

similar manner, so the proper antenna heights can be calculated from the existing 4,000-mc data. Nevertheless it seems advisable to conduct some tests to obtain further knowledge about matters pertaining to this type of terrain and the effect of, for instance, salt storms on the path attenuation. Often the antenna at one end of a path is located several hundred feet higher than the one at the low end, which is only a few feet above ground, to keep the direct radio path within the limits of the first Fresnel zone. In such cases it may become physically impossible to mount a 6,000- or 11,000-mc antenna low enough.

Paths with multiple reflecting surfaces

present a more complicated problem. The loss to the direct signal contributed by even Fresnel zone reflections from each surface separately may not be severe, but if the antenna heights are not properly selected the reflections from two or more surfaces will no doubt coincide during some bending condition. The combined effect can be severe enough to make the circuit unreliable. To predict without testing how the 6,000- and 11,000-mc transmissions will behave at present antenna heights is very difficult, if not impossible. It is necessary to have a complete and accurate knowledge of the paths with respect to the location and elevation of all reflecting surfaces, as well as obstruc-

tions. Such information cannot be obtained from maps and, until further tests have been made, it does not seem advisable to rely on the available 4,000-mc data. The surest way is to make antenna height versus path loss tests at all proposed frequencies. The 4,000-mc antenna heights were selected so as to avoid coinciding reflections from the various surfaces. With the closer spacings between the Fresnel zones at the higher frequencies it becomes even more of a problem to find antenna heights which will prevent coinciding reflections. Of course, a common antenna may be used if the reflections are blocked by a nearby obstruction, which in itself is not a reflecting surface.

The Stability and Compensation of Saturating Sampled-Data Systems

FRANCIS J. MULLIN
ASSOCIATE MEMBER AIEE

Synopsis: Sampled-data feedback systems which contain a saturating element are investigated in this paper. A condition which must be satisfied if these systems are to be stable is developed; it is shown that this is a necessary condition for stability and an intuitive argument is presented to indicate that this condition is probably sufficient. Finally, a method to obtain a digital filter for compensating and/or stabilizing these systems is presented.

THE SYSTEM to be investigated in this paper is shown in Fig. 1. In particular it is wished to determine the effect of the saturating element on the response of the system when a step function is applied to the input.

Saturating sampled-data systems have been investigated previously by Nease¹ and Kalman.² In his study Nease con-

sidered pulse-width modulators in which the duration of the pulse applied to the system at the n th sampling instant was related to the error signal at the n th sampling instant. Since the pulse width can be no larger (in time) than the sampling period, a saturation effect between the pulse width and the sampled-error signal was necessarily introduced in his analysis. Kalman, on the other hand, discussed saturating systems in which the pulse width of sampled-error signal was small compared to the time constants of the system in which case the pulse could be accurately replaced by an impulse. This same assumption is made when the z -transform^{3,4} or modified z -transform⁵ are used to analyze linear sampled-data systems. Kalman was not primarily concerned with the analysis of these systems but rather with the problems involved in controlling such systems in some optimum fashion. Consequently, he talked about these systems in rather general terms and presented some concepts which can be used in finding this controlling function.

In this paper the study will be these

saturating sampled-data systems using the difference equations which describe the sampled-error signal as the starting point of the analysis. This approach to the problem differs from that used by Kalman and the objectives of this study are also different. The first investigation will concern the stability requirements of these systems and then a method will be given to obtain a digital filter which will force these systems to follow step or ramp inputs when saturation is present. Two illustrative examples will be given to clarify the ideas involved.

Requirements for Stability

For this study the most convenient manner to describe the system is in terms of the sampled-error signal, $E(n)$, and its various derivatives. For notational convenience we shall use the following definitions (see Fig. 1):

$$\begin{aligned} E(n) &= e(t)|_{t=nT} & C(n) &= c(t)|_{t=nT} \\ \dot{E}(n) &= \left. \frac{d[e(t)]}{dt} \right|_{t=nT} & \dot{C}(n) &= \left. \frac{d[c(t)]}{dt} \right|_{t=nT} \\ \ddot{E}(n) &= \left. \frac{d^2[e(t)]}{dt^2} \right|_{t=nT} & \ddot{C}(n) &= \left. \frac{d^2[c(t)]}{dt^2} \right|_{t=nT} \\ & \dots & & \\ R(n) &= r(t)|_{t=nT} \end{aligned} \quad (1)$$

T is the sampling period of the system and n , the independent variable of the difference equations, numbers the sampling period.

To form the difference equations which

Paper 59-52, recommended by the AIEE Computing Devices Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., February 1-6, 1959. Manuscript submitted October 30, 1958; made available for printing November 25, 1958.

FRANCIS J. MULLIN is with the California Institute of Technology, Pasadena, Calif.

This research was supported by the U. S. Air Force through the Air Force Office of Scientific Research, Air Research and Development Command under Contract no. AF 18(600)-1521 at the University of California, Berkeley. The author would like to extend his personal thanks to Professors E. I. Jury and A. M. Hopkin of the University for their help and encouragement during the course of this research.

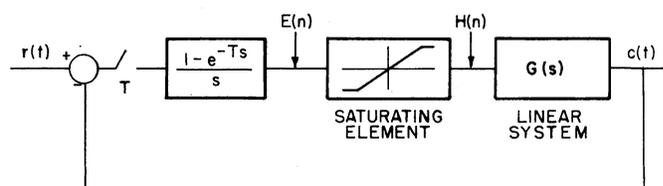


Fig. 1. A saturating sampled-data feedback system

describe the sampled-error signal, proceed as follows. The linear differential equation relating $c(t)$ to $H(n)$, which is constant over any sampling interval, is solved for $c(t)$ during the n th sampling interval and differentiated $(p-1)$ times where p is the order of $G(s)$ in Fig. 1. If these equations are evaluated for $t=(n+1)T$, then p first-order difference equations which describe the sampled output in terms of $H(n)$ are obtained. By substituting $C(n)=R(n)-E(n)$, $\dot{C}(n)=\dot{R}(n)-\dot{E}(n)$, etc., into these equations and noting that $H(n)=f[E(n)]$, the p first-order difference equations which involve only the sampled input and sampled-error signals are formed. The sampled error is used as the dependent variable because the input and error terms separate. This would not be the case if the sampled output is used as the dependent variable since $H(n)=f[R(n)-C(n)]$. This method of determining the difference equation is completely general and does not depend upon the type of nonlinear element which follows the hold circuit in Fig. 1. However, in this paper the concern will be only with a single-valued saturating nonlinearity which might be represented in a piecewise linear fashion, for example, by equation 2.

$$H(n)=f[E(n)]=\begin{cases} KE_0; & E(n)>KE_0 \\ KE(n); & -KE_0<E(n)<KE_0 \\ -KE_0; & E(n)<-KE_0 \end{cases} \quad (2)$$

K is the gain of the linear region of the saturating element. Assume that the linear part of the system, i.e., $G(s)$ in Fig. 1, has at least one but not more than $(p-1)$ integrators. This means that the step-function response of the linear part of the system of Fig. 1 may contain terms involving t, t^2, t^3, \dots, t^K , but $K < (p-1)$. Then, under these conditions, for any finite sampling period, it can be shown⁶ that the magnitude of the sampled-error signal is bounded. This means that the sampled-error signal must have one of two properties: 1. it can approach a constant value, i.e., an equilibrium point in the phase space, and 2. it can approach a periodic solution in the phase space, i.e., a limit cycle.

Due to the assumption that $G(s)$ possesses at least one integrator, it is known that if the sampled-error signal approaches a constant value, this constant value is zero. Also known is the fact that if any oscillations are to occur then these oscillations must take place about the origin. Again, the reason for this is that there is at least one integrator in $G(s)$ and feedback from output to input.

