Another Reason to Eschew Pole-Zero Cancellation

Robert N. Clark

ABSTRACT: It is shown that, if series compensation by pole-zero cancellation is used in a system that is linear for small signals but has a large signal nonlinearity in the loop, an observable mode, corresponding to the canceled pole, will appear in the system response when large transient signals occur in the loop.

Introduction

In the design of series compensators for servomechanism-like systems, one may be tempted to use pole-zero cancellation to achieve a desired input-output transfer function. Some textbooks warn against this temptation and provide some reasons for these warnings [1]-[4]. These reasons are based on the linear dynamic behavior of the closed-loop system as it is affected, for example, by noise in the error channel or by inevitable inexact pole-zero cancellation. We are cautioned against right-half-plane cancellation, against canceling too many plant poles with compensator zeros, and so forth.

In this paper, a further reason to avoid series compensation by pole-zero cancellation is demonstrated. This reason is based on the dynamic behavior of a system that is linear for small amplitude signals in its control loop but is nonlinear for larger signals, a common situation in feedback control systems.

Example

An example will illustrate the point of this paper. Consider the simple system shown in Fig. 1. The nonlinear element is a saturating device, with unity gain in its linear range.

\[ m(t) = e(t) \quad \text{for} \quad |e(t)| < L \]
\[ = L \quad \text{for} \quad e(t) > L \]
\[ = -L \quad \text{for} \quad e(t) < -L \]  

In the linear mode of behavior, the troublesome pole of the plant transfer function at \( s = -\alpha \) is canceled by a zero of the compensator transfer function, while the pole at \( s = -P \) is placed to satisfy the response requirements of the system.

It is instructive to write a state-space model for this system, valid for \( e(t) \) remaining in its linear range. One such model follows, where \( x \) is a \( 3 \times 1 \) matrix representing the state vector; \( y \) the scalar output; \( u \) the scalar input; \( A, B, \) and \( C \) appropriate size matrices; and superscript \( T \) indicates transpose.

\[
\dot{x} = Ax + Bu, \quad y = Cx \quad (2)
\]

\[
A = \begin{bmatrix} -P & 0 & K_c(P - \alpha) \\ K_p & -\alpha & -K_p \\ 0 & 1 & 0 \end{bmatrix}
\]

\[
B^T = [-K_c(P - \alpha), K_p, 0]
\]

\[
C = [0, 0, 1] 
\]

State variable \( x_2(t) \) is associated with the compensator, while \( x_3(t) \) and \( x_2(t) \) are the plant state variables with output \( y(t) \) equal to \( x_3(t) \). Taking the Laplace transform of Eq. (2) and rearranging to get transfer functions yields the following, where \( x_0 \) equals \( x(0) \), the initial value of the state vector.

\[
Y(s) = C[sI - A]^{-1}BU(s)
\]
\[
+ C[sI - A]^{-1}x_0 
\]

\[
x_0^T = [x_{10}, x_{20}, x_{30}] 
\]

Figure 2 represents Eq. (4) in block diagram form and clearly shows the distinction between the zero state response and the zero-input response of the system. We now substitute the following illustrative numerical values into \( A, B, \) and \( C \):

\[
K_c = 10 \\
\alpha = 1 \\
K_p = 600 \\
P = 100
\]

With these values, the two transfer-function matrices in Fig. 2 become

\[
\frac{Y(s)}{U(s)} \bigg|_{s = 0} = \frac{1}{D(s)} [6000 (s + 100)]
\]

\[
\frac{Y(s)}{U(s)} \bigg|_{s = 0} = \frac{6000}{{(s^2 + 101s + 1000)}}
\]

\[
\frac{Y(s)}{U(s)} \bigg|_{s = 0} = \frac{6000}{{s^3 + 100s^2 + 6000}} 
\]

where \( D(s) = (s + 1)(s^2 + 100s + 6000) \).

