq. (4-13). However, by establishing the fact that the covariance matrix of the output noise is nearly diagonal it can be argued that the results for this case are approximately equal to the results obtained when white noise is introduced at the output.

The external noise correlation matrix, $\underline{\Phi}_n$, is obtained by sampling the continuous noise correlation function. If it can be shown that

$\phi_n\left(\frac{m}{2W_x}\right)$ is much smaller than $\phi_n(0)$ for all $m \geq 1$ it follows that all elements of Φ_n off the major diagonal are much smaller than the major diagonal elements.

Consider an external noise with an autocorrelation function of the form

$$\frac{N\beta e^{-\beta|\tau|}}{2} \tag{6-16}$$

introduced within the feedback loop of a system which is being identified by the sampling technique. As β approaches infinity, the autocorrelation function approaches $N\,\delta(\tau)$, i.e., the noise becomes white. For this case the power spectral density of the equivalent output noise expressed in terms of the complex frequency, s, is

$$\Phi_n(s) = \frac{N\beta^2[1 - G(s)][1 - G(-s)]}{2\,(s+\beta)\,(-s+\beta)} \tag{6-17}$$

$$= \frac{N\beta^2 \; P(s)\,P(-s)}{2\,(s+\beta)\,(-s+\beta)\,Q(s)\,Q(-s)}$$

where P(s) and Q(s) are polynominals in s. The equivalent output noise autocorrelation function, $\phi_n(\tau)$, can be obtained by taking the inverse Laplace transform of Eq. (6-17). $\phi_n(\tau)$ will be composed of a sum of terms, the relative amplitudes of these terms are determined by the residues associated with the poles of $\Phi_n(s)$. Evaluation of these residues shows that the residues associated with the poles located at $s = \pm\beta$ approach β times a constant as β approaches infinity, whereas the residues associated with the roots of Q(s) and Q(-s) simply approach a constant as β approaches infinity. Thus in the limit as the external

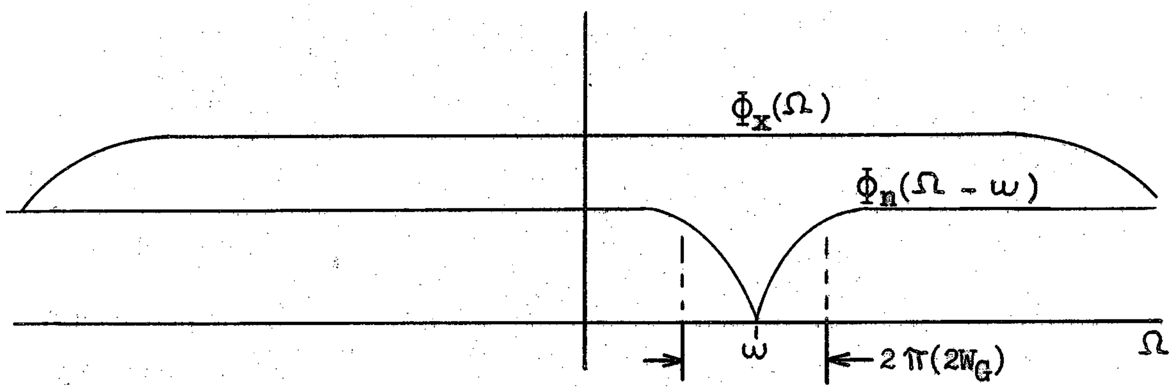


Fig. 6-10

Convolution of $\Phi_n(\omega)$ and $\Phi_x(\omega)$

disturbance becomes white $\phi_n\left(\frac{m}{2W_x}\right)$ becomes much smaller than $\phi_n(0)$ so that the covariance matrix of the external disturbance is nearly numerically equal to that obtained when white noise is introduced at the output of the system. It follows from this, that $\measuredangle_{\hat{g}}$ and the identification time will be nearly equivalent in the two cases also.

