1961-05-01

Analysis and design of control systems by means of time domain matrices

Dorf, Richard Carl

Monterey, California: U.S. Naval Postgraduate School

https://hdl.handle.net/10945/11592

ANALYSIS AND DESIGN OF CONTROL SYSTEMS BY MEANS OF TIME DOMAIN MATRICES

RICHARD C. DORF

5/7 5-4

ANALYSIS AND DESIGN OF CONTROL SYSTEMS
BY MEANS OF TIME DOMAIN MATRICES

\+ + + + +

RICHARD C. DORF

ANALYSIS AND DESIGN OF CONTROL SYSTEMS

BY MEANS OF TIME DOMAIN MATRICES

by

Richard Carl Dorf

Submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

United States Naval Postgraduate School
Monterey, California

May, 1961

ANALYSIS AND DESIGN OF CONTROL SYSTEMS

BY MEANS OF TIME DOMAIN MATRICES

by

Richard Carl Dorf

This work is accepted as fulfilling the

Dissertation requirement for the degree

DOCTOR OF PHILOSOPHY

from the

United States Naval Postgraduate School

ABSTRACT

The aim of this dissertation is to present a new method of engineering analysis and design for complex control systems. This method is the time domain infinite matrix method. The formulation of the infinite matrix follows from the convolution summation of sampled data systems. The mathematical basis of the time domain matrix formulation is presented in a discussion of the applicable concepts of infinite matrices and sequence spaces. This method of analysis and design is applicable to both continuous data and sampled data systems. For continuous systems it is necessary to introduce a fictitious sampler and hold of sufficient sampling rate to effect an accurate approximation.

The time domain matrix method is presented and illustrated as a method of analysis and design of linear, nonlinear, and time varying systems of the continuous or sampled data class. Sampled data, time varying systems may not be conveniently investigated by any other existing method. Furthermore the investigation of nonlinear systems is greatly simplified by the time domain approach. Multiloop systems may be treated with ease and the signals at intermediate points throughout the loops are readily available. Also, systems with multiple nonlinearities may be investigated, for which there is not a presently available method of analysis and design.

Two methods of design of a discrete compensator for a sampled data system are presented. These methods are accomplished directly in the time domain and allow for a compromise of specifications in the time domain. Also the response between sampling instants is

accounted for in one of the two design procedures.

The time domain matrix method may be readily programmed on a digital computer and therefore provides a rapid analysis and design technique.

TABLE OF CONTENTS

TABLE OF CONTENTS

[illegible]

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS

LIST OF TABLES

TABLE OF SYMBOLS AND ABBREVIATIONS

Symbol	Description
s	The Laplace complex variable
z	z-transform variable ($=e^{sT}$)
T	The sampling period or interval
r(t), R(s), R(z)	Input Signal
C(t), C(s), C(z)	Output Signal
D(z)	Transfer function of discrete compensator
G(s), G(z)	Transfer function of the plant or system
k,n	An interger
K	Gain
[G]	System Transfer Matrix
[D]	Discrete Compensator Matrix
R]	Input sequence column matrix
C]	Output sequence column matrix
$r^*(t)$, $R^*(s)$	Sampled Input Signal
	A sampler

CHAPTER I

INTRODUCTION

During the last two decades, feedback control systems have become increasingly important to our technolcgical civilization, and have particularly contributed to national defense. Simultaneously, there has been increased interest and effort in the investigation of automatic control systems with respect to their analysis and design with various excitation signals.

The I.R.E. defines a feedback control system as a control system comprising one or more feedback control loops, which combines functions of the controlled signals with functions of the commands to tend to maintain prescribed relationships between the commands and the controlled signals. Feedback control systems are a large class of systems which include many subclasses of which a few are linear systems, nonlinear systems, multivariable systems, time varying systems, sampled-data systems, and adaptive systems. The active and dynamic elements are not limited and may be electronic, electromechanical, hydraulic, or pneumatic.

An important subclass of control systems, sampled-data control systems, is a dynamical system which operates with sampled or quantized information. That is, the information is present as a sequence of discrete numbers in time, in contrast to continuous data systems for which the controlling information is monitored continuously in time. Ordinarily, the information is carried in the amplitude of the samples and may be considered pulse amplitude modulation. The block diagram of a typical sampled-data control system is shown in figure 1-1, where the sampled-controller may be a special or general purpose digital computer.

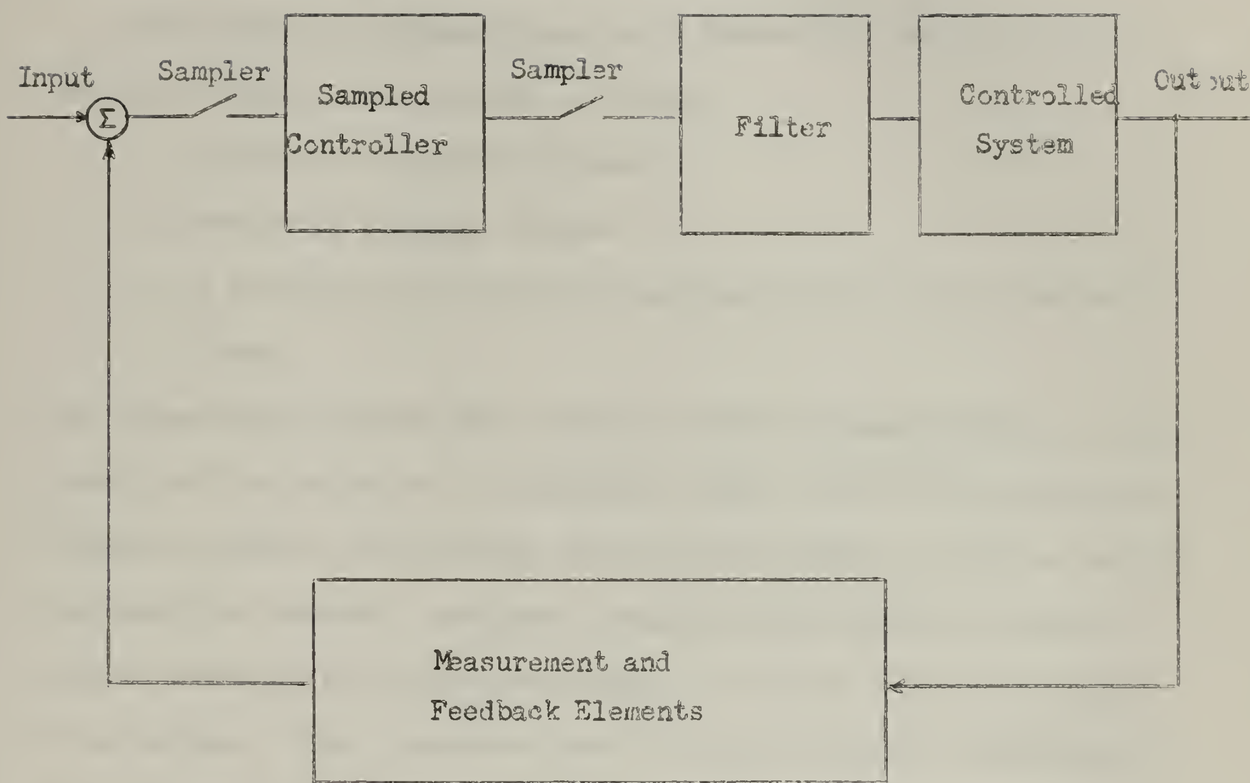


Figure 1-1. Sampled Data Control System Block-Diagram

The wide use of digital computers today has further stimulated investigation of sampled-data systems. The seemingly unlimited possibilities of the use of a digital computer as an active element allow the possibility of considering these systems as a beginning towards the realization of the much discussed adaptive and learning control systems which may imitate the human brain and nervous system.

There exist at present three major methods for the analysis and design of sampled-data systems as follows:

1. Difference equation approach
2. Frequency response methods
3. Z transform and modified transforms and the use of the root locus.

All these methods assume that sampling occurs instantaneously or therefore, that the pulses are of negligible width. The first method uses classical methods of difference equations and yields a solution only at the sampling instants. Frequency response methods are an attempt to extend concepts from continuous systems and suffer from the attendant disadvantages. The z transform uses a complex-variable transformation and determines the performance of the system by location of roots on loci in the complex z plane, and by inverse transformation.

It is the purpose of this dissertation to present a method of engineering analysis and design for complex control systems. This method is the use of infinite matrices in the time domain. Analysis and design for this method takes place directly in the time domain and avoids the necessity of transformations in complex-variables. This method has proven to be potentially useful for the investigation of a variety of systems.

Time domain analysis and design has the important advantage of affording investigation directly in the domain of interest and direct evaluation of performance, therefore avoiding use of correlation theorems which are complex and may be inaccurate. This method avoids the difficulty of solution that is present for higher order systems using the z-transformation. In fact, the amount of work necessary for investigation is approximately the same for a first order as for any n^{th} order system.

This method can be applied to many classes of systems which either do not lend themselves or are not possibly investigated by frequency or z-transform techniques. Therefore, beyond linear system investigation, one may analyze and design nonlinear and time-varying systems. Also the investigation of continuous systems is made possible through the introduction of a fictitious sampler and hold of suitably high sampling rate. The error introduced by this approximation of a continuous system by a sampled-data system can be reduced to a negligible amount with an attendant increase in labor of calculation.

This dissertation is presented in three parts. The first part comprising Chapters 2 and 3 presents the mathematical background and formulation of the time domain matrix equations. The second part comprised of Chapters 4 and 5 presents the analysis and design methods and verification of the theory by application to various types of systems. The third part comprised of Chapters 6, 7, and 8 is concerned with the design and realizibility of the digital compensator in the time domain, and the final conclusions and possibilities for future investigation.

CHAPTER 2

THE TIME DOMAIN MATRIX

2-1 Introduction

The method of investigation of sampled-data control systems most commonly used today is the z-transformation, as is correspondingly the s-transform for continuous systems. The z-transformation converts the difference equations to a set of simultaneous algebraic equations which may be solved for the unknown variable and then by means of the inverse transform yields the response in the time domain. Formulation of the problem directly in the time domain allows one to avoid this transformation and inverse transformation.

2-2 A Sampled-data System and the z Transform

A block diagram of a simple open loop sampled-data system is shown in figure 2-1. The definition of the z-transform of $x^*(t)$ is [1]

$$X(z) = X^*\left(\frac{1}{T} \ln z\right) = \sum_{n=0}^{\infty} x(nT) z^{-n} \tag{2-1}$$

where $z = e^{-sT} = u + jv$

and T = sampling period

Then it can be shown that the output at the sampling instants $Y(z)$ is:

$$Y(z) = G(z)\, X(z) \tag{2-2}$$

where $G(z)$ is the z transform of $G(s)$.

In order to obtain the response between the sampling instants Jury[2] introduced the modified z-transform where the response between the samples is:

$$Y(z,m) = G(z,m)\, X(z) \tag{2-3}$$

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

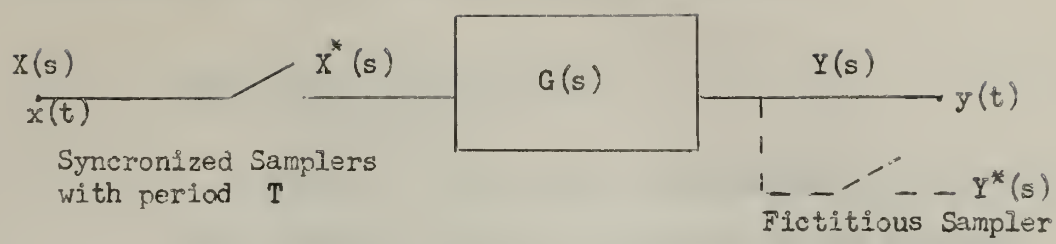


Figure 2-1. Open Loop Sampled System

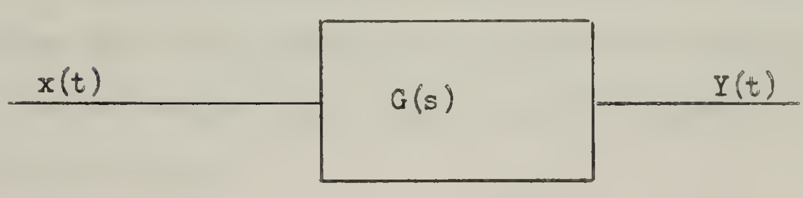


Figure 2-2. An Open Loop Continuous System

In order to obtain the response in the time domain, the inverse transformation must be utilized of which one form is:

$$y(nT) = \frac{1}{2\pi j} \oint_{\Gamma} G(z) z^{n-1} \, dz \qquad (2\text{-}4)$$

where Γ is the path of integration in the z-plane which encloses all the singularities of the integrand.

It is found that the usefulness of the transform lies with the use of the root loci on the z-plane. However, one desires to obtain the response directly in the time domain and that will be the subject of the next section.

2-3 The Convolution Summation

The time response of a continuous system G(s) as shown in figure 2-2 to a continuous input x(t) is given by the convolution integral:

$$y(t) = \int_{-\infty}^{t} g(t-\lambda)\, x(\lambda)\, d\lambda \qquad (2\text{-}5)$$

where λ = dummy variable time delay

and $g(t-\lambda)$ = the delayed impulse response of the system G(s).

This method of obtaining the response is usually avoided in favor of the Laplace transform due to the difficulty of evaluation of the integral.

However, for the sampled data system as shown in figure 2-1, the input signal may be written:

$$x^*(t) = \sum_{n=0}^{\infty} x(nT)\, \delta(t-nT) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \qquad (2\text{-}6)$$

where $[x]$ is a column matrix of n elements where n extends to infinity

Now, in contrast to equation 2-5, one may write an equation for the continuous output response as a convolution summation:

$$y(t) = \sum_{k=0}^{n} g(nT - kT)\, x(kT) \tag{2-7}$$

$$\text{or} \quad y_n = \sum_{k=0}^{n} g_{n-k} x_k \quad \text{or} \quad y_n = \sum_{k=0}^{n} g_k x_{n-k} \tag{2-8}$$

where g_{n-k} is often referred to as the weighting sequence.

The values at the sampling instants of the impulse response, g_n, are related to g(t) the impulse response as:

$$g_n = g(nT) = g(t)\Big|_{t = nT} \tag{2-9}$$

Therefore, when the input is an ideal impulse of unity height, then:

$$x_{n-k} = \begin{cases} 1 & \text{for } k = n \\ 0 & k \neq n \end{cases}$$

and $y_n = g_n$ the weighting sequence.

If equation 2-8 is expanded, one obtains:

$$\begin{array}{llllllllll}
n = 0 & y_0 & = & g_{0-0} & x_0 & & & & & \\
n = 1 & y_1 & = & g_{1-0} & x_0 & + & g_{1-1} & x_1 & & \\
n = 2 & y_2 & = & g_{2-0} & x_0 & + & g_{2-1} & x_1 & + & g_{2-2} \; x_2 \\
\vdots & \vdots & & \vdots & & & & & & \\
n = n & y_n & = & g_{n-0} & x_0 & + & g_{n-1} & x_1 & + & g_{n-2} \; x_2 + \cdots \\
 & & & & & & & & + & g_{n-n} \; x_n
\end{array} \tag{2-10}$$

And equation 2-10 may be written as:

$$
\begin{aligned}
y_0 &= g_0 \; x_0 \\
y_1 &= g_1 \; x_0 + g_0 \; x_1 \\
y_2 &= g_2 \; x_0 + g_1 \; x_1 + g_0 \; x_2 \\
&\;\vdots \\
y_n &= g_n \; x_0 + g_{n-1} \; x_1 + \ldots g_0 \; x_n
\end{aligned}
\qquad (2\text{-}11)
$$

It can then be seen that this set of equations 2-11 may be written in matrix form as:

$$
\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}
=
\begin{bmatrix}
g_0 & 0 & 0 & 0 & 0 & \cdots & 0 \\
g_1 & g_0 & 0 & 0 & 0 & & 0 \\
g_2 & g_1 & g_0 & 0 & 0 & & 0 \\
g_3 & g_2 & g_1 & g_0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
g_n & g_{n-1} & g_{n-2} & g_{n-3} & \cdot & \cdot & \cdot g_0
\end{bmatrix}
\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ \\ x_n \end{bmatrix}
$$

or $\quad Y] = [G] \, X] \qquad (2\text{-}12)$

where $\quad Y]$ and $X]$ are column matrices of order n

and $\quad [G]$ is a square matrix of order n called the system transfer matrix.[3]

For the case of a time-varying system one then obtains instead of equation 2-7 and 2-8:

$$
y(nT) = \sum_{k=0}^{n} g(nT,kT) \, x(kT) \qquad (2\text{-}13)
$$

and the transfer matrix form for the system G(s) would be:

$$[G] = \begin{bmatrix} g(0,0) & 0 & 0 & 0 & \cdots & 0 \\ g(1,0) & g(1,1) & 0 & 0 & & 0 \\ g(2,0) & g(2,1) & g(2,2) & 0 & & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ g(n,0) & g(n,1) & g(n,2) & g(n,3) & \cdots & g(n,n) \end{bmatrix} \tag{2-14}$$

For example, in order to obtain the output response for a unit step input to a system which is simply an integrator, we have to determine the X and G matrices. For a unit step input, the input at the sampling instants is always one or:

$$X] = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

To determine the system matrix one must determine the weighting values at the sampling instants. If the period T is equal to one second then one simple method of determining the sequence is to find $G(z)$ and to divide:

$$G(z) = \frac{z}{z-1} = 1 + z^{-1} + z^{-2} + z^{-3} + \cdots \tag{2-15}$$

Another, more generally useful method is to determine the impulse response $g(t)$ and find the values at the sampling instants. For an integrator, $G(s) = 1/s$ or:

$g(t) = u(t)$ a unit step.

Therefore, the system matrix is:

$$[G] = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & & 0 & 0 \\ 1 & 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \tag{2-16}$$

Then, the output can be obtained using matrix multiplication:[3]

$$Y] = \left[G\right] X]$$

$$Y] = \begin{bmatrix} 1 & 0 & 0 & \\ 1 & 1 & 0 & \cdots \\ 1 & 1 & 1 & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ 1 & 1 & 1 & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ \cdot \\ \cdot \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ \vdots \end{bmatrix} \qquad (2\text{-}17)$$

2-4 The Evaluation of the System Transfer Matrix of a Transfer Function

In order to determine the output response of a system, the system transfer matrix must be available. This is determined with the most facility and accuracy by determining the $g(t)$ by taking the inverse transform of $G(s)$. In order to evaluate the matrix values substitute $t = nT$. This method of evaluation of the matrix is illustrated in Appendix A and Table A-1 gives the values for some representive systems.

In many cases, a system has an undetermined $G(s)$ or frequency response and the $G(s)$ must be determined experimentally. In this case, it would be as convenient to determine the impulse response of the system directly and therefore the g_n values. For a large percentage of the sampled systems, a hold circuit filters the output of the sampler as shown in figure 2-3. In this case, the unit impulse is converted to a unit pulse of one sampling period width. Therefore, the z-transform equation for the output $Y(z)$ is:

$$Y(z) = G_hG(z)X(z)$$

$$\text{where } G_hG(z) = \mathcal{Z}\left\{G_hG(s)\right\} \qquad (2\text{-}18)$$

Therefore, we are interested in determining the matrix $\left[G_hG\right]$, and experimentally the values of this matrix may be determined by exciting the system $G(s)$ with a unit pulse of period T.

This method has been verified experimentally and yielded values within three percent of the expected elements of the matrix. This method of determination of the pulse response of the controlled system is easy to accomplish, and quite useful in investigating components with unknown transfer functions.

2-5 Response Between the Sampling Instants

In any but the most well behaved system, the response between the sampling instants is of interest and must be determined. For this purpose the modified z transform was introduced in the z domain, and an analogous method must be determined for the time domain matrix. Writing equation (2-9) one has:

$$g_n = g(nT) = g(t)\Big|_{t = nT} \tag{2-19}$$

for the values at the sampling instants. Therefore, if the values at half-way between each sample are to be determined ($m = 1/2$ in the modified transform) one has:

$$g_n(m = 1/2) = g(nT + \tfrac{1}{2}T) = g(t)\Big|_{t = nT + \frac{1}{2}T} \tag{2-20}$$

The symbol m was chosen to be consistent with the modified z-transform and is defined as the percentage of the period from the sample point, as shown in figure 2-4. The index m can assume a value 0 to 1 and is at the n^{th} sample when $m = 0$ and the $n-1^{th}$ sample when $m = 1$.

For example, if the output between the sampling instants is required for the system shown in figure 2-5, then one writes the equation for the output as:

$$\left.Y(m)\right] = \left[G(m)\right] \left.X\right] \tag{2-21}$$

The impulse response of the system is:

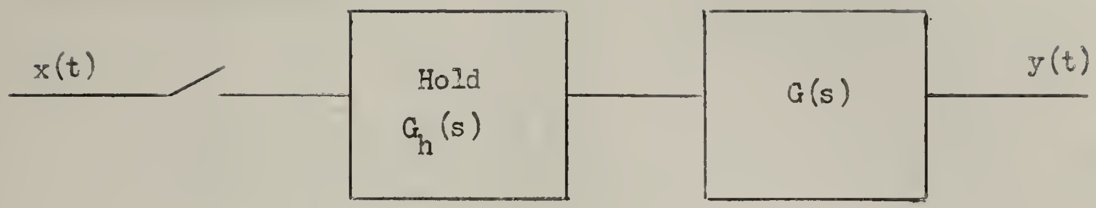


Figure 2-3. Open Loop Sampled System With Hold

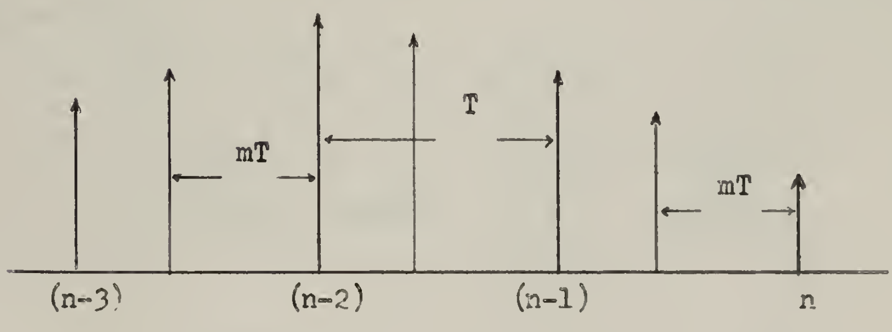


Figure 2-4. The Intersample Response

$$g(t) = e^{-t}$$

Then, the values of the transfer matrix at the sampling instants (m = 0) is:

$$g_n = e^{-n}$$

and the matrix is:

$$\left[G\right] = \begin{bmatrix} 1 & 0 & 0 \\ e^{-1} & 1 & 0 \\ e^{-2} & e^{-1} & 1 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ e^{-n} & & \end{bmatrix} \qquad (2\text{-}22)$$

In order to determine the values at half-way between the sample points one must determine $\left[G(m)\right]$ where m = 1/2. Therefore

$$g_n(1/2) = e^{-(n + 1/2)}$$

$$\text{and} \quad \left[G(1/2)\right] = \begin{bmatrix} e^{-1/2} & 0 & 0 & \cdot\ \cdot \\ e^{-1.5} & e^{-.5} & 0 & \\ e^{-2.5} & e^{-1.5} & e^{-.5} & \cdot\ \cdot \\ & \cdot & & \\ & \cdot & & \end{bmatrix} \qquad (2\text{-}23)$$

Then, the output at the sampling instants and at the mid-point of the sampling period is:

$$\left.c\right] = \begin{bmatrix} 1 & 0 & 0 \\ (.3679) & 1 & 0 \\ (.1353) & (.3679) & 1 \\ \cdot & & \\ \cdot & & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} 1 \\ 1.3679 \\ 1.5032 \\ \\ \end{bmatrix} \qquad (2\text{-}24)$$

and

$$\left.c(1/2)\right] = \begin{bmatrix} .6065 & 0 & 0 \\ .2231 & .6065 & 0 \\ .0821 & .2231 & .6065 \\ \cdot & & \cdot \\ \cdot & & \cdot \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ \\ \end{bmatrix} = \begin{bmatrix} .6065 \\ .8296 \\ .9117 \\ \\ \end{bmatrix} \qquad (2\text{-}25)$$

The determination of these values is discussed further in Appendix A and values are given for the systems considered.

2-6 The Formulation of the Matrix Equation for Closed Loop Systems

Before considering the closed loop system, one must consider the two block open loop system as shown in figure 2-6 and examine the matrix algebra.

Then, the necessary equations are:

$$B\,] = [\,G_1\,]\; X\,] \quad \text{and} \quad Y\,] = [\,G_2\,]\; B\,] \tag{2-26}$$

Therefore, $Y\,] = [\,G_2\,]\,[\,G_1\,]\; X\,]$ (2-27)

and, in general, since matrix multiplication is not commutative it is incorrect to write the transfer matrices in the reverse order, that is:

$$Y\,] = [\,G_2\,]\,[\,G_1\,]\; X\,] \neq [\,G_1\,]\,[\,G_2\,]\; X\,] \tag{2-28}$$

Now, for a closed loop control system as shown in figure 2-7 one has:

$$e(t) = r(t) - (c)t \quad \text{and} \quad e^*(t) = r^*(t) - c^*(t)$$

In matrix form one obtains:

$$E\,] = R\,] - C\,] \quad \text{and} \quad C\,] = [\,G_2\,]\,[\,G_1\,]\; E\,] \tag{2-29}$$

Therefore, one has:

$$E\,] = R\,] - [\,G_2\,]\,[\,G_1\,]\; E\,] \quad \text{and} \quad E\,] = \left\{[\,I\,] + [\,G_2\,][\,G_1\,]\right\}^{-1} R\,] \tag{2-30}$$

Therefore

$$C\,] = [\,G_2\,]\,[\,G_1\,] \left\{[\,I\,] + [\,G_2\,]\,[\,G_1\,]\right\}^{-1} R\,] \tag{2-31}$$

where $[\,I\,]$ = identity matrix $= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & & \\ 0 & \cdot & & 0 & 1 \end{bmatrix}$

and $[A]^{-1}$ = inverse of matrix A. Thus, the solution for the output response involves matrix multiplication, addition, and inversion. Fortunately, one finds the inversion process is simplified by the fact

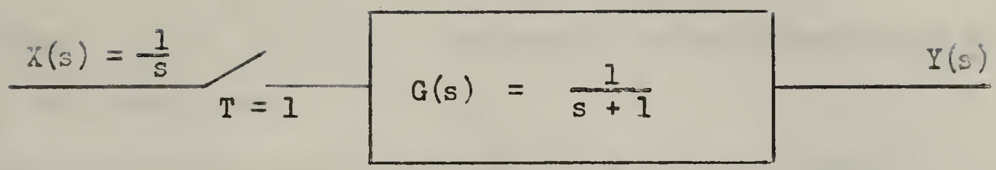


Figure 2-5. First Order Sampled System

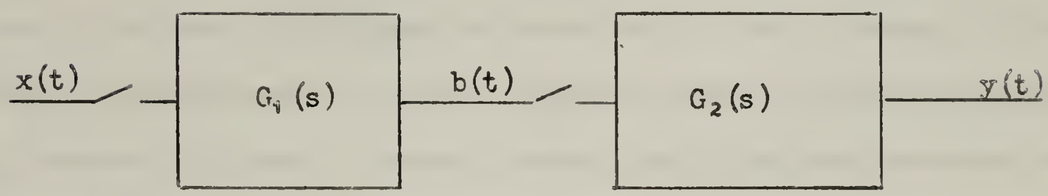


Figure 2-6. Two Block Open Loop System

that the matrices are all lower trangular matrices; that is, all the elements above the diagonal are zero. Also, one finds equation 2-31 may be written as:

$$C] = [K] R] \quad \text{where} \quad [K] = [A]\{[I] + [A]\}^{-1} \tag{2-32}$$

where $[A] = [G_2][G_1]$ = the product of the transmission matrices in the forward path.

Post multiplying equation 2-32 by $[I] + [A]$ one obtains:

$$[K]\{[I] + [A]\} = [A] \tag{2-33}$$

Adding $[I] - [I] = [0]$ to both sides, then:

$$[K]\{[I] + [A]\} = [A] + [I] - [I] = \{[I] + [A]\} - [I]$$

Therefore, postmultiplying by the inverse of $[I] + [A]$, one obtains:

$$[K] = [I] - \{[I] + [A]\}^{-1} \tag{2-34}$$

which substitutes subtraction for more difficult multiplication necessary in equation (2-32).

2-7 The Solution for the Response of a Simple System

In order to evaluate the response of a closed-loop system, the inverse of the matrix $[I]+[A]$ must be determined. Since a physical system always is represented by a lower triangular matrix, one powerful method for inversion is given by Frazer, Duncan, and Collar and presented in Appendix B.[4] Consider a simple approximate first order transfer function system where the sampling period is one second and $a \gg 1$, so that a neglible delay will be introduced, but there will be no output at the $n = 0$ sampling instant. The system with a step input is shown in figure 2-8.

The system matrix is determined in Appendix A and given in Table A-1 and rewritten here:

$$[G] = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & & \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & \vdots & & \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix} \tag{2-35}$$

then

$$[I] + [G] = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 1 & 1 & 0 & \\ 1 & 1 & 1 & \\ & & \vdots & \end{bmatrix}$$ and

$$\{[I] + [G]\}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 \cdots & 0 \\ \vdots & \vdots & \vdots & \\ 0 & 0 & 0 \cdots & 1 \end{bmatrix} \tag{2-36}$$

Therefore,

$$[I] - \{[I] + [G]\}^{-1} = \begin{bmatrix} 0 & 0 & 0 & & \\ +1 & 0 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & & & & 0 \end{bmatrix} \tag{2-37}$$

so that

$$c] = \begin{bmatrix} 0 & 0 & & 0 & \\ 1 & 0 & & 0 & \\ 0 & 1 & & 0 & \\ 0 & 0 & & 1 & \cdots \\ & & \vdots & & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{2-38}$$

The values between the sampling instants after the first period would be the same as at the sampling instants, since $[G(m = 0)] = [G(m)]$ for $n > 1$.

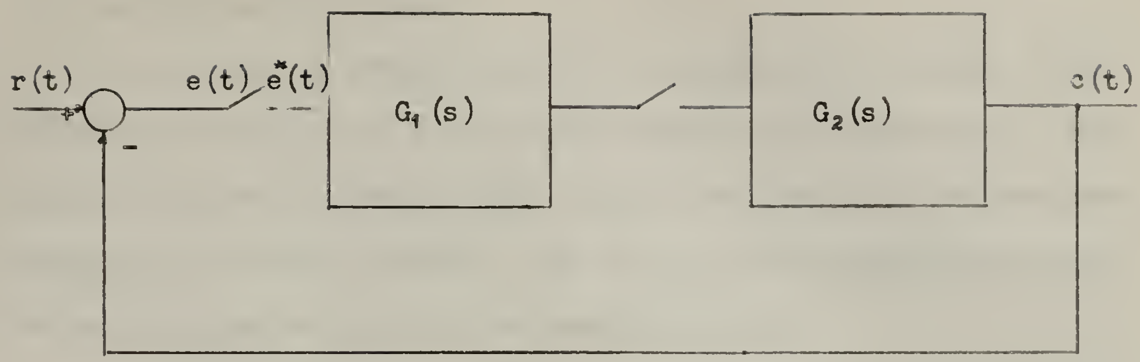


Figure 2-7. Closed Loop Sampled System

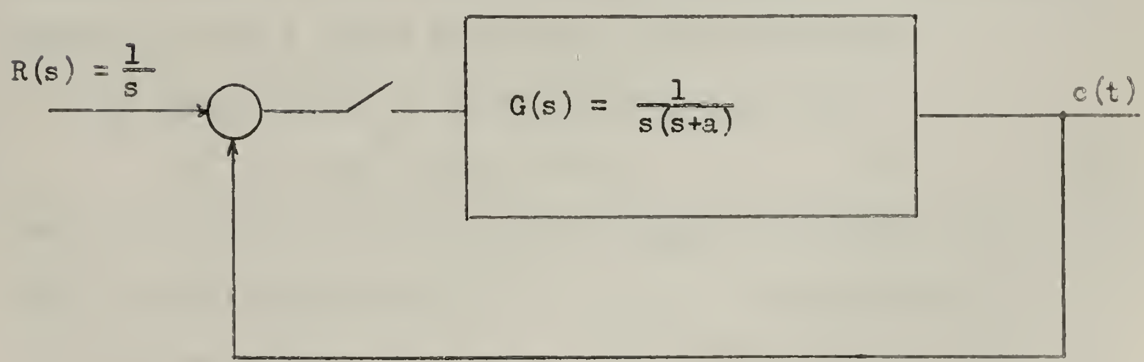


Figure 2-8. Closed Loop System with A Step Input

CHAPTER 3

THE MATHEMATICAL THEORY OF INFINITE MATRICES

3-1 The Mathematical Theory

It was shown in chapter two that the output time response could be obtained with the use of a matrix formulation. Furthermore, these matrices consist of elements whose values are those of a time response evaluated at discrete instants. For an open loop sampled-data system the matrix equation 2-12 will be rewritten here:

$$Y\rfloor = [G]X\rfloor \tag{3-1}$$

The column matrices $Y\rfloor$ and $X\rfloor$ contain the values of the discrete response and input respectively. If the response is to be determined for all time, then the order of the column matrix, and the square system matrix, is infinite. Of course, even if one does not need to evaluate the infinite number of response values, the system matrix can be considered as an infinite matrix; that is, of large order. The content of this chapter is a discussion of infinite matrices and infinite sequences, and the calculation of the time response of a sampled-data system utilizing a matrix formulation in the time domain.

The system matrix G is an infinite matrix since

$$[A] = A = (a_{ij}) \quad (i,j = 1,2,3, \cdot \cdot \cdot n \ldots \infty) \tag{3-2}$$

that is, the matrix A is an array of elements of infinite order. Henceforth, in this chapter, let $[A] = A$ and the column matrices $X\rfloor = x$ and $Y\rfloor = y$ for convenience in notation. Cooke[5] discusses, at length, the characteristics of infinite matrices. Only the characteristics of immediate importance shall be mentioned here. In general, the theory of infinite matrices is connected with mathematical analysis and the theory of functions.

The operations of interest for the infinite matrices are:

$$A + B = (a_{ij} + b_{ij})$$

$$AB = \sum_{k=1}^{\infty} a_{ik} b_{kj} \tag{3-3}$$

$$\lambda A = \lambda a_{ij}$$

where λ is a scalar.

The input and output time responses are expressed as an infinite sequence of discrete values. These discrete values constitute the elements of the column vectors x and y. A vector x may be considered a sequence space. A definition of sequence space is:[5]

A set S of sequences is called a sequence space when it contains the origin, and is such that, for every x and y in S and for every (complex) scalar c, x+y and cx are in S.

The sequence space of interest is called σ, the space of all sequences. Then, the matrix equation

$$y = Ax \qquad \text{or} \qquad y_n = \sum_{k=0}^{n} a_{nk} x_k \tag{3-4}$$

is a linear transformation of σ on itself, that is, the system with matrix A transforms the input sequence into another sequence, the output sequence. When

$$a_{nk} = a_{n-k} \tag{3-5}$$

the system is time invariant as previously discussed in section 2-3.

The infinite, positive time, matrices for physical systems are always lower triangular matrices (L.T.M.), that is:

$a_{ij} = 0$ when $j > i$.

Also, for time invariant systems the elements are equal along the upper left to lower right diagonal. This matrix is called a diagonally invariant matrix (D.I.M.) and exists when:

$$a_{ij} = a_{i-j} \quad \text{for every } i,j. \tag{3-6}$$

It is important to note, that a matrix multiplication is non-commutative for the class of systems which are time-varying (T.V.). These time-varying systems, for the operations defined in equation 3-3, are characterized by a non-Abelian (non-commutative) algebra. The time invariant systems may be called a sub-class (T.I.) of the large class T.V. and are characterized by a Abelian algebra. That is, for time invariant systems, matrix multiplication is commutative. Thus for linear constant coefficient systems the order of the matrix multiplication may be reversed, while for nonlinear or time varying systems they may not.

Inversion of the infinite matrix A is often necessary and it is shown in reference 3, that if there is no i where $a_{ii} = 0$, and A is nonsingular, then a unique inverse of A exists. The evaluation of the inverse of the L.T.M. is discussed in Appendix B.

3-2 Convergence and Stability in the Sequence Space.

In an automatic control system it is of great importance to determine if the overall closed-loop system is stable and therefore the output time response is bounded. The output response is a sequence in the sequence space and it must be determined if the sequence is bounded. Therefore, if a step function test signal is applied to the system it is necessary to determine if the output sequence converges to a final value. If the system response diverges, the system is considered unstable.

