

AD No. 14812
ACT A

DEPARTMENT OF ELECTRICAL ENGINEERING

COLLEGE OF ENGINEERING

UNIVERSITY OF WASHINGTON

OPTIMUM ADJUSTMENT OF A
TORQUE-SATURATED SERVOMECHANISM

By

M. H. EKLUND, D. M. OLSON, and T. M. STOUT

TECHNICAL REPORT

PREPARED UNDER CONTRACT Nonr-477(02)

OPTIMUM DESIGN OF DELIBERATELY NONLINEAR SERVOMECHANISMS

(NR 374 371)

For

OFFICE OF NAVAL RESEARCH
EE DEPARTMENT REPORT NO. 12

January 12, 1953



OK
12-1-53
12-1-53
12-1-53

OPTIMUM ADJUSTMENT OF A TORQUE-SATURATED SERVOMECHANISM

by

M. H. EKLUND, D. M. OLSON, and T. M. STOUT

Technical Report
Prepared under Contract Nonr-477(02)
Optimum Design of Deliberately Nonlinear Servomechanisms
(NR 374 371)

for

Office of Naval Research

EE Department Report No. 12
January 12, 1953

Department of Electrical Engineering
University of Washington

ACKNOWLEDGMENT

The work described in this report was made possible by the support of the Office of Naval Research under Contract Nonr-477(02), Project NR 374 371, Optimum Design of Deliberately Nonlinear Servomechanisms.

ABSTRACT

The behavior of a torque-saturated servomechanism characterized by viscous friction and inertia has been studied. An idealized saturation curve is assumed and the step function response is computed piecewise; the response curves are given in dimensionless form, and a number of useful auxiliary curves are included.

The integrals of the absolute and squared error have been determined and plotted, and are employed in assessing the importance of saturation. It is found that there is an optimum damping ratio which decreases with the size of the step input, that the system gain in the unsaturated region can be increased to advantage, and that the performance of the saturating system is similar to that of an optimum off-on system for large inputs but definitely inferior for small inputs.

OPTIMUM ADJUSTMENT OF A TORQUE-SATURATED SERVOMECHANISM

In the study of simple positioning systems, it is customary to assume that the torque developed by the servomotor is directly proportional to the error between the input and output shafts. It is further assumed, tacitly or otherwise, that this proportionality applies for errors of any magnitude. In any physical system, however, the maximum torque is always limited.

The available torque may be limited by saturation of the amplifier controlling the motor, either accidentally as a result of inherent limitations of the amplifier or deliberately as a means of protecting the motor from excessive voltages. Magnetic saturation in the motor itself is another possible cause of a torque limitation.

Since torque saturation is probable in any physical system, two questions should be investigated: how does a saturated system behave (analysis), and how can saturation be either circumvented or used to advantage (synthesis)? The analysis of saturated servomechanisms has received some attention¹⁻⁵, but the important question of design has been largely neglected⁶. The following problems have therefore been investigated and are discussed in this report:

- (1) For a given size input, what is the loss in transient response due to saturation?
- (2) How can a saturating system be designed to take best advantage of the saturation?
- (3) How does an optimum saturated servomechanism compare with an optimum relay servomechanism?

Families of curves showing the transient response of a basic servomechanism for various input magnitudes and system parameters are presented, using a set of dimensionless quantities descriptive of the system and the input. Response criteria are applied to determine optimum parameter combinations, and the results are used in a comparative study of several methods for controlling a particular servomotor.

ANALYSIS

To facilitate the analysis, a particularly simple sort of torque-error relation is considered, shown in Fig. 1. Torque

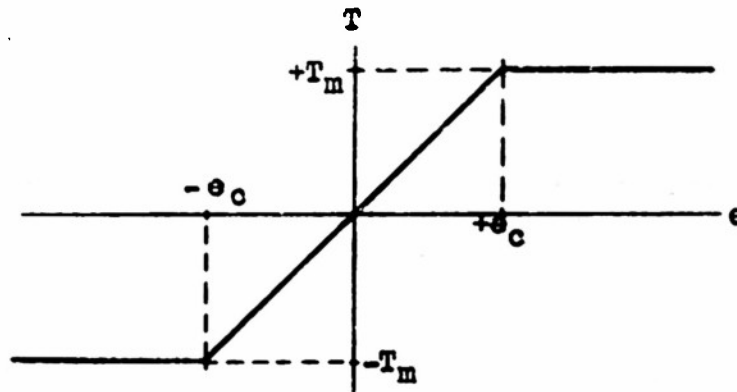


Figure 1. Torque-Error Relation

saturation occurs whenever the error exceeds a critical value, e_c , called the critical error. For errors greater than this critical error, the torque is constant ($\pm T_m$); for errors less than the critical error, the torque is proportional to the error ($K e$).

We also assume that the system is characterized by inertia and friction, and that there are no time lags ahead of the saturating element. A block diagram of the system under

consideration is given in Fig. 2. In this diagram and in the analytical work which follows, "s" may be regarded as the usual

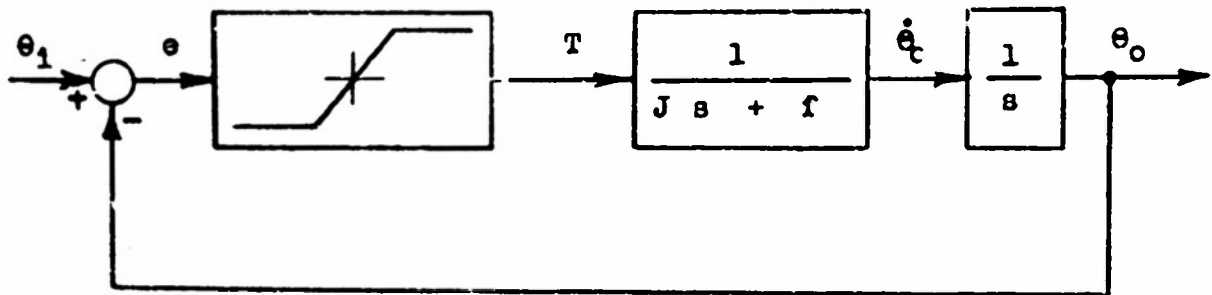


Figure 2. Block-Diagram of Torque-Saturated Servomechanism

Laplace transform variable or as an abbreviation for d/dt , the differential operator; θ_1 , θ_o , and e are the input, output, and error, respectively.

Only step inputs are considered and the system is assumed to be at rest when the step input is applied.

The differential equations describing the response of this servomechanism are

$$\begin{aligned} J \ddot{\theta}_o + f \dot{\theta}_o &= K e_c & e > e_c \\ &= K e & -e_c \leq e \leq e_c \\ &= -K e_c & e \leq -e_c \end{aligned} \quad (1)$$

where the dots denote differentiation with respect to time. If these equations are divided by f , new equations are obtained

$$\begin{aligned} \tau \ddot{\theta}_o + \dot{\theta}_o &= K_v e_c & e > e_c \\ &= K_v e & -e_c \leq e \leq e_c \\ &= -K_v e_c & e \leq -e_c \end{aligned} \quad (2)$$

in which $\tau = \frac{J}{f}$, the time constant, seconds;

$K_v = \frac{K}{f}$, the velocity constant, seconds⁻¹.

Using the definition of the error, $e = \theta_1 - \theta_c$, and the fact that a step input is being considered, Eq. (2) can be reduced to

$$\begin{aligned} &= -K_v e_c & e &\geq e_c \\ \tau \ddot{e} + \dot{e} &= -K_v e & -e_c &\leq e \leq e_c \\ &= K_v e_c & e &\leq -e_c \end{aligned} \quad (3)$$

Before solving these equations and discussing the solutions, it will be advantageous to list some additional parameters which can be used to describe the system:

$\Omega = \frac{T_m}{f} = K_v e_c$, the limiting velocity, radians/second;

$\omega_0 = \frac{K}{J} = \frac{K_v}{\tau}$, the undamped natural frequency, radians/second;

$\zeta = \frac{f}{2\sqrt{JK}} = \frac{1}{2\sqrt{K_v\tau}}$, the damping ratio, dimensionless;
 $\frac{e}{e_c}$, a measure of the error, dimensionless;

$P = \frac{e(0)}{e_c}$, a measure of the initial error, dimensionless;

$Q = \frac{\dot{e}}{\Omega}$, a measure of the error rate, dimensionless;

$\omega_0 t$, a measure of time, dimensionless.

We have chosen to use ζ , ω_0 , and e_c to describe a given system and have plotted e/e_c as a function of $\omega_0 t$ with ζ and P as parameters.

In the variables chosen, the equations become

$$\begin{aligned} &= - \frac{\omega_0}{2\zeta} e_c & e > e_c \\ \frac{1}{2\zeta\omega_0} \ddot{e} + \dot{e} &= - \frac{\omega_0}{2\zeta} e & -e_c \leq e \leq e_c \\ &= \frac{\omega_0}{2\zeta} e_c & e \leq -e_c \end{aligned} \quad (4)$$

General solutions for these equations are developed in the Appendix.

The character of the response to a step input is strongly influenced by the damping ratio, ζ , and the initial error, $P e_0$, and several different cases can be identified:

- (a) If P is less than unity, the system remains within the unsaturated or linear region at all times.
- (b) If P is greater than unity and ζ is greater than 0.287, the system is initially in the saturated region, reaches the linear region at a time t_1 , and then remains in the linear region.
- (c) If P is much greater than unity and ζ is less than 0.287, the system may go from the saturated region into the linear region and then return to the saturated region.

The three cases may be considered separately. Only positive values of P need to be treated, since the response for a negative P is identical except for sign with the response for the same positive P .

In general, the response is determined piece-wise, applying the differential equation appropriate to the region and taking advantage of the continuity of the error and error rate at the time of transition from one region to another.

- (a) If the initial error is less than the critical error ($P < 1$), the system response is given for all time by

$$\frac{e}{e_c} = P e^{-\zeta \omega_0 t} \left[\cos \sqrt{1 - \zeta^2} \omega_0 t + \frac{\zeta}{\sqrt{1 - \zeta^2}} \sin \sqrt{1 - \zeta^2} \omega_0 t \right] \quad (5)$$

which is obtained from Eq. (A6) of the Appendix by using the initial condition

$$q_0 = q(0) = 0 \quad (6)$$

Curves based on Eq. (5) are available⁷ and can easily be adjusted to account for different values of P .

- (b) If the initial error is greater than the critical error ($P > 1$), the initial portion of the system response is given by

$$\frac{e}{e_c} = P - \frac{\omega_0}{2\zeta} t + \frac{1}{4\zeta^2} (1 - e^{-2\zeta \omega_0 t}) \quad (7)$$

obtained from Eq. (B3) of the Appendix and the initial condition of Eq. (6) above. Curves based on this equation are presented in Fig. 3. Equation (7) applies until a time t_1 at which the error is equal to the critical error; this time can be found graphically from Fig. 3 by construction of a horizontal line at

$$\frac{e}{e_c} - P = \frac{1}{1 - P} \quad (8)$$

The dashed line shown in Fig. 3 is such a line for $P = 6$.

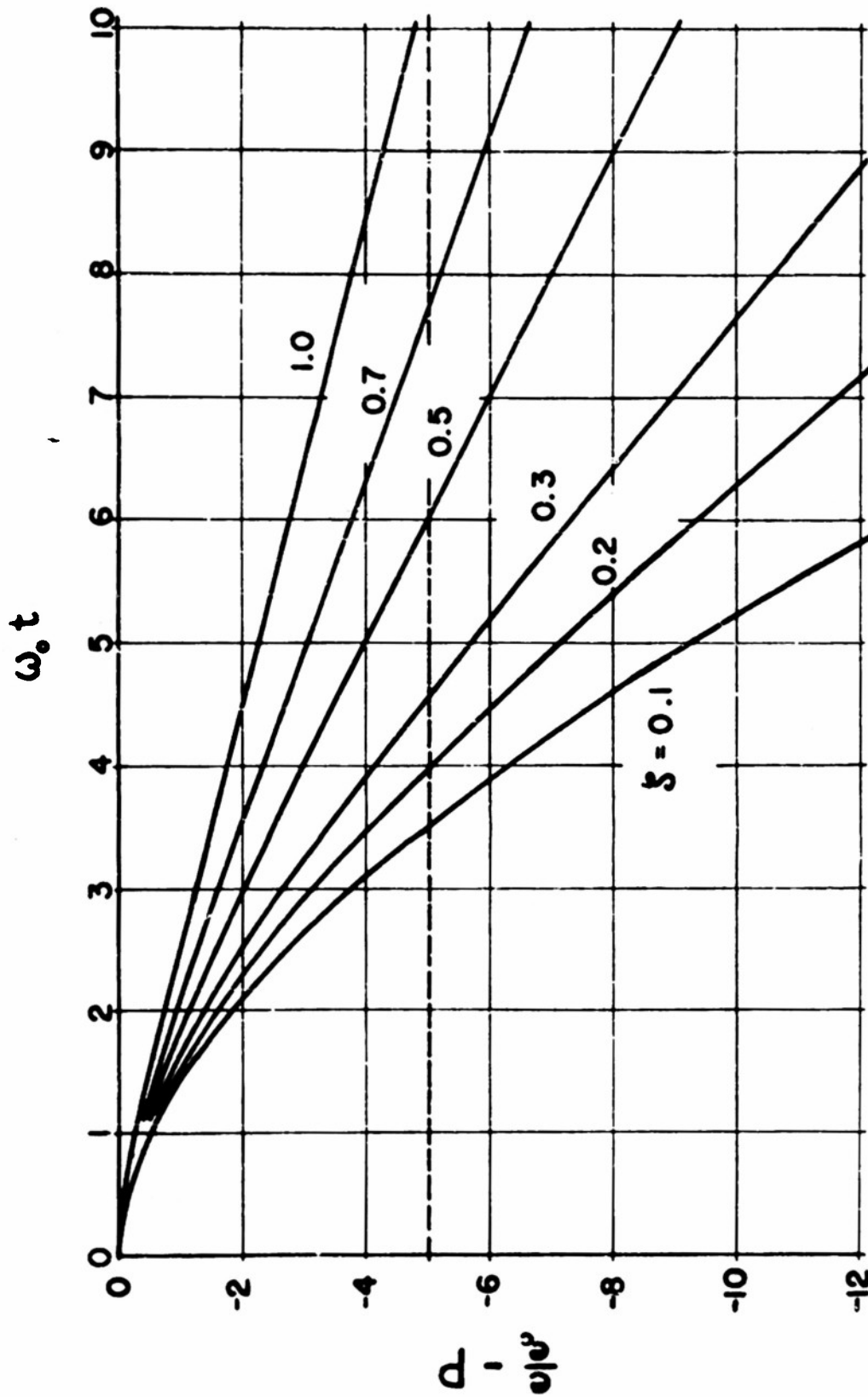


FIGURE 3
RESPONSE IN SATURATED REGION (ZERO INITIAL VELOCITY)

Knowing t_1 , the error rate when the system enters the linear region can be computed using Eq. (B4) with $Q_0 = 0$; results of such computations are summarized in Fig. 4, in which

$$Q_1 = \frac{2\zeta}{\omega_0 e_0} \dot{e}(t_1) \quad (9)$$

the ratio of the actual error rate to the maximum error rate.

