

GAIN-MAXIMIZING CONTROL IN HIGH-ORDER PROCESSES

RYOHEI TANUMA

*Fuji Electric Corporate Research and Development Ltd.,
2-2-1, Nagasaka, Yokosuka 240-01*

Key Words: Process Control; Adaptive Control; Nonlinear Process; Gain-Maximizing Control; High-Order Process; Broken-Line Fitting

This paper describes the study of the gain-maximizing (GMAX) control system in high-order processes. The GMAX system is applied to the process where a high-order linear dynamic element is followed by a static sigmoid function. It regulates the process output at the sigmoid curve's inflection point, where the static gain becomes maximum. Adaptive broken-line fitting based on a new high-order difference equation is developed. New feedback compensation is also proposed to improve control performance. Computer simulations show that the process output settles quite rapidly to the maximum-gain level, and reliable GMAX control can be achieved.

Introduction

The author and co-workers have developed a gain-maximizing (GMAX) control system for dissolved oxygen (DO) control in the activated sludge process.⁴⁾ In this process, the DO concentration is related to the aeration air flow rate by a sigmoid curve, and the gain of DO concentration with respect to the air flow rate maximizes at the inflection point (IP) of the sigmoid curve. The GMAX control automatically adjusts the DO level so as to maximize the gain, and the DO level is thereby regulated at the IP.

Similar sigmoid relations exist in various processes, particularly in chemical and biological processes. A typical example can be seen in a reduction-oxidation (redox) process, where the redox potential jumps at the equivalence point, thus varying along a sigmoid curve. The pH value in an acid-base reaction behaves in almost the same manner. In these processes, the IPs usually correspond to optimum conditions to be maintained. Although the IP levels often vary with unknown factors such as sensor contamination and environmental disturbances, the GMAX action will automatically hold the output at the IP levels, thereby maintaining optimum conditions. Hence, similar GMAX methods would be quite useful in these processes. However, most existing sigmoid processes exhibit high-order dynamic behavior, whereas the GMAX method used in DO control is only for a first-order process.

The author has proposed a new GMAX method which is applicable to high-order processes.⁵⁾ This paper describes the improved version of this new method. Adaptive broken-line fitting for high-order

processes is first developed and incorporated into the GMAX method. To improve control performance, new feedback compensation is introduced. Computer simulations verify the excellent performance of the new method.

2. Adaptive Broken-Line Fitting

Nonlinear processes can usually be represented by two types of generalized models: a linear dynamic element followed by a nonlinear static element (DS series model) and a nonlinear static element followed by a linear dynamic element (SD series model).¹⁾

In this paper, we deal with the sigmoid process described by the DS series model. This is expressed as the block diagram shown in Fig. 1, where F is the input, C_L is the output of the linear block, and C is the output of the nonlinear block. The gain of the linear block is taken as unity. Let us assume that the linear part is stable and nonvibrational. The sigmoid curve in the nonlinear block is illustrated in Fig. 2. For the GMAX control, we consider the broken-line fitting to the curve as shown in this figure. The adaptive broken-line fitting (ABLF) described below is the parameter identification technique to determine the broken line.

Using the zero-average variables

$$\begin{aligned}u &= F - \bar{F} \\x &= C_L - \bar{C}_L \\y &= C - \bar{C}\end{aligned}$$

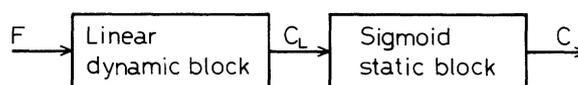


Fig. 1. Sigmoid process described by the DS series model

* Received October 30, 1990. Correspondence concerning this article should be addressed to R. Tanuma.

