SRC-TR-87-66

CHEMICAL PROCESS SYSTEMS LABORATORY

Short-Cut Operability Analysis
Part III - Short-Cut Methodology for the Assessment of Process Control Designs

by

Maria Marino-Galarraga
Thomas J. McAvoy
Thomas E. Marlin

RESEARCH REPORT

CHEMICAL PROCESS SYSTEMS ENGINEERING LABORATORY

SHORT-CUT OPERABILITY ANALYSIS

PART III - SHORT-CUT METHODOLOGY FOR THE ASSESSMENT OF PROCESS CONTROL DESIGNS

Maria Marino-Galarraga
Thomas J. McAvoy
Thomas E. Marlin

A CONSTITUENT LABORATORY OF THE SYSTEMS RESEARCH CENTER

THE UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742

Short-Cut Operability Analysis

Part III - Short-Cut Methodology for the Assessment of Process Control Designs

Maria Marino-Galarraga*

Thomas J. McAvoy*

Thomas E. Marlin**

*Department of Chemical Engineering
University of Maryland
College Park, MD 20742

and

**Stone and Webster Engineering Company
245 Summer Street
Boston, MA 02107

ABSTRACT: In this paper a methodology for the evaluation and comparison of several control strategies at different stages of the design is developed. The two parts of this methodology extend the previously developed short-cut operability analysis. The first part of the methodology is based on steady state information. It uses the Relative Disturbance Gain (RDG) in conjunction with the Relative Gain Array (RGA) to distinguish between clearly unfavorable control schemes and those which might result in acceptable transient performance. Also, it uses a steady state approach to decide whether a decoupling controller is beneficial and what its structure should be. The second part of the methodology makes use of approximate dynamic information to evaluate which control structures meet the process specifications. Internal Model Control (IMC) with the assumption of a perfect process model is used as an upper bound for the performance obtainable with advanced control.

I. INTRODUCTION

Practicing control engineers are often faced with the challenge of designing control systems using several single loop PID controllers or advanced controllers. In carrying out this design to determine, for example, loop pairing, interaction among the loops and the nature of the disturbances affecting the system should be considered. At the design stage of a process only steady state information is typically available. The use of dynamic simulations for evaluating the performance of different designs or control schemes requires dynamic process models. The cost of developing these models is frequently not justified. In some cases experience with the control of similar units can be used. However, there exists a need for a technology to select among a number of possible control schemes when only steady state and limited dynamic information is available.

The Relative Gain Array (RGA) (Bristol, 1966) is one method that has been widely used to evaluate control systems and decide on loop pairing. A shortcoming of the RGA is that it does not consider the nature of the disturbances affecting the system. In Part I of this series (Stanley et al., 1985) a new, steady state operability index, the Relative Disturbance Gain (RDG), was introduced. The RDG is similar to the RGA in that it is a dimensionless, scale independent index which does not depend on controller design. As such it is an ideal tool for evaluating control system operability. One of the most important uses of the RDG is to decide if interaction resulting from a disturbance is favorable or unfavorable. In Addition, Marino-Galarraga et al. (1985) defined a dynamic RDG and showed that for 2x2 systems this dynamic RDG is a fundamental parameter in determining the transient performance of a control design.

The dynamic response of control systems depends on the disturbances occurring in the plant. In this paper, the disturbances considered are step upsets. The reason for this choice is that step upsets place a great demand on the plant design. Some of the typical step disturbances in process control systems are changes in upstream units, such as feed quality and production changes, utility changes such as steam upsets, and equipment changes such as boiler and compressor trips. In our analysis the step upsets pass through a dynamic block before affecting the process output.

In this paper a methodology for the evaluation of the dynamic performance of multiloop control systems is developed. In principle this methodology is applicable to nxn systems, but to date only 2x2 control systems have been studied. The procedure is divided into two parts: a preliminary procedure based on steady state information, and a more detailed procedure based on approximate dynamic information. The steady state methodology uses the RDG in conjunction with the RGA to yield considerable insight into the behavior of multiloop control systems. This steady state approach provides a screening tool for the evaluation of different control schemes at early stages of the design. The methodology developed allows the designer to obtain estimates of the transient performance of the system as well as an indication of the control structure required to meet the control specifications. The RDG can be used to determine the need for decoupling and to decide on the decoupling structure.

If approximate dynamic information about the process is available, further evaluation can be carried out in order to determine the best control structure for the process under study. The Internal Model Control (IMC) structure with the assumption of perfect process model is used in this study as an upper bound for the performance obtainable with advanced control. Garcia and Morari (1982) have shown that the Internal Model Control structure (IMC) provides a unified basis for advanced control techniques that have proven useful in solving industrial control problems. These techniques are dynamic matrix control (DMC) (Cutler and Ramaker, 1980; Prett and Gillette, 1980; Garcia, 1984) and model algorithmic control (MAC) (Mehra et al., 1981). Two different procedures for the IMC controller design are considered. The method originally presented by Garcia (1982) is briefly reviewed. This method provides optimal tuning guidelines for the case when decoupling is beneficial. Then, a new method is presented in which, by making use of the interaction effects between the variables for a given disturbance, considerable improvement over Garcia's approach is obtained for the examples treated. The performance of the ideal IMC controllers is compared to the performance of classical controllers (PI). If the best possible response obtained with the IMC controller is not significantly superior to that obtained with classical controllers, then it is concluded that advanced control is not the best choice for the process under study. On the other hand, if the best IMC performance is considerably superior to that of classical control schemes, advanced control holds promise of increasing system performance. Further simulations could be performed in order to evaluate how much improvement can be expected during actual implementation.

II. BRIEF REVIEW OF THE RELATIVE DISTURBANCE GAIN (RDG)

In Part I (Stanley et al., 1985) the Relative Disturbance Gain (RDG) was defined, and its properties were discussed. The RDG is an nxn index. For the 2x2 system shown in Figure 1 the RDG for loop 1 is defined for a step disturbance as:

$$\beta_1 = \frac{\left.\frac{\partial m_1}{\partial d}\right|_{y_1\ y_2}}{\left.\frac{\partial m_1}{\partial d}\right|_{y_1\ m_2}} = \frac{1}{1 - \frac{K_{12}K_{21}}{K_{11}K_{22}}} \left(1 - \frac{K_{F2}K_{12}}{K_{F1}K_{22}} \right) \tag{1}$$

The RDG, β_1, is similar to the Relative Gain in that it can be calculated from steady state information. The RDG is also independent of controller design and it is scale independent.

The most important property of the RDG is its relationship to the integral of the error occurring after step forcing. Consider the case where an upset occurs and only loop 1 is operational. The integral of the error in this case is:

$$IE_{SL,1} = \int_0^\infty \varepsilon_1 \, dt \tag{2}$$

where the subscript SL refers to single loop, and the subscript 1 indicates which error is integrated. Now consider the case where both loops are operational and the same upset occurs. The integral of the error in this case is:

$$IE_{ML,1} = \int_0^\infty \varepsilon_1 \, dt \tag{3}$$

where the subscript ML refers to multiple loop operation. As shown in Part I (Stanley et al., 1985) β_1 is related to $IE_{ML,1}$ as

$$IE_{ML,1} = IE_{SL,1}\ f_1\ \beta_1 \tag{4}$$

Equation 4 is the key result of Part I. Equation 4 shows that $IE_{ML,1}$ is related to $IE_{SL,1}$ through the multiplication of two dimensionless parameters, f_1 and β_1. The f_1 parameter contains information on how much the multiloop controllers have to be detuned relative to the single loop controllers. A simple approach to estimating f_1 from steady state information for 2x2 systems is given in Part II (Marino-Galarraga et al., 1985a). The parameter β_1 provides an indication of whether or not interaction resulting from a disturbance is favorable or unfavorable. β_1 can also be calculated from steady state information. Thus, two of the three terms on the right hand side of equation 4 can be easily calculated. The third term, $IE_{SL,1}$, is also straightforward to calculate. Consider the case where the process G_{11} is first order with dead-time. A number of different PID tuning rules have been published for such system (eq. Ziegler-Nichols, 1942; Lopez, 1968). If one uses one of these tuning methods it is straightforward to show that:

$$IE_{SL,1} = \frac{T_{R1}K_{F1}}{K_{c1}K_{11}}\ d \tag{5}$$

where K_{c1} and T_{R1} are the controller gain and reset time, d is the size of the step upset, and K_{F1} and K_{11} are steady state process gains. In order to calculate K_{c1} and T_{R1}, an estimate of the process time constant and dead-time is required.

Equation 4 can be used to compare alternative control schemes. As discussed in Part I, a necessary condition for good transient performance is that $IE_{ML,1}$ be small. This condition is not sufficient however, since positive and negative areas may cancel. In comparing alternative control schemes all three terms on the right hand side of equation 4 are important, and they should be calculated if possible. For example, if the $IE_{SL,1}$ differs drastically for two schemes, then the $IE_{ML,1}$ will be profoundly affected. Shinskey (1985) gives examples for dual composition control where the individual schemes have widely differing $IE_{SL,1}$'s and f_1's. In the discussion below, it is assumed that the individual $IE_{SL,1}$'s and f_1's are roughly equal. This assumption allows us to focus on the contribution of β_1 to $IE_{ML,1}$.

III. STEADY STATE OPERABILITY ANALYSIS

a) Assessment of the Performance of Classical Control Systems

The RDG, used in conjunction with the RGA, provides valuable information about the dynamic behavior of classical multiloop (PI) control systems. The main objective of the steady state analysis is to distinguish between clearly unfavorable control schemes and those which could result in acceptable transient performance for the expected disturbances. This analysis provides a first screening tool, allowing the designer to reduce the number of control schemes requiring further evaluation. Several numerical examples for distillation and chemical reactor control have been studied (Stanley et al. 1985, Marino-Galarraga et al. 1985).

Further insight into the shape of closed-loop responses can be obtained by analyzing the effects of interaction. Two parameters are used in this analysis; they are the RGA and the RDG. Previous studies (Shinskey, 1979) have demonstrated that systems with large RGA's have long transients and systems with small RGA's respond rather quickly when both are compared to the single-loop transients. This behavior is for each controlled variable and any disturbance since the RGA is a general parameter for the system. It has also been demonstrated (Stanley, et al. 1985) that systems with large RDG's have large integral errors and those with small RDG's have integral errors similar in magnitude to the single variable values. The integral error is specific to each controlled variable and disturbance.

This qualitative information can be used to predict the shape of transient responses. The general 2x2 system in Figure 1 is considered; the system is subjected to a disturbance d which can affect both controlled variables. Since the system is linear, the overall response to the upset is the sum of the responses from the disturbance through F_1 (with $F_2 = 0$) and the response from the disturbance through F_2 (with $F_1 = 0$). By analyzing each disturbance path separately, the effects of interaction can be clearly ascertained.

A sample transient is given in Figure 2. The response of the controlled variable is the sum of the two individual upset paths. In this case, the RGA is large, which causes each individual upset response to be sluggish at large times. However, the RDG is small which indicates that the two individual responses cancel to yield a small integral error. Specific physical systems are now analyzed to cover the possible combinations of RGA and RDG.