The system is defined stable if, to a bounded input, there corresponds a bounded output, and the system is not capable of maintaining self-sustained oscillations. It has been indicated that the error signal (or output) is bounded and we would now like to develop a condition which must be satisfied if no limit cycle is to exist in the saturated system. For the systems investigated the difference equations which describe the sampled-error signal are:

$$E(n+1)=[f_{11}(E(n),T)]E(n)+f_{12}(T)\dot{E}(n)+f_{13}(T)\ddot{E}(n)+\dots \quad (3A)$$

$$\dot{E}(n+1)=[f_{21}(E(n),T)]E(n)+f_{22}(T)\dot{E}(n)+f_{23}(T)\ddot{E}(n)+\dots \quad (3B)$$

$$\ddot{E}(n+1)=[f_{31}(E(n),T)]E(n)+f_{32}(T)\dot{E}(n)+f_{33}(T)\ddot{E}(n)+\dots \quad (3C)$$

In these equations the nonlinear terms appear only as coefficients of $E(n)$ since there is an amplitude-dependent nonlinearity; the other coefficients are functions of only T , the sampling period. Also, the equations are not symmetrical, i.e., $f_{ij}(T) \neq f_{ji}(T)$. If the system is to be stable, then there exists a steady-state value of the error signal (which for systems with integrators is zero) called $E_{ss}(n)$, which can be obtained from equation 3(A) with $E(n+1)=E(n)$ and all derivative terms set equal to zero. Imagine that the system has reached this steady-state value. Then, if $E_{ss}(n)$ is to be an equilibrium point, any perturbation about $E_{ss}(n)$ must result in decaying oscillations; that is, the operating point of the system must eventually return to $E_{ss}(n)$ if this operating point is changed slightly from its steady-state value. Now, if the perturbed values of $E(n)$, $\dot{E}(n)$, etc., are denoted by $e(n)$, $\dot{e}(n)$, etc., then to a first approximation, the difference equations for the perturbed error are:¹

$$e(n+1)=\bar{f}_{11}(T, E_{ss})e(n)+f_{12}(T)\dot{e}(n)+f_{13}(T)\ddot{e}(n)+\dots \quad (4A)$$

$$\dot{e}(n+1)=\bar{f}_{21}(T, E_{ss})e(n)+f_{22}(T)\dot{e}(n)+f_{23}(T)\ddot{e}(n)+\dots \quad (4B)$$

$$\ddot{e}(n+1)=\bar{f}_{31}(T, E_{ss})e(n)+f_{32}(T)\dot{e}(n)+f_{33}(T)\ddot{e}(n)+\dots \quad (4C)$$

where

$$\bar{f}_{ij}=\frac{\partial}{\partial E(n)} [f_{ij}(E(n),T)E(n)]|_{E(n)=E_{ss}} \quad j=1,2,\dots,p \quad (5)$$

Equations 4(A-C) are p first-order difference equations which can be solved for the perturbed error, $e(n)$, in a straightforward manner.⁷ If the solution for $e(n)$ shows that the perturbation decreases with increasing n , then $E_{ss}(n)$ is a stable equilibrium point. On the other hand,

if the solution shows that the perturbation does not decay with increasing n , then $E_{ss}(n)$ is an unstable point and a limit cycle will exist. This is certainly a necessary condition for the existence of a limit cycle.

We would now like to indicate the plausibility of this being a sufficient condition although this cannot be proved in a rigorous fashion. In the equations for the perturbed error signal 4(A-C) the quantities $f_{ij}(T, E_{ss})$ give the slope of the saturation curve at E_{ss} . This means that the equations for the perturbed error signal are simply the nonlinear difference equations which describe the system, linearized in the vicinity of E_{ss} . Also, the equations for the perturbed error signal are the same as those of the completely linearized system where this system is one in which the saturation effect is ignored and the saturating element is assumed to have a gain of K over an indefinite range of input magnitudes. If the actual saturation curve has its maximum slope or highest gain in the vicinity of the origin, then the equations for the perturbed error signal are linearized in the vicinity of the highest gain. Now, if the completely linearized system is stable for this maximum gain, the equations for the perturbed error signal indicate decaying oscillations since these equations are the same as those of the completely linearized system. From a stability point of view the completely linearized system is "worse" than the saturated system because there are no bounds placed on the magnitude of the error signal. Consequently, it is felt that if the system is stable when the error signal is not bounded, it should be stable when this signal is bounded. And because the equations for the perturbed error signal are the same as the equations for the completely linearized system, it is believed that the previously stated condition is also a sufficient condition for stability.

In other words, the original nonlinear system has been replaced with a linear system having a gain equal to the maximum gain of the saturating system. Then if the linear system, which places no bound on the magnitude of the various signals in this system, is stable, it seems as though the saturated system, which restricts the magnitude of the various signals, should also be stable since the equations for the completely linearized system are the same as those for the perturbed error of the saturating system. The definite exceptions to this statement are those systems which are conditionally stable, in which case a root-locus plot of the poles of the completely linearized sys-

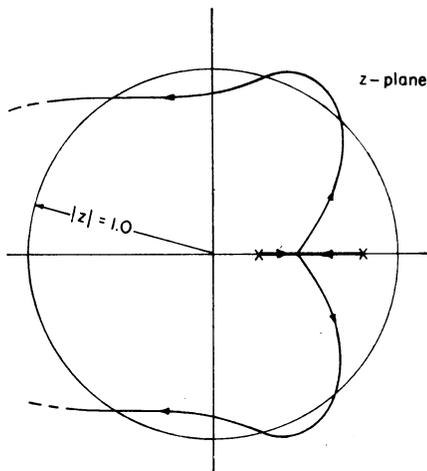


Fig. 2. A root locus plot of two poles of a conditionally stable sampled-data system. Arrows indicate the direction of increasing gain

tem is outside of the unit circle in the z -plane for certain values of gain but inside the unit circle for larger values of gain. Such a root-locus plot might be as shown in Fig. 2 and it seems likely that such a condition would occur only for very small values of T .

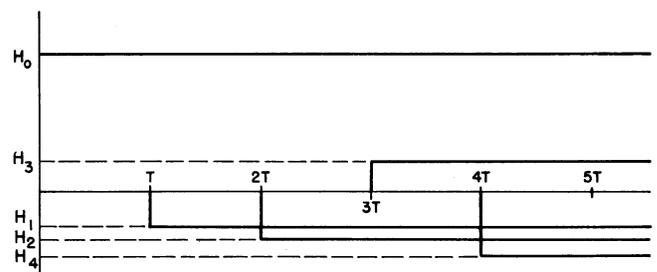
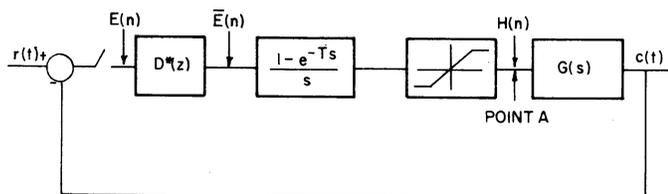
The summary of the stability criterion for these systems might be stated in the following manner. A necessary condition for saturating sampled-data systems which contain at least one integrator to be stable (no limit cycle exists) is that the equations for the perturbed error signal show that the perturbation decreases with increasing n . It seems plausible that this is a sufficient condition but this could not be proved rigorously.

Compensation Procedure

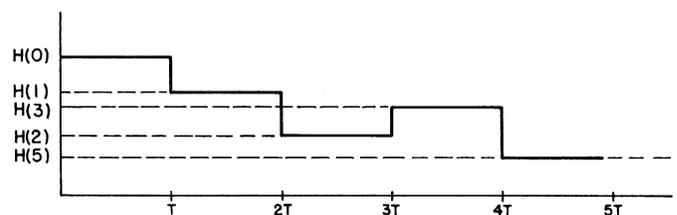
The system to be compensated is that of Fig. 1. It is wished to find a digital filter or discrete compensator which will operate on the sampled-error signal in such a fashion as to force the error signal to zero for inputs which are step functions of restricted magnitude or ramps with restricted slopes. In block diagram terminology this means that we want to find the

Fig. 3 (below). A block diagram showing the position of the discrete compensator, $D^*(z)$

Fig. 4 (right). (A) The compensating step-function sequence at point A. (B) The equivalent pulse sequence at point A, $H(0)=H_0$, $H(1)=H_0+H_1$, $H(2)=H_0+H_1+H_2$, etc.



