621-503. 5

On the Optimizing Control by a Continuous Correlation Method*

By Seizo FUJII** and Naomi KANDA***

This paper is concerned with the theoretical analysis of the optimizing control system by a continuous correlation method.

First, detailed studies are carried out on the patterns and the instability of the step response of the system to a displacement of the extremum depending upon the initial conditions, the parameters of the system, and the magnitude of the displacement.

Second, the integral of the absolute error and the hunting loss are considered as the performance criteria for the transient and steady state, respectively, of the system contaminated by a noise. The contributions of each element of the system to these criteria are clarified.

Finally, a reasonable design method is proposed on the basis of the results obtained.

Introduction

Optimizing control, as is well known, is one which finds automatically the extremum of a performance criterion (e.g. efficiency, profit, loss) of a certain system and keeps it in the immediate neighbourhood of the extremum in spite of changes in circumstances. It is classified broadly into the model method and the direct method. The continuous correlation method which is included in the direct method is treated in this paper. Although several researches^{(2)~(12)} have been reported on this method since Draper and Li⁽¹⁾ first contrived it, little is known about the most basic matters such as the system stability, the transient response characteristics, because of the difficulties in analysis due to the nonlinearity of the equations describing the systems.

In this paper, the most basic system of single variable optimizing control is considered, and the response characteristics and a reasonable design method are investigated in detail.

1. Optimizing control system

Consider a single variable system which contains an extremum (assumed as the minimum hereafter). A sinusoidal test signal is superimposed on the input of the controlled system. Here, when correlating its output with the sinusoidal signal (correlation

signal), which is the same frequency as that of test signal, by multiplication and integration, if the input exists at the optimum, the correlation vanishes; and if it is separate from the optimum, the correlation is positive or negative depending on whether the offset is positive or negative. The optimizing control system is constructed by introducing the negative feedback loop in order to reduce the correlation. Figure 1 shows a block diagram of the basic configuration of the optimizing control system by the method stated above. By multiplier (3) and integrator (5) in the figure, the correlation is always observed and the continuous control action is performed. The output of the controlled system (1) consists of not only a signal necessary for minimum-seeking but also constant component plus higher harmonics components, and noise in general. Therefore a band-pass filter (2) and a low-pass filter (4) are inserted before and after the multiplier. In many cases, instead of the low-pass filter, the integrator includes its action. On the other hand, as a substitute for the band-pass filter,

Fig. 1 Basic configuration of the optimizing control system by a continuous correlation method

* Received 24th July, 1967.

** Assistant Professor, Faculty of Engineering, Nagoya University, Chikusa-ku, Nagoya.

*** Lecturer, Faculty of Engineering, Nagoya University.

a high-pass filter can be used. Besides these, in order for the control system to act correctly, it is necessary that the phase of the fundamental component of the band-pass filter output resulting from test signal coincide with that of correlation signal. Furthermore, as a test signal and a correlation signal, an arbitrary periodic signal can be used including a sine wave and a square wave. In practice, a square wave can be simply generated and is most often used.

According to the above, the system shown in Fig. 2 is considered in this paper. That is, a controlled system is characterized as the parabola and first-order lag, the integrator includes the action of the low-pass filter, and the band-pass filter is substituted by the high-pass filter (s denotes Laplace operator in the figure).

Referring to Fig. 2, the following relations between variables of the system are given:

$$\left. \begin{aligned} y^* &= k(x_{\text{opt}} - x + E \sin \omega_p t)^2 + y_{\text{min}} \\ T(dy/dt) + y &= y^* \\ T_f(dy_f/dt) + y_f &= T_f(dy/dt) + T_f(dn/dt) \\ dx/dt &= K y_f A \sin(\omega_p t + \theta) \\ \theta &= \text{Arc tan}\{1/(\omega_p T_f)\} - \text{Arc tan}(\omega_p T) \end{aligned} \right\} \quad (1)$$

In Eq. (1) x_{opt} , y_{min} and n denote the optimum value of the input, the minimum of the output and the noise, respectively. Equation (1) becomes Eq. (2) by the introduction of dimensionless variables:

$$\left. \begin{aligned} N(dY/d\tau) + Y &= (U_1 - X + \sin \tau)^2 + U_2 \\ M(dY_f/d\tau) + Y_f &= M(dY/d\tau) + M(dV/d\tau) \\ dX/d\tau &= B Y_f \sin(\tau + \theta) \\ \theta &= \text{Arc tan}(1/M) - \text{Arc tan } N \end{aligned} \right\} \quad (2)$$

where the dimensionless variables are defined as follows:

$$\left. \begin{aligned} N &= \omega_p T, & M &= \omega_p T_f, & \tau &= \omega_p t, \\ X &= x/E, & Y &= y/(kE^2), & Y_f &= y_f/(kE^2), \\ V &= n/(kE^2), & U_1 &= x_{\text{opt}}/E, & U_2 &= y_{\text{min}}/(kE^2), \\ B &= kKA E/\omega_p. \end{aligned} \right\} \quad (3)$$