Four different cases are possible, depending on the size of the RDG and RGA. The range of values for these cases is somewhat arbitrary since desirable control performance varies from problem to problem. For this paper, large RGA values will be taken as 5 or higher and large RDG 2 or higher. The reason for these choices is that large RGA values are associated with sluggish responses of the individual disturbance paths, and this effect is more pronounced for values of RGA greater than 5. Similarly, large RDG values indicate that interaction is unfavorable, since the individual disturbance paths do not cancel. The deterioration of the transient performance is most noticeable for

RDG values greater than 2. Similarly, small RDG values will be taken as those smaller than 2, and small RGA values will be assumed to be close to unity, since smaller RGA values (i.e., less than 0.5) indicate poor pairing of the controlled and manipulated variables. The four different combinations for the RGA and RDG are discussed below for the case of 2x2 systems.

Case 1: Large RGA and small RDG.

When the RGA is a large positive number, the response to the individual disturbance paths is very sluggish. However, a small RDG value indicates that the effect of interaction among the loops is favorable, and that the individual paths for the disturbance are of opposite direction and cancel each other. To illustrate this case an example published by Marino-Galarraga et al. (1984) is used. In their paper a simplified model for a two-bed catalytic reactor originally studied by Foss et al. (1980) was developed. The system has an RGA value of 13.7, and when it is subjected to a step disturbance in m_1, the resulting RDG values are $\beta_1 = 1$ and $\beta_2 = 0$. The response of the simplified model to this disturbance is shown in Figure 3. Since in general the two disturbance paths, F_1 and F_2, will have different dynamics, one can usually expect an initial transient in the form of a peak followed by a long tail. The peak is produced by the initial response of one of the individual paths before the other path can respond due to its dead-time. The long tail occurs as a consequence of the cancellation of the individual sluggish transients at large times. If the RDG is close to zero the tail will have a sign opposite to that of the initial peak. Conversely, if the RDG value for the system is close to unity, the long tail will have the same sign as the initial transient. This behavior has been observed in many systems with large RGA values subjected to disturbances resulting in small RDG values.

Stanley (1986) studied several distillation towers and control schemes via non-linear simulation. The response of one of the towers studied by Stanley, tower 2, to a step disturbance in the feed composition ($\beta_1 = 0.21$ and $\beta_2 = 0.79$) is reproduced in Figure 4. This control system had an RGA of 38. As can be seen, the response of the system is of the form of a sharp peak followed by a long tail. The large RGA value indicates that the individual disturbance paths are very sluggish, and the small RDG values show that the two paths cancel resulting in a small area under the response curve.

One of the properties of the RDG is that disturbances entering the system through one of the manipulated variables will always result in RDG values of zero and one (Marino-Galarraga et al., 1985). Therefore, systems with large RGA values subjected to disturbances in the manipulated variables will always show the type of transient behavior indicated in Figure 3. For example, one common disturbance for distillation towers are upsets in the condenser and reboiler. When an energy balance scheme is used for control (i.e., distillate composition controlled by manipulation of reflux flow, and bottoms composition controlled by manipulation of vapor flow), these disturbances are proportional to changes in the manipulated variables. Therefore, these two disturbances will always result in RDG values of zero and one, producing transient responses of the type shown in Figures 3 and 4.

Case 2: Large RGA and large RDG.

When both the RGA and the RDG are large, the interaction effects among the loops for the disturbance being considered are clearly unfavorable. The large RGA value indicates that the response to the individual disturbance paths is very slow. In addition, the large RDG value indicates that interaction is unfavorable since the individual disturbance paths do not cancel each other. In this case one can expect a very sluggish response.

An example of this type of system is the case of distillation towers controlled using an energy balance scheme subjected to step disturbances in the feed flow. Energy balance control results in large RGA values, and disturbances in the feed flow have large RDG values for this control strategy. The response of tower 2 (λ = 38) to a step in the feed flow (Stanley, 1986) is shown in Figure 5. The RDG values for this disturbance are $\beta_1 = -5$ and $\beta_2 = -5.3$. As can be seen, the response of the system is extremely sluggish.

Another case in which both the RGA and RDG values are large is when a system with a large RGA value is subjected to setpoint changes in one loop. For example, an on-line optimizer could send set point changes to a control system and some changes may only involve one variable. As discussed in Part I (Stanley et al., 1985), the RDG value for the loop subjected to a setpoint change is equal to the RGA. Therefore, systems with large RGA values subjected to single setpoint changes will show sluggish responses. An example of this case is the two-bed catalytic reactor with an RGA value of 13.7, discussed before (Marino-Galarraga et al., 1984). Figure 6 shows the response of the simplified model of the system for a unit setpoint change in the first loop. In this case the RDG for the disturbed loop, β_1, is equal to the RGA. As predicted by the RDG, the response of the system is very sluggish, and the integral of the error calculated relative to the final steady state is large. Although slow, this response is about the best possible with two PI controllers; any significant change in tuning to speed up the response leads to instability. One can conclude that a system with large RGA and RDG values is clearly unfavorable in terms of its transient response. For cases where more than one setpoint is changed it is straightforward to calculate the RDG and carry out the analysis. This particular case has important implications for on-line optimization since it may not be feasible when setpoint changes are frequent and the RGA is large.

In this case the RGA value indicates that the responses of the individual disturbance paths are rapid. However, a large RDG value indicates that the interaction effects are not favorable, since the individual paths are in the same direction and do not cancel.

One example of this type of system is the distillation tower model presented by McAvoy and Weischedel (1981) and labeled tower A by these authors. When a material balance control scheme is used, this system has a RGA value of 0.611. When a disturbance in the feed flow enters the system, the RDG value for the first loop, β_1, is equal to 2.01, indicating that the interaction effects for this loop are unfavorable. The response of the first loop is reproduced in Figure 7, with the corresponding individual paths shown as dashed lines. As can be seen, the individual paths are in the same direction, resulting in a high peak. In this case, even though the individual responses are well behaved, the large RDG value indicates a deterioration in the transient response due to interaction since the individual paths are in the same direction and result in a large area under the response curve.

Case 4: Small RGA and small RDG.

In this case the small RGA value indicates that the individual disturbance paths are rapid. In addition, the small RDG value indicates that the interaction is favorable and that the individual paths cancel each other and one can expect good transient behavior.

One example of this case is the distillation tower reported by Wood and Berry (1973). The RGA value for this system is 2.01, and for a unit step disturbance in feed flow, the RDG for the first loop is $\beta_1 = -0.5$. The transient response of the first loop is reproduced in Figure 9. As predicted by the steady state analysis, the individual paths cancel each other producing a well behaved transient. Another example of systems in this category are control systems with small RGA values subjected to disturbances in the manipulated variables, since these disturbances result in RDG values of zero and one. However, the transient response of the system depends on the way in which the individual disturbance paths cancel each other. For example, Figure 8 shows a plot of the response of two control schemes to a step disturbance for the case of a small RGA value and an RDG of zero. For both schemes the steady state interaction is favorable. However, for scheme 2 the dynamic interaction effects are more favorable than for scheme 1. In order to characterize the dynamic behavior of the system, further analysis has to be carried out.

b) Evaluation of the Control Structure

In addition to predicting the control performance of multiloop PID systems, the RDG can be used to decide if decoupling should be used in a particular system. A steady state analysis can provide an indication of the need for decouplers and the required decoupler structure. As shown by equation 4 the RDG can be related to the net area under the error response curves for the cases of multiloop and single loop control. The reduction in controller parameters, f_1, can be estimated from steady state information using the approach given in Part II (Marino-Galarraga et al., (1985a)). If the calculated area ratios ($|f_1 \beta_1|$) for the important disturbances are small, interaction is favorable for the ith loop, and that loop should not be decoupled from the other loop. If the area ratio is large ($|f_1 \beta_1| > 1$), the opposite conclusion holds. To illustrate the use of the operability method for deciding about the decoupler structure, a distillation model published by Wood and Berry (1973) is considered. The column was used for methanol-water separation and was found to be well described by the transfer function model:

$$y(s) = G(s)\ m(s) + F(s)\ d(s)$$

The transfer function matrices were determined by experimentally pulse testing the tower (Wood and Berry, 1973):

$$G(s) = \begin{bmatrix} \dfrac{12.8e^{-s}}{16.7s+1} & \dfrac{-18.9e^{-3s}}{21.0s+1} \\ \dfrac{6.6e^{-7s}}{10.9s+1} & \dfrac{-19.4e^{-3s}}{14.4s+1} \end{bmatrix} \tag{7}$$

$$F(s) = \begin{bmatrix} \dfrac{3.8e^{-8.1s}}{14.9s+1} \\ \dfrac{4.9e^{-3.4s}}{13.2s+1} \end{bmatrix} \tag{8}$$

From the steady state gains the RGA and RDG values can be calculated as:

$$\lambda = 2.01 \tag{9}$$

$$\beta_1 = -0.5 \tag{10}$$

$$\beta_2 = 1.2 \tag{11}$$

The reductions of the multiloop controller settings with respect to the SISO settings were calculated using the steady state short-cut tuning method presented in Part II (Marino-Galarraga et al., 1985a) and they are:

$$f_1 = 1.55 \tag{12}$$

$$f_2 = 1.55 \tag{13}$$

From these f_i's the following area ratios can be estimated:

$$f_1\ \beta_1 = -.773 \tag{14}$$

$$f_2\ \beta_2 = 1.85 \tag{15}$$

These $f_i\ \beta_i$ values indicate that the performance of loop 2 can be improved ($|f_2 \beta_2| > 1$) via a one-way decoupler from loop 1. When such a decoupler is used, then the decoupled system behaves like a single loop system, and it is straightforward to show:

$$f_2 \beta_2 = 1.0 \tag{16}$$

Figure 10 shows the block diagram of a one-way decoupler from loop 1 to loop 2. The equation for the decoupler D_{21} is:

$$D_{21} = -G_{21}/G_{22} \tag{17}$$

Figure 11 compares the performance of the one-way decoupled and non-decoupled systems. As can be seen the performance of loop 2 is improved in agreement with the area ratio calculation. Now consider where loop 1 is decoupled from loop 2. In this case the arrows on the decoupler in Figure 10 are reversed and the transfer function implemented is:

$$D_{12} = \frac{-G_{12}}{G_{11}} \tag{18}$$

The area ratio, $f_1 \beta_1 = -.773$, indicates that the decoupler will not improve the x_D response and in fact the area will increase since $f_1 \beta_1$ for the decoupled system will be 1.0. Figure 12 shows the response of this decoupler compared to the original multiloop system. As can be seen the decoupled response is indeed inferior, in agreement with the steady state $f_1 \beta_1$ calculation. In this case one can obtain valuable information about control system configuration based only on steady state information.

The same approach can be applied to the question of whether two way simplified decouplers should be used. For two way simplified decoupling one has to note that the decoupled loop transfer functions are no longer G_{ii} but $G_{ii}/\lambda(s)$ (McAvoy 1983). Thus, in calculating f_i's for two way decoupled loops one has to account for this change in transfer function. While the addition of one decoupler to a loop could help a loop, the addition of a second decoupler away from a loop, which alters the loop's transfer function, may actually be deleterious.

The results presented in this section show the power of the RDG in analyzing control system performance. In addition, the limited cases studied to date seem to indicate that for disturbances resulting in small RDG values, the behavior of classical control systems (PI) is similar to that of the ideal IMC controller with the assumption of perfect model. This point will be developed further in the following section.