Within the linear region, the response will be given by

$$\begin{aligned} \frac{e}{e_c} = & e^{-\zeta\omega_0 x} \left[\cos\sqrt{1-\zeta^2}\omega_0 x + \frac{\zeta}{\sqrt{1-\zeta^2}} \sin\sqrt{1-\zeta^2}\omega_0 x \right] \\ & + Q_1 \frac{1}{2\zeta\sqrt{1-\zeta^2}} e^{-\zeta\omega_0 x} \sin\sqrt{1-\zeta^2}\omega_0 x \end{aligned} \quad (10)$$

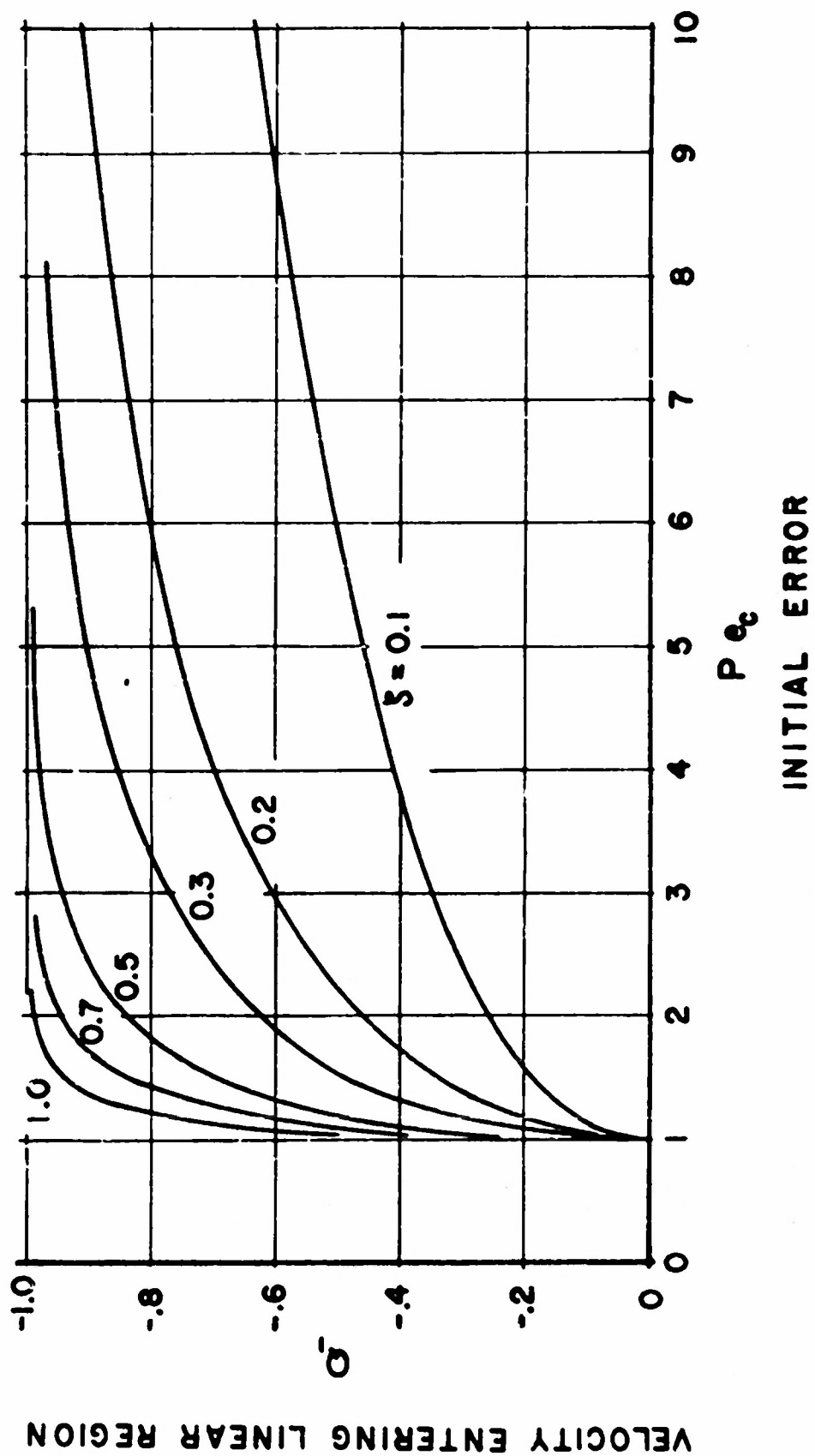
where $x = t - t_1$ and Q_1 may be obtained from Fig. 4. It may be noted this response can be plotted from the previous Eq. (5) by setting $P = 1$ and adding a term to account for the initial error rate.

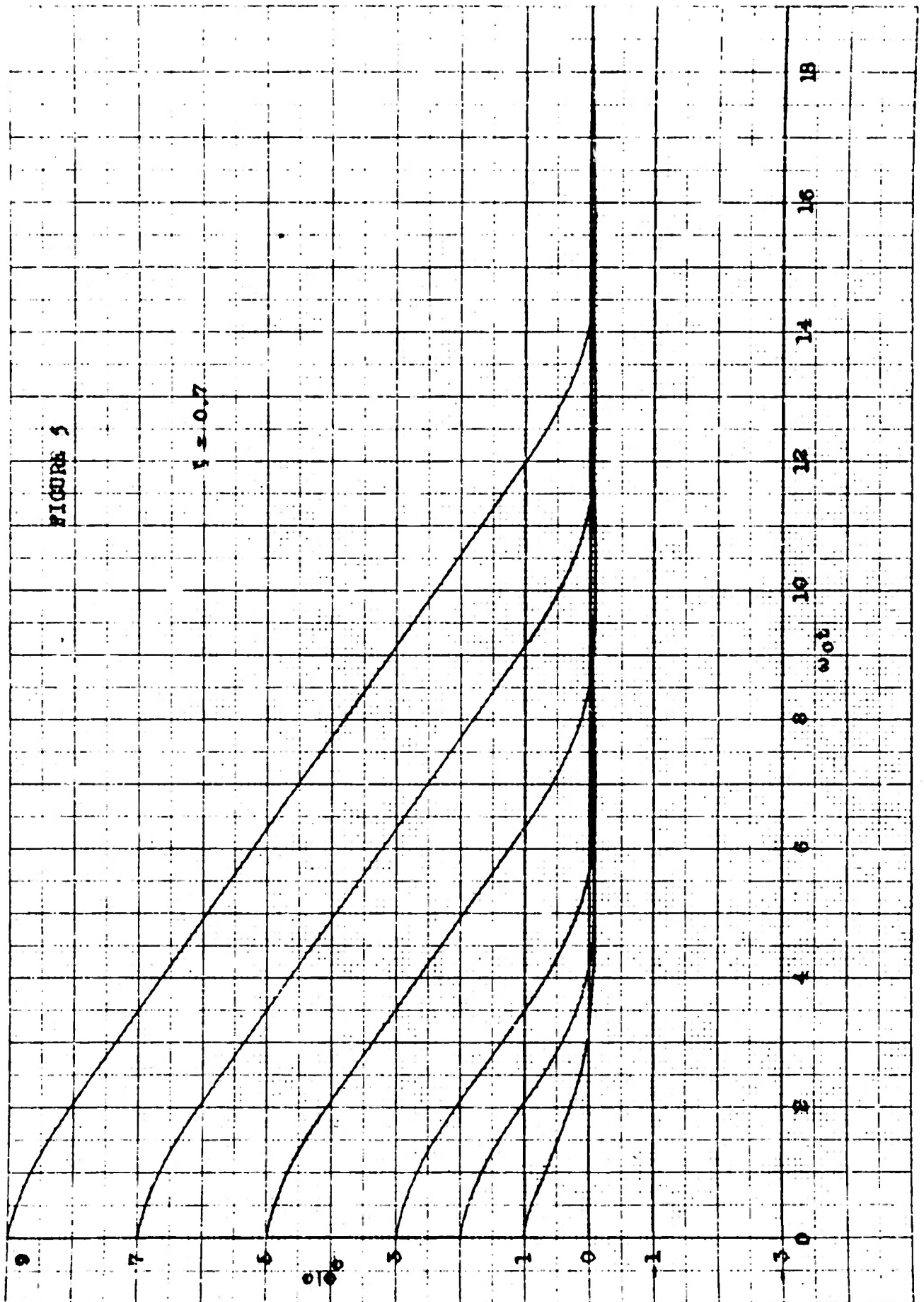
It has already been shown^{1,3} that, if $K_v\tau < 3.03$, the system will remain within the linear region after the initial period in the saturated region. In the notation of this report, the requirement is

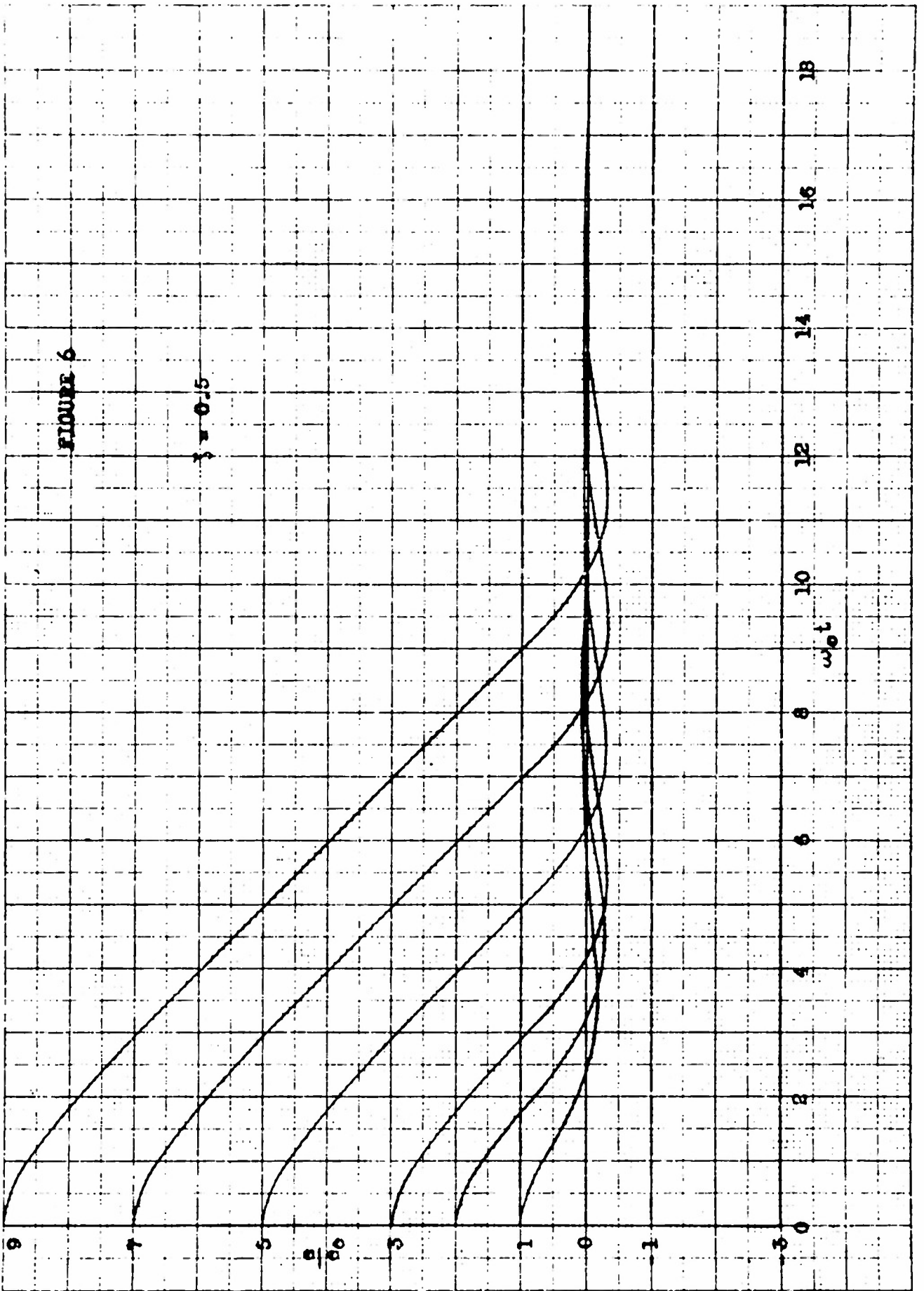
$$\zeta > \frac{1}{2\sqrt{3.03}} = 0.287 \quad (11)$$

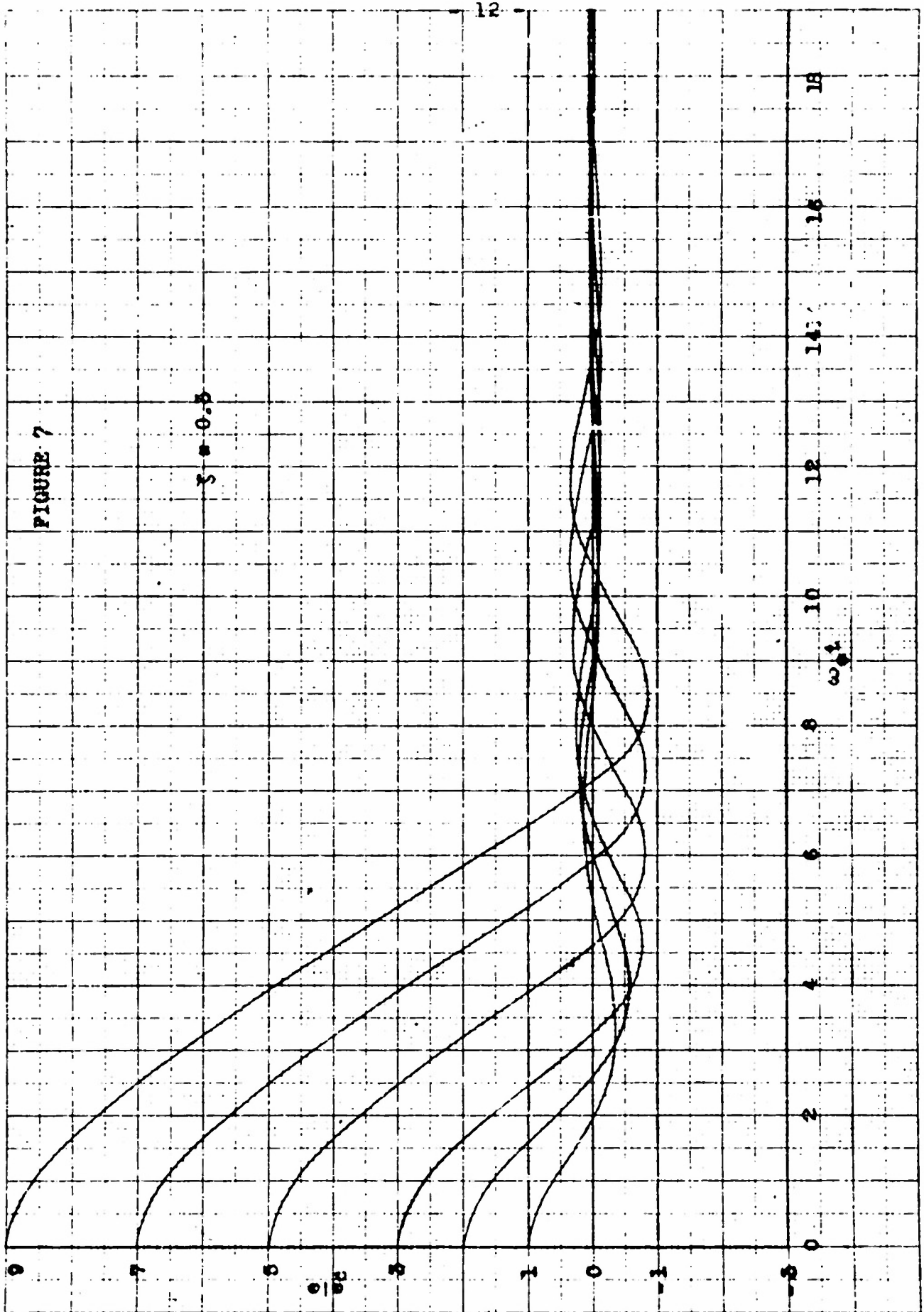
from the relations on p. 4. Thus for $\zeta = 0.3, 0.5, \text{ or } 0.7$, the system does not re-enter the saturated region and the response curves can be calculated from Eqs. (7) and (10). Families of response curves for these damping ratios and several values of P are given in Figs. 5-7.

FIGURE 4









(c) If P is large enough and ξ is less than 0.287, say 0.2 or 0.1, the system will cross the linear region and continue into the saturated region, past $e = -e_c$. The time at which the system re-enters the saturated region, t_2 , will be a function of Q_1 and ξ which is best determined graphically. A plot of $\omega_0(t_2 - t_1)$, the time required to cross the linear region, is given in Fig. 8.

Knowing the time spent in the linear region, the velocity at time t_2 can be calculated from Eq. (A7) using $P = 1$ and $Q_0 = Q_1$; the results may be expressed in terms of

$$Q_2 = \frac{2\xi}{\omega_0 e_c} \dot{e}(t_2) \quad (12)$$

and plotted as functions of Q_1 and ξ , as shown in Fig. 9. This figure shows, as does Fig. 8, that there is no re-entry into the saturated region for $\xi = 0.2$ if Q_1 is less than 0.55, or if Q_1 is less than 0.18 for $\xi = 0.1$. (From Fig. 4, the corresponding values of P are 2.55 and 1.4, respectively.) In such cases, as in the cases for $\xi > 0.287$, the response is calculated in two parts from Eqs. (7) and (10).

