ANALOG COMPUTER SIMULATION

OF

SAMPLED-DATA SYSTEMS

by

Byron Lee Dennison

Thesis submitted to the Graduate

Faculty of the Virginia

Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

Electrical Engineering

June, 1962

Blacksburg, Virginia

TABLE OF CONTENTS

I INTRODUCTION AND BACKGROUND

History

During World War II the serious need for automatic control of various devices came to the attention of engineers. At that time most interest was in building automatic devices to regulate some quantity in a system, such as a voltage or an angular velocity, causing it to remain as nearly as possible at some specific value set into the system. Of equal or perhaps greater importance was the necessity of slaving two devices together such that the second (slave) device would reproduce with a minimum of error the motions of the first (master) device. The eventual solution of both of these problems was based on the concept of comparing the actual output of the device or system with the desired output, or reference input, and making necessary corrections to minimize the difference or error. This, of course, amounted to a feedback system and such systems came to be known as servomechanisms. In recent years the use of the feedback principle has been greatly expanded and systems such as those described as well as other related systems are now commonly referred to as automatic control systems or feedback control systems.[1]

Probably the first serious work in feedback control systems.

was done at the Bell Telephone Laboratories about 1941 in connection with a radar controlled anti-aircraft-gun. It was a natural tendency for this work to fall into the hands of electrical engineers, although the devices were largely hydraulically controlled. This was a result of the fact that the phenomena encountered, such as instability, was similar to that observed in feedback amplifiers. These amplifiers had been announced by the Bell Telephone Laboratories in 1934[2] and by the time work was started on control systems their behavior and methods of analysis were familiar to electrical engineers. Today the theory of continuous feedback control systems is widely known and much has been written on the subject; therefore, the familiarity of the reader with such systems is assumed.

In the usual servomechanism all of the variables involved, such as input, output, and error, are continuous functions of time and conventional mathematical techniques can be applied in their analysis. As systems became more involved, in certain situations the variables, or signals, were no longer continuous but changed in discrete steps. Three common examples of this situation are as follows:

Example 1. Relay controlled servomechanisms. In certain

types of systems the actuating signal is either at some fixed value or at zero.[3,4] The controlled output is forced to the desired value of application of pulses of energy in the direction necessary to minimize the system error. These pulses of energy and their direction are controlled by a relay. Methods of analysis and design of relay servomechanisms have been developed and presented in the literature (see Bibliography).

Example 2. Control systems with time shared elements. Situations sometimes exist in more elaborate control systems in which certain elements are time shared between two or more portions of the same system[5], or even with other systems. The share element is often a computer or some other form of controller which is used to provide the input signal for various portions of the system on a time shared basis. The computer might first determine the proper signal to be delivered to a first subsystem (A) and then go on to compute the signal for some other subsystem (B) and from there perhaps to still other portions of the overall system but eventually returning to the original subsystem (A). The various subsystems are then receiving signals periodically which are in the form of pulses and in the limit, for a large number of subsystems, the pulses approach impulses. In many practical situations the amplitude of the trailing edge of these pulses is stored in some type of memory device, possibly a

capacitor, and the operation of that particular subsystem continues based on the stored value as its input. Eventually, however, the computer comes back and again feeds the subsystem directly for a short period of time and so the cycle repeats indefinitely. It is seen that each of the various subsystems is actually being controlled by samples of the continuous input signal which would exist if the computer were not time shared.

Example 3. Control system involving digital computers. In recent years many automatic control systems involve the use of digital computers to determine the signal for a portion of the system or perhaps an entire system[6]. The nature of a digital computer is such that a finite time is required to arrive at a desired solution after the data has been entered into the computer. The usual function of a digital computer in a control system is to, at some instant of time, take the values of the continuous signals being fed from the system into the computer, operate on them in some prescribed manner, and after an interval of time arrive at a solution which is then fed forward into the remainder of the control system. The process is then repeated based on the values being fed into the computer at some later time and, after some delay, the output of the computer will change to the new solution which is also fed into the remainder of the system. The output of

the digital computer might be considered as samples of the signal that would exist at that point if the computer were of a continuous type.

Examples 2 and 3 are similar to the extent that the signals being fed into the control system at certain points are actually samples of the signals which would ideally exist at these points. The two types of systems are referred to collectively as sampled-data systems. Much has been written on these systems in the approximately ten years that they have been in use, some of which is included in the bibliography. A general description and mathematical analysis of sampled-data systems is not attempted in this paper. It is important to note, however, that sampled quantities are more digital in nature than analog, since the quantity changes in discrete steps. In a conventional continuous system the quantities are all of an analog nature. This paper deals primarily with sampled-data systems.

Two other types of quantities are closely related to sampled data. They are referred to as Quasi-Sampled data and Quantized Data[7].

The signal resulting in example 2 above, in which a computer or controller is time shared, might more properly be considered quasi-sampled if the shared device is analog in nature. This results from the fact that during the time the computer is actually feeding

a signal into the system the data is continuous but at all other times the signal is based on a sample. However, if the time shared element is digital in nature the resulting signal would be true sampled-data. In this paper both true sampled-data and quasi-sampled are considered.

A sampled signal consists of values of a continuous signal taken at discrete intervals of time. Thus the sampled-data can have any value but can change value only at certain times. A quantized signal consists of discrete values related to a continuous signal which can change at any time. The discrete values are in effect rounded off values of the continuous signal. Quantized data is truly digital in nature and for that reason is sometimes referred to as digitized data. Such data will always result whenever a digital device, such as a digital computer, is present in an automatic control system. If, however, the difference in value between steps is small, quantized data approximates continuous data and this approximation is assumed to hold in the work presented in this paper.

Throughout this report various phenomena associated with sampled-data systems will be discussed, so it is therefore necessary to define certain terms peculiar to these systems. The necessary terms are defined with reference to figure 1. An arbitrary contin-

uous variable, y(t), which is to be subjected to a sampling-holding process is illustrated in Figure 1a. The result of periodically sampling y(t) is shown in Figure 1b, this being the signal appearing at the output of a simple sampling device. Throughout this paper the superscript * will be used to indicate a sampled quantity or any other discretely changing signal; hence, the sampled variable of Figure 1b is identified as y^*. The time interval "A" is referred to in this paper and in the literature variously as sampling time, sampling interval, and pulse width. The time interval "B" is referred to as the sample period or sample interval. The reciprocal of the interval "B" is referred to as the sampling frequency or sampling rate and is abbreviated f_s. In most, but not all practical cases the sampling rate is constant.

A common situation occuring in practical sampled-data systems is illustrated in Figure 1c. The value of the sampled quantity existing at the end of a sampling time is held until the start of the next sampling period. The time intervals "A" and "B" are defined as above, but, in addition, the interval "C" is known as the hold time or hold period.

As sampled-data systems evolved it became desirable to be able to predict with some degree of accuracy their expected

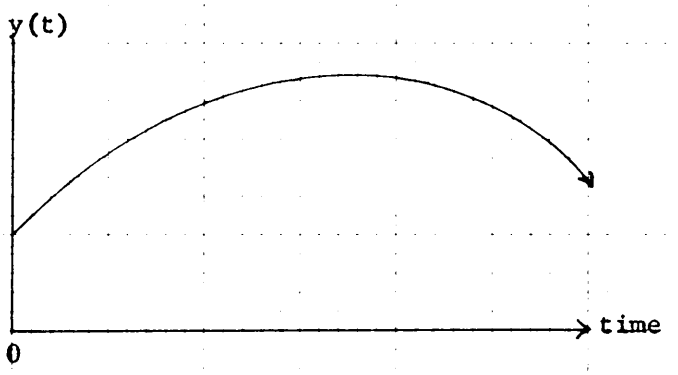


Figure 1a - Continuous Data

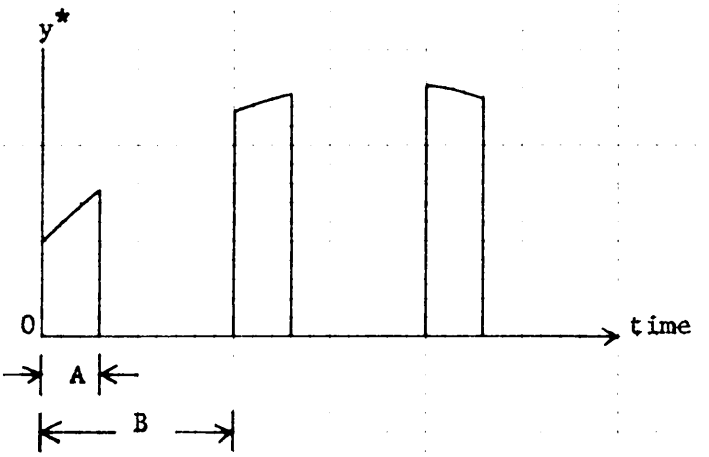


Figure 1b - Sampled Data

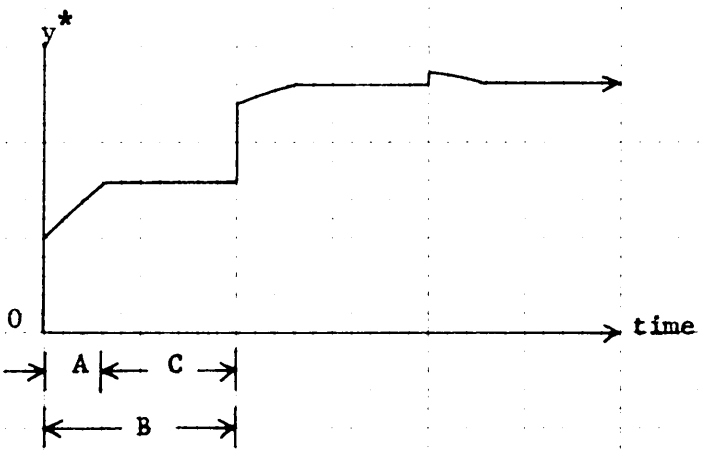


Figure 1c - Sampled Data with Hold

performance. By 1950, when sampled-data systems were first coming into use, the theory for linear continuous systems was well advanced. Methods of design and analysis were well defined and widely known. It was soon learned that sampled-data systems are, in general, less stable than equivalent continuous systems. The result of the sampling process is to introduce time and phase lags with an effect somewhat like that of an additional time constant introduced into the system. It was also learned that if the sample period is short compared to all significant time constants in the system, the behavior of the sampled system approaches that of the equivalent continuous system. If these conditions exist in any particular system, that system can be analyzed by conventional methods with reasonable accuracy. It is however, desirable to be able to analyze cases where the effect of sampling is pronounced and ordinary methods of analysis do not hold.

Mathematical methods of analyzing sampled-data systems are complicated by the discontinuous nature of the variables. In time, however, various analytical methods of predicting the performance of these systems evolved. The first major contribution to analytical analysis was made by Linvill[8] in 1951. Linvill's method involves a comparison of the sampling process

with amplitude modulation and subsequent application of well known methods of modulation analysis.

In 1952, a paper was published by Ragazzini and Zadeh of Columbia University entitled, "The Analysis of Sampled-Data Systems", which is regarded by most authorities today as being the most significant work in this area.[9] In this work the now well-known Z transform operator is developed and applied to the analysis of sampled-data systems. The Z operator is a complex operator and is related to sampled-data systems in much the same way as the Laplace operator is related to continuous systems. The method of Ragazzini and Zadeh is limited in many respects. It is based on the assumption that the sampling is instantaneous (i.e., sampling time is zero) and the resulting system responses hold only at the instants of sampling. Additional practical limitations are that the method is laborious and also that the level of mathematics involved is above that familiar to most engineers in industry. Nevertheless, the paper was a significant advance in the understanding of sampled-data systems.

Several other investigators applied the Z transform and attempted to modify the method of Ragazzini and Zadeh to eliminate some of the restrictions and approximations imposed. In 1956,

Jury published a paper[10] introducing a modified Z transform with which the response of sampled systems can be determined. Jury's method differed from the original in that the system response can be found at any instant and not just at the instant of sampling. The method still assumes zero sampling time and is quite tedious for all but the most simple systems.

Still another significant contribution was by Farmanfarma, who was at the time a graduate student at the University of California at Berkeley. In a paper[11] published in January, 1957, he introduced the P transform, essentially an extension of the Z transform. Using the P transform Farmanfarma demonstrated a method for the analysis of sampled-data systems in which the effect of finite sampling time can be considered. The method is again quite tedious for all but the most simple systems. In a later paper[12] Farmanfarma extended his work to present a method of analysis for multiple sampler systems.

Many other investigators have worked on the analytical approach to sampled-data system analysis; however, these mentioned above probably represent the most significant contributions. In 1957, Murphy and Ormsby published a paper[13] comparing the various techniques available for sampled-data system analysis and also gave an extensive bibliography. A more recent and lengthy

bibliography was given by Freeman and Lowenchuss[14] in 1958. Unfortunately, all analytical methods which have been developed have the disadvantages of involving approximations to various degrees and require a level of mathematics, in their development, if not in their application, above that of many engineers. Furthermore, the methods are tedious to apply in all but the most simple systems, and therefore, are to some extent more useful for theoretical investigations than as a design tool. This paper does not, in general, deal with analytical methods of analysis and these will not be pursued further, but have been discussed primarily as background information for the material to follow. Several books have recently become available in which analytical methods are presented in detail (see bibliography).

During World War II, and the period following, practical electronic analog computers (differential analyzers) were developed. This occurred concurrently with the rapid development of servomechanisms previously mentioned. During the war, Bell Telephone Laboratories developed a high gain direct coupled amplifier (operational amplifier) which was used as a component part of an anti-aircraft-gun director. It was found that these amplifiers could be used in conjunction with various passive networks to, in effect, perform certain mathematical operations, such as summation and integration, on electrical signals. The first general purpose

analog computer using operational amplifiers was constructed at Columbia University[15] and described in 1947. The first commercially produced computers of this type became available about 1950 and by 1955 were in fairly common use by engineers as a design tool. In particular, computers were useful as simulation devices and enjoyed wide application in the design of continuous automatic control systems.

Explanation of Problems and Objectives of Investigations

Because of the complexity of the analytical design methods for sampled-data automatic control systems it is highly desirable to be able to simulate such systems by means of computers. A limited amount of sampled-data simulation has been done using digital computers[16], however, this is not necessarily the most desirable approach to the problem. Since even in a sampled system most, but of course not all, of the variables are continuous it follows that it would be natural to simulate most of the system on an analog computer. In its entirety a sampled-data automatic control system can be thought of as a continuous system operating in conjunction with one or more sampling and possibly holding devices. The only major problem involved in simulating an entire system on the analog computer is the development of a method of simulating the

sampling and holding processes. The continuous portion can be simulated using conventional simulation techniques described in many books on analog computation (see bibliography). Simulation of the sampling and holding processes involves the generation of discontinuous signals which are, to a certain extent, digital in nature. It is necessary that special simulation techniques be developed to represent these processes since in normal usage analog computers generate only continuous signals.

The purpose of the investigation described in this paper was to develop methods for the simulation of entire sampled-data systems on an analog computer. It is desirable that the methods be relatively simple and, if possible, involve only equipment usually included on a general purpose machine. It is also desirable to represent various sampling rates and finite sampling time. To be of practical value it should be possible to perform the simulation with a minimum of internal modifications to the computer.

In most sampled-data systems there is some type of hold device associated with the sampler. The most simple case is that shown in Figure 1c in which the signal fed forward to the continuous portion of the system is either the continuous input signal itself, y(t), or the value of that signal at the end of the last sampling period. This situation is known as a zero order hold. In more involved

system the signal fed forward might depend not only on a present sample but also on the previous sample, a situation referred to as a first order hold. The concept can be extended to earlier samples resulting in still higher order holds. It is, of course, desirable to simulate the hold process on the computer, but this requires some type of memory device to retain the samples as the sampling process continues. This is the sort of process usually associated with a digital computer or device. Methods have been devised for simulation of both the sampling and holding processes on an analog computer and are described in this paper.

In certain types of sampled-data systems, the most recent sample and various previous samples are combined in a manner designed to improve (compensate) the overall performance of the system. This is normally done by means of some type of computer, usually digital, and the resulting computed output is fed forward into the remainder of the system. Often this computation amounts to an attempt to predict or extrapolate the value of the quantity being sampled at some future time. This extrapolation can be of two forms:

1. To predict the value of some future sample from the present and past samples.
2. To continuously predict from present and past samples

the value of the quantity itself at some definite time in the future.

In either event the signal being fed forward to the continuous portion of a sampled-data system usually amounts to a weighted summation of the value of the most recent sample and a finite number of past samples. It is, of course, desirable that this situation be simulated on the analog computer but because of the digital nature of the process, the simulation requires special techniques. In a practical system, particularly if digital devices are involved, a finite time delay occurs during the computation of the signal to be fed forward, which usually causes some deterioration in the performance of the overall system. It is also desirable that this delay be represented, if appreciable, in any simulation of an overall system.

The object of the investigation discussed in this paper was to devise practical methods of simulating all of the phenomena described above associated with sampled-data systems, evaluate their limitations, and to actually use them on a computer. The techniques are described in some detail in this report, in addition, representative results of actual simulations are included to illustrate the effectiveness of the techniques. In order to summarize the results of the investigation a representative

sampled-data automatic control system has been simulated using some of the methods developed and is presented at the end of the report. Curves showing actual data obtained from a computer are included and briefly evaluated as to the quality of simulation. No attempt is made, however, to study sampled-data systems as such.

A limited amount of work has already been reported in the area of analog simulation of sampled-data systems[17-22]. In general the simulated systems involved only a zero order hold device, although Chestnut et al.[22] discuss briefly the simulation of first order holds. Little has been reported in the way of extrapolation of future samples or continuous prediction of a quantity from samples, although it is known that the problem has been considered elsewhere[23]. Most of those techniques reported by other investigators require extensive special equipment in addition to that normally found on a general purpose analog computer. More details on the previous approaches to the problem are presented later in this paper where those methods are compared with the methods described by this author.

In 1957 the author was involved in an analog simulation of a sampled-data system while employed in the Government and

Industrial Division of the Philco Corporation at Philadelphia, Pennsylvania. It appears that the work at Philco was somewhat advanced over that which has been reported in the literature. A significant feature is that a reasonable simulation of a second order hold was accomplished. The work was done in connection with a proposed system for military requirements and, unfortunately, was not reported. It is emphasized, however, that the work in the present investigation is somewhat an advancement over that done at Philco.

In the discussion to follow it is assumed that the reader is already familiar with the principles of analog computation and to some extent familiar with the DC analog computer. A knowledge of simple linear programming is also assumed, but nonlinear techniques are discussed as needed. As previously stated a basic knowledge of continuous automatic control systems is assumed but no specific knowledge of sampled-data systems, beyond that discussed in this report, is necessary.

II SIMULATION OF SAMPLING AND HOLDING DEVICES

Storage Circuits for the Analog Computer

If a sampled-data system is to be simulated on the analog computer it is, of course, necessary that a method of simulating the sampling and holding processes be devised. A limited amount of work has already been done in this area and two distinct methods of simulating these processes have been developed.

1. The storage of a sampled voltage on a large low leakage capacitor[18].
2. The use of an integrating operational amplifier in the hold condition as a storage device[19-22].

Some difficulty is encountered with the first method due to the fact that when the voltage is sensed and fed into the remainder of the computer circuit some of the charge is necessarily drained from the capacitor causing its voltage to change. This necessitates the use of large capacitors and a sensing device with a very high input impedance.

The second method, using an operational amplifier, is preferable since the stored voltage is unaffected by the load on the circuit. The storage time is essentially determined only by the leakage resistance of the feedback capacitor associated with the

amplifier. On most analog computers it is possible to operate integrators in the hold condition without the use of external circuitry. Using this technique it is necessary to use only equipment normally present on computers to simulate the sampling and holding processes.

An operational amplifier and its associated input and feedback impedances is shown in Figure 2. The transfer function of this circuit can be shown to be[24]

$$\frac{E_{OUT}(s)}{E_{IN}(s)} = -\frac{Z_f(s)}{Z_i(s)} \quad . \tag{1}$$

It is assumed in this relation that the open loop gain of the operational amplifier is very large, that the amplifier input impedance is infinite, and that the output impedance is zero. These conditions are very nearly satisfied in the operational amplifiers usually employed in analog computers.

A special case of Figure 2, in which the input impedance is simply a resistance and the feedback impedance is a parallel resistance-capacitance combination, is illustrated in Figure 3. This circuit is of particular interest since it can be adapted to serve as a storage or memory device in an analog computer and is, therefore, examined in some detail.

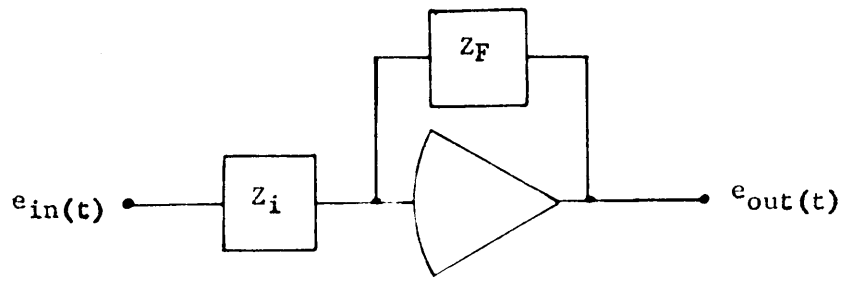


Figure 2

Operational Amplifier with Input and Feedback Impedances. General Case.

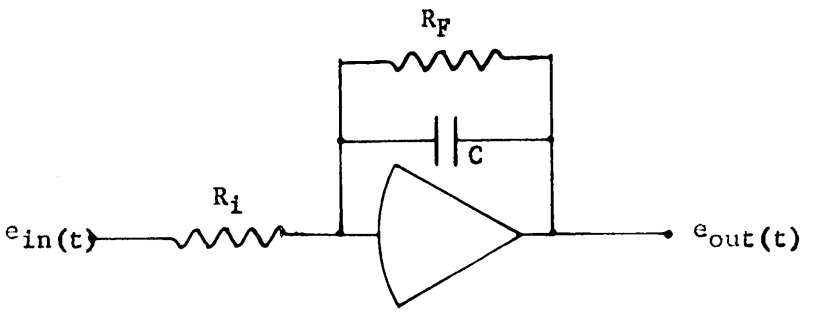


Figure 3

Operational Amplifier with Input and Feedback Impedance. Special Case of First Order Lag Circuit.

The input impedance is expressed simply as

$$Z_{i(s)} = R_i . \tag{2}$$

The feedback impedance is

$$Z_{f(s)} = \frac{R_f \times \frac{1}{sC}}{R_f + \frac{1}{sC}} = \frac{R_f}{R_f C s + 1} . \tag{3}$$

Substitution into equation 1 results in the transfer function

$$\frac{E_{OUT(s)}}{E_{IN(s)}} = -\frac{\frac{R_f}{R_f C s + 1}}{R_i} = -\frac{R_f}{R_i}\frac{1}{R_f C s + 1} . \tag{4}$$

The implied differential equation indicating the stimulus response relationship of the circuit as functions of time is

$$R_f C \frac{d(e_{OUT})}{dt} + e_{OUT} = -\frac{R_f}{R_i} e_{IN} . \tag{5}$$

Consider the response of the circuit to the application, at $t = 0$, of a step voltage E with the capacitor initially charged to the voltage E_o. Returning to the Laplace notation but including the initial value of e_{out} the differential equation can be expressed as[25]

$$R_f C (s E_{OUT(s)} - E_o) + E_{OUT(s)} = -\frac{R_f}{R_i} E_{IN(s)} . \tag{6}$$

which, when rearranged gives

$$E_{OUT(S)} = -\frac{R_f}{R_i}\frac{E_{IN(S)}}{R_fCs+1} + \frac{R_fCE_o}{R_fCs+1} . \qquad (7)$$

But for this case it is known that

$$e_{IN(t)} = E \quad \text{OR} \quad E_{IN(S)} = \frac{E}{S} \qquad (8)$$

which upon substitution into equation 7 yields

$$E_{OUT(S)} = -\frac{R_f}{R_i}\frac{E}{s(R_fCs+1)} + \frac{R_fCE_o}{R_fCs+1} . \qquad (9)$$

Use of the partial fraction expansion on equation 9 results in

$$E_{OUT(S)} = -\frac{R_f}{R_i}E\left[\frac{1}{S} - \frac{R_fC}{R_fCs+1}\right] + \frac{R_fCE_o}{R_fCs+1} \qquad (10)$$

which can be arranged into the more familiar form

$$E_{OUT(S)} = -\frac{R_f}{R_i}E\left[\frac{1}{S} - \frac{1}{S+\frac{1}{R_fC}}\right] + \frac{E_o}{S+\frac{1}{R_fC}} . \qquad (11)$$

The resulting output voltage in the time domain is obtained by application of the inverse transform[26] and is

$$e_{OUT(t)} = -\frac{R_f}{R_i}E\left(1 - e^{-\frac{t}{R_fC}}\right) + E_o e^{-\frac{t}{R_fC}} . \qquad (12)$$

Equation 12 can be rearranged to give the more useful form

$$e_{our}(t) = \left[E_o + \frac{R_f}{R_i} E\right] e^{-\frac{t}{R_f C}} - \frac{R_f}{R_i} E. \qquad (13)$$

Study of equation 13 reveals that the response of the circuit of Figure 3 to a step input is exponential in nature starting at E_o at $t = 0$ and after an infinite time, eventually reaching a value of $-\frac{R_f}{R_i}E$. The time constant is R_fC seconds. Figure 4 shows graphically the behavior of the output voltage if the initial value is less than the final value and Figure 5 shows the results if the initial value is greater than the final value. Note that it is possible for initial and final values to be of opposite signs as indicated in Figure 5.

In order to adapt the circuit of Figure 3, known as a first order lag circuit[27], to the simulation of the sampling process it is necessary that the control relays be added as shown in Figure 6. Assume for the present that relay contacts "B" are always open. The circuit resulting with relay contacts "A" closed is a first order lag circuit identical to that shown in Figure 3 and having the transfer function expressed by equation 4. When relay contacts "A" are open the circuit becomes that of a conventional integrator in the hold condition[28] which is described in most books on analog

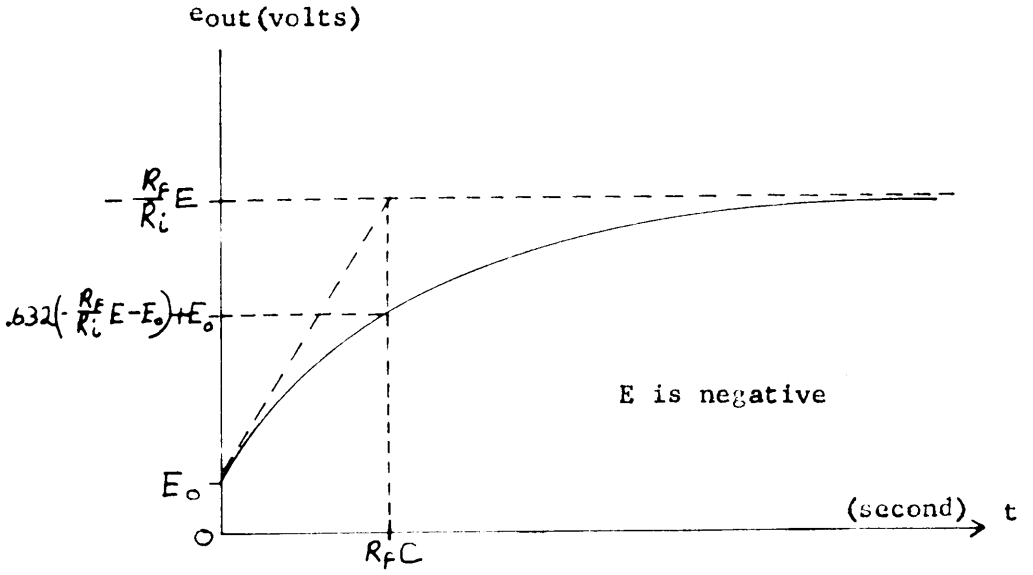


Figure 4

First Order Lag Circuit with Step Excitation. $-\frac{R_f}{R_i}E > E_o$

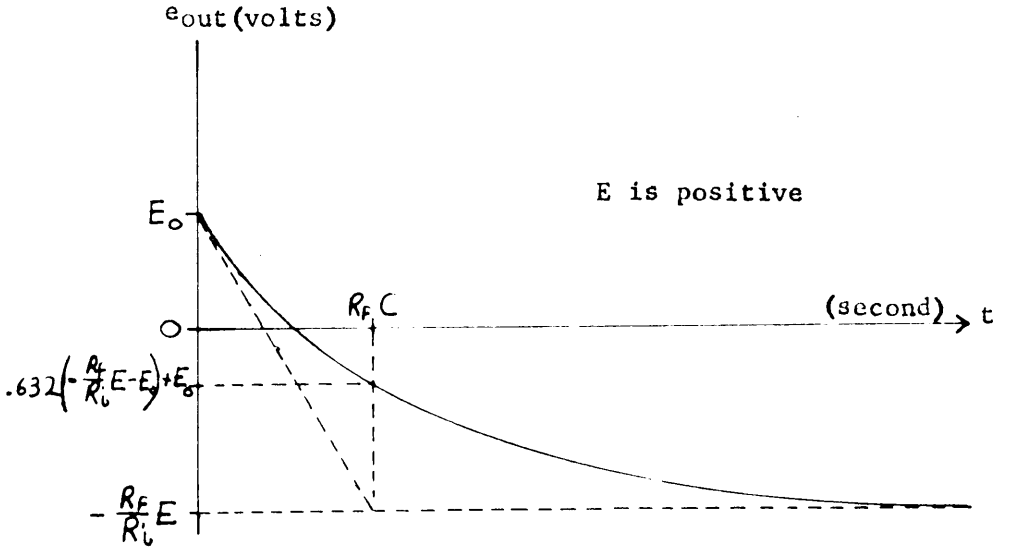


Figure 5

First Order Lag Circuit with Step Excitation $E_o > -\frac{R_f}{R_i}E$

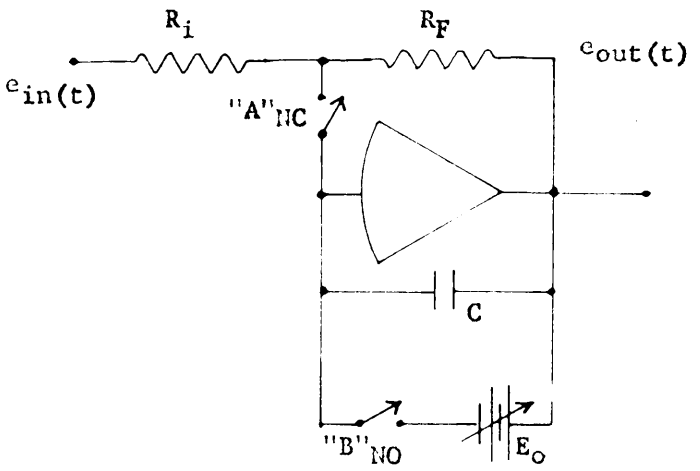


Figure 6a

Diagram of Storage Element Showing Control Circuitry

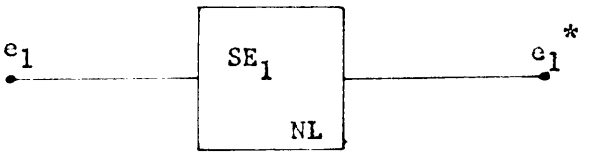


Figure 6b

Block Diagram Symbol for Storage Element

computation. The output voltage of the amplifier will be held indefinitely at the value existing at the time contacts "A" were opened. Relay "B" can be used when desired to set and hold the output voltage to the value E_o. The voltage E_o can be obtained from one of the floating power supplies available on many computers and can be of either polarity. In many cases it is only desired to set the output voltage to zero in which case the floating power supply would be eliminated and relay "B" would simply discharge the capacitor to zero.

Inspection of equation 13 reveals that if the input voltage remains constant at some value E and if $R_I = R_F$, the output voltage of the amplifier will become -E after all transients decay. If relay "A" is then opened, the output voltage will remain at -E indefinitely even though the input voltage changes.

In order to further evaluate the first order lag circuit of Figure 3 the response to a ramp input voltage is examined. For simplicity it is assumed that the ramp starts from zero at t = 0 and is expressed mathematically as

$$e_{in}(t) = At \tag{14}$$

where A is a constant and is indicated graphically in Figure 7a.

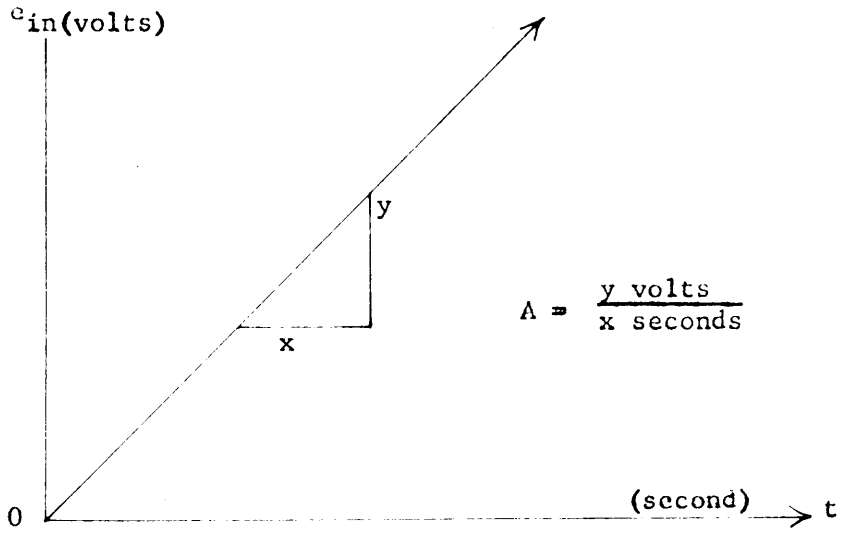


Figure 7a - Ramp Input Voltage $e_{in(t)} = At$

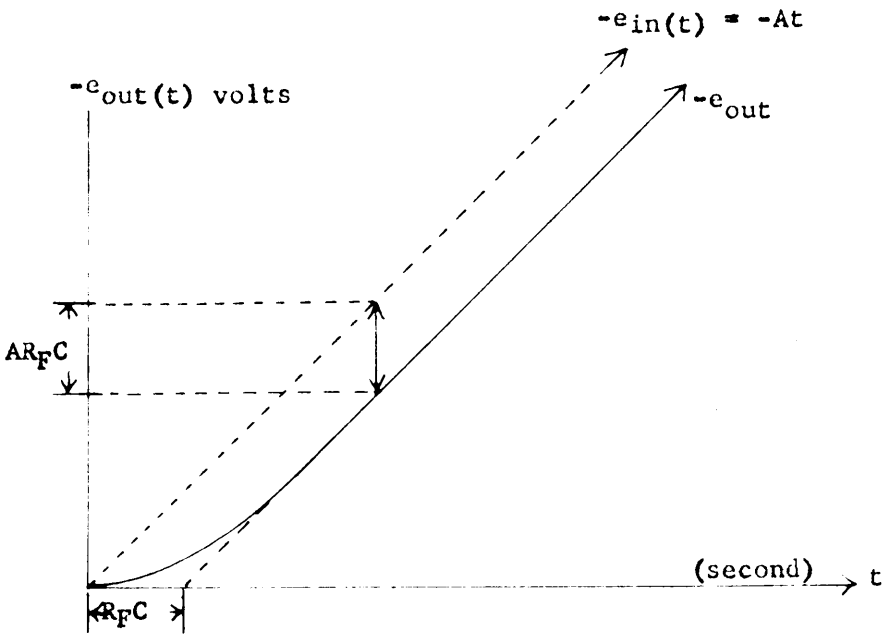


Figure 7b - Response to Ramp Input Voltage $e_{in(t)} = At$

$(R_F = R_i)$

It is assumed that there is no initial charge on the capacitor such that $e_{o(t)} = 0$ when $t = 0$.

It has been previously shown that the output voltage of the operational amplifier is

$$E_{OUT(s)} = -\frac{R_f}{R_i}\left(\frac{1}{R_f C s + 1}\right) E_{IN(s)}. \tag{15}$$

In the present case of a ramp input the forcing function is[26]

$$E_{IN(s)} = \frac{A}{s^2} \tag{16}$$

which upon substitution into equation 15 yields

$$E_{OUT(s)} = -\frac{R_f}{R_i}\left(\frac{A}{s^2(R_f C s + 1)}\right). \tag{17}$$

Application of the partial fraction expansion results in

$$E_{OUT(s)} = -A\frac{R_f}{R_i}\left(\frac{R_f^2 C^2}{R_f C s + 1} + \frac{1 - R_f C s}{s^2}\right) \tag{18}$$

which can be rewritten in the more familiar form

$$E_{OUT(s)} = -A\frac{R_f}{R_i}\left(\frac{R_f C}{s + \frac{1}{R_f C}} + \frac{1}{s^2} - \frac{R_f C}{s}\right). \tag{19}$$

Taking the inverse transform to obtain the output as a function of time gives[26]

$$e_{OUT(t)} = -A \frac{R_F}{R_i} \left[R_F C \, e^{-\frac{t}{R_F C}} + t - R_F C \right]. \qquad (20)$$

It is seen that the response to a ramp input has a transient term in the form of a decaying exponential with time constant R_fC as well as a steady state term which is, as would be expected, in the form of a ramp. After all transients have decayed the output voltage follows the relationship

$$e_{OUT(t)} = -\frac{R_F}{R_i} A \left(t - R_F C \right). \qquad (21)$$

The response of a first order lag circuit to the ramp in Figure 7a is shown graphically in Figure 7b. For purposes of comparison $-e_{out}(t)$ is shown rather than $+e_{out}(t)$. Note that an extension of the steady state output ramp does not pass through the origin but instead intersects the time axis at $t = R_fC$, or at one time constant.

If again the special case of $R_i = R_F$ is considered, it is seen that for positive values of A the output voltage will, at all times, be less, in magnitude, than the output voltage. After the

transient decays the difference is expressed as

$$|e_{IN}(t)| - |e_{OUT}(t)| = AR_fC. \qquad (22)$$

This relationship can also be interpreted to indicate that the output voltage follows the input voltage but lags in time by an amount R_fC seconds. Note that the time lag is constant but the voltage lag or difference is a function of the rate of change of the input voltage.

With the relays shown in Figure 6 it is possible for a first order lag circuit to perform the function of sampling the input voltage and then holding the sample indefinitely. This is explained using the curves of Figure 7. If the input voltage is changing at a rate of A volts per second it has been shown that under steady state conditions the output voltage will change at the same absolute rate with the relay "A" closed and relay "B" open. If at t = T relay "A" is opened the output voltage of the amplifier will be held indefinitely at $e_{out}(T)$. If $R_f = R_i$ the output voltage is

$$e_{OUT}(t) = -(e_{IN}(t) - AR_fC) \quad \text{for } t \gg T \qquad (23)$$

It is seen that the voltage which is held is essentially of the same magnitude as the input voltage which existed at $t = T$ provided the rate of change of the input voltage, A, and/or the circuit time constant, R_fC, are small. If these conditions are satisfied the process can be thought of as sampling the input voltage with the sample being stored as long as relay "A" remains open. There will, however, always be some error in any practical circuit where the quantity AR_fC cannot be made zero. The product of R_fC cannot, of course, be zero; however, if the input voltage to the amplifier was non-varying there would be no error in the sample, assuming that all transients had decayed prior to opening relay "A". The percent of error in the sample is greater when the magnitude of the input voltage is small. The amount of error, however, remains constant.

From the analysis of the first order lag circuit it is seen that it is desirable to have the time constant, R_fC, as small as practical for two reasons:

1. The transient response of the circuit will decay faster. During transient conditions the output voltage will be different in magnitude from the input voltage, assuming $R_f = R_i$, even though the

input voltage is not time varying.

2. If the input voltage is time varying there will be a closer correspondence between the magnitudes of input and output voltage under steady state conditions.

In analog computers using vacuum tubes it is impractical to use input resistors much below 100KΩ because of the loading effect on other computer elements and because of high power dissipation in the resistor itself. Feedback capacitors less than .01μf are seldom used in real time computers because of the very fast transients that result and which may contain frequencies above the upper cutoff point of the operational amplifier. As a result of these limitations the smallest practical time constant is on the order of .001 seconds; that resulting from the use of R_f = 100KΩ and C = .01μf. Since for use as a sample-hold device $R_f = R_i$ in a first order lag circuit it follows that the input resistance compatable with the values above is also 100KΩ. Precision resistors should be used since the accuracy of the samples depends on the equality of the resistors. The capacitor tolerance, however, is not critical since only the time constant is affected.

Using the values suggested above substitution into equation

22 indicates that the sample of an input voltage with a rate of change of not more than 1000 volts/second will not be in error by more than one volt. In addition, it is seen from equations 13 and 20 that for practical purposes all transients will be negligible after about 5 milliseconds of decay. It is concluded that in the circuit of Figure 6 the signal to be sampled should be present for at least 5 milliseconds prior to the opening of the "A" relay if an accurate sample is to be obtained.

In general, the input voltage would not be changing at a constant rate but the rate, A, would also be time varying. If over a period of several time constants the rate can be assumed constant the preceeding analysis for a ramp input is indicative of the response of the circuit to any form of time varying input voltage. For the constants suggested the rate of change or slope of the input voltage should be relatively constant over a period of at least 5 milliseconds if the sample error is to be computed using equation 22.

A common and important type of signal encountered in sample-data automatic control systems is the sinusoid. This special case is investigated in further detail. The steady state response of the first order lag circuit (Figure 3) with

sinusoidal input voltage is obtained by letting $s = j\omega$[29] in equation 4 resulting in

$$E_{OUT}(j\omega) = -\frac{1}{j\omega R_f C + 1} E_{IN}(j\omega). \tag{24}$$

if

$$e_{IN}(t) = E \sin \omega t \tag{25}$$

the steady state output voltage can be shown to be[30]

$$e_{OUT}(t) = -\frac{E}{\sqrt{\omega^2 R_f^2 C^2 + 1}} \sin(\omega t - \theta) \tag{26}$$

where

$$\theta = \tan^{-1} \omega R_f C. \tag{27}$$

The response indicated by equation 26 and 27 is shown graphically in Figure 8.

From Figure 8 it is seen that at any given time the input and output voltages do not correspond in magnitude for two reasons.

1. The input and output waveforms differ in amplitude. This alone, at a given frequency, might be compensated for by use of unequal input and feedback resistances, i.e., $R_f \neq R_i$.
2. The waveforms of input and output voltage are displaced

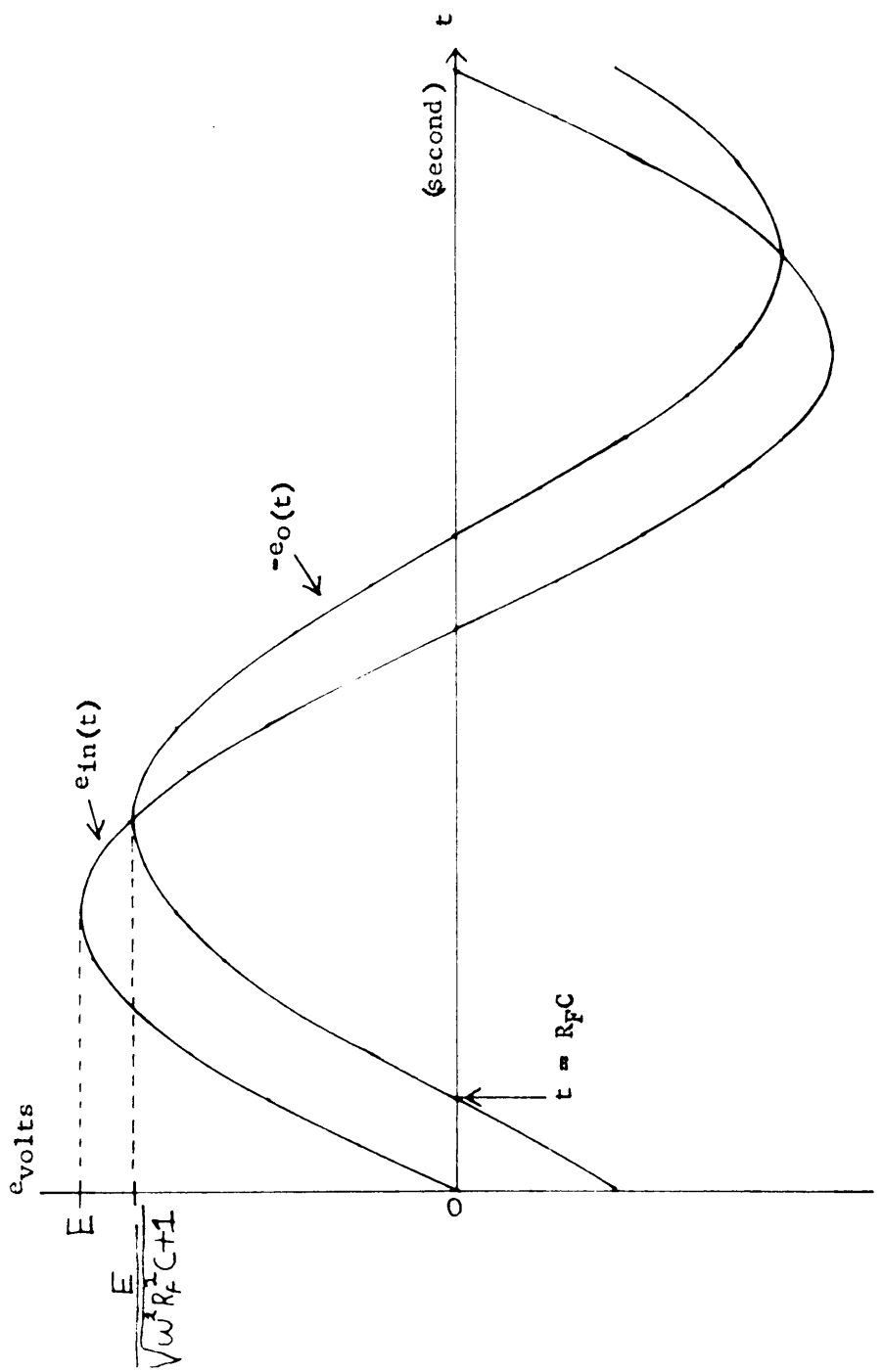


Figure 8

Response of First Order Lag Circuit to Sinusoidal Input. ($R_F = R_1$)

in phase relative to each other. Depending on the portion of the cycle considered the instantaneous magnitude of the output voltage can be larger or smaller than that of the input voltage.

It is seen that, in general, if it is attempted to sample the input voltage by placing the amplifier in the hold condition some error would exist in the sample. Furthermore, it is seen that, depending on the portion of the cycle that the sample is taken, the resulting sample could be larger or smaller than the true value.

Inspection of equation 24 reveals that if $\omega R_f C \ll 1$ the sample errors become quite small. As an illustration it can be shown, using the component values discussed previously ($R_f = R_i = 100K\Omega$, $C = .01\mu f$), that if the angular frequency does not exceed 100 radians per second (f = 15.92 cps) the phase lag will not exceed 5.75 degrees and the output amplitude will be greater than 99.01% of the input amplitude. The maximum possible error can be shown to be less than 9.9% of the input amplitude and occurs if the sample happens to be taken as the input signal passes through zero. In practical control systems it is unlikely that frequency components as high as 100 radians per second would exist.

To summarize it is seen that the first order lag circuit

illustrated in Figure 6 could serve as a sampling and holding device in an analog computer simulation of a sampled-data system provided the rate of change of the signals to be sampled is not too great. Using the component values suggested above it is seen that appreciable sampling errors would not occur with the time constants and natural frequencies normally encountered in practical automatic control systems. If, however, errors are present because of excessive rates of change they can be reduced in either of two ways.

1. Time scale the computer program to slow the problem thereby reducing the rate of change of all variables. It is emphasized, however, that if the rates of change approach the values discussed above it would probably be desirable, if not necessary, to slow the solution based on considerations other than simulation of sampling.
2. Use smaller values for R_f, R_i, and C in the sampling device. This, however, might require the use of computing components not usually standard on computers and is not a desirable solution.

An apparent problem with the circuit of Figure 6 is the period of time that the sampled voltage can be held without

changing appreciably. In theory the voltage would hold indefinitely, with relay "A" open, because there is no path through which the capacitor can discharge. In practice, however, this is not the case and the voltage will not hold indefinitely for three reasons.

1. The capacitor dielectric has a finite resistance allowing the capacitor to discharge itself. This resistance in a good quality capacitor is on the order of 10^{12} ohms or greater[31].
2. Resistance in parallel with the capacitor due to imperfect insulation in the computer patchboard and wiring.
3. The finite input resistance at the summing junction of the operational amplifier resulting in a small grid current tending to discharge the feedback capacitor. This current is typically less than 1 microampere[32].

Considering the effect of all of these, the equivalent resistance shunting the feedback capacitor is several hundred megohms resulting in a long discharge time for the circuit. A measurement on a hold circuit using a Donner Type 3930 .01μf capacitor with the Heath model ES-400 computer indicated a

discharge time constant of about 3 hours. In the sampling operations to be described it is unnecessary for the hold condition to exist more than a few seconds at most, therefore, capacitor discharge in the hold condition is of no practical significance.

The first order lag circuit shown in Figure 6, which has been analyzed in considerable detail, can be used in various combinations to simulate the sampling and holding process of a sampled-data automatic control system. It will hereafter be referred to as a storage element and represented in block diagram form as illustrated in Figure 6b. When the "A" relay is open the storage element is said to be in the store (S) condition or state; when the "A" relay is closed it is in the load (L) condition. The symbol e_1^* is used to indicate sampled or stored values of the input voltage, e_1, fed into storage element one (SE_1). The notation NL is included in the block to indicate that the circuit is normally in the load condition. This corresponds to Figure 6a in which relay "A" is normally closed and relay "B" is normally open. The notation NS is used to indicate storage elements in which both relays "A" and "B" are normally open.

In the terminology of sampled-data systems sampling devices

are classified by the method in which the sampled outputs are obtained. They are defined as follows[33]:

1. A zero order hold device has an output determined by the most recent sample received and is constant between sampling periods. The output is based on the assumption that the continuous input will remain constant.
2. A first order hold device has an output determined from the most recent sample received and the previous sample. The output is constant between sampling periods and is based on the assumption that the first derivative of the continuous input is constant.
3. A second order hold device has an output determined from the most recent sample received as well as the two previous samples. The output is constant between sampling periods and is based on the assumption that the second derivative of the continuous input is constant.

In general, the output of any hold device can be expressed as a weighted summation of the most recent sample and a finite number of past samples such as the following:

$$Y^* = AX_o^* + BX_1^* + CX_2^* + \text{------} \qquad (28)$$

where Y* is the output of the hold device.

X_0^* is the most recent sample of the continuous input signal.

X_1^* is the sample previous to X_0^*.