(A)



(B)

$D^*(z)$ of Fig. 3 which will force the error to zero, subject to the condition that $D^*(z)$ be physically realizable. By physically realizable³ is meant that the roots of the denominator of $D^*(z)$ must lie inside the unit circle in the z -plane and that the highest power of z in the denominator of $D^*(z)$ must be equal to, or greater than, the highest power of z in the numerator.

For simplicity, assume that what is wanted is to find a discrete compensator which will force the system to follow step functions with a restricted range of magnitudes. Apply a step function equal to the saturation level of the amplifier to the linear part of the system and find the time required for its output to reach a value equal to the maximum step function the system is to follow. This time places a lower bound on the time or number of sampling periods required for the error to go to zero for the entire range of input magnitudes and determines the minimum number of sampling periods which must elapse before the output and input can be equal for all values of the input. Now since this minimum time was determined by applying the maximum step function possible to the linear part of the system, it is obvious that the output of the system cannot be made equal to the input for all time in this minimum time. The time required to do this must be greater than this minimum time since a negative signal must be applied to the linear part of the system to keep the output from overshooting its desired value. In other words, if the system is initially driven in one direction with a series of positive pulses, then the only means of bringing the output to some constant value is to apply a series of negative pulses to the system later. If

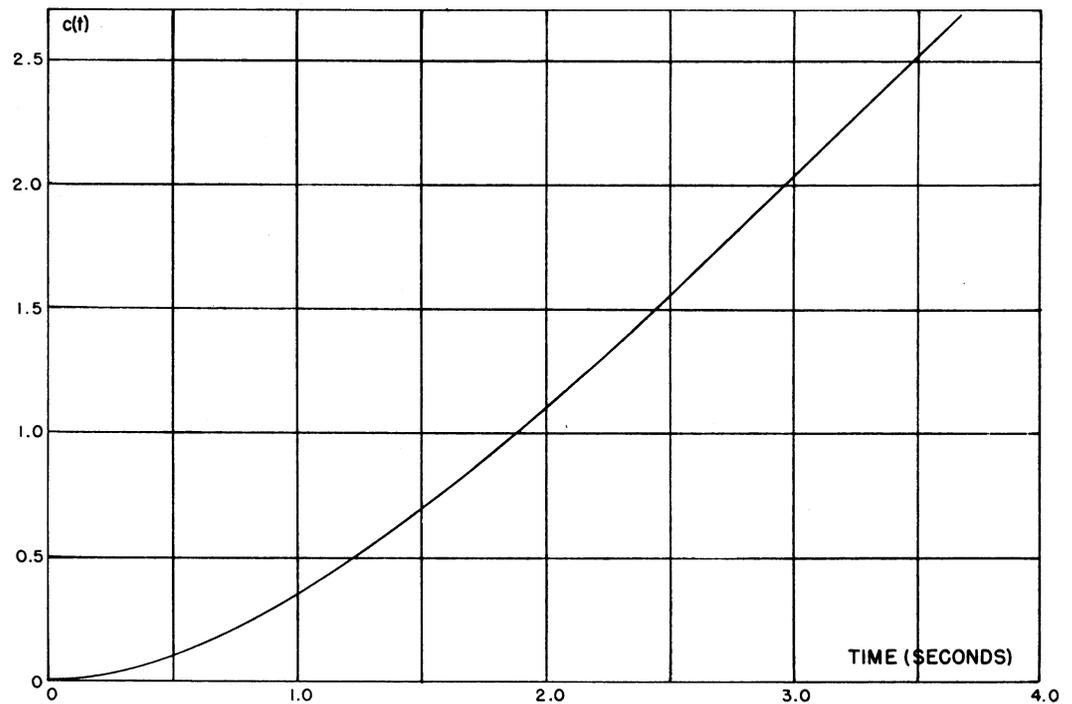
these later pulses are properly chosen it is possible to force the output to be equal to the input for all time after the second pulse sequence is applied. Thus, the initial problem which must be solved is to find a series of pulses of duration T which will cause the output and input to be the same within a reasonable length of time. The duration of these pulses must be T because the error signal of the system can change only at the sampling instants, and the magnitudes of these pulses must be less than the saturation level of the amplifier. This second condition is the one which causes the difficulty and makes this problem different from that of finding a discrete compensator for a linear system.

The series of pulses required to force the output to equal the input can be found as follows. Consider the sequence of pulses which are applied to the input of the linear element to be composed of a series of delayed step functions as shown in Fig. 4. Then, using the delayed step functions as the input to the linear part of the system, the output of the linear element can be written:

$$c(t) = \sum_{i=0}^K H_i h(t-iT) \quad (6)$$

where $h(t)$ is the step-function response of the linear system. The number of step functions, $(K+1)$, is determined from a consideration of the range-input magnitudes which the system should follow and it is possible to determine the range of input magnitudes for a given K . This will be pointed out in example II. This equation for $c(t)$ will contain terms which involve t and terms which do not involve t . The quantities which involve t are equated to zero for $t > KT$ and the sum of the constant terms is equated to the

Fig. 5. The unit step-function response of $G(s) = 1/s(s+1)$



value of the largest input step function the system will receive. This gives the relationship between the H_i 's necessary to force the output to be equal to the input. There will always be more unknown H_i 's than there are equations indicating that some of the H_i 's can be arbitrarily specified. For convenience the H_i 's will be arbitrarily specified from H_0 onward. For example, if three values of H_i can be specified, then the three values chosen will be assigned to $H_0, H_1,$ and H_2 . Then solve for the unknown H_i 's in terms of those specified as arbitrary.

The word arbitrary has been used loosely here since the condition of restricted magnitude must be imposed on the amplitude of the impulses. To find what must be done to satisfy this condition, let V denote the saturated output of the amplifier. Then, the restricted magnitude condition can be stated:

$$|H_0| \leq V \tag{7A}$$

$$|H_0 + H_1| \leq V \tag{7B}$$

$$|H_0 + H_1 + \dots + H_K| \leq V \tag{7K}$$

Since all of the unspecified H_i 's have been expressed in terms of those we are at liberty to specify, all of these $(K+1)$ equations can be written in terms of only those H_i 's which are arbitrary. These equations take care of the magnitude restrictions and then simply find the unspecified H_i 's which satisfy these magnitude conditions. After obtaining them, compute the other H_i 's and obtain the desired sequence of pulses. This can be done in a straightforward manner.