6.5 Consideration of a Narrowband Noise Process

Up to now primary consideration has been given to problems where the bandwidth of the external noise is wide compared to the system being identified. If the bandwidth of the external disturbance is narrow compared to the unknown system the identification time can be reduced in some cases. If the disturbance occurs at frequencies that are well outside of the pass-band of $g(\lambda)$ the effects of the noise upon the identification could be almost completely eliminated by filtering the observed signal before performing any identification operations. (See Fig. 6-11) The filter could be designed with a stop-band centered about the noised frequencies if they are known, or, if exact knowledge concerning the frequencies of the disturbance is lacking, all signals above some lower cutoff frequency could be eliminated provided the cut-off frequency is high enough to permit the fine structure of $g(\lambda)$ to be reproduced.

When the noise occurs at frequencies within the pass-band of the unknown system the problem becomes somewhat more complicated. If the disturbance bandwidth is sufficiently narrow compared to the transmission spectrum of the system being identified, and if the range of noise frequencies is known, then the unwanted signals can be eliminated by a suitable filter as shown in Fig. 6-11. The filter will also suppress the signal which carries information about the unknown system, but if the

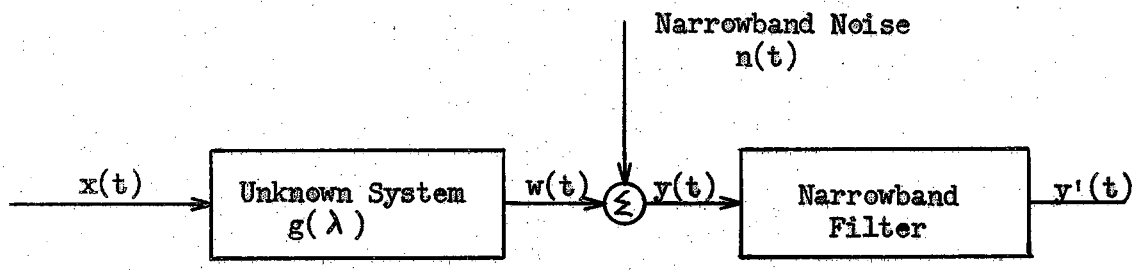


Fig. 6-11

Elimination of Narrowband Noise by Filtering filter has a narrow stop-band the loss of signal information over this band of frequencies will not in general have a serious effect upon the impulse response estimate. When the noise can be filtered in this manner the identification time becomes the same order of magnitude as the significant duration of the impulse response.

The noise cannot be filtered out before performing the identification operations if the frequency location of the disturbance is not known, and the techniques of the previous chapters must be used. The identification time may still be reduced, however, if because of the narrowband properties of the noise, a larger output signal-to-noise ratio can be used.

CHAPTER 7

SUMMARY OF RESULTS, CONCLUSIONS, AND RELATED PROBLEMS

This final chapter discusses the significance of the ideal identifier and its relation to practical identification techniques. The crosscorrelation, sampling, and matched filter identification methods are compared from a mathematical, noise immunity, and operational standpoint. The importance of the results of this research for adaptive systems is discussed, and some suggestions are given for future work on the problem of system identification.

7.1 Significance of the Ideal Identifier

The concept of the ideal identifier provides a common basis for comparing all conceivable identification techniques. The problem of system identification can be stripped of all considerations of hardware and implementation and considered solely from the point of view of statistical parameter estimation by introducing the ideal identifier. The ideal identifier has been defined in such a manner that no *a priori* knowledge concerning the unknown impulse response is required. In effect, this means that each sample value of $g(\lambda)$ is completely independent of all other sample values. As a result, the expression obtained for the identification time is independent of the particular impulse response that is being measured, and depends only upon the measurement environment parameters and the degree of accuracy that is desired. The measurement environment is specified by the power spectral density of the external noise and the mean square value of the test signal. The overall accuracy of the impulse response estimate is determined by the variance $\sigma^2_{\hat{g}}$ and the sampling interval $\Delta\lambda$. The variance is a direct measure of the errors resulting from external noise, and a trade off between variance and identification time can be made. Although the sampling rate is not a direct measure of the error introduced by approximating $g(\lambda)$ by a finite set of parameters, it determines a cutoff frequency above which information about $g(\lambda)$ is lost. When a priori information about the unknown system's bandwidth is available $\Delta\lambda$ can be chosen so that this cutoff frequency is large compared to the system bandwidth.