IV. OPERABILITY ANALYSIS USING APPROXIMATE DYNAMIC INFORMATION

If an approximate dynamic model is available, a more detailed operability procedure can be used to determine the best control structure for the process. In some cases approximate dynamic information can be easily developed using short-cut procedures. Kapoor et al. (1986) have developed analytical expressions for the evaluation of dynamic models for distillation towers. Using these approximate dynamic models the need for advanced control can be evaluated.

The methodology presented in the previous sections allows the designer to obtain estimates of the transient performance of the system when classical control (PI) and decouplers are used. In this section IMC control is also evaluated. The Internal Model Control (IMC) structure with the assumption of perfect process model is used as a rough upper bound for the performance obtainable with advanced control.

Internal Model Control (Garcia and Morari, 1982) is a recently developed process control technique. This new control structure addresses specific issues such as regulatory behavior, servo behavior, robustness, and the ability to deal with constraints on the controlled and manipulated variables.

The structure of the IMC controller is shown in Figure 13. The basic theoretical principles and design procedures of IMC are described in detail by Garcia (1982). The multivariable IMC controller is an alternative to using classical multiloop PID controllers. The stability analysis for IMC controllers is very simple, and the design procedure consists in the development of proper factorizations of the process model (Morari, 1983).

In this work it is assumed that the short-cut dynamic model of the process is perfect. The resulting controller represents the best performance that can be obtained with the IMC structure. In the general case the actual performance will show some deterioration due to modeling errors. Additional deterioration in the performance results from simplifications in the design, since the controller is realizable but not always stable. Making the controller stable during actual implementation results in some deterioration in the performance.

Two different procedures for IMC controller design for 2x2 systems are presented here. The method originally presented by Garcia (1982) is briefly reviewed. This method provides optimal tuning guidelines for the case when decoupling is beneficial. Then, a new method is presented which makes use of the interaction effects between the variables for a given disturbance to improve transient performance.

General IMC Design Procedure

The output vector of a 2x2 IMC system, y(s), can be obtained from the block diagram of the IMC structure (Figure 13):

$$y(s) = G(s)\ G_c(s)\ [I + (G(s) - G_m(s))\ G_c(s)]^{-1}\ [y_s - G_F(s)\ d] + G_F(s)\ d \tag{19}$$

If the model is assumed to be perfect $G(s) = G_m(s)$, the output of the system is:

$$y(s) = G_m(s)\ G_c(s)\ [y_s - G_F(s)\ d] + G_F(s)\ d \tag{20}$$

For perfect control the IMC controller must be equal to the inverse of the process model:

$$G_m(s)\ G_c(s) = I \tag{21}$$

If the controller is designed such that:

$$G_m(0)\ G_c(0) = I \tag{22}$$

the system will have no offset (Garcia, 1982).

In the general case, the controller cannot be implemented as the exact inverse of the process model. If the model contains time delays, the resulting controller will be unrealizable. In addition, if the model has zeros in the right half plane, the controller will be unstable. The following factorization is suggested by Garcia (1982) and Morari (1983):

$$G_m(s) = G_+(s)\ G_-(s) \tag{23}$$

with

$$G_+(0) = I \tag{24}$$

for zero offset. The purpose of this factorization is to make the controller;

$$G_c(s)=G_-^{-1}(s) \tag{25}$$

stable and realizable. In general, the factorization is difficult to perform, and guidelines have not been reported. The difficulties arise because of the presence of right half plane zeroes and dead times in the process model matrix. Each situation is analyzed separately in the following sections.

<u>Factorization of Zeroes</u>

From equations 23 and 25 the controller transfer function is:

$$G_c(s) = G_-^{-1}\ (s) = \frac{\mathrm{adj}[G_m(s)]\ \ G_+(s)}{\det[G_m(s)]} \tag{26}$$

where,

$$\det[G_m(s)] = G_{11}(s)\ G_{22}(s) - G_{12}(s)\ G_{21}(s) \qquad (27)$$

If all the time delays are equal, they can be factored out of the process model. For an open-loop stable process, $\det[G_m s]$ contains all the right half plane zeroes as factors. Then, $G_+(s)$ should be selected to cancel the right half plane zeroes of the system in equation 26, as well as the transmission zeroes. Holt and Morari (1985) presented some guidelines for the selection of $G_+(s)$ when all delays are equal. However, in the general case the time delays of the process will all be different. As can be seen in equation 26, each of the elements of the controller matrix is divided by the determinant of the process matrix. The presence of the different time delays in each of the elements of the process matrix makes the calculation of the zeroes very difficult if not impossible. Because of the difficulties in the factorization of the transmission zeroes, Garcia (1982) suggests a modified factorization of the process model:

$$G_m(s) = G'_+(s)\ G'_-(s) \qquad (28)$$

with

$$G'_+(0) = I \qquad (29)$$

Using this factorization gives the controller:

$$G_c(s) = [G'_-(s)]^{-1} \qquad (30)$$

which will be realizable but not necessarily stable. In actual implementation $G_c(s)$ must be approximated in order to obtain a stable controller.

<u>Factorization of Time Delays</u>

The factorization of the time delays can be done in different ways. Holt and Morari (1985a) proposed the following theorem to determine if a diagonal G'_+ (in equation 28) is optimal:

> The optimal $G'_+(s)$ matrix is diagonal if and only if the rows and/or columns of $G_m(s)$ can be rearranged such that the smallest numerator delay in every row occurs along the diagonal.

The use of the term "optimal $G'_+(s)$ matrix" by Holt and Morari refers to the fact that the $G'_+(s)$ matrix obtained using their factorization procedure allows the minimum amount of dead-time in the control system response. The authors indicate that the dead time of the response of each loop represents a lower bound for the settling time.

It is important to point out that when the $G'_+(s)$ matrix is diagonal, the outputs of the system, y_1 and y_2, are dynamically decoupled. This can be easily seen by substituting equations 29 and 30 into 20. When $G'_+(s)$ diagonal, the output vector of the system is:

$$y(s) = G'_+(s)\ y_s + [I - G'_+(s)]\ G_F(s)\ d \tag{31}$$

Holt and Morari (1985a) propose using a diagonal $G'_+(s)$ matrix with the smallest dead times such that the controller is realizable. This matrix is of the form:

$$G'_+(s) = \mathrm{diag}\ [e^{-\alpha_{11}s}, e^{-\alpha_{22}s}] \tag{32}$$

where,

$$\alpha_{ij} = \max_i \max\ [0, (q_{ij} - p_{ij})] \tag{33}$$

and,

p_{ij} = minimum delay in numerator of element ij of $G_m^{-1}(s)$

q_{ij} = minimum delay in denominator of element ij of $G_m^{-1}(s)$

Interacting System

In the previous sections it was shown that in the case of PI controllers interaction can improve the overall performance of some systems. Further, in some cases the use of decouplers resulted in worse transient behavior for some of the disturbances. In this section the same concepts are applied in order to demonstrate that, for some disturbances, an interacting IMC system (i.e., non-diagonal $G'_+(s)$ matrix) results in better transient behavior than a decoupled IMC system or a classical PI system. The proposed non-diagonal $G'_+(s)$ matrix is of the form:

$$G'_+(s) = \begin{bmatrix} e^{-\alpha_{11}s} & [e^{-\tau_1 s} - e^{-\alpha_{12}s}] \\ [e^{-\tau_2 s} - e^{-\alpha_{21}s}] & e^{-\alpha_{22}s} \end{bmatrix} \tag{34}$$

this form satisfies the requirement:

$$G'_+(0) = I \tag{35}$$

The first step of the design procedure is the calculation of the minimum time delays of the $G'_+(s)$ matrix such that the resulting controller:

$$G_c(s) = [G'_-(s)]^{-1} \tag{36}$$

is realizable. To accomplish this, α_{ij} must be chosen such that:

$$\alpha_{ij} = \max_k \max\ [0,\ (q_{ki} - p_{ki})] \tag{37}$$

and

$$\tau_i \geq \alpha_{ij} \quad \text{for } i \neq j \tag{38}$$

where,

p_{ki} = minimum delay in numerator of element ki of $G_m^{-1}(s)$

q_{ki} = minimum delay in denominator of element ki of $G_m^{-1}(s)$

As can be seen, if the minimum values for the off-diagonal elements are chosen (i.e., $\tau_i = \alpha_{ij}$ in equation 38), $G'_+(s)$ will be diagonal and the controller will result in a dynamically decoupled response. It is important to point out that the values obtained by the above procedure are the minimum delays necessary to obtain a realizable controller. Greater time delays can be specified, if necessary, in order to improve the overall performance of the system. One approach for the calculation of the dead-times is to pose the design procedure as an optimization problem in which a performance index is minimized. Any optimization criterion could be used for the design. The integral of the Square Error (ISE) is used here because it penalizes large deviations, resulting in smaller overshoots.

The output vector of a 2x2 IMC system for the case of unit step load disturbance can be calculated from equation 31 with $y_s = 0$:

$$y(s) = [I - G'_+(s)]G_F(s) \ \frac{1}{s} \tag{39}$$

using the non-diagonal $G'_+(s)$ matrix of the form given in equation 32, equation 39 can be rewritten as:

$$\begin{bmatrix} y_1(s) \\ \\ y_2(s) \end{bmatrix} = \begin{bmatrix} 1-e^{-\alpha_{11}s} & -[e^{-\tau_1 s} - e^{-\alpha_{12}s}] \\ \\ [e^{-\tau_2 s} - e^{-\alpha_{21}s}] & 1 - e^{\alpha_{22}s} \end{bmatrix} \begin{bmatrix} \frac{G_{F1}(s)}{s} \\ \\ \frac{G_{F2}(s)}{s} \end{bmatrix} \tag{40}$$

A suitable optimization method is used to determine the values of the delays that result in the minimum ISE for both loops. It is important to point out that this method results in the minimum ISE for a particular disturbance, and that the design will not be optimum for other disturbances affecting the system.

To illustrate the use of the IMC design procedures they will be applied to the binary distillation column discussed by Wood and Berry (1973) (equation 7 and 8). The disturbance considered is a step in the feed flow rate of magnitude 0.34 lb/min.

Holt and Morari (1985a) studied this system, and since the smallest dead-times in every row occurs along the diagonal their theorem states that the optimal $G'_+(s)$ matrix is diagonal. Using equation 33, they calculated the following $G'_+(s)$ matrix:

$$G'_+(s) = \mathrm{diag}[e^{-s}, e^{-3s}] \tag{41}$$

If a non-diagonal $G'_+(s)$ of the form given by equation 34 is chosen instead, the minimum values for the delays can be calculated from equations 37 and 38 as:

$$\alpha_{11} = \alpha_{12} = 1 \tag{42}$$

$$\tau_1 \geq 1 \tag{43}$$

$$\alpha_{22} = \alpha_{21} = 3 \tag{44}$$

$$\tau_2 \geq 3 \tag{45}$$

The time delay values required to obtain an optimum ISE are calculated using the Nelder-Mead method (1964). The resulting $G'_+(s)$ matrix is:

$$G_+(s) = \begin{bmatrix} e^{-1.00s} & [e^{-5.16s} - e^{-5.89s}] \\ [e^{-3.00s} - e^{-5.26s}] & e^{-3.00s} \end{bmatrix} \tag{46}$$

The optimum ISEs for the loops are:

$$ISE_1 = 0.00114 \tag{47}$$

$$ISE_2 = 2.45 \tag{48}$$

For comparison, the ISEs for the loops using the decoupled IMC system ($G'_+(s)$ given in equation 41) are:

$$ISE_1 = .237 \tag{49}$$

$$ISE_2 = 3.80 \tag{50}$$

Figure 14 shows the response of the system to a disturbance of 0.34 lb/min in the feed flow rate using the interacting IMC controller designed for minimum ISE. For comparison, the decoupled IMC system and PI responses are also shown. As can be observed in the figure, by taking advantage of interaction, an interacting IMC controller results in much better transient behavior than that of a decoupled IMC controller or PI controllers.