X_2^* is the sample previous to X_1^*.

A, B, C, etc. are constants which can be either positive or negative.

In the analog computer the output and samples will appear as voltages. The relation for a hold simulator in terms of voltages is therefore

$$e_{out}^* = Ae_0^* + Be_1^* + Ce_2^* + \cdots\cdots \quad (29)$$

where the subscripts have the same significance as in equation 28.

Methods for simulation of each type of hold device up to second order have been developed using the storage elements which have been described. The storage elements can be connected in series or in a series of parallel pairs to represent different order hold devices. Both methods are considered and evaluated in detail.

Zero Order Hold Simulation Using Parallel Connected Storage Elements

The most simple of the hold devices to simulate and also the most commonly encountered is the zero order hold. This can be simulated using two storage elements connected in series or in parallel. The parallel circuit is the most obvious and is

discussed first.

A parallel connected zero order hold simulator is shown schematically in Figure 9a and in block diagram form in Figure 9b. The multiple contact "A" relay is one of those normally used to set initial conditions into integrators in the Heath Model ES-400 computer. The multiple contact "B" relay is one of those normally used as a hold relay for integrators in that computer. The Heath ES-400 analog computer was used in this study because of the easily accessible relay contacts which are available for use in non-conventional circuits such as the storage elements used here.

Assuming ideal storage elements the operation of the zero order hold simulator can be explained by reference to the timing diagram Figure 10. The assumption of an ideal storage element implies a time constant of zero and an output in the load state which is equal in magnitude to the input voltage, but of opposite polarity.

An arbitrary time varying voltage, $e_{in(t)}$ is shown in Figure 10a. The initial values of e_1^* and e_2^* are assumed zero. For the present discussion it is also assumed that the "B" relay contacts remain open. The sampling process is started by alternately energizing and releasing relay "A" as indicated in Figure 10b, resulting in the output voltages indicated in Figures 10c,

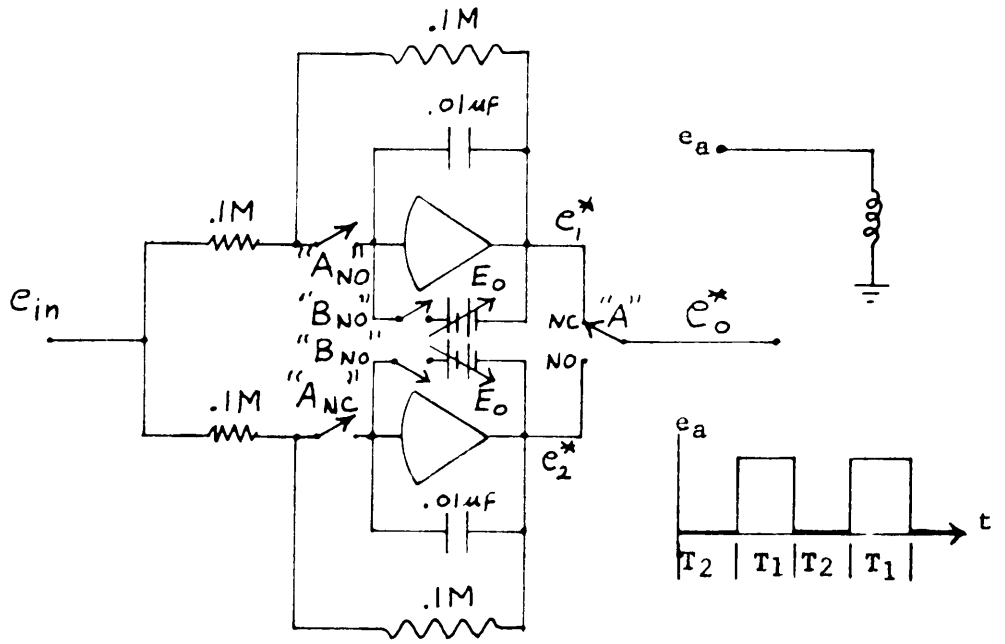


Figure 9a

Schematic Diagram of Parallel Connected Zero Order Hold Simulator

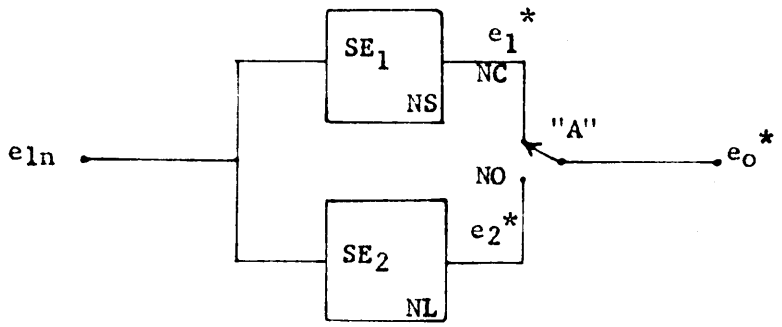


Figure 9b

Block Diagram of Parallel Connected Zero Order Hold Simulator

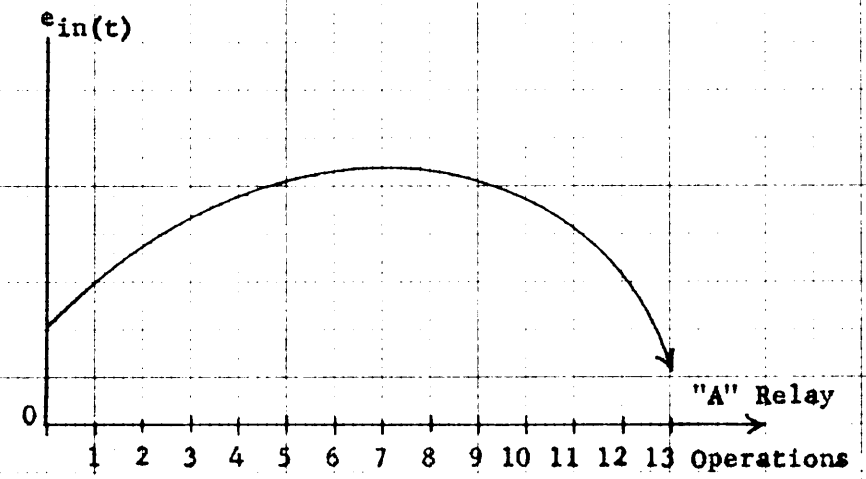


Figure 10a - Arbitrary Input Voltage

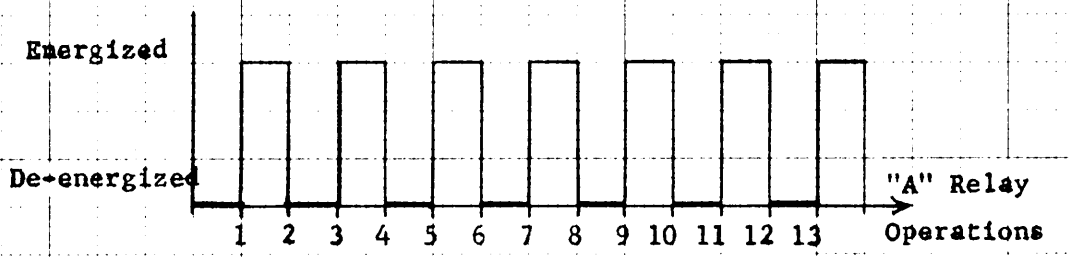


Figure 10b - State of "A" Relay

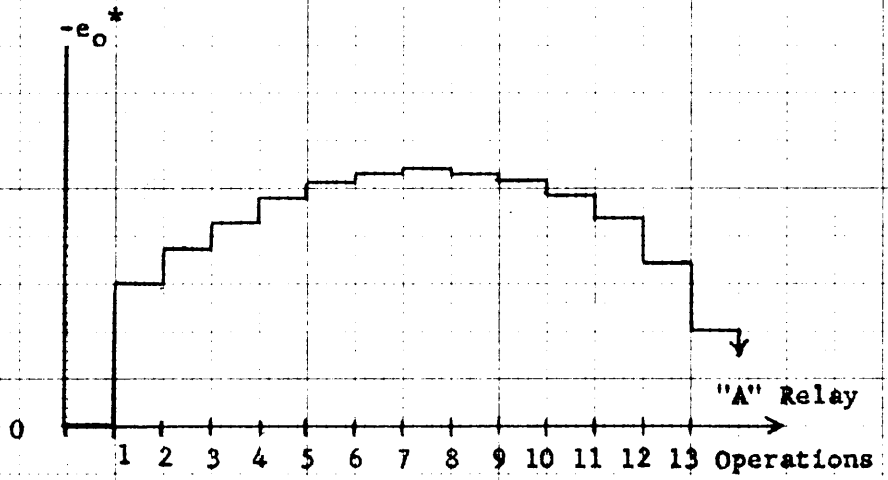


Figure 10c

Output Voltage of Parallel Connected Zero Order Hold Simulator

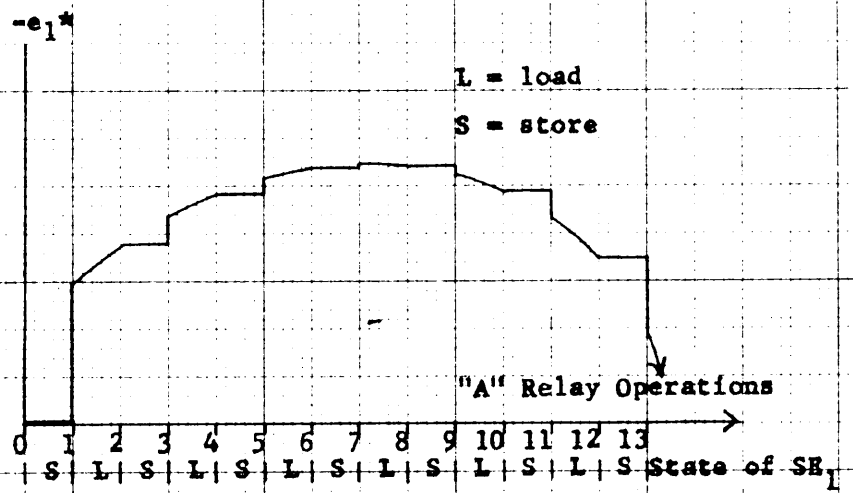


Figure 10d - Output of Storage Element One

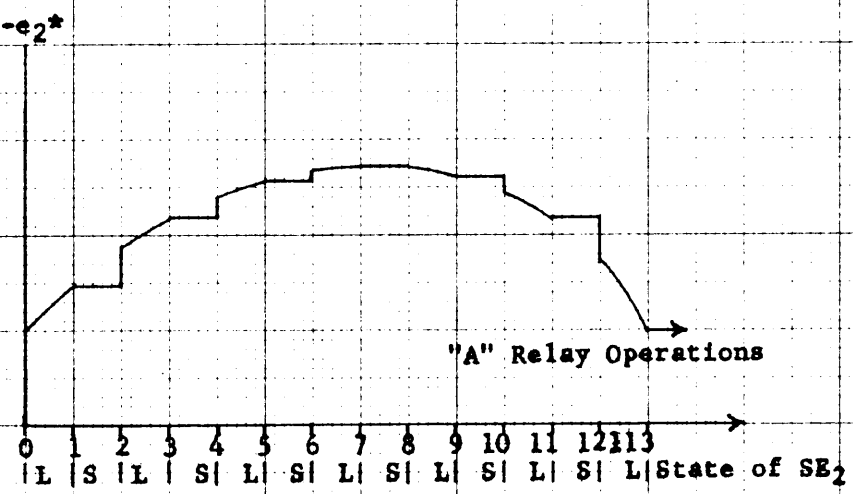


Figure 10e - Output of Storage Element Two

10d, and 10e. Note that the negative of the output voltages is plotted to enable easy comparison with the input voltage. Storage element one alternately stores and follows $e_{in}(t)$ as seen in Figure 10d. The output of the second storage element is similar to that of the first; however, close examination of the circuit and Figure 10e shows that during the time SE_1 is loading, and hence following, the input signal SE_2 is storing and vice versa. The ultimate output of the simulator is always switched to the storage element which is in the store condition resulting in the output shown in Figure 10c. Comparison of Figures 10a and 10c reveals that the conditions for a zero order hold are satisfied and, therefore, the circuit of Figure 9 can be used to simulate a zero order hold device on an analog computer. The output is simply a sample of the input signal and that sample is held until it is replaced by a more recent, and more correct, sample.

The parallel connected zero order hold simulator shown in Figure 9 was set up on a Heath Model ES-400 computer using the following values in the storage elements.

Input resistor (R_1)	100KΩ
Feedback resistor (R_f)	100KΩ
Feedback capacitor (C)	.01μf

The circuit was excited with a sinusoidal input voltage of

10.0 volt amplitude at a frequency of .10 cps which was sampled at a uniform rate of 1.0 samples/second. The actual input and the resulting output are shown in Figure 11. The time constant of the storage elements was .001 seconds which is much less than the sample period of 1.0 seconds, assuring accurate samples. Study of Figure 11 indicates that a reasonably good simulation of a zero order hold is obtained except that during the transition from one sample to the next the output voltage drops to zero very briefly. This period was measured to be about 1.2 milliseconds for the ES-400 computer and is due to the finite throwover time of the break before make "A" relay used at the output of the simulator.

In the case of the zero order hold simulator the relay throwover time would not, in general, introduce appreciable error in the simulation of a complete sampled-data control system. This is because, in a practical system, the relay throwover time would usually be small in comparison with the sample period and also because the time constants of the continuous portion of the system are large compared to the throwover time. It is conceivable, however, that unusual situations could occur where these conditions would not be satisfied and errors would result. Relay throwover time is of greater importance in the simulation of higher order hold

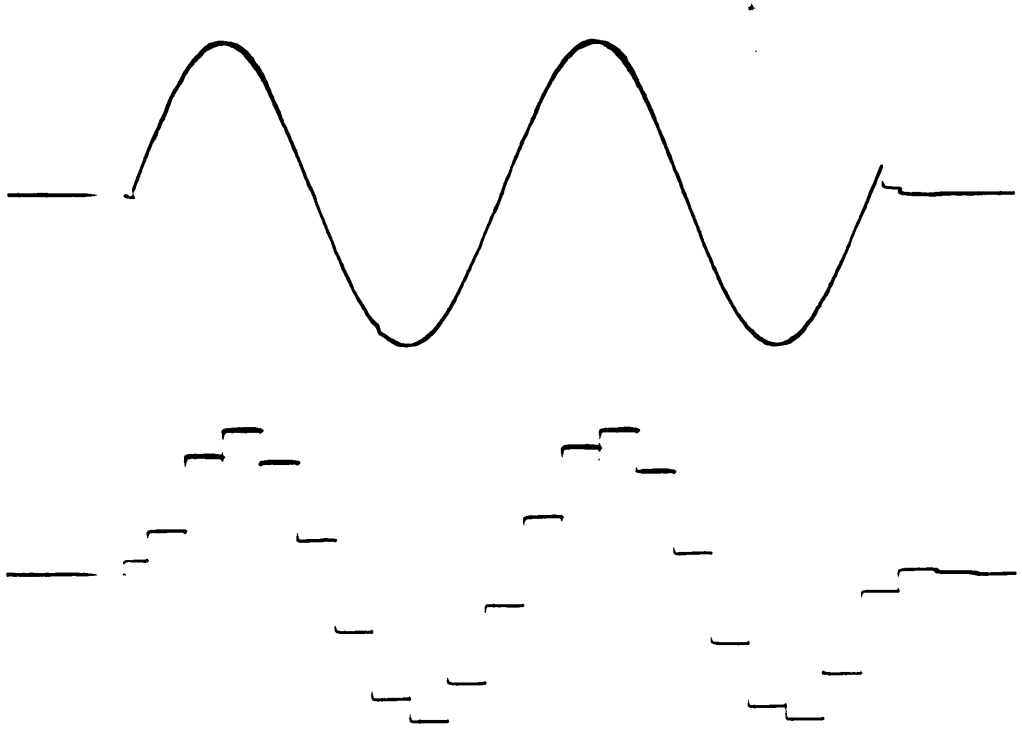


Figure 11 Output of Parallel Connected Zero Order Hold Circuit

$R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper Speed 5mm/second

devices and is treated in more detail below.

In order to properly simulate a hold device it is important that the relays associated with the storage elements be properly timed with respect to the sampled-data system being represented. They must, in addition, be properly synchronized with each other. In most practical sampled-data systems the sampling rate, and therefore the sample period, is uniform. In certain cases, however, the sample period may be random or follow some other non-uniform schedule. Only uniform sampling rates are considered in this report although most of the work could be adapted to the simulation of non-uniform rates.

As mentioned previously the Heath Model ES-400 analog computer was chosen for use in this investigation because of its versatility and particularly because all of the control relay contacts are terminated on the computer problem board and can be used in any desired manner. In order to explain the simulation techniques developed in this investigation an understanding of the control circuits of the ES-400 computer is essential. These are illustrated schematically in Figure 12.

As indicated in the figure there are a total of 16 SPDT contacts available on the computer. Eight of the contacts are associated with Relays "A" which can be energized by switch "A"

located on the computer control panel. In addition after, and only after, relays "A" are energized relays "B" can be energized by switch "B" operating the remaining eight contacts. Note that if switch "B" remains closed the "A" and "B" relays operate identically. In normal computer usage the relays "A" are used to set in initial conditions and relays "B" are used as hold relays. A remote control jack is in parallel with switch "A" in effect allowing that switch to be operated from a remote location. In normal usage two or more computers can be slaved together by connecting a two conductor cable between the remote control jacks on all computers. When slaved in this manner closure of switch "A" in any one computer will energize the "A" relays in all computers.

The contacts of relays "A" in the computer (Figure 12) served as the "A" relays in the storage elements (Figure 9) and likewise computer relays "B" served as the "B" relays. This is consistent with the symbols already used in Figures 6 and 9 and is followed throughout this paper. A careful study of the timing diagram (Figure 10) for the parallel connected zero order hold simulator (Figure 9) reveals that the "A" relays must change state at the sampling instant and also that the relay goes through one complete cycle in two sample periods. Also it is seen

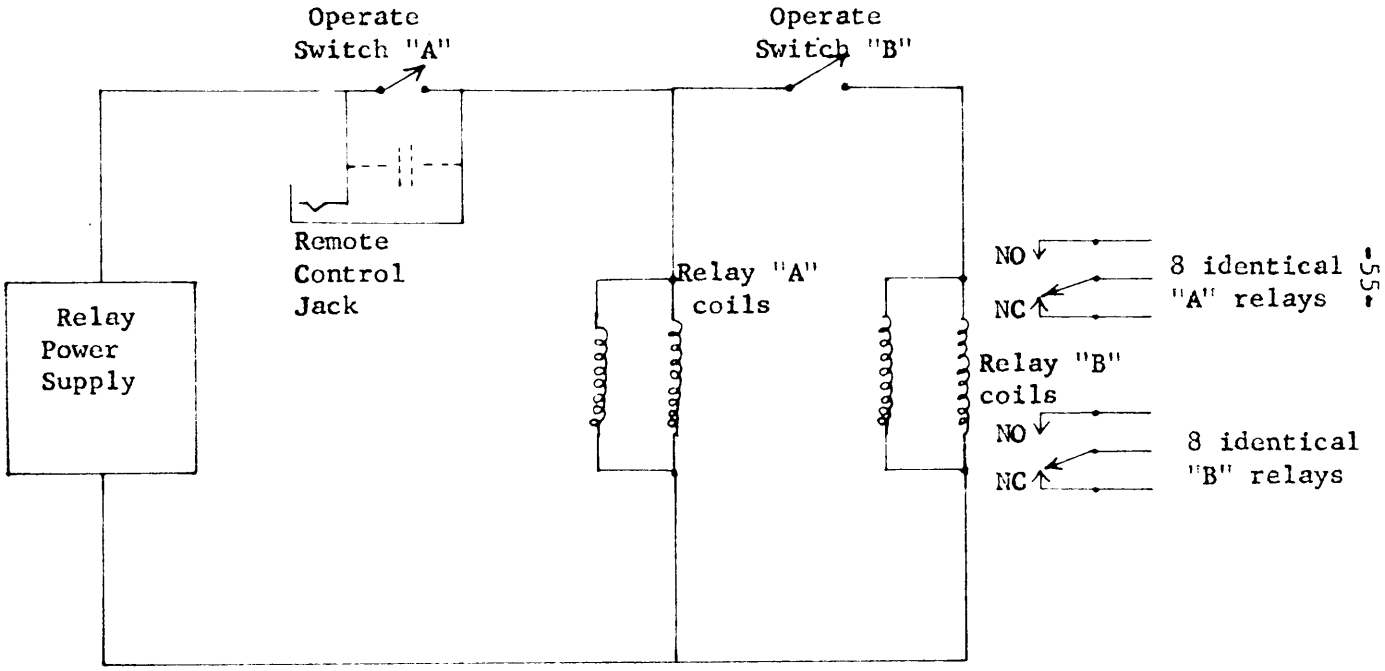


Figure 12

Control Circuits on Heath Model ES-400 Analog Computer

(Simplified)

that if the sample periods are to be equal, i.e., uniform sampling rate, the energized and de-energized periods of the "A" relays must be equal and furthermore the relays must cycle at a constant rate. This is equivalent to operating switch "A" on the computer on and off at a constant rate and with equal on and off times. This can be done automatically, however, through the use of a master timing oscillator. This oscillator operates a set of auxiliary contacts which, via the remote control jack, in turn operate the "A" relays. In a practical sampling situation the rate of operation of the relays would not exceed a few cycles per second. It is seen that the master timing oscillator would be of relatively low frequency for an electronic oscillator.

Although not the only possibility, a multivibrator circuit suggests itself as being suited for use as a timing oscillator for the following reasons:

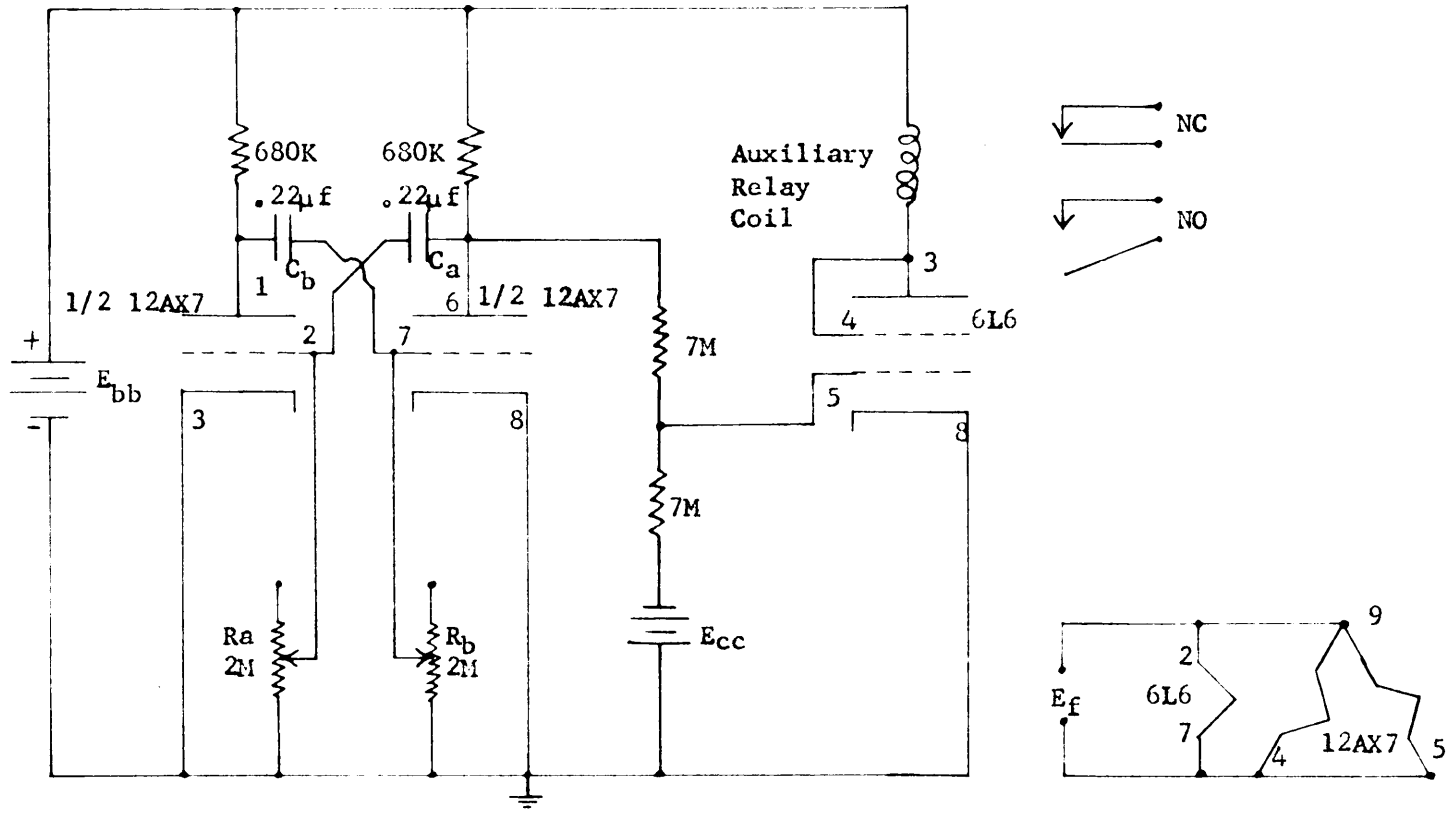
1. It can be made to operate at low frequencies with little difficulty.
2. Since the oscillator is to operate a relay the natural well defined two state operation of a multivibrator is desirable.
3. The time in each state of the auxiliary relay and thus the "A" relays is easily controlled.

The multivibrator circuit shown in Figure 13 was designed

and constructed for use as a master timing oscillator in this investigation. The two triode sections of the 12AX7 and associated components constitute a free running multivibrator circuit.[34] The plate output of one side of the multivibrator is direct coupled to a 6L6 beam power tube used as an amplifier. A relay coil, which operates a pair of auxiliary contacts, is placed in the plate circuit of the 6L6. A discarded telephone relay was adapted to serve as the plate relay. The amplifier stage was necessary to raise the power level of the multivibrator output sufficiently to drive the relay which was available. The bias voltage, E_{cc}, was adjusted to give positive relay action when the multivibrator changed states.

The time constants of the multivibrator, and thus its frequency, are determined by the products R_aC_a and R_bC_b, (see Figure 13). The time constants can be varied independently by the setting of R_a and R_b. If $R_a = R_b$ and all other circuit parameters are balanced the multivibrator will remain in each state an equal time. Also, if R_a and R_b are changed simultaneously the frequency will change, decreasing for increasing resistances. It is, of course, desirable to be able to change the multivibrator frequency because in the simulation of a sampled-data system this is equivalent to varying the sampling

Figure 13



Circuit Diagram of Master Timing Oscillator

E_{bb} = 250V DC
E_{cc} = 100V DC
E_f = 6.3V AC

rate.

Since R_a and R_b are individually variable it is possible to vary the time that the multivibrator remains in each state, however, for most applications it is desirable that the circuit remain in each state an equal time. The individual adjustments can be used to compensate for unequal values in corresponding components in each side of the multivibrator which would cause the circuit to remain in each state unequal amounts of time. This, of course, sacrifices the simplicity of a single adjustment to obtain better accuracy in the master timing oscillator.

A convenient method of setting R_a and R_b in the multivibrator is necessary. It is practical to operate a Veeder Root counter by a pair of contacts on the auxiliary plate relay. The number of cycles of the timing oscillator can be counted and timed with a stopwatch allowing the sampling rate to be computed. The counter will advance one count for each cycle of the multivibrator, corresponding to two samples for each count, in the parallel connected zero order hold simulator described above. By counting over a long period of time, typically in the order of one minute, the frequency of the oscillator can be set quite accurately, and if regulated plate and bias supplies are used to power the circuit, the frequency will remain resonably stable.

A more difficult problem, however, is to set R_a and R_b

such that the time in each state is equal. This can be done conveniently using the auxiliary computer circuit illustrated in Figure 14a. The +100 volt reference supply on the computer is fed to the armature of one of the "A" relays on the computer. The "A" relays are then operated by the master timing oscillator via the remote control jack as previously described. Thus the +100 volts alternately appears at the normally open and normally closed contacts of the relay but the voltage at one of these contacts is changed to -100 volts by an inverting amplifier. The resulting pulsating wave trains are then summed and integrated by a second operational amplifier and the results read on the computer output meter. The waveform resulting after summation, but prior to integration, is shown in Figure 14b. Study of this waveform reveals that if the time in each state is equal, i.e., $T_1 = T_2$, the output of the integrator will remain constant but otherwise it will drift. Using this circuit the multivibrator adjustments can be set such that the output voltage on the integrator remains fixed while at the same time the oscillator is operating at the desired frequency as indicated by the counter. There is, unfortunately, some interaction between these adjustments and considerable care must be taken to obtain the desired combination of frequency and timing.

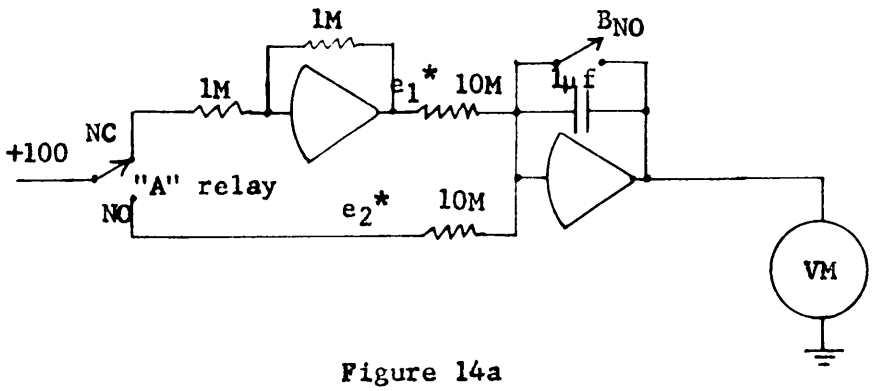


Figure 14a

Circuit for Adjustment of Master Timing Oscillator

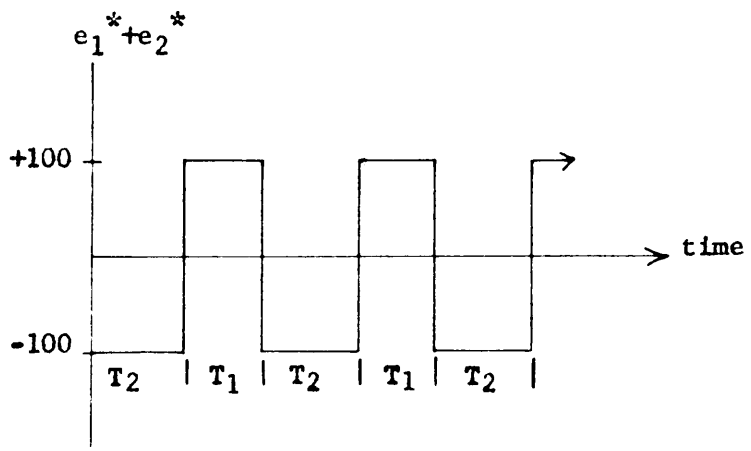


Figure 14b - Pulsating Wave e_1*+e_2*

The master timing oscillator which was constructed for use in this investigation, had the controls for R_a and R_b brought out to arbitrary calibrated dials. This enabled calibration curves to be drawn for the unit. One such curve is shown in Figure 15 for the particular case of symmetrical operation ($T_1 = T_2$), the case of most interest. The proper setting for each dial can be read to give any desired sampling rate.

An unexpected difficulty was encountered in the simulation of the zero order hold described above as well as in other sampling circuits which were attempted. In normal computer usage the computer relays would only be operated at the start and stop of a problem. However, in a sampled-data simulation problem the relays operate continuously while the problem is running. It was found that transient voltages were introduced into the problem circuitry each time the relays operated causing small errors to occur in the problem solutions. It was found that these transients could be greatly reduced by placing a small capacitor across the remote control jack, which is electrically in parallel with the auxiliary relay contacts. This capacitor is indicated by dotted lines in Figure 12. It is cautioned, however, that if the capacitor is made too large it will introduce a time lag between

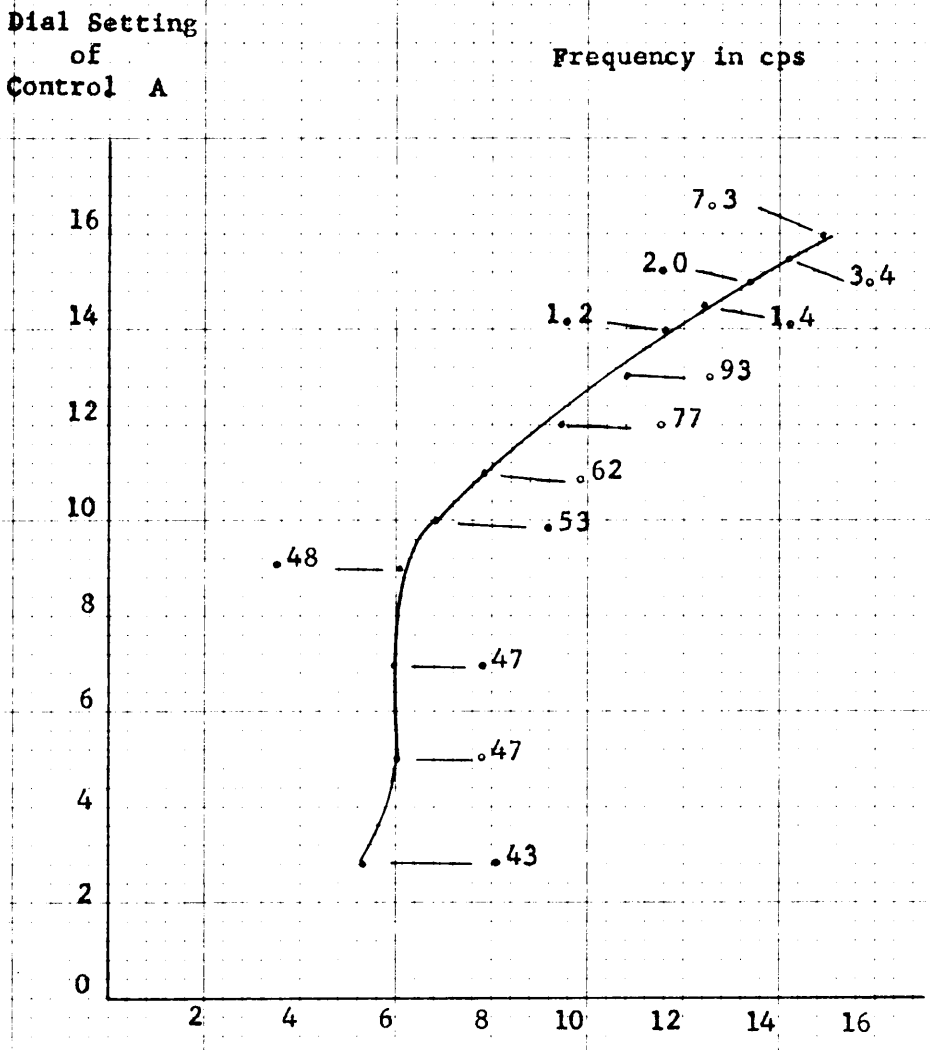


Figure 15

Calibration Curve for Master Timing Oscillator

(Symmetrical Operation)

the opening of the auxiliary relay on the timing oscillator and the release of the "A" relays. A .01μf capacitor was found to perform satisfactorily.

Referring to the circuit diagram of a storage element, Figure 6a, it has been mentioned that the element can be reset to the value E_o by closure of contacts "B". On the ES-400 computer a pair of normally open contacts on Relay "B" can be used for this purpose. If the "A" Relays were cycling, or if switch "A" were closed, and switch "B" were then closed the voltage E_o would be set into the storage elements. On the ES-400 an initial condition floating power supply would be used for E_o. In most practical cases it is desired to set the storage elements to zero so that all that is necessary is to let the "B" contacts simply short out the capacitor eliminating the need for a floating power supply.

Upon further consideration it is seen that, provided the initial condition of the storage element is to be set to zero, no special provision need be made for setting initial values and the "B" contacts can also be eliminated. All that is required is to start the "A" relays cycling before starting the remainder of the system being simulated. If under these conditions the input to the hold simulator is zero the storage elements will automatically be set to zero. For the parallel connected zero order

hold simulator of Figure 9 it will require at most two sample periods to set all storage elements to zero, after which the remainder of the system can be started and the actual problem solution begun. Unless otherwise stated initial conditions were set in this manner in all of the work described below.

There are certain disadvantages associated with zero order hold simulation, using parallel connected storage elements. Probably the least desirable feature is that the relay operation must be very closely timed if uniform sampling periods are to be represented. As a result of this considerable effort was expended to insure that the multivibrator remains in each state an equal amount of time. Another undesirable feature is the fact that momentarily the output of the simulator becomes zero, during the relay throwover time, introducing errors into the automatic control system being simulated. The possibility of filtering the output of the simulator suggests itself and was investigated. Either a low pass filter using passive elements or one using an operational amplifier connected as a first order lag circuit could be used (see Figure 3 and equation 24). Unfortunately the filtering reduces the response of the simulator in addition to filtering out the spike due to relay switching time; however, in many cases a reasonable compromise between the

two effects can be attained. The effect of relay throwover time is more pronounced in the simulation of higher order holds and is treated in more detail in the discussion of those simulations.

Zero Order Hold Simulation Using Series Connected Storage Elements.

Most of the difficulties encountered in the simulation of zero order hold devices with parallel connected storage elements can be eliminated by connecting the storage elements in series. A schematic diagram for a zero order hold simulator using this system is shown in Figure 16a and the corresponding block diagram in Figure 16b. For certain types of simulation, however, other problems are introduced by the series connected circuits.

The operation of the series connected circuit is explained with the aid of the timing diagrams in Figure 17. Figure 17a shows an arbitrary input voltage $e_{in(t)}$. For purposes of explanation it is assumed that just prior to t = 0 the output voltage of both storage elements has been set to zero by closure of the "B" relays. It is also assumed that the "A" relays are cycling and remain in each state for an equal period of time. At t = 0 the "B" relays are de-energized and remain so, and the "A" relays, in the cycling process, have just switched to the de-energized or

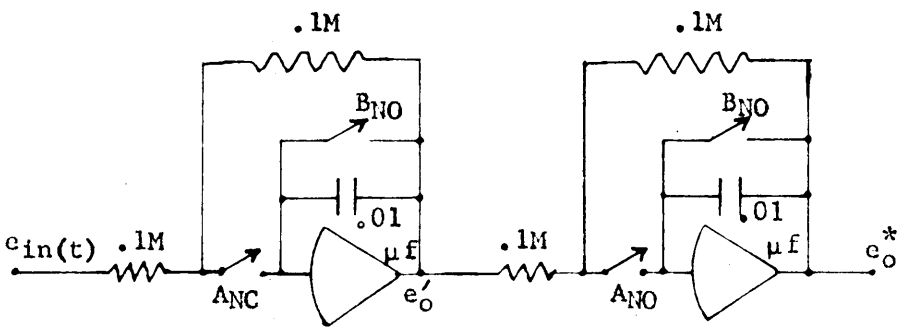


Figure 16a - Circuit Diagram of Series Connected Zero Order Hold Simulator

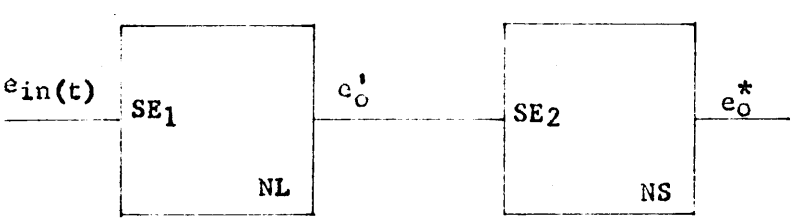


Figure 16b

Block Diagram of Series Connected Zero Order Hold Simulator

normal state.

The output of the first storage element is shown in Figure 17b. The negative of the output is shown for purposes of comparison. When the "A" relays are in their normal (de-energized) condition, the output of storage element one is following the input signal, $e_{in(t)}$, except for sign. However, after the "A" relays become energized the output is held constant at the value existing at the instant the relays changed state. Ideally the output again instantly assumes a value equal to the magnitude of the input voltage when the "A" relays return to the de-energized state. In practice, however, this readjustment takes a short time because of the finite time constant associated with the storage element as discussed previously. The output of the first storage element, e_o', is actually a quasi-sampled signal.

The second storage element is performing the same function as the first with two exceptions:

1. The second storage element is always in the opposite state from the first, however, the two elements change states at the same time.
2. The input to the second storage element is a quasi-sampled rather than a continuous signal.

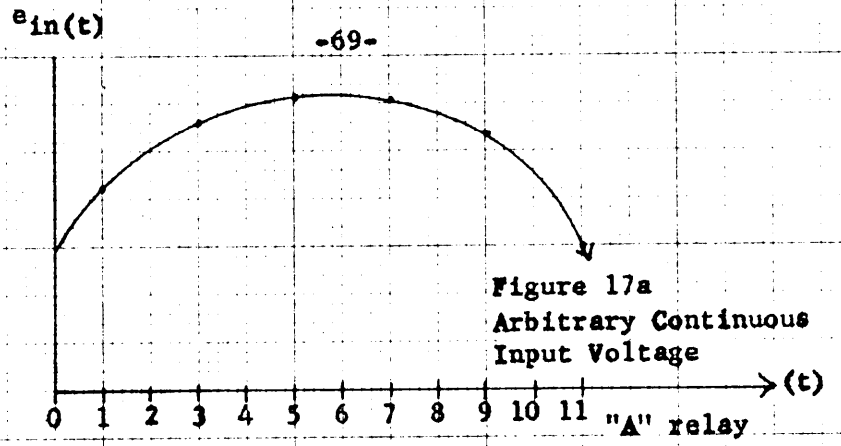


Figure 17a
Arbitrary Continuous Input Voltage

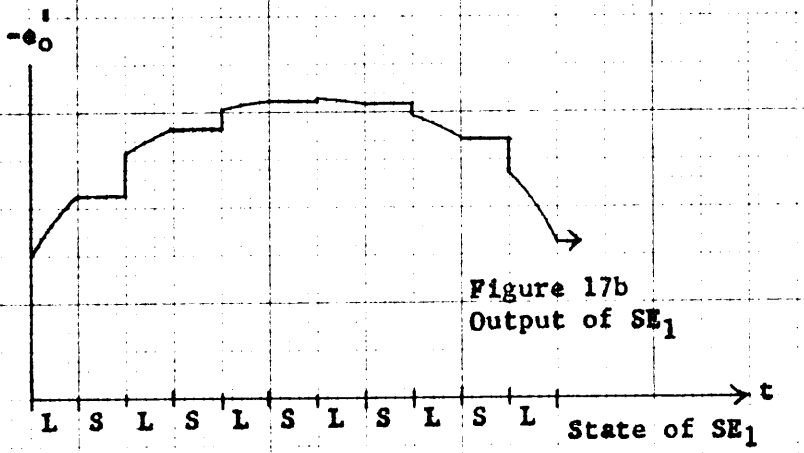


Figure 17b
Output of SE_1

L = Load
S = Store

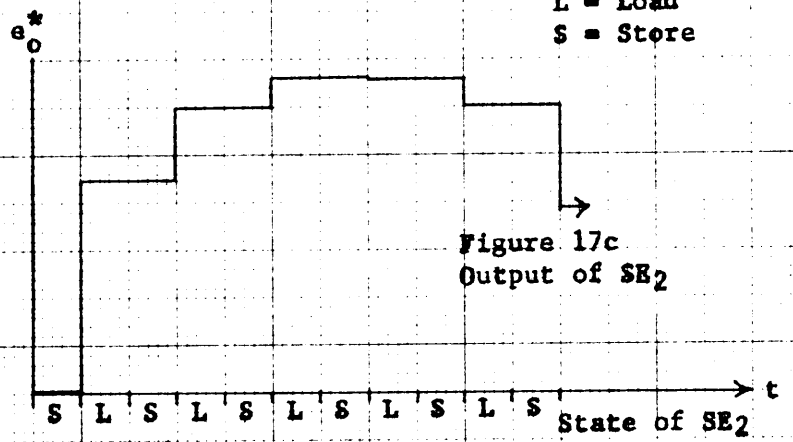


Figure 17c
Output of SE_2

Figure 17

Timing Diagram for Series Connected Zero Order Hold Simulator

The voltage resulting at the output of the second storage element is illustrated in Figure 17c. It is seen that this output, e_1^*, is true sampled data with the first sample being taken at the first operation of the "A" relays (t = 1). The voltage e_1^* is held constant at the value of e_{in} existing at t = 1 from t = 1 until t = 3 at which time it is replaced by the value of e_{in} at t = 3 which is then held until t = 5 and so on. The output of storage element two can be thought of as a series of succssive samples of $e_{in(t)}$, which is the requirement for a zero order hold simulation circuit. Close examination of Figure 17 reveals that actually the sample that starts at a given time, say t = 1, is the value of $e_{in(t)}$ just previous to t = 1. This, of course, is because of the finite operating time of the relays and the finite time constants of the storage elements.

The series connected zero order hold simulator shown in Figure 16 was set up on a Heath Model ES-400 analog computer using the same component values in the hold circuits that were used previously. The circuit was excited with a sinusoidal input voltage of 10.0 volts amplitude at a frequency of .10 cycles per second which was sampled at a uniform rate of 1.0 sample per second. The actual input and resulting output

signal are shown in Figure 18. The time constant of the storage elements was .001 second, which is much less than the one second sample period, assuring reasonably accurate samples. Close inspection of the curves shows that the circuit of Figure 16 is a reasonably good simulation of a zero order hold device.

To a first approximation the effect of the series connected simulator of Figure 16 on the input signal is identical to the effect obtained with the parallel connected zero order hold simulator shown in Figure 9. Closer examination of the output voltages reveals, however, that they differ in two important respects which, for most purposes, makes the series simulator superior.

1. In the zero order hold simulator using parallel connected storage elements the output voltage momentarily goes to zero during the throwover time of the output relay. The effects of this have been discussed previously and were shown to be undesirable. Inspection of the timing diagrams for the series simulator, Figure 17, indicates that a similar condition cannot exist with series connected storage elements. During the transient between samples the output voltage e_o^* will be in error in that it will be neither at the level of the new or the old sample but

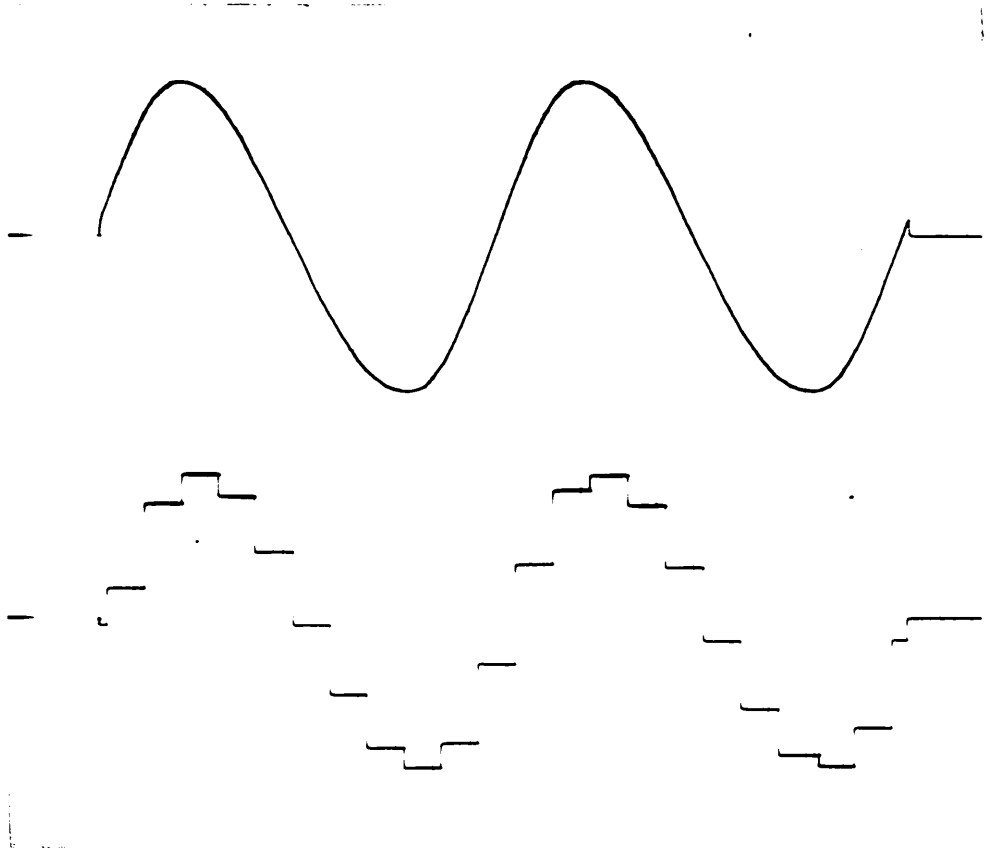


Figure 18 Output of Series Connected Zero Order Hold Circuit

$R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper Speed 5mm/second

will be somewhere between the value of these samples. This, of course, is far more desirable than having the voltage level fall to zero momentarily. Unless the sampling rate is extremely fast the transient time between samples is an insignificant portion of the sample period and introduces negligible error into the overall simulation.

2. The operation of the series connected zero order hold simulator (Figure 16) is such that the "A" relays go through one complete cycle in representing one sample period. In the parallel connected simulator (Figure 9) two sample periods are represented by each cycle of the "A" relays and furthermore, if uniform sampling rate is to be simulated, considerable effort is required to get the master timing oscillator properly adjusted. Therefore, in the series simulator the "A" relays cycle at twice the rate that they would in the parallel circuit while representing the same sampling rate.

An important feature of the series simulator is the fact that it is not necessary that the "A" relays remain an equal time in each state in order to simulate a uniform sampling rate as was required in the parallel simulator. To simulate a uniform sampling rate with this circuit it is necessary only that the "A" relays

cycle at a constant frequency but the duration in each state can be somewhat unequal, the amount depending on the sampling rate. This has major practical advantages in that the adjustment of the master timing oscillator is made considerably easier. It is also noted that the counter on the oscillator will now count samples directly rather than samples divided by two as in the parallel case.

In the series connected zero order hold simulator two transients occur during each sample period, one in each storage element. Both of these transients must have sufficient time to decay before the "A" relays change state or magnitude errors will exist in the sampling process. The maximum sampling rate that can exist in the computer is, therefore, limited by the finite switching time of the relays, the electrical time of the relays and the electrical time constants of the storage elements. In the parallel connected simulator these transients occur only once each sample period, and also for a given sampling rate, the computer relays operate at half the frequency that they would in the equivalent series simulator. For these reasons the parallel simulator may have some advantages over the series simulator in representing high sampling rates. For most applications, however, the series simulator is superior.