The identification time as determined by the ideal identifier is a conservative estimate of the time that would be required to identify an unknown system in a practical situation. It is conservative in two respects. First, ideal identification is performed in an environment of white noise, and thus represents a condition where it is impossible to use the autocorrelation properties of the noise to reduce the identification time. In practice the noise is never truly white, and the correlation properties of the noise could theoretically be used to reduce the identification time. This was illustrated for the special case of a narrowband noise in Chapter 6. The problem of how to best use the autocorrelation properties of the noise to reduce identification time has not been considered. Second, ideal identification yields a conservative estimate of identification time because it assumes no a priori knowledge about the system, when in practice at least some prior knowledge is usually available. Proper exploitation of this a priori knowledge should result in a reduced identification time.

Although the identification time of the ideal identifier represents a greatest lower bound on the identification time obtainable by practical techniques it is an important result because many of the methods in use, or proposed in the literature, are equivalent to the ideal identifier; it is, therefore, applicable to a wide class of current problems. In addition, as mentioned previously, the result of the ideal identifier serves as a base point to which other identification methods can be compared.

7.2 Equivalence of Crosscorrelation, Sampling, and Matched Filter Identification

There is an inherent mathematical similarity between the three practical identification techniques that have been considered in this work. That this is the case is not surprising because, after all, each technique has the same objective - that of identifying, in terms of the impulse response, an unknown linear system. The mathematical unity of the three methods is provided by the Wiener-Hopf equation

$$\phi_{xy}(\tau) = \int_0^\infty \phi_x(\tau - \lambda)\, g(\lambda)\, d\lambda \qquad (7\text{-}1)$$

The "solution" of this equation is accomplished in crosscorrelation identification by using a test signal with an impulse-like autocorrelation function and measuring the value of the input-output crosscorrelation function at the desired values of delay. Matched filter identification also uses a white test signal and approximates a solution to Eq. (7-1) in a similar manner. The two techniques differ in that the matched filter output signal displays an estimate of $g(\lambda)$ as a function of real time, whereas the crosscorrelation method presents its information as a function of the crosscorrelator delay parameter, τ.

The identification technique based on samples of input-output data approaches the solution of the Wiener-Hopf equation in a different manner. In this case the integral in Eq. (7-1) is approximated by a finite sum, and a set of linear algebraic equations is formed with sample points of $g(\lambda)$ as the variables. An estimate of $g(\lambda)$ is obtained by solving this set of algebraic equations. The solution is greatly simplified if the test signal is white, for then the equations become independent.

While the underlying mathematical equivalence of these identification techniques may be apparent the statistical equivalence is not. In crosscorrelation identification the external noise is multiplied by the delayed test signal resulting in a noise term in the signal at the output of the multiplier. Since the information-bearing part of this signal is the average value, noise reduction is accomplished by means of a low-pass averaging filter. The external noise is for the large part unaffected by the matched filter in that identification technique; the variance in the estimate is reduced in this case by averaging over several individual estimates of the impulse response. This type of noise reduction may be achieved by using a comb filter.

Sample points of the estimate of $g(\lambda)$ in crosscorrelation identification may be thought of as being measured in parallel, each channel of the correlator providing an estimate of a single sample point. This parallel type of operation allows the use of low-pass filters to reduce the effects of the external noise. The matched filter method provides a continuous estimate of $g(\lambda)$ as a function of real time; estimates of $g(\lambda)$ may be thought of as being measured in series. In this case it is only by the use of a repetitive test signal and averaging over a number of estimates, that the effects of the external noise can be reduced.

