A New Approach to the Stability and Control of Nonlinear Processes

J. F. LEATHRUM, E. F. JOHNSON, and L. LAPIDUS
Princeton University, Princeton, New Jersey

In the analysis and design of control systems one must recognize the virtual inevitability of known or unknown disturbances on the system. It is the function of the controller to maintain the system in the desired state in the face of such disturbances. The present work is a study of control-system behavior for large disturbances where the disturbances are always step-changes in one or more of the state variables. The need for considering such large disturbances arises from two fundamental facts. (1) Even if the disturbances appear to be small, there is no suitable criterion of smallness available. (2) Quite large disturbances do occur in real control systems. Therefore, if a controller cannot maintain the desired state of a system when subjected to large disturbances, it may often be considered a failure.

The need for considering nonlinear phenomena in control systems is obvious from their universal occurrence (9). Linear approximations fail to predict many types of system behavior which could only be classified as failures of the control system. Only by considering the nonlinearities can one hope to be assured that a particular process-controller combination will not fail.

In the present work the procedure has been to use all the information available about a process to derive a mathematical model for its dynamic behavior. In general the mathematical model will be nonlinear, and the type of system of interest is one which is characterized by a set of ordinary differential equations of the form

$$\dot{x}_i = f_i(x_1, x_2, \ldots x_n) \quad i = 1, 2, 3, 4, \ldots n$$

The $f_i(x_1, \ldots x_n)$ functions are composed of all the contributions to the rates of change of the state variables, x_i. The process and the controller both contribute to the rates of change of the state variables. Since time does not occur explicitly in the $f_i(x_1, \ldots x_n)$, and the x_i are independent of space position, the system is autonomous and of lumped parameter form.

The existence and uniqueness of solutions to the system equations have been discussed extensively by Nemytskii and Stepanov (*15*). It suffices here to note only that existence and uniqueness are assured if the $f_i(x)$ are continuous and have continuous first partial derivatives. The existence and uniqueness of the solution do not, however, assure that the solution is available. Indeed, in all but a few special cases the analytical solution is not available.

A school of mathematicians beginning with Poincaré (*16*) developed the qualitative theory of differential equations to treat the type of problem outlined above. The fundamental idea is that one admits that a solution in closed form is not available. As much information as possible about the solution is then sought from the nature of the functions $f_i(x_1, \ldots x_n)$. This approach to the problem will prevail throughout this paper.

The two fundamental problems of control are to assure stability and to assure an acceptable level of control quality (*12*). The stability problem requires an absolute answer. Either the system returns to the desired state when it is disturbed or it does not. Once stability has been assured, then one must assure that the desired state is reached in an acceptable amount of time. This is the problem of control stability. For nonlinear systems a statement of stability or control quality must always carry with it a statement of the size of the disturbance. The behavior of the system may be entirely different for very large disturbances than for small ones.

THEORY

Lyapunov's Concepts

The application of the qualitative theory of differential equations to stability problems was first attempted by Poincaré in the 1880's (*16*). His classification of equilibrium points is still followed as one of the first steps in the analysis of nonlinear systems. In the decade following Poincaré's work, Lyapunov developed a new branch of applied mechanics using qualitative theory (*13*). It was Lyapunov's goal to obtain stability criteria for real physical systems. His results are embodied in two distinct methods of stability analysis. The first method involves the use of the analytical solution to the problem. His second method (or direct method) is a qualitative approach to stability analysis of those systems where an analytical solution is not available. The details of the direct method will be discussed in the next section. It is the fundamental idea of Lyapunov's approach which has had the most profound effect upon control theory in Russia.

The most important contribution of Lyapunov's work is that it unifies the analysis of linear and nonlinear problems. The same methods are used for each, in recognition of the fact that in most cases an analytical solution is not available. Geometric considerations led him to stability criteria which are applicable to both linear and nonlinear systems. This concept of a geometric approach to stability problems did not ease the nonlinear problem appreciably. As one might expect, these concepts produce the most useful results for linear systems only.