Expression of the system by Eq. (2) is together

Fig. 2 Block diagram of the optimizing control system to be considered (symbols in brackets show the dimensionless representation)

shown with bracketed symbols in Fig. 2 (s' denotes s/ω_p in the figure). The following arguments will be made mostly referring to the dimensionless expression Eq. (2).

2. Analysis of system response

2.1 Basic characteristics of response

Consider the most basic system in which the dynamics of the controlled system, high-pass filter and noise is neglected. In this section, the characteristics of the response due to the step disturbance are investigated. The step disturbance represents, as shown in Fig. 3, the displacement of the static characteristic curve (displacement of the optimum value of input and the minimum value of output, U_1 and U_2). However, it is assumed that the curvature of the static characteristic curve is unchanged. Moreover, it is assumed for convenience of calculation that the test signal and the correlation signal are both square waves, and the phase difference between them vanishes since the dynamics is neglected (if sine wave is assumed, the exact solution cannot be obtained).

2.1.1 System equation and solution

Replacing the sine wave in Eq. (2) by the square wave, the dimensionless equation of the system becomes

$$dZ/d\tau = B(Z-1)^2 + BU_2, \quad (4)$$

where for $2n\pi \leq \tau < (2n+1)\pi$, ($n=0, 1, 2$),

$$\left. \begin{aligned} Z &= X - U_1, \text{ test signal: } E, \\ &\text{correlation signal: } A, \end{aligned} \right\} \quad (5.a)$$

for $(2n+1)\pi \leq \tau < 2(n+1)\pi$,

$$\left. \begin{aligned} Z &= U_1 - X, \text{ test signal: } -E, \\ &\text{correlation signal: } -A. \end{aligned} \right\} \quad (5.b)$$

After this, the time intervals corresponding to Eqs. (5.a) and (5.b) are called the intervals I_i and II_i for $n=i$, respectively. Now, by putting $Z=Z_0$ at

Fig. 3 Displacement of the static characteristic of the controlled system due to the step disturbance (displacement of the optimum of input and the minimum of output)

the time τ_0 when disturbances U_1 and U_2 occur, the solutions of Eq. (4) are given as follows:

(1) for $U_2 \geq 0$ (upward displacement of static characteristics),

$$Z = 1 + \sqrt{U_2} \tan[B\sqrt{U_2}(\tau - \tau_0)] + \text{Arc tan}\{(Z_0 - 1)/\sqrt{U_2}\}, \dots\dots\dots(6)$$

(2) for $U_2 \leq 0$ (downward displacement),

(2-1) if $U_2 \geq -(Z_0 - 1)^2$,

$$Z = 1 + \sqrt{-U_2} \coth[\text{arc coth}\{(Z_0 - 1)/\sqrt{-U_2}\} - B\sqrt{-U_2}(\tau - \tau_0)], \dots\dots\dots(7.a)$$

(2-2) if $U_2 \leq -(Z_0 - 1)^2$,

$$Z = 1 + \sqrt{-U_2} \tanh[\text{arc tanh}\{(Z_0 - 1)/\sqrt{-U_2}\} - B\sqrt{-U_2}(\tau - \tau_0)], \dots\dots\dots(7.b)$$

As seen from Eqs. (5.a) and (5.b), Z is separately defined for every time interval of π , and hence Eqs. (6) and (7) express partially the solutions in the interval including the initial time τ_0 . Time τ_0 when the disturbance occurs can be limited to $0 \leq \tau_0 \leq 2\pi$ without loss of generality. Solution over several intervals X is given by connecting every solution by equating the final value of the interval to the initial value of the following interval.

