2-1-1961

A Class of Predictive Adaptive Controls

John E. Gibson
Purdue University

J. S. Meditch
Purdue University

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

FINAL REPORT VOL. I
PROJECT 8225
TASK 82181

PURDUE UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

A Class of Predictive Adaptive Controls

J. E. Gibson, Principal Investigator

J. S. Meditch

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 I

CONTRACT AF 33(616) - 6890

Project 8225

Task 82181

PRF 2358

A CLASS OF PREDICTIVE ADAPTIVE CONTROLS

for

U. S. AIR FORCE

WRIGHT AIR DEVELOPMENT DIVISION

WRIGHT-PATTERSON AIR FORCE BASE

DAYTON, OHIO

by

J. E. Gibson, Principal Investigator

J. S. Meditch

School of Electrical Engineering

Purdue University

Lafayette, Indiana

February 1, 1961

PREFACE

This report is the first 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.

This volume presents the development and analysis of a particular class of adaptive controls under the assumption of the availability of identification information. The 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 and that covered by this report is the unnecessary restriction of the adjustment procedure to incremental or continuous adjustment of physical parameters. This is the parameter adjustment solution to the control signal modification problem. The more general procedure, discussed here, lies in control signal synthesis, 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 at Purdue is the identification problem. In Volume 2 of this final report Cooper and Lindenlaub report on their study of the speed and accuracy of various identification schemes which do not require a priori information concerning the plant.

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

CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF SYMBOLS

LIST OF SYMBOLS (continued)

		Page
$\delta(t)$	unit impulse function	24
$c_i(t)$	dynamic process response due to initial conditions	24
$u(t)$	external disturbance	24
ϵ	parameter independent of time	29
$m_o(t)$	optimum control variable	29
$m_\epsilon(t)$	arbitrary variation of m(t)	29
m_n	series coefficients	35
n	index on summation	35
N	upper index on summation	35
$p_n(t)$	arbitrary member of class of orthonormal polynomials	35
$[\quad]^*$	estimate of future value of function within brackets	40
$K_k(t)$	time-varying gain	40
$C_o(s)$	Laplace transform of $c_o(t)$	46
$C(s)$	Laplace transform of c(t)	46
$E(s)$	Laplace transform of system error	46
$G_c(s)$	controller transfer function	46
$W(s)$	dynamic process transfer function	46
$Y_n(s)$	Laplace transform of pre-multiplier output	48
$I_n(s)$	Laplace transform of reset integrator output	49
$R_n(s)$	Laplace transform of sampler output	49
$Q_n(s)$	Laplace transform of zero-order hold output	50
$M_n(s)$	Laplace transform of post-multiplier output	50
$\tilde{K}_o$	equivalent linear gain	54

LIST OF SYMBOLS (continued)

		Page
K	dynamic process parameter, constant	56
a	dynamic process parameter, variable	56
λ_o	a constant	
z	z-transform independent variable	57
e	base of natural logarithm system	57
$x(t + T)$	desired output of predictor	62
$y(t)$	actual output of predictor	62
$e_p(t)$	instantaneous prediction error	62
$\overline{e_p^2(t)}$	mean-square prediction error	62
f	a function	62
α_i	signal parameters	62
β_j	predictor parameters	62
T	prediction interval	62
L	a constant	62
g	a function	64
$\Phi(s)$	spectral density	65
$\approx$	"is approximately equal to"	67
Φ^+	factor of spectral density with its poles and zeros in upper half of complex plane	67
$H_{opt}(s)$	transfer function of optimum predictor	68
b	a constant	68
a	a constant	68
$\left.\overline{\epsilon^2(t)}\right\|_{min}$	minimum mean-square error	68
A	a constant	68

LIST OF SYMBOLS (continued)

		Page
ln	natural logarithm	70
α	a constant	72
T_1	a time constant	72
T_2	a time constant	72
R_1	a resistance	72
R_2	a resistance	72
C_1	a capacitance	72
A	a constant	73
B	a constant	73
γ	a constant	73
a(t)	dynamic process parameter	77
b(t)	dynamic process parameter	82
ω	radian frequency	92
R	remainder term	97
t'	independent variable, time	124
$P_i(t)$	transformed Legendre polynomials	124
$\phi_{xx}(\tau)$	autocorrelation function of x(t)	130
$\phi_{xz}(\tau)$	crosscorrelation function of x(t) and z(t)	130
$h_o(t)$	unit impulse response of optimum filter	130
x(t)	filter input	131
y(t)	filter output	131
h(t)	impulse response of filter	131
z(t)	desired output of filter	131
ϵ (t)	instantaneous error in filter output	131

LIST OF SYMBOLS (continued)

		Page
$F(s)$	complex Fourier transform of the time function	132
$\Phi_{xx}(s)$	complex Fourier transform of $\phi_{xx}(\tau)$	132
$\Phi_{xz}(s)$	complex Fourier transform of $\phi_{xz}(\tau)$	132
$\Phi^{+}_{xx}(s)$	factor of $\Phi_{xx}(s)$ with its poles and zeros in upper half of complex plane	132
$\Phi^{-}_{xx}(s)$	factor reamining in $\Phi_{xx}(s)$ after removal of $\Phi^{+}_{xx}(s)$	132
$H_{opt}(s)$	complex Fourier transform of $h_o(t)$	132
$\phi_{cc}(\tau)$	autocorrelation function of $c(t)$	133
$\Phi_{cc}(s)$	complex Fourier transform of $\phi_{cc}(\tau)$	133
w	complex variable $u + jv$	133
$\overline{\epsilon^2(t)}$	mean-square prediction error	134

ABSTRACT

A new class of control systems termed predictive adaptive controls is developed and the performance characteristics are investigated analytically and experimentally.

The concepts of signal prediction, interval control, and synthesis of the control variable by a sum of orthonormal polynomials in t are introduced and developed in relation to adaptive control. A modified least squares integral index of performance is formulated and used as the criterion for system optimization. Control of dynamic processes is subdivided into intervals of a specified length T and prediction is used to obtain estimates of future values of system error.

Minimization of the index of performance leads to a family of control laws which specify the structure of the controller. The resulting control configuration is optimum in a specific mathematical sense and is readily realizable with available physical components. The adaptive capability is achieved through time-varying gains which are specific functions of the unit impulse response of the dynamic process being controlled.

Predictor design is presented in terms of the classical Wiener-Lee theory, and a relationship for control interval length as a function of prediction accuracy is developed.

Preliminary design of the controller is considered from the viewpoints of relative weighting of system error and control effort, control interval length T, and the number of terms needed in the orthonormal polynomial sum approximation of the control variable. A method of obtaining an engineering estimate of the latter quantity is developed and illustrated by three examples, two of which are investigated experimentally.

Two applications of predictive adaptive control are investigated on an analog computer. The two dynamic processes used are a first-order process whose parameter varies over a range of ten to one and a second-order process whose parameter varies in such a manner that the process is unstable at one extremum and heavily damped at the other. The results of three basic experiments which evaluate the steady-state adaptability, transient response, and statistical signal response of the two systems are reported. It is found that all three aspects of system performance improve with decreasing control interval length, but that the minimum value of the interval length which can be used is limited by the accuracy of the time-varying gain and controller circuitry. Improved performance which can be achieved by increasing the relative weighting of system error and control effort, is limited by saturation considerations. Theoretical results that point to the need for keeping the control interval length short to preserve stability, prediction accuracy, and loss of control due to process parameter drift are substantiated by the experimental results. For the two systems investigated it is found that satisfactory control is achieved if the interval length is chosen so that process parameter drift is no more than 4% per control interval. A figure of 5% was estimated originally.

A one-term approximation of the control variable is used to control the first-order process and is found to give satisfactory performance. A four-term approximation is found to give adequate control of the second-order process whereas the three-term approximation does not. These results bear out the predictions made in the theoretical analyses.

CHAPTER 1

INTRODUCTION

The need for precise control of dynamic processes has stimulated interest in the development of theories and methods for optimizing control systems. Hazen [1] in 1934 and Hall [2] in 1943 initiated what is today termed the conventional design of feedback control systems. Their work was followed by that of Wiener [3] in 1949 which forms the foundation of classical analytical design theory of optimum controls. As originally formulated, Wiener's methods are applicable only to linear, time-invariant dynamic processes which are to be optimized with respect to a least-squares figure of merit or performance index. Usually, the optimization amounts to specifying a compensation scheme which maximizes, minimizes, or gives a particular value to the specified index of performance. Booton [4], in 1952, extended Wiener's work further by using ensemble averages instead of time averages. His results permit the optimization of linear, time-varying dynamic processes subjected to stochastic signals possessing either time-invariant or time-varying statistics. This is in contrast to Newton's [5] methods which are restricted to time-invariant dynamic processes with deterministic and/or stochastic signals having time-invariant statistics.

Mathews and Steeg [6], and Booton [7], in 1956, studied the response characteristics of terminal, or final-value controls. Their work presents the analytical design of a class of non-linear systems but is restrictive because only one point of the response, the terminal point, is considered.

1.1 Adaptive Controls

More recently, considerable interest has centered about a new class of control systems termed adaptive [8] or self-adaptive controls [9].

The extent of this interest is indicated in a lengthy bibliography on the subject compiled by Stromer [10].

An adaptive control system is defined here as a control system which is capable of monitoring its own performance with respect to a given index of performance or optimum condition and modifying its behavior by closed-loop action in such a manner as to optimize the index of performance or approach the optimum condition. The necessity of such systems is apparent in the control of dynamic processes whose operating characteristics vary over a wide range during normal operation. For example, such dynamic processes as high-speed aircraft, space vehicles, and chemical plants experience wide variations in their environments throughout their course of operation. This places heavy demands on their control systems which cannot be met in a completely satisfactory manner by conventional controllers. The reason for this is clear when one recalls that conventional designs are based on satisfying one or more design criteria assuming the dynamic process is linear and time-invariant throughout its performance envelope. An example in point here is the minimization of the integral-square-error of a positional control system for a ramp input subject to a constraint on the mean-square noise power in the output. At best, this problem could be treated by conventional methods only if a complete knowledge of the time-invariant or time-varying character of both the fixed elements and the signals is available *a priori*. Unfortunately, the dynamic processes mentioned above are called upon to function in environments which are at most only partially known *a priori*. Hence, the information needed to effect a conventional design for such a process is not available until the process has begun functioning. As a result, the use of control systems, capable of *monitoring*, *evaluating*, and *modifying* their performance to meet the demands of control dictated by a changing environment is mandatory for such dynamic processes. Moreover, as a result of changing environment, the goal or task of the control system may change and the weighting of system error may become more or less important.

In summary, example situations where the use of adaptive control is warranted may be classified broadly as follows:

1. The characterization of the dynamic process is an unknown function of the environment to which the dynamic process is subjected.
2. The goal or task of the control system changes with environment. For example, the task of a chemical process controller during normal operation is to maintain such process parameters as temperatures, pressures, flow rates, and product qualities at their desired values. On the other hand, during startup the controller must change the process variables as rapidly as possible to achieve the desired steady-state.
3. The index of performance used to evaluate the performance of the dynamic process changes with time. For example, small deviations from the desired trajectory of a ballistic missile must be weighted more heavily during the terminal phase of the trajectory than they are during the earlier phases of the flight path.

1.2 Statement of the Adaptive Control Problem

The definition of an adaptive control system implies three functions which the system must be capable of performing [11]:

1. Provide information about the character of the dynamic process, i.e., <u>identify</u> the dynamic process.
2. Evaluate the performance of the dynamic process with respect to an index of performance and make a <u>decision</u> on how to achieve optimum performance.

3. Initiate modification of signals and/or dynamic process parameters in order to realize optimum performance.

Hence, the general problem of adaptive control divides logically into three basic problems: identification, decision, and modification. Each of these basic problems in itself represents a complete area of research. However, any research effort concerned with one of these can proceed logically only if the other two aspects of the over-all problem are kept in mind. A block diagram depicting the subdivision of the adaptive control problem into its three logical phases is shown in Fig. 1-1.

This research is concerned with a new method for achieving modification assuming that identification information is available continuously, and that an index of performance has been specified. The index of performance to be used in this research is formulated in Chapter 2.

Since this research deals with an approach to the modification problem, only a summary of the salient features of the identification and decision phases of the over-all problem is given here.

1.3 The Identification Problem

The identification problem is the problem of obtaining a description of the relationship between the input m(t) and the output c(t) of an unknown dynamic process as shown in Fig. 1-2. Mathematically, the problem is one of determining the functional transformation G between the variables m(t) and c(t) given by

$$c(t) = G\left[m(t)\right] \qquad (1\text{-}1)$$

where t is the independent variable, time.

Two basic requirements of any identification procedure are:

1. It must perform the identification function without excessively disturbing normal operation of the dynamic process.

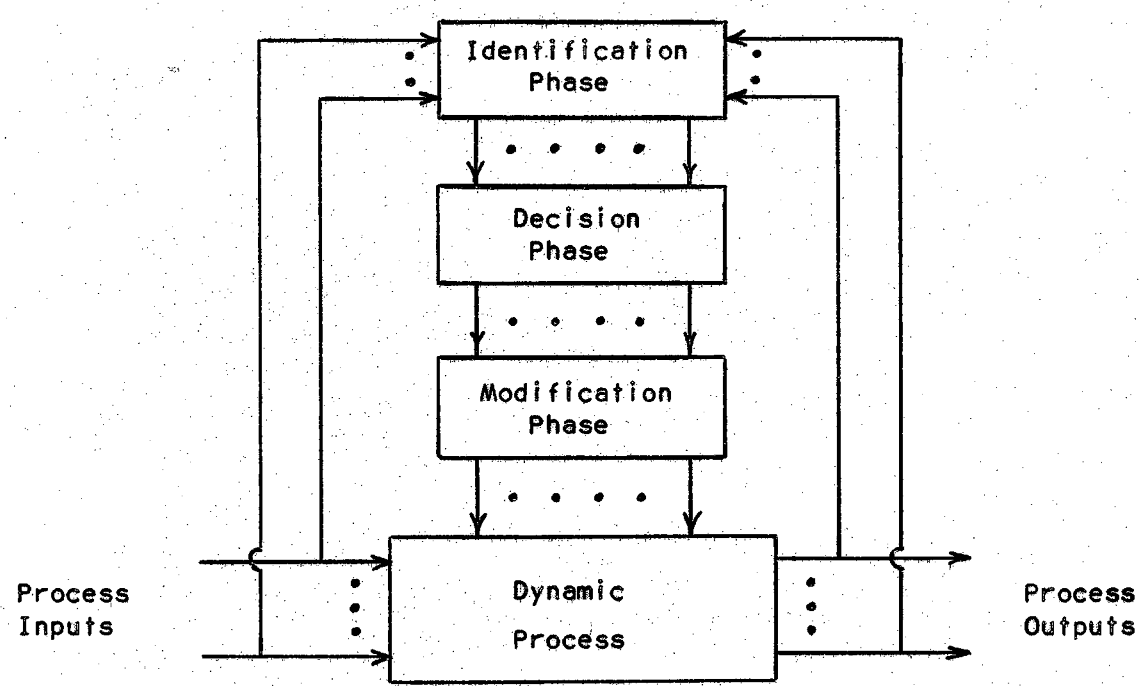


Fig. 1-1

Block Diagram Formulation of Adaptive Control Problem.

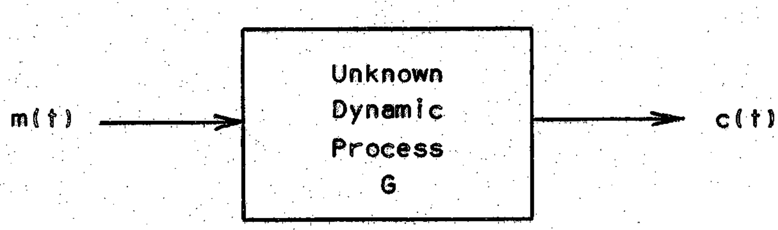


Fig. 1-2

Unknown Dynamic Process To Be Determined by Identification Procedures.

2. It must perform the identification function in an interval of time comparable to the interval of time for which the significant identification information is valid.

Both requirements are essential in order to perform adaptation continuously without recourse to halting system operation, taking measurements, and spending considerable effort in computation in order to obtain identification information.

For the impulse response representation the functional transformation G of Eq. 1-1 assumes the form

$$G = \frac{a_m p^m + a_{m-1} p^{m-1} + \dots + a_1 p + a_o}{b_n p^n + b_{n-1} p^{n-1} + \dots + b_1 p + b_o} \qquad (1\text{-}2)$$

where p is the operator $\frac{d}{dt}$, $n \geq m$ is required for physical realizability, and the a_i and b_i are constants or slowly varying functions of time t.

Various identification schemes have been investigated by Braun [12], Kalman [13], Turin [14], and Joseph, et al. [15]. These methods will not be reviewed here because they are not relevant to the work which follows. However, the approach to the identification problem given by Levin [16] could be used with the solution of the modification problem given in this research to form a complete adaptive control system. Levin's procedure for identification is outlined below.

The method proposed by Levin employs sampling of the input and output signals of the dynamic process, and requires no special test signal at the input to the process being identified. This latter property permits identification of dynamic processes within control loops, a feature which is needed in the adaptive controls developed in this research. Since the procedure can be repeated periodically, it is applicable to linear, slowly time-varying processes. The scheme is similar to cross-correlation [17] since the result of each is a set of sample points of the unit impulse response of the dynamic process.

The model assumed is indicated in Fig. 1-3. The process input is denoted by m(n) and the resulting output by c(n) where n denotes the number of the sampling instant. The sampling instants are assumed to be separated by some time interval t_a so that the nth sampling instant corresponds to time $t = nt_a$. In order to develop a realistic identification procedure, Levin assumed the presence of uncertainty in the measured output. This is denoted by the disturbance u(n) which is assumed to be a stationary, Gaussian, white noise signal with zero mean.

In the discrete formulation, the sequence of output values of the assumed model becomes

$$c(n) = \sum_{p=o}^{\infty} w(p)\ m(n-p) + u(n) \qquad (1-3)$$

for $n \geq p$.

Physically only a finite number of the w(p) can be determined and, hence, the impulse response is approximated by a finite set of values, w(0), w(1), ... , w(P) where P is chosen such that $w(p) \approx 0$ for $p > P$. This approximation is usually valid for most physical systems.

A typical set of input and output observations are shown in Fig. 1-4 to indicate the relation of one to the other. The following assumptions were made by Levin in the derivation of the set of algebraic equations whose solution gives the w(n):

1. $w(p) = 0$ for $p > P$ for some $P > 0$.

2. m(n) is observed for $0 \leq n \leq N$ and is not identically zero in this interval.

3. c(n) is observed for $0 \leq n \leq N + P$.

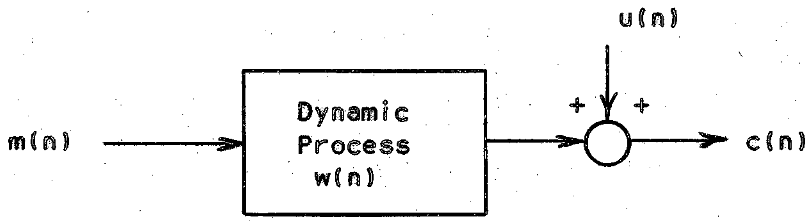


Fig. 1-3

Model for Dynamic Process Identification.

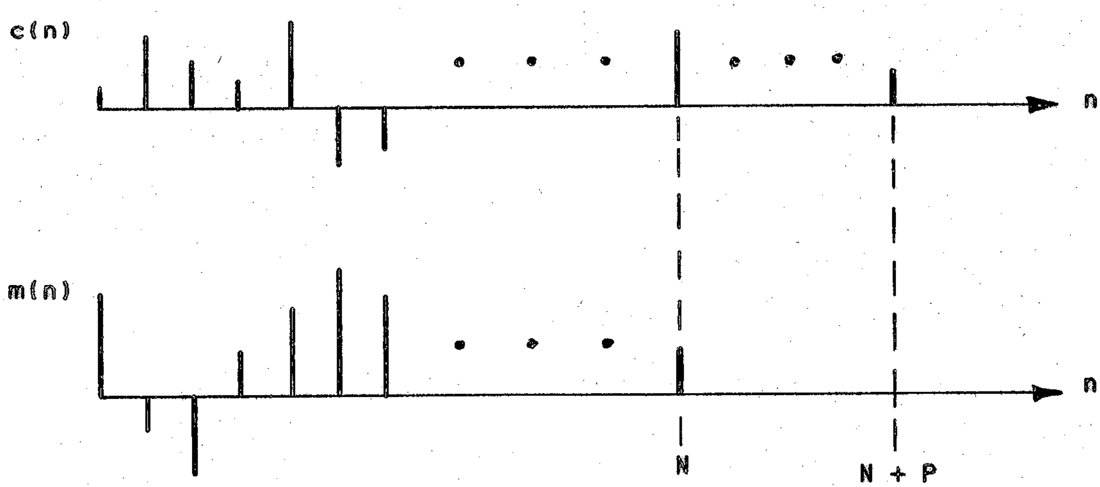


Fig. 1-4

Set of Input-Output Observations.

The results of Levin's derivation may be summarized readily if the following matrix notation is employed:

$$[c] = \begin{bmatrix} c(P) \\ c(P+1) \\ \vdots \\ c(P+N) \end{bmatrix} \qquad (1\text{-}4)$$

$$[w^*] = \begin{bmatrix} w^*(0) \\ w^*(1) \\ \vdots \\ w^*(P) \end{bmatrix} \qquad (1\text{-}5)$$

and

$$[m] = \begin{bmatrix} m(P) & m(P+1) & \cdots & m(P+N) \\ m(P-1) & m(P) & \cdots & m(P+N-1) \\ \vdots & \vdots & \cdots & \vdots \\ m(0) & m(1) & \cdots & m(N) \end{bmatrix} \qquad (1\text{-}6)$$

where $w^*(n)$ is the best mean-square estimate of $w(n)$, the latter being the exact value of the impulse response $w(t)$ at time $t = nt_a$.

Then according to Levin's development, the $w^*(n)$ satisfy the normal equations

$$[m]\,[m]'\,[w^*] = [m]\,[c] \qquad (1\text{-}7)$$

where $[m]'$ is the transpose of the matrix $[m]$.

A small special-purpose digital computer could be designed and pre-programmed to solve the set of equations indicated in Eq. 1-7. This information could then be utilized by the modification portion of the adaptive system.

1.4 The Decision Problem

The decision problem deals with the development and specification of analytical methods by which dynamic process performance can be evaluated and from which a strategy to achieve adaptation can be evolved. The most common method of process evaluation utilizes the notion of an index of performance. An index of performance is defined as a functional relationship involving dynamic process characteristics in such a manner that the optimum operating characteristics can be determined from it.

Numerous indices of performance have been treated in the literature [18, 19, 20]. Hence, only the concepts which underlie the index of performance to be developed in Chapter 2 and used in this research are given here.

The most common indices of performance used in present day control technology are those which employ some arbitrary function of system error. In this context system error is defined to be the difference between the desired value of the process state and the actual value of the process state. Symbolically,

$$I = F\left[e(t)\right] \qquad (1\text{-}8)$$

where I = index of performance

$e(t)$ = system error

F = some arbitrary functional operation.

In applying the concepts of dynamic programming to the optimization of control processes, Bellman [21] postulated a rather broad class of indices of performance in terms of cost functions. Consider the dynamic process shown in Fig. 1-5 and let the state of the process be characterized by a vector* $\bar{c}(t)$ and let $\bar{m}(t)$ be the input or control vector. Further,

---

*A vector as used here is defined as a column matrix.

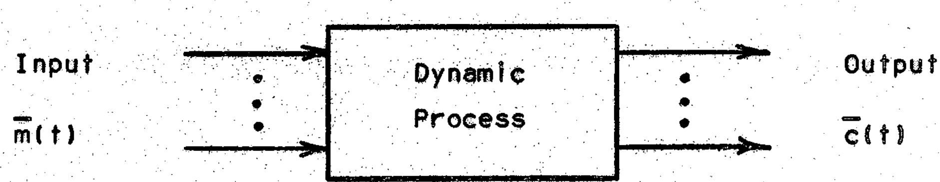


Fig. 1-5

Multi-dimensional Dynamic Process.

let $c_o(t)$ represent the desired state of the process, $G\left[\bar{c}_o(t) - \bar{c}(t)\right]$ be a function measuring the cost of deviation of $\bar{c}(t)$ from $\bar{c}_o(t)$, and $H\left[\bar{m}(t)\right]$ be a function measuring the cost of control. Then the total cost function or index of performance, denoted $J\left[\bar{c}(t), \bar{m}(t)\right]$, becomes

$$J\left[\bar{c}(t), \bar{m}(t)\right] = G\left[\bar{c}_o(t) - \bar{c}(t)\right] + H\left[\bar{m}(t)\right] \quad (1\text{-}9)$$

Observe that the total cost function consists of two parts. The first is actually a measure of system error as discussed earlier, while the second is a measure of the amount of control effort exerted in driving the process from its present state to the desired state. While dynamic programming concepts are not used in this research, the formulation of Eq. 1-9 and its interpretation as a compounded cost function is basic for the work to follow.

1.5 The Modification Problem

After the identification and decision problems have been solved, the adaptive loop must adjust or modify the dynamic process to bring it to the desired state. Modification is usually based on the following information:

1. The desired state of the dynamic process.
2. The present state of the dynamic process.
3. The character of the input-output relationship of the dynamic process.
4. The index of performance chosen as the measure of system performance.

Conceptually, the modification phase of the adaptive control problem may be viewed as computer control of the dynamic process as shown in Fig. 1-6. Typical operations which might be required of the computer controller include evaluation of the index of performance, generation of control signals for the adjustment of parameters, and/or generation of new signals to be applied directly to the input of the dynamic process.

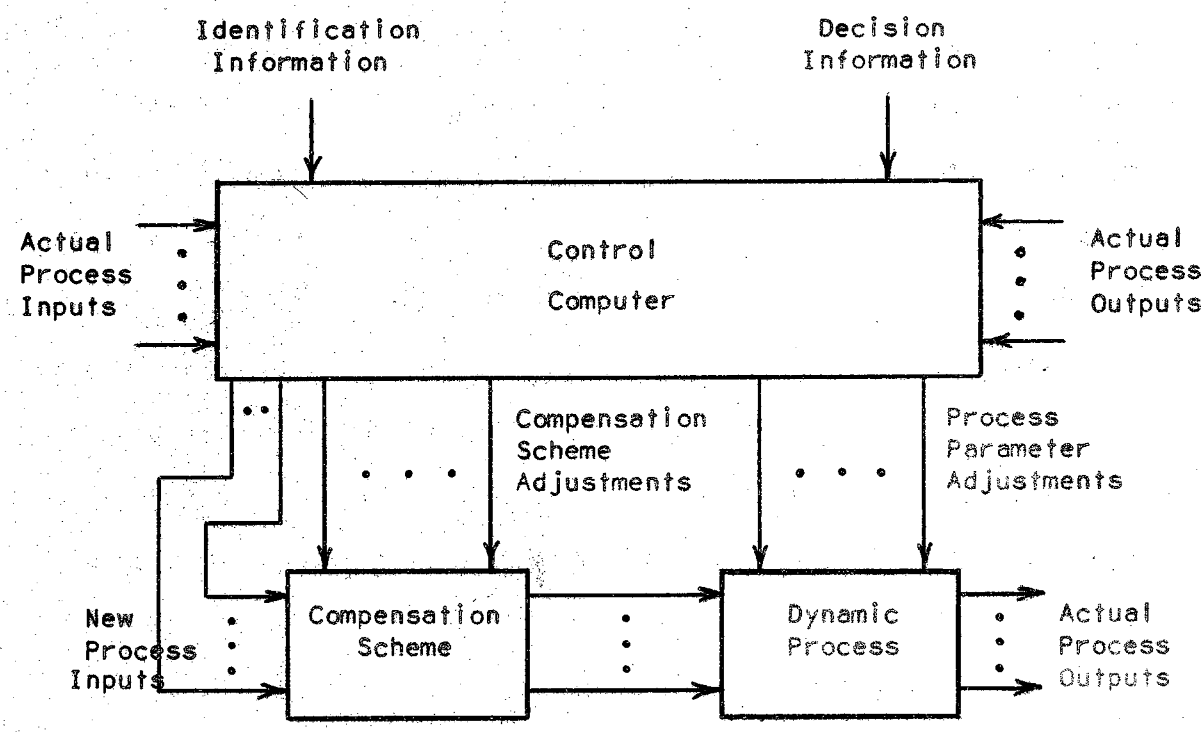


Fig. 1-6

Block Diagram Formulation of Modification Problem.

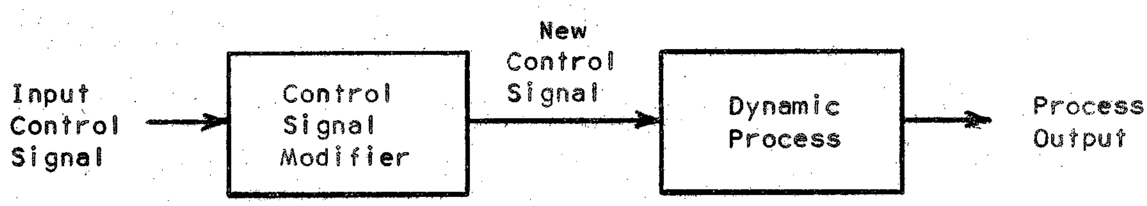


Fig. 1-7

Control Signal Modification.

The present approach to the modification problem utilizes a concept called control signal modification as shown in Fig. 1-7. Control signal modification is defined as the application of linear, time-varying and/or nonlinear operations on the actual system input to derive a control signal which actuates the dynamic process. This approach lends itself to two interpretations which are termed parameter adjustment and control signal synthesis.

Parameter Adjustment. This method performs modification by adjusting the parameters of the dynamic process and/or a compensation scheme to satisfy the index of performance. See Fig. 1-8. Since the control requirements vary with time due to changes in process dynamics and process signals, the adjustment of the parameters is a time-varying operation. Clearly this approach achieves modification by direct recourse to the shaping of the dynamic process transient response. The work of Anderson, et al. [22] is one of the more interesting applications of the parameter adjustment method. The system, which is shown in Fig. 1-9, utilizes the impulse-response-area ratio as the index of performance. The technique provides a means for the system to adjust its parameters for optimum dynamic response by using a null-type index of performance.

The parameter adjustment approach modifies the control signal indirectly by manipulating the parameters of the elements employed in the over-all system. Its primary function is to shape the dynamic process transient response in accordance with the dictates of the index of performance.

Control Signal Synthesis. Rather than modify the control signal indirectly, this approach utilizes the identification and decision information to synthesize a new control signal which is then used to actuate the dynamic process. The scheme is shown in block diagram for in Fig. 1-10.

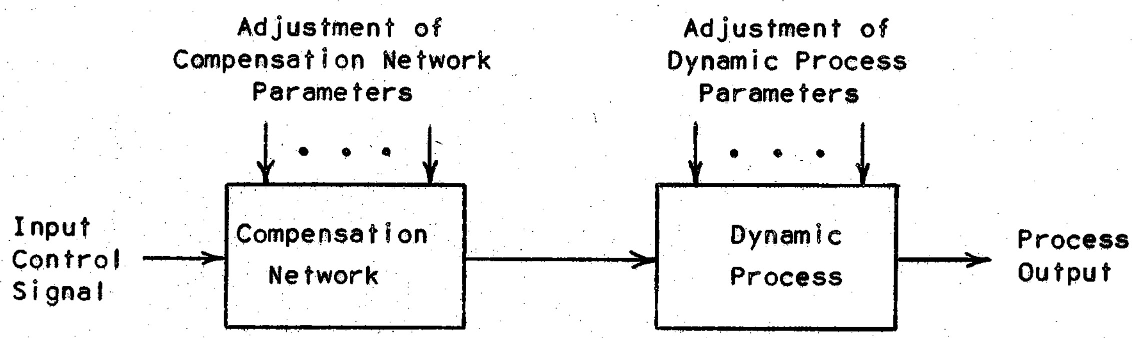


Fig. 1-8

Control Signal Modification by Parameter Adjustments.

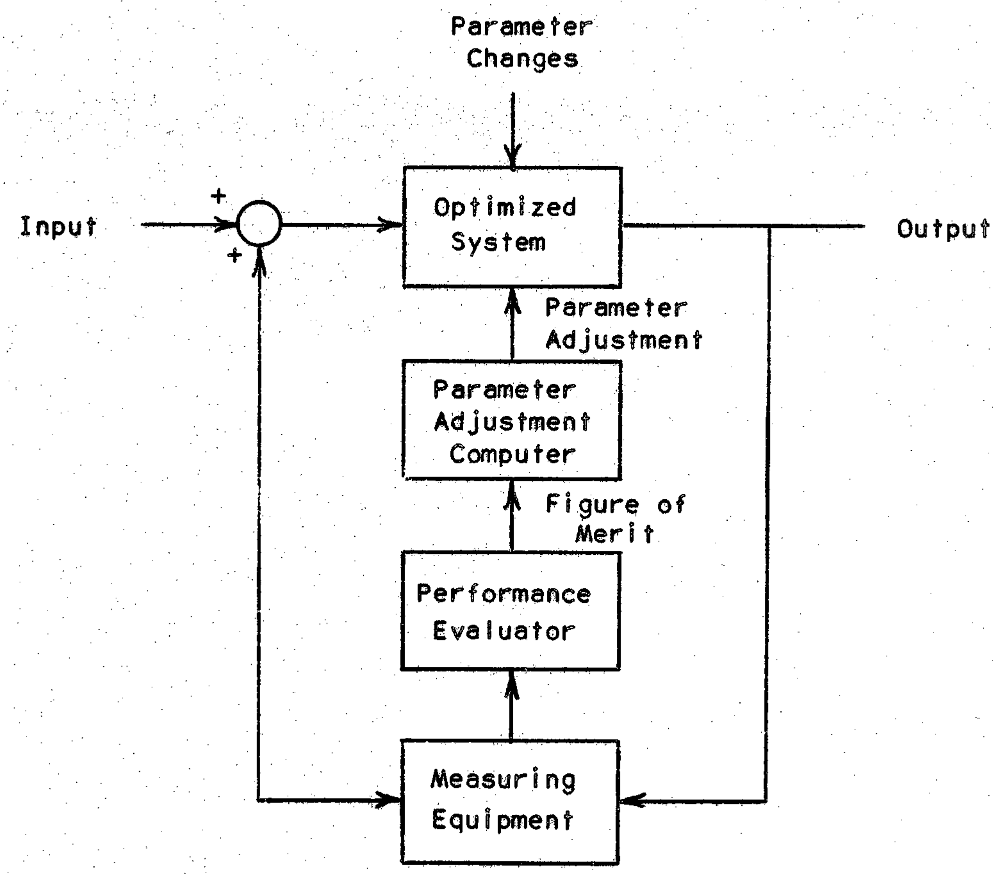


Fig. 1-9

Parameter Adjustment Scheme of Anderson, et al. [22].

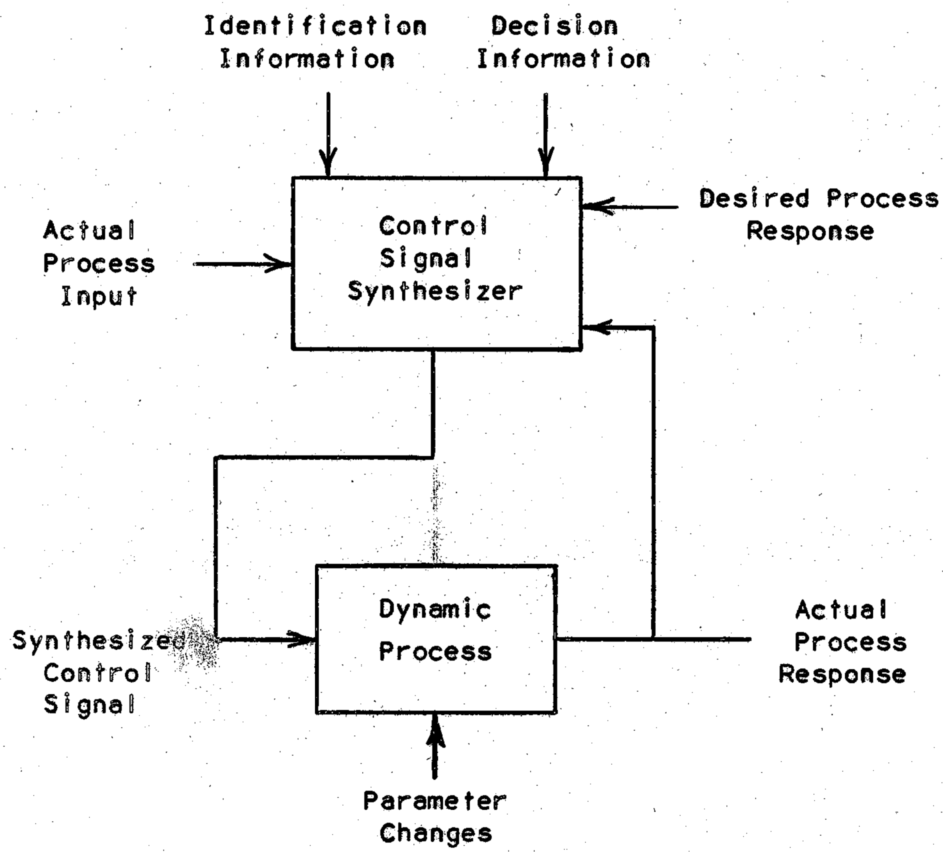


Fig. 1-10

Control Signal Synthesis Approach to the Modification Problem.

Basically, this approach is concerned with obtaining an optimum approximation to the desired response by operating on the information contained in the index of performance to derive the actuating signal.

In certain applications it may be impossible to alter the character of the dynamic process or of a compensation scheme in order to achieve optimum operation. This situation will arise in those cases where process parameters must be controlled indirectly because the process has no physical adjustments available.