The IMC system designed for optimum ISE will result in optimal responses only for feed flow disturbances. To evaluate the performance of the controller for other disturbances, Figures 15 and 16 show the responses of the interacting IMC controller to a unit step disturbance in the reflux flow (first manipulative variable) and in the bottoms steam flow (second manipulative variable) respectively. Again, the responses of the decoupled IMC controller and the PI controller are shown for comparison purposes. As can be seen the interacting IMC controller results in better transient behavior than the decoupled IMC controller. However, for these two additional disturbances, the transient performance of the PI controller is similar to that of the IMC controllers. The results of this example and a number of systems studied seem to indicate that the transient performance of PI systems with small RDG values will compare favorably to the transient performance of decoupled IMC systems. The IMC controller designed for the minimization of the ISE for a specific disturbance will always result in a better transient than a decoupled IMC controller for the same disturbance.

By using the methodology discussed here the need for advanced control can be evaluated. The Internal Model Control (IMC) structure with the assumption of perfect process model is used as an upper bound for the performance obtainable with advanced control. This methodology can be used as a screening tool to compare the performance of advanced and classical controllers at early stages of the design. Since the comparison is based on the performance of the ideal IMC controller, if the best IMC performance is considerably superior to that of classical control schemes, advanced control holds promise of increasing system performance. Further simulations could be performed in order to evaluate just how much improvement can be expected during actual implementation.

IV. CONCLUSIONS

A short-cut operability methodology for the evaluation and comparison of several control strategies at an early design stage has been developed. A preliminary steady state procedure, based on an RDG and RGA analysis, can be used at early stages of the design for the screening of the control schemes that result in acceptable transient performances. The proposed methodology also allows the designer to determine the need for decoupling and to decide on the decoupling structure. In addition, the need for advanced control can be evaluated. The IMC structure is used as an upper bound for the performance obtainable with advanced control. A new IMC design method which makes use of the interaction effects between the variables for a given disturbance is presented. The short-cut methodology presented provides valuable information for the evaluation of the performance of multiloop control systems and for the determination of the complexity of the control structure needed to achieve the control objectives.

REFERENCES

Bristol, E., IEEE Trans. Autom. Control, 1966, AC-II, 133.

Cutler, C.R. and B.L. Ramarker, "Dynamic Matrix Control - A Computer Control Algorithm," AIChE 86th National Meeting, Houston, TX, April, 1979. (also in Joint Automatic Control Conference Proceedings, 1980)

Foss, A.S., Edmunds, J.M. and B. Kouvaritakis, Ind. Eng. Chem. Fundam., 1980, 19, 109.

Garcia, C.E., Ph.D. Dissertation, University of Wisconsin, Madison, WI, 1982.

Garcia, C. and M. Morari, "Internal Model Control. 1. A Unifying Review and Some New Results,: Ind. Eng. Chem. Process Des. Dev., 1982, 21, 308.

Garcia, C.E., "Quadratic/Dynamic Matrix Control of Nonlinear Processes. An Application to a Batch Reaction Process," AIChE Annual Meeting, San Francisco, CA, 1984.

Holt, B.R. and M. Morari, Chem. Eng. Sci., 1985, 40, 59.

Holt, B.R. and M. Morari, Chem. Eng. Sci., 1985a, 40, 1229.

Kapoor, N., McAvoy, T.J. and T.E. Marlin, AIChE J., 1986, 32, 411.

Lopez, A.M., Ph.D. Thesis, Louisiana State University, Baton Rouge, LA, 1968.

McAvoy, T.J. and K. Weischedel, Proceedings of VIII IFAC Congress, Kyoto, Japan, 1981, paper 107.2.

McAvoy, T.J., "Interaction Analysis Principles and Applications," Instrument Society of America, Research Triangle Park, NC, 1983.

Marino-Galarraga, M., McAvoy, T.J. and T.E. Marlin, Proceedings American Control Conference, San Diego, CA, 1984.

Marino-Galarraga, M., Marlin, T.E. and T.J. McAvoy, Proceedings American Control Conference, Boston, MA, 1985.

Marino-Galarraga, M., T.J. McAvoy and T.E. Marlin, Ind. Eng. Chem. Process Des. Dev., 1985a, submitted.

Mehra, R.K., Rouhani, R., Eterno, J., Richalet, J. and A. Rault, Engr. Foundation Conf. on Chemical Process Control II, Sea Island, GA, 1981.

Morari, M., Chem. Eng. Sci., 1983, 38, 1881.

Nelder, J.A. and R. Mead, Computer J. 1964, 7, 308.

Prett, D.M. and Gillette, R.D., Joint Automatic Control Conference Proceedings, San Francisco, CA, 1980.

Shinskey, F.G., Proceedings American Control Conference, Boston, MA, 1985.

Stanley, G., M.S. Dissertation, University of Maryland, College Park, MD, 1986.

Stanley, G., Marino-Galarraga, M. and T.J. McAvoy, Ind. Eng. Chem. Process Des. Dev., 1985, 24. 1181.

Wood, R. and M. Berry, Chem. Eng. Sci., 1973, 28, 1707.

Ziegler, J. and Nichols, N., Am. Soc. Mech. Eng. Trans., 1942, 64, 759.

NOTATION

d = Arbitrary disturbance

D_{ij} = Decoupler elements

f_i = Detuning parameter defined by equation 7

F_i = Disturbance transfer function for loop i

G_c = IMC controller

G = Process transfer function matrix

G_F = Disturbance transfer function matrix

G_m = Process model

G_+ = Non-invertible part of the model

G_- = Invertible part of the model

IE = Integral of the error under the response curve

ISE = Integral of the square error

K_{ci} = Controller gain for loop i

K_{Fi} = Disturbance steady state gain

K_{ij} = Process steady state gain

m_i = Manipulated variable for loop i

s = Laplace variable

T_{Ri} = Reset time for loop i

y_i = Controlled variable for loop i

y_s = Setpoint vector

p_{ij} = Minimum delay in numerator of element ij of G_m^{-1}

q_{ij} = Minimum delay in denominator of element ij of G_m^{-1}

Greek Letters

α_{ij} = Dead time elements of the G+ matrix

β_i = RDG element for loop i

ε = Error under the response curve

λ = Steady-state RGA element

τ_i = Dead time elements of the G+ matrix

FIGURE CAPTIONS

Figure 1: General 2x2 System.

Figure 2: Effect of Individual Disturbance Paths on Transient Performance.

Figure 3: Response of Simplified Catalytic Reactor Model to a Unit Step Disturbance in m_1.

Figure 4: Response of Tower 2 to a Feed Composition Disturbance (Stanley, 1985).

Figure 5: Response of Tower 2 to a Feed Flow Disturbance (Stanley, 1985).

Figure 6: Response of Simplified Catalytic Reactor Model to a Unit Step in Setpoint of First Loop.

Figure 7: Response of Tower A (loop 1) to a Feed Flow Disturbance.

Figure 8: Control Schemes with Same RDG value and Different Dynamic Behavior.

Figure 9: Response of Wood and Berry Tower (loop 1) to a Feed Flow Disturbance.

Figure 10: One-way Decoupling Scheme.

Figure 11: Response of Wood and Berry Distillation Tower to a Feed Flow Disturbance.
———Loop 2 Decoupled
- - -Multiloop Performance

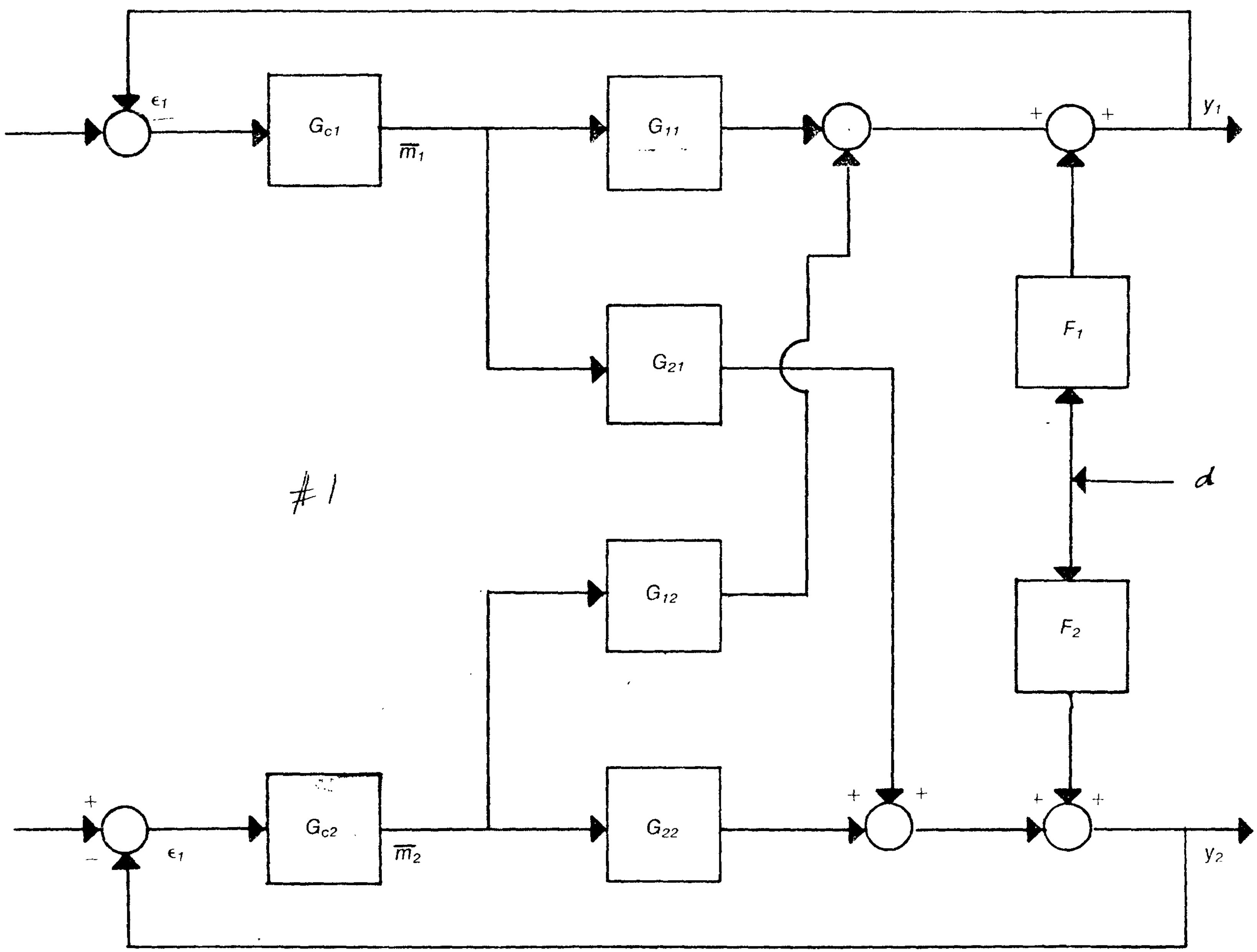
Figure 12: Response of Wood and Berry Distillation Tower to a Feed Flow Disturbance.
———Loop 1 Decoupled
- - -Multiloop Performance