The "B" relays are used in the series simulator of Figure 16 in the same manner as they are used in the parallel simulator of Figure 9, that is to set the initial voltages of the storage elements prior to starting the problem. The process is identical to that previously discussed for the parallel simulator and the same limitations and simplifications exist. If, in the series simulator, the storage elements were initially set to some voltage prior to starting the problem and the "A" relays are initially de-energized, the first sample of $e_{in(t)}$ would not be taken until the instant that the "A" relays become energized. In general however, this is of no practical consequence. In the usual situation it is desired to set all of the storage elements to zero prior to starting the simulation problem. It is, however, not necessary that "B" relays be used to accomplish this. If the "A" relays are started cycling prior to the start of the problem and the input voltage to the simulator is made zero all storage elements will automatically be set to zero after one complete cycle of the "A" relays. In a practical simulation problem it is usually convenient to do this allowing the "B" relays to be eliminated from the circuit.

First Order Hold Simulation Using Parallel Connected Storage Elements

The simulation of some types of sampled-data systems on the analog computer may require the representation of a first order hold device. As previously stated the output of such a hold device is in the form of a constant quantity, the value of which depends on the most recant sample of a continuous input signal as well as the previous sample. The output of the hold device changes each time the value of the samples change but remains constant between sampling times. For the case of the first order hold device equation 28 reduces to

$$Y^* = AX_0^* + BX_1^* \quad . \tag{30}$$

Since in the analog computer all of the variables are represented by voltages, a simulator for such a device must obey the relation

$$e_{OUT}^* = A e_0^* + B e_1^* \tag{31}$$

which is, of course, a special case of equation 29.

In order to mechanize Equation 31 on the analog computer it is necessary to store or remember the previous sample. This is a process normally associated with digital computers. It is possible, however, to simulate a first order hold device on an

analog computer by an extension of the methods previously described for simulating a zero order hold device. This again can be done using storage elements connected in series or in a series of parallel connected pairs. In most respects, the series connection was shown to be superior.

The circuit shown in Figure 19 consisting of a series of two pairs of parallel connected storage elements can be used to simulate a first order hold device. The combination of SE_1, SE_2, and A_{1a} considered together constitute a parallel connected zero order hold simulator as illustrated in Figure 9, the operation of which has been described. An arbitrary continuous input signal, $e_{in}(t)$, is shown in the timing diagram of Figure 20a. It is assumed that the storage elements are initially set to zero and, in the process of cycling, the "A" relays are initially de-energized. The output of the first pair of storage elements, e_o^*, is shown in Figure 20b. Note that e_o^* is identical to the output of a parallel connected zero order hold simulator (see Figure 10c) and is seen to be a sample of the continuous input $e_{in}(t)$, taken at the end of the previous sample period. In order to facilitate easy comparison with the input voltage $-e_o^*$ is plotted rather than the quantity itself.

The output of the first pair of storage elements is fed into

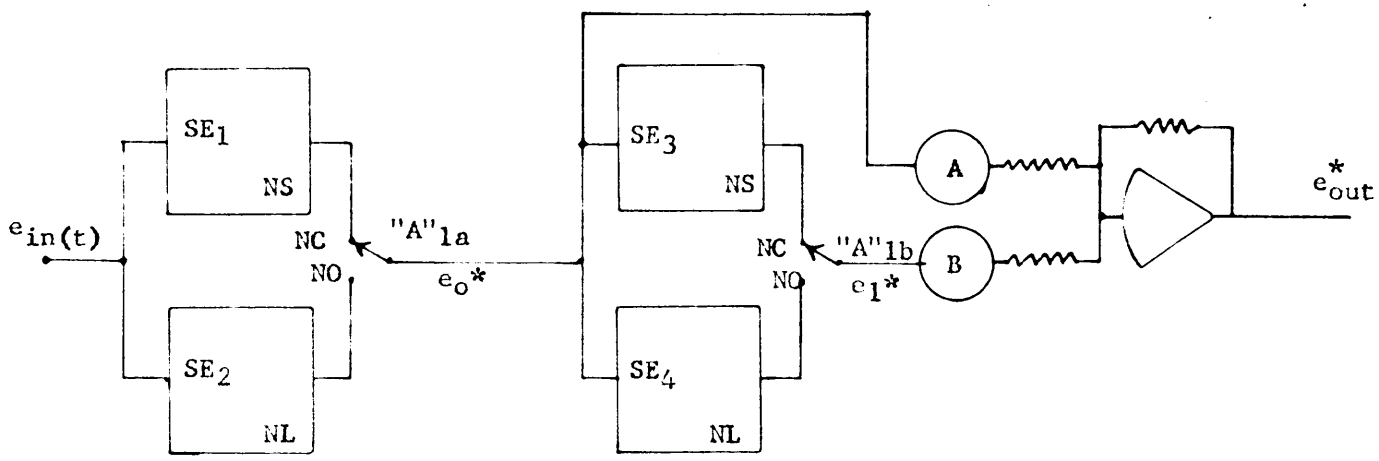


Figure 19

First Order Hold Simulator Using Parallel Connected Storage Elements

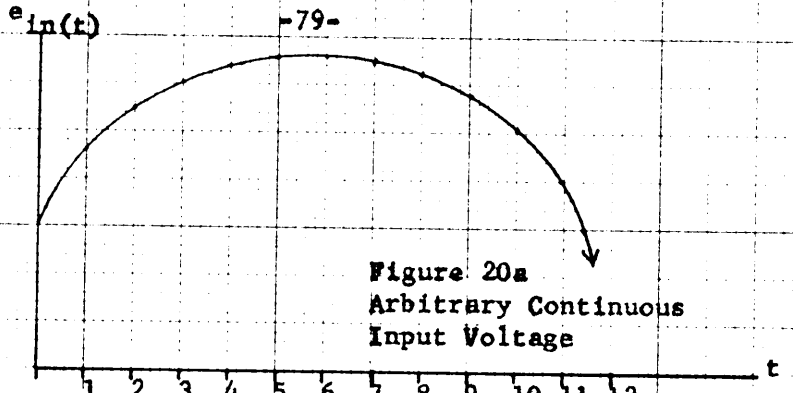


Figure 20a
Arbitrary Continuous Input Voltage

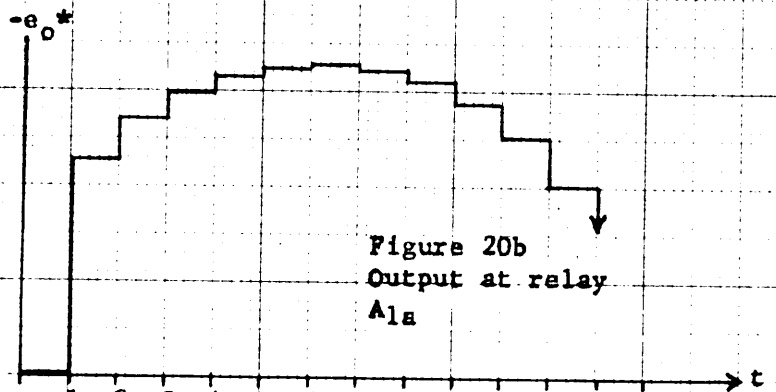


Figure 20b
Output at relay A_{1a}

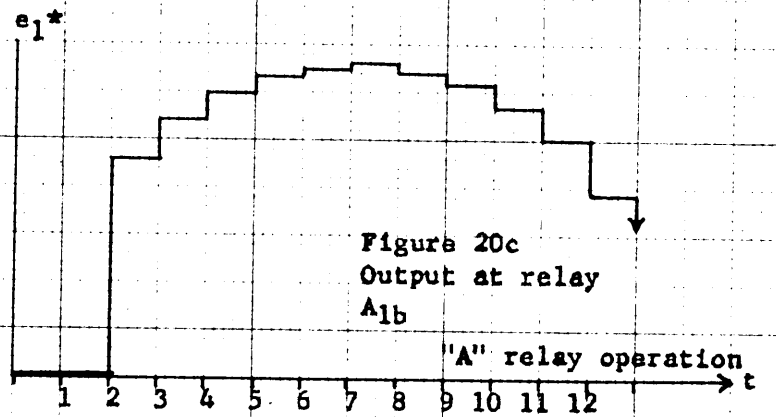


Figure 20c
Output at relay A_{1b}

Figure 20 Theoretical Timing Diagram for First Order Hold Simulator Using Parallel Connected Storage Elements.

another identical pair of storage elements, SE_3 and SE_4. The voltage ideally resulting at the output of the entire simulator circuit, e_1^*, is shown in Figure 20c. The voltage can be thought of as a sample of the output of the first pair of storage elements, e_o^*, which is itself a sample. Relative to the input signal $e_{in(t)}$ the output e_1^* is the value of $e_{in(t)}$ at the start of the previous sample period. For example, the output e_1^* existing from t = 3 to t = 4 is actually the value of $e_{in(t)}$ existing at t = 2. The voltages e_o^* and e_1^* can be multiplied by constants, the signs changed as necessary, and then added by a summing amplifier as shown in Figure 19. The output of the summing amplifier is described by equation 31 so it is concluded that the circuit illustrated in Figure 19 can, at least ideally, simulate the action of a first order hold device.

Certain practical problems, however, make the parallel representation of a first order hold device difficult to use. The most obvious problem is the difficulty of adjusting the master timing oscillator to cause the "A" relays to remain an equal time in each state, that being a necessary condition for the simulation of a uniform sampling rate. This situation is identical to that encountered in the zero order hold simulator of Figure 9 and the method of adjustment, using an auxiliary

computer circuit, is the same as that described for the zero order circuit. Two additional difficulties are encountered when the storage elements used have very short time constants.

The input voltage to storage elements SE_3 and SE_4 (see Figure 19) briefly drops to zero during the throwover time of relay "A_{1a}". Ideally the "A" relays in the storage elements will switch in unison with relay "A_{1a}" and all of the storage elements will be in the store condition during the throwover period. If the relay contacts do not break simultaneously it is possible that the output voltage of the storage element which is switching from the load to store state might drop somewhat before its "A" relay opens resulting in an erroneous sample. If the time constant of the storage elements is long compared to the relay throwover time satisfactory operation will result. Storage element time constants of .001 seconds or greater were found to give satisfactory results. Another solution is to insert a low pass filter, either passive or active, after relay "A_{1a}".

Another more pronounced difficulty was observed when the outputs of SE_1 and SE_2 were fed through relay "A_{1a}" into a second pair of storage elements. It was observed that with $e_{in(t)}$ constant the outputs of SE_1 and SE_2 were of different

levels and also that the levels changed after each change of state of the "A" relays. These uneven voltages were, of course, passed on to SE_3 and SE_4 resulting in unsatisfactory operation of the overall simulator. When the load, SE_3 and SE_4, was disconnected at relay "A_{1a}", satisfactory operation of SE_1 and SE_2 was observed. When larger feedback capacitors were used in these storage elements the performance was improved although increasing the value of the resistors resulted in no change in performance. It is thought that this phenomena is a result of the finite frequency response of the operational amplifier.

The load on each of the first pair of storage elements alternates between 100KΩ and infinite resistance as relay "A1a" cycles. Theoretically, the output voltage of the storage elements should remain constant when the element is in the store condition even though the load on the amplifier changes. In this circuit the load is suddenly applied causing the output voltage and the summing function voltage to drop. Because of the finite frequency response of the operational amplifier the drop in summing junction voltage does not result in the instantaneous compensating rise in output voltage normally expected. In the meantime during the transient the capacitor discharges somewhat so that

the steady state voltage eventually reached is less than that existing before the load was applied. Since it is unlikely that the transient response in any pair of amplifiers would be identical the resulting output levels on the different storage elements would not be the same, causing the voltage at the output of the relay "A_{1a}" to alternate in level.

The best solution to the problem of erratic output voltages was obtained by increasing the value of the capacitors in the storage elements subjected to switching loads. The larger capacitor cannot change voltage as fast during the transient time resulting in a smaller voltage drop. A .1μf capacitor performed reasonably well but a .01μf capacitor introduced considerable error. An alternate solution is to increase the value of the resistors in the storage elements resulting in less loading of the preceeding storage elements. In either case the time constant of the storage element is increased causing some deterioration in the performance of the storage elements as sampling devices. Perhaps another possible solution would be to use a small choke in series with the armature of relay "A_{1a}" to prevent the load current from increasing too suddenly.

First Order Hold Simulation Using Series Connected Storage Elements

Many of the disadvantages of the first order hold simulator using parallel connected storage elements (Figure 19) can be overcome by using the storage elements connected in series such as illustrated in Figure 21. Comparison of Figures 16 and 21 reveals that the first order simulator essentially consists of two of the zero order circuits in tandem.

The operation of the series connected circuit of Figure 21 can be explained using the theoretical timing diagrams shown in Figure 17 and Figure 22. In these diagrams it is assumed that all storage elements are initially set to zero. It is, furthermore, assumed that at t = 0 the cycling "A" relays have just become de-energized and therefore each storage element is in its normal condition. An arbitrary continuous input signal $e_{in(t)}$ is shown in Figures 17a and 22a. The first two storage elements, SE_1 and SE_2, constitute a series connected zero order hold simulator. The operation of this circuit has been discussed previously and the voltage appearing at the output of SE_2 is shown in Figure 17c and repeated for convenience in Figure 22b. This output represents samples of the continuous input signal $e_{in(t)}$.

The theoretical outputs of storage elements SE_3 and SE_4 are

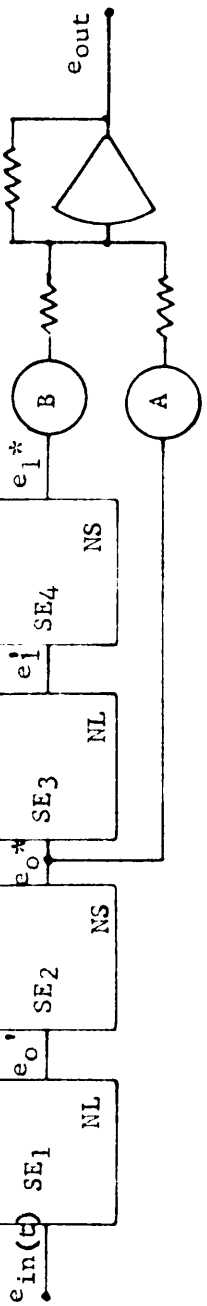


Figure 21 First Order Hold Simulator Using Series Connected Storage Elements

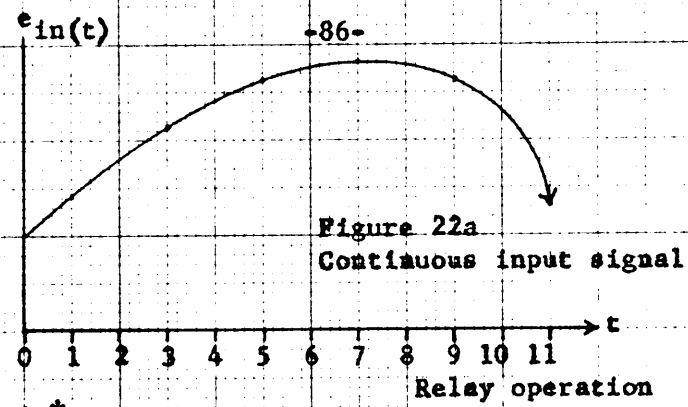


Figure 22a
Continuous input signal

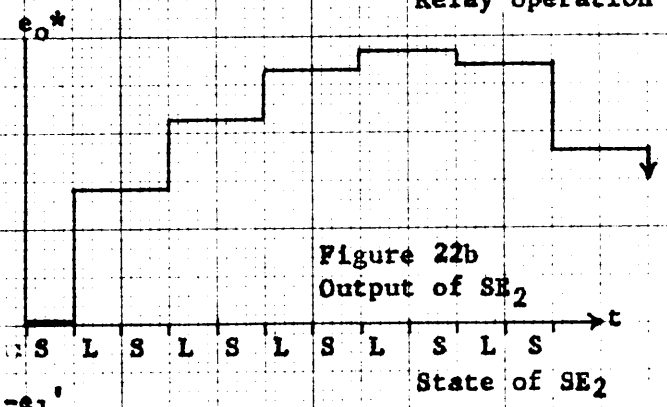


Figure 22b
Output of SE_2

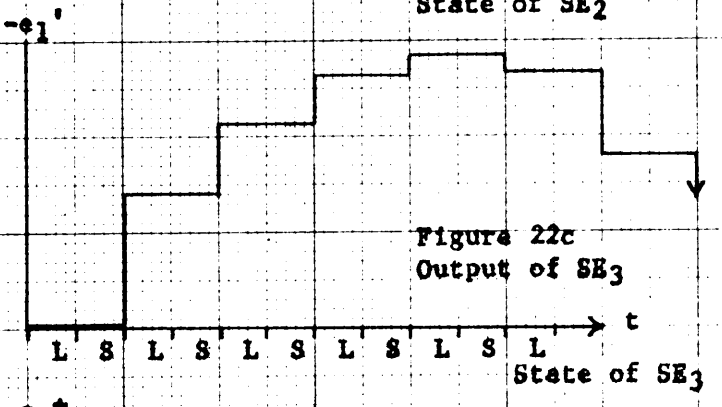


Figure 22c
Output of SE_3

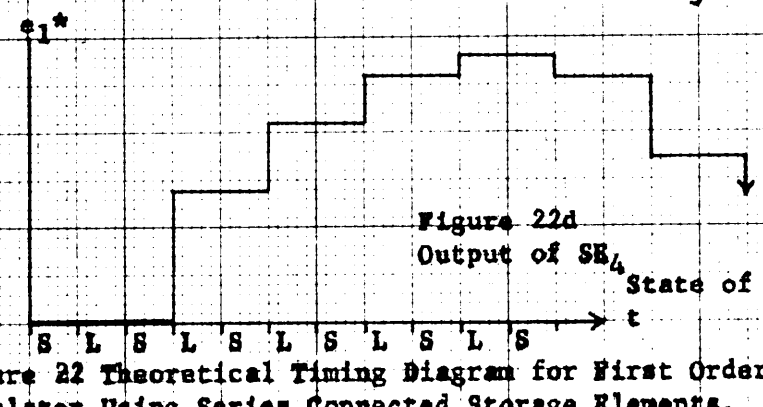


Figure 22d
Output of SE_4

Figure 22 Theoretical Timing Diagram for First Order Hold Simulator Using Series Connected Storage Elements.

shown in Figures 22c and 22d respectively. The function of SE_3 is to sample and store the output of SE_2 at some time between sampling instants. The output of SE_3 thus represents a delayed sample of the input voltage $e_{in}(t)$. The delay time is less than one sample period and ideally is one half sample period. The output of SE_4 is essentially a sample of SE_3 taken at exactly the instant $e_{in}(t)$ is resampled. The overall effect of the four series connected storage elements is that of a digital delay line in which $e_{in(t)}$ is sampled at regular intervals and the sample is transferred from one storage element to another along the delay line followed by other later samples. The delay line is said to be digital in nature because the samples move in discrete steps down the line. Study of Figure 22 indicates that the output of SE_2 is essentially a periodic sample (e_o^*) of $e_{in}(t)$ and the output of SE_4 is the previous sample (e_1^*).

The voltages representing the present and past samples can be multiplied by weighting constants, signs changed if necessary, and summed as indicated in Figure 21. The voltage appearing at the output of the summing amplifier can be expressed mathematically as follows:

$$e_{out}^* = A e_o^* + B e_1^* \qquad (32)$$

This is the same as equation 31 and is the special case of equation

29 for the first order hold device. Since these voltages are analogous to variables in a sampled-data system it is concluded that the circuit of Figure 21 is an alternate method of simulating a first order hold device on the analog computer.

An interesting feature of the series connected storage elements is evident from the ideal timing diagrams of Figure 22. It is noticed that a sampled signal is delayed by an amount exactly equal to one sample period in passing through each pair of series connected storage elements. This fact is of practical significance in some of the more involved simulation processes to be described below.

In the series first order hold simulator the "A" relays go through one complete cycle each sample period. This is different from the parallel connected simulator of Figure 19 in which the "A" relays go through one complete cycle in two sample periods. In the series simulator it is not important that the relays remain in each state an equal amount of time as long as the time in each state is long compared to the time constant of the storage elements. This, of course, makes adjustment of the master timing oscillator somewhat less difficult. In some of the more advanced simulation processes described below there is some advantage to operating the relays such that the period in each state is equal to the sample period. Under

such conditions the parallel connected simulator may be more desirable than the series connected circuit.

As was true in the simulation of zero order hold devices the difficulties associated with the relay throwover time in the parallel connected circuit do not appear with the series connection. None of the storage element voltages can change in the series circuit during throwover time since all elements are in the store condition until the relays have made positive closure.

The curves of Figure 23 illustrate the performance of an actual series connected first order hold simulator such as that shown in Figure 21. Shown on these recordings are the original .10 cps sinusoidal input voltage, $e_{in(t)}$, the most recent sample e_o^*, and the previous sample e_1^*. The sampling rate is 1.0 sample/second and the time constant of all storage elements is .001 seconds. Note that there is no noticeable error introduced by the storage element time constants and that the curves are free from transients during intermediate switching periods. The curves are exaggerated in order to illustrate the sampling process, since in most practical sampled-data systems the frequency of the input signal would be much less than the sampling frequency. From the curves of Figure 23 it is seen that the simulation of the delay line is reasonably satisfactory. The voltages shown, e_o^* and e_1^*, could have been combined with a

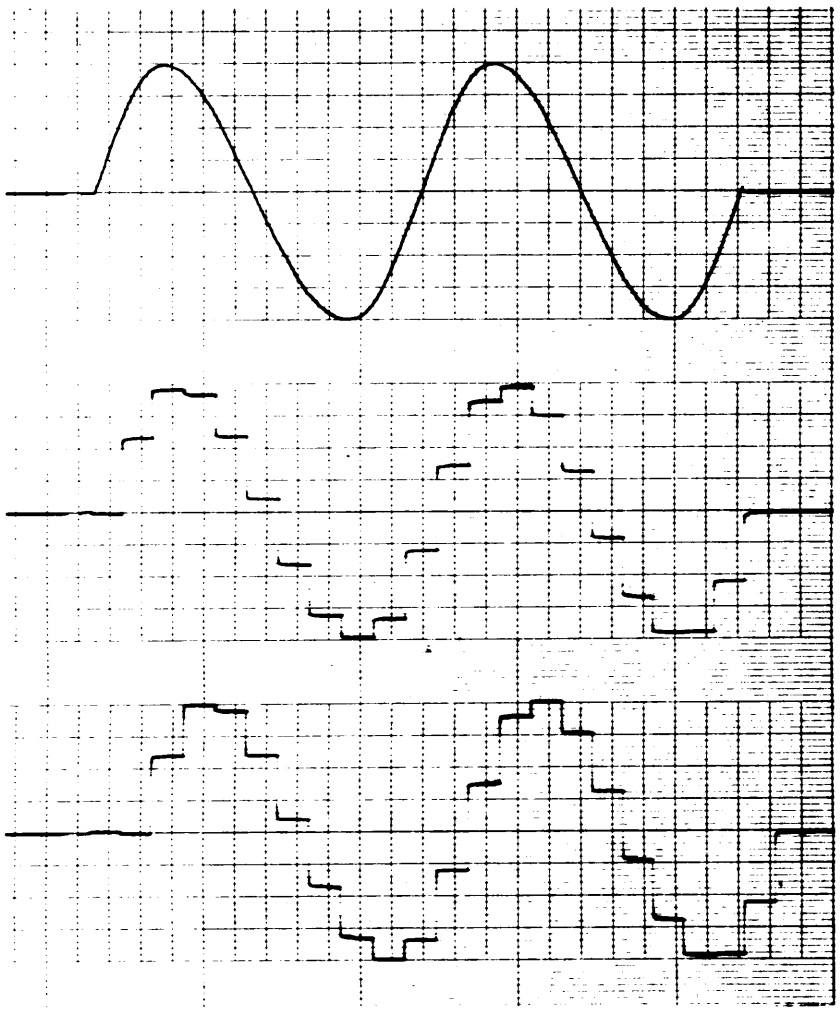


Figure 23 Outputs of Series Connected First Order Hold Circuits

$R_f = R_i = 100K\Omega$ $C = .01\mu f$ 1.0 sample/second

Paper speed 5mm/second

summing amplifier to give the complete simulation of the first order hold device.

In the simulation of first order hold devices it is sometimes necessary to set the output voltage of each storage element prior to running the problem. In general, it is possible to set the initial value of each storage element using floating power supplies and the "B" relays as indicated in Figure 6a. If it is desired only to set the initial voltages to zero this can be accomplished by simply forcing the input voltage of the delay line $e_{in(t)}$ to zero and cycling the "A" relays. For the first order hold simulator this will require two complete cycles of the relays after which the system problem can be started by allowing $e_{in(t)}$ to assume its normal value.

Second Order Hold Simulation Using Parallel Connected Storage Elements

The results from a search of the literature indicate that there has been nothing published describing the simulation of a second order hold device on an analog oomputer. The methods for simulating a first order hold device can be extended to provide simulation of second order hold devices. To enable some types of sampled-data systems to be represented on the computer it is necessary to simulate second and sometimes higher order hold devices.

As stated previously the output of a second order hold device depends on the most recent sample of the continuous input data and the two previous samples. The output remains constant during the sample period but changes each time a new sample is taken. It is seen that the output itself is discontinuous and appears in the form of a sampled quantity. In the special case of a second order hold the equation 28 becomes

$$Y^* = AX_0^* + BX_1^* + CX_2^* . \qquad (33)$$

In the analog computer, however, all variables appear as voltages so the simulation circuit for a second order hold device must obey the relation

$$e_{OUT}^* = Ae_0^* + Be_1^* + Ce_2^* \qquad (34)$$

this being a special case of equation 29. As might be expected from the simulation of zero and first order hold devices a second order hold can be simulated either by using a series of storage elements connected in parallel pairs or simply connected in series.

A second order hold simulator using parallel connected storage elements is shown in Figure 24 and its operation illustrated by the theoretical timing diagram of Figure 25. The operation of the device can be considered in three parts. The first pair of storage elements, SE_1 and SE_2, sample the continuous input voltage

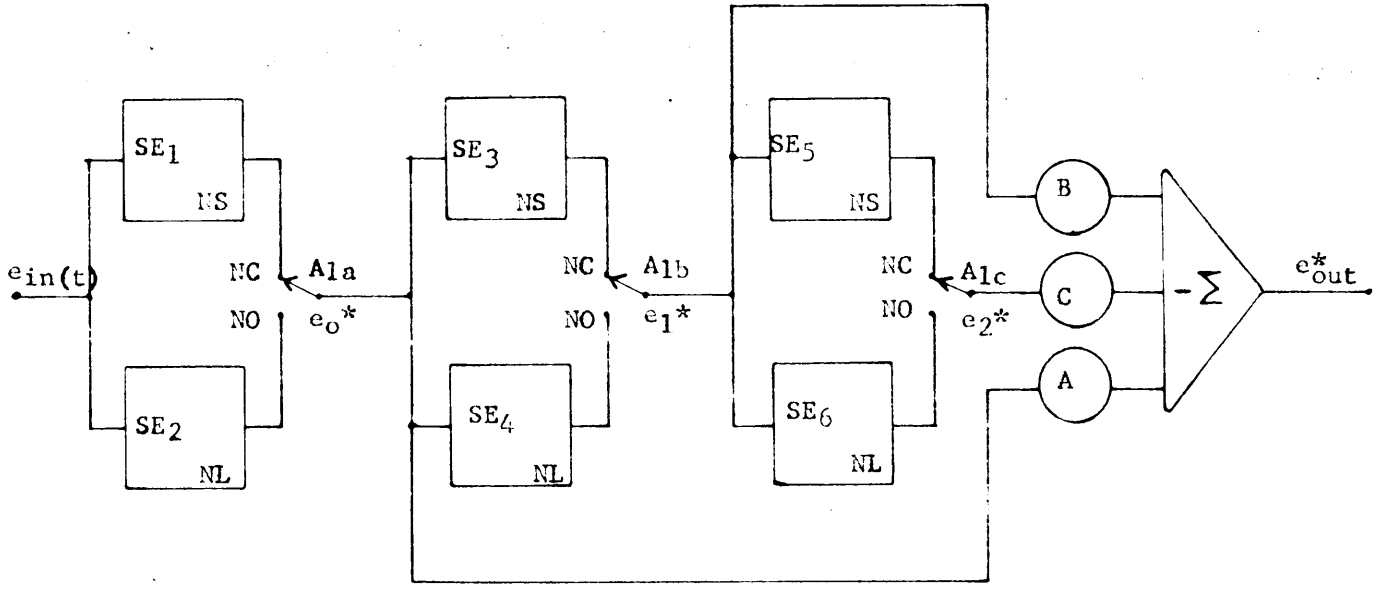


Figure 24 Second Order Hold Simulator Using Parallel Connected Storage Elements

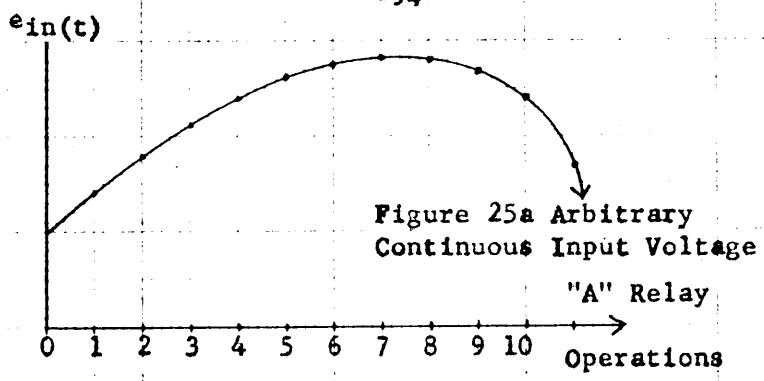


Figure 25a Arbitrary Continuous Input Voltage

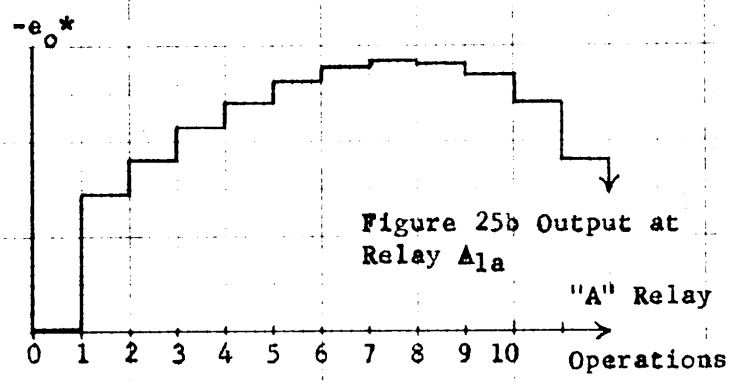


Figure 25b Output at Relay A_{1a}

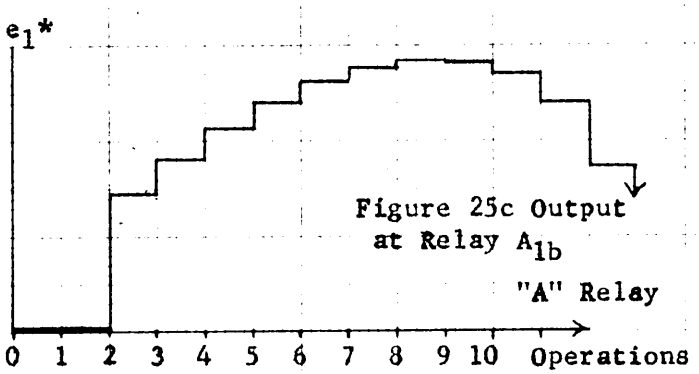


Figure 25c Output at Relay A_{1b}

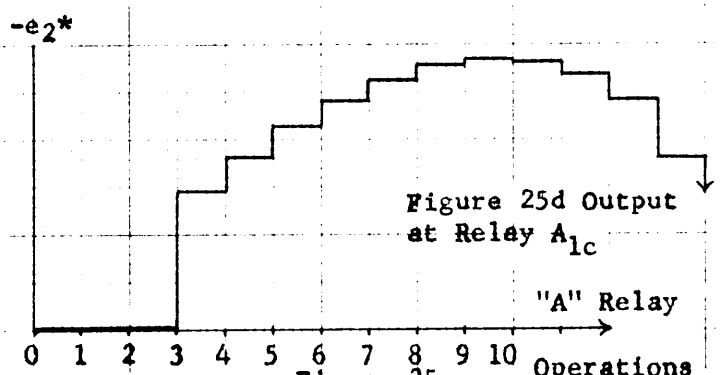


Figure 25d Output at Relay A_{1c}

Figure 25
Timing Diagram for Parallel Connected Second Order Hold Simulator

$e_{in(t)}$ (Figure 25a) with the sample e_o* appearing on the armature of relay A_{1a} (Figure 25b). The function of the second pair of storage elements, SE_3 and SE_4, is to delay the voltage at their input, which is a sample of $e_{in(t)}$ by exactly one sample period. The output e_1* appearing at the armature of relay "A_{1b}" is actually the previous sample of $e_{in(t)}$ as indicated by Figure 25c. In the simulation of a second order hold device it is also necessary to obtain the second previous sample. This can be obtained by using a third pair of storage elements, SE_4 and SE_5, to delay e_1* by exactly one sample period. The output e_2* appearing at the armature of relay A_{1c} represents the second previous sample of $e_{in(t)}$ and is shown in Figure 25d. The resulting circuit is a simple extension of the parallel connected first order hold simulator illustrated in Figure 19 and amounts to a three section delay line.

The outputs of each pair of storage elements can be multiplied by a constant, sign changed if necessary and the results summed as indicated in Figure 24 resulting in an output of the form

$$e_{out}^{*} = A e_o^{*} + B e_1^{*} + C e_2^{*} \qquad (35)$$

Comparison of this equation with equation 34 shows that the circuit of Figure 24 does indeed simulate the effect of a second order hold device. In the summing process, consideration must be given to the fact that the voltages e_o^* and e_2^* are actually the negative of the sample of $e_{in(t)}$ while e_1^* is of the correct sign. This is, of course, because of the inherent sign change in the operational amplifiers used in the storage elements.

All of the various practical difficulties associated with parallel representation of the first order hold simulator of Figure 19 are also present in the second order hold simulator illustrated in Figure 24. For this reason, the parallel simulator was not employed in the simulations described below so no experimental data on this circuit is included. The practical difficulties can be resolved, however, using the methods discussed for the first order simulator.

Second Order Hold Simulation Using Series Connected Storage Elements

It is also possible to simulate the effect of a second order hold device using series connected storage elements by a simple extension of the method used for first order hold simulation. A block diagram of a series connected second order hold

simulator is shown in Figure 26 and its operation can be explained using the theoretical timing diagram of Figure 27.

The function of the first two storage elements, SE_1 and SE_2, is to obtain periodic samples of the continuous input signal $e_{in}(t)$ as indicated in Figure 27b. The next pair of storage elements, SE_3 and SE_4, serve to sample the original sample, e_0^*, and delay it by one sample period as indicated in Figure 27c. The output of SE_4 has a value equal to the previous value of e_0^* and can be interpreted as the previous sample. The last two storage elements, SE_5 and SE_6, sample e_1^* and delay it by one sample period as indicated by Figure 27d resulting in an output of SE_6 which corresponds to the value of e_0^* two sample periods previously. The six storage elements constitute a digital delay line in which a sample value is passed from one storage element to the next with each operation of the "A" relays.

If the outputs of SE_2, SE_4, and SE_6 are multiplied by constants, signs changed as required, and summed as indicated in Figure 26 the resulting output obeys the relation

$$e_{out}^* = A e_0^* + B e_1^* + C e_2^* . \qquad (36)$$

It has been previously shown (equation 34) that this is the require-

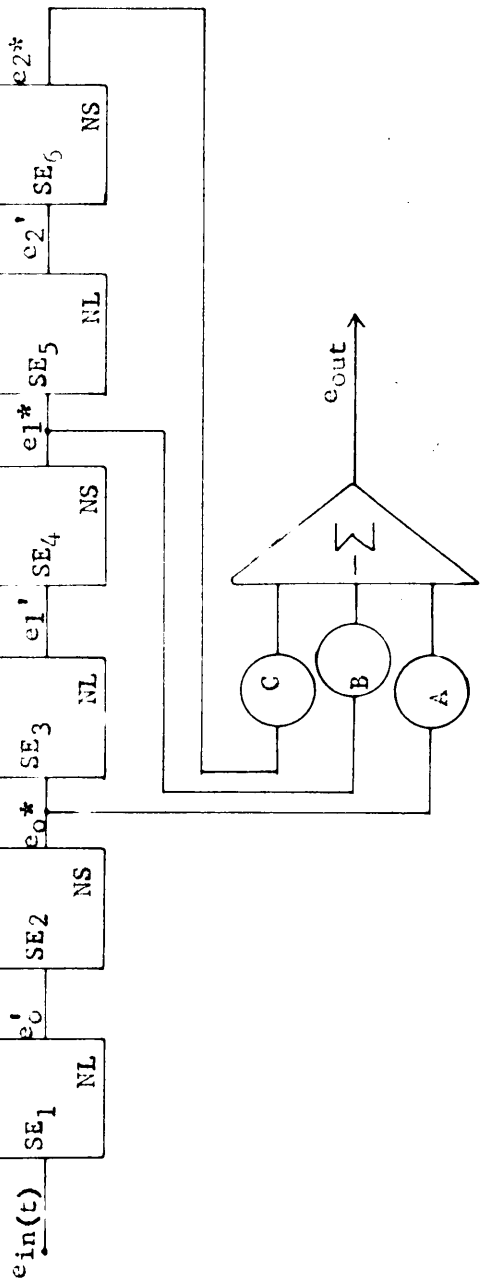


Figure 26 Second Order Hold Simulator Using Series Connected Storage Elements

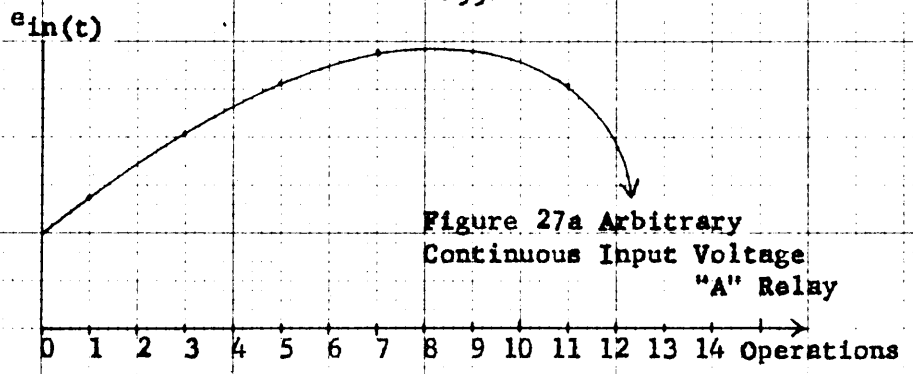


Figure 27a Arbitrary Continuous Input Voltage "A" Relay

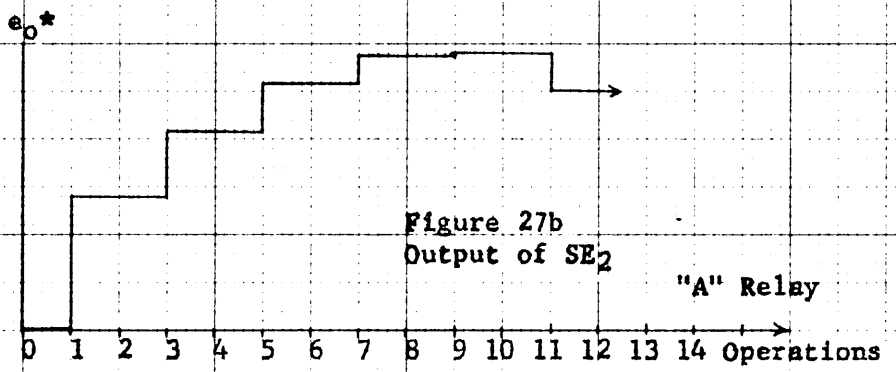


Figure 27b Output of SE_2 "A" Relay

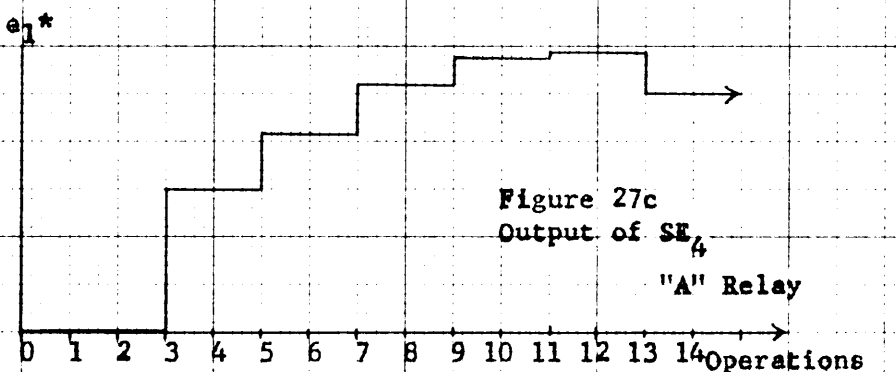


Figure 27c Output of SE_4 "A" Relay

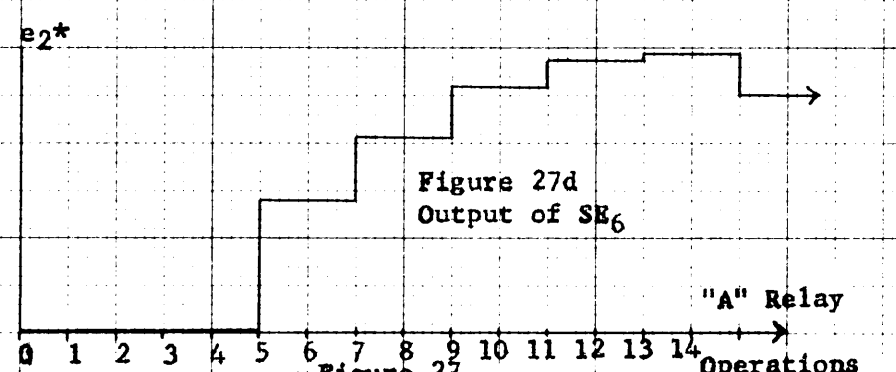


Figure 27d Output of SE_6 "A" Relay

Figure 27

Timing Diagram for Series Connected Second Order Hold Simulator

ment for the simulation of a second order hold device so it is concluded that the circuit illustrated in Figure 26 also can be used to represent such devices.

Note that in the series simulator the sign on all samples is the same as the sign of the continuous input signal at the time the sample was taken. This is different from the parallel simulator (Figure 24) in which the sign of the sample depends on the sample observed.

The "A" relays go through one complete cycle each sample period as was the case in the series simulations of zero and first order hold devices. Also it is not necessary that the "A" relays remain for an equal time in each state, resulting in less difficult adjustment of the master timing oscillator. The relays in the series connected circuit operate twice as fast as the relays in the equivalent parallel connected circuit while representing the same sampling rate. The maximum sampling rate which could be represented using the series circuit would, therefore, be less than that for the parallel circuit because of practical limits on relay operating speed. The relay switching time and the time constant of the storage elements should, of course, be short compared to the sample period for effective simulation. If the sampled-data system being represented in a

practical problem is properly time scaled there should be no necessity for operating the relays at excessive speeds.

In certain problems it may be necessary to set the initial voltages on the storage elements. This can be done using the "B" relays and floating power supplies but in the more usual case where it is desired to simply set the storage elements to zero, a better method is available. If, in the case of a series second order hold simulator, the input voltage to the delay line, $e_{in(t)}$, is made zero and the "A" relays are cycled three or more times all storage elements will be set to zero. In most problems the initial voltages are of no interest and the problem can be started without first going through the above reset procedure.

Actual data obtained from a series connected second order hold simulator such as that illustrated in Figure 26 is shown in Figure 28. The exciting voltage is a .10 cps sine wave of amplitude 10.0 volts which is subjected to a sampling rate of 1.0 sample per second. The resulting values of e_o^*, e_1^*, and e_2^* are shown as well as the input voltage $e_{in(t)}$. Careful study of these curves reveals that the circuit of Figure 26 performs reasonably well as a simulator of a second order hold device. At high sampling frequencies the quality of the simulation would be somewhat reduced because of the effect of relay switching time and storage element

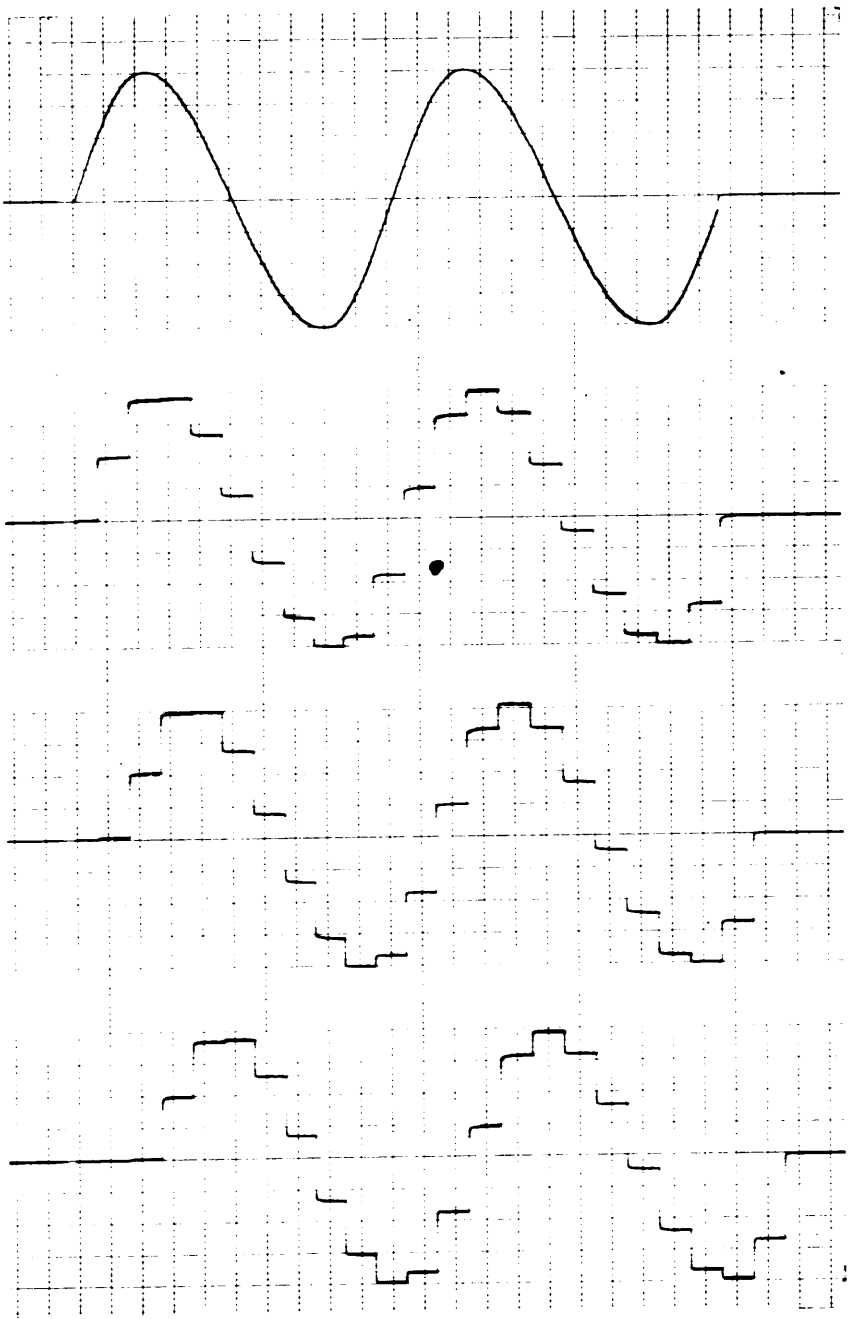


Figure 28 Outputs of Series Connected Second Order Hold Circuit

$R_f = R_i = 100K\Omega$ $C = .01\mu f$ 1.0 sample/second

Paper speed 5mm/second

time constants.

Simulation of Finite Sampling Time

In a practical sampled-data system it is physically impossible to obtain a sample in zero time. The period of time from the end of one sample to the start of the next sample is referred to as the sampling period. It should be emphasized that the sampling period is different from the sample period. The meaning of both terms is indicated in Figure 1. In many cases the sampling period is quite short in comparison with the sample period and can be neglected. The effect of finite sampling time is neglected in the simulation methods for sampling and hold devices discussed above. In the early analytical approaches[8,9,10] to the analysis of sampled-data systems the sampling process was considered to occur instantaneously. In later work[11,12] methods for approximately analyzing the effect of finite sampling time were introduced, however, the methods were quite complex and are of doubtful value in routine design work. It is, therefore, desirable to devise a method of representing finite sampling time on the analog computer which is compatable with the simulation methods already presented.

The method of finite sampling time simulation which is

presented below is designed to be used with the series connected hold simulation circuits already presented. The concept can, however, be adapted for use with the parallel connected hold simulators.

A circuit for the simulation of finite sampling time is illustrated in Figure 29a. Note the circuit is similar to that of an ordinary storage element except that an additional normally closed relay, "D_{NC}" is in series with the usual "A" relay. The finite sampling time simulator is normally used as the first of a series of storage elements in a hold simulator so in keeping with the conventions previously established in Figure 6, a normally closed "A" relay is used and the storage element is considered as a normally load (NL) element. A block diagram of the circuit of Figure 29a is shown in Figure 29b with the notation FST indicating that the element incorporates finite sampling time simulation.