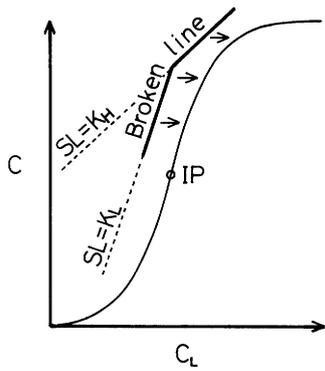


Fig. 2. Sigmoid curve and the broken line

where \bar{F} , \bar{C}_L and \bar{C} are the time averages of F , C_L and C , respectively, we can express the linear part as

$$x_k = a_1 x_{k-1} + a_2 x_{k-2} + \dots + a_n x_{k-n} + b_1 u_{k-1} + b_2 u_{k-2} + \dots + b_n u_{k-n} \quad (1)$$

where $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ are time-invariant coefficients. If we introduce the broken-line model, the nonlinear part can be expressed as

$$K(y_{k-i}) = \begin{cases} K_L & \text{for } y_{k-i} \leq 0 \\ K_H & \text{for } y_{k-i} > 0 \end{cases} \quad (2)$$

$(i=0, 1, 2, \dots, n)$

where K_L and K_H respectively correspond to the slopes (SLs) of the lower and upper linear parts of the broken line (Fig. 2). To obtain an ABLF algorithm, the shift equation (1) must be converted into a difference form where the gain appears explicitly. However, the difference equation which includes high-order difference terms²⁾ is not suitable for parameter identification because the high-order terms may enhance high-frequency noise.

To avoid this problem, the author proposes a new difference form where all the difference terms are of first order:

$$x_k + \alpha'_1 \Delta x_{k-1} + \alpha'_2 \Delta x_{k-2} + \dots + \alpha'_n \Delta x_{k-n} = \beta'_1 u_{k-1} - \beta'_2 \Delta u_{k-2} - \dots - \beta'_n \Delta u_{k-n} \quad (3)$$

where

$$\begin{aligned} \Delta x_{k-1} &= x_k - x_{k-1}, & \Delta x_{k-2} &= x_{k-1} - x_{k-2}, & \dots \\ & \dots, & \Delta x_{k-n} &= x_{k-n+1} - x_{k-n} \\ \Delta u_{k-2} &= u_{k-1} - u_{k-2}, & \Delta u_{k-3} &= u_{k-2} - u_{k-3}, & \dots \\ & \dots, & \Delta u_{k-n} &= u_{k-n+1} - u_{k-n} \end{aligned}$$

Hereafter, the differences for the variables y , C , F , λ , and χ are defined similarly. The time-invariant coefficients $\alpha'_1, \alpha'_2, \dots, \alpha'_n, \beta'_1, \beta'_2, \dots, \beta'_n$ are related to the coefficients in Eq. (1) as

$$\left. \begin{aligned} \alpha'_1 &= (a_1 + a_2 + \dots + a_n) / \gamma \\ \alpha'_2 &= (a_2 + a_3 + \dots + a_n) / \gamma \\ & \dots \\ \alpha'_n &= a_n / \gamma \\ \beta'_1 &= (b_1 + b_2 + \dots + b_n) / \gamma \\ \beta'_2 &= (b_2 + b_3 + \dots + b_n) / \gamma \\ & \dots \\ \beta'_n &= b_n / \gamma \end{aligned} \right\} \quad (4)$$

where

$$\gamma = 1 - a_1 - a_2 - \dots - a_n$$

and $\beta'_1 (=1)$ is the gain. Note that

$$\gamma > 0 \quad (5)$$

because of the assumption for the linear part. Substituting Eq. (2) into Eq. (3), we obtain

$$y_k + \alpha_1 \Delta y_{k-1} + \alpha_2 \Delta y_{k-2} + \dots + \alpha_n \Delta y_{k-n} = K(y_k) u_{k-1} - \beta_2 \Delta u_{k-2} - \dots - \beta_n \Delta u_{k-n} \quad (6)$$