If the system re-enters the saturated region, the next part of the response is calculated from Eq. (B5) which becomes

$$\frac{e}{e_c} = -1 + \frac{\omega_0}{2\xi} y + \frac{1}{4\xi^2} (1 - e^{-2\xi\omega_0 y}) + \frac{Q_2}{4\xi^2} (1 - e^{-2\xi\omega_0 y}) \quad (13)$$

where $y = t - t_2$. The relative velocity Q_2 can be obtained from Fig. 9.

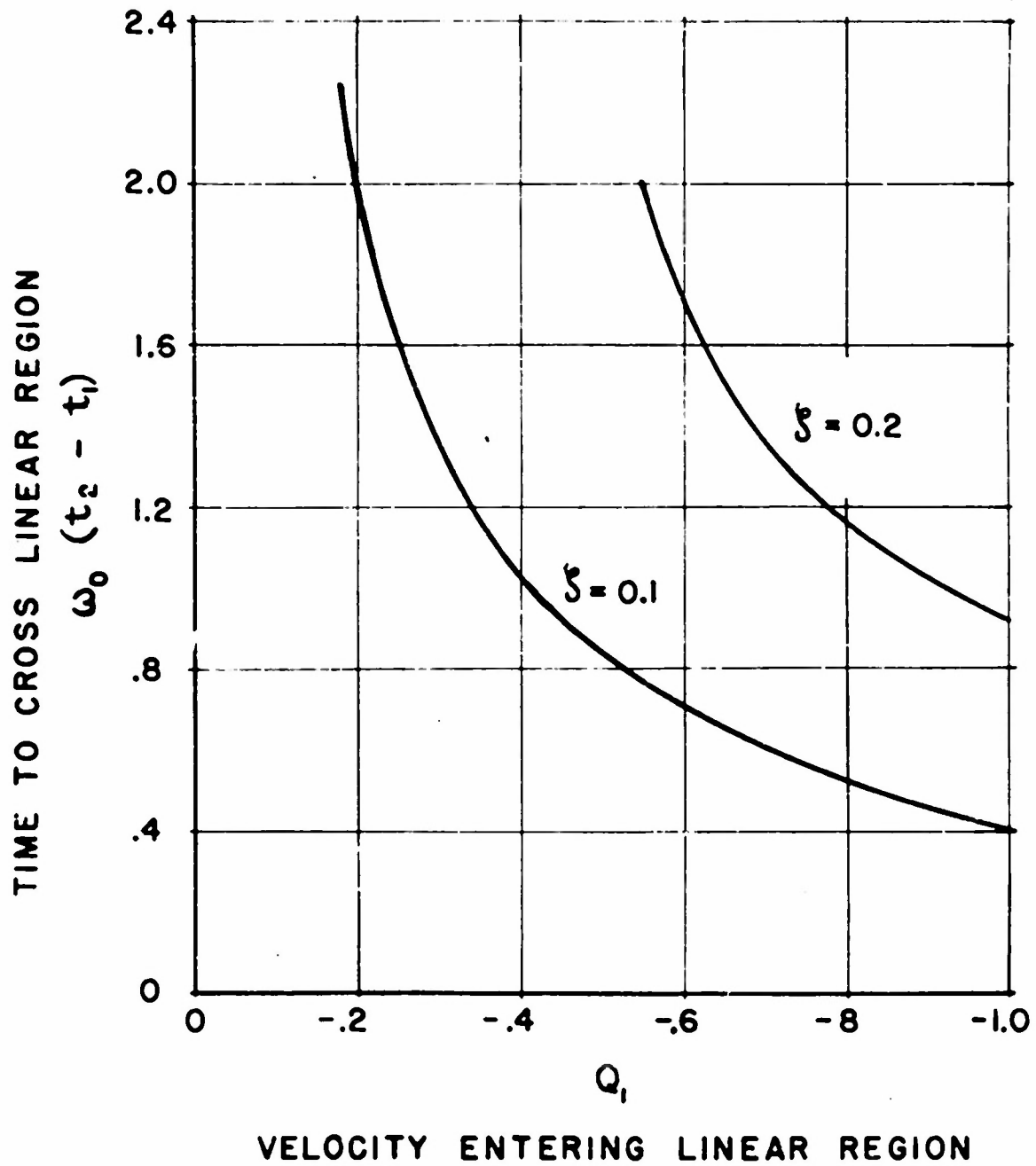


FIGURE 8

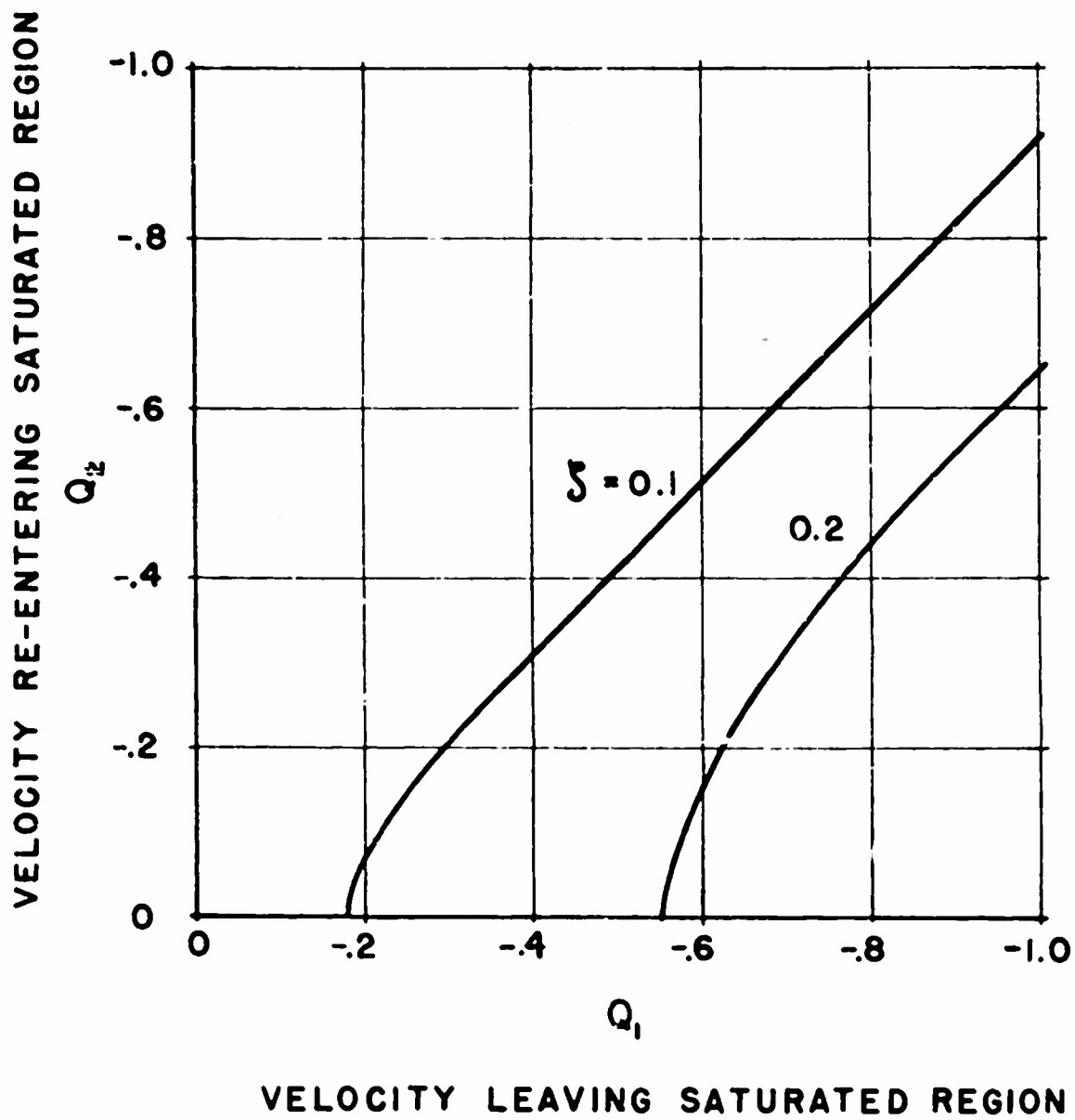


FIGURE 9

Response curves computed from Eq. (13), like those shown in Figs. 10 and 11, can be used to determine the time interval, $(t_3 - t_2)$, spent in the saturated region.

The error rate at the time of re-entry into the linear region, specified by

$$Q_3 = \frac{2\zeta}{\omega_0 \theta_0} \dot{\theta}(t_3) , \quad (14)$$

can then be computed and plotted as a function of ζ and either Q_2 , Q_1 , or P . The most useful combination is probably Q_3 as a function of Q_1 , shown in Fig. 12.

With the curves already presented, sufficient information is available to permit efficient calculation of the response for any combination of initial error and system parameters. The procedure may be understood by reference to Fig. 13.

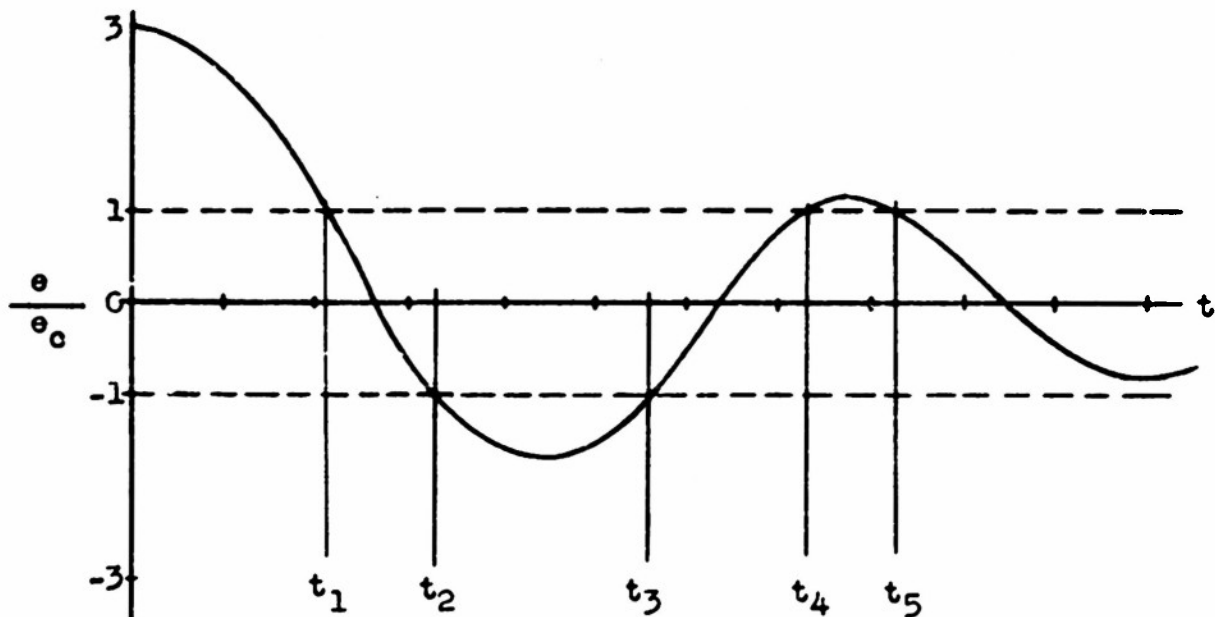


Figure 13. Typical Torque-Saturated Servomechanism Response

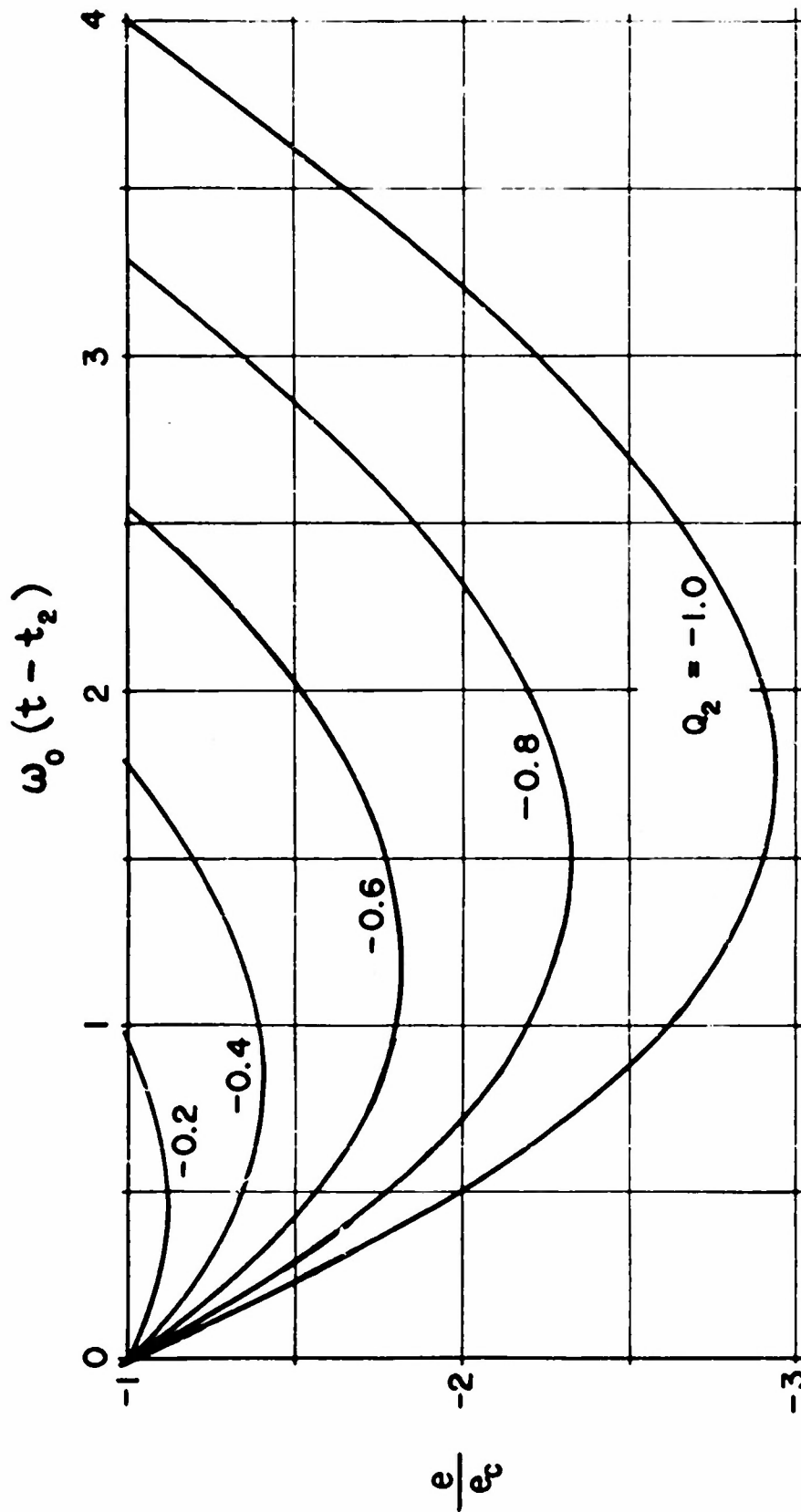


FIGURE 10
RESPONSE IN SECOND SATURATED REGION
($\xi = 0.2$)