For a closed loop control system, the output response may be written

$$y = Ax \tag{3-7}$$

where A is the closed loop system matrix and

$$A = G(I + G)^{-1} \tag{3-8}$$

The G is an open loop system matrix as shown in figure 3-1

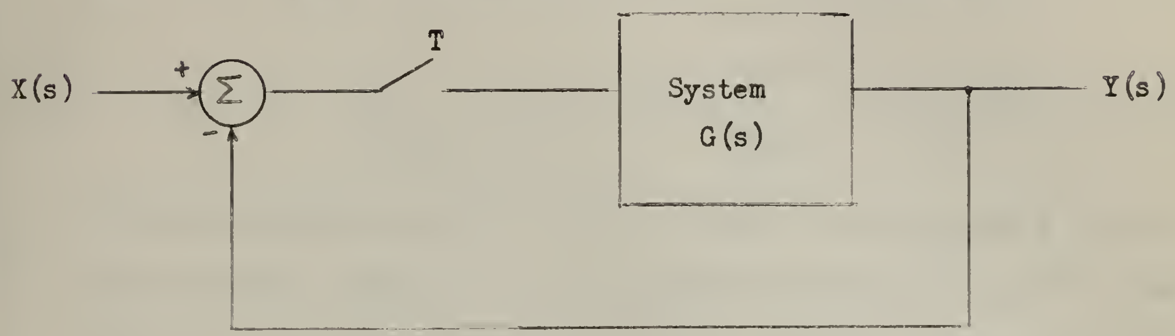


Figure 3-1. Closed Loop Sampled Data System

Then, given an input signal of a convergent nature, that is, of a bounded nature, the problem is to determine if the output sequence y is of a bounded nature. Mathematically, the problem is, given a convergent sequence x, under the transformation A, does a convergent sequence y result ? A theorem of fundamental importance concerning this problem is that of Kojima-Schur.[5]

Kojima-Schur theorem: the necessary and sufficient conditions that

$$y = Ax \quad \text{or} \quad y_n = \sum_{k=1}^{\infty} a_{nk} x_k \tag{3-9}$$

should tend to a finite limit as $n \to \infty$ whenever x_k is convergent are that

$$\text{(a)} \quad \sum_{k=1}^{\infty} \left| a_{nk} \right| \leq M \text{ for every } n \tag{3-10}$$

$$\text{(b)} \quad \lim_{n \to \infty} a_{nk} = \alpha_k \text{ for every fixed } k, \tag{3-11}$$

(c) $\lim_{n \to \infty} \sum_{k=1}^{\infty} a_{nk} = \alpha$ (3-12)

Moreover, if $\lim_{k \to \infty} x_k = \bar{x}$ then the final value is

(d) $\bar{y} = \lim_{n \to \infty} y_n = \alpha \bar{X} + \sum_{k=1}^{\infty} \alpha_k (x_k - \bar{X})$ (3-13)

A matrix satisfying a,b,c is called a k-matrix with α_k and α for its characteristic numbers. For, the special condition of $\alpha_k = 0$ in condition (b), one defines the matrix A as a T_α matrix. These are the conditions for convergence and clearly define the stability of a control system in the time domain. By examination of the impulse response of stable systems a physical understanding of these conditions results. Consider, for an example, a time invariant open loop system with a transfer function

$$G(s) = \frac{1}{S} \text{ and } T = 1 \text{ as discussed in section 2-3.}$$

Then $a_{nk} = a_{n-k}$ and $a_m = 1.00$ where $m = n - k$. It is obvious that condition (b) yields $\alpha_k = 1.00$. Furthermore, it can be seen that this system does not satisfy conditions (a) and (c). That is, examining condition (a) one finds for n ,

that $\sum_{K=1}^{\infty} \left| a_{nk} \right| \to \infty$. This result is as expected since the response is diverging as was determined in section 2-3, equation 2-17. Therefore condition (a) requires that the series $\sum \left| a_{nk} \right|$ converge in order for the system to be stable. For a closed loop system of figure 3-1, it is important at this point to recall that $A = G \left\{ I + G \right\}^{-1}$. Therefore, A

may satisfy the condition (a) while G may not.

Theorem: If A is the matrix for a time-varying system (T.V.), the system is stable if and only if it is a k_r matrix.

Corollary: If A is the matrix for a time invariant system, then it is a D.I.M. and $A = a_{i-j}$, and A is stable if and only if A is a T_α matrix.

For the large class of time invariant systems, the corollary implies that for stability it is necessary for

$$\sum_{n=0}^{\infty} \left| a_n \right| < \infty \qquad \text{where } n = i - j \tag{3-14}$$

For stable T.I. systems, one finds that $\alpha_k = 0$, that is the matrix is a T_α matrix. For condition (c) one obtains $\alpha = 1.00$ for a type I servo system. Then, by the use of relation (d) one obtains, as expected, the final value as:

$$\bar{y} = \alpha \bar{X} = \bar{X} \tag{3-15}$$

The transformation on σ, accomplished by A as a T_α matrix is called a regular transformation. The property of importance is that as in equation 3-15, the T_α matrices have the property of consistency. That is, every convergent sequence is transformed by such a matrix into another convergent sequence with the same limit when n approaches infinity. In addition, T_α or k matrices will frequently transform divergent sequences into convergent sequences. On this point it may be stated:[5]

Corresponding to each unbounded divergent sequence $\{ S_n \}$ and each bounded sequence $\{ y_n \}$, there is a general (square) T_α matrix which carries $\{ S_n \}$ into $\{ Y_n \}$.

In this chapter, the mathematical basis for the infinite matrices

has been discussed, and the operations defined. Furthermore, the theory of stability and convergence in the sequence space has been discussed. It is not to detract from this, that usually the conditions a), b), c), d) are investigated simultaneously with the evaluation of the actual output sequence for a given system.

CHAPTER 4

ANALYSIS OF CLOSED LOOP CONTROL SYSTEMS

4-1 A Method of Evaluating the Response of a Closed Loop System Without Inversion of Matrices

Consider the simple, error sampled, closed loop system as discussed in chapter two and shown in figure 4-1.

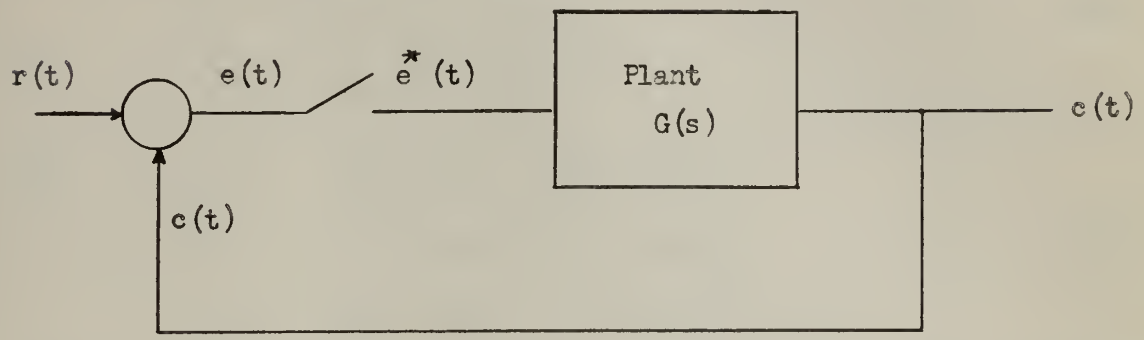


Figure 4-1. Error Sampled System

It was shown that the output response time sequence may be written as:

$$C] = [G]\left\{[I] + [G]\right\}^{-1} R] = [I] - \left\{[I] + [G]\right\}^{-1} R] \quad (4\text{-}1)$$

Therefore, the evaluation of the output response involves the inversion of $[I] + [G]$ as discussed in Appendix B.

However, there is a simple method of evaluating the output response which avoids the inversion of a matrix. The output response column matrix may be written as:

$$C] = [G]\ E] \qquad (4\text{-}2)$$

where $E]$ = the error sequence in time.

The error sequence may be written as

$$E] = R] - C] \qquad (4\text{-}3)$$

It can be seen that the error matrix can be easily evaluated by subtraction. Then the error matrix multiplied by the system matrix will yield

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

the output response. Of course, the error matrix cannot be evaluated until the output sample values are available. However, it may be seen that the output sequence may be evaluated in a step by step procedure by examining the expanded form of equations 4-2 and 4-3 for the first three sampling instants. For a linear, time invariant system, one has:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} g_0 & 0 & 0 \\ g_1 & g_0 & 0 \\ g_2 & g_1 & g_0 \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \end{bmatrix} \tag{4-4}$$

and

$$\begin{bmatrix} e_0 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix} - \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} (r_0 - c_0) \\ (r_1 - c_1) \\ (r_2 - c_2) \end{bmatrix} \tag{4-5}$$

Expanding equation 4-4 with row by row matrix multiplication one obtains the following results.

For the first sampling instant, the time origin, one has:

$$c_0 = g_0 e_0$$

and

$$e_0 = r_0 - c_0 \tag{4-6}$$

Therefore, solving for e_0 one obtains:

$$e_0 = \frac{r_0}{1 + g_0} \tag{4-7}$$

However, for all physical systems, there can not be an output immediately; that is, there is always a time delay of small, but real magnitude. Therefore, $g_0 = 0$ for systems of interest. Then one has

$$e_0 = r_0$$

and

$$c_0 = 0 \tag{4-8}$$

For the second sampling instant one has:

$$c_1 = g_1 e_0 + g_0 e_1 = g_1 e_0 = g_1 r_0 \tag{4-9}$$

and

$$e_1 = r_1 - c_1$$

Since g_1 and r_0 are known, then c_1 may be evaluated and then e_1 may be evaluated. Now, for the third sampling instant one obtains:

$$c_2 = g_2 e_0 + g_1 e_1 + g_0 e_2 = g_2 e_0 + g_1 e_1$$

and

$$e_2 = r_2 - c_2 \qquad (4\text{-}10)$$

Therefore, it can be seen that one may proceed to evaluate the time response sample by sample in time. The equations such as 4-10 become lengthy for larger n and it is simpler to use the matrix form. The matrix form may be written, with $g_0 = 0$, as:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ g_1 & 0 & 0 & \\ g_2 & g_1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \\ g_n & g_{n-1} & g_{n-2} & \end{bmatrix} \begin{bmatrix} r_0 \\ (r_1 - c_1) \\ (r_2 - c_2) \\ \\ (r_n - c_n) \end{bmatrix} \qquad (4\text{-}11)$$

For a linear time invariant system, the system matrix is diagonally invariant. Therefore, multiplication of $[G]\,e]$ as a row-column multiplication may be replaced by a column-column multiplication. This is possible since, for example, the third row is identical to the first column from the third element upwards. That is, one may write the 3rd multiplication as:

$$c_2 = \begin{bmatrix} g_2, & g_1, & g_0 \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} g_2, & g_1, & 0 \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \end{bmatrix} \qquad (4\text{-}12)$$

or alternately as:

$$c_2 = \begin{bmatrix} 0 \\ g_1 \\ g_2 \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \end{bmatrix} = g_2 e_0 + g_1 e_1 \qquad (4\text{-}13)$$

Multiplication is accomplished by multiplying the first element of E] by the last element of [G] of interest.

Then, in general it is possible to write:

$$\left[C\right] = \begin{bmatrix} 0 \\ g_1 \\ g_2 \\ g_3 \\ g_4 \\ \cdot \\ \cdot \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ e_2 \\ e_3 \\ e_4 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} 0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ \cdot \\ \cdot \end{bmatrix} \qquad (4\text{-}14)$$

If a table of the form of equation 4-14 is established, a step by step evaluation procedure may be used as follows:

1) Evaluate $e_0 = r_0$, and $c_0 = 0$

2) Then for c_1, multiply the G and E matrices by starting at g_1 in the G matrix. Then one obtains $c_1 = g_1 e_0$. Now evaluate $e_1 = r_1 - c_1$ and fill in c_1 and e_2 in the computation table.

3) Now evaluate c_2 by multiplying the G and E matrices by starting at g_2. Then one obtains $c_2 = g_2 e_0 + g_1 e_1$. Then evaluate $e_2 = r_2 - c_2$ and place the values in the table.

4) Continue this procedure for each sample point of interest, typically until the system settles to a final value.

An example will best illustrate this procedure. Let the system matrix of figure 4-1 be:

$$\left[G\right] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \\ .4 & 0 & 0 & 0 & 0 & \\ .8 & .4 & 0 & 0 & 0 & \\ 1.00 & .8 & .4 & 0 & 0 & \\ 1.0 & 1.0 & .8 & .4 & 0 & \cdot\ \cdot \\ \cdot & \cdot & & & & \\ \cdot & \cdot & & & & \\ \cdot & \cdot & & & & \end{bmatrix} \qquad (4\text{-}15)$$

Assume a step input signal $R\big] = \big[1,1,1,\ \cdot\ \cdot\ \cdot\ \cdot\ 1\big]$. Then, establishing a table as in equation 14 one calculates each value of c_n

and e_n step by step as outlined. This method is illustrated for the first two calculations by use of dashed lines to indicate the flow of the calculations as follows:

$$
\begin{array}{cccc}
 & G & E & C \\
C\Big] = & \begin{bmatrix} 0 \\ .4 \\ .8 \\ 1.0 \end{bmatrix} & \begin{bmatrix} e_0 = 1.0 \\ e_1 = r_1 - c_1 = 1 - .40 = .60 \\ e_2 \\ e_3 \end{bmatrix} = & \begin{bmatrix} c_0 = 0 \\ c_1 = .40 \\ c_2 \\ c_3 \end{bmatrix}
\end{array}
\qquad (4\text{-}16)
$$

Then, the next step is illustrated schematically as:

$$
\begin{array}{cccc}
 & G & E & C \\
C\Big] = & \begin{bmatrix} 0 \\ .4 \\ .8 \\ 1.0 \end{bmatrix} & \begin{bmatrix} e_0 = 1.0 \\ e_1 = .60 \\ e_2 = 1 - c_2 = -.040 \\ e_3 \end{bmatrix} = & \begin{bmatrix} c_0 = 0 \\ c_1 = .40 \\ c_2 = 1.040 \\ c_3 \end{bmatrix}
\end{array}
\qquad (4\text{-}17)
$$

If this procedure is carried on step by step, one obtains for the response at the first six sample points:

$$
\begin{array}{cccc}
 & G & E & C \\
C\Big] = & \begin{bmatrix} 0 \\ .4 \\ .8 \\ 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} & \begin{bmatrix} 1 \\ .6 \\ -.040 \\ -.464 \\ -.3824 \\ -.0358 \end{bmatrix} = & \begin{bmatrix} 0 \\ .4 \\ 1.040 \\ 1.464 \\ 1.3824 \\ 1.0358 \end{bmatrix}
\end{array}
\qquad (4\text{-}18)
$$

To evaluate the intersample response one proceeds as discussed in section 2-5. Rewriting equation 2-21 one has:

$$
C(m)\Big] = \Big[G(m)\Big]\, E\Big] \qquad (4\text{-}19)
$$

If the G(m) matrix for the midpoint of the sampling period is found to be:

$$\left[G(m)\right] = \begin{bmatrix} .20 & 0 & 0 & 0 & 0 \\ .60 & .20 & 0 & 0 & 0 \\ .90 & .60 & .20 & 0 & 0 \\ 1.00 & .90 & .60 & .20 & 0 \end{bmatrix} \tag{4-20}$$

Then one may calculate the intersampling response as:

$$\begin{array}{ccccc} & & G(m) & E & C \\ C(1/2)\Big] & = & \begin{bmatrix} .20 \\ .60 \\ .90 \\ 1.00 \\ 1.00 \end{bmatrix} & \begin{bmatrix} 1 \\ .600 \\ -.040 \\ -.464 \\ -.3824 \end{bmatrix} & = \begin{bmatrix} .200 \\ .720 \\ 1.252 \\ 1.423 \\ 1.209 \end{bmatrix} \end{array} \tag{4-21}$$

It is valuable to note at this point, that the calculations in the case considered in this section were simplified by the unity feedback condition which yields $E] = R] - C]$. Single loop systems with other than unity feedback shall be considered in the next section.

4-2 Evaluation of the Response of A Closed Loop System with Other than Unity Feedback.

In this section the closed-loop system shall be considered with other than unity feedback as shown for one case in figure 4-2. The samplers are synchronized and with the same sampling period. Then the matrix equations may be written as:

$$C] = [G] E] \tag{4-22}$$

$$E] = R] - B] \qquad \text{where} \qquad B] = [H] C]$$

Therefore $$E] = R] - [H] C] \tag{4-23}$$

Now, to evaluate the output response one uses equations 4-22 and 4-23 and evaluates the c_n and e_n step by step as outlined in the previous section. The only difference is in this case, in order to evaluate $E]$, one must evaluate $[H] C]$ and then substract b_n, the sample value, from r_n to obtain the error e_n at the particular sample. It must be pointed

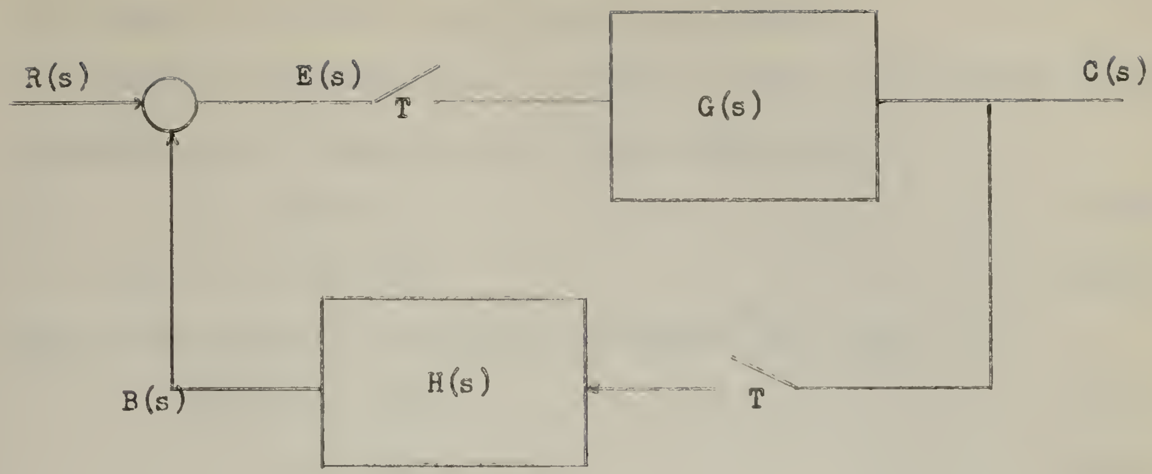


Figure 4-2. Two Sampler Feedback System

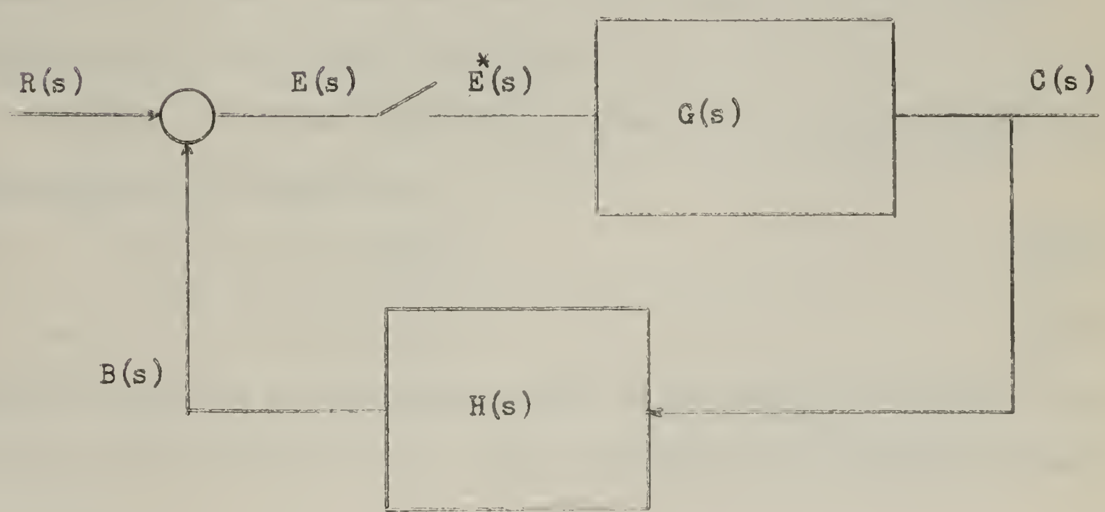


Figure 4-3. Error Sampled System

out that the solution in this case depended on the fact that the transfer blocks G(s) and H(s) were separated by samplers.

Consider the system shown in figure 4-3 which does not have samplers separating both transfer blocks. Then one may write:

$$C\rfloor = [G] E\rfloor \tag{4-24}$$

and
$$E\rfloor = R\rfloor - B\rfloor \tag{4-25}$$

But, noting that $B(s) = H(s)C(s) = H(s)G(s)\,E^*(s)$, then:

$$E\rfloor = R\rfloor - HC\rfloor \tag{4-26}$$

$$= R\rfloor - [HG]\, E\rfloor \tag{4-27}$$

Therefore, it can be seen that $E\rfloor$ is not readily available in equation 4-26 since $HC\rfloor$ is not usually available.

Then, rearranging equation 4-27, one obtains:

$$\left\{ [I] + [HG] \right\} E\rfloor = R\rfloor \tag{4-28}$$

or
$$E\rfloor = \left\{ [I] + [HG] \right\}^{-1} R\rfloor \tag{4-29}$$

Therefore, in order to evaluate the output response for this system, inversion of a matrix cannot be avoided.

Finally, if one was analyzing a system such as shown in figure 4-4, the equations of interest are:

$$E\rfloor = R\rfloor - [H]\, C\rfloor \tag{4-30}$$

and
$$C\rfloor = [G_1]\,[G_2]\, E\rfloor \tag{4-31}$$

If the response at an intermediate point in the system is desired, such as the output of $G_1(s)$, it is readily available by writing the following equation:

$$E_2\rfloor = [G_1]\, E\rfloor \tag{4-32}$$

The response at intermediate points in a system is of great importance in nonlinear systems which will be discussed in section 4-5.

Section 4-3 Analysis of Multiloop Control Systems.

The Introduction of more than one feedback loop is necessary or inherently present in many control systems. The investigation of a multiloop sampled-data control system is complicated by the presence of samplers in some loops, and the absence of samplers in others. Therefore, all the feedback signals are not of the same form throughout the system. Consider at first a two loop system as shown in figure 4-5, which has samplers separating all the transfer blocks.

One may write a set of simultaneous sampled-data equations, where the starred notation indicates a sampled signal or transform, as follows:

$$\begin{aligned} E^*(s) &= R^*(s) - H^*(s)C^*(s) - M^*(s) \\ M^*(s) &= G_a^*(s)E^*(s) \qquad\qquad (4\text{-}33) \\ C^*(s) &= G_b^*(s)M^*(s) = G_b^*(s)G_a^*(s)E^*(s) \end{aligned}$$

Solving these equations simultaneously, one obtains for the sampled-output:

$$C^*(s) = \frac{G_b^*(s)G_a^*(s)R^*(s)}{1 + G_a^*(s) + G_b^*(s)G_a^*(s)H^*(s)} \qquad (4\text{-}34)$$

The z-transformed equation then follows as:

$$C(z) = \frac{G_b(z)G_a(z)R(z)}{1 + G_a(z) + G_b(z)G_a(z)H(z)} \qquad (4\text{-}35)$$

One may then write the matrix equations directly from equations 4-34 and 4-35, or alternatively one may derive the matrix equations directly from the system signals. In either case, one obtains:

$$C\Big] = \Big\{\Big[I\Big] + \Big[G_a\Big] + \Big[G_b\Big]\Big[G_a\Big]\Big[H\Big]\Big\}^{-1}\Big[G_b\Big]\Big[G_a\Big]\,R\Big] \qquad (4\text{-}37)$$

and $$C(m)\Big] = \Big\{\Big[I\Big] + \Big[G_a\Big] + \Big[G_b\Big]\Big[G_a\Big]\Big[H\Big]\Big\}^{-1}\Big[G_b(m)\Big]\Big[G_a(m)\Big]\,R\Big] \qquad (4\text{-}38)$$

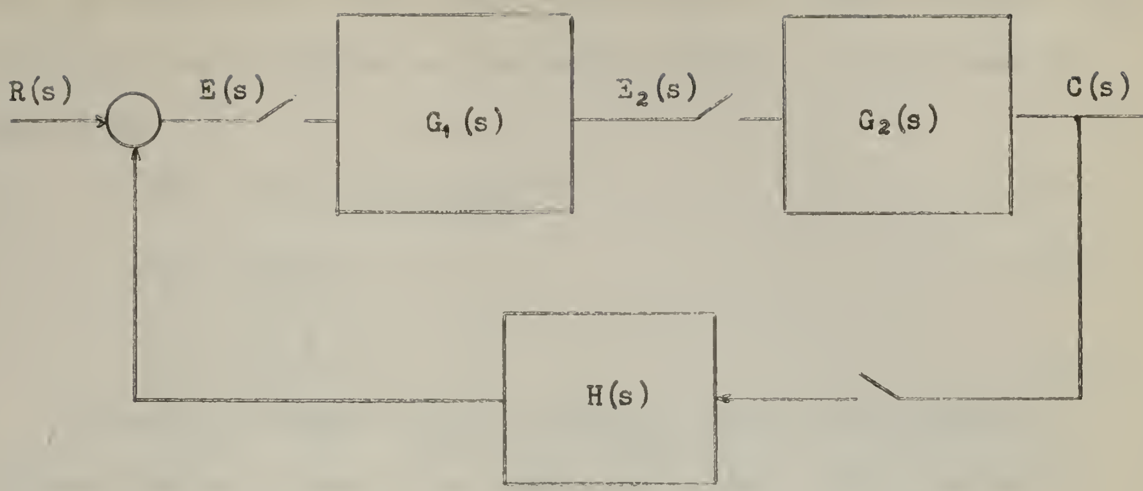


Figure 4-4. Three Sampler Single Loop System

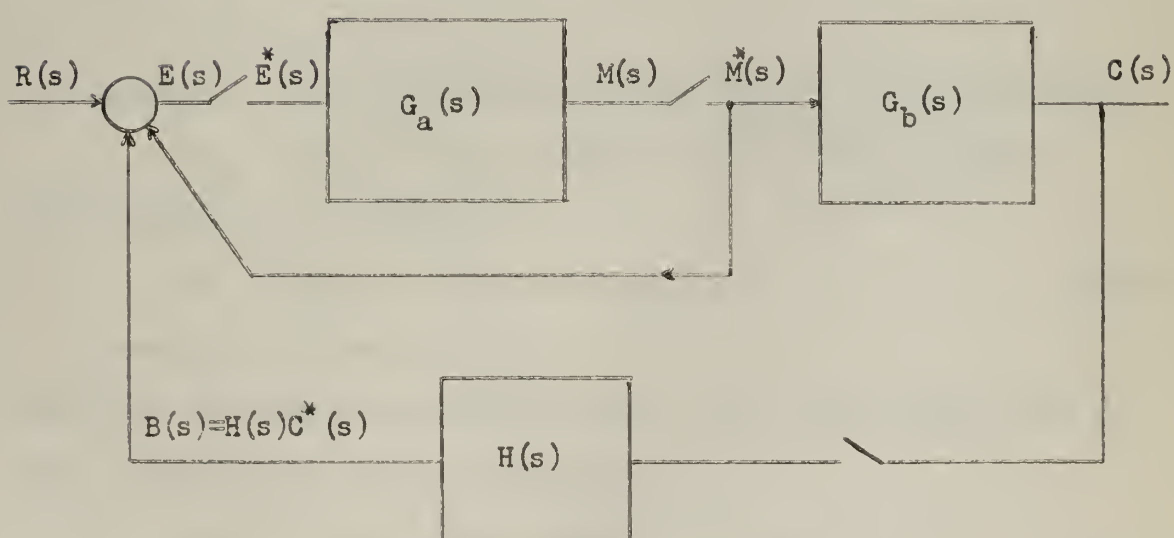


Figure 4-5. Multiloop Sampled System

If it was desired to avoid the inversion necessary in equations 4-37 and 4-38, one may write a set of simultaneous equations which provide a means of step by step evaluation of intermediate signals. In this case one would write:

$$E] = R] - M] - [H] C] \qquad (4\text{-}39)$$

$$M] = [G_a] E] \qquad (4\text{-}40)$$

$$C] = [G_b] M] \qquad (4\text{-}41)$$

In order to evaluate $E]$, $M]$, and $C]$ a step by step procedure is followed similar to that of the single loop method. First, assuming no immediate output for either G_a or G_b, one may write:

$$m_0 = 0$$

$$c_0 = 0 \qquad (4\text{-}42)$$

and $e_0 = r_0$

Then, for the next sample one obtains:

$$m_1 = g_{1a} e_0 = g_{1a} r_0$$

where g_{1a} = the first impulse response value of the G_a block and

g_{na} = the n^{th} impulse response value of the G_a block.

Then, continuing, one obtains:

$$c_1 = g_{1b} m_0 + g_{0b} m_1 = g_{1b} m_0 = 0 \qquad (4\text{-}43)$$

since $g_{0b} = 0$ and $m_0 = 0$.

Therefore, one uses equations 4-39, 4-40, 4-41 in that order, step by step. For the next sample one obtains:

$$e_1 = r_1 - m_1 - (h_0 c_1 + h_1 c_0) = r_1 - g_{1a} r_0 - 0 \qquad (4\text{-}44)$$

since $c_0 = c_1 = 0$

and $m_2 = g_{2a} e_0 + g_{1a} e_1$ (4-45)

and finally

$$c_2 = g_{2b}\, m_0 + g_{1b}\, m_1 = g_{1b}\, m_1 \qquad \text{since } m_0 = 0$$

As is the usual case, it is actually easier and more methodical to use the matrix form throughout the calculations. Therefore, one will establish a table similar to that of section 4-1 using equations 4-39, 4-40, 4-41. The flow of calculation is from $E]$ to $M]$ to $C]$ and then back through the feedback loops again to $E]$. It is obvious, that actually one is simply following the signal around the closed loops.

Finally, consider a multiloop system which possesses only one sampler in the control loops. The simple sampler typically is placed in the error channel as is shown in figure 4-6.

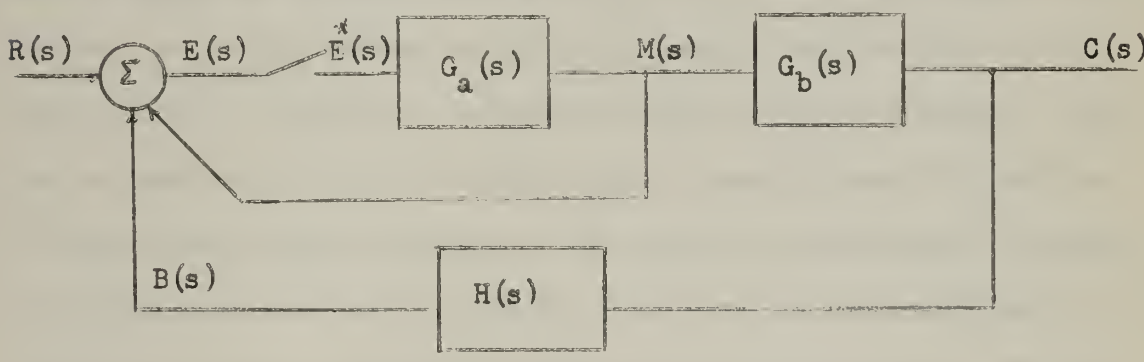


Figure 4-6. A Multiloop Sampled System

One may write the equations for the signals as follows:

$$\begin{aligned} E^* &= R^* - M^* - B^* \\ M &= G_a E^* \\ C &= G_b M = G_b G_a E^* \\ B &= HC \end{aligned} \tag{4-46}$$

Therefore, one obtains for the closed loop sampled error and output:

$$E^* = \frac{R^*}{1 + G_a^* + \overline{HG_bG_a}^*} \tag{4-47}$$

and
$$C^* = \frac{\overline{G_aG_b}^* \; R^*}{1 + G_a^* + \overline{HG_bG_a}^*} \tag{4-48}$$

One may then write the matrix equations directly as:

$$E] = \left\{ [I] + [G_a] + [HG_bG_a] \right\}^{-1} R] \tag{4-49}$$

and
$$C] = [G_aG_b] \left\{ [I] + [G_a] + [HG_bG_a] \right\}^{-1} R] \tag{4-50}$$

The equation for the intersample response is then:

$$C(m)] = [G_aG_b(m)] \left\{ [I] + [G_a] + [HG_bG_a] \right\}^{-1} R] \tag{4-51}$$

Therefore, whenever one desires to calculate the intersample response and the inversion of the matrix $I + G_a + HG_bG_a$ has been already accomplished, one simply evaluates the $G_aG_b(m)$ matrix and carries out the multiplication. Therefore, multiloop systems with one sampler or samplers separating all transfer blocks may be equally treated by the use of time domain infinite matrices. The response at any sampler location is readily available and is usually of interest in the investigation of multiloop systems, particularly with nonlinearities present. The introduction of nonlinearities into a control loop is treated in section 4-5.

Section 4-4. Analysis of Time-varying Control Systems

The analysis of time-varying sampled-data control systems may be accomplished using the time domain matrix method. A control system may have as one of the transfer function blocks, a component whose parameters are changing with time. This effect is present with high altitude jet aircraft, where the dynamic characteristics of the aircraft change with altitude, and therefore time.

Consider the simple open loop system shown in figure 4-7. The impulse response of the time varying component is changing with time.

Figure 4-7. Sampled Open Loop System

Therefore the matrix equation for the output may be written as in section 2-3, equation 2-13 and rewritten here:

$$C\Big] = \Big[G(n,k)\Big] \; R\Big] \tag{4-52}$$

where

$$\Big[G(n,k)\Big] = \begin{bmatrix} g(0,0) & 0 & 0 & \\ g(1,0) & g(1,1) & 0 & \cdots \\ g(2,0) & g(2,1) & g(2,2) & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

In order to substitute numbers in the G(n,k) matrix, the time variation of the system G must be known or determined. As an illustration, consider an abrupt change of

$$G(s) = \frac{(1 - e^{-sT})}{s^2(s + a)}$$ from $a = 1.0$ to $a = 2.0$ at the third sampling instant $(k = 2)$ where $T = 1$ second. Then, using the values of g_n from table A-1 one has for the system matrix:

$$\Big[G(n,k)\Big] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \cdots \\ .3679 & 0 & 0 & 0 & 0 & 0 \\ .7675 & .3679 & 0 & 0 & 0 & 0 \\ .9145 & .7675 & .2838 & 0 & 0 & 0 \\ .9685 & .9145 & .4707 & .2838 & 0 & 0 \cdots \\ .9884 & .9685 & .4960 & .4707 & .2838 & 0 \\ \vdots & \vdots & & \vdots & & \end{bmatrix}$$

Then, of course if this component had feedback introduced as shown in figure 4-8, one would have the following closed loop equation:

$$C] = [G(n,k)] \left\{ [I] + [G(n,k)] \right\}^{-1} R] \qquad (4\text{-}53)$$

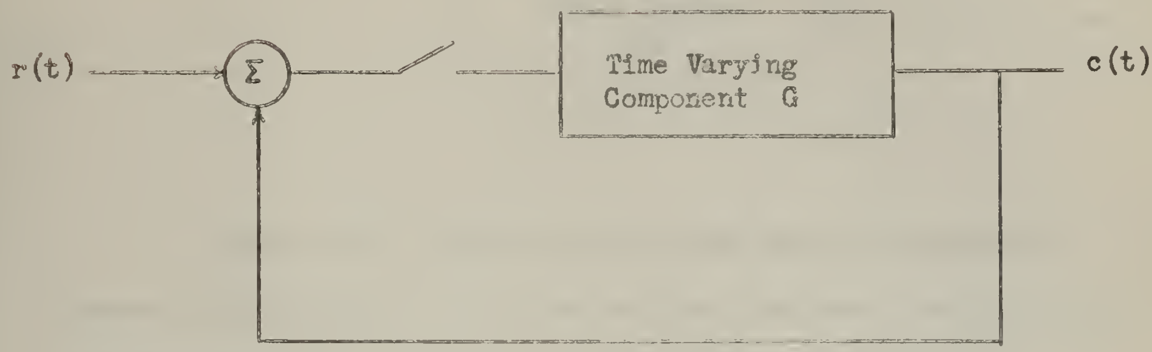


Figure 4-8. Time Varying Closed Loop System

In order to obtain the intersample response, the equation for the output response may be written:

$$C(m)] = [G_{n,k}(m)] \left\{ [I] + [G(n,k)] \right\}^{-1} R] \qquad (4\text{-}54)$$

Analysis of a feedback system with more than one time varying element follows the same approach as for the time invariant systems. The matrix equations are found to be the same as for time invariant systems with the time variation of a transfer function only affecting the system matrix itself. Therefore, for the system shown in figure 4-9, the following equation is obtained:

$$C] = [G(n,k)] \left\{ [I] + [G(n,k)] \; [H(n,k)] \right\}^{-1} R] \qquad (4\text{-}55)$$

It can be seen that it is not necessary for samplers to separate the time varying component from all other transfer blocks. Therefore, if H(s) was time invariant, it would not be necessary to have a sampler

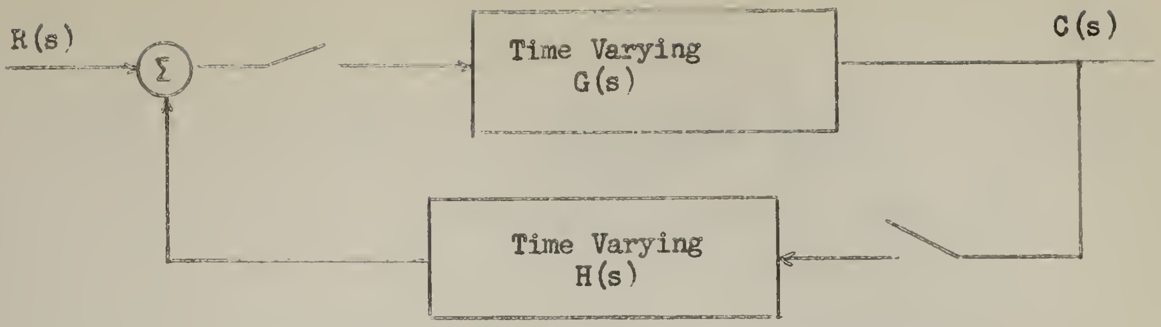


Figure 4-9. System with Two Time Varying Components

between G(s) and H(s) for the use of the time domain matrices. For the system shown in figure 4-10, one obtains the equation:

$$C] = [G(n,k)] \left\{ [I] + [GH(n,k)] \right\}^{-1} R] \tag{4-56}$$

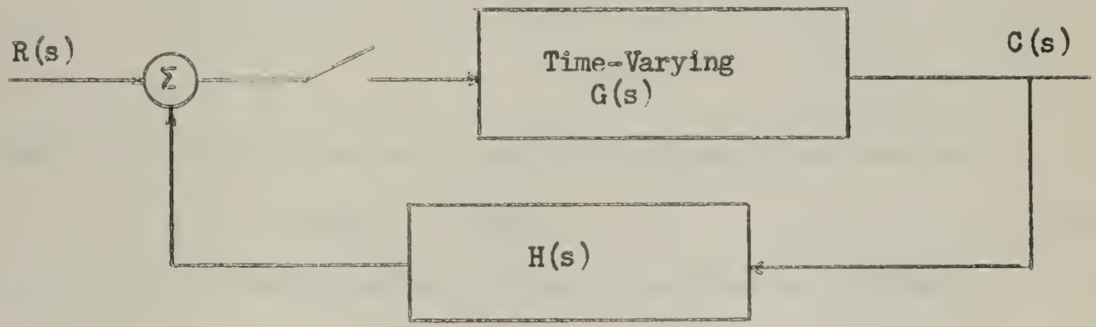


Figure 4-10. Time Varying System with A Feedback Component

A control system, where one of the components is a time-varying amplifier is worthy of consideration. Consider the system as shown in figure 4-11.