Once the pulse sequence has been deter-

mined, the discrete compensator which will produce this sequence of pulses can very easily be found. Knowing the pulse sequence, the sampled output of the system and the signal $\bar{E}(n)$ of Fig. 3 may be found as also the sampled error, $\bar{E}(n)$, since the sampled input and output are known. $E(n)$ is the input to $D^*(z)$ and $\bar{E}(n)$ is the output of $D^*(z)$. Both

are discrete signals which are zero after a finite time. This means we can take the z -transform of these two signals and obtain

$$D^*(z) = \frac{\bar{E}^*(z)}{E^*(z)} \tag{8}$$

$D^*(z)$ will always be a stable device since it has a bounded output, $\bar{E}^*(z)$, for a

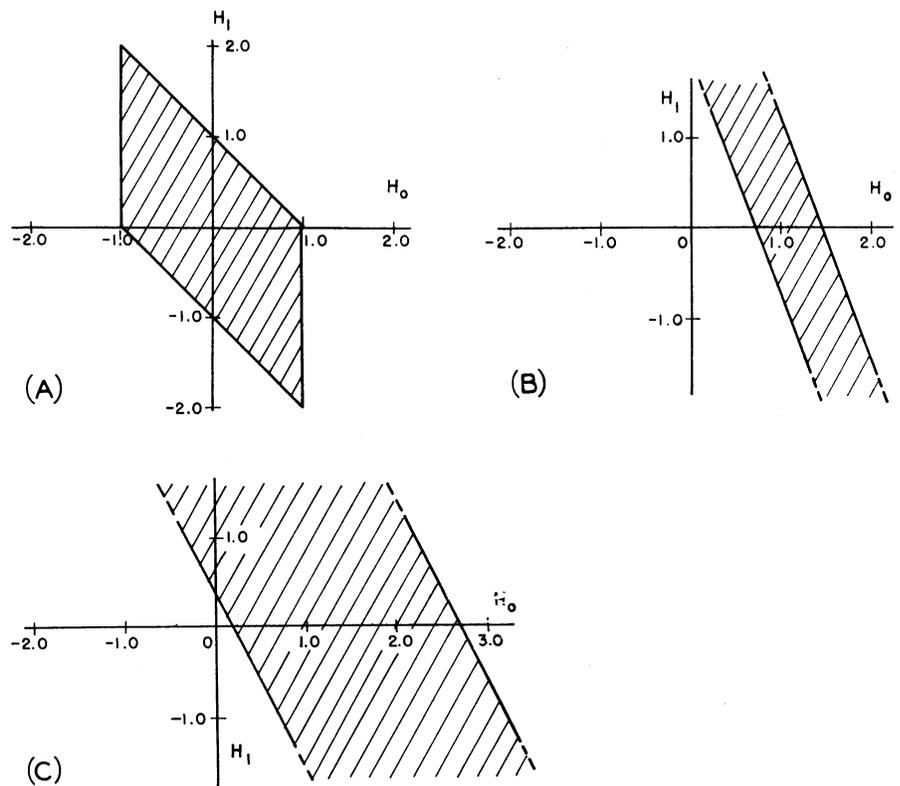


Fig. 6. (A) Magnitude restrictions imposed by equations 21 and 22. (B) Magnitude restriction imposed by equation 23. (C) Magnitude restriction imposed by equation 24

bounded input, $E^*(z)$. It is physically realizable since its output is zero before the input is applied.

In this process, the system has actually been compensated for the largest step the system will receive and in doing this the amplifier has really been forced to operate in its linear region. This might not be evident from the preceding discussion but it will be made clear in the examples which follow. Consequently, the system will follow all step functions within the range considered. In addition a system has been arrived at which is guaranteed to be stable since there is a bounded output for a bounded input.

Example I

In Fig. 3 let $G(s) = 1/s(s+1)$, $T = 1.0$ second, and for convenience assume the saturation curve is given by equation 2 with $K = 1.0$. A discrete compensator is to be found which will allow the system to follow step inputs whose magnitudes may be as large as two units.

First find the response of $G(s)$ to the largest step function this linear network can receive. In this example it will be a unit step and for a unit step input, the output of $G(s)$ is

$$c(t) = t - 1 + e^{-t} \quad (9)$$

This response is shown in Fig. 5 and it is seen that it takes 2.95 seconds for $c(t)$ to reach 2.0 volts. This is the minimum time required to have the output equal to the input for all inputs up to and including 2 volts. In other words, it will require at least three sampling periods to do this but four sampling periods will allow a little freedom in finding the discrete compensator. If the output is to equal the input after four sampling periods, then $K = 4$ in equation 6 and the output of the linear element can be written:

$$c(t) = \sum_{i=0}^4 H_i [(t-iT) - 1 + e^{-(t-iT)}] \quad (10)$$

This equation may be written in the following form with $T = 1.0$

$$c(t) = t \sum_{i=0}^4 H_i - \sum_{i=0}^4 i H_i t - \sum_{i=0}^4 H_i + e^{-t} \times \sum_{i=0}^4 H_i e^i \quad (11)$$

If $c(t)$ is to equal 2.0 for all $t > 4.0$, then the first and third terms of this equation require that

$$\sum_{i=0}^4 H_i = 0 \quad (12)$$

and the last term requires that

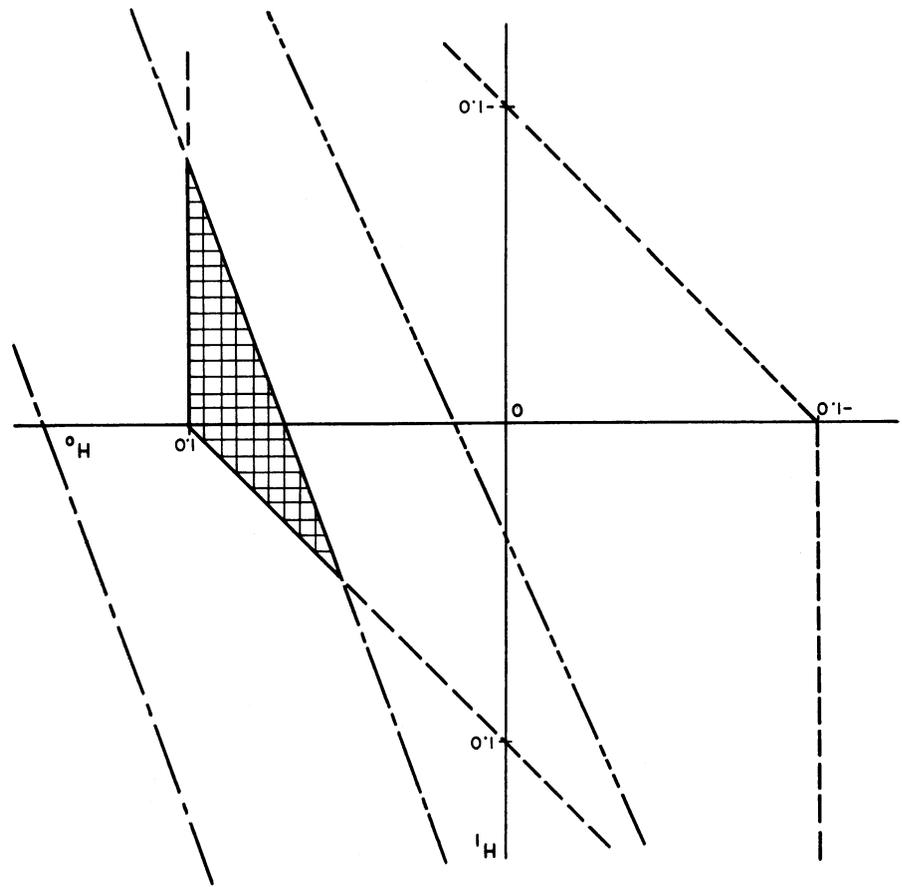


Fig. 7. The superposition of all magnitude restrictions imposed on H_0 and H_1 . Shaded region shows permissible values

$$\sum_{i=0}^4 H_i e^i = 0 \quad (13)$$

The second term, which does not involve time, is equated to 2.0 giving us the third condition which must be satisfied. It is:

$$\sum_{i=0}^4 i H_i = -2.0 \quad (14)$$

If these equations are satisfied, then $c(t)$ will equal 2.0 for all $t > 4.0$. Writing out these equations shows:

$$H_0 + H_1 + H_2 + H_3 + H_4 = 0 \quad (15)$$

$$H_0 + 2.718H_1 + 7.39H_2 + 20.09H_3 + 54.60H_4 = 0 \quad (16)$$

$$H_1 + 2H_2 + 3H_3 + 4H_4 = -2.0 \quad (17)$$

As has been previously mentioned, we have three equations and five unknowns. To solve these equations assume that H_0 and H_1 are known and solve for H_2 , H_3 , and H_4 in terms of H_0 and H_1 . Then