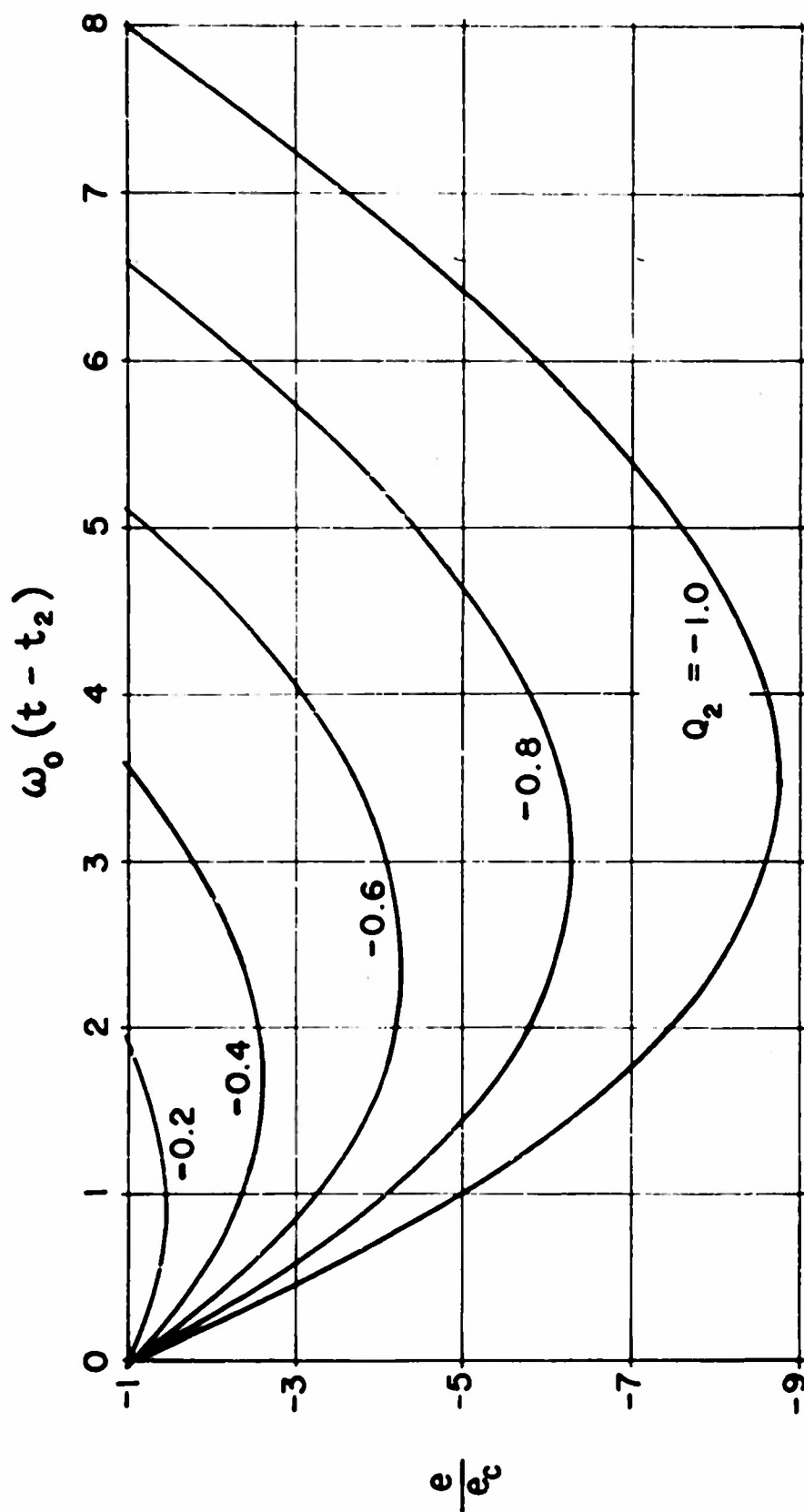
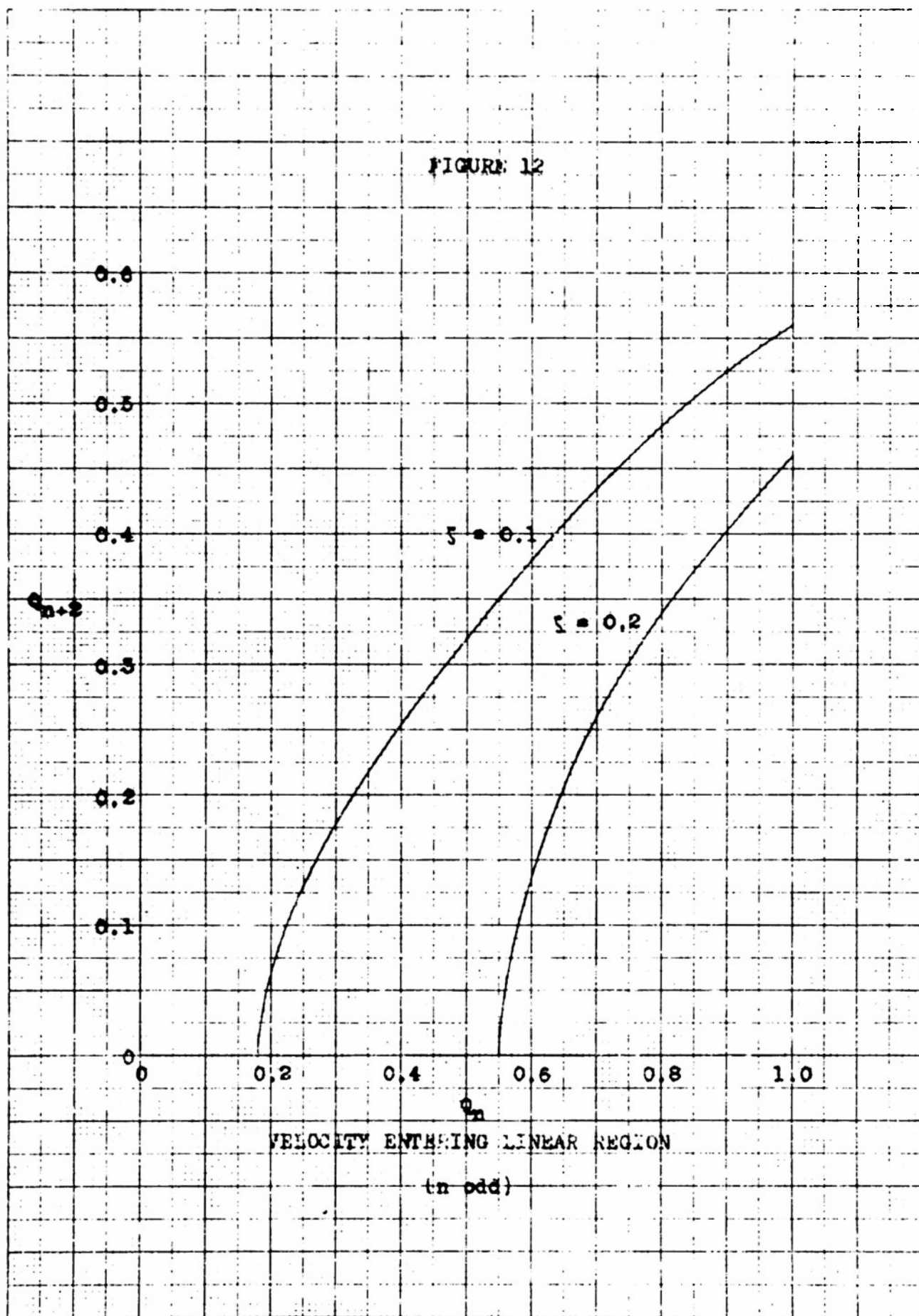


FIGURE 11
RESPONSE IN SECOND SATURATED REGION
($\xi = 0.1$)

FIGURE 12

VELOCITY RE-ENTERING LINEAR REGION



VELOCITY ENTERING LINEAR REGION

(in odd)

The calculations previously outlined give the response for $0 < t < t_3$. If Q_3 is small enough, the system remains in the linear region for $t > t_3$ and Eq. (10) can be used to compute the remainder of the response by substituting Q_3 for Q_1 and reversing the signs of both terms.

If there are additional re-entries into the saturated regions, as shown in Fig. 13, the response for $t_3 < t < t_4$ is computed in the same way as the response for $t_1 < t < t_2$, using Eq. (10) with appropriate sign changes. Similarly, the response for $t_4 < t < t_5$ can be computed from Eq. 13 by changing the sign of all terms. The error rates at t_4 and t_5 , specified in terms of Q_4 and Q_5 , can be found from Q_3 by using Figs. 9 and 12. The same procedure can be employed as many times as necessary, until the system leaves the saturated regions for the last time, when Eq. (10) can be used to compute the remainder of the response.

Figure 12 can be used to determine the number of times that a given system will enter the saturated regions in response to a specified input. Suppose, for example, that $\zeta = 0.1$ and $F = 5$. From Fig. 4, we find $Q_1 = -0.47$. From Fig. 12, the error rates at subsequent entries into the linear region are $Q_3 = 0.30$, and $Q_5 = -0.175$; Q_7 does not exist. This system will therefore re-enter the saturated region twice after first leaving the linear region.

Figure 12 can also be used to determine the maximum number of times that the system re-enters the saturated regions for a given ζ by assuming that F is large enough to make $Q_1 = -1$.

Successive applications of Fig. 12 show that the maximum possible number of re-entries for ζ equal to 0.2 and 0.1 are, respectively, one and four.

Response curves for ζ equal to 0.2 and 0.1 are given in Figs. 14 and 15 for several values of P . These curves, with those of Figs. 5-7, have been used to determine the effects of varying the parameters of the system, to check the results of experimental work with an analog computer, and to find optimum parameter values.

Since the torque-error curve employed in these calculations might be considered somewhat unrealistic, due to the abrupt transition from linearity to saturation, some study was made of a system in which the torque is

$$T = K e_0 \left[1 - e^{-\left(\frac{e}{e_0}\right)} \right] \quad (e > 0) \quad (15)$$

and

$$T(-e) = - T(e) \quad (16)$$

This expression for the torque reduces to $K e$ for small errors and approaches $K e_0$ for large errors, like the idealized characteristic used in the previous calculations. The response curves for this system, computed by a step-by-step method⁸, are similar to those presented in this report and it was felt that no major errors would result from use of the simpler relation.

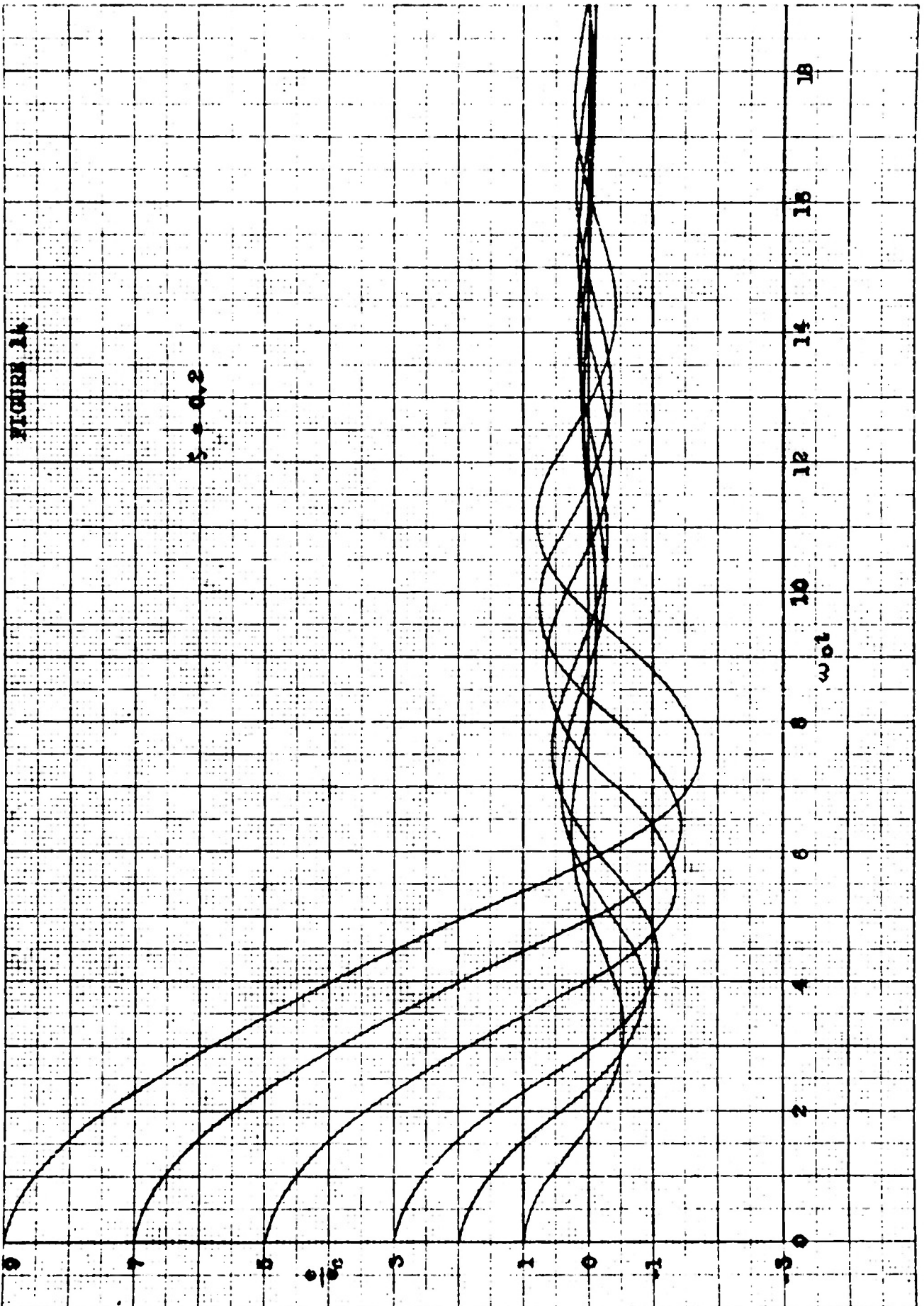
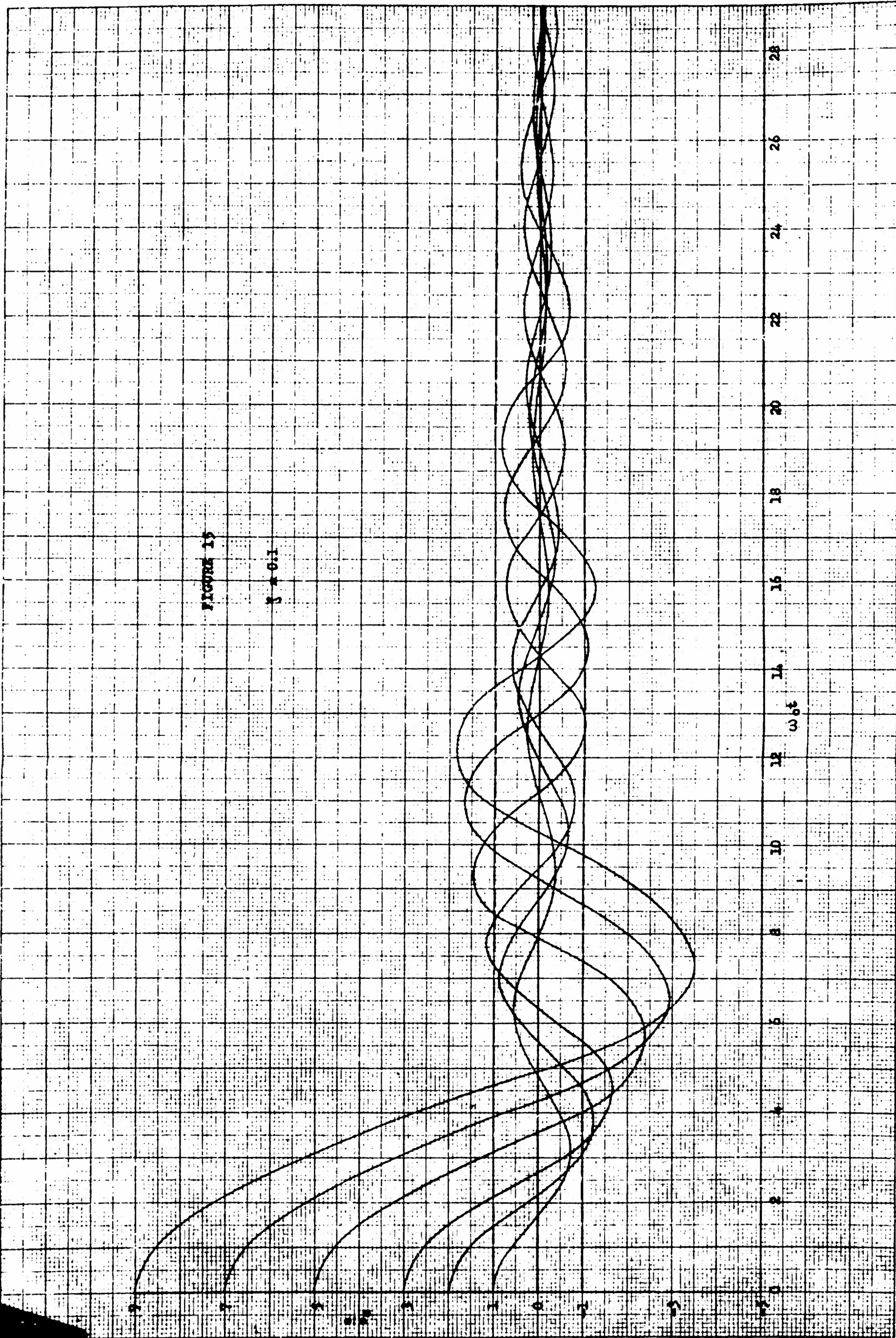


FIGURE 13

$\gamma = 0.1$



RESPONSE CRITERIA

In order to determine whether saturation is helpful or harmful under any specified conditions, it is necessary to use some criteria by which the "goodness" of the response can be measured.