The gain a(t) is changing with time and therefore has a different magnitude at each sampling instant. The equation for the output may be written:

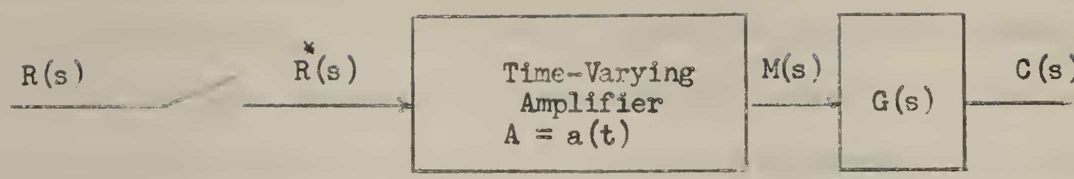


Figure 4-11. Open Loop Time Varying System

$$C] = \left[AG(n,k)\right] R] \tag{4-57}$$

or

$$\begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \\ a_0g_1 & 0 & 0 & 0 & 0 & \cdots \\ a_0g_2 & a_1g_1 & 0 & 0 & 0 & \\ a_0g_3 & a_1g_2 & a_2g_1 & 0 & 0 & \\ a_0g_4 & a_1g_3 & a_2g_2 & a_3g_1 & 0 & \cdots \\ \cdot & & & \cdot & \cdot & \\ \cdot & & & \cdot & \cdot & \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \\ r_4 \\ \\ \\ \end{bmatrix} \tag{4-58}$$

Then, the output at the third sampling instant is as expected:

$$C_2 = a_0g_2r_0 + a_1g_1r_1 \tag{4-59}$$

It can be seen that equation 4-58 may be written as follows:

$$C] = [G]\,[A]\,R] = \begin{bmatrix} 0 & 0 & 0 & 0 & \\ g_1 & 0 & 0 & 0 & \cdots \\ g_2 & g_1 & 0 & 0 & \\ g_3 & g_2 & g_1 & 0 & \end{bmatrix} \begin{bmatrix} a_0 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 \\ 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & a_3 \end{bmatrix} R] \tag{4-60}$$

The operation of multiplication of the A and G matrices is not commutative as was previously discussed. That this is so is obvious from the form of the matrices. Now, consider an open system identical to that of figure 4-11 except that a hold circuit immediately follows the sampler. Then, if the amplifier changes gain during the sample period,

this must be accounted for in the system matrix G, which is a function of m. In most cases it is reasonable to make the simplifying assumption that the gain changes only at the sampling instants. In other words, the assumption is that if the gain is changing continuously with time, the time constant of this change is much greater than the sampling period. Therefore, one is approximating the gain by a staircase function as shown in figure 4-12.

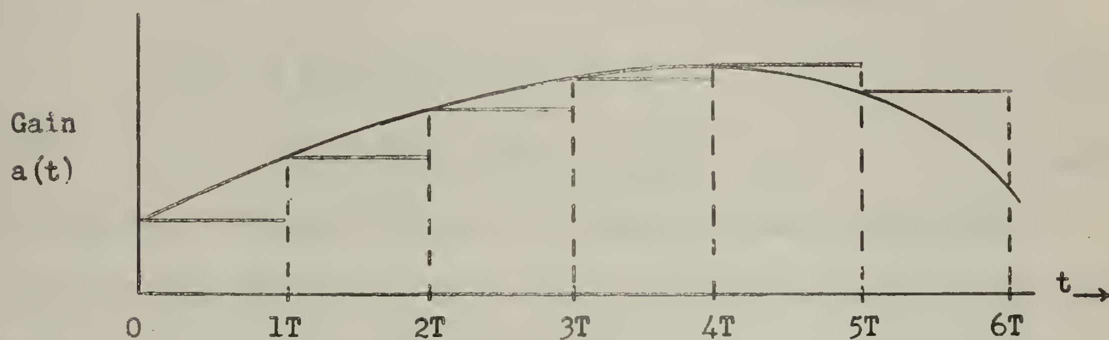


Figure 4-12. A Time Varying Gain

Then, on this basis, one may analyze closed loop systems with a time varying gain. For the system shown in figure 4-13, one may write the equations for the system signals as:

$$E] = R] - [H]\ C] \tag{4-61}$$

and $$C] = [G]\ [A]\ E] \tag{4-62}$$

Equations 4-61 and 4-62 may be solved by the step by step procedure previously outlined for the time-invariant systems. Using this step by step method, the gain change at each interval is clearly displayed to the investigator. Alternately, one may solve for the output response and error response by means of inversion of the matrix in the following equations:

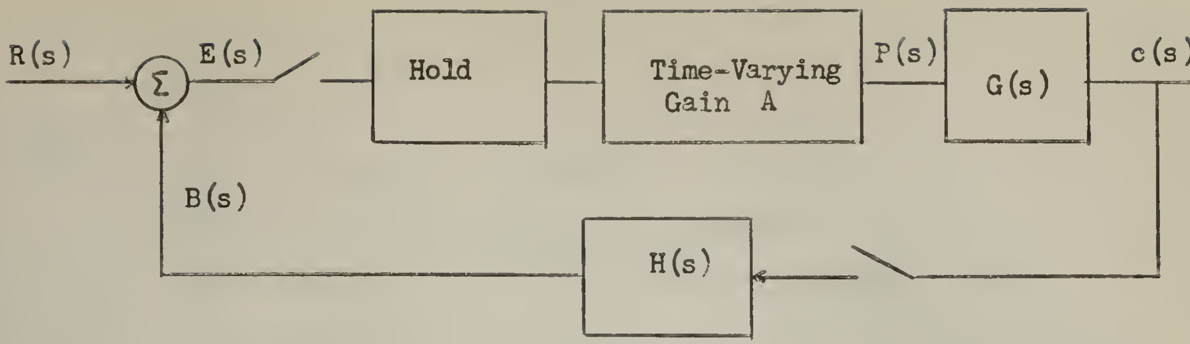


Figure 4-13. System with A Time-Varying Gain

$$E] = \left\{[I]+[H][G][A]\right\}^{-1} R] \qquad (4\text{-}63)$$

and

$$C] = [G][A]\left\{[I]+[H][G][A]\right\}^{-1} R] \qquad (4\text{-}64)$$

If there was no sampler between the output and the feedback block H(s), then one obtains the equation for the error sequence in the same manner as carried out in section 4-2. Therefore, one obtains for the error sequence:

$$E] = \left\{[I] + [HG]\ [A]\right\}^{-1} R] \qquad (4\text{-}65)$$

As an example, consider a system where $H(s) = 1$ as shown in figure 4-13. Let $R(s) = 1/s$, and the gain function be $a(t) = e^{-\frac{1}{2}t}$, with a sampling period of one second. In this the time constant of the gain change is only double the sampling period, but the results are instructive with this marked gain change as shown in figure 4-14.

Use equations 4-61 and 4-62 where $[H] = [I]$ and one may write:

$$E] = R] - C] \qquad (4\text{-}66)$$

$$C] = [G][A]E] = [G]P] \qquad (4\text{-}67)$$

where $P] = [A]E]$ the amplified magnitude of the error pulse

and

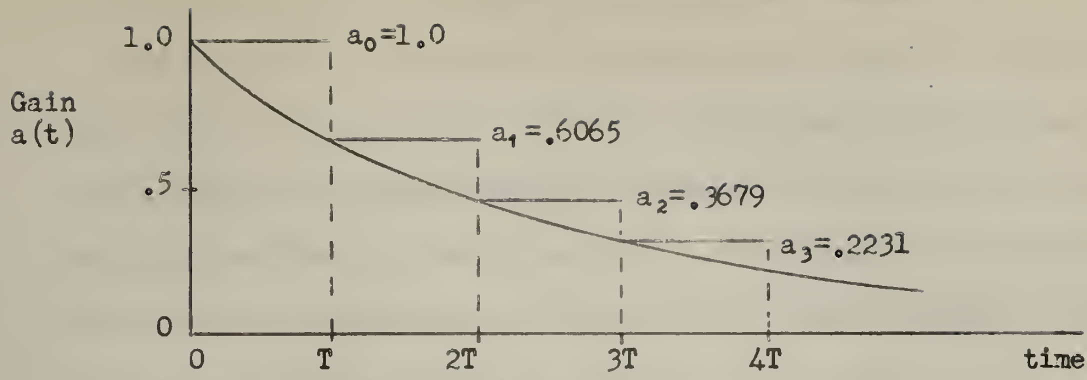


Figure 4-14. Time Varying Gain Example

$$[G] = \begin{bmatrix} 0 & 0 & 0 & \cdot\ \cdot \\ .4 & 0 & 0 & \\ .8 & .4 & 0 & \\ 1.0 & .8 & .4 & \cdot\ \cdot \\ \cdot & & \cdot & \\ \cdot & & \cdot & \end{bmatrix}$$

Then, one obtains:

$$C] = \begin{array}{ccc} G & A & E \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ .4 & 0 & 0 & 0 \\ .8 & .4 & 0 & 0 \\ 1.0 & .8 & .4 & 0 \\ 1.0 & 1.0 & .8 & .4 \end{bmatrix} & \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0 & .6065 & 0 & 0 \\ 0 & 0 & .3679 & 0 \\ 0 & 0 & 0 & .2231 \\ & & & \end{bmatrix} & \begin{bmatrix} 1 \\ .600 \\ .0544 \\ -.2991 \\ \ \end{bmatrix} \end{array}$$

$$= \begin{array}{cc} G & P \\ \begin{bmatrix} 0 \\ .4 \\ .8 \\ 1.0 \\ 1.0 \\ \cdot \\ \cdot \end{bmatrix} & \begin{bmatrix} 1.0 \\ .3639 \\ .0200 \\ -.06673 \\ \cdot \\ \cdot \\ \ \end{bmatrix} \end{array} = \begin{bmatrix} 0 \\ .4 \\ .9456 \\ 1.2991 \\ 1.3532 \\ \cdot \\ \cdot \end{bmatrix} \qquad (4\text{-}68)$$

The availability of the amplified magnitude of the error pulse as P] is often useful in the investigation of nonlinear systems which shall be discussed in the next section.

4-5. Analysis of Control Systems with Nonlinear Components.

The analysis of sampled-data control systems with nonlinear components may be accomplished by the use of time-domain infinite matrices. The methods most commonly used for continuous nonlinear systems are the describing function and the phase-plane. These methods are somewhat limited in their application to sampled-data control systems. Consider a single loop nonlinear system as shown in figure 4-15.

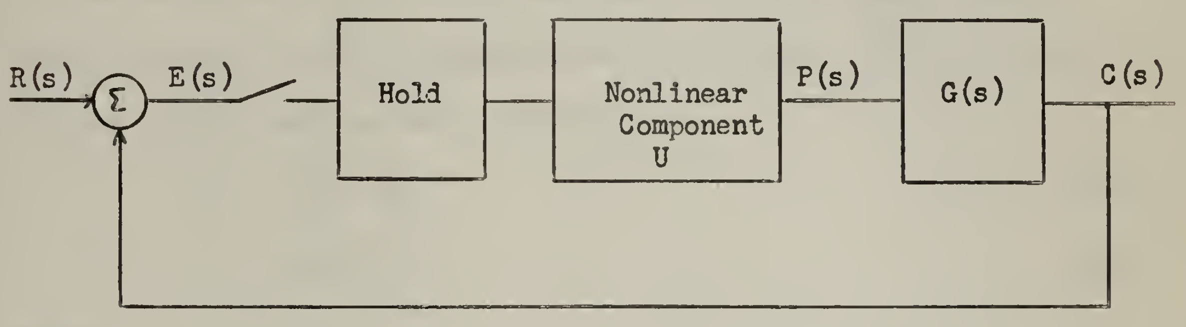


Figure 4-15. Single Loop Nonlinear System

Then, it can be seen that the nonlinear component has replaced the variable gain amplifier of figure 4-13. The nonlinearities considered are those sensitive to the magnitude of the signal input; that is, they have a nonlinear amplitude response such as shown in figure 4-16 for a saturating amplifier.

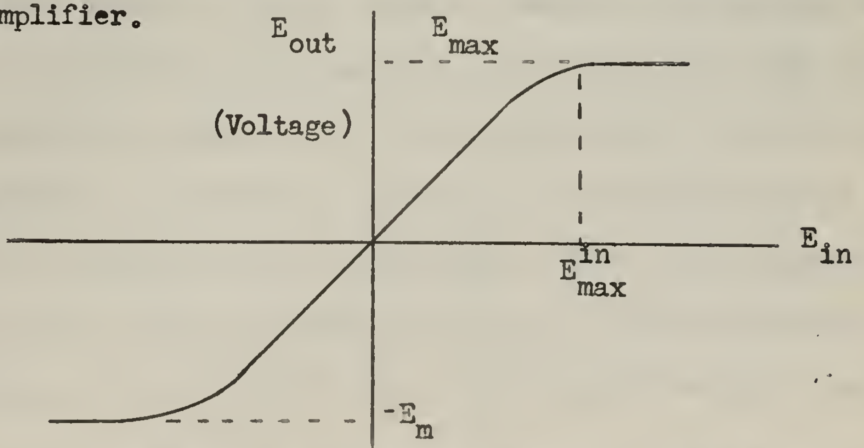


Figure 4-16. Saturating Amplifier Characteristic

Therefore, as the input voltage varies in magnitude, the gain of the amplifier changes. Since the input to the nonlinear element is the held error magnitude, this magnitude determines the gain. However, the error is evaluated directly in the time domain and therefore a gain calculated for a calculated error input is also known as the gain at the specific sample instant. Therefore, the nonlinear element may be treated as a time-varying gain. This statement is true for hysteresis, a relay servo, and other single nonlinearities following a sampler. Then the output of the nonlinearity is the sequence $P]$. One may write for the output response and error response:

$$E] = R] - C] \tag{4-69}$$

$$C] = [G]\,[U]\,E] \tag{4-70}$$

where

$$U] = \begin{bmatrix} u_0 & 0 & 0 & 0 \cdots \\ 0 & u_1 & 0 & 0 \\ 0 & 0 & u_2 & 0 \cdots \\ 0 & 0 & 0 & u_3 \\ \vdots & & & \vdots \quad \cdots \end{bmatrix}$$

The values of u_n depend upon the error magnitude and are determined in the step by step solution. As an example, consider a saturating amplifier with a characteristic curve as shown in figure 4-17, which is the nonlinear component of figure 4-15. If the input signal is a unit step, then the amplifier is expected to saturate. This amplifier has a maximum gain of two throughout the linear region. Then, using equation 4-69, at the first sampling instant $(n = 0)$ $e_0 = 1.0$ and therefore $u_0 = 1.0$, and $p_0 = 1.0$, a gain of one. The G matrix for this example is:

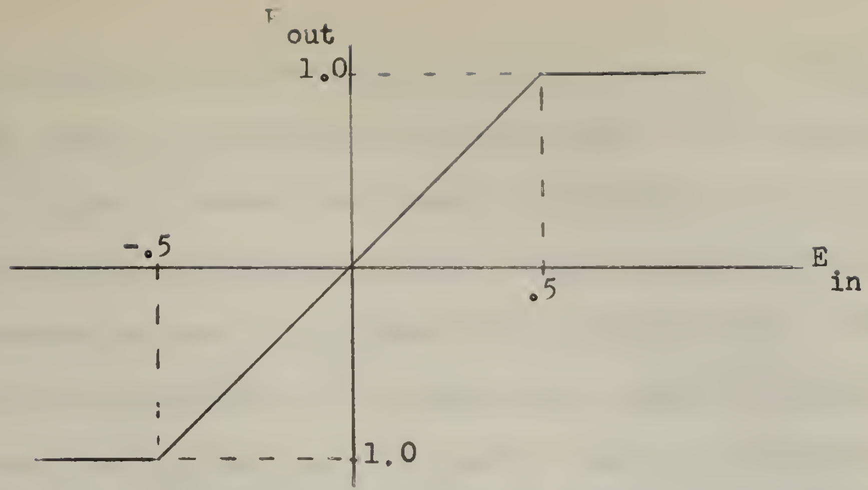


Figure 4-17. Saturating Amplifier Characteristic Example

$$[G] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ .4 & 0 & 0 & 0 \ldots \\ .8 & .4 & 0 & 0 \\ 1.0 & .8 & .4 & 0 \\ 1.0 & 1.0 & .8 & .4 \end{bmatrix}$$

Therefore, $c_1 = g_1 e_0 = .40$ and $e_1 = .60$. This magnitude continues to saturate the amplifier and $p_1 = 1.0$, and $u_1 = 1.667$. Continuing in this manner, one obtains the following matrix solution:

$$C] = [G]\,[U]\,E] = [G]\,P] \tag{4-71}$$

$$= \begin{bmatrix} 0 & \cdots \\ .4 & \cdots \\ .8 & \cdots \\ 1.0 & \\ 1.0 & \\ 1.0 & \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.667 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.56 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .6 \\ -.2 \\ -.64 \\ -.28 \\ +.424 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ .4 \\ .8 \\ 1.0 \\ 1.0 \\ 1.0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -.4 \\ -1 \\ -.56 \\ +.828 \end{bmatrix} = \begin{bmatrix} 0 \\ .4 \\ 1.20 \\ 1.64 \\ 1.28 \\ .576 \end{bmatrix}$$

The analysis of systems with a nonlinear component which is also a storage device follows along in the same manner. For example, if the nonlinearity was hysteresis in a magnetic amplifier, then the gain is dependent upon the past value of input voltage to determine which side of the hysteresis loop is applicable to the present signal. This can be seen clearly from figure 4-18 which shows a simple hysteresis loop. If the input was E_1, there are two possible output magnitudes depending upon the magnitude of the input of the previous sample. If the previous sample had a magnitude of E_0, then the output magnitude will be E_b.

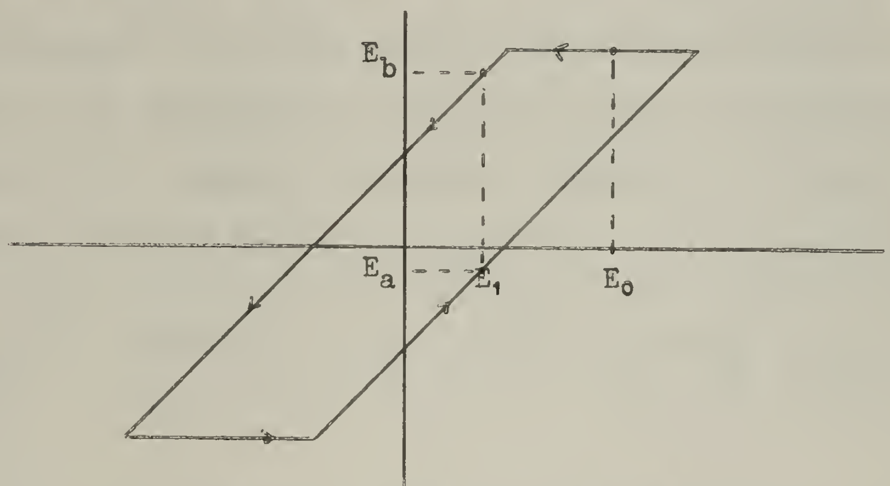


Figure 4-18. Hysteresis Characteristic

Furthermore, any nonlinearity may be treated in this manner since the relation

$$P] = [U] E] \tag{4-72}$$

applies to any nonlinearities whose characteristics are known or can be approximated.

The analysis of systems with a nonlinear component which does not have a sampler immediately preceding it cannot be treated by this method. However, as an approximation a fictitious sampler may be placed before the nonlinearity. The accuracy of this approximation depends upon the

[illegible]

[illegible]

[illegible]

[illegible]

sampling rate and the time constants of the system. This approximation shall be discussed further in the next section.

If there is more than one nonlinearity present in the system and samplers appear before every nonlinearity, then the investigation of the system by time domain matrices is entirely possible. In fact, for a system with many nonlinearities, analysis is possible with the introduction of samplers with a high sampling rate, wherever a nonlinearity exists. The necessary sampling rate shall be discussed in the next section.

Furthermore, the signal sequence magnitudes are available at intermediate points throughout the system and can aid an investigator in the analysis of the component requirements. Consider for example the system of figure 4-19 which has two nonlinearities. Then the equations for the

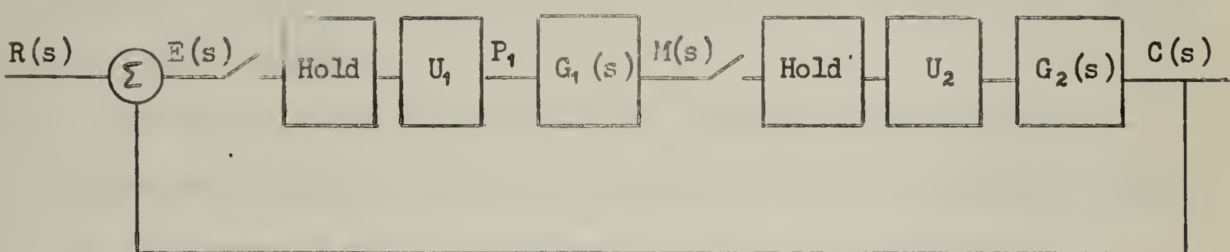


Figure 4-19. System with Two Nonlinear Components

error, the output sequence of G_1, and the output sequence of G_2 is:

$$E] = R] - C] \qquad (4\text{-}73)$$

$$M] = [G_1]\,[U_1]\,E] \qquad (4\text{-}74)$$

$$C] = [G_2]\,[U_2]\,M] \qquad (4\text{-}75)$$

A phase plane portrait of the error and the derivative of the error often aids an investigator in the analysis of a nonlinear system. By the use of the backward difference formulas of numerical methods, the

derivative of the error may be obtained using the past values of the discrete error. The derivative of the error may be calculated by use of the present and one past value as:[6]

$$e'_n = \frac{1}{T}(e_n - e_{n-1}) + 0(T) \qquad (4\text{-}76)$$

where the inaccuracy is of the order of magnitude of the sampling period T. If the present value and two past values of the error are to be used, one has:[6]

$$e'_n = \frac{1}{2T}(3e_n - 4e_{n-1} + 4e_{n-2}) + 0(T^2) \qquad (4\text{-}77)$$

Use of three values of error has reduced the inaccuracy of the approximation to within the order of the sampling period squared. In order for this calculation to be accurate, the sampling period must be short with respect to the time constants of the system. This requirement is not restricting since this condition is necessary for stability in a sampled-data control system.

If it was useful to postulate the phase space and determine the derivatives of higher orders it is possible to use the following formulas, the choice of formula depending on the accuracy required.[6]

$$e''_n = \frac{1}{T^2}(e_n - 2e_{n-1} + e_{n-2}) + 0(T) \qquad (4\text{-}78)$$

$$e'''_n = \frac{1}{T^3}(e_n - 3e_{n-1} + 3e_{n-2} - e_{n-3}) + 0(T) \qquad (4\text{-}79)$$

or

$$e''_n = \frac{1}{T^2}(2e_n - 5e_{n-1} + 4e_{n-2} - e_{n-3}) + 0(T^2) \qquad (4\text{-}80)$$

Formulas for higher derivatives using backward differences are readily available.[7]

Writing the above formulas in a matrix form results in the following matrices with an accuracy of the order of the sample period.

[illegible]

$$\overset{\prime}{E}\Big] = \left(\frac{1}{T}\right)\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \\ -1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & -1 & 1 & 0 & 0 & \\ 0 & 0 & -1 & 1 & 0 & \cdots \\ \vdots & & \vdots & & \vdots & \end{bmatrix} E\Big] \tag{4-81}$$

$$\overset{\prime\prime}{E}\Big] = \left(\frac{1}{T^2}\right)\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ -2 & 1 & 0 & 0 & \\ 1 & -2 & 1 & 0 & \\ 0 & 1 & -2 & 1 & \cdots \\ \vdots & & & & \end{bmatrix} E\Big] \tag{4-82}$$

The matrices with an accuracy of the order of the square of the sample period are:

$$\overset{\prime}{E}\Big] = \left(\frac{1}{2T}\right)\begin{bmatrix} 3 & 0 & 0 & 0 & 0 & \cdots \\ -4 & 3 & 0 & 0 & 0 & \\ 4 & -4 & 3 & 0 & 0 & \\ 0 & 4 & -4 & 0 & 0 & \\ 0 & 4 & -4 & 3 & 0 & \cdots \\ \vdots & & & & \vdots & \end{bmatrix} E\Big] \tag{4-83}$$

$$\overset{\prime\prime}{E}\Big] = \left(\frac{1}{T^2}\right)\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & \cdots \\ -5 & 2 & 0 & 0 & 0 & \\ 4 & -5 & 2 & 0 & 0 & \cdots \\ -1 & 4 & -5 & 2 & 0 & \\ 0 & -1 & 4 & -5 & 2 & \cdots \\ \vdots & & & & \vdots & \end{bmatrix} E\Big] \tag{4-84}$$

4-6. Analysis of Continuous Control Systems

The analysis of continuous control systems can be accomplished with the use of time-domain infinite matrices. This is possible through the introduction of a fictitious (mathematical) sampler or samplers in the

continuous closed loop. The approach is basically that of the use of numerical analysis methods in the solution of a differential equation. Therefore, the accuracy of this approximation depends upon the sampling rate of the fictitious sampler. However, the fact that it is not necessary to calculate the closed-loop roots or to invert the closed-loop response equation in order to obtain the time response, is of great importance.

Consider a continuous system to be investigated, which is a single loop system as shown in figure 4-20.

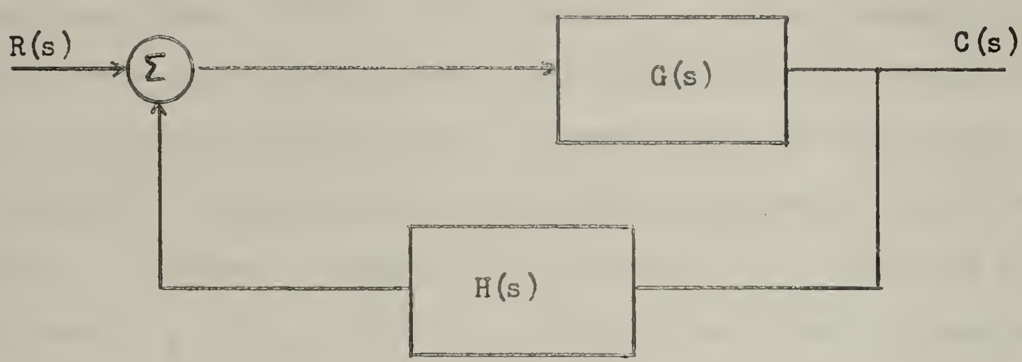


Figure 4-20. Continuous Feedback Control System

The approximation introduced by the fictitious sampler will depend upon:[2]

1) The location of the fictitious sampler

2) The form of the fictitious hold circuit

3) The frequency of the sampling

Since it is important for the sampler frequency to be many times greater than the highest frequency of the input signal to the sampler, the location of the sampler is usually chosen to be in the feedback loop before the H(s) as shown in figure 4-21. This location takes advantage of the filtering of the input signal R(s) by the plant G(s).

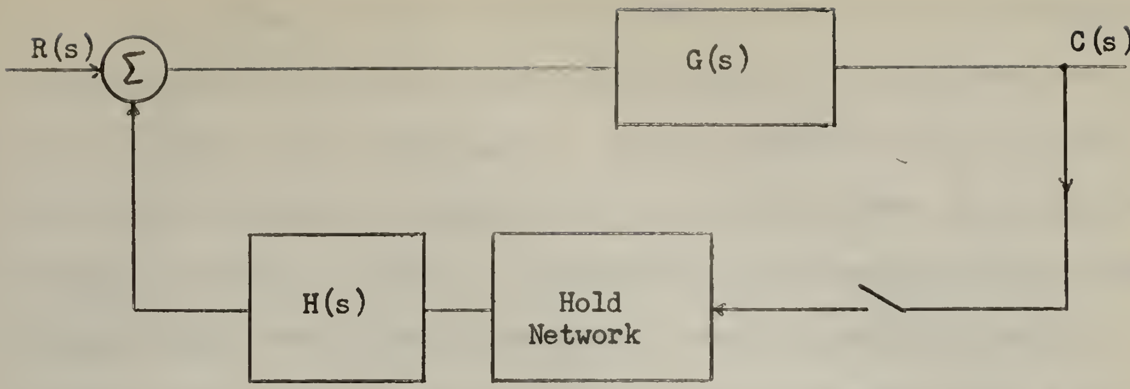


Figure 4-21. Location of the Fictitious Sampler

The fictitious hold network is present in order to reconstruct the continuous signal from the sampled signal. In actual sampled-data systems a zero-order hold is usually used for its practical realizability and for stability considerations. However, for a fictitious hold the possibility of straight line and parabolic approximations should be considered. Consider a straight line approximation hold which can be achieved by a triangular hold circuit which has a transfer function:[2]

$$G_{hold}(s) = \frac{(1 - e^{-sT})^2 \, e^{sT}}{s^2 T}$$

The approximation of a time function is shown in figure 4-22. This hold is not physically realizable, but is very useful for mathematical approximations.

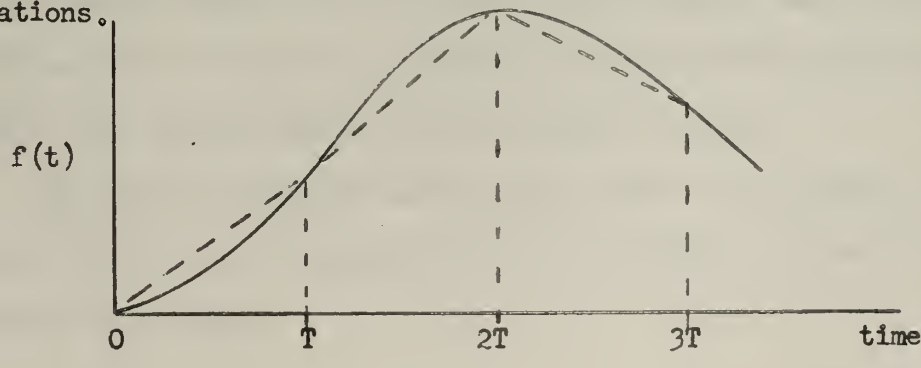


Figure 4-22. The Approximation of A Time Function

The sampling frequency determines the accuracy of the mathematical approximation and must be balanced by the amount of calculation that will be acceptable to the investigator. As the sampling period approaches zero, the inaccuracy is approaching zero, but the number of calculations rapidly approaches infinity. Fortunately, it has been determined experimentally that there exists a reasonable sampling rate for closed loop systems which yields less than 5% error in approximation. For a type I, second order system, this approximation holds when the sampling frequency is ten times the magnitude of the pole of the plant G(s). As an example, for a plant with a transfer function:

$$G(s) = \frac{K}{s(s + 1)}$$

the sampling frequency should be 10 radians per second or the sampling period is approximately one-half second. The calculations of previous sections were carried out with a period of one second and a type one system. Therefore, the number of calculations necessary to determine the approximate response to a continuous-data system are not impractical. The simplest method of determining the necessary sampling rate is trial and error. One calculates a few points on the output response using a trial sampling rate, then recalculates these points using twice the sampling rate. If there is a neglible change in the results, then the former rate was sufficient for the desired accuracy.

This method of approximation may be used on non-linear or time-varying systems and therefore all the previously discussed methods of analysis will apply.

[illegible]

[illegible]

[illegible]

[illegible]

CHAPTER 5

INTRODUCTION TO THE DESIGN OF CLOSED LOOP CONTROL SYSTEMS

5-1. Introduction

The method of time-domain infinite matrices may be used successfully for design of closed loop control systems. The design or development of control systems for a specific application is a practical problem of great importance. The designer may readily apply the analysis methods of the previous chapter to the design problem. They are used to evaluate the performance of the system under the specified operating conditions. Also, with the aid of experience in calculating the response directly in the time-domain, the designer may determine what system parameters must be adjusted and in what manner.

Given the basic specifications of the system and the performance requirements, the designer often must determine a compromise between conflicting requirements. In the design of closed loop continuous systems, many designers rely on the open and closed loop frequency response curves of the system to indicate the performance of the system. These frequency response techniques are not very useful in sampled-data systems. Correlation of the time response of the system with the frequency response is not readily achieved since a transformation of the z variable, into a new complex frequency variable w, is necessary to map the unit circle of the z-plane into the entire left half of the w plane. Therefore, the use of frequency response characteristics, such as the height and frequency of the resonant peak, and the bandwidth, is very limited. Design by use of the root locus method in the z-plane is limited since 1) the design only considers the response at the sampling

[illegible]

instants, and the intersampling response may be wholly unacceptable.
2) Correlation theorems between the time domain and the z-plane are accurate only under certain conditions.

The time domain performance indices such as rise time, settling time, over-shoot, and number of oscillations are readily applied to sampled-data systems using the time domain matrix method. Furthermore, this method allows one to design directly in the time domain and determine the compensator necessary for the specific application. This method of design of a digital compensator shall be discussed in chapter 7.

For the design of continuous systems, the designer has all the standard techniques at his disposal for the selection of a system adjustment or compensator. The methods of Bode, Nichol, and the use of the root locus all may be applied in the design of the continuous system. Then, the sampled data approximation may be introduced and the design evaluated directly in the time domain by means of the time domain matrix method. Therefore, all the background of previous determined theory and methods will apply profitably. The designer may use these techniques to determine the type and location of the compensator, and the parameter values.

The designer may determine the necessary compensator directly in the time domain by use of the digital compensator design technique discussed in Chapter 7. If the system is a continuous system approximated by a sampled-data system, the digital compensator may be converted to a continuous compensator by time domain network synthesis techniques.

In addition to these design problems, one may consider the design of adaptive systems. Due to the great interest of these systems, they

are reserved for discussion in Chapter 6. This chapter shall be concerned with the use of standard design techniques and the use of time domain matrices to design various types of control systems such as nonlinear and time varying systems.

5-2. The Design of A System Utilizing Time Varying Gain Compensation.