The "D" relay is controlled in a manner such that it will become energized a predetermined time after the "A" relays assume their de-energized condition. The "D" relay becomes de-energized at the time the "A" relays are energized. Assuming that the "A" relays are continuously cycling, the relay timing is indicated

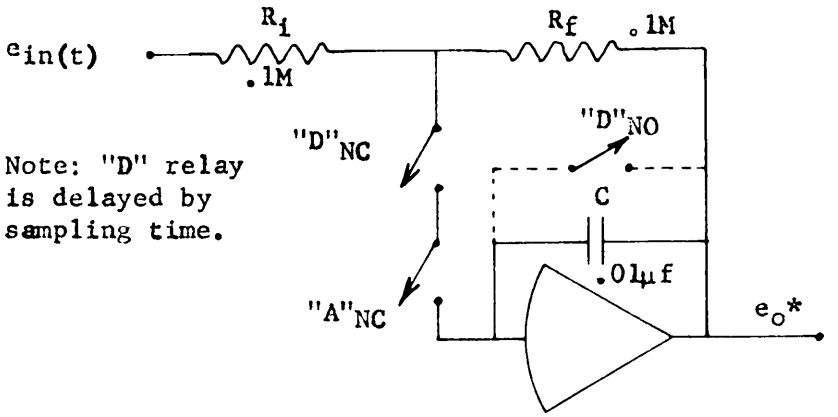


Figure 29a Circuit Diagram of Finite Sampling Time-Zero Order Hold Simulator

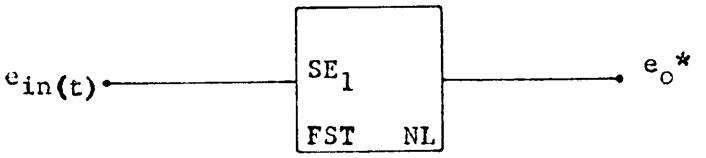


Figure 29b Block Diagram Finite Sampling Time-zero Order Hold Simulator

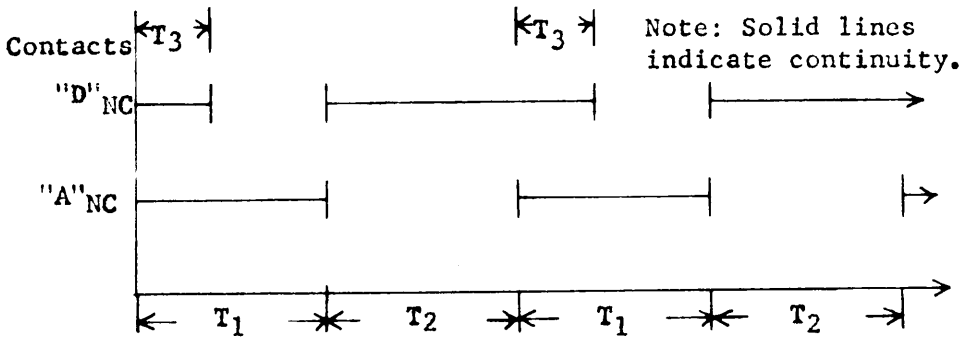


Figure 29c Relay timing Diagram for Finite Sampling Time-zero Order Hold Simulator

graphically in Figure 29c. Two complete cycles are shown starting with the "A" relay initially de-energized. It is not necessary that $T_1 = T_2$ although in normal operation this will be approximately the case. Note that the "D" relay is delayed by the period T_3 after closure of the "A_{NC}" contacts by an appropriate delay device.

The overall operation of the finite sampling time simulator of Figure 29a is indicated in the theoretical timing diagram of Figure 30 for an arbitrary input voltage $e_{in(t)}$. During the period T_3 the output of the amplifier e_o^* is following $e_{in(t)}$ but at all other times e_o^* remains constant at the value of $e_{in(t)}$ existing at the end of the previous sampling period. Comparison of Figure 30 with Figure 17, the timing diagram for the series connected zero order hold simulator, indicates that as the delay time, T_3, of relay "D" approaches zero the two output voltages become identical. This suggests that actually the circuit of Figure 29 does more than represent finite sampling time but, in addition, simulates the effect of a zero order hold device following the sampler. In most practical systems the sampler is followed by a hold device, therefore the circuit is useful as shown.

Only a minor modification of the circuit of Figure 29 is required to represent finite sampling time without a hold effect. This can be accomplished by connecting a pair of "D_{NO}" contacts

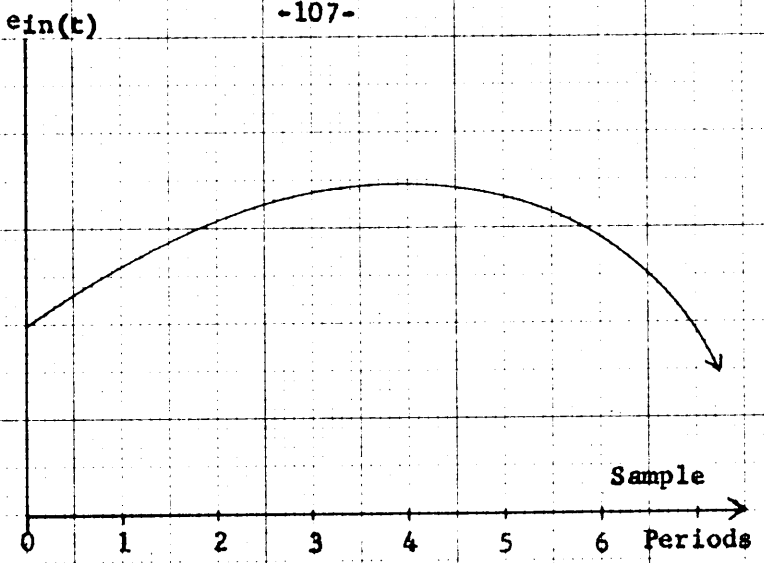


Figure 30a Arbitrary Continuous Input Signal

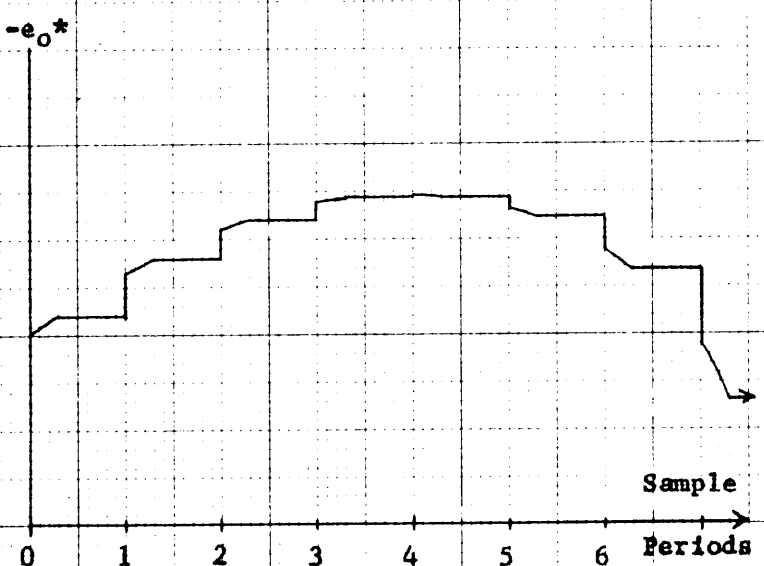


Figure 30b Output of Finite Sampling Time-zero Order Hold Simulator

Figure 30 Theoretical Timing Diagram for Finite Sampling Time-zero Order Hold Simulator

across the feedback capacitor as indicated by the dotted lines in Figure 29a. At the end of the sampling period the output of the storage element would be forced to zero resulting in an output of the form illustrated in Figure 1b. In fact, the effect of finite sampling time without hold can be simulated simply with a set of "D_{NC}" contacts in series with a set of "A_{NC}" contacts without the necessity of an operational amplifier. This can be verified by study of the relay timing diagram of Figure 29c.

Certain practical limitations exist with the finite sampling time-zero order hold circuit illustrated in Figure 29. The most significant restriction is that the sampling time cannot exceed the de-energized period (T_1) of the "A" relays. However, the period T_1 can be made well over half of the total sample period by proper adjustment of the master timing oscillator. In most practical situations the sampling period would be considerably less than half the sample period. It is necessary that the time constant of the storage element be short compared to the sampling period otherwise e_o^* will not follow the input signal, $e_{in(t)}$, during sampling. It is also necessary that the switching time of the relays be short compared to the sampling period otherwise it is difficult to obtain a predictable pulse width.

If a sampled-data system involves finite sampling time as well as a first or higher order hold device it can be simulated as before by following the zero order hold simulator by an appropriate number of storage elements acting as a delay line. The zero order hold would, of course, be represented using the circuit of Figure 29a. The outputs along the delay line, e_1^*, e_2^*, etc., would be ordinary samples since the finite sampling time would have no effect on the delayed samples either in the simulator or in the actual system.

It is noted that the zero order hold simulator with finite sampling time (Figure 29) is different from the conventional series representation of a zero order hold in that only one operational amplifier is involved. This causes the value of e_o^* and all later samples generated to be the negative of the value of $e_{in}(t)$ existing at the time the sample was taken. In the conventional series connected hold simulator this change of sign does not occur. It is necessary that the computer operator be aware of these differences in order to properly perform the summation required in the simulator first and higher order hold devices.

In order for the finite sampling time-zero order hold simulator of Figure 29 to operate properly it has been shown that the operation of the "D" relay must be synchronized with the "A"

relays and hence with the master timing oscillator. It is, of course, desirable to do this without the use of any equipment not already present on a general purpose analog computer. It is not practical to use the normal integrator control relays (the "A" and "B" relays) for the delayed relay, "D", since this would require some internal wiring changes in the computer control circuitry. The possibility of using one of the function relays present on many of the larger computers is promising.

The ES-400 computer used in this investigation has no internal function relays, however, some discarded telephone type relays were adapted for this use. Most function relays in computers are polarized but the relays available were not. This presented no particular difficulty since the non-polarized relay can be connected in series with a diode, which is available on the ES-400 computer, to give the effect of a polarized relay.

In the simulation of finite sampling time it is necessary that the computer operator be able to set the sampling time to a predetermined value. The circuit shown in Figure 31 can be used to energize a second relay a predetermined time after the operation of a first relay. In terms of the finite sampling time simulator (Figure 29) the circuit can be used to open the "D_{NC}" contacts a predictable time after closure of the "A_{NC}" contacts. The same

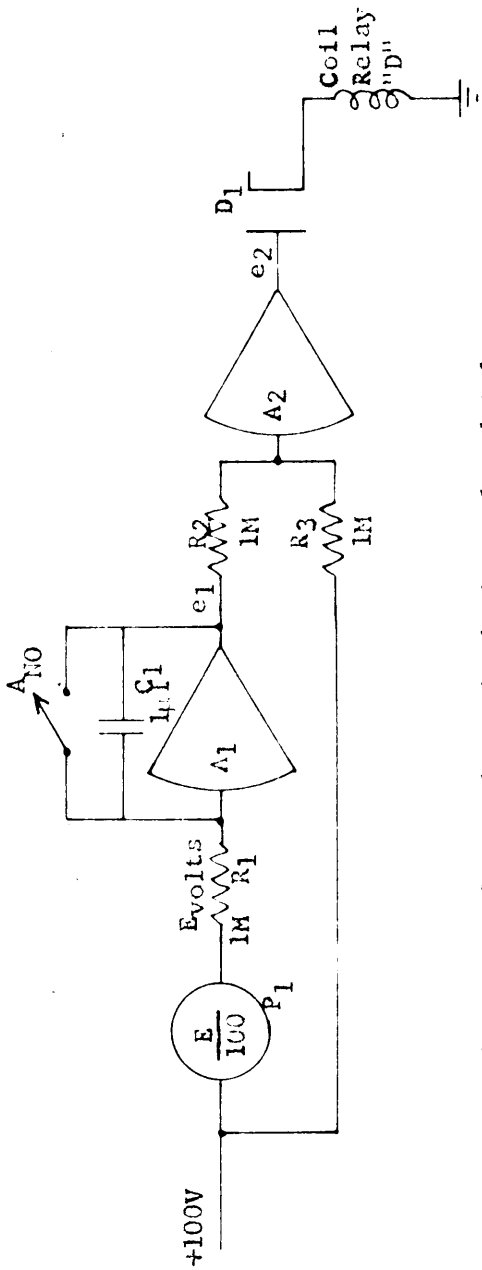


Figure 31 Circuit for Predetermined Time Delayed Relay

technique is used in the simulation of other phenomena to be described below.

The diode, D_1, is inserted in series with the time delayed relay (telephone type relay "D") to produce the effect of a polarized relay. When e_2 becomes positive the diode is biased in the forward direction and relay "D" becomes energized; however, it will remain un-energized for all negative values of e_2. Operational amplifier A_2 is used as a comparator device with $R_2 = R_3$. A fixed voltage, normally the computer reference voltage, is applied at the input of R_3 and for purposes of explanation, will be assumed positive. For all values of e_1 less negative than -100 volts and for all positive values of e_1 the summing junction of amplifier A_2 is positive causing the output to be saturated at a high negative voltage and relay "D" remains de-energized. For all values of e_1 more negative than -100 volts the summing junction is negative and the amplifier becomes saturated at a high positive voltage causing relay "D" to become energized. Because of the high gain of operational amplifier A_2 the output voltage, e_2, changes very fast as e_1 goes through -100 volts giving positive action of relay "D". The sole purpose of the comparator circuit is to insure that there is a well defined voltage at which relay "D" becomes energized and de-energized.

Amplifier A_1 is used as an ordinary integrator which is fed

by a constant E volts. This voltage is determined by the setting of the coefficient potentiometer P_1 which is in turn fed by the +100V computer reference supply. The output voltage of the integrator is expressed as

$$e_1 = -\frac{1}{R_1 C_1} \int_0^t e_{IN}\, dt \,. \qquad (37)$$

But in this case $e_{in} = E$, a constant, and equation 37 becomes

$$e_1 = -\frac{1}{R_1 C_1} \int_0^t E\, dt = -\frac{Et}{R_1 C_1} \,. \qquad (38)$$

Equation 38 indicates that the output of the integrator is a negative ramp. When e_1 passes through -100 volts relay "D" is suddenly energized. The initial conditions of the integrator are set and held at zero by the "A_{NO}" relay during the portion of the cycle in which the "A" relays are energized. At the instant the "A" relays become de-energized, the integrator begins generating a negative ramp which will eventually pass through -100V causing relay "D" to become energized. The slope of the ramp is seen from equation 38 to depend partially on R_1 and C_1 but can be set to any desired value, within practical limits, by the potentiometer.

It is important that the delay between the instant that the "A" relays become de-energized and the instant relay "D" becomes energized be predictable. Since it is known that relay "D" becomes

energized at the instant e_1 is -100 volts, equation 38 can be solved for this particular condition yielding

$$-100 = -\frac{ET}{R_1C_1} \tag{39}$$

where T is the resulting delay time. Solving equation 39 for the delay time results in

$$T = \frac{100R_1C_1}{E} \tag{40}$$

With T in seconds

R_1 in megohms

C_1 in μf

E in volts.

By proper choice of integrator components the delay time can be set to any desired value.

Note that relay "D" is always de-energized when the "A" relays are energized. This fact together with the delayed feature of relay "D" results in a relay timing diagram for "A_{NC}" and "D_{NC}" contacts identical to that shown in Figure 29c for the finite sampling time-zero order hold simulator previously discussed. The delay time is shown as T_3 in this diagram. It is concluded that the circuit of Figure 31 can be used in conjunction with the circuit of Figure 29

to simulate finite sampling time with zero order hold. The sampling time is determined by the delay time which can be computed using Equation 40.

A finite sampling time-zero order hold simulator such as that described above was actually run on the computer using the circuit of Figure 31 to control the delayed relay. The performance of this circuit when excited by a sinusoidal waveform is shown in Figure 32. The input voltage $e_{in(t)}$ is shown as well as the resulting output e_o^*. The values of components used are those indicated in Figures 29 and 31. The ratio of system frequency to sampling frequency is somewhat higher than that normally encountered in order to more clearly illustrate the performance of the circuit. From this data it is seen that the representation of finite sampling time with sample hold is reasonably satisfactory.

Simulation of Quasi-sampled Signals

Occasionally in the simulation of sampled-data systems it becomes necessary to generate a quasi-sampled signal. In the quasi-sampled situation the data or signal is sampled during certain periods but is continuous at all other times. An example of this type of signal is shown in Figure 33. Note that the signal is continuous during period T_1 but is based on a sample during period T_2. Two methods for

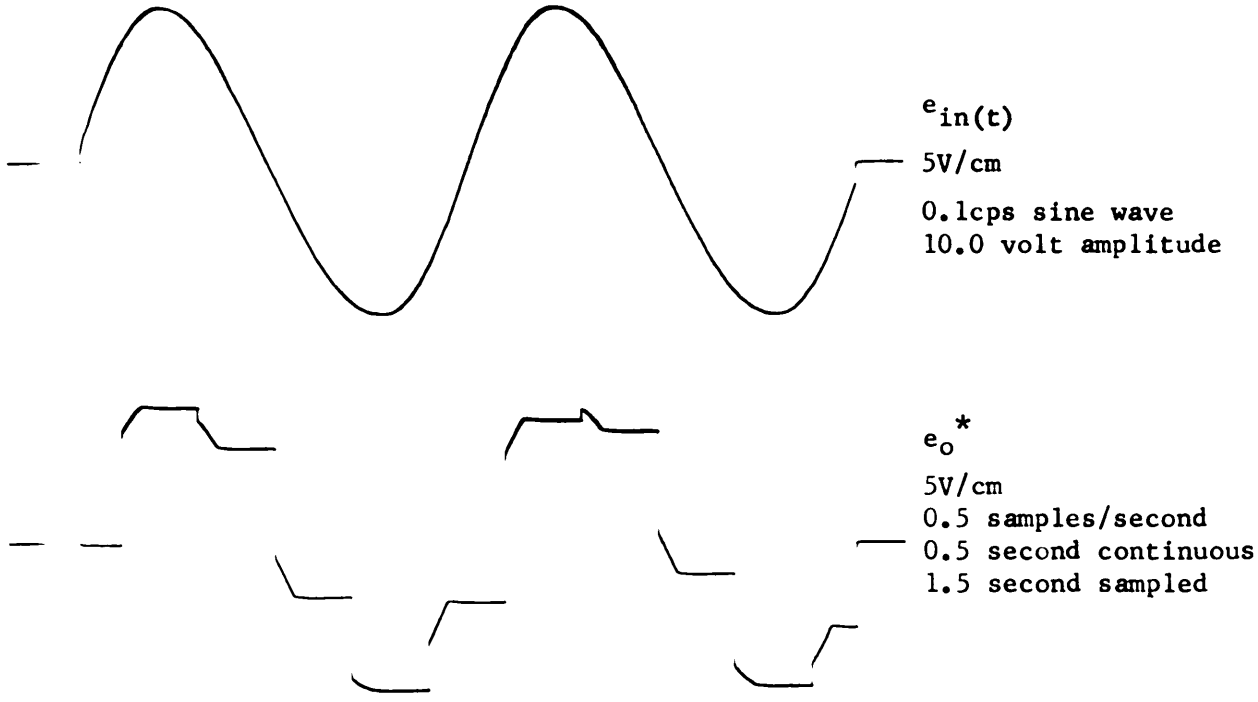


Figure 32 Performance of Finite Sampling Time Simulator

$R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper speed 5mm/second

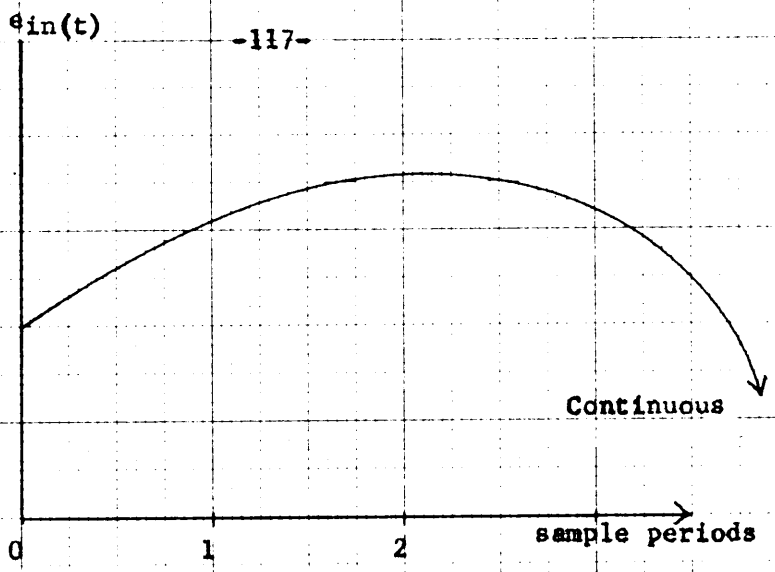


Figure 33a Arbitrary Continuous Input Signal

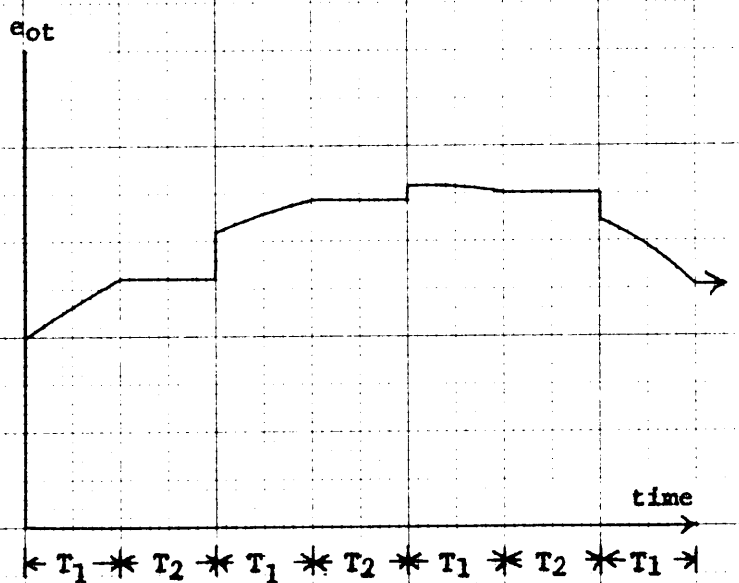


Figure 33b Quasi-sampled Signal

simulating the quasi-sampling process are suggested by the simulation techniques described above.

Comparison of the quasi-sampled signal in Figure 33 with the signal resulting from finite sampling time with zero order hold shown in Figure 30 reveals that the two situations are identical. Actually the same response can be thought of in two different ways. In the quasi-sampled system the continuous response is deliberate and may be of approximately the same duration as the sampled response. In the true sampled-data system with finite sampling time the continuous portion is often considered undesirable and is usually of short duration in comparison with the sample period.

Since the quasi-sampled situation is identical with the finite sampling time-zero order situation, it follows that the same simulator circuit (Figure 29) can be used to represent both cases. The duration of the continuous portion of the signal is equal to the delay time of relay "D". This can, of course, be set to a predetermined value by a potentiometer adjustment. The continuous portion of the signal cannot exceed the de-energized period of the "A" relays, as indicated by the relay timing diagram of Figure 29c. If the energized and de-energized periods of the "A" relays are equal, which is usually approximately the case, the continuous

portion of the signal cannot exceed 50% of the sample period. This difficulty can be relieved somewhat by intentionally adjusting the master timing oscillator such that the "A" relays are de-energized for considerably longer period than they are energized, allowing the continuous portion of the output signal to exceed half the sample period. The degree of asymmetry in timing oscillator operation is limited, particularly at high frequencies, by the time constant of the storage elements. The curves of Figure 32 can be considered as the experimental results of a simulation of the quasi-sampling process using the technique just described. It is seen that the quality of the simulation is reasonably good.

Inspection of Figure 17b, the output voltage of a single storage element, suggests that this circuit might also be used to generate a quasi-sampled signal. The continuous portion of the output signal occurs during the period that the "A" relay in Figure 6a is de-energized and the sampled portion during the period when the storage element is in the store condition. The only way of controlling these periods is by adjusting the master timing oscillator for the necessary asymmetrical operation. It has been shown that it is difficult, although not impossible, to precisely control the time in each state of the oscillator.

In general, the quasi-sampled simulator based on the use of a finite sampling time-zero order hold simulator (Figure 29) is the

most preferable of the two methods because of the ease of adjustment. The simulation using a single storage element requires some less equipment since the circuitry associated with the delayed relay is not needed. In either case the quality of simulation deteriorates at high sampling rates, and also when the ratio of continuous period to sampled period approaches zero or one. This deterioration in performance results primarily from the effect of the storage element time constant and relay switching time.

General Conclusions and Comments

The circuits discussed above will simulate most of the basic effects encountered in sampled-data systems including some of the imperfections in the sampling and holding processes. In contrast to most of the work done by other investigators, very little equipment is involved other than that normally included on a general purpose analog computer. The various simulation circuits can be combined in a number of ways to simulate quite complex sampled-data systems. In the next section of this report some of the more advanced simulation processes are described, however, at this point a comparison of the work described above with that of other investigators is appropriate.

In 1955 Wadel[19] described a method for the analysis of combined sampled and continuous data systems on the analog computer.

He suggested the construction of a delay line using a series of integrators alternating between the operate and hold conditions which bears some resemblance to the delay line using series connected storage elements described in this paper. Wadel placed the continuous portion of the system in hold while setting the integrators in the delay line to new values. This was necessary because of the relatively long time constants of the integrating amplifiers used in the delay line and resulted in a distorted solution on strip chart recordings. A multiple stepping relay was used to control the all integrating amplifiers in the program. Simulation of zero and first order hold devices was described, however, the method can be extended to higher order hold devices.

Also in 1955 a paper was published by R. C. Klien[18] describing an entirely different approach to the problem of analog simulation of sampled-data systems. The samples were stored on large capacitors which were alternately connected and disconnected to the continuous signal by means of several relays. No operational amplifiers were involved in the sampling and holding processes. The relays were controlled by thyratrons which were driven by a signal generator. The method as described requires a 400 cps power supply. Only zero order hold devices can be simulated, however, it is possible to represent computing time.

In 1957 Metzger[20] submitted an M.S. thesis to the University of California at Berkeley which describes a method of simulating sampled-data with finite sampling time but without hold. This was done by means of an elaborately controlled relay which was inserted in an analog computer simulation of a continuous system at the point of sampling. The operating frequency of the relay and the ratio of the closed to open periods could be controlled. The relay and associated control circuitry were furnished as a unit by the Donner Scientific Company. No hold circuits were described by Metzger.

The results of an analog study of a serial type digital data system were published by O. I. Elgerd[21] of the University of Florida in 1958. The work was performed, however, at the General Electric Company in Roanoke, Virginia. Elgerd used a single integrator with very fast time constant, alternating between the hold and operate conditions, as a sampling and storage device. This integrator was similar to a single storage element of the type used in the present investigation. The storage element was operated in the load condition for a very short period but was in the store condition during most of the sample period. The switching in the storage element was accomplished by means of a rotating commutator driven by a variable

speed D-C motor. Elgerd's techniques are limited to the simulation of zero order hold devices.

The most recent work known to this writer in the area of analog computer simulation of sampled-data systems was published in 1959 by Chestnut[22] et.al. of the General Electric Company. Their method was similar to that described by Elgerd except that a high speed mercury wetted relay was used to control the storage element instead of a rotating commutator. A rather elaborate computer program, driven by a low frequency oscillator, was used to operate the relay. The same limitations apply to this method that applied to the work of Elgerd.

The methods of simulating various phenomena associated with sampled-data systems presented in this paper have some resemblance to some of the methods described by previous investigators. They are in general, however, much more flexible and are capable of representing situations which cannot be represented using the previous techniques. In addition, the methods resulting from the present investigation require very little in the way of equipment not present on a good quality general purpose analog computer.

The sampling rate of all of the simulators described in this report can be set to any desired value, within practical limits, by proper adjustment of the master timing oscillator. The maximum

sampling rate which can exist in the computer is limited by the switching time of the relays and the time constant of the storage elements. Both should be short compared to the sample period for effective simulation and both can be made in the order of a few milliseconds. The minimum sampling rate is set by the leakage time constant of the capacitors in the storage elements. These are in the order of several minutes and hence offer no practical limitations. In most practical sampled-data systems the sampling rate is on the order of several samples/second. This offers no difficulty in analog simulation because the problem can be time scaled such that the sampling rate occuring in the computer is within practical limits.

In many physical systems the sampling rate is much greater than the frequencies existing in the system responses (i.e., of the continuous variable being sampled). If the sampling rate is slowed to a frequency suitable for computer simulation the frequencies of the continuous system are reduced by the same factor.

If the ratio of sampling frequency to system frequency is large in the physical system the continuous signals will vary, in the computer at very low rates. Under such conditions the effects of amplifier unbalance, integrator drift, and noise can introduce appreciable error into the problem solution because of the long computing time involved. It should be noted, however, that if the

sampling rate in the system is large the performance of the system approaches that of the equivalent continuous system and it is unlikely that such a system would be treated as a sampled-data problem on the computer.

In most of the work presented earlier in this report, it was to some extent inferred that the system error signal would be subjected to the sampling and holding process. While this is the usual situation it is not always the case. The sampling and hold simulators described above can be used at any point in a continuous system without modification and in fact, two or more simulation circuits can be used simultaneously if sampling occurs at more than one point in the continuous system. Simulation of multiple sampling would not be difficult to accomplish provided all samples were taken simultaneously. If the sampling occured at the same rate at all points, but at different times, the simulation would be only slightly more difficult. If, however, different sampling rates occured at each point, simulation could be accomplished, but only at the expense of considerable duplication of equipment.

III SIMULATION OF COMPENSATION AND EXTRAPOLATION TECHNIQUES IN SAMPLED-DATA SYSTEMS

Compensation in Sampled-Data Systems

In general, sampled-data systems are less stable than equivalent continuous systems because of the inherent time lag introduced by the sampling process. A continuous system which is less stable than desired can be compensated by introducing appropriate networks which alter the open loop response in a manner that results in satisfactory performance of the closed loop system[35]. The effect of such networks is not difficult to represent on an analog computer. It is also possible to improve the performance of sampled-data systems in the same manner by introducing the networks into the continuous portion of the system. In sampled-data systems other methods of compensation are available and offer some advantages over the use of networks. These methods are sometimes referred to as digital compensation and are based on performing various operations on the samples with a digital computer or device and the resulting signal is fed forward into the continuous portion of the system. The operations performed on the samples are such that the performance of the overall system is more nearly optimum than that resulting if the samples were simply fed directly into the continuous system. In general, the natural method of compensating continuous

systems involves the use of analog computers or devices and the natural method of compensation for sampled-data systems is by means of digital computers or devices[36].

In most practical sampled-data automatic control systems the sampling is actually, or in effect, on the error signal[37] and the sampled error is fed forward into the continuous portion of the system. For this reason the investigation to be described will refer to error sampling; however, the concepts can be extended to sampling at other points.

In general, it is desired in any automatic control device to have the actual system output equal to the desired output, or reference input, at all times. This is, of course, impossible to do in a practical system. The error signal is usually proportional to the difference between the actual output and the desired output and ideally would remain zero. The magnitude of the error signal is an indication of the performance of the system and the rate of change of error indicates that the performance will become more optimum or less optimum in the immediate future. It is possible to compute a signal to feed into the forward loop of the system, instead of the error itself, which will improve the overall system performance. The computed signal, or effective error, is based on the magnitude and sign of the actual error and its derivatives. This process is known as series compensation or series equalization[38].

In continuous systems series compensation takes the form of adding, at a point of low power level, some type of analog computer (most commonly an electrical network) in the forward loop of the system. In a sampled-data system, however, compensation usually takes the form of some type of digital computer (usually a type of hold device) which processes the samples according to some predetermined rule and feeds the resulting signal into the forward loop of the system. The forward loop of the system, would in general, be continuous in nature and because of the finite time constants involved, it would provide a smoothing action on the samples. The latter process is known as digital compensation[39]. The object of the investigation described below was to devise methods of simulating digital compensation on the analog computer.

It is shown below that the output of a digital compensator is usually in the form of a weighted summation of the most recent sample and a finite number of previous samples, and can be expressed as follows:

$$Y^* = AX_o^* + BX_1^* + CX_2^* + \text{- - - - -} \qquad (41)$$

where

Y^* is the output of the digital compensator.

X_o^* is the most recent sample of the continuous input signal (error).

X_1^* is the sample previous to X_o^*.

X_2^* is the sample previous to X_1^*.

A, B, C, etc. are constants which can be either positive or negative.

It is noted that equation 41 is not an infinite series but contains a finite number of X^* terms which in practice seldom exceed three.

Comparison of equations 38 and 41 reveals that a digital compensator amounts to a hold device, several of which have been discussed above, and for which methods of simulation have been described. It is only necessary to determine the proper values of constants in equation 41. A detailed discussion of digital compensation is not attempted in this report since much has been written on this subject. Certain processes are discussed here, however, in order to explain the techniques which have been developed for representing digital compensation on an analog computer.

In general, when an automatic control system is subjected to dynamic operation the output, or controlled quantity, will lag the desired output, or reference input, resulting in system errors. This lag can be attributed to various causes but in a continuous system, it is primarily because of the time constants, both electrical and mechanical, of the system. In addition, in a sampled-data system there is an additional time lag inherent in the sampling process. The

time lag can be reduced by increasing the gain in the forward loop of the system, magnifying the error, but this is likely to result in instability since the system will tend to overshoot the desired steady state value. A more desirable situation would be to magnify the apparent error when the actual error is large but as the output approaches the desired value, reduce the apparent error to minimize overshoot. This, in effect, amounts to increasing the system gain when the error is large resulting in a large restoring force, tending to reduce the error, but reducing the gain when the error is small tending to reduce or eliminate overshoot. In other words, the forward gain becomes a function of the error resulting in a non-linear system.

This concept can be extended to more elaborate methods in which the rate of change of error is sensed as well as the magnitude. In a situation in which the error is increasing with time, it is desirable to raise the effective gain of the system to prevent further increase and preferably to cause the error to decrease, particularly if the magnitude of error is large. Also, if the error is decreasing rapidly, it is desirable that the gain be reduced to prevent overshoot, particularly if the magnitude of error is small. It is, therefore, desirable that the effective error which is computed and fed into the forward loop of the system, effectively changing the

forward gain, depend on the rate of change of error (i.e., the first derivative) as well as the error itself. The concept can be extended, if desired, to still higher derivatives of error resulting in better performance of the closed loop system.

It has been demonstrated in the preceeding discussion that the optimum effective error to be fed into the forward loop of a system depends on the existing error as well as its derivatives. An alternate way of interpreting the effective error is to consider it as a prediction of the actual error at some future time. If the actual error is large and increasing rapidly, the predicted error, or effective error, would be considerably larger and the restoring force would be larger than that which would exist without compensation. In contrast, if the error were small and decreasing rapidly, the effective error would be still smaller and perhaps even of opposite sign, which would cause the output rate of change to be smaller than that which would exist without compensation and possibly even cause a braking action tending to prevent overshoot.

In theory error prediction can be made at any time in the future if the present error and all of its derivatives are known. If only a finite number of derivatives are known, the prediction becomes less accurate and, as might be expected, more error occurs in the prediction if the extrapolation is made into the more distant future. The familiar lead networks used in series compensation of

continuous systems are, in effect, approximate attempts to extrapolate error based on the value of the error and its first derivative.

In sampled-data systems in which the error is the sampled quantity, the effective error is computed by a digital computer or hold device, and is itself discontinuous in form. The principle described above involving computing the effective error from the existing error and its derivatives can be applied to the sampled-data case. The rate of change of error can be approximated from samples in the following manner:

$$\frac{dX}{dt} \approx \frac{X_0^* - X_1^*}{T} \tag{42}$$

where T is the sample period and, in this case X is the error. The second derivative can be approximated by the difference between two successive first derivatives as follows:

$$\frac{d^2X}{dt^2} \approx \frac{\left(\frac{X_0^* - X_1^*}{T}\right) - \left(\frac{X_1^* - X_2^*}{T}\right)}{T} \tag{43}$$

The concept can, of course, be extended to higher order derivatives. Note that the accuracy of the approximations is enhanced if the error is changing slowly and/or at high sampling rates.

Two possible approaches can be used in predicting or extrapolating future error from the present and past samples.

1. A future error sample can be predicted from the present and past samples.
2. A continuous prediction of the error can be made based on the present and past samples.

In the investigation described in this paper, both prediction of error samples and continuous prediction of error are considered. The most complex case considered is based upon the use of a second order hold device in which the present and two previous samples are considered; however, the methods can be extended to any degree of complexity desired. The methods presented are not restricted to error prediction, but can be applied at any point in an automatic control system or at more than one point if desired. In the analog computer, all variables appear as voltages, so the prediction process is essentially that of predicting the future value of some voltage based on the present and past samples of that voltage.

Linear Extrapolation of a Future Sample

A plot of an arbitrary time varying voltage $e_{(t)}$ is shown in Figure 34a. The result of periodically sampling this signal is shown as e^* in Figure 34b. The samples are of short duration, and the effect of any hold device, if present, is not shown. In Figure 34c, extrapolated values of future samples are indicated by the dotted lines. It is seen that the prediction is based on the assumption

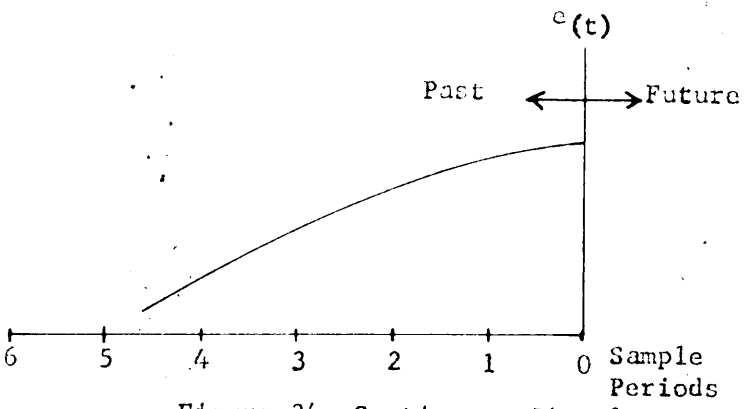


Figure 34a Continuous Signal

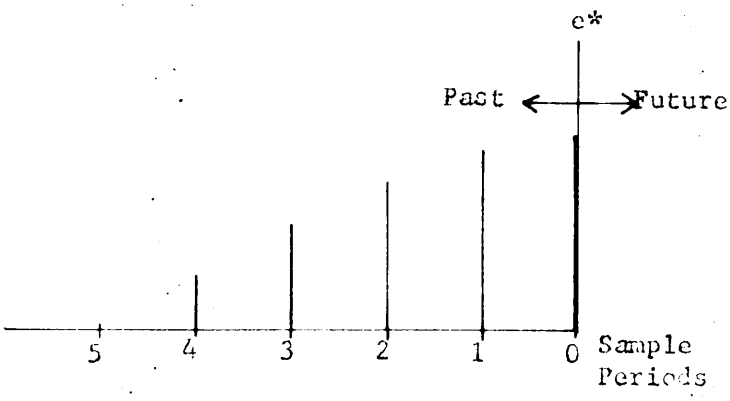


Figure 34b Sampled Signal with Short Sampling Time (Without Hold)

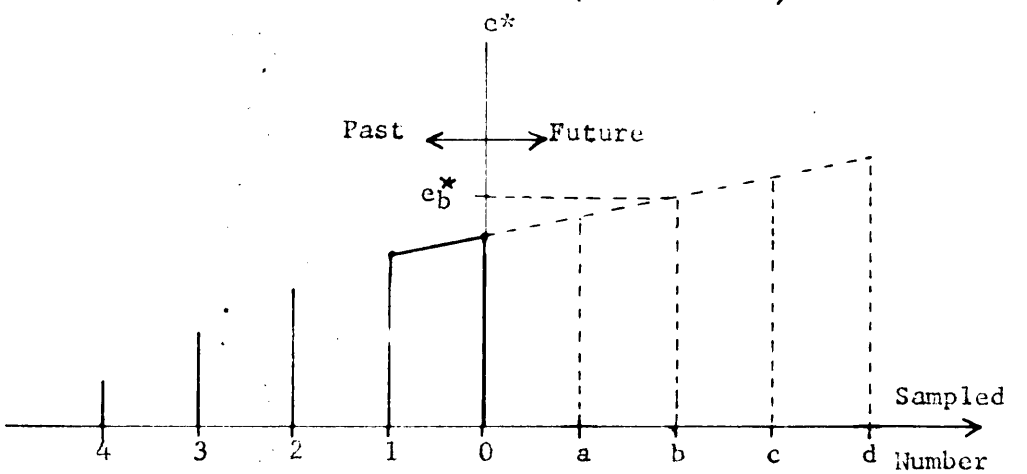


Figure 34c Predicted Samples Based on Constant First Derivative

that the rate of change (first derivative) of the voltage will remain constant. This form of prediction is referred to as first order[33] or linear[40] extrapolation. In the practical case this assumption will not always be true resulting in some error in the predicted value. In a sampled-data system the predicted value of the sample is fed, as the effective error, into the forward loop of the system. It is desirable to simulate the prediction process on an analog computer but in order to accomplish this it is necessary to express the predicted value in the form of an equation.

In linear prediction the predicted value of a future sample is based on the assumption that the rate of change of the signal will remain constant indefinitely. Under this assumption any future sample will fall along a straight line which can be expressed mathematically as

$$e_m^* = Mt + e_o^* \tag{44}$$

where

e_m^* is the predicted value of a sample to be taken at m.

M is the present slope of $e_{(t)}$.

e_o^* is the present value of $e_{(t)}$.

t is the interval of extrapolation.

The present rate of change of $e_{(t)}$ can be estimated from the two most recent samples, e_o^* and e_1^* as follows:

$$M = \frac{d\,e(t)}{dt} \approx M^* = \frac{e_0^* - e_1^*}{T} \qquad (45)$$

where

M* is the slope computed from samples.

T is the sample period as defined in Figure 1.

Substitution of M* for M in equation 44 results in

$$e_m^* = \left(\frac{e_0^* - e_1^*}{T}\right) t + e_0^* . \qquad (46)$$

It is convenient to express the extrapolation time in terms of sample periods by making the substitution

$$t = mT . \qquad (47)$$

It is not necessary that m be an integer. Substitution of equation 47 into equation 46 results in

$$e_m^* = \left(\frac{e_0^* - e_1^*}{T}\right) mT + e_0^* \qquad (48)$$

which reduces to

$$e_m^* = (m+1)\, e_0^* - m\, e_1^* . \qquad (49)$$

In most practical sampled-data systems the values of the samples are held until they are replaced by more recent samples. At the instant of sampling the present sample becomes the past sample and is replaced by a new present sample from which a new predicted sample is computed. Under these conditions it is seen that, at the instant of sampling, samples as well as the predicted sample change simultaneously and, instantly after which time they remain constant throughout the sample period. A minor exception is the case when finite sampling time is considered in which the present sample quantity is actually quasi-sampled causing the predicted value to change over part of the sampling period. The sampling time is usually small, however, in comparison with the period of the sample.

Comparison of Equations 31 and 49 reveals that a linear prediction device actually amounts to a first order hold device. The value of the constant A in equation 31 is m+1 and the value of B is -m. Both the holding and sampling process can be simulated by the circuit of Figure 19 or preferably by that of Figure 21. It is seen that linear prediction amounts to a weighted summation of the most recent and a finite number (in this case one) of past samples as expected.

In linear prediction it is assumed that the first derivative of the signal is a constant and, therefore, the predicted value is

only correct when the input signal varies linearly. The predicted value will be correct for ramp or step inputs except for the first sample period after application of the signal. Also predicted values of a triangular wave will be correct except at the peak of the wave where overshoot will occur for m sample periods after the peak, after which a correction is made and the predicted values will again be correct until the next peak occurs. For all other types of inputs, including sinusoidal, the extrapolated values are in error.

A first order hold circuit with constants selected to predict two sample periods in advance is shown in Figure 35. This is the special case for m = 2 and the predicted quantity is shown in Figure 34c as e_b^*. For this case equation 49 becomes

$$e_b^* = 3e_o^* - 2e_1^* \qquad (50)$$

Data obtained from the circuit of Figure 35 is illustrated in Figure 36 for a .05 cps triangular input voltage sampled at a rate of 1.0 sample/second. The effect of finite sampling time was not included. The continuous input voltage (e_t), the present sample (e_o^*), and the predicted sample (e_b^*) are shown to enable the quality of sampling and extrapolation to be evaluated. The predicted value of the sample and the value actually resulting two sample periods later are tabulated for each sample period in

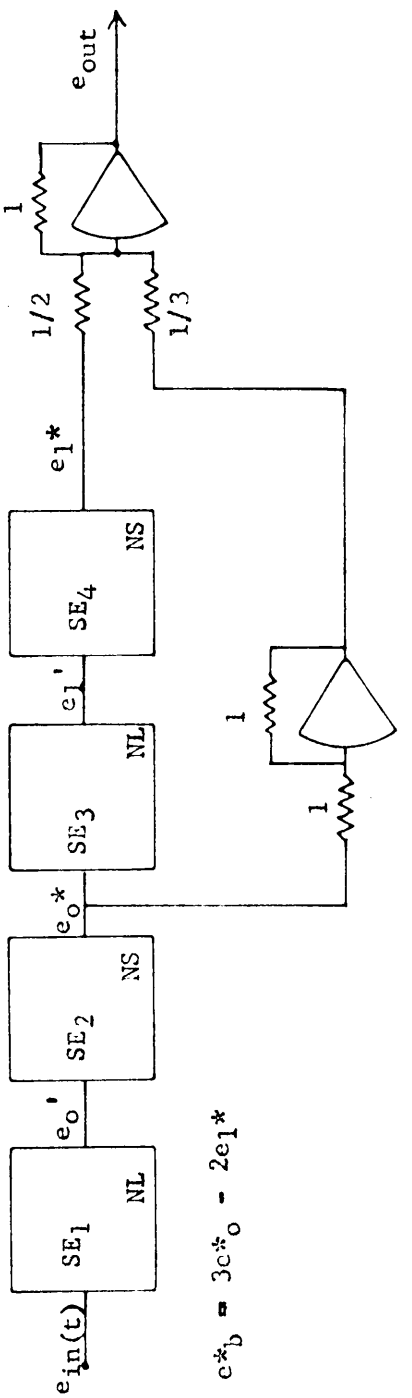


Figure 35 Circuit for Linear Prediction of Future Sample. m = 2.

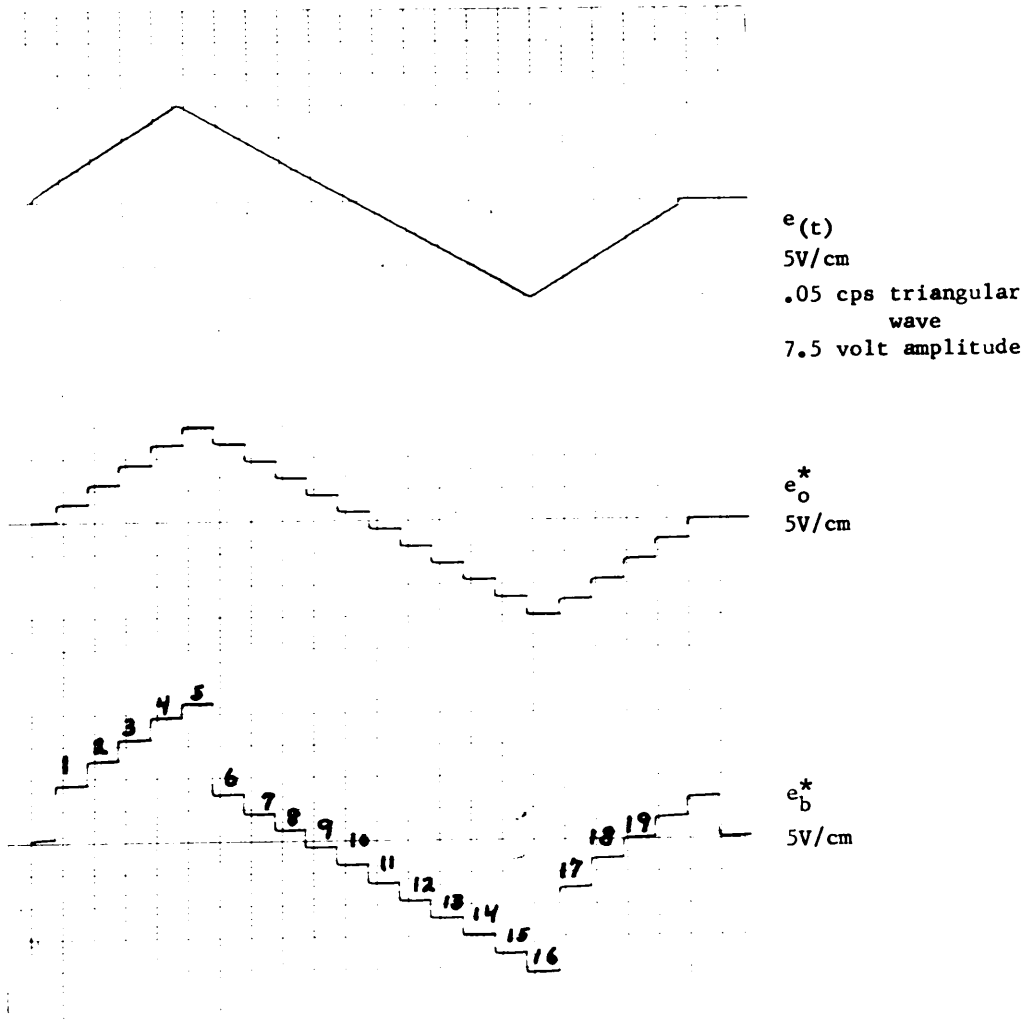


Figure 36 Performance of Linear Prediction Circuit with Triangular Input

m = 2 $e_b^* = 3e_o^* - 2e_1^*$ Paper speed 5mm/second

1.0 sample/second $R_f = R_i = 100K\Omega$ $C = .01\mu f$

Table 1. The error, defined as the predicted value minus the value actually resulting, is also tabulated. Careful study of the table reveals that, in general, the predicted values are quite close to the values resulting two sample periods later. At the peaks, however, the extrapolated value overshoots for two sample periods resulting in large errors after which a pronounced correction occurred. As expected, however, the prediction is reasonably satisfactory as long as the slope of the input voltage remains constant, otherwise errors result.

The curves of Figure 37 are similar to those of Figure 36 except that the input to the first order prediction circuit (Figure 35) is a .05 cps sine wave. The predicted values of the samples, the actual value resulting two sample periods later, and the error for this input are also tabulated in Table I. The property of a sinusoid which is of particular interest is that its slope is continuously changing and the rate of change of the slope (second derivative) is also changing reaching a maximum at the peak of the sine wave. A careful study of Table I reveals that predicted values are reasonably accurate when the slope is changing slowly (i.e., when the sine wave goes through zero) but a greater error is evident in the form of overshoot when the slope is changing rapidly.