$$H_2 = -3.82H_0 - 2.03H_1 + 3.12 \quad (18)$$

$$H_3 = 3.67H_0 + 1.705H_1 - 4.25 \quad (19)$$

$$H_4 = -0.86H_0 - 0.397H_1 + 1.14 \quad (20)$$

Now impose the magnitude restrictions. This is most easily done by drawing an H_0 , H_1 plane and plotting the equations

which must be satisfied if the magnitude of the pulses are to be limited. Refer now to Fig. 6. The first two magnitude conditions are:

$$|H_0| \leq 1.0 \quad (21)$$

$$|H_0 + H_1| \leq 1.0 \quad (22)$$

Equation 21 eliminates that part of H_0 , H_1 plane to the right of the line $H_0 = 1$ and the left of the line $H_0 = -1.0$; equation 22 eliminates the section above the line $H_0 + H_1 = +1.0$ and below the line $H_0 + H_1 = -1.0$. This is shown in Fig. 6(A). The third magnitude condition is:

$$|H_0 + H_1 + H_2| \leq 1.0$$

Using equation 18 this becomes:

$$|-2.82H_0 - 1.03H_1 + 3.12| \leq 1.0 \quad (23)$$

Equation 23 eliminates the section to the right of the line $(-2.82 H_0 - 1.03 H_1 + 3.12 = 1.0)$ and to the left of the line $(-2.82 H_0 - 1.03 H_1 + 3.12 = -1.0)$ and this restriction is shown in Fig. 6(B). The fourth condition is:

$$|H_0 + H_1 + H_2 + H_3| \leq 1.0$$

And using equations 18 and 19 this becomes:

$$|0.86H_0 + 0.397H_1 - 1.14| \leq 1.0 \quad (24)$$

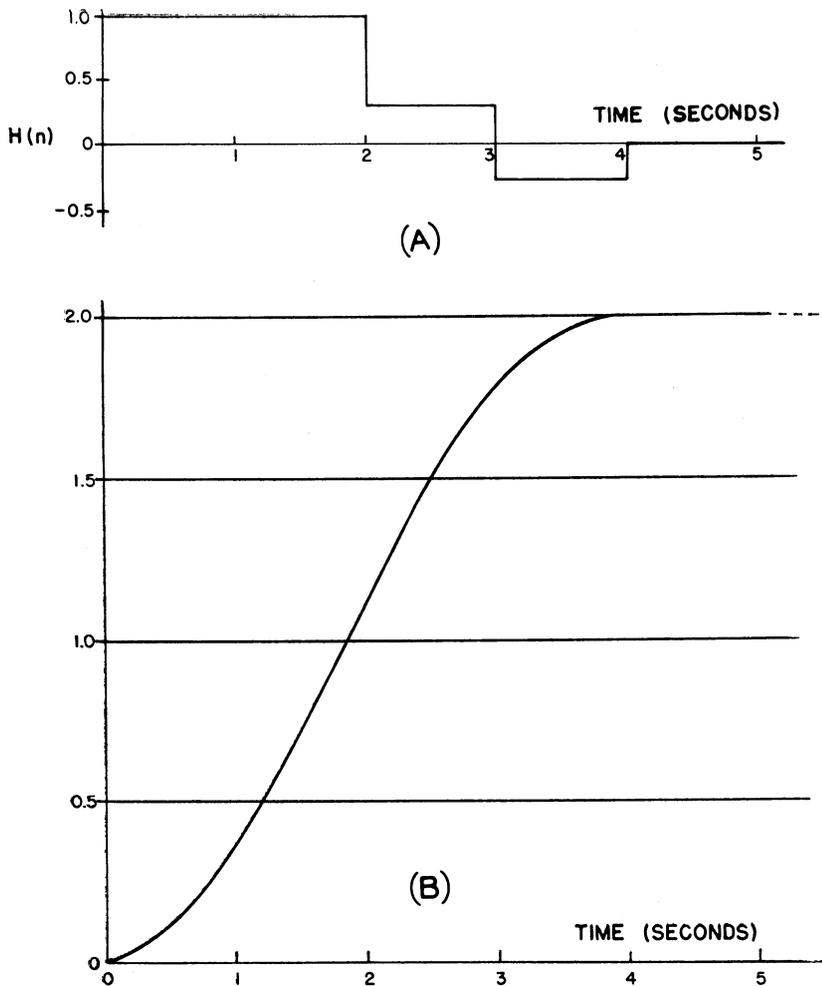


Fig. 8. (A) The pulse sequence to compensate $G(s) = 1/s(s+1)$ with $T = 1.0$ second. (B) The output of the system when this pulse sequence is applied to the input of $G(s)$

Equation 24 eliminates no additional region of H_0, H_1 plane as is seen in Fig. 6(C). The fifth magnitude condition is automatically satisfied since we have already imposed the condition $\sum H_i = 0$ in determining the individual H_i 's.

The result of all the magnitude restrictions is completely shown in Fig. 7. It is seen that any values of H_0 and H_1 bounded by the three lines ($H_0 = 1.0$), ($H_0 + H_1 = 1.0$), and ($-2.82 H_0 - 1.13 H_1 + 3.12 = -1.0$), will satisfy the magnitude restrictions and also force the output to equal the input in 4 seconds. To find the sequence of pulses let us arbitrarily select $H_0 = 1.0$ and $H_1 = 0$. Then $H_2 = -0.70$; $H_3 = -0.58$; and $H_4 = 0.28$. For these values of H_i , the desired pulse sequence is 1, 1, 0.30, -0.28 , 0, 0, . . . This pulse sequence and the output of the linear system are shown in Fig. 8.

Since the output has been allowed to reach a value of 2.0, this is the response the system will have when the input is 2.0. Determining the discrete compensator which will give this response when the input is 2.0 units is straightforward now that the proper sequence of pulses

has been found. If the input is 2.0 units and the output of the system is as shown in Fig. 8(B), then the sampled error is:

$$E(n) = \begin{cases} 2.0 & n=0 \\ 1.632 & n=1 \\ 0.865 & n=2 \\ 0.210 & n=3 \\ 0.0 & n \geq 4 \end{cases} \quad (25)$$

Knowing the pulse sequence which must be applied to the linear part of the system, that is, at the output of the saturating amplifier, next find the input to the amplifier required to give this output. In this example assume the gain of the linear region of the amplifier to be unity and so:

$$\bar{E}(n) = \begin{cases} 1.0 & n=0 \\ 1.0 & n=1 \\ 0.30 & n=2 \\ -0.28 & n=3 \\ 0.0 & n \geq 4 \end{cases} \quad (26)$$

Using equation 8 it is finally found:

$$D^*(z) = \frac{1.0 + 1.0z^{-1} + 0.30z^{-2} - 0.28z^{-3}}{2.0 + 1.632z^{-1} + 0.865z^{-2} + 0.21z^{-3}}$$

or

$$D^*(z) = \frac{1.0z^3 + 1.0z^2 + 0.30z - 0.28}{2.0z^3 + 1.632z^2 + 0.865z + 0.21} \quad (27)$$

This is the discrete compensator which will force the output to follow a step-function input whose magnitude is less than or equal to 2.0 units.

In addition to showing the procedure for determining the discrete compensator, this example has also indicated several other characteristics of the method. The most obvious one is that an identical scheme can be used to find a discrete compensator which will follow ramp inputs. To do this all that must be done is to place different specifications on the H_i 's of equation 11. For example, if it is wished to find a discrete compensator to follow ramp inputs whose slope was less than one, then impose the conditions

$$\begin{aligned} \sum_{i=0}^4 H_i &= 1.0 \\ -\sum_{i=0}^4 iH_i - \sum H_i &= 4.0 \\ \sum H_i e^{i\theta} &= 0 \end{aligned}$$

and find the H_i 's as before, subject to the magnitude restrictions. With the H_i 's determined, the discrete compensator is found exactly as it was for a step input.

