H. M. Sangster

THE UNIVERSITY OF MISSOURI

Bulletin

ENGINEERING REPRINT SERIES
Reprint Number 37

❖ ❖ ❖

Engineering Experiment Station
Columbia, Missouri

AUTOMATIC CONTROL -- THE FUNDAMENTALS

GLADWYN LAGO
Associate Professor of
Electrical Engineering

Reprinted from
Proceedings of the Second Annual Conference on Automatic Control
University of Oklahoma April 29-30, 1957

COLLEGE OF ENGINEERING
THE ENGINEERING EXPERIMENT STATION

The Engineering Experiment Station was organized in 1909 as a part of the College of Engineering. The staff of the Station includes all members of the Faculty of the College of Engineering, together with Research Assistants supported by the Station Funds.

The Station is primarily an engineering research institution engaged in the investigation of fundamental engineering problems of general interest, in the improvement of engineering design, and in the development of new industrial processes.

The Station desires particularly to co-operate with industries of Missouri in the solution of such problems. For this purpose, there is available not only the special equipment belonging to the Station but all of the equipment and facilities of the College of Engineering not in immediate use for class instruction.

Inquiries regarding these matters should be addressed to

The Director,
Engineering Experiment Station
University of Missouri
Columbia, Missouri

THE UNIVERSITY OF MISSOURI BULLETIN

VOL. 59, NO. 32 ENGINEERING REPRINT SERIES, NO. 37

Published by the University of Missouri at Room 102, Building T-3, Columbia, Missouri. Entered as second-class matter, January 2, 1914, at post office at Columbia, Missouri, under Act of Congress of August 24, 1912. Issued four times monthly October through May, three times monthly June through September.

500
September 20, 1958

AUTOMATIC CONTROL - THE FUNDAMENTALS

GLADWYN LAGO

UNIVERSITY OF MISSOURI

AUTOMATIC CONTROL - THE FUNDAMENTALS

Gladwyn Lago

Since the beginning of recorded time, man has sought to make his life easier by producing devices that would relieve him of the drudgery of many tasks. The industrial revolution in the 19th century relieved man's muscles of many tasks by using energy obtained from coal, oil, other chemicals and from flowing water.

The development of automatic control devices in the last few years has been referred to by some as the second industrial revolution. The first industrial revolution relieved man's muscles whereas the second industrial revolution relieves man's brains as well as his muscles.

Automatic control cuts across all fields and in the process creates a new field of its own. Automatic control encompasses such things as: positioning of mechanical objects, controlling fluid flow, viscosity, temperature, humidity, pressure, chemical reactions, handling of parts, assembly of parts, data handling, computing and inventory control among others.

Automatic control in the broadest sense of the terms signifies an entire group of operations. For example, raw materials may be supplied to the input of a plant and the finished product taken from the output without being touched by a human operator. Inside this plant many distinct sets of operations may be involved each of which may make use of one or more automatic control devices. The operation of the entire plant can be understood once each of the automatic control devices itself is understood. The building block of all these devices is the feedback control loop or the closed-loop system. This paper is a review of many methods of analysis and design of feedback control systems.

A simplified version of a closed-loop system is shown in Figure 1. This figure should be referred to for the following discussion. The symbol $r(t)$ is the time function that describes the reference input or input signal. The $c(t)$ is the controlled variable or output of the closed-loop system. The difference between the output and the input is called the actuating signal (often referred to as the error signal) and $e(t)$ stands for this signal. The $g_2(t)$ refers to the plant or that portion of the system that the feedback loop is attempting to control. Many times the characteristics of the plant must be modified before the loop can be closed and this is done by the compensation elements referred to as $g_1(t)$. The signal $e(t)$ is modified by the compensation elements, and the modified signal $m(t)$ is applied to the plant.

Although many mathematical tools can be used in the analysis of feedback systems, the Laplace transform is the method used most extensively today. The direct Laplace transform of a time function is defined by use of an improper integral and the integral transforms the function of time into a function of a complex variable s. On Figure 1 R(s), E(s), $G_1(s)$, M(s), $G_2(s)$ and C(s) are the Laplace transforms of r(t), e(t), $g_1(t)$, M(t), $g_2(t)$ and c(t) respectively.

It can be shown that C(s) in terms of the other transformed quantities is

$$C(s) = \frac{R(s)\, G_1(s)\, G_2(s)}{1 + G_1(s)\, G_2(s)} \tag{1}$$

For convenience let

$$G(s) = G_1(s)\, G_2(s) \tag{2}$$

and the ratio of C(s) to R(s) can be written as

$$\frac{C(s)}{R(s)} = \frac{G(s)}{1+G(s)} \tag{3}$$

The designer of an automatic control device has himself in the peculiar position of wanting a system with zero error, yet he is working with a system that will not function unless there is an error. The answer to this apparent dilemma is that everything in the world is done only within certain tolerances. The same is true of a control system and as long as a closed-loop system can perform a given task within certain tolerances, the operation of the system is satisfactory.

One of the criteria that has been developed to determine if the performance of a system is satisfactory, is the response of the system to a unit-step function input. This type of input specifies that the system is at rest and all of a sudden the input signal r(t) is given a value of unity. The response of the output c(t) is observed and is plotted as shown in Figure 2. The quantities most often used to discuss a response of this type are: percent overshoot, rise time, settling time and delay time. These quantities are indicated on Figure 2 and discussed in the following paragraph.

The percent overshoot is the percentage by which the output exceeds the input. The rise time is the time it takes the system to go from 10 percent to 90 percent of the outputs final value. Settling time is a measure of how long it takes the system to settle down to within 5 percent of the final value. Delay time is the time is takes the system to reach 50 percent of the fianl value. Sometimes different percentage values are used in defining these terms (as for rise time and settling time) but the meaning of the definitions remain the same.

One method for the analysis of control systems is the writing and solving of the differential equations for the system with some standard type of input signal such as the unit-step function. This method leads to the transient response of the system in a straightforward manner. However, if the response obtained by this analysis is not the desired response, the system must then be redesigned and the transient response method of analysis is not satisfactory as a design tool. This is shown briefly through the use of an example.

Let it be supposed that the equations that describe a system are the following

$$1.425\, e(t) + 0.03 \frac{de(t)}{dt} + 7.6 \int e(t)\, dt = 0.0025 \frac{d^2c(t)}{dt^2} + 0.01 \frac{dc(t)}{dt} \tag{4}$$

and $$e(t) = r(t) - c(t) \tag{5}$$

When the Laplace transforms of equations 4 and 5 are taken and the results solved for the ratio of C(s) and R(s), the following equation is obtained.

$$\frac{C(s)}{R(s)} = \frac{12\ [s^2+47.4s+254]}{[s^3+16s^2+569s+3054]} \tag{6}$$

Before the solution of equation 6 can be found for more specific input signal, the denominator must be factored. When this is done the equation 6 can be written as

$$\frac{C(s)}{R(s)} = \frac{12\quad s+6.2\quad s+41.2}{[s+6]\ [s+5+j22]\ [s+5-j22]} \tag{7}$$

All of the methods for factoring high degree polynomials involve some sort of an approximation procedure. Therefore, the relationship that exists between the system coefficients and the roots of the characteristic equation (the denominator of the Laplace transform) are so devious that design essentially becomes a matter of trail and error. After a few terms are defined, this matter will be discussed further.

Equation 7 is a function of a complex variable and as s takes on most values, the magnitude of this equation behaves in a regular (analytic) manner except when s approaches a value that makes the denominator equal to zero. For example, as s approaches -6, the magnitude of the equation behaves in a rather singular manner in that it increases without bound. Therefore, any value of s that makes the denominator equal to zero is known as a singularity of the function of s. Actually, the type of singularities present in

equation 7 are known more specifically as poles. In a similar manner, any value of s that makes the numerator equal to zero is known as a zero of the function. For this particular example, the poles and zeros are located on the complex s-plane, as shown in Figure 3. The poles are indicated by the x-marks and the zeros by the small circles.

Now let us return to the discussion of the limitations of the transient method of response. For this particular example, suppose the transient response found for this system is not satisfactory and suppose the designer wants to move the poles at $s = -5 \pm j22$ farther to the left in the s-plane to say $s = -10 \pm j22$. What can be done to achieve the desired results. The devious nature of factoring makes it impossible to know what to do except to make a guess and repeat the entire analysis.