The time required for the error to reach any specified value appears to be an unsatisfactory measure of the response for systems of the type considered here. Using the time needed to reduce the error to, say, one-tenth of the original value⁴ seems undesirable, since such a criterion places too much emphasis on a rapid reduction of the large initial error and too little emphasis on adequate damping. A more satisfactory criterion, the time required for the error to become small or negligible, would be somewhat difficult to apply and would require agreement on the meaning of "small" and "negligible".

Better measures of the response are provided by the integrated absolute error

$$\Lambda_1 = \int_0^{\infty} |e| \, dt$$

and the integrated squared error

$$\Lambda_2 = \int_0^{\infty} e^2 \, dt$$

which, in this report, are determined for a step change in the system input. An optimum system is defined as one which minimizes the integral employed.

These integrals provide objective measures of response, rigidly defined and offering no opportunity for individual interpretation, but at the same time leading to optimum system

adjustments in satisfactory agreement with those obtained by intuitive methods. The integrated absolute error criterion, as previously noted⁹, favors systems with slightly more damping than the integrated squared error criterion and is somewhat easier to use.

For initial errors which do not produce saturation, the system remains at all times within the linear region and the integrated absolute error can be computed from*

$$A_1 = \frac{2e_0}{\omega_0} \left[\zeta + \frac{e^{-\sigma \theta_1}}{1 - e^{-\sigma \pi}} \right] \quad e_0 < e_c \quad (17)$$

where $\sigma = \frac{\zeta}{\sqrt{1 - \zeta^2}}$

$$\theta_1 = \tan^{-1}(1/\sigma) \quad (\pi/2 < \theta_1 < \pi)$$

The integrated squared error is¹¹

$$A_2 = \frac{e_0^2}{\omega_0} \left[\frac{1 + 4\zeta^2}{4\zeta} \right] \quad e_0 < e_c \quad (18)$$

For initial errors greater than the critical error, which do produce saturation, the integrals can be evaluated either by numerical integration using the calculated response curves or by direct measurement using an analog computer. The values of A_1 plotted in Fig. 15 were obtained from the calculated responses curves using a planimeter, while the values of A_2 in Fig. 17 were obtained experimentally.** Both sets of curves have been

*See Appendix II and Reference 10.

**A detailed description of the experimental procedure is given in Reference 12.

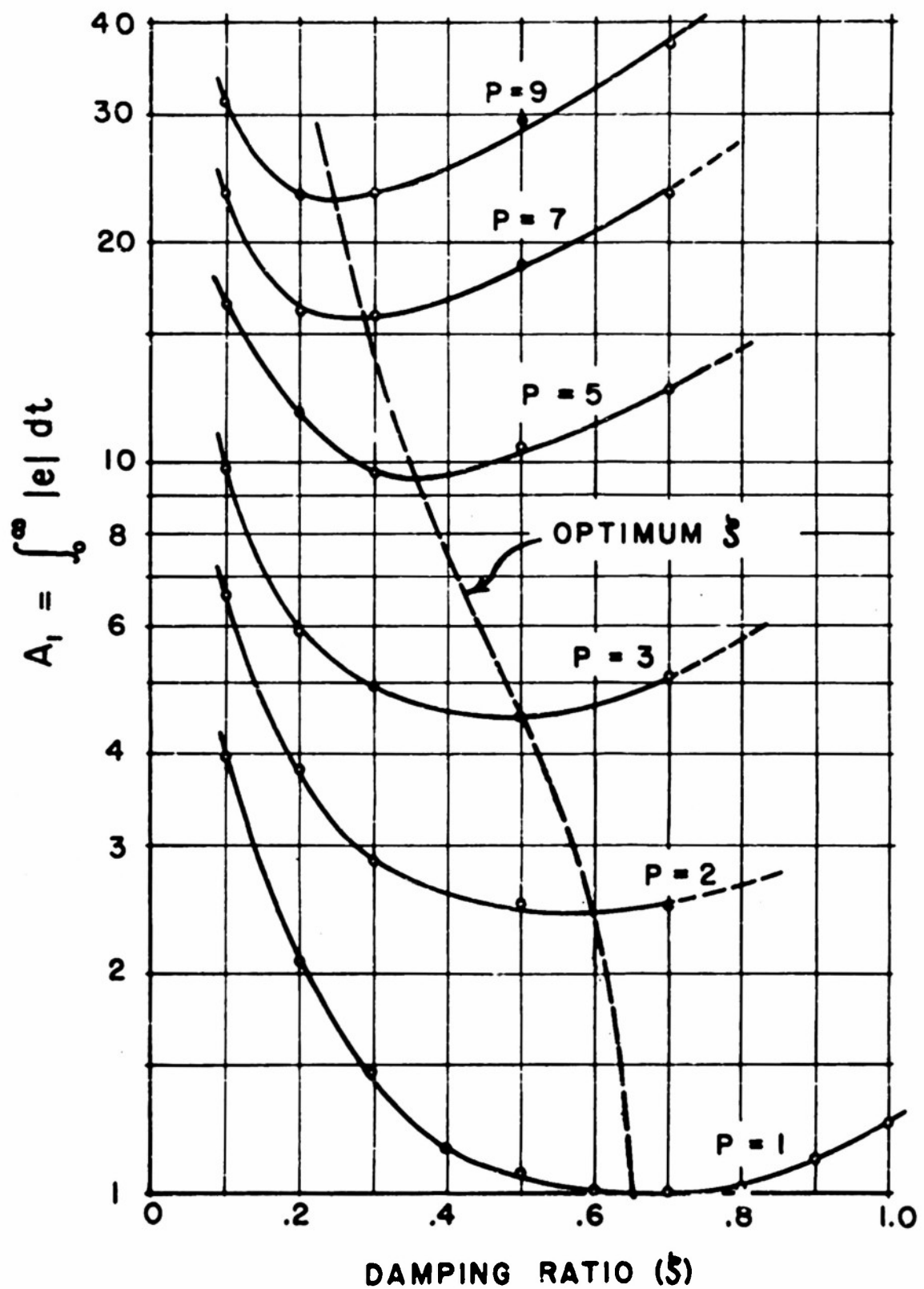


FIGURE 16

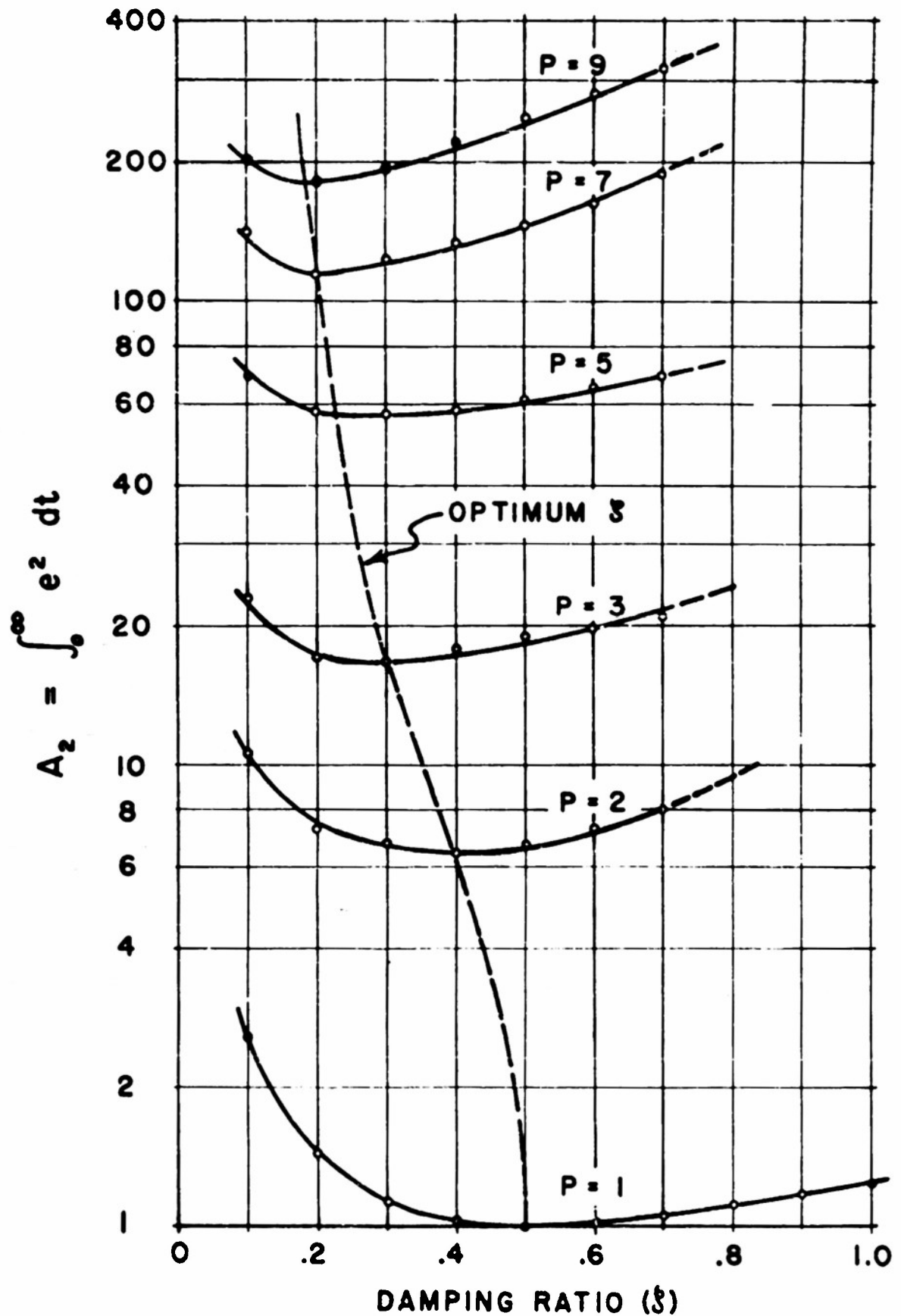


FIGURE 17

plotted to show the integrated absolute or squared error with respect to the values for $e_0 = e_c$ or $P = 1$; in particular, the vertical scales have been adjusted to produce a minimum value of unity for $P = 1$. Since the true values can be calculated for this case from Eqs. (17) and (18), the actual values for any other case can be determined from the graphs by a simple proportion. To facilitate such comparisons, computed values of A_1 and A_2 are given in Table I.

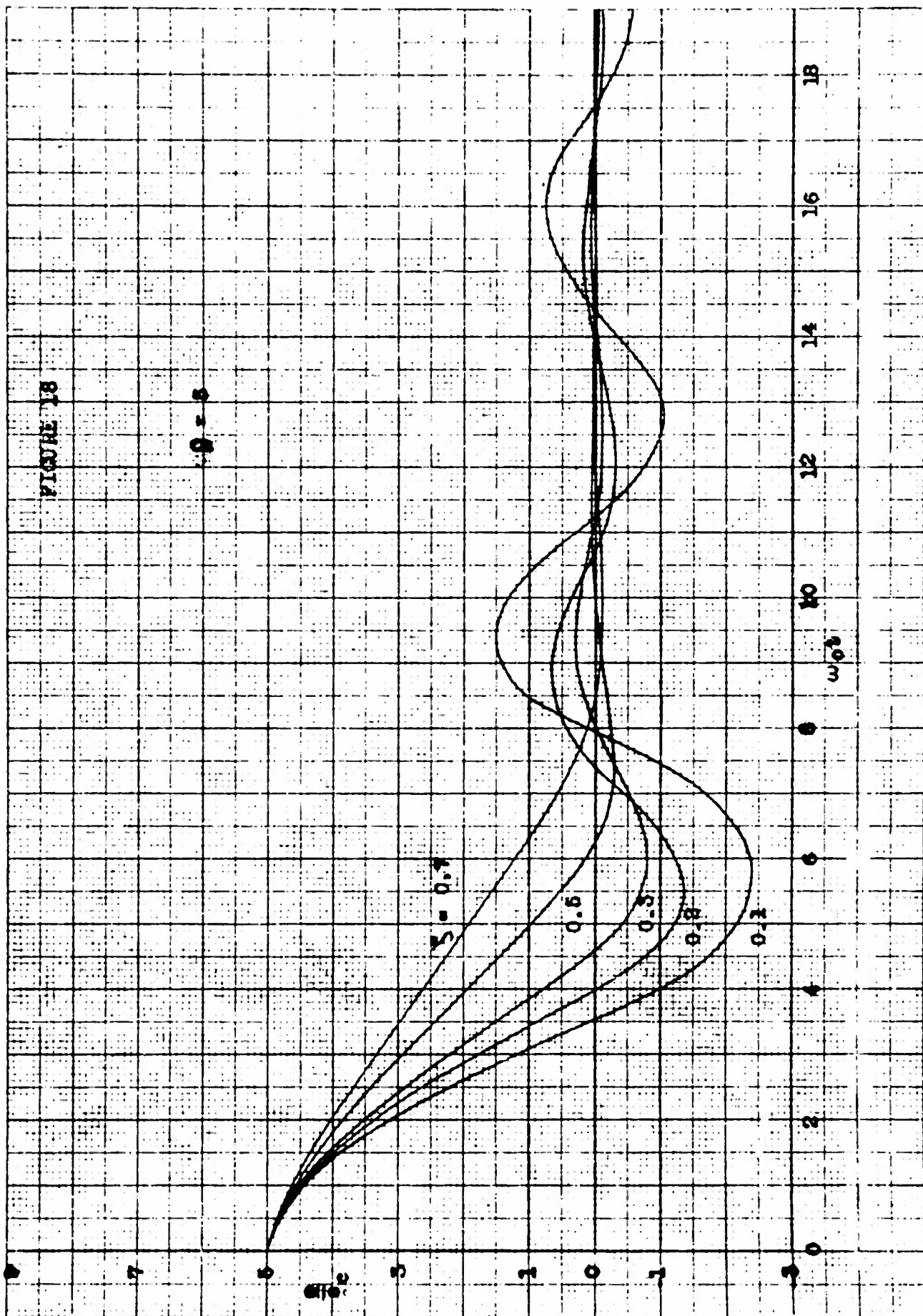
Damping Ratio	$A_1 \omega_0$	$A_2 \omega_0$
0.1	6.384	2.600
0.2	3.341	1.450
0.3	2.364	1.133
0.4	1.924	1.025
0.5	1.712	1.000
0.6	1.618	1.017
0.65	1.606	
0.7	1.608	1.057
0.8	1.631	1.113
0.9	1.804	1.178
1.0	2.000	1.250

Table I. Error Integrals for $e_0 = e_c$ ($P = 1$)

For any initial error and using either criterion, there is an optimum damping ratio which represents a compromise between a rapid reduction of the error in the saturated region and adequate damping after the system is within the linear region.

The optimum damping ratio is constant for linear operation of the system ($P \leq 1$) and decreases with P when saturation occurs. From Fig. 16, the optimum damping ratio determined using the integrated absolute error criterion is seen to be about 0.65 for $P \leq 1$, decreasing to about 0.25 for $P = 9$. Using Fig. 17 and the integrated squared error criterion, the corresponding values of damping are 0.50 and approximately 0.20.

The correlation between the optimum values of damping ratio determined from Figs. 16 and 17 and those obtained by empirical methods can be studied with the help of Fig. 18, which shows the response curves for several different damping ratios and an initial error which is five times the critical error ($\bar{P} = 5$). According to Figs. 16 and 17, the recommended damping ratio is either 0.35 or 0.25, depending on the criterion employed. An intuitive judgment would probably result in a damping ratio between 0.3 and 0.5.



APPLICATIONS

The data contained in Figs. 16 and 17 can be applied in a variety of ways to study the advantages or disadvantages of saturation.

The first question considered in the loss in performance as a result of saturation, measured by the increase in the integrated absolute or squared error. The standard for this comparison is a hypothetical linear system which is identical with the saturating system for small errors and which has unlimited torque. The torque-error characteristics of these systems are given in Fig. 19.