TABLE I

PERFORMANCE DATA FOR LINEAR PREDICTION CIRCUIT

	Triangular Input			Sinusoidal Input		
Sample Period	Predicted Value e_b^*	Resulting Value e_o^*	Error $e_b^*-e_o^*$	Predicted Value e_b^*	Resulting Value e_o^*	Error $e_b^*-e_o^*$
1	9.0	9.0	0.0	3.0	9.5	6.5
2	12.5	12.0	0.5	11.0	12.5	1.5
3	16.0	15.0	1.0	19.0	14.5	4.5
4	19.5	12.0	7.5	20.0	15.5	4.5
5	21.5	9.5	12.0	19.5	14.0	5.5
6	7.5	7.0	0.5	16.5	12.0	4.5
7	4.5	4.0	0.5	13.0	9.0	4.0
8	2.0	1.5	0.5	8.0	5.0	3.0
9	-1.0	1.0	0.0	3.0	1.0	2.0
10	-4.0	-4.0	0.0	-2.0	-3.0	1.0
11	-7.0	-7.0	0.0	-7.0	-7.0	0.0
12	-9.5	-9.5	0.0	-11.5	-10.5	1.0
13	-12.0	-12.0	0.0	-15.0	-13.0	2.0
14	-15.0	-15.0	0.0	-17.0	-14.5	2.5
15	-18.0	-12.5	5.5	-18.0	-15.0	3.0
16	-21.0	-9.5	11.5	-17.5	-14.0	3.5
17	-7.5	-6.0	1.5	-15.5	-11.5	4.0
18	-3.0	-3.0	0.0	-12.0	-8.0	4.0
19	0.0	0.0	0.0	-7.0	-4.0	3.0

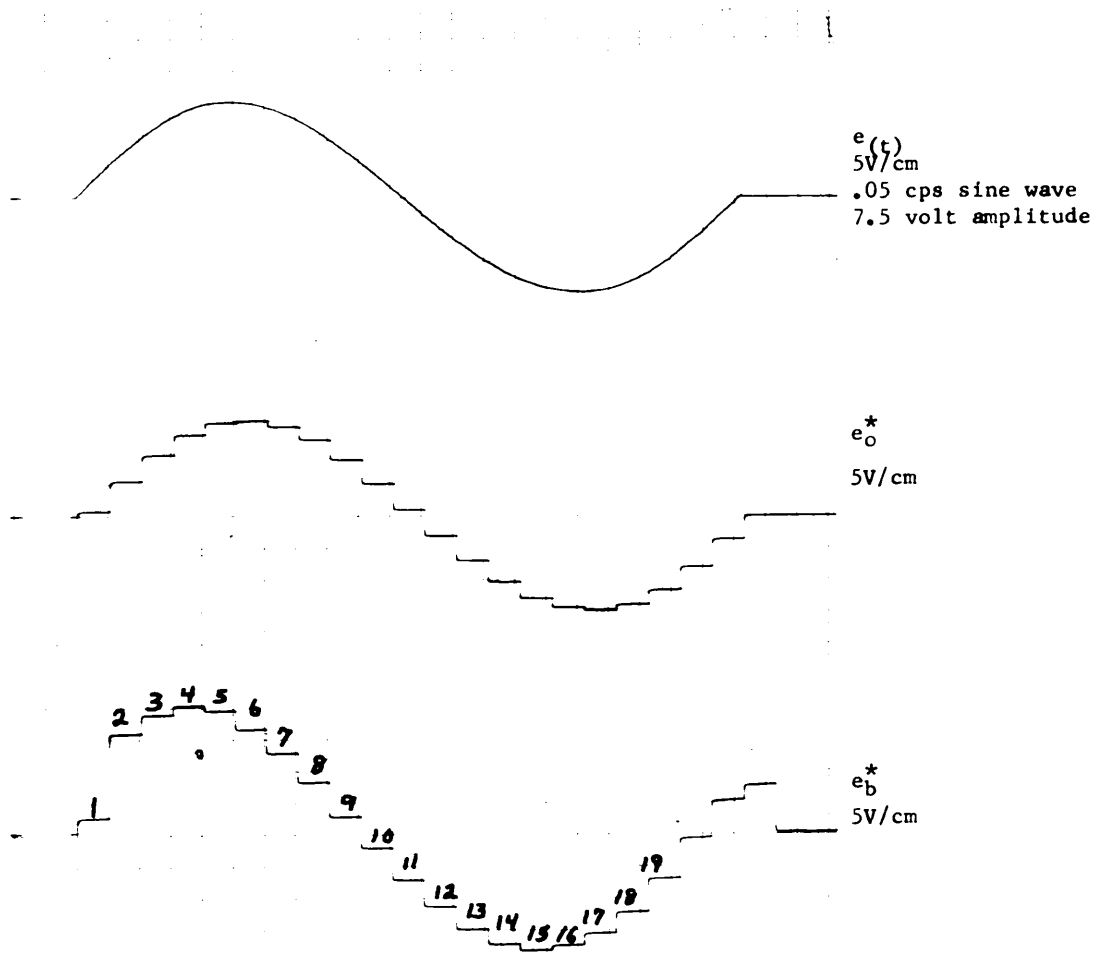


Figure 37 Performance of Linear Prediction

Circuit with Sinusoidal Input

m = 2 $e_b^* = 3e_o^* - 2e_1^*$ Paper speed 5mm/second

1.0 sample/second $R_f = R_i = 100K\Omega$ $C = .01\mu f$

Quadratic Extrapolation of a Future Sample

The concept of linear extrapolation can be extended to consider higher order derivatives resulting in more accurate prediction of future samples with complex input signals. The next logical step in this direction is now considered.

Figure 38a shows an arbitrary time varying voltage $e_{(t)}$. When $e_{(t)}$ is subjected to sampling, but without hold, the result will be as illustrated in Figure 38b if the sampling time is very short in comparison with the sample period. Extrapolated values of future samples are indicated by the dashed lines in Figure 38c. The predicted values in this case are based on the assumption that the rate of change of the slope (i.e., the second derivative) of $e(t)$ will remain constant at the present value. This type of extrapolation is referred to as quadratic prediction[40]. In many practical situations the assumption of constant second derivative does not hold and the predicted value will be in error, however, in general, the error will be smaller than that which would exist if linear extrapolation were used.

It is desirable to be able to simulate quadratic prediction using an analog computer. The predicted values of error can be used as the effective system error which can then be fed into the computer representation of the continuous portion of the system under study resulting in a computer simulation of the complete sampled-data

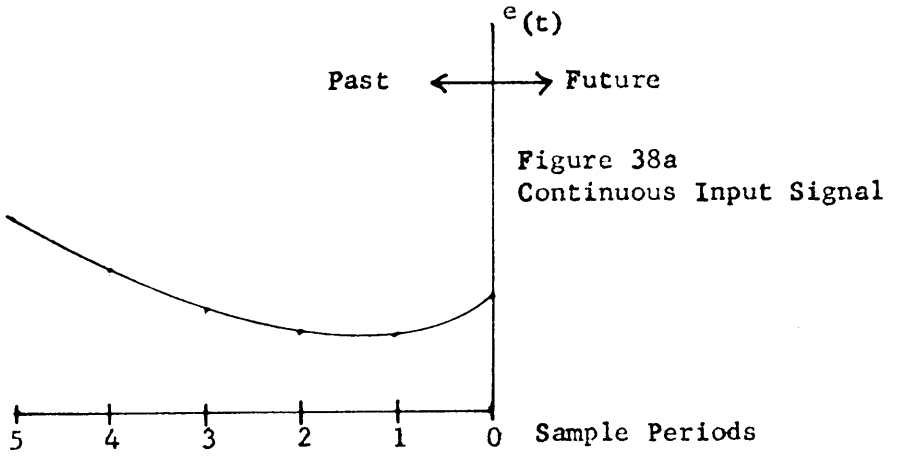


Figure 38a
Continuous Input Signal

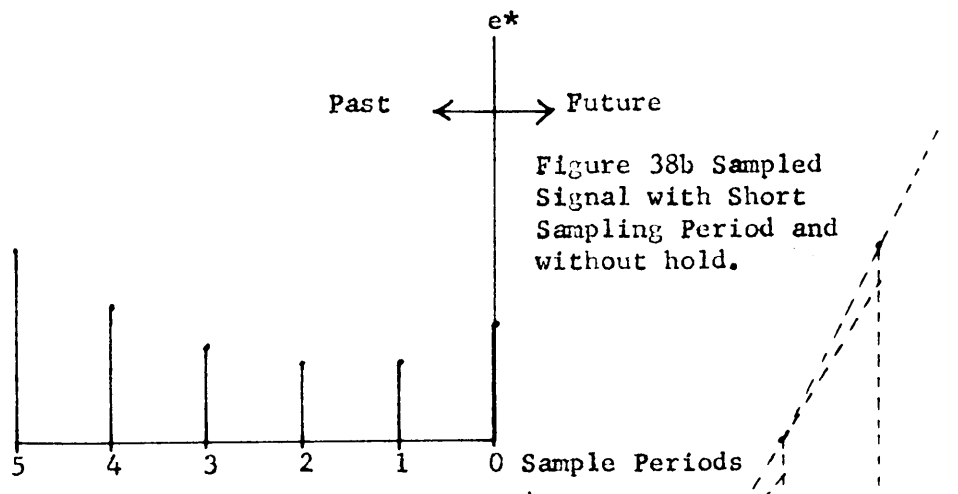


Figure 38b Sampled Signal with Short Sampling Period and without hold.

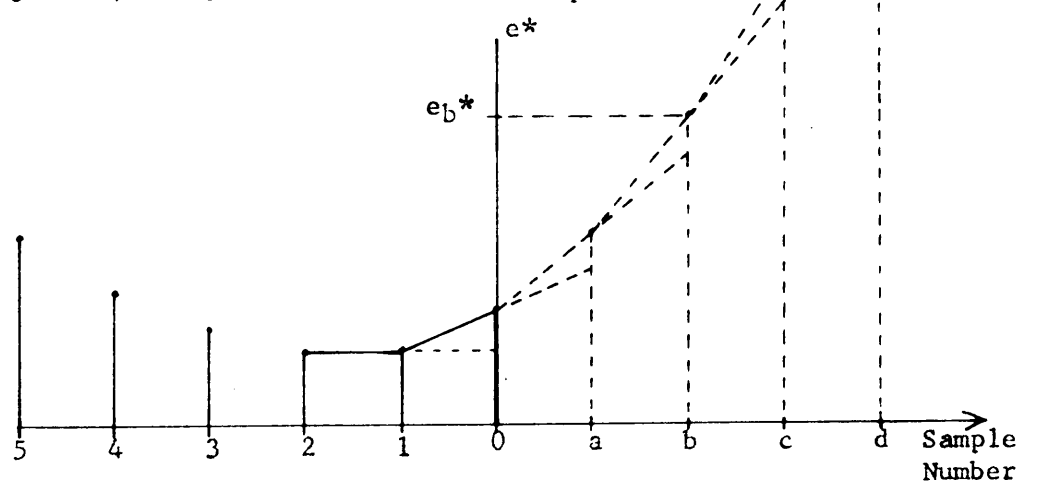


Figure 38c Predicted Sample Based on Constant Second Derivative

system. In order to simulate the prediction process on the computer, it is necessary that the predicted value by expressed in the form of an equation.

In quadratic extrapolation it is assumed that the second derivative of the input voltage is constant. Under this assumption any future sample e_n^* will fall along a curve which can be expressed mathematically as

$$e_n^* = K_1 t^2 + K_2 t + e_o \qquad (51)$$

where e_n^* is the predicted value of a sample to be taken at an arbitrary point n. e_o is the present value of the input signal. t is the interval of extrapolation. K_1 and K_2 are constants.

An approximate predicted value for e_n^* can be obtained from samples by considering the curve implied by equation 51 as consisting of straight line segments with the breaks occuring at the sampling instants as indicated in Figure 38c. The value of e_n^* can be expressed as

$$e_n^* = M_{m,n} t' + e_m^* \qquad (52)$$

where $M_{n,m}$ is the slope of the segment from arbitrary point to n. t' is the time from m to n. e_m^* is the predicted value of the sample at m.

Time m is the last complete sample period before time n and can be defined mathematically as

$$m = \frac{t}{T} \text{ with the positive remainder } t'. \tag{53}$$

where t is the interval of extrapolation and T is the sample period.

The quantity m must be an integer, however, n need not be since t' can have any value less than the sample period T. The quantity t' is always positive and can be expressed as

$$t' = t - mT \tag{54}$$

It is seen from equation 52 that the slope $M_{m,n}$ must first be computed in order to determine the value of e_n^*. The slope of the segment from 0 to a can be expressed as

$$M_{0,a} = M_{1,0} + \Delta M \tag{55}$$

where Δ M is the change in slope from segment 1,0 to segment 1,a. Since it is assumed that the second derivative is constant, it follows that Δ M is also constant and at each sampling instant the slope between adjacent segments will change by this amount. It, therefore, follows that the slope from a to b can be written as

$$M_{a,b} = M_{0,a} + \Delta M = (M_{1,0} + \Delta M) + \Delta M = M_{1,0} + 2\Delta M \tag{56}$$

The procedure can be extended to the general case for the slope from m to n as follows

$$M_{m,n} = M_{1,o} + (m+1)\,\Delta M \,. \tag{57}$$

It is necessary that e_m^* be known in order to solve equation 52 for e_n^*. This can be computed using the relationship

$$e_m^* = (M_{m-1,m})T + e_{m-1}^* \,. \tag{58}$$

Using the same procedure e_{m-1}^* can also be expressed as

$$e_{m-1}^* = (M_{m-2,m-1})T + e_{m-2}^* \tag{59}$$

which when substituted into equation 58 yields

$$e_m^* = (M_{m-1,m})T + (M_{m-2,m-1})T + e_{m-2}^* \,. \tag{60}$$

It is seen that e_{m-2}^* must next be evaluated. Similar substitutions can be repeated until eventually the remaining sampled term becomes the present sample e_o^* which is known. The resulting general expression for e_m^* becomes

$$e_m^* = M_{o,a}T + M_{a,b}T + M_{b,c}T + \cdots\cdots M_{m-1,m}T + e_o^* \tag{61}$$

which when re-arranged yields

$$e_m^* = (M_{o,a} + M_{a,b} + M_{b,c} + \cdots\cdots + M_{m-1,m})T + e_o^* \,. \tag{62}$$

Each slope in equation 62 can be expressed in terms of the known slope, $M_{1,o}$, and the expected change of slope, ΔM, by use of equation 57 resulting in

$$e_m^* = \left[(M_{1,o} + \Delta M) + (M_{1,o} + 2\Delta M) + \cdots + (M_{1,o} + m(\Delta M))\right]T + e_o^* . \quad (63)$$

This can be re-arranged to give

$$e_m^* = \left[m M_{1,o} + (1+2+3+\cdots+m)\Delta M \right]T + e_o^* . \quad (64)$$

Equation 64 is next substituted for e_m^* in equation 52 yielding

$$e_n^* = M_{m,n} t' + \left[m M_{1,o} + (1+2+\cdots+m)\Delta M \right]T + e_o^* . \quad (65)$$

The slope $M_{m,n}$ can be expressed by equation 57 and equation 65 becomes

$$e_n^* = \left[M_{1,o} + (m+1)\Delta M \right] t' + \left[m M_{1,o} + (1+2+\cdots+m)\Delta M \right]T + e_o^* . \quad (66)$$

The quantities $M_{1,o}$ and ΔM can be computed from the present and previous two samples. The slope $M_{1,o}$ can be expressed as

$$M_{1,o} = M_{1,o}^* = \frac{e_o^* - e_1^*}{T} \approx \frac{de(t)}{dt} . \quad (67)$$

The change in slope can be obtained in the following manner:

$$\Delta M = M_{1,0} - M_{2,1}$$

$$\Delta M = \Delta M^* = \frac{e_0^* - e_1^*}{T} - \frac{e_1^* - e_2^*}{T}$$

$$\Delta M^* = \frac{e_0^* - 2e_1^* + e_2^*}{T} \qquad (68)$$

Note that $\frac{\Delta M^*}{T} \approx \frac{d^2 e(t)}{dt^2}$

which is assumed constant in quadratic prediction. This implies that ΔM^* is a constant as was assumed above.

Equation 64 can be stated in a more useful form by expressing $M_{1,0}$ and ΔM in terms of sampled quantities resulting in

$$e_n^* = \left[\frac{e_0^* - e_1^*}{T} + (m+1)\frac{e_0^* - 2e_1^* + e_2^*}{T}\right]t' + \left[m\left(\frac{e_0^* - e_1^*}{T}\right) + (1+2+\cdots+m)\left(\frac{e_0^* - 2e_1^* + e_2^*}{T}\right)\right]T + e_0^* \qquad (69)$$

which can be reduced to

$$e_n^* = \left[e_0^* - e_1^* + (m+1)(e_0^* - 2e_1^* + e_2^*)\right]\frac{t'}{T} + m(e_0^* - e_1^*) + (1+2+\cdots+m)(e_0^* - 2e_1^* + e_2^*) + e_0^* \qquad (70)$$

$$e_n^* = \left[(m+2)e_0^* - (2m+3)e_1^* + (m+1)e_2^*\right]\frac{t'}{T} + m(e_0^* - e_1^*) + (1+2+\cdots+m)(e_0^* - 2e_1^* + e_2^*) + e_0^*. \qquad (71)$$

If n is an integer, implying that the voltage to be predicted is an integer number of sample periods in the future, $t' = T$. Equation 69 can now be reduced considerably by making the substitution $n = m + 1$ resulting in

$$e_n^* = (e_0^* - e_1^*) + n(e_0^* - 2e_1^* + e_2^*) + (n-1)(e_0^* - e_1^*) + (1+2+\cdots\cdots+(n-1))(e_0^* - 2e_1^* + e_2^*) + e_0^* \qquad (72)$$

$$e_n^* = n(e_0^* - e_1^*) + (1+2+\cdots\cdots+(n-1)+n)(e_0^* - 2e_1^* + e_2^*) + e_0^* . \qquad (73)$$

To further simplify let

$$s = (1+2+3 - - - - - - + n) \qquad (74)$$

Equation 73 now reduces to

$$e_n^* = n(e_0^* - e_1^*) + s(e_0^* - 2e_1^* + e_2^*) + e_0^* \qquad (75)$$

$$e_n^* = (n+s+1)e_0^* - (2s+n)e_1^* + s e_2^* . \qquad (76)$$

Inspection of equation 76 reveals that the predicted value of a future sample can be expressed as a weighted summation of the present sample, e_0^* and two previous samples, e_1^* and e_2^*. Comparison

of this equation with equation 36 indicates that a quadratic predictor is actually a second order hold device with the following coefficients.

$$A = n + s + 1$$

$$B = -(2s + n)$$

$$C = s$$

This sampling and holding process can be represented on the analog computer by the circuit of Figure 24 or preferably the circuit of Figure 26 which have been described above.

If the samples are held from one sampling instant to the next, which usually occurs in practice, the predicted value will also be held throughout the sampling period and will change abruptly at the sampling instant. A minor exception occurs when finite sampling time is considered as was true with linear prediction.

In the development of the quadratic prediction circuit it was assumed that the second derivative (i.e., rate of change of slope) of the input would remain constant at the value existing shortly before the prediction is made. The predicted value will, of course, be in error if this assumption is not true. Except for transient conditions the prediction will be correct for step, ramp, and triangular inputs as was the case with the linear prediction circuit. The quadratic prediction circuit will, in general, also predict

correctly future values of a parabolic wave since the second derivative of such a waveform is constant. Some error will occur, however, as the parabolic wave passes through zero since the second derivative will change sign at this point although the magnitude remains constant. For all other types of inputs, including sinusoidal, the extrapolated value will be in error. If the sampling rate is high, the errors tend to become smaller since the second derivative of the input signal does not change appreciably from one sample to the next. In general, however, for a given sampling rate and extrapolation time the errors in predicted values will be less with quadratic extrapolation than with linear extrapolation.

A second order hold circuit with constants chosen to predict the value of a sample two sample periods in advance (n = 2) is shown in Figure 39. The same circuit (Figure 26) can be used to predict to other times in the future simply by changing the potentiometer settings and/or the value of the summing amplifier input resistors to the values implied by equation 76. For prediction to two sample periods into the future (n = 2) this equation reduces to

$$e_b^* = 6e_o^* - 8e_1^* + 3e_2^* \qquad (77)$$

and the circuit of Figure 39 results. In the circuit as shown, instantaneous sampling is assumed although this assumption is not

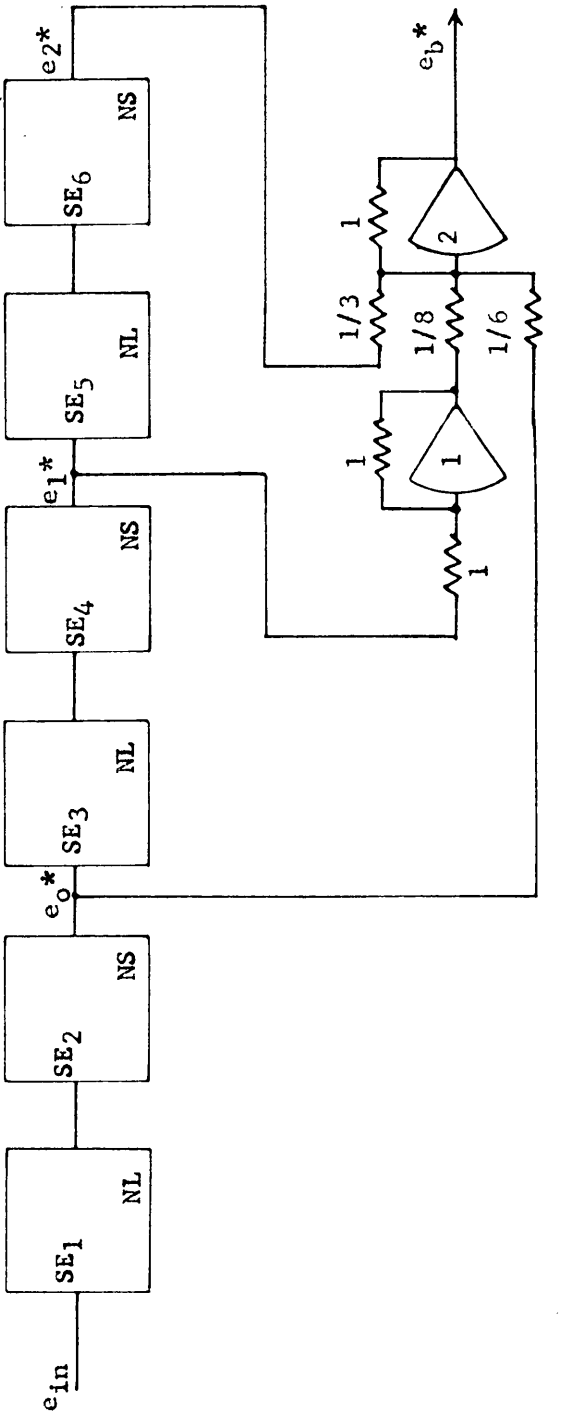


Figure 39 Circuit for Quadratic Prediction of Future Sample n = 2

necessary.

Curves showing the performance of the quadratic extrapolation circuit of Figure 39 with a .05 cps triangular wave input are shown in Figure 40. Plotted are the unsampled input voltage, the present sample, and the predicted value. The input signal is identical with that previously used with the linear prediction circuit of Figure 35 resulting in the response shown in Figure 36. The predicted sample values for the quadratic circuit and the resulting errors are listed in Table 2. It is seen that, in general, the predicted value is quite close to the value actually resulting except at the peaks of the triangular wave where considerable error exists since the second derivative of the triangular wave is not constant at this point. Comparison of Figures 36 and 40 shows that the behavior of the two circuits at the peaks is quite different with the quadratic circuit giving, if anything, the inferior performance. This is a result of the abrupt change in the slope of the triangular wave causing the prediction circuit to over compensate since it assumes that the slope will again change by this amount at the next sampling instant.

The performance of the quadratic predictor is best shown by using an input voltage with a parabolic waveform. A generator which would produce this waveform was not available, but was constructed using an operational amplifier driven by a triangular

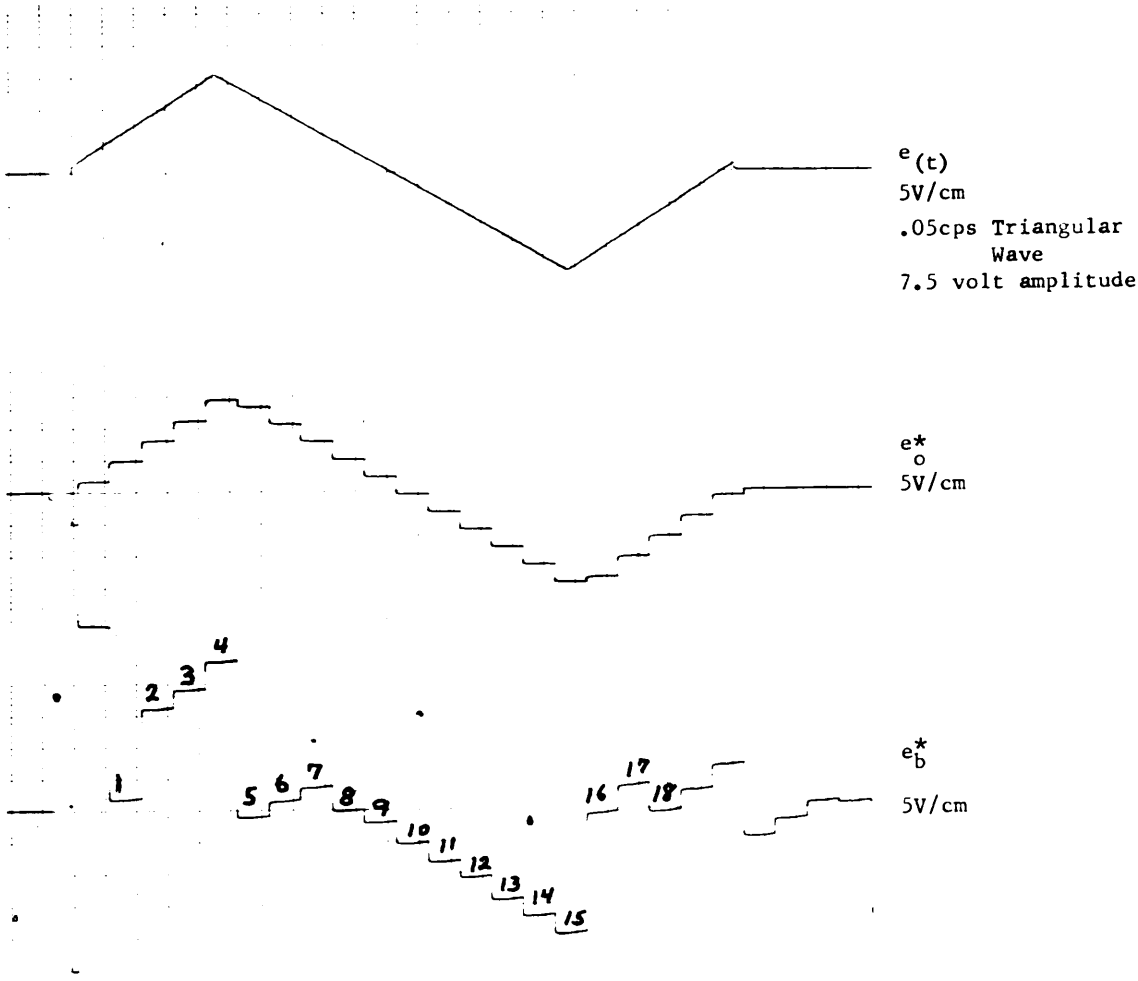


Figure 40 Performance of Quadratic Prediction Circuit

with Triangular Input

$n = 2 \quad e_b^* = 6e_o^* - 8e_1^* + 3e_2^*$ 1.0 sample/second

$R_f = R_i = 100K\Omega \quad C = .01\mu f$ Paper speed 5mm/second

TABLE II

PERFORMANCE DATA FOR QUADRATIC PREDICTION CIRCUIT

	Triangular Input			Parabolic Input			Sinusoidal Input		
Sample Period	Predicted Value e_b^*	Resulting Value e_o^*	Error $e_b^*-e_o^*$	Predicted Value e_b^*	Resulting Value e_o^*	Error $e_b^*-e_o^*$	Predicted Value e_b^*	Actual Value e_o^*	Error $e_b^*-e_o^*$
1	2.0	6.0	4.0	25.5	14.0	11.5	27.0	18.0	9.0
2	11.0	14.5	3.5	16.0	15.5	0.5	17.5	14.5	3.0
3	19.0	13.0	6.0	16.0	15.5	0.5	17.0	14.5	2.5
4	23.5	10.5	13.0	15.0	14.5	0.5	15.0	13.5	1.5
5	-1.0	8.0	9.0	13.0	12.5	0.5	11.5	11.0	0.5
6	1.5	5.0	3.5	9.5	9.0	0.5	8.0	7.5	0.5
7	3.5	2.0	1.5	5.0	5.0	0.0	3.5	4.0	0.5
8	0.0	-0.5	0.5	0.0	-0.5	0.5	-1.0	0.0	1.0
9	-2.0	-3.5	1.5	-6.5	-5.5	1.0	-4.5	-4.5	0.0
10	-5.5	-6.0	0.5	-9.0	-9.0	0.0	-9.0	-8.0	1.0
11	-8.5	-9.0	0.05	-9.5	-12.0	2.5	-12.5	-11.0	1.5
12	-11.0	-11.5	0.5	-13.5	-13.0	0.5	-14.0	-13.5	0.5
13	-14.5	-14.5	0.0	-14.0	-14.0	0.0	-15.5	-14.5	1.0
14	-16.0	-13.5	2.5	-13.5	-8.0	5.5	-15.0	-14.5	1.0
15	-19.5	-11.5	8.0	-12.0	-6.5	5.5	-13.5	-13.5	0.0
16	-1.0	-7.5	6.5	-9.5	-9.0	0.5	-11.0	-11.0	0.0
17	3.5	-4.0	0.5	-6.0	-6.0	0.0	-7.0	-7.0	0.0
18	-0.5	-1.0	0.5	-1.5	-1.5	0.0	-0.5	-2.5	2.0

waveform which was readily available from a Model 202A Hewlett-Packard Low Frequency Function Generator. In theory, a parabolic waveform can be obtained by integrating a triangular waveform, however, certain practical problems make this not feasible. The circuit diagram of a more useful parabolic generator is shown in Figure 41.

It is difficult to completely eliminate a DC component of voltage in the triangular wave produced by the function generator. The result of integrating this displaced triangular wave is a parabolic wave superimposed on a ramp which would eventually cause the integrating amplifier to become saturated. The use of a first order lag circuit such as illustrated in Figure 3 is a better solution since it will, with a properly chosen time constant, essentially integrate the triangular wave and simply multiply any DC component by a constant.

The transfer function of the first order lag circuit has been derived above (equation 4) and is

$$\frac{E_{OUT}(s)}{E_{IN}(s)} = -\frac{R_f/R_i}{R_f C s + 1} \tag{78}$$

which for sinusoidal output waveforms has been shown (see equation 24) to reduce to

$$\frac{E_{OUT}(j\omega)}{E_{IN}(j\omega)} = -\frac{R_f/R_i}{j\omega R_f C + 1} \quad . \tag{79}$$

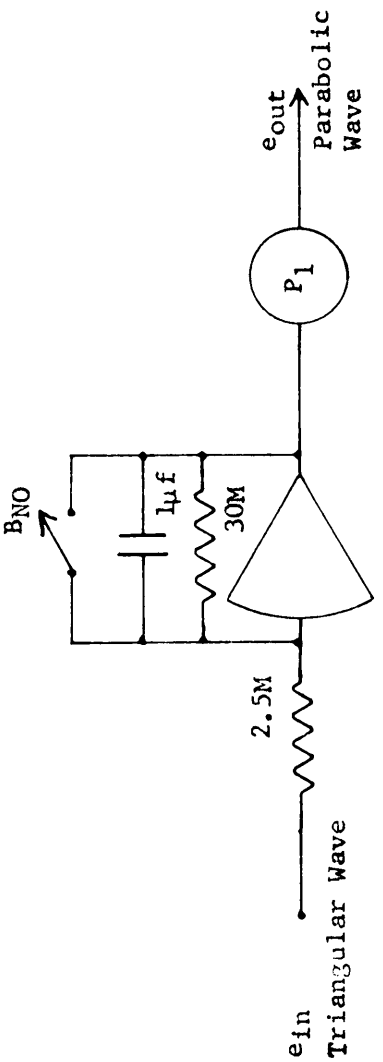


Figure 41 Circuit Diagram of Parabolic Wave Generator

The lowest frequency component in the parabolic wave is the same as the frequency of the triangular input wave. If $\omega R_f C >> 1$ the transfer function becomes

$$\frac{E_{OUT}(s)}{E_{IN}(s)} \approx -\frac{R_f/R_i}{R_f C s} \tag{80}$$

which is essentially integration. The transfer function for the zero frequency (DC) component of the triangular input voltage is simply

$$\frac{E_{OUT}(0)}{E_{IN}(0)} = -\frac{R_f}{R_i} . \tag{81}$$

A parabolic wave generator with the following input signal and circuit parameters was used to evaluate the quadratic prediction circuit.

Input signal - .05 cps triangular wave.

R_f = 30.0 MΩ

R_i = 2.5MΩ

C = 1.0μf

Under these conditions the product R_f C ω is only 1.5 so the output waveform is only approximately parabolic. The actual waveform resulting is shown in Figure 42. The DC component, if any, is simply multiplied by 12 and is not integrated so the problem of

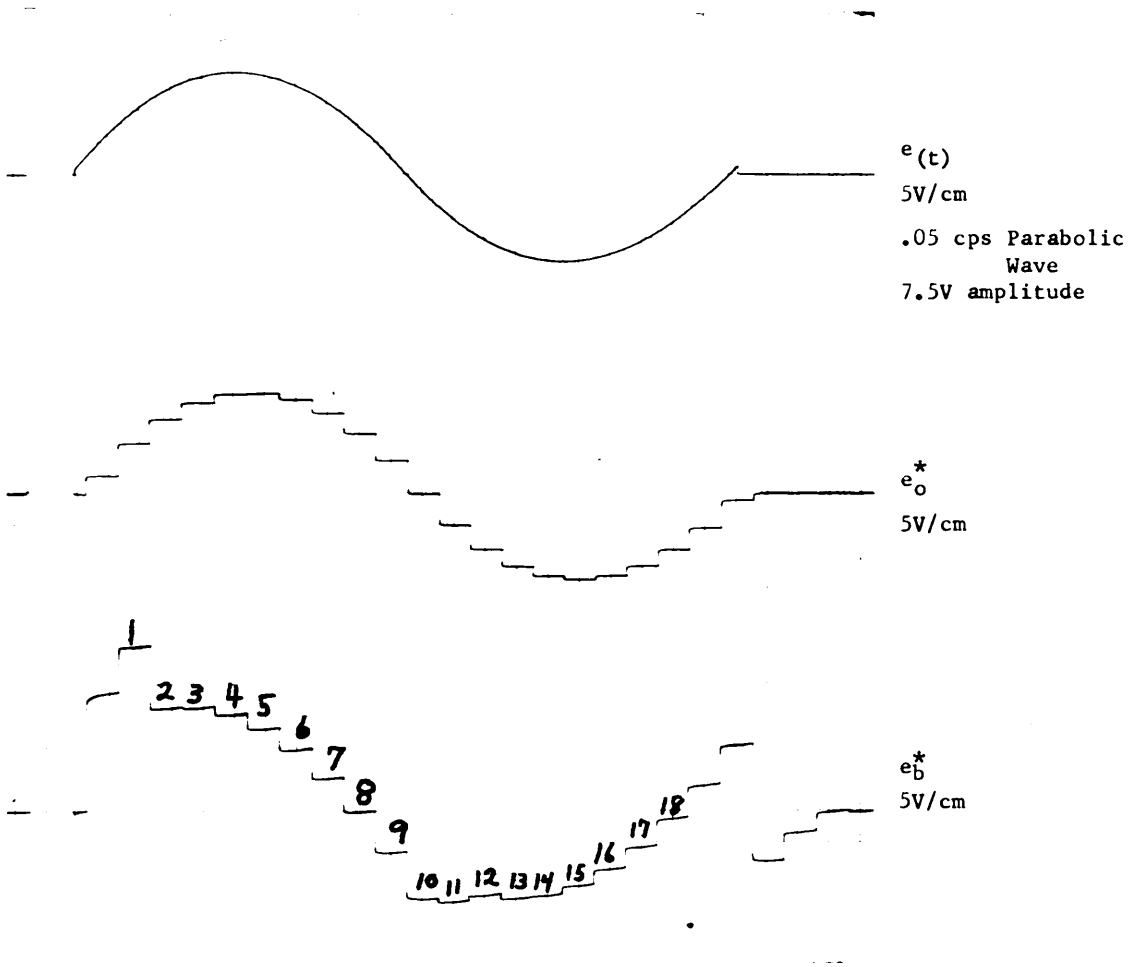


Figure 42 Performance of Quadratic Prediction Circuit

with Parabolic Input

$n = 2 \quad e_b^* = 6e_o^* - 8e_1^* + 3e_2^* \quad 1.0$ sample second

$R_f = R_i = 100K\Omega \quad C = .01\mu f$ Paper speed 5mm/second

1/T potentiometers adjusted experimentally

amplifier saturation does not exist. It is seen that the parabolic wave resulting at the output of the circuit in Figure 41 is superimposed on a DC level rather than on a ramp as would be the case if an ordinary integration circuit were used.

The performance of the quadratic prediction circuit of Figure 39 when subjected to a parabolic input is shown in Figure 42. The predicted values are compared with the resulting values in Table 2. It is seen that, in general, the predicted values agree quite closely with the resulting values as they should. It is concluded that the quadratic prediction circuit will give a more accurate prediction of a future value of a continuous waveform with curvature than that which can be obtained from the equivalent linear prediction circuit.

The performance of the quadratic prediction circuit with sinusoidal input is illustrated in Figure 43. The input waveform is the same as that used to obtain the curves of Figure 37 for the linear extrapolator. The predicted values and resulting values for sinusoidal excitation are also indicated in Table 2. Except for the initial transients it is seen that the predicted values are reasonably close to the resulting values even though the second derivative is not constant. A comparison of the data from Table 2 for quadratic prediction with the data from Table 1 for linear

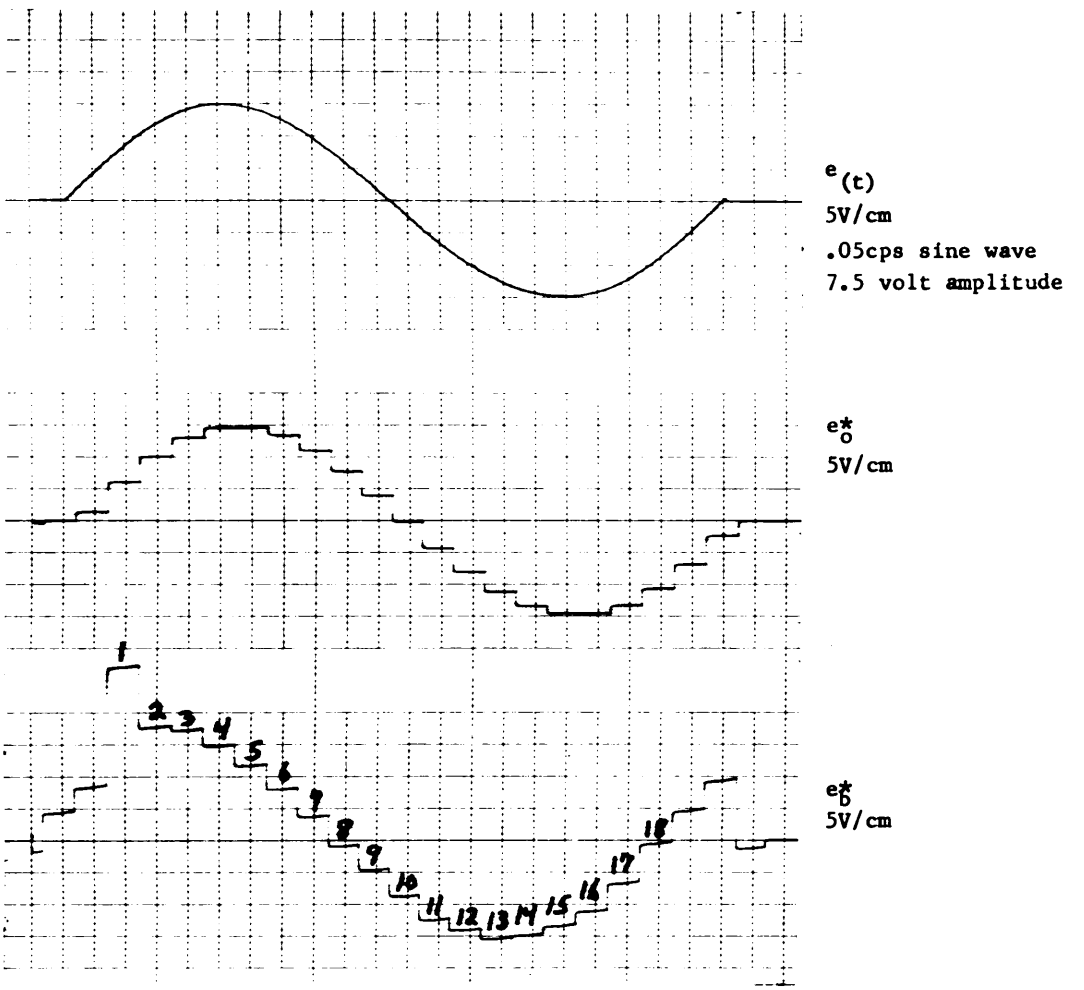


Figure 43 Performance of Quadratic Prediction Circuit with Sinusoidal Input

$n = 2 \quad e_b^* = 6e_o^* - 8e_1^* + 3e_2^*$

1.0 sample/second $R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper speed 5mm/second

prediction reveals that the quadratic prediction is much superior as would be expected.

The object of the investigation discussed in this report is not, however, to evaluate the effectiveness of the various methods of extrapolation but to simply develop methods of simulating these processes on the analog computer. If the extrapolation in a particular system is based on the assumption that a derivative of higher order than the second is constant the extrapolation process can be represented on the computer if desired. This would involve the simulation of higher order hold devices which could be done by an extension of the techniques described above. This would, of course, require considerably more computation equipment. In practice however, hold devices of higher than second order are seldom encountered.

Continuous Linear Extrapolation of a Signal

The methods which have been presented for the prediction of future samples can be extended to give a continuous prediction of the value of the signal itself at a definite time in the future. For example, the predictor might continuously compute the signal expected 1.0 second from the present rather than simply compute a future sample as was done above. Occasionally situations of this type exist in automatic control systems and it is desirable that

methods for simulating the process on the analog computer be developed.

Essentially the method of continuous prediction developed in this investigation is based on predicting from samples the rate of change of the variable (the first derivative) at some future time and then integrating the rate to obtain the value of the variable itself. Several practical complications arise, however, in implementing this operation on the analog computer. As might be expected the predicted value becomes less accurate as the interval of extrapolation is made longer. Also, if higher order derivatives of the signal are considered the predicted signal becomes more accurate. Prediction based on the use of zero, first, and second order hold devices are considered in this investigation and methods of simulating these processes are presented.

The output of a zero order hold device can be thought of as a continuous prediction of the input signal. It has already been stated that the output of a zero order hold device is proportional to the most recent sample received and remains constant until another sample is obtained. Methods for the simulation of such hold devices have already been discussed and are shown in Figures 9 and 16. These circuits will now be considered as a form of a continuous prediction device.

The zero order hold continuously predicts future values of the input signal based on the assumption that the value of the signal will not change from the value that existed at the last sampling instant. This is, of course, a very crude form of continuous extrapolation and unless the input signal is of constant amplitude the predicted values will be in error. Zero order hold devices are quite commonly used in practice because they are relatively simple, inexpensive, and raise the effective gain of the system. In general, zero order hold devices are not thought of as hold devices but are considered as such now in order to compare them with the more elaborate continuous prediction devices to be discussed.

The accuracy of continuous prediction using a zero order hold device can be improved upon considerably by using a first order hold device. The technique can be explained with reference to Figure 44. In this method the rate of change of the input variable (i.e., the first derivative) is computed from the two most recent samples. The value of some arbitrary future sample, at m, is computed using the methods previously discussed. The interval of extrapolation in this case is mT where m need not be an integer. It is desired to continuously predict the value of the input voltage from m to n, a duration of one sample period.

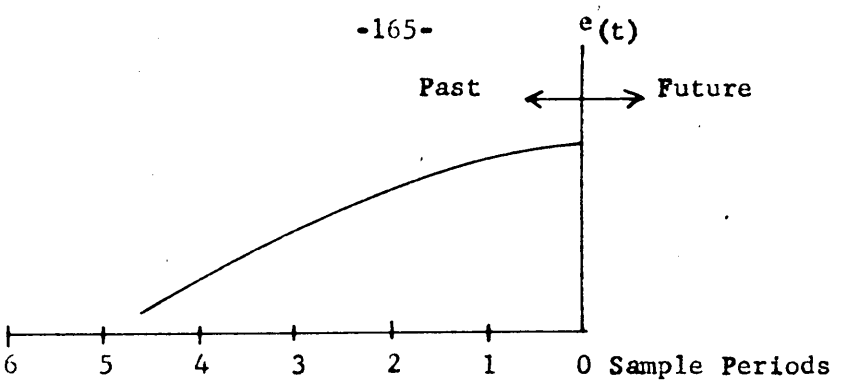


Figure 44a Arbitrary Continuous Signal

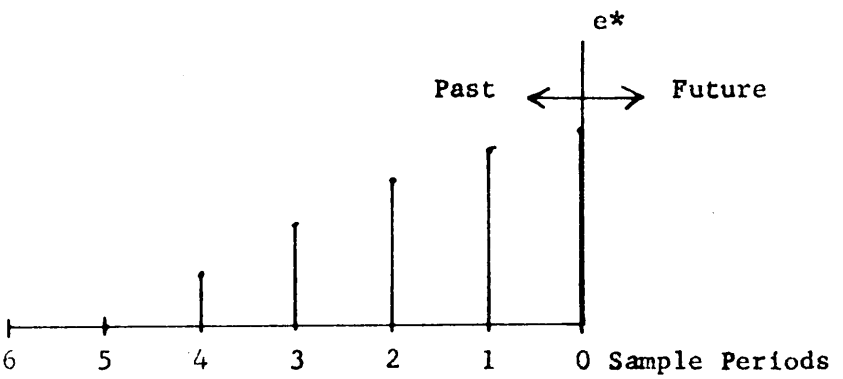


Figure 44b Sampled Signal with Short Sampling Time (without hold)

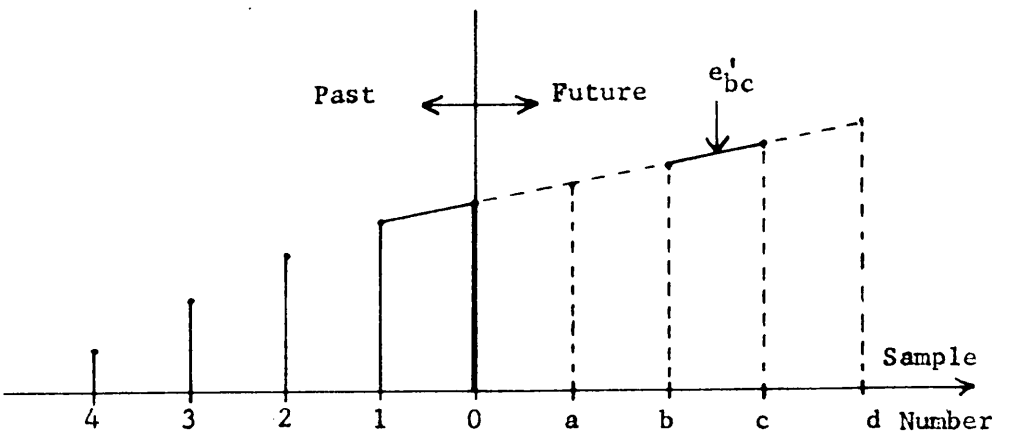


Figure 44c Predicted Continuous Signal Based on Constant First Derivative

This can be computed as follows.

$$e'_{m,n} = e^*_m + \int_0^t M_{m,n}\, dt \qquad (82)$$

where

$e'_{m,n}$ is the continuously extrapolated value of the signal from m to n.

e^*_m is the predicted value of the sample to occur at m.

$M_{m,n}$ is the predicted slope of the signal from m to n.

The value of e_m* can be computed using the linear prediction technique previously described by satisfying equation 49. In linear prediction it is assumed that the slope of the signal is such that $M_{m,n} = M_{1,o}$. $M_{1,o}$, however, can be computed from the present and one previous sample as indicated from equation 42. This can be substituted for $M_{m,n}$ in equation 82 resulting in

$$e'_{m,n} = e^*_m + \int_0^t \frac{e_o^* - e_1^*}{T}\, dt. \qquad (83)$$

However, e_m* can be expressed in terms of equation 49 yielding

$$e'_{m,n} = (m+1)\, e_o^* - m e_1^* + \frac{1}{T}\int_0^t (e_o^* - e_1^*)\, dt. \qquad (84)$$

At the end of each sample period both $M_{1,o}$ and e_m* are computed from new samples and the predicted value is corrected. Since in linear extrapolation techniques it is assumed that the first

derivative of the signal is constant the continuously predicted values will only be correct for signals that vary linearly with time. It is not, however, necessary that the input signal remain constant to obtain accurate prediction as was true in the zero order case.

In order to simulate linear continuous prediction it is necessary to perform the mathematical operations indicated by equation 84. In theory the representation of this equation on the analog computer is not difficult; however, several practical problems are encountered. A circuit for linear continuous prediction is shown in Figure 45. The theoretical operation of this circuit will first be explained and then some of the practical imperfections are discussed.