For these reasons, other design methods have been developed. A procedure that has received extensive attention is to execute the design in terms of the frequency response of the system. By this is meant that a sinusoidal input is applied to the system and left long enough for the system to reach steady-state conditions. After this the magnitude and phase of the output as compared with input is measured as the frequency is varied over a wide range. The curves obtained in this manner are used to determine if a system is satisfactory in somewhat the same manner that the step-function response curves are used. Mathematically, s can be replaced by $j\omega$ in the equations for C(s) over R(s) and the frequency response of the closed-loop system is obtained. A typical frequency response curve is shown in Figure 4. The magnitude of this ratio is defined as M. The maximum value of M is referred to here as M_M and the value of the angular frequency ω at which M is equal to 0.707 is called the band width of the system.

Graphical techniques have been developed that allow the designer to work in terms of the frequency response of the open-loop instead of the closed-loop systems. One of the things that this technique does is to solve the problem of absolute stability. This is one of the problems inherent when a feedback system is used. The open-loop system may be stable but when the feedback path is closed, the resulting closed-loop may be unstable. An unstable system is one in which the input no longer has any control over the output. If equation 3 is written as a function of ω by substituting $j\omega$ for s, the result is

$$\frac{C(j\omega)}{R(j\omega)} = \frac{G(j\omega)}{1 + G(j\omega)} \tag{8}$$

By inspection, it can be observed that

$$G(j\omega) = -1 \tag{9}$$

the ratio of $C(j\omega)$ to $R(j\omega)$ is infinite thus representing an unstable system. As a generalization of this technique, it has been shown that if $G(j\omega)$ is plotted on the complex plane and if as the resulting curve is traversed in the direction of increasing frequency, the minus one point is to the left, then the system will be stable when the feedback loop is closed. Conversely, if the minus one point is to the right, the system will be unstable when the feedback loop is closed. Figure 5 indicates a stable and an unstable system.

The plot of $G(j\omega)$ can also be made to yield much more information than whether a system is stable or not. A family of constant M curves can easily be placed on the $G(j\omega)$ plot as shown on Figure 6a. By observing the frequency at which $G(j\omega)$ crosses one of these curves, the closed-loop frequency characteristics can be determined from the open-loop frequency plot. In this example $G(j\omega)$ crosses the $M = 1.1$ curve at $\omega = \omega_1$ and this yields a point on the closed-loop plot of part b. At $\omega = \omega_2$, $G(j\omega)$ crosses the $M=1.3$ curve and yields another point in part b. From the plot of part b, the value of M_M and band width can be determined.

A third facet of the use of the $G(j\omega)$ plots is shown in Figure 7 and this has to do with the problem of compensation. The uncompensated system $G_2(j\omega)$ as shown yields an unstable system if the feedback loop is closed around only this portion of the system because of the location of the minus one point. If a compensation network $G_1(j\omega)$ can be placed in series with the $G_2(j\omega)$ so that the resulting curve

$$G(j\omega) = G_1(j\omega)\ G_2(j\omega) \tag{10}$$

is shifter as shown, a stable system will result when the feedback path is closed around the combination. This compensation of a $G(j\omega)$ plot to reshape it until it has more desirable characteristics is perhaps the most important part of the use of these plots.

The procedure of designing a system in terms of the frequency response has brought up the problem of correlating the frequency response and the transient response of a system. For example, a desired transient response may be the final result but the designer chooses to work in terms of the system frequency response. One rule of thumb that has been used extensively states that if the frequency response of the closed-loop system has a value of $M_m = 1.3$, the transient response will be satisfactory. However, many systems exist that have this same 1.3 value and yet the systems transient response may be utterly different. Two such systems are indicated on Figure 8. Another consideration is that even though this rule of thumb has been verified in a large number of cases, there is always a feeling of doubt when such a rule is applied in a new situation.

The need for correlating the frequency response and the transient response is the primary reason for development of the root-locus method

which has been used extensively. This method studies the movement of the poles of the closed-loop system as the gain of the open-loop system is changed. Attention is focused upon the entire complex s-plane and from this the frequency response and transient response can be deduced. A more complete discussion of this concept will be presented after the root-locus method itself is discussed further.

The root-locus plots behave in a very definite manner and follow a set of rules that are relatively easy to apply. Although a complete discussion of this subject is beyond the scope of this paper. Figure 9 gives an idea of the behavior of these plots. For this figure, the open-loop G(s) function is

$$G(s) = \frac{K(s+0.05)}{s(s+0.5)\ (s+1)\ (s+2)} \tag{11}$$

The gain constant K is the variable and takes on values from zero to infinity. The location of the poles of closed-loop system is the quantity depicted as a function of the gain K on the root-locus plot. In general, the loci move from the open-loop poles to the open-loop zeros. Therefore, in Figure 9, when the gain K is zero, the poles of the closed-loop system start at 0, -0.5, -1, and -2. As the gain K is increased, one of the poles of the closed-loop system moves from 0 to -0.05 along the negative real axis. Another of the poles moves from -2 to $-\infty$ along the negative real axis. The other two poles start at -0.5 and -1 and first move toward each other along the negative real axis and finally meet at some intermediate point. After this, they break away from the negative real axis and move toward their respective asymptotes. The location of the break-away point and of the asymptotes can easily be calculated. These plots can also be calibrated in terms of K. That is for a certain value K, each of the loci will be at a certain point. The approximate location of the loci for some value of K equal to K_1 is shown by the arrows on Figure 9.

The zeros of the open-loop and closed-loop systems are the same. The location of the poles and zeros for the system shown on Figure 9 for K set at a value of K_1 is shown on Figure 10. If this system is driven by a unit step function, the Laplace transform of the output is given in general terms by

$$C(s) = \frac{K_1\ (S+Z)}{S(S+P_1)\ (S+P_2)\ (S+\sigma_1+j\omega_1)\ (S+\sigma_1-j\omega_1)} \tag{12}$$

The inverse transform of equation 12 (which is the step function response) is of the form

$$c(t) = k_o+k_1\epsilon^{-P_1 t}+k_2\epsilon^{-P_2 t}+k\epsilon^{-\sigma_1 t}\ \mathrm{Sin}\ (\omega_1 t+J) \tag{13}$$

The terms containing P_1 and P_2 are of the exponential decaying type. The farther P_1 and P_2 are to the left of the complex s-plane, the more rapidly these terms will decay. Also, the more nearly $-P_1$ approaches the zero at $-Z$, the smaller will be the coefficient k_1. The last term of equation 13 represents and xponentially decaying sinusoidal component of the response. The farther the complex conjugate pair of poles are to the left in the complex plane, the more rapidly this term will decay. The farther these complex pair of poles are away from the horizontal axis, the higher the frequency of oscillation of this component.

Therefore, the skilled designer can observe the location of the root-locus plots and almost by inspection can determine the approximate step function response the system will have for a certain value of K.

Not only is the step function response apparent from these plots, but in a somewhat different manner, the frequency response can also be obtained. As stated earlier, the frequency response of the closed-loop system can be obtained mathematically by replacing S by $j\omega$ in ratio of C(s) over R(s). When this is done for the system being discussed, the result is

$$\frac{C(j\omega)}{R(j\omega)} = \frac{K_1(j\omega+Z)}{(j\omega+P_1)(j\omega+P_2)(j\omega+\sigma_1+j\omega_1)(j\omega+\sigma_1-j\omega_1)} \qquad (14)$$

By inspection of Figure 11, it can be seen that each of the quantities such as $(j\omega-Z)$ is a vector (or phasor) from the particular zero or pole in question to the point on the vertical axis determined by the value of the ω under consideration. Therefore, the frequency response of the system at a particular value of ω is the product of K_1 times the proper number of vectors in the numerator corresponding to the number of zeros of the system divided by the proper number of vectors in the denominator corresponding to the number of poles of the system. In order to determine how the frequency response varies as a function of frequency, it is only necessary to visualize how these vectors swing as ω moves up the vertical axis.

The analysis of a system as K is varied is straightforward by use of the root-locus method. However, if no value of K gives the desired response, then the root-locus plots must be reshaped or in other words compensation is required. As a simple example of this sort, let it be supposed that it is desired to have the loci that move into the right half plane and do so at a higher value of K. This can be done by adding a compensation network which itself adds poles and zeros to the open-loop system. The zeros may be added in such a manner as to attract the loci that formally moved into the right half-plane and by inserting two new poles farther to the left on the negative real axis. The resulting loci is indicated on Figure 12.