with the approximation

$$\Delta x_{k-i} \approx \Delta y_{k-i} / K_i$$

$$K_i = [K(y_{k-i}) + K(y_{k-i-1})] / 2 \quad (i=1, 2, \dots, n)$$

and the time averaging

$$\alpha_i = \overline{\alpha'_i K(y_k) / K_i} \quad (i=1, 2, \dots, n)$$

$$\beta_i = \overline{\beta'_i K(y_k)} \quad (i=2, \dots, n)$$

Using the parameter vector

$$p = [\alpha_1, \dots, \alpha_n, \beta_2, \dots, \beta_n, K_L, K_H]$$

and the signal vector

$$\phi_k = [\Delta y_{k-1}, \dots, \Delta y_{k-n}, \Delta u_{k-2}, \dots, \Delta u_{k-n}, h(y_k, u_{k-1})]^T$$

we can rewrite Eq. (6) as

$$y_k + p \phi_k = 0 \quad (7)$$

where

$$h(y_k, u_{k-1}) = \begin{cases} [-u_{k-1}, 0] & \text{for } y_k \leq 0 \\ [0, -u_{k-1}] & \text{for } y_k > 0 \end{cases}$$

The equation error e is defined as

$$e = y_k + \hat{p}_k \phi_k \quad (8)$$

where \hat{p}_k is the adjustable parameter vector,

$$\hat{p}_k = [\hat{\alpha}_{1k}, \dots, \hat{\alpha}_{nk}, \hat{\beta}_{2k}, \dots, \hat{\beta}_{nk}, \hat{K}_{Lk}, \hat{K}_{Hk}]$$

Subtracting Eq. (7) from Eq. (8), we obtain the error equation,

$$e = (\hat{p}_k - p)\phi_k \quad (9)$$

and consequently, the identification algorithm using a fixed-trace scheme is

$$\left. \begin{aligned} e_o &= y_k + \hat{p}_{k-1}\phi_k \\ \hat{p}_k &= \hat{p}_{k-1} - \frac{\mathbf{G}_{k-1}\phi_k e_o}{1 + \phi_k^T \mathbf{G}_{k-1} \phi_k} \\ \mathbf{G}_k &= \frac{1}{v_k} \left(\mathbf{G}_{k-1} - \frac{\mathbf{G}_{k-1}\phi_k \phi_k^T \mathbf{G}_{k-1}}{\mu + \phi_k^T \mathbf{G}_{k-1} \phi_k} \right) \end{aligned} \right\} \quad (10)$$

where e_o is the *a-priori* value of e , \mathbf{G}_k is the gain matrix, μ is a constant, and the value of v_k is updated at each step of the identification in such a way that the trace of \mathbf{G}_k remains constant.³⁾ The parameters \hat{K}_{Lk} and \hat{K}_{Hk} , the last two members of \hat{p}_k , are estimates of K_L and K_H , respectively. Hence, adjusting the process output level to equalize \hat{K}_{Hk} and \hat{K}_{Lk} , we can maintain the output at the IP level. We refer to the above identification algorithm as the ABLF.

The difference form is essential in the ABLF. Consider that a test signal fluctuates y and u over an appropriate range as shown in Fig. 3. To obtain the distinct difference between \hat{K}_{Lk} and \hat{K}_{Hk} , they should be determined in the outer parts of the fluctuating zones (the crosshatched parts). The difference form permits this kind of weighting because the static variables, y_k and u_{k-1} , automatically dominate over the dynamic variables, $\Delta y_1, \dots$ and $\Delta u_2, \dots$, in the outer parts. This consideration also shows that the period of the test signal must be long enough to make the static variables predominant.

3. Control System

Figure 4 shows the GMAX control system using the high-order ABLF. Control loop Lp1 manipulates adjustable set point C_r to maximize the gain, and control loop Lp2 controls the output C around C_r . The high-pass filters convert input F and output C to zero-average signals u and y , respectively. The filtering algorithm is

$$\zeta_k = \eta_{k-m/2} - \frac{\eta_k + \eta_{k-1} + \dots + \eta_{k-m+1}}{m} \quad (11)$$