2.1.2 Stability of system

It is noted from Eqs. (6) and (7.a) that, depending on the values of the system parameters, solution Z (or X) may possibly diverge to infinity within the interval including τ_0 ; that is, the system has a possibility to show the unstable response with the so-called finite escape time (hereafter, the term "unstable" is used to represent such a phenomenon). The conditions for solution not to diverge within

the interval including τ_0 , are derived from those that the final value of the interval remains limited, and are given as follows:

$$(1) \left. \begin{aligned} Z_0 < 1 + \sqrt{U_2} \cot B\sqrt{U_2}(p\pi - \tau_0), \\ 0 \leq B\sqrt{U_2} < 1, \quad U_2 \geq 0, \end{aligned} \right\} \dots\dots\dots(8)$$

$$(2) \left. \begin{aligned} Z_0 < 1 + \sqrt{-U_2} \coth B\sqrt{-U_2}(p\pi - \tau_0), \\ U_2 \leq 0, \end{aligned} \right\} \dots\dots\dots(9)$$

where $p=1$ for the interval I_0 and $p=2$ for the interval II_0 . For the given values of B and U_2 , the above two inequalities define the stable region, in which the solution is stable, of the initial deviation Z_0 ($U_1 - X_0$ or $X_0 - U_1$) as a function of τ_0 . The boundaries of the stable region of the initial deviation ($U_1 - X_0$) against τ_0 are illustrated in Fig. 4 as the curves [(1), (1')] and [(2), (2')] corresponding to Eqs. (8) and (9), respectively.

Next, in order for the solution to be stable in the interval following the interval including τ_0 , it is necessary that the initial value of the following interval exist in the stable region defined by Eq. (8) or (9), according to the sign of U_2 . The boundaries of ($U_1 - X_0$) corresponding to Eqs. (8) and (9) against τ_0 are also illustrated in Fig. 4 as the curves [(3), (3')] and [(4), (4')], respectively. As the result of the above, it becomes apparent that the regions of ($U_1 - X_0$) with stable response are the shaded areas in Fig. 4. In addition, the regions of ($U_1 - X_0$) where the response is stable for arbitrary τ_0 (whenever a disturbance may occur) are in the range between A and B or between A' and B', as are represented by the arrows, according to the sign of U_2 . These results are obtained by substituting $\tau_0=0$ and $\tau_0=\pi$ into Eqs. (8) and (9):

$$(1) \text{ for } U_2 \geq 0, \\ |Z_0| = |U_1 - X_0| < 1 + \sqrt{U_2} \cot(\pi B\sqrt{U_2}), \left. \begin{aligned} 0 \leq B\sqrt{U_2} < 1, \end{aligned} \right\} \dots\dots\dots(10)$$

$$(2) \text{ for } U_2 \leq 0, \\ |Z_0| = |U_1 - X_0| < 1 + \sqrt{-U_2} \coth(\pi B\sqrt{-U_2}). \dots\dots\dots(11)$$

Thus, so far as the above conditions are satisfied, the responses are always stable. Figure 5 shows the relations between the magnitudes of disturbance U_2 and $|U_1 - X_0|$ satisfying the above

Fig. 4 Relationship between the initial deviation ($U_1 - X_0$) and the time when the disturbance occurs τ_0 , in order for the response to be stable

Fig. 5 Relationship between the initial deviation $|U_1 - X_0|$ and the disturbance U_2 in order for the response to be stable

ditions for various values of loop gain B . The shaded areas in the figure indicate the stable regions. As seen in the figure, in case of $U_2 > 0$ the system is ready to fall into instability as compared with the case of $U_2 < 0$. This trend is significant especially for a greater value of B . Some typical response patterns are illustrated in Fig. 6. Curve (1) represents an unstable response, and curves (2)~(5) are stable ones. However, curve (2) shows a remarkable reverse movement. It can be seen from Eqs. (6) and (7) that this phenomenon occurs. Curves (2) and (3) indicate that the different responses are caused for different values of τ_0 in spite of the same amount of initial deviation. It is readily understood that the above reverse movement becomes the most significant in the case where a disturbance occurs at the boundary of time interval ($\tau_0 = 0$ or π). Furthermore, if the reverse movement is raised so extremes due to the square characteristic of the controlled system, the response diverges into infinity.

Now, in this section it has been stated that this system has a remarkable character in that the response diverges into infinity within a finite time if either Eq. (8) or (9) is unsatisfied. This character originates, however, in the fact that the static characteristic curve of the controlled system has been assumed to be parabola, and if it has been assumed to be piecewise linear such as $|X^*|$, there is no solution to diverge within a finite time. Even in such a case, it is a matter of course that the pattern of the response is qualitatively similar, and that large deviation can be caused under certain conditions. Still, it is reasonable to presume that the solution diverging within a finite time will exist, if $\alpha > 1$, when $(X^*)^*$ is assumed.