The inputs, e_1^* and e_o^*, to the linear continuous prediction circuit are obtained from a two section delay line (first order hold simulator) such as those previously described and illustrated in Figure 19 and Figure 21. The average slope between the last two samples ($M_{1,o}$) is obtained from the solution of equation 42 using amplifier 2 and potentiometer p_1. The potentiometer setting can be determined by calculating 1/T or experimentally as described below. Amplifier 3 or 4 integrates the slope giving a ramp $(M_{1,o})t$ which, based on the assumption of constant first derivative, is the predicted value of the continuous signal, except for a constant, at any

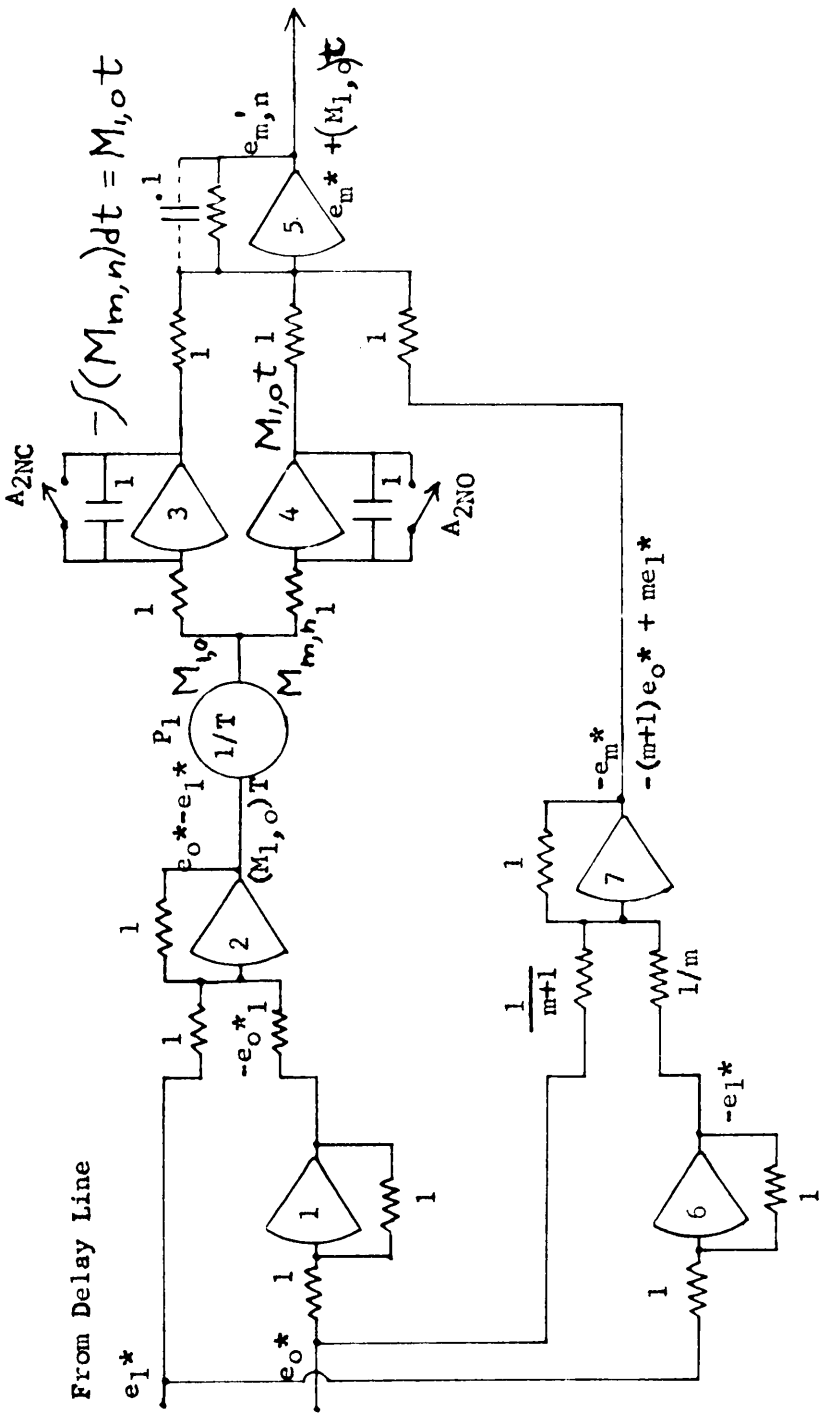


Figure 45 Diagram of Linear Continuous Prediction Circuit

future time. The predicted value of a sample at mT seconds in the future is added to the output of the integrator using amplifier 5. The voltage resulting at the output of this amplifier satisfies equation 84 and is a continuous prediction of values of the signal being sampled at a time mT seconds in the future. The prediction is, of course, based on the assumption that the first derivative of the continuous input signal will remain constant.

The predicted value of a sample mT seconds in the future can be obtained using the techniques for linear extrapolation of a future sample which have been described above and summarized by equation 49. This equation is solved in amplifier 7 and the results added to the outputs of the integrators as described above.

The situation as described to this point exists only for one sample period after which e_o^* becomes e_1^* and a new value of e_o^* appears. A new slope and predicted initial value are computed based on the more recent samples and the previous prediction is corrected. The cycle of a continuous prediction followed by a corrected prediction continues indefinitely.

It is seen from Figure 45 that two integration circuits are incorporated in the linear prediction circuit; however, at any given time only one is integrating, the output of the other being clamped to zero. The necessity for two integrators is

because of practical difficulties involved in setting the integrators to zero and re-starting them within a very short time. If a single integrator were used it would be necessary for relay A_2 to close and then re-open very rapidly at the instant of sampling and during the finite time that the relay was closed errors would be introduced into the prediction. These difficulties are reduced somewhat by using two integrators each of which performs the required integration during alternate sample periods. This is easily accomplished using a single relay, A_2, to reset both integrators by using the normally closed contacts on one integrator and the normally open contacts on the other as shown in Figure 45. With this arrangement there is no necessity for a relay to close and then immediately re-open which would require rather critical timing.

The operation of relay "A_2" must be synchronized with the sample periods such that it will change state at precisely the sampling instant. The cyclic rate of relay "A2" is, therefore, exactly half the sampling rate. This lends itself very nicely to use with the parallel connected delay line of Figure 19 since the timing conditions imposed on relay "A_2" are the same as those for the "A" relays of the delay line. It is necessary to simply use a set of normally open and a set of normally closed contacts on

the same relay that is controlling the delay line to set the initial conditions on the integrating amplifiers in Figure 45.

It has been previously shown that there is considerable advantage in using series connected storage elements in a delay line as compared with the use of parallel connected elements. It has also been shown that the "A" relays in the series connected delay lines operate at the sampling frequency. This is, of course, twice the rate that it is desired to operate the "A_2" relay in the continuous prediction circuit, so the method of setting initial conditions into the integrators is not as simple as when using the parallel connected delay line. It is seen that for reasons of simplicity it is preferable to use a parallel connected delay line in conjunction with a continuous prediction circuit provided the inferior performance of the delay line itself can be tolerated.

If it is desired, however, to use a series connected delay line because of its superior performance it is necessary that relay A_2 in the prediction circuit operate at exactly one half the cyclic rate of the "A" relays in the delay line. It is, furthermore, necessary that relay "A_2" change state simultaneously with the start of each cycle of "A" relay operation. This necessitates the use of a frequency divider device which is synchronized with the relays in the delay line. It is, of course, desirable that the device use

only standard analog computer components. One such circuit which has been used successfully is shown in Figure 46a.

The operation of this circuit can be explained using the timing diagram of Figure 46b. Assume for purposes of explanation that the output of integrating amplifiers 2 and 4 is initially set to zero. Relay "A_1" is one of the "A" relays in a computer which is programmed as a series connected delay line and is cycling at a rate of one cycle per sample period. Further assume that relays I and II are initially de-energized and also that in the process of cycling the "A" relays, and hence relay A_1, have just returned to the de-energized state. This is the situation indicated at t = 0 in Figure 46b. The effective input voltage to amplifier 2 at this time is +50 volts and because of the small RC product the output of the amplifier increases rapidly in the negative direction until it becomes saturated after which time the output voltage remains constant at a high negative value (about -200 volts on the ES-400 computer). With this condition existing diode D_1 conducts and relay I becomes energized almost immediately after relay "A_1" has become de-energized. Presently however, at t = 1 relay A_1 becomes energized and the input voltage to amplifier 4 becomes +50 volts since relay I is now energized. The output of this amplifier then rises rapidly to negative saturation causing relay II to become energized. In the meantime, the effective

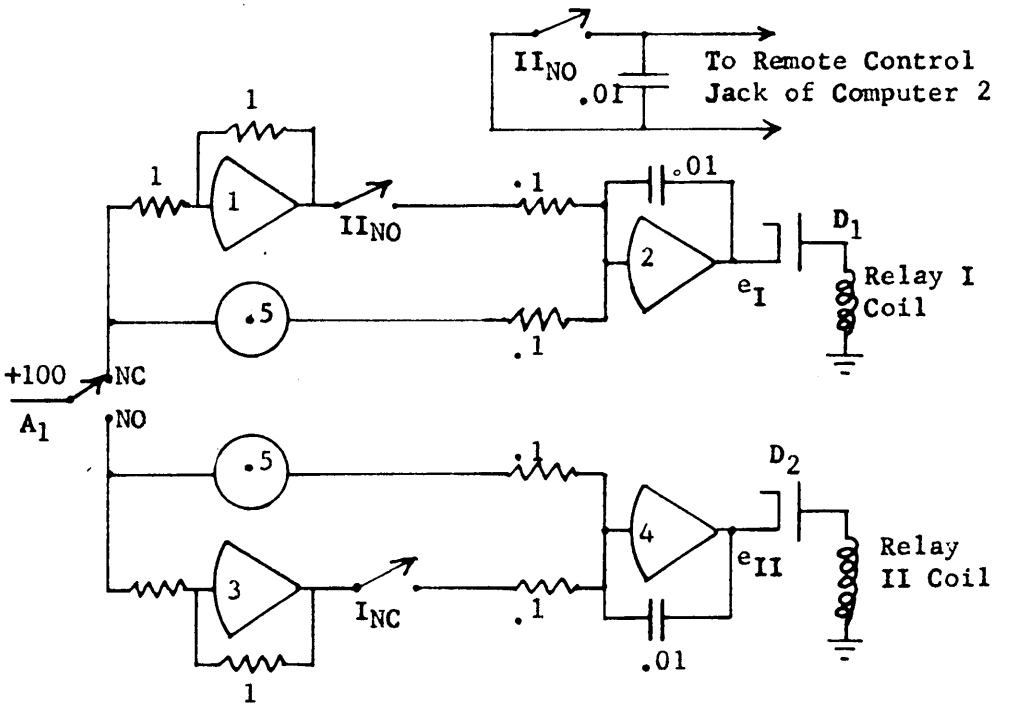


Figure 46a Circuit Diagram for Synchronized Frequency Divider

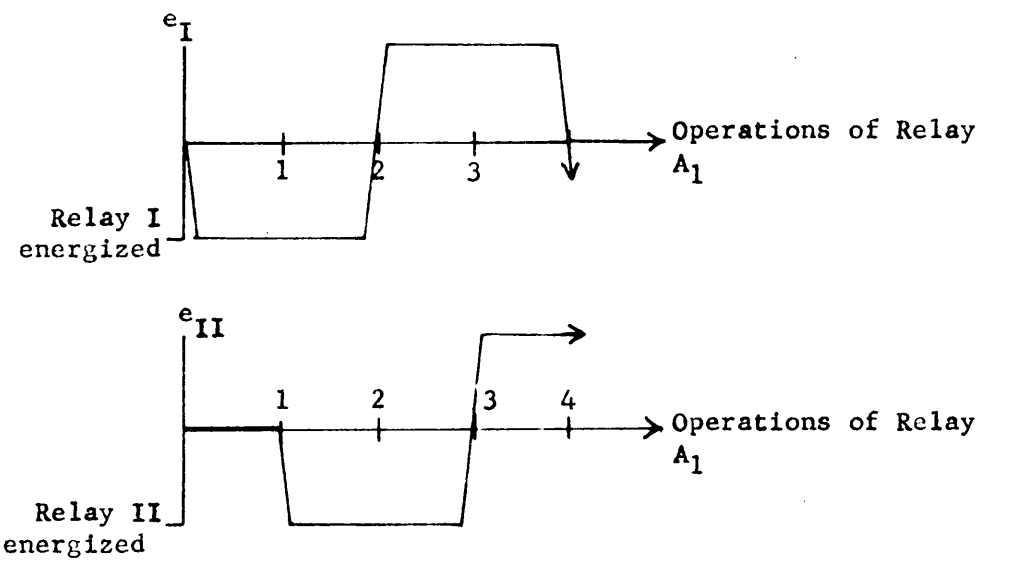


Figure 46b Timing Diagram for Synchronized Frequency Divider

input voltage to amplifier 2 is zero, since relay A_1 is now energized, such that its output voltage does not change and relay I remains energized.

Eventually relay A_1 begins a second cycle becoming de-energized at t = 2. Since relay II is now energized the effective input voltage to amplifier 2 is now -50 volts and it begins rapidly integrating from negative saturation through zero and in a short time reaches positive saturation. Since the diode D_1 is now reverse biased no current will exist and relay I becomes de-energized. In the meantime, from t = 2 to t = 3, the effective input voltage to amplifier 4 is zero causing its output to remain constant and relay II to remain energized.

At t = 3 relay A_1 again becomes de-energized causing the effective input voltage of amplifier 4 to become -50 volts and, in turn, causing relay II to become de-energized.

At t = 4, relay "A_1" begins a third cycle. The circuit conditions are similar to those that existed at t = 0 except that the initial voltage of amplifier 4 is now about +250 volts rather than zero. This has no appreciable effect, however, since the gain of the integrator is quite high. In fact, it is now seen that the original assumption of zero initial conditions was not necessary and no provision for setting the initial condition of amplifiers 2 and 4 need be incorporated into the circuit.

The relay timing diagram of Figure 46b shows that relays I and II operate at exactly half the frequency of the "A" relays in the delay line and that the operation is synchronized such that relay I changes state each time relay "A_1" becomes de-energized. Relay II, however, changes state each time relay "A_1" becomes energized. The timing of relays I and II is exactly that required of relay "A_2" in the continuous prediction circuit of Figure 45.

An undesirable feature of the frequency divider circuit of Figure 46 is the brief time delay occuring between operation of relay "A_1" and the subsequent operation of relays I or II. In order to minimize this lag it is necessary to use integrator circuits with high gain (i.e., a small RC product). The values indicated in Figure 46, giving a gain of 1000, are about the smallest available on a general purpose analog computer. Using these values, considering the amplifier saturation voltage as 200 volts and the relay pull in voltage as 70 volts, will result in a maximum electrical time delay of about 5 milliseconds. This is not significant at the sampling rates which are likely to be encountered in the computer. Probably more significant are the mechanical and electrical time constants of the relay itself.

The operational relays often included in general purpose analog computers could serve as relays I and II in the frequency divider circuit. The Heath ES-400 Computers used in this investigation

do not have such relays so it was necessary to adapt two telephone type relays as a substitute. If polarized operational relays were available it would not be necessary to use the diodes D_1 and D_2 illustrated in Figure 46. Another simplification which could be employed is the elimination of the two potentiometers and replacing the .1M amplifier input resistor following each potentiometer by one of .2M.

In the investigation described in this report actually two Heath ES-400 computers were used for convenience. The first computer was used to simulate the delay line, the "A" relays being controlled by the master timing oscillator as previously discussed. One of the "A" relays in that computer also served as relay "A_1" in the frequency divider circuit of Figure 46. A set of normally open contacts on relay II were connected across switch "A" (see Figure 12) in the second computer via the remote control jack. Again it was found that a .01 μf capacitor connected across the jack was necessary to minimize transients induced into the computing circuits by the switching operations. The "A" relays in computer B then operate nearly simultaneously with relay II in the frequency divider circuit and in turn almost simultaneously with the "A" relays in computer A but at half the cyclic rate. An "A" relay in computer B was then used as relay "A_2" in the continuous prediction circuit of Figure 45. The remaining amplifiers

in both computers are available for simulation of the continuous portion of the sampled-data system. The "B" relays on both computers can be used to set the initial conditions into the continuous system.

The correct setting of the 1/T potentiometer can be obtained by calculation or experimentally. For a given sample period, T, the potentiometer setting can be calculated and the value set using the usual method of setting coefficient potentiometers. A better approach is to set the potentiometer experimentally, which will account for errors in the sampling rate and inaccuracies in the computer components.

It has been shown above that theoretically a linear continuous prediction circuit will have no error if the waveform being sampled has a constant slope. Such waveforms include ramps and, except at the peaks, triangular waves. The 1/T potentiometer can be adjusted experimentally by feeding a triangular waveform into the linear continuous prediction circuit while viewing the output voltage with a recorder, and adjusting the potentiometer until a smooth curve results. The output waveform will, of course, be discontinuous at the peaks, however, this is ignored during adjustment of the potentiometer.

The performance of an actual linear continuous prediction circuit such as that illustrated in Figure 45 is shown in Figure

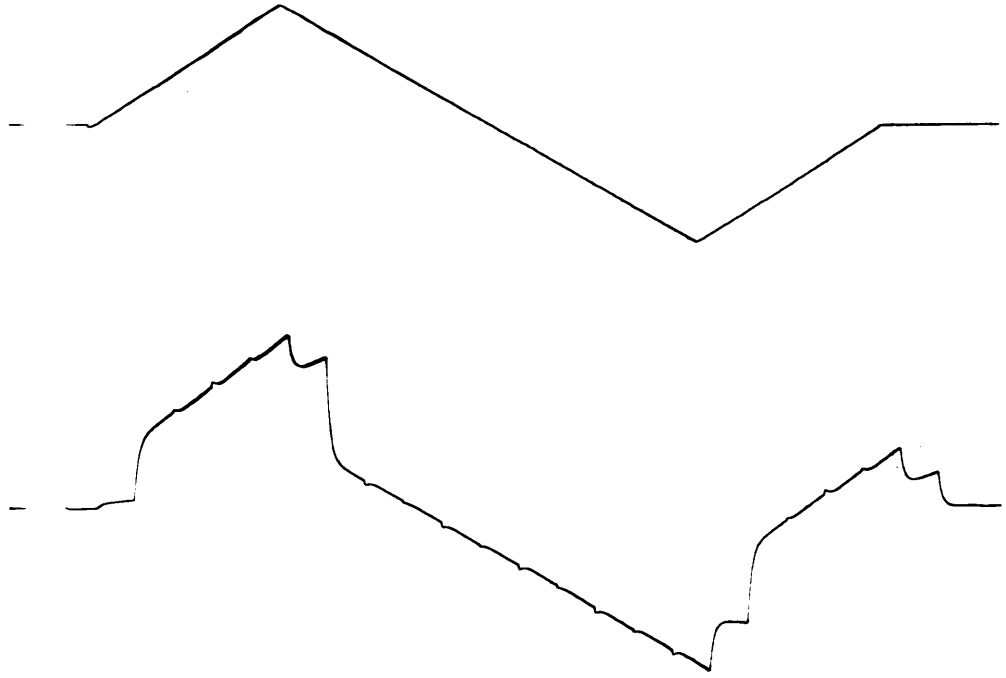


Figure 47 Performance of Linear Continuous Prediction Circuit with Triangular Input

1.0 sample-second $R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper speed 5mm/second $\frac{1}{T}$ potentiometer adjusted experimentally

47. The continuous input voltage was triangular in form and the circuit parameters were selected to continuously predict the value of the input voltage exactly two sample periods in advance. The 1/T potentiometer was adjusted experimentally. Close examination of these curves indicates that the performance of the circuit is quite satisfactory except during starting transients and at the peak of the triangular wave. This is expected, however, since the assumption of a constant first derivative is not valid at these points.

A transient of short duration occurs at each sampling instant. This is a result of the fact that all relays in the simulation circuit are not operating simultaneously as they theoretically should. The transient spikes can be either upward or downward depending on the relative timing of the relays. They are of such short duration that they would have little effect on the overall system performance, however, for best accuracy it is desirable that they be eliminated by suitable filtering. The filtering can be and was done using several techniques, but the most simple is to shunt the feedback resistor in the output amplifier of the continuous prediction circuit with a small capacitor. This capacitor is shown by dotted lines around amplifier 5 in Figure 45. The amplifier becomes, in effect, a low pass filter eliminating the

spikes, smoothing out the predicted value, and unfortunately, tending to limit the frequency response of the system. A .1μf capacitor was found to be a suitable compromise although the value is not critical. Such a capacitor was incorporated in the circuit when the data shown in Figure 47 was obtained.

The effect of misadjustment of the 1/T potentiometer is shown in the theoretical curves of Figure 48. It is seen that a discontinuity will result in the predicted value, assuming ramp or triangular input, if the potentiometer is misadjusted in either direction. This results from the fact that the integrators in the continuous prediction circuit are integrating either too fast or too slow such that the value at the output of the circuit at the end of one sample period (i.e., the end of the integration process) is not equal to the next predicted sample. In cases where the slope of the continuous input is not constant, discontinuities will exist in the output waveform since at the start of each sample period a necessary correction is made in the predicted value.

The response of the linear continuous prediction circuit to sinusoidal excitation is illustrated in Figure 49. The circuit parameters were identical to those existing when the triangular data of Figure 47 was obtained. It is seen that the prediction was quite satisfactory during the portion of the cycle where the

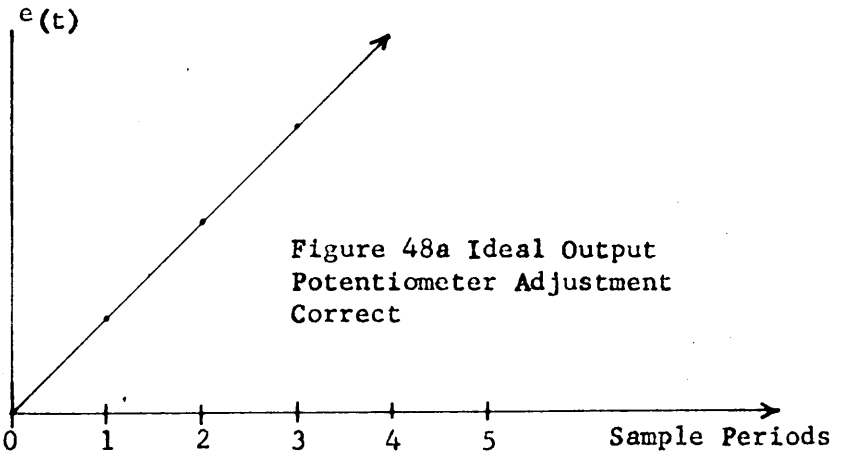


Figure 48a Ideal Output
Potentiometer Adjustment
Correct

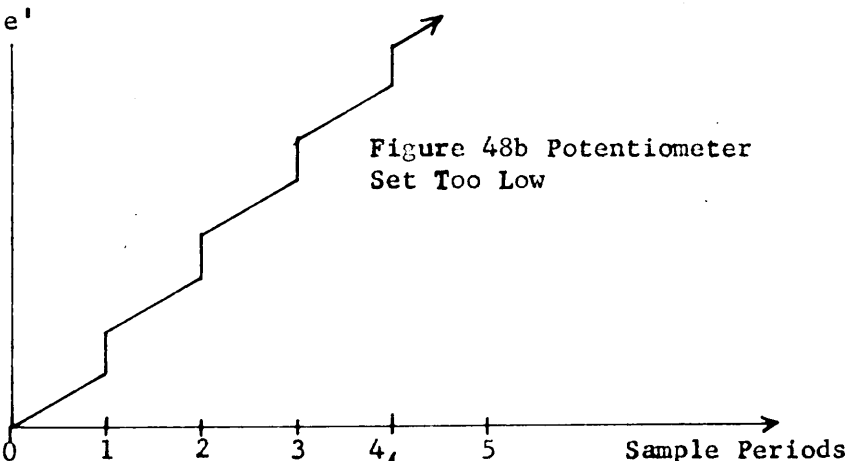


Figure 48b Potentiometer
Set Too Low

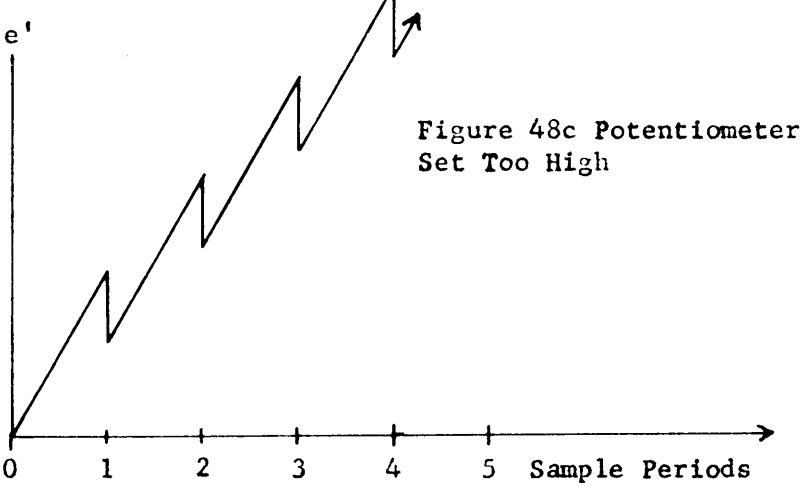


Figure 48c Potentiometer
Set Too High

Figure 48 Adjustment of 1/T Potentiometer in Linear Continuous Prediction Circuit. (Ramp or Triangular Input)

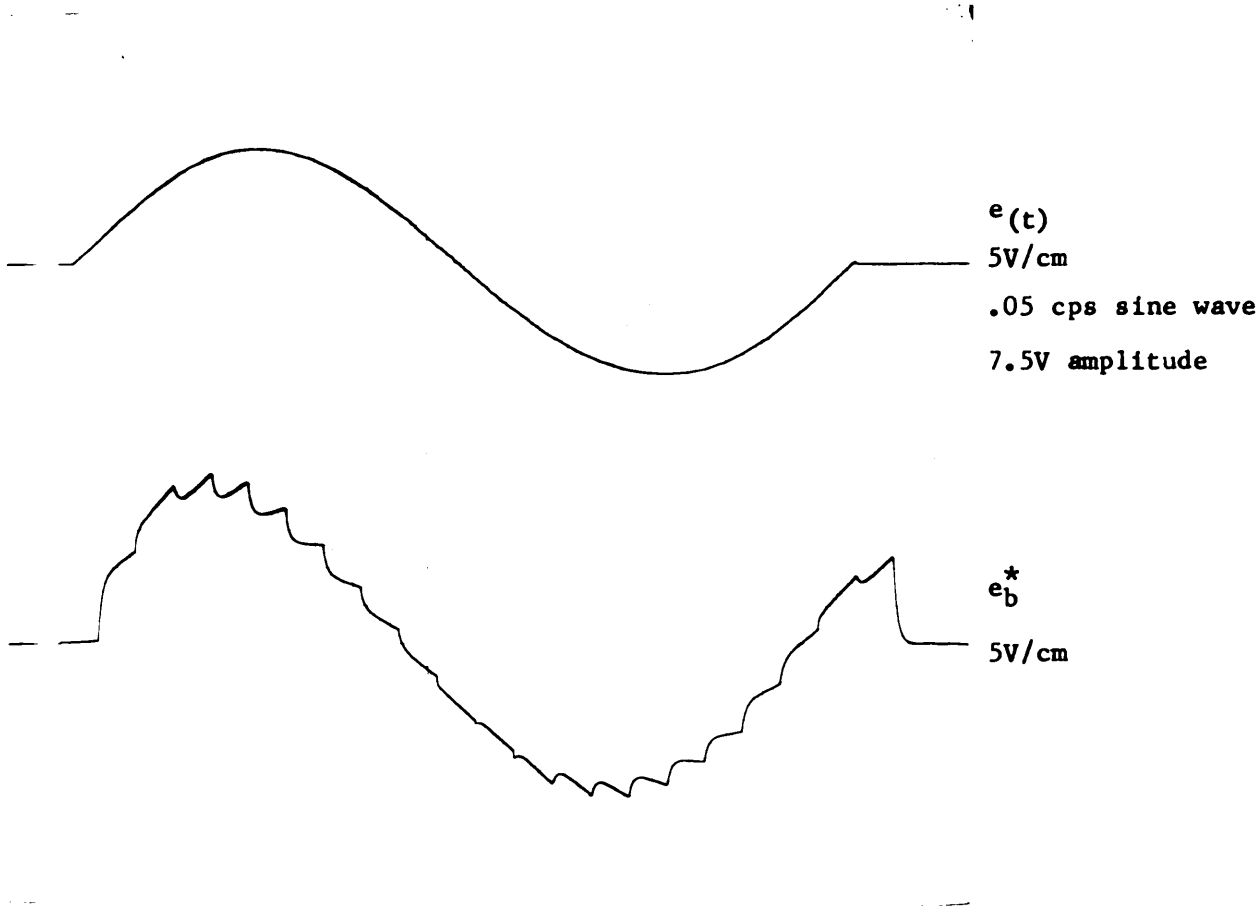


Figure 49 Performance of Linear Continuous Prediction Circuit with Sinusoidal Input

1.0 sample/second $R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper speed 5mm/second

$\frac{1}{T}$ Potentiometer adjusted experimentally

slope was changing slowly, however, considerable error existed at the peaks of the wave where the slope was changing rapidly. Note that the prediction is corrected at the start of each sample period. The .1μf capacitor around amplifier 5 was again used to smooth the transients at the output of the prediction circuit.

Continuous Quadratic Extrapolation of a Signal

It has been demonstrated that continuous data extrapolated from a first order hold circuit, considering two samples, is more accurate than that obtained from a zero order hold device considering only one sample. The technique can be extended to quadratic continuous prediction, using three samples, by assuming that the second derivative of the signal will remain constant at the average value existing through the last two sample periods. The value of the rate of change of slope (i.e., the second derivative) can be computed from samples as indicated by equation 68 and which was previously done in the quadratic prediction of future samples.

The technique of quadratic linear prediction can be explained with reference to the theoretical curves of Figure 50. It is desired to continuously predict the value of the continuous input voltage over a period of one sample period, from m to n, at a time mT seconds in the future. The prediction is to be based on the assumption that

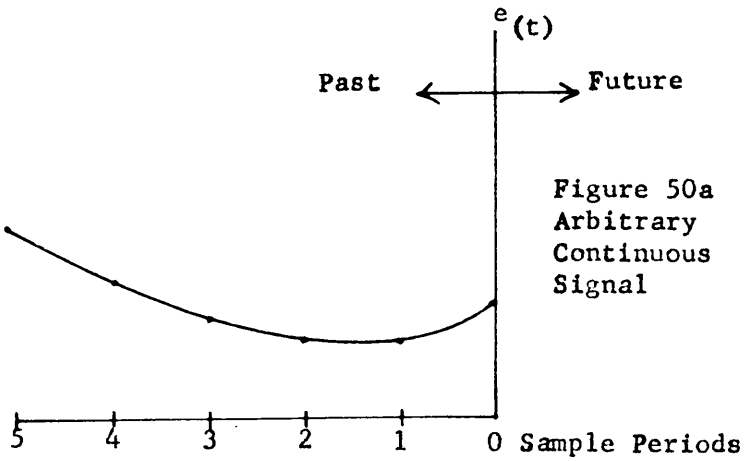


Figure 50a Arbitrary Continuous Signal

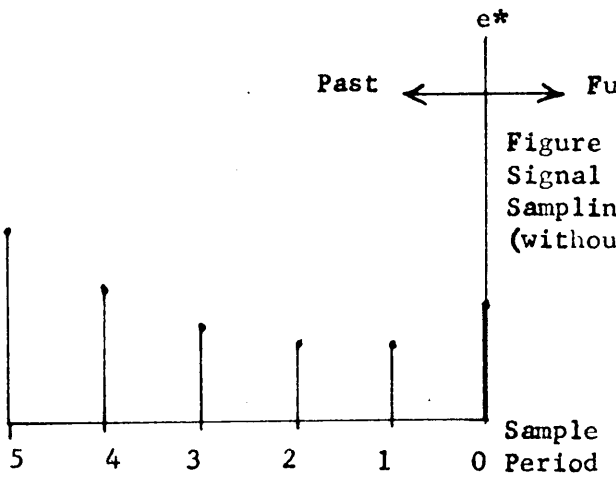


Figure 50b Sampled Signal with Short Sampling Time. (without hold)

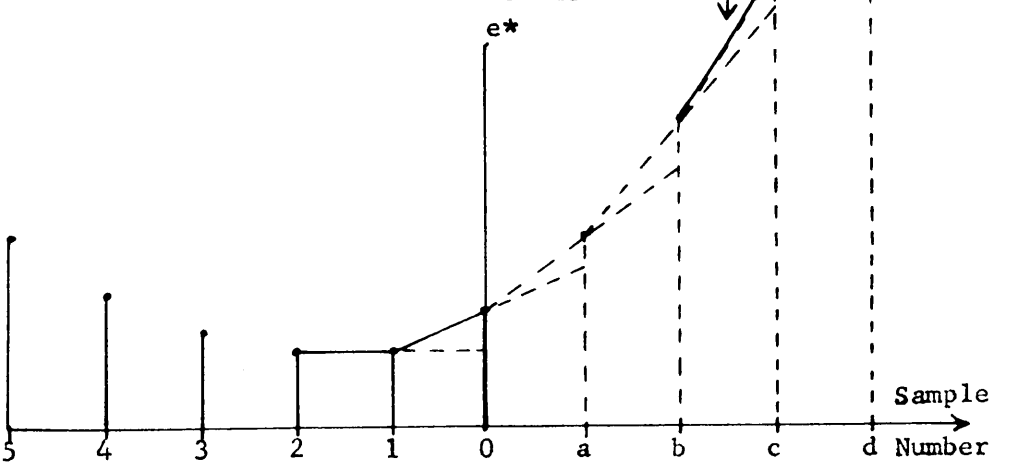


Figure 50c Predicted Continuous Signal Based on Constant Second Derivative

the rate of change of slope (i.e., the second derivative) of the input signal will remain constant at its present value. The approach is similar to that used in continuous linear extrapolation and previously indicated by equation 82. The equation is restated as

$$e'_{m,n} = e^*_m + \int_o^t M_{m,n}\, dt . \tag{85}$$

In the case of linear extrapolation it was assumed that $M_{m,n}$ was a constant and equal to $M_{1,o}$ which was computed from two samples. In quadratic continuous prediction the slope $M_{m,n}$ is computed as follows

$$M_{m,n} = \int_o^t M'_{m,n}\, dt + M^*_m \tag{86}$$

where $M_{m,n}$ is the continuously predicted slope from m to n.

$M'_{m,n}$ is the rate of change of slope from m to n (i.e. the second derivative).

M^*_m is the predicted value of the slope at time m.

In quadratic prediction it is assumed that the second derivative will remain constant at approximately the present value, $M'_{2,o}$, such that the following relation is true.

$$M'_{m,n} = M'_{2,o} \tag{87}$$

Since $M'_{2,o}$ can be obtained from samples as indicated by equation 68 it is possible to write equation 87 as

$$M'_{m,n} = \frac{e_o^* - 2e_1^* + e_2^*}{T^2} \quad . \tag{88}$$

This when substituted into equation 86 yields

$$M_{m,n} = \frac{1}{T^2} \int_o^t (e_o^* - 2e_1^* + e_2^*)\, dt + M_m^* \quad . \tag{89}$$

The value of M* can be computed using equation 56 as follows.

$$M_{m,n}^* = M_{1,o} + m(\Delta M) \tag{90}$$

which in terms of the sample voltages becomes

$$M_{m,n}^* = (m+1)\, e_o^* - (2m+1)\, e_1^* + m\, e_2^* \quad . \tag{91}$$

The value of e_m^* can also be computed from samples using equation 76 which becomes for this case

$$e_m^* = (m+S+1)\, e_o^* - (2S+m)\, e_1^* + S\, e_2^* \tag{92}$$

where S = (1+2+3+ ---------- +m)

A circuit for implementing continuous quadratic prediction based on the simultaneous solution of equations 85, 86, 89, 90, and

92 is illustrated in Figure 51. The circuit is essentially two of the continuous linear prediction circuits working together, one continuously predicting the rate of change (solving equation 86) and the second predicting the value of the variable itself (solving equation 85). Note that the amount of computing equipment required increases rapidly with the order of the prediction and number of samples considered. The theoretical operation of the circuit is explained below; however, the same practical problems existing with the linear continuous prediction circuit exist in the quadratic circuit and are to some extent compounded.

The inputs to the prediction circuit e_o^*, e_1^*, and e_2^* are obtained from a three section delay line (second order hold) such as that described previously and illustrated in Figures 24 and 26. The average slope period product between the last two samples, e_o^* and e_1^*, is obtained from the solution of equation 65 with amplifier 3 in Figure 51. In a similar manner, the slope period product between samples e_2^* and e_1^* is obtained with amplifier 4. The average rate of change of slope (i.e., the second derivative) multiplied by T^2 is computed with amplifier 6 by solving equation 88. The sample period is divided out by potentiometer P_1 resulting in $T\,M'_{2,o}$. Integrating amplifiers 7 and 8 are used in the same manner as amplifiers 3 and 4 in the

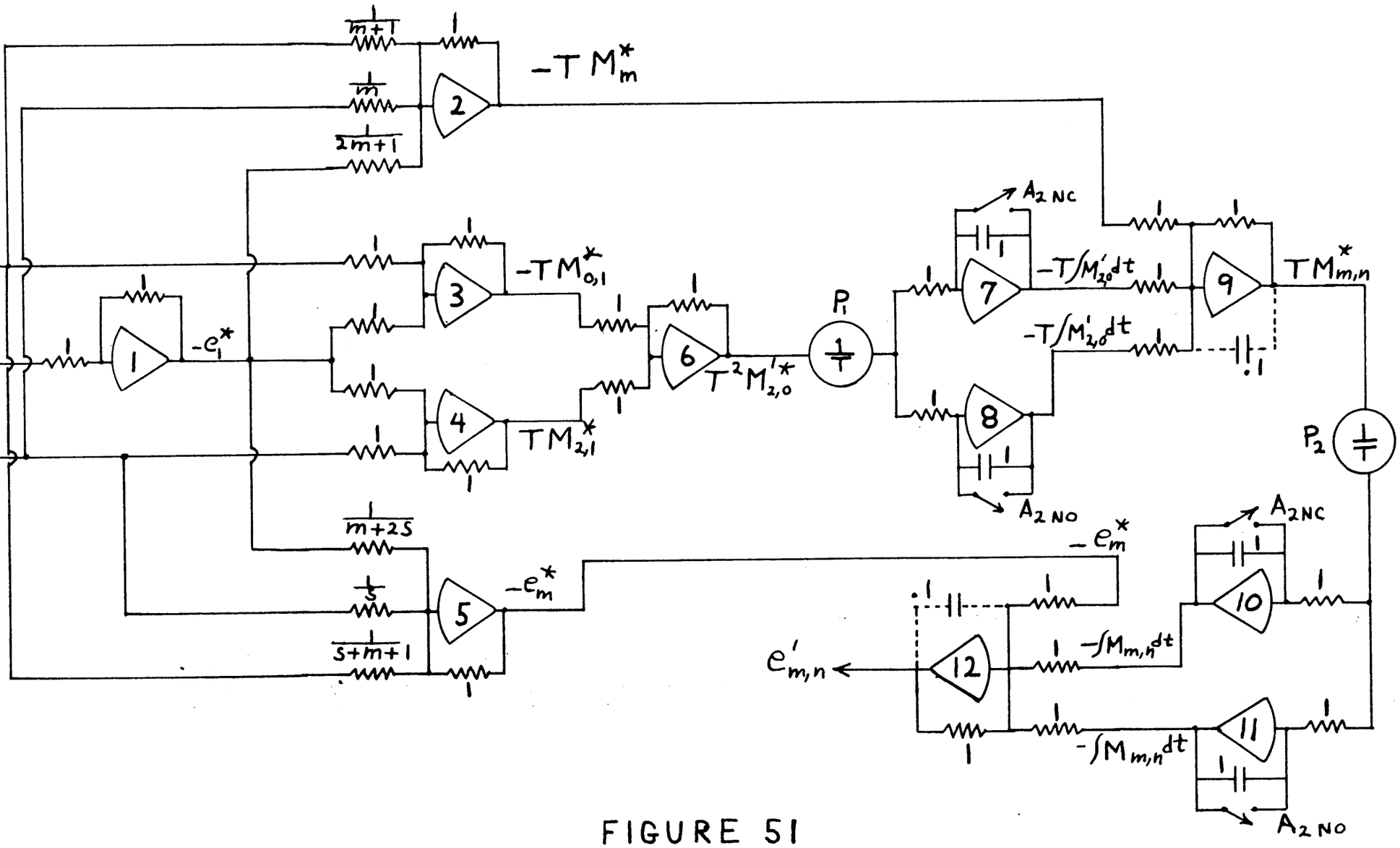


FIGURE 51

QUADRATIC CONTINUOUS PREDICTION CIRCUIT

linear prediction circuit of Figure 45 to generate the time varying component of the slope period product.

Equation 86 is solved with amplifier 9 and potentiometer P_2 resulting in a continuous prediction of the slope based on the assumption of a constant second derivative. The constant portion of the slope period product is computed with amplifier 2 by solving equation 90 and then added to the time varying portion obtained from integrators 7 and 8 with amplifier 9.

The remainder of the circuit is essentially the same as that of the linear continuous prediction circuit of Figure 45. The time varying portion of the predicted output is generated by integrating amplifiers 10 and 11. The constant portion of the output voltage is obtained from the solution of equation 92 with amplifier 5 and is added to the time varying portion by amplifier 12 satisfying equation 85. The output voltage of amplifier 12 is a continuous prediction of the value of the continuous voltage being sampled at a time mT sample periods in the future.

At the end of each sample period e_o^* becomes e_1^*, e_1^* becomes e_2^*, and a new value of e_o^* appears. The new values are used to correct the prediction and this cycle of a continuous prediction followed by a corrected prediction continues indefinitely. Since the predictions are based on the assumption of a constant first

derivative they will be accurate for step, ramp, and triangular input waveforms, except for transient conditions, but, in addition, will be accurate for parabolic waveforms for which the first derivative is not constant. Since the quadratic prediction circuit can follow curved waveforms it is apparent that for arbitrary waveforms the quadratic continuous prediction circuit will give superior performance as compared with linear continuous prediction techniques.

The function of integrating amplifiers 7, 8, 10, and 11 in the quadratic extrapolation circuit of Figure 51 is similar to that of amplifiers 3 and 4 in the linear prediction circuit of Figure 45. All of the "A_2" relays in the quadratic circuit must operate at half the cyclic rate of the "A" relays in the delay line. The practical problems and limitations that exist in the linear circuit involving the control and operation of the "A_2" relays also exist in the quadratic extrapolation circuit and are resolved in the same manner as before.

In the investigation described in this report the delay line was programmed on one ES-400 computer and the quadratic continuous prediction circuit on a second computer. The "A" relays in the second computer, controlled by a frequency divider circuit, were used as the "A_2" relays in the prediction circuit of Figure 51. The details of controlling the relays in the second

computer are identical with those described above for the linear continuous prediction circuit.

The setting of the 1/T potentiometer can be determined by calculation or experimentally. Each potentiometer is set using a slightly different procedure.

Potentiometer P_1 is set with a parabolic waveform fed into the prediction circuit. The parabolic signal can be generated from a triangular waveform using the circuit illustrated in Figure 41 which was described above. The derivative of the parabolic wave is, of course, a triangular wave. A voltage proportional to the predicted derivative appears at the output of amplifier 9 in the continuous extrapolation circuit of Figure 51. This voltage is observed on a recorder and potentiometer P_1 is adjusted to give a smooth triangular waveform, except at the peaks. If the potentiometer is not adjusted correctly a discontinuity will appear in the output waveform at each sampling instant.

Potentiometer P_2 must be adjusted after P_1 is in correct adjustment. A triangular waveform is fed into the continuous prediction circuit and the predicted continuous signal (output of amplifier 12) is observed on a recorder. Potentiometer P_2 is adjusted to give a smooth triangular waveform, except at the peaks. It is seen that the adjustment procedure for potentiometer P_2 in the quadratic prediction circuit of Figure 51 is similar to that

for potentiometer P_1 in the linear prediction circuit of Figure 45.

Small spikes appear at the output of the quadratic continuous prediction circuit (amplifier 12) at each sampling instant as was true in the case of the linear continuous prediction circuit. These transients are, as previously discussed, attributed to the fact that all of the relays involved do not operate simultaneously. Similar spikes appear on the continuously predicted slope at the output of amplifier 9. In both cases the spikes are of such small amplitude and of such short duration that they would have a negligible effect on the operation of the overall system. If desired, however, they can be reduced considerably by placing a small capacitor in parallel with the feedback resistors of amplifiers 9 and 12 as indicated by dotted lines in Figure 51.

The actual performance of a quadratic continuous prediction circuit of the form indicated in Figure 51 is shown in Figure 52. The input signal is a parabolic wave and the constants of the circuit are selected to continuously predict the output waveform 2.0 seconds in advance (i.e., $m = 2$). It is seen that the quality of the prediction is quite satisfactory except during transient conditions and when the parabolic wave passes through zero. Under these conditions the second derivative is not constant and, as expected, this results in an erroneous prediction. Potentiometers P_1 and P_2

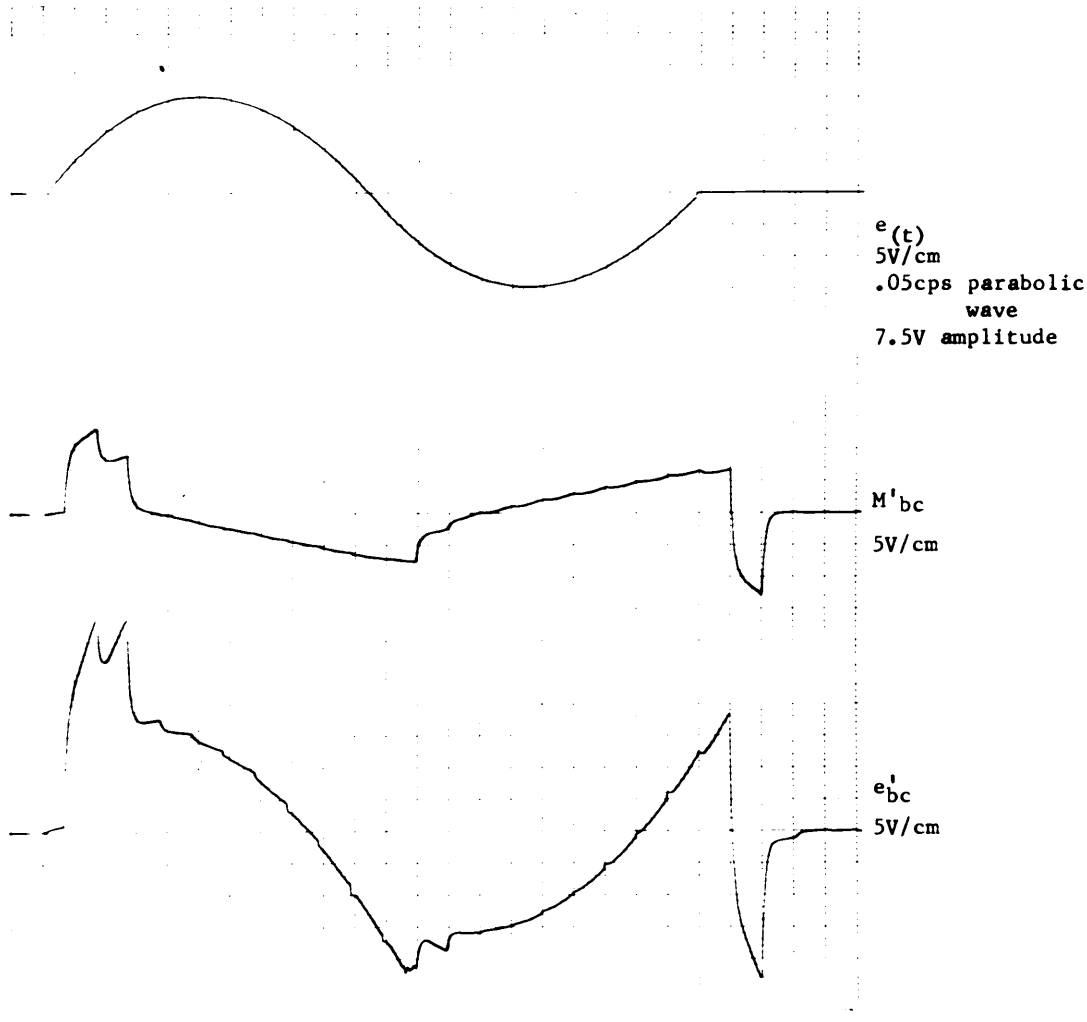


Figure 52 Performance of Quadratic Continuous Prediction Circuit with Parabolic Input

1.0 sample/second Paper speed 5mm/second $R_f = R_i = 100K\Omega$ $C = .01\mu f$

were adjusted experimentally as discussed above. Also shown is the output of amplifier 9 ($M_b'c$) which should be of triangular waveform; however, discontinuities result at the times when the second derivative of the input signal is not constant. Capacitors of .1μf were used around amplifiers 9 and 12 as discussed above to filter the brief transients that occur at the sampling instant.

The performance of the same quadratic prediction circuit but with a sinusoidal input voltage is illustrated by the curves of Figure 53. The input voltage and period of extrapolation are the same as that used with the linear prediction circuit of Figure 45 and resulting in the curves of Figure 49. It is seen that the performance of the quadratic prediction circuit is superior to that of the linear prediction circuit. Some error in prediction does occur, however, since the second derivative of a sinusoidal wave is not constant. Note that the derivative waveform is essentially a cosine wave as expected. Considerable error appears at the start and stop of the sine wave since the second derivative changes rapidly at these points.

The techniques used in continuous prediction can be expanded to third and higher order hold devices, considering four or more samples, if desired. The amount of computing equipment involved increases rapidly however.

It is apparent from a search of the literature that very little

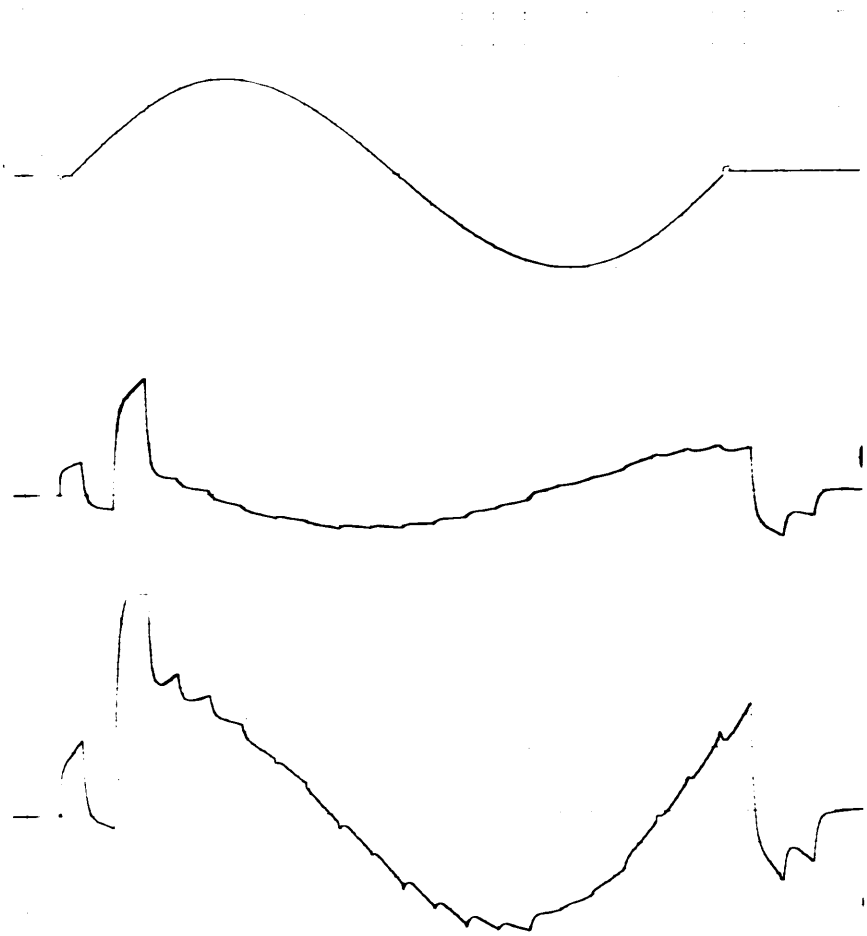


Figure 53 Performance of Quadratic Continuous Prediction Circuit with Sinusoidal Input

1.0 sample/second $R_f = R_i = 100K\Omega$ $C = .01\mu f$ Paper speed = 5mm/second

work has been previously done with continuous extrapolation of sampled signals on an analog computer. Only recently, however, (January 1962) a letter[23] appeared in the IRE Transactions indicating that some work is now going on in this area at Syracuse University under contract from the United States Air Force. It appears that the approach by these investigators is somewhat different from that presented in this report. In general, the method is less difficult to implement on a computer but also somewhat less flexible.