Consider the design of a closed-loop, unity feedback sampled-data system as shown in figure 5-1. The basic design steps are:

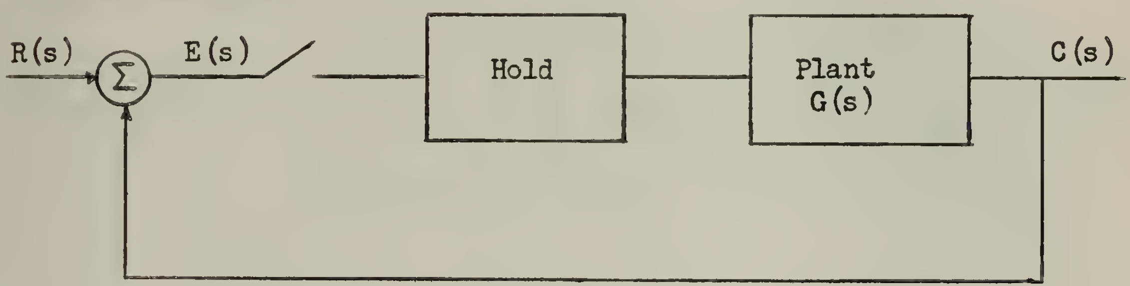


Figure 5-1. Unity Feedback Sampled-Data Control System

(1) Determine the response for the given system G(s) and compare with the required performance specifications.

(2) If the performance is not satisfactory, adjust the gain or choose another plant if possible. Otherwise, introduce a compensating element or block in the closed-loop. Choose the compensating component on the basis of factors such as design criteria and design experience and the time response obtained for the uncompensated system.

(3) Evaluate the compensated system response and readjust the system parameters if necessary.

In this section it shall be assumed that 1) and 2) have been accomplished and it has been decided to attempt to compensate using a time-

[illegible]

[illegible]

[illegible]

varying amplifier in the forward transmission path as shown in figure 5-2. The variation of the gain must be chosen to provide the desired output response performance. Changing the gain of the amplifier does not alter the system dynamics and therefore the achievable results are limited. The designer may alter the response, but it is only possible to compromise between desired performance indexes. For example, it is possible to reduce the rise time, but only with a resulting increasing maximum overshoot.

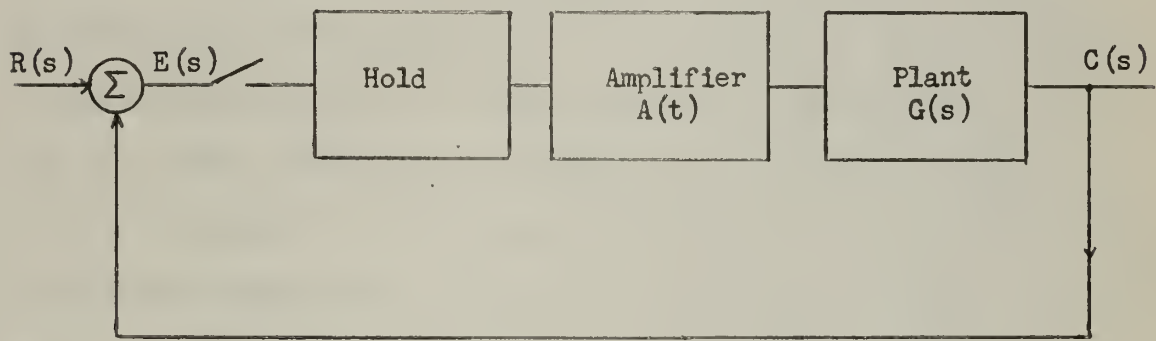


Figure 5-2. The Compensated Control System

An example will best illustrate the possibilities. Consider a system with a transfer function of

$$G(s) = \frac{1}{s(s + 1)} \tag{5-1}$$

and a zero order hold and sampling period of one second. The input signal for this example will be a unit step. Then the uncompensated output at the sampling instants is found to be:

$$\begin{array}{cc} & \begin{array}{ccc} \quad G \quad & \quad E \quad & \quad C \quad \end{array} \\ C] = & \begin{bmatrix} 0 & 1 & 0 \\ .3679 & .6321 & .3679 \\ .7675 & .000 & 1.00 \\ .9145 & -.400 & 1.40 \\ .9685 & -.400 & 1.40 \\ .9853 & -.150 & 1.15 \end{bmatrix} \end{array}$$

Therefore, the uncompensated system has a rise time of two seconds and an overshoot of greater than 40% occurring between the third and fourth sample. If the specifications call for a maximum overshoot of 30% and a rise time less than four seconds, the required response can be achieved with a time-varying gain. The designer learns, from the experience of calculating the response directly in the time domain, that the first two error samples largely determine the magnitude of the maximum overshoot. Therefore, the designer choses an amplifier with a gain of one-half at the first two samples, and a gain of one thereafter. This amplifier can be practically realized by constructing an amplifier which switches to a gain of .5 when a step input is applied, and switches to a gain of 1.0 after two seconds (n = 2). This system will give a more desirable response than simply lowering the gain to one-half for all time. This fact shall be verified later in this section. Calculating the compensated response one obtains:

$$C] = [G]\,[A]\,E] = [G]\,P] \qquad (5\text{-}3)$$

and therefore

$$C] = \begin{bmatrix} 0 \\ .3679 \\ .7675 \\ .9145 \\ .9685 \\ .9884 \end{bmatrix} \begin{bmatrix} .5 & 0 & 0 & 0 & \cdots \\ 0 & .5 & 0 & 0 & \\ 0 & 0 & 1 & 0 & . \\ 0 & 0 & 0 & 1 & . \\ . & & & & . \\ . & & & & \end{bmatrix} \begin{bmatrix} 1 \\ .81606 \\ . \\ . \\ . \end{bmatrix} \qquad (5\text{-}4)$$

$$= \begin{bmatrix} 0 \\ .3679 \\ .7675 \\ .9145 \\ .9685 \\ .9884 \\ .9957 \end{bmatrix} \begin{bmatrix} .5 \\ .4080 \\ .4662 \\ .0581 \\ -.2365 \\ -.2733 \\ \end{bmatrix} = \begin{bmatrix} 0 \\ .1839 \\ .5338 \\ .9419 \\ 1.2365 \\ 1.2733 \\ 1.1238 \end{bmatrix} \qquad (5\text{-}5)$$

$$\begin{matrix} . \\ . \end{matrix}$$

The intersample response is found as follows:

$$C(m)\Big] = \Big[G(m)\Big]\ P\Big] = \begin{bmatrix} .10653 & .5 \\ .6166 & .4080 \\ .8590 & .4662 \\ .9481 & .0581 \\ .9809 & -.2365 \\ .9930 & -.2733 \end{bmatrix} = \begin{bmatrix} .0533 \\ .3517 \\ .7307 \\ 1.1817 \\ 1.2884 \\ 1.2137 \end{bmatrix} \qquad (5\text{-}6)$$

It can be seen that the maximum overshoot is about 29% for the compensated system and the rise time is 3.15 seconds. The response of the uncompensated, compensated, and the system with a gain of one-half for all time are shown on figure 5-3. A comparison of the rise time and maximum overshoot is given in table 5-1.

System	Rise Time (seconds)	Maximum Overshoot (Percent)
Uncompensated (Gain = 1)	2.0	45
Time-varying Gain	3.15	29
Gain = .5 for all time	3.70	15

Further discussion of the application of the time domain matrix method to the design of control systems is presented in the next section.

5-3. Application of the Time Domain Matrix Method to the Design of Various Types of Control Systems.

In this section it is intended to show the great scope of possibilities of design of various types of control systems using the time-domain matrix method.

A) Single Loop Linear Sampled-Data Control System

When the preliminary analysis of a linear sampled data system reveals that the overall transient performance is inadequate, compensation techniques must be employed in order to improve the system performance. For a single loop system, the simplest and most direct step is to change the

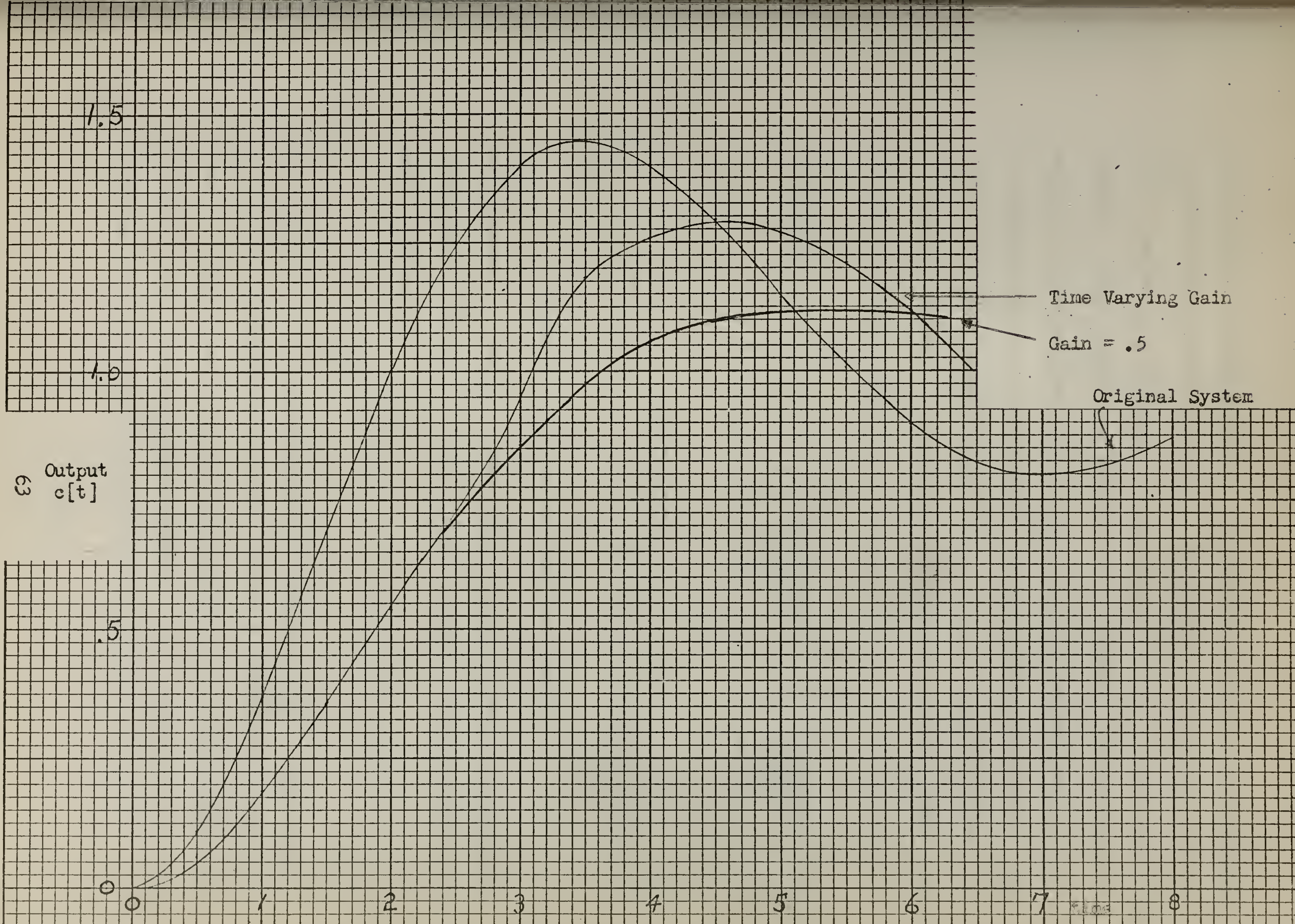


Figure 5-3. Step Response of Uncompensated and Compensated System.

system gain. Usually, however, this adjustment alone is not sufficient to satisfy the design requirements. Therefore, it is necessary to insert a compensating network in the system in order to achieve the desired response. This compensating network may be cascaded in the forward or feedback channel or inserted as a minor feedback or feedforward loop. The sampling process in the control loop complicates the choice of the location for the compensating network. For error sampled systems, it is found advantageous to operate on the sampled error by a cascade compensator as shown in figure 5-4.

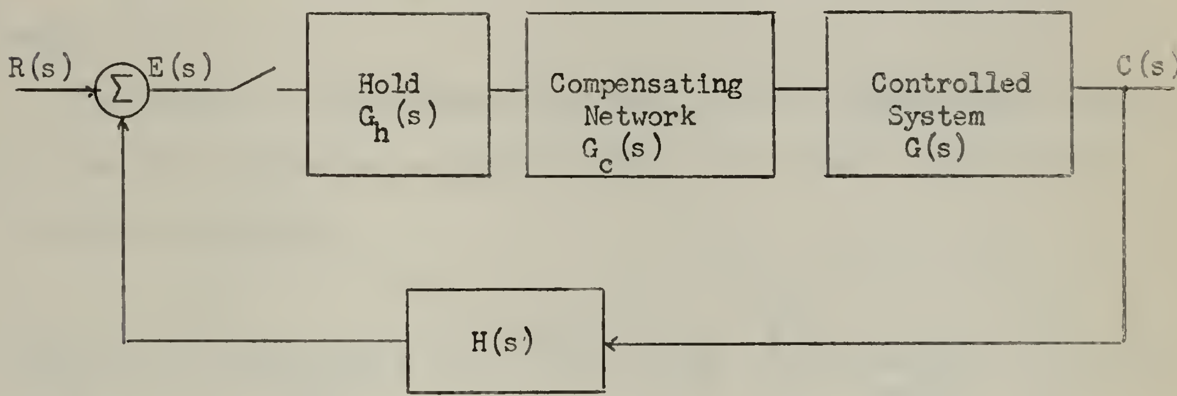


Figure 5-4. Continuous-Data Compensation

This compensator operates on the sampled and held error and yields continuous data information to the controlled system. Furthermore, for the output response one may write:

$$C\rfloor = \lfloor G_h G_c G \rfloor \; E\rfloor \tag{5-7}$$

where $\lfloor G_h G_c G \rfloor$ is the system matrix for $G_h(s)G_c(s)G(s)$. Investigation of this form of compensation reveals that it is difficult to stabilize a sampled-data control system containing higher-order integration with the use of linear continuous-data networks. The stabilization and compensation of sampled-data systems by means of continuous cascade

compensation is further complicated by the calculations necessary in order to find the matrix for the overall system for every trial compensator. This is not a disadvantage peculiar to this method, but also results when using the z-transform. The use of a continuous-data network for compensation is not to be excluded. These networks are simple R-C networks coupled with amplifiers and are easy to realize. Furthermore, for simple systems, they are sufficient and therefore probably desirable. However, for more complex systems it is often necessary to evaluate many trials in order to arrive at a reasonable compensation design.

Therefore, it often becomes desirable to use a sampled-data network as a cascade compensator as shown in figure 5-5. Then, for the output response one obtains:

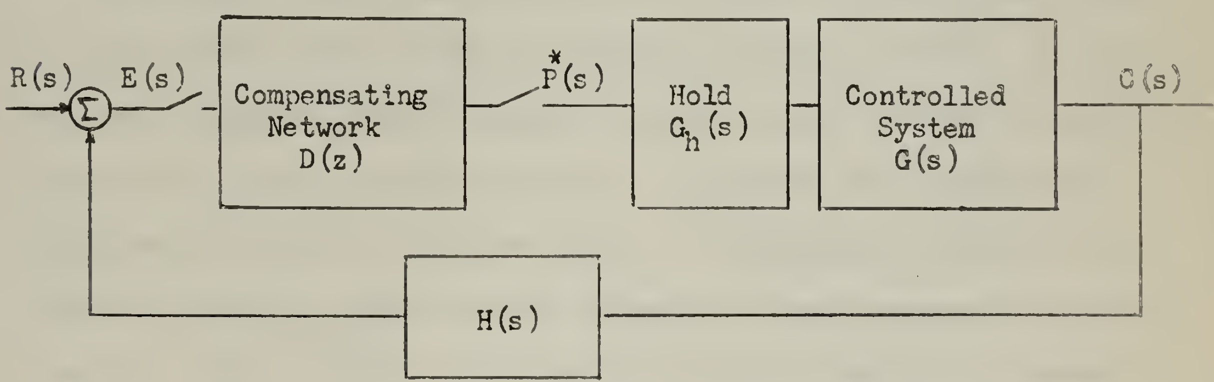


Figure 5-5. The Digital Compensator in A Sampled-Data System

$$C] = [G_hG]\,[D]\,E] = [G_hG]\;P] \tag{5-8}$$

where $P] = [D]\,E]$ (5-9)

Therefore, the sampled input to the controlled system is the output of the compensating device which is a result of a transformation of the error sequence. As was stated in Chapter 3, there is always a $[D]$ which

will transform a diverging sequence $E]$ into a stable, converging sequence $P]$. Furthermore, the design of $[D]$ is relatively easier than the design of a continuous data compensator since the $[G_hG]$ matrix need be evaluated only once. The sampled-data compensating network can be realized as a program of a digital computer or more simply as a data-processing network. It has been shown, that a digital processing unit may be realized by operational amplifiers and electronic samplers.[8] Therefore, a digital compensator may be part of a large computer program for missile control, or simply a controller for a DC motor. The realizability and the synthesis of digital compensators shall be discussed further in Chapter 7.

B) Multiloop Sampled-Data Control Systems.

The design of multiloop sampled-data control systems is more complex and difficult than design of single loop control system. The use of linear continuous data compensation networks has the same limitations as discussed in the previous paragraph. Therefore, the sampled-data compensating network is used more often. A fundamental problem in the design of multiloop systems is the selection of the location of the compensators. This problem is difficult to solve in continuous systems as it is in sampled-data systems. One advantage of the time domain matrix method is the availability, in the calculations, of the response at intermediate points throughout the multiloops, wherever a sampler exists. Knowledge of each response allows the designer to use this information to adjust the parameters of the compensator. If there is only one sampler present in the multiloop system, then this advantage is not present. If a pulsed-data network is used as the compensator, then often

feedforward control will aid in the elimination of disturbance inputs. Consider figure 5-6 which shows a feedback control system with a disturbance input and feedforward control. Then the output due to the input signal C, and that due to the disturbance signal C_u can be considered separately as:

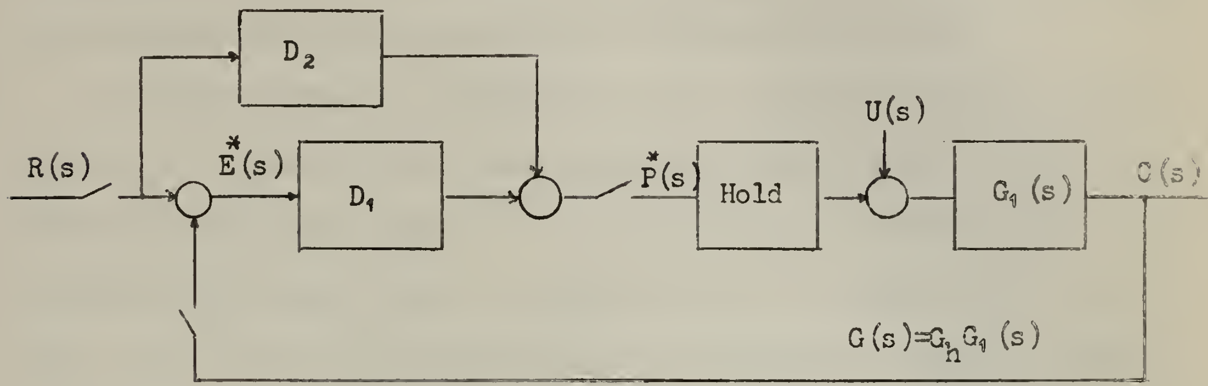


Figure 5-6. Control System with A Disturbance Input

$$C] = [G]\ P] \tag{5-10}$$

$$C_u] = \left\{[I] + [D_1]\ [G]\right\}^{-1}\ UG] \tag{5-11}$$

where $\quad P] = [D_2]\ R] + [D_1]\ R]$

Then, the output due to the reference may be written as:

$$C] = [G]\left\{[D_1] + [D_2]\right\}\left\{[I] + [D_1]\ [G]\right\}^{-1}\ R] \tag{5-12}$$

Therefore, the output due to the disturbance may be minimized by means of D_1, and the digital network D_2 used to design the system for the output performance with respect to the signal input. Multiloop systems will be considered further in the next chapter, especially the conditional feedback system.

C) Nonlinear Control Systems.

The techniques used in analyzing nonlinear systems as presented in

the previous chapter are very useful in the design of nonlinear systems. The nonlinear component is considered as a time-varying amplifier in the system where the gain is dependent upon the input magnitude. The design of linear as well as nonlinear compensators for the nonlinear systems is possible. Also it is usually an advantage to have available the signal magnitude at various intermediate points in the system.

The first method of compensation to be considered would be the insertion of a linear compensating network at a point in the system located before the nonlinear element. If this proved to be unsatisfactory, then a pulsed data network could be inserted before the nonlinearity in the loop. Also, since the nonlinearity is considered as a magnitude sensitive device and treated as a time varying amplifier, an interesting possibility for compensation would be the use of a time-varying amplifier with essentially complementary characteristics to that of the nonlinear device. If the amplifier and the nonlinear element were cascaded, then the magnitude of the signal into the nonlinear element could be kept within the linear region of the component. In the case where the nonlinear element has storage of energy, this could not be achieved by a time-varying amplifier. Therefore, in most cases a pulsed-data network is used as a compensator and designed on the basis of performance criteria and evaluation.

The design of a nonlinear system usually involves some trial and error steps in order to evaluate effective design changes. The design of a relay servo, for example, may require a number of trials in order to arrive at the proper output voltage on the relay and acceptable values of dead-zone and hysteresis. The use of a nonlinear element as a compensator also may require many design trials in order to arrive at an

acceptable design for all the desired criteria. If there is a reason for using a nonlinear device as a compensator, the calculations are as easily accomplished as for the inherently nonlinear system. Furthermore, the shape of the nonlinearity necessary to give the required output response may be determined. There are no general restrictions imposed on the design of nonlinear compensators. This is a distinct advantage over present methods of design of nonlinear compensators. The method may be applied to the solution of systems with multiple nonlinearities with the same ease.

D) Continuous Data Control System.

The design of continuous data control systems may be achieved by the use of the time domain matrix by the introduction of the approximation of the fictitious sampler and hold. The approximation and the attendant error is discussed in section 4-6. Basically, the sampling rate must be sufficient and a fictitious hold introduced in order to produce a continuous input to the controlled system. A continuous-data compensation network may be selected by any of the standard design techniques and then introduced into the feedforward channel and its compensating effects evaluated by the time domain matrix method. A single loop continuous system with fictitious sampler and hold is shown in Figure 5-7. The fictitious sampler and hold are usually inserted in the feedback loop in order to take advantage of the filtering action of the forward channel on the input signal. The use of lead and lag networks as compensating filters may be investigated directly in the time domain by this method.

If it proves desirable to investigate the use of a compensator in

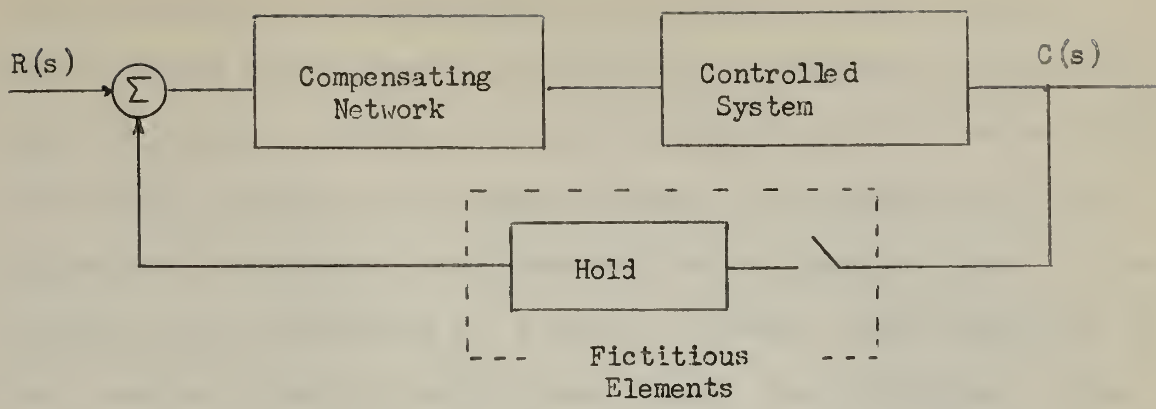


Figure 5-7. Continuous Control System With Fictitious Sampling.

the feedback channel, the introduction of a second fictitious sampler and hold will allow for compensation by a D matrix in the matrix equations. Then the discrete compensator D found to be desirable may be approximated by a continuous network in the time domain. By techniques of approximation in the time domain a network can be synthesized to yield a prescribed output for a prescribed input.[2,9] The time necessary for the calculation of the digital compensator and evaluation of a suitable continuous network may be considerably less than that for a design carried out for the continuous data system by standard s-plane or frequency response techniques.

E) Time-Varying Sampled Data Control Systems.

The design of time varying sampled-data control systems is possible with the use of the method of time domain matrices. The standard methods of analysis and design for non-time-varying systems, that is the z-transform, and frequency response methods, are not applicable. The time varying system is represented by the G(n,k) matrix and allows the general design procedures of the preceding paragraphs to be applied. Furthermore, the design of a continuous data time-varying system is

possible through the introduction of a fictitious sampler and hold. This approach offers great possibilities to the designer, in the analysis and design of systems of which the dynamics vary with some variable. An important example is the high performance jet aircraft, where the dynamics of the aircraft vary with altitude and therefore time. A compensator may be introduced and a design evaluated. This compensator may be a continuous data or sampled data network. The limitations, advantages, and purposes of each type of compensator are essentially the same as those for the non-time varying systems discussed in the preceding paragraphs.

If the variation of the time varying element is relatively large, it often is impossible to compensate with a non-time varying compensator. In this case it is necessary to use a time varying compensator such as a time varying amplifier, or perhaps a time varying network. The systems of this class are often termed adaptive systems. It is useful to consider systems with time varying dynamics (poles and zeros of the system transfer function), time varying gain, and time varying sampling rate. Research on these concepts has been carried out.[10] The area of adaptive sampled-data systems shall be considered further in the next chapter.

Finally, it is possible to use a time varying compensator in a non-time varying system. For example, it is possible to compensate a system for specified criteria by the use of a time varying amplifier. This possibility was illustrated by an example in the previous section.

CHAPTER 6

THE ANALYSIS AND DESIGN OF ADAPTIVE CONTROL SYSTEMS

6-1. Introduction

An important class of control systems are those for which parameter values of the system change as functions of some independent variable during the period of operation. An important example is the change in the system dynamics of a supersonic aircraft with altitude. Another example is that of a chemical process where a parameter may change as a function of the ambient temperature. Usually, a fixed invariable compensator will be adequate over a restricted range of operating conditions. If the system is expected to operate outside this region, then some other form of compensation is necessary. The compensator often used, varies or adapts to the changed operating conditions.

At the present time there is no widely accepted fundamental definition of an adaptive system, although several are advanced in the literature.[11,12,13] A system which changes or adapts a parameter of the controlled system to drive the actual performance towards the desired performance shall be considered adaptive. The two basic segments of an adaptive system are 1) the identification of the dynamics of the controlled system, directly or by means of a related variable; 2) the generation of an appropriate actuating signal for the controlled system. Identification of the dynamics may be accomplished, for example, by measuring the impulse response, by measuring the response to white noise, or measuring a related variable such as the output to a known input signal. The adaptive actuator may be a nonlinear, time-varying, or a digital device. For complex systems, it is common to use a digital

computer in which case the system must be treated as a sampled-data system.

6-2. The Use of A Model in an Adaptive System.

One philosophy of design of adaptive systems incorporates the use of a system model in the input channel as shown in figure 6-1 . The output of the model is then compared with the actual output and the error used to drive the system towards the desired output. Another form of a closed loop system using a model is shown in figure 6-2. This configuration has been called a conditional feedback system, the feedback of a control signal being conditional on the result of the error between the desired and actual system output. The necessary conditions for the desired performance are determined by the model and the Adaptive computer provides the actuating signal. The model may be a physical simulation or analog of the process or a mathematical abstraction manifested as a set of equations stored in a computer. The Adaptive computer is often a special purpose digital computer, but may be a nonlinear or time varying controller.

The application of time domain matrix methods to the analysis and design of adaptive systems follows the methods discussed in the previous chapters. The investigation of a sampled-data conditional feedback system shall illustrate the basic approach. Consider a basic system where $G_1 = 1$, the adaptive computer is a direct connection, and the feedback $H(s) = 2$, as shown in figure 6-3. The sampling period will be one second and the plant is a type I, 2^{nd} order system for which the system matrix is set forth in Appendix A. The goal of the design shall be to obtain a system output response for shifting system dynamics which

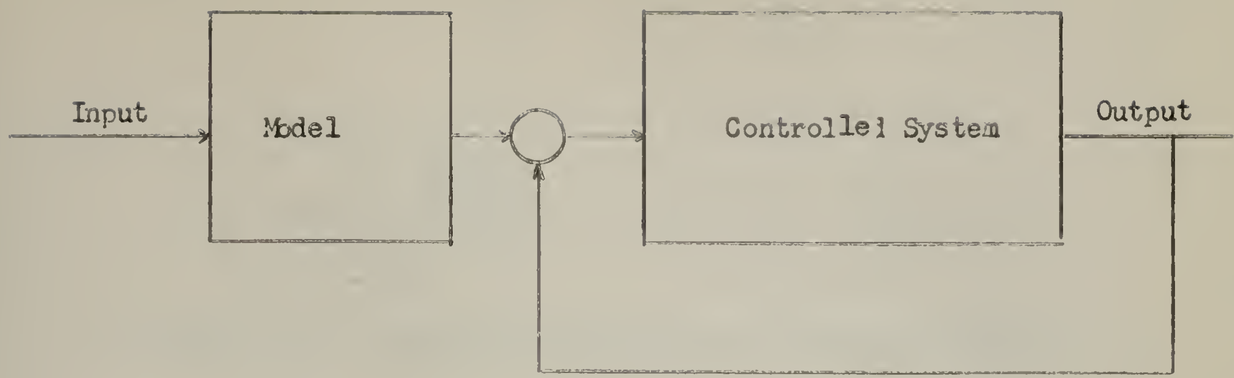


Figure 6-1. The Use of A Model to Shape the Input Signal

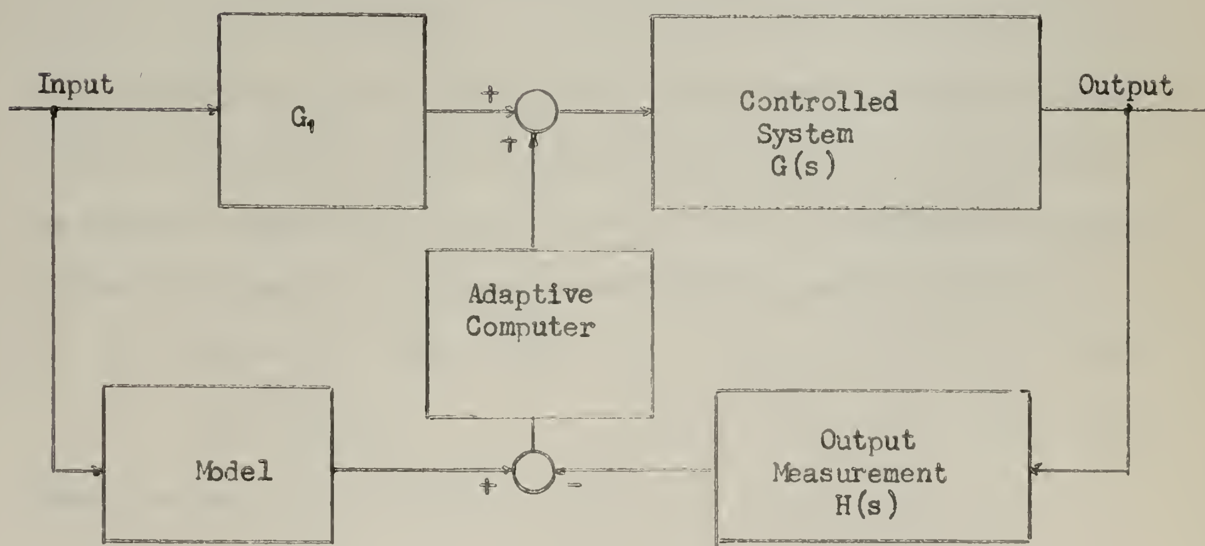


Figure 6-2. The Use of A Model in A Conditional Feedback System

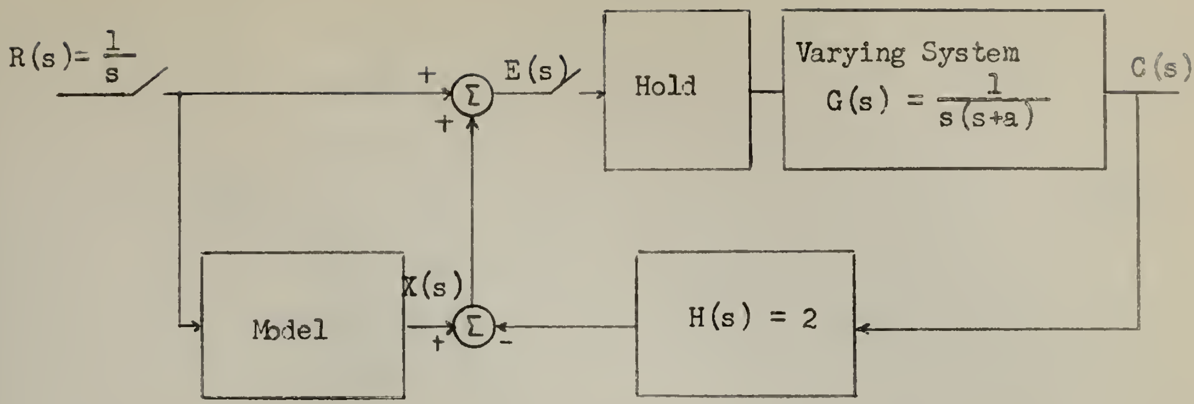


Figure 6-3. A Sampled-Data Conditional Feedback System

approaches the output response for the desired system. The desired response is present in the loop as the output of the model, $X(s)$. This desired response is compared with the actual and the difference used to drive the output towards the desired response. For example, if $a = 1$ for the unvaried system, then the output for a step input would be:

$$C] = \text{Column}\{0, .3679, 1.00, 1.40, 1.40, 1.15, .894, .798, \cdots\}$$

where $\{C\}$ is the transposed column vector. This response is shown on figure 6-4 as curve number 1. Furthermore, if the varying system manifests its variation in a shift in the pole a, over the range $a = .50$ to $a = 2.0$, the simple single loop system output would vary in the limit as shown in figure 6-4 as curve 2 and 3. For the conditional feedback system of figure 6-3, one obtains the following matrix equations:

$$E] = R] + X] - 2C] \tag{6-1}$$

$$C] = [G] E] \tag{6-2}$$

Then, one may write equation 6-1 as:

[illegible]

[illegible]

[illegible]

$$E] = \begin{bmatrix} 1 - 2C_0 \\ 1.3679 - 2C_1 \\ 2.000 - 2C_2 \\ 2.40 - 2C_3 \\ 2.40 - 2C_4 \\ 2.15 - 2C_5 \\ 1.894 - 2C_6 \\ 1.798 - 2C_7 \end{bmatrix} \tag{6-3}$$

For the system with the model in the limit at a = 2.0, one obtains:

$$C] = \begin{array}{c} G \\ \begin{bmatrix} 0 \\ .2838 \\ .4708 \\ .4961 \\ .4995 \\ .500 \\ .500 \\ \cdot \\ \cdot \end{bmatrix} \end{array} \cdot \begin{array}{c} E \\ \begin{bmatrix} 1 \\ 1.226 \\ .3626 \\ .0478 \\ -.1838 \\ -.631 \\ -.3064 \\ \cdot \\ \cdot \end{bmatrix} \end{array} = \begin{bmatrix} 0 \\ .2838 \\ .8187 \\ 1.1761 \\ 1.2919 \\ 1.2625 \\ 1.0522 \\ .843 \\ \cdot \\ \cdot \end{bmatrix} \tag{6-4}$$

This output response for the system with the model is shown as curve number 4 on figure 6-4 and the response at the limit for the pole a = .50 is shown as curve number 5. It can be seen, that the system with the model gives a response more closely approximating the desired than would the uncompensated system for either the limiting pole magnitude of a = .50 or a = 2.0. The simple conditional feedback system illustrated here, yields a definite advantage for varying parameter systems. If it was necessary to duplicate the desired response more perfectly, it would be necessary to use an adaptive controller in the feedback loop as shown in figure 6-2. In a sampled data system this would usually be a digital

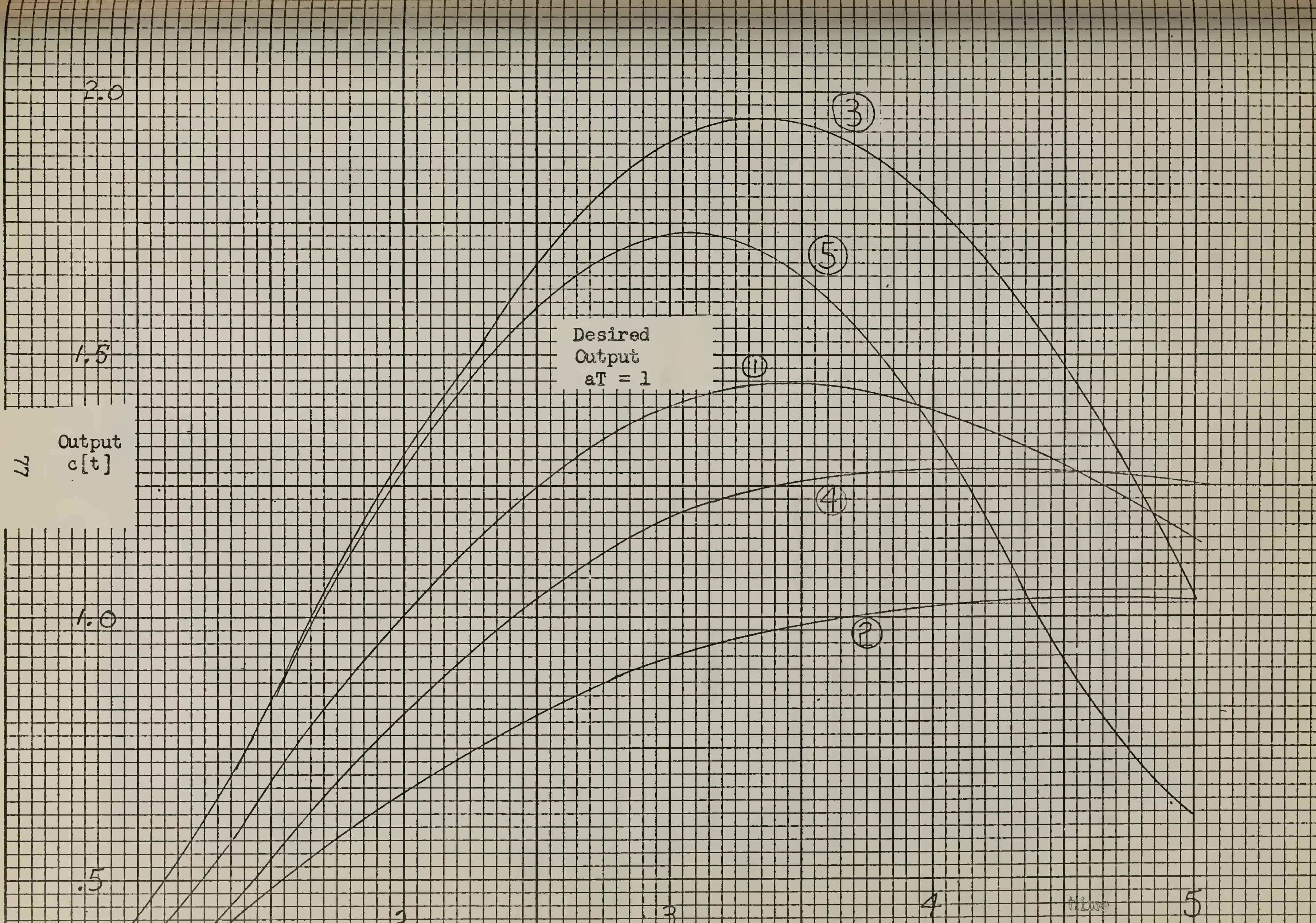


Figure 6-4. Output Response for A Conditional Feedback System

network or D block. The design of digital networks shall be discussed in the next chapter.