Developments in the West in the analysis of nonlinear systems occurred much later and were much less fruitful than Lyapunov's methods (*14*). In these works the nonlinear problem is usually considered to be almost linear or a problem of searching for approximate solutions of some form. The unifying concept of qualitative theory did not reach Western control literature until the important papers of Bertram and Kalman were published (*3*). In the short time since the publication of those papers, a phenomenal interest has developed in the West in Lyapunov's direct method.

The goal of stability analysis is to determine whether the size of the disturbance on a system will affect the stability. Consider for example a second-order autonomous system

$$\dot{x} = P(x, y)$$

$$\dot{y} = Q(x, y)$$

where P and Q satisfy the conditions necessary for existence and uniqueness of solutions. Multiple equilibrium states are distinct possibilities in such a system. If an undesired equilibrium results from a particular disturbance, then the system is certainly unstable. The location of the equilibrium states may be determined algebraically from

$$P(x_e, y_e) = Q(x_e, y_e) = 0$$

The stability of a system is also affected by the existence of limit cycles. Limit cycles are continuous oscillations which represent distinct changes in the stability of the system. In a second-order system, a limit cycle is simply a closed solution in a plane. For higher order systems the concept of a limit cycle may be generalized by referring to it as an invariant limiting set (*11*).

The existence and location of limit cycles are of great concern in predicting the success of a control system in handling large disturbances. Poincaré offers the greatest insight into the existence of limit cycles wherein the results are in the form of necessary conditions which must be fulfilled for a limit cycle to exist. Other forms of necessary conditions for the existence of limit cycles are embodied in Lyapunov's direct method and in Bendixon's criterion (*1*). Once these necessary conditions are fulfilled, one still has no assurance that limit cycles do in fact occur.

The important problem of stability analysis remains one of determining conclusively the existence of invariant limiting sets. The first goal of improving control quality is to eliminate the undesirable invariant limiting sets from the behavior pattern of the system.

Lyapunov's Direct Method

Consider an autonomous system represented by Equation (1)

$$\dot{x}_i = f_i(x_1, x_2, \ldots x_n) \quad i = 1, 2, \ldots n \qquad (1)$$

where now the origin is one of the equilibrium states

$$f_i(0) = 0 \quad i = 1, 2, \ldots n \qquad (2)$$

The search for stability properties of the solution $x_i(x_{10}, x_{20} \ldots x_{n0}, t_0, t)$ is reduced to the search for a scalar function, $V(x_1, x_2, \ldots x_n)$, with the following properties:

1. Outside the origin: $V(x_1, x_2, \ldots, x_n) > 0$
2. $V(0) = 0$
3. $V(x_1, x_2, \ldots, x_n)$ is continuous and has continuous first partial derivatives in an open region Ω about the origin (*11*).
4. $\dot{V} = f_1 \dfrac{\partial V}{\partial x_1} + f_2 \dfrac{\partial V}{\partial x_2} + \ldots f_n \dfrac{\partial V}{\partial x_n} \leqq 0$ in Ω

A function with the properties given above is a Lyapunov function. If such a function exists, the system is "stable in the sense of Lyapunov." If condition 4 is modified as follows

4*a*. $\dot{V} < 0$ in Ω outside the origin

4*b*. $V(0) = 0$

then the stability is asymptotic. These statements are a direct result of Lyapunov's theorems which are proven from geometric considerations in n-dimensional Euclidean space.

The existence of a Lyapunov function is a sufficient condition for stability. Lyapunov's direct method involves a search for such a function. The fact that Lyapunov functions have not been found for many systems does not necessarily indicate that those systems are unstable. In fact the search for a Lyapunov function which is both a necessary and sufficient condition for stability has been likened to the search for an analytical solution (*12*).

Often the region, Ω, may not include the entire region of interest. It is then very useful to determine the extent of asymptotic stability by Lyapunov's direct method. If a bounded region, Ω_L, is defined by

$$V(x_1, x_2, \ldots, x_n) < L$$

and within Ω_L

$$V > 0$$

$$V(0) = 0$$

$$\dot{V} < 0 \text{ for } x \neq 0$$

$$\dot{V}(0) = 0$$

then the origin is asymptotically stable, and any solution starting in Ω_L tends to the origin as $t \to \infty$ (*11*). This concept is especially useful for determining those regions prohibited to invariant limiting sets other than the origin.