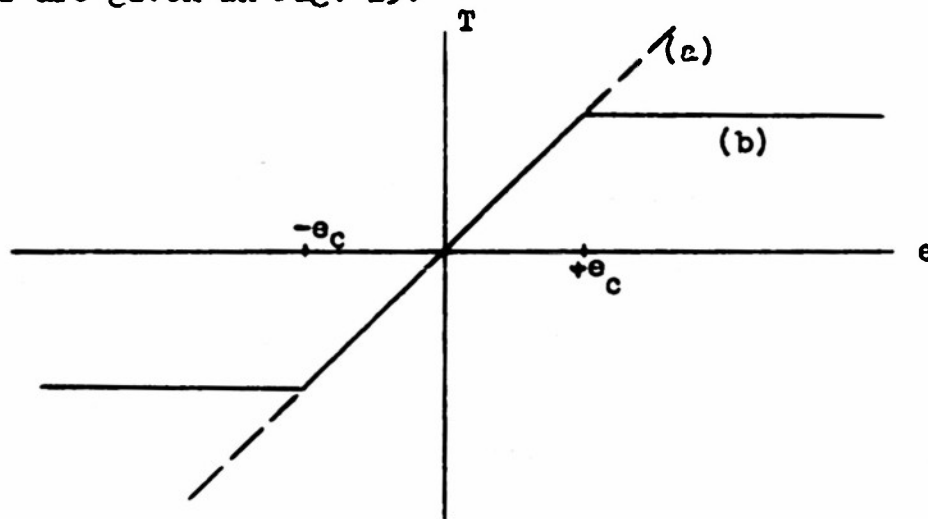


Figure 19. Torque-Error Characteristics of (a) Linear System and (b) Saturating System

The comparison is facilitated by plotting, say, the values of A_1 as a function of e_0/e_c , as shown in Fig. 20. In the linear region, the curves are straight lines with unit slope, as may be seen from Eq. (17). For initial errors larger than the critical error, the integrated absolute error for the saturating system increases rapidly and finally attains a slope of two, which is

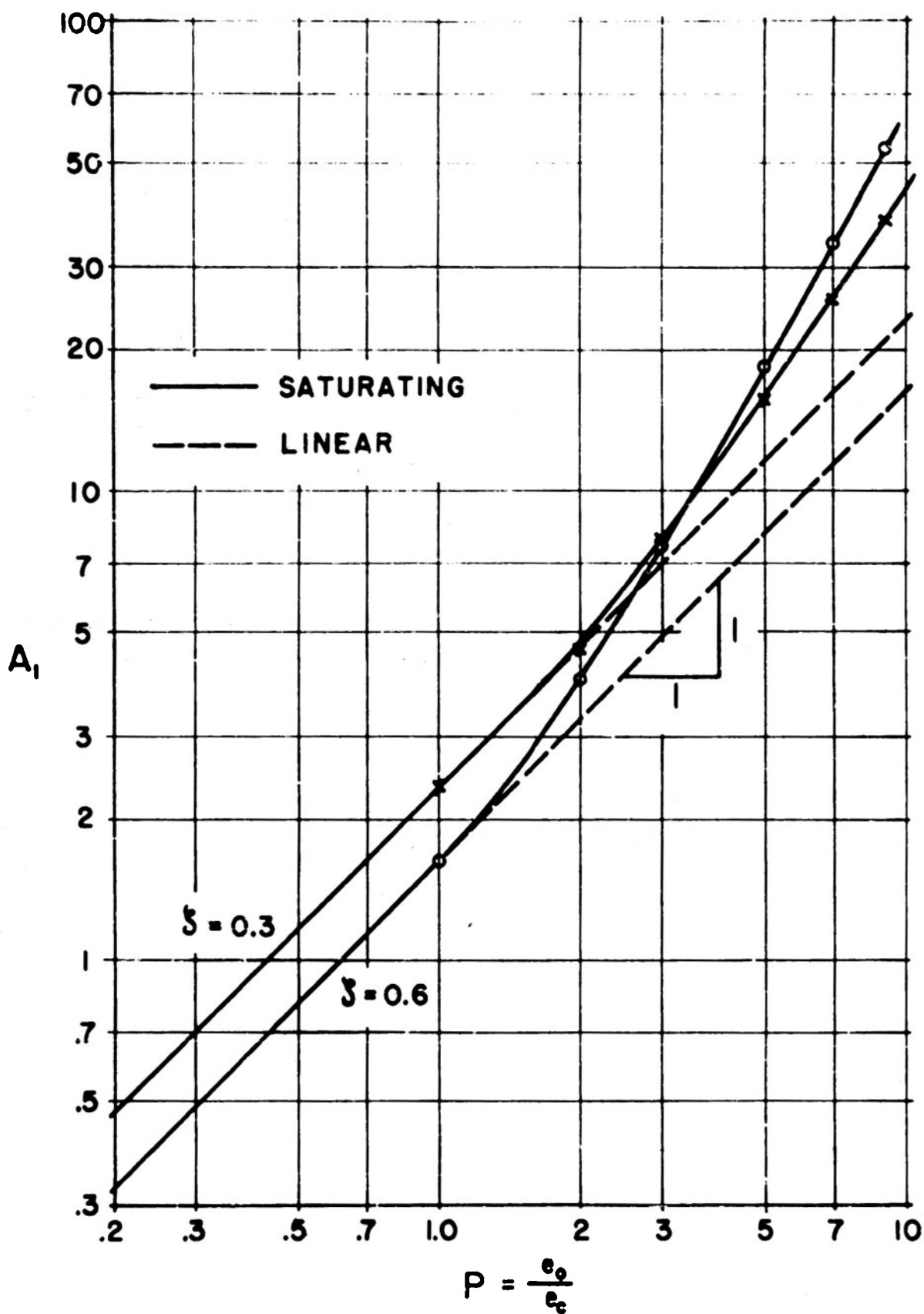


FIGURE 20

characteristic of rate-limited systems¹³. For $\zeta = 0.6$ and $P = 9$, where the integrated absolute error would be expected to be $(9)(1.618) = 14.56$ units for a linear system, the value of A_1 for the saturating system is about 53 units, or an increase of approximately 3.6 times. For $\zeta = 0.3$, the corresponding increase is from $(9)(2.364) = 21.3$ to 38, or about 1.8 times. The poorer response of the saturating system can be traced, of course, to the limitation on the torque available for correcting errors.

The information in Figs. 16 and 17, plotted in the form used in Fig. 20, can also be employed for a determination of the best value of K , the slope of the torque-error characteristic in the linear region. As an example, we might consider a system for which $K = 1$, $\omega_0 = 1$, $\zeta = 0.6$, and $e_c = 1$, shown by the solid curve in Fig. 21. Since the maximum available torque is fixed, changing K merely changes e_c , ω_0 , and ζ , while leaving the limiting velocity and time constant unchanged.

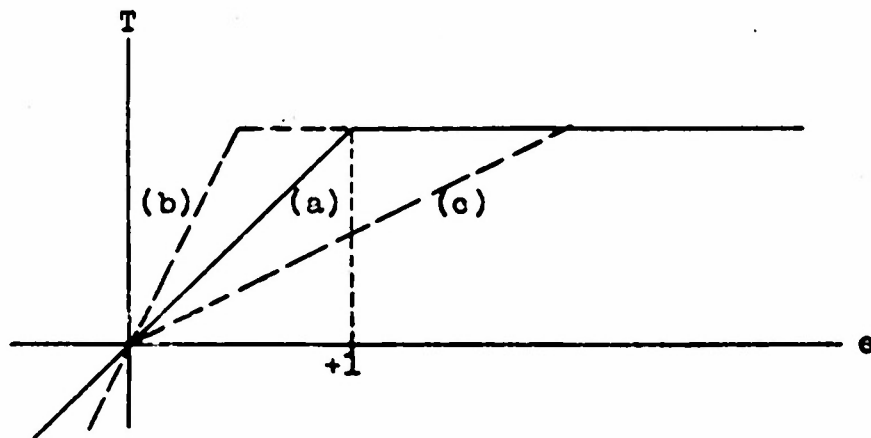


Figure 21. Torque-Error Characteristics of Saturating System for Three Values of K : (a) $K = 1$, (b) $K = 2$, and (c) $K = 0.5$

The important system parameters are given in Table II for various values of K .

K	0.25	0.50	1	2	4	9
ω_0	0.50	0.71	1.00	1.41	2.00	3.00
ξ	1.20	0.85	0.60	0.425	0.30	0.20
e_0	4.00	2.00	1.00	0.50	0.25	0.11

Table II

In plotting the curves of integrated absolute error as a function of initial error, we proceed as follows. The curve for $K = 1$ may be traced directly from the curve for $\xi = 0.6$ in Fig. 20. To obtain the curve for, say, $K = 4$, we may first plot a straight line through the point $e_0 = 1$, $A_1 = 1.18$ (computed from Eq. (17) or simply take $A_1 = 2.364$ for $\xi = 0.3$ from Table I and divide by 2 to account for the change in ω_0) with unit slope; this line describes the variation of A_1 with e_0 for initial errors less than 0.25. The curve for initial errors greater than 0.25 is obtained by tracing the curve for $\xi = 0.3$ from Fig. 20, after shifting the curve to the left and down so that the point at $F = 1$ coincides with the new value of A_1 at $e_0 = 0.25$. These curves, and a number of similar curves, are shown in Fig. 22.

Inspection of Fig. 22 reveals two interesting properties of this system. First, for errors which are too small to produce saturation, the integrated absolute error is a continuously decreasing function of gain, as expected⁹. Second, for errors much larger than the critical error, the integrated absolute error appears to be independent of K . This result may best be understood by study of a family of response curves for various

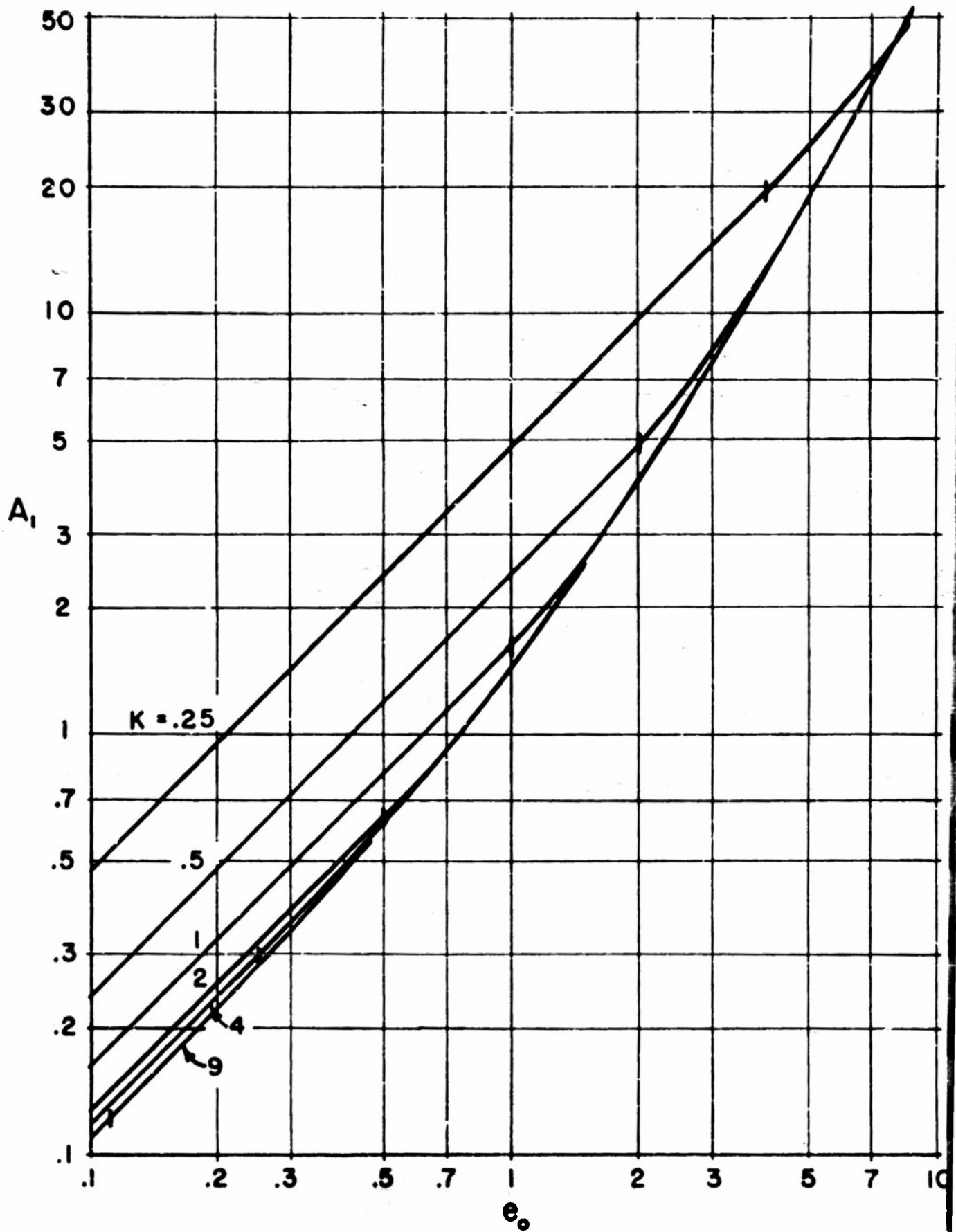


FIGURE 22

values of K , shown in Fig. 23. In drawing these curves, it was assumed that the initial error was just sufficient to saturate the system for $K = 1$, so that the response is entirely within the linear region. As K is increased, the critical error becomes smaller and the time within the saturated region increases. For the larger values of K , the initial portion of the response is practically identical for the various curves. In the limit, as K approaches infinity, the system would effectively become a viscous-damped relay system whose integrated absolute error is a definite, non-zero quantity. This behavior is in distinct contrast with that of a linear system, whose integrated absolute error approaches zero as K approaches infinity.

Evidently, then, it is advantageous to make K as large as possible. In a physical system, saturation can be produced for small errors by the insertion of a high-gain amplifier between the error-sensing device and the saturating element, which may be the amplifier driving the servo motor or the motor itself. This type of control provides some of the advantages of relay control, while permitting an adequately damped linear response for small errors.

For a motor characterized by inertia and friction, the best method of control is an adaptation¹³ of the off-on system proposed by McDonald¹⁴. In this system, full accelerating torque is applied to the output shaft until the error is reduced to a specified value (dependent on the initial error); decelerating torque is then applied to bring the error and error rate to zero simultaneously, when the torque is removed. Switching is

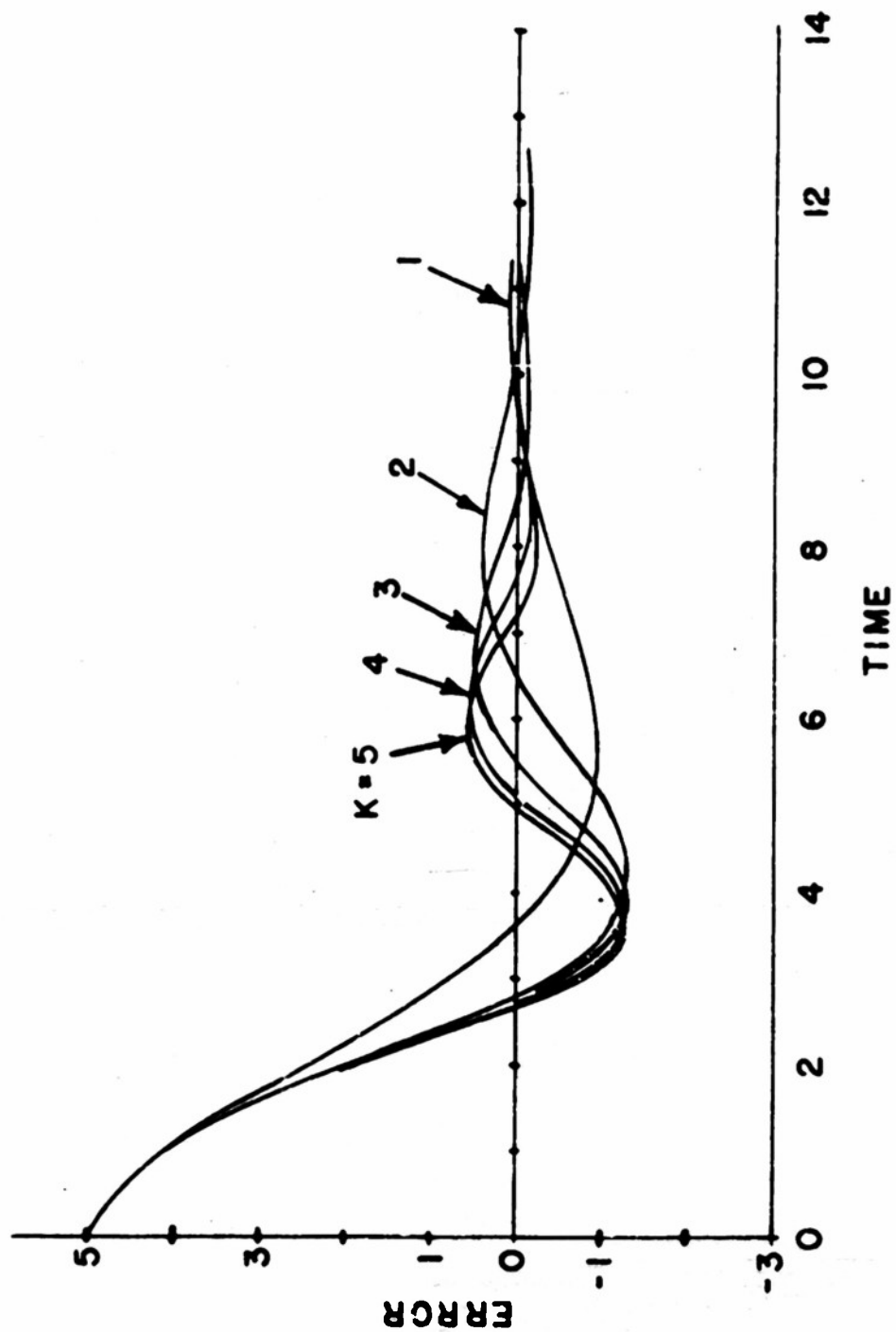


FIGURE 23

initiated by a nonlinear lead network, which compares the error with a nonlinear function of the error rate to determine whether positive or negative torque is required. It will be instructive to compare the saturating system with this optimum relay system.