6-3 An Adaptive Gain System.

One possible method of adaption for a system is to set a parameter of the closed loop system dependent on a variable related to a performance index. One proposal advanced in the literature considers the variation of a compensator pole based on a measured figure of merit.[12] The calculation of the figure of merit, such as ITAE (Integrated Product of Time and Absolute Error), is usually accomplished by a special purpose computer and consideration of the signals as sampled-data usually follows.

One form of an adaptive system would use the magnitude of the sampled error to control the gain of the system. There is no signal storage involved in a simple gain adjustment system and therefore the compensation improvement is limited. However, this system illustrates the possibilities of this approach. Consider the system shown in figure 6-5. The input shall be considered a step function, and the amplifier gain a function of the error. Then, for the equations of the system one obtains:

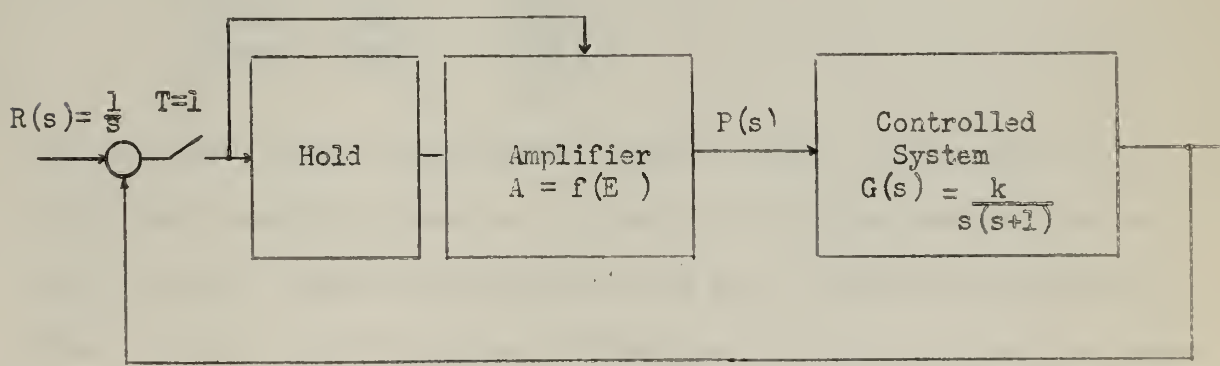


Figure 6-5. Adaptive System With A Variable Gain

$$E] = R] - C] \qquad (6\text{-}5)$$

$$[A] = \begin{bmatrix} a_0 & 0 & 0 & \cdot\ \cdot \\ 0 & a_1 & 0 & \\ 0 & 0 & a_2 & \cdot\ \cdot \\ \circ & & & \\ \circ & & & \end{bmatrix} = \text{function of the sampled error}$$

$$P] = [A]\ E] \qquad (6\text{-}6)$$

$$C] = [G]\ P] \qquad (6\text{-}7)$$

For a system where K = 1 and the amplifier gain is directly proportional to the magnitude of the error with a minimum gain of .20, one obtains:

$$P] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & & & 0 \\ 0 & .632 & 0 & 0 & 0 & & & 0 \\ 0 & 0 & .20 & 0 & 0 & & & 0 \\ & & 0 & .2277 & 0 & & & 0 \\ & & & 0 & .3283 & 0 & 0 & 0 \\ & & & & 0 & .312 & 0 & 0 \\ & & & & & 0 & .243 & 0 \\ & & & & & & 0 & .20 \\ & & & & & & & \end{bmatrix} \begin{bmatrix} 1 \\ .6321 \\ .0854 \\ -.2277 \\ -.3283 \\ -.312 \\ -.243 \\ -.1684 \\ -.1135 \end{bmatrix} = \begin{bmatrix} 1 \\ .40 \\ .0171 \\ -.0519 \\ -.1078 \\ -.0973 \\ -.059 \\ -.0337 \\ -.0227 \end{bmatrix} \qquad (6\text{-}8)$$

and

$$C] = \begin{bmatrix} 0 \\ .3679 \\ .7675 \\ .9145 \\ .9685 \\ .9884 \\ .9957 \\ .9984 \\ .9994 \end{bmatrix} \begin{bmatrix} 1 \\ .40 \\ .0171 \\ -.0519 \\ -.1078 \\ -.0973 \\ -.059 \\ -.0337 \\ -.0227 \end{bmatrix} = \begin{bmatrix} 0 \\ .367 \\ .9147 \\ 1.228 \\ 1.328 \\ 1.312 \\ 1.242 \\ 1.168 \\ 1.113 \\ 1.076 \end{bmatrix} \qquad (6\text{-}9)$$

The response for the uncompensated system is curve 1 on figure 6-6, while the compensated response is curve 2. The compensated system has less overshoot, shorter settling time and also a somewhat greater rise time. For this type of system, the designer has the choice of the minimum gain and the function of error that controls the gain of the amplifier. If the pole of the plant is expected to change within the range

$a = 1$ to $a = .5$, then the designer may evaluate the response at $a = .5$ for the compensated system. The response for the compensated system is shown on figure 6-6 as curve 3 and the uncompensated response is shown as curve 4. It can be seen that, again, the compensated system will have a smaller overshoot, less settling time and approximately the same rise time as the uncompensated system.

The change in the output response effected by this compensation scheme may not be satisfactory for many purposes. It is not the intent of this chapter to treat exhaustively adaptive systems, but rather to illustrate the application of the time domain matrix method to adaptive systems. Another powerful approach would involve relating the gain of the amplifier to the derivative of the error. The derivative of the error can be generated from the error sample magnitudes by the method of backward differences. The use of the error derivatives to control the gain or any other system parameter will yield a more desirable response than that using the error magnitude. A system which uses the error derivative to control the sampling rate is discussed in the next section.

6-4. An Adaptive Sampling Frequency System

Sampled-data control systems usually have fixed sampling frequencies which must be set high enough to give satisfactory performance for all anticipated conditions. It is useful to reduce the sampling frequency whenever possible in order to extend component life and allow time sharing of the digital components, particularly the digital computers. It is usually desired to have an efficient sampler. That is, over a given time interval, fewer samples are needed with the variable

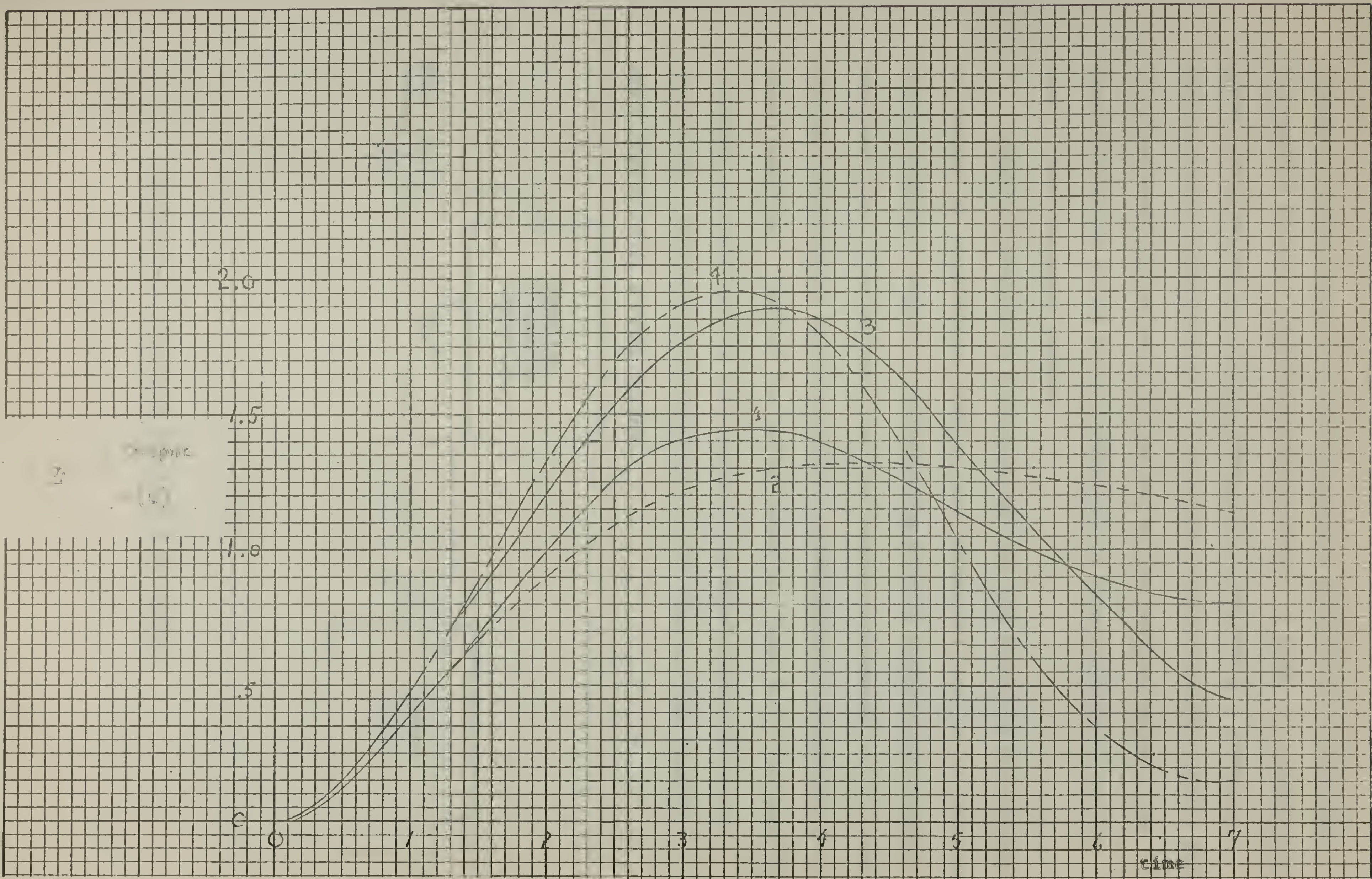


Figure 6-6. Compensated and Uncompensated System Response

frequency system than with a fixed frequency system while maintaining essentially the same response characteristics. A study of an adaptive system which varies the sampling frequency by measuring a system parameter has been accomplished.[10] It was shown experimentally, that a sampler whose sampling period is controlled by the absolute value of the first derivative of the error signal will be a more efficient sampler than a fixed frequency sampler.

Analytical methods of investigation were not available and in order to compare the experimental results with calculated results, it was necessary to develop a method of investigating sampled systems with varying sampling rates. Fortunately, the time domain matrix method with a time varying system matrix may be extended to investigation of variable frequency sampling. Consider the system shown in figure 6-7. The sampling frequency is controlled by a function of the error, and could be controlled by a function of the sampled error if this was the available error signal as in a radar system.

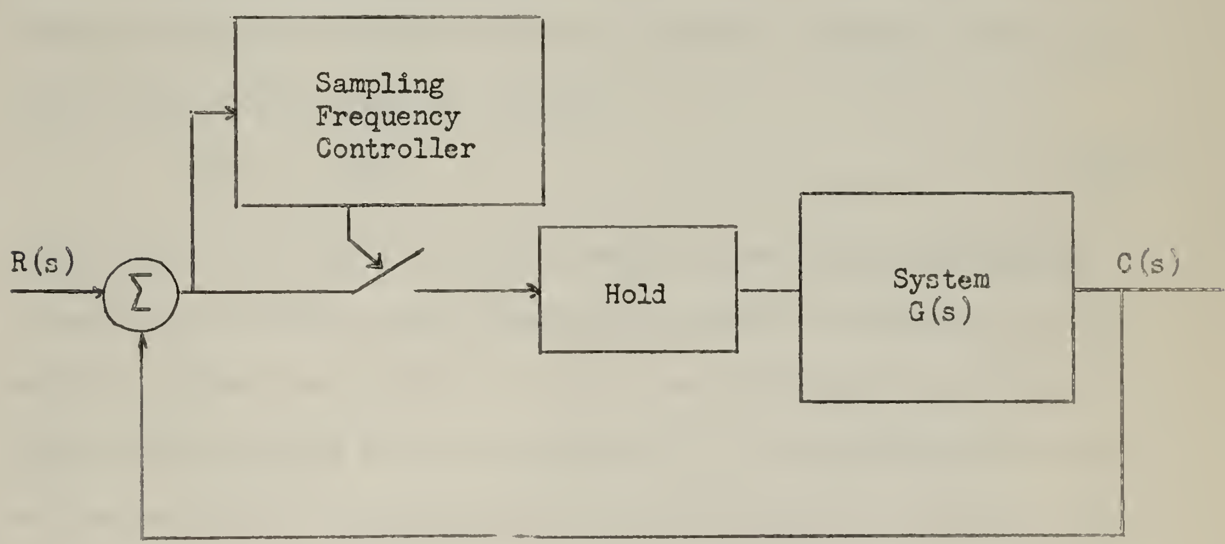


Figure 6-7. Variable Sampling Rate System

As was shown in the previous section, the derivatives of the error may be obtained from the sampled error by backward difference formulas. The output response may be written as:

$$C(k)\Big] = \Big[G(n,k)\Big]\ E(k)\Big] \tag{6-10}$$

where $\Big[G(n,k)\Big]$ is the time varying system matrix and $C(k)\Big]$ and $E(k)\Big]$ are time varying column matrices. That is, with the period of sampling changing with time, one cannot correctly write

$$C\Big] = \Big\{C_0, C_1, C_2, C_3, C_4, \ . \ . \ . \Big\} \tag{6-11}$$

where the output samples are equally spaced T seconds apart. When the sampling rate changes, samples occur at various times and in this case k equals the number of seconds elapsed from the time origin.

As an example, consider a system with two values of sampling period, T = 1 second and T = 2 seconds. Obviously, it is more efficient to have a sampling period of two seconds when the error is changing slowly, therefore the first derivative of the error may be used to switch the sampling period from one second to two seconds. Consider a step input and a system with a transfer function:

$$G(s) = \frac{1}{s(s+1)} \tag{6-12}$$

Furthermore, for illustration, the sampling rate shall be considered to switch at the first sample when the derivative of the error is zero, which is at the first overshoot peak. Considering line A of table 6-1 which is the response for this system with T = 1 second for all time, one can expect the same values of response for the first three seconds. At the output peak overshoot the sampling period switches to T = 2 seconds and one obtains for equation 6-10:

$$
C(k)\Big] = \begin{bmatrix}
0 & 0 & 0 \cdots & & \\
g(1,0) & 0 & 0 & & \\
g(2,0) & g(2,1) & 0 & & \\
g(3,0) & g(3,1) & g(3,2) & 0 & \\
g(5,0) & g(5,1) & g(5,2) & g(5,3) & 0 \\
g(7,0) & g(7,1) & g(7,2) & g(7,3) & g(7,5) \\
 & & & & \vdots
\end{bmatrix}
\begin{bmatrix} e(0) \\ e(1) \\ e(2) \\ e(3) \\ e(5) \\ e(7) \\ e(9) \end{bmatrix}
\qquad (6\text{-}13)
$$

Therefore, using the values for the g_n given in appendix A for T = 1 second up to n = 3 and then for T = 2 seconds one obtains the following, where the star indicates the T = 2 condition:

$$
C(k)\Big] = \begin{bmatrix}
0 & 0 & & & & \\
.3679 & 0 & & & & \\
.7675 & .3679 & 0 & & & \\
.9145 & .7675 & .3679 & 0 & & \\
.9884 & .9685 & .9145 & 1.1353^* & & \\
.9984 & .9957 & .9884 & 1.8830^* & 1.1353^* & \\
.9998 & .9994 & .9984 & 1.9842^* & 1.8830^* & 1.1353^*
\end{bmatrix}
\begin{bmatrix} 1.0 \\ .6321 \\ .000 \\ -.400^* \\ -.1465^* \\ +.2917^* \end{bmatrix}
\qquad (6\text{-}14)
$$

Therefore, the output response is:

$$
C(k)\Big] = \begin{bmatrix}
C(0) = 0 \\
C(1) = .3679 \\
C(2) = 1.000 \\
C(3) = 1.400 \\
C(5) = 1.1465 \\
C(7) = .7083 \\
C(9) = .8931
\end{bmatrix}
\qquad (6\text{-}15)
$$

It can be seen that there will be fewer samples necessary for this variable frequency and it is therefore more efficient. The response of the variable sampling frequency system is given in table 6-1, and it is essentially the same as the fixed frequency system with respect to overshoot.

time (seconds)	0	1	2	3	5	7	9
A Output-Fixed T	0	.368	1.00	1.40	1.15	.802	.994
B Output-VariableT	0	.368	1.00	1.40	1.15	.708	.893

Table 6-1. Response With Fixed and Variable Sampling Period.

rise time, and settling time. This subject is treated further in reference 10, where it was shown experimentally that a reduction in the number of samples of twenty-five to fifty percent may be accomplished.

6-5. The Use of the Time-Domain Matrix for Analysis and Design of Adaptive Systems.

Adaptive systems may be treated as time-varying control systems in almost all cases. Therefore, continuous or sampled-data adaptive systems may be profitably investigated by the use of time domain matrix methods. In fact, there may be no other method than can be used except solution by numerical methods ultilizing the digital computer.

One of the necessary steps in the adaptive process is the identification of the system, that is, the impulse response or the transfer function. With the use of time domain matrices it is possible to determine the pulse response when the system is at rest as outlined in chapter two. If the system is not at rest at any time, then the identification of the dynamics must be accomplished by some other means.[13]

A predictor or learning system may be designed by the use of backward and forward differences. The evaluation of a forward difference equation allows the system to predict its next few values and adjust for optimum conditions. For example, the Gregory-Newton forward interpolation formula may be programmed in a digital network for the predicted output as:

$$c_{n+1} = 3c_n - 3c_{n-1} + c_{n-2} \tag{6-16}$$

using the present value and two past values requiring two storage elements. The accuracy of prediction increases with an increase in the number of storage elements and if three past values were to be used

in the calculation, then the formula to be programmed would be:

$$c_{n+1} = 4c_n - 6c_{n-1} + 4c_{n-2} - c_{n-3} \qquad (6\text{-}17)$$

On the basis of the predicted value of the output response, the adaptive system may adjust a system parameter or add a signal to drive the actual output towards the desired output. The accuracy of this method improves with an increasing sampling frequency as would be expected. This adaptive system using equation 6-16 and 6-17 may be implemented by the use of digital logic networks or operational amplifiers and electronic samplers.[8] Equations 6-16 and 6-17 may be written in matrix form to aid in the calculation of the predicted values.

The use of a special or general purpose computer in an adaptive system can be investigated in general by the method of time domain matrices. One adaptive system uses the integrated Product of Time and Absolute Error as the figure of merit and obtains the magnitude of this figure of merit by calculations on the sampled signals. These calculations are accomplished by a programmed numerical method and can be written in matrix form. Furthermore, the method of steepest descent may also be calculated by the matrix approach. In general, the use of numerical methods in the calculation of the figure of merit and adjustment necessary in adaptive systems may readily be accomplished through the use of time domain matrix methods.

CHAPTER 7

THE DESIGN OF DISCRETE-DATA COMPENSATORS

7-1. Introduction

In the design of control systems it is desired to satisfy a set of specifications and response requirements. In many cases, this requires the introduction of a compensator in the control loop. The design process may be carried out by the use of the time domain matrix for sampled-data systems and for continuous data systems by the introduction of the fictitious sampler.

The introduction of a discrete-data compensator in the control loop will allow the designer to achieve a required output response. The compensator operates on the discrete-data input and yields a discrete-data output. This compensator may be realized by a special or general purpose digital computer, a logic network, or a circuit composed of operational amplifiers and electronic samplers.

The process to be controlled is usually a continuous process and therefore it is necessary to investigate the response of the system between sampling instants while carrying out the design. Design by z plane or root locus methods does not include this possibility and often results in unsatisfactory intersample system response. The use of time domain design techniques has two basic advantages:

1) the design is carried out directly in the time domain

2) the intersample response is accounted for in the design.

A sampled data system with a discrete data compensator is shown in figure 7-1. The $D^*(s)$ is the sampled data transfer function of the compensator.

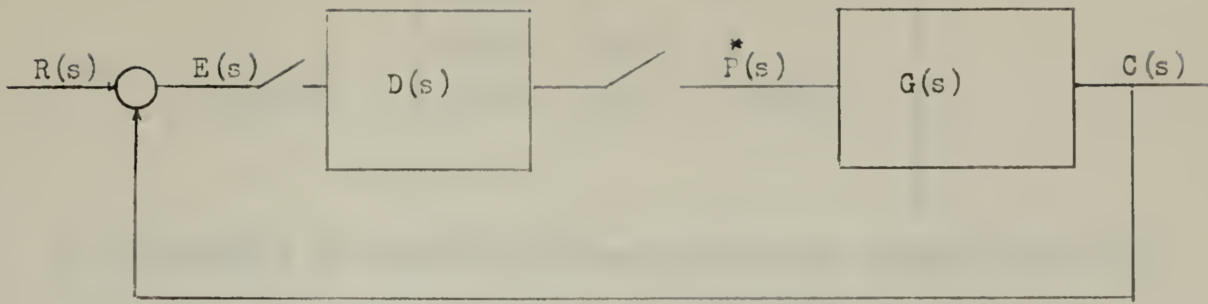


Figure 7-1. A Sampled-Data System with A Discrete-Data Compensator.

7-2. The Physical Realizability of Discrete-Data Compensators.

The digital compensator operates on the input signal and yields a transformed discrete time sequence output. The z-transform equation for this operation may be written for figure 7-1 as:

$$P(z) = D(z)E(z) \tag{7-1}$$

Then, it can be seen that this operation may be written as a matrix equation as follows:

$$P] = [D]\ E] \tag{7-2}$$

where

$$[D] = \begin{bmatrix} d_0 & 0 & 0 & 0 & \\ d_1 & d_0 & 0 & 0 & \\ d_2 & d_1 & d_0 & 0 & \cdots \\ \cdot & & & & \\ \cdot & & & & \\ d_n & & & & \end{bmatrix}$$

when the compensator is time invariant. If the digital compensator is time varying, then one may write:

$$P] = [D(n,k)]\ E]$$

where

$$
\begin{bmatrix} D(n,k) \end{bmatrix} = \begin{bmatrix} d(0,0) & 0 & 0 & \\ d(1,0) & d(1,1) & 0 & \\ d(2,0) & d(2,1) & d(2,2) & \cdots \\ \cdot & & \cdot & \\ \cdot & & \cdot & \\ \cdot & & \cdot & \end{bmatrix}
$$

It is necessary to determine what restrictions are imposed upon the D matrix by the requirement of physical realizability.

A system is said to be physically realizable, if the output signal of the system does not depend upon future information of the input signal. In the matrix equation 7-1, this is expressed as the requirement that all the elements of the D matrix above the main diagonal be zero. The truth of this statement can be seen from the following matrix equation.

$$
P] = \begin{bmatrix} d_0 & d_a & 0 & 0 & \\ d_1 & d_0 & d_a & 0 & \cdots \\ d_2 & d_1 & d_0 & d_a & \\ \cdot & & & & \\ \cdot & & & & \end{bmatrix} E] \qquad (7\text{-}3)
$$

Expanding the second row $(n = 1)$, one obtains:

$$
p_1 = d_1 e_0 + d_0 e_1 + d_a e_2 \qquad (7\text{-}4)
$$

Therefore, p_1 would depend upon the future input signal e_2 which is not available at the second sampling instant $(n = 1)$. Therefore, it is necessary for $d_a = 0$, and all elements above the main diagonal to be equal to zero.

Furthermore, for a stable compensator, it is necessary for:

$$
\lim_{n \to \infty} \left| \frac{d_{n+1}}{d_n} \right| \leq 1 \qquad (7\text{-}5)
$$

If this limit is equal to one then the D block is a steady-state oscillator.

The discrete data compensators may be realized by:[1]

1) Digital Programming

2) Delay Line Networks

3) Discrete-data RC networks

4) Analog Computer Elements.[8]

7-3. The Sensitivity Matrix of the Sampled Data System with A Discrete Compensator.

A logical quantative measure of the control property of a feedback system is the sensitivity which is defined as the relative change in the system transfer function T, divided by the relative change in the plant G.[14] In the z transform notation, this may be expressed as:

$$S = S_G^T(z) = \frac{dT(z)/T(z)}{dG(z)/G(z)} \tag{7-6}$$

where $T(z) = \dfrac{C(z)}{R(z)}$

For figure 7-1, one obtains:

$$T(z) = \frac{G(z)D(z)}{1 + G(z)D(z)} \tag{7-7}$$

and

$$S = \frac{1}{1 + G(z)D(z)} \tag{7-8}$$

This system possesses two degrees of freedom, which permits independent realization of T and S by means of the G and D components. Equations 7-7 and 7-8 may be written in matrix form as:

$$[T] = [G]\,[D]\left\{[I] + [G]\,[D]\right\}^{-1} \tag{7-9}$$

and

$$[S] = \left\{[I] + [G]\,[D]\right\}^{-1}$$

Then the sensitivity specifications fix the sensitivity matrix and therefore,

$$[G][D] = [S]^{-1} - [I] \tag{7-10}$$

The required output specifications fix the T matrix and since;

$$[T] = [G][D][S] \tag{7-11}$$

one obtains:

$$[G] = [T][S]^{-1}[D]^{-1} \tag{7-12}$$

Therefore, equations 7-12 and 7-10 may be used to obtain the required system matrices G and D. If the plant G is fixed by power and load considerations, then a new compensator, D_1, must be introduced in the control loop.

The sensitivity of an uncompensated and the compensated system is shown in figure 7-2. The uncompensated system is a type one, second order system with a zero order hold. The compensation reduces the overshoot for a step input, to half the uncompensated value while reducing the rise time by half a sampling period. The magnitude of the sensitivity response of figure 7-2 is greatest during the first three sampling periods and settles out most rapidly for the compensated system.

The use of the sensitivity matrix or response, aids the designer in understanding the effects of compensation by a discrete data network. Furthermore, the variation of the system response with a system changing with time may be studied for adaptive systems by considering the sensitivity.

7-4. Design of a Closed Loop Discrete Compensator by Means of An Open Loop Discrete Compensator.

For the single closed loop sampled-data system shown in figure 7-3, a design of a discrete compensator is usually required in order

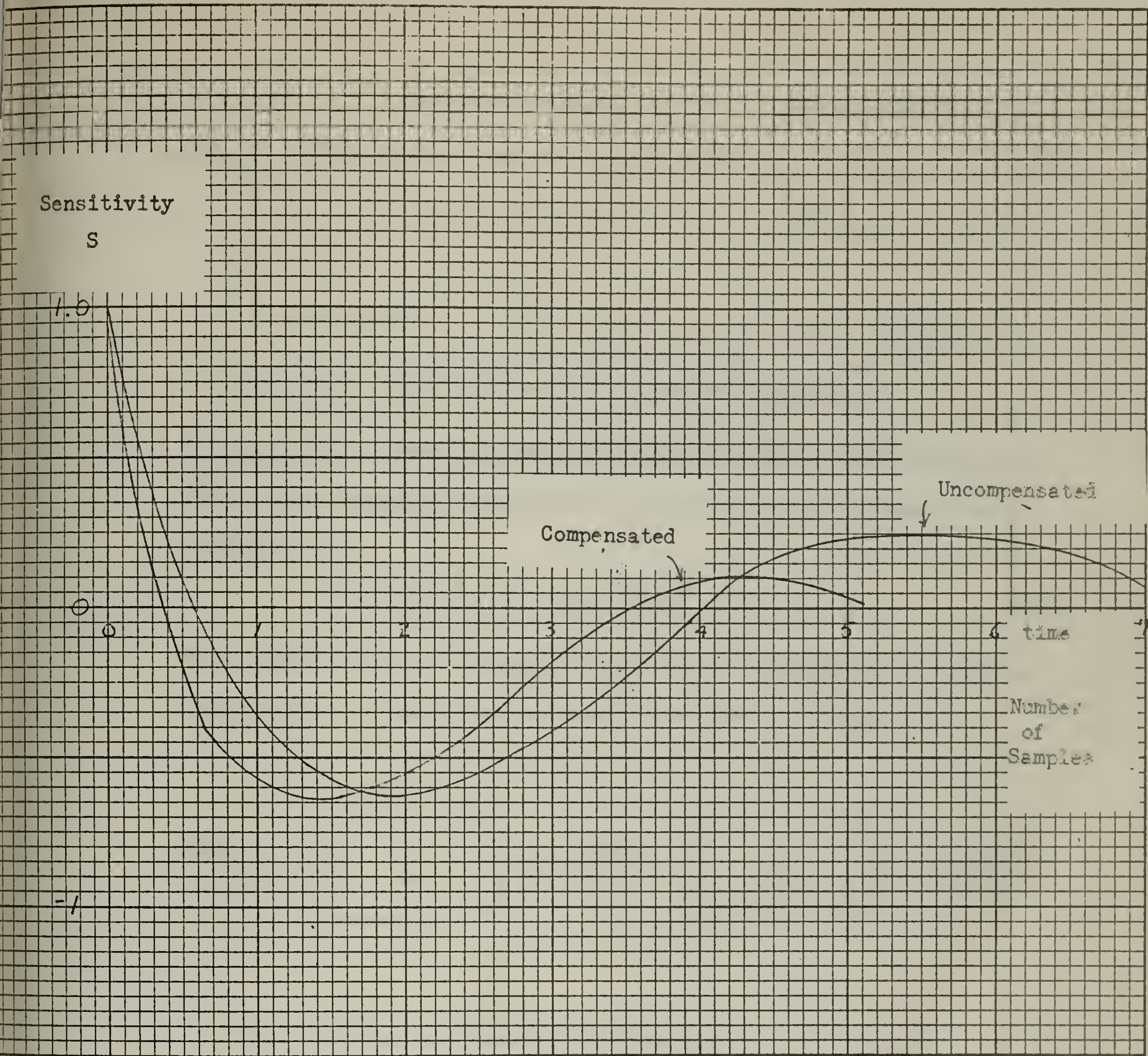
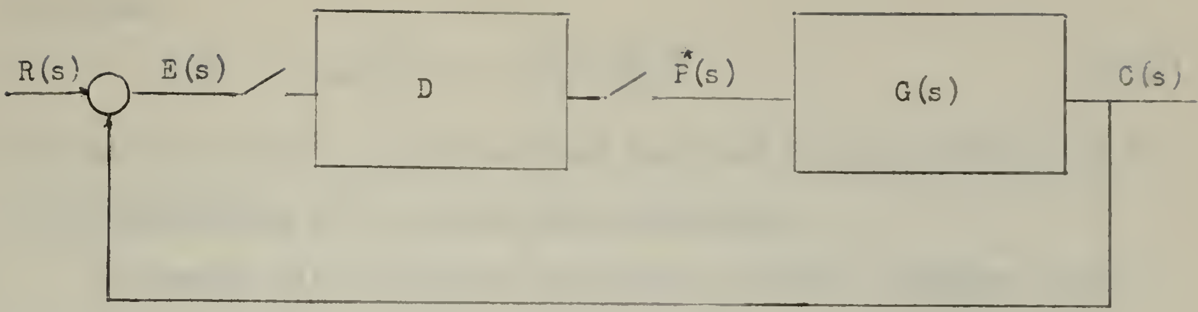


Figure 7-2. The Sensitivity of the Compensated and Uncompensated System

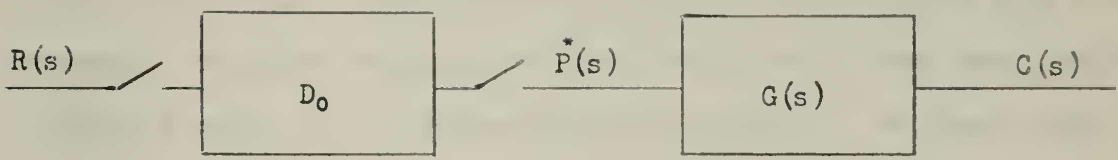


7-3. Error Sampled System with A Discrete Compensator

for the system to meet the necessary specifications. The output response at the sampling instants may be written in a time domain matrix equation as:

$$C] = [G]\,[D]\left\{[I] + [G]\,[D]\right\}^{-1} R] \qquad (7\text{-}13)$$

This complex relation involving D causes the investigator to consider the design of an open loop compensator as shown in figure 7-4.[2]



7-4. An Open Loop Discrete Compensator

Then, the matrix equation for the sampled output response may be written as:

$$C] = [G]\,[D_0]\;R] \qquad (7\text{-}14)$$

If an open loop compensator may be determined, then a closed loop compensator may be found. To find the relation between the open and closed loop compensator, equations 7-13 and 7-14 are equated and solved for D

yielding:

$$[D] = [D_0]\left\{[I] - [D_0][G]\right\}^{-1} \tag{7-15}$$

It can be seen that this method takes the plant G into consideration in the determination of the closed loop compensator.