where the pair of ζ and η is that of y and C , or u and F , respectively, and m is the number of data, which is taken as an even number. These high-pass filters can effectively reject unwanted low-frequency components.⁴⁾ On receiving u and y , the ABLF determines the parameter vector \hat{p}_k , which is used not only for gain maximization, but also for feedback compensation and adaptive PI control in Lp2. The controlled variable in Lp1 is c_v , obtained from \hat{p}_k as

$$c_{vk} = (\hat{K}_{Hk} - \hat{K}_{Lk}) / \hat{K}_k \quad (12)$$

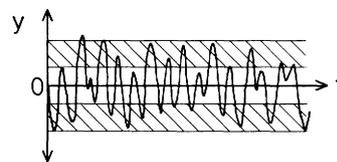


Fig. 3. Fluctuation of y and the zones where K_L and K_H should be determined

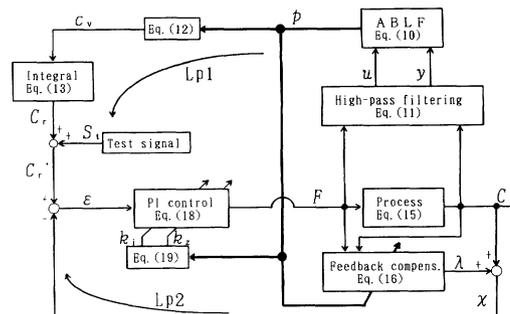


Fig. 4. Control system

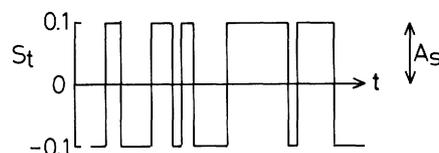


Fig. 5. Test signal

where

$$\hat{K}_k = (\hat{K}_{Lk} + \hat{K}_{Hk}) / 2$$

The c_v values are limited in $-1 \leq c_v \leq 1$ to exclude extraordinary values. Since $c_{vk} \geq 0$ and < 0 approximately correspond to $d^2C/dF^2 \geq 0$ and < 0 , respectively, controlling c_{vk} at zero by manipulating C_r results in gain maximization. For this control, the integral action

$$C_{rk} = C_{rk-1} + k_z c_{vk} \quad (13)$$

where k_z is the integral gain is used because c_v is often noisy. For C to fluctuate over an appropriate range, the test signal S_t is applied to C_r , and C tracks the perturbed set point $C'_r = C + S_t$. In this system, S_t is a pseudo-random binary signal, which is a rectangular wave with randomized periods as shown in Fig. 5. It is specified by the amplitude A_s , the average period τ_p , and the standard deviation S_d of the period.

To keep the GMAX control in high performance, the following two conditions must be satisfied: C must fluctuate over a constant range (**Cond. A**), and C and F fluctuating ranges must be divided equally into lower and upper regions (**Cond. B**).

Cond. B is well satisfied with the high-pass filters described above. For details, refer to the previous paper.⁴⁾

Simple adaptive PI control has been used to satisfy Cond. A in the GMAX DO control.⁴⁾ Because it is

only for a first-order process, fixed-parameter PI control was used preliminarily in the previous high-order GMAX system, and consequently the C range could not be fixed.⁵⁾

For the new high-order GMAX, the author has developed the feedback compensation shown in Fig. 4. Since the compensated feedback signal χ behaves as the output of a first-order process as described below, we can use the same adaptive PI control. The linear model of the high-order process is obtained by replacing $K(y_k)$ in Eq. (6) with constant gain K :

$$y_k + \alpha_1 \Delta y_{k-1} = K u_{k-1} - (\alpha_2 \Delta y_{k-2} + \dots + \alpha_n \Delta y_{k-n}) - (\beta_2 \Delta u_{k-2} + \dots + \beta_n \Delta u_{k-n}) \quad (14)$$

Using C and F , we can rewrite Eq. (14) as

$$C_k + \alpha_1 \Delta C_{k-1} = K F_{k-1} - \mathbf{p}' \boldsymbol{\phi}' + \text{const.} \quad (15)$$

where

$$\mathbf{p}' = [\alpha_2, \dots, \alpha_n, \beta_2, \dots, \beta_n]$$

$$\boldsymbol{\phi}' = [\Delta C_{k-2}, \dots, \Delta C_{k-n}, \Delta F_{k-2}, \dots, \Delta F_{k-n}]^T$$

This means that C_k behaves as the output of the first-order element whose input is the right-hand side of Eq. (15). It is apparent that this first-order element is stable because all difference terms in Eq. (15) converge to zero for a constant F values.