Also, if the system has a time delay, this same method can be used both for step and ramp inputs. And, if it is desired, some of the H_i 's which were left unspecified can be used to place restrictions on the derivatives of the response to insure, for example, a continuously increasing response. Thus, having arbitrary values of H_i at our disposal is good in that it allows additional freedom in meeting simultaneous specifications on the response and its derivatives.

Example II

In example I the discrete compensator was determined in a relatively easy fashion but no investigation was made in any detail on the limitations which arise when the time at which we wish the output to equal the input is selected. In this example it will be shown how this time, or the number of pulses in the pulse sequence used for compensation, is related to the range of input step functions the system will follow.

In Fig. 3 let $G(s) = (s+0.5)(s+2)/s(s^2+2s+s)$ and $T = 1.0$ second. Also let the input-output characteristics of the amplifier be given again by equation 2 with $K = 1.0$. If a unit step function is applied at the input of $G(s)$, then the output of $G(s)$ is:

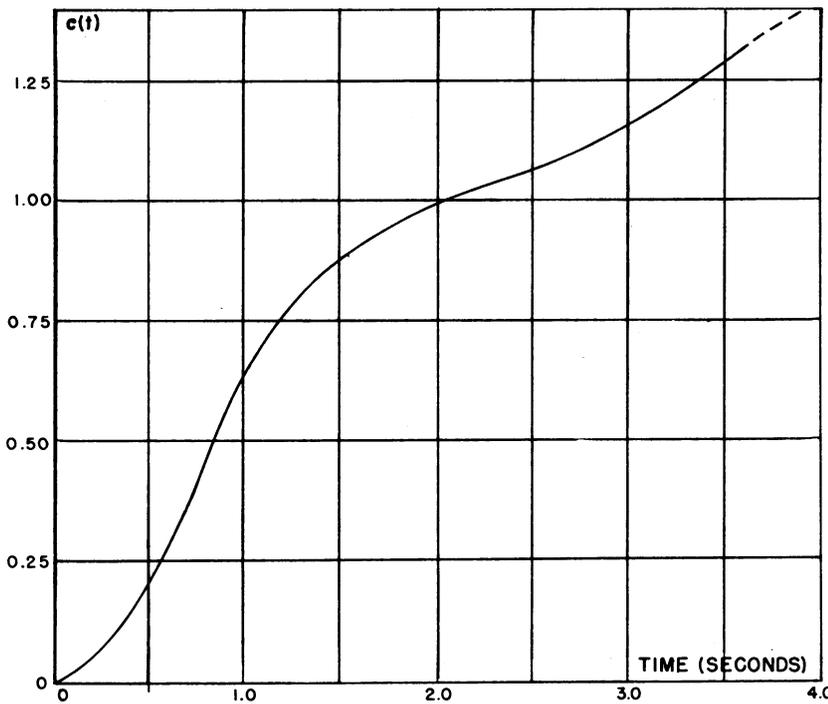


Fig. 9. The unit step-function response of $G(s) = \frac{(s+0.5)(s+2)}{s^2+2s+5}$

$$c(t) = 0.25 + 0.42 - 0.521e^{-t} \times \sin(2t + 53.6^\circ \text{ (degrees)}) \quad (28)$$

This response is shown in Fig. 9.

Use a sequence of six pulses to compensate the system and investigate exactly what can be done with those six pulses. With this input to $G(s)$ its output is:

$$c(t) = \sum_{i=0}^6 H_i [0.25(t-i) + 0.42 - 0.521e^{-(t-i)} \sin[2(t-i) + 53.6^\circ]]$$

or

$$c(t) = 0.25t \sum_{i=0}^6 H_i - 0.25 \sum_{i=0}^6 iH_i + 0.42 \times \sum_{i=0}^6 H_i - 0.521e^{-t} \sum_{i=0}^6 e^i H_i \times [\cos 2i \sin(2t + 53.6^\circ) - \sin 2i \cos(2t + 53.6^\circ)] \quad (29)$$

Now suppose that $c(t) = A$, a constant, for $t > 6.0$ and try to determine the maximum value of A . Then from equation 29 it can be seen that for this condition to be satisfied

$$\sum_{i=0}^6 H_i = 0 \quad (30)$$

$$-0.25 \sum_{i=0}^6 iH_i = A \quad (31)$$

$$\sum_{i=0}^6 H_i e^i \cos 2i = 0 \quad (32)$$

$$\sum_{i=0}^6 H_i e^i \sin 2i = 0 \quad (33)$$

Here are four equations and seven unknowns. Consider $H_0, H_1,$ and H_2 as known and solve for the other H_i 's in terms of $H_0, H_1, H_2,$ and A . Doing this it is found that the unknown H_i 's can be expressed as:

$$H_3 = -1.693H_2 - 2.345H_1 - 3.051H_0 + 2.773A \quad (34)$$

$$H_4 = -1.847H_3 - 2.622H_2 - 3.412H_1 - 4.216H_0 + 3.20A \quad (35)$$

$$H_5 = 2.0H_4 - 3.0H_3 - 4.0H_2 - 5.0H_1 - 6.0H_0 + 4.0A \quad (36)$$

$$H_6 = -H_5 - H_4 - H_3 - H_2 - H_1 - H_0 \quad (37)$$

In these equations specify $H_0, H_1,$ and H_2 . Then take $H_0 = 1.0$ to try to obtain a good initial response but let H_1 and H_2 remain arbitrary and assume that whatever these values are, they satisfy the magnitude conditions imposed by the saturation curve. This means that:

$$|1 + H_1| \leq 1.0 \quad (38)$$

$$|1 + H_1 + H_2| \leq 1.0 \quad (39)$$

The fourth magnitude condition is:

$$|H_0 + H_1 + H_2 + H_3| \leq 1.0$$

$$|-2.051 - 1.345H_1 - 0.693H_2 + 2.773A| \leq 1.0 \quad (40)$$

Now find values of H_1 and H_2 which will maximize A subject to equations 38 and 39. Since H_1 must be < 0 , the best choice of H_1 is $H_1 = 0$. Similarly if $H_1 = 0$, then $H_2 < 0$ and also is set $H_2 = 0$. This means:

$$|-2.051 + 2.773A| \leq 1.0 \quad (41)$$

So the maximum value of A is given by:

$$-2.051 + 2.773A = 1.0 \quad (42)$$

or

$$A = 1.10 \quad (43)$$

Thus, using six pulses the best that can be done is to obtain a discrete compensator which will force the system to follow step inputs up to 1.10 times the voltage required to saturate the amplifier. This is because the maximum constant value the output can attain with six pulses of restricted magnitude is 1.00. Whether or not this can be done depends upon the restrictions imposed by the remaining magnitude conditions.

If a sequence of eight pulses had been used to compensate the system, equations 30-33 would have been the same, except that the summations would be from $i=0$ to $i=8$. In this case there would be four equations and nine unknowns and if it is solved for the last four H_i 's in terms of the first five, it would be found that:

$$H_5 = -(1.578H_4 + 2.157H_3 + 2.837H_2 + 3.489H_1 + 4.133H_0) + 2.583A \quad (44)$$

$$H_6 = -(2.235H_5 + 3.262H_4 + 4.321H_3 + 5.395H_2 + 6.466H_1 + 7.534H_0) + 4.275A \quad (45)$$

$$H_7 = -(-2.0H_6 - 3.0H_5 - 4.0H_4 - 5.0H_3 - 6.0H_2 - 7.0H_1 - 8.0H_0) + 4A \quad (46)$$

$$H_8 = -(H_7 + H_6 + H_5 + H_4 + H_3 + H_2 + H_1 + H_0) \quad (47)$$

Then, as was done with the six pulses, start finding the restrictions on the H_i 's which the saturation effect causes. The maximum value of A will result if $H_0 = 1.0, H_1 = H_2 = H_3 = H_4 = 0$ and then the sixth magnitude condition reduces to:

$$|H_0 + H_5| = |1 + H_5| \leq 1.0$$

or,

$$|-3.133 + 2.583A| \leq 1.0$$

This shows that the maximum value of A is given by:

$$A = \frac{4.133}{2.583} = 1.60 \quad (48)$$

Thus, using an 8-pulse sequence, the best that can be done is obtain a discrete compensator which will force the system to follow step inputs up to 1.60 times the voltage at which the amplifier saturates. The remaining magnitude conditions will determine whether or not this is possible.