An example will illustrate the design procedure. Consider a system with a transfer function $G(s) = \frac{1}{s(s+1)}$, a zero order hold, and a sampling period of one second. A design of a compensator shall be accomplished for the best compromise for a step and ramp input. The open loop compensator is chosen to yield a system output with a final value of one for a step input. Then, one has:

$$D_0(z) = (1 + d_1 z^{-1})\ (1 - z^{-1})\ \left(\frac{1}{1 + d_1}\right)$$

$$= \left(\frac{1}{1 + d}\right)\ (1 + (d_1 - 1)z^{-1} - d_1 z^{-2}) \tag{7-16}$$

This compensator will have a minimum complexity since the use of further delays such as $d_2 z^{-2}$, increases the complexity of the closed loop compensator. The final value in the time domain, for a step input, yields a value of unity for the system output as expected. The final value theorem is written and evaluated as follows:

$$\underset{t\to\infty}{\text{limit}}\ C] = \underset{t\to\infty}{\lim}\ [G]\ [D_0]\ R] \tag{7-17}$$

$$= \underset{t\to\infty}{\text{limit}} \begin{bmatrix} 0 & 0 & 0 \\ g_1 & 0 & \\ \cdot & \cdot & \\ \cdot & \cdot & \\ \cdot & \cdot & \\ g_n & g_{n+1} & \\ & & \end{bmatrix} \left(\frac{1}{1+d_1}\right) \begin{bmatrix} 1 & 0 & 0 \cdots \\ (d_1-1) & 1 & 0 \\ -d_1 & (d_1-1) & -1 \\ 0 & -d_1 & (d_1-1)\cdots \\ \cdot & & \cdot \\ \cdot & & \cdot \\ & & \cdot \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \\ \\ \\ \end{bmatrix}$$

$$= \underset{n\to\infty}{\text{limit}} \quad \frac{g_{n+1} + d_1 g_n}{1 + d_1} = 1$$

The output response for the sampling instants may be written for a step input as:

$$C] = \left(\frac{1}{1+d_1}\right) \begin{bmatrix} 0 & 0 & 0 \cdots \\ g_1 & 0 & 0 \\ g_2 & g_1 & 0 \cdots \\ g_3 & g_2 & g_1 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \end{bmatrix} \left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ (d_1-1) & 1 & 0 & 0 \\ -d_1 & (d_1-1) & 1 & 0 \\ 0 & -d_1 & (d_1-1) & 1 \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \end{bmatrix} \begin{matrix} 1 \\ 1 \\ 1 \\ 1 \\ \cdot \\ \cdot \\ \ \end{matrix} \right]$$

$$= \left(\frac{1}{1+d_1}\right) \begin{bmatrix} 0 & 0 & 0 \cdots \\ g_1 & 0 & 0 \\ g_2 & g_1 & 0 \\ g_3 & g_2 & g_1 \cdots \\ \cdot & & \\ \cdot & & \\ \cdot & & \end{bmatrix} \begin{bmatrix} 1 \\ d_1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \ \end{bmatrix} \tag{7-18}$$

Therefore, the output at the sampling instants is:

$$c_n = \frac{g_n + d_1 g_{n-1}}{1 + d_1} , \quad n \geq 1 \tag{7-19}$$

The response between the sampling instants may be evaluated from:

$$C(m)] = \left[G(m)\right] \left[D_0\right] R] \tag{7-20}$$

Then, for half way between the sampling instants $(m = 1/2)$, one obtains for the output response:

$$c_n(1/2) = \frac{g_{n-1/2} + d_1 \ g_{n-1.5}}{1 + d_1} , \quad n \quad 2 \tag{7-21}$$

For a ramp input, the output response may be written as:

$$C] = \left(\frac{1}{1+d_1}\right) \left[G\right] \left[D\right] R] \tag{7-22}$$

where $R] = \text{column} \left\{ 0, 1, 2, 3, 4, \cdots \right\}$

Then, one obtains:

$$c_0 = c_1 = 0$$

$$c_n = \frac{g_{n-1}}{1 + d_1} + g_{n-2} + g_{n-3} + \cdots, \quad n \quad 2 \qquad (7\text{-}23)$$

The output response is then easily evaluated for the step and ramp inputs and the results for this example are shown on figures 7-5 and 7-6. The designer would be able to choose a value of d on the basis of the specifications for a step and ramp input. In this case, as a compromise, the designer might choose $d_1 = -.60$. Then, the matrix for the closed loop compensator D may be evaluated from equation 7-15. Since the system G(s) is second order, the closed loop compensator required will be third order.

If more control over the output response is desired by the designer, then one may use a higher order D_0. For example, a higher order compensator might be:

$$D_0(z) = \left(\frac{1}{1+d_1+d_2}\right)(1 + d_1 z^{-1} + d_2 z^{-2})(1 - z^{-1}) \qquad (7\text{-}24)$$

With this form of compensator, a deadbeat response to a step input is possible. In general, this method of compensation in the open loop, gives the designer strong control over the system output response. The method may be applied to linear sampled-data and continious systems, and time varying sampled-data and continuous systems.

7-5. Design of The Closed Loop Compensator By Time Domain Evaluation.

A simple method of the design of the closed loop discrete compensator should not be overlooked; that is, the arbitrary choice of a D(z) and an evaluation of the output response by the use of time domain matrices. The evaluation of the output response between the sampling instants avoids the possibility of hidden oscillations. The choice of

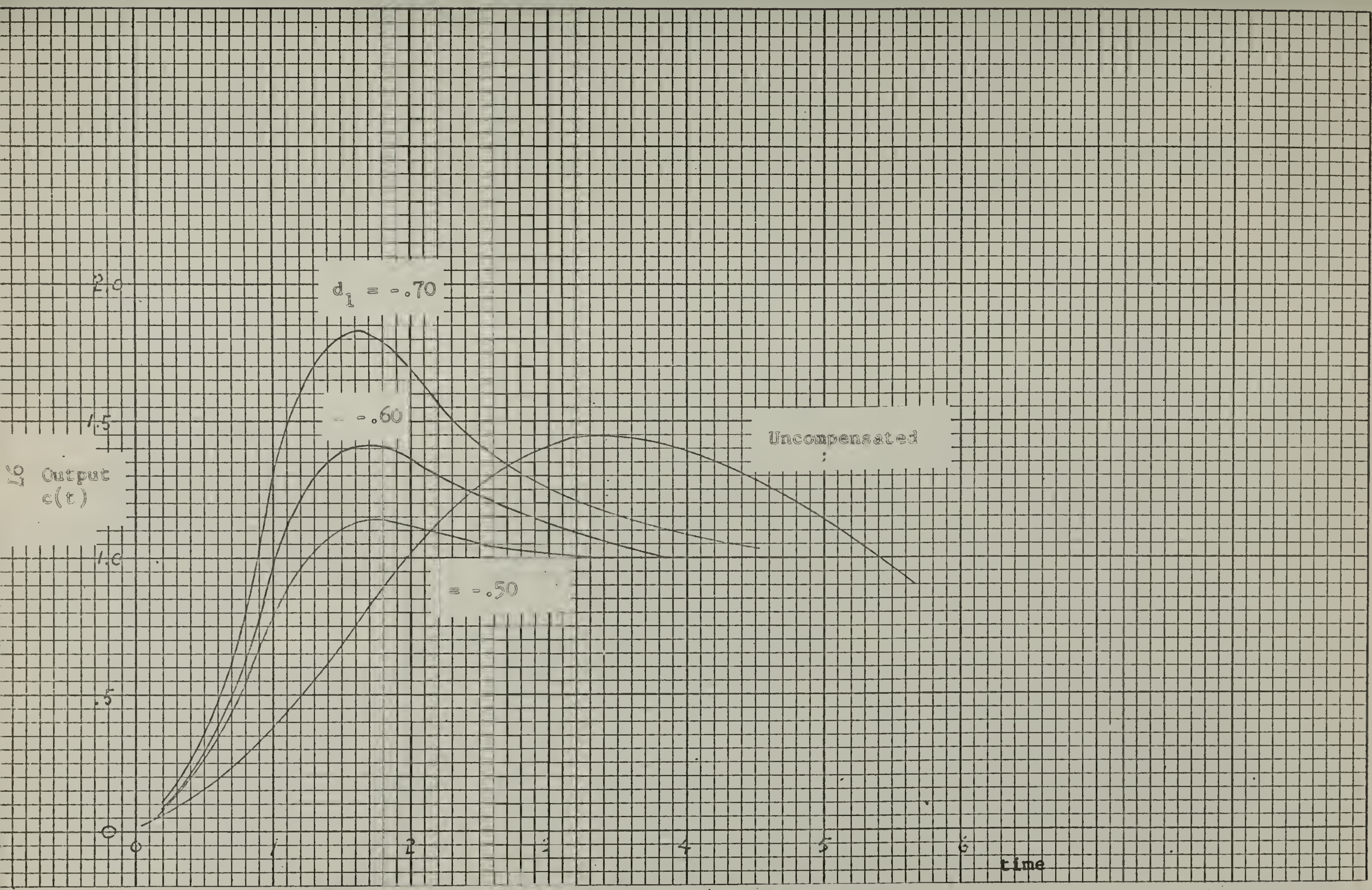


Figure 7-5. Response of the System to A Step Input

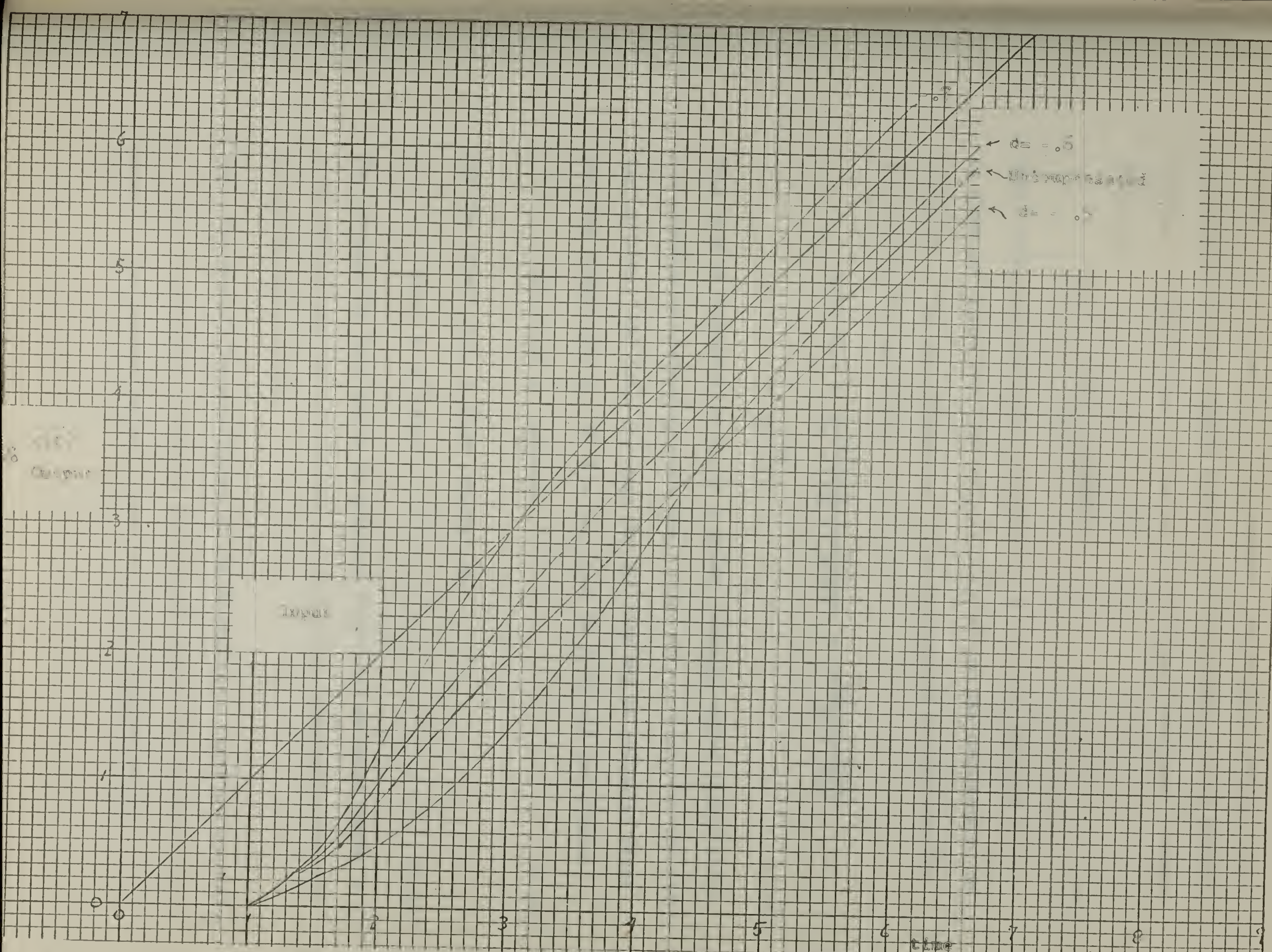


Figure 7-6. Response of the System to A Ramp Input

the D(z) may be aided by all the standard z-transform techniques, and evaluated by the time domain matrix. For the example, of the previous section, one might choose on the basis of z-plane techniques a compensator as:

$$D(z) = \frac{K(z - .36788)}{(z + d)} \qquad (7\text{-}25)$$

Then, evaluating the output response for K = 2, d = .50 and d = .72, as shown in figure 7-7, the designer would choose the compensator on the basis of the specifications. It is interesting to note that for K = 2, d = .72, one obtains

$$G(z)\,D(z) = \frac{(.3679)(z + .72)}{(z - 1)(z - .3679)} \times \frac{2(z - .3679)}{(z + .72)}$$

$$= \frac{.7358}{(z - 1)} \qquad (7\text{-}26)$$

This type of compensation in the z plane, commonly called cancellation compensation is misleading since the z-plane only accounts for the sampling instants. The total response for all time may be evaluated using the time domain matrix. If K is set equal to 2.718 and one obtains:

$$G(z)\,D(z) = \frac{1}{(z - 1)} \qquad (7\text{-}27)$$

and the closed loop, z plane response is:

$$C(z) = z^{-1} R(z) \qquad (7\text{-}28)$$

The response for this system is shown on figure 7-7 as curve number 3. It is obvious that it is necessary to consider the intersample response in most design problems.

7-6. The General Characteristics of the Closed Loop Discrete Compensator.

It is worthwhile to look at the general form of the discrete compensator D. The compensator may be written in the z-transform as;

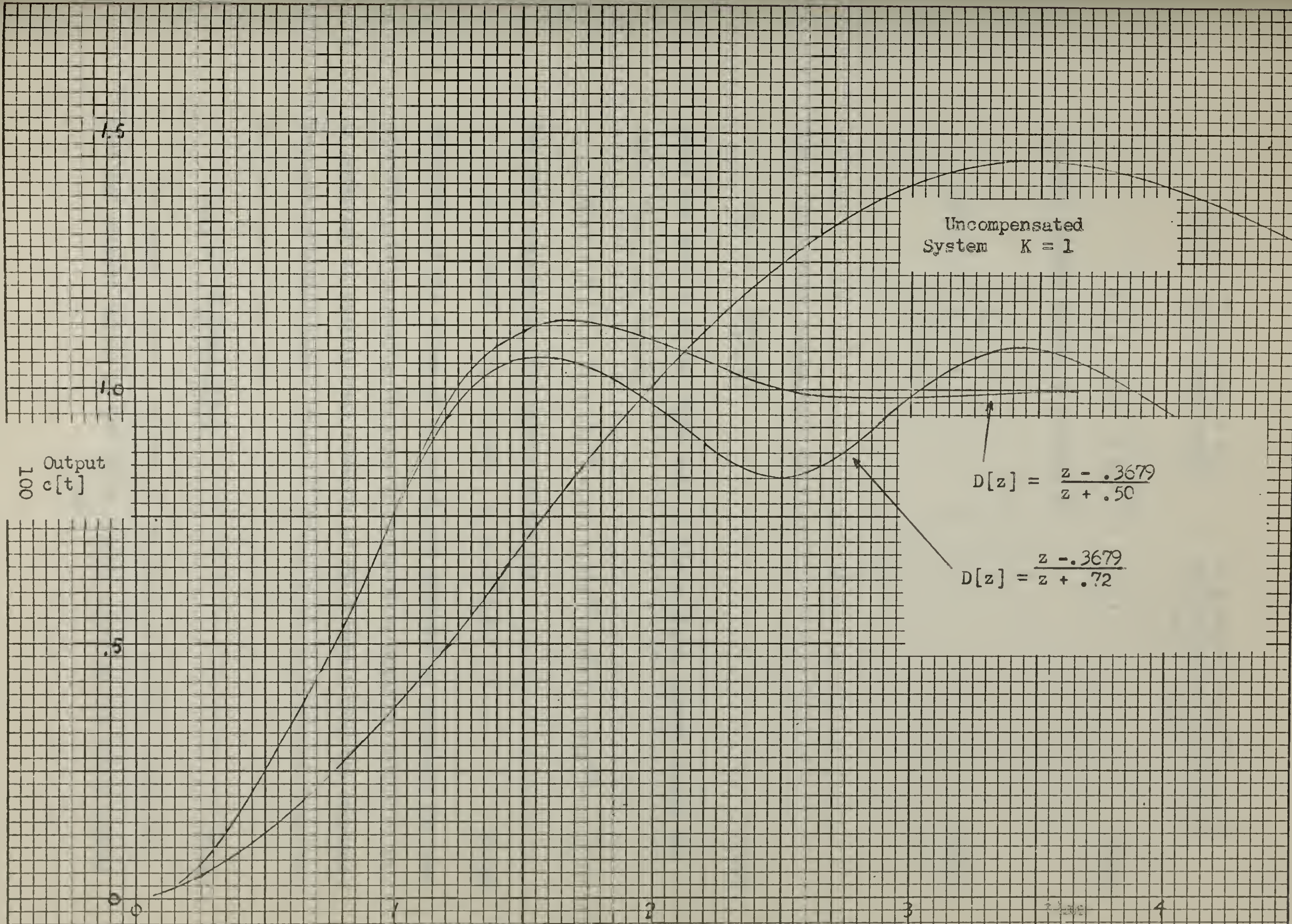


Figure [illegible]. System Response to a Unit Step Input

$$D(z) = \frac{K(z + z_1)(z + z_2)\cdots}{(z + p_1)(z + p_2)\cdots} = \frac{K\left[a_0 + a_1 z^{-1} + a_2 z^{-2} + \cdots\right]}{1 + b_1 z^{-1} + b_2 z^{-2} + \cdots}$$

$$= K\left\{1 + d_1 z^{-1} + d_2 z^{-2} + \ldots\right\} \tag{7-29}$$

The time domain matrix is written as:

$$[D] = K \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ d_1 & 1 & 0 & 0 & 0 & \\ d_2 & d_1 & 1 & 0 & 0 & \\ d_3 & d_2 & d_1 & 1 & 0 & \cdots \\ \cdot & & & & \cdot & \\ \cdot & & & & \cdot & \\ \cdot & & & & \cdot & \end{bmatrix} \tag{7-30}$$

For a first order compensator one has:

$$D(z) = \frac{K(z - a)}{(z + b)} \tag{7-31}$$

where the zero is in the right half of the z plane, almost always the case. Then, one obtains:

$$D(z) = K\left\{1 - (a + b)z^{-1} + b(a + b)z^{-2} - b^2 (a + b)z^{-3} + \ldots\right\} \tag{7-32}$$

For this compensator, the matrix values are:

$$d_1 = -(a+b)\ , \quad d_2 = b(a+b)\ , \quad d_3 = -b^2(a+b)\ , \ \ldots\ldots \tag{7-33}$$

Therefore, if the values of d_1 and d_2 are obtained as necessary for compensation, the necessary $D(z)$ may be realized as equation 7-31. If it is necessary to specify four elements of the D matrix then the compensator must be second order. That is, if d_1, d_2 d_3, d_4 are specified, then $D(z)$ must be of the form:

$$D(z) = \frac{K\ (z-a)\ (z+c)}{(z+b)\ (z+e)} \tag{7-34}$$

The sampled impulse response of the discrete network weights the past and present values of the input signal, yielding the output signal

at the sampling instants. There are three possible sampled impulse response time series for a stable discrete compensator. These possibilities are illustrated by figure 7-8 which also lists the necessary first order compensator and its z-plane pole and zero locations. An unstable discrete compensator occurs when $b > 1$ in case c, where then the alternating series is a diverging time series.

It is worthwhile to investigate further, the alternating weighting series of the compensator of case c. This discrete compensator would normally be used to stabilize the system while maintaining a desired rise and settling time. From equation 7-32 one can observe that if $(a+b)$ is greater in magnitude than one, the second term of the converging series is greater than the constant term of one. Therefore, it is actually possible to weigh past sampled information with a greater value than present information. This condition may be used, although usually to weigh past data with the greater value would have an unstabilizing influence. A picture of the sampled impulse response of the discrete compensator is useful in designing a compensator as shown in the next section.

7-7. Time Domain Design of A Discrete Compensator.

The design methods discussed in the previous sections depend either on standard techniques on the z-plane or on the choice of a trial compensator and evaluation in the time domain. These methods are powerful and are the ones used predominantly in practice. However, it would be of value to the designer if one could design the compensator directly in the time domain. Two methods of compensation design in the time domain by the use of time domain matrices have been developed by the author

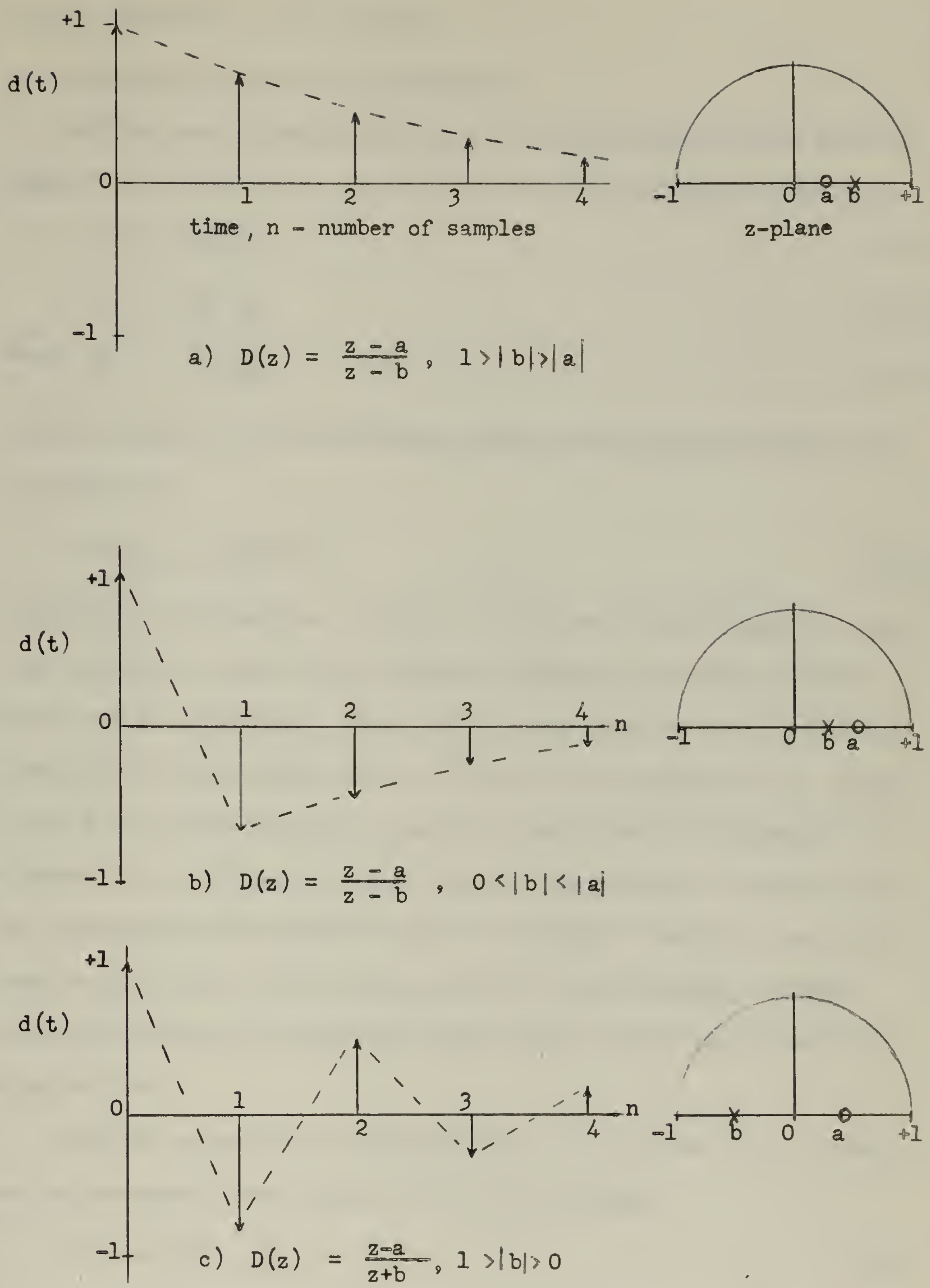


Figure 7-8. The Stable Sampled Impulse Response of The Discrete Compensator

and are presented in this section.

A. Synthesis at the Sampling Instants.

For the error compensated single loop sampled-data system shown in figure 7-3, one may write the following time domain matrix equations:

$$C] = [G]\,P] \tag{7-35}$$

$$P] = [D]\,E] \tag{7-36}$$

where

$$E] = R] - C] \tag{7-37}$$

Also, the equation for the response between the sampling instants may be written as:

$$C(m)] = [G(m)]\,P] \tag{7-38}$$

Therefore, if the designer specifies the values of the C matrix in equation 7-35 on the basis of the required response, the values of the P matrix may be determined. Then, with a known input R, and a specified C matrix, the error matrix may be evaluated from equation 7-37. Since, P and E are then known, the required D matrix may be evaluated from equation 7-35. It must be noted that the designer has no control over the intersample response when using this method. When the P matrix is found on the basis of the desired output C at the sampling instants, the output between the sampling instants C(m) is then calculated from equation 7-38.

Once the values of the sampled output are specified, the P matrix may be determined from equation 7-35, or by writing:

$$P] = [G]^{-1}\,C] \tag{7-39}$$

It is usually convenient to use equation 7-35 and avoid the inversion of the G matrix. It is possible to specify an output response at the

sampling instants which the compensated system will be unable to yield, in which case the calculated compensator, $D(z)$, will be unrealizable. For most design problems, it is sufficient to find the first few elements of the D matrix and determine a $D(z)$ compensator on this basis from equation 7-32 or 7-34.

An example will illustrate the procedure and the possibilities of compensation. Consider the system of figure 7-9 which has been considered in the previous sections. The output response of the uncompensated system is shown as curve number 1 in figure 7-10.

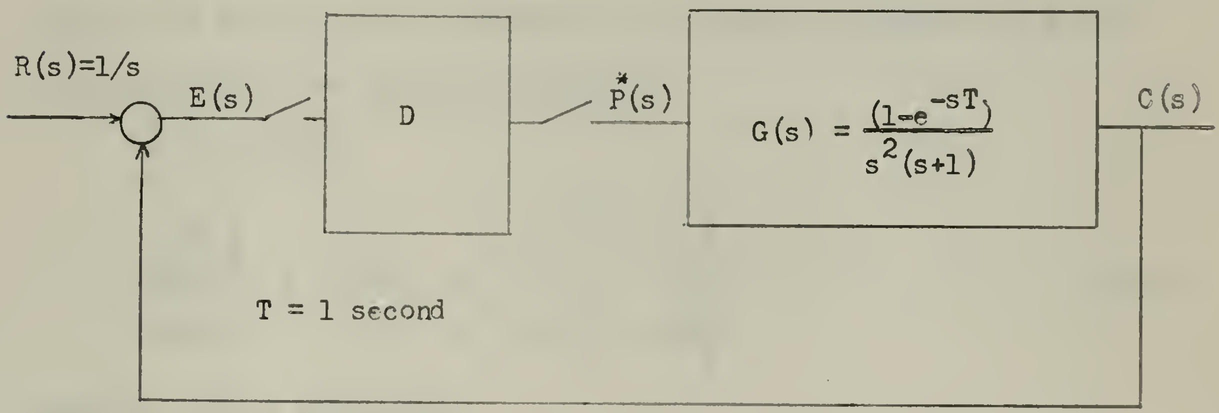


Figure 7-9. Compensated Sampled Data Control System.

It is desired to lower the rise time (for 90 % of the final value) while lowering the maximum overshoot to 10 percent. Then, on this basis, set:

$$C\big] = \text{column}\left\{0,\ .50,\ 1.0,\ 1.10,\ \cdot\ \cdot\ \cdot\right\} \tag{7-40}$$

and using equation 7-35 evaluate the necessary values in the P matrix. For the first four sampling instants, one may write:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ .3679 & 0 & 0 & 0 \\ .7675 & .3679 & 0 & 0 \\ .9145 & .7675 & .3679 & 0 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix} \tag{7-41}$$

Using this matrix, the p_n may be evaluated step by step as follows:

$$p_0 = \frac{c_1}{.3679} = \frac{.50}{.3679} = 1.359$$

$$p_1 = \frac{c_2 - g_2 p_0}{g_1} = \frac{1.0 - (.7675)(1.359)}{.3679} = -.1170 \tag{7-42}$$

$$p_2 = \frac{c_3 - g_3 p_0 - g_2 p_1}{g_1} = -.1441$$

Then, the D matrix may be evaluated on the basis of equation 7-36 and the calculated error values as follows:

$$\Big]P = [D]E]$$

$$\begin{bmatrix} 1.359 \\ -.1170 \\ -.1441 \end{bmatrix} = K \begin{bmatrix} d_0 & 0 & 0 \\ d_1 & d_0 & 0 \\ d_2 & d_1 & d_0 \end{bmatrix} \begin{bmatrix} 1.0 \\ .50 \\ .00 \end{bmatrix} \tag{7-43}$$

Therefore, the d values are;

$$Kd_0 = 1.359 \ , \quad Kd_1 = -.7965 \ , \quad Kd_2 = +\ .2542$$

and the D matrix may be written as:

$$K[D] = 1.359 \begin{bmatrix} 1 & 0 & 0 \\ -.5861 & 1 & 0 \\ +.1870 & -.5861 & 1 \end{bmatrix} \tag{7-44}$$

If a first order compensator is used to realize this compensator, equation 7-31 through 7-33 will apply. Using equation 7-33, the compensator is found from:

$$(a+b) = .5861 \ , \quad b = \frac{d_2}{(a+b)} = \frac{.1870}{.5861} = .3191$$

Therefore, $a = .2670$ and the compensator $D(z)$ may be written as:

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

$$[illegible]$$

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

[illegible]

$$D(z) = \frac{1.359(z - .2670)}{z + .3191} \quad (7\text{-}45)$$

$$= 1.359\,(1 - .5861z^{-1} + .1870z^{-2} - .0597z^{-3} + .0191z^{-4} - \ldots$$

For this compensator, the total response may be evaluated using equations 7-35 and 7-38. The response of the uncompensated and compensated systems are shown on figure 7-10, as curves number one and two respectively. The output response between sampling instants is not directly taken into account in this design approach. Therefore, if too stringent requirements are imposed on the response at the sampling instants, the intersample response may become unacceptable. The designer will usually evaluate the intersample response for a design choice by using equation 7-38. For example, if the specifications called for a maximum overshoot of five percent for the previous example, one might attempt to set c_3 equal to unity rather than 1.10 as was previously done. This design change will not effect the values of p_0 and p_1 or d_0 and d_1 that have been determined. Therefore;

$$p_0 = 1.359 \quad , \quad p_1 = -.1170$$

and
$$p_2 = \frac{c_3 - g_3 p_0 - g_2 p_1}{g_1} = -.4159 \quad (7\text{-}46)$$

Then, the gain and pole and zero of the first order compensator may be determined as follows;

$$Kd_2 = p_2 - .5Kd_1 = -.4159 - .5(-.7965) = -.01765 \quad (7\text{-}47)$$

where $K = 1.359$ and $Kd_1 = -.7965$ as previously.

Therefore, from equation 7-33;

$$(a+b) = \left| d_1 \right| = .5861$$

and

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

$$b = \frac{d_2}{(a+b)} = \frac{-.01299}{.5861} = -.02216 \tag{7-48}$$

$$a = .6083$$

Then, the compensator may be written as:

$$D(z) = \frac{1.359(z - .6083)}{(z - .0222)} \tag{7-49}$$

$$= 1.359\ (1 - .5861z^{-1} - .01299z^{-2} - .000287z^{-3} - \ldots)$$

The output halfway between the sampling instants may be calculated from equation 7-38 as:

$$C(1/2)\Big] = \Big[G(1/2)\Big]\Big[D\Big]\ E\Big] \tag{7-50}$$

$$C(1/2)\Big] = \left|\begin{array}{c|c} .1065 & 1.359 \\ .6166 & -.7965 \\ .859 & -.0177 \\ .9481 & -.0004 \end{array}\right| \quad E\Big]$$

$$= \left|\begin{array}{c|c} .1447 & 1 \\ .7531 & .5 \\ .6744 & 0 \\ .5933 & 0 \end{array}\right] = \left.\begin{array}{r} .145 \\ .8255 \\ 1.0509 \\ .9305 \end{array}\right]$$

This response would have an overshoot of five percent between the sampling instants and the design would marginally meet the specifications. This response is shown as curve number 3 on figure 7-10. If the designer chose to set c_3 equal to 1.050, then the complete output response for a compensator calculated for this new set of response values at the sampling instants is shown as curve number 4 in figure 7-10. The maximum overshoot for this compensation is 6.5 percent. A comparison of the rise time, maximum overshoot, and settling time for the uncompensated and compensated systems is presented in table 7-1. The designer would

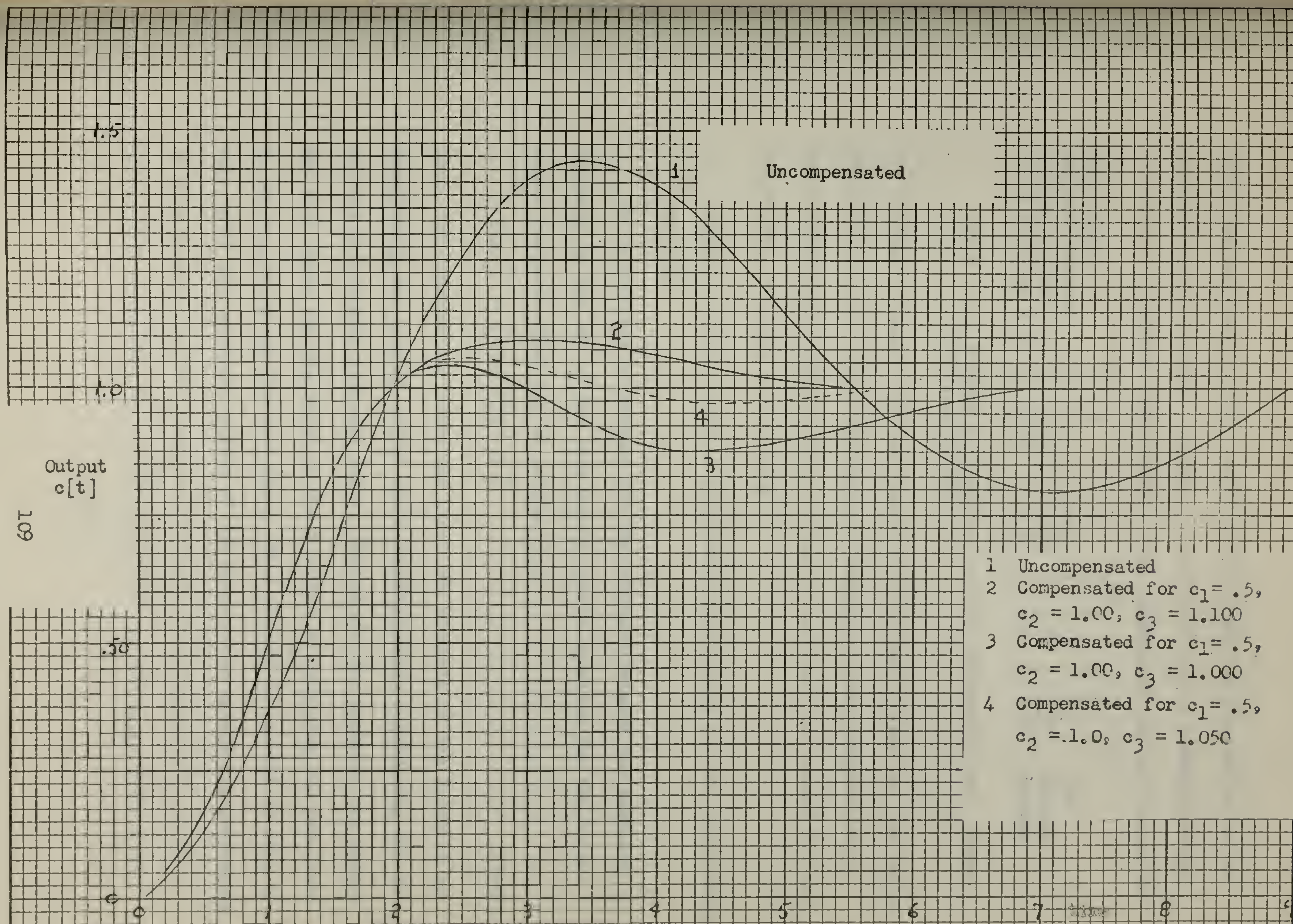


Figure 7-10. Compensated and Uncompensated System Response

chose the compensation that provided the closest approximation to the original specifications. It can be seen that it is a powerful method of design of discrete compensators. The compensator may be located in

Curve Number	1	2	3	4	
Maximum Overshoot	45	10.0	5.0	6.5	%
90 % Rise Time	1.82	1.65	1.65	1.65	seconds
2.5 % Settling Time	16.0	4.8	6.4	5.2	seconds

Table 7-1. Indices of Response For the Compensated and Uncompensated Systems.

the feedback channel or in one loop in a multiloop system and the design method loses none of its usefulness. However, if the designer has a set of stringent specifications, it may be necessary to account for the intersample response in the design.