Let us consider the feedback compensation signal λ_k that is generated by

$$\lambda_k + \hat{\alpha}_{1k} \Delta \lambda_{k-1} = \hat{\mathbf{p}}'_k \boldsymbol{\phi}' \quad (16)$$

where

$$\hat{\mathbf{p}}'_k = [\hat{\alpha}_{2k}, \dots, \hat{\alpha}_{nk}, \hat{\beta}_{2k}, \dots, \hat{\beta}_{nk}]$$

As shown in Fig. 4, λ_k is added to C_k to give the feedback signal χ_k . Now we assume that the identification is completed; that is $\alpha_{1k} = \alpha_1$ and $\hat{\mathbf{p}}'_k = \mathbf{p}'$. Adding Eqs. (15) and (16), we obtain

$$\chi_k + \alpha_1 \Delta \chi_{k-1} = K F_{k-1} + \text{const.} \quad (17)$$

where

$$\chi_k = C_k + \lambda_k$$

That is, χ_k behaves as the output of the first-order system. Hence, using χ_k as the feedback signal, we can use the same adaptive PI control as that used in the first-order GMAX system.⁴⁾ The PI control algorithm used here is expressed as

$$e_k = C'_{rk} - \chi_k$$

$$F_k = F_{k-1} + k_p (e_k - e_{k-1}) + k_i e_k \quad (18)$$

where k_p is the proportional gain and k_i is the integral gain. The PI parameters are determined by

$$\left. \begin{aligned} k_i &= (1-d)/\hat{K}_k \\ k_p &= \hat{\alpha}_{1k} \cdot k_i \end{aligned} \right\} \quad (19)$$

where d is the desired pole value of the closed loop. In Eqs. (16) and (19), the values of $\hat{\alpha}_{1k}$ and \hat{K}_k are limited in $\hat{\alpha}_{1k} \geq -0.4$ and $\hat{K}_k \geq 1$, respectively, to exclude extraordinary values. If the PI parameters are thereby determined, the closed loop transfer function $G_L(z)$ becomes⁴⁾

$$G_L(z) = \frac{1-d}{z-d} \quad (20)$$

That is, we can specify the fluctuation range of χ .

Now let us consider how the χ range is related to the C range. The denominator of the process transfer function can be expressed as

$$A_p(z) = z^n - a_1 z^{n-1} - a_2 z^{n-2} - \dots - a_n$$

$$= (z - \rho_1)(z - \rho_2) \dots (z - \rho_n) \quad (21)$$

where $\rho_1, \rho_2, \dots, \rho_n$ are the poles of the process to be controlled. Referring to Eqs. (4) and (21), we obtain

$$\alpha_1 = \frac{1-\gamma}{\gamma} = \frac{1-A_p(1)}{A_p(1)}$$

$$= \frac{1-(1-\rho_1)(1-\rho_2) \dots (1-\rho_n)}{(1-\rho_1)(1-\rho_2) \dots (1-\rho_n)} \quad (22)$$

The parameter α_1 can also be expressed as

$$\alpha_1 = \frac{1-(1-\rho)}{1-\rho} \quad (23)$$

where ρ is the pole of the first-order delay which χ_k obeys. Comparing Eqs. (22) and (23), we obtain

$$0 < 1-\rho = (1-\rho_1)(1-\rho_2) \dots (1-\rho_n) < 1 \quad (24)$$

which shows

$$\rho_i < \rho < 1 \quad (i=1, 2, \dots, n) \quad (25)$$

That is, ρ is larger than the maximum pole of the system. Hence, taking the test-signal period long enough for the dynamic gain of χ nearly to saturate, we can let the C range be almost the same with the χ range which is regulated to be constant.