In summary, both approaches are concerned with altering the nature of the control signal which actuates the dynamic process being controlled. However, the first method achieves this goal indirectly by acting through the adjustable system parameters, whereas the second does it directly by creating a new control signal. While the parameter adjustment method shapes transient response directly, the control signal synthesis scheme treats it indirectly since the process impulse response will invariably appear in the formulation of the index of performance. In a sense, the two approaches are similar with the roles of transient response shaping and control signal generation inter-changed. However, it is useful to separate the two in an operational sense.

1.6 Research Objectives

The first objective of this research is to develop a new class of adaptive controls. The ultimate result will be a control configuration which is optimum in a specific mathematical sense and is readily realizable with available physical components. The concepts of prediction and interval control, which are defined in Chapter 2, will be employed to achieve this objective. The focal point of the first research objective is modification by control signal synthesis.

The second objective of this work is to evaluate the performance characteristics of this new class of adaptive controls. Analytical and experimental methods are employed to achieve this second objective. The results of the two methods are compared and used to evaluate the class of adaptive controls developed.

CHAPTER 2

DEVELOPMENT OF THE MODIFICATION PROBLEM

The purpose of this chapter is to develop the modification problem in terms of the concepts of prediction and interval control, and to formulate the index of performance to be used in this research.

2.1 Prediction in Adaptive Control

A number of researchers [23, 24] have investigated the use of prediction in conventional communication and control systems with reasonable success. It is to be expected, then, that the incorporation of prediction in adaptive controls might aid the over-all system in combating erratic and undesirable behavior in the dynamic process. By anticipating wide variations of the actual response from the desired response, the adaptive loop is given "lead" time to synthesize, with the aid of a specified index of performance, the control signal which will offset the effects of these variations. Hence, prediction appears to be a desirable feature in adaptive controls.

2.2 Concept of Interval Control

Prediction must be based on the past history of the function being predicted. Also, well-known results from prediction theory [25] indicate that prediction accuracy deteriorates with an increasing prediction interval length. Hence, a finite prediction interval length T must be used to maintain a specified prediction accuracy.

Because of this prediction requirement, a reasonable engineering approach to the modification problem is to divide the process control into intervals of length T as shown in Fig. 2-1. Then, information gathered during the interval $-T \leq t \leq 0$ can be used to achieve optimum control over the interval $0 \leq t \leq T$. By letting $t = 0$ be the present

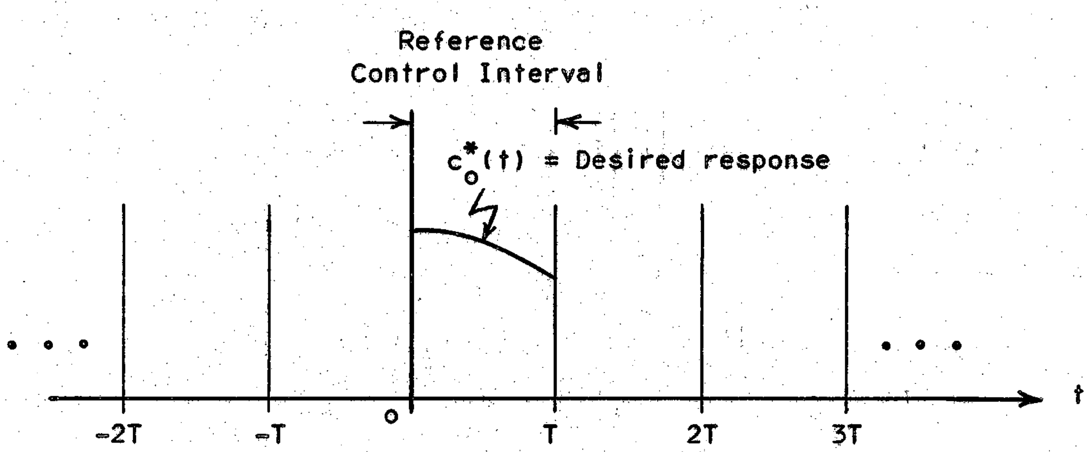


Fig. 2-1

Subdivision of Process Control into Intervals of Length T.

time, the interval $0 \leq t \leq T$ may be chosen as the reference interval over which the process is to be optimized. Hence, with respect to actual system time, the point $t = 0$ corresponds to the beginning of a control interval of length T into the future. By using a fixed prediction interval length T and operating on data as they occur in the interval $-T \leq t \leq 0$, a prediction of the desired response and actual response of the dynamic process for the interval $0 \leq t \leq T$ can be obtained during the former interval. This result then permits the adaptive loop to take action at $t = 0$ to optimize dynamic process performance during the reference control interval $0 \leq t \leq T$.

This subdivision of the optimization into intervals will permit the use of the classical z-transform method [26] to analyze certain response characteristics of the class of controls developed.

2.3 Formulation of an Index of Performance

Consider the single-dimensional dynamic process shown in Fig. 2-2 having the input variable $m(t)$, the output variable $c(t)$, and external disturbance $u(t)$, and the unit impulse response $w(t,\tau)$ which is time-varying as a function of environment E. The unit impulse response $w(t,\tau)$ is defined here as the response of the dynamic process at time t to an impulse applied at time τ. A modified least squares index of performance will be formulated for this process by considering an interval of length T in the future, where $t = 0$ is taken as the present time. It will differ from conventional least squares indices of performance in the following ways:

1. The process will be optimized over a future interval of time T and no errors before $t = 0$ will be weighted.
2. Provision will be made for unequal weighting of system error during the control interval.

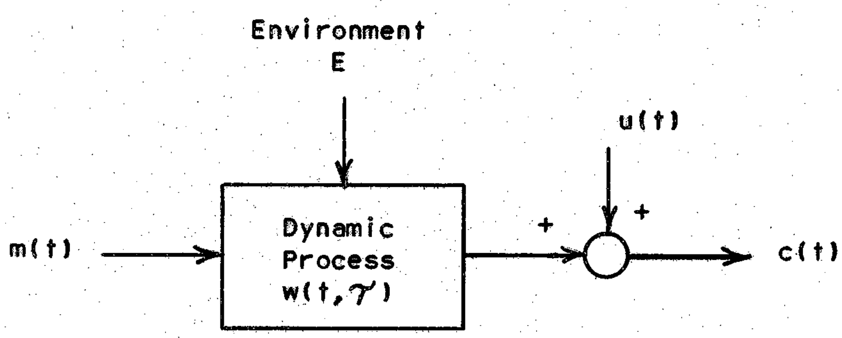


Fig. 2-2

Dynamic Process with Time-varying Impulse Response w(t,τ).

3. The existence of a model characterizing the desired input-output transformation of the dynamic process will be assumed.

4. Prediction will be used to establish the future values of the desired response.

5. The control variable m(t) will be manipulated in the present (t = 0) to optimize process response in the future.

The reason for not weighting system errors in the past is because no control can be effected in the present or future to reduce these errors.

In attempting to optimize the dynamic process of Fig. 2-2 over all time, the classical index of performance is the integral-square-error given by

$$I = \int_{-\infty}^{\infty} \left[c_o(t) - c(t)\right]^2 dt \tag{2-1}$$

where

$c_o(t)$ = desired process response

$c(t)$ = actual process response

t = dummy variable of integration, time.

However, since optimization is to be executed on a per interval basis and $c_o(t)$ is available only for $0 \leq t \leq T$, Eq. 2-1 becomes

$$I = \int_{0}^{T} \left[c_o(t) - c(t)\right]^2 dt \tag{2-2}$$

In order to provide for unequal weighting of response errors over the control interval, an arbitrary weighting factor $\lambda(t)$ is introduced into the integrand of Eq. 2-2 to give

$$I = \int_{0}^{T} \lambda(t) \left[c_o(t) - c(t)\right]^2 dt \tag{2-3}$$

This weighting factor is obtained from engineering considerations based on the goals or objectives of control. For example, if response errors are important only at the end of each control interval, the choice $\lambda(t) = \delta(t - T)$ is made where $\delta(t)$ is the unit impulse function. If equal weighting is to be given to response errors, then $\lambda(t)$ is simply a constant. An important restriction on $\lambda(t)$ which is necessary to give meaningful engineering results is $\lambda(t) > 0$ for $0 \leq t \leq T$.

Finally, a cost term accounting for the amount of control resources utilized to achieve modification is added to Eq. 2-3 to give

$$I = \int_0^T \lambda(t) \left[c_o(t) - c(t)\right]^2 dt + \int_0^T m^2(t)\, dt \qquad (2\text{-}4)$$

Clearly, Eq. 2-4 is a member of the general class of indices of performance defined by Bellman in Eq. 1-15.

The actual response c(t) of Eq. 2-4 is comprised of three components. The first is due to the initial energy stored in the dynamic process at t = 0 and accounts for excitations prior to t = 0. This term is denoted by $c_i(t)$. The second component of c(t) is that due to the disturbance u(t) and the third is caused by the new excitation m(t), $0 \leq t \leq T$, and is given by the convolution integral

$$\int_0^t m(\tau)\, w(t,\tau)\, d\tau \qquad (2\text{-}5)$$

where τ is the dummy variable of integration. Hence, the actual output c(t) is given by

$$c(t) = c_i(t) + u(t) + \int_0^t m(\tau)\, w(t,\tau)\, d\tau \qquad (2\text{-}6)$$

Substitution of Eq. 2-6 into Eq. 2-4 yields the final form of the index of performance,

$$I = \int_0^T \left\{\lambda(t) \left[c_o(t) - c_i(t) - u(t) - \int_0^t m(\tau)\, w(t,\tau)\, d\tau\right]^2 + m^2(t)\right\} dt \qquad (2\text{-}7)$$

where:

$c_o(t)$ = desired process response during the control interval.

$c_i(t)$ = component of process response during the control interval due to initial conditions at the beginning of the control interval.

$u(t)$ = component of process response during control interval due to external disturbance.

$m(t)$ = process input control variable to be chosen to minimize Eq. 2-7.

$w(t,\tau)$ = process unit impulse response for the control interval.

$\lambda(t)$ = arbitrary system error weighting factor.

Eq. 2-7 is an index of performance comprised of two cost functions. The first term represents a measure of the deviation of the actual dynamic process response from the desired dynamic response. On the other hand, the second term measures the amount of control effort which is exerted. The weighting factor $\lambda(t)$ provides considerable flexibility which is not a property of most integral indices of performance. Not only does it provide for unequal weighting of response errors on the control interval, but it also permits a relative weighting between the two terms of the index of performance. Moreover, depending on the control situations to be encountered, a judicious choice of $\lambda(t)$ will provide response superior to that of conventional indices of performance. As a result, the presence of $\lambda(t)$ provides the design engineer with considerable latitude in seeking optimum designs.

2.4 The Complete System

The complete modification problem as developed in this chapter may be visualized in block diagram form as shown in Fig. 2-3. The function of the control unit and signal synthesizer is to utilize the indicated input information to generate the optimum control signal m(t). Clearly, such a task could be accomplished by a large digital computer. However,

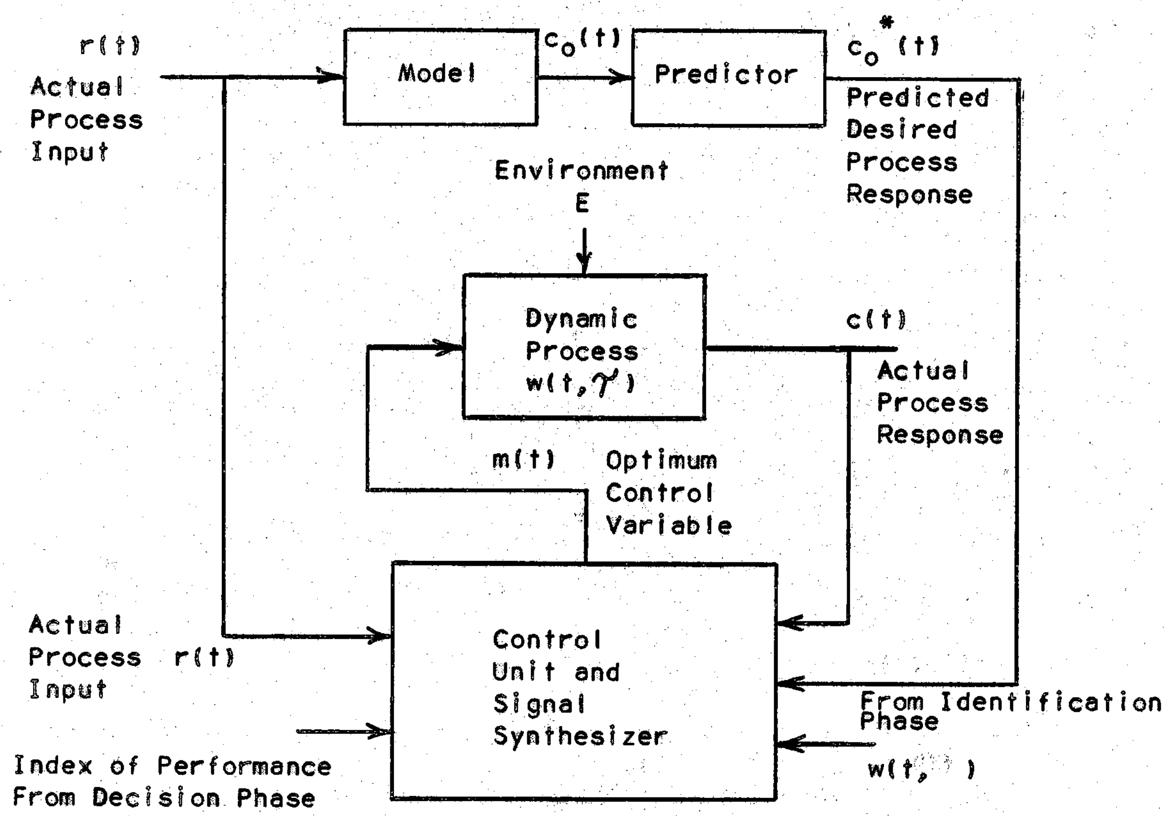


Fig. 2-3

Block Diagram of Complete System for Performing Modification Function.

engineering factors such as size, weight, and cost often dictate the need for small, special-purpose computers to perform the control task. One of the objectives of this research is to develop a class of adaptive controls which can be realized readily from physical components. This objective requires keeping system complexity at a minimum.

In order to keep complexity and cost at a minimum, operation of the control unit in real-time is highly desirable. If real-time operation can be achieved, there will be no need for high-speed computing devices with their inherently complex input-output accessory equipment. Computations in the control unit could then be performed by analog components, e.g., multipliers, integrators, summing amplifiers, and diode function generators operating at the same rate as the dynamic process.

2.5 General Considerations

Basically, the entire adaptive control process as developed here may be viewed as a sequence of decisions to be made every T units of time. This decision for each interval T is based upon the present state of the dynamic process being controlled and upon the desired behavior of that process over a future interval of time as obtained from a prediction operation.

Since all possible control signals m(t) are not acceptable because of physical limitations imposed on the control problem, the actual response of the dynamic process cannot, in general, be expected to agree exactly with the desired response. Hence, an index of performance was developed to be used in selecting the optimum member from the class of acceptable control signals.

CHAPTER 3

THE OPTIMIZATION PROBLEM

The optimization problem is concerned with the selection of a physically realizable control variable m(t) which will minimize the index of performance, Eq. 2-7. In other words, the problem of determining optimum control deals with the minimization of a particular integral over a fixed interval.

A number of minimization techniques are presented here as background material for the work which is to follow. Another purpose of this chapter is to point out the computational difficulties which arise when optimization of adaptive controls is considered. The minimization techniques treated are:

1. Calculus of variations.
2. Approximation of m(t) by discrete segments.
3. Approximation of m(t) by a sum of orthonormal polynomials.

In order to simplify the mathematics and still indicate the concepts underlying the first two approaches, let the disturbance $u(t) = 0$ in Eq. 2-7.

3.1 Calculus of Variations

A fundamental problem in the calculus of variations is to determine a function such that a particular definite integral involving that function and certain of its derivatives assumes a maximum or a minimum value [27]. The application of this mathematical tool to the optimization of control systems was a major step in the development of analytical control theory as shown in Newton [5, p. 143].

The application of this technique will be considered for the class of adaptive controls discussed in Chapter 2 and conclusions will be drawn as to the feasibility of the method for this class.

With the substitution $u(t) = 0$ Eq. 2-7 becomes:

$$I = \int_0^T \left\{ \lambda(t) \left[c_o(t) - c_i(t) - \int_0^t m(\tau)\, w(t,\tau)\, d\tau \right]^2 + m^2(t) \right\} dt \tag{3-1}$$

In order to determine the optimum control variable, it is assumed that a solution does exist and is denoted by $m_o(t)$. A variation of $m_o(t)$ is then constructed by letting

$$m(t) = m_o(t) + \epsilon\, m_\epsilon(t) \tag{3-2}$$

where ϵ is a parameter independent of t and $m_\epsilon(t)$ is the variation of $m(t)$. If $m_o(t)$ is the optimum control variable which therefore minimizes Eq. 3-1, then any variation of ϵ from zero in Eq. 3-2 must cause an increase in the value of Eq. 3-1 from its minimum. Hence, if Eq. 3-2 is substituted into Eq. 3-1, the derivative of the resultant equation with respect to ϵ for ϵ set equal to zero must be zero.

Substituting Eq. 3-2 into Eq. 3-1 and differentiating with respect to ϵ gives

$$\frac{\partial I}{\partial \epsilon} = 2 \int_0^T \left\{ \lambda(t) \left[c_o(t) - c_i(t) - \int_0^t \left[m_o(\tau) + \epsilon\, m_\epsilon(\tau) \right] w(t,\tau)\, d\tau \right] \cdot \left[- \int_0^t m_\epsilon(\tau)\, w(t,\tau)\, d\tau \right] + \left[m_o(t) + \epsilon\, m_\epsilon(t) \right] m_\epsilon(t) \right\} dt \tag{3-3}$$

From the argument given above, if $\epsilon = 0$, the right hand side of Eq. 3-3 must be zero. Hence,

$$\left.\frac{\partial I}{\partial \epsilon}\right|_{\epsilon=0} = \int_0^T \left\{ \lambda(t) \left[c_o(t) - c_i(t) - \int_0^t m_o(\tau)\, w(t,\tau)\, d\tau \right] \cdot \left[- \int_0^t m_\epsilon(\tau)\, w(t,\tau)\, d\tau \right] + m_o(t)\, m_\epsilon(t) \right\} dt = 0 \tag{3-4}$$

Although Eq. 3-4 expresses the condition for a minimum, it is not in a form in which the variation $m_\epsilon(t)$ is separable. The solution of variational problems of this type is usually expressed in the form of a differential equation (commonly termed the Euler equation) with boundary conditions. For an Nth order dynamic process it is necessary to integrate Eq. 3-4 by parts N times to obtain the Euler relation. Hence, without a knowledge of the order of the dynamic process, solution of Eq. 3-4 is impossible. Moreover, the presence of boundary conditions, a natural consequence of this type of variational problem, poses additional difficulties. In particular, for an Nth order dynamic process, there will be N natural boundary conditions which the solution must satisfy.

In most physical situations the order of the dynamic process is known. Nevertheless, the solution of Eq. 3-4 will give no insight into the structural form of the adaptive loop other than to indicate the need for a complex, high-speed digital computer for the generation of $m_o(t)$. In addition, the presence of boundary conditions will not permit sequential computations, but will require trial and error calculations for solution of $m_o(t)$.

Clearly, the calculus of variations approach imposes heavy demands on the computational ability of the adaptive loop in order to optimize the dynamic process response. Extensive numerical computations are necessary which will obscure the relationship between adaptive and non-adaptive controls. As a result, this approach is not tractable from either an analytical or experimental viewpoint for the purposes discussed in Chapter 2.

More recently Bellman [28, Ch. 9] has developed a new approach to calculus of variations problems in terms of dynamic programming. Although the method offers some hope for the application of the calculus of variations to the adaptive control problem, it is computationally cumbersome.

3.2 Discrete Segment Approximation

In this section the continuous control optimization problem which was discussed in the last section will be replaced by an approximately equivalent discrete formulation. The interval from $t = 0$ to $t = T$ is partitioned by a sequence of points $(t_0, t_1, \ldots, t_N)$ separated by a distance Δ where $\Delta = \frac{T}{N}$ and N equals the number of partition points. The control variable is then approximated by a sequence of discrete levels m_n as shown in Fig. 3-1. The integral

$$c(t) = \int_0^t m(\tau)\, w(t,\tau)\, d\tau \qquad (3\text{-}5)$$

is approximated by the sum

$$c_n = \Delta \sum_{j=1}^{N} w_{nj}\, m_n \qquad n \geqslant j \qquad (3\text{-}6)$$

and

$$c_n = 0 \qquad j > n \qquad (3\text{-}7)$$

because $w_{nj} = 0$ for $j > n$; that is, the system does not have access to its future values. Hence, the output becomes a sequence of values $(c_1, c_2, \ldots, c_N)$.

The component of the output due to initial conditions $c_i(t)$, and the system error weighting factor $\lambda(t)$ are also approximated by sequences of values, $(c_{i_1}, c_{i_2}, \ldots, c_{i_N})$ and $(\lambda_1, \lambda_2, \ldots, \lambda_N)$, respectively. The same is done for the desired response $c_o(t)$.

Using the above definitions and approximating the index of performance, Eq. 3-1, by a sum gives

$$I = \Delta \sum_{n=1}^{N} \lambda_n \left[c_o - c_i - c_n\right]^2 + m_n^2 \qquad (3\text{-}8)$$

The optimization problem for this case deals with the choice of the m_n such that Eq. 3-8 is a minimum. Therefore, for any integer k, the condition for a minimum value of Eq. 3-8 is

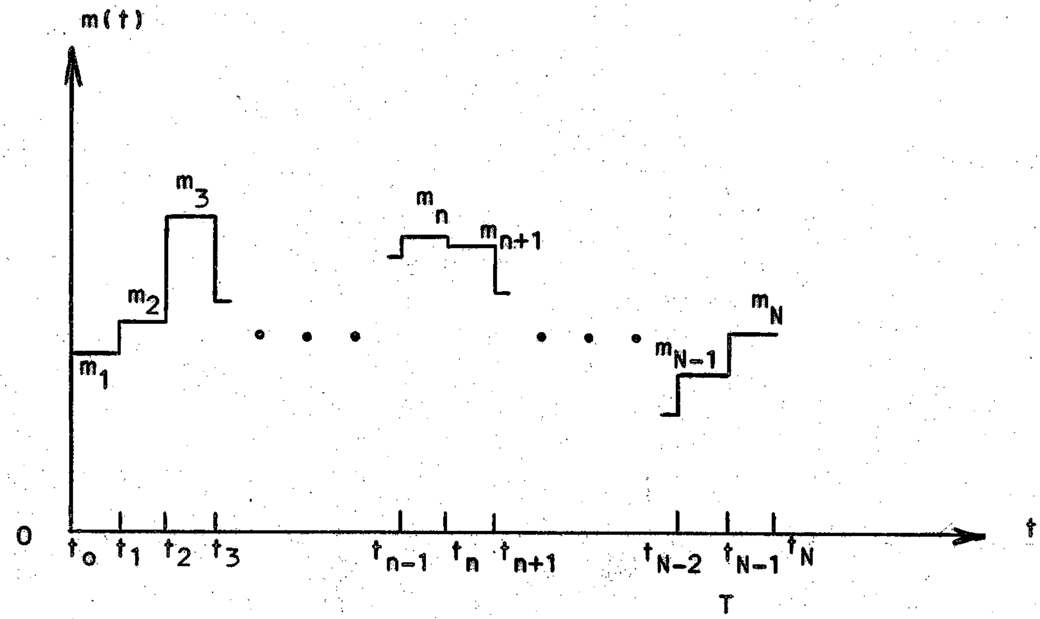


Fig. 3-1

Discrete Segment Approximation of Control Variable m(t).

$$\frac{\partial I}{\partial m_k} = 0 \qquad (3\text{-}9)$$

for all k, k = 1, 2, ... , N. Performing the indicated differentiation and setting the result equal to zero gives

$$\frac{\partial I}{\partial m_k} = \Delta \sum_{n=1}^{N} \left\{ \lambda_n \left[c_{o_n} - c_{i_n} - c_n \right] \frac{\partial c_n}{\partial m_k} - m_n \frac{\partial m_n}{\partial m_k} \right\} = 0 \qquad (3\text{-}10)$$

for k = 1, 2, ... , N.

Since

$$\frac{\partial c_n}{\partial m_k} = \Delta \cdot w_{nk} \qquad n \geq k$$

$$\qquad (3\text{-}11)$$

$$\frac{\partial c_n}{\partial m_k} = 0 \qquad n < k$$

and

$$\frac{\partial m_n}{\partial m_k} = 1 \qquad n = k$$

$$\qquad (3\text{-}12)$$

$$\frac{\partial m_n}{\partial m_k} = 0 \qquad n \neq k$$

the condition for a minimum becomes

$$\Delta \sum_{n=1}^{N} \left\{ \lambda_n \left[c_o - c_i - c_n \right] w_{nk} - m_k \right\} = 0 \qquad (3\text{-}13)$$

with

$$c_n = \Delta \sum_{j=1}^{N} w_{nj}\, m_n \qquad n \geq j \qquad (3\text{-}14)$$

Eq. 3-13 actually represents N linear algebraic equations in N unknowns. The equations must be solved simultaneously at the beginning of each control interval to give the optimum control variable as a sequence of values $(m_1, m_2, \ldots, m_N)$ for that control interval. This formulation is more amenable to digital computation than the first method considered, but still obscures any real insight which one may hope to gain about the structure of the adaptive loop.

An approach to the simultaneous solution of Eq. 3-13 is obtained by considering the last member of this set $(k = N)$ which is

$$\Delta \lambda_N \left[c_{oM} - c_{iM} - w_{NN} m_N \right] w_{NN} - m_N = 0 \qquad (3\text{-}15)$$

Since m_N is the only unknown, Eq. 3-15 is easy to solve. The solution of Eq. 3-15 may then be substituted into Eq. 3-13 for $k = N - 1$, and the resulting equation solved for its only unknown, m_{N-1}. Hence, the solution of the set of equations given by Eq. 3-13 propagates backward through the set. The use of high-speed digital computation is again mandatory to determine the optimum control variable. Here again no insight into the real nature of adaptive control can be gained.

3.3 Orthonormal Polynomial Sum Approximation

For a large class of adaptive control problems the use of a high-speed digital computing facility is undesirable. Such factors as size, weight, and cost are paramount in practical applications. Unfortunately, the necessity of high-speed digital computation has been a natural consequence of the two optimization procedures considered thus far. While these mathematical procedures for optimization are well-defined, the end results do not lend themselves to a well-defined engineering interpretation. The only interpretation has been that large-scale digital computation is necessary.

What is really sought here is a set of reasonable assumptions based on engineering considerations which will simplify the optimization procedure, keep the complexity of the adaptive loop at a minimum, give reasonable over-all system performance, and be consistent with the objectives of adaptive control as discussed in Chapter 2.

First, the optimum control signal m(t) should be one that is physically realizable. That is, it should not consist of impulses or higher order signularity functions which will invalidate the assumption that the dynamic process can be characterized by a linear, time-varying, weighting function $w(t, \tau)$. Secondly, m(t) should be relatively simple to synthesize during normal operation of the system. This second factor implies simplicity of the adaptive loop. Thirdly, the mathematical formulation of m(t) should lend itself readily to an optimization procedure which is simple and which gives physical insight into the form of the adaptive loop.

The approximation of m(t) by an N-term sum of orthonormal polynomials in t is considered in this section. This approximation is defined by

$$m(t) = \sum_{n=0}^{N} m_n p_n(t) \qquad (3\text{-}16)$$

where the m_n are the coefficients which are to be determined, and the $p_n(t)$ are polynomials in t which are orthonormal over the interval $[0,T]$. In other words, the set of polynomials satisfies the following two conditions:

(a) $p_n(t)$ is a polynomial in t of degree n.

(b) $$\int_0^T p_k(t)\, p_n(t)\, dt = \begin{cases} 1 & k = n \\ 0 & k \neq n \end{cases} \qquad (3\text{-}17)$$

where T is the control interval length. These polynomials are the Legendre polynomials with their usual interval of orthonormality $[-1, 1]$ transformed into the interval $[0, T]$.

The input signal thus becomes a polynomial in t whose degree is dictated by the particular N chosen. The coefficients m_0, m_1, m_2, etc., will be generated by the adaptive loop in response to changes in process dynamics and the desired response.

The motivation for using orthonormal polynomials in t, rather than a Taylor series expansion as in Braun [12], for m(t) is the hope that the coefficients m_n can be generated independently for each control interval in the former case. If this can be done, the signal synthesis portion of the adaptive loop can assume the form shown in Fig. 3-2.

It will be shown that independent generation of the m_n is possible in real-time by means of time-varying gains and integrators. Clearly, such a scheme will avoid the necessity of complex high-speed digital computation, and will offer considerable simplicity in system design. The detailed treatment of this approach is presented in Chapter 4.

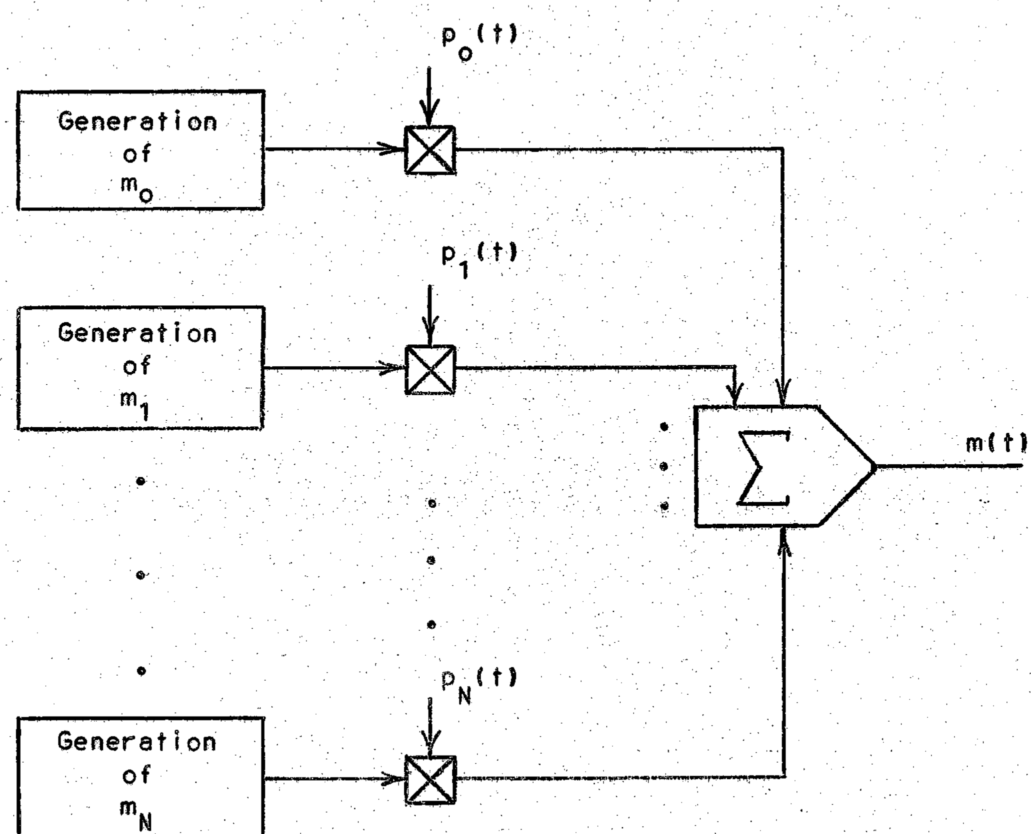


Fig. 3-2

Signal Synthesis for Independent Generation of Control Signal Coefficients.

CHAPTER 4

DERIVATION AND ANALYSIS OF THE OPTIMUM CONTROL CONFIGURATION

The purpose of this chapter is to investigate the last approach to the optimization problem which is given in Chapter 3. The control equations are developed and the optimum control configuration is derived.

In addition, a theoretical analysis of certain characteristics of the class of adaptive controls developed is presented. A theoretical system transfer function is derived and applied to a stability analysis. Limitations of the transfer function approach are also discussed. Finally, since this class of adaptive controls employs prediction, accuracy requirements in terms of the prediction operations are discussed briefly.

4.1 General Condition for Optimum

The three equations from which the general condition for optimum control will be derived are repeated below:

$$I = \int_0^T \left\{ \lambda(t) \left[c_o(t) - c(t) \right]^2 + m^2(t) \right\} dt \qquad (2\text{-}4)$$

$$c(t) = c_i(t) + u(t) + \int_0^t m(\mathcal{T})\, w(t,\mathcal{T})\, d\mathcal{T} \qquad (2\text{-}6)$$

and

$$m(t) = \sum_{n=0}^{N} m_n\, p_n(t) \qquad (3\text{-}16)$$

for $0 \leqslant t \leqslant T$. The terms in these equations have been defined previously. It is assumed here that the disturbance is a stationary, Gaussian, white noise process which is independent of the input signal m(t).

The values of the various m_n, n = 0, 1, ... , N, needed to minimize Eq. 2-4 are obtained by differentiating the equation with respect to

m_k, $k = 0, 1, \ldots, N$, and setting the result equal to zero. Thus,

$$\frac{\partial I}{\partial m_k} = 2 \int_0^T \left\{ \lambda(t) \left[c_o(t) - c(t)\right] \left[- \frac{\partial c(t)}{\partial m_k} \right] + m(t) \cdot \frac{\partial m(t)}{\partial m_k} \right\} dt = 0 \tag{4-1}$$

for $k = 0, 1, \ldots, N$.

From Eq. 3-16

$$\frac{\partial m(t)}{\partial m_k} = p_k(t) \tag{4-2}$$

and from Eq. 2-6,

$$\frac{\partial c(t)}{\partial m_k} = \int_0^t \frac{\partial m(\tau)}{\partial m_k} w(t,\tau)\, d\tau \tag{4-3}$$

for $k = 0, 1, \ldots, N$.

Substituting Eq. 4-2 into Eq. 4-3 yields

$$\frac{\partial c(t)}{\partial m_k} = \int_0^t p_k(\tau)\, w(t,\tau)\, d\tau \tag{4-4}$$

for $k = 0, 1, \ldots, N$.

From Eqs. 3-16, 3-17, and 4-2, the last term of Eq. 4-1 becomes

$$\int_0^T m(t) \frac{\partial m(t)}{\partial m_k} dt = \int_0^T \left[\sum_{n=0}^{N} m_n\, p_n(t) \right] p_k(t)\, dt \tag{4-5}$$

which simplifies to

$$\int_0^T m(t) \frac{\partial m(t)}{\partial m_k} dt = m_k \tag{4-6}$$

because of the orthonormality of the polynomials for $k = 0, 1, \ldots, N$.

Substituting Eqs. 4-4 and 4-6 into Eq. 4-1 gives

$$\int_0^T \lambda(t) \left[c_o(t) - c(t)\right] \left[- \int_0^t p_k(\tau)\, w(t,\tau)\, d\tau \right] dt + m_k = 0 \tag{4-7}$$

for $k = 0, 1, \ldots, N$. Eq. 4-7 actually represents N linear algebraic equations in as many unknowns since $c(t)$ is also a function of the m_k as seen from Eqs. 2-6 and 3-16. These equations can be solved explicitly for the m_k providing $c_i(t)$, $u(t)$, and $w(t,\tau)$ are known. Such a solution would again require the use of a high-speed digital computer in the adaptive loop in order to control the dynamic process. However, an approximate solution can be obtained by considering an estimate or prediction of the quantity $\left[c_o(t) - c(t)\right]$.

The coefficients m_k which are needed for a particular control interval must be available at the beginning of that interval according to Eq. 3-16. But Eqs. 2-6 and 4-7 indicate the m_k depend upon the responses $c_o(t)$ and $c(t)$ during the same interval. However, if the quantity $\left[c_o(t) - c(t)\right]$ can be estimated T units of time in advance, it is possible to employ Eq. 4-7 directly to estimate the values of the m_k for the succeeding interval. That is, the coefficients m_k for the Pth interval can be generated by real-time operations during the (P-1)th interval.

In order to apply the notion of prediction, define

$$\left[c_o(t) - c(t)\right]^* = \text{best available estimate of } \left[c_o(t) - c(t)\right] \text{ T units of time in advance}$$

Eq. 4-7 can then be solved directly for the estimated m_k to give

$$m_k = \int_0^T \left\{ \lambda(t) \left[c_o(t) - c(t)\right] \left[\int_0^t p_k(\tau)\, w(t,\tau)\, d\tau\right]\right\} dt \qquad (4\text{-}8)$$

for $k = 0, 1, \ldots, N$.

Let a time-varying gain $K_k(t)$ be defined by

$$K_k(t) = \lambda(t) \int_0^t p_k(\tau)\, w(t,\tau)\, d\tau \qquad (4\text{-}9)$$

for $k = 0, 1, \ldots, N$ where $0 \leq \tau \leq t \leq T$.

When Eq. 4-9 is substituted into Eq. 4-8, the result is

$$m_k = \int_0^T K_k(t) \left[c_o(t) - c(t)\right]^* dt \qquad (4\text{-}10)$$

for k = 0, 1, ..., N. With a change in the index of summation, Eq. 3-16 becomes

$$m(t) = \sum_{k=0}^{N} m_k \, p_k(t) \qquad (4\text{-}11)$$

for $0 \leq t \leq T$.

Eq. 4-10 is the general condition for the optimum. The combination of Eqs. 4-10 and 4-11 constitutes the control laws for a class of predictive adaptive controls. Eq. 4-10 indicates how the coefficients m_k for any control interval P can be obtained by real-time computation during the preceding control interval, (P-1). Eq. 4-11 indicates how these coefficients are combined with their corresponding polynomials $p_k(t)$ to generate the optimum control variable. The fact that the optimization procedure presented here renders the index of performance Eq. 2-4 a minimum is demonstrated in Appendix B.

Eq. 4-10 suggests a formal synthesis procedure for the generation of the m_k which is shown in Fig. 4-1. Each of the m_k is then multiplied by its corresponding $p_k(t)$ and the results summed to give m(t) according to Eq. 4-11. The complete control configuration then assumes the form given in Fig. 4-2. The configuration of Fig. 4-2 is optimum on a per interval basis. Hence, the time-varying gains and integrators must be reset at the end of each control interval to initiate computation for the next interval. The function of the sample and hold devices is to read out the values of the various m_k at the end of each interval and to maintain these values throughout the new interval.