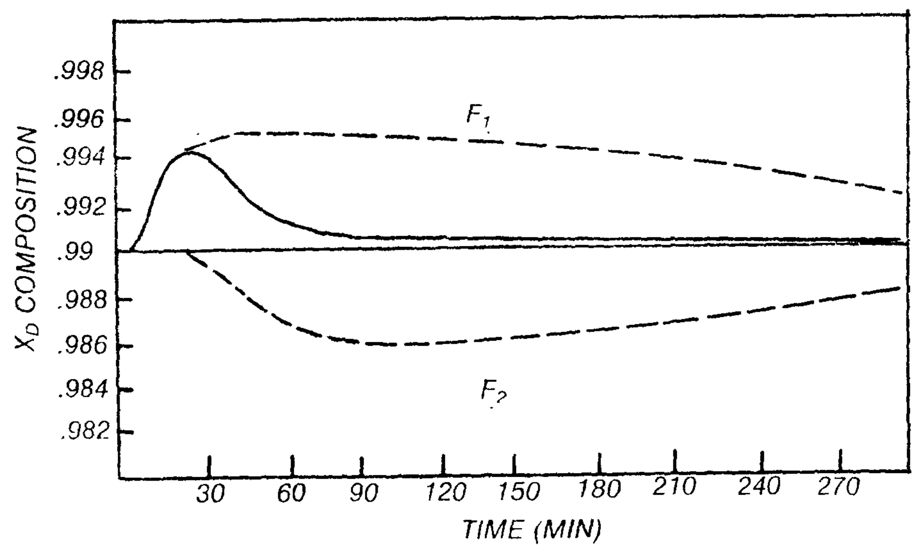
Figure 13: IMC Structure

Figure 14: Response of Wood and Berry Tower to a Feed Flow Disturbance.
———Interacting IMC
—●—●—Decoupled IMC
- - -PI

Figure 15: Response of Wood and Berry Tower to a Step Disturbance in Reflux Flow.
———Interacting IMC
—●—●—Decoupled IMC
- - -PI

Figure 16: Response of Wood and Berry Tower to a Step Disturbance in Steam Flow.
——Interacting IMC
-•-•-Decoupled IMC
- - -PI

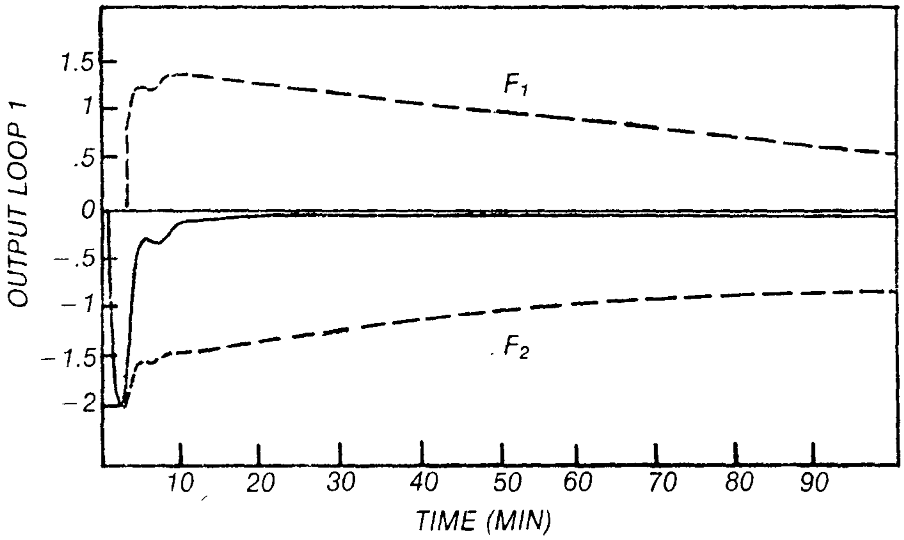

ε1
Gc1
m1
G11
G21
G12
G22
Gc2
m2
ε1
F1
F2
d
y1
y2
#1

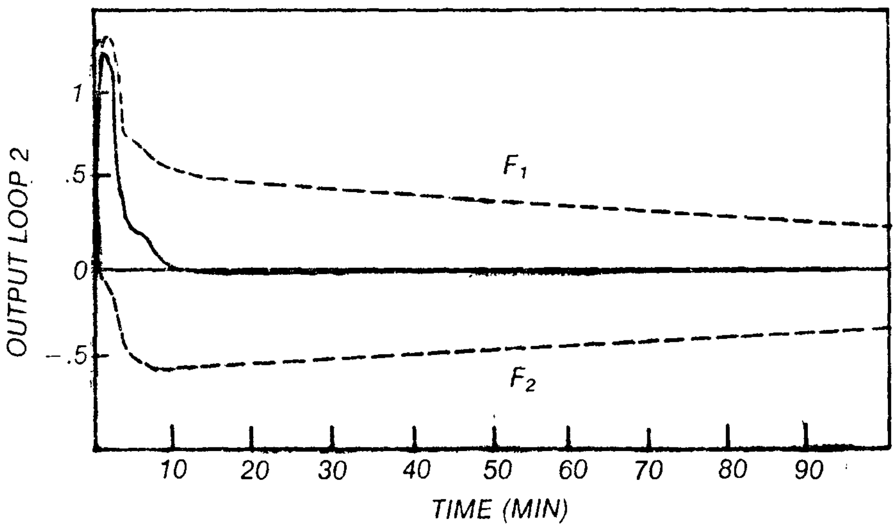


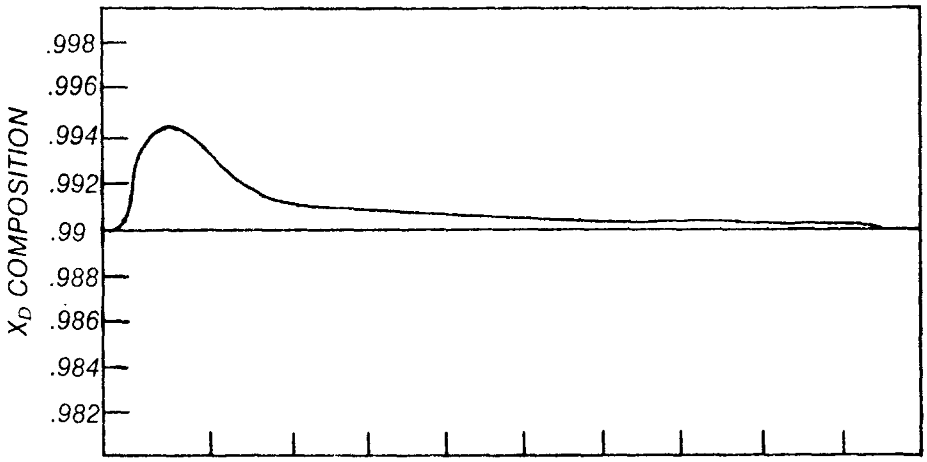
.998
.996
.994
.992
.99
.988
.986
.984
.982
X_D COMPOSITION
F_1
F_2
30
60
90
120
150
180
210
240
270
TIME (MIN)

#2

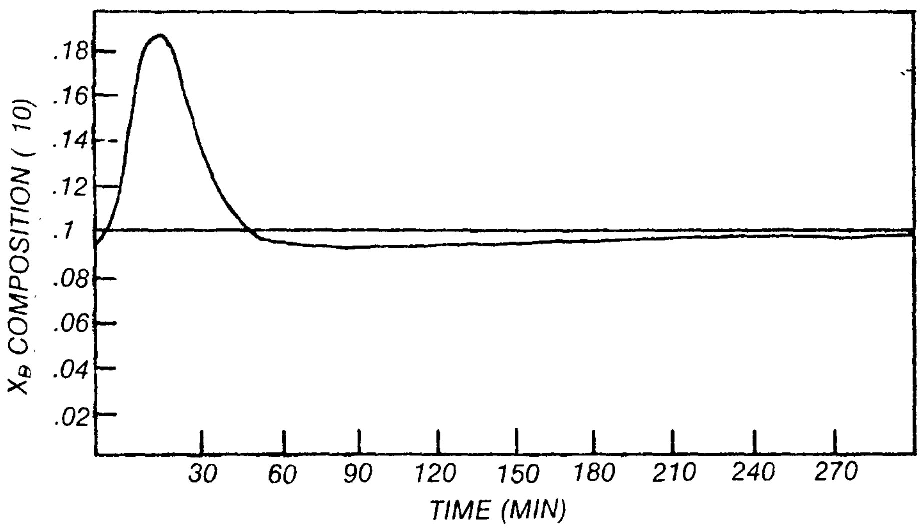
#3

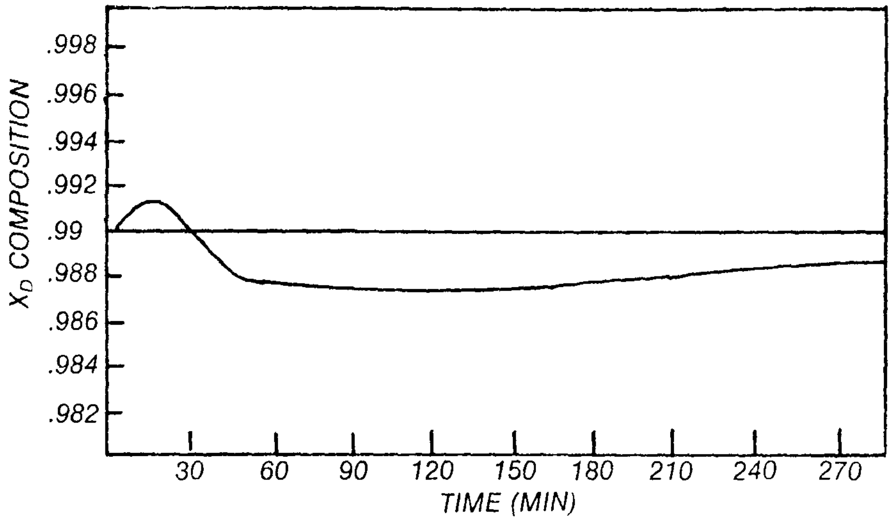

1.5
1
.5
0
−.5
−1
−1.5
−2
OUTPUT LOOP 1
F_1
F_2
10
20
30
40
50
60
70
80
90
TIME (MIN)


1
.5
0
−.5
OUTPUT LOOP 2
F_1
F_2
10
20
30
40
50
60
70
80
90
TIME (MIN)