The constants of a typical two-phase 60-cycle servo motor will be used in this comparison, in order to illustrate how the data contained in the general curves can be applied to a specific case. The manufacturer's catalog supplies, in addition to a speed-torque curve at rated voltage, the following information:

Poles	4
No-load Speed	1680 RPM
Rated Voltage	
Control field	115 volts
Reference field	115 volts
Stall Torque at	
rated voltage	13.7 oz-in
Moment of Inertia	0.66 oz-in ²
	(0.00171 oz-in-sec ² /rad)

At zero speed and rated control voltage, the slope of the speed-torque curve is such that the friction coefficient is⁷

$$f = 0.0215 \text{ oz-in-sec/rad} \quad (19)$$

The maximum possible motor acceleration is

$$\begin{aligned} \alpha_m &= \frac{13.7}{0.00171} \\ &= 8000 \text{ rad/sec}^2 \end{aligned} \quad (20)$$

If the speed-torque characteristics are "linearized" as shown in Fig. 24, the no-load speed at rated control voltage would become

$$\begin{aligned} \Omega_m &= \frac{13.7}{0.0215} \\ &= 637 \text{ rad/sec} \quad (6080 \text{ RPM}) \end{aligned} \quad (21)$$

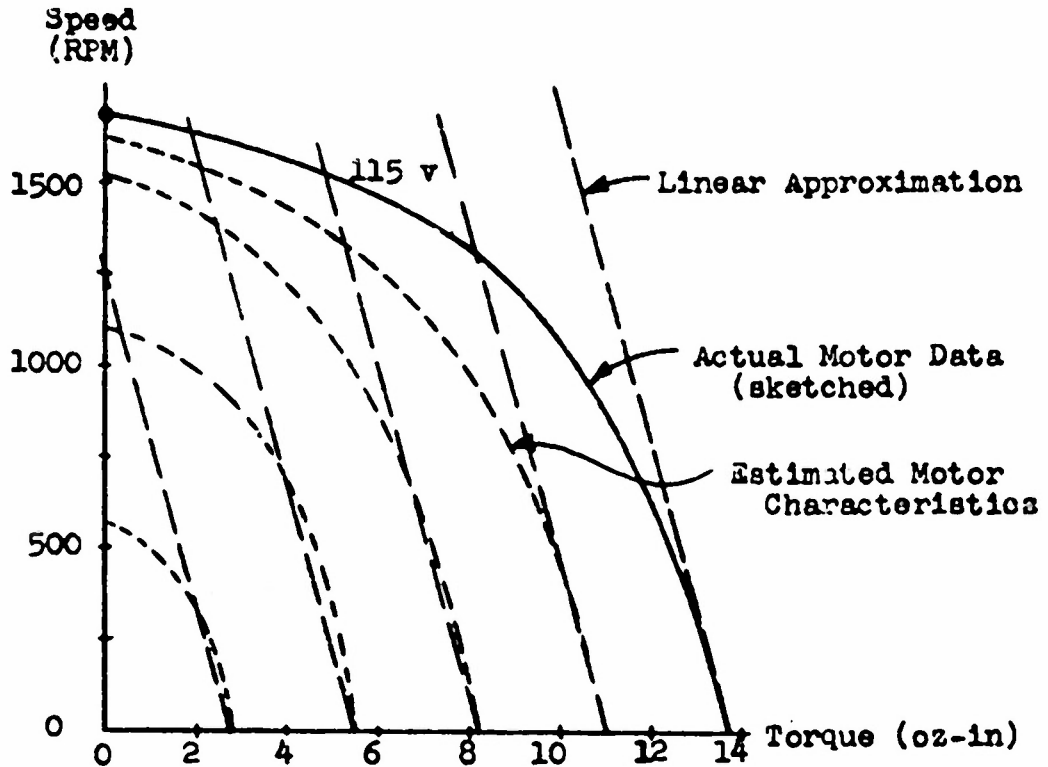


Figure 24. Typical Speed-Torque Characteristics and Linear Approximation

which would correspond to a motor constant

$$K_m = \frac{637}{115} = 5.55 \text{ rad/sec/volt} \quad (22)$$

The time constant of the motor, based on the value of f computed above, would be

$$\tau = \frac{0.00171}{0.0215} \approx 0.08 \text{ sec} \quad (23)$$

A time constant calculated in this way is known to be optimistic and the actual performance will not be as good as this figure would indicate¹⁵⁻¹⁷. Likewise, the no-load speed of 6080 RPM is too high. The linear approximation, however, is reasonably good

at low speeds, as suggested by Fig. 24, and provides a basis for a fair estimate of the performance of a servo system in which this motor is used.

Suppose now that a servomechanism is designed using this motor, coupled to an output shaft which adds no friction or inertia through an ideal gear train having a ratio of 100, and that the data system supplies a signal of 1 volt/degree or 57.3 volts/radian. A block diagram of this system, assuming that the motor is controlled by a linear amplifier, is given in Fig. 25.

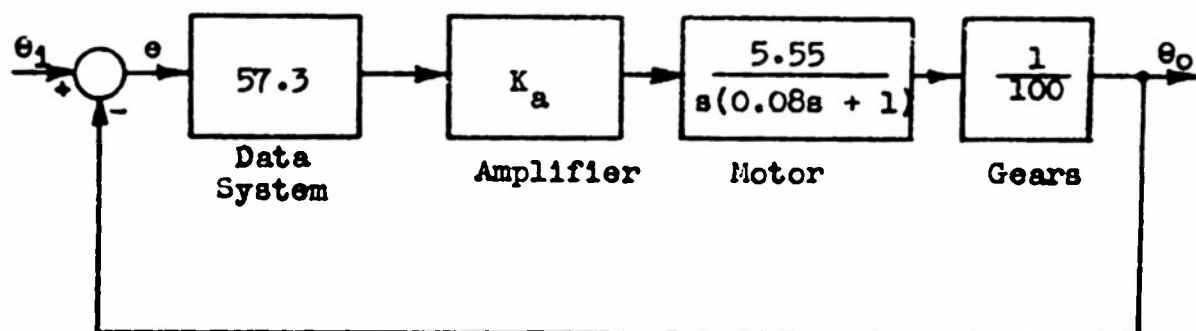


Figure 25. Block Diagram of a Linear Servomechanism

Since the damping ratio is

$$\zeta = \frac{1}{2 \sqrt{K_v \tau}}, \quad (24)$$

the velocity constant of the system will be

$$K_v = \frac{1}{4 \zeta^2 \tau} \quad (25)$$

$$= (57.3)(K_a)(5.55)\left(\frac{1}{100}\right), \quad (26)$$

where $\tau = 0.08$ seconds.

If it is specified that $\xi = 0.5$, it turns out that the velocity constant and amplifier gain must be

$$K_v = 12.5 \text{ sec}^{-1} \quad (27)$$

$$K_a = 3.93 \text{ volts/volt} \quad (28)$$

With this amplifier gain, the system has an undamped resonant frequency which is

$$\begin{aligned} \omega_o &= \sqrt{\frac{K_v}{\tau}} \\ &= 12.5 \text{ rad/sec} \end{aligned} \quad (29)$$

or approximately 2 cycles/second.

Using Eq. (17) and Table I, the integrated absolute error as a function of the initial error for a step input is

$$\begin{aligned} A_1 &= \frac{1.712}{12.5} e_o \\ &= 0.137 e_o \end{aligned} \quad (30)$$

provided that saturation does not occur. If limiters are employed so that the motor voltage cannot exceed 115 volts, saturation may be considered to result for errors greater than

$$\begin{aligned} e_o &= \frac{115}{(57.3)(3.93)} \\ &= 0.51 \text{ rad} \end{aligned} \quad (31)$$

or approximately 29 degrees. Thus for initial errors less than 0.51 radians, the integrated absolute error is given by Eq. (30), which gives a straight line of unit slope on log-log coordinates.

For initial errors greater than 0.51 radians, the integrated absolute error is larger than the linear theory predicts, as indicated in Fig. 20. A curve like those of Fig. 20 was plotted for $\xi = 0.5$ and used to obtain the curve for the saturating system shown in Fig. 26.

To plot the corresponding curves for this same motor, with the same gear train but operated by an optimum off-on controller of the type considered earlier¹³, it is necessary to compute the maximum available output speed and acceleration. Using the same assumptions as before, the maximum output speed would be

$$\Omega = \frac{637}{100} = 6.37 \text{ rad/sec} \quad (32)$$

and the maximum output acceleration, from Eq. (20), is

$$\alpha = \frac{8000}{100} = 80 \text{ rad/sec}^2 \quad (33)$$

As shown in a previous study¹³, the integrated absolute error cannot be less than

$$A_1 = \frac{e_o^2}{2\Omega} = \frac{e_o^2}{12.74} \quad (34)$$

or

$$A_1 = \sqrt{\frac{1}{\alpha}} e_o^{3/2} = \frac{e_o^{3/2}}{8.94}, \quad (35)$$

whichever is greater. Equation (34) was developed by neglecting the inertia, while (35) is based on zero friction; the integrated absolute error in the presence of both inertia and friction is somewhat greater than the values given by Eqs. (34) and (35), which are the asymptotes of the correct curve for very small or very large initial errors.

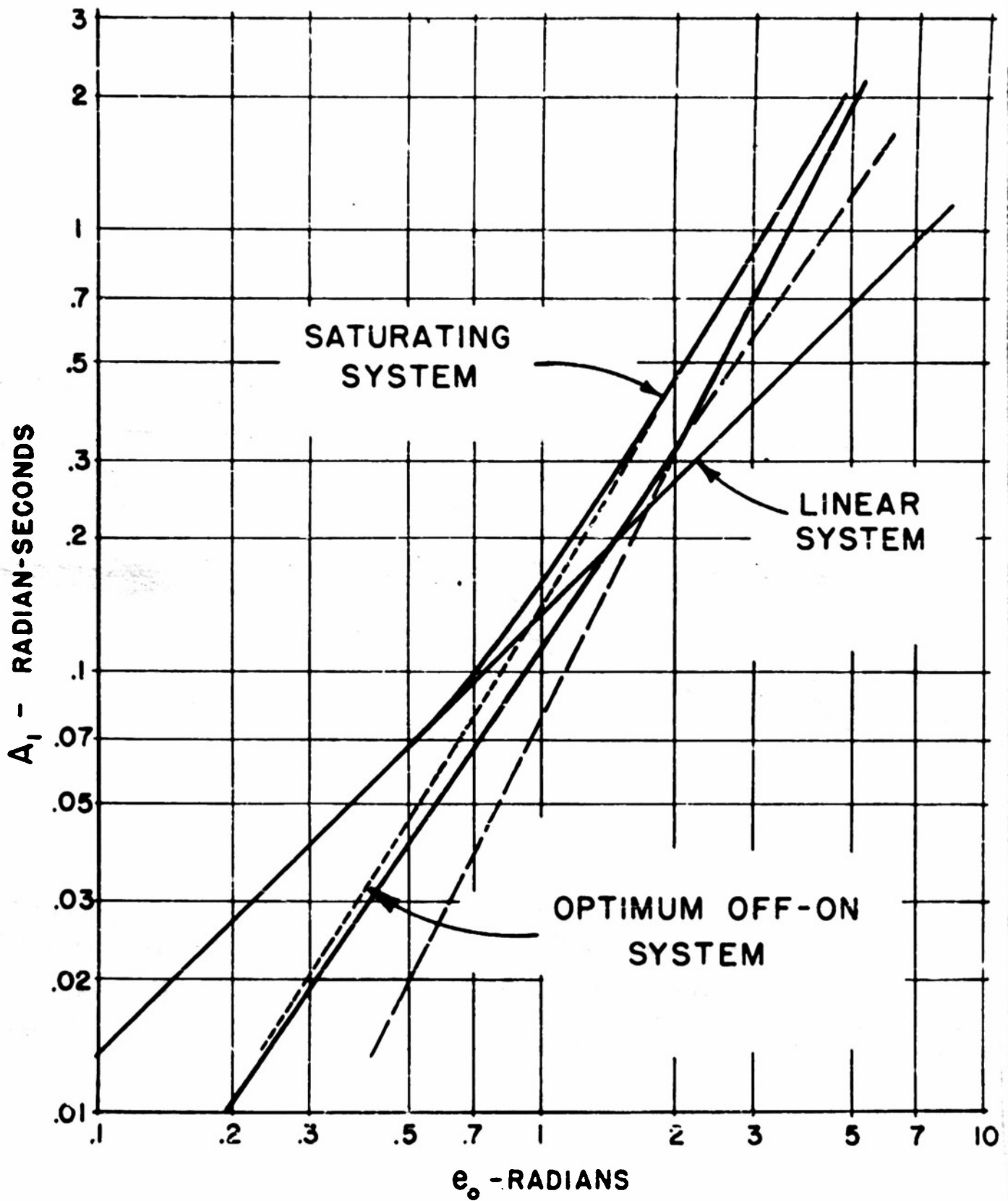


FIGURE 26

The straight-line asymptotes for the off-on system can be drawn quickly by noting that the curves intersect at an initial error given by

$$\frac{e_o^{3/2}}{8.94} = \frac{e_o^2}{12.74} \quad (36)$$

or
$$e_o^{1/2} = 1.425$$

$$e_o = 2.03 \text{ rad} \quad (37)$$

At this point, the integrated absolute error is

$$A_1 = 0.326 \text{ rad-sec} \quad (38)$$

from either Eq. (34) or (35). Knowing one point on each curve and the slopes, the dashed lines in Fig. 26 are constructed. The dotted curve represents the actual integrated absolute error for a system with both friction and inertia, and is taken from Fig. 6E of a previous report¹³.

An optimum off-on system is no better than a saturating but otherwise linear system for initial errors of 2 radians or more, but is significantly better at small errors. It should be remembered that all of the curves in Fig. 26 are calculated for a linear approximation to the motor curves and that consideration of the curvature in the speed-torque characteristics would lead to a new set of performance curves. However, the trends shown in Fig. 26 are representative.

The advantage of the off-on system at small errors would be decreased by use of a lower damping ratio and higher amplifier gain, as indicated by Fig. 22.

Similar results, differing chiefly in detail, are obtained by use of the integrated squared error. The slopes of the curves are less distinctive, the integral is relatively insensitive to small variations in the system response, and the range of the integral is greatly increased. The necessary curves are developed in the same way as the curves of integrated absolute error used in this report.

CONCLUSIONS

The statement has been made that "it is very easy to synthesize controllers which appear theoretically to have very superior performance but when actually installed prove to be very ordinary, because of system saturation"³. It is probably more appropriate to state that a saturated system is almost as good as the best system that can be built and that, if saturation is avoided, a loss of performance is necessarily entailed.

Hopkin¹⁸ and McDonald¹⁹ have described "dual-mode" systems, in which a nonlinear lead network and relay are used for large errors and a linear amplifier is used for small errors. The saturating system with viscous friction is similar in some respects, but has the disadvantage that the saturating amplifier must be capable of supplying the full motor power.

Comparison of the responses of linear and nonlinear systems should be based on a thorough investigation for a wide range of input magnitudes. Log-log plots of integrated absolute or squared error, based on either analytic or experimental data, seem to provide a graphic method of comparison.