This example has shown how one can determine the range of input which can be followed for a given pulse sequence. This means that once the range of inputs

which the system will receive is known, it is possible to design a discrete compensator to force the system to follow these step inputs in a straightforward manner. There is no restriction on the order of the transfer function but this transfer function must have at least one integrator. If the system exhibits a limit cycle without the compensator, then the discrete compensator both stabilizes and compensates the system.

Conclusions

Saturating sampled-data systems have been examined and a stability criterion for such systems was developed. Also, a method to determine a digital filter or

discrete compensator to force these systems to follow step and ramp inputs was presented and it was also shown how the range of input signal magnitudes which the system will follow is related to the complexity of the digital filter. The procedure presented to obtain this filter was developed for sampled-data systems but it can be equally well applied to find a digital filter for continuous systems which have a saturating element.

References

1. ANALYSIS AND DESIGN OF NONLINEAR SAMPLED-DATA CONTROL SYSTEMS, R. F. Nease. *WADC Technical Note 57-162*, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., June 1957.
2. OPTIMAL CONTROL OF SATURATING SYSTEMS BY

INTERMITTENT ACTION, R. E. Kalman, *WESCON Convention Record*, Institute of Radio Engineers, New York, N. Y., pt. 4, 1957, pp. 130-35.

3. THE ANALYSIS OF SAMPLED-DATA SYSTEMS, J. R. Ragazzini, L. A. Zadeh. *AIEE Transactions*, vol. 71, pt. II, Nov. 1952, pp. 225-34.

4. SAMPLED-DATA CONTROL SYSTEMS, (book), E. I. Jury. John Wiley & Sons, Inc., New York, N. Y., 1958.

5. SYNTHESIS AND CRITICAL STUDY OF SAMPLED-DATA CONTROL SYSTEMS, Eliahu I. Jury, *AIEE Transactions*, vol. 75, pt. II, July 1956, pp. 141-51.

6. THE ANALYSIS AND COMPENSATION OF NON-LINEAR SAMPLED-DATA FEEDBACK SYSTEMS, F. J. Mullin. *Ph.D. Thesis*, University of California, Berkeley, Calif., Sept. 1958.

7. A NOTE ON THE OPERATIONAL SOLUTION OF LINEAR DIFFERENCE EQUATIONS, E. I. Jury, F. J. Mullin. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 266, no. 3, Sept. 1958, pp. 189-205.

8. DISCRETE COMPENSATION OF SAMPLED-DATA AND CONTINUOUS CONTROL SYSTEMS, Eliahu I. Jury, William Schroeder. *AIEE Transactions*, vol. 75, pt. II, 1956 (Jan. 1957 section), pp. 317-25.

Discussion

C. A. Desoer and S. Kodama (University of California, Berkeley, Calif.): The author has presented an interesting paper on nonlinear sampled-data systems. He deserves commendation, for this is a difficult subject on two counts: first, it is nonlinear, and second, the presence of the sampler makes it time varying. We would like, however, to offer some comments* particularly on the section of the paper entitled, "Requirements for Stability."

From the paper it seems that the author believes the following to be true: For a saturating sampled-data servo such as the one shown in Fig. 1 of the paper, when the transfer function $G(s)$ is of the order p and contains at least one but not more than $(p-1)$ integrators then, for any finite sampling period, the magnitude of the sampled-error signal is bounded.

First, this theorem as it stands is misleading and may suggest incorrect conclusions, and second, a study of the author's proof (reference 6 of the paper) discloses some gaps in the reasoning.

That the theorem is misleading is best shown by a counter example:¹ Let the saturating amplifier have unity gain over its linear region and let $G(s) = 1/s^2(s+1)$, then it can be shown that, if the input $r(t)$ remains equal to zero, any output phase-space trajectory which starts from the subspace defined by $c(nT) = E_0$, $\dot{c}(nT) \geq 5.5 E_0$, $E_0 \geq \check{c}(nT) > 0$, becomes unbounded regardless of the sampling period. More precisely, it can be shown that the output behaves in an oscillatory manner; its amplitude and half-cycle period increase with time at least logarithmically.

Since the author's theorem leads to erroneous conclusions it should be useful to indicate where the proof's logic (reference 6 of the paper) is faulty. The author's method of proof consists of the following steps: 1. he assumes that all the derivatives of $e(t)$ of order $p-1$ and higher are bounded, 2. he uses equations 3(A), 3(B), 3(C), etc., to express $E(n+1)$ as a linear combination of

the $(p-1)$ th derivatives evaluated at nT , $(n+1)T$, $(n+2)T$, etc., and 3. the form of the linear combination makes it appear as if the assumed boundedness of these derivatives implies the boundedness of $E(n+1)$ for all n . This last step is erroneous because 1. in some cases some of the coefficients of equations 3 are identically zero and the proposed elimination cannot be carried out (this is the case, for example, where $G(s)$ has more than one integrator), 2. some of the coefficients of equations 3 (and, hence, some of those of the linear combination giving $E(n+1)$) actually depend on $E(n)$. In particular it can be shown that as a consequence of the saturation characteristic of the amplifier, the coefficients $f_{21}(E(n), T)$, $f_{31}(E(n), T)$, . . . all go to zero as $E(n)$ becomes arbitrarily large. Since these coefficients appear in the denominator of the linear combination giving $E(n+1)$, the third step of the author's reasoning is incorrect.

In view of these considerations it should be worthwhile knowing a correct statement of sufficient conditions for boundedness of the output together with a proof that satisfies the objections just mentioned. First the statement:

For a sampled-data servo with a saturating amplifier such as that shown in Fig. 1, with the input $r(t)$ equal to zero for all $t > 0$ and an arbitrary set of initial conditions (by "arbitrary set of initial conditions" is meant any set of initial conditions that can be reached by applying to the servo any continuous input $r(t)$), the output $c(t)$ will be bounded for all t (including sampling instants) if 1. $G(s)$ has left half-plane poles and only one simple pole at $s=0$ with a positive residue, and 2. the degree of the numerator polynomial is at least one less than that of the denominator polynomial.

Note that the statement does not make any particular assumption regarding the shape of the amplifier-gain characteristic except that it is single valued and that it saturates, i.e., as $E(n)$ becomes large the amplifier output $h(t)$ becomes equal to $\pm H_0$, the saturating level of the amplifier, more precisely when $|e(t)| \geq E_0$, $h(t) = e(t)/|e(t)| H_0$.

The proof follows. If $g(t)$ is the unit impulse response of $G(s)$, then

$$c(t) = \int_{-\infty}^{+\infty} h(\tau)g(t-\tau)d\tau$$

Let p be the degree of the denominator polynomial of $G(s)$. Differentiating this expression k times with respect to t ,

$$c^k(t) = \int_{-\infty}^{+\infty} h(\tau)g^k(t-\tau)d\tau \quad k=1, 2, \dots, p-1$$

where $g^k(t)$ denotes the k th derivative of $g(t)$ taken in the distribution sense² so that it may contain impulses, doublets, etc., at $t=0$. As a result of the piecewise continuous character of $h(\tau)$, the aforementioned equality may not have a meaning where $k \geq 2$ for $t = nT$ because distribution theory does not define the integral of an impulse times a function that is discontinuous at the impulse. In view of the assumptions on $G(s)$, $g(t)$ is of the form

$$g(t) = A + \sum_k \left(\sum_l A_{kl} \frac{t^l}{l!} \right) e^{p_k t} \quad (49)$$

where $Re(p_k) < 0$ for all k 's, hence all its derivatives are absolutely integrable, in particular,

$$\int_{-\infty}^{+\infty} |g^k(t)| dt < M_k \quad (k=1, 2, \dots, p-1)$$

where M_k is an appropriate positive constant. Since for all t , $|h(t)| \leq H_0$ we have

$$|c^k(t)| \leq H_0 \int_{-\infty}^{+\infty} g^k(t-\tau)d\tau$$

and

$$|c^k(t)| < H_0 M_k \quad (k=1, 2, \dots, p-1)$$

the inequality being valid for all t if $k=1$ but may not be valid at sampling instants for $k \geq 2$. It should be stressed that the uniform boundedness of all derivatives of the output $c(t)$ (except possibly at sampling instants) follows from the fact that $G(s)$ contains a single integrator and that its other poles lie inside the left-half plane.