The sampling method of system identification incorporates both series and parallel type of operations. In general the input and output signals are observed for a length of time that is equal to an integral number of periods of the test signal. Thus, in computing the empirical correlation functions a real time or series type of averaging is performed. The solution of the set of algebraic equations provides estimates of each sample point simultaneously, a parallel type of operation.

Although each identification scheme is based upon the solution of the Wiener-Hopf equation the external noise enters the problem differently in each method, and different techniques are used to reduce the variance of the impulse response estimate; nevertheless, each type of identifier yields exactly the same results for the identification time. While the mutual equivalence of these practical identification operations is important, of even greater importance is the result that these methods are equivalent to an ideal identifier; they represent the best that can be dome when no a priori knowledge of the unknown system is available.

7.3 Operational Similarities and Relative Advantages of the Various Identification Methods

There is in addition to the similarities of the identification techniques mentioned above an operational similarity which is particularly evident when the external noise is white and the test signals are optimized. In order to guarantee a satisfactory estimate of $g(\lambda)$ both crosscorrelation and matched filter identification techniques require a wideband test signal, one that, from the standpoint of the unknown system, approximates a white spectrum. The optimum test signal for the sampling technique is white only when the external disturbance is white noise. While the sampling procedure offers the advantage of being able to operate with non-white test signals the computational advantages gained by using a white test signal often outweigh any advantages that might be gained by optimizing the test signal for a non-white external noise.

Another similarity regarding the test signals used by the various identification methods is that in each case it is advantageous to use a periodic test signal. For crosscorrelation the use of a periodic test signal, and an ideal finite-memory integrator as an averaging filter, eliminates the noise terms in the impulse response estimate that would normally result from the random character of the test signal. Averaging over a number of independent estimates is easily accomplished in the matched filter method if the test signal is periodic. The advantage of using a periodic test signal, and observing $x(t)$ and $y(t)$ for an integral number of periods, in the identification technique using sampling is that the test signal correlation matrix is independent of the particular time at which the sequence of sample points begins. This property makes it possible to permanently store the test signal correlation information, thereby simplifying the computational problem.

Each identification method offers some unique feature with regard to the way the impulse response estimate is presented. The matched filter method offers the advantage of producing a continuous estimate of $g(\lambda)$, and is particularly useful when analog computations are to be performed. The sampling technique presents a set of equally spaced sample points which lend themselves to digital computation. However, the sampling method requires some digital computation in order to obtain the estimates whereas both the matched filter and crosscorrelation methods yield impulse response estimates directly. The output of the averaging filter for each channel in the crosscorrelation identifier is a continuous signal representing the estimate of a single sample point of the unknown impulse response. Output data in this form can be used directly for certain types of analog computations, and it is simple to convert to digital form. Another advantage of crosscorrelation is that the distance between sample points is determined by the delays in each channel and need not be equally spaced. This property permits such things as grouping a large number of sample points where $g(\lambda)$ is large or expected to be changing rapidly with respect to λ, and placing fewer sample points where the value of $g(\lambda)$ is expected to be overshadowed by noise or slowly varying.

It is extremely difficult to assess the relative advantages of the identification techniques from the standpoint of equipment complexity unless the assessment is made in relation to a particular application. For instance, the construction of a suitable matched filter and comb filter may seem much more complicated than a crosscorrelator unit. However, if a continuous visual display of the impulse response estimate is needed in a certain application it can be obtained directly with the matched filter identifier and an ordinary oscilloscope, whereas the crosscorrelation method would require additional electronic equipment to transform its output data into a continuous display of $g(\lambda)$. In the absence of a well-defined application no general conclusions regarding the relative merits of the various methods can be made from the equipment standpoint.

7.4 Significance of Identification Time Results for Adaptive Systems

The results for the identification time of a linear system that have been established in this work indicate that any of the identification schemes that have been shown to be equivalent to the ideal identifier are practical for use in adaptive systems provided the environmental conditions, or system parameters, are slowly varying with respect to the significant length of the impulse response. Although the identification time is highly dependent upon the output signal-to-noise ratio and the variance of the estimate, the identification time is of the order of 10 to 100 times the significant duration of the impulse response. As a result, these identification techniques will not provide accurate data for rapidly varying systems.