2.1.3 Existence of periodic solution and convergence of stable solution

Representing the solution Z at time τ as $Z(\tau)$, it can be proved from Eq. (6) or (7) that there is a solution satisfying $Z(\tau) = Z(\tau + 2\pi)$ under the

Fig. 6 Several typical responses in the basic system ($B=0.1$, $U_2=0$)

condition (8) or (9). Assuming $\tau_0=0$, for convenience, the initial value of the periodic solution Z_0 is given as follows:

$$(1) \text{ for } U_2 \geq 0, \\ Z_0 = Z(0) = \{\sqrt{U_2} \cot(\pi B \sqrt{U_2}) - \sqrt{1+U_2} \operatorname{cosec}^2(\pi B \sqrt{U_2}), \dots \dots \dots (12)$$

$$(2) \text{ for } U_2 \leq 0, \\ Z_0 = Z(0) = \{-\sqrt{-U_2} \coth(\pi B \sqrt{-U_2}) - \sqrt{1-U_2} \operatorname{cosech}^2(\pi B \sqrt{-U_2}), \dots \dots \dots (13)$$

Representing the periodic solution as Z_s , it can also be proved that $|Z(\tau) - Z_s(\tau)| > |Z(\tau + 2\pi) - Z_s(\tau + 2\pi)|$, and that all solutions approach periodic solution as time τ increases. An example of the periodic solution is illustrated together in Fig. 6 [curve (5)]. As seen in the figure, the periodic solution is no more than a periodical fluctuation in the neighbourhood of the optimum. So far as the condition of stability is satisfied, the system achieves the periodical steady state after the transient behavior, and the object of control can be attained.

2.2 Effect of high-pass filter

When the first order high-pass filter is inserted into the system discussed in the previous section, the system equations are given in the following dimensionless form

$$\left. \begin{aligned} Y &= (Z-1)^2 + U_2, \\ M(dY_f/d\tau) + Y_f &= M(dY/d\tau), \\ dZ/d\tau &= BY_f, \end{aligned} \right\} \dots \dots \dots (14)$$

using the same symbols as in Fig. 2.

Integrating the above equations, the following equivalent equations are obtained:

$$Y = \{Z-1-1/(2BM)\}^2 + c, \quad dZ/d\tau = BY_f, \quad \dots (14)'$$

where c denotes the integral constant.

Equation (14) is a second order system, and therefore its solution has two integral constants determined by the initial values of Z and Y_f . That is, applying the suffix 0 to the value at time τ_0 ,

$$Y_{f0} = \{Z_0-1-1/(2BM)\}^2 + c. \dots \dots \dots (15)$$

From Eq. (14)', the following solutions are obtained:

$$(1) \text{ for } c \geq 0, \\ Z = \sqrt{c} \tan[B\sqrt{c}(\tau-\tau_0) + \operatorname{Arc tan}\{[Z_0-1-1/(2BM)]/\sqrt{c}\}] + [1+1/(2BM)], \dots \dots \dots (16)$$

$$(2) \text{ for } c \leq 0, \\ (2-1) \text{ if } c \geq -[Z_0-1-1/(2BM)]^2, \\ Z = \sqrt{-c} \coth[\operatorname{arccoth}\{[Z_0-1-1/(2BM)]/\sqrt{-c}\} - B\sqrt{-c}(\tau-\tau_0)] + [1+1/(2BM)], \dots \dots (17 \cdot a)$$

$$(2-2) \text{ if } c \leq -[Z_0-1-1/(2BM)]^2, \\ Z = \sqrt{-c} \tanh[\operatorname{arctanh}\{[Z_0-1-1/(2BM)]/\sqrt{-c}\} - B\sqrt{-c}(\tau-\tau_0)] + [1+1/(2BM)], \dots \dots (17 \cdot b)$$

Similarly to the case of Eqs. (6) and (7), the above equations express partially the solutions in

the interval including τ_0 , and the solution over several intervals is given by connecting all solutions at the boundaries of intervals. In this case, however, it is necessary to satisfy the following two conditions that: at the boundaries of intervals, the output of the integrator is continuous and the step displacements of the input and output of high-pass filter are identical (note that the square wave changes its sign at the boundary of interval). Representing the final value of interval as i and the initial value of the interval as $i+1$ with double suffix (i, e) and ($i+1, s$) respectively, the above two conditions can be described as follows:

$$\left. \begin{aligned} X_{i+1,s} &= X_{i,e} \\ (Y_f)_{i+1,s} &= (Y_f)_{i,e} + (Y)_{i+1,s} - (Y)_{i,e} \\ &= (Y_f)_{i,e} - 2Z_{i+1,s} + 2Z_{i,e} \end{aligned} \right\} \dots\dots\dots (18)$$