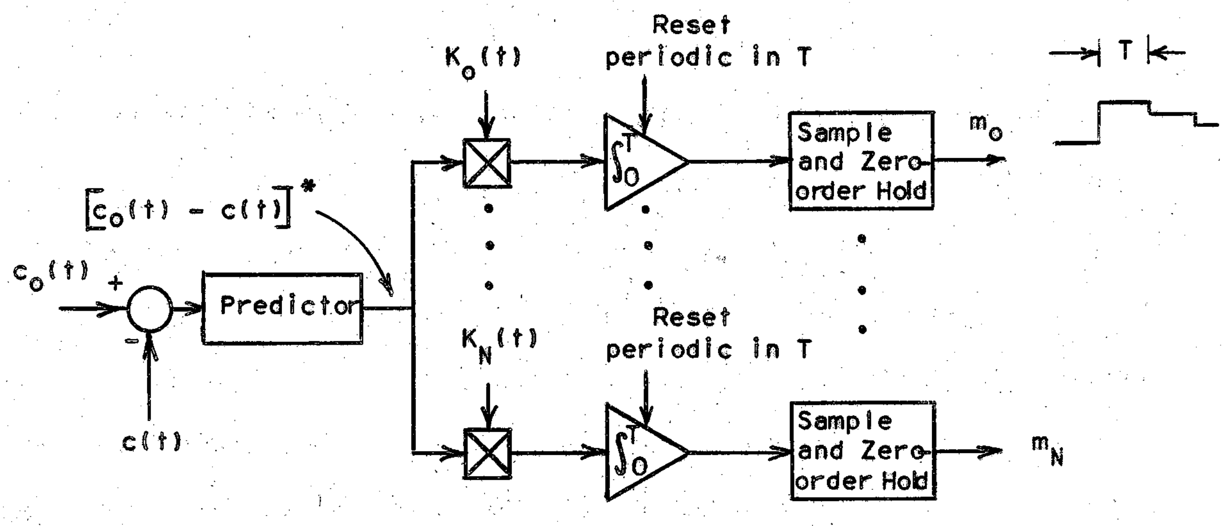


Fig. 4-1

Synthesis of Control Coefficients.

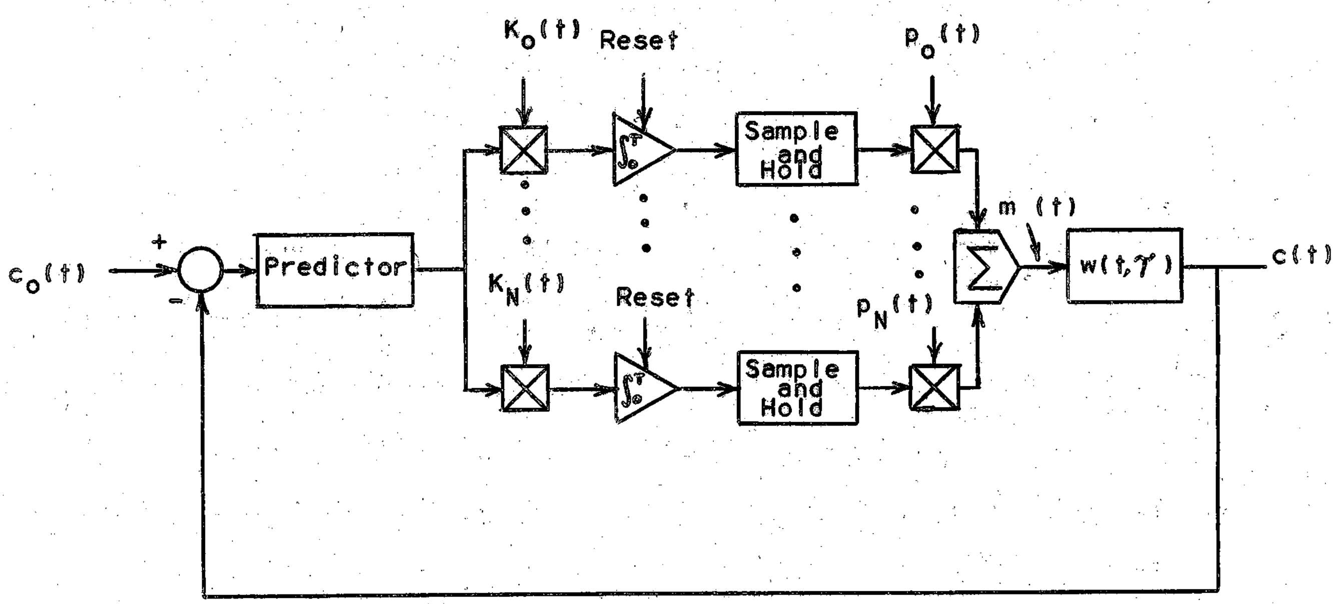


Fig. 4-2

Complete Control Configuration.

In terms of the information flow in the over-all system, it is clear that the controller operates on the quantity

$$\left[c_o(t) - c(t)\right]^* \tag{4-12}$$

to derive the control variable m(t). The quantity in (4-12) is simply the predicted system error. In other words, the function of the predictor is to present the controller with an estimate of the <u>future</u> system error. The task of the controller is then to synthesize a control signal which will minimize the <u>actual</u> system error in the succeeding control interval. Hence, over-all system operation may be viewed as data processing of the predicted system error to derive the optimum control signal. The character of the data processing changes to accommodate changes in the dynamic process $w(t,\tau)$, changes in the desired response $c_o(t)$, and changes in the index of performance which are governed by $\lambda(t)$, the system error weighting factor.

An important feature of this class of adaptive controls is the nature of the time-varying gains which are given by Eq. 4-9,

$$K_k(t) = \lambda(t) \int_0^t p_k(\tau)\, w(t,\tau)\, d\tau \tag{4-9}$$

for k = 0, 1, ..., N. Since the polynomials $p_k(t)$ are linear combinations of the singularity functions, i.e., the step, ramp, and parabolic functions, etc., the time-varying gains are simply the products of the error weighting factor $\lambda(t)$ and the response of the process $w(t,\tau)$ to linear combinations of these same singularity functions. Therefore, the time-varying gains $K_k(t)$ are easy to generate given a knowledge of the dynamic process impulse $w(t,\tau)$.

The adaptive nature of the optimum control configuration is clear from Eqs. 4-9 and 4-10. The time-varying gains $K_k(t)$ are related directly to the error weighting factor $\lambda(t)$ and the dynamic process unit impulse response $w(t,\tau)$. Hence, the adaptive loop has a means of changing the index of performance by changing $\lambda(t)$ as the goals of control change, and is also capable of accounting for changes in process dynamics $w(t,\tau)$ all through the time-varying gains $K_k(t)$.

Examination of the control laws and the control configuration reveals three important features of this class of adaptive controls:

1. The controller can be realized using simple analog components.
2. The controller operates in real-time.
3. Complex computational operations have been avoided.

These three items satisfy the original goals which were established in the formulation of the modification problem (Chapter 2). Item 3 is actually an outgrowth of the first two, but is included for emphasis.

In conclusion, the control laws of Eqs. 4-10 and 4-11, and the control configuration of Fig. 4-2 completely specify the class of adaptive controls to be studied in this work. Their derivation has been based on the specified index of performance, the assumption of the general functional form of the optimum control variable, and on the use of prediction to obtain the predicted error signal.

4.2 Theoretical System Transfer Function

In the analysis of feedback control systems, it is often desirable to determine the system transfer function if it exists. In this section, it will be shown that a system transfer function does exist theoretically, but is impossible to obtain in general. However, a slight modification of the results developed in this section will permit the derivation of stability results for a particular sub-class of these systems.

Under the assumption of a linear, time-invariant dynamic process and ideal prediction, the block diagram of Fig. 4-2 becomes that shown in Fig. 4-3. The assumption that the controller portion of the system

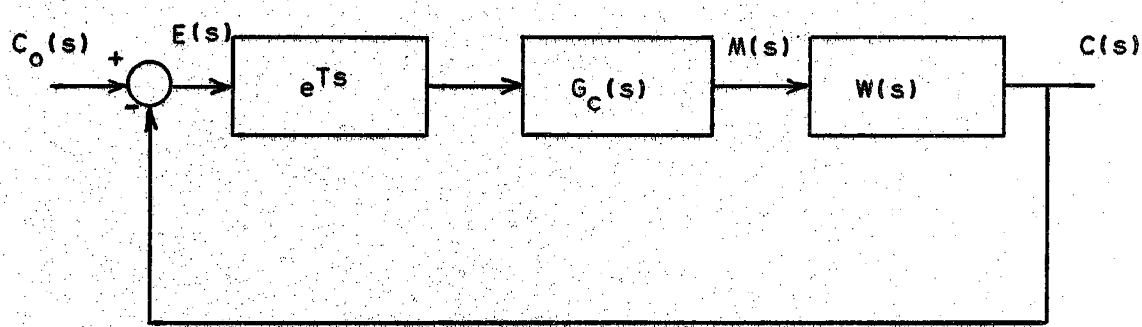


Fig. 4-3

Model of Complete System with Ideal Prediction and Time-invariant Dynamic Process.

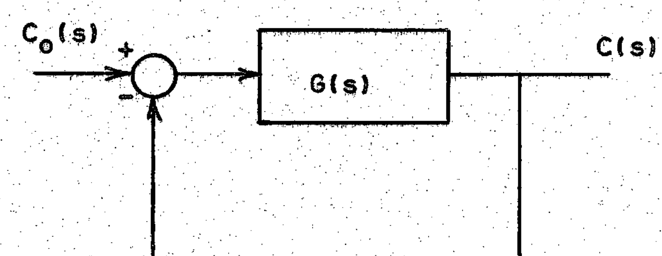


Fig. 4-4

Simplified Block Diagram of System Model.

can be characterized by a transfer function $G_c(s)$, in principle at least, will be demonstrated later in this section.

Using standard block diagram reduction techniques, the diagram of Fig. 4-3 reduces to that shown in Fig. 4-4 where G(s) is given by

$$G(s) = e^{Ts} G_c(s) W(s) \tag{4-13}$$

The closed-loop transfer function is then defined by the relation

$$\frac{C(s)}{C_o(s)} = \frac{e^{Ts} G_c(s) W(s)}{1 + e^{Ts} G_c(s) W(s)} \tag{4-14}$$

This equation may also be written

$$\frac{C(s)}{C_o(s)} = \frac{G_c(s) W(s)}{e^{-Ts} + G_c(s) W(s)} \tag{4-15}$$

Attention is now directed to the development of the transfer function $G_c(s)$ characterizing the controller. Since the input to each channel of the controller of Fig. 4-2 which computes the coefficients m_k, $k = 0, 1, \ldots, N$, is $\left[c_o(t) - c(t)\right]^*$, and the outputs of all the channels are summed to form m(t), it is necessary to consider only the nth channel, n an integer, such that $0 \leq n \leq N$, and sum the transfer functions of the N channels to obtain $G_c(s)$. The nth channel may be represented as in Fig. 4-5.

The channel of Fig. 4-5 will be subdivided into three parts for purposes of analysis: (1) the pre-multiplier, (2) the integrator, sampler and zero-order hold, and (3) the post-multiplier.

Letting $e(t) = \left[c_o(t) - c(t)\right]^*$ be the input and $y_n(t)$ the output of the pre-multiplier (Fig. 4-6), the output is given by

$$y_n(t) = K_n(t) \cdot e(t). \tag{4-16}$$

Since multiplication in the time domain corresponds to convolution in the frequency domain, the Laplace transform of the pre-multiplier output $Y_n(s)$ is determined by the relation [29, p. 275],

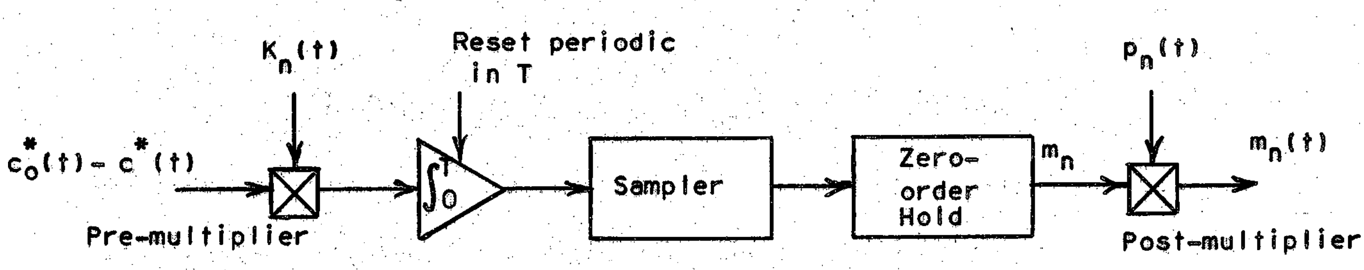


Fig. 4-5

nth Channel of Controller

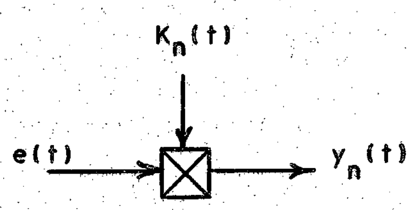


Fig. 4-6

Pre-multiplier of Controller.

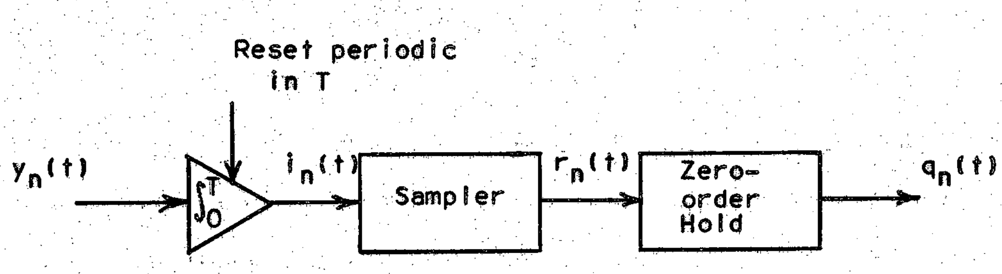


Fig. 4-7

Integrator, Sampler, and Zero-order Hold of Controller.

$$Y_n(s) = \frac{1}{2\pi j} \int_{c_2 - j\infty}^{c_2 + j\infty} E(s-w)\, K_n(s)\, dw \qquad (4\text{-}17)$$

where max $(\sigma_{a_1}, \sigma_{a_2}, \sigma_{a_1} + \sigma_{a_2}) < \sigma$, $\sigma_{a_2} < c_2 < \sigma - \sigma_{a_1}$ in which c_2 is a real constant, $\sigma = \text{Re}[s]$, and $\sigma_{a_1}, \sigma_{a_2}$ are the abscissas of absolute convergence of the time functions $e(t)$ and $K_n(t)$, respectively. Eq. 4-17 indicates the basic difficulty associated with obtaining the transfer function of the controller. The presence of the controller input $E(s)$ within the complex convolution of Eq. 4-17 makes obtaining the transfer function $G_c(s)$ impossible in the general case.

The combination of integrator, sampler, and zero-order hold is given in Fig. 4-7 where the inputs and outputs of each device have been defined. As pointed out in Section 4.1, since the optimization process is executed on a per interval basis, the integration in the controller channel must be reset to zero at the end of a given control interval and the beginning of the following interval. That is, at the end of the kth control interval, the output of the integrator must be

$$i_n(kt) = \int_{(k-1)T}^{kT} y_n(t_1)\, dt_1 \qquad k = 1, 2, \ldots \qquad (4\text{-}18)$$

It is possible to view this process of reset integration on a continuous integration basis because the sampler is synchronized with the time-varying gains $K_n(t)$. Thus, if the output of the integration process at any time t, $t \geq 0$ is

$$i_n(t) = \int_{t-T}^{t} y_n(t_1)\, dt_1 \qquad (4\text{-}19$$

then at the kth sampling instant $t = kT$ Eq. 4-19 reduces to Eq. 4-18. Eq. 4-19 may also be written

$$i_n(t) = \int_0^t y_n(t_1)\, dt_1 - \int_0^{t-T} y_n(t_1)\, dt_1 \qquad (4\text{-}20)$$

Letting

$$f_n(t) = \int_0^t y_n(t_1)\, dt_1 \qquad (4\text{-}21)$$

Eq. 4-20 becomes

$$i_n(t) = f_n(t) - f_n(t - T) \qquad (4\text{-}22)$$

The Laplace transform of Eq. 4-22 is simply

$$I_n(s) = (1 - e^{-Ts})\, F_n(s) \qquad (4\text{-}23)$$

But

$$F_n(s) = \frac{1}{s}\, Y_n(s) \qquad (4\text{-}24)$$

where $Y_n(s)$ is the Laplace transform of $y_n(t)$. Substituting Eq. 4-24 into Eq. 4-23 gives the transfer function of the reset integration

$$\frac{I_n(s)}{Y_n(s)} = \frac{1}{s}\,(1 - e^{-Ts}) \qquad (4\text{-}25)$$

From Truxal [30, p. 503] the transfer function between $r_n(t)$ and $i_n(t)$ is expressed by

$$R_n(s) = \sum_{\mu = -\infty}^{\mu = +\infty} I_n(s + j\mu\omega_s) \qquad (4\text{-}26)$$

where μ is an integer and $\omega_s = \frac{2\pi}{T}$. Substituting $I_n(s)$ from Eq. 4-25 into Eq. 4-26 gives

$$R_n(s) = \sum_{\mu = -\infty}^{\mu = +\infty} \frac{1}{s + j\mu\omega_s}\left[1 - e^{-T(s + j\mu\omega_s)}\right] Y_n(s + j\mu\omega_s) \qquad (4\text{-}27)$$

The transfer function for the zero-order hold, Truxal [30, p. 507], is

$$\frac{Q_n(s)}{R_n(s)} = \frac{1}{s}\,(1 - e^{-Ts}) \qquad (4\text{-}28)$$

Combining Eqs. 4-27 and 4-28 and recalling $\omega_s = \frac{2\pi}{T}$ there results

$$Q_n(s) = \frac{1}{s}\,(1 - e^{-Ts}) \sum_{\mu = -\infty}^{\mu = +\infty} \frac{1}{s + j\mu\omega_s}\left[1 - e^{-j2\pi\mu}\, e^{-Ts}\, Y_n(s + j\omega_s)\right] \qquad (4\text{-}29)$$

which simplifies to

$$Q_n(s) = \frac{1}{s}\,(1 - e^{-Ts})^2 \sum_{\mu = -\infty}^{\mu = +\infty} \frac{1}{s + j\mu\omega_s}\, Y_n(s + j\mu\omega_s) \qquad (4\text{-}30)$$

Eq. 4-30 represents the transfer function of the integrator, sampler, and zero-order hold combination.

The post-multiplier with its inputs $q_n(t)$ and $p_n(t)$, and its output $m_n(t)$ is shown in Fig. 4-8. By definition the output of this multiplier is

$$m_n(t) = p_n(t) \cdot q_n(t) \qquad (4\text{-}31)$$

Again, since multiplication in the time domain corresponds to complex convolution in the frequency domain, the Laplace transform of the multiplier output is expressed by

$$M_n(s) = \frac{1}{2\pi j}\int_{c_2 - j\infty}^{c_2 + j\infty} P_n(s-w)\, Q_n(s)\, dw \qquad (4\text{-}32)$$

where $\max(\sigma_{a_1}, \sigma_{a_2}, \sigma_{a_1} + \sigma_{a_2}) < \sigma$, $\sigma_{a_2} < c_2 < \sigma - \sigma_{a_1}$ in which c_2 is a real constant, $\sigma = \mathrm{Re}[s]$, and σ_{a_1}, σ_{a_2} are the abscissas of absolute convergence of the time functions $p_n(t)$ and $q_n(t)$, respectively, [29, p. 275].

Eqs. 4-17, 4-30, and 4-32 are the three basic relations for determining the transfer function

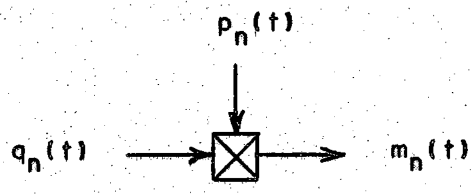
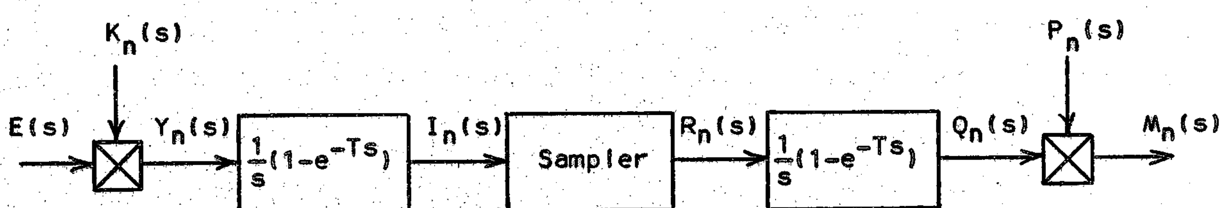


Fig. 4-8

Post-multiplier of Controller.



with $$R_n(s) = \sum_{\mu=-\infty}^{\mu=\infty} I_n(s + j\mu\omega_s) \qquad \omega_s = \frac{2\pi}{T}$$

Fig. 4-9

Model of nth Channel of Controller.

$$G_c(s) = \sum_{n=0}^{N} \frac{M_n(s)}{E(s)} \tag{4-33}$$

of the controller. The inherent difficulties in obtaining $G_c(s)$ are brought out by these three equations. The presence of complex convolu- in Eqs. 4-17 and 4-32, and infinite sum of Eq. 4-30 render solution for the general case impossible as mentioned earlier. The main obstacle to the use of the transfer function approach for this class of adaptive controls appears to be the presence of the pre-multiplier having inputs e(t) and $K_n(t)$. This multiplication operation forces the Laplace transform of the input variable, E(s), to appear under a complex convolution.

The results developed in this section will be modified slightly in the next section and applied to a stability analysis of a particular sub-class of the class of adaptive controls under investigation in this research.

The work of this section is summarized in the block diagram given in Fig. 4-9.

4.3 Some Stability Results

In some adaptive control applications it may be possible to use only one channel in the controller and still get satisfactory performance. For this sub-class of predictive adaptive controls, in which a one-term approximation of the control variable m(t) is employed, it is possible to utilize the results of the preceding section to effect an analytical stability analysis.

The basic block diagram for the system using a one-term approximation of the control variable is shown in Fig. 4-10. From Appendix A the external input to the post-multiplier is

$$p_0(t) = \sqrt{\frac{1}{T}} \tag{4-34}$$

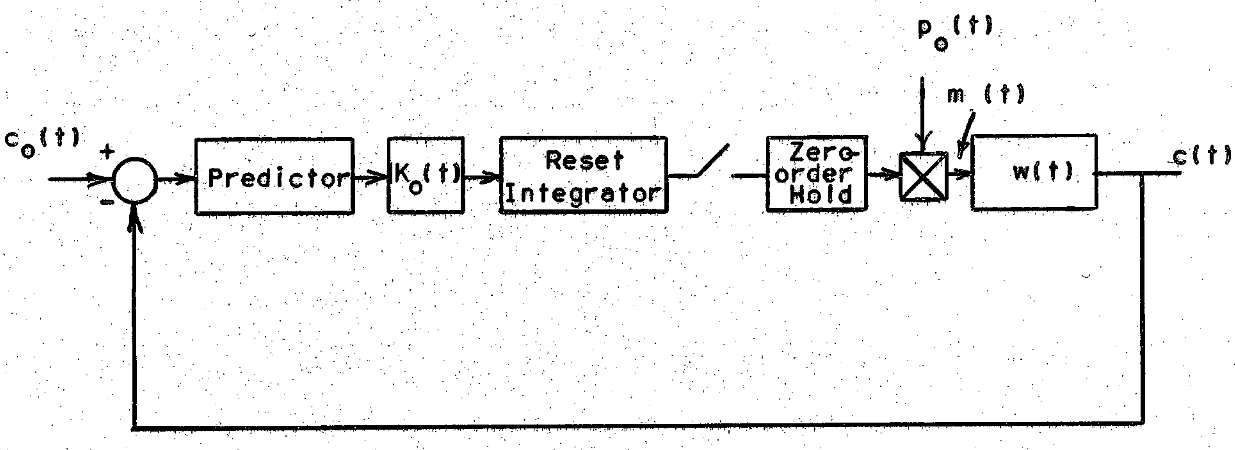


Fig. 4-10

Block Diagram of Predictive Adaptive Control for One-term Approximation of Control Variable.

for all t. Hence, the post-multiplier is replaced by a gain of $\sqrt{\frac{1}{T}}$.
From Eq. 4-9

$$K_o(t) = \lambda(t) \int_0^t p_o(\tau)\, w(t,\tau)\, d\tau \qquad (4\text{-}35)$$

where all of the symbols have been defined previously.

The analysis which follows is not exact. It is based on a linearization of the controller in order to determine bounds on $K_o(t)$ for stability. The time-varying gain $K_o(t)$ is replaced by a constant gain $\tilde{K}_o$ and the resulting system is analyzed to determine the range of values on $\tilde{K}_o$ for which the closed loop system is stable. $K_o(t)$ is then constrained to lie within this range for each control interval $0 < t < T$. That is, the actual range on $K_o(t)$ is compared with the required range on $\tilde{K}_o$ to establish requirements on the system parameters and/or the control interval length for closed-loop system stability.

Utilizing the results of Section 4.2 and the simplications discussed above, and replacing the ideal prediction operation e^{Ts} by the approximation $1 + Ts$, the frequency domain block diagram becomes that shown in Fig. 4-11. After a few simple block diagram manipulations, Fig. 4-11 reduces to Fig. 4-12.

Since the system employs sampling, use of the z-transform instead of complex frequency s will facilitate the analysis considerably and will, therefore, be used in the work which follows. Letting

$$G_1(s) = \frac{\tilde{K}_o(Ts + 1)(1 - e^{-Ts})}{s} \qquad (4\text{-}36)$$

and

$$G_2(s) = \frac{(1 - e^{-Ts})\, W(s)}{\sqrt{T}\; s} \qquad (4\text{-}37)$$

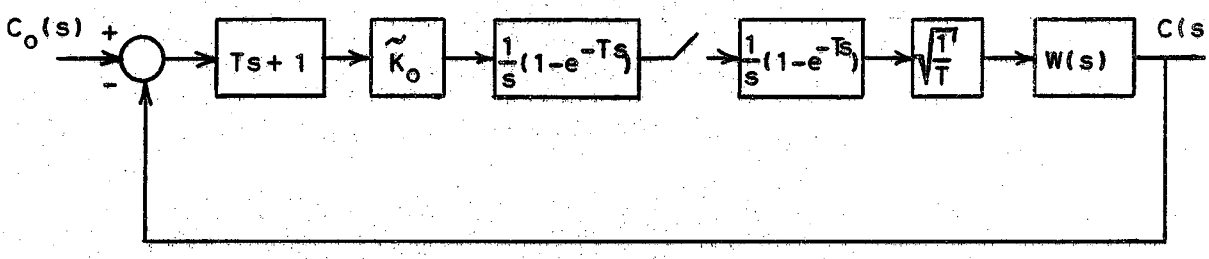


Fig. 4-11

Block Diagram for Stability Analysis.

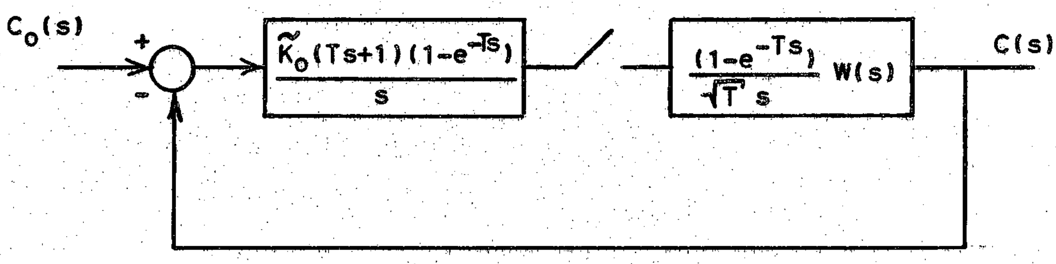


Fig. 4-12

Simplified Block Diagram for Stability Analysis.

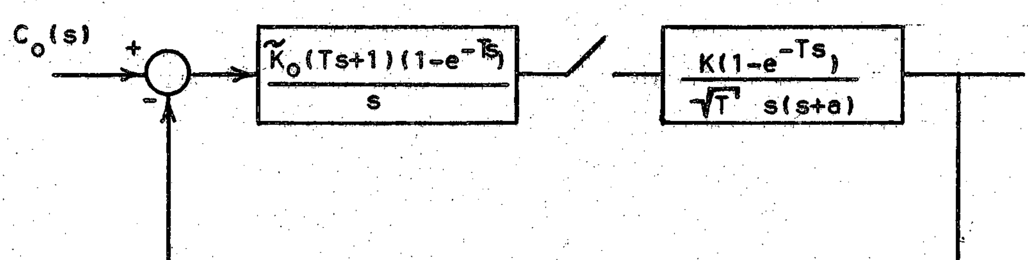


Fig. 4-13

Block Diagram for Stability Analysis of Example.

it is known that the stability [26, Ch. 6] of the configuration of Fig. 4-12 is governed by the locations of the zeros of $1 + G_1 G_2(z)$, where $G_1 G_2(z)$ is the z-transform of $G_1(s) \cdot G_2(s)$. In particular, if the zeros of $1 + G_1 G_2(z)$ lie within the unit circle of the complex z-plane, the closed loop configuration will be stable. It remains then to determine the conditions on $\widetilde{K}_o$ in order that the zeros of $1 + G_1 G_2(z)$ lie within the unit circle.

The exact procedure is best clarified by a specific example.

Example. The transfer function of the dynamic process is assumed to be of the form

$$W(s) = \frac{K}{s + a} \tag{4-38}$$

where K and a are the process parameters. It is further assumed that K and a are both positive with K fixed, but a variable. Assuming also that system error is weighted uniformly over each interval, that is,

$$\lambda(t) = \lambda_o = \text{constant}, \tag{4-39}$$

the time-varying gain $K_o(t)$ is given by

$$K_o(t) = \frac{\lambda_o}{a\sqrt{T}}(1 - e^{-at}) \tag{4-40}$$

for each interval.

For this example, Fig. 4-12 assumes the form given in Fig. 4-13. The next step in the stability analysis is to determine the z-transform of

$$G_1(s)\, G_2(s) = \frac{\widetilde{K}_o K}{\sqrt{T}} (1 - e^{-Ts})^2 \frac{Ts + 1}{S(s + a)} \tag{4-41}$$

Expanding $G_1(s)\, G_2(s)$ in a partial fraction expansion and employing a table of z-transforms [30, p. 511] to identify the corresponding z-transforms gives

$$G_1G_2(z) = \frac{\tilde{K}_oK}{\sqrt{T}}(1 - 2z^{-1} + z^{-2})\left[\frac{aTz}{(z-1)^2} - \frac{(1-aT)z}{z-1} + \frac{(1-aT)z}{z-e^{-aT}}\right] \qquad (4\text{-}42)$$

Simplifying, Eq. 4-42 yields

$$G_1G_2(z) = \frac{\tilde{K}_oK}{a^2\sqrt{T}}\;\frac{(2aT - 1 + e^{-aT} - aTe^{-aT})z - (aT - 1 + e^{-aT})}{z(z - e^{-aT})} \qquad (4\text{-}43)$$

Adding unity to Eq. 4-43 and combining terms gives the result

$$1 + G_1G_2(z) = \frac{z^2 + \frac{1}{a^2\sqrt{T}}\left[\tilde{K}_oK(2aT - 1 + e^{-aT} - aTe^{-aT}) - a^2\sqrt{T}e^{-aT}\right]z}{z(z - e^{-aT})} + \frac{\frac{1}{a^2\sqrt{T}}\tilde{K}_oK(aT + e^{-aT} - 1)}{z(z - e^{-aT})} \qquad (4\text{-}44)$$

Since the numerator of Eq. 4-44 is quadratic, the Schur-Cohn test [30, p. 523] will be utilized to determine the conditions for stability. Let $p(z)$ be the numerator polynomial of $1 + G_1G_2(z)$. The Schur-Cohn test then requires:

(1) $|p(0)| < 1$

(2) $p(1) > 0$

(3) $p(-1) > 0$

where the coefficient of the z^2 term of $p(z)$ is unity. For this example,

$$p(z) = z^2 + \frac{1}{a^2\sqrt{T}}\left[\tilde{K}_oK(2aT - 1 + e^{-aT} - aTe^{-aT})\right] z - \frac{\tilde{K}_oK}{a^2\sqrt{T}}(aT + e^{-aT} - 1) \qquad (4\text{-}45)$$

The three conditions of the Schur-Cohn test are now examined.

<u>Condition (1).</u> $\left|p(0)\right| < 1$. This condition becomes

$$\left| - \frac{\tilde{K}_o K}{a^2 \sqrt{T}} (aT + e^{-aT} - 1) \right| < 1 \tag{4-46}$$

From the problem specifications K is positive and so is $\frac{1}{a^2\sqrt{T}}$. Also, it is clear that

$$(aT + e^{-aT} - 1) > 0 \qquad \text{for } aT > 0. \tag{4-47}$$

Hence, condition (1) may be written

$$\left|\tilde{K}_o\right| < \frac{a^2\sqrt{T}}{K(aT + e^{-aT} - 1)} \tag{4-48}$$

The actual time-varying gain $K_o(t)$ given by Eq. 4-40 is

$$K_o(t) = \frac{\lambda_o K}{a\sqrt{T}} (1 - e^{-aT}) \qquad 0 \leq t \leq T \tag{4-49}$$

which is always positive. This gain is a simple exponential which reaches its maximum value at $t = T$.

$$K_o(t)\Big|_{max} = K_o(t)\Big|_{t=T} = \frac{\lambda_o K}{a\sqrt{T}} (1 - e^{-aT}) \tag{4-50}$$

Hence, condition (1) will be satisfied if

$$\frac{\lambda_o K}{a\sqrt{T}} (1 - e^{-aT}) < \frac{a^2\sqrt{T}}{K(aT + e^{-aT} - 1)} \tag{4-51}$$

Eq. 4-51 may be simplified to give

$$\lambda_o K^2 (1 - e^{-aT})(aT + e^{-aT} - 1) < a^3 T \tag{4-52}$$

This is a transcendental inequality which can be solved to determine a condition on T if λ_o and the bounds on a and K are known. A specific numerical example will be considered after the other two conditions have been examined.

Condition (2). Substituting $z = 1$ into Eq. 4-45 for condition (2) gives

$$1 + \frac{1}{a^2\sqrt{T}}\left[\tilde{K}_oK(2aT - 1 - aTe^{-aT} + e^{-aT}) - a^2\sqrt{T}\,e^{-aT}\right] - \frac{1}{a^2\sqrt{T}}\tilde{K}_oK(aT + e^{-aT} - 1) > 0 \tag{4-53}$$

After some simplification, Eq. 4-50 becomes

$$\tilde{K}_oK\, aT(1 - e^{-aT}) + a^2\sqrt{T}(1 - e^{-aT}) > 0 \tag{4-54}$$

Since $1 - e^{-aT}$ is greater than zero for all aT greater than zero, Eq. 4-54 is solved for the condition on $\tilde{K}_o$ to give

$$\tilde{K}_o > -\frac{a}{K\sqrt{T}} \tag{4-55}$$

Therefore, since $K_o(t)$ as given by Eq. 4-40 is always positive, the second condition of the Schur-Cohn test is automatically satisfied.

Condition (3). $p(-1) > 0$. Substituting $z = -1$ into Eq. 4-45 and applying condition (3) gives the requirement

$$1 - \frac{1}{a^2\sqrt{T}}\left[\tilde{K}_oK(2aT - 1 + e^{-aT} - aTe^{-aT}) - a^2\sqrt{T}\,e^{-aT}\right] - \frac{1}{a^2\sqrt{T}}\tilde{K}_oK(aT + e^{-aT} - 1) > 0 \tag{4-56}$$

Eq. 4-56 can be simplified to the inequality

$$\tilde{K}_o < \frac{a^2\sqrt{T}(1 + e^{-aT})}{K(3aT - 2 + 2e^{-aT} - aTe^{-aT})} \tag{4-57}$$

As in condition (1), this condition imposes an upper bound on $K_o(t)$ in order to assure stability. Hence, in terms of the maximum value of $K_o(t)$ Eq. 4-57 becomes

$$\frac{\lambda_o K}{a\sqrt{T}}(1 - e^{-aT}) < \frac{a^2\sqrt{T}(1 + e^{-aT})}{K(3aT - 2 + 2e^{-aT} - aTe^{-aT})} \tag{4-58}$$

Rewriting Eq. 4-58 into the form of a transcendental inequality yields

$$\lambda_o K^2 (1 - e^{-aT})(3aT - 2 + 2e^{-aT} - aTe^{-aT}) < a^3 T(1 + e^{-aT}) \qquad (4\text{-}59)$$

Thus, for this example, Eqs. 4-52 and 4-59 are the two inequalities which must be satisfied to insure closed loop stability. Given bounds on a and K, and the value of λ_o, it is necessary to determine the values of T, if they exist, which will satisfy these two inequalities.

Assume for the system considered here that $\lambda_o = 2$, $K = 2$, and a varies between 2 and 8. Since Eqs. 4-52 and 4-59 are transcendental, a range of values of T which will satisfy them both for the extreme variations of a must be found by trial and error. Considering first the value of $a = 2$, it is found after a number of trials that for

$$T = 0.4 \qquad (4\text{-}60)$$

Eqs. 4-52 and 4-59 are

$$\text{Condition (1):} \quad 4.13 < 4.64 \qquad (4\text{-}61)$$

$$\text{Condition (3):} \quad 1.10 < 3.2 \qquad (4\text{-}62)$$

It was found that for $T < 0.4$, the two conditions were also satisfied. Considering next the other extreme value of a, $a = 8$, it was found that the two conditions were also satisfied for $T \leq 0.4$. In particular, for $T = 0.4$, the two conditions were

$$\text{Condition (1):} \quad 5.79 < 213.5 \qquad (4\text{-}63)$$

$$\text{Condition (2):} \quad 17.2 < 204.8 \qquad (4\text{-}64)$$

Therefore, the closed-loop system will be stable for variations of a in the range [2, 8] if a control interval length less than or equal to 0.4 is used.