The concept of using the output signal-to-noise ratio as a criterion for establishing the mean square value of the test signal is a reasonable one if the external noise bandwidth is at least of the order of magnitude of the system bandwidth. In such situations the effects of the noise upon the system can be used as a basis for assigning a value of γ so that the test signal will not unduly disturb the normal operation of the system. However, if the external disturbance is confined to a relatively small frequency band the signal-to-noise ratio concept must be used with more care. The frequency distribution of the output signal due to $x(t)$ is determined by the transmission characteristics of the unknown system; hence, when the noise is narrowband, the signal energy is distributed over a wider band of frequencies than the noise. This condition may result in an allowable value of γ that is appreciably different from that which would be expected for wideband noise. In cases of low intensity wideband noise, or when the noise is narrowband, it may be possible to increase the output signal-to-noise ratio enough to reduce the identification time to the order of magnitude of the impulse response's significant length. In the absence of a priori knowledge this is the best that can be done by any measurement technique.

7.5 Related Problems

This research has studied the problem of identifying an unknown linear system by means of its impulse response function. Estimates of the system transfer function or other transforms (such as the Z-transform) which completely describe the system may be obtained from the impulse response estimate. However, there is no guarantee that the transform of an impulse response estimate will be a good estimate of the true transform [12, Guillemin, p. 662]. Consequently, a problem worthy of consideration is the identification time required to obtain estimates of the system transfer function directly.

In addition to the impulse response, transfer function, and Z-transform, a complete description of the unknown system can be obtained by specifying the coefficients of a series expansion of the impulse response. The representation of $g(\lambda)$ by a series of orthogonal functions is of particular interest because the practical identification techniques that have been considered are readily adapted to the measurement of the coefficients of such a series.

As an example, the crosscorrelation identification method can be modified to measure the coefficients of an orthogonal expansion of $g(\lambda)$ by replacing the ideal delay in Fig. 3-1 by a filter with an impulse response $h_i(\lambda)$. With this change the expected value of the multiplier output signal is

$$\int_0^\infty \int_0^\infty E[x(t - \lambda_1) \quad x(t - \lambda_2)] \; h_i(\lambda_2) \;\; g(\lambda_1) \;\; d\lambda_1 \;\; d\lambda_2 \tag{7-2}$$

and if the test signal is white this becomes proportional to

$$\int_0^\infty h_i(\lambda_1) \; g(\lambda_1) \;\; d\lambda_1 \tag{7-3}$$

Comparing Eq. (7-3) to Eq. (1-9) shows that if $h_i(\lambda)$ is made equal to $\psi_i(\lambda)$ the modified crosscorrelation technique can be used to measure the coefficients of an orthonormal series expansion of $g(\lambda)$. Of course, each coefficient would be measured by a different channel of the correlator.

No change in the mechanization is required to adapt the sampling identification method to estimate the coefficients of an orthonormal series expansion of the impulse response. The necessary modifications occur only in the manner in which the data is processed. If g(p) is represented

$$g(p) = \sum_{i=1}^{Q} \alpha_i \, \psi_i(p) \tag{7-4}$$

Where the ψ_i are members of a set of orthogonal functions, the normal equations, Eq. (4-8), can be put into the form

$$t_a (x)'(x) \quad (\psi) \quad (\hat{\alpha}) = (x)'(y) \tag{7-5}$$

where

$$(\psi) = \begin{bmatrix} \psi_1(0) & \psi_2(0) & \cdots & \psi_Q(0) \\ \psi_1(1) & \psi_2(1) & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \psi_1(P) & \psi_2(P) & & \psi_Q(P) \end{bmatrix} \qquad (7-6)$$

and

$$(\hat{\alpha}) = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\alpha}_Q \end{bmatrix} \qquad (7-7)$$

In order to assure the existence of a unique solution for the $\hat{\alpha}_i$ the $\psi_i(p)$ must be linearly independent and $Q = P + 1$.