B. Synthesis of the Discrete Compensator Accounting for the Intersample Response.

It is possible to account for the response between the sampling instants by the use of the time domain matrix equation for the intersample response which may be written as:

$$C(m)\Big] = \Big[G(m)\Big]\Big[D\Big]\;E\Big] \qquad (7\text{-}51)$$

Then, the designer may chose the first few elements of the D matrix by considering the sampled output while simultaneously considering the resultant intersample response value. Therefore, for example, the designer will chose the Kd_0 value of the D matrix by examining the resultant c_1 and $c_{.5}$ values at the same time. This process is carried on step by step until all necessary values of the compensator matrix are determined. In this manner, the next step would involve the determination of Kd_1 by a choice of c_2 and $c_{1.5}$. The choice of c_n and $c_{n-\Delta}$, where Δ is

usually equal to one-half to yield the response midway between samples, is facilitated by the determination of an expression relating the two variables and its graphical presentation. The three basic matrix equations are:

$$C\rbrack = \lbrack G\rbrack \lbrack D\rbrack E\rbrack = \lbrack D\rbrack \lbrack G\rbrack E\rbrack \tag{7-52}$$

$$C(m)\rbrack = \lbrack G(m)\rbrack \lbrack D\rbrack E\rbrack = \lbrack D\rbrack \lbrack G(m)\rbrack E\rbrack \tag{7-53}$$

$$E\rbrack = R\rbrack - C\rbrack \tag{7-54}$$

where order of matrix multiplication may be interchanged for time invariant systems. Then, for the typical system with $g_0 = 0$, and a unit step input, one obtains:

$$c_1 = Kd_0 g_1 \tag{7-55}$$

and

$$c_{.5} = Kd_0 g_{.5} \tag{7-56}$$

It is usually necessary to choose Kd_0 such that the value of the first sampled response is less than the final value in order to achieve small maximum overshoot. In the previous example, for instance, Kd_0 was set equal to 1.359. Then, one may solve for $c_{.5}$ in terms of c_1, as:

$$c_{.5} = \left(\frac{g_{.5}}{g_1} \right) c_1 \tag{7-57}$$

Equations 7-55 through 7-57 allow the designer to choose the value of Kd_0 while taking into account the intersample response. When the value of Kd_0 is chosen, the value of e_1 is determined and then since $d_0 = 1$ by the definition of equation 7-29, one obtains:

$$C\rbrack = \left| \begin{array}{c|c|c} K & 0 & 1 \\ Kd_1 & g_1 & e_1 \\ Kd_2 & g_2 & \end{array} \right\rbrack \tag{7-58}$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

[illegible]

$$[illegible]$$

and

$$C(1/2)\Big] = \begin{vmatrix} K & g_{.5} & 1 \\ Kd_1 & g_{1.5} & e_1 \end{vmatrix} \tag{7-59}$$

Therefore,

$$c_2 = Kd_1 g_1 + K(g_2 + g_1 e_1) \tag{7-60}$$

and

$$c_{1.5} = Kd_1 g_{.5} + K(g_{1.5} + g_{.5} e_1) \tag{7-61}$$

where the g_n, K, and e_1 are known. Solving for $c_{1.5}$ in terms of c_2 one obtains the general form of:

$$c_{1.5} = \left(\frac{g_{.5}}{g_1}\right) c_2 + Q \tag{7-62}$$

where Q is a constant. In general, the linear relation may be written as:

$$c_{n-\Delta} = \left(\frac{g_{1-\Delta}}{g_1}\right) c_n + Q_n \tag{7-63}$$

It is often useful to show this linear relationship in the form of a graph.

Consider the example of the previous section A, where the specified maximum overshoot was five percent. The gain constant K shall again be set equal to 1.359, so that $c_1 = .50$ as in the previous example. The design procedure in section A did not account for the intersample response and it was found that this resulted in the responses presented in figure 7-10. The design of section A to satisfy the overshoot requirement would have a response as in curve number 4 on figure 7-10. Now, in order to carry out the design while accounting for the intersample response, one obtains the equations for c_2 and $c_{1.5}$ in terms of

Kd_1 from:

$$C\Big] = \begin{vmatrix} 1.359 & 0 & 1 \\ Kd_1 & .3679 & .50 \\ Kd_2 & .7675 & e_2 \end{vmatrix} \qquad (7\text{-}64)$$

and

$$C(1/2)\Big] = \begin{vmatrix} 1.359 & .1065 & 1 \\ Kd_1 & .6166 & .5 \\ Kd_2 & .859 & e_2 \end{vmatrix} \qquad (7\text{-}65)$$

Therefore, one has:

$$c_2 = .3679Kd_1 + 1.2931 \qquad (7\text{-}66)$$

and

$$c_{1.5} = .1065Kd_1 + .9104 \qquad (7\text{-}67)$$

Then $c_{1.5}$ is related to c_2 as:

$$c_{1.5} = .2896c_2 + .5360 \qquad (7\text{-}68)$$

and this relation is shown on figure 7-11. Considering the equations 7-66 through 7-68 and the figure 7-11, the designer might choose $c_2 = .980$ and $c_{1.5} = .8234$ as a compromise. Then, one finds $Kd_1 = -.851$, and $d_1 = -.626$. The next step is to write the equation for c_3 and $c_{2.5}$ noting that the maximum overshoot for this design will occur in this interval. One obtains from equations 7-64 and 7-65:

$$c_3 = .3679Kd_2 + .9648 \qquad (7\text{-}69)$$

$$c_{2.5} = .1065Kd_2 + 1.0194 \qquad (7\text{-}70)$$

and

$$c_{2.5} = .2896c_3 + .740 \qquad (7\text{-}71)$$

Equation 7-71 is plotted on figure 7-11, and this curve shows that the overshoot may be limited to five percent. If c_3 is chosen as 1.040,

then $c_2 = 1.041$ and $Kd_2 = +.2045$. A compensator is then determined from equation 7-33 as:

$$D(z) = \frac{1.359\,(z - .386)}{(z + .240)} \qquad (7\text{-}72)$$

The total output response is then determined and is plotted in figure 7-12 as curve 5. The uncompensated system response is again curve number 1, and curve 3 is the response for the design of section A which did not account for the intersample response. The rise time and overshoot are equal for each design, while the 2.5 % settling time has been reduced 50 percent and there is no undershoot present.

As another example, consider the single loop system as shown in figure 7-9, where the sampling period is again one second, the input is a unit step, and the system has a third order transfer function. The transfer function and zero order hold are written as:

$$G(s) = \frac{4(1 - e^{s})}{s^2(s+1)(s+2)} \qquad (7\text{-}73)$$

and then the system transfer matrix is given in appendix A, table A-1 section II. The uncompensated response of this system is unstable, and the designer might try reducing the gain of the system in order to achieve stability. If the gain is reduced to two, or one-half of the unstable gain, the output response is shown in figure 7-13. This system has a maximum overshoot of 75 percent and a settling time of greater than 10 seconds. If the designer desired to maintain the same rise time while reducing the overshoot to 20 percent, it would be necessary to introduce a compensator. Using the discrete compensator design procedure, for a system gain of 4 as originally specified, one writes:

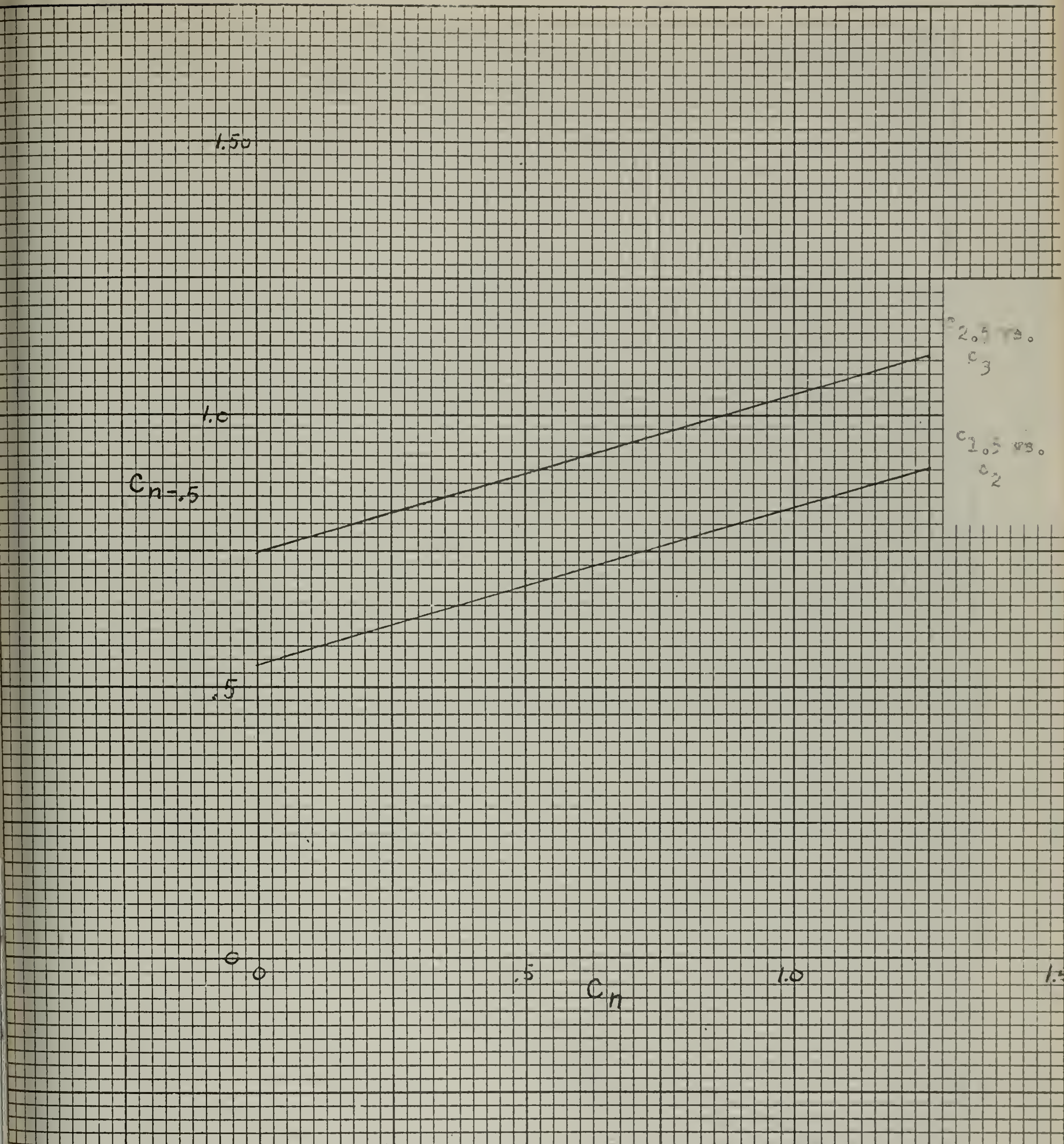


Figure 7-11. The Relation of the Intersample Response and the Sampled Response

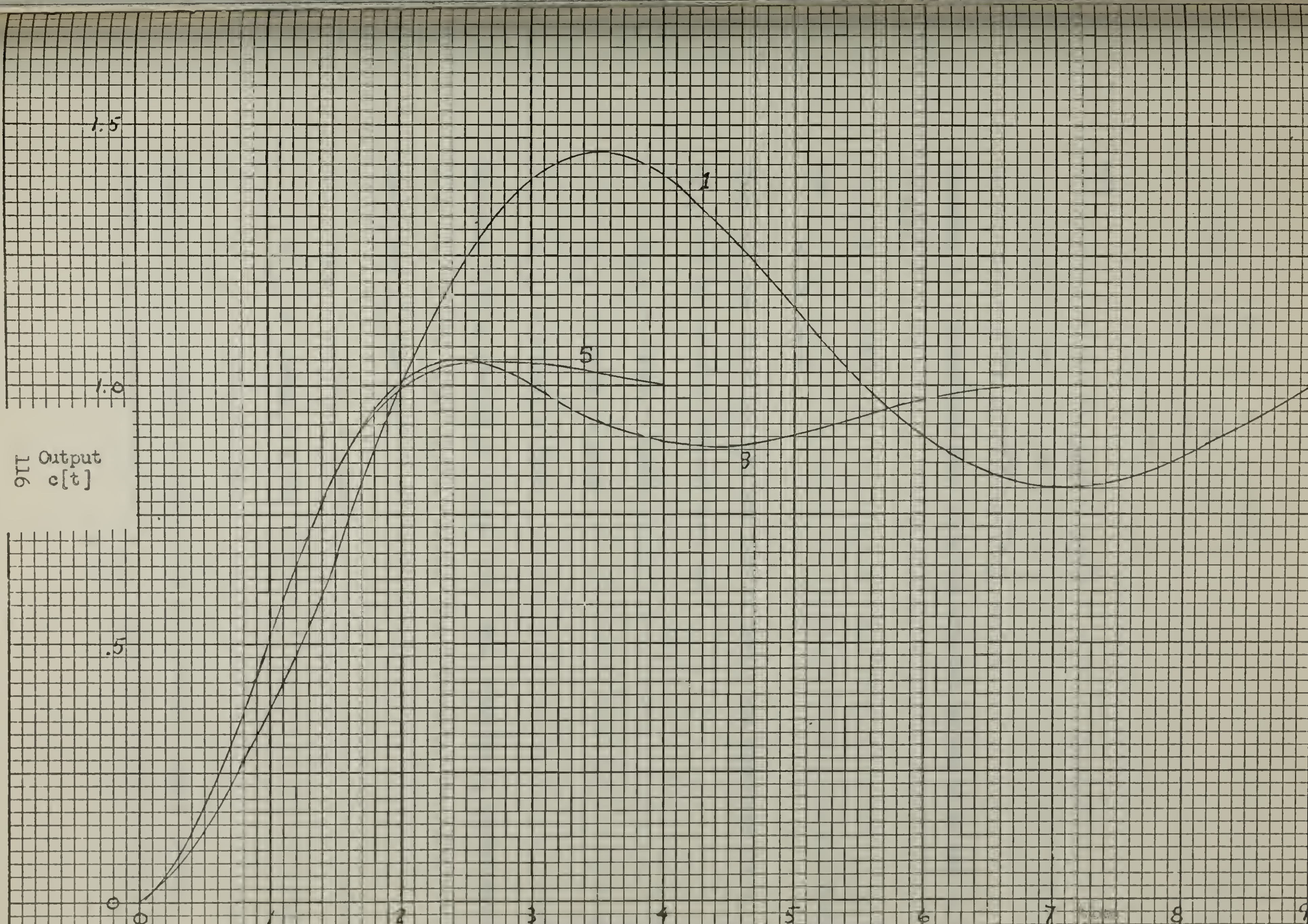


Figure [illegible]-12. Compensated and Uncompensated System Response.

$$c\Big] = \left|\begin{array}{c|c|c} K & 0 & 1.0 \\ Kd_1 & .3362 & e_1 \\ Kd_2 & 1.1869 & e_2 \\ & 1.6736 & \end{array}\right] \quad (7\text{-}74)$$

and

$$c(1/2)\Big] = \left|\begin{array}{c|c|c} K & .0582 & 1.0 \\ Kd_1 & .7845 & e_1 \\ Kd_2 & 1.4789 & e_2 \end{array}\right] \quad (7\text{-}75)$$

Choosing K = 1.0 and therefore using the original system of 4, one obtains:

$$c_1 = .3362$$

and (7-76)

$$e_1 = .6638$$

Then, one uses equations 7-74 and 7-75 to obtain the equations for c_2, $c_{1.5}$, and the relation between them, as:

$$c_{2.0} = .3362d_1 + 1.410 \quad (7\text{-}77)$$

$$c_{1.5} = .0582d_1 + .8231 \quad (7\text{-}78)$$

$$c_{1.5} = .1732c_2 + .5789 \quad (7\text{-}79)$$

Equation 7-79 is plotted on figure 7-14. Now, choosing c_2 approximately equal to 1.0, one might use $d_1 = -1.0$. Then, obtaining the relations for c_3 and $c_{2.5}$ one has:

$$c_{3.0} = .3362d_2 + 1.0266 \quad (7\text{-}80)$$

$$c_{2.5} = .0582d_2 + 1.172 \quad (7\text{-}81)$$

$$c_{2.5} = .1732c_3 + .9942 \quad (7\text{-}82)$$

Equation 7-82 is plotted on figure 7-14 and from this figure one may choose c_3 approximately equal to $c_{2.5}$ at about 17 percent overshoot.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Then it is found that d_2 may be set at .5. A first order compensator is determined for these values and one obtains:

$$D(z) = \frac{(z + .5)}{(z - .5)} = (1 - z^{-1} + .5z^{-2} - .25z^{-3} + .125z^{-4} - . .) \quad (7\text{-}83)$$

The response for the compensated system is shown in figure 7-13. The maximum overshoot is 20 percent, the rise time is reduced to 1.8 seconds and the settling time is reduced to 6.4 seconds.

These design methods are equally applicable for other system input signals. This will be shown by designing a compensator for the system considered in this section subjected to a unit ramp signal input. Again the sampling period is one second and the system and hold transfer function is:

$$G(s) = \frac{(1 - e^{-s})}{s^2(s + 1)} \quad (7\text{-}84)$$

The input signal and the uncompensated system response is shown on figure 7-14. The first two samples have a zero value, and one may write for the output response:

$$c] = \left|\begin{array}{c|c|c} & 0 & 0 \\ K & .3679 & 1.0 \\ Kd_1 & .7675 & e_2 \\ Kd_2 & .9145 & e_3 \end{array}\right] \quad (7\text{-}85)$$

and

$$c(1/2)] = \left|\begin{array}{c|c|c} K & .1065 & 0 \\ Kd_1 & .6166 & 1.0 \\ Kd_2 & .859 & e_2 \end{array}\right] \quad (7\text{-}86)$$

Then, choosing K = 2 so that c_2 = .7358, one may write for the next sample interval:

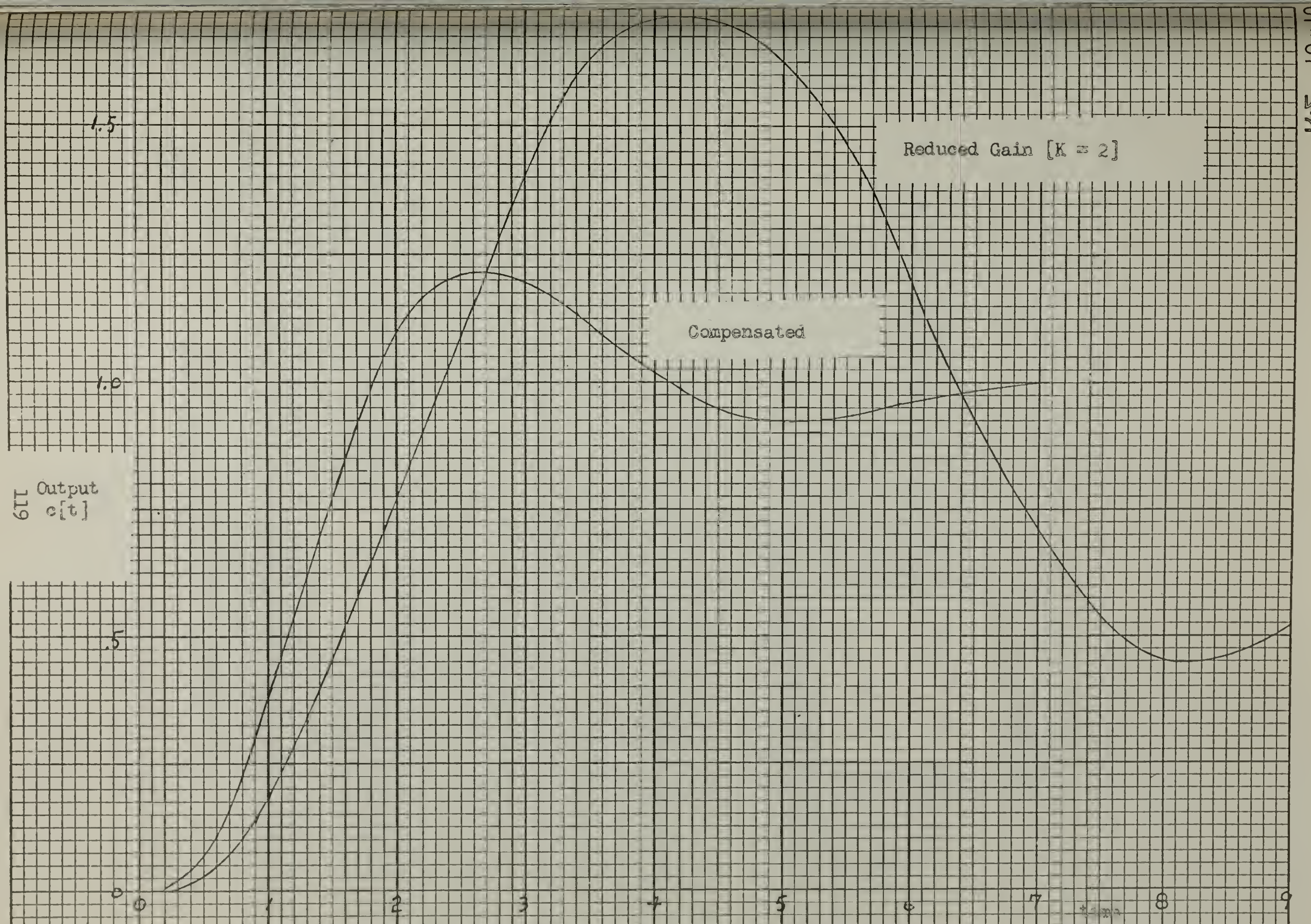


Figure 7-13. Compensated and Reduced Gain System Response

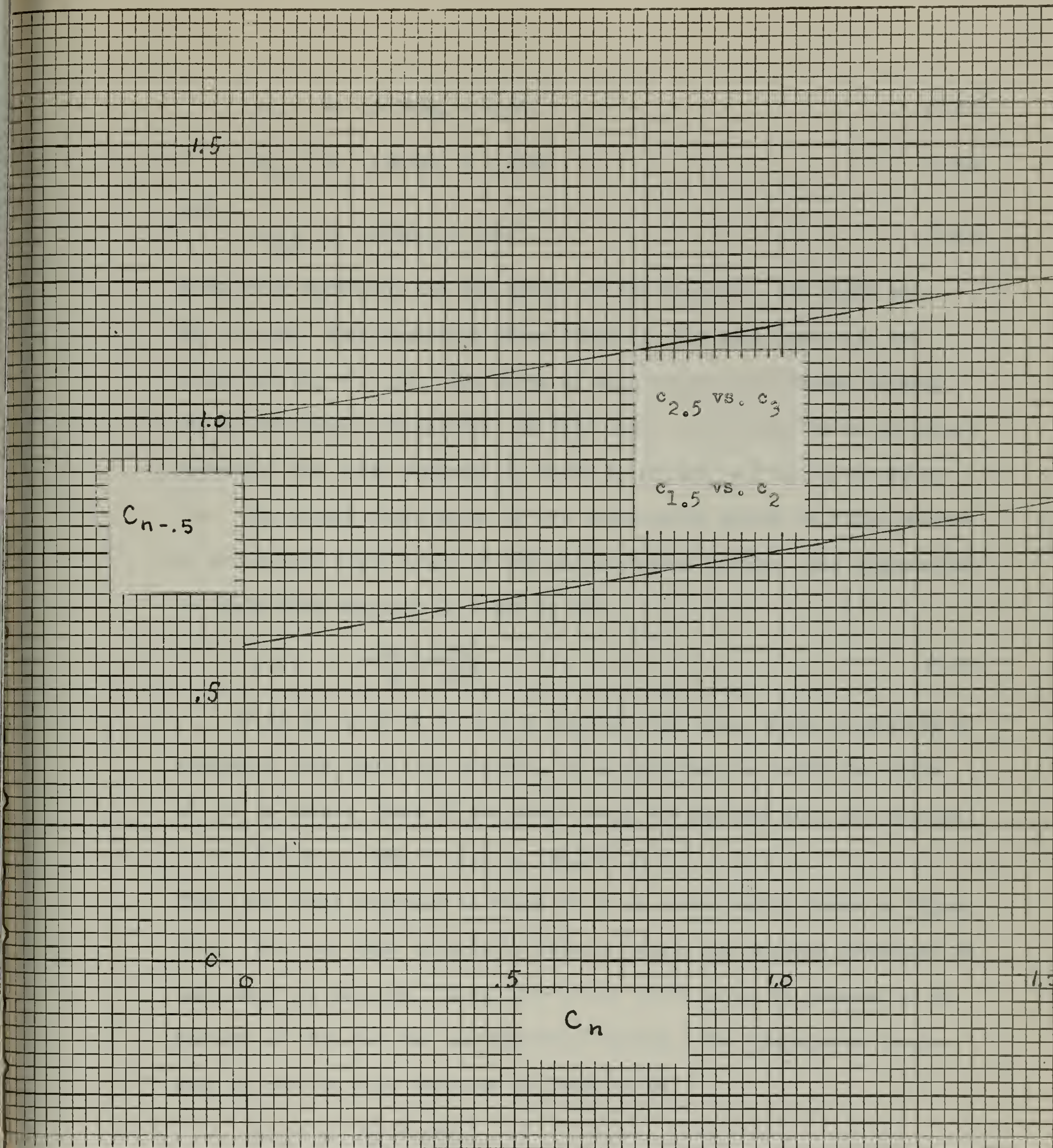


Figure 7-14. The Relation of Intersample Response and The Sampled Response

$$c_3 = .3679Kd_1 + 2.465 \tag{7-87}$$

$$c_{2.5} = .1065Kd_1 + 1.5026 \tag{7-88}$$

and

$$c_{2.5} = .2896c_3 + .7977 \tag{7-89}$$

From these equations and the nature of the response, one might set $Kd_1 = -.50$. Then repeating these steps for the next interval, one might let $Kd_2 = +.20$. The value of Kd_2 chosen would depend largely on the overshoot limitation. In this case, there would be no overshoot present. The total response is shown in figure 7-14 as curve number two. The steady state error of the compensated system is sixty percent of the steady state error of the uncompensated system. The compensator to yield this response is found to be:

$$D(z) = \frac{2(z + .15)}{(z + .40)} \tag{7-90}$$

If it is required to have a shorter rise time, while permitting an overshoot less than five percent, it is necessary to raise the value of K to three. Then if the usual steps are carried out, one obtains:

$$K = 3 \quad , \quad Kd_1 = -1.0 \quad , \quad Kd_2 = +.5$$

Then, the total response is found to be as curve number three in figure 7-15. The rise time is greatly reduced, while the maximum overshoot is less than two percent. Also, the steady state error is reduced to 42 percent of that for the uncompensated system. The compensator necessary to realize this response is found to be:

$$D(z) = \frac{3(z + .167)}{(z + .50)}$$

If the input signal was expected to alternate between a ramp and step signal, it would be necessary to design a compromise compensator with

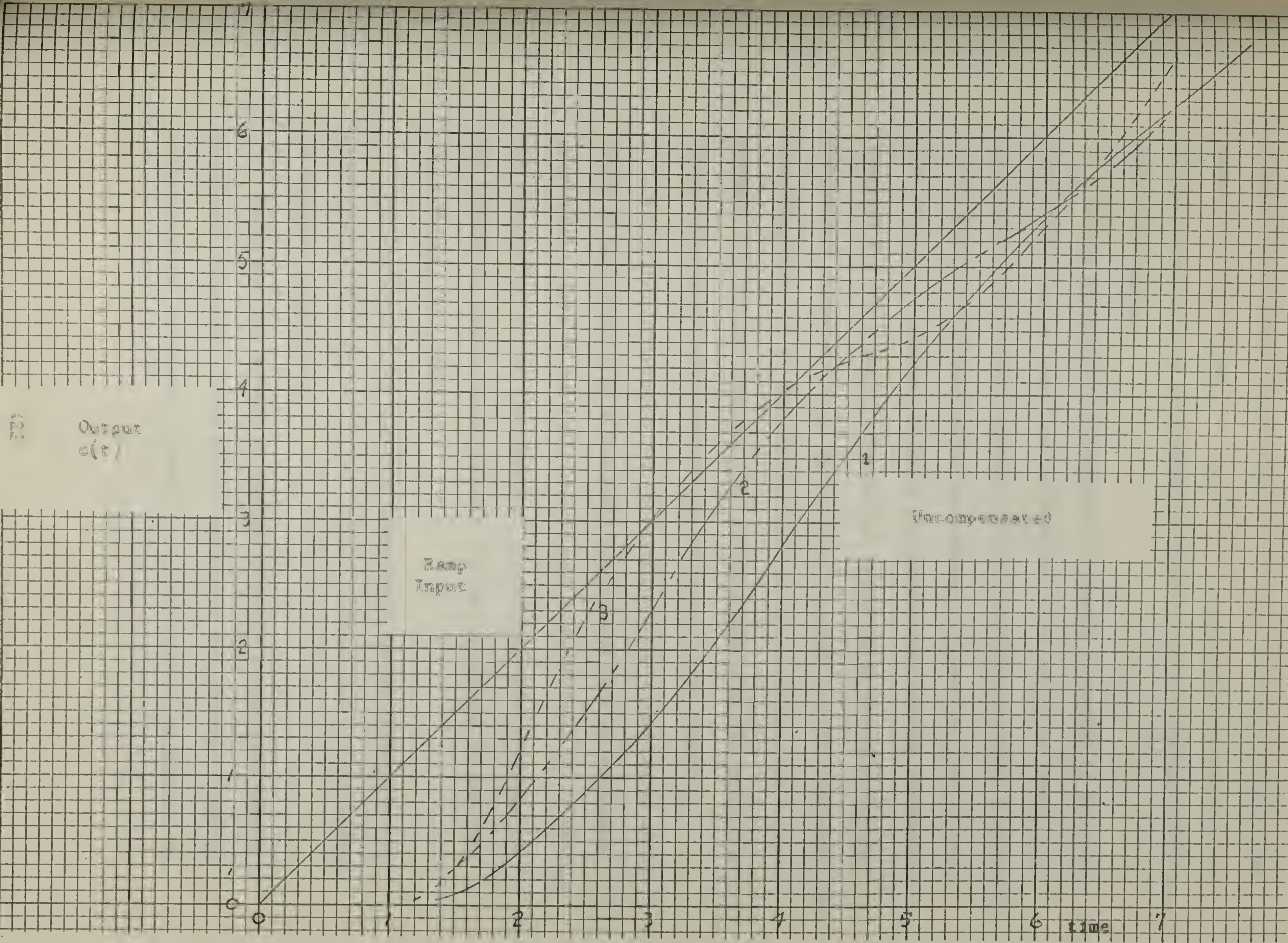


Figure 7-15. System Response to A Ramp Input

respect to the design requirements. This is necessary, since a compensator to be optimum with respect to a given set of conditions, cannot be expected to be optimum for a changed set of operating conditions and requirements.

It can be seen that the design approaches presented in this section allows the designer a wide latitude of the choice of specifications, and affords direct control over the output response. Furthermore, the time domain approach gives the designer a direct picture of the time response, therefore avoiding inaccurate and unwieldy correlation theorems for the z plane. The design method which accounts for the intersample response allows the designer to wield control over the total time response. It is obvious that this design procedure may be programmed in a digital computer, the designer supplying the specifications and system transfer matrix as the input to the computer.

A discrete compensator may be determined for a continuous system which is approximated by a fictitous sampler and hold as discussed previously. Then the calculated D(z) may be synthesized by a continuous compensator in the time domain.[9] Therefore, this technique is not limited to sampled-data systems, but may be applied to any type of continuous data system.

The location of the discrete compensator is not limited to the error channel and may be located in any arbitrary loop in the system as feedback compensation.

Therefore, two methods of design, directly in the time domain, have been presented. These design procedures are entirely flexible in application, accurate in calculation, and rapid in solution. The

design in the time domain gives the designer a complete insight into the total time response of the system. No other existing design procedure can be carried out directly in the time domain, nor can any existing design procedure provide the flexibility and accuracy of the time domain matrix method.

CHAPTER 8

CONCLUSIONS

8-1. Summary of Results

The aim of this dissertation was to present a new method of engineering analysis and design for complex control systems. This method is the time domain infinite matrix method. The formulation of the infinite matrix follows from the convolution summation of sampled data systems. The mathematical basis of the time domain matrix formulation is presented in a discussion of the applicable concepts of infinite matrices and sequence spaces. This method of analysis and design is applicable to both continuous data and sampled data systems. For continuous systems it is necessary to introduce a fictitious sampler and hold of sufficient sampling rate to effect an accurate approximation.

It is possible to analyze and design linear, nonlinear, and time varying systems of the continuous or sampled data class. Sampled data, time varying systems may not be investigated by any other existing method. Furthermore the investigation of nonlinear systems is greatly simplified by the time domain approach. Multiloop systems may be treated with ease and the signals at intermediate points throughout the loops are readily available. Also, systems with multiple nonlinearities may be investigated, for which there is not a presently available method of analysis and design.

Two methods of design of a discrete compensator for a sampled data system are presented. These methods are accomplished directly in the time domain and allow for a compromise of specifications in

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

the time domain. Also the response between sampling instants is accounted for in one of the two design procedures.

The time domain matrix method may be readily programmed on a digital computer and therefore provides a rapid analysis and design technique.

8-2. Further Conclusions

It is an important advantage that the design and analysis of systems using the time domain matrix method takes place directly in the time domain and not in any transformed complex variable domain. This advantage aids the designer in understanding and controlling the time response of the system under study.

The availability of the intersample response is also an advantage to the designer, so that a design may be accomplished which accounts for the total time response and not that solely of the sampling instants. Therefore, the time domain matrix method has strong advantages over the commonly used z-transform on both of these points.

Since adaptive systems may be treated as time varying systems, the time domain matrix method may be applied in general. Furthermore, a learning control system with a predictor may be realized and investigated by the use of time domain matrices. The investigation of conditional control systems incorporating a model was also accomplished.

For nonlinear systems, the time domain matrices provide a useful analysis and design technique. The characteristic of a deliberately introduced nonlinear compensator may be determined and investigated; that is, it may be specified by the design approach. Furthermore,

the derivatives of the error or any other desired signal may be evaluated readily by matrix methods and used to determine the phase space response to aid the designer in the investigation. The use of a fictitous sampler and hold in a continuous data system permits these techniques of investigation and design to be applied.

The design of the discrete compensator by means of the time domain matrix method provides the designer direct control of the time response of the control system. Therefore, the designer may specify the required system response and determine directly, the necessary discrete compensator. The realizability, stability, and sensitivity of the discrete compensator was investigated, to provide the designer with an insight into the general characteristics of the discrete compensator.

In comparison with other methods of investigation and design, the time domain matrix method is more accurate due to its numerical formulation. Furthermore, the accuracy and ease of calculation for the matrix method remains the same regardless of the order of the system transfer function, while the accuracy decreases and the difficulty of calculation increases for other methods, such as the z-transform. The time necessary for calculation of the time response of a control system is considerably less than that of standard complex frequency techniques. Actually, the time response may be determined as rapidly as the determination of instability or stability by standard techniques such as Routh's method.

The transfer matrix of the system may be easily determined by experimental pulse techniques to an accuracy of at least three

percent. This method of system characterization may be used for the identification process in adaptive systems.

Therefore, it has been shown that the time domain matrix method is applicable to a wide range of control system problems. Techniques have been developed which permit the solution of problems which are not solveable by any other method. Examples of such problems are the analysis and design of systems containing more than one nonlinear element, and the analysis and design of time varying systems, for continuous or sampled data.

A method has been developed which permits the design of a compensator for continuous or sampled data systems, without the use of trial and error methods. It also has been determined, that the difficulty of the application of this method and the labor involved, is considerably less than for other known methods when applied to systems of reasonable complexity. The advantages inherent in this method are sufficiently great, that it should find wide application in the engineering analysis and design of systems.

8-3. Future Work

The time domain matrix method may be applied to the complete spectrum of control problems. Therefore, the possibilities for future investigation utilizing this method are boundless.

For all classes of control systems, it would be valuable to investigate the problem of the location of the discrete compensator in the control loops, and determine and classify the relocation of the compensator on the time performance indices.

For nonlinear systems, the investigation of the use of the error

derivatives and the signals available at intermediate points throughout the loops, would be very worthwhile.

For time varying and adaptive system, this method would bring a new insight into the very difficult problems. It is possible to use the time domain matrix approach to investigate the method of steepest descent for adaptive systems, as one example.