The point to be emphasized from this example is that there is no clear cut method of determining where to add the poles and zeros of the compensation network. The loci can be reshaped by using any number of different compensation networks and the determination of which of these to use is essentially a trial and error procedure.

The final method of design of feedback control systems to be discussed is the method originally suggested by E. A. Guillemin. The Guilllemin's procedure removes the trial and error process necessary in the other methods discussed thus far by first focusing attention on the poles and zeros of the closed-loop system and then determining the poles and zeros of the open-loop system. Once these are known, the $G_1(s)$ for the compensation network can be found from which the compensation network can be synthesized. As a very brief example of this, suppose the plant portion of the system is given by

$$G_2(s) = \frac{K}{S(S+P_1)\ (S+P_2)} \tag{15}$$

The demands placed on the system are studied and it is determined that the following ratio of C(s) to R(s) will satisfy these demands

$$\frac{C(s)}{R(s)} = \frac{G_2(s)\ G_2(s)}{1+\ G_1(s)\ G_2(s)}$$

$$= \frac{K_2(s+Z)}{(s+P_3)(s+P_4)(s+\sigma+j\omega)(s+\sigma-j\omega)} \tag{16}$$

This equation can be chosen to satisfy both frequency response and transient response requirements. From this the product of $G_1(s)$ $G_2(s)$ can be determined as

$$G_1(s)\ G_2(s) = \frac{K_3(s+Z)}{S(s+P_5)(s+P_6)(s+P_7)} \tag{17}$$

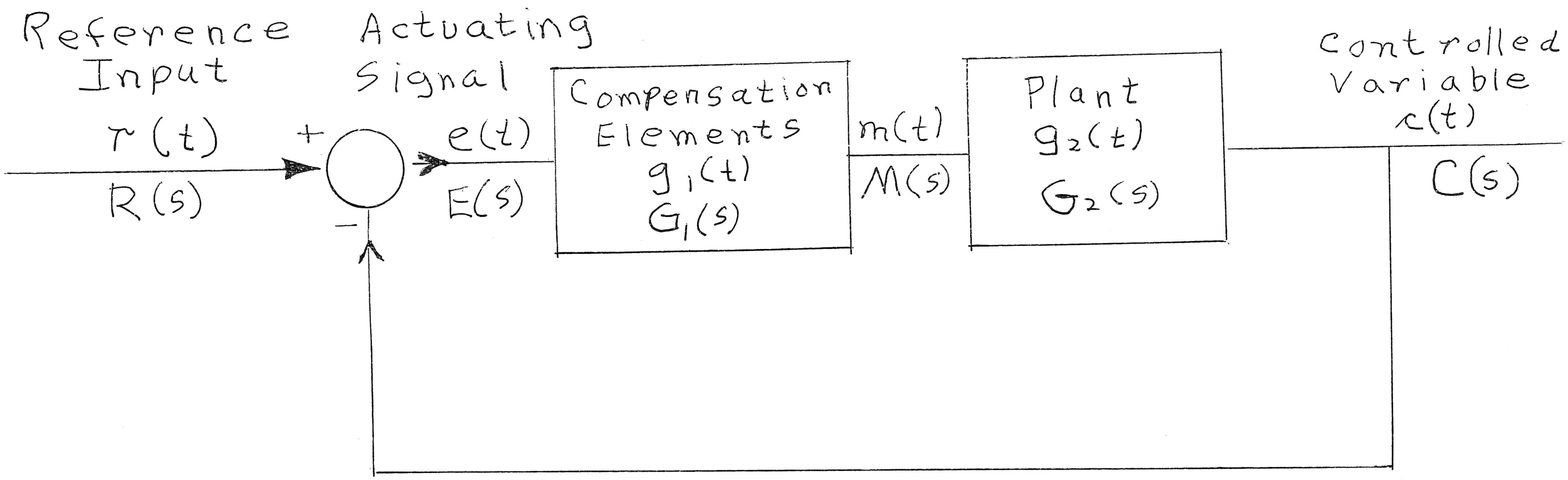
and the equation for the compensation network can be determined as

$$G_1(s) = \frac{G_1(s)\ G_2(s)}{G_2(s)} = \frac{K_4(s+Z)(s+P_1)(s+P_2)}{(s+P_5)(s+P_6)(s+P_7)} \tag{18}$$

From this equation the compensation can be designed.

This method is essentially a direct synthesis procedure and eliminates the trial and error process inherent in the other methods discussed. This method of design essentially says that we want the closed loop system to behave in a certain manner so let this be the starting point. Then the compensation network is found to make the system function in the desired manner.

This paper has presented a very brief and sketchy survey of the methods now in common use for the design of feedback control systems. It is admitted that the coverage is far from complete. The purpose of this paper is rather to show how the pieces fit together with the hope of increasing the readers interest in this fascinating field.



Reference Input
r(t)
R(s)
+
−
Actuating Signal
e(t)
E(s)
Compensation Elements
$g_1(t)$
$G_1(s)$
m(t)
M(s)
Plant
$g_2(t)$
$G_2(s)$
Controlled Variable
c(t)
C(s)