It is the nature of Lyapunov's direct method that it is easy to apply to linear systems where it is of little value. The application of the method to nonlinear systems requires a great deal of ingenuity which is acquired only by experience. Unfortunately, the theorems of Lyapunov offer no clues on how to find a Lyapunov function.

An important starting point in deriving a Lyapunov function is to recognize that a Lyapunov function of quadratic form may always be found for linear systems (3). An important contribution to applications of Lyapunov's direct method was made by Letov in studies of controlling linear systems with nonlinear controllers (*12*). He was able to find a Lyapunov function which was a sum of a quadratic form and an integral of the controller action:

$$V = \Phi + \int_0^\sigma f(\sigma) d\sigma$$

where

$$\Phi = \text{quadratic form}$$

$$\sigma = \text{control variable} \qquad (3)$$

$$f(\sigma) = \text{controller action}$$

$$\sigma f(\sigma) > 0 \quad \text{for} \quad \sigma \neq 0.$$

Letov recognized that solutions to the stability problem obtained from this Lyapunov function were only sufficient conditions for stability. Because of this limitation, one might expect that at times the solutions would be physically unrealizable. This does not detract, however, from the contributions made by Letov particularly in cases where one of the characteristic roots of the linear system may be zero. The application of Lyapunov's direct method to such a problem was also demonstrated by Higgins in a stability analysis of a nuclear reactor (7). Letov also extended his technique of stability analysis to the control quality problem. By a change of variables he was able to reduce the control quality problem to a conventional stability problem.

Other interesting discussions and applications of Lyapunov's direct method have appeared but are too numerous to discuss in detail here. The results of Krasovskii (10) and Ingwerson (8) deserve particular mention here, however. Krasovskii derived the conditions for which a quadratic form in the derivatives, $\dot{x}_i$, may be a Lyapunov function. Unfortunately, Krasovskii's results have more rigor than practicability. One interesting application of Krasovskii's result to control systems has been discussed by Chang (5). Ingwerson discussed extensions of the quadratic form of Lyapunov functions to nonlinear systems. The method is especially powerful in the stability analysis of systems which have nonlinearities in only one of the set of functions, $f_i(x)$.

A GEOMETRIC STABILITY CRITERION FOR SECOND-ORDER AUTONOMOUS SYSTEMS

Consider once again a second-order autonomous system

$$\dot{x} = P(x, y)$$

$$\dot{y} = Q(x, y)$$

where $P(0, 0) = Q(0, 0)$. The functions P and Q are in general continuous and nonlinear. Without loss of generality they may be separated into their linear and nonlinear parts

$$\dot{x} = ax + by + P_2(x, y)$$

$$\dot{y} = cx + dy + Q_2(x, y) \tag{4}$$

where P_2 and Q_2 are of degree two or higher. This separation is performed merely to define the linear parameters, a, b, c, and d.

Nemytskii and Stepanov suggested that sufficient information for stability analysis of such a system may be contained in the two loci (15)

$$P(x, y) = 0 \tag{5}$$

$$Q(x, y) = 0 \tag{6}$$

They did not, however, suggest any formal procedure for obtaining the desired results. It is the basic proposition of this paper that the information contained in those loci plus two additional ones is sufficient for stability analysis. The two additional loci are the loci of maximum and minimum radius:

$$\dot{R} = 0 \tag{7}$$

where

$$R = \sqrt{x^2 + m^2y^2} \tag{8}$$

$$m^2 = -b/c \tag{9}$$

It will be assumed that the origin is an isolated equilibrium point and the linear parameters are restricted thus:

$$ad - bc > 0 \tag{10}$$

$$-b/c > 0 \tag{11}$$

Beginning at an arbitrary point in the (x, y) plane it is now possible to construct an approximate integral according to the following policy.

1. When the radius is increasing, the radius is held to its minimum (that is, for $\dot{R} > 0$; $R = R_{\min}$).
2. When the radius is decreasing, the radius is held to its maximum (that is, for $\dot{R} < 0$; $R = R_{\max}$).