.998
.996
.994
.992
.99
.988
.986
.984
.982
X_D COMPOSITION

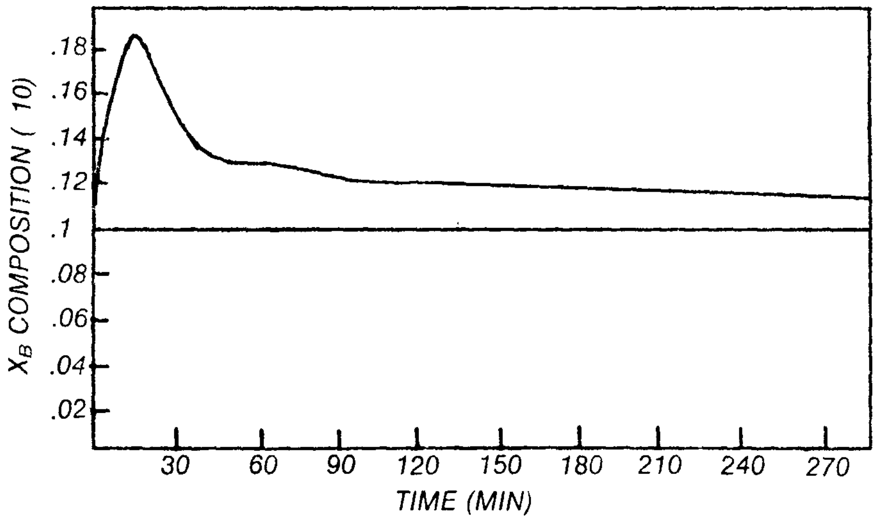
#4


.18
.16
.14
.12
.1
.08
.06
.04
.02
X_B COMPOSITION (10)
30
60
90
120
150
180
210
240
270
TIME (MIN)

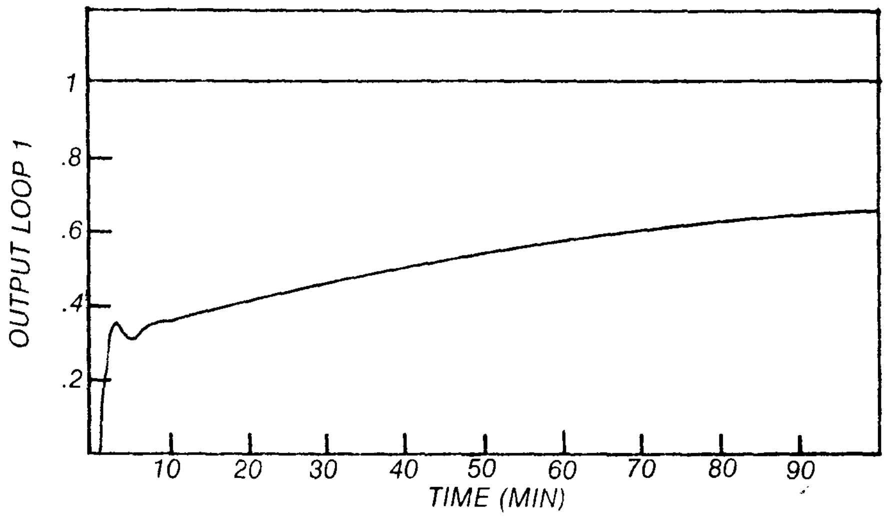


.998
.996
.994
.992
.99
.988
.986
.984
.982
X_D COMPOSITION
30
60
90
120
150
180
210
240
270
TIME (MIN)

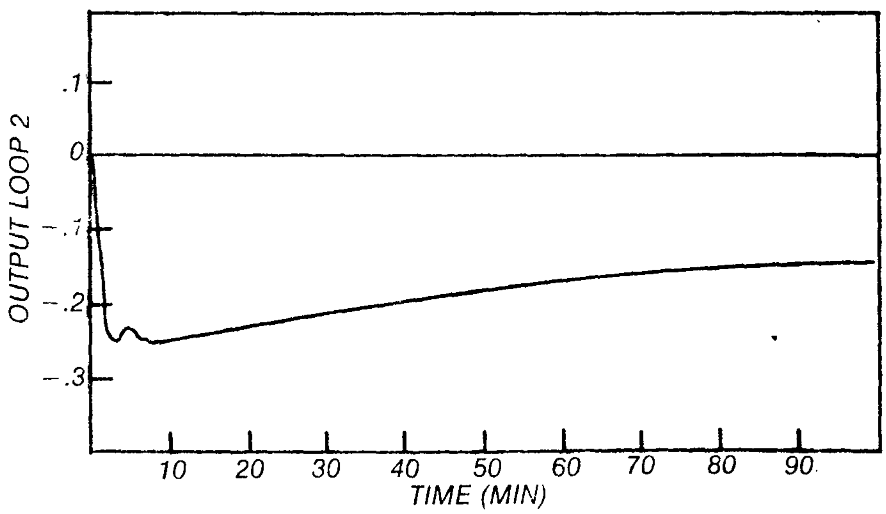
5

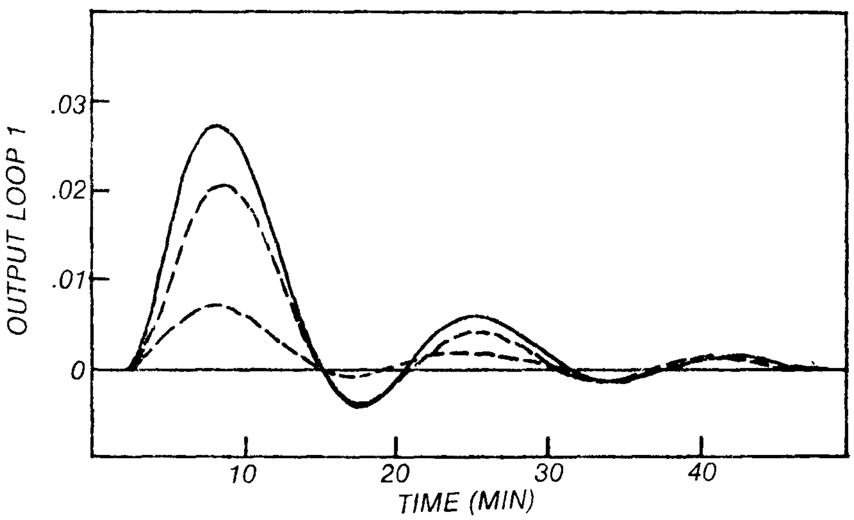


.18
.16
.14
.12
.1
.08
.06
.04
.02
X_B COMPOSITION (10)
30
60
90
120
150
180
210
240
270
TIME (MIN)

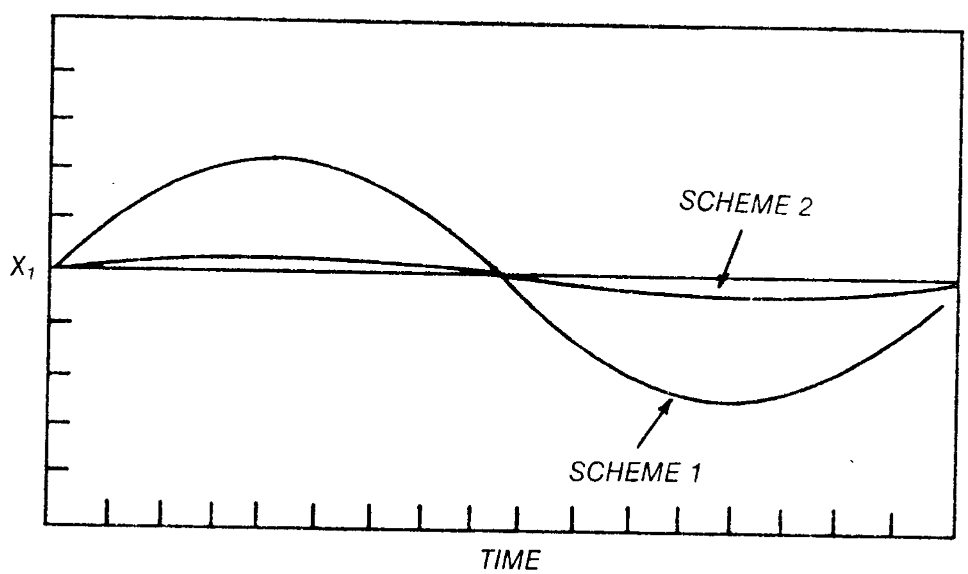

OUTPUT LOOP 1
1
.8
.6
.4
.2
10
20
30
40
50
60
70
80
90
TIME (MIN)

#6


OUTPUT LOOP 2
.1
0
−.1
−.2
−.3
10
20
30
40
50
60
70
80
90
TIME (MIN)