Figure 1

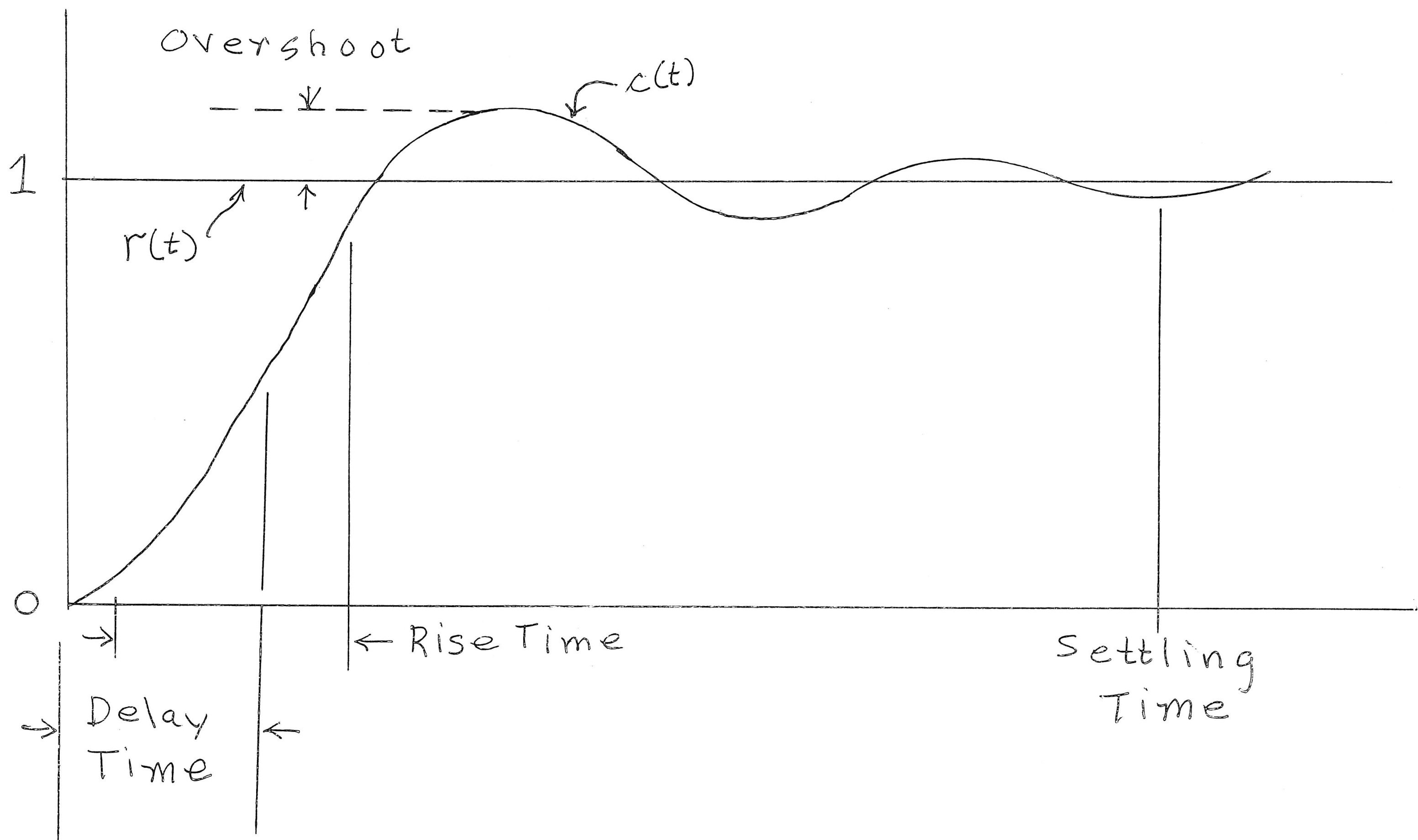


Figure 2

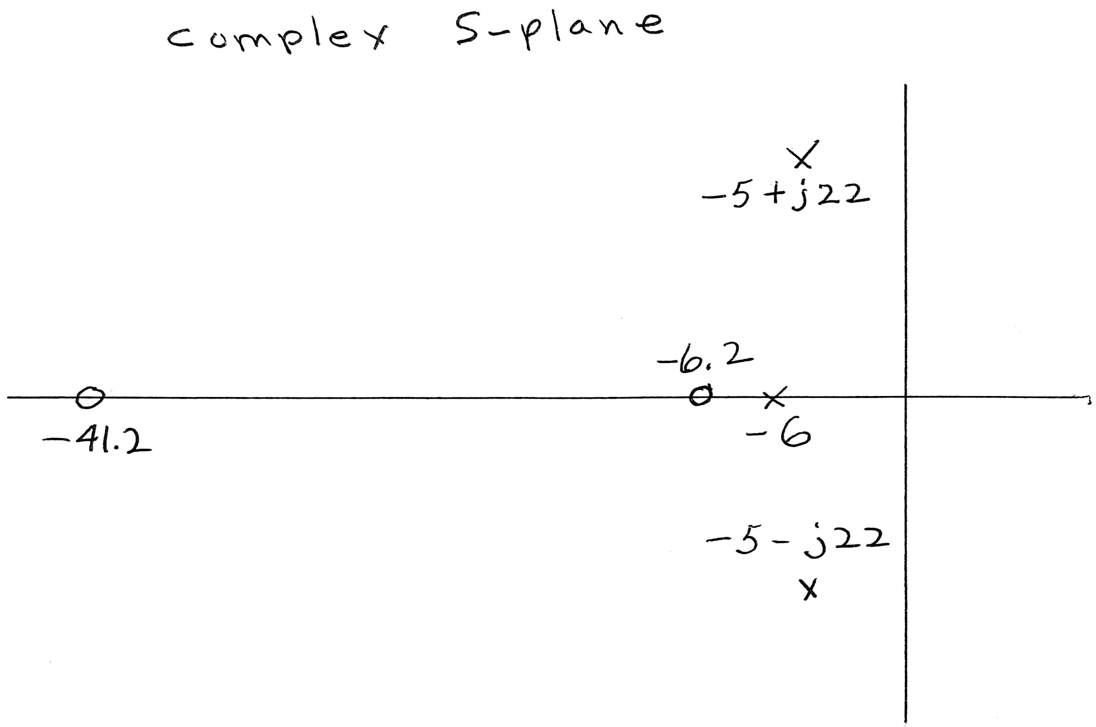
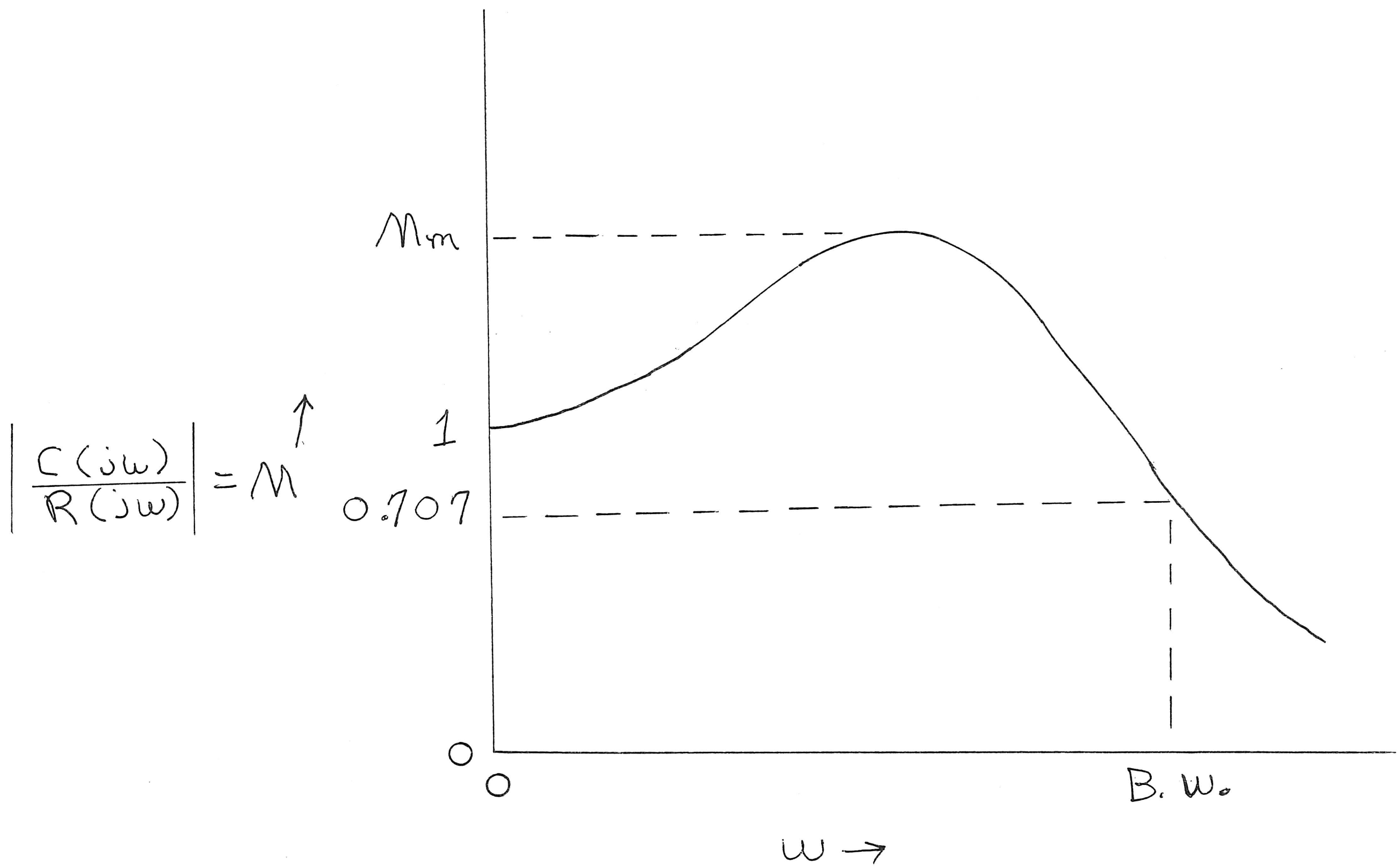


Figure 3



$\left|\frac{C(j\omega)}{R(j\omega)}\right| = M$
M_m
1
0.707
0
0
B. W.
$\omega \rightarrow$