Moreover, the condition for the solutions (16) and (17-a) not to diverge in the interval including τ_0 can be obtained by the same procedure as the previous one. That is, the condition for the response to be stable for arbitrary τ_0 is given as follows:

$$\left. \begin{aligned} (1) \quad c' &\geq 0, \\ |Z_0| &< 1 + 1/(2BM) + \sqrt{c'} \cot(\pi B \sqrt{c'}), \\ c' &= Y_{f0} - \{ |Z_0| - 1 - 1/(2BM) \}^2, \dots\dots\dots (19) \\ 0 &< B \sqrt{c'} < 1, \end{aligned} \right\}$$

$$(2) \quad c' \leq 0, \\ |Z_0| < 1 + 1/(2BM) + \sqrt{-c'} \coth(\pi B \sqrt{-c'}). \dots (20)$$

Furthermore, the periodic solution is given as follows: by use of periodicity conditions $Y_f(0) = Y_f(2\pi)$ and $Z(0) = Z(2\pi)$, the condition of the connection of Eq. (18) and the relation of Eq. (15), the following result is obtained:

$$Y_f = 0, \quad Z = 0.$$

The above result means that the system is fixed at the optimum in a steady state. In addition, it can be proved that all stable responses approach the above steady state as time passes. Figure 7 shows the stable region of $(U_1 - X_0)$ against the magnitude of disturbance U_2 for various values of M (that is, T_f), which is obtained from Eqs. (19) and (20) in the case where the system is in the optimum before disturbance occurs. The shaded

Fig. 7 Relationship between the initial deviation $|U_1 - X_0|$ and the disturbance U_2 in order for the response to be stable (parameter M)

areas indicate the stable regions. As seen in the figure, the system becomes more stable as the value of M decreases. This is due to the fact that the component necessary for the optimizing control action in the output of the controlled system is damped out in passing through the high-pass filter and consequently the loop gain is reduced. Thus, it can be seen that the high-pass filter has the effect of stabilizing the system and of fixing the optimum in a steady state (but the system fluctuates somewhat in the neighbourhood of the optimum except the case of the square test signal).

2.3 Effect of dynamic lag of the controlled system

In this section, the effect of dynamic lag, namely its time constant N , of the controlled system on the response is considered, when the high-pass filter is ignored. In such a case, it is, however, difficult to seek the solution analytically. Hence the transient response due to the step disturbance, where U_1 alone is considered, is investigated by means of numerical calculus. Some typical time responses are illustrated in Fig. 8. Based on the observations of them, the following are deduced:

(1) The system becomes more stable as the time constant of the dynamic lag, N , increases. This is the same phenomenon as observed in the effect of the high-pass filter in the previous section. Thus, the system loop gain is reduced by dynamic lag.

(2) The periodic solution exists, and the stable solution approaches it.

(3) The solution which diverges into infinity within the interval including τ_0 exists (see $N=0.2$ in Fig. 8). And even if it is bounded in that interval, it is possible to diverge at the following interval (see $N=0.5$).

Moreover, it was observed that the system was ready to fall into instability in the neighbourhood of $\tau_0 = \pi$ or 0 depending on whether $(U_1 - X_0)$ was

Fig. 8 Several typical responses in the case of the controlled system with dynamic lag ($U_1=5.5$, $U_2=0$, $B=0.1$, $X_0=0$, $Y_{f0}=1$, $\tau_0=0$)

positive or negative.

3. Transient and steady state performance

In the previous chapter, the basic characteristics of the system response and the effect of each element of the system have been discussed mainly from the qualitative point of view. In this chapter, the system shown in Fig. 2 is considered, and, under the assumption that the curvature k and time constant T remain constant, the effect of the system parameters (although a square wave was used in previous chapter, a sign wave will be used hereafter) upon the performance of transient and steady state response is investigated quantitatively. Since it is, however, difficult to get an exact solution analytically, a numerical method by digital computer simulation (in the following section such solution by the numerical method will be called an exact solution) and an approximate analytical method are applied.

3.1 Transient state performance

As the performance index for transient response the integral of absolute nondimensionalized error of the controlled system output (hereafter it is called integral of error), as defined in the following, is considered:

$$S = (1/N) \int_{\tau_0}^{\infty} |Y - (U_2 + 1/2)| d\tau. \quad (21)$$

In the above equation the desired value (final value) of the output Y is not only U_2 , but also $1/2$, which denotes the average value of periodic solution to which the transient response approaches, is added to it. Here, the initial condition is assumed as follows:

$$X_0 = 0, Y_0 = 1/2, Y_{f0} = 0. \quad (22)$$

This means that the system has been in the optimum before disturbance occurs, and that approximate average values in such cases are adopted as initial values. Then, it is provided that the disturbance U_1 is positive and the value of τ_0 is π or 0. It is evident from the previous discussions that, if the value of τ_0 is considered in the range from zero to 2π , all solutions are discussed without exception even if it is assumed that $U_1 > 0$ only. It was previously recognized that S in Eq. (21) became the maximum and the system was the most unstable in the neighbourhood of $\tau_0 = \pi$, and vice versa in the neighbourhood of $\tau_0 = 0$.