It should be pointed out that the results obtained above are conservative. By constraining T to be less than some specified value, $K_o(t)$

has been held to the range of values for which the poles of the closed-loop system are within the unit circle of the z-plane. However, since the system is time-varying, it may be possible for the system to be stable even if $K_o(t)$ exceeds the bounds imposed above. In terms of the complex plane, this means the poles of the system may move outside the unit circle during a portion of the time of system operation. As long as these poles do not remain outside of the unit circle, however, it is still possible for the closed-loop system to be stable.

This section has indicated a method for analytical stability analysis of the sub-class of predictive adaptive controls in which a one-term approximation of the control variable is used. It is clear that other methods such as the Nyquist criterion could also have been used, and the presentation here is by no means exhaustive.

For the general class of predictive adaptive controls, however, no known analytical methods of stability analysis are applicable. The difficulty lies primarily in the fact that it is not possible to obtain a transfer function for the controller portion of the system. Hence, in a particular application where more than a one-term approximation of the control variable is used, analog or digital computer studies may be employed to study stability characteristics.

4.4 Prediction Accuracy Limitations

Reference to the optimum control configuration of Fig. 4-2 indicates it is necessary to consider another factor in addition to stability to establish the control interval length. This second factor is prediction accuracy.

The basic function of the controller is to generate the control variable by operating upon an estimate of future system error. Hence, the accuracy of this estimate is a primary consideration in system design.

While the subjects of prediction and prediction accuracy are treated in detail in Chapter 5, the salient features of prediction accuracy will be discussed here since control interval length is related to prediction accuracy as well as to stability as shown in Section 4.3.

For purposes of illustration, only a functional solution of the prediction accuracy requirement will be given in this section with the details left to Chapter 5. Consider the predictor in Fig. 4-14 which has an input x(t), an actual output y(t), and the desired output x(t + T), where T is the prediction interval length. The instantaneous error in prediction is defined by

$$e_p(t) = x(t + T) - y(t) \tag{4-65}$$

The mean-square prediction error is then given by

$$\overline{e_p^2(t)} = \overline{[x(t + T) - y(t)]^2} \tag{4-66}$$

where the bar indicates the averaging operation. If the input signal x(t) can be characterized by a finite number of parameters $(\alpha_1, \ldots, \alpha_R)$, R an integer, the output signal y(t) will also depend upon these parameters, and, in addition, upon the set of parameters $(\beta_1, \ldots, \beta_Q)$, Q an integer, characterizing the prediction operation, and upon the prediction interval length T [25, p. 432]. Thus, the mean-square prediction error will be some function of these same quantities and will be defined by some relation

$$\overline{e_p^2(t)} = f(\alpha_1, \ldots, \alpha_R, \beta_1, \ldots, \beta_Q, T) \tag{4-67}$$

Let L be an upper bound on the amount of mean-square error which can be tolerated in the predictor output,

$$\overline{e_p^2(t)} \leq L \tag{4-68}$$

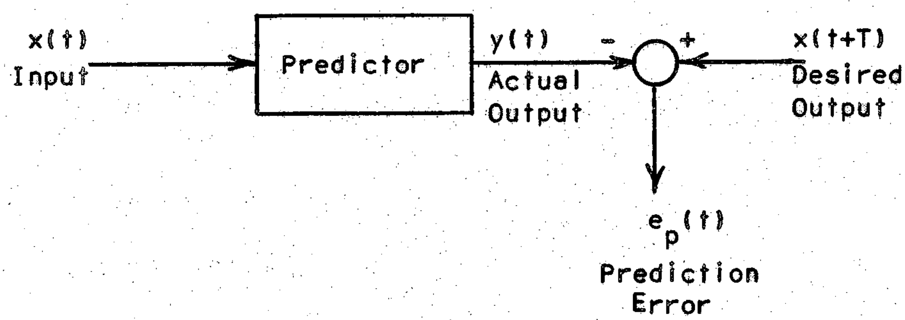


Fig. 4-14

Prediction Operation Block Diagram.

Then, from Eq. 4-67, the prediction accuracy requirement becomes

$$f(\alpha_1, \ldots, \alpha_R, \beta_1, \ldots, \beta_Q, T) \leq L \qquad (4\text{-}69)$$

With the input signal parameters ($\alpha_1, \ldots, \alpha_R$) and the predictor parameters ($\beta_1, \ldots, \beta_Q$) known, Eq. 4-69 represents a relation with one unknown quantity T. If this relation can be solved to find values of T for which it is satisfied, it is clear this solution will in general depend upon the parameters ($\alpha_1, \ldots, \alpha_R$), ($\beta_1, \ldots, \beta_Q$) and upon L. Usually functions of the form given in Eq. 4-67 are monotonic non-decreasing functions of T for signals encountered in practice. Therefore, the solution of Eq. 4-69 can be indicated formally by

$$T \leq g(\alpha_1, \ldots, \alpha_R, \beta_1, \ldots, \beta_Q, L) \qquad (4\text{-}70)$$

where g is some function such that

$$f\left[\alpha_1, \ldots, \alpha_R, \beta_1, \ldots, \beta_Q, g(\alpha_1, \ldots, \alpha_R, \beta_1, \ldots, \beta_Q, L)\right] \leq L \qquad (4\text{-}71)$$

Eq. 4-70 places an upper bound on the control interval length in order to satisfy the prediction accuracy requirement. In any design problem it is necessary to consider both stability and prediction accuracy in selecting the control interval length.

CHAPTER 5

PREDICTOR AND CONTROLLER DESIGN CONSIDERATIONS

This chapter presents the salient features of the predictor and controller designs for predictive adaptive controls. Predictor design is outlined on the basis of the classical Wiener-Lee [25] theory and linear extrapolation. Controller design is presented in terms of the fundamental controller parameters which are: (1) the system error weighting factor $\lambda(t)$, (2) the control interval length T, and (3) the order N of the polynomial sum approximation of the control variable m(t).

5.1 Predictor Design

For statistical input signals the design of the predictors needed for the class of adaptive controls developed in this work will be based on the classical Wiener-Lee theory.

Since Wiener-Lee prediction theory leads to the design of linear predictors, the operations of prediction and difference commute and the predicted error signals may be written

$$\left[c_o(t) - c(t)\right]^* = c_o^*(t) - c^*(t) \qquad (5-1)$$

Hence, the block diagram of Fig. 4-2 may be redrawn as in Fig. 5-1.

In order to effect the design of the predictors in terms of Wiener-Lee theory, the spectral densities of the two signals to be predicted must be known _a priori_. In any practical design problem involving statistical signals, the spectral density of the desired response $c_o(t)$ will be known. Let it be denoted by $\Phi_{c_oc_o}(s)$. However, the spectral density $\Phi_{cc}(s)$ of the dynamic process response is not know _a priori_. Since c(t) is the controlled variable and the primary function of controlling it is to make the difference $\left[c_o(t) - c(t)\right]$ as small as possible

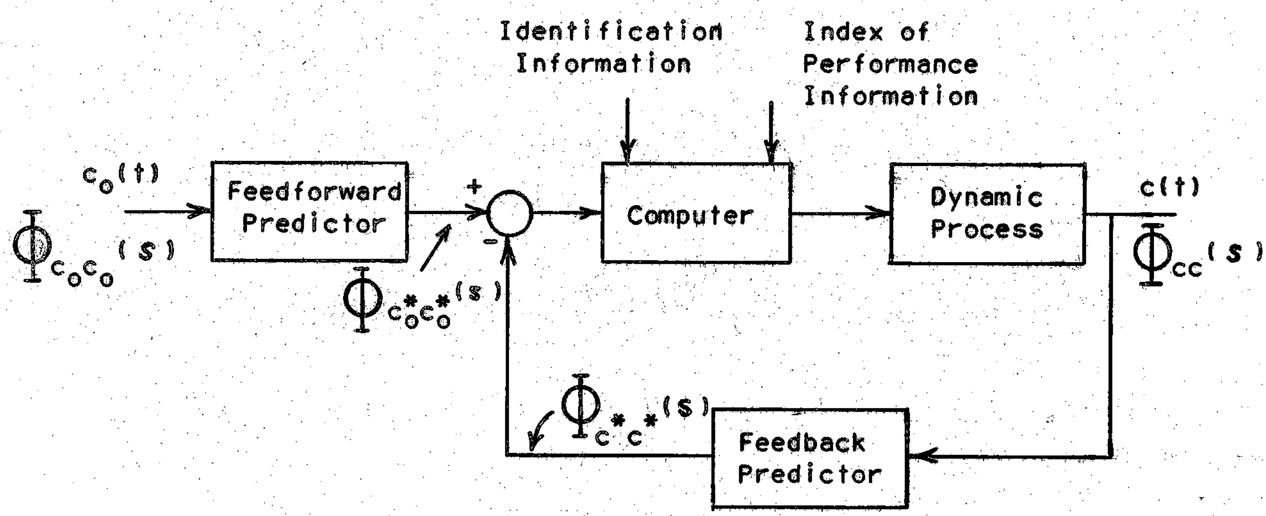


Fig. 5-1

Optimum Control Configuration.

over a long interval of time, a reasonable first assumption on $\Phi_{cc}(s)$ is

$$\Phi_{cc}(s) \approx \Phi_{c_o c_o}(s). \tag{5-2}$$

This assumption will permit a first design of the feedback predictor of the adaptive system. Then, as experience is gained with the system, normal operating records may be employed to obtain better information about the spectral properties of the dynamic process response c(t), [25, Ch. 10] and re-design of the feedback predictor may then be based on this new information. Of course, if operating records of the dynamic process to be controlled are available a priori, these should be employed to carry out the first design.

A review of the design of Wiener-Lee predictors is given in Appendix C with all of the necessary equations. These results will now be applied to obtain the design of the predictors to be used in the experimental work of Chapter 6.

The spectra to be used in this research are of the form

$$\Phi(s) = \frac{a^2}{b^2 + s^2} \tag{5-3}$$

where s is the complex variable $\sigma + j\omega$. By spectral factorization

$$\Phi^+(s) = \frac{a}{b + js} \tag{5-4}$$

Substituting Eq. 5-4 into Eq. C-16 gives

$$\psi(t + T) = \int_{\infty - jv_1}^{\infty + jv_1} \frac{a}{b + js} e^{j(t+T)w} \, dw \tag{5-5}$$

which when evaluated becomes

$$\psi(t + T) = \begin{cases} 2\pi a e^{-b(t+T)} & t > -T \\ 0 & t < -T \end{cases} \tag{5-6}$$

Substitution of Eq. 5-6 into Eq. C-15 yields

$$H_{opt}(s) = \frac{1}{2\pi \frac{a}{b + js}} \int_0^{\infty} 2\pi a e^{-b(t+T)} e^{-jst} dt \qquad (5\text{-}7)$$

$$H_{opt}(s) = e^{-bT} \qquad (5\text{-}8)$$

Hence, the optimum predictor for the spectra to be used in the experimental studies is a simple attenuator.

To determine how prediction error varies with control interval length T, the mean-square prediction error will also be evaluated here utilizing the results given in Appendix C.

From Eq. 5-6,

$$\psi^2(t) = 4\pi^2 a^2 e^{-2bt} \qquad t > 0 \qquad (5\text{-}9)$$

Substituting Eq. 5-9 into Eq. C-27 gives the minimum mean-square prediction error. Thus,

$$\overline{\epsilon^2(t)}\Big|_{min} = 2\pi a^2 \int_0^T e^{-2bt} dt$$

$$= 2\pi a^2 (1 - e^{-2bT}) \qquad (5\text{-}10)$$

Eq. 5-10 is plotted in Fig. 5-2 to show the variation of prediction error with control interval length. The necessity of keeping the control interval length small is obvious.

Eq. 5-10 will now be used with the results of Section 4.4 to determine control interval length in terms of prediction accuracy. Assume it is desired that the mean-square prediction error be less than some number A. This condition then places an upper bound on the prediction error which is expressed by

$$2\pi a^2 (1 - e^{-2bT}) \leqslant A \qquad (5\text{-}11)$$

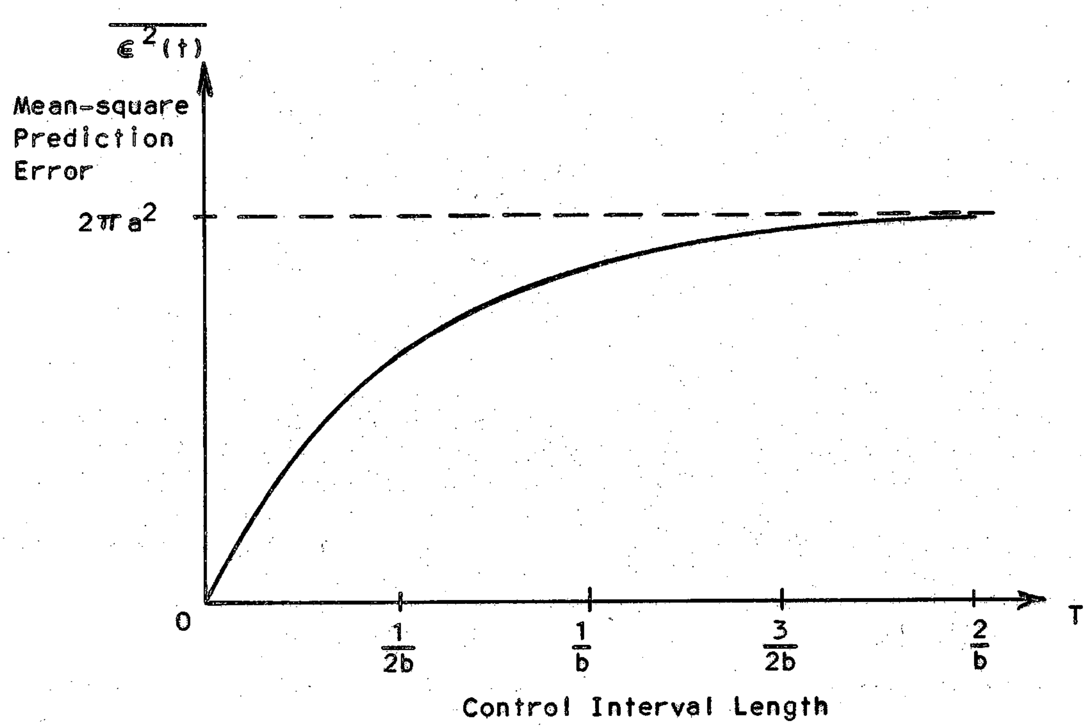


Fig. 5-2

Mean-square Prediction Error as a Function of Control Interval Length.

which readily simplifies to

$$e^{-2bT} \geq 1 - \frac{A}{2\pi a^2} \tag{5-12}$$

Taking the natural logarithm of both sides gives

$$-2bT \geq \ln\left[1 - \frac{A}{2\pi a^2}\right] \tag{5-13}$$

or

$$2bT \leq -\ln\left[1 - \frac{A}{2\pi a^2}\right] \tag{5-14}$$

Assuming the signal spectrum parameters a and b are known, the condition on the control interval length T becomes

$$T \leq -\frac{1}{2b}\ln\left[1 - \frac{A}{2\pi a^2}\right] \tag{5-15}$$

5.2 Extrapolation

In some design applications, the spectral density of $c_o(t)$ may not be available a priori. Moreover, it may be known that $c_o(t)$ is a polynomial type signal. In such cases, extrapolation may be used to obtain $\left[c_o(t) - c(t)\right]^*$.

Since an ideal lead having the transfer function e^{Ts} is not physically realizable, it is necessary to employ an approximation to this ideal lead. If the control interval length is kept small and the frequencies of the signals within the system are low such that $|Ts| << 1$ where $s = \sigma + j\omega$, the first two terms of the expansion [31, p. 100]

$$e^{Ts} = 1 + \sum_{n=1}^{\infty} \frac{(Ts)^n}{n!} \tag{5-16}$$

may be used as the approximation.

The transfer function of the first two terms of Eq. 5-16 can be approximated by a passive lead network such as shown in Fig. 5-3. The actual transfer function of this network is

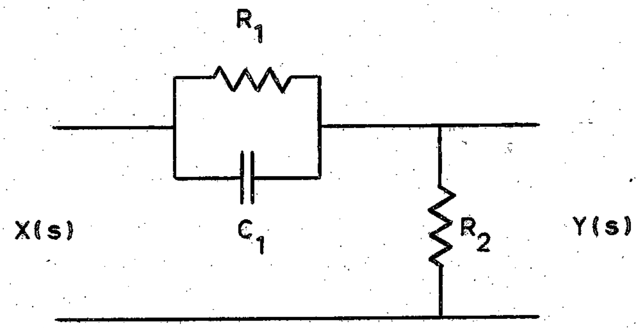


Fig. 5-3

Lead Network Approximation of Extrapolator.

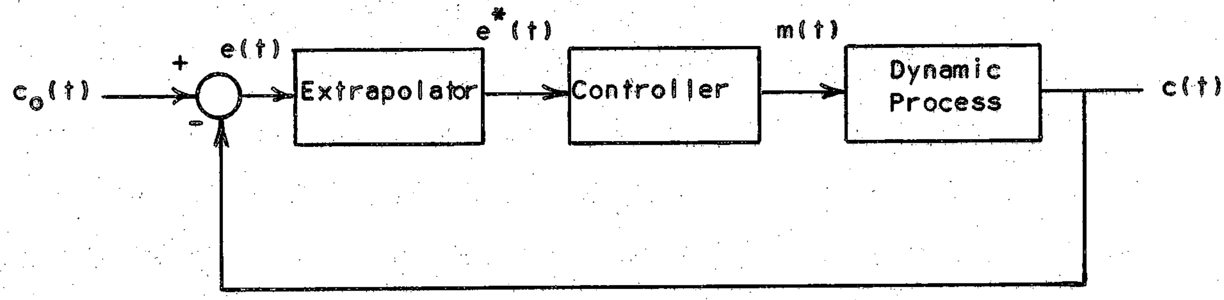


Fig. 5-4

Control Configuration Employing Extrapolation.

$$\frac{Y(s)}{X(s)} = \alpha \frac{1 + T_1 s}{1 + T_2 s} \qquad T_1 > T_2 \tag{5-17}$$

where $\alpha = \frac{R_2}{R_1 + R_2}$, $T_1 = R_1 C_1$, and $T_2 = \frac{R_1 R_2}{R_1 + R_2} C_1$. Since the d-c gain of the network is less than unity, a gain of $\frac{1}{\alpha}$ must be introduced to compensate for the attenuation. If $T_1 >> T_2$ is chosen, Eq. 5-17 becomes approximately

$$\frac{Y(s)}{X(s)} \approx 1 + T_1 s \tag{5-18}$$

which is the desired transfer function.

When the network of Fig. 5-3 is used in the over-all system, the control configuration assumes the form depicted in Fig. 5-4.

An important factor to consider in using extrapolation is the difficulty which arises when there is appreciable noise present in the control loop. Eq. 5-16 indicates that the approximation of the ideal lead produces an extrapolated signal comprised of the original signal plus T times the first derivative of the original signal. The presence of the differentiation will always worsen the noise conditions in the system and may even cause amplifier saturation.

The use of extrapolation has been presented here as an alternative to Wiener-Lee predictor design. The purpose of this section has been to give a treatment of the predictor design in terms of extrapolation, and to point out the difficulty associated with its use. For the experimental investigations to be given in the next chapter, it will be assumed that the spectrum of $c_0(t)$ is known. Hence, extrapolation will not be investigated experimentally.

5.3 Controller Design

With the predictor design known, the basis of the synthesis procedure is clear and the over-all system assumes the configuration of

Fig. 4-2. However, before the synthesis of the system can be completed, it is necessary to consider the selection of three fundamental parameters of the controller. These parameters are:

1. The system error weighting factor $\lambda(t)$.
2. The control interval length T.
3. The order N of the approximation of the control variable.

The purpose of this section is to present a qualitative and quantitative discussion of how these parameters can be selected.

<u>System Error Weighting Factor</u>. Examination of the control equations, Eqs. 4-9 - 4-11, and the optimum control configuration, Fig. 4-2, reveals that the choice of $\lambda(t)$ is somewhat arbitrary. The only quantitative restriction, which was given in Section 2.3, is that $\lambda(t) > 0$.

While the designer has some freedom in the choice of $\lambda(t)$, his selection should be governed primarily by the aims or goals of control. (See Sections 1.1 and 2.3). For example, if errors occurring near the end of each control interval are more important than those near the beginning of the interval, then $\lambda(t)$ could assume the forms

$$\lambda(t) = At^n \tag{5-19}$$

or

$$\lambda(t) = Be^{\gamma t} \tag{5-20}$$

where A, B, and γ are positive constants. Linear and nonlinear combinations of Eqs. 5-19 and 5-20 are also possible. Because of the infinity of combinations which exist as choices of $\lambda(t)$, only one will be selected for use in the experimental work which follows. System error will be weighted uniformly over each control interval by taking

$$\lambda(t) = \lambda_0 \tag{5-21}$$

where λ_0 is a constant. Response characteristics for different values of λ_0 will then be investigated.

Control Interval Length. Quantitative determination of the control interval length T to be used in a particular design application has been presented in detail in Sections 4.3 and 4.4. The two basic considerations used in the analyses presented were stability and prediction accuracy, respectively. A third factor which depends on parameter drift is discussed here.

This third factor which comes to bear on the problem of selecting the control interval length is the drift rate of the process parameters. Since the control coefficients m_k are generated during one interval for use at the beginning of and throughout the succeeding interval, the sampling instants are actually the points in time at which adaptation occurs. Hence, the choice of T governs the frequency of adaptation. If the process parameters change considerably during a control interval, it is clear the adaptation which occurred at the beginning of that interval will be inadequate for the parameter changes. Deciding how much parameter drift should be tolerated during a given control interval is, as in the case of choosing $\lambda(t)$, somewhat subjective. However, it seems reasonable that T should be chosen small enough so that parameter drift is less than 5% per control interval.

Number of Terms in Polynomial Approximations. The optimization procedure given in Section 4.1 provides no means of choosing the order N of the control variable approximation

$$m(t) = \sum_{k=0}^{N} m_k p_k(t) \quad . \qquad (4-11)$$

Intuitively, one would expect a higher-order dynamic process to require more channels in the controller than would a lower-order process.

Actually, Eq. 4-11 represents an infinite series and the optimization of Section 4.1 is valid only if the series Eq. 4-11 converges absolutely. Thus, the question of absolute convergence is a basic consideration in the design of the controller.

The answer to the question of absolute convergence of the series, Eq. 4-11, is not at all obvious for the general case from the control equations

$$K_k(t) = \lambda(t) \int_0^t p_k(\tau)\, w(t,\tau)\, d\tau \qquad (4\text{-}9)$$

and

$$m_k = \int_0^T K_k(t) \left[c_o(t) - c(t)\right]^* dt \qquad (4\text{-}10)$$

for $k = 0, 1, \ldots, N$. The problem is compounded further by the increasing complexity of the polynomials $p_k(t)$ for increasing values of k. (See Appendix A.)

A method for determining an approximate value of the number of controller channels needed for a particular control application will be presented here and illustrated with examples. Two of these examples will be investigated experimentally. The objective of the method is to obtain an engineering estimate of the number of terms needed in Eq. 4-11 in order to achieve adequate control.

The method is based on a direct application of Eqs. 4-9 - 4-11 and the following assumptions:

1. The dynamic process is assumed to be at the extremum of its characteristics corresponding to the most unstable process configuration.
2. The predicted system error, $\left[c_o(t) - c(t)\right]^*$, is assumed to be bounded by some number A during the control interval.

In practice, these two assumptions correspond to a step function input of desired response at the same time that the poles of the dynamic process transfer function are in the right-half plane or on the $j\omega$-axis.

In applying the method, the coefficients m_k, $k = 0, 1, \ldots, N$, are evaluated under assumptions 1 and 2, and the series Eq. 4-11 expanded in a power series in t to examine the behavior of the latter series' coefficients. Examples for a first-order, a second-order, and a third-order dynamic process are presented below.

Example 5.1

Consider the dynamic process characterized by the differential equation

$$\frac{dc}{dt} + a(t)\, c(t) = K\, m(t)$$

where K is a constant and $0 \leq a(t) \leq 1$. The process is on the verge of instability when $a(t) = 0$.

Assuming system error is weighted uniformly over each control interval so that $\lambda(t) = \lambda_o$, a constant, and employing Eqs. 4-9 and 4-10 for $k = 0, 1, 2, 3, 4$ gives the first five coefficients:

$$m_o = \frac{1}{2}\lambda_o AKT^{\frac{3}{2}}$$

$$m_1 = -\frac{\sqrt{3}}{6}\lambda_o AKT^{\frac{3}{2}}$$

$$m_2 = 0$$

$$m_3 = 0$$

$$m_4 = 0$$

These results indicate a two term approximation in Eq. 4-11 should be sufficient to control the first-order process. Expanding Eq. 4-11 gives

$$m(t) = (\frac{1}{2}\lambda_o AKT + \frac{1}{2}\lambda_o AKT)t^o - \lambda_o AKt^1$$

where the two terms in the coefficient of t^o are due to m_o and m_1, respectively. These two terms are equal in magnitude indicating that m_1 is as important as m_o in generating the control signal. In terms of classical control system design, however, it is known that this first-order process can be compensated by a pure gain which can be provided by using only a one-term approximation of the control signal.

Example 5.2

Consider the second-order dynamic process characterized by the differential equation

$$\frac{d^2c}{dt^2} + a(t)\,\frac{dc}{dt} + 4c(t) = 4m(t)$$

where $0 \leqslant a(t) \leqslant 8$. The process is on the verge of instability when $a(t) = 0$ which corresponds to zero damping. The locus of the poles of the transfer function of the process is given in Fig. 5-5.

Again assuming uniform weighting of system error and employing Eqs. 4-9 and 4-10 under assumptions 1 and 2 above for k = 0, 1, 2, 3, 4, gives the first five coefficients:

$$m_o = \frac{\lambda_o A}{\sqrt{T}}(T - \tfrac{1}{2}\sin 2T)$$

$$m_1 = \frac{\lambda_o A\sqrt{3}}{2\sqrt{T}}\left[\frac{1}{T}(\cos 2T-1) + \sin 2T\right]$$

$$m_2 = \frac{6\lambda_o A\sqrt{5}}{T^2\sqrt{T}}\left(\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right) - \frac{6\lambda_o A\sqrt{5}}{T\sqrt{T}}\left[\frac{T^2}{2} + \frac{1}{4}(\cos 2T-1)\right]$$
$$+ \frac{\lambda_o A\sqrt{5}}{\sqrt{T}}(T - \tfrac{1}{2}\sin 2T)$$

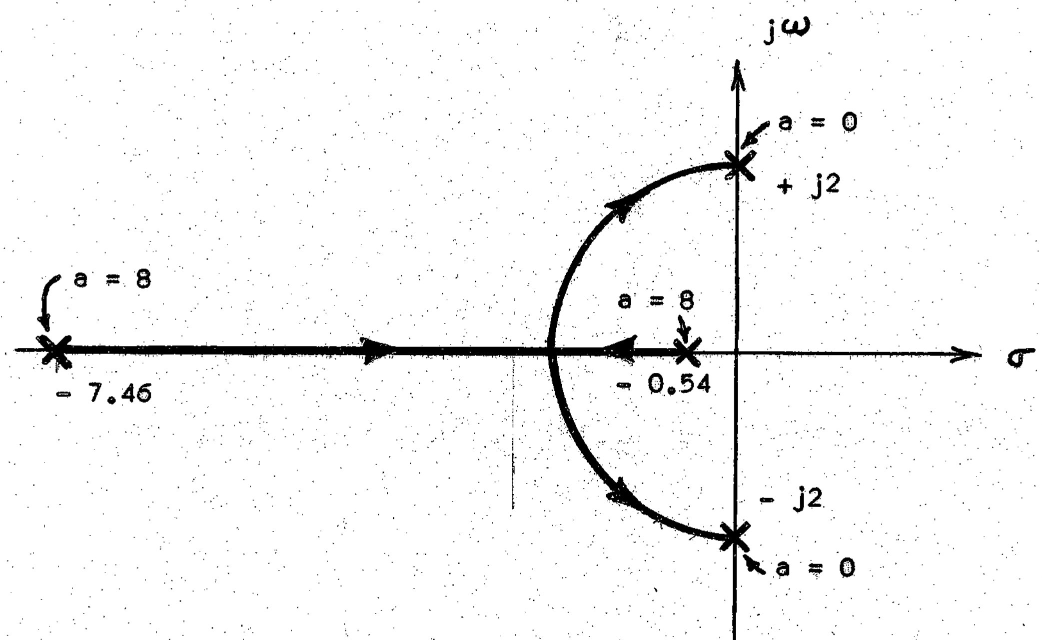


Fig. 5-5

Locus of Poles of Second-order Process as a Function of Process Parameter.

$$m_3 = \frac{60\lambda_o A\sqrt{7}}{T^3\sqrt{T}}\left[\frac{T^4}{12} - \frac{T^2}{4} - \frac{1}{8}(\cos 2T - 1)\right] - \frac{30\lambda_o A\sqrt{7}}{T^2\sqrt{T}}\left(\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right)$$

$$+ \frac{12\lambda_o A\sqrt{7}}{T\sqrt{T}}\left[\frac{T^2}{2} + \frac{1}{4}(\cos 2T - 1)\right] - \frac{\lambda_o A\sqrt{7}}{\sqrt{T}}\left(T - \frac{1}{2}\sin 2T\right)$$

$$m_4 = \frac{630\lambda_o A}{T^4\sqrt{T}}\left(\frac{T^5}{15} - \frac{T^3}{3} + \frac{T}{2} - \frac{1}{4}\sin 2T\right) - \frac{1260\lambda_o A}{T^3\sqrt{T}}\left[\frac{T^4}{12} - \frac{T^2}{4} - \frac{1}{8}(\cos 2T - 1)\right]$$

$$+ \frac{270\lambda_o A}{T^2\sqrt{T}}\left(\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right) - \frac{60\lambda_o A}{T\sqrt{T}}\left[\frac{T^2}{2} + \frac{1}{4}(\cos 2T - 1)\right]$$

$$+ \frac{3\lambda_o A}{\sqrt{T}}\left(T - \frac{1}{2}\sin 2T\right)$$

Because of the complexity of these expressions, it is not possible to draw any conclusions about the behavior of coefficients for all values of T. However, because of stability and prediction accuracy requirements as discussed earlier, only the shorter control interval lengths are of interest. Therefore, making the assumption $T << 1$, the coefficients are given approximately by the expressions:

$$m_o = \frac{2}{3}\ \lambda_o A T^{\frac{5}{2}}$$

$$m_1 = -\frac{\sqrt{3}}{3}\ \lambda_o A T^{\frac{5}{2}}$$

$$m_2 = \frac{\sqrt{5}}{15}\ \lambda_o A T^{\frac{5}{2}}$$

$$m_3 = \frac{\sqrt{7}}{105}\ \lambda_o A T^{\frac{9}{2}}$$

$$m_4 = -\frac{1}{315}\ \lambda_o A T^{\frac{9}{2}}$$

Observe that the coefficients m_o, m_1, and m_2 are of order $T^{\frac{5}{2}}$, whereas the coefficients m_3 and m_4 are of order $T^{\frac{9}{2}}$. Since the above expressions are valid for $T << 1$, it is clear the m_3 and m_4 are significantly less than the first three coefficients.

By substituting the above set of coefficients into Eq. 4-11, expanding, and summing the coefficients of like powers of t, the power series expansion for m(t) based on a five-term approximation is obtained. It is then possible to determine the contribution of each m_k, k = 0, 1, 2, 3, 4 to the power series for m(t). The results are summarized in Table 5-1 where the contribution of each m_k to each coefficient of t^n, n = 0, 1, 2, 3, 4 is made clear. Thus, the coefficient of t^o is the sum of the terms in the first column, that of t^1 the sum of the terms in the second column, and so on.

Since $T << 1$, observe that only the top three terms of column 1 are significant in the coefficient of t^o, the top two terms of column 2 are significant in the coefficient of t^1, and only the top term of column 3 is significant in the coefficient of t^2. Observe also that the terms in the fourth and fifth columns are multiplied by t^3 and t^4, respectively, where $0 \leq t \leq T$, and, therefore, their maximum and minimum values are of order T^4, whereas the above coefficients have maxima and minima of order T^2.

These results indicate a three-term approximation, N = 2, in Eq. 4-11 will give adequate control for this second-order process. However, assumption 2 assumes a constant predicted error and does not account for rapid but continuous changes in predicted error in the control interval. Hence, the result N = 2 is a conservative figure and it is to be expected that N = 3, i.e., four channels in

Control Coefficient	Contribution of control coefficient to total coefficient in power series of m(t)				
	t^o	t^1	t^2	t^3	t^4
m_o	$\frac{2}{3}\lambda_o AT^2$				
m_1	$\lambda_o AT^2$	$-2\lambda_o AT$			
m_2	$\frac{1}{3}\lambda_o AT^2$	$-2\lambda_o AT$	$2\lambda_o A$		
m_3	$-\frac{1}{15}\lambda_o AT^4$	$\frac{4}{5}\lambda_o AT^3$	$-2\lambda_o AT^2$	$\frac{4}{3}\lambda_o AT$	
m_4	$-\frac{1}{105}\lambda_o AT^4$	$\frac{4}{21}\lambda_o AT^3$	$-\frac{6}{7}\lambda_o AT^2$	$\frac{4}{3}\lambda_o AT$	$-\frac{2}{3}\lambda_o A$

TABLE 5-1

Tabulation of the Contribution of Each Control Coefficient to the Total Coefficient in the Power Series of the Control Variable m(t) for $T << 1$.

the controller, will be needed. This fact will be demonstrated experimentally.

Example 5.3

Consider the third-order dynamic process characterized by the differential equation

$$\left[\frac{d}{dt} + a(t)\right]\left[\frac{d^2}{dt^2} + b(t)\frac{d}{dt} + 4\right] c(t) = 4m(t)$$

where $0 \leq a(t) \leq 10$ and $0 \leq b(t) \leq 8$. The process is on the verge of instability when $a(t) = b(t) = 0$. The movement of the poles of the dynamic process transfer function may be assumed as in Fig. 5-6.