REFERENCES

1. Tou, J. T., Digital and Sampled Data Control Systems, McGraw Hill Book Company, Inc., 1959

2. Jury, E. I., Sampled Data Control Systems, John Wiley and Sons, Inc., 1958.

3. Stoll, R. R., Linear Algebra and Matrix Theory, McGraw Hill Book Co., Inc., 1952

4. Frazer, R. A., Duncan, W. J., Collar, A. R., Elementary Matrices, Cambridge University Press, 1952.

5. Cooke, R. G., Infinite Matrices and Sequence Spaces, Macmillan and Co., 1950.

6. Salvadori, M. G., Baron, M. L., Numerical Methods in Engineering, Prentice Hall, Inc., 1952.

7. Kunz, K. S., Numerical Analysis, McGraw Hill Book Company, 1957.

8. Enos, R. L., "An Analog Simulation of A Discrete Compensator For A Sampled-Data System," Masters Thesis, U.S.N. Postgraduate School, December, 1960.

9. Ba Hli, F., "A General Method For Time Domain Network Synthesis," I.R.E. Trans. on Circuit Theory, Vol. CT-1, Sept. 1954, pp. 21-28.

10. Farren, M. C., Phillips, C.A., "Adaptive Sampling Frequency For Sampled Data Control Systems," Masters Thesis, U.S.N. Postgraduate School, December, 1960.

11. Thaler, G. J., "Self Adaptive Control Systems," U.S.N.P.G.S. Tech. Report No. 18, April, 1960.

12. Gibson, J. E., "Generalizing the Adaptive Principle," Control Engineering, December, 1960, pp. 93-96.

13. Proceedings of the Self Adaptive Flight Control Systems Symposium, WADC Tech. Report 59-49, March, 1959.

14. Horowitz, I., "Fundamental Theory of Automatic Linear Feedback Control Systems," I.R.E. Trans. on Automatic Control, December 1959, pp. 5-19.

BIBLIOGRAPHY

Ashby, W. R., An Introduction to Cybernetics, Wiley and Sons, 1956.

Birkhoff, G., MacLane, S., A Survey of Modern Algebra, The MacMillian Co., 1953

Friedland, B., "A Technique For the Analysis of Time Varying Sampled Data Systems," Trans. AIEE, Vol. 76, Pt. II, January 1957, pp. 407-413.

Hildebrand, F. B., Methods of Applied Mathematics, Prentice-Hall, Inc., 1952.

Johnson, R., First Course in Abstract Albegra, Prentice Hall, Inc. 1953.

Thaler, G. J., Brown, R. G., Analysis and Design of Feedback Control Systems, Second Edition, McGraw Hill Book Co., 1960.

Truxal, J. G., Automatic Feedback Control System Synthesis, McGraw Hill Book Co., 1955.

APPENDIX A

THE SYSTEM TRANSFER MATRIX

A-1. Determination of the System Transfer Matrix Using the Z-Transform.

As discussed in chapter two, the determination of the system transfer matrix G is an important first step in the use of the time domain matrix method. It is necessary to evaluate the values of g_n, the values of the impulse response at the sampling instants. If the system transfer function is given in the Laplace variable, it is possible, by standard methods, to transform to the z complex variable where $z = e^{sT}$. Then the values of the impulse response are obtained by inverting the z-equation to the time domain. The simplest method to accomplish this is use of division to yield the response at the sampling instants. That is, G(z) may be expanded as:

$$G(z) = g_0 + g_1 z^{-1} + g_2 z^{-2} + \cdots g_n z^{-n} \qquad \text{(a-1)}$$

This series is the constant term and principal part of the Laurent expansion of G(z).

As an example, consider a system as shown in Figure A-1.

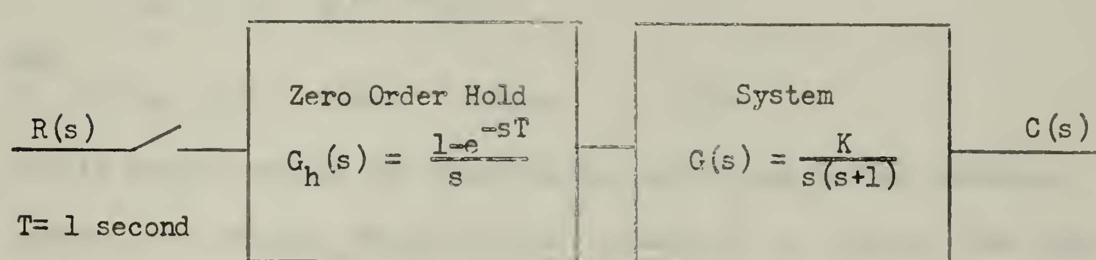


Figure A-1. Open Loop Sampled System

The z transform of this system may be written as:

$$G_hG(z) = \mathcal{Z}\left\{\frac{K(1 - e^{-sT})}{s^2(s + 1)}\right\} \qquad \text{(a-2)}$$

$$= \frac{1}{z-1} - \frac{(1-e^{-1})}{z-.3679} = \frac{K(.3679z^{-1} + .2642z^{-2})}{1 - 1.3679z^{-1} + .3679z^{-2}}$$

Then dividing we have

$$\frac{G_nG(z)}{K} = .3679z^{-1} + .7675z^{-2} + .9145z^{-3} + \cdots \qquad \text{(a-3)}$$

This method is tedious in that one must transform into the z variable and then divide. It also involves an inaccuracy introduced by the round-off in division. Jury[2] presents a method of inversion avoiding division which uses $G(z)$ in the form:

$$G(z) = \frac{p_0 + p_1 z^{-1} + p_2 z^{-2} + \cdots p_n z^{-n}}{1 + q_1 z^{-1} + q_2 z^{-2} + \cdots q_m z^{-m}} \qquad \text{(a-4)}$$

where $m \geq n$.

Then, it can be shown that:

$$g_0 = p_0$$

$$g_1 = p_1 - q_1 g_0 \qquad \text{(a-5)}$$

$$g_2 = p_2 - q_1 g_1 - q_2 g_0$$

and

$$g_n = p_n - q_1 g_{n-1} - q_2 g_{n-2} - \cdots - q_n g_0$$

Jury's method reduces the inaccuracies introduced by the division process, but remains inaccurate for a value of n greater than four due to the roundoff error of multiplication. For a higher order system, the use of the z-transform and inversion is unwieldly and inaccurate.

A-2. The use of the Impulse Response to Evaluate the System Transfer Matrix.

A simple and more direct method of evaluation of the system transfer matrix is the use of the impulse response of the system. If the impulse response $g(t)$ is determined, then the elements of the system matrix are found by substituting $t = nT$. That is,

$$g_n = g(t)\Big|_{t=nT} = g(nT) \tag{a-6}$$

Since, invariably the impulse response is made up of a sum of time functions, the arithmetic operation involved in the evaluation is addition instead of division or multiplication. Therefore, the roundoff error of calculation is reduced as well as the difficulty of manipulation. Now reconsider the example of the previous section. The system transfer function is:

$$G_hG(s) = \frac{1 - e^{-sT}}{s^2(s+1)} \tag{a-7}$$

and the impulse response is:

$$g(t) = \mathcal{L}^{-1}\left\{\frac{1}{s^2(s+1)}\right\} - \mathcal{L}^{-1}\left\{\frac{e^{-sT}}{s^2(s+1)}\right\} \tag{a-8}$$

The e^{-sT} simply implies a delay of one sampling period. Then the inverse is:

$$g(t) = \left[t - (1 - e^{-t})\right] - \left[(t - T) - (1 - e^{-(t - T)})\right] u(t-T)$$

where T = 1 second as on figure A-1. (a-9)

Then the values of interest at the sampling instants when $t = nT = n$, are:

$$g_n = T + e^{-nT} - e^{-(nT - T)}$$

$$= 1 + e^{-n} - e^{-(n - 1)} \qquad n \geqslant 1 \tag{a-10}$$

Then, it is found:

$$g_0 = 0$$

$$g_1 = e^{-1} = .36788 \qquad \text{(a-11)}$$

$$g_2 = 1 + e^{-2} - e^{-1} = .76746, \text{ etc.}$$

Therefore, the values of the system matrix are rapidly and accurately evaluated using a mathematics table. This calculation is a great improvement in accuracy over that of division in equation a-3.

A-3. Evaluation of the System Transfer Matrix for Intersample Response Using the Modified-z-Transform.

In order to determine the response of a system between the sampling instants, it is necessary to obtain the system transfer matrix for values of time between the samples. In order to accomplish this with standard notation, let $m = 1 - \Delta$, where Δ is the percentage delay from the sampling instants as shown in figure A-2.

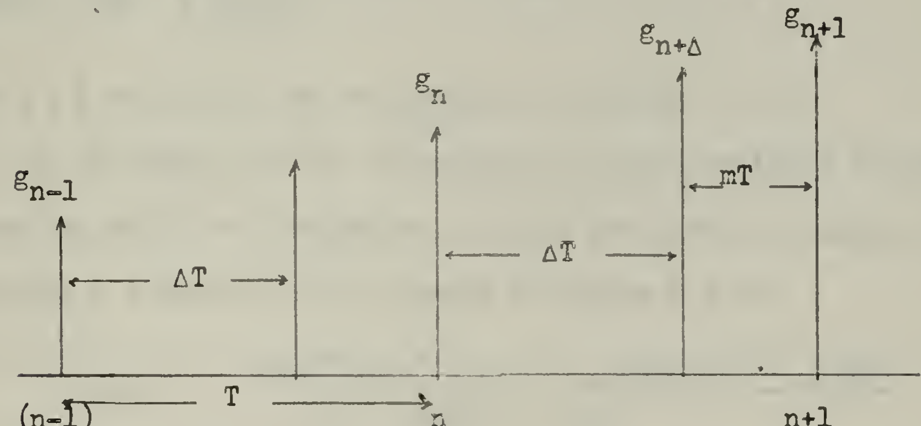


Figure A-2. Delayed Sampling

Then, writing the transformation $t = (n - \Delta)T$ one has:

$$t = (n - \Delta)T = (n - 1 + m)T \qquad 0 < m < 1$$

and the modified z transform is:

[illegible]

[illegible]

[illegible]

[illegible]

$$G(z,m) = z^{-1} \sum_{k=0}^{\infty} g(kT + mT)z^{-k} \qquad \text{(a-12)}$$

The output response of a system between the samples Y(z,m) may then be determined for an imput X(z) as:

$$Y(z,m) = G(z,m)\,X(z) \qquad \text{(a-13)}$$

G(z,m) may be determined and inverted by division or an alternate method to obtain the elements of the system matrix for intersampling response. Then the intersample output response is:

$$\left.Y(m)\right] = \left[G(m)\right] \left.X\right] \qquad \text{(a-14)}$$

where

$$\left.G(m)\right] = \begin{bmatrix} g_0(m) & 0 & 0 & \cdots \\ g_1(m) & g_0(m) & 0 & \\ g_2(m) & g_1(m) & g_0(m) & \cdots \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdots \end{bmatrix}$$

Since $G(z) = \underset{m \to 0}{zG(z,m)}$ (a-15)

the $\left[G\right]$ matrix may be determined by allowing $m \to 0$.

As an example of the determination of the modified system matrix using the modified z-transform consider the previous example. The modified z transform of the system is figure A-2 is:

$$G(z,m) = \frac{(m+e^{-m}-1)z^2 + p_1 z + (.3679m+e^{-m} - .7358)}{z^2 - 1.3679z + .3679} \qquad \text{(a-16)}$$

where $p_1 = 2.3679 - m(1.3679) - 2e^{-m}$

Expansion by the method of section A-1 yields

$$g_0(m) = p_0(m) = m + e^{-m} - 1$$

$$g_1(m) = p_1(m) - p_0(m)q_1 \qquad \text{(a-17)}$$

$$g_2(m) = p_2(m) - p_1(m)q_1 - p_0(m)\left(q_2 - q_1^{\,2}\right)$$

and so on.

A-4. Evaluation of the System Transfer Matrix for Intersample Response Using the Impulse Response.

The values of the modified matrix are simple to evaluate by the impulse response method. One has:

$$g_n(m) = g(nT + \Delta T) = g(t)\Big|_{t = nT + \Delta T} \qquad \text{(a-18)}$$

$$= g(nT + T - mT)$$

Then, for the example previously discussed one has:

$$g_0(m) = (t - 1 + e^{-t})\Big|_{t = nT + \Delta T}, \quad T = 1$$

$$= \Delta T - 1 + e^{-\Delta T} \qquad \text{(a-19)}$$

$$= \Delta - 1 + e^{-\Delta} \qquad \Delta = 1 - m$$

Then, for example when $\Delta = 1/4$, $m = 3/4$ one has:

$$g_0(3/4) = e^{-1/4} - 3/4 = .0288$$

For $n > 1$ one has

$$g_n(m) = T + e^{-t} - e^{-(t-T)}\Big|_{t = nT + \Delta T}, \quad T = 1$$

$$= 1 + e^{-(n+\Delta)} - e^{-(n+\Delta-1)} \qquad \text{(a-20)}$$

Therefore,

$$g_1(m) = 1 + e^{-(1+\Delta)} - e^{-\Delta}$$

$$g_2(m) = 1 + e^{-(2+\Delta)} - e^{-(1+\Delta)} \qquad \text{(a-21)}$$

and for $m = 1/2$, $\Delta = 1/2$

$$g_1(1/2) = .61660 = g_{1.5}$$

$$g_2(1/2) = .85895 = g_{2.5} \qquad \text{(a-22)}$$

where $g_{n+\Delta}$ is another form of notation.

A-5. The System Matrix for A Type I Second Order System with A Hold.

The system transfer function for a type I second order system with a hold may be written as:

$$G_hG(s) = \frac{K(1 - e^{-sT})}{s^2(s + a)} \tag{a-23}$$

Then using the inverse Laplace transform the impulse response is:

$$g(t) = \frac{K}{a^2}\left(at - (1 - e^{-at})\right) \qquad \text{for } t \leqslant T \tag{a-24}$$

and

$$g(t) = \frac{K}{a^2}\left\{\left(at - (1-e^{-at})\right) - \left(a(t-T) - (1-e^{-a(t-T)})\right)\right\} \quad t \geqslant T$$

$$= \frac{K}{a^2}\left[aT + e^{-at} - e^{-a(t-T)}\right] \tag{a-25}$$

Therefore:

$$g_n = \frac{K}{a^2}\left[aT + e^{-anT} - e^{-aT(n-1)}\right] \qquad n \geqslant 1 \tag{a-26}$$

and

$$g_n(\Delta) = \frac{K}{a^2}\left[aT + e^{-aT(n+\Delta)} - e^{-aT(n+\Delta-1)}\right] \tag{a-27}$$

The matrix values for this second order system are tabulated in Table A-1 for three values of the constant aT. These matrix values may be used for any second order system with a hold which possesses the same aT value, making the proper allowance for the $1/a^2$ factor appearing in equation a-26. For example, a second order system with a pole at 10, a sampling period of .10 second, and a gain of 100 will possess the matrix given in Table A-1 for the aT = 1 system with a gain of one.

A-6. The System Matrix for a Type II Second Order System With A Hold Network.

The transfer function of a type two system of interest may be written as:

$$G_hG(s) = \frac{K(s + a)(1 - e^{-sT})}{s^3} \qquad \text{(a-28)}$$

Then using the inverse Laplace transform one obtains:

$$G(t) = K\left\{(t + \frac{a}{2}t^2) - \left[(t-T) + \frac{a}{2}(t-T)^2\right] u(t-T)\right\} \qquad \text{(a-29)}$$

Then, at the sampling instants,

$$g_n = K\left\{T + \frac{a}{2}\left[(nT)^2 - \left((n-1)T\right)^2\right]\right\} \quad n \geqslant 1 \qquad \text{(a-30)}$$

The values for the matrix of this system are given in Table A-1.

TABLE A-1

ELEMENT VALUES FOR THE SYSTEM TRANSFER MATRIX FOR VARIOUS SYSTEMS

Ia Type I Second Order System With A Hold Network

$$\frac{G_h G(s)}{K} = \frac{1}{s^2(s+1)} , \quad T = 1 \text{ second} , \quad aT = 1$$

n	g_n	$n+\Delta$	$g_{(n+\Delta)}$
0	0	.5	.10653
1	.36788	1.5	.61660
2	.76746	2.5	.85895
3	.91445	3.5	.94812
4	.96853	4.5	.98091
5	.98842	5.5	.99298
6	.99574	6.5	.99741
7	.99843	7.5	.99905
8	.99943		
	1.00000		

$(n+\Delta)$	$g_{n+\Delta}$	$(n+\Delta)$	$g_{n+\Delta}$
.25	.0288	.75	.22237
1.25	.5077	1.75	.7014
2.25	.8189	2.75	.89016
3.25	.9334	3.75	.9596
4.25	.975	4.75	.9852

Ib Type I Second Order System With A Hold Network

$$\frac{G_h G(s)}{K} = \frac{1 - e^{-sT}}{s^2(s + 1)} , \quad T = 2 \text{ seconds}, \quad aT = 2.0$$

n	g_n	$n+\Delta$	$g_{n+\Delta}$
0	0	.5	.36788
1	1.13534	1.5	1.6819
2	1.8830	2.5	1.95695
3	1.98416	3.5	1.99417
4	1.99786	4.5	1.99921
5	1.99971		
	2.0000		

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Ic Type I Second Order System with a Hold Network

$$\frac{G_h G(s)}{K} = \frac{1-e^{-sT}}{s^2(s+1)} \quad , \; T = .5 \text{ seconds} \quad , \; aT = .5$$

n	g_n	$n+\Delta$	$g_{n+\Delta}$
0	0	.5	.0288
1	.10653	1.5	.19357
2	.26137	2.5	.31413
3	.35523	3.5	.38727
4	.41221	4.5	.43163
5	.44674	5.5	.45853
6	.46771		
7	.48041		
8	.48812		
9	.49279		
10	.49563		

II A Type I Third Order System With A Hold Network

$$\frac{G_h G(s)}{K} = \frac{4(1-e^{-sT})}{s^2(s+1)(s+2)} \quad , \; T = 1 \text{ second}$$

n	g_n	$n+\Delta$	$g_{n+\Delta}$
0	0	.5	.05824
1	.33616	1.5	.7845
2	1.18686	2.5	1.47885
3	1.67364	3.5	1.7983
4	1.87626		
5	1.9540		
6	1.9830		
7	1.99372		
∞	2.0000		

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

IIIa Type II Second Order System With A Hold Network

$$\frac{G_hG(s)}{K} = \frac{15(s+10)(1-e^{-sT})}{s^3} \quad , \quad T = .05 \text{ seconds}$$

n	g_n
0	0
1	.9375
2	1.3125
3	1.6875
4	2.0625
5	2.4375
6	2.8125

IIIb Type II Second Order System With A Hold Network

$$\frac{G_hG(s)}{K} = \frac{15(s+10)(1-e^{-sT})}{s^3} \quad , \quad T = .10 \text{ seconds}$$

n	g_n	$n+\Delta$	$g_{n+\Delta}$
0	0	.5	.9375
1	2.25	1.5	3.000
2	3.75	2.5	4.500
3	5.25	3.5	6.000
4	6.75		
5	8.25		
n	$g_{n-1} + 1.50$		

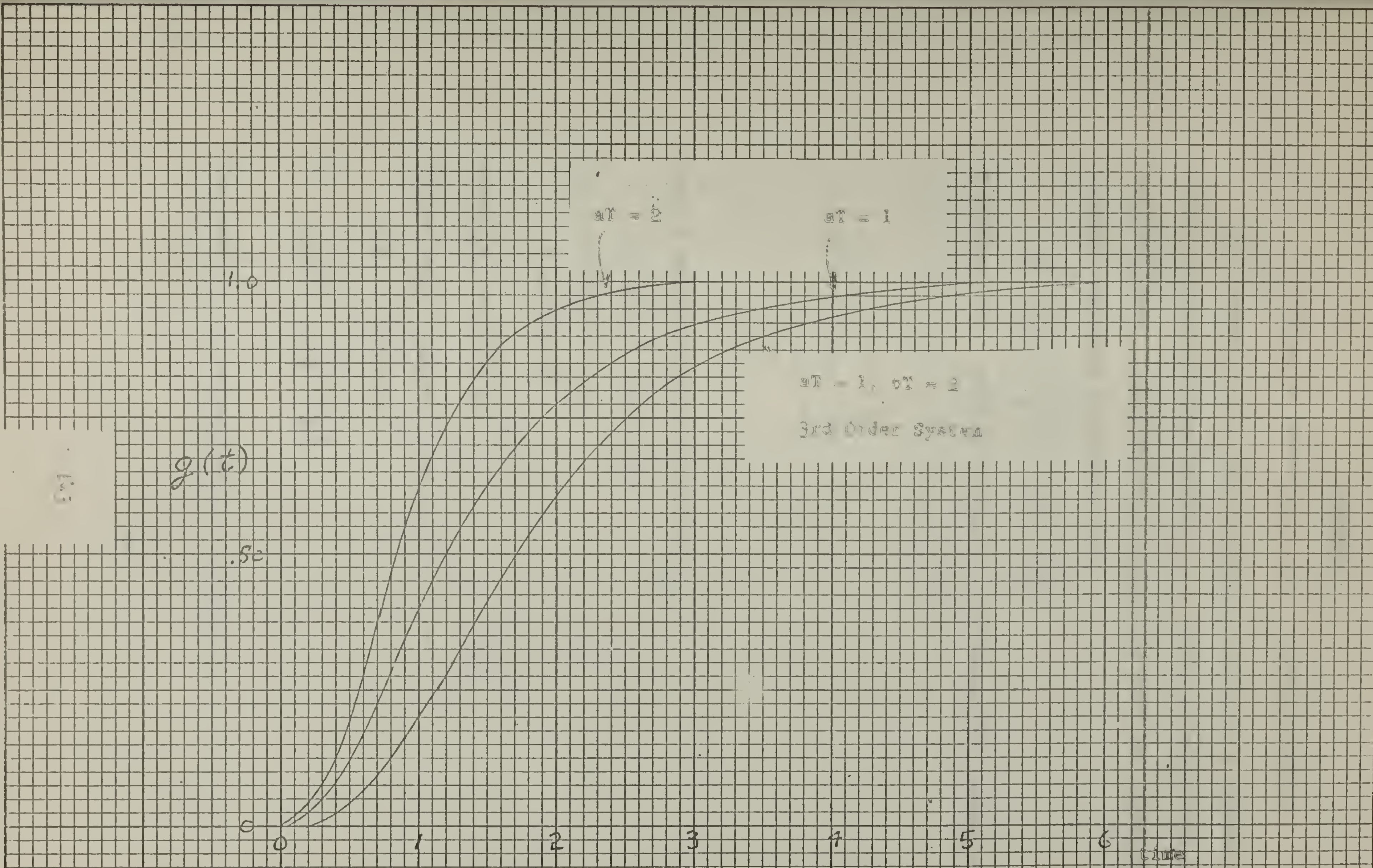
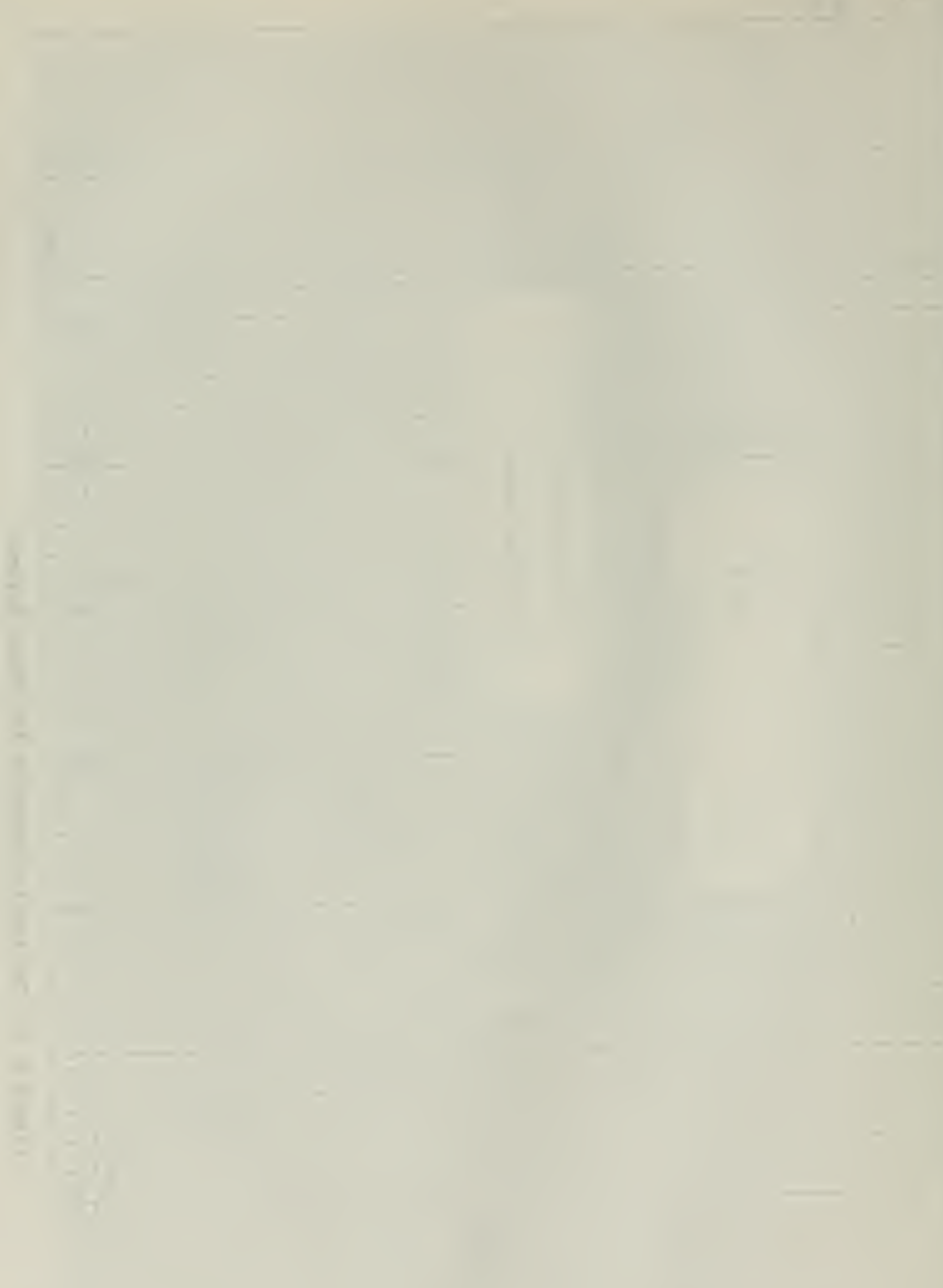


Figure A-3. The Impulse Response For Three Systems.



APPENDIX B

A METHOD OF INVERTING A LOWER TRIANGULAR MATRIX

The inversion of a lower triangular matrix is a necessary step in the evaluation of the response of a system by means of time domain matrices. It was shown in section 2-6 that the output time response was given by equation 2-34:

$$C] = \left\{ [I] - \left[[I] + [G] \right]^{-1} \right\} R] \qquad \text{(b-1)}$$

Therefore, it is usually necessary to calculate the inverse of $[I] + [G] = [A]$. If $[A]$ is written:

$$[A] = \begin{bmatrix} a_{11} & 0 & 0 & \cdots \\ a_{21} & a_{22} & 0 & \cdots \\ a_{31} & a_{32} & a_{33} & \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

Then a matrix equation may be written:[4]

$$[A] [\delta]^{-1} Y] = H] \qquad \text{(b-2)}$$

The solution for Y is:

$$Y] = [B] H] \qquad \text{(b-3)}$$

where $[B] = \left\{ [A] [\delta]^{-1} \right\}^{-1}$

By the rules of matrix inversion:

$$[B] = [\delta] [A]^{-1} \qquad \text{(b-4)}$$

or $$[A]^{-1} = [\delta]^{-1} [B] \qquad \text{(b-5)}$$

The equation b-2 may written for clarity as:

$$
\begin{aligned}
&y_1 & & & &= h_1 \\
&\left(\frac{a_{21}}{a_{11}}\right)y_1 &+ y_2 & & &= h_2 \qquad \text{(b-6)} \\
&\left(\frac{a_{31}}{a_{11}}\right)y_1 &+ \frac{a_{32}}{a_{22}} y_2 &+ y_3 & &= h_3
\end{aligned}
$$

etc.

Then it can be seen that

$$
[\delta] = \begin{bmatrix} a_{11} & 0 & 0 & 0 & \\ 0 & a_{22} & 0 & 0 \cdots & \\ 0 & 0 & a_{33} & 0 & \\ 0 & 0 & 0 & a_{44} & 0 \\ & \vdots & & 0 & a_{nn} \end{bmatrix}
$$

and

$$
[\delta]^{-1} = \begin{bmatrix} (1/a_{11}) & 0 & 0 & 0 & 0 \\ 0 & (1/a_{22}) & 0 & 0 & 0 \\ 0 & 0 & (1/a_{33}) & 0 \cdots & \\ & \vdots & & \vdots & \\ 0 & 0 & & & (1/a_{nn}) \end{bmatrix}
$$

Hence, $[A]^{-1}$ can be obtained by multiplication of the successive rows of $[B]$ by $(1/a_{11})$, $(1/a_{22})$, etc. Then it can be seen that the elements in the first column of $[B]$ are the values of $Y]$ when

$$
H] = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix} = \{1, 0, 0, \ldots\}
$$

The elements in the second column of $[B]$ are the values of $y]$ when $H] = \{0, 1, 0, 0, \ldots\}$ and so on.

When $H] = \{1, 0, 0, \cdot \cdot\}$ one obtains the first column as:

$$y_1 = 1$$

$$y_2 = \left[- \frac{a_{21}}{a_{11}}\right] \{1\}$$

$$y_3 = \left[- \frac{a_{31}}{a_{11}}, - \frac{a_{32}}{a_{22}}\right] \{y_1, y_2\} \tag{b-7}$$

$$y_4 = \left[- \frac{a_{41}}{a_{11}}, - \frac{a_{42}}{a_{22}}, - \frac{a_{43}}{a_{33}}\right] \{y_1, y_2, y_3\}$$

This solution suggests a method of computation which allows these calculations to be easily carried out. The elements of $[\delta]^{-1}$ by which the rows of $[B]$ are to be multiplied, are recorded on the extreme right. This useful scheme, in the form of a table, is shown below for a 4x4 matrix.

--	--	--	--	1	0	0	0	$\frac{1}{a_{11}}$
$-(\frac{a_{21}}{a_{11}})$	--	--	--	b_{21}	1	0	0	$\frac{1}{a_{22}}$
$-(\frac{a_{31}}{a_n})$	$-\frac{a_{32}}{a_{22}}$	--	--	b_{31}	b_{32}	1	0	$\frac{1}{a_{33}}$
$-(\frac{a_{41}}{a_{11}})$	$-\frac{a_{42}}{a_{22}}$	$-\frac{a_{43}}{a_{33}}$	--	b_{41}	b_{42}	b_{43}	1	$\frac{1}{a_{44}}$

The rules for setting up this table are.

i) To form the left-hand array enter blanks in, and to the right of the principal diagonal. Derive the remaining elements from the A elements as shown.

ii) Commence the right hand array by entering units in the principal diagonal and zeros to the right of the diagonal

iii) Calculate the remaining elements of $[B]$ in succession by the following method. To obtain b_{ij} postmultiply the row of the left-hand array level with b_{ij} by the part column of $[B]$ standing above b_{ij}. Blank elements of the left-hand array are to be disregarded.

iv) To find $[A]^{-1}$ multiply the rows of $[B]$ by the scalar factors on the right.

For example, the first column of $[B]$ is completed below:

$$b_{21} = \left[-\frac{a_{21}}{a_{11}} \right] \{1\}$$

$$b_{31} = \left[-\frac{a_{31}}{a_{11}}, \; -\frac{a_{32}}{a_{22}} \right] \begin{bmatrix} 1 \\ b_{21} \end{bmatrix} \tag{b-9}$$

$$b_{41} = \left[-\frac{a_{41}}{a_{11}}, \; -\frac{a_{42}}{a_{22}}, \; -\frac{a_{43}}{a_{33}} \right] \begin{bmatrix} 1 \\ b_{21} \\ b_{31} \end{bmatrix}$$

As an illustration of this method consider the inversion of the matrix $[A]$ given as:

$$[A] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 3 & -11 & -36 & 0 \\ 1 & -6 & -14 & \frac{22}{9} \end{bmatrix} \tag{b-10}$$

One then sets up the computation table following the given rules.

$$\left|\begin{array}{cccc|cccc|c} - & - & - & - & 1 & 0 & 0 & 0 & 1 \\ 0 & - & - & - & b_{21} & 1 & 0 & 0 & -1 \\ -3 & -11 & - & - & b_{31} & b_{32} & 1 & 0 & -\frac{1}{36} \\ -1 & -6 & -\frac{7}{18} & - & b_{41} & b_{42} & b_{43} & 1 & \frac{9}{22} \end{array}\right|$$

Then, the b_{ij} are calculated as follows:

$$b_{21} = \begin{bmatrix} 0 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = 0$$

$$b_{31} = \begin{bmatrix} -3 , & -11 \end{bmatrix} \begin{bmatrix} 1 \\ b_{21} \end{bmatrix} = \begin{bmatrix} -3, & -11 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -3 \tag{b-11}$$

$$b_{41} = \begin{bmatrix} -1 , & -6 , & -\frac{7}{18} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} = +\frac{1}{6}$$

$$b_{32} = \begin{bmatrix} -3, & -11 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -11$$

$$b_{42} = \begin{bmatrix} -1, & -6, & -\frac{7}{18} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ -11 \end{bmatrix} = -6 + \frac{77}{18} = -\frac{31}{18}$$

$$b_{43} = \begin{bmatrix} -1, & -6, & -\frac{7}{18} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -\frac{7}{18}$$

Then, one obtains

$$\begin{bmatrix} B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & -11 & 1 & 0 \\ \frac{1}{6} & -\frac{31}{18} & -\frac{7}{18} & 1 \end{bmatrix} \tag{b-12}$$

and since $\begin{bmatrix} A \end{bmatrix}^{-1} = \begin{bmatrix} \delta \end{bmatrix}^{-1} \begin{bmatrix} B \end{bmatrix}$ one has:

$$\begin{bmatrix} A \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ \frac{1}{12} & \frac{11}{36} & -\frac{1}{36} & 0 \\ \frac{3}{44} & -\frac{31}{44} & -\frac{7}{44} & \frac{9}{22} \end{bmatrix} \tag{b-13}$$

For a physical system it was shown in appendix A that invariably the element g_0 of the system matrix has a value of zero. For these physical systems one is usually interested in evaluating the inverse of $[I] + [G]$ and therefore the matrix $[\delta]^{-1}$ is simply equal to $[I]$.

As an illustration consider a type I system with a hold that has the following transfer function

$$G_hG(s) = \frac{(1-e^{-sT})}{s^2(s+1)} \quad , \text{ where } K = 1 \text{ and } T = 1$$

It was shown in appendix A that the system matrix is:

$$[G] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ .3679 & 0 & 0 & 0 & 0 & 0 \\ .7675 & .3679 & 0 & 0 & 0 & 0 \\ .9145 & .7675 & .3679 & 0 & 0 & 0 \\ .9685 & .9145 & .7675 & .3679 & 0 & 0 \\ .9884 & .9685 & .9145 & .7675 & .3679 & 0 \end{bmatrix} \qquad \text{(b-14)}$$

Then, one obtains for $[I] + [G]$, the principal diagonal containing element values of unity. Completing the table for inversion, one has:

--	--	--	--	--
-.3679	--	--	--	--
-.7675	-.3679	--	--	--
-.9145	-.7675	-.3679	--	--
-.9685	-.9145	-.7675	-.3679	--
-.9884	-.9685	-.9145	-.7675	-.3679

1	0	0	0	0	0	1
-.3679	1	0	0	0	0	1
-.6321	-.3679	1	0	0	0	1
-.400	-.6321	-.3679	1	0	0	1
+.001	-.400	-.6321	-.3679	1	0	1
+.2528	+.001	-.400	-.6321	-.3679	1	1

Furthermore, for non-time varying systems the $[G]$ matrix always contains equal values along any upper left to lower right diagonal. Therefore, inversion of $[I] + [G]$ need be accomplished for only one column as is obvious from the example. This class of matrix called a diagonally invariant matrix (D.I.M.) in chapter three, lends itself to the use of a recursion formula for the elements of the inverse matrix. For the first column of the matrix $[B]$ one may write:

$$b_{k1} = -(a_{k1} + a_{k-1,1}\, b_{k\,1} + \cdots + a_{21}\, b_{k2}) \qquad \text{(b-15)}$$

For example, in the previous solution:

$$b_{31} = -(a_{31} b_{33} + a_{21} b_{32}) \qquad \text{(b-16)}$$

Since, the elements are invariant on the diagonal, one has:

$$b_{33} = b_{11} = 1$$

$$\text{and} \quad b_{32} = b_{21} \qquad \text{(b-17)}$$

Therefore,

$$b_{31} = -(a_{31} + a_{21} b_{21}) = -(.7675 + .3679(-.3679)) = -.6321$$

In order to obtain accurate results, it is necessary to use either a desk calculating machine or an automatic digital computer. This is necessary in order to avoid the roundoff errors of multiplication. For the usual requirements of one percent accuracy, the desk calculator is completely adequate.