3.1.1 Results and considerations

Figure 9 represents the effects of frequency of a test signal N and a loop gain B upon S in case where the filter time constant M , the disturbance (U_1, U_2) are constant. It can be seen from the

figure that S decreases as N increases, that is, the transient performance is improved. This is owing to the fact that the rate of the control action is intensified as N increases. Moreover, in case $\tau_0 = \pi$, S reaches a minimum as B increases for each value of N , but goes to infinity with a further increase of B . This should be noted when B is determined so as to minimize S . On the other hand, in case $\tau_0 = 0$, S is apt not to increase even if B increases, but this situation can not be expected always in practical. In addition, similar results were obtained when both U_1 and U_2 were considered.

Next, the effect of M upon S was investigated. According to the result, in case of disturbance U_1 only, S decreases as M decreases, and the opposite was true in case of U_2 only. After all, the effect of M on S varies in a complicated way depending on the disturbance. Therefore, a sweeping statement can not be made. It is, however, deduced that it is better for the value of M to be rather small, from the fact that the system is stabilized as M decreases.

Fig. 10 represents the plot of integral of error against loop gain for various values of U_1 , but it differs from the previous figures in that the ordinate and abscissa are replaced by nondimensional parameters S/U_1^2 and BU_1 , respectively, both of which excludes a parameter E . It can be seen from the figure that the transient performance is improved according as the ratio of the amplitude of the test signal E to the magnitude of disturbance

Fig. 9 Effect of B and N on the integral of error S

X increases (U_1 decreases). This means that the rate of control action is intensified. On the other hand, there exists such a value of BU_1 as to minimize S/U_1^2 (hence S), and its value is almost fixed independently of U_1 . This fact will be applicable at the time of design.

Finally, the relation between S and $\Delta\theta$, which is the difference in phase angle of correlation signal from the desired value [see Eq. (2)] was investigated. As a result, it was found that the best result was not always obtained when $\Delta\theta=0^\circ$, and that $\Delta\theta=\pm 45^\circ$ was still allowable from the practical viewpoint.

3.1.2 Comments on the approximate method

In the system in Fig. 2, when X deviates slightly from the optimum compared with the amplitude of the test signal and fluctuates slowly, the above system is usually approximated by the system shown in Fig. 11, where κ is given as follows:

Fig. 10 Relation between gain BU_1 and the integral of error S/U_1^2 excluding the amplitude of test signal E

Fig. 11 Approximate representation of the system in Fig. 2 by equivalent linearization

$$\kappa = M/\sqrt{(1+N^2)(1+M^2)} \dots\dots\dots (23)$$

Figure 12 shows a transient response obtained from the above approximated system (hereafter called the approximate solution) under comparison with the exact solution. Since the approximate system is made without taking account of the time when the disturbance is applied, it is impossible to express correctly the transient response. Nevertheless the approximate solution comes between two exact solutions corresponding to $\tau_0=0$ and π , and agrees nearly with them after a sufficient passage of time in the response. Therefore, if the settling time is applied as an index of the rapidity of response, approximate solution can be used without serious error. But it must be noted that the approximate system produces a gross error in the transient state, and that since it is always stable, it gives no information upon the stability shown in Chapter 1.

3.2 Steady state performance

3.2.1 Hunting loss due to test signal

In a steady state, the output of the controlled system hunts around the optimum due to the continued searching action, and a so-called hunting loss is created. It was found by the exact solution that the hunting loss was almost equal to $1/(2kE^2)$, that is, the fluctuation of X due to test signal might be ignored.

3.2.2 Hunting loss in the presence of noise

Noise included in the output of the controlled system causes the fluctuation of its input and, consequently, an increase of the hunting loss. In such a case, the hunting loss is given as follows:

$$(y^* - y_{\min}) = k(-x_E - x_n + E \sin \omega_p t)^2 \dots\dots\dots (24)$$

where x_E : fluctuation due to test signal

x_n : fluctuation due to noise

As stated in the previous section, x_E is so small that it can be ignored. Then, by taking account of the absence of correlation between noise and test signal, and non-dimensionalizing by the standard deviation σ , the following result is obtained:

$$(y^* - y_{\min})/\sigma = kE^2/(2\sigma) + kx_n^2/\sigma \dots\dots\dots (25)$$