Assuming $\lambda(t) = \lambda_o$ and utilizing Eqs. 4-9 and 4-11 under assumptions 1 and 2 for k = 0, 1, 2, 3, 4, 5 gives the first six coefficients:

$$m_o = \frac{\lambda_o A}{2\sqrt{T}}\left[T^2 + \frac{1}{2}(\cos 2T - 1)\right]$$

$$m_1 = \frac{\lambda_o A\sqrt{3}}{\sqrt{T}}\left[\frac{T^3}{3} - \frac{1}{2} + \frac{1}{4T}\sin 2T - \frac{T^2}{2} - \frac{1}{4}(\cos 2T - 1)\right]$$

$$m_2 = \frac{3\lambda_o A\sqrt{5}}{2T^2\sqrt{T}}\left[\frac{T^4}{3} - T^2 - \frac{1}{2}(\cos 2T - 1)\right] - \frac{3\lambda_o A\sqrt{5}}{T\sqrt{T}}\left(\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right)$$
$$+ \frac{\lambda_o A\sqrt{5}}{\sqrt{T}}\left[\frac{T^2}{2} + \frac{1}{4}(\cos 2T - 1)\right]$$

$$m_3 = \frac{15\lambda_o A\sqrt{7}}{T^3\sqrt{T}}\left[\frac{T^5}{15} - \frac{T^3}{3} + \frac{T}{2} - \frac{1}{4}\sin 2T\right] - \frac{15\lambda_o A\sqrt{7}}{2T^2\sqrt{T}}\left[\frac{T^4}{3} - T^2 - \frac{1}{2}(\cos 2T - 1)\right]$$
$$+ \frac{6\lambda_o A\sqrt{7}}{T\sqrt{T}}\left[\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right] - \frac{\lambda_o A\sqrt{7}}{2\sqrt{T}}\left[T^2 + \frac{1}{2}(\cos 2T - 1)\right]$$

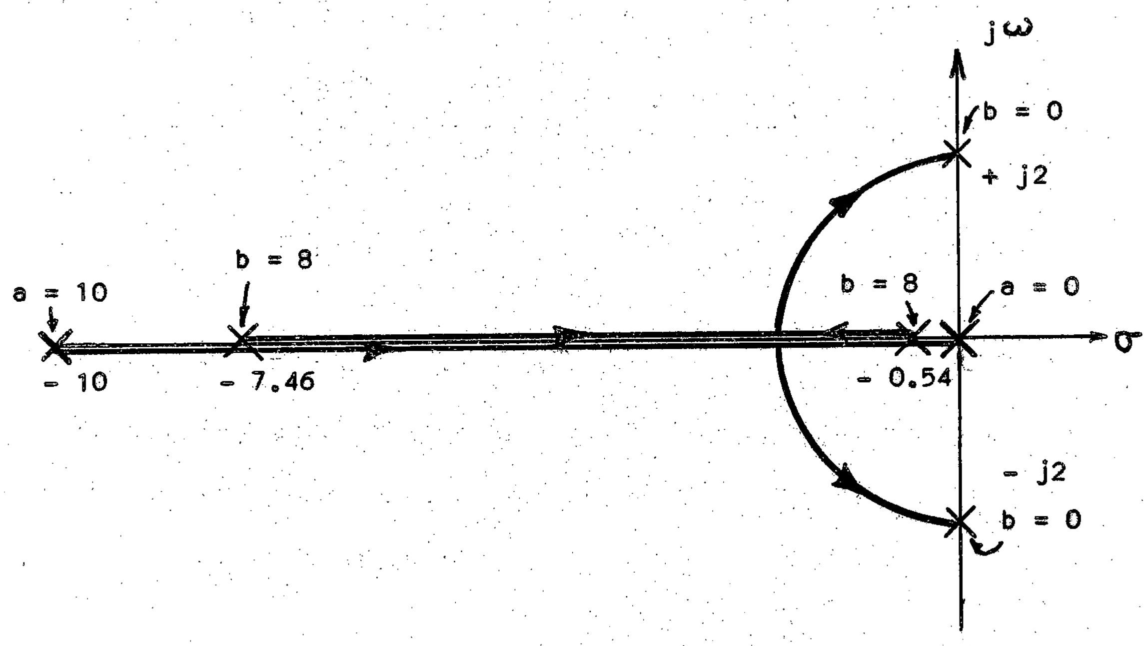


Fig. 5-6

Locus of Poles of Third-order process
as a Function of Process Parameters.

$$m_4 = \frac{315\,\lambda_o A}{T^4\sqrt{T}}\left[\frac{T^6}{45} - \frac{T^4}{6} + \frac{T^2}{2} + \frac{1}{4}(\cos 2T - 1)\right] - \frac{315\,\lambda_o A}{T^3\sqrt{T}}\left[\frac{T^5}{15} - \frac{T^3}{3} + \frac{T}{2} - \frac{1}{4}\sin 2T\right]$$

$$+ \frac{135\,\lambda_o A}{2T^2\sqrt{T}}\left[\frac{T^4}{3} - T^2 - \frac{1}{2}(\cos 2T - 1)\right] - \frac{30\,\lambda_o A}{T\sqrt{T}}\left[\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right]$$

$$+ \frac{3\,\lambda_o A}{2\sqrt{T}}\left[T^2 + \frac{1}{2}(\cos 2T - 1)\right]$$

$$m_5 = \frac{1890\,\lambda_o A\sqrt{11}}{T^5\sqrt{T}}\left[\frac{T^7}{315} - \frac{T^5}{30} + \frac{T^3}{6} - \frac{T^2}{4} + \frac{1}{8}\sin 2T\right] - \frac{945\,\lambda_o A\sqrt{11}}{T^4\sqrt{T}}\left[\frac{T^6}{45} - \frac{T^4}{6} + \frac{T^2}{2} + \frac{1}{4}(\cos 2T - 1)\right]$$

$$+ \frac{420\,\lambda_o A\sqrt{11}}{T^3\sqrt{T}}\left[\frac{T^5}{15} - \frac{T^3}{3} + \frac{T}{2} - \frac{1}{4}\sin 2T\right] - \frac{55\,\lambda_o A\sqrt{11}}{T^2\sqrt{T}}\left[\frac{T^4}{3} - T^2 - \frac{1}{4}(\cos 2T - 1)\right]$$

$$+ \frac{15\,\lambda_o A\sqrt{11}}{T\sqrt{T}}\left[\frac{T^3}{3} - \frac{T}{2} + \frac{1}{4}\sin 2T\right] - \frac{\lambda_o A\sqrt{11}}{2\sqrt{T}}\left[T^2 + \frac{1}{2}(\cos 2T - 1)\right]$$

As for the second-order case, these expressions do not permit one to draw any conclusions about the behavior of the coefficients for all values of the control interval length T. However, for the faster sampling rates, $T << 1$, the above expressions simplify to:

$$m_o = \frac{1}{6}\,\lambda_o A T^{\frac{7}{2}}$$

$$m_1 = -\frac{\sqrt{3}}{10}\,\lambda_o A T^{\frac{7}{2}}$$

$$m_2 = \frac{\sqrt{5}}{30}\,\lambda_o A T^{\frac{7}{2}}$$

$$m_3 = \frac{\sqrt{7}}{210}\,\lambda_o A T^{\frac{7}{2}}$$

$$m_4 = \frac{-1}{630}\,\lambda_o A T^{\frac{11}{2}}$$

$$m_5 = \frac{\sqrt{11}}{20790}\,\lambda_o A T^{\frac{11}{2}}$$

Observe that the coefficients m_o, m_1, m_2, and m_3 are of order $T^{\frac{7}{2}}$, whereas m_4 and m_5 are of order $T^{\frac{11}{2}}$. Hence, with $T << 1$ only the first four coefficients are significant. The contribution of each m_k, k = 0, 1, 2, 3, 4, 5 to each coefficient of the power series of m(t) is given in Table 5-2.

Arguing as in the case of the second-order process, only the top four terms of column 1, the top three terms of column 2, the top two terms of column 3, and the topmost term of column 4 are significant in the formation of the coefficients of the power series of m(t) since $T << 1$. Hence, Table 5-2 gives the conservative value of N = 3, i.e., four channels in the controller. Again, because of assumption 2 and experience with the second-order case, it is to be expected that five channels in the controller will be needed to assure adequate control.

It must be emphasized that the method illustrated in the above three examples is not a technique for determining the value of N which is necessary and sufficient to insure absolute convergence of the series Eq. 4-11. Rather it is a method by which it is possible to obtain an engineering estimate of the number of channels needed in the controller to give adequate control of a given dynamic process.

Control Coefficient	Contribution of control coefficient to total coefficient in power series of m(t)					
	t^0	t^1	t^2	t^3	t^4	t^5
m_0	$\frac{1}{6}A\lambda_oT^3$					
m_1	$\frac{3}{10}A\lambda_oT^3$	$-\frac{3}{5}A\lambda_oT^2$				
m_2	$\frac{1}{6}A\lambda_oT^3$	$-A\lambda_oT^2$	$A\lambda_oT$			
m_3	$\frac{1}{30}A\lambda_oT^3$	$-\frac{2}{5}A\lambda_oT^2$	$A\lambda_oT$	$-\frac{2}{3}A\lambda_o$		
m_4	$-\frac{1}{210}A\lambda_oT^5$	$\frac{2}{21}A\lambda_oT^4$	$-\frac{3}{7}A\lambda_oT^3$	$\frac{2}{3}A\lambda_oT^2$	$-\frac{1}{3}A\lambda_oT$	
m_5	$-\frac{1}{1890}A\lambda_oT^5$	$\frac{1}{63}A\lambda_oT^4$	$-\frac{1}{9}A\lambda_oT^3$	$\frac{8}{27}A\lambda_oT^2$	$-\frac{1}{3}A\lambda_oT$	$\frac{2}{15}A\lambda_o$

TABLE 5-2

Tabulation of the Contribution of Each Control Coefficient to the Total Coefficient in the Power Series of the Control Variable m(t) for $T << 1$.

CHAPTER 6

EXPERIMENTAL STUDIES

The function of the present chapter is to investigate experimentally the response characteristics of some typical control systems employing predictive adaptive control. Various aspects of predictive adaptive control systems' behavior are presented graphically to depict certain limitations as well as advantages of this class of controls.

6.1 Outline of Procedure

The control of a first-order and a second-order dynamic process using predictive adaptive control will be investigated with the aid of an analog computer. The exact nature of the process parameter variations will be given as each system is considered. Also, in order to emphasize the results rather than the details of the simulations, the circuitry necessary to perform the operations of reset integration and sample-and-hold, as well as the complete analog computer diagrams, will be given in Appendix D.

Three basic experiments are performed on each of the systems to evaluate the quality of predictive adaptive control in essentially three different control situations. These experiments are outlined below and measure the following three aspects of control: (1) the ability of the adaptive system to maintain the output at a predetermined constant level, (2) the quality of system transient response for step functions of desired response, and (3) the ability of the system to follow statistical signals, all in the presence of extreme variations of the dynamic process parameters.

1. The first experiment is performed by making the desired response $c_o(t)$ a constant and observing the deviation of the output $c(t)$

from the desired value as the dynamic process parameters vary between their extreme values. The per cent deviation of the output from the desired value is defined by the relation

$$\% \text{ deviation} = \frac{\text{extreme value of } c(t) - \text{desired value of } c_0(t)}{\text{desired value of } c_0(t)} \times 100\% \tag{6-1}$$

and places a measure on the ability of the system to cope with process parameter variations in the steady-state. This type of control is important for chemical processes where it is desired to maintain the quality of output products constant within prescribed limits.

2. The second experiment is performed by applying a step function of desired response $c_0(t)$ to the system and evaluating the character of the transient response in terms of rise time and per cent overshoot. Since process parameters vary during the operations, each test is performed at least three times in order to obtain an average behavior.

   Rise time is defined as the total elapsed time from the application of the step to the time at which the response first reaches the desired level. Per cent overshoot is defined by the relation

$$\% \text{ overshoot} = \frac{\text{maximum value of } c(t) \text{ during transient} - \text{desired value of } c_0(t)}{\text{desired value of } c_0(t)} \times 100\% \tag{6-2}$$

   These two quantities measure the ability of the controller to drive the dynamic process from one equilibrium state to another in the presence of parameter variations.

3. The third experiment is performed by shaping the output of a noise generator to obtain a signal with a known spectrum for

$c_o(t)$. Typical response records are presented to indicate system response for a statistical input signal.

As mentioned in Chapter 5 it will be assumed that system error is to be weighted uniformly over each control interval so that $\lambda(t) = \lambda_o$, λ_o a constant, will be used in the experiments. Reference to Eq. 4-11 reveals that λ_o will then be a scale factor in each of the time-varying gains. Therefore, various values of λ_o will be used to point out how the response characteristics of predictive adaptive controls depend upon this factor which governs the relative weighting of system error with respect to control effort in the index of performance, Eq. 2-4.

An important consideration in the design of predictive adaptive controls is the control interval length T. To demonstrate the effects of this design parameter, the data of experiments 1 and 2 will be presented graphically as a function of T. The error weighting factor λ_o will then be used as a parameter in the presentation of these data.

Since the basic work of this research deals with the modification problem, the identification portion of the complete system is simulated using a model of the process from which the time-varying gains are derived. A block diagram is given in Fig. 6-1 to show the flow of information in the stimulation studies.

6.2 First-order Dynamic Process

The first system to be considered is one involving the control of a dynamic process characterized by the differential equation

$$\frac{dc}{dt} + a(t)\, c(t) = m(t) \qquad (6\text{-}3)$$

The parameter $a(t)$ varies between 1 and 0.1, a range of 10 to 1, in a sawtooth manner at a frequency of 0.08 cps as shown in Fig. 6-2.

The predictor is designed on the assumption that the spectrum of the input $c_o(t)$ is of the form

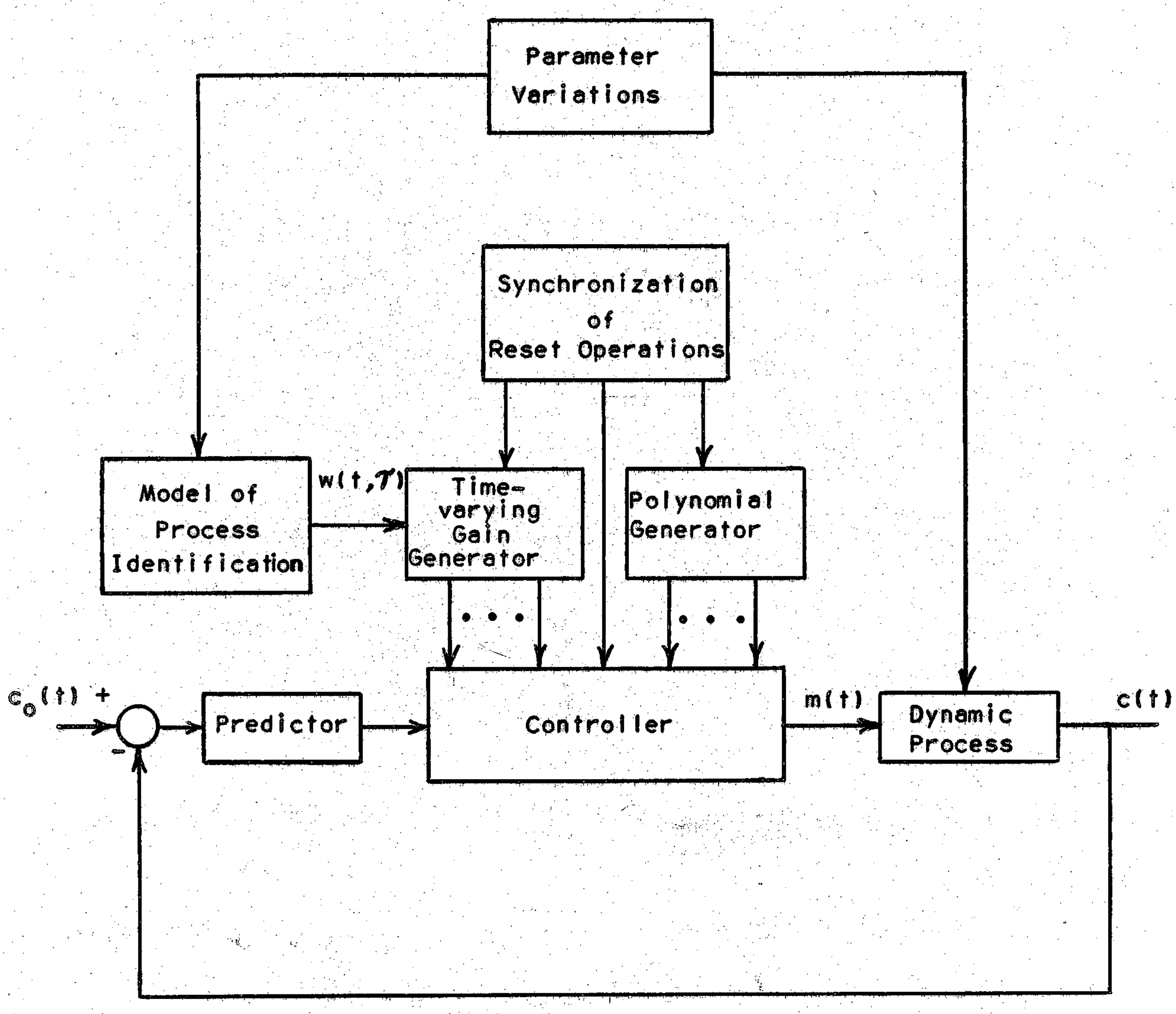


Fig. 6-1

Information Flow Diagram for Simulation Studies.

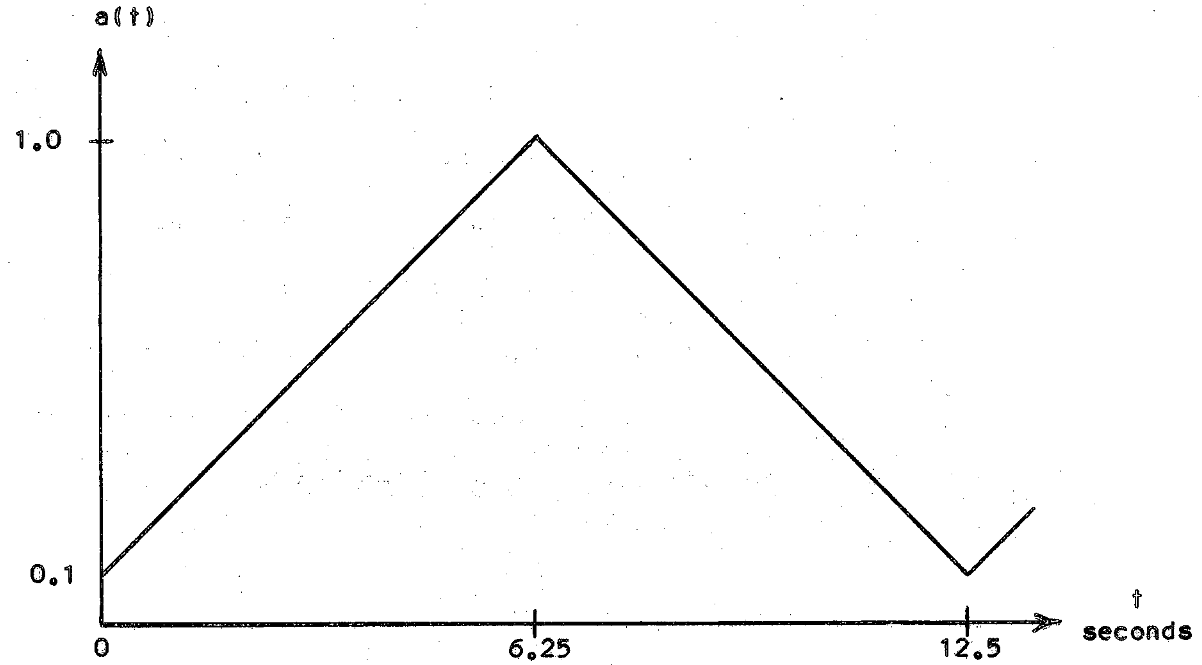


Fig. 6-2

Parameter Variation for First-order Dynamic Process.

$$\Phi(\omega) = \frac{1}{\omega^2 + (\frac{1}{2})^2} \tag{6-4}$$

giving the predictor impulse response

$$h_p(t) = e^{-\frac{T}{2}} \delta(t) \tag{6-5}$$

where T is the control interval length and $\delta(t)$ is the unit impulse function. A spectrum of the form Eq. 6-4 is used in the statistical signal measurements. Moreover, Lee [26, Ch. 8] has shown that a Poisson square wave with an average zero crossing frequency of $\frac{1}{2}$ also has a spectrum of the form Eq. 6-4. A Poisson square wave is defined here as a waveform which alternates between two values E and -E at event points which are statistically independent. The probability of finding n event points in an interval τ is given by the Poisson distribution [26, p. 221]. Thus, the step functions applied to the system can be considered as segments of such a waveform.

Using a one-term approximation of m(t),

$$m(t) = m_o p_o(t) \tag{6-6}$$

data were obtained for control interval lengths $T = \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$ sec., and for $\lambda_o = 2, 3, 4, 6$. The results for these values for the first two experiments are shown in Figs. 6-3, 6-4, and 6-5.

Since it is possible to compensate the first-order process with a pure gain, the curves of Figs. 6-3 and 6-4 indicate improved system performance with increasing λ_o for all values of control interval length T. However, the data for per cent overshoot is not as well-behaved and indicates the need for making the control interval length less than $\frac{1}{4}$ sec. to keep the overshoot for a step input below 20%.

The amount of deviation in the output with a constant input is excessively large for the lower values of λ_o and the control interval

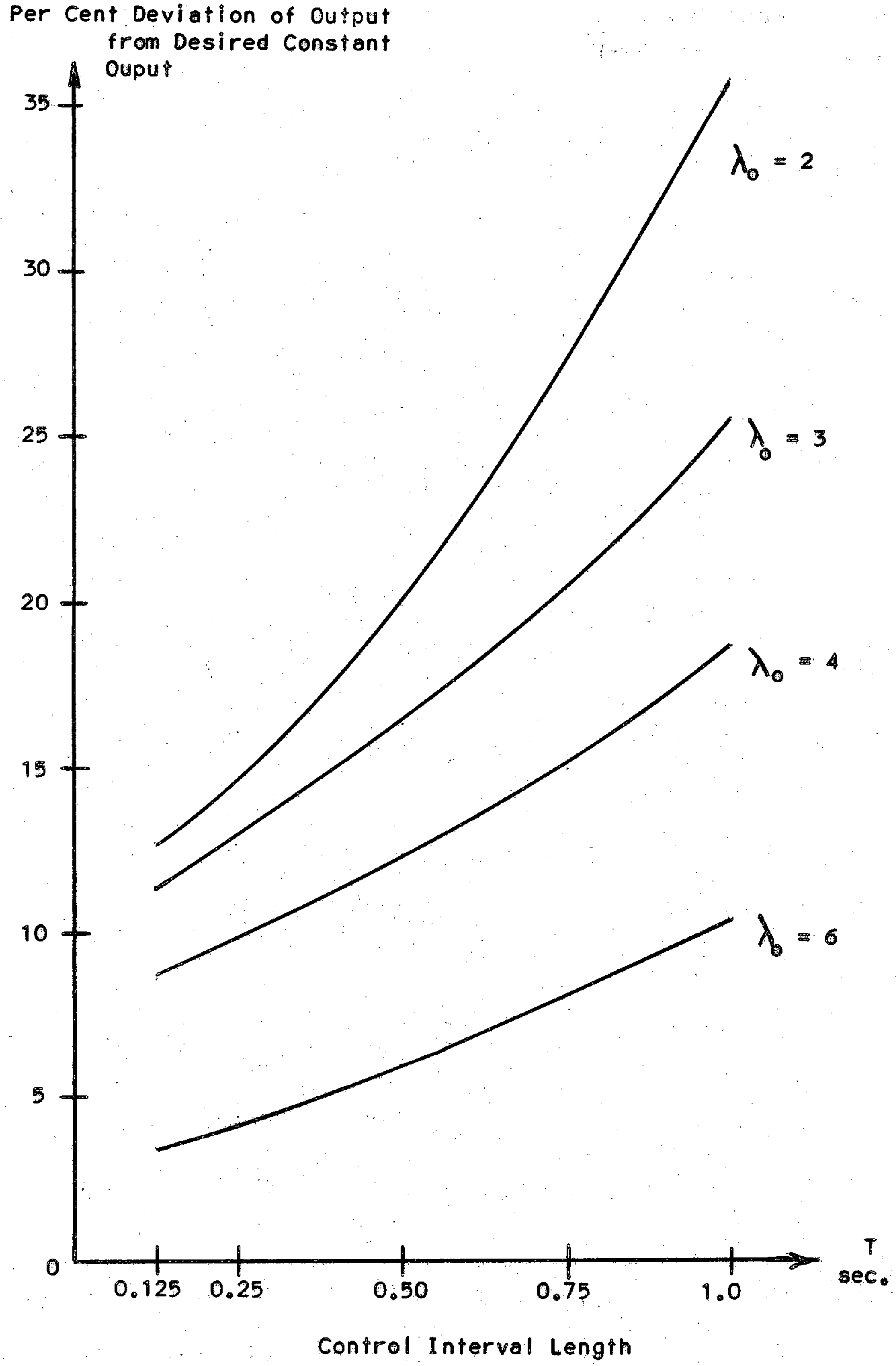


Fig. 6-3

Steady-state Adaptability of First-order Dynamic Process for One-term Approximation of Control Variable.

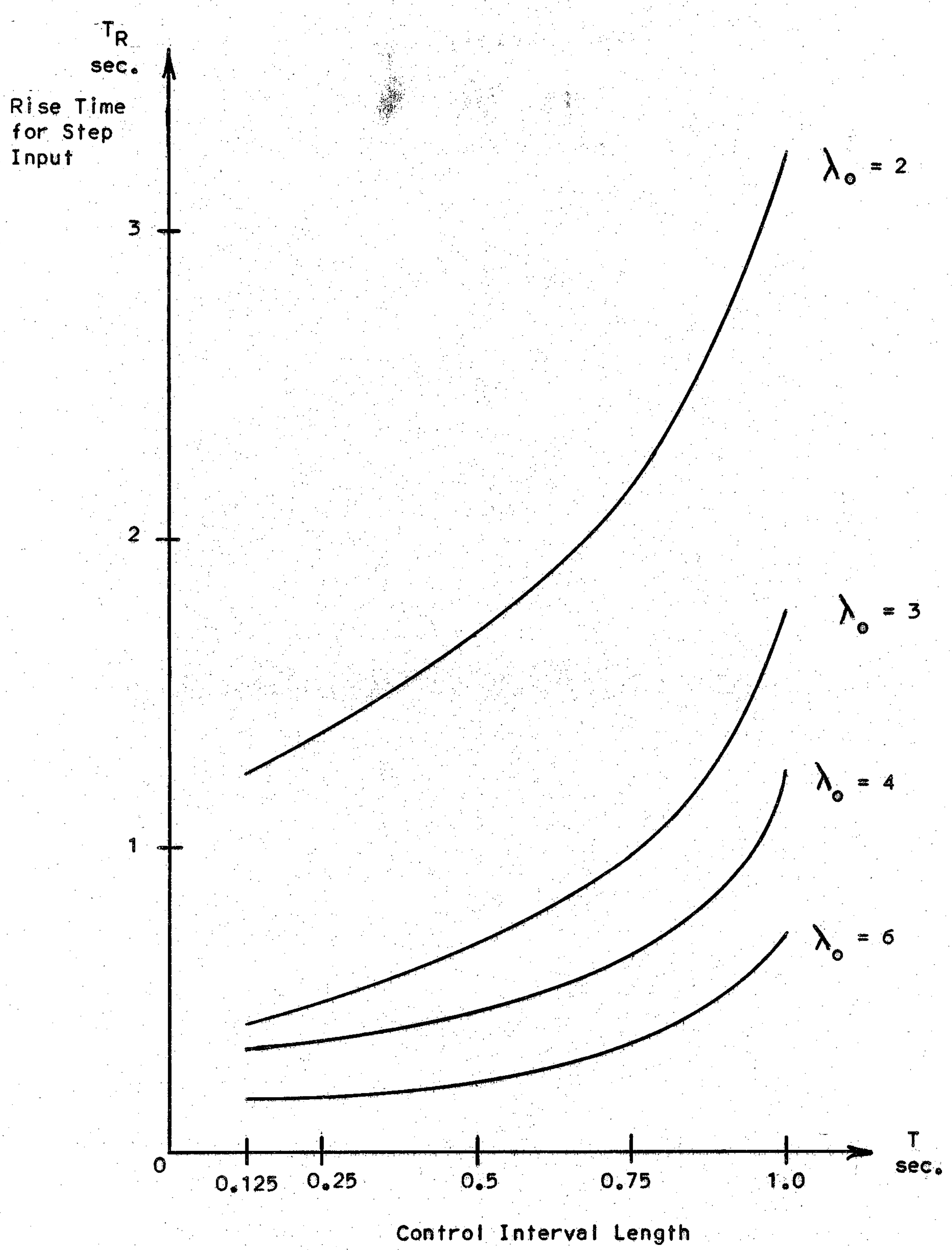


Fig. 6-4

Rise Time for First-order Dynamic Process
for One-term Approximation of Control Variable.

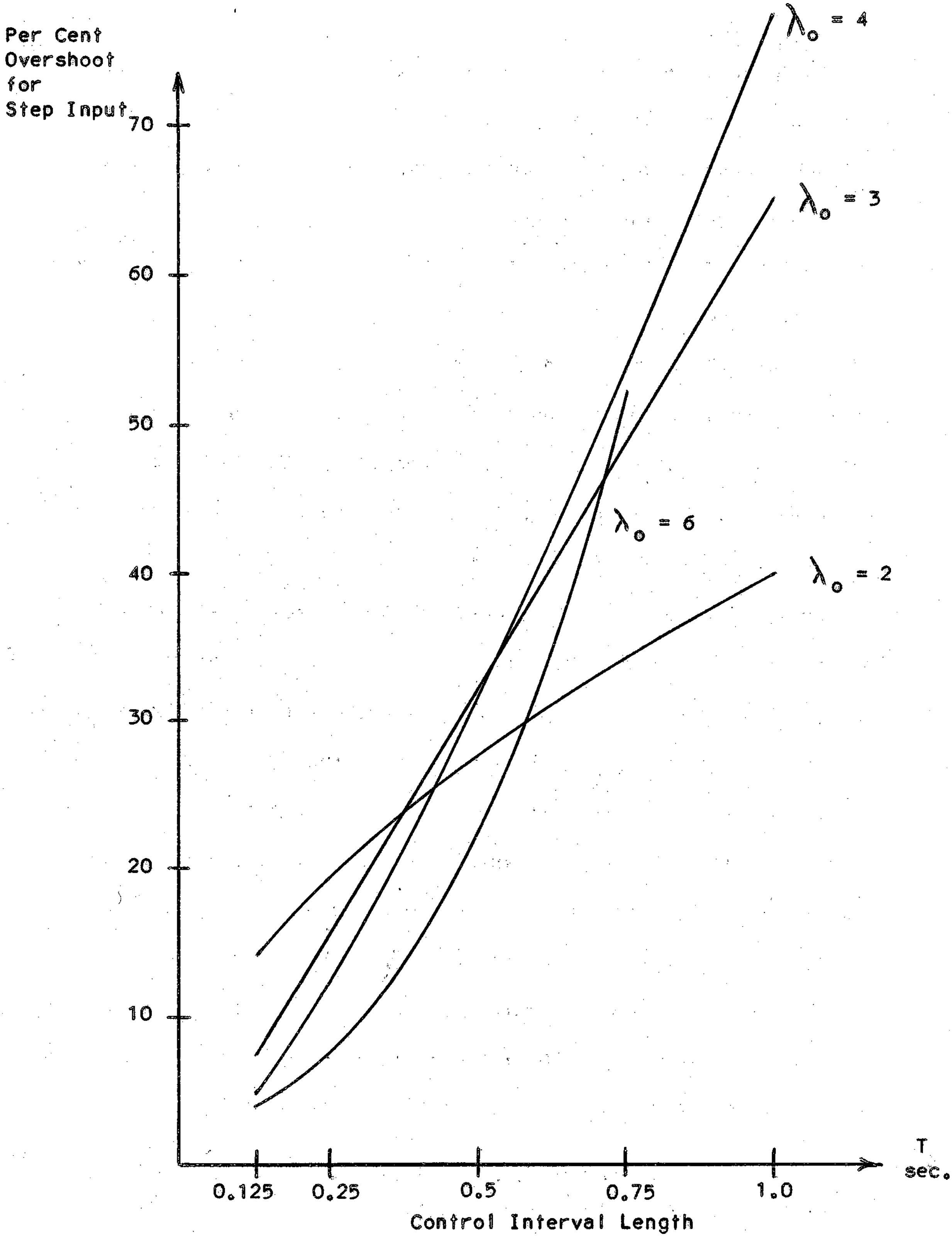


Fig. 6-5

Per Cent Overshoot for First-order Dynamic
Process for One-term Approximation of Control Variable.

lengths above $\frac{3}{4}$ sec. The reason for this was discussed in Section 5.3 where it was pointed out that the choice of control interval length is dependent upon the parameter drift rate. For example, reference to Fig. 6-3 indicates that for $\lambda_o = 2$ the control interval length must be less than 0.5 sec. in order that the output not deviate by more than 20%. Since the parameter drifts between its extrema in 6.25 sec. (Fig. 6-2), a control interval length of 0.5 sec. corresponds to letting the parameter drift 8% between adaptation points. To keep the output from deviating more than 10%, however, values of $\lambda_o \geq 4$ and $T \leq \frac{1}{4}$ sec. are necessary. Values of $T \leq \frac{1}{4}$ sec. correspond to a parameter drift less than or equal to 4% per control interval.

Examination of Figs. 6-3, 6-4, and 6-5 reveals that the quality of control continually improves with increasing λ_o and decreasing T. From a theoretical viewpoint this is gratifying, but from a practical viewpoint it is misleading. Arbitrarily increasing λ_o, which is a factor in the time-varying gain, will cause saturation at the input to the dynamic process. Thus, there exists a practical limitation on the value of λ_o which will depend upon the range of input values of m(t) for which the dynamic process is linear.

The minimum value of T which may be used is governed by the accuracy of the components used in the time-varying gain generator. For a given value a of the parameter a(t), the time-varying gain is given by

$$K_o(t) = \frac{\lambda_o}{a\sqrt{T}}\,(1 - e^{-at}) \qquad (6\text{-}7)$$

for $0 \leq t \leq T$. Using Taylor's formula with Lagrange's form of the remainder [32, p. 114], Eq. 6-7 is

$$K_o(t) = \frac{\lambda_o}{a\sqrt{T}}\left[at - \frac{a^2 e^{-a\tau}}{2!}\,t^2\right] \qquad (6\text{-}8)$$

where $0 \leq \tau \leq t \leq T$. Simplifying Eq. 6-8 gives

$$K_e(t) = \frac{\lambda_o}{\sqrt{T}}\, t + R \qquad (6\text{-}9)$$

where R is the remainder term

$$R = -\frac{a\lambda_o}{2\sqrt{T}}\, e^{-a\tau}\, t^2 \qquad (6\text{-}10)$$

for $0 \leq \tau \leq t \leq T$. Since $e^{-a\tau}$ and t^2 are positive in the interval of interest, the magnitude of the remainder term is bounded from above by

$$|R| \leq \frac{a\lambda_o}{2\sqrt{T}}\, T^2 \qquad (6\text{-}11)$$

Note from Eqs. 6-9 and 6-10 that only the remainder term depends on the process parameter a. Thus, if the components used to generate and detect $K_o(t)$ are insensitive to this remainder term, the controller will be unable to detect variations in the dynamic process and the adaptive capability will be lost. An upper bound on the per cent accuracy required in the equipment may be determined by taking the ratio of the maximum value of $|R|$ and the maximum value of the first term to give

$$\%\ \text{accuracy required} \leq \frac{aT}{2} \times 100\% \qquad (6\text{-}12)$$

For a nominal value of $a = 0.5$ and $T = \frac{1}{16}$ sec. Eq. 6-12 gives

$$\%\ \text{accuracy required} \leq 1.56\% \qquad (6\text{-}13)$$

Two typical step responses for the first-order dynamic process are given in Figs. 6-6 and 6-7 for $\lambda_o = 2$, $T = \frac{1}{4}$ sec., and $\lambda_o = 6$, $T = \frac{1}{16}$ sec., respectively. The desired response $c_o(t)$ and the parameter variation $a(t)$ are also included. Fig. 6-6 shows the large deviations which occur in the response as a result of parameter variation in the steady-state. Fig. 6-7 indicates system response for a number of step changes in the desired response.

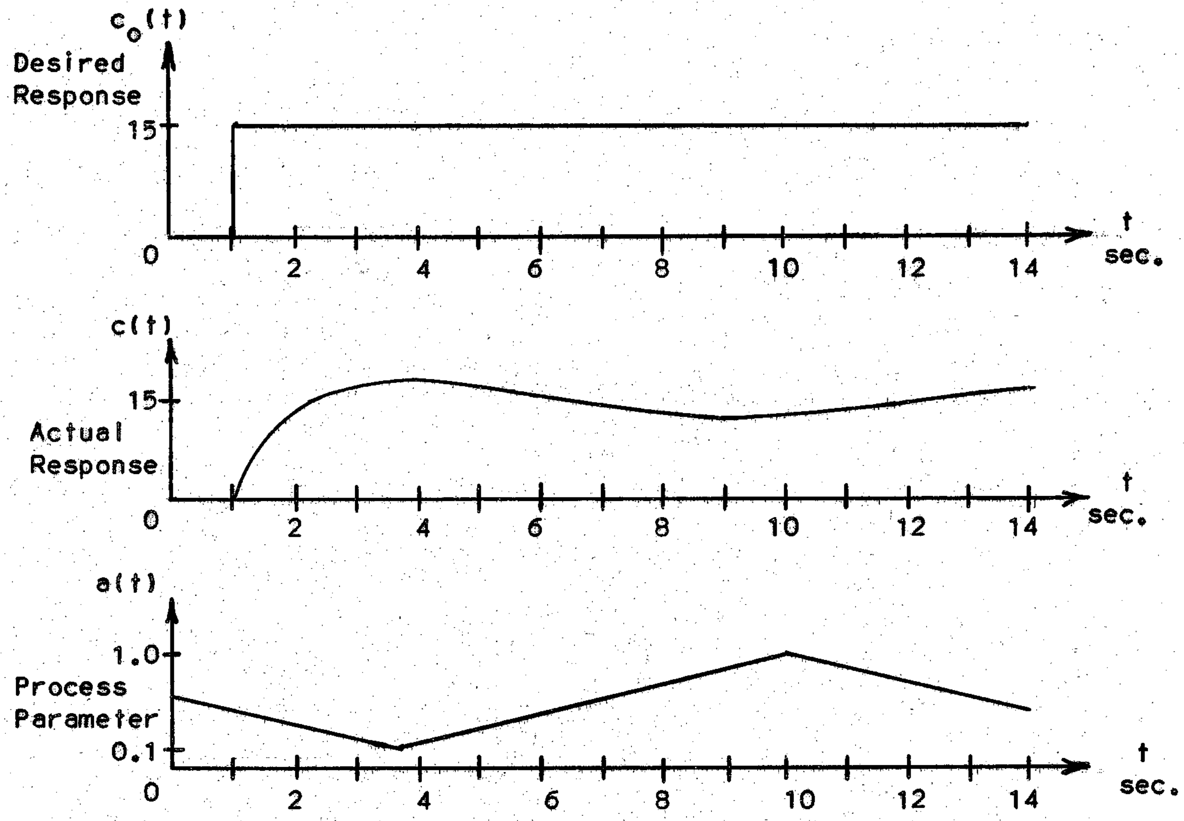


Fig. 6-6

Typical Step Response of First-order Dynamic Process for $\lambda_o = 2$ and $T = \frac{1}{4}$ sec.

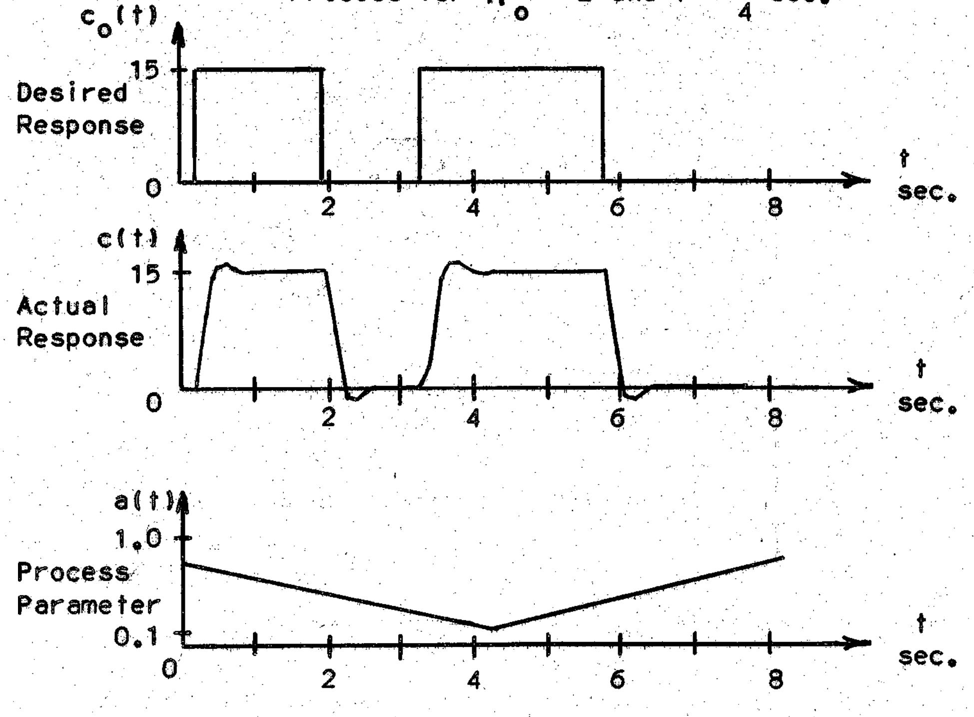


Fig. 6-7

Typical Step Responses of First-order Dynamic Process for $\lambda_o = 6$ and $T = \frac{1}{16}$ sec.