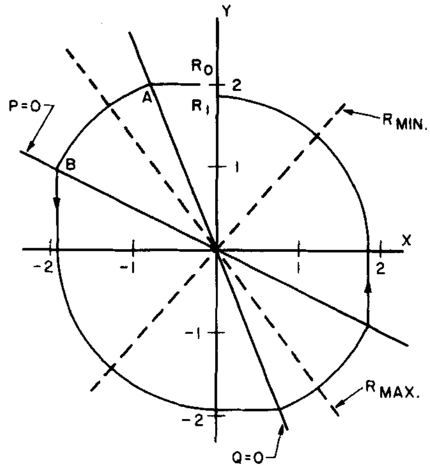

Fig. 1. Linear system.

When this alternating extreme radius path (AERP) is constructed, the direction of rotation about the origin is determined by the signs of P and Q. Using the information from the four loci of Equations (5), (6), and (7), one may construct the AERP. It is composed of straight lines and arcs of circles with centers at the origin.

Is there any relationship between the stability of the AERP and the stability of the actual integral? Since there is no mathematical proof of such a relationship, it will be demonstrated in the following examples that a relationship does seem to exist.

PROBLEMS FROM CLASSICAL MECHANICS

Linear Systems

The construction of an alternating extreme radius path (AERP) will first be illustrated for the linear system

$$\dot{x} = ax + by$$
$$\dot{y} = cx + dy \tag{12}$$

The four pertinent loci are all straight lines in the (x, y) plane (Figure 1).

$$P(x, y) = ax + by = 0 \tag{13}$$

$$Q(x, y) = cx + dy = 0 \tag{14}$$

$$\dot{R} = 0; \; y = \pm \left(\sqrt{\frac{ac}{-}}\right)x \tag{15}$$

Starting on the y-axis at R_0, the correct path according to the alternating extreme radius policy is a horizontal straight line. Any path with negative slope would cause the radius not to be a minimum. Any path with positive slope would violate the positive sign on Q. At point A the correct path is the arc of the circle with its center at the origin. Any path exterior to this one would not be a minimum radius path. An interior path to this arc would violate the positive sign of $\dot{R}$. After the locus of $R_{\max}$ is crossed, the correct path is the arc of the same circle. Now with $\dot{R} < 0$, any interior path would not be a maximum radius path. Any exterior path would violate the negative sign of $\dot{R}$. At point B, a vertical straight line is the path of maximum radius. In like manner the complete cycle of the AERP may be constructed. It should be noted here that once the general locations of the $\dot{R} = 0$ loci are ascertained, their exact location is not important. For this reason these loci of $\dot{R} = 0$ will not be included in many of the examples which follow. The importance of these loci in the reasoning above is obvious, however.

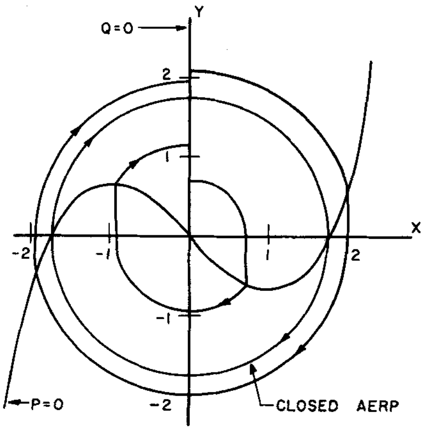

Figure 2. van der Pol's equation
$\epsilon = 1$.

For linear systems the radius after one cycle of the AERP may be expressed analytically in terms of the initial radius:

$$R_1 = R_0 \left(\frac{a^2}{bc} - 1\right)^{\frac{a}{|a|}} \left(\frac{d^2}{bc} - 1\right)^{\frac{d}{|d|}} \tag{16}$$

The difference in the squares of the radii is then

$$R_1^2 - R_0^2 = R_0^2 \left[\left(\frac{a^2}{bc} - 1\right)^{\frac{2a}{|a|}} \left(\frac{d^2}{bc} - 1\right)^{\frac{