4. Simulations

Computer simulations have been performed to evaluate the new method. The process model for the simulations is as follows:

[Linear part]

$$G_L(s) = \frac{1}{(1+0.2s)(1+0.3s)(1+0.1s)(1+0.05s)} \quad (26)$$

[Nonlinear part]

$$C = \begin{cases} 1 - 0.5 \exp[-10(C_L - C_o)] & \text{for } C_L \geq C_o \\ 0.5 \exp[10(C_L - C_o)] & \text{for } C_L < C_o \end{cases} \quad (27)$$

where C_0 is the C_L value at the inflection point. Equation (27) is graphically shown in Fig. 6. The maximum slope is 5 at $C=C_0=0.5$.

The parameter values for the simulations were as follows:

$$\begin{array}{ll} \mu=0.7 & G_0=2 \times 10^4 \mathbf{I} \\ T=0.1 & m=20 \\ L=mT=2 & \tau_p=2 \\ A_s=0.1 & S_d=0.5 \\ d=0.905 & k_z=0.005 \end{array}$$

Note that τ_p is equalized to L for the high-pass filtering.⁴⁾ Two simulations, I and II, were performed. In Simulation I, the integral action in Lp1 was fixed, and C_r was increased stepwise. In Simulation II, the system was evaluated in normal action.

5. Results and Discussion

Figure 7 shows the results of Simulation I. The set point C_r was increased from 0.2 to 0.8 in six steps (the C_r values are shown in the figure). The input F and the output C increase with C_r . The fluctuations of F and C are caused by the test signal. As C_r increases, \hat{K}_L and \hat{K}_H first increase, maximize at $C_r=0.5$, and then decrease; correspondingly, \hat{K}_H is larger than \hat{K}_L at $C_r < 0.5$, they are equalized at $C_r=0.5$, and then \hat{K}_L becomes larger than \hat{K}_H at $C_r > 0.5$. Note that the two gains and the other parameters, $\hat{\alpha}_{1-4}$ and $\hat{\beta}_{2-4}$, become stable as C_r reaches the IP level. This may be because the process can be well approximat-

ed by a linear model around the IP level. Although the C range is a little larger in the lower and upper region than around the IP level, it is almost constant in the range of $C_r=0.3$ to 0.7.

Figure 8 shows the results of Simulation II. The set point was first fixed at $C_r=0.2$ until time $t=10$, and was then released. The gains, \hat{K}_L and \hat{K}_H , are first as low as 1, and $\hat{K}_H > \hat{K}_L$. As soon as it is released, C_r rapidly increases and settles at the IP level. The setting time (~ 20) is about half that in the previous high-order GMAX with a fixed-parameter PI control.⁵⁾

At $t=250$, C_0 was changed from 0.5 to 0.7. Although a great disturbance is caused here, it settles soon, and hence the deviation of C_r from the IP level is trivial.

The results of the two simulations show that the new algorithm has high GMAX performance. The ABLF performs quite satisfactorily. The two gains are clearly separated, and their relation well represents the curvature of the sigmoid curve. The adaptive PI