A typical system response for a statistical signal $c_o(t)$ having a spectrum of the form Eq. 6-4 is given in Fig. 6-8 for $\lambda_o = 6$, $T = \frac{1}{16}$ sec. The parameter variations are the same as before. Although there is considerable smoothing, the ability of the system to follow rapid variations in the desired response such as in the sample of Fig. 6-8 appears good. The actual response lags the desired response by approximately one control interval. Larger values of T which were tried yielded poorer response giving more smoothing and missing the sharper peaks in the statistical signal.

6.3 Second-order Dynamic Process

The second system investigated deals with the control of a second-order dynamic process whose differential equation is

$$\frac{d^2c}{dt^2} + a(t)\frac{dc}{dt} + 4c(t) = 4m(t) \qquad (6\text{-}14)$$

The parameter a(t) is assumed to vary between 0 and 8 in a sawtooth manner at a frequency of 0.08 cps as shown in Fig. 6-9.

The predictor is designed under the same assumptions as for the first-order dynamic process and is given by Eq. 6-5.

The results of Section 5.3 indicate a four-term approximation of m(t),

$$m(t) = \sum_{k=0}^{3} m_k p_k(t) \qquad (6\text{-}15)$$

is needed to give adequate control. Both a four-term and a three-term approximation are used to obtain a comparison between their abilities to give adequate control. Since the analysis of Section 5.3 is valid for the shorter control intervals, it is to be expected that neither approximation will be adequate for control interval lengths greater than $\frac{1}{4}$ sec.

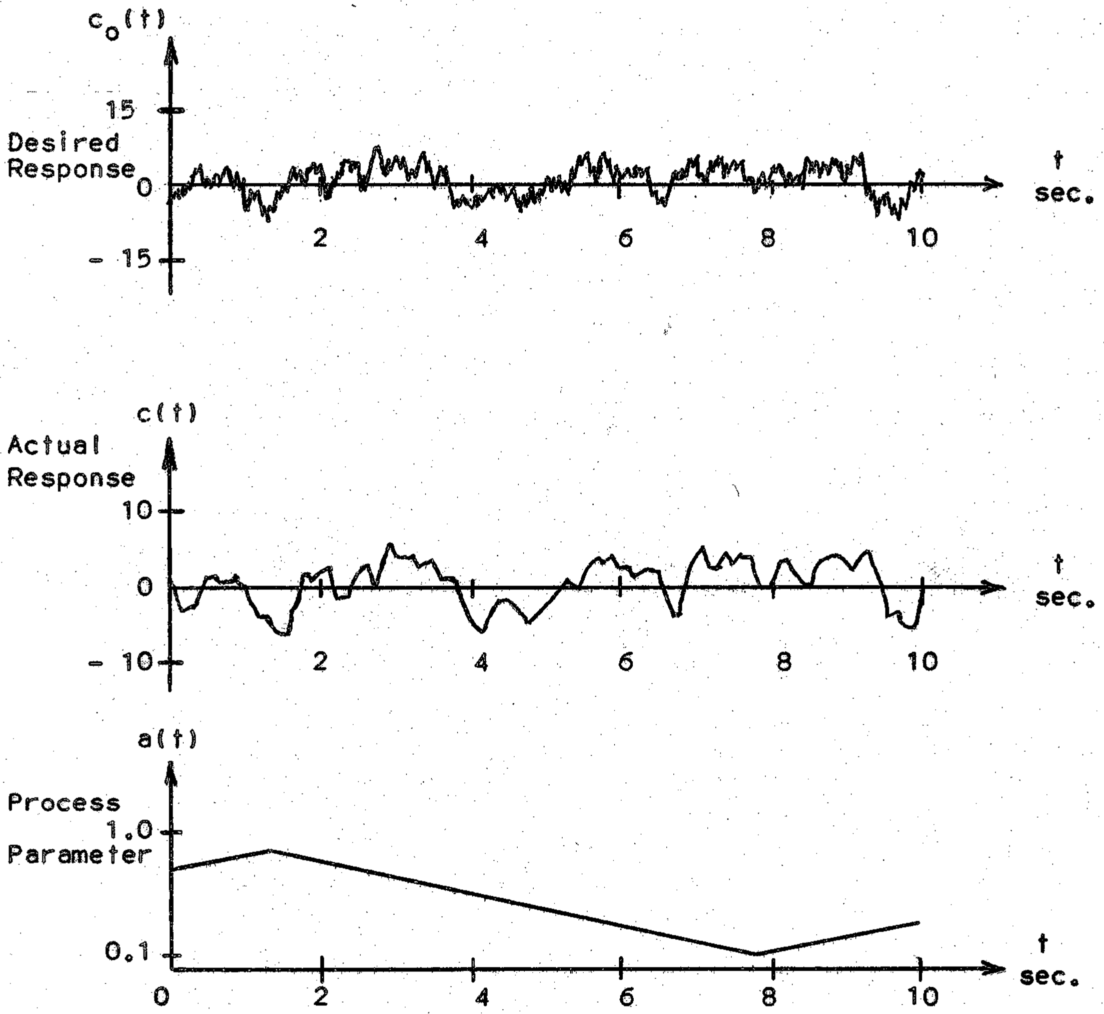


Fig. 6-8

Statistical Signal Response of First-order Dynamic Process for $\lambda_o = 6$ and $T = \frac{1}{16}$ sec.

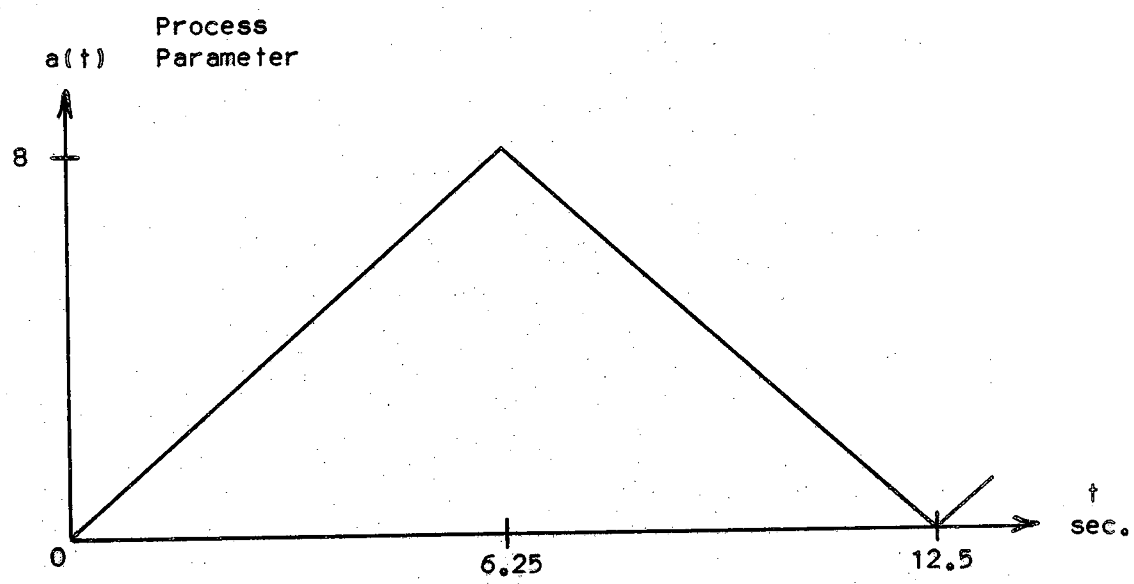


Fig. 6-9

Parameter Variation for Second-order Dynamic Process.

For the four-term approximation of m(t) data were obtained for $T = \frac{1}{8}, \frac{1}{4}, \frac{3}{8}, \frac{1}{2}$ sec. and $\lambda_0 = 4, 6, 8, 10$. The results for these values for the first two experiments are presented in Figs. 6-10, 6-11, 6-12, and 6-13. Fig. 6-11 is presented to illustrate more clearly steady-state adaptability for $T = \frac{1}{8}, \frac{1}{4}$, and $\frac{3}{8}$ sec.

Figs. 6-10 and 6-11 clearly indicate the improved steady-state adaptability for decreasing T and increased λ_0. The adaptive capability is, however, completely lost for $T > \frac{3}{8}$ sec. This is attributed to two factors. First, as mentioned above, the four-term approximation of m(t) is valid only for the shorter control intervals. Second, the dynamic process is known to become unstable during the course of its parameter variation. Thus, in the vicinity of this unstable state, the frequency of adaptation must be fast so that the process being controlled has less time to manifest its instability before correction occurs. Since the parameter drifts between its extrema in 6.25 sec. (Fig. 6-9) and the data of Fig. 6-10 indicate the control interval must be less than $\frac{3}{8}$ sec. to keep the output from drifting more than 20%, this means the control interval length must be chosen so that parameter drift is less than or equal to 6% per control interval.

The transient data also indicates a trend toward incontrollability for $T \geq \frac{3}{8}$ sec. with the four-term approximation. Examination of the rise time data of Fig. 6-12 without reference to the per cent overshoot data of Fig. 6-13 would lead one to believe that transient response improves with increasing T. However, for $T \geq \frac{3}{8}$ sec. only the system employing $\lambda_0 = 4$ can be considered to give satisfactory step response if it is desired that per cent overshoot be kept below 20%. In fact, Fig. 6-13 clearly denotes a rather sharp degradation of control for $T > \frac{1}{4}$ sec. A comparison of Fig. 6-5 for the first-order process with Fig. 6-13 for

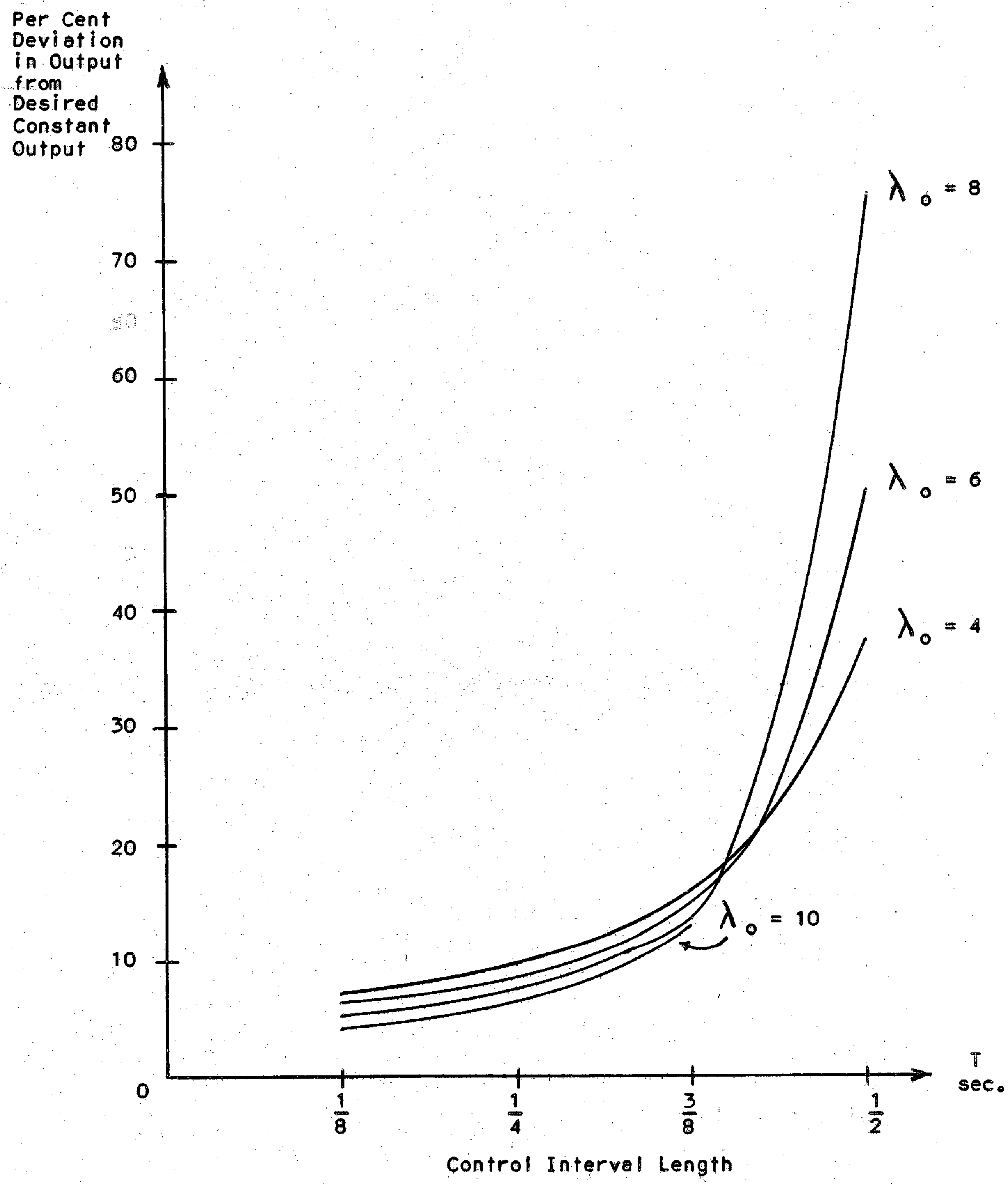


Fig. 6-10

Steady-state Adaptability of Second-order Dynamic Process for Four-term Approximation of Control Variable.

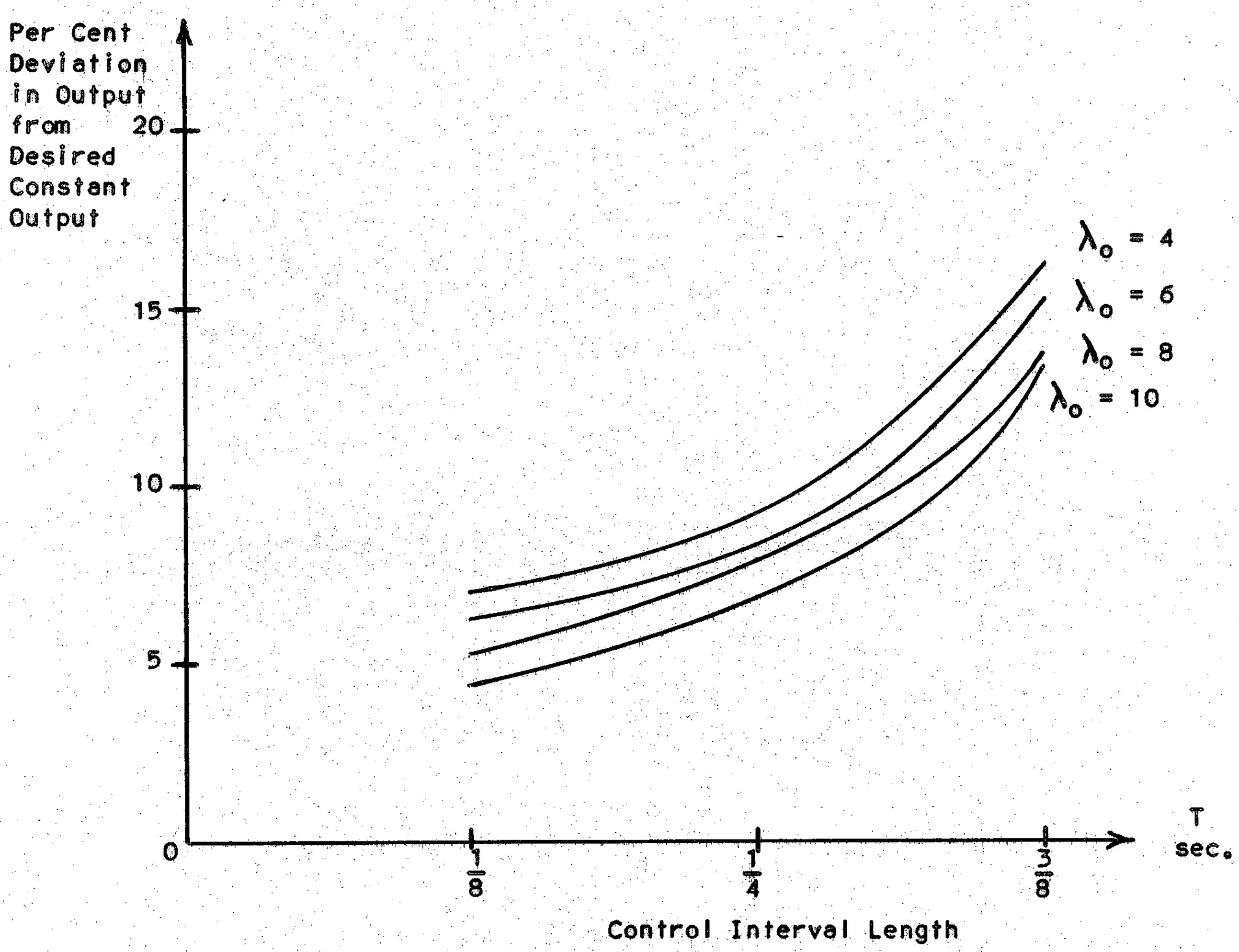


Fig. 6-11

Steady-state Adaptability of Second-order Dynamic Process for Four-term Approximation of Control Variable.

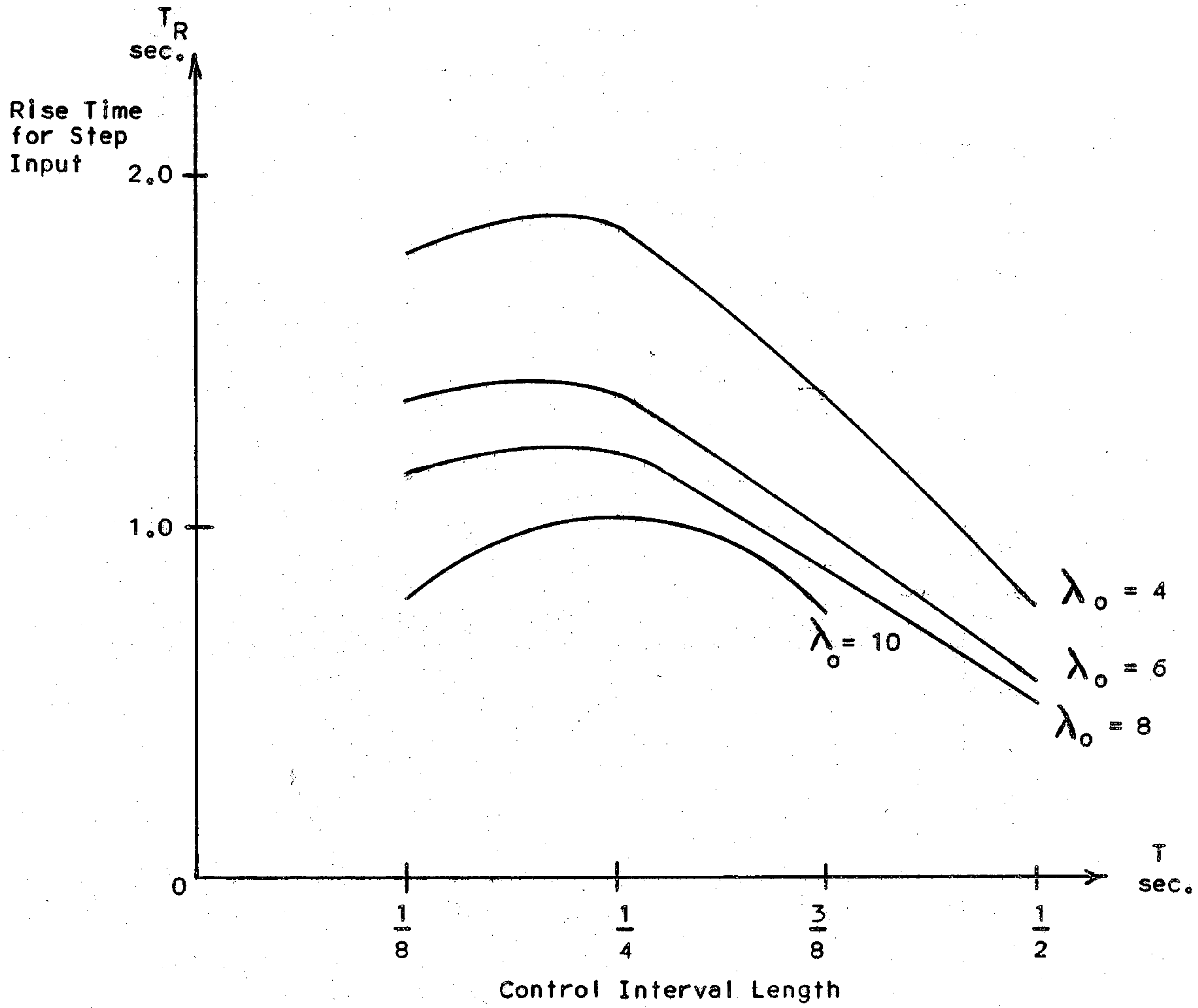


Fig. 6-12

Rise Time for Second-order Dynamic Process for Four-term Approximation of Control Variable.

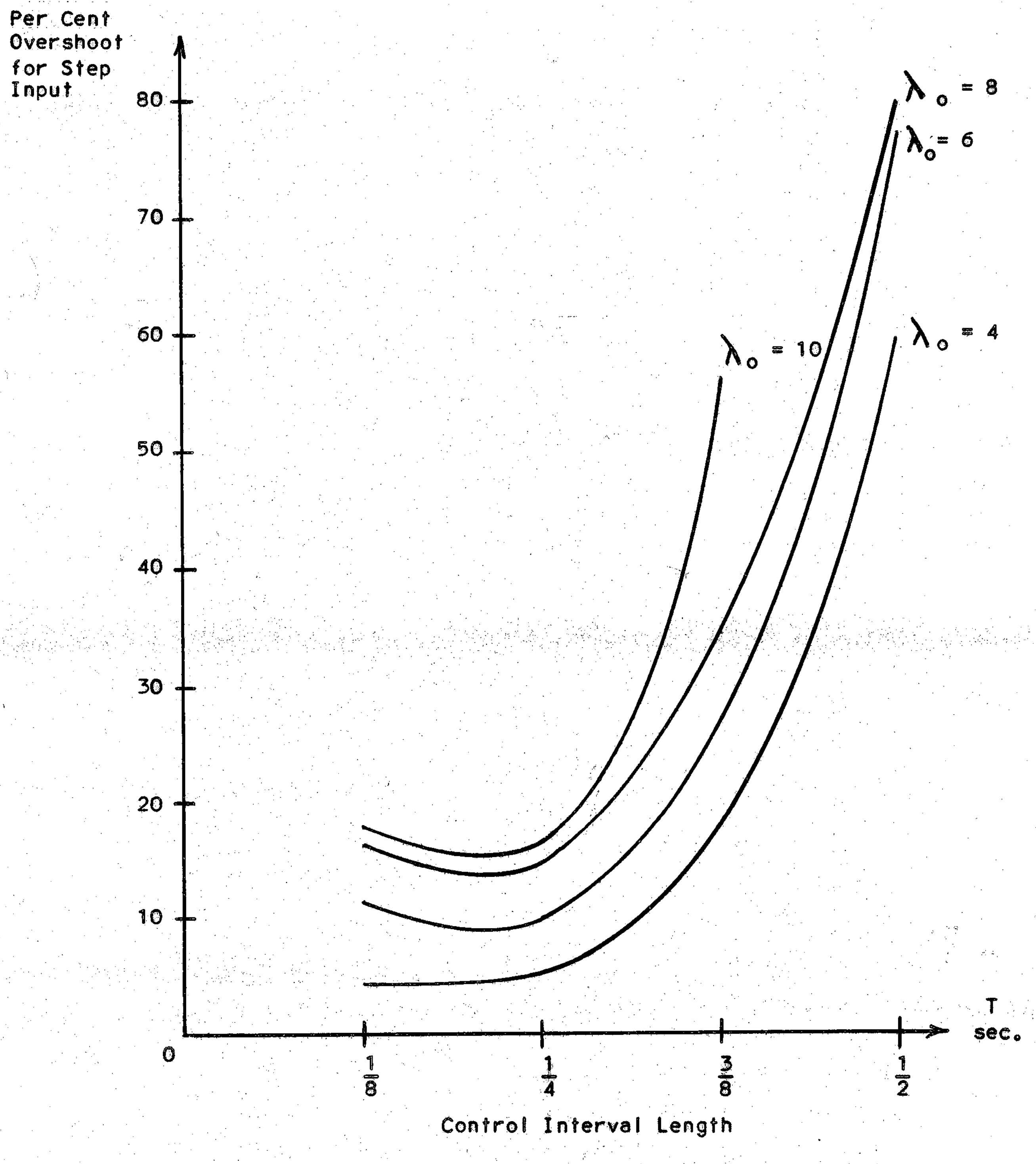


Fig. 6-13

Per Cent Overshoot for Second-order Dynamic Process
for Four-term Approximation of Control Variable.

the second-order process points out the need for a control interval length $T \leqslant \frac{1}{4}$ sec. for satisfactory transient response for both systems. This figure amounts to letting the parameter drift 4% between adaptation points for the control of these dynamic processes.

A typical step response for one of the better behaved systems using a four-term approximation of m(t) with $\lambda_o = 8$ and $T = \frac{1}{8}$ sec. to control the second-order process is shown in Fig. 6-14. The parameter variation a(t) is shown to indicate the behavior of the process parameter during the transient.

The step response for the same system using the same value of $\lambda_o = 8$, but a control interval three times as long, $T = \frac{3}{8}$ sec., is given in Fig. 6-15. The quality of the response has clearly degenerated as a result of tripling the interval length. Not only is the per cent overshoot large, but the ripple in the output after the transient has subsided is of the order of 5%.

A typical response for a statistical signal having a spectrum of the form of Eq. 6-4 is shown in Fig. 6-16 for $\lambda_o = 10$ and $T = \frac{1}{8}$ sec. The smoothing introduced by the second-order process is much greater than that for the first-order process and the former system is only able to follow the slower, well-defined variations in $c_o(t)$. Even for the slower variations in $c_o(t)$, the response c(t) lags by approximately four control intervals.

In order to show the inadequacy of the three-term approximation of m(t) and to obtain a comparison with the four-term approximation, experiments 1 and 2 were repeated for the second-order case using the three-term approximation

$$m(t) = \sum_{k=0}^{2} m_k p_k(t). \tag{6-16}$$

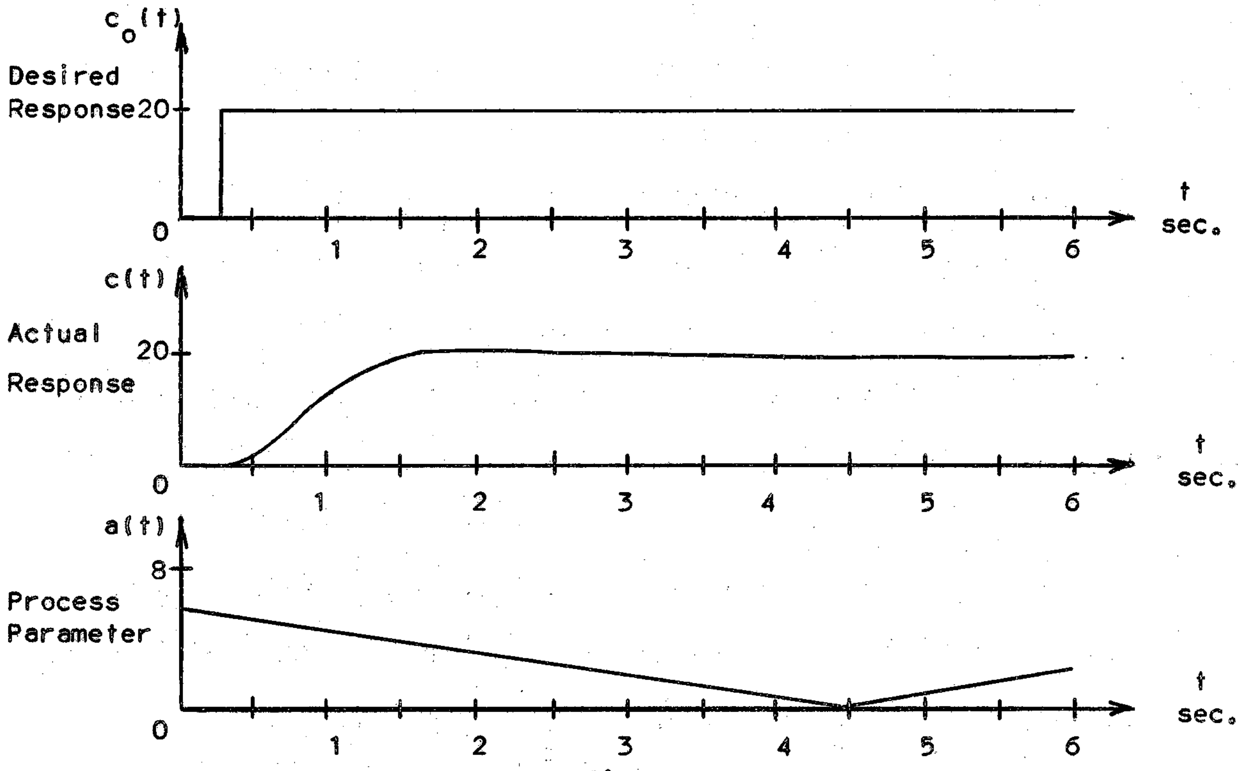


Fig. 6-14

Typical Step Response of Second-order Dynamic Process for $\lambda_o = 8$ and $T = \frac{1}{8}$ sec.

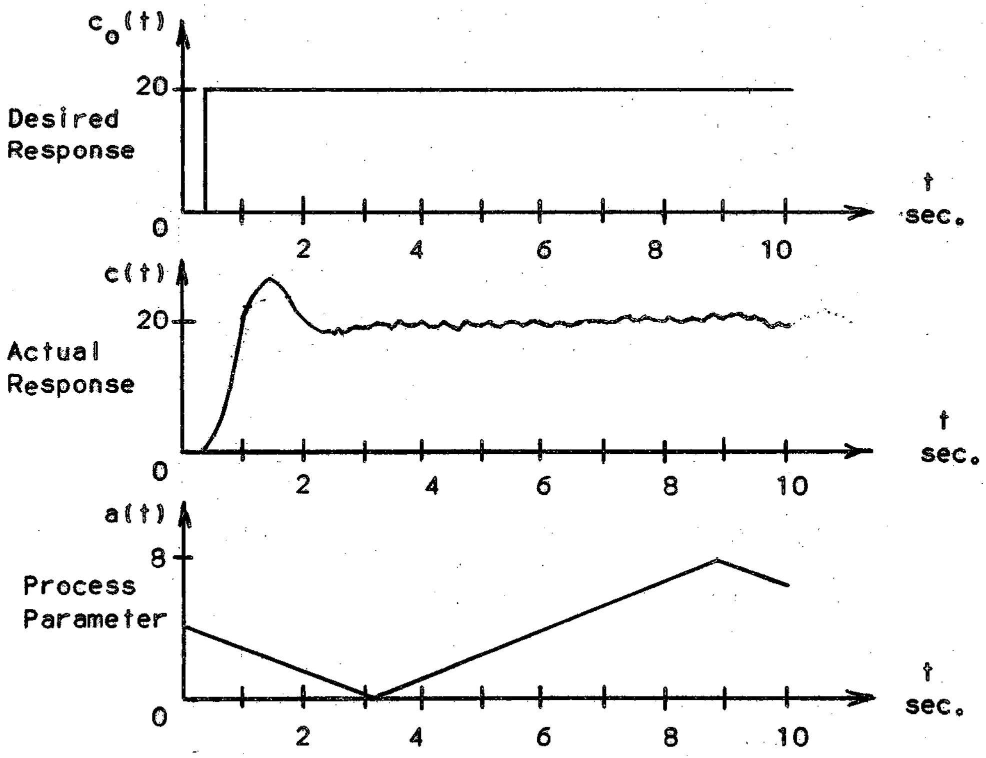


Fig. 6-15

Typical Step Response of Second-order Dynamic Process for $\lambda_o = 8$ and $T = \frac{3}{8}$ sec.

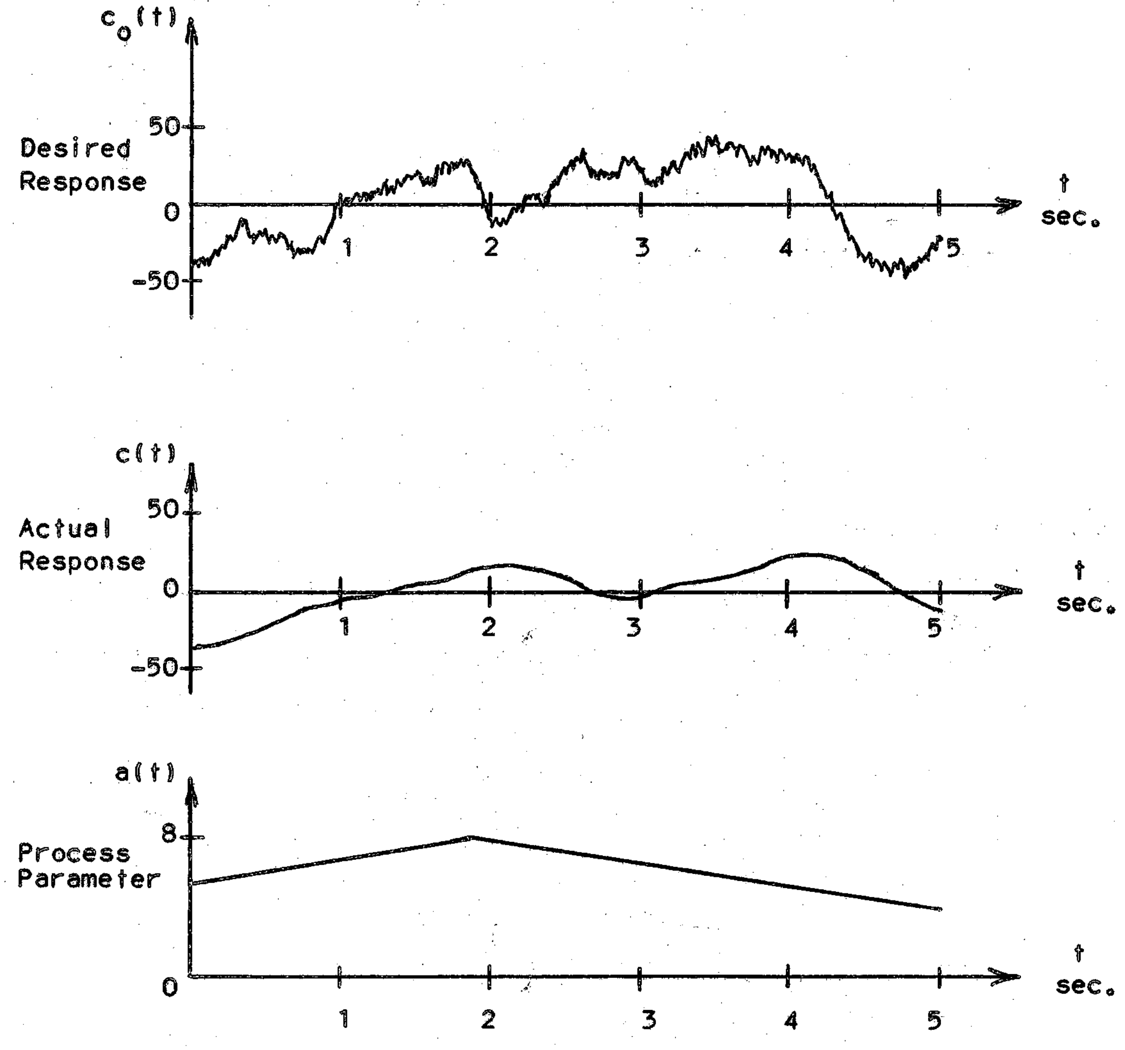


Fig. 6-16

Statistical Signal Response of Second-order Dynamic Process for $\lambda_o = 10$, $T = \frac{1}{8}$ sec.

The data are shown in Figs. 6-17, 6-18, and 6-19. Data were not obtained for $T = \frac{1}{2}$ sec. because the resulting systems were unstable. Despite the reasonable behavior of the steady-state adaptability and the rise time characteristics, the lack of control for the three-term approximation is brought out clearly by the per cent overshoot characteristics. None of the systems investigated exhibits a per cent overshoot for a step input less than 35%. In most cases there was a tendency of the system to become unstable for transient inputs. It is clear that on the basis of transient response, the three-term approximation is completely inadequate even for the shortest control interval length, $T = \frac{1}{8}$ sec. A comparison of the per cent overshoot for a step input for the three-term and the four-term approximations is given in Table 6-1 for $T = \frac{1}{8}$ sec.

6.4 Summary and Conclusions

A number of response characteristics were found to be common to the two systems investigated. The most important of these is the continued improvement of performance with decreasing control interval length. This feature was anticipated theoretically and found to be limited in practice by the information handling capabilities of the components used in the time-varying gain generator and the controller.

For the sawtooth type of parameter variations used it is found that adequate control of both processes is realized by choosing the control interval length T such that the process parameter drifts by no more than 4% per control interval.

The factor λ_o governs the relative weighting of system error and control effort. It was found that steady-state adaptability improves with increasing λ_o. Hence, as λ_o is increased, the controller places more emphasis on system error than on control effort, and therefore, has less regard for the problem of saturation.