The matched filter technique using a periodic test signal can be adapted to measuring the coefficients of an orthogonal series expansion of the unknown impulse response by feeding the output signal of the matched filter into a spectrum analyzer. A simple spectrum analyzer consisting of a bank of narrowband band-pass filters may be used to obtain the Fourier series coefficients of the matched filter output signal.

The effects of external noise upon the measurement of orthogonal function coefficients has not been considered in detail. However, because such an expansion can be interpreted as a change in coordinates it does not seem likely that the identification time associated with the estimation of the coefficients will differ appreciably from the results obtained for estimating points of the impulse response function.

The extension of the results of this research to multidimensional systems, i.e., systems with several input signals and several output signals, can be achieved by using a suitable matrix notation. Identification of multidimensional systems by sampling techniques has been discussed by Woodrow [31].

The limits on the identification time that have been established by this work were obtained under the assumption that no a priori knowledge about the system was available. In practice the engineer usually has some knowledge of the properties of the system he is working with, even if it is only an estimate.

If an appreciable reduction in the identification time is possible it is felt that it will be obtained only if the available a priori knowledge about the system is used in an optimum manner. Unfortunately this statement raises more questions than it answers. What type of a priori knowledge about the unknown system will be most useful in reducing the identification time? Can the a priori knowledge be expressed in a useful mathematical manner? How is the a priori knowledge to be incorporated in the identification technique? All of these questions are at this point unanswered and should serve to stimulate future research in the area.

Several practical identification techniques that require a limited amount of a priori knowledge about the system, usually the order, have been proposed in the literature. Kalman has suggested an identification technique that estimates the coefficients of the numerator and denominator polynominals of the system's pulse transfer function [13, Kalman]. This method requires a priori knowledge about the order of the unknown system.

A number of identification techniques which employ a model of the physical system have been suggested. Margolis and Leondes [17] propose the use of a "learning model" for system identification, and Whitaker, et.al. discuss an adaptive flight control system employing a model of the system to be identified [28]. The general approach using the model technique is the following: if the order of the system to be measured is known, a model of the same order is chosen; if the order is not known, the engineer decides to represent the unknown system by an n^{th} order system where n is based upon some a priori knowledge about the system and perhaps a certain amount of engineering judgment. A block diagram of a typical identification technique employing a model is shown in Fig. 7-1. The difference between the output of the system under test and the output of the model is a measure of the degree of "goodness" for the model. When the model is an exact replica of the unknown system the error signal will be zero. A parameter adjustment computer adjusts the parameters of the model until some function of the error signal is satisfied. The nature of the parameter adjustment computer varies with the application.

The effects of external noise upon the identification time required by techniques which make use of some a priori knowledge have not been studied. An analysis of this problem would perhaps provide a clue to the savings in identification time that could be achieved by optimally utilizing a priori information about the system.

The aim of the identification techniques that have been presented in this work has been to obtain a complete description of the input-output relationships of a linear system. A very important and basic question arises at this point. In the applications, particularly in adaptive control applications, is a complete description of the system necessary?