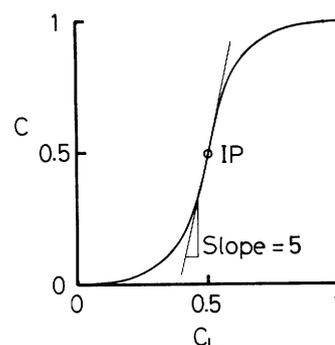


Fig. 6. Sigmoid curve for the simulations

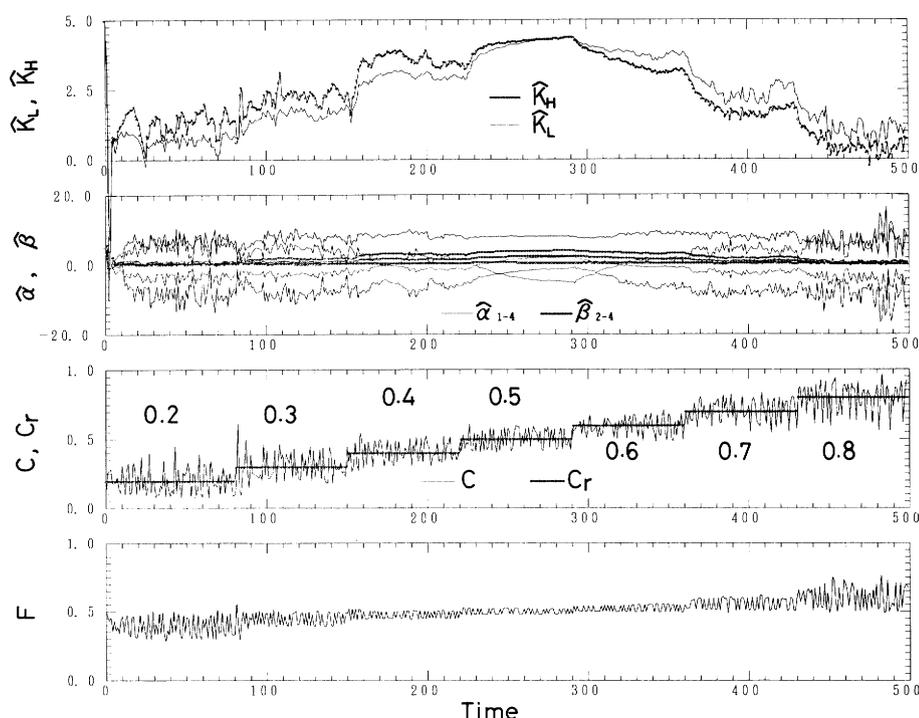


Fig. 7. Simulation I: the integral action in Lp1 was fixed, and C_r was increased stepwise