APPENDIX I

A. Linear Region

In the unsaturated or linear region, where $-e_c \leq e \leq e_c$, the differential equation is [from Eq. (4)]

$$\ddot{e} + 2 \zeta \omega_0 \dot{e} + \omega_0^2 e = 0 \quad (A1)$$

subject to the initial conditions

$$e(0) = P e_c \quad (A2)$$

$$\dot{e}(0) = Q_0 \frac{\omega_0 e_c}{2 \zeta} \quad (A3)$$

Applying the Laplace transformation to Eq. (A1) and rearranging, we obtain

$$E(s) = P e_c \frac{s + 2 \zeta \omega_0}{s^2 + 2 \zeta \omega_0 s + \omega_0^2} + Q_0 \frac{\omega_0 e_c}{2 \zeta} \frac{1}{s^2 + 2 \zeta \omega_0 s + \omega_0^2} \quad (A4)$$

$$= P e_c \left[\frac{s + \zeta \omega_0}{(s + \zeta \omega_0)^2 + \omega_0^2 (1 - \zeta^2)} + \frac{\zeta \omega_0}{(s + \zeta \omega_0)^2 + \omega_0^2 (1 - \zeta^2)} \right] \quad (A5)$$

$$+ Q_0 \frac{\omega_0 e_c}{2 \zeta} \frac{1}{(s + \zeta \omega_0)^2 + \omega_0^2 (1 - \zeta^2)}$$

The inverse transformation gives

$$\frac{e(t)}{e_c} = P e^{-\zeta \omega_0 t} \left[\cos \sqrt{1 - \zeta^2} \omega_0 t + \frac{\zeta}{\sqrt{1 - \zeta^2}} \sin \sqrt{1 - \zeta^2} \omega_0 t \right] \quad (A6)$$

$$+ Q_0 \frac{1}{2 \zeta \sqrt{1 - \zeta^2}} e^{-\zeta \omega_0 t} \sin \sqrt{1 - \zeta^2} \omega_0 t$$

By direct differentiation, the error rate in the linear region is

$$\begin{aligned} \frac{\dot{e}(t)}{e_c} = & -P\omega_0 \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_0 t} \sin \sqrt{1-\zeta^2} \omega_0 t \\ & + \frac{Q_0\omega_0}{2\zeta} e^{-\zeta\omega_0 t} \left[\cos \sqrt{1-\zeta^2} \omega_0 t - \frac{\zeta}{\sqrt{1-\zeta^2}} \sin \sqrt{1-\zeta^2} \omega_0 t \right] \end{aligned} \quad (A7)$$

B. Saturated Region

In the saturated region for which $e \geq e_c$, the differential equation is [from Eq. (4)]

$$\ddot{e} + 2\zeta\omega_0 \dot{e} = -\omega_0^2 e_c \quad (B1)$$

subject to the initial conditions of Eqs. (A2) and (A3). Applying the Laplace transformation and rearranging, we obtain

$$E(s) = \frac{Pe_c}{s} - \frac{\omega_0^2 e_c}{s^2(s + 2\zeta\omega_0)} + Q_0 \frac{\omega_0 e_c}{2\zeta} \frac{1}{s(s + 2\zeta\omega_0)} \quad (B2)$$

The inverse transformation gives

$$\frac{e(t)}{e_c} = P - \frac{\omega_0}{2\zeta} t + \frac{1}{4\zeta^2} (1 - e^{-2\zeta\omega_0 t}) + \frac{Q_0}{4\zeta^2} (1 - e^{-2\zeta\omega_0 t}) \quad (B3)$$

from which, by differentiation, we obtain

$$\frac{\dot{e}(t)}{e_c} = -\frac{\omega_0}{2\zeta} + \frac{\omega_0}{2\zeta} e^{-2\zeta\omega_0 t} + Q_0 \frac{\omega_0}{2\zeta} e^{-2\zeta\omega_0 t} \quad (B4)$$

In the other saturated region for which $e \leq -e_c$, the right-hand side of Eq. (B1) becomes $+\omega_0^2 e_c$; the only difference in the solutions is therefore in the signs of the terms which do not involve the initial conditions.

The error is thus

$$\frac{e(t)}{e_c} = P + \frac{u_0}{2\zeta} t - \frac{1}{4\zeta^2} (1 - e^{-2\zeta\omega_0 t}) + \frac{Q_0}{4\zeta^2} (1 - e^{-2\zeta\omega_0 t}) \quad (B5)$$

and the error rate is

$$\frac{\dot{e}(t)}{e_c} = \frac{\omega_0}{2\zeta} - \frac{u_0}{2\zeta} e^{-2\zeta\omega_0 t} + Q_0 \frac{\omega_0}{2\zeta} e^{-2\zeta\omega_0 t} \quad (B6)$$

APPENDIX II

Although calculation of the integrated absolute error for even a linear second-order system appears formidable at first glance, the process is straightforward and the final result is remarkably simple. If the system is assumed to be initially at rest, the error for a unit step function input is [from Appendix I, Eq. (A6)]

$$e(t) = e^{-\zeta\omega_0 t} \left[\cos\sqrt{1-\zeta^2} \omega_0 t + \frac{\zeta}{\sqrt{1-\zeta^2}} \sin\sqrt{1-\zeta^2} \omega_0 t \right] \quad (C1)$$

which can also be written as

$$e(\theta) = e^{-\sigma\theta} \left[\cos \theta + \sigma \sin \theta \right] \quad (C2)$$

where

$$\sigma = \frac{\zeta}{\sqrt{1-\zeta^2}} \quad (C3)$$

$$\theta = \sqrt{1-\zeta^2} \omega_0 t \quad (C4)$$

The integral of the absolute value of the error is then

$$I = \int_0^\infty |e(t)| dt \quad (C5)$$

$$= \frac{1}{\omega_0 \sqrt{1-\zeta^2}} \int_0^\infty |e(\theta)| d\theta \quad (C6)$$

In the oscillatory case, the error is alternately positive and negative, being zero for

$$\theta = \tan^{-1}(-1/\sigma) + n\pi \quad (C7)$$

$$= \theta_1 + n\pi \quad (C8)$$

where $\pi/2 < \theta_1 < \pi$, and $n = 0, 1, 2, \dots$. Equation (C6) may thus be written as the sum of a number of integrals with alternating signs, as follows:

$$I_{\omega_0} \sqrt{1 - \zeta^2} = \int_0^{\theta_1} e(\theta) d\theta - \int_{\theta_1}^{\theta_1 + \pi} e(\theta) d\theta + \int_{\theta_1 + \pi}^{\theta_1 + 2\pi} e(\theta) d\theta - \dots \quad (C9)$$

$$= -E(0) + 2 E(\theta_1) - 2 E(\theta_1 + \pi) + 2 E(\theta_1 + 2\pi) - \dots \quad (C10)$$

where $E(\theta)$ is the indefinite integral

$$E(\theta) = \int e(\theta) d\theta \quad (C11)$$

$$= \frac{e^{-\sigma\theta}}{\sigma^2 + 1} \left[(1 - \sigma^2) \sin\theta - 2\sigma \cos\theta \right] \quad (C12)$$

Now, since $\cos(\theta_1 + n\pi) = (-1)^n \cos\theta_1$ and $\sin(\theta_1 + n\pi) = (-1)^n \sin\theta_1$, Eq. (C10) can be put in the form

$$I_{\omega_0} \sqrt{1 - \zeta^2} = \frac{2\sigma}{\sigma^2 + 1} + 2 \sum_{n=0}^{\infty} \frac{e^{-\sigma\theta_1} e^{-n\pi}}{\sigma^2 + 1} \left[(1 - \sigma^2) \sin\theta_1 - 2\sigma \cos\theta_1 \right] \quad (C13)$$

$$= \frac{2\sigma}{\sigma^2 + 1} \left\{ 1 + e^{-\sigma\theta_1} \left[\frac{1 - \sigma^2}{\sigma} \sin\theta_1 - 2 \cos\theta_1 \right] \sum_{n=0}^{\infty} e^{-n\pi} \right\} \quad (C14)$$

The summation can be written in finite form by use of the formula for the sum of a geometric series; the result is

$$I_{m_0} \sqrt{1 - \zeta^2} = \frac{2\sigma}{\sigma^2 + 1} \left\{ 1 + \frac{e^{-\sigma\theta_1}}{1 - e^{-\sigma\pi}} \left[\frac{1 - \sigma^2}{\sigma} \sin\theta_1 - 2 \cos\theta_1 \right] \right\} \quad (C15)$$

Further simplification is possible in the interests of ease of computation. Recalling that $\tan \theta_1 = -1/\sigma$, it turns out that

$$\cos \theta_1 = -\zeta \quad (C16)$$

$$\sin \theta_1 = \sqrt{1 - \zeta^2} \quad (C17)$$

With these substitutions and leaving $\sigma = \zeta (1 - \zeta^2)^{-1/2}$ wherever convenient, the final expression is

$$I = \frac{2}{\omega_0} \left[\zeta + \frac{e^{-\sigma\theta_1}}{1 - e^{-\sigma\pi}} \right] \quad (C18)$$

The integrated absolute error for a step function of magnitude P is therefore

$$A_1 = \frac{2P}{\omega_0} \left[\zeta + \frac{e^{-\sigma\theta_1}}{1 - e^{-\sigma\pi}} \right] \quad (C19)$$

The simplicity of Eq. (C19) is somewhat misleading, as σ and θ_1 are still functions of the damping ratio, ζ ; nevertheless, the result is simpler than might have been expected.

REFERENCES

1. "Servos with Torque Saturation", W. Hurewicz and N. B. Nichols. Radiation Laboratory Report 555, May 1, 1944.
2. "Servos with Torque Saturation", W. Hurewicz. Radiation Laboratory Report 592, September 28, 1944.
3. Automatic Feedback Control (book), W. R. Ahrendt and J. F. Taplin. McGraw-Hill Book Company, New York, 1951, pp. 209-213.
4. "A Study of Amplifier Saturation and Magnetic Saturation in a Servomechanism", C. Leondes. AIEE Miscellaneous Paper 52-198, May, 1952.
5. "Some Saturation Phenomena in Servomechanisms", E. Levinson. AIEE Technical Paper 53-110, December, 1952.
6. "Compensation of Feedback-Control Systems Subject to Saturation", G. C. Newton. Jour. Frank. Inst., Vol. 254, October, 1952, pp. 281-296; November, 1952, pp. 391-413.
7. Principles of Servomechanisms (book), G. S. Brown and D. P. Campbell. John Wiley and Sons, New York, 1948.
8. "Analysis of General Dynamic Systems Including One Nonlinear Element", T. M. Stout. Electrical Engineering Department Report No. 11, University of Washington, December 19, 1952.
9. "Analogue Methods for Optimum Servomechanism Design", F. C. Fickelsen and T. M. Stout. Trans. A.I.E.E., Vol. 71, Part II, 1952, pp. 244-250.
10. "A Differential Amplifier Study of Certain Nonlinearly-Damped Servomechanisms", R. R. Caldwell and V. C. Rideout. AIEE Technical Paper 53-107, 1953.
11. The Analysis and Synthesis of Linear Servomechanisms (book), A. C. Hall. Technology Press, Cambridge, Mass., 1947.
12. Torque-Saturated Second-Order Servomechanism Response, D. M. Olson. Thesis for MS(EE) Degree, University of Washington, 1953.
13. "A Study of Some Nearly Optimum Servomechanisms", T. M. Stout. Electrical Engineering Department Report No. 10, University of Washington, November 28, 1952.
14. "Nonlinear Techniques for Improving Servomechanism Performance", D. McDonald. Proc. Nat. Elect. Conf., Vol. 6, Chicago, 1950, pp. 400-421.

15. "Operating Characteristics of Two-Phase Servomotors", R. J. N. Koopman. Trans. A.I.E.E., Vol. 68, 1949, pp. 319-329.
16. "Transient Response of Small 2-Phase Induction Motors", A. M. Hopkin. Trans. A.I.E.E., Vol. 70, 1951, pp. 881-886.
17. "Transfer Function for a 2-Phase Induction Servo Motor", L. O. Brown. Trans. A.I.E.E., Vol. 70, 1951, pp. 1890-1893.
18. "A Phase-Plane Approach to the Compensation of Saturating Servomechanisms", A. M. Hopkin. Trans. A.I.E.E., Vol. 70, 1951, pp. 631-639.
19. "Multiple Mode Operation of Servomechanisms", D. McDonald. Rev. Sci. Inst., Vol. 23, January, 1952, pp. 22-30.

University of Washington
Distribution List for Technical Reports
Nonr-477(02)
NR 374 371

Chief of Naval Research Department of the Navy Washington 25, D. C. 4 Code 427	Chief, Bureau of Aeronautics Department of the Navy Washington 25, D. C. 2 Code KL
Director Naval Research Laboratory Washington, D. C. 4 Tech. Inf. Off.	Chief, Bureau of Ordnance Department of the Navy Washington 25, D. C. 4 Re
1 Director ONR Special Devices Center Fort Washington, New York	1 Naval Ordnance Laboratory White Oaks, Maryland
1 Office of Naval Research Branch Office 150 Causeway Street Boston 14, Mass.	Chief of Naval Operations Department of the Navy Washington 25, D. C. 3 Op-20
1 Office of Naval Research Branch Office 346 Broadway New York, N. Y.	1 Director Naval Electronics Laboratory San Diego, California
1 Office of Naval Research Branch Office 844 N. Rush Street Chicago 11, Ill.	U. S. Naval Post Graduate School Monterey, Cal. 1 Physics Dept. 1 Electronics Dept.
2 Office of Naval Research Branch Office 1000 Geary Street San Francisco 9, Cal.	1 Commanding Officer U. S. Naval Air Development Center Johnsville, Pennsylvania
1 Office of Naval Research Branch Office 1050 E. Green Street Los Angeles, Cal.	2 British Joint Services Mission (Technical Services) P. O. Box 680 Benjamin Franklin Station Washington, D. C.
2 Officer-in-Charge Office of Naval Research Navy #100, Fleet Post Office New York, N. Y.	1 British Commonwealth Scientific Office 1785 Massachusetts Avenue NW Washington, D. C.
Chief, Bureau of Ships Department of the Navy Washington 25, D. C. 2 Code 810	1 Carnegie Institute of Technology E. E. Department Pittsburgh, Pennsylvania (Prof. R. L. Bright)
	1 Stanford University Electronics Research Laboratory Stanford, California (L. Farrell McGhie)

- 1 University of California
Division of Electrical Engineering
Berkeley 4, California
(Prof. P. L. Morton)
- 1 National Bureau of Standards
Basic Instrumentation Section
Washington 8, D. C.
- 1 Polytechnic Institute of Brooklyn
Brooklyn, New York
(Prof. E. Weber)
- 1 Bell Aircraft Corporation
Niagara Falls, New York
(Mr. L. M. Glass)
- 1 Boeing Airplane Company
Seattle 14, Washington
(Mr. N. H. Nelson)
- 1 Consolidated Vultee Aircraft Co.
San Diego 12, Cal.
(Mrs. D. B. Burke)
- 1 Douglas Aircraft Company
3000 Ocean Blvd.
Santa Monica, Cal.
(Mr. E. F. Burton)
- 1 Rand Corporation
1500 Fourth Street
Santa Monica, Cal.
(Mr. R. H. Best)
- 1 Fairchild Engine and Airplane Corp.
Guided Missiles Division
Farmingdale, N. Y.
(Mr. S. M. Truman)
- 1 Raytheon Manufacturing Co.
Guided Missiles and Radar Division
Waltham, Mass.
- 1 Glenn L. Martin Co.
Baltimore 3, Md.
(Mrs. Mary Ezzo)
- 1 Goodyear Aircraft Co.
Culver City, Cal.
(Dr. E. Amstein)
- 1 Gruman Aircraft Engr. Corp.
Bethpage, New York
(Mr. T. Schwaelder)
- 1 Hughes Aircraft Company
Culver City, Cal.
(Mr. B. T. Milek)
- 1 North American Aviation, Inc
12214 Lakewood Blvd.
Downey, Cal.
(Aieophysics Library)
- 1 Sperry Gyroscope Company
Great Neck, L. I., N. Y.
(Librarian)
- 1 Mass. Institute of Technology
Research Lab. of Electronics
Cambridge 39, Mass.
(Dr. J. B. Wiesner)
- 2 Harvard University
Cruft Laboratory
Cambridge, Mass.
(Dr. R. P. King)