Simulation of Computational Time Delays

It has been previously shown that the output of a hold device is a weighted summation of the most recent sample and perhaps a finite number of past samples. In a practical automatic control system the weighting and summation is performed by some type of digital computer. It has been shown in the introduction of this report that any system involving a digital computer is, in effect, a sampled-data system. The nature of a digital computer is such that a time delay exists between the time the data is entered into the machine and the time that the resulting output or answer appears. This, of course, introduces a fixed time delay which, in general, adversely affects the system performance.

It is desirable that methods be developed for simulating this time delay on an analog computer. This delay can be represented by two distinct methods.

1. Delay the operation of certain relays in the sample-hold simulator and, if involved, the continuous prediction circuit.
2. Use a standard dead time simulator circuit.

The first method gives the superior performance but is more difficult to construct on the analog computer. This method, using relays, results in a reasonably true simulation of delay time while the second method is approximate and somewhat artificial.

The circuit details of delay simulators depend somewhat on how and where they are to be used. This can be explained by referring to Figure 54. The continuous prediction section shown in the figure may or may not be present depending on the particular sampled-data system to be simulated. Since all of the blocks are in series the time delay can be introduced at any point in the series. The exact method of simulating the delay will depend on where it is introduced. The nature of the signal appearing at the various points indicated in Figure 54 is different as follows:

Point A - A continuous function if time.

Points B and C - A series of held samples (staircase).

Point D - A continuously predicted signal which is, in general,

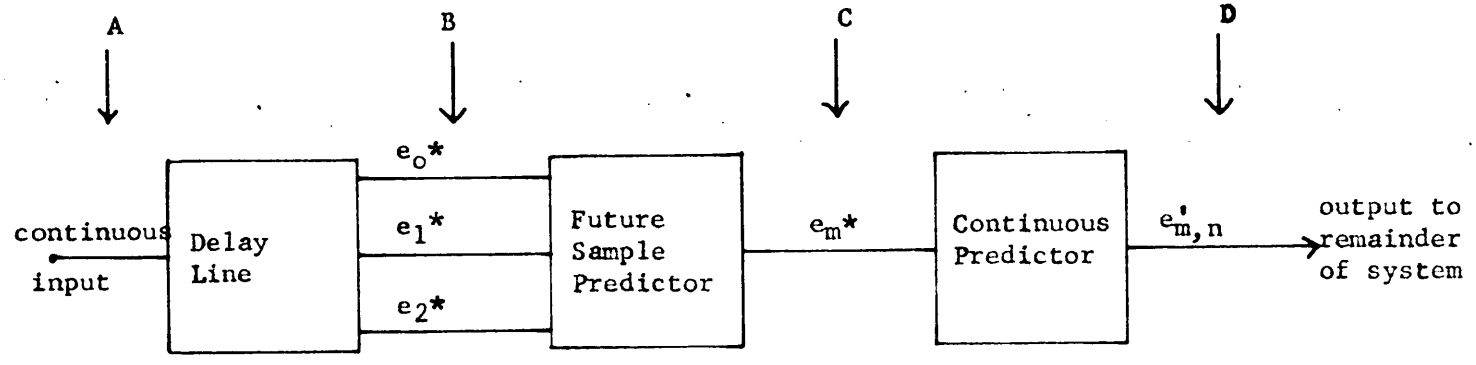


Figure 54 Block Diagram of the Simulation of a Sampling and Prediction Device

discontinuous at the sampling instant.

Probably the most simple case is to delay a sample by specified time. In other words, a held sample is fed into a sample delay circuit for one sample period and appears at the output in the same form but delayed in time. Such a device could be inserted in each line at point B in Figure 54 delaying each sample or at C delaying the weighted summation of samples. A circuit which will perform this function is illustrated in Figure 55a.

The operation of this sample delay circuit is based upon the use of a time delayed relay indicated as relay I in Figure 55a. It is used in conjunction with a delay line using series connected storage elements. The circuit used to control the delayed relay is shown in Figure 31 and is identical with that used in the simulation of finite sampling time which has been previously described. The delay time can be computed using equation 40. It has been shown that relay I becomes energized a predetermined time after relay "A" (see Figure 31) becomes de-energized and becomes de-energized at the instant the relay "A" becomes energized. The delay simulator shown in Figure 55a would normally follow a delay line or perhaps an entire hold simulator.

The operation of this circuit can be explained with the aid of the theoretical timing diagram of Figure 55b. Assume that a new

NOTE: Delayed Relay I controlled by circuit illustrated in Figure 31.

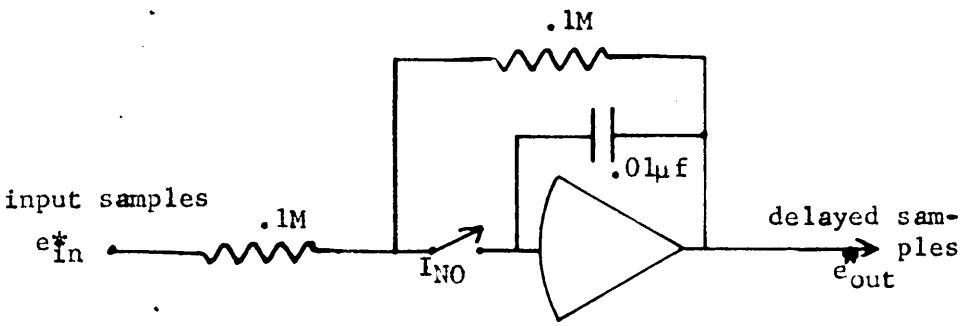


Figure 55a Diagram of Sample Delay Circuit

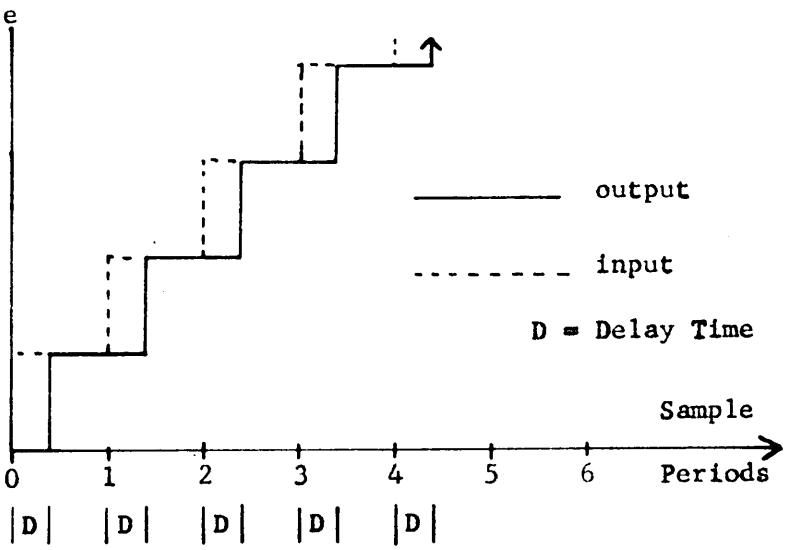


Figure 55b Theoretical Timing Diagram of Sample Delay Circuit

sample will appear at the input of the delay circuit each time the "A" relays in the delay line become de-energized. These samples are indicated by dotted lines in Figure 55b. At the instant the "A" relays become de-energized and for a pre-determined time afterwards the storage element in the delay circuit is in the store condition, storing the previous sample. After a time delay, however, relay I becomes energized and the storage element loads the new sample as indicated by the solid lines in Figure 55b. The polarity of the sample has, of course, been changed but the magnitude is unchanged. As the "A" relays continue to cycle eventually relay "A" again becomes energized and relay I opens returning the storage element to the store condition and, therefore, ready to accept a new sample. The overall effect is that the input sample has been delayed by an amount equal to the delay time of relay I, which can be set to any value within practical limits. If the new sample appears when the "A" relay becomes energized, instead of de-energized as assumed above, it is necessary to replace the "A_{NO}" contact in the delayed relay control circuit (Figure 31) with "A_{NC}" contacts.

The sample delay circuit of Figure 55a has certain practical limitations. It is limited to the delay of a previously sampled signal and this circuit alone will not perform the sampling operation. Also a careful study of the timing diagram of Figure 55b reveals that

the delay time is limited to somewhat less than one sample period since relay I is de-energized at the instant the "A" relays in the delay line become energized. The usual limitations involving the time constant of the storage element must be considered. If the delay circuit as shown is to be used with a delay line having parallel connected storage elements it is necessary to modify the control circuit of the delay relay.

The effectiveness of the sample delay circuit of Figure 55 is illustrated by the curves of Figure 56 which were recorded from an operating circuit. A 0.1 cps sine wave was sampled at a rate of 1 sample/second and fed into the sample delay circuit with constants chosen to give a 0.3 second delay time. Both the sampled input and the delayed sample appearing at the output are shown. The polarity of the output waveform has been reversed to enable easy comparison of the time relations. A careful study of these curves indicates that the sample delay circuit performs quite satisfactorily.

It has been shown that in order to use the circuit of Figure 55a it is first necessary to use one of the previously described hold simulators to generate the samples which are then fed to the sample delay circuit. The circuit illustrated in Figure 57 can be used as a combined zero order hold-sample delay simulator. In other words, the sample is taken and then delayed with the same circuit.

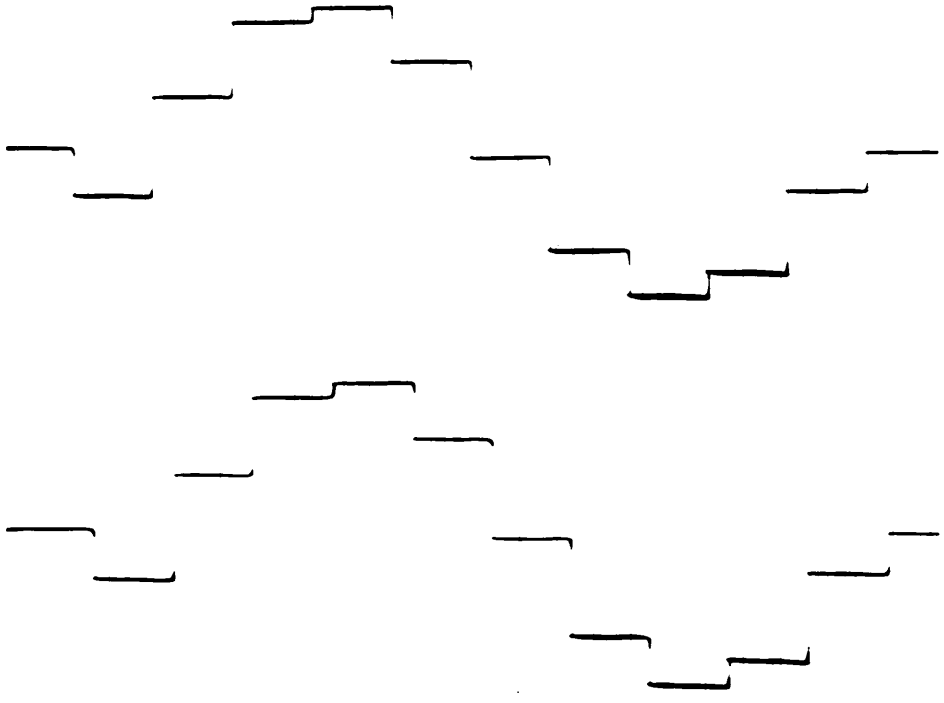


Figure 56 Performance of Sample Delay Circuit

Input signal: 0.1cps sine wave of 10 volt amplitude sampled at 1.0 sample/second

Delayed relay circuit: R = 1.0M C = 0.1μf Potentiometer setting .333

Paper speed 10mm/second

NOTE: Delayed Relay I is controlled by the circuit of Figure 31 except that the A_{NO} contacts in that circuit are replaced by A_{NC} contacts.

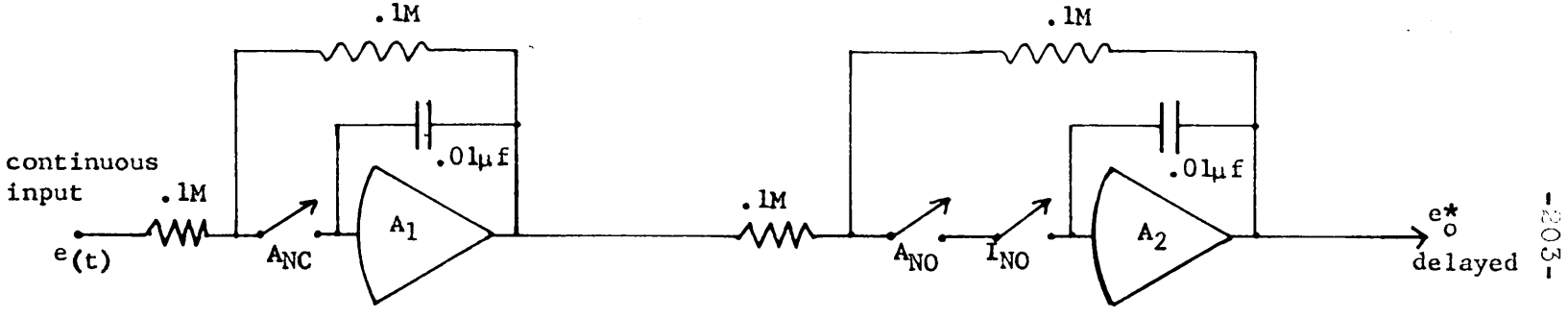


Figure 57 Zero Order Hold Simulator with Delayed Sample Output

This is accomplished by using the sample delay circuit of Figure 55a as the second storage element in the series connected zero order hold simulator of Figure 16a. Relay I is a time delayed relay operated by the control circuit illustrated in Figure 31 except that the "A_{NO}" contacts in the control circuit are replaced by "A_{NC}" contacts.

The operation of this circuit can be explained using the theoretical timing diagram shown in Figure 58. The curve illustrated in Figure 58a represents an arbitrary continuous input signal. Amplifier A_1 operates in exactly the same manner as SE_1 in the series connected zero order hold simulator illustrated in Figure 16a. The output of amplifier 1 is a quasi-sampled signal as indicated in Figure 58b and 17b. Relay I closes a predetermined time, the delay time, after relay "A" becomes energized resulting in an output from amplifier 2 shown in Figure 58c. Amplifier 2 returns to the store condition when relay "A" becomes de-energized at which time amplifier 1 goes to the load condition. The cycle then repeats indefinitely as the "A" relays continue to cycle. The net result is that the continuous signal is sampled but the samples do not appear at the output until a finite time after they were obtained. The "A_{NO}" relay is used in the second storage element to insure that amplifier 2 enters the store condition simultaneously as amplifier 1 enters the load condition. There is

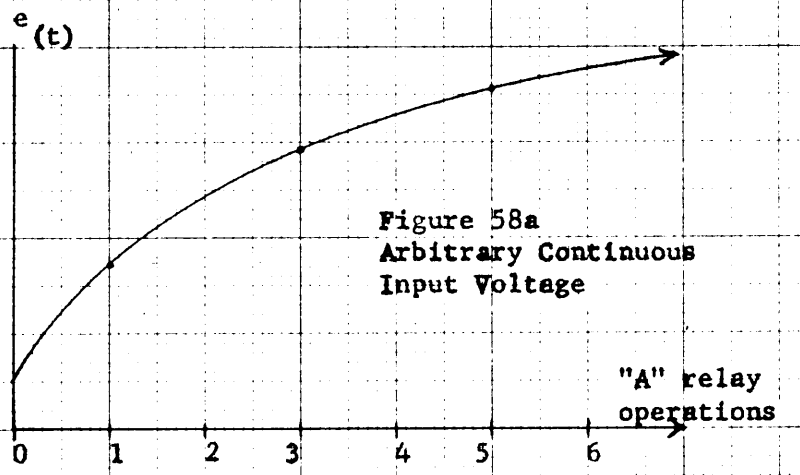


Figure 58a
Arbitrary Continuous Input Voltage

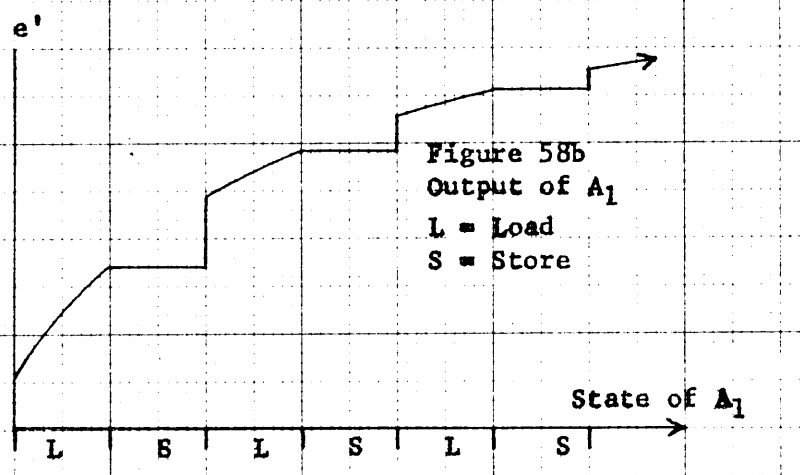


Figure 58b
Output of A_1
L = Load
S = Store

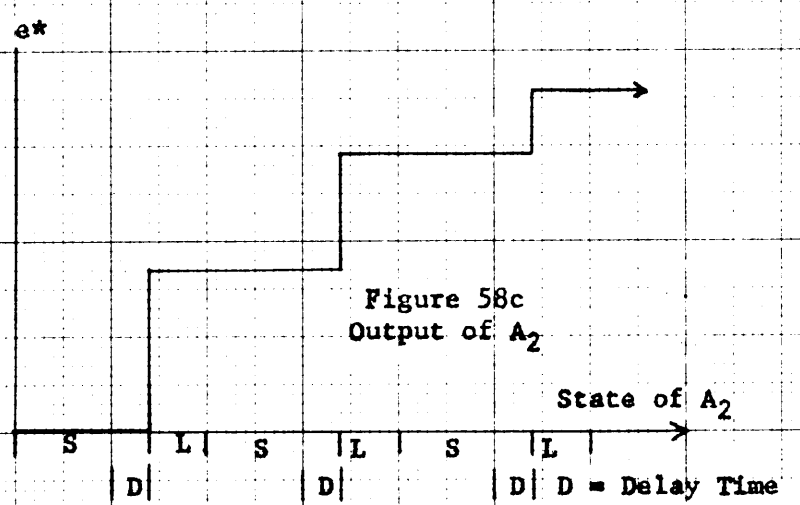


Figure 58c
Output of A_2

Figure 58 Theoretical Timing Diagram for Zero Order Hold Simulator with Delayed Sample Output.

a slight unwanted delay between the time the "A" relays are de-energized resulting in poor performance of the circuit that can be overcome by including the "A_{NO}" relay.

The quality of performance of the delayed sample circuit of Figure 57 is indicated by the recordings of Figure 59. The input signal was a 0.1 cps sine wave which was sampled at a rate of 1.0 sample/second. The constants of the circuit were chosen to result in a delay time of 0.3 seconds. Close examination of the input and output waveforms indicates that the time delay is near the value expected and that, in general, the performance of the circuit is quite satisfactory.

The circuit illustrated in Figure 57 would normally be used at point A in Figure 54 while the circuit of Figure 55 would normally be used at point B or C. The delayed sample circuit of Figure 57 can be used as a sample delay circuit; however, it is preferable to use the circuit of Figure 55 for this purpose since it involves less computing equipment.

In certain situations it may be desired to use a continuous prediction circuit such as illustrated in Figures 45 and 51 with delayed samples. It is necessary that the "A_2" relays in those circuits be synchronized with the delayed samples. If delayed samples are used the sampling instant (i.e., change of sample value) does not occur simultaneously with the operation of the "A" relays

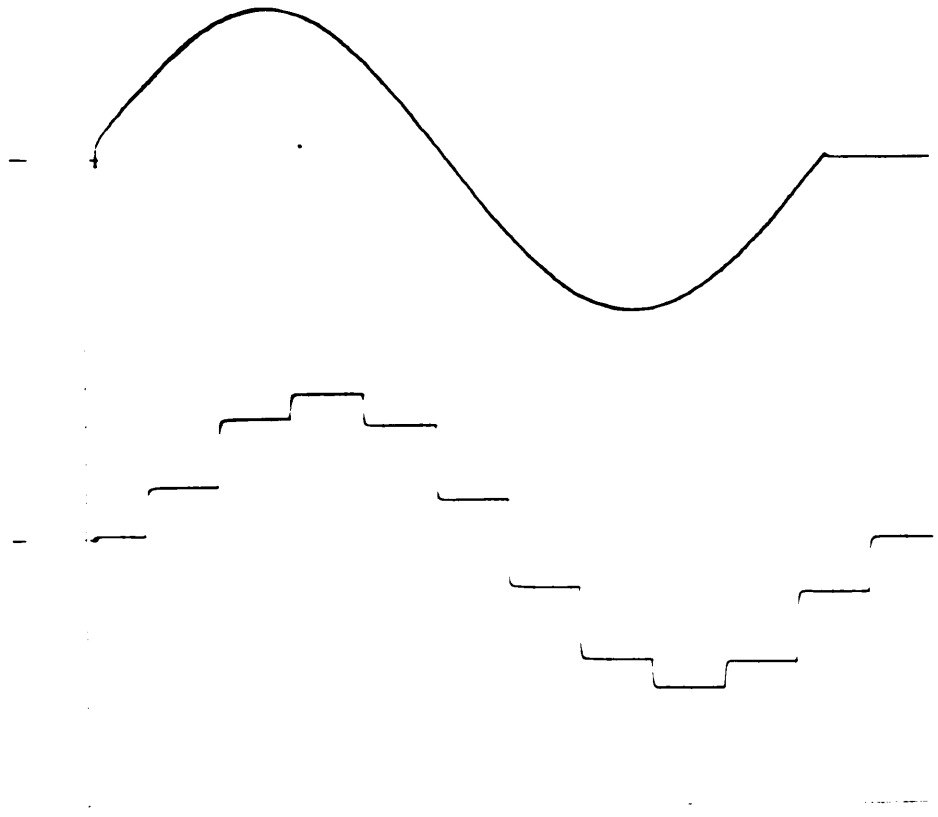


Figure 59 Performance of Delayed Sample Circuit

1.0 sample/second Paper speed 10mm/second

Delayed relay circuit R = 1.0M C = 0.1μf Potentiometer setting .333

in the delay line. It is impossible to use the synchronized frequency divider circuit of Figure 46a to operate the "A_2" relays since the operation of these relays must be delayed by the samp period as the samples. The "A_2" relays must change state at a predetermined time after the start of each cycle of the "A" relays in the delay line. Essentially what is required is a delayed frequency divider circuit. Such a circuit is illustrated in Figure 60.

The operation of this circuit is somewhat involved and can best be described by referring to the relay timing diagram of Figure 61. Assume that all relays are initially in the de-energized state. Relay "A_1" is one of the "A" relays in the computer simulating the delay line and is cycling at a rate of one cycle-sample period determined by the master timing oscillator. At time T_o relay A_1 becomes energized and a voltage of +100T, where T is the potentiometer setting, is fed through the A_{2NC} contact at "f" to integrating amplifier 1 resulting in the generation of a negative ramp at the output of this amplifier. After a certain time delay, which can be computed from equation 40, the ramp passes through -100 volts causing amplifier 2 to become saturated at a high positive voltage and, in turn, energizing relay I. Relays I and II are function relays although in the work described here telephone type relays were adapted for this use.

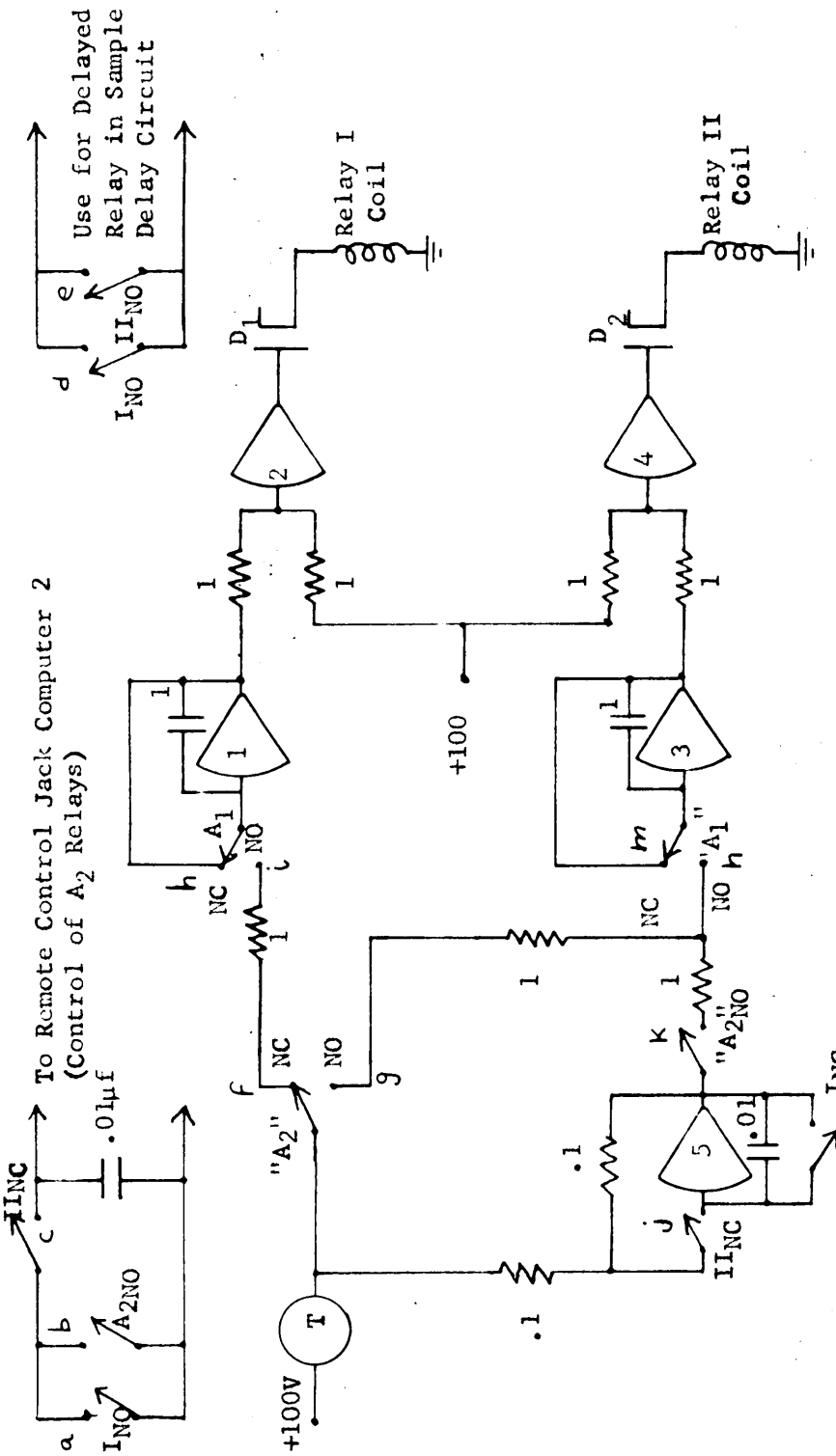


Figure 60 Circuit Diagram of Delayed Frequency Divider

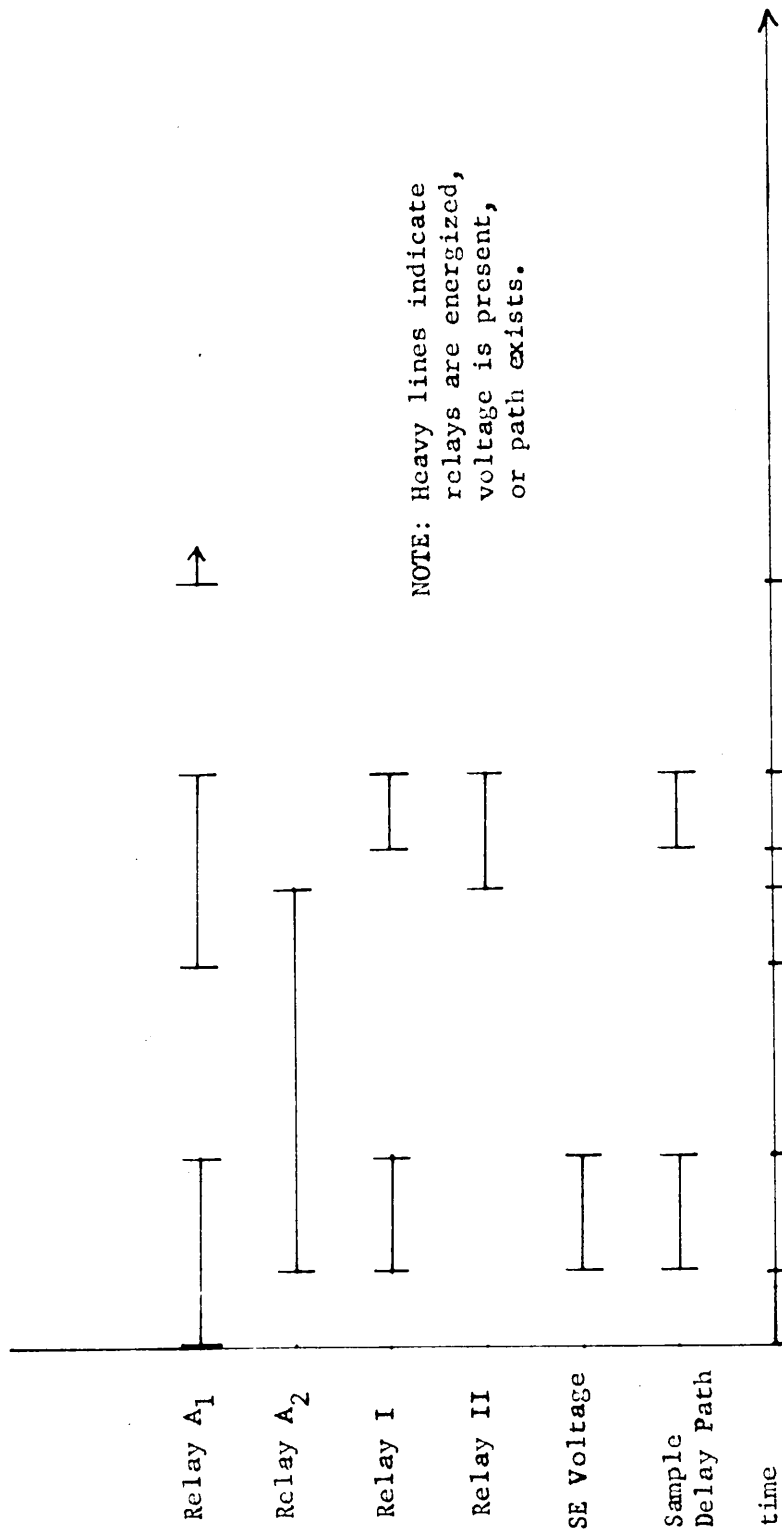


Figure 61 Relay Timing Diagram for Delayed Frequency Divider Circuit

Closure of the I_{NO} contacts at "b" causes relay "A_2" to be energized via the remote control jack at time T_1. The "A_{2NO}" contacts at "b" are closed creating a parallel path around the contacts at "a" thereby keeping relay A_2 energized as long as the II_{NC} contacts at "c" remain closed. Prior to the operation of relays I and "A2" the output voltage of amplifiers 3 and 5 was held at zero because of the closed contacts at points "l and m". However, after relays "A_2" and I become energized integrating amplifier 3 and the storage element (amplifier 5) are released and the output of the storage element assumes a value of -100T volts almost immediately. The storage element voltage is fed through the "A_{2NO}" contacts at "k" and summed with the +100T volts routed through the "A_{2NO}" contacts at point "g" causing the net input voltage to amplifier 3 to be zero. The output of integrating amplifier 3 thus remains at zero and relay II remains de-energized.

Evantually at time T_2 relay "A_1" becomes de-energized and the outputs of amplifier 1 and 3 are forced to zero by the "A_{1NC}" contacts at "h" and "m" releasing relay I which, in turn, causes the output voltage of the storage element to be set to zero by the I_{NC} contacts at "l". Relay II still remains de-energized and relay "A_2" continues to be energized through the "A_{2NO}" contacts at "b" and the II_{NC} contacts at "c".

At time T_3 relay A_1 again becomes energized. A voltage of +100T volts is then fed to amplifier 3 via the "A_{2NO}" contacts at "g" and the A_{1NO} contacts at "n" causing a negative ramp to be generated at the output of amplifier 3. After a predetermined time the ramp passes through -100 volts causing relay II to be energized by amplifier 4 at time T_4. The II_{NC} contacts at "c" are opened causing relay "A_2" to become de-energized. An input voltage of +100T appears at the input of integrating amplifier 1 via the "A_{2NC}" contacts at "f" and the "A_{1NO}" contacts at "i" and after a time delay relay I again becomes energized at T_5. The output voltage of the storage element immediately builds up to a value of -100T volts but since the "A_{2NO}" contacts at "n" are now open this voltage has no effect.

At time T_6 relay "A_1" is again de-energized causing the output voltage of amplifiers 1 and 3 to be forced to zero and, in turn, causing both relays I and II to become de-energized. The output voltage of the storage element also goes to zero at this time. Relay "A_2" remains de-energized since both the "A_{2NO}" contacts at "b" and the I_{NO} contacts at "a" are open. The conditions existing after time T_6 are exactly those assumed prior to T_o, all relays being de-energized, such that when relay "A_1" again becomes energized at T_7 the entire cycle will repeat.

Examination of the timing diagram of Figure 61 shows that relay "A_2" cycles at exactly half the rate of relay "A_1" and, therefore, the circuit can serve as a frequency divider as required. Furthermore, the time delay from the operation of relay "A_1" until the subsequent operation of relay "A_2" can be set to any desired value, within practical limits, by proper adjustment of potentiometer T. The resulting time delay can be computed using equation 40. An additional set of parallel connected I_{NO} and II_{NO} contacts at points "d" and "e" in Figure 60 can be used as the delayed relay in the sample delay circuit of Figure 55 or the delayed sample circuit of Figure 57. It is not necessary that the master timing oscillator, and consequently relay "A_1", remain in each state an equal time. The maximum delay available in the delayed frequency divider is limited to somewhat less than one sample period.

The performance of the delayed frequency divider circuit illustrated in Figure 60 is shown by the recordings of Figure 62. The circuit constants and potentiometer setting were chosen to give a delay time of 0.3 seconds. Close examination of these curves indicates that the circuit is performing the function intended.

A somewhat artificial method simulating sample computation time on the analog computer involves the use of a standard Pade

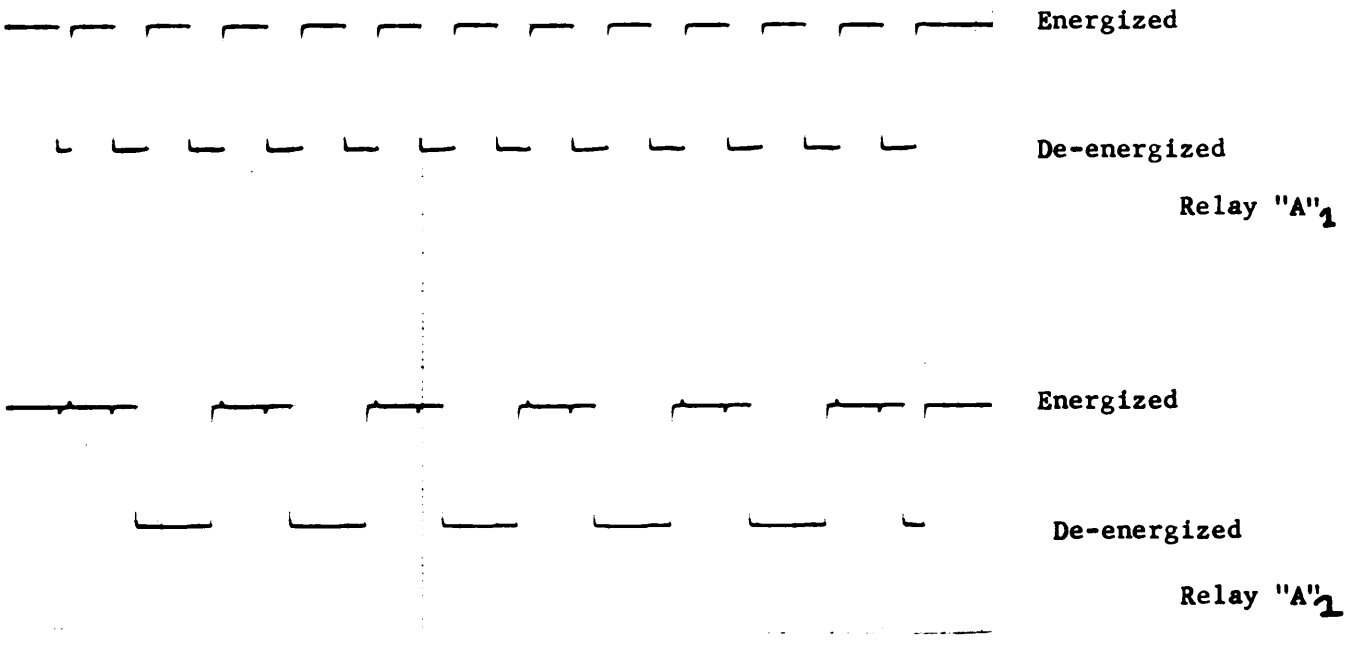


Figure 62 Performance of Delayed Frequency Divider Circuit

Sample rate 1.0 sample/second 0.3 second delay time

Delay potentiometer setting .333

delay or dead time circuit, which is discussed in many books on analog computing techniques.[41,42,43] These circuits are only approximate representations of time delay, however, they are easy to program on a computer since they involve only summers, integrators, and potentiometers. Several circuits exist of various orders. The higher order circuits tend, however, to give a better representation of delay time but involve more computing equipment.

If the delay time is constant a plot of phase lag versus frequency for sinusoidal excitation is a straight line. Unfortunately, the phase characteristics of a Pade delay circuit deviate from this ideal characteristic at high frequencies tending to give insufficient phase lag. The higher order Pade circuits tend to follow the ideal characteristics to higher frequencies however.

It is seen that the Pade circuit will not respond well to waveforms, such as a step or square wave, which have a high harmonic constant. An additional phase lag can be introduced at high frequencies by using the first order lag circuit illustrated in Figure 63. This circuit is actually a low pass filter which has been shown (equation 24) to have a break point at $\omega = \frac{1}{R_f C}$. The break point can be placed to extend the linear portion of

the phase characteristics of a Pade circuit to higher frequencies.

The component values for the compensation circuit can be selected in the following manner[44] using the Pade phase characteristics shown in Figure 64a. It is seen that both the second and fourth order Pade approximations deviate from the ideal phase characteristics. The fourth order circuit is considerably better than the second order circuit, however, although it involves three more amplifiers. The difference between the ideal characteristics and the actual characteristics are illustrated in Figure 64b for both the second and fourth order Pade circuits. A filter following the Pade circuit (having the phase characteristics shown in the figure) will yeild ideal time delay simulation.

It is seen that the required correction is approximately that which can be obtained from a first order lag circuit such as that shown in Figure 63. The time constant of the filter is chosen such that the break frequency occurs at the frequency where the correction curve passes through 45 degrees.

As an example, if the correction curve passes through 45 degrees at $\omega = 4$ the time constant of the compensation network should be such that

$$\frac{1}{R_f C} = \omega = 4 \quad . \tag{93}$$

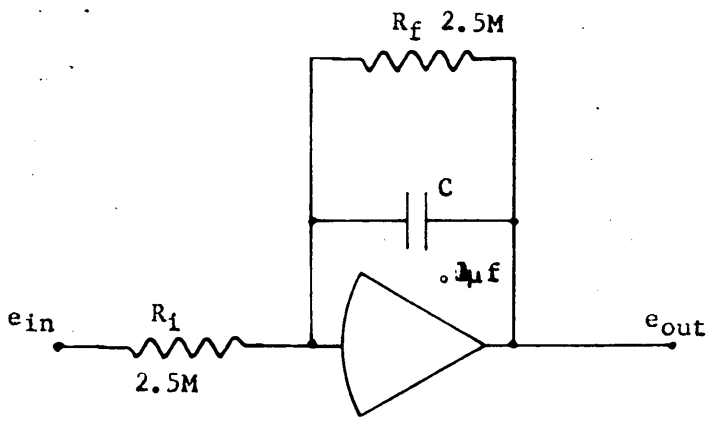


Figure 63 Low Pass Filter for Fourth Order Pade Delay Circuit

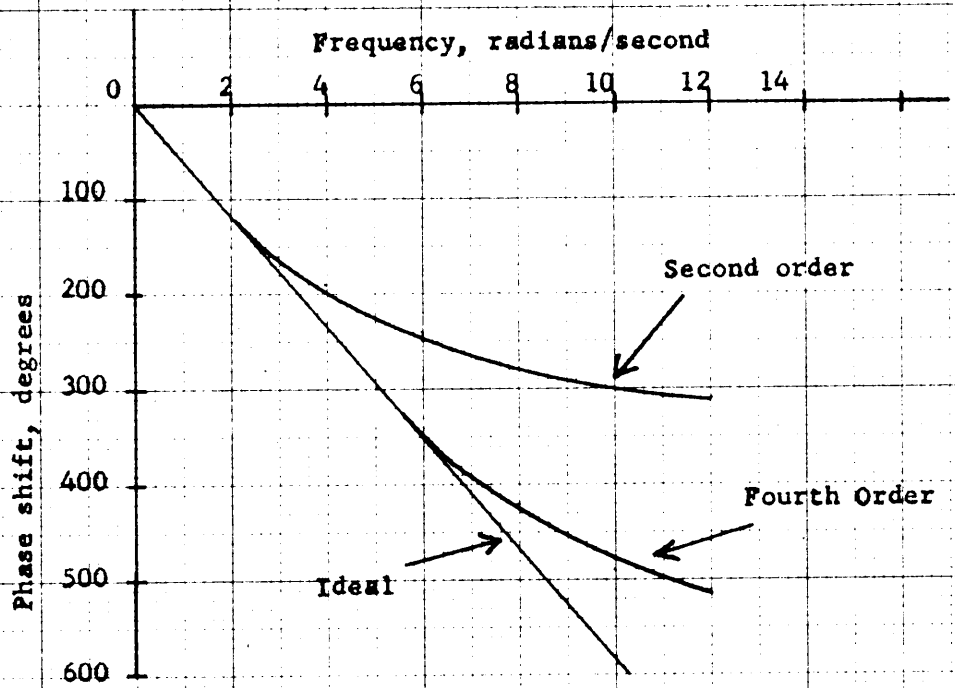


Figure 64a Characteristics of Pade Delay Circuits

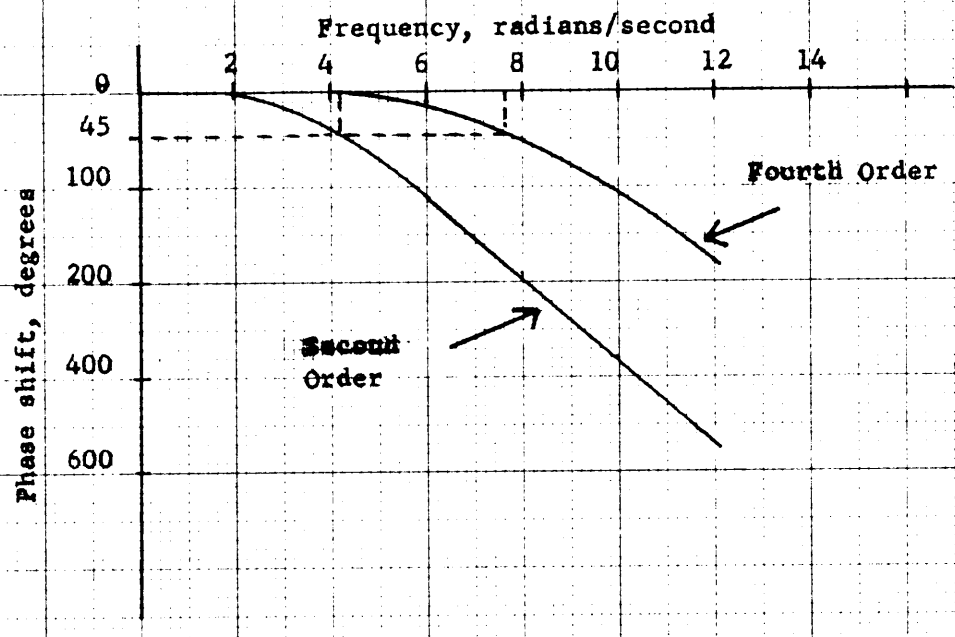


Figure 64b Correction Curves for Pade Delay Circuit

A choice of C = 0.1μf and R_f = 2.5M would give the desired result. In order to have unity gain at low frequencies, R_i in the circuit is also chosen as 2.5M. Since the phase characteristic of the low pass filter is not exactly that indicated by the correction curve the performance of the Pade delay circuit-low pass filter combination will not be ideal although it is improved considerably by the addition of the filter.

A fourth order Pade delay time simulator is shown in Figure 65. The actual potentiometer settings[45] and choice of amplifier gains are chosen to result in the following effective potentiometer settings, for a desired delay time T.

$$\begin{aligned} P_1 &= P_2 = \frac{1680}{T^4} \\ P_3 &= P_4 = \frac{840}{T^3} \\ P_5 &= P_6 = \frac{180}{T^2} \\ P_7 &= P_8 = \frac{20}{T} \end{aligned} \tag{94}$$

The curves of Figure 66 illustrate the response of the fourth order Pade delay circuit illustrated in Figure 65 with square wave excitation. The constants were adjusted to give a 0.5 second time delay. It is seen that the performance of the circuit is not very satisfactory since square waves contain significant high order harmonics which cause transient oscillations

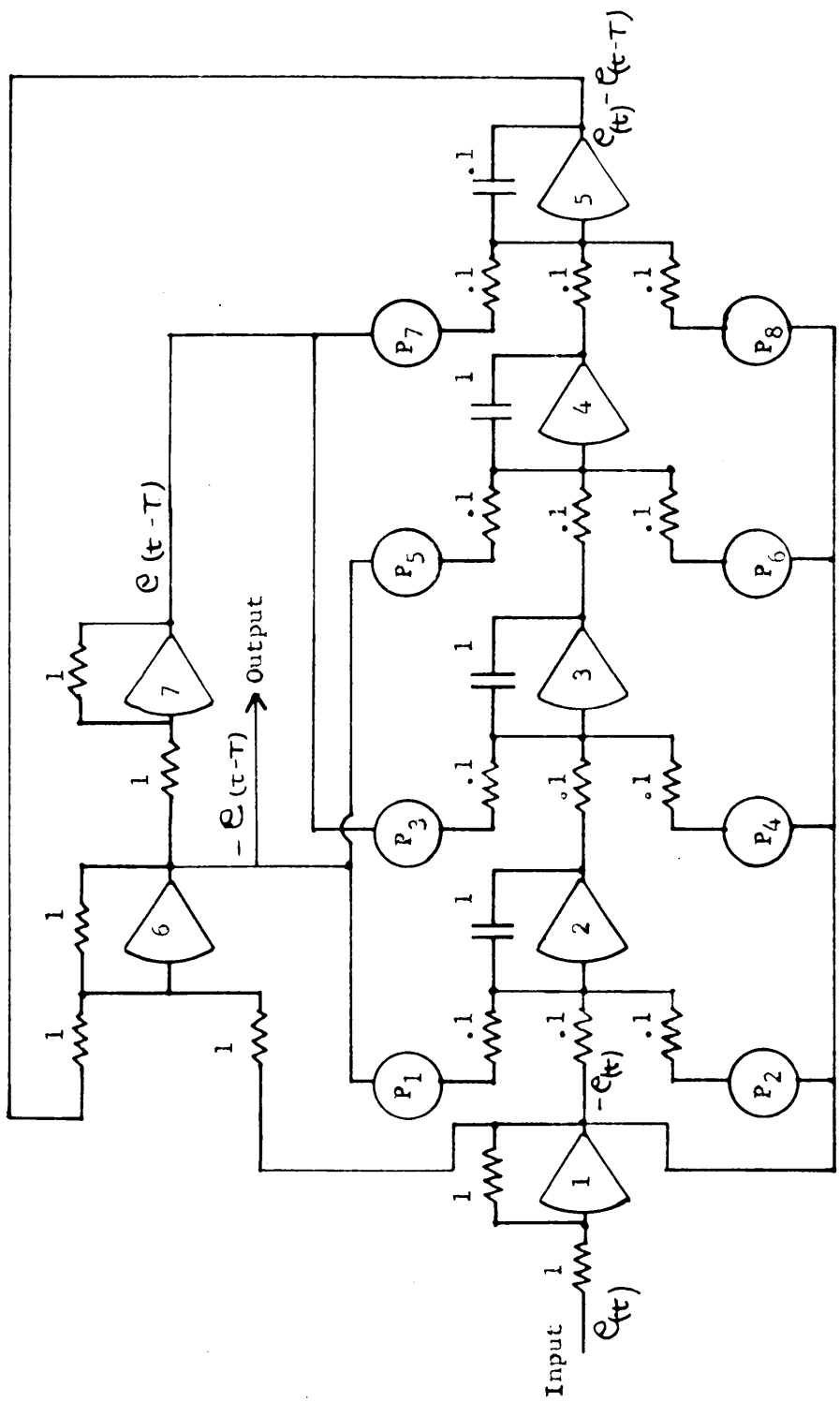


Figure 65 Circuit Diagram of Fourth Order Pade Delay Circuit

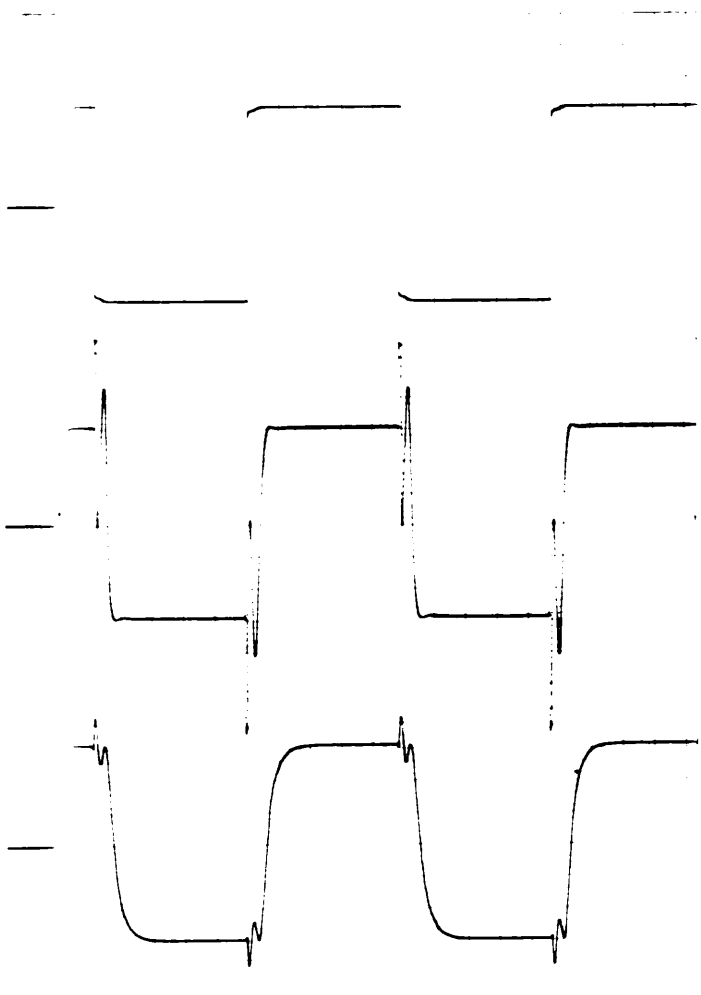


Figure 66 Response of Pade Delay Circuit with Low Pass Filter to Square Wave Excitation

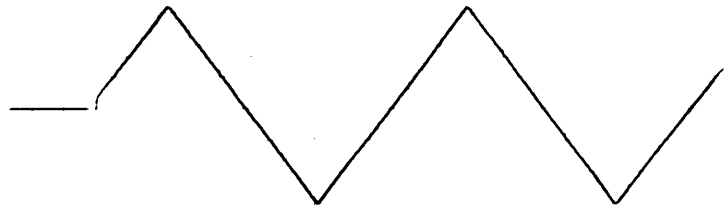
0.5 second delay time Paper speed 5mm/second

Filter circuit: $R_f = R_i = 2.5M\Omega$ $C = 0.1\mu f$

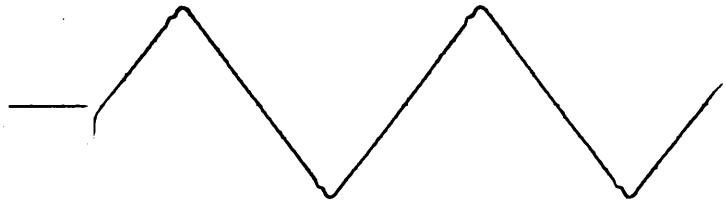
as desired but also to greatly increase the rise time which is undesirable.