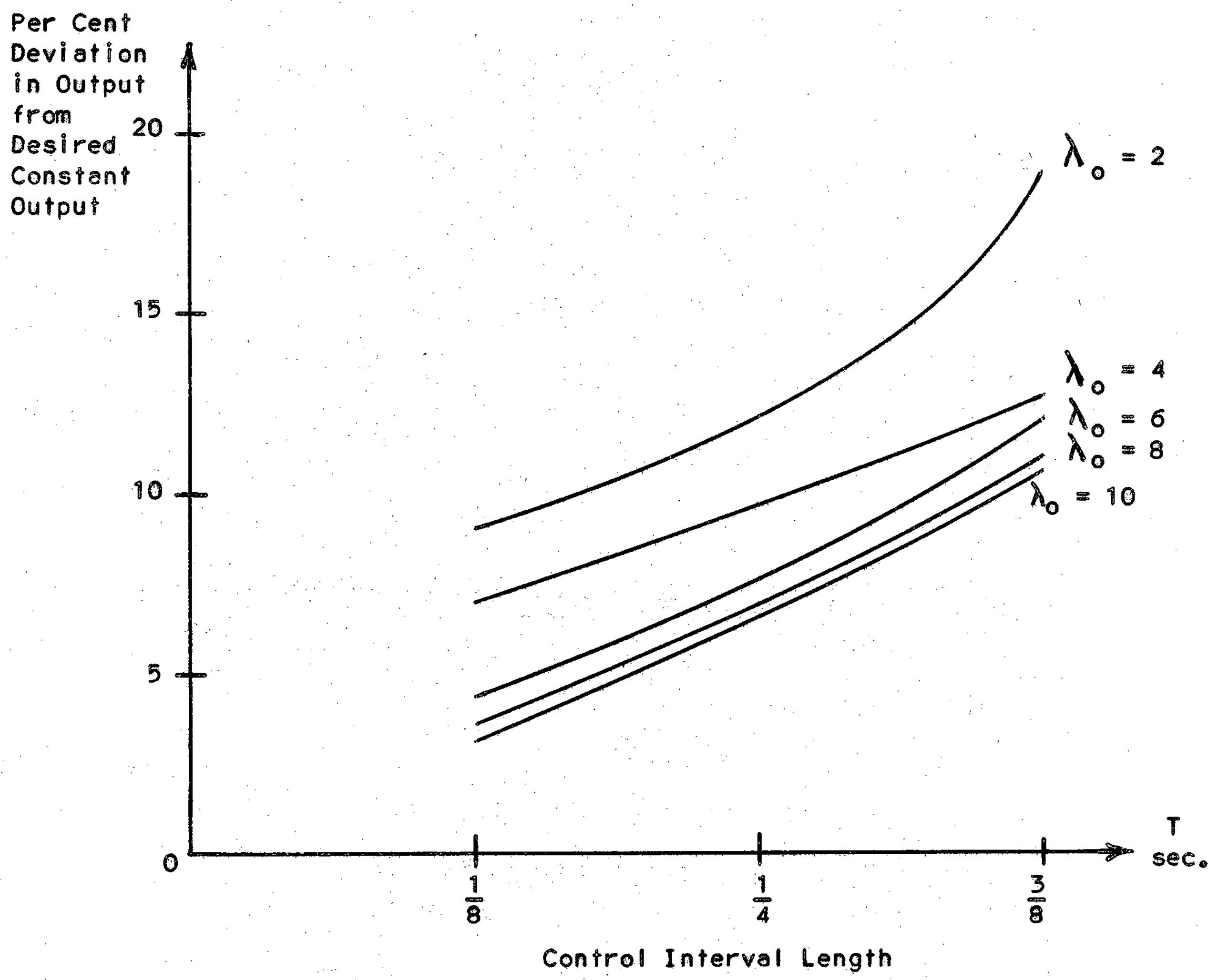


Fig. 6-17

Steady-state Adaptability for Second-order Dynamic Process for Three-term Approximation of Control Variable.

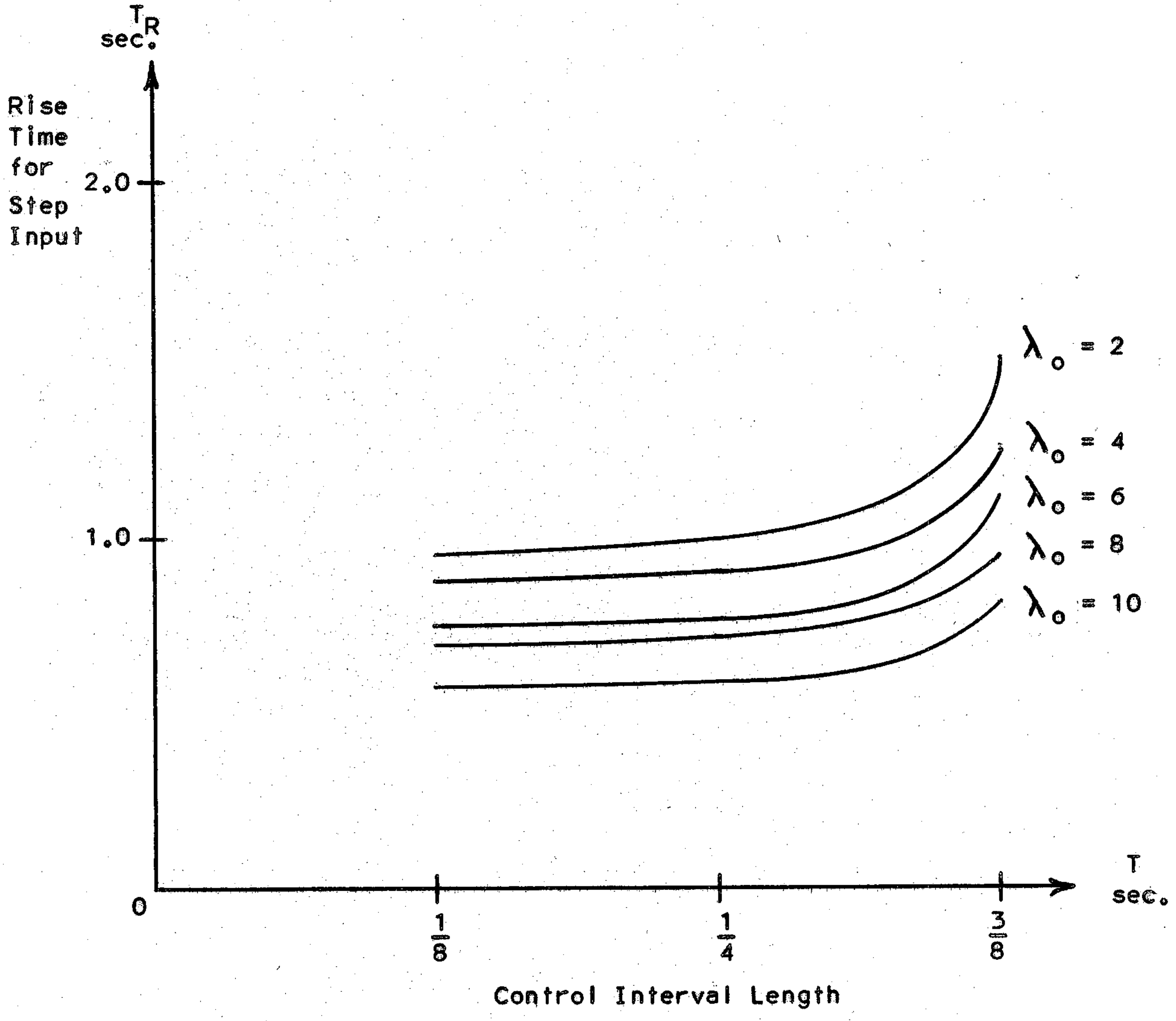


Fig. 6-18

Rise Time for Second-order Dynamic Process for Three-term Approximation of Control Variable.

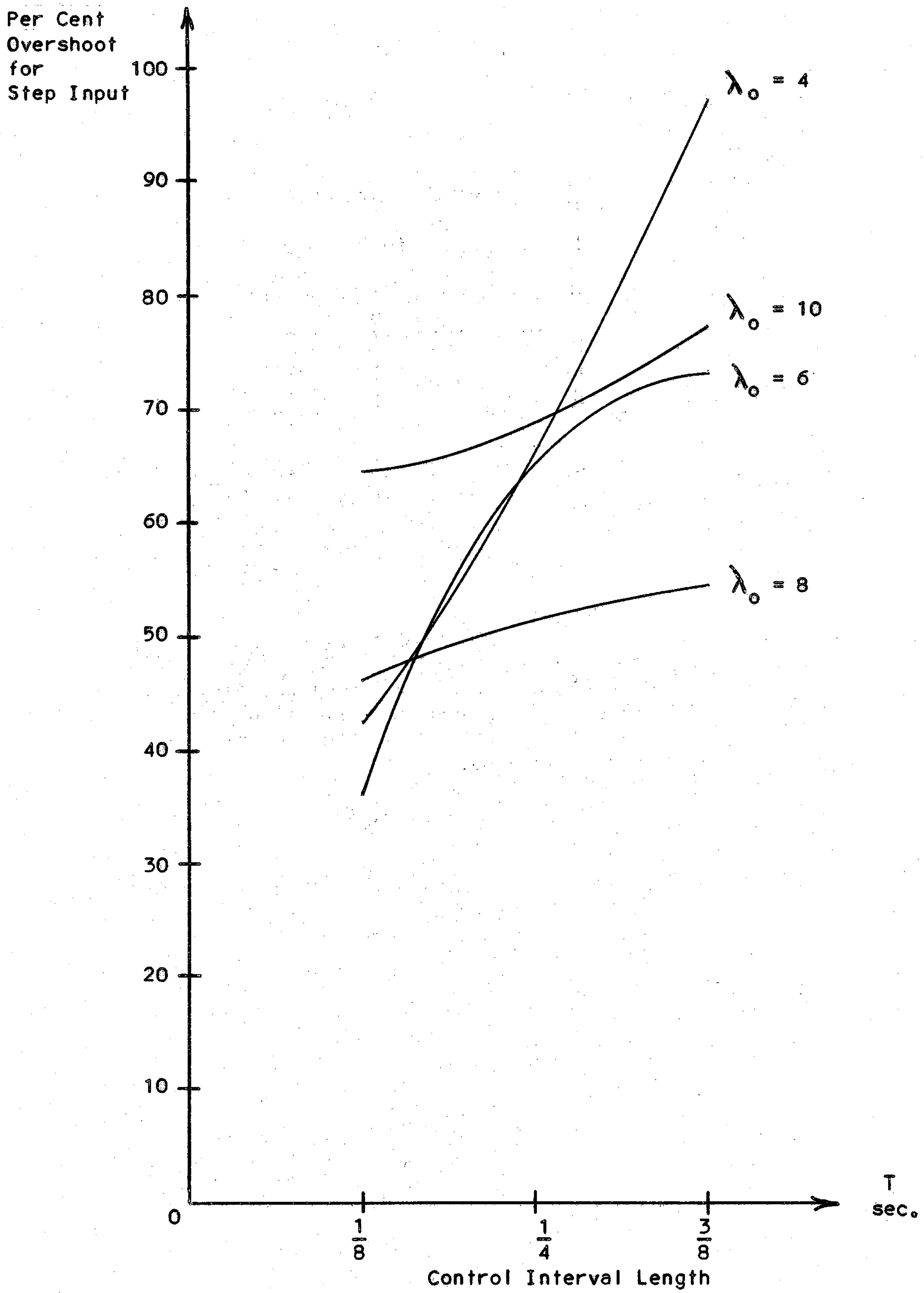


Fig. 6-19

Per Cent Overshoot for Second-order Dynamic Process for Three-term Approximation of Control Variable.

λ_o	Per Cent Overshoot for Three Terms	Four Terms
4	41.5	4.5
6	35	11.25
8	45	15.7

TABLE 6-1

Comparison of Per Cent Overshoot for Three and Four-term Approximations of Control Variable for Second-order Dynamic Process.

The results obtained here indicate the method presented in Section 5.3 for estimating the number of terms needed in $m(t)$ is sound. The statement that the method is valid only for the faster adaptation frequencies has also been substantiated.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 Summary of Results

This work has presented the development and investigation of a new class of control systems termed predictive adaptive controls. The development was based on the assumptions of prediction, interval control, and synthesis of the control variable by a sum of orthonormal polynomials in t. The optimization procedure led to the formation of a family of control laws from which the synthesis of the optimum controller was specified. It was shown that while the transfer function of the controller could not be derived in practice, a quasi-linear model of the controller could be used to obtain a semi-quantitative stability analysis.

Predictor design was presented in terms of the classical Wiener-Lee theory, and relationships for control interval length in terms of prediction accuracy were developed. Preliminary controller design was considered from the viewpoints of system error weighting factor, control interval length, and the number of terms needed in the orthonormal polynomial sum approximation of the control variable. A method for obtaining an engineering estimate of the latter quantity was developed and illustrated by three examples, two of which were investigated experimentally.

Control of first-order and second-order dynamic processes was investigated on an analog computer. Three basic experiments which evaluated the steady-state adaptability, transient response, and the statistical signal response of the two systems in the presence of extreme parameter variations were performed. In general, it was found that all three aspects of performance improved with decreasing control interval length, but that the minimum value of interval length which could be used was limited by the accuracy of the time-varying gain and controller circuitry. Improved performance which could be obtained by increasing the system error weighting factor was limited by the useful linear range of the dynamic process input. For the two systems investigated it was pointed out that the control interval length should be chosen so that the process parameters do not drift by more than 4% per interval in order to assure adequate control. Theoretical results pointing to the need for keeping the control interval length short to preserve stability, prediction accuracy, and loss of control due to process parameter drift were substantiated by the simulation results.

Investigation of the number of terms needed in the control variable revealed that the four-term approximation was adequate for control of a second-order process whereas the three-term approximation was not. This result was anticipated by the preliminary design of the controller.

One of the unique features of the class of adaptive controls presented here was that explicit evaluation of the index of performance in order for the controller to effect a control policy was not necessary. The system did not execute a hunting procedure to perform system optimization. Instead, the optimization was performed directly by generating time-varying gains. The time-varying gain circuitry required as its input the unit impulse response of the dynamic process being controlled. This information must be supplied by a suitable identification procedure. Hence, it is clear that the decision step was built into the controller from the optimization of the index of performance from which the control laws were specified.

7.2 Recommendations

A number of interesting problems which merit further research have arisen as a result of this work.

The index of performance used to develop the class of controls investigated in this research dealt with process optimization over the immediate future, i.e., the interval $[0,T]$. This approach may possibly be extended to include an optimization over the entire future by a slight alteration of the index of performance. The new index of performance would assume the form

$$I = \int_{\tau=nT}^{\infty} \left\{ \lambda(t) \left[c_o(t) - c(t)\right]^2 + m^2(t) \right\} dt \qquad (7\text{-}1)$$

for $n = 0, 1, \ldots,$ where T is the control interval length. The optimization would then deal with specifying the controller to generate the control coefficients m_k of

$$m_k(t) = \sum_{k=0}^{N} m_k p_k(t) \qquad (3\text{-}16)$$

at each sampling instant, $\tau = nT$. Control is still executed on a per interval basis but the coefficients m_k specified at each sampling or adaptation point would be optimum for all time in the future instead of only for the immediate future $[0,T]$. One of the problems associated with this new formulation is prediction of system error, $\left[c_o(t) - c(t)\right]$. Since prediction accuracy usually becomes poorer as the prediction interval increases, it may be advisable to place an arbitrary weighting on the prediction operation in which the distant future is weighted less heavily than the immediate future.

Another problem worthy of consideration is the choice of the class of polynomials used in the control variable m(t) given by Eq. 3-16. The Legendre polynomials were chosen for this research primarily because they are polynomials in t. Therefore, the resulting control signal was itself a polynomial in t which is a common type of driving signal for dynamic processes. For statistical signals, of course, the m_k were random variables. In general, there exists no method for choosing the particular class of polynomials which would be optimum, in some prescribed sense, for a given application. Some work has been done by Lee [33] on the synthesis of networks in terms of orthonormal polynomials, and more recently, some new results in the representation of signals have been presented by Lerner [34]. However, very little effort has been devoted to the development of criteria for optimum synthesis of signals or classes of signals especially as applied in control systems.

The stability analysis given in Section 4.3 is restricted to systems employing a one-term approximation of the control variable. Hence, further work is needed to determine more precisely stability requirements for the more general class of systems which use an N-term ($N \gg 1$) approximation of the control variable.

Finally, a comparison of the class of adaptive controls developed here with equivalent non-adaptive systems would be desirable. The non-adaptive system could be designed by classical methods [5] for the nominal values of the dynamic process parameters, and the response characteristics of the resulting system investigated for the extrema of the process parameters.

BIBLIOGRAPHY

1. Hazen, H. L., "Theory of Servomechanisms," J. Franklin Institute, Vol. 218, No. 3, pp. 279-330, Sept., 1934.

2. Hall, A. C., Analysis and Synthesis of Linear Servomechanisms, Technology Press, Cambridge, Mass., 1943.

3. Wiener, N., The Extrapolation, Interpolation, and Smoothing of Stationary Time Series, Technology Press, Cambridge, Mass., 1948.

4. Booton, R. C., Jr., "An Optimization Theory for Time-Varying Linear Systems with Nonstationary Statistical Inputs," Proc. IRE, Vol. 40, No. 8, pp. 977-981, August, 1952.

5. 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.

6. Mathews, M. V., and Steeg, C. W., "Terminal Controller Synthesis," ASME Princeton Symposium on Nonlinear Control Systems, Princeton, N. J., March, 1956.

7. Booton, R. C., Jr., "Optimum Design of Final-Value Control Systems," Proceedings of the Polytechnic Institute of Brooklyn Symposium on Nonlinear Control Systems, Brooklyn, N. Y., April 25-27, 1956.

8. Miskin, E., and Braun, L., Jr., Adaptive Control Systems, McGraw-Hill Book Co., Inc., New York, N. Y., 1961.

9. Gregory, P. C., Editor, Proceedings of the Self-Adaptive Flight Control Systems Symposium, WADC Report 59-49, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, March, 1959.

10. Stromer, P. R., "Adaptive or Self-Optimizing Control Systems - A Bibliography," IRE Transactions - PGAC, Vol. AC-4, No. 1, May, 1959.

11. Cooper, G. R., 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, School of Electrical Engineering, Purdue University, Lafayette, Indiana, July 1, 1960.

12. Braun, L., Jr., "On Adaptive Control Systems," IRE Transactions - PGAC, Vol. AC-4, No. 2, Nov., 1959.

13. Kalman, R. E., "Design of Self-Optimizing Control Systems," ASME Transactions, Vol. 80, pp. 468-478, Feb., 1958.

14. 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.

15. Joseph, P., Lewis, J., and Tou, J., "Plant Identification in the Presence of Disturbances and Application to Digital Adaptive Systems, AIEE Transactions Paper No. 61-73, January 29, 1961.

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. Anderson, G. W., Buland, R. N., and Cooper, G. R., "The Aeronutronic Self-Optimizing Automatic Control System," Proceedings of the Self-Adaptive Flight Control Systems Symposium, WADC Report 59-49, Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, March, 1959.

18. Graham, D., and Lathrop, R. C., "The Synthesis of Optimum Transient Response Criteria and Standard Forms," Transactions AIEE, Vol. 72, pt. 2, 1954.

19. Gibson, J. E., Leedham, C. D., McVey, E. S., and Rekasius, Z. V., "Specifications and Data Presentation in Linear Control Systems," AF Technical Report No. 1, Contract AF 29(600)-1933 PRF 2030, Holloman Air Force Base, New Mexico, July, 1959.

20. Murphy, G. J., and Bold, N. T., "Optimization Based on a Square-Error Criterion with an Arbitrary Weighting Function," IRE Transactions - PGAC, Vol. AC-5, No. 1, January, 1960.

21. Bellman, R., "On the Application of the Theory of Dynamic Programming to the Study of Control Processes," The Rand Corporation, Santa Monica, Calif., 1956.

22. Anderson, G. W., Aseltine, J. A., Mancini, A. R., and Sarture, C. W., "A Self-Adjusting System for Optimum Dynamic Performance," 1958 IRE National Convention Record, pt. 4, pp. 182-190, March, 1958.

23. Steeg, C. W., "The Design of Optimum Filters and Predictors," 1957 IRE Convention Record, Part IV, March, 1957.

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

25. Lee, Y. W., Statistical Theory of Communication, John Wiley and Sons, Inc., New York, N. Y., 1960.

26. Tou, J. T., Digital and Sampled-Data Control Systems, McGraw-Hill Book Company, Inc., New York, N. Y., 1959.

27. Hildebrand, F. B., Methods of Applied Mathematics, Prentice-Hall, Inc., New York, N. Y., 1952.

28. Bellman, R., Dynamic Programming, Princeton University Press, Princeton, N. J., 1957.

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

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

31. Churchill, R. V., Introduction to Complex Variables and Applications, McGraw-Hill Book Company, Inc., New York, N. Y., 1948.

32. Taylor, A. E., Advanced Calculus, Ginn and Company, Boston, Mass., 1955.

33. Lee, Y. W., "Synthesis of Electric Networks by Means of the Fourier Transforms of Laguerre's Functions," Journal of Mathematics and Physics, M.I.T., Vol. 11, 1932.

34. Lerner, R. M., "The Representation of Signals," IRE Transactions - PGCT, Vol. CT-6, May, 1959.

35. Sansone, G., Orthogonal Functions, Vol. IX, Interscience Series on Pure and Applied Mathematics, Interscience Publishers, Inc., New York, N. Y., 1959.

APPENDIX A

A CLASS OF ORTHONORMAL POLYNOMIALS

In this research the optimum control variable was approximated by a sum of polynomials in t

$$m(t) = \sum_{k=0}^{N} m_k p_k(t) \qquad \text{(A-1)}$$

for $0 \leqslant t \leqslant T$ where the $p_k(t)$ are the polynomials. By making the polynomials orthonormal over the interval $[0,T]$, considerable simplification resulted in the final control equations. The orthonormal property of the polynomials $p_k(t)$, $k = 0, 1, \ldots,$ is given by

$$\int_0^T p_k(t)\, p_n(t)\, dt = \begin{cases} 0 & k \neq n \\ 1 & k = n \end{cases} \qquad \text{(A-2)}$$

where $p_k(t)$ is a polynomial in t of degree k.

The class of polynomials satisfying Eq. A-2 and forming a complete orthonormal system with respect to functions integrable on $[0,T]$ is the class of Legendre polynomials [35]. The Legendre polynomials are usually defined on the interval $[-1,1]$, but, by the change of variable

$$t = \frac{2}{T}\, t' - 1 \qquad \text{(A-3)}$$

where t' is the independent variable of the original polynomials, become the class of polynomials needed in this work.

Making the change of variable Eq. A-3 in the original Legendre polynomials, the $p_k(t)$ are given by the relation

$$p_k(t) = \sqrt{\frac{2K + 1}{T}}\; P_k(t) \qquad \text{(A-4)}$$

for $0 \leq t \leq T$ where the first six $P_k(t)$ are

$$
\begin{aligned}
P_0(t) &= 1 \\
P_1(t) &= \frac{2}{T}\, t - 1 \\
P_2(t) &= \frac{6}{T^2}\, t^2 - \frac{6}{T}\, t + 1 \\
P_3(t) &= \frac{20}{T^3}\, t^3 - \frac{30}{T^2}\, t^2 + \frac{12}{T}\, t - 1 \\
P_4(t) &= \frac{70}{T^4}\, t^4 - \frac{140}{T^3}\, t^3 + \frac{90}{T^2}\, t^2 - \frac{20}{T}\, t + 1 \\
P_5(t) &= \frac{252}{T^5}\, t^5 - \frac{630}{T^4}\, t^4 + \frac{560}{T^3}\, t^3 - \frac{210}{T^2}\, t^2 + \frac{30}{T}\, t - 1\ .
\end{aligned}
\tag{A-5}
$$

Eqs. A-5 are plotted in Figs. A-1, A-2, and A-3.

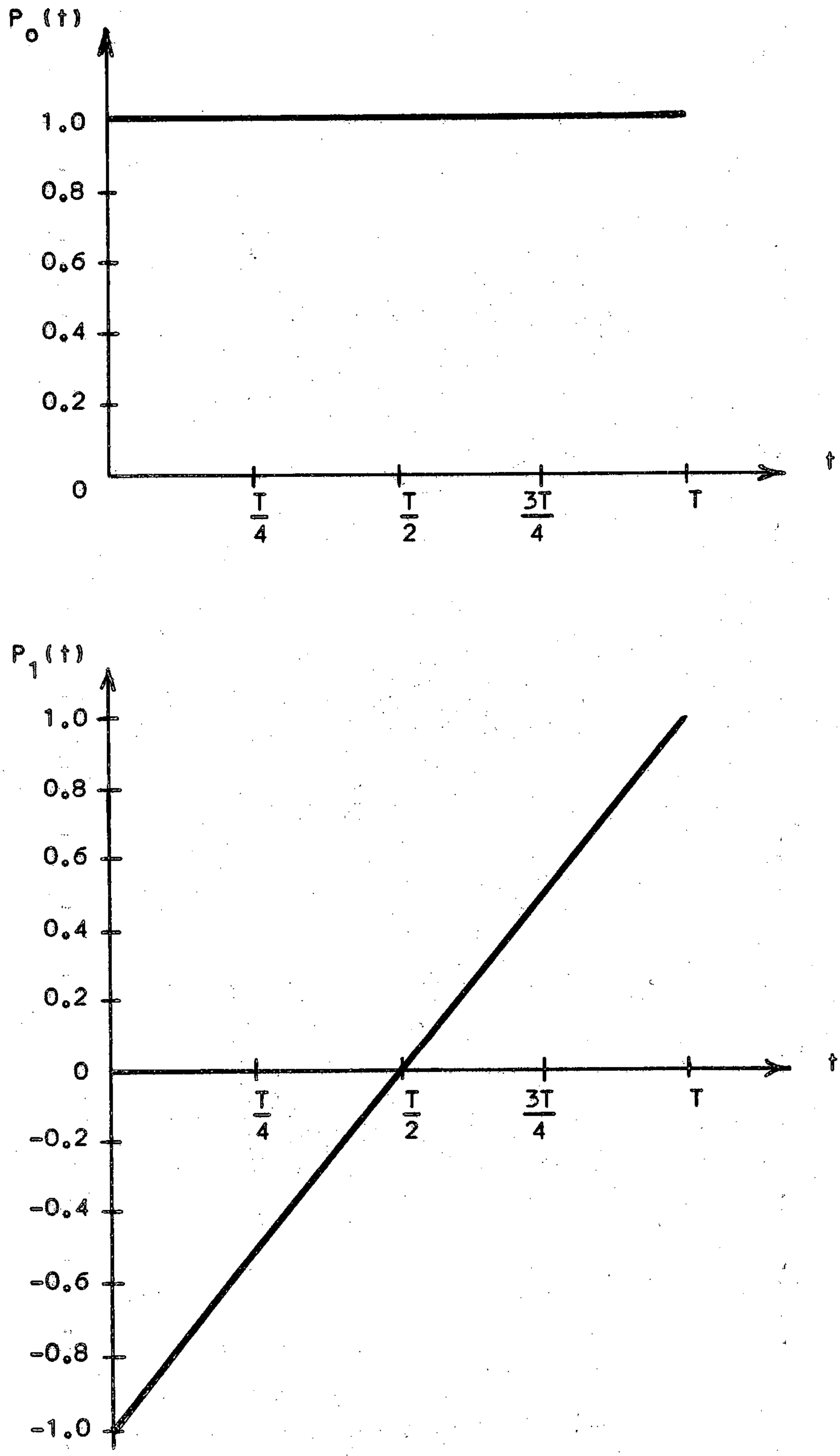


Fig. A-1

Transformed Legendre Polynomials $P_o(t)$ and $P_1(t)$.

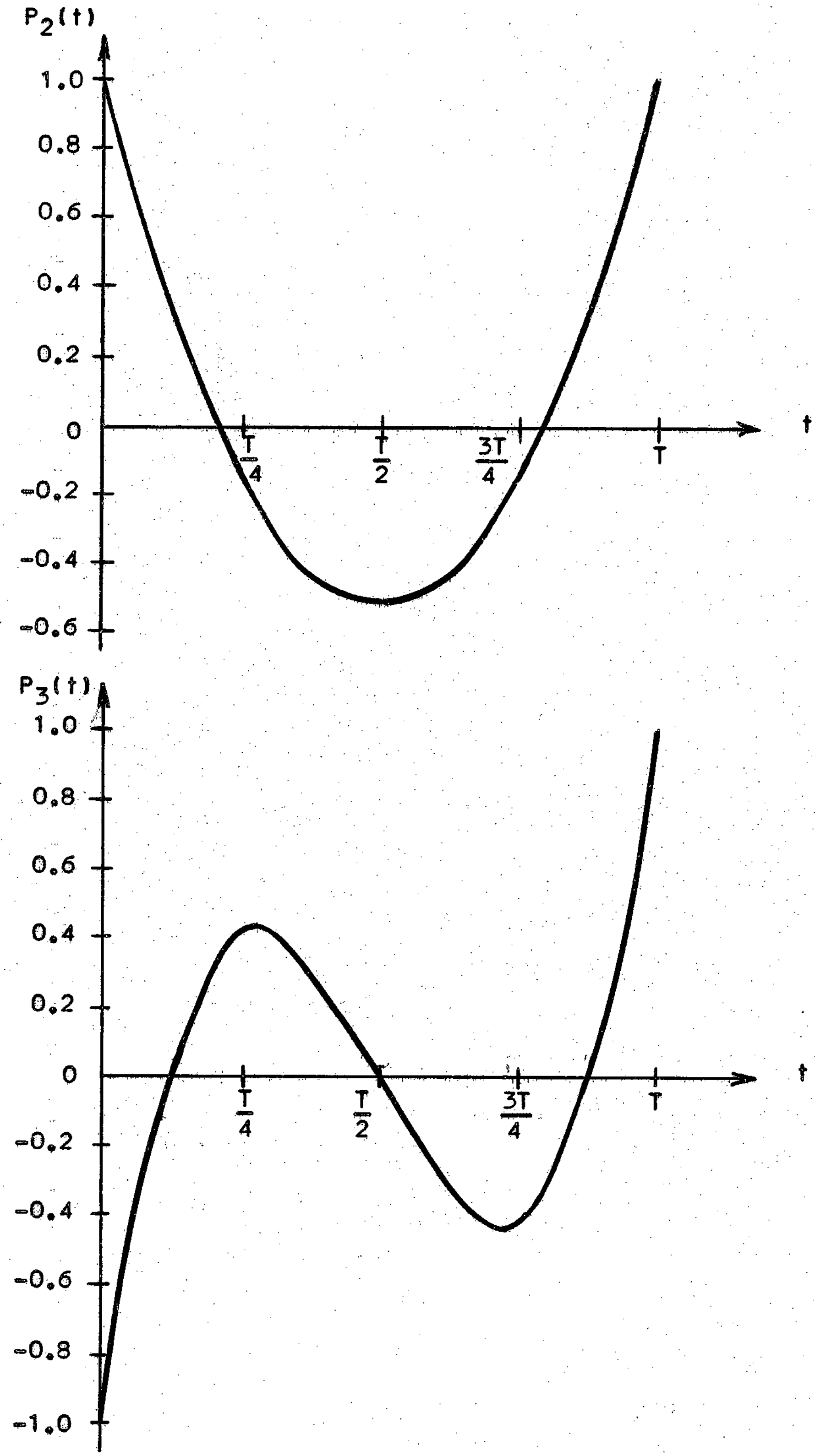


Fig. A-2

Transformed Legendre Polynomials $P_2(t)$ and $P_3(t)$.

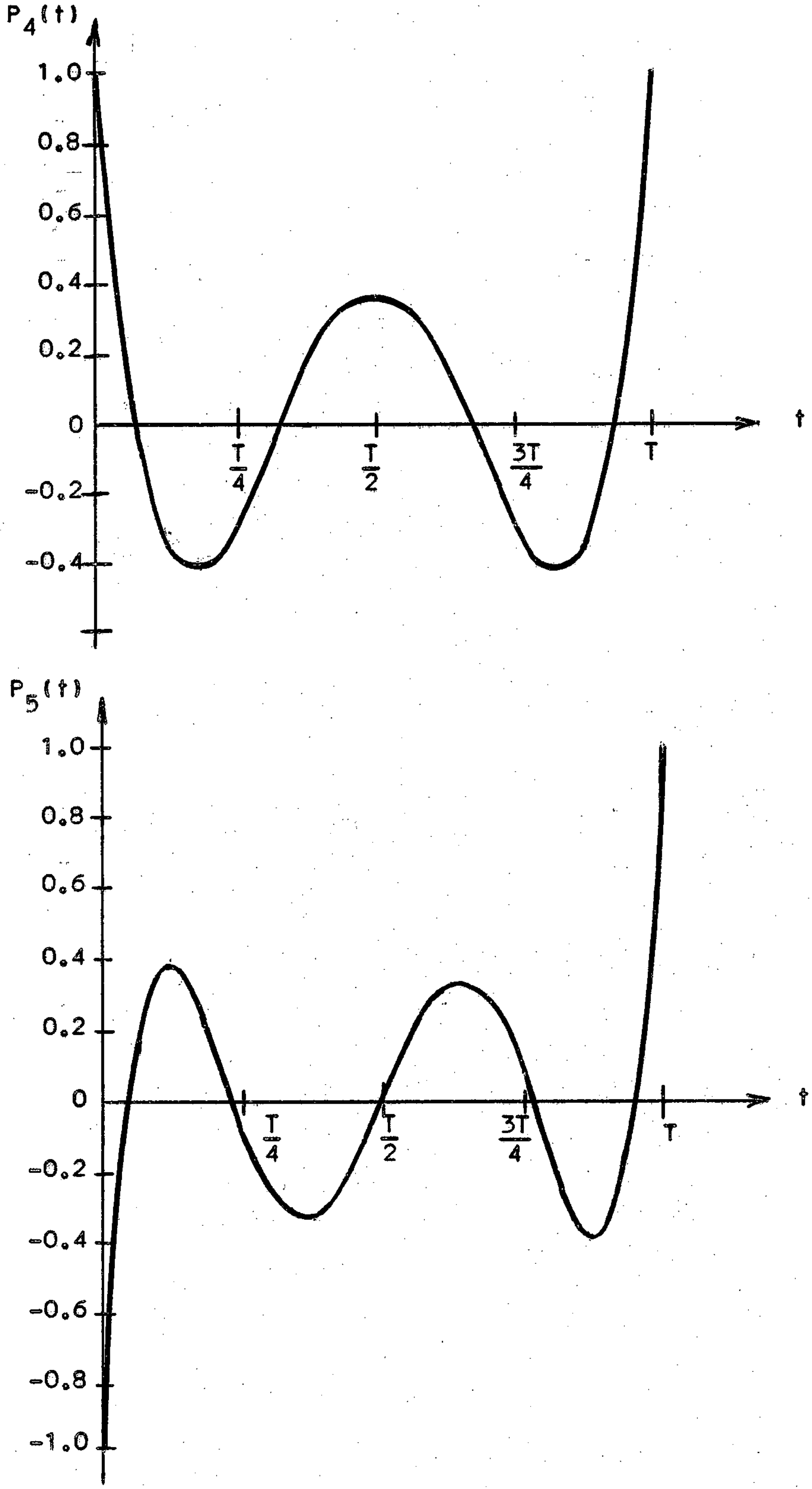


Fig. A-3

Transformed Legendre Polynomials $P_4(t)$ and $P_5(t)$.

APPENDIX B

GUARANTEE OF OPTIMUM SOLUTION

Strictly speaking, the solution of the set of equations given by Eq. 4-7 merely produces a stationary value of the index of performance, Eq. 2-4. Intuitively, if a solution exists, it would seem that it must render Eq. 2-4 a minimum since the latter equation can be made a maximum by choosing the m_k arbitrarily large. For the sake of completeness, however, an analytical argument is presented below to show the m_k of Eq. 4-7 do minimize the index of performance.

The requirement that Eq. 2-4 be a minimum is

$$\frac{\partial^2 I}{\partial m_k^2} > 0 \qquad \text{(B-1)}$$

for k = 0, 1, ..., N.

Differentiating Eq. 4-7 with respect to m_k gives

$$\frac{\partial^2 I}{\partial m_k^2} = \frac{\partial}{\partial m_k}\left\{\int_0^T \lambda(t)\,\left[c_o(t) - c(t)\right]\left[-\int_0^t p_k(\tau)\,w(t,\tau)\,d\tau\right] dt + m_k\right\}$$

$$= \int_0^T \left\{\lambda(t)\left[-\frac{\partial c(t)}{\partial m_k}\right]\left[-\int_0^t p_k(\tau)\,w(t,\tau)\,d\tau\right]\right\} dt + 1 \qquad \text{(B-2)}$$

for k = 0, 1, ..., N.

From Eq. 4-4

$$\frac{\partial c(t)}{\partial m_k} = \int_0^t p_k(\tau)\,w(t,\tau)\,d\tau \qquad \text{(4-4)}$$

for k = 0, 1, ..., N.

Hence, substituting Eq. 4-4 into Eq. B-2 yields

$$\frac{\partial^2 I}{\partial m_k^2} = \int_0^T \lambda(t) \left[\int_0^t p_k(\tau)\; w(t,\tau)\; d\tau \right]^2 dt + 1 \qquad \text{(B-3)}$$

for k = 0, 1, ..., N.

Since $\lambda(t) > 0$ was specified in Chapter 2, the first term of Eq. B-3 has the property

$$\int_0^T \lambda(t) \left[\int_0^t p_k(\tau)\; w(t,\tau)\; d\tau \right]^2 dt \geq 0 \qquad \text{(B-4)}$$

Therefore, the right-hand side of Eq. B-3 is always greater than zero giving the result

$$\frac{\partial^2 I}{\partial m_k^2} > 0 \qquad \text{(B-5)}$$

for k = 0, 1, ..., N and thereby insuring a minimum.

APPENDIX C

SUMMARY OF WIENER-LEE PREDICTION THEORY [25]

The prediction problem may be viewed as a linear filtering problem as shown in Fig. C-1. The filter is characterized by its unit impulse response h(t), and has an input x(t), an output y(t), and a desired output z(t). The filter h(t) is to be determined so that the mean-square difference between the actual output y(t) and the desired output z(t) is a minimum. In terms of Fig. C-1 the problem is one of finding a physically realizable filter h(t) such that

$$\overline{\epsilon^2(t)} = \text{minimum} \tag{C-1}$$

where the bar denotes an averaging over all time.