In this section the magnitude of the hunting

Fig. 12 An example of the comparison between the exact solution and the approximate solution of step response ($U_1=10/3$, $U_2=0$, $N=M=1$, $B=0.3$)

loss is calculated from Eqs. (2) and (25) exactly (by numerical calculus) and the effect of noise on the hunting loss is discussed. In addition, the results from the approximate system are compared with exact ones and its accuracy is checked. Noise is assumed as a Gaussian random process with zero mean value, and the following two power spectral densities are provided: $4T\sigma^2/(1+\omega^2T^2)$ [noise (1)], and $8\zeta\omega_n^3\sigma^2/[(\omega_n^2-\omega^2)^2+4\zeta\omega_n^2\omega^2]$ [noise (2), $\zeta=0.2$, $\omega_n=1/T$].

The approximate solution is obtained as follows: Under the assumption that the fluctuation of x due to noise is small, the system shown in Fig. 2 can be approximated by the system in Fig. 11. In this case the mean square value of the fluctuation of x can be calculated. For example, in case of noise (1) the following result is obtained.

$$\frac{k\bar{x}_n^2}{\sigma} = \frac{BNM^2}{2\kappa(M^2-N^2)} \left[\frac{(1+\kappa BN)\{(1-\kappa BN)^2+N^2\}}{\{1-N^2-(\kappa BN)^2\}^2+4N^2} - \frac{(1+\kappa BM)\{(1-\kappa BM)^2+M^2\}}{\{1-M^2-(\kappa BM)^2\}^2+4M^2} \right] \times \left(\frac{\sigma}{kE^2} \right) \equiv \lambda^2 \left(\frac{\sigma}{kE^2} \right) \quad (26)$$

In Fig. 13 both exact and approximate solutions are together presented for various values of N and $B\sqrt{\sigma/(kE^2)}$. In this figure, kE^2/σ is plotted on the abscissa and $(y^*-y_{min})/\sigma$ on the ordinate. As seen from the figure, the hunting loss increases with a decrease of $B\sqrt{\sigma/(kE^2)}$, but the accuracy of the approximation is reduced. However, so far as $B\sqrt{\sigma/(kE^2)}$ is small, the approximation is tolerably accurate. In the figure the chain line indicates the hunting loss of a noise-free system for comparison, and it can be seen that noise has a considerable effect upon the hunting loss on certain conditions. Similar results are obtained in case of noise (2).

Fig. 13 Comparison between the exact solution and the approximate one of hunting loss in case noise exists

4. Principle and procedure of design

In the previous chapter the integral of the transient error and the hunting loss were considered as the transient and steady state performance. In this chapter, the minimum of each performance index and the relationship between them are investigated in connection with the requirement from the viewpoint of design that they should be made as small as possible. This relationship is, of course, very complicated and it is impossible to discuss exactly. Therefore, an approximate method is employed depending on the conditions.

First, from Eqs. (25) and (26) concerned with the approximate system, the minimum of the hunting loss $[(y^*-y_{min})/\sigma]_{min}$ is given as follows:

$$\text{for } kE^2/\sigma = \sqrt{2}\lambda, \quad (y^*-y_{min})/\sigma]_{min} = \sqrt{2}\lambda. \quad (27)$$

The above result shows that the following functional relation exists:

$$(y^*-y_{min})/\sigma]_{min} = kE^2/\sigma = F(\kappa BN, M, N) \dots (28)$$

Next, consider the minimum of the integral of transient error S . Figure 10 in section 3.1.1 shows that S is minimized for an almost equal value of BU_1 (but U_2 remains constant). The value of B so as to minimize S decreases, as seen in Fig. 9, almost proportionally to N . On the other hand, the value of κ increases with a decreasing N as shown in Eq. (22). From the above considerations it can be expected that S would be minimized, independently of N , for an almost equal value of non-dimensional parameter κBN . In the general case of considering both U_1 and U_2 , using $\sqrt{U_1^2+U_2^2}$, which is an equivalent representation of U_1 and U_2 in U_1 side, instead of U_1 and plotting S against $\kappa B\sqrt{U_1^2+U_2^2}$, it is found that S is minimized for the almost equal value shown in Eq. (29) of $\kappa B\sqrt{U_1^2+U_2^2}$:

$$\kappa B\sqrt{U_1^2+U_2^2} = 0.35 \sim 0.4 \dots (29)$$

Thus, the above assumption is supported. By use of dimensional quantities, Eq. (29) is rewritten as follows:

$$\frac{(\kappa BN)}{N} \sqrt{\left(\frac{kx_{opt}^2 + y_{min}}{\sigma} \right)} \left(\frac{\sigma}{kE^2} \right) = 0.35 \sim 0.4 \quad (30)$$