The response of the same Pade delay circuit, with and without filter to a 0.1 cps triangular wave, is shown by the recordings of Figure 67. Since the harmonic content of this waveform is much less than that of a square wave the performance of the delay circuit is considerably better. The addition of the low pass filter is seen to improve the performance slightly.

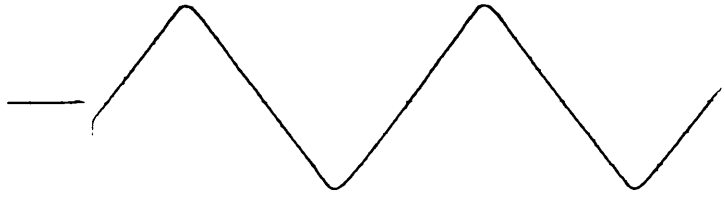
Since the response of the Pade delay circuit is rather poor with step excitation, it follows that this circuit would not, in general, be suitable for the delay of samples. If the sampling rate were very low, however, such that the settling time of the delay circuit was short compared to the sample period, it would be possible to use the Pade circuit at points "B" and "C" in Figure 54. The Pade circuit would, in general, however, be more acceptable at point "D" where the waveform of the signal is usually such that the harmonic content is relatively small.



$e_{in}(t)$
5V/cm
0.1 cps triangular wave
7.5 volts amplitude



Output of Pade Circuit
5V/cm



Output of Pade Circuit
with low pass filter
5V/cm

Figure 67 Response of Pade Delay Circuit with Low Pass Filter to Triangular Wave Excitation

0.5 second delay time Paper speed 5mm/second

Filter circuit: $R_f = R_i = 2.5M\Omega$ $C = 0.1\mu f$

IV ANALOG SIMULATION OF A SAMPLED-DATA POSITION CONTROL SYSTEM

The simulation techniques which have been presented in the preceeding sections can be combined in various manners to simulate complete sampled-data automatic control systems. The simulation circuits required depend, of course, on the particular system to be simulated.

The position control system illustrated in Figure 68, which involves sampled-data, was simulated as an example of an analog simulation of a complete sampled-data system. The simulation of this system involves many, but not all, of the techniques described earlier in this report. No attempt has been made to analyze the results or to evaluate the system performance since the computing techniques were the primary interest in this investigation.

The system is a very ordinary position control servomechanism except for the fact that the error signal is sampled. The resulting samples are used to continuously predict the future value of the error which, in effect, partially compensates for time delays introduced by the system time constants and sampling process. Both finite sampling time and computation time are simulated. The complete specifications of the system are given below.

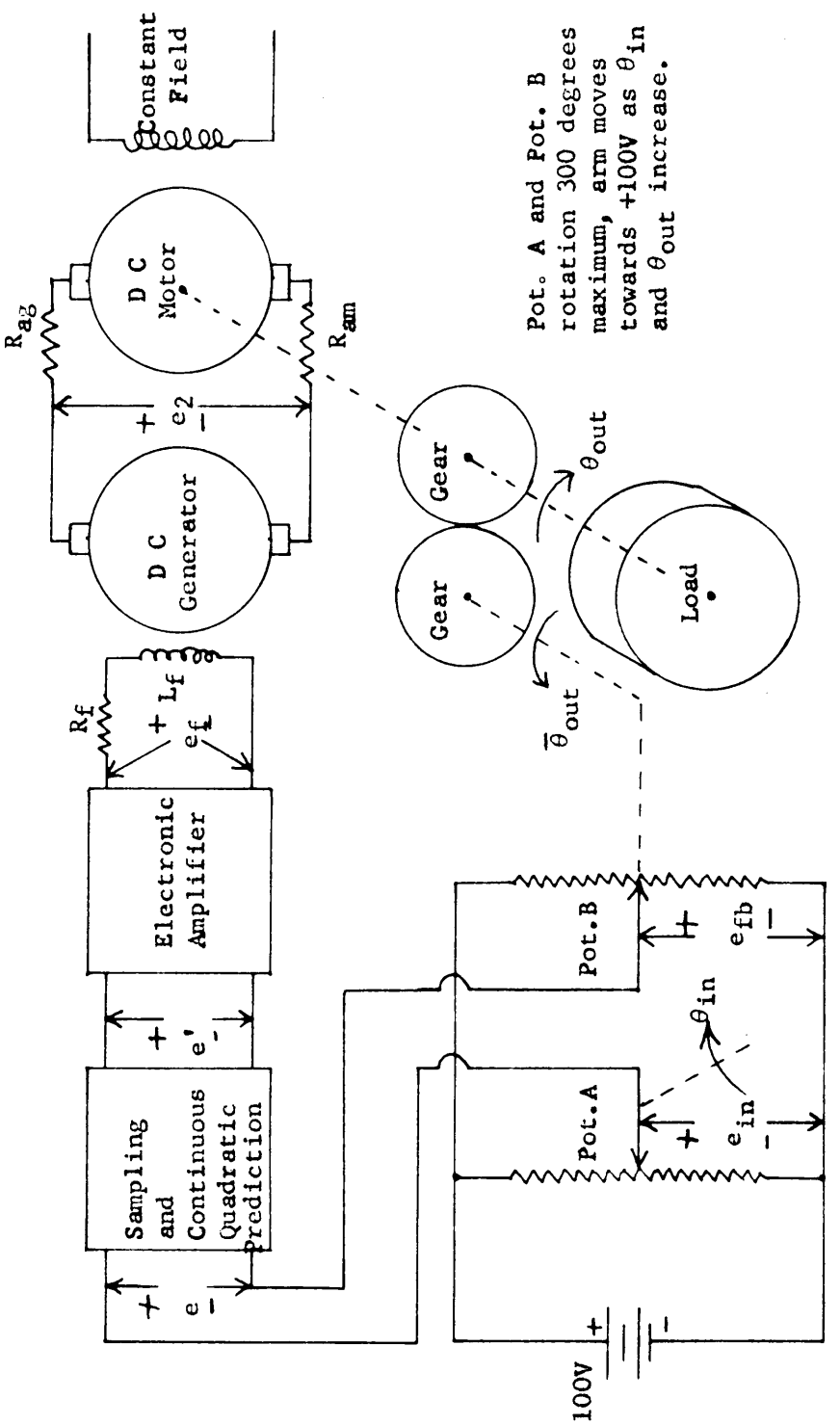


Figure 68 Diagram of Example Position Control System

Specifications for Example Position Control System

Generator Constants:

K_g = 37.0 Volts/ampere, R_f = 15.5Ω, L_f = 4.88 hy, R_{ag} = .002Ω

Motor Constants:

Emf Constant = K_e = 4.0 Volts/rpm = 38.2 Volts/radian/second

Torque Constant = K_t = 28.2 foot pounds/ampere

R_a = .010Ω

Total equivalent inertia at motor armature = J_L = 4438 slug-feet^2

Gear ratio 1/1 (backlash neglected)

Electronic Amplifier Gain: Adjustable

Sampling Rate: 10.0 samples/second

Sampling Time: 0.01 seconds

Effective Error Computation Time 0.05 second

Transfer Functions

Generator:

$$\frac{E_{2(s)}}{E_{f(s)}} = \frac{\frac{K_g}{R_f}}{\frac{L_f s}{R_f} + 1} = \frac{2.39}{.3155 + 1} \quad \text{volts/volt}$$

Motor

$$\frac{\text{Angular Velocity } (\omega(s))}{E_{2(s)}} = \frac{1}{K_e\left[\frac{(R_{am} + R_{ag})\,J_L}{K_t K_e} s + 1\right]} = \frac{.0262}{.0495 s + 1} \quad \frac{\text{radians/sec}}{\text{volt}}$$

A block diagram of the system, indicating the transfer function, is illustrated in Figure 69. The details of obtaining the transfer

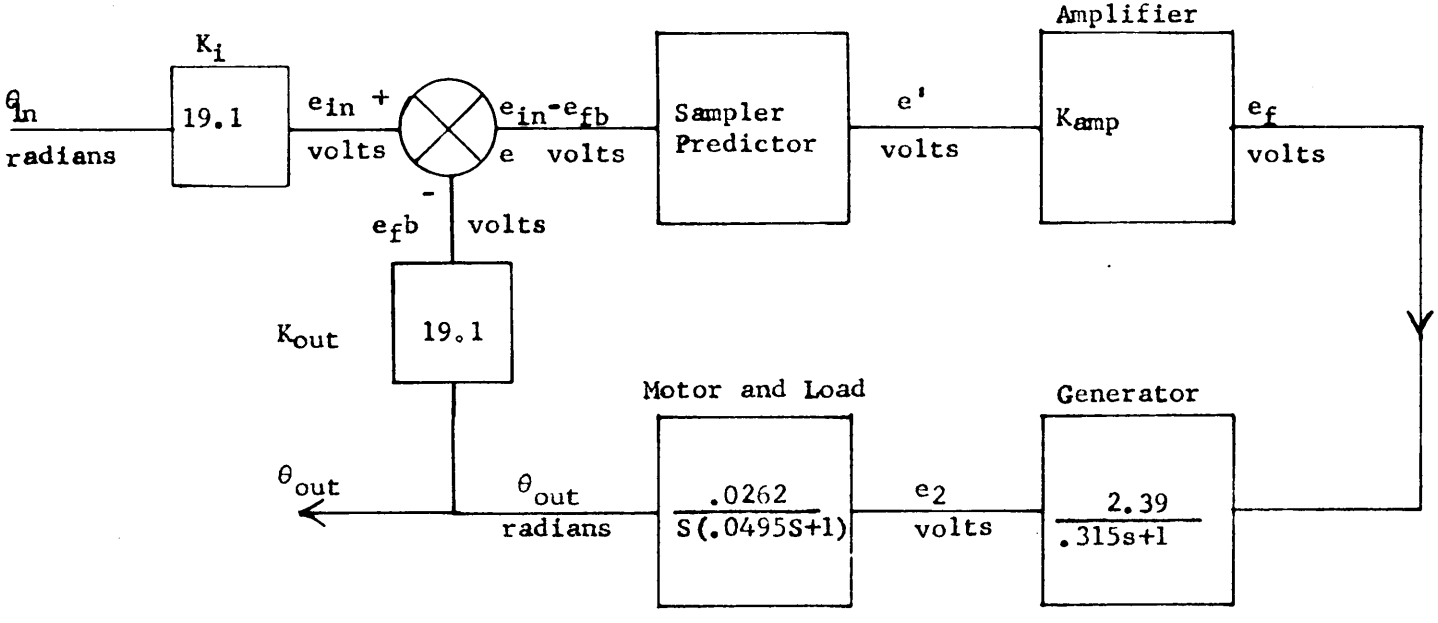


Figure 69 Block Diagram of Example Position Servo System

function are not pertinent to this investigation.

A block diagram of the computer simulation of this example system is shown in Figure 70. A finite time for sampling of 0.01 seconds is represented using the circuit shown in Figure 29a. This circuit is followed by a two stage series connected delay line. The resulting samples of the error signal e_0*, e_1*, and e_2* are used to continuously predict the value of the error two sample periods (0.2 seconds) in advance using the quadratic continuous prediction circuit shown in Figure 51. A fourth order Pade delay circuit is used to delay the predicted error by 0.05 seconds simulating the computation time in the system. A low pass filter of the form shown in Figure 63 follows the delay circuit to partially correct for the imperfect performance of the delay simulator. The continuous system is, of course, represented using conventional analog computing techniques. The complete simulation circuit diagram for the example system, including computer control circuits, is shown in Figure 71. (Inside rear cover). The points A through F indicated in the block diagram of Figure 70 are also indicated on the computer diagram. The solution time of the problem has been slowed by a factor of 10 such that the sample rate in the computer occurs at 1.0 samples/second rather than 10.0 samples/second resulting in

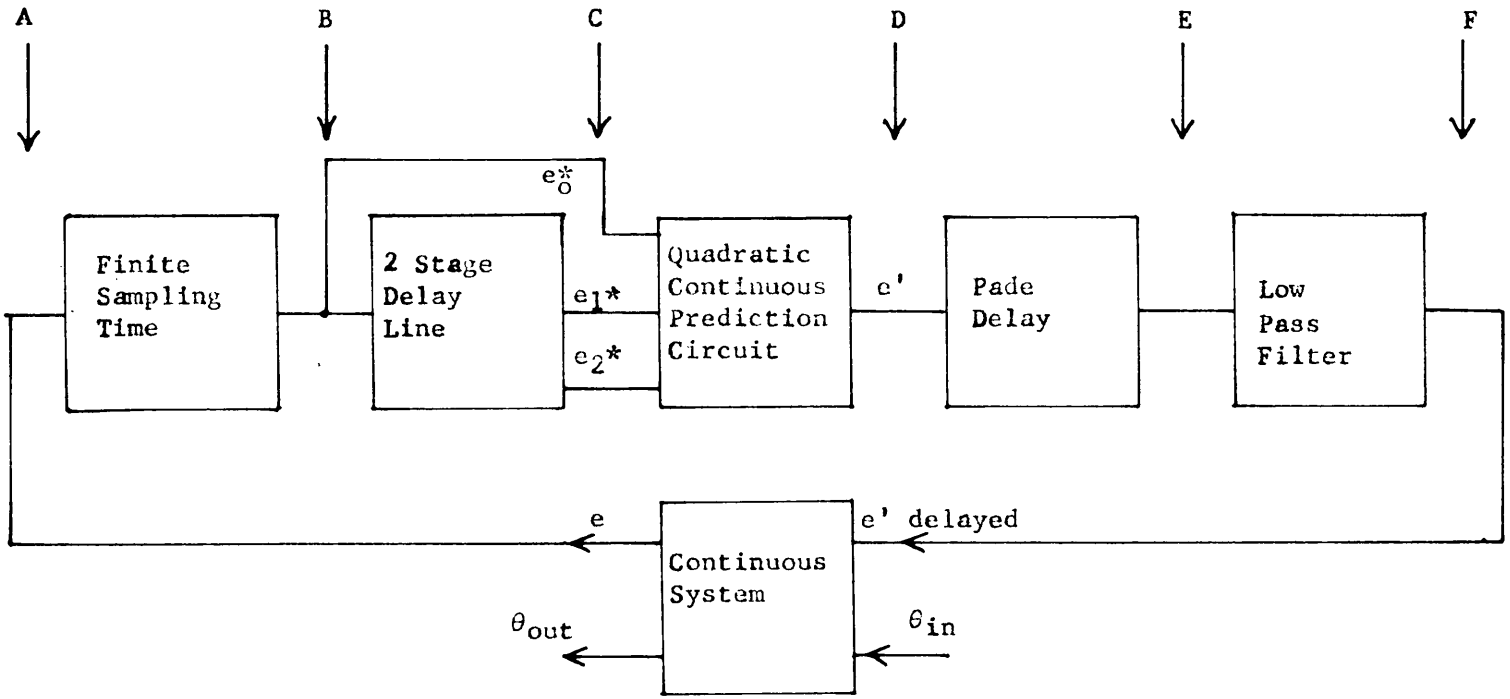


Figure 70 Block Diagram for Computer Simulation of Example Position Control System

more accurate simulation of the sampling phenomena.

Actual data obtained from the computer is shown in Figures 72, 73, and 74. The response is shown for a step input at θ_{1n} of 21.0 degrees with an electronic amplifier gain of 7.75. The curves of Figure 72 illustrate the performance without sampling. The sampler-predictor block shown in Figure 69 was bypassed while this data was obtained. It is seen that the system is moderately stable with a tendency to oscillate at approximately one cycle per second. The curves shown in Figure 73 and Figure 74 illustrate the behavior of the system with the sampling and prediction process described above. It is seen that the system has become unstable oscillating at about two cycles per second, although the amplitude of the oscillations is limited, this being a characteristic of a non-linear system. The actual output of the system is shown as curve I in both Figures 73 and 74, with the corresponding error signal shown as curve II. The sampled error is illustrated as curve III in Figure 73. The effect of the .01 second sampling time is seen to be almost negligible. The continuously predicted error (with sign reversed) is shown in Figure 73, curve IV and 74 curve III. It is seen that the predicted values are considerably in error since the value of the second derivative changes appreciably from one sample period

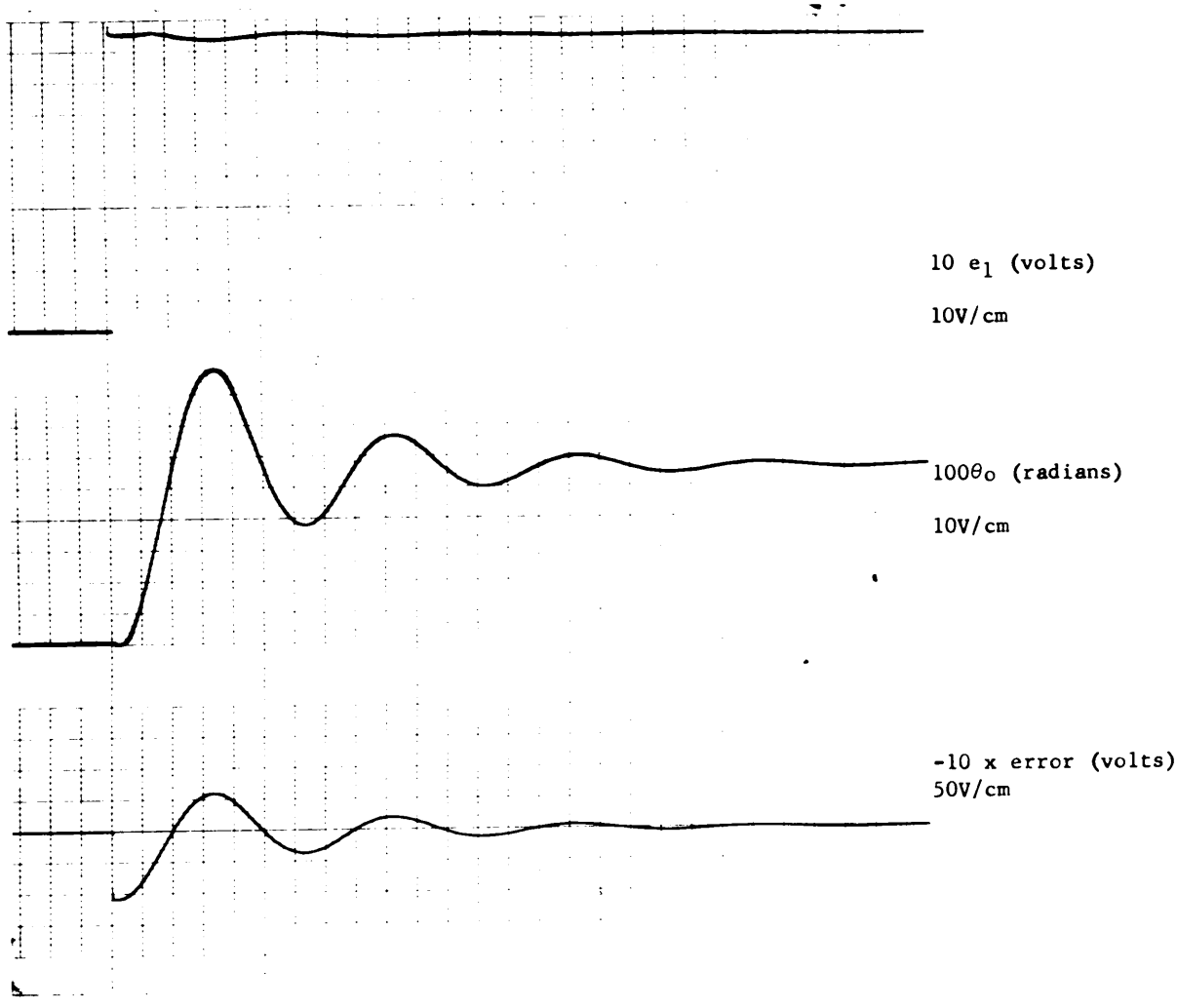


Figure 72 Response of Continuous System to Step Input

Paper speed 2.5mm/second Problem slowed by factor of 10

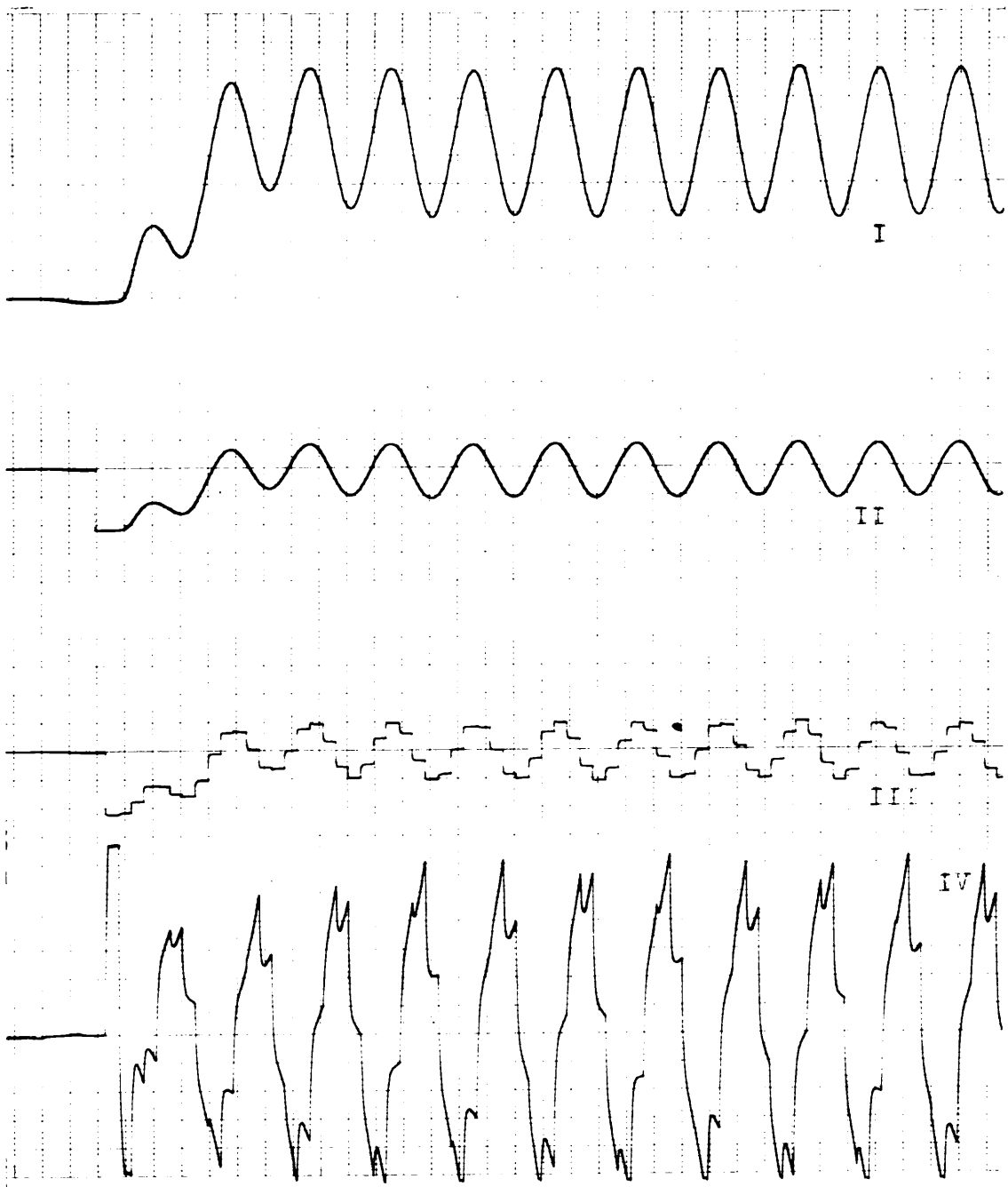


Figure 73 Response of Sampled-Data System to Step Input

Paper speed 2.5mm/second Problem slowed by factor of 10

I $100\theta_o$, 10V/cm

II -10 x error, 50V/cm

III error sample, 50V/cm

IV continuously predicted error, 50V/cm

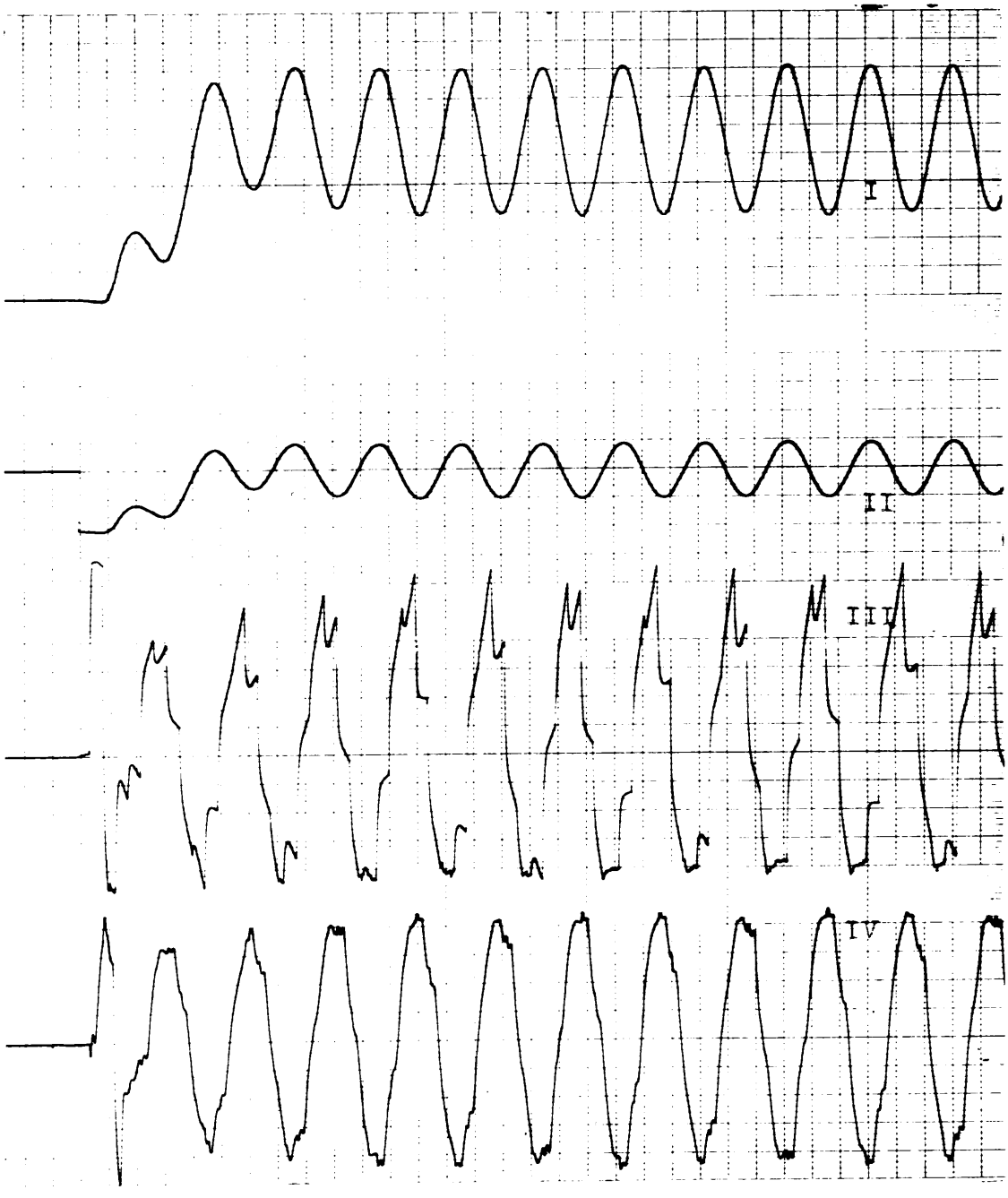


Figure 74 Response of Sampled-Data System to Step Input

Paper speed 2.5mm/second Problem slowed by factor of 10

I $100\theta_o$, 10V/cm

II -10 x error, 50V/cm

III continuously predicted error, 50V/cm

IV output of delay circuit 50V/cm

to the next. The effect of the Pade delay circuit and associated filter is evident from a comparison of curves III and IV in Figure 74. The computation time is represented as desired, however, the predicted waveform has been smoothed considerably.

The recordings in Figure 73 and Figure 74 were obtained from the following points in the computer diagram of Figure 71.

Figure	Point
73-I, 74-I	θ_{out}
73-II, 74-II	A
73-III	B
73-IV, 74-III	D
74-IV	F

The points are also illustrated in the computer block diagram of Figure 70. In some cases the signals have been inverted for ease of comparison.

In conclusion, it appears that the effect of sampling has been to make a rather marginal system completely unstable. Examination and explanation of the system performance is not, however, within the scope of the present investigation.

V CONCLUSIONS AND SUGGESTED EXTENSIONS

The object of the investigation which has been described in this report was to develop methods of simulating various phenomena associated with sampled-data automatic control systems on the analog computer. Methods for simulating these phenomena have been developed, explained, and evaluated above. In addition, experimental data has been presented to illustrate the quality of simulation techniques. In many cases alternate methods for simulating a particular process have been presented. The performance resulting from each of the simulation circuits was reasonably satisfactory when the circuits were operated within the practical limitations which have been outlined.

The investigation which has been described can be usefully extended along several lines. Some of the more obvious of these are listed below.

1. Use the techniques presented in this report to study the behavior and performance of representative sampled-data automatic control systems. A procedure for the design of such systems using the computer could be developed.

2. Use the techniques which have been outlined to verify and illustrate analytical methods of sampled-data system analysis.

3. Develop hold simulators and extrapolation circuits which use the present and three or more past samples. At most only two past samples have been considered in the investigation described here.

4. Adapt the simulation techniques discussed in this report to analog computers other than the Heath ES-400 computer which was used in the present investigation.

5. Devise methods for controlling the delay lines and prediction circuits using electronic rather than relay switching. A limited amount of work has been described in which electronic circuits have been used to control operational amplifiers in an analog computer[46]. The work was not, however, in the area of sampled-data simulation.

It is hoped that these suggestions will prove of interest to other engineers and that the work outlined in this report will serve as a background for additional investigations.

VI ACKNOWLEDGMENTS

The author wishes to express his appreciation to Professor Homer M.Powell, his thesis advisor, for the help, cooperation, and suggestions which were given.

The author is especially grateful for the untiring efforts of , who under somewhat trying conditions, proofread, corrected, and typed the manuscript.

A word of appreciation must certainly go to who spent many hours assisting the author to obtain the final data to be included in the thesis. The author is also indebted to , who called his attention to some recent work[23] related to the investigation described in this report.

VII BIBLIOGRAPHY

Literature Cited

1. Murphy, Gordon J., Control Engineering, D. Van Nostrand Co., Inc., Princeton, N. J., 1959, p. 1.

2. Black, H. S., "Stabilized Feedback Amplifiers", Electrical Engineering, Vol. 53, 1934, pp. 114-120.

3. Thaler, G. J. and Brown, R. G., Analysis and Design of Feedback Control Systems Second Edition, McGraw-Hill Book Company, Inc., New York, 1960, Chapter 14.

4. Murphy, Gordon J., Control Engineering, D.Van Nostrand Co., Inc., Princeton, N. J., 1959, pp. 322-326.

5. Thaler, G. J. and Brown, R. G., Analysis and Design of Feedback Control Systems Second Edition, McGraw-Hill Book Company, Inc., New York, 1960, pp. 389-390.

6. Truxal, John G., Control System Synthesis, McGraw-Hill Book Company, Inc., New York, 1955, p. 501.

7. Leiby, David W., "Digital Speed-Control Analysis by Analog-Computer Methods", Electrical Manufacturing, Vol. 65, p. 179, May 1960.

8. Linvill, William K., "Sampled-Data Control Systems Studied Through Comparison of Sampling With Amplitude Modulation",

Transactions of A.I.E.E., Vol. 70, Pt. II, 1951, pp. 1779-1788.

9. Ragazzini, J. R. and Zadeh, L. A., "The Analysis of Sampled-Data Systems", Transactions of A.I.E.E., Vol. 71, Pt. II, Nov., 1952, pp. 225-232.

10. Jury, Eliahu I., "Synthesis and Critical Study of Sampled-Data Control Systems", Transactions of A.I.E.E., Vol. 75, Pt. II, July 1956, pp. 141-151.

11. Farmanfarma, G., "Analysis of Linear Sampled-Data Systems with Finite Pulse Width: Open Loop", Transactions of A.I.E.E., Vol. 75, Pt. I, 1956 (Jan. 57 section) pp. 808-819.

12. Farmanfarma, G., "Analysis of Multiple Sampler Systems with Finite Pulse Width: Open Loop", Transactions of A.I.E.E., Vol. 77, Pt. II, March 1958, pp. 20-28.

13. Murphy, G. J., Ormsby, R. D., "A Survey of the Techniques for Analysis of Sampled-Data Control Systems", Transactions of the Professional Group on Automatic Control I.R.E., Vol. AC-2, Feb. 1957, pp. 79-90.

14. Freeman, H., Lowenchuss, O., "Bibliography of Sampled-Data Control Systems and Z Transform Applications" Transactions, Professional Group on Automatic Control I.R.E., Vol. PGAC-4 March 1958, pp. 28-30.

15. Ragazzini, J. R., Randall, R. H., and Russell, F. A., "Analysis of Problems in Dynamics by Electronic Circuits", Proceedings of the I.R.E., Vol. 35, May, 1947, pp. 444-452.

16. Tostanoski, B. M., "The Analysis of Sampled-Data Servomechanisms Performed on the IBM Type 650", Transactions of A.I.E.E., Vol. 75, Pt. i, September, 1956, pp. 446-450.

17. Johnson, G. W., and Lindorff, D. P., "Transient Analysis of Sampled-Data Control Systems", Transactions of A.I.E.E., Vol. 73, Pt. II, pp. 147-153.

18. Klien, R. C., "Analog Simulation of Sampled-Data Systems", I.R.E. Transactions on Telemetry and Remote Control, Vol. TRC-1 No. 2, May, 1955, pp. 2-7.

19. Wadel, L. B., "Analysis of Combined Sample-Data Systems and Continuous Data Systems on an Electronic Analog Computer", IRE Convention Record, Part 4, 1955, pp. 3-7.

20. Metzger, I. W., Experimental Study of a Sampled Feedback System with Finite Pulse Width, M.S. Thesis, University of California at Berkeley, May 1957.

21. Elgerd, O. I., "An Analog Computer Study of the Transient Behavior and Stability Characteristics of Serial Type Digital Data Systems", Transactions of A.I.E.E., Vol. 77, Pt. I, July 1958, pp. 358-366.

22. Chestnut, H., Dabul, A., and Leiby, D., "Analog Computer Study of Sampled Data Systems", Transactions of A.I.E.E., Vol. 77, Pt. II, January 1959, pp. 634-640.

23. Brule', J. D., "Polynomial Extrapolation of Sampled Data with an Analog Computer", I.R.E. Transactions on Automatic Control, Vol. AC-7, January 1962, pp. 76-77.

24. Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956, p. 10.

25. Gardner, M. F. and Barnes, J. L., Transients in Linear Systems, Volume I, John Wiley and Sons, Inc., New York, 1942, pp. 127-129.

26. Gardner, M. F. and Barnes J. L., Transients in Linear Systems, Volume I, John Wiley and Sons, Inc., New York, 1942, p. 120.

27. Smith, George W., and Wood, Roger C., Principles of Analog Computation, McGraw-Hill Book Company, Inc., New York, 1959, p. 70.

28. Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956, pp. 195-198.

29. Gardner, M. F. and Barnes, J. L., Transients in Linear Systems, Volume I, John Wiley and Sons, Inc., New York, 1942, p. 105.

30. Fich, Sylvan, Transient Analysis in Electrical Engineering, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1951, pp. 131-132.

31. Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956, p. 187.

32. Korn, Granio A. and Korn, Theresa M., Electronic Analog Computers, McGraw-Hill Book Company, Inc., New York, 1956, pp. 193-194.

33. Thaler, G. J. and Brown, R. G., Analysis and Design of Feedback Control Systems, Second Edition, McGraw-Hill Book Company, Inc. New York, 1960, p. 406.

34. Millman, Jacob, and Taub, Herbert, Pulse and Digital Circuits, McGraw-Hill Book Company, Inc., New York 1956, p. 197-201.

35. Savant, C. J. Jr., Basic Feedback Control System Design, McGraw-Hill Book Company, Inc., New York, 1958, p. 172.

36. Murphy, Gordon J., Control Engineering, D. Van Nostrand Company Inc., Princeton, N. J., 1959, p. 264.

37. Thaler, G. J. and Brown, R. G., Analysis and Design of Feedback Control Systems Second Edition, McGraw-Hill Book Company, Inc., New York, 1960, p. 392.

38. Chestnut, Harold and Mayer, Robert W., Servomechanisms and Regulating System Design, Volume I, Second Edition, John Wiley and Sons, Inc., New York, 1959, p. 270.

39. Murphy, Gordon J., Control Engineering, D. Van Nostrand Company, Inc., Princeton, N. J., 1959, pp. 262-266.

40. Truxal, John G., Control System Synthesis, McGraw-Hill Book Company, Inc., New York, 1955, p. 544.

41. Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956, pp. 127-132.

42. Jackson, Albert S., Analog Computation, McGraw-Hill Book Company, Inc., New York, 1960, pp. 255-266.

43. Smith, George W. and Wood, Roger C., Principles of Analog Computation, McGraw-Hill Book Company, Inc., New York, 1959, pp. 113-121.

44. Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956, p. 132.

45. Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956, p. 131.

46. Diamantides, Nick O., "A Multipurpose Electronic Switch for Analog Computer Simulation and Autocorrelation Applications", IRE Transactions on Electronic Computers, Vol. EC-5, Number 4, December 1956, pp. 197-202.

Literature Examined

Black, H. S., "Stabilized Feedback Amplifiers", Electrical Engineering, Vol. 53, 1934, pp. 114-120.

Brule', J. D., "Polynomial Extrapolation of Sampled Data with an Analog Computer", I.R.E. Transactions on Automatic Control, Vol. AC-7, January 1962, pp. 76-77.

Chestnut, H., Dabul, A., and Lieby, D., "Analog Computer Study of Sampled Data Systems", Transactions of A.I.E.E., Vol. 77, Pt, II, January, 1959, pp. 634-640.

Chestnut, Harold and Mayer, Robert W., Servomechanisms and Regulating System Design, Volume I, Second Edition, John Wiley and Sons, Inc., New York, 1959.

Diamantides, Nick D., "A Multipurpose Electronic Switch for Analog Computer Simulation and Autocorrelation Applications", IRE Transactions on Electronic Computers, Vol. EC-5 Number 4, December 1956, pp. 197-202.

Elgerd, O. I., "An Analog Computer Study of the Transient Behavior and Stability Characteristics of Serial Type Digital Data Systems", Transactions of A.I.E.E., Vol. 77, Pt. I, July 1958, pp. 358-366.

Farmanfarma, G., "Analysis of Linear Sampled Data Systems with Finite Pulse Width: Open Loop", Transactions of A.I.E.E., Vol. 75, Pt. I 1956, (Jan, 1957 section) pp. 808-819.

Farmanfarma, G., "Analysis of Multiple Sampler Systems with Finite Pulse Width: Open Loop", Transactions of A.I.E.E., Vol. 77, Pt. II, March 1958, pp. 20-28.

Fich, Sylvan, Transient Analysis in Electrical Engineering, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1951.

Freeman, H., Lowenchuss, O., "Bibliography of Sampled-Data Control Systems on Z Transform Applications", Transactions, Professional Group on Automatic Control IRE, Vol. PGAC-4, March, 1958, pp. 28-30.

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

Jackson, Albert S., Analog Computation, McGraw-Hill Book Company, Inc., New York, 1960.

Johnson, Clarence L., Analog Computer Techniques, McGraw-Hill Book Company, Inc., New York, 1956.

Johnson, G. W. and Lindorff, D. P., "Transient Analysis of Sampled-Data Control Systems", Transactions of A.I.E.E., Vol. 73, Pt. II, July 1954, pp. 147-153.

Jury, E. I., "Synthesis and Critical Study of Sampled-Data Control Systems", Transactions of A.I.E.E., Vol. 75, Pt. II, July, 1956, pp. 141-151.

Klien, R. C., "Analog Simulation of Sampled-Data Systems", I.R.E. Transactions on Telemetry and Remote Control, Vol. TRC-1, No. 2, May, 1955, pp. 2-7.

Lieby, David W., "Digital Speed-Control Analysis by Analog Computer Methods", Electrical Manufacturing, Vol. 65, p. 179, May, 1960.

Korn, Granio A. and Korn, Theresa M., Electronic Analog Computers, McGraw-Hill Book Company, Inc., New York, 1956.

Lago, Gladwyn V., "Additions to Z Transformation Theory for Sampled-Data Systems", A.I.E.E. Transactions, Vol. 73, Pt. II, (January 1955 section), p. 403.

Lago, Gladwyn V., Truxal, John G., "The Design of Sampled-Data Feedback Systems", A.I.E.E. Transactions, Vol. 73, Pt. II, November, 1954, p. 247.

Linvill, William K., "Sampled-Data Control Systems Studied Through Comparison with Amplitude Modulation", Transactions of A.I.E.E., Vol. 70, Pt. II, 1951, pp. 1779-1788.

Metzger, I. W., Experimental Study of a Sampled Feedback System with Finite Pulse Width, M.S. Thesis, University of California at Berkeley, May, 1957.

Millman, Jacob and Taub, Herbert, Pulse and Digital Circuits, McGraw-Hill Book Company, Inc., New York, 1958.

Murphy, Gordon J., Control Engineering, D. Van Nostrand Company, Inc., Princeton, N. J., 1959.

Murphy, Gordon J., Kennedy, Harold B., "Closed-Loop Analysis of Sampled-Data Systems with Appreciable Pulse Width", A.I.E.E.

Transactions, Vol. 77, Part II, pp. 659-665 (January 1959).

Murphy, G. J., Ormsby, R. D., "A Survey of the Techniques for Analysis of Sampled-Data Control Systems", Transactions of the Professional Group on Automatic Control I.R.E., Vol. AC-2, February, 1957, pp. 79-90.

Nesline, F. W., Jr., "Optimum Response of Discontinuous Feedback Control Systems", A.I.E.E. Transactions, Vol. 77, Part II, pp. 651-658. (January, 1959).

Ragazzini, J. R., Randall, R. H., and Russell, F. A., "Analysis of Problems in Dynamics by Electronic Circuits", Proceedings of the I.R.E., Vol. 35, May, 1947, pp. 444-452.

Ragazzini, J. R. and Zadeh, L. A., "The Analysis of Sampled-Data Systems", Transactions of A.I.E.E., Vol. 70, Pt. II, 1951, pp. 1779-1788.

Savant, C. J., Jr., Basic Feedback Control System Design, McGraw-Hill Book Company, Inc., New York, 1958.

Smith, George W., and Wood, Roger C., Principles of Analog Computation, McGraw-Hill Book Company, Inc., New York, 1959.

Thaler, G. J. and Brown, R. G., Analysis and Design of Feedback Control Systems, Second Edition, McGraw-Hill Book Company, Inc. New York, 1960.

Truxal, John G., Control System Synthesis, McGraw-Hill Book Company, Inc., New York, 1955.

Tostanoski, B. M., "The Analysis of Sampled-Data Servomechanisms Performed on the IBM Type 650", Transactions of A.I.E.E., Vol. 75, Pt. I, September, 1956, pp. 446-450.

Tou, Julius, "Analysis of Sampled-Data Control Systems with Finite Sampling Duration", Proceedings National Electronics Conference, Vol. 13, 1957, pp. 561-573.

Wadel, L. B., "Analysis of Combined Sample-Data Systems and Continuous Systems on the Electronic Analog Computer", I.R.E. Convention Record, Part 4, 1955, pp. 3-7.

The vita has been removed from the scanned document

ABSTRACT

The electronic analog computer is widely used to simulate most types of automatic control systems. Only a limited amount of work has been reported, however, involving analog computer simulation of sampled-data systems. This is to be expected since such systems are essentially digital in nature.

The purpose of the work described in this thesis was to develop methods of simulating various phenomena associated with sampled-data systems. The techniques which have been developed are described and evaluated in the report. In addition, experimental data is presented to illustrate the performance of the various simulation circuits.

As an illustration of the techniques which have been developed, the simulation of a representative sampled-data system is described. Data obtained from this simulation is included in the report.

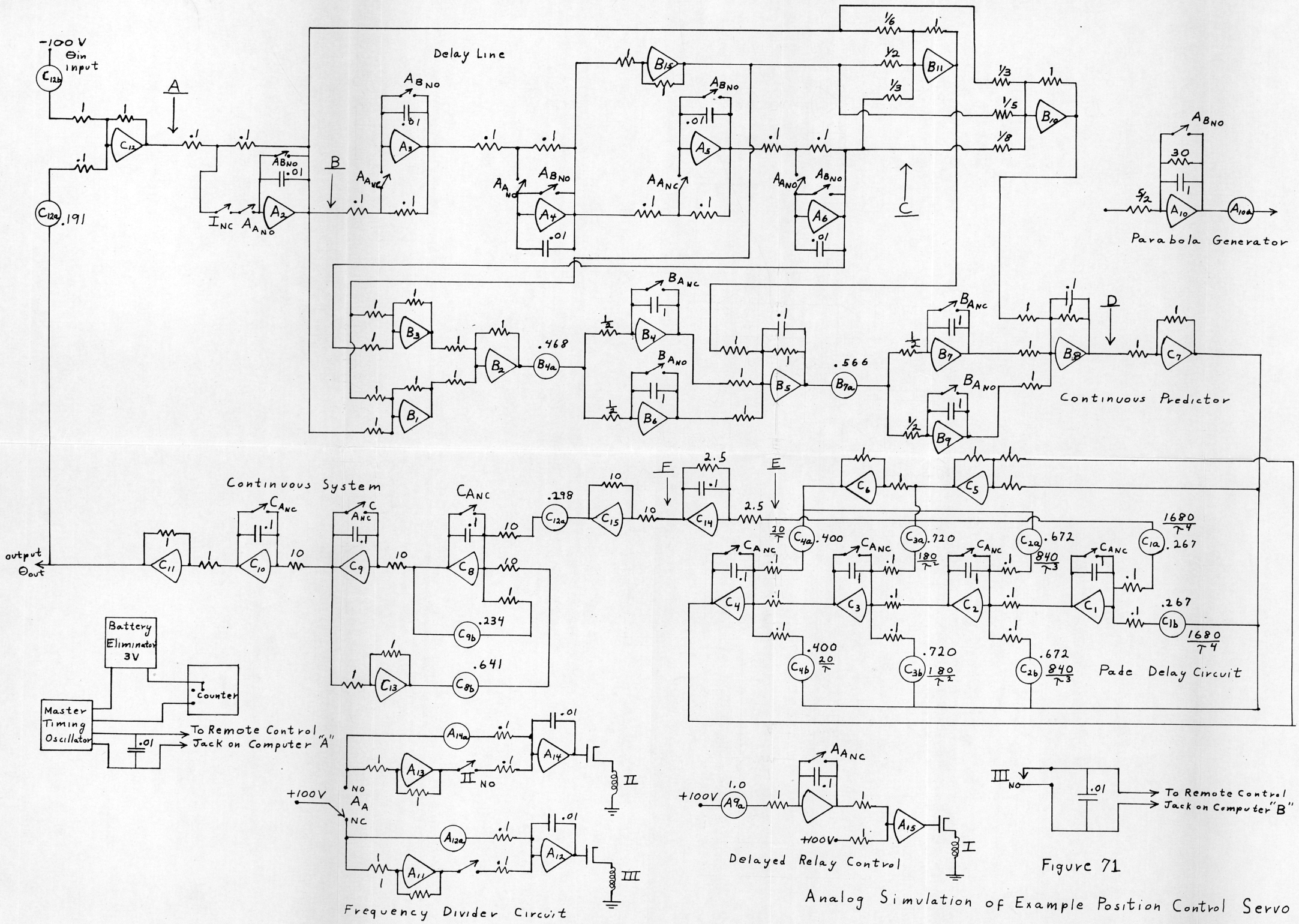


Figure 71

Analog Simulation of Example Position Control Servo