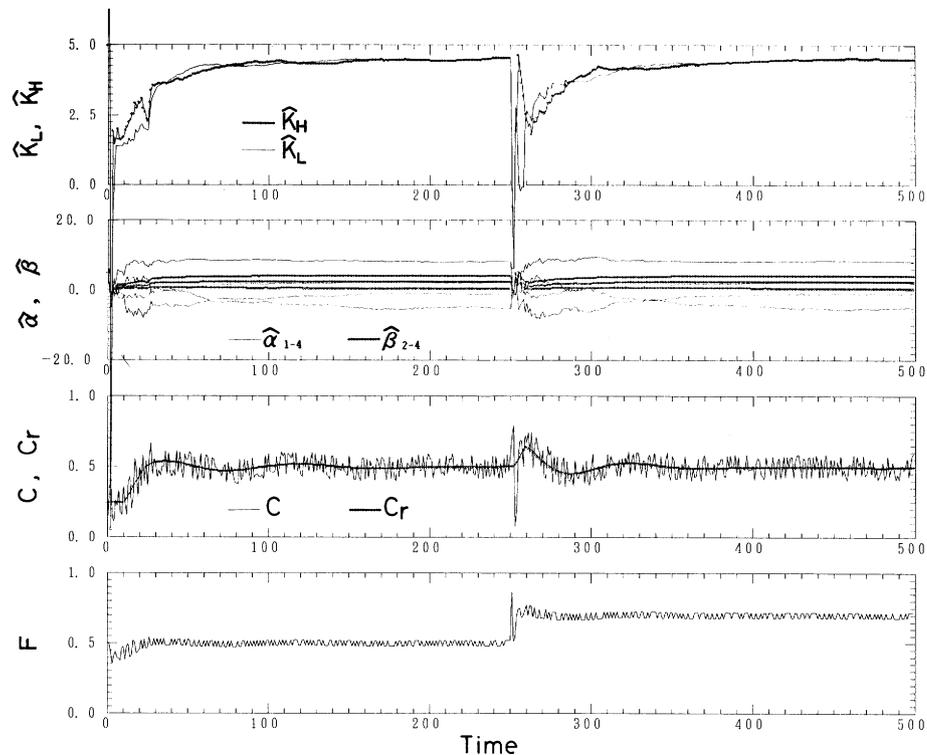


Fig. 8. Simulation II: normal action

control with the new feedback compensation maintains the C range nearly constant in a wide zone around the IP level. Although the C range is a little greater in the outer region, this is rather desirable to accelerate the GMAX action. The rapid settling to the IP position is owing to the controlled C range.

6. Conclusions

The simulation study shows that the new GMAX system performs quite satisfactorily in the high-order process. The high performance is owing to the ABLF based on the new difference equation, and the adaptive PI control with the new feedback compensation.

Nomenclature

A_s	= amplitude of the test signal	[—]
$a_1, \dots, a_n, b_1, \dots, b_n$	= coefficients of the shift equation	[—]
C	= process output	[—]
C_L	= output of the nonlinear block	[—]
C_r	= set point	[—]
C_r'	= perturbed set point	[—]
C_o	= C_L at the inflection point	[—]
c_v	= controlled variable in Lp1	[—]
d	= desired pole value	[—]
e	= equation error	[—]
e_o	= <i>a priori</i> value of e	[—]
F	= process input	[—]
G	= gain matrix	[—]
$h(y_k, u_{k-1})$	= switching vector in the signal vector	[—]
$K(y_{k-i})$	= switching gain	[—]
K_L	= lower part gain	[—]
\hat{K}_L	= adjustable lower part gain	[—]
K_H	= upper part gain	[—]
\hat{K}_H	= adjustable upper part gain	[—]

k_i	= integral gain for the PI control	[—]
k_p	= proportional gain for the PI control	[—]
k_z	= integral gain in Lp1	[—]
L	= period of the moving average	[—]
m	= number of data for the moving average	[—]
p	= parameter vector	[—]
\hat{p}	= adjustable parameter vector	[—]
S_d	= standard deviation of the period of the test signal	[—]
T	= sampling time	[—]
t	= time	[—]
u	= filtered process input	[—]
x	= zero-average output of the linear block	[—]
y	= filtered process output	[—]

$\alpha_1, \dots, \alpha_n, \beta_2, \dots, \beta_n$	= coefficients of the dynamic terms in the process model	[—]
$\hat{\alpha}_1, \dots, \hat{\alpha}_n, \hat{\beta}_2, \dots, \hat{\beta}_n$	= adjustable coefficients of the dynamic terms in the process model	[—]
$\alpha'_1, \dots, \alpha'_n, \beta'_1, \dots, \beta'_n$	= coefficients of the linear block model	[—]
γ	= $1 - a_1 - a_2 \dots - a_n$	[—]
μ, v	= parameters in the identification algorithm	[—]
ζ	= output of the high-pass filter	[—]
η	= input of the high-pass filter	[—]
λ	= feedback compensation signal	[—]
ρ	= pole of the first-order delay which χ obeys	[—]
ρ_1, \dots, ρ_n	= poles of the system to be controlled	[—]
τ_p	= average period of the test signal	[—]
ϕ, ϕ'	= signal vectors	[—]
χ	= feedback signal	[—]

Literature Cited

- Hotta, K.: "Dynamic Characteristics of Processes" (in Japanese), p. 82, Baifukan (1975).

- 2) Kitamori, T.: *J. Soc. Instrument and Control Engineers*, **22**, 599 (1983).
- 3) Landau, I. D. and M. Tomizuka: "Theory and Practice of Adaptive Control Systems" (in Japanese), Ohmu-sha (1981).
- 4) Tanuma, R, K. Sasaki and I. Matsunaga: *Trans. of the Society of Instrument and Control Engineers*, **26**, 528 (1990).
- 5) Tanuma, R.: Proceedings of 7th SICE Symposium on Adaptive Control, 29 (1987).