Figure 4

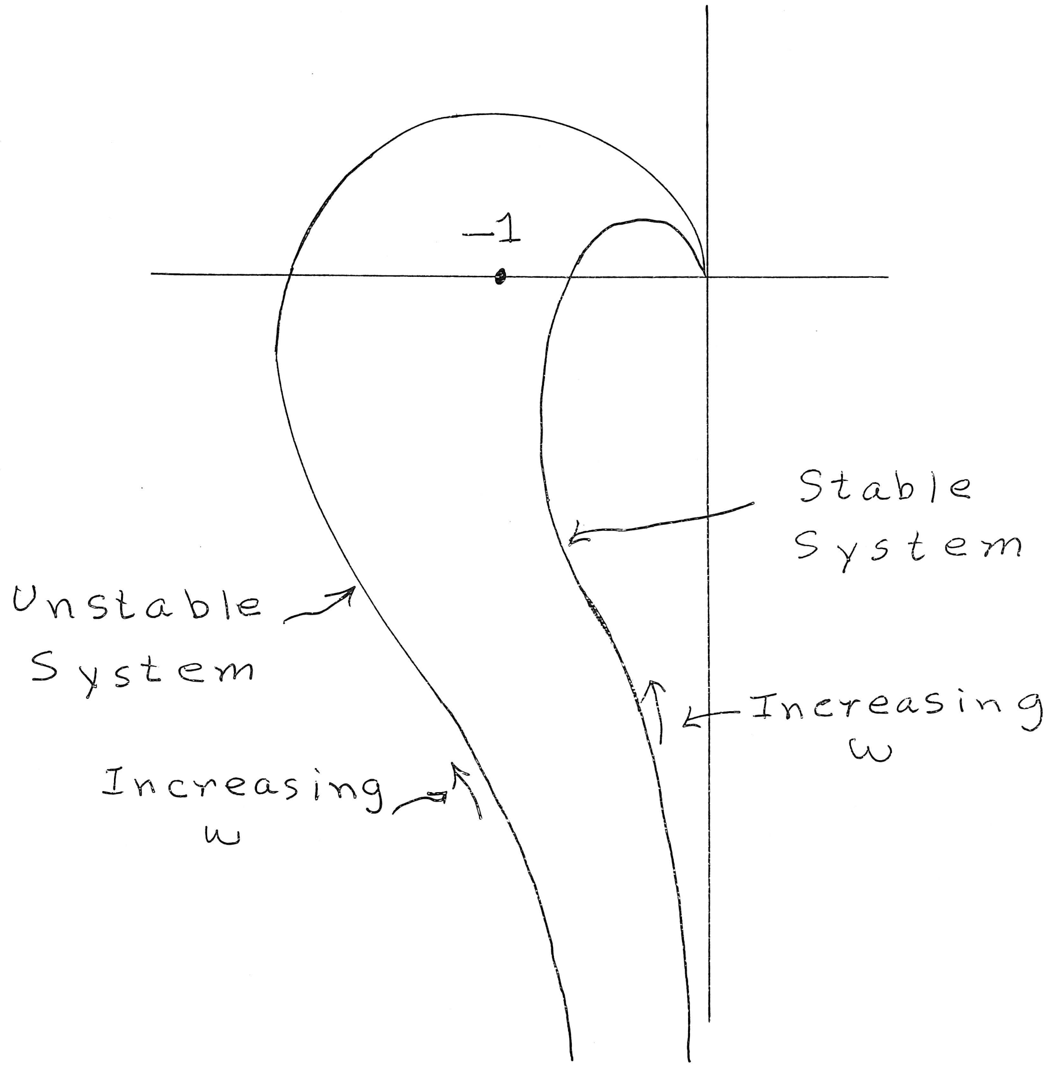


Figure 5

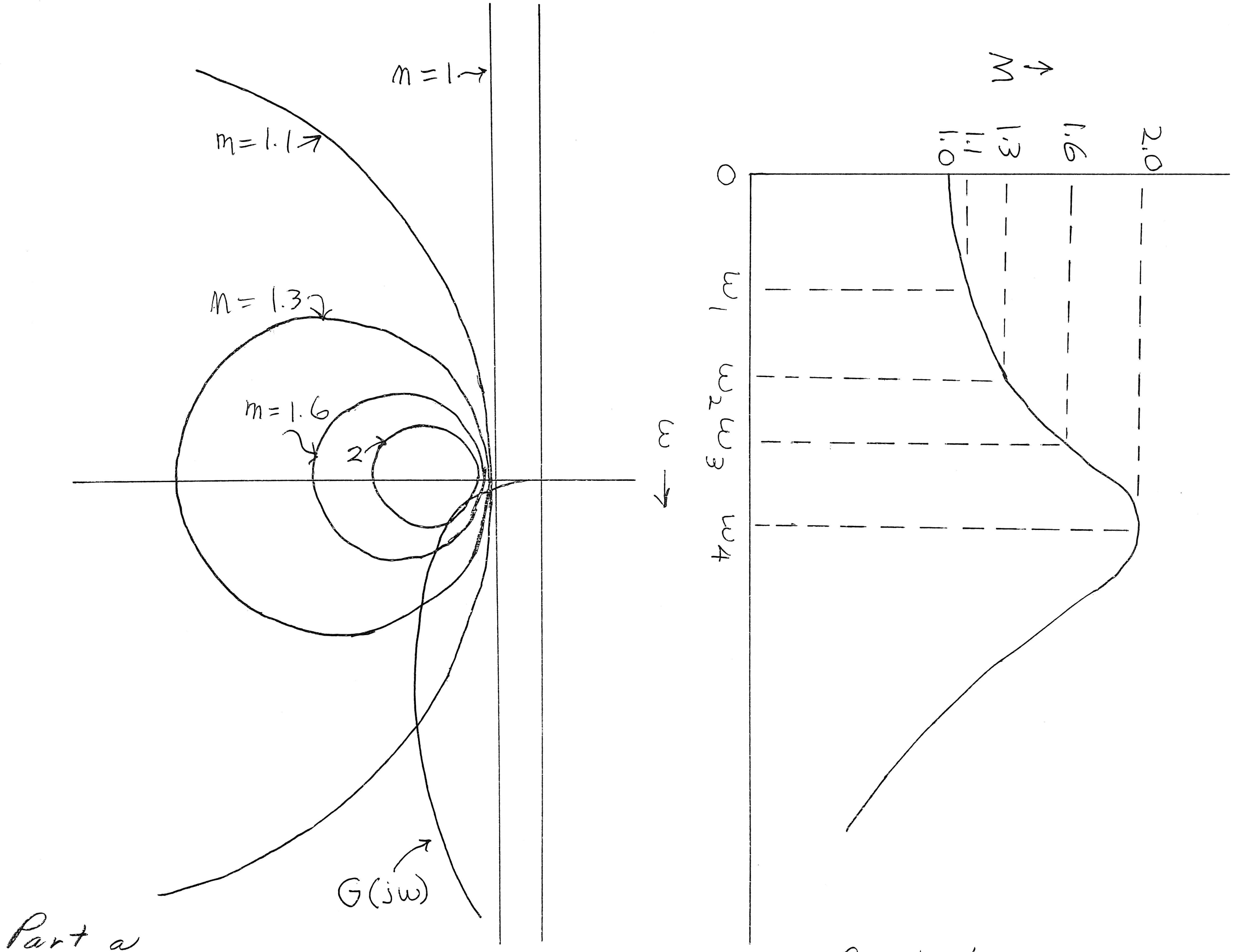


Figure 6

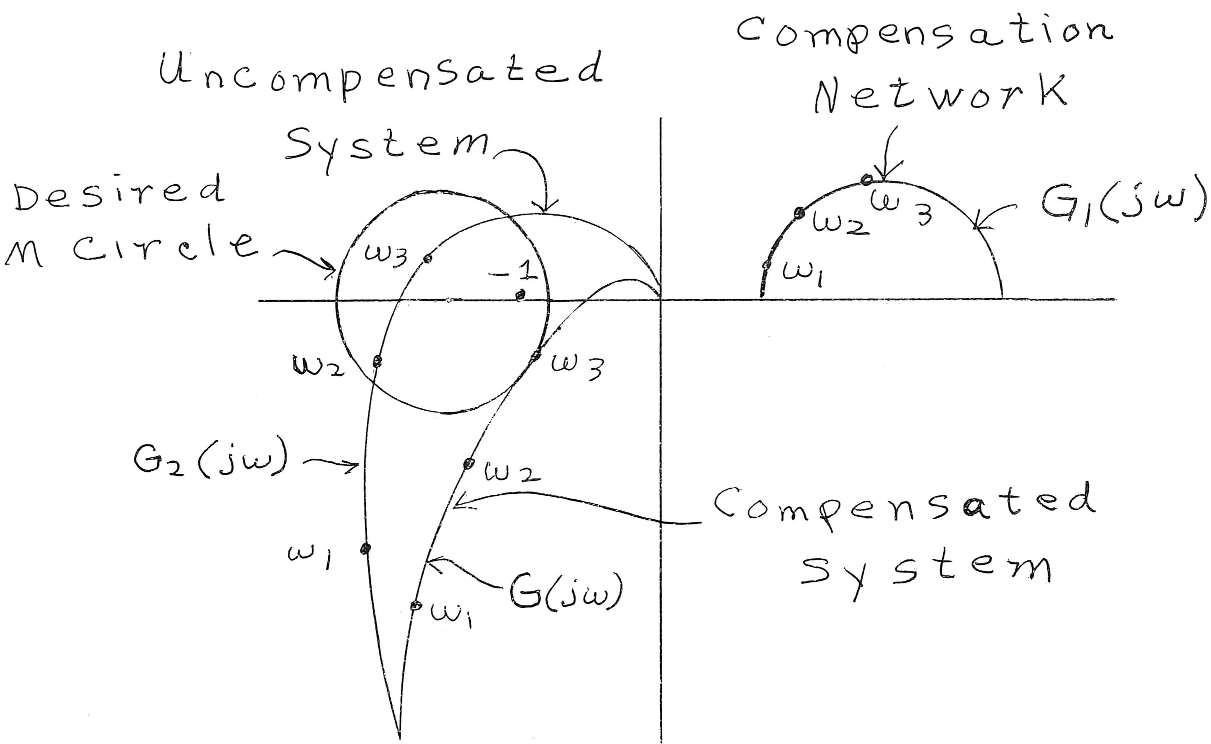

Uncompensated
System
Desired
M Circle
ω_3
-1
ω_2
ω_3
$G_2(j\omega)$
ω_2
ω_1
$G(j\omega)$
ω_1
Compensation
Network
ω_3
ω_2
ω_1
$G_1(j\omega)$
Compensated
System