It remains to show that $c(t)$ itself is uniformly bounded. If it can be proved that the servo cannot remain in saturation for a time longer than T_1 (where T_1 is a constant independent of the initial conditions and dependent only on the characteristics of the servo) then, the already proved boundedness of $\dot{c}(t)$ implies that $c(t)$ is bounded. In particular

* The results reported here came from research supported by the U. S. Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract no. AF18(600)-1521.

$$|c(t)| < E_0 + H_0 M_1 \frac{T_1}{2}$$

This inequality follows from the following geometric argument: When the servo enters saturation, $c(t) = E_0$. (Here we assume positive saturation; in case of negative saturation $c(t) = -E_0$ but the remainder of the argument follows with a few trivial changes.) The existence of T_1 means that during the following T_1 seconds the servo has come out of saturation. Therefore the largest excursion of $c(t)$ would occur if $c(t)$ would increase at the maximum rate for $T_1/2$ seconds and decrease at the maximum rate for the remaining $T_1/2$ seconds.

It remains to establish the existence of time T_1 . Suppose the servo enters saturation ($c(t) > E_0$) in the interval $(m-1)T < t \leq mT$ and is still in saturation at $t = mT$, then $c(mT) < E_0 + H_0 M_1 T$ since the slope of the curve $c(t)$ is always smaller than $H_0 M_1$. Thus, as long as the servo remains in saturation the following situation exists: 1. the input to $G(s)$ is $-H_0$, and 2. the initial values just after the sampling instant mT are such that $E_0 < c(mT) < E_0 + H_0 M_1 T$, $|c^k(mT+)| < H_0 M_k$ ($k=1, 2, \dots, p-1$). The solution of the differential equation indicates that as long as the servo is in saturation and $t > mT$

$$c(t) = \sum_{k=0}^{p-1} c^k(mT+) \psi_k(t-mT) - H_0 \int_{mT}^t g(\tau-mT) d\tau \quad (50)$$

The pole locations of $G(s)$ require that the transients $\psi_k(t-mT)$ are of the same form as the right hand side of equation 49.

Therefore the function

$$f(t) = \sum_{k=0}^{m-1} c^k(mT+) \psi_k(t-mT)$$

is uniformly bounded for all $t > mT$ by a constant M which is independent of the initial conditions since the latter are uniformly bounded. Now, equation 50 cannot hold for arbitrarily large t since the integral $\int_{mT}^t g(\tau-mT) d\tau$, for large enough t , increases linearly with t . As a consequence $c(t)$ given by equation 50 will eventually become smaller than E_0 and the servo will come out of saturation. The time the servo will remain in saturation will be bounded above by $T_1 = T_0 + T$ where T_0 is the smallest number x such that

$$\int_0^x g(\tau) d\tau = \frac{M - E_0}{H_0}$$

From this expression it is apparent that T_1 depends exclusively on the characteristics of the servo and is independent of the initial conditions.

In short, it has been shown that, except maybe at the sampling instants $t = nT$, the phase-space point $[c(t), \dot{c}(t), \ddot{c}(t), \dots, c^{p-1}(t)]$ remains inside the parallelepiped defined by the following inequalities

$$|c(t)| < E_0 + H_0 M_1 \frac{T_1}{2}$$

$$|c^k(t)| < H_0 M_k \quad (k=1, 2, \dots, p-1)$$

These inequalities hold whatever might be the initial conditions. The only thing that was proved was the boundedness of $c(t)$ and its $(p-1)$ first derivatives. This does not imply that the servo is asymptotically stable which would require proving that $c(t)$ and its $(p-1)$ first derivatives go to zero, as $t \rightarrow \infty$.

REFERENCES

1. C. A. Desoer, S. Kodama. (available from the authors).
2. PRINCIPLES AND TECHNIQUES OF APPLIED MATHEMATICS (book), B. Friedman. John Wiley & Sons, Inc., New York, N. Y., 1956, chap. III.

Francis J. Mullin: The author would like to thank Professor Desoer and Mr. Kodama for their discussion and certainly agrees with their results: namely, that if the sampled-data system shown in Fig. 1 contains a saturating element and $G(s)$ is a transfer function with *one* integrator and all of its remaining poles in the left-half s -plane, then the continuous output of the system is bounded for all time. This is in contrast to the statement in the paper which permitted $G(s)$ to contain $(p-1)$ integrators and still have a bounded output. This statement in the paper of the restrictions on $G(s)$ is incorrect. $G(s)$ may contain only a single integrator if the output of the saturating system is to be bounded.

The proof for the boundedness of the output which the discussers have presented is both understandable and clever. The idea of showing that the system cannot remain indefinitely in saturation is the key to their argument and the fact that this is true, definitely shows the boundedness of the output.

The argument which was originally used to show the boundedness of the output was based on the difference equations which describe the sampled-error signal. The sampled-error signal was expressed as a combination of sampled derivatives of the error signal and since all derivatives are bounded, it was believed the sampled-error signal was bounded. This is true only if $G(s)$ is a stable transfer function with a single integrator. If this condition is not satisfied, then it is not possible to express the sampled-error signal in terms of the derivative quantities because some coefficients of the difference equations are identically zero. This fact was overlooked in reference 6. At any rate, the question on the boundedness of the output of the system in Fig. 1 has been cleared up by the discussion and the author appreciates the efforts of the discussers and this contribution.

Synchronous Methods of Operation for Private Telegraph Networks

B. S. DIAMOND
MEMBER AIEE

THERE ARE MANY private telegraph networks in the United States serving industrial, commercial, financial, and transportation organizations, and used for the transmission of intracompany message traffic, data transmission, and telemetering. In size there is a wide range, from small networks with a home office and a single branch, to large and highly complex ones with tens of thousands of miles of facilities, hundreds of operating points, and weekly message files running into six figures. The majority of private networks lease both equipment and facilities. However some

own their operating equipment and lease facilities, while others own both equipment and the plant that provides the facilities.

Most trunk circuits are routed via voice-frequency channels with either 150 or 170 cps (cycles per second) allocated to each channel. There are, however, a number of radio circuits and a few physical lines of moderate length. Local drops are usually metallic loops.

The standard method of operation is start-stop at either 60, 75, or 100 wpm (words per minute). The exceptions include some leased multiplex channels, an

occasional morse circuit left over from the past, telautographs or facsimile for certain types of local service, and a few wide-band installations for data transmission. The advantages of start-stop operation are obvious. Installations are easy to make, convenient methods of switching are available, maintenance and planning problems are relatively simple, and, when traffic loads are light, ordinary typists can take the place of professional telegraphists.

In areas where lines or facilities are cheap, plentiful, and efficient, start-stop is undoubtedly the preferred method of operation. On the other hand, where circuits are costly, and accuracy and

Paper 59-61, recommended by the AIEE Radio Communication Systems Committee and approved by the AIEE Technical Operations Department for presentation at the AIEE Winter General Meeting, New York, N. Y., February 1-8, 1959. Manuscript submitted October 14, 1958; made available for printing November 25, 1958.

B. S. DIAMOND is a telegraph consultant, Madison, N. J.