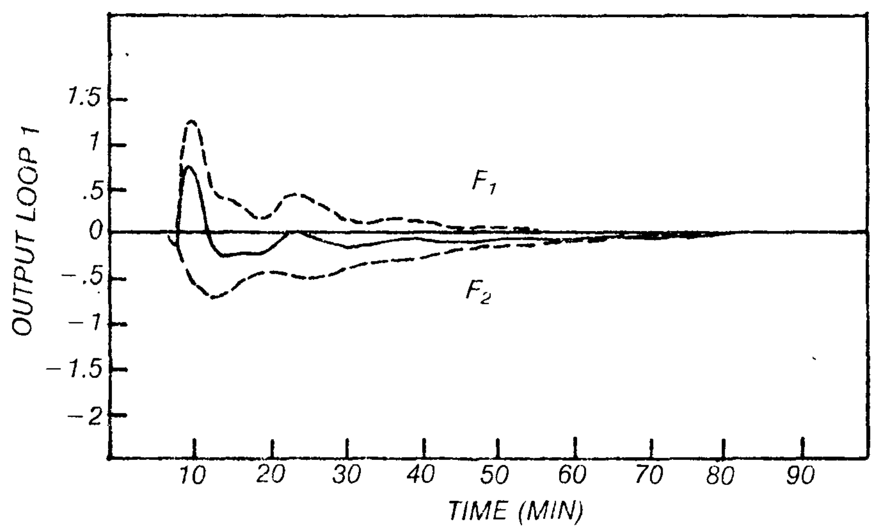


7

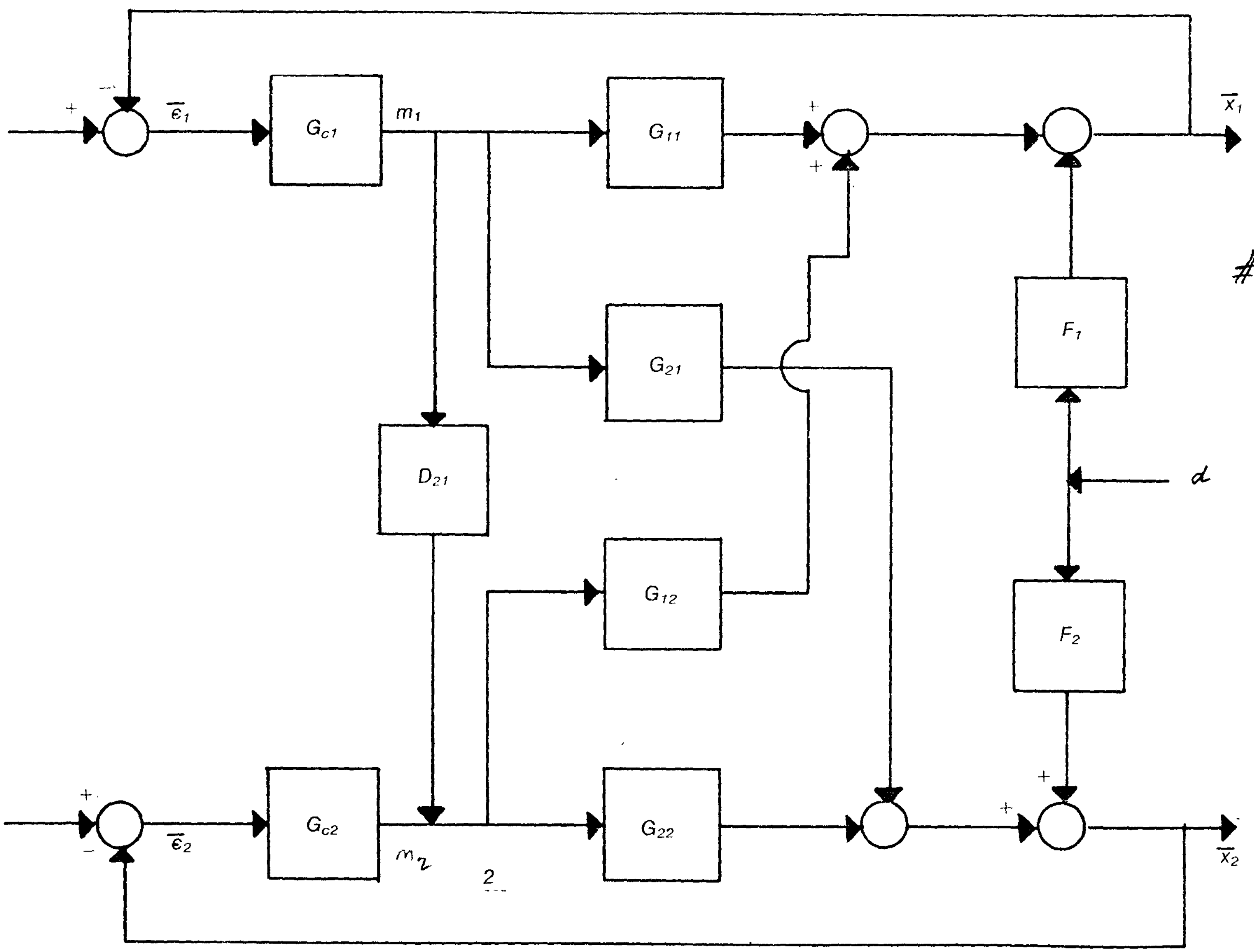


8

#9



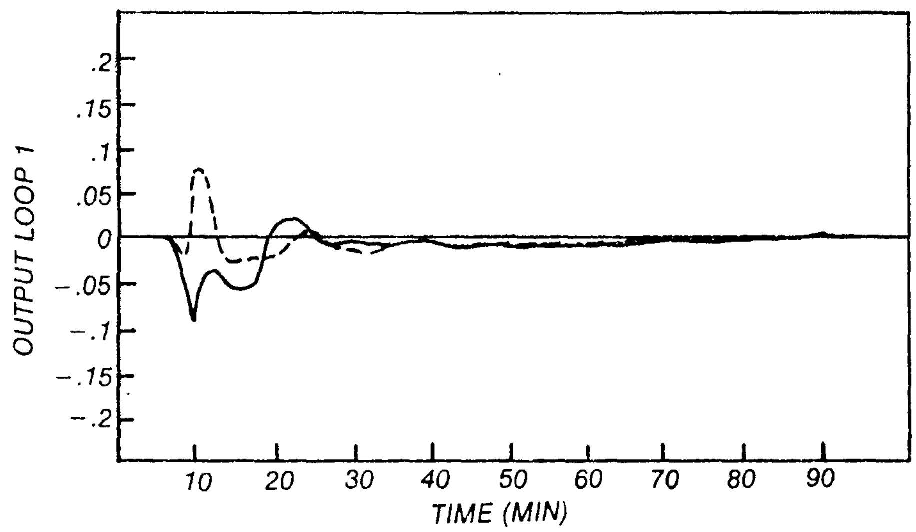
OUTPUT LOOP 1
1.5
1
.5
0
-.5
-1
-1.5
-2
F1
F2
10
20
30
40
50
60
70
80
90
TIME (MIN)

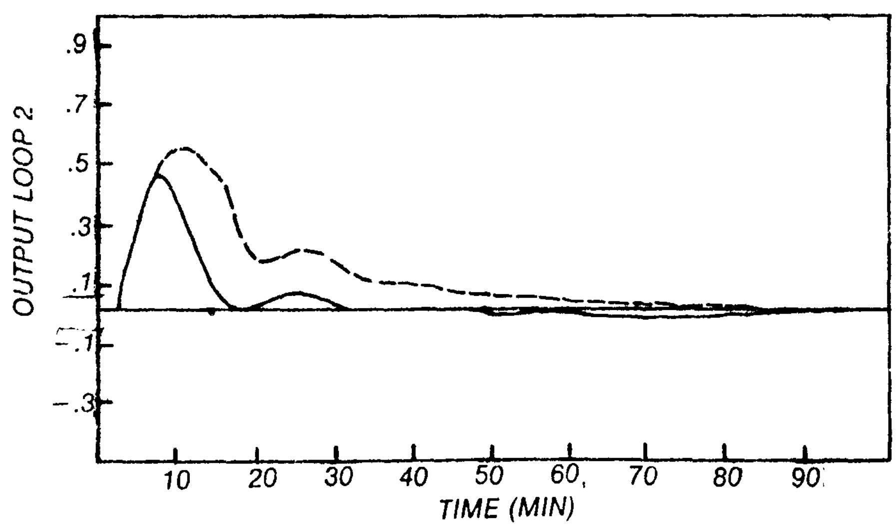


$\bar{\epsilon}_1$
G_{c1}
m_1
G_{11}
G_{21}
D_{21}
G_{12}
G_{22}
F_1
F_2
d
$\bar{x}_1$
$\bar{x}_2$
$\bar{\epsilon}_2$
G_{c2}
m_2
2

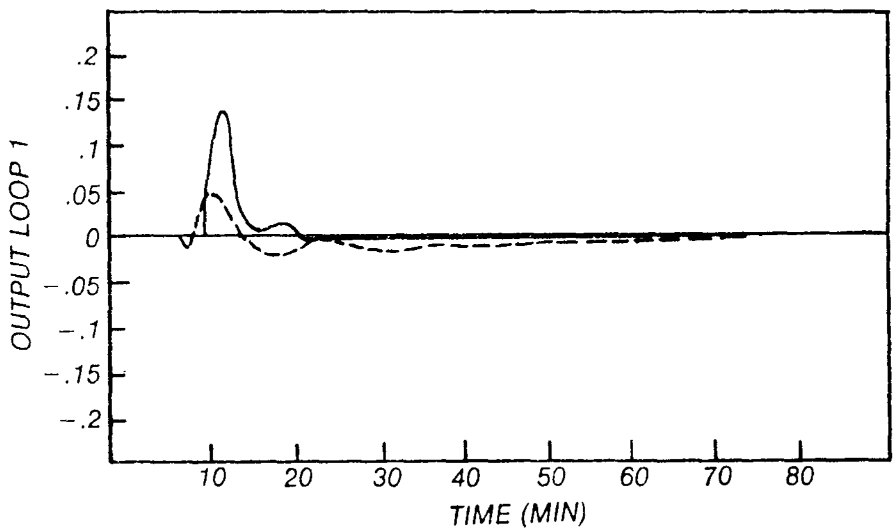
#10

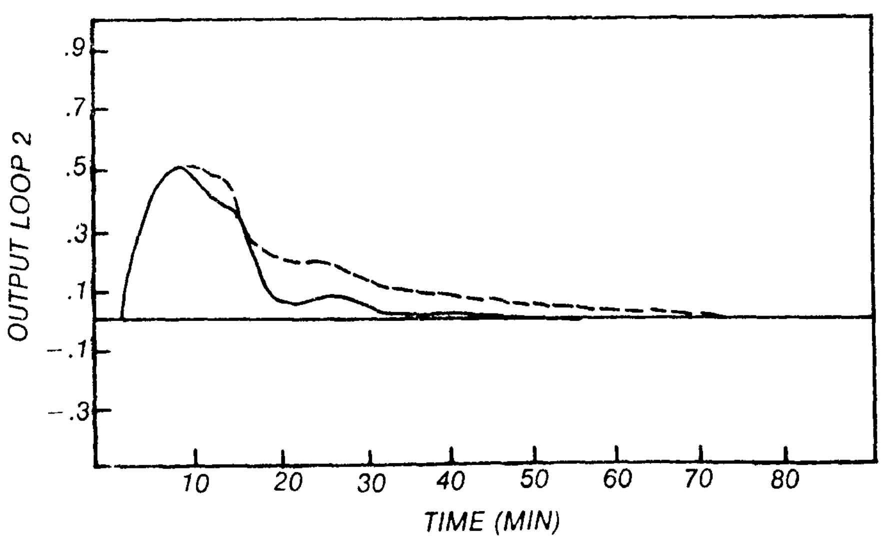
#11


.2
.15
.1
.05
0
-.05
-.1
-.15
-.2
OUTPUT LOOP 1
10
20
30
40
50
60
70
80
90
TIME (MIN)

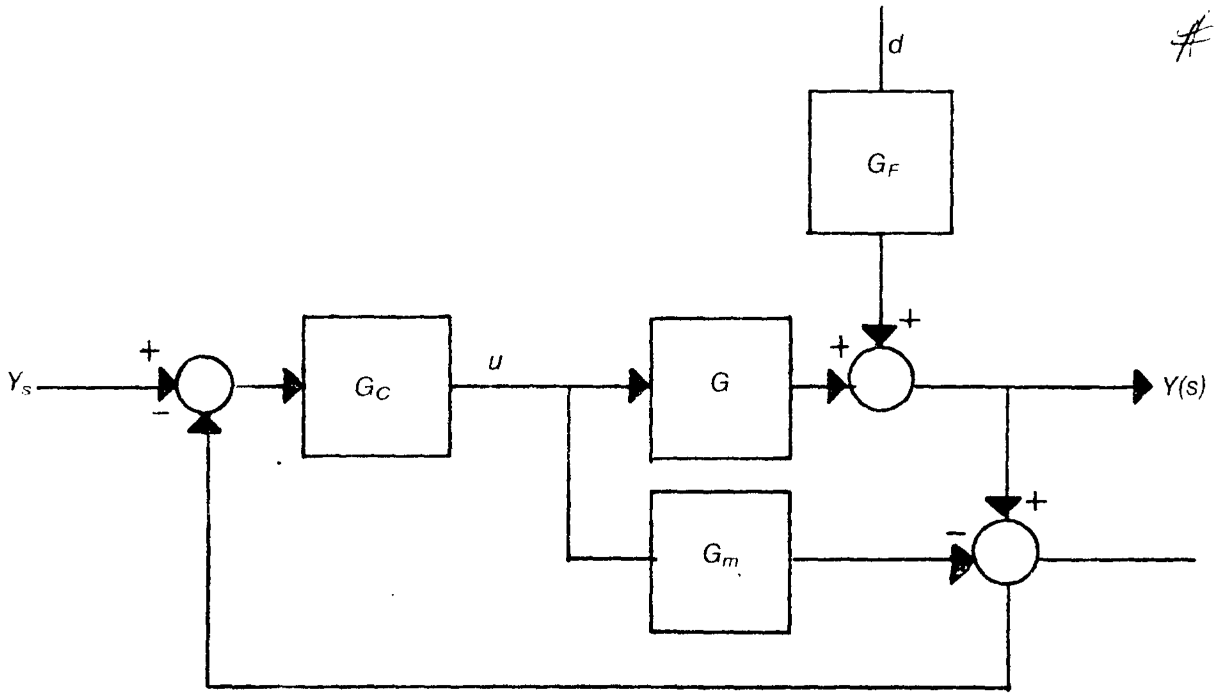

.9
.7
.5
.3
.1
-.1
-.3
OUTPUT LOOP 2
10
20
30
40
50
60
70
80
90
TIME (MIN)