Figure 7

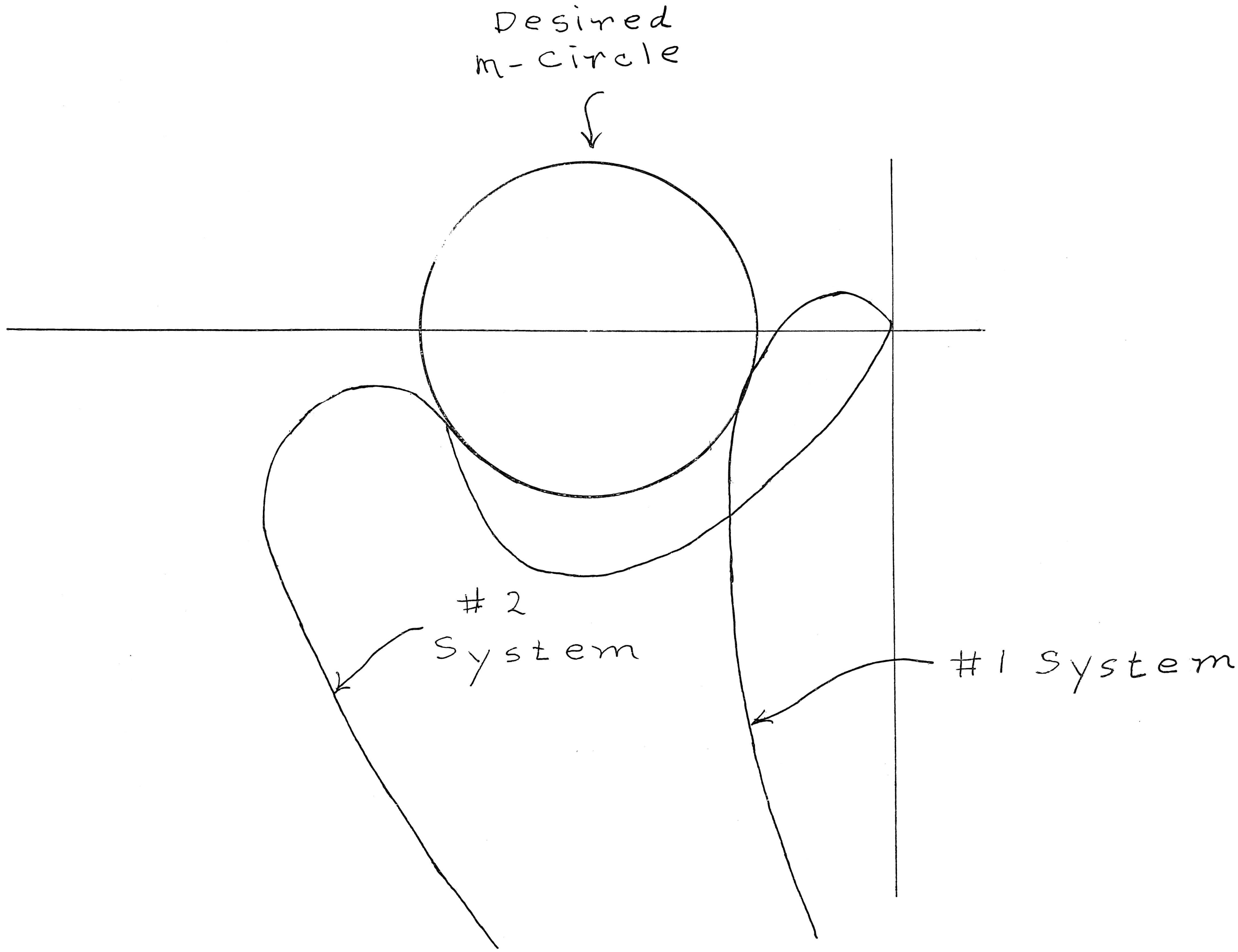


Figure 8

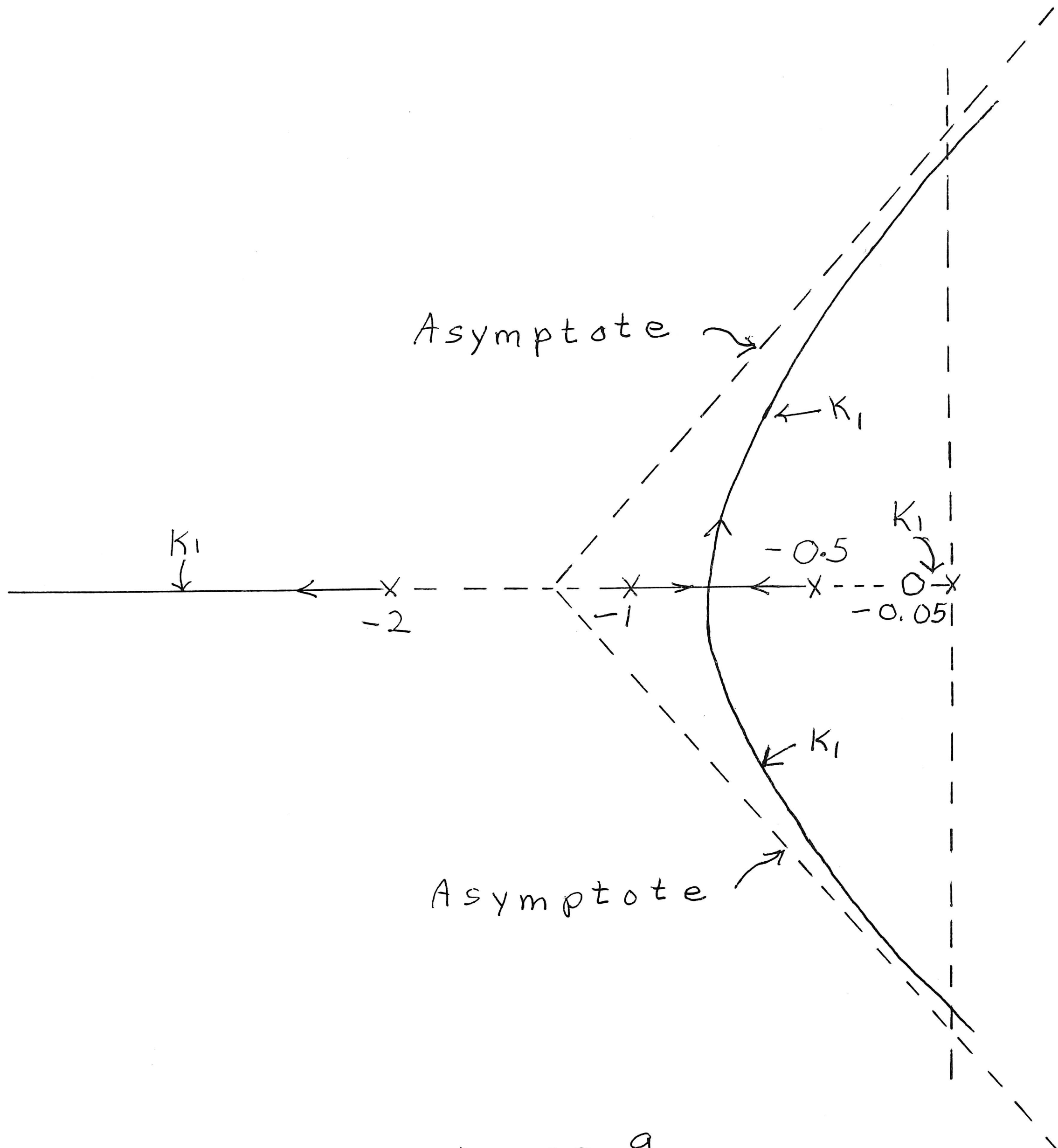


Figure 9

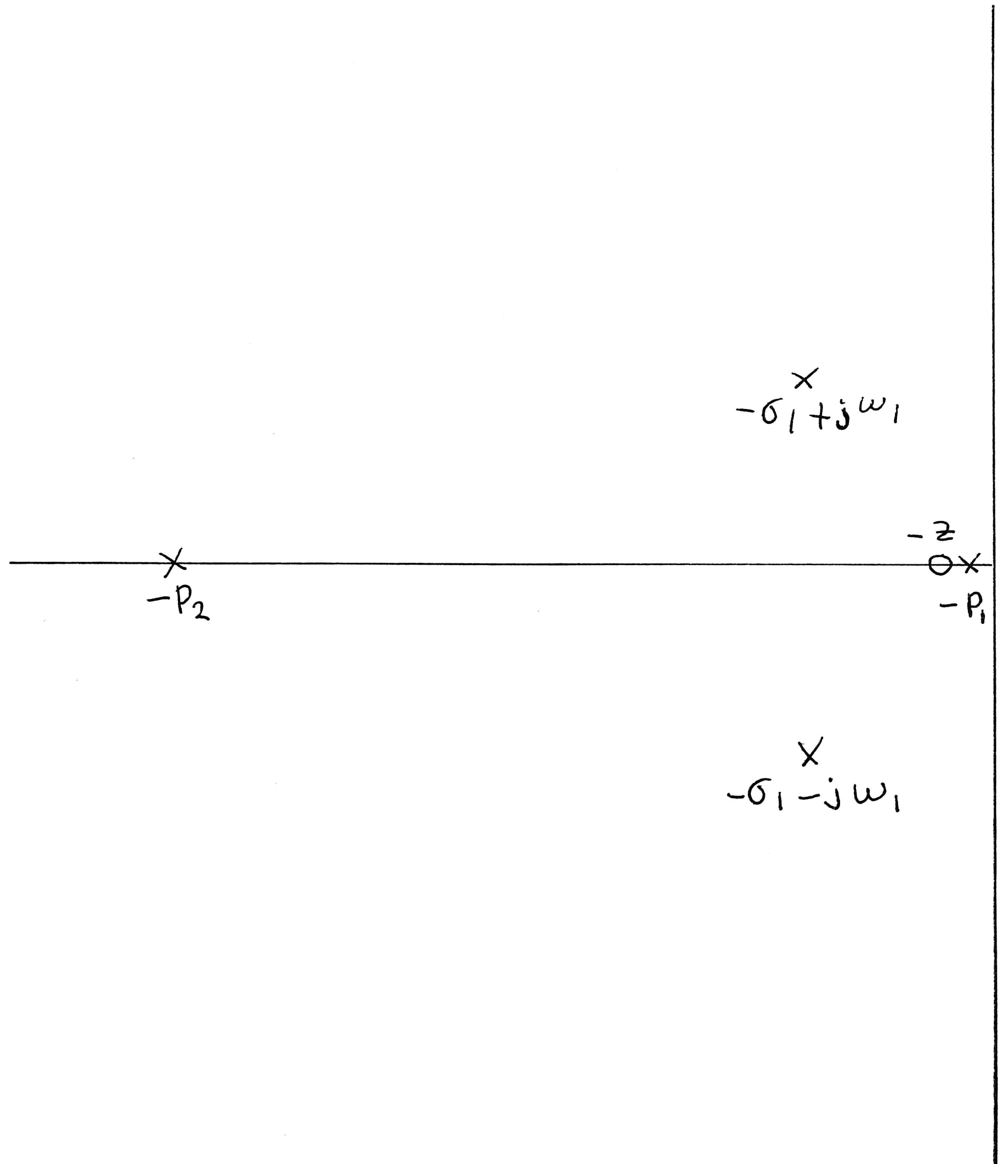


Figure 10

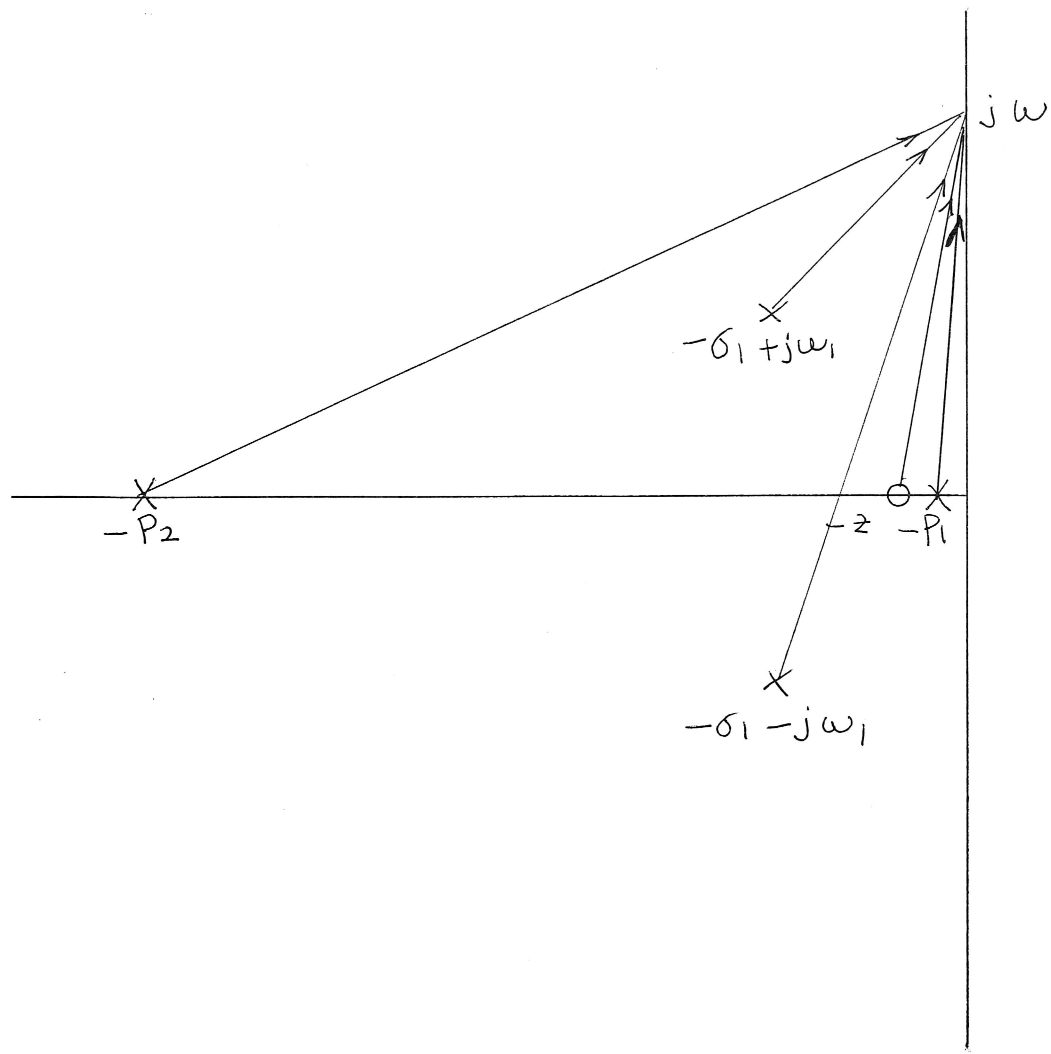


Figure 11

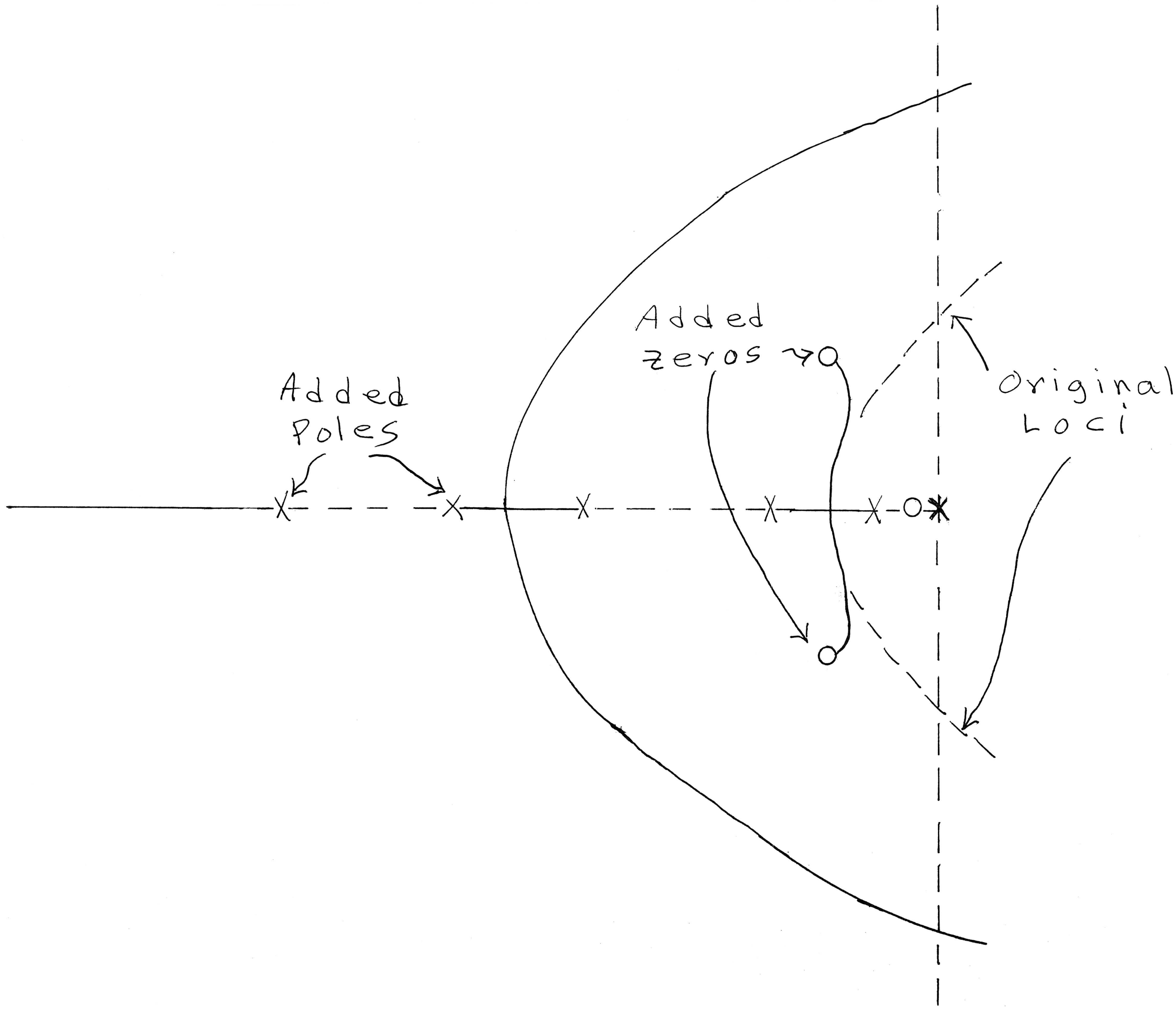


Figure 12

PUBLICATIONS OF THE ENGINEERING REPRINT SERIES

Copies of the complete list of publications may be secured from the Director of the Engineering Experiment Station, University of Missouri

Reprint No.

20. Autotransformer Betters Motor Phase Conversion by Joseph C. Hogan, Associate Professor of Electrical Engineering. Reprinted from Electrical World, Vol. 144, p. 120, October 17, 1955.

21. Sequence Summation Factors by Adrian Pauw, Professor of Civil Engineering. Reprinted from the Proceedings of the American Society of Civil Engineers. Vol. 81, Paper No. 763, August, 1955.

22. Pulsed Eddy Currents Gage Plating Thickness by Donald L. Waidelich, Associate Director, Engineering Experiment Station. Reprinted from Electronics, Vol. 28, p. 146, November, 1955.

23. Relay Protection for Lines Being Sleet-Melted by the Short-Circuit Method by J. C. Hogan, Associate Professor of Electrical Engineering and C. G. Pebler, Commonwealth Associates, Inc. Reprinted from Transactions of the American Institute of Electrical Engineers, Vol. 74, December, 1955.

24. Supplemental Irrigation....Careful Planning is Essential by Harry Rubey, Professor of Civil Engineering. Reprinted from What's New in Crops and Soils, Vol. 7, August-September, 1955.