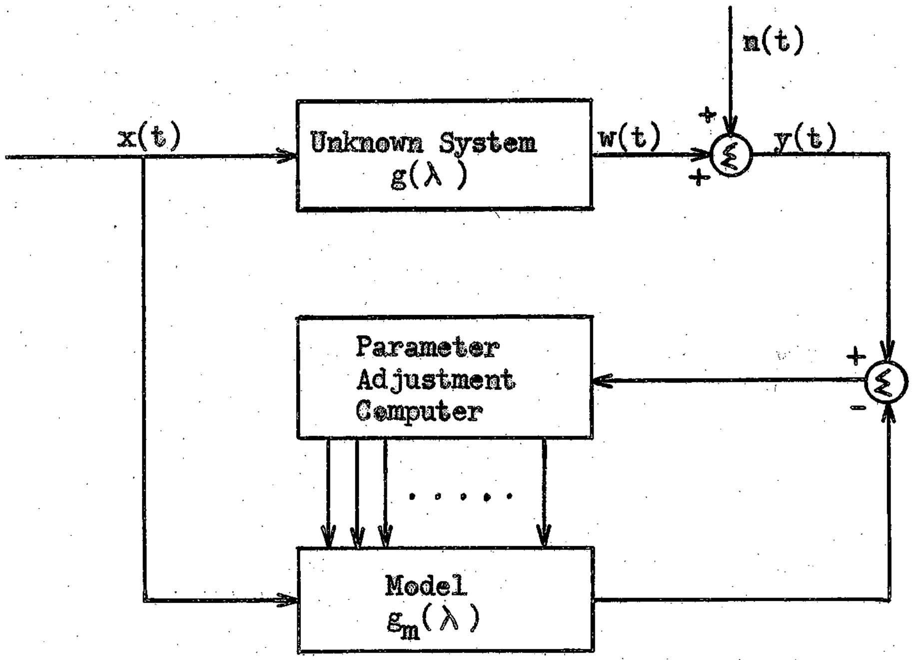


Fig. 7-1

Identification by Means of a Model

True, knowing the impulse response or transfer function enables the engineer to compute any other properties of the system he might desire, such as gain, rise time, or overshoot. Perhaps, however, it would be easier and faster to measure the other quantities directly. The general problem of identification time requirements and measurement techniques for partial identification of systems has not been investigated, and it is felt that this problem warrants attention. Insights gained from the consideration of partial identification of linear systems may open the way to solving the identification problem of rapidly time-varying or non-linear systems where complete identification becomes impractical.

BIBLIOGRAPHY

1. Anderson, G. W., Buland, R. N., and Cooper, G. R. "Use of Crosscorrelation in an Adaptive Control System", Proceedings of the National Electronics Conference, Vol. 15, October, 1959.

2. Aseltine, J. A., Mancini, A. R., and Sarture, C. W., "A Survey of Adaptive Control Systems", IRE Transactions PGAC - 6, December 1958.

3. Bello, P., "Joint Estimation of Delay, Doppler and Doppler Rate," IRE Transactions - PGIT, Vol. IT - 6, No. 3, June, 1960.

4. Braun, L., Jr. "On Adaptive Control Systems," IRE Transaction - PGAC, Vol. AC-4, No. 2, November, 1959.

5. Cooper, G. R., and Gibson, J. E., et al., "Survey of the Philosophy and State of the Art of Adaptive Systems," Technical Report No. 1, Contract AF 33(616) - 6890, PRF 2358, Purdue University, July 1960.

6. Courant, R., Differential and Integral Calculas, Vol. I, Interscience Publishers, Inc., New York, N. Y., 1937.

7. Cramer, H., Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., 1946.

8. Davenport, W. B., and Root, W. L., An Introduction to the Theory of Random Signals and Noise, McGraw Hill Book Company, Inc., New York, N. Y., 1958.

9. Doob, J. L., Stochastic Processes, John Wiley and Sons, Inc., New 1953.

10. Gardner, M. F., and Barnes, J. L., Transients in Linear Systems, John Wiley and Sons, Inc., New York, N. Y., 1942.

11. Guillemin, E. A., The Mathematics of Circuit Analysis, John Wiley and Sons, Inc., New York, N. Y., 1949.

12. Guillemin, E. A., Synthesis of Passive Networks, John Wiley and Sons, Inc., New York, N. Y., 1957.

13. Kalman, R. E., "Design of A Self-Optimizing Control System," ASME Transactions, Vol. 80, February, 1958.

14. Laning, J. H., and Battin, R. H., Random Processes In Automatic Control, McGraw-Hill Book Company, Inc., New York, N. Y., 1956.

15. Lee, Y. W., "Applications of Statistical Methods to Communications Problems," M.I.T. Research Laboratory for Electronics. Technical Report 181, September, 1951.

16. Levin, M. J., "Optimum Estimation of Impulse Response in the Presence of Noise," IRE Transactions - PGCT, Vol. CT - 7, No. 1, March 1960.