We note that the input-output transfer function (or the zero state response transfer function) has a pole-zero cancellation due to the pole-zero cancellation in the forward path of the system. This indicates that the closed-loop system is uncontrollable, the mode having eigenvalue equal to \(-1\) being the uncontrollable mode, which does not appear in the zero state response of the system. However, the closed-loop system is observable from the output so that all three modes appear in the zero-input response. The troublesome mode (with eigenvalue equal to \(-1\)) appears in those components of the zero-input response due to \( x_{10} \) and \( x_{20} \), as is evident from Eq. (7). Nevertheless, if \( x_{10} \) and \( x_{20} \) are zero, and if \( e(t) \) does not exceed its linear range, the canceled mode will not appear in...

Fig. 1. Block diagram for illustrative system.

Robert N. Clark is Professor of Electrical Engineering and Adjunct Professor of Aeronautics and Astronautics, University of Washington (FS-10), Seattle, WA 98195. He is also a consulting analyst at Boeing Aerospace Company, Seattle, WA 98124. The author observed the phenomenon described here during development work at the Boeing Aerospace Company. The simulations were done using the TUTSIM computer simulation package.

April 1988
Fig. 2. Zero state response and zero-input response transfer-function matrices for illustrative system.

Fig. 3. Error responses for step inputs of one, two, and four units.

Fig. 4. Error response for step input of four units.

y(t). An example of such a response is given in Fig. 3, the lower curve, which shows the error signal e(t) obtained in response to a step input u(t) of unit amplitude and with zero initial state. The error promptly goes to (essentially) zero in about 0.09 sec, as is expected from the linear analysis.

If the saturation limit of the nonlinear device L is 10 and the step input magnitude exceeds unity, the signal m(t) will go into saturation and remain there until the feedback signal can manage to drive e(t) below its saturation level. Figure 3, the middle curve, shows the error response for a step of amplitude 2. In this case, m(t) comes out of saturation at t = 0.0061 sec, and remains out of saturation subsequently. The upper curve in Fig. 3 is the error response for an input step of amplitude 4, and, in this case, m(t) comes out of saturation at t = 0.0115 sec, and it too remains out of saturation. In both cases, where m(t) has entered saturation temporarily, the error does not approach zero promptly but appears in Fig. 3 to assume almost a constant, nonzero value in spite of the integrating property of the plant. What is happening here is made apparent if the transient is observed for longer than 0.2 sec. Figure 4 takes the error response out to 1 sec, which shows that the error actually converges toward zero exponentially, with a time constant of 1 sec—the time constant of the 'canceled' mode. When the system comes out of saturation at t = 0.0115 sec, the state vector is nonzero:

\[
x(0.0115)^T = [-25.91, 68.605, 0.39521]
\]  

(8)

The response of the system for t > 0.0115 sec is the same as it would be if both the step input and the nonzero state given by Eq. (8) were applied together at t = 0.0115 sec, hence, the observable mode with eigenvalue equal to -1 appears because it is present in the 'zero-input' part of the response.

Therefore, it is preferable to employ a compensation scheme that does not depend on pole-zero cancellation in the forward path. There are many possibilities, one is the parallel compensation scheme shown in Fig. 5. The zero state transfer function in this case is shown here, and the error in response to a step input of amplitude 4 is shown in Fig. 6.

\[
\frac{Y(s)}{U(s)} = \frac{3600(s + 200)}{(s + 86.91)(s^2 + 114s + 8284)}
\]  

(9)

Here m(t) is in saturation until t = 0.0171 sec, but all of the modes of the closed-loop system have small time constants, compared to 1 sec, so the error converges promptly to zero.

References


Robert N. Clark received the B.S.E.E. (1950) and the M.S.E.E. (1951) degrees from the University of Michigan. He was with Honeywell, Inc. Since 1957, he has been on the Faculty of Electrical Engineering at the University of Washington teaching automatic control, mechanics, circuits, and systems. His research and consulting experience lies in the application of control theory to engineering problems; he is the author of an introductory textbook on automatic control. From 1966 to 1968, he was at Stanford University as a NSF Fellow and a Lecturer in Electrical Engineering, earning the Ph.D. in 1969.