25. Analysis of Single-Phase-to-Three-Phase Static Phase Converters by J. C. Hogan, Associate Professor of Electrical Engineering. Reprinted from Transactions of the American Institute of Electrical Engineers, Vol. 74, p. 403, January, 1956.

26. Enrollment and Incomes in Civil Engineering can be Increased by Harry Rubey, Professor of Civil Engineering. Reprinted from Journal of Engineering Education, Vol. 46, p. 236, November, 1955.

27. A Synthesis Procedure for Sampled-Data Systems by G. V. Lago, Associate Professor of Electrical Engineering. Reprinted from Proceedings of the National Electronics Conference, Vol. 11, p. 251, 1955.

28. Design of Optimum Phase-Shift Oscillators by Donald L. Waidelich, Associate Director, Engineering Experiment Station. Reprinted from Proceedings of the National Electronics Conference, Vol. 11, p. 222, 1955. This article also appeared in Electronics Equipment, Vol. 4, p. 38, April, 1956.

29. Investigation Concerning Polarization in Barium Titanate Ceramics by G. W. Marks, U. S. Navy Electronics Laboratory, Donald L. Waidelich, Associate Director Engineering Experiment Station, University of Missouri and L. A. Monson, U. S. Navy Electronics Laboratory. Reprinted from Transactions of the American Institute of Electrical Engineers, Vol. 75, Part I, p 469, 1956.

30. The Influence of Shank Area on the Tensile Impact Strength of Bolts by John Love, Jr., General Electric Company and O. A. Pringle, Associate Professor of Mechanical Engineering. Reprinted from Transactions of the American Society of Mechanical Engineers, Vol. 78, p 1489, October, 1956.

31. Measurement of Coating Thicknesses by Use of Pulsed Eddy Currents by Donald L. Waidelich, Associate Director, Engineering Experiment Station. Reprinted from Nondestructive Testing, Vol. 14, p 14, May-June 1956.