It has been shown that there exists a unique, physically realizable filter $h_o(t)$ which will render $\overline{\epsilon^2(t)}$ a minimum [25, Ch. 14]. This filter is determined from a solution of the Wiener-Hopf equation

$$\int_{-\infty}^{\infty} h_{opt}(\tau)\, \phi_{xx}(\tau - \tau_1)\, d\tau_1 - \phi_{xz}(\tau) = 0 \qquad \tau \geq 0 \tag{C-2}$$

where $\phi_{xx}(\tau)$ is the autocorrelation function of the input x(t), and $\phi_{xz}(\tau)$ is the crosscorrelation function of the input x(t) and the desired response z(t). These two functions are defined by the relations

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

and

$$\phi_{xz}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t-\tau)\, z(t)\, dt \tag{C-4}$$

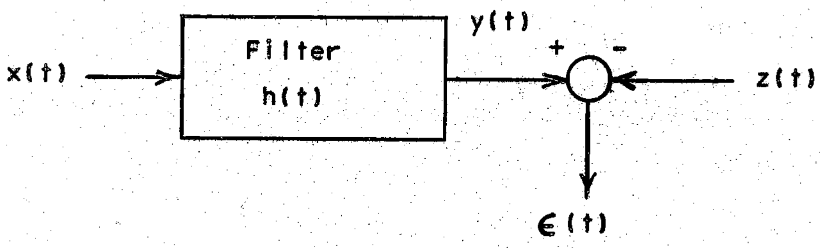


Fig. C-1

Block Diagram for Linear Filtering Problem.

Solution of Eq. C-2 can be effected readily in the frequency domain provided $\phi_{xx}(\tau)$ and $\phi_{xz}(\tau)$ are Fourier transformable. The technique was developed by Wiener and is termed spectrum factorization [25, p. 376]. The results are summarized below where appropriate definitions have been given to each of the functions used.

The complex Fourier transform pair is defined by

$$F(s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t)\, e^{-jst}\, dt \qquad \text{(C-5)}$$

and

$$f(t) = \int_{-\infty + j\sigma_1}^{+\infty + j\sigma_1} F(s) e^{jts}\, ds \qquad \text{(C-6)}$$

Now, let

$\Phi_{xx}(s)$ = complex Fourier transform of $\phi_{xx}(\tau)$, and

$\Phi_{xz}(s)$ = complex Fourier transform of $\phi_{xz}(\tau)$.

Also, let

$\Phi^{+}_{xx}(s)$ = any factor of $\Phi_{xx}(s)$ which contains all the poles and zeros of $\Phi_{xx}(s)$ which lie in the upper half of the complex plane, and

$\Phi^{-}_{xx}(s)$ = the remaining factor of $\Phi_{xx}(s)$ which contains all the poles and zeros of $\Phi_{xx}(s)$ which lie in the lower half of the complex plane.

Then, the Fourier transform $H_{opt}(s)$ of the optimum filter $h_o(t)$ which satisfies Eq. C-2 is given by

$$H_{opt}(s) = \frac{1}{2\pi \Phi^{+}_{xx}(s)} \int_{0}^{\infty} \psi(t)\, e^{-jst}\, dt \qquad \text{(C-7)}$$

where

$$\psi(t) = \int_{-\infty + jv_1}^{+\infty + jv_1} \frac{\Phi_{xz}(w)}{\Phi_{xx}^{-}(w)} e^{jwt}\, dw \qquad (C\text{-}8)$$

in which the complex variable of integration w is w = u + jv with u and v the independent real variables [25, p. 392].

This result will now be specialized for the cases of pure prediction.

Pure Prediction.

Consider the situation where c(t), which is the signal to be predicted, is relatively free from noise contamination. Then, in terms of the notation of Fig. C-1

$$x(t) = c(t) \qquad (C\text{-}9)$$

and

$$z(t) = c(t + T) = c^{*}(t) \qquad (C\text{-}10)$$

where T is the prediction interval length. For this case

$$\begin{aligned}\phi_{xx}(\tau) &= \overline{c(t)\, c(t + \tau)} \\ &= \phi_{cc}(\tau)\end{aligned} \qquad (C\text{-}11)$$

where the bar denotes the averaging operation of Eq. C-3. Hence,

$$\phi_{xx}(\tau) = \phi_{cc}(\tau) \qquad (C\text{-}12)$$

The input-desired output autocorrelation function is

$$\begin{aligned}\phi_{xz}(\tau) &= \overline{c(t)\, c(t + \tau + T)} \\ &= \phi_{cc}(\tau + T)\end{aligned} \qquad (C\text{-}13)$$

where the bar denotes the averaging operation of Eq. C-4. This result gives

$$\Phi_{xz}(s) = \int_{-\infty}^{\infty} \Phi_{cc}(\tau + T)\, e^{-js\tau}\, d\tau$$

$$= e^{jsT}\Phi_{cc}(s) \qquad \text{(C-14)}$$

Hence, the optimum Wiener predictor is given by the relations

$$H_{opt}(s) = \frac{1}{2\pi\Phi_{cc}^{+}(s)} \int_{0}^{\infty} \psi(t + T)\, e^{-jst}\, dt \qquad \text{(C-15)}$$

where

$$\psi(t + T) = \int_{-\infty + jv_1}^{\infty + jv_1} \frac{e^{jwT}\Phi_{cc}(w)}{\Phi_{cc}^{-}(w)}\, e^{jtw}\, dw$$

$$= \int_{-\infty + jv_1}^{\infty + jv_1} \Phi_{cc}^{+}(w)\, e^{j(t + T)w}\, dw \qquad \text{(C-16)}$$

Prediction Errors.

Because the entire adaptation process is based on the predicted error signal, an analysis of prediction accuracy is a paramount consideration in the design of predictive adaptive controls. Hence, the equations necessary to determine mean-square prediction accuracy are reviewed below.

Lee [25, p. 429] has shown that the minimum mean-square error for the optimum Wiener filter given in Eqs. C-7 and C-8 is

$$\left.\overline{\epsilon^2(t)}\right|_{min} = \phi_{xx}(0) - \frac{1}{2\pi} \int_{0}^{\infty} \psi^2(t)\, dt \qquad \text{(C-17)}$$

where $\psi(t)$ is given by Eq. C-8 and $\phi_{xx}(0)$ is the value of the autocorrelation function $\phi_{xx}(\tau)$ for $\tau = 0$. For pure prediction with prediction interval length T, Eq. C-17 becomes

$$\left.\overline{\epsilon^2(t)}\right|_{min} = \phi_{cc}(0) - \frac{1}{2\pi} \int_0^\infty \psi^2(t + T)\, dt \qquad \text{(C-18)}$$

where $\psi(t + T)$ is given by Eq. C-16 and $\phi_{cc}(0)$ is the value of the auto-correlation function $\phi_{cc}(\tau)$ for $\tau = 0$. After a change of variable in the second term, Eq. C-18 becomes

$$\left.\overline{\epsilon^2(t)}\right|_{min} = \phi_{cc}(0) - \frac{1}{2\pi} \int_T^\infty \psi^2(t)\, dt \qquad \text{(C-19)}$$

It has been shown [25, p. 434] that

$$\phi_{cc}(0) = \frac{1}{2\pi} \int_0^\infty \psi^2(t)\, dt \qquad \text{(C-20)}$$

Hence, substituting Eq. C-20 into Eq. C-19 gives

$$\left.\overline{\epsilon^2(t)}\right|_{min} = \frac{1}{2\pi} \int_0^T \psi^2(t)\, dt \qquad \text{(C-21)}$$

for pure prediction where $\psi(t)$ is given by Eq. C-16 for T = 0.

APPENDIX D

DESCRIPTION OF EXPERIMENTAL APPARATUS

The results presented in Chapter 6 were obtained using the Berkeley EASE Model 1032 Analog Computer and standard simulation techniques.

The operations of resetting the integrators and of sampling were performed with relays. Two 650 ohm DPST relays which were driven by the transistor circuit of Fig. D-1 were used to drive two larger relays whose contacts were used for resetting and sampling. The control interval length was changed by varying the frequency of the square wave input to the transistor drive circuit. The reset operation was then achieved by using a pair of relay contacts in series with a 1000 ohm resistor between grid and output of the integrator. The gain of the reset integrator shown is ten. This is needed to compensate partially for the attenuation in the multipliers since the output of each multiplier is 0.01 times the product of the two input signals. In order to simplify the computer diagrams given below, reset integrators will be shown as conventional integrators but will be marked "reset".

The sample and hold circuit used is shown in Fig. D-3. A second pair of contacts which are normally closed were used to provide proper sequencing so that the output of the reset integrator was sampled before the integrator was reset. The interconnection of relays and contacts is shown in Fig. D-4. Because R_1 is energized first, the sampling circuit is closed just as R_2 is energized. The pull-in time which R_2 requires to close the reset contacts is long enough so that sampling is completed before reset occurs. To simplify the complete simulation diagrams further, the sample and hold circuit will be indicated by a block where it is understood the circuit in the block is that of Fig. D-3.

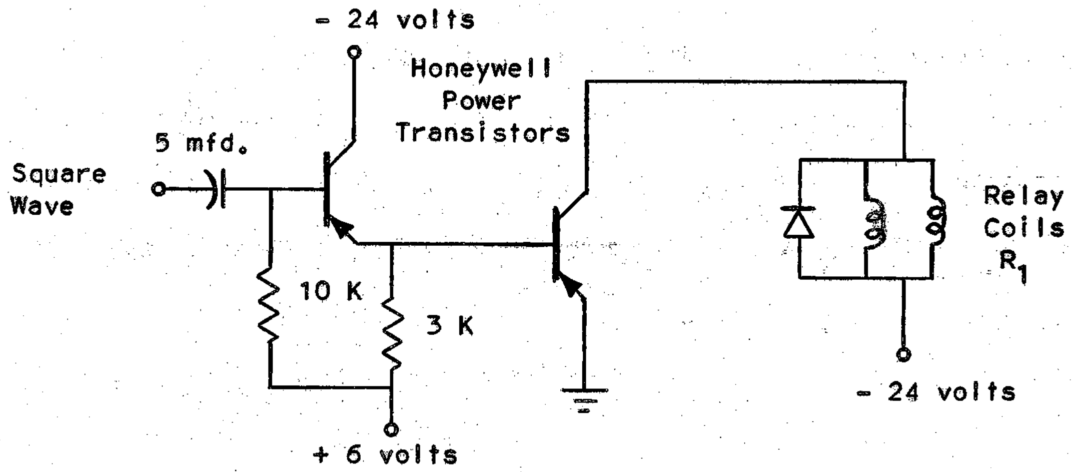


Fig. D-1

Transistor Drive Circuit.

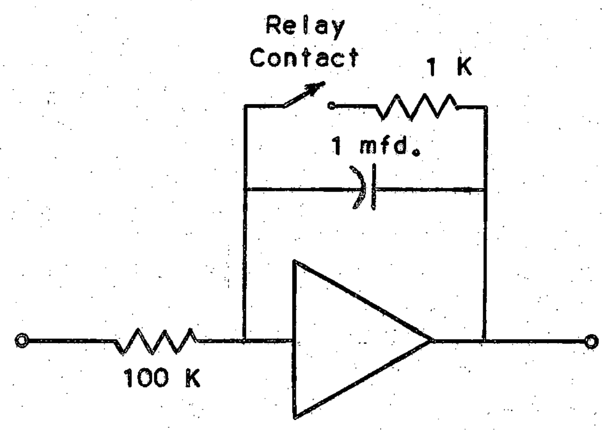


Fig. D-2

Reset Integrator.

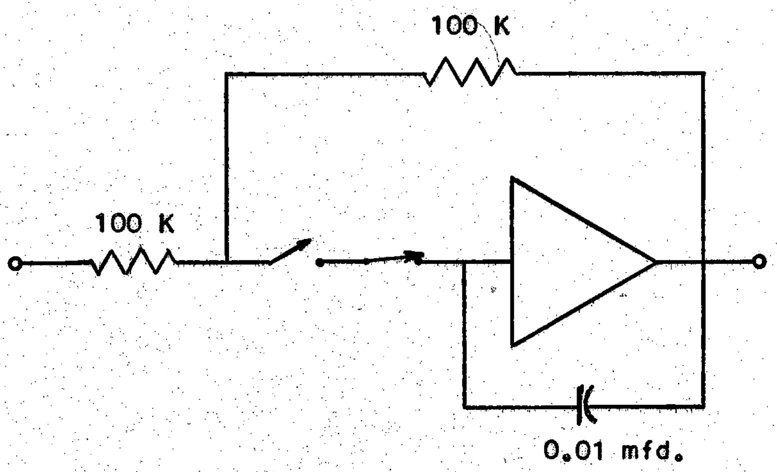


Fig. D-3

Sample and Hold Circuit.

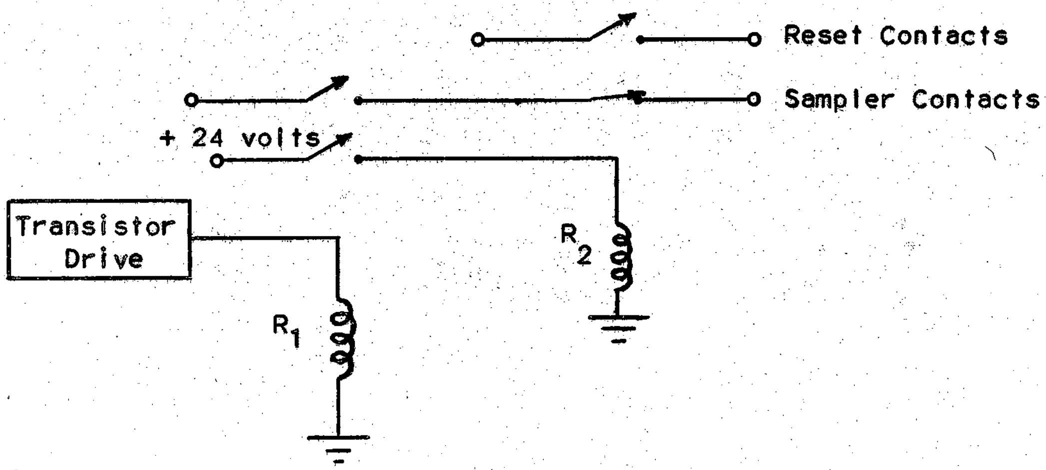


Fig. D-4

Interconnection of Relays and Relay Contacts.

In order to avoid exceeding the frequency response limits of the relays, the systems simulated were time scaled to operate at 1/8 of real time. That is, if t is real time and τ is simulation time, the relation between the two time scales is $t = \frac{\tau}{8}$.

The polynomial generator used is given in Fig. D-5. Since the first polynomial $p_0(\tau)$ is a constant, it is supplied in the first channel of the controller by a gain adjustment.

The controller for the first-order dynamic process using a two-term approximation of the control variable is shown in Fig. D-6. The one-term approximation is obtained by breaking the upper channel at the input to the summing amplifier.

The complete simulation diagram for control of the first-order process is given in Fig. D-7. The identification operation is simulated by using a model indentical to the process. The parameter of the model and the process are driven by the same source with the output of the model as the input to the time-varying gain generator. The controller, given in Fig. D-6, is indicated as a block with its external inputs $-10K_0(\tau)$, $-10K_1(\tau)$, and $-10p_1(\tau)$. As mentioned above, the extra factors of 10 are needed to compensate for multiplier attenuation.

The time-varying gain generator and the controller for the second-order dynamic process with a four-term approximation of the control variable are given in Figs. D-8 and D-9, respectively. A model of the dynamic process is again used to simulate identification.

With the polynomial generator, time-varying gain generator and controller indicated by blocks, the complete second-order system assumes the form of Fig. D-10.

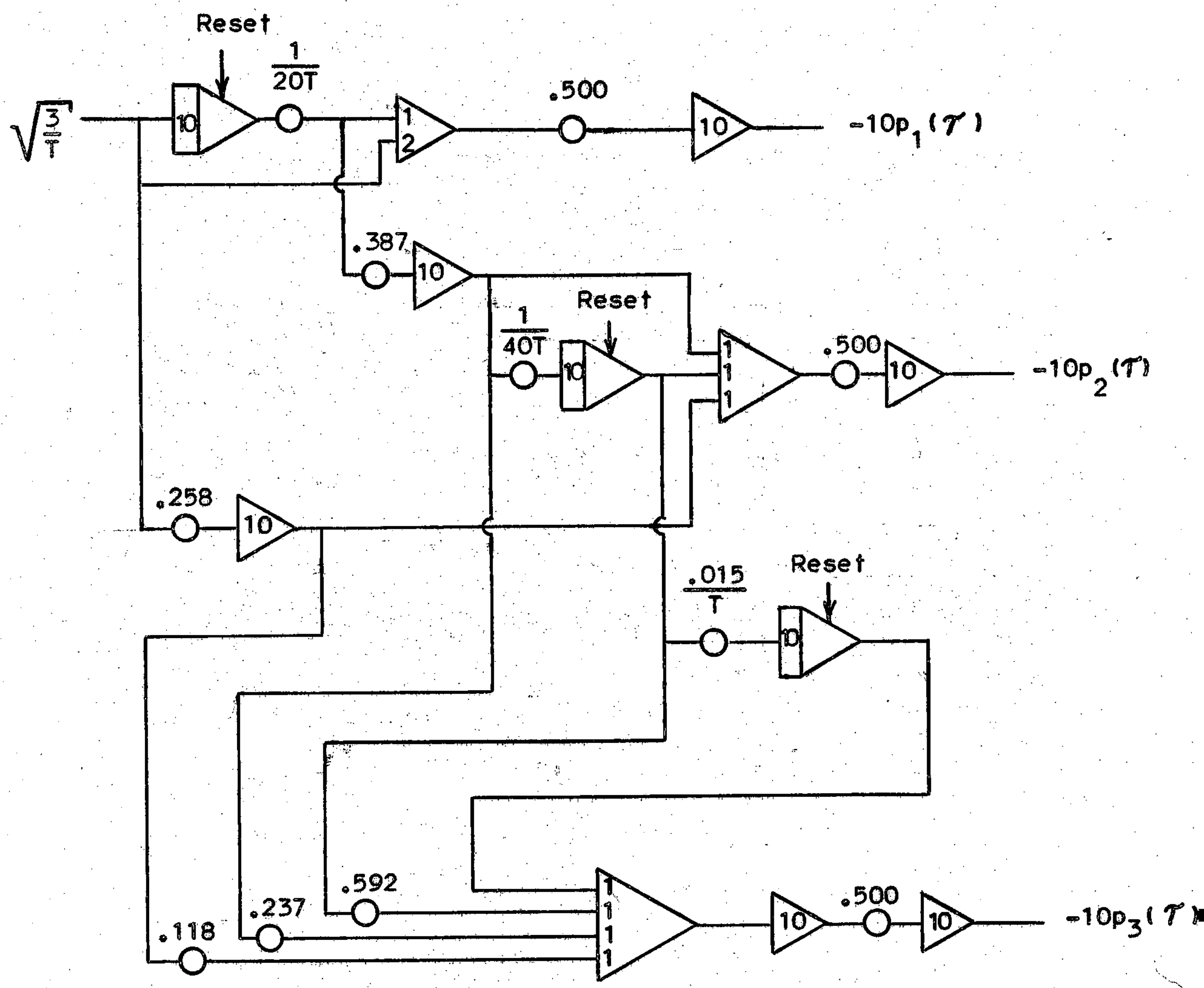


Fig. D-5

Polynomial Generator.

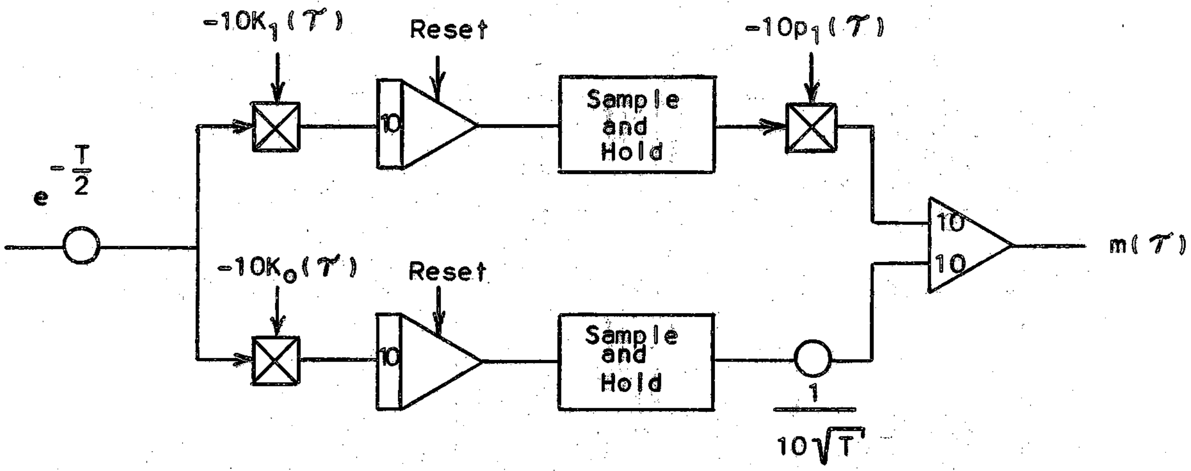


Fig. D-6

First-order Process Controller.

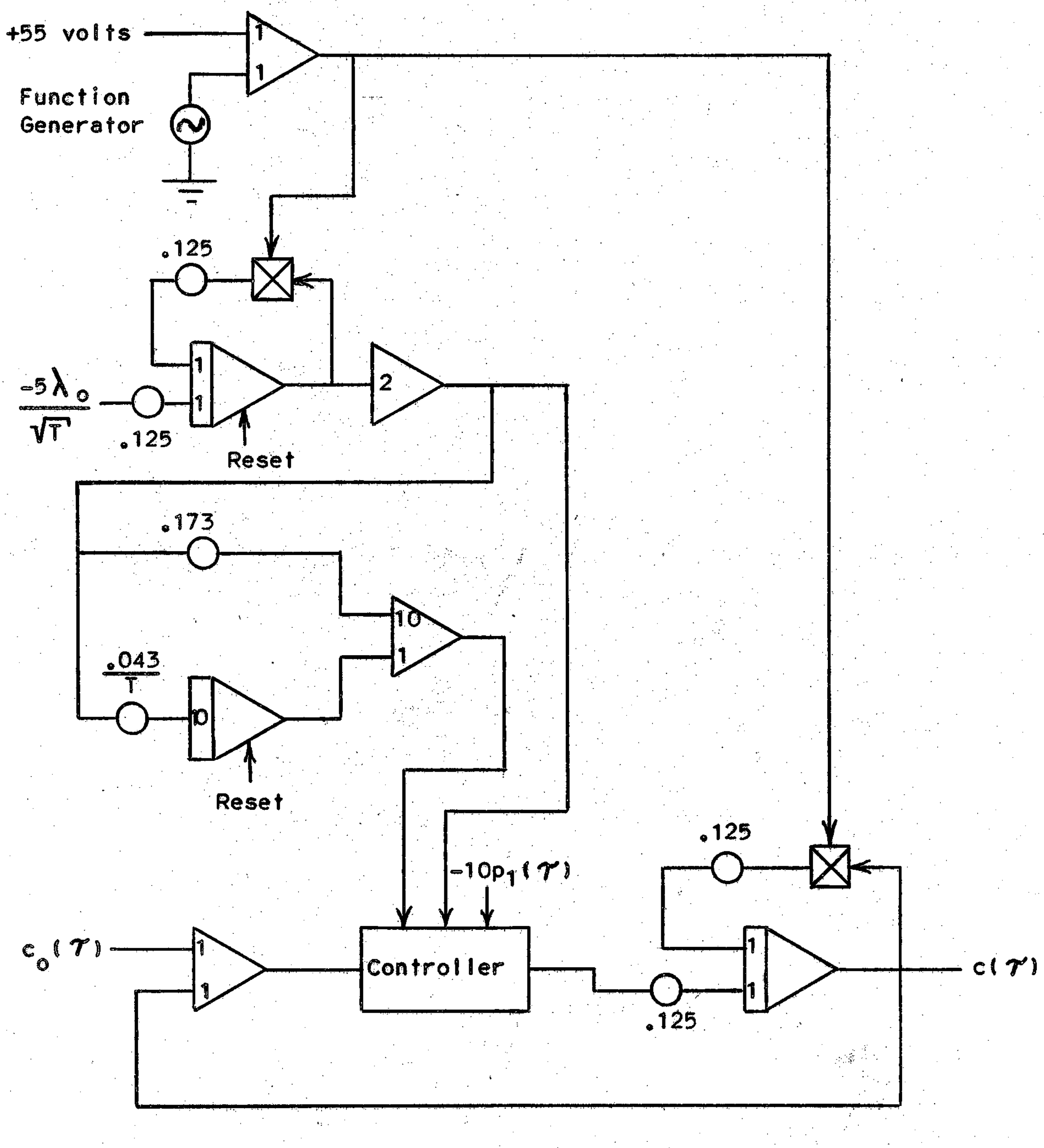


Fig. D-7

Complete First-order System.

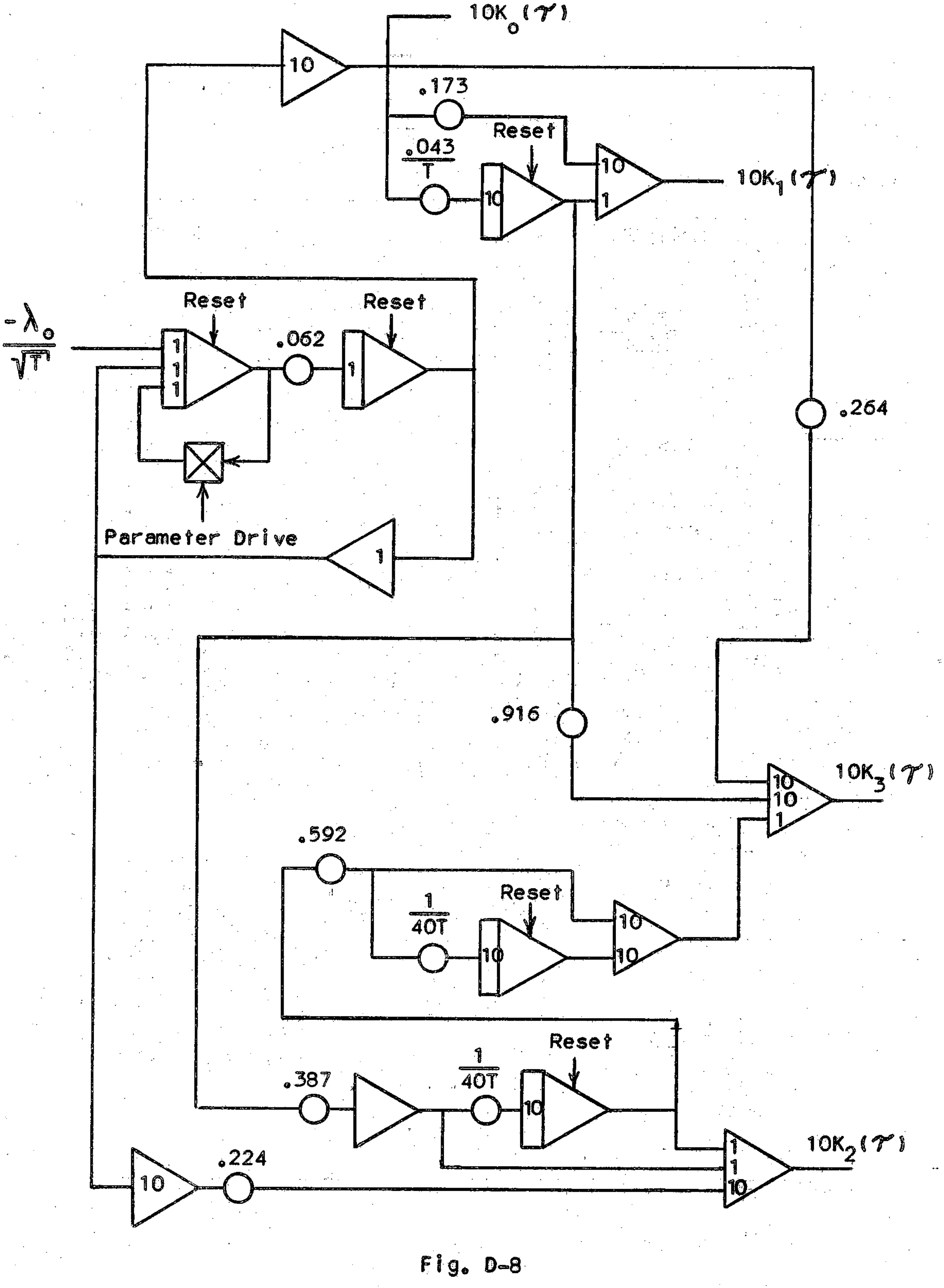


Fig. D-8

Second-order Dynamic Process Time-varying Gain Generator.

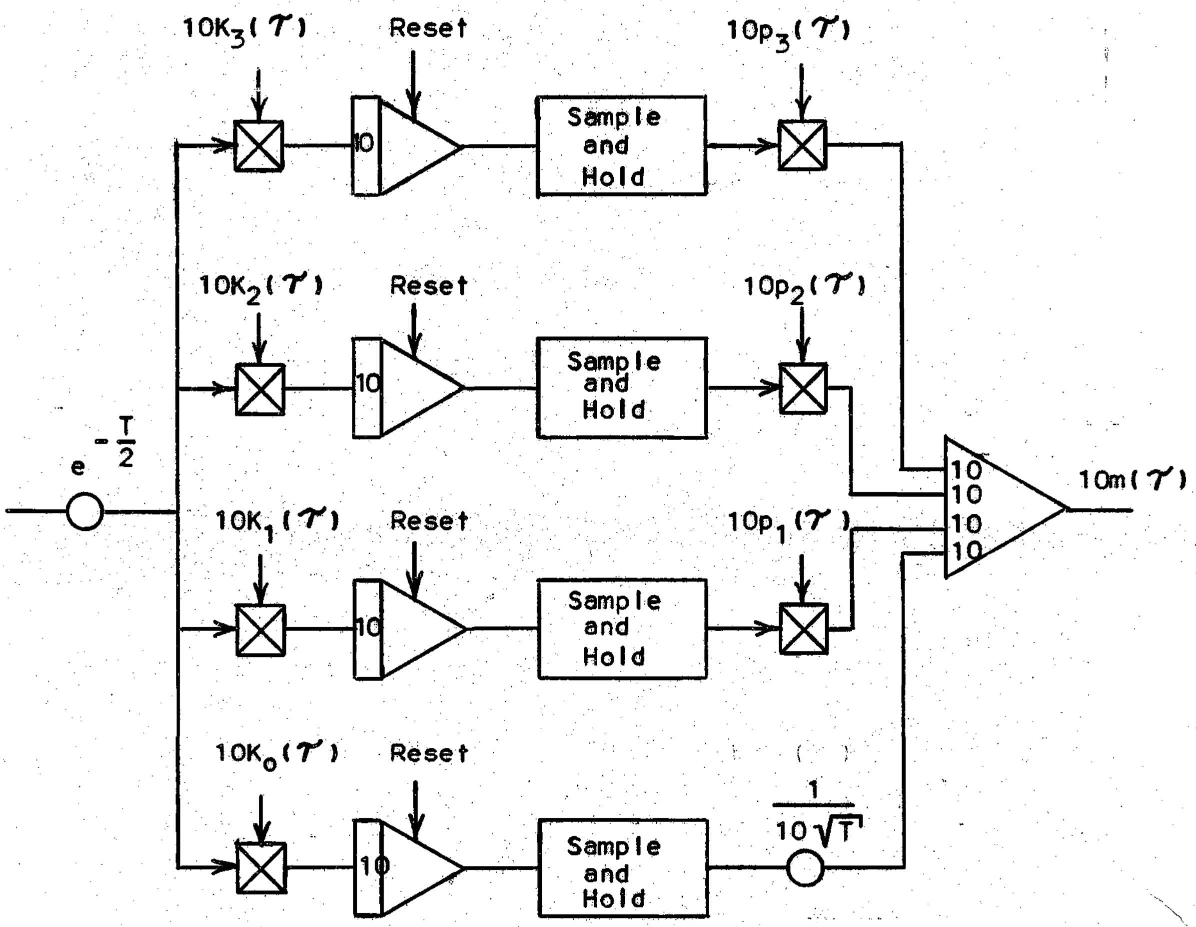


Fig. D-9

Second-order Process Controller.

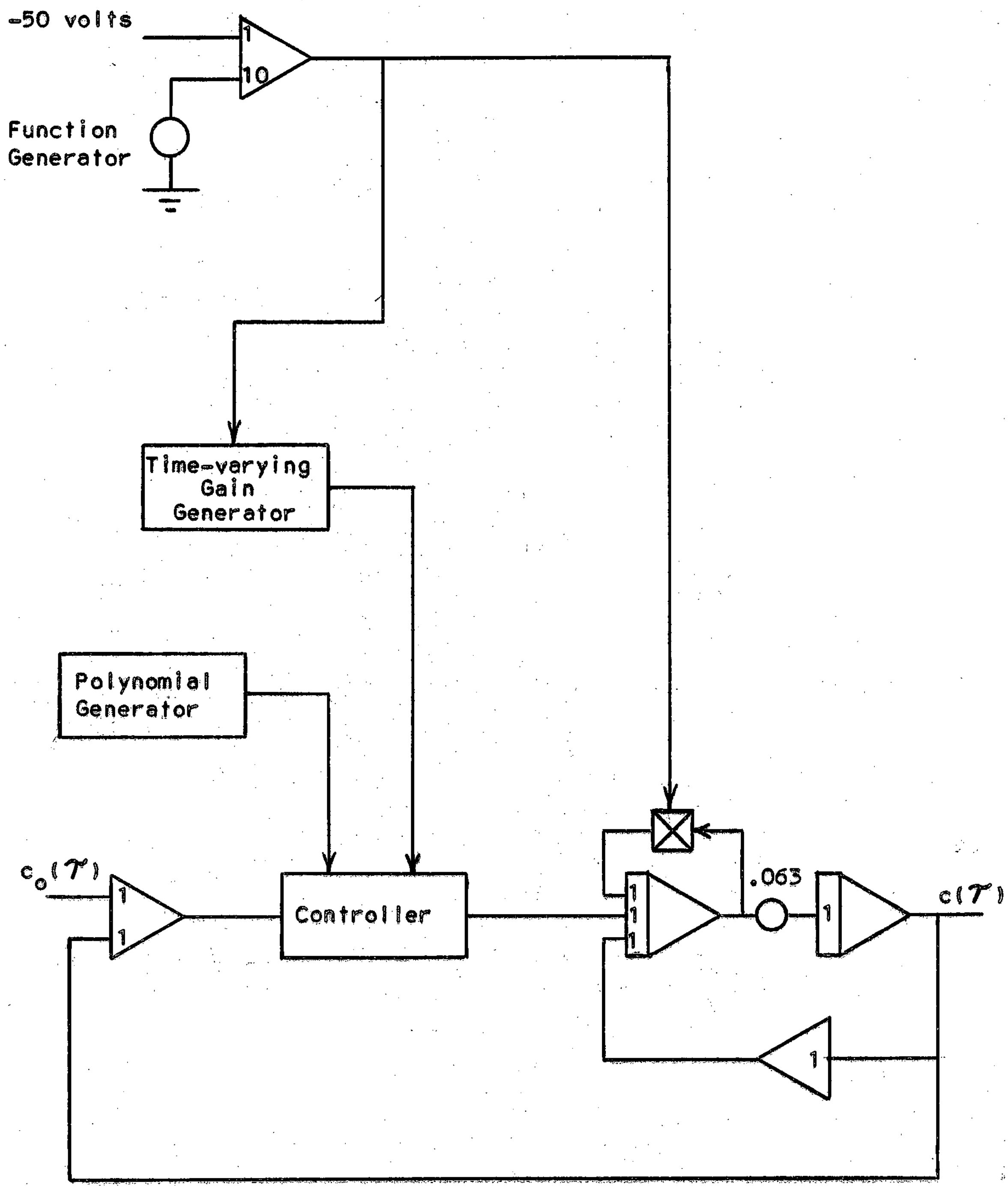


Fig. D-10

Complete Second-order System.

ERRATA

Technical Report No. 2
Vol. 1

Contract AF 33(616)-6890 PRF 2358

Page	Line	Should read	Instead of
14	Bottom	form	for
24	13	Eq. 1-9	Eq. 1-15
26	lower right of figure	$w(t,\tau)$	$w(t,\ )$
31	Eq. 3-8	$\left[c_{o_n} - c_{i_n} - c_n\right]^2$	$\left[c_o - c_i - c_n\right]^2$
33	Eq. 3-13	$\left[c_{o_n} - c_{i_n} - c_n\right]$	$\left[c_o - c_i - c_n\right]$
34	Eq. 3-15	$\left[c_{o_N} - c_{i_N} - w_{NN}m_N\right]$	$\left[c_{oM} - c_{iM} - w_{NN}m_N\right]$
43	4 from bottom	impulse response $w(t,\tau)$	impulse $w(t,\tau)$
48	Eq. 4-18	$i_n(kT) =$	$i_n(kt) =$
48	Eq. 4-19	(4-19)	(4-19
50	Eq. 4-29	$\left[1 - e^{-j2\pi\mu}e^{-Ts}\right]Y_n(s+j\omega_s)$	$\left[1 - e^{-j2\pi\mu}e^{-Ts}Y_n(s+j\omega_s)\right]$
65	3 from bottom	known <u>a priori</u>	know <u>a priori</u>
68	12	Eq. C-21	Eq. C-27
89	6	Eq. 4-9	Eq. 4-11
89	9 from bottom	simulation	stimulation
92	7	Lee [25, Ch. 8]	Lee [26, Ch. 8]
92	12	[25, p. 221]	[26, p. 221]
118	Eq. 3-16	$m(t) =$	$m_k(t) =$
120	Reference 8	Mishkin, E.	Miskin, E.
133	6	case	cases
133	5 from bottom	crosscorrelation	autocorrelation
139	15	identical	indentical