#12

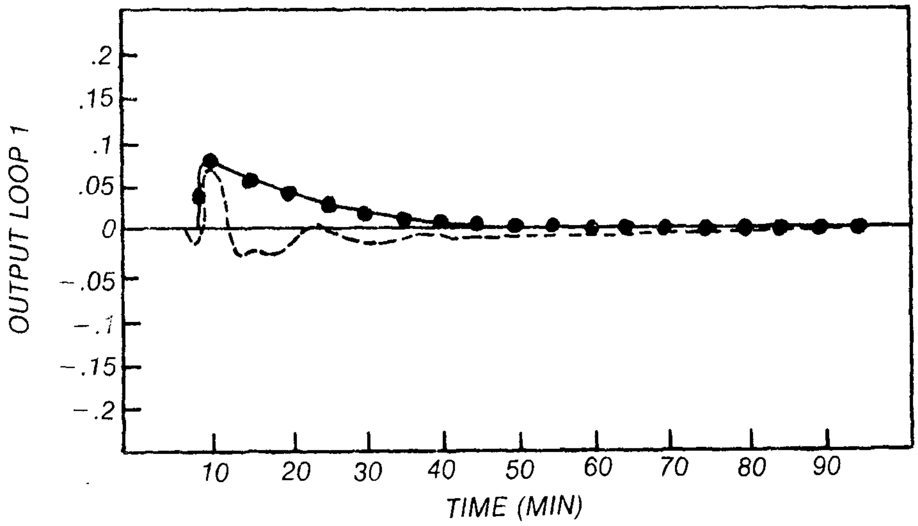

.2
.15
.1
.05
0
−.05
−.1
−.15
−.2
OUTPUT LOOP 1
10
20
30
40
50
60
70
80
TIME (MIN)


.9
.7
.5
.3
.1
−.1
−.3
OUTPUT LOOP 2
10
20
30
40
50
60
70
80
TIME (MIN)

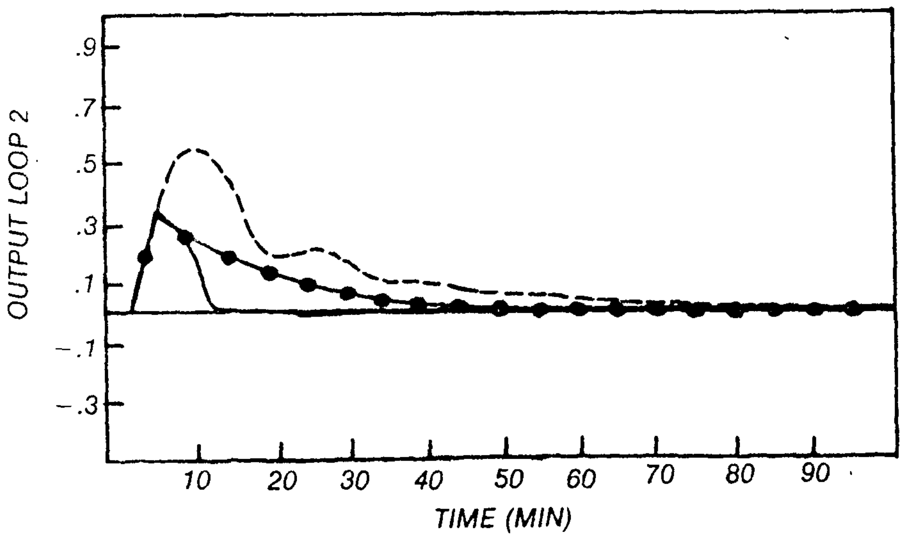
#13



#14

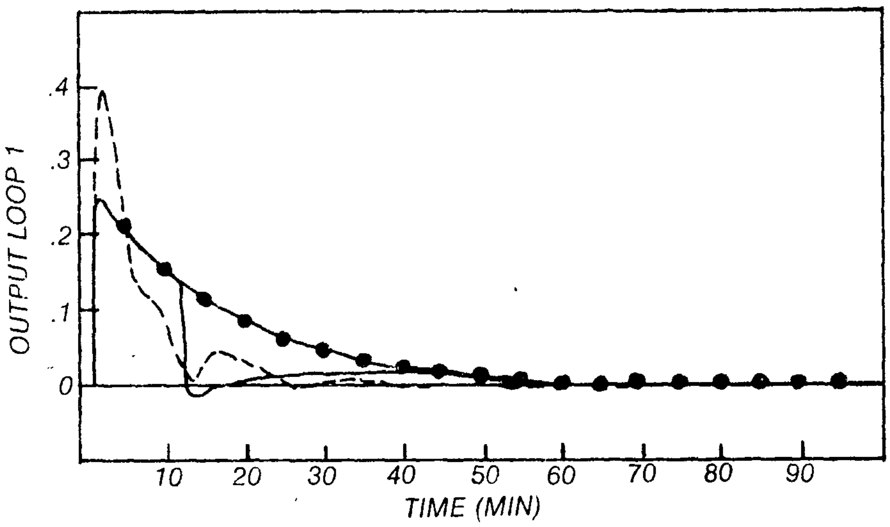


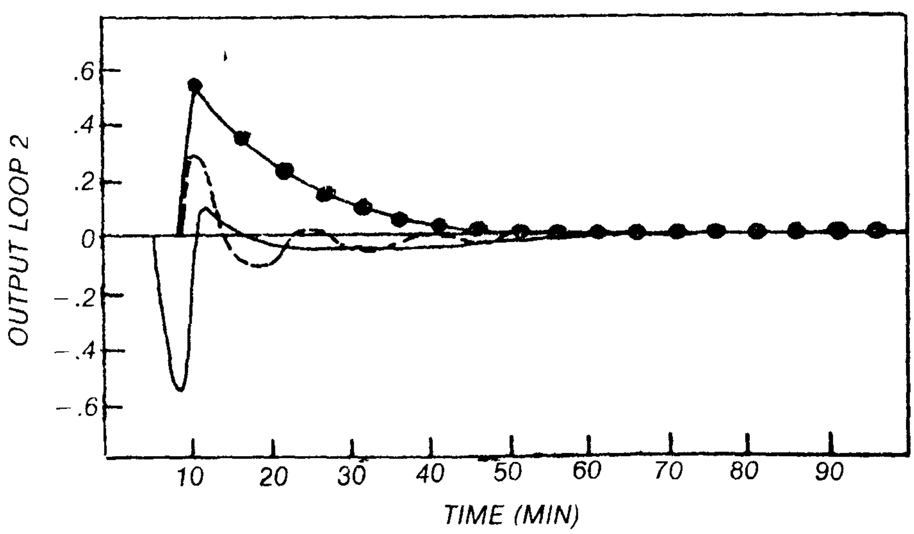
OUTPUT LOOP 1
.2
.15
.1
.05
0
−.05
−.1
−.15
−.2
10
20
30
40
50
60
70
80
90
TIME (MIN)



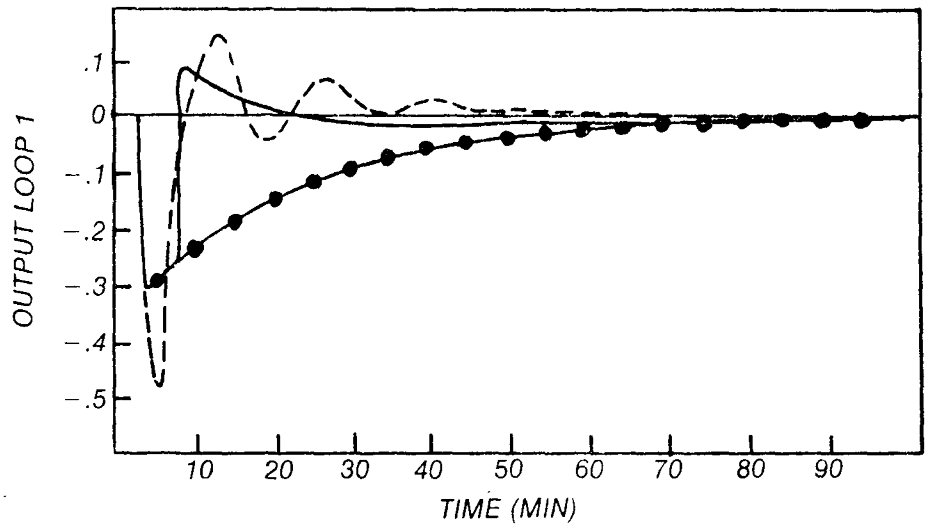
OUTPUT LOOP 2
.9
.7
.5
.3
.1
−.1
−.3
10
20
30
40
50
60
70
80
90
TIME (MIN)

#15

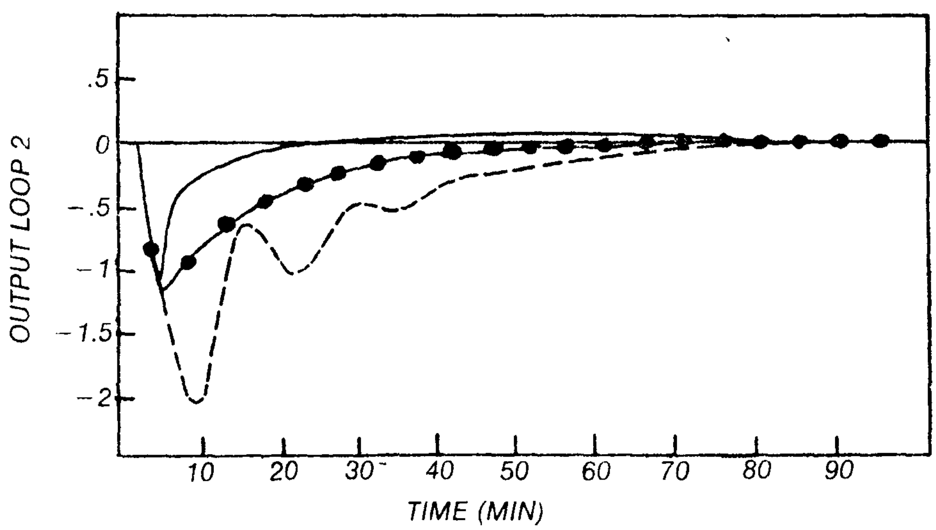

.4
.3
.2
.1
0
OUTPUT LOOP 1
10
20
30
40
50
60
70
80
90
TIME (MIN)


.6
.4
.2
0
−.2
−.4
−.6
OUTPUT LOOP 2
10
20
30
40
50
60
70
80
90
TIME (MIN)

#16



OUTPUT LOOP 1
.1
0
−.1
−.2
−.3
−.4
−.5
10
20
30
40
50
60
70
80
90
TIME (MIN)



OUTPUT LOOP 2
.5
0
−.5
−1
−1.5
−2
10
20
30
40
50
60
70
80
90
TIME (MIN)