32. Head Losses in Storm Drain Junction Boxes by Horace W. Wood, Professor of Civil Engineering. Reprinted from Highway Research Board Proceedings, Vol. 35, p 177, 1956.

33. Stability of Laminar Flow in Curved Channels by Chia-Shun Yih, Associate Professor of Engineering Mechanics, University of Michigan and W. M. Sangster, Associate Professor of Civil Engineering, University of Missouri. Reprinted from The Philosophical Magazine, Volume 2, Eighth Series, Page 305, March 1957.

34. Viscosity of Suspensions of Spherical and Other Isodimensional Particles in Liquids by Andrew Pusheng Ting, Chemical Construction Corporation and Ralph H. Luebbers, Professor of Chemical Engineering, University of Missouri. Reprinted from the American Institute of Chemical Engineers Journal, Volume 3, Page 111, March, 1957.

35. Irrigation-Drainage-Climatology for Flat Humid Land by Harry Rubey Professor Emeritus of Civil Engineering. Reprinted from the Proceedings of the American Society of Civil Engineers, Volume 83, Paper No. 1253, May 1957.

36. The Impedance of a Coil Near a Conductor by D. L. Waidelich, Associate Director, Engineering Experiment Station and C. J. Renken, Jr., Argonne National Laboratory. Reprinted from the Proceedings of the National Electronics Conference, Volume 12, Page 188, 1956.

37. Automatic Control - The Fundamentals by Gladwyn Lago, Associate Professor of Electrical Engineering. Reprinted from the Proceedings of the Second Annual Conference on Automatic Control, University of Oklahoma, April 29-30, 1957, Page 9.

*Out of Print.