17. Margolis, M., and Leondes, C. T., "A Parameter Tracking Servo For Adaptive Control Systems," IRE Transactions - PGAC, Vol. AC-4, No. 2, November, 1959.

18. Middleton, D., "An Introduction to Statistical Communication Theory," McGraw-Hill Book Company, Inc., New York, N. Y., 1960.

19. Moore, E. F., "Gendonken Experiments on Sequential Machines," Automata Studies, Princeton University Press, Princeton, N. J., 1955.

20. Newton, G. C., Jr., Gould, L. A., and Kaiser, J. F., Analytical Design of Linear Feedback Controls, John Wiley and Sons, Inc., New York, N. Y., 1957.

21. Shannon, C. E., "Communication in the Presence of Noise," Proceedings of the IRE, Vol. 37, No. 1, January 1949.

22. Truxal, J. G., Automatic Feedback Control System Synthesis, McGraw-Hill Book Company, Inc., New York, N. Y., 1955.

23. Truxal, J. G., "Trends in Adaptive Control Systems," Proceedings of the National Electronics Conference, Vol. 15, October, 1959.

24. Tompkins, D. N., Jr., "Codes With Zero Correlation," Ph.D. Thesis, Purdue University, June, 1960.

25. Turin, G. L. "On the Estimation in the Presence of Noise of the Impulse Response of a Random Linear Filter," IRE Transactions - PGIT, Vol. IT-3, No. 1, March, 1957.

26. Turin, G. L. "An Introduction to Matched Filters," IRE Transactions - PGIT, Vol. IT - 6, No. 3, June, 1960.

27. W.A.D.D. Technical Report 60-201, "A Study to Determine the Feasibility of a Self-Optimizing Automatic Flight Control System", Aeronutronic, A Division of the Ford Motor Company, June, 1960.

28. Whitaker, H. P., Yamron, J., and Kezer, A., "Design of Model-Reference Adaptive Control Systems for Aircraft," Report R-164 M.I.T. Instrumentation Laboratory, September, 1958.

29. Westman, H. P., editor, Reference Data for Radio Engineers, International Telephone and Telegraph Corporation, New York, N. Y., 1956.

30. Woodrow, R. A., "Data Fitting with Linear Transfer Functions," Journal of Electronics and Control, Vol. VI, No. 5, May, 1959.

31. Woodrow, R. A., "On Finding a Best Linear Approximation to System Dynamics from Short Duration Samples of Operating Data," Journal of Electronics and Control, Vol. VII, No. 2, August, 1959.

32. Woodward, P. M., Probability and Information Theory, with Applications to Radar, McGraw-Hill Book Company, Inc., New York, N. Y., 1953.

33. Zadeh, L. A., "On the Identification Problem," IRE Transactions-PGCT, Vol. CT-3, No. 4, December, 1956.

ERRATA

Technical Report No. 2

Vol. II

Contract AF 33(616)-6890 PRF 2358

Page	Line	Should read	Instead of
8	Eq. (1-1)	$m \geq n$	$m \leq n$
8	14	$m \geq n$	$m \leq n$
17	Eq. (2-8)	$\left.\frac{\partial w}{\partial q_j}\right\rvert_{\{q_{j_a}\}}$	$\left.\frac{\partial w}{\partial q_j}\right\rvert_{\{q_{i_a}\}}$
27	1	CROSSCORRELATION	CROSSCORREOATION
27	11	CROSSCORRELATION	CORSSCORRELATION
52	bottom	function	functions
59	Eq.(5-7)	$e^{j(\lambda_1 - t + \Delta)\omega}$	$e^{j(\lambda_1 - t + \Delta)}$
60	Eq.(5-8)	$e^{j(\lambda_1 - t + \Delta)\omega}$	$e^{j(\lambda_1 - t + \Delta)}$
94	16	noise	noised
101	18	done	dome
106	above Eq.(7-4)	represented by	represented
106	below Eq.(7-4)	where	Where