Equations (28) and (30) mean that, when M and N are regarded as parameters, $(kx_{opt}^2 + y_{min})/\sigma$ and $(y^*-y_{min})/\sigma]_{min} = kE^2/\sigma$ are related through a variable BN under the condition of the minimum integral of transient error. Here, from the fact that the system falls into instability for too large value of M , while the signal to noise ratio at the output of the high-pass filter is reduced for too small value of M , the possible range for the value

of M is given as follows:

$$M = \omega_p T_f = 0.5 \sim 1. \quad (31)$$

Thus, if the value of M is fixed, the parameter in Eqs. (28) and (30) is N alone.

Now, from the above, the following design procedure is derived. Suppose that the curvature and time constant of the controlled system k , T , standard deviation of noise σ , predicted maximum magnitude of disturbance x_{opt} , y_{min} are known, then the parameters to be determined are the amplitude and angular frequency of test signal E , ω_p , time constant of high-pass filter T_f , and product of gain of integrator and amplitude of correlating signal KA (since the amplitude of correlation signal A is always presented as the product KA , it can be included in K). In the following description, the case of noise (1) is taken as an example for convenience of illustration. As stated before, if the value of N is determined, $(y^* - y_{min})/\sigma$ and simultaneously kE^2/σ are given for various κBN from Eq. (28), and $(kx_{opt}^2 + y_{min})/\sigma$ from Eq. (30). Figure 14 shows the relations stated above. Suppose that the predicted maximum of disturbance is expressed as (a) in the figure. If N is chosen 1,

Fig. 14 Chart for determining the control system design (in the case of noise (1))

then kE^2/σ (simultaneously the hunting loss $(y^* - y_{min})/\sigma$) and κBN are determined as (b) and (c) in the figure (the integral of error is minimized for such a pair of parameter values). Subsequently, three unknown parameters ω_p , E , and KA can be determined by using three equations $N = \omega_p T$, kE^2/σ and $\kappa BN = \{M/\sqrt{(1+N^2)(1+M^2)}\} \times (kKAET)$. On the other hand, T_f is determined from Eq. (31). Thus, the design is completed. In the above procedure, there is a certain degree of freedom in the choice of N . Hence, it is necessary to choose the most adequate one taking account of the hunting loss. In addition, it is desired for the sake of stability to choose a comparatively small value of KA (consequently B) rather than the one given from the above procedure (see 3.1.1). This results in the decrease of the hunting loss, but contradicts the requirement of rapidity [because the index of the settling time $1/(\kappa BN)$ increases]. Therefore, it is necessary to make some compromise between them.

Concluding remarks

In this paper the characteristics of the response of optimizing control system by a continuous correlation method were investigated in detail and a reasonable design method was proposed.

Grateful acknowledgment is extended to Professor Tadayo Ito for his invaluable assistance and guidance in this investigation.

References

- (1) C. S. Draper and Y. T. Li: *Principles of Optimizing Control Systems and an Application to the Internal Combustion Engine*, (1951), ASME.
- (2) G. Vasu: *Trans. ASME*, Vol. 79, No. 3 (1957), p. 481.
- (3) R. J. McGrath et al.: *IRE Trans. of Aut. Cont.*, AC-6 (1961), p. 154.
- (4) J. M. Nightingale: *Proc. IEE*, Part C, Vol. 109, No. 16 (1962), p. 539.
- (5) P. H. Hammond and M. J. Duckenfield: *Automatika*, Vol. 1, No. 2 (1963-8), p. 147.
- (6) V.W. Eveleigh: *Automatic and Remote Control*, (1963), p. 472, Butterworths.
- (7) S. Fujii and N. Kanda: *Automatic and Remote Control*, (1963), p. 380, Butterworths.
- (8) S. Fujii and T. Koga: *Trans. Japan Soc. Mech. Engrs.*, Vol. 26, No. 161 (1960), p. 129.
- (9) K. Fukunaga and Y. Inoue: *Jour. Japan Assoc. Aut. Con. Engrs.*, Vol. 6, No. 4 (1962), p. 207.
- (10) Y. Nishimura: *Preprint 2nd Meet. Japan Soc. Instr. and Cont. Engrs.*, (1963), p. 147.
- (11) A. Ichikawa and H. Yamazaki: *Preprint 3rd Meet. Japan Soc. Instr. and Cont. Engrs.*, (1964), p. 363.
- (12) H. Hatanaka et al.: *Preprint 42th Ann. Meet. Japan Soc. Mech. Engrs.*, No. 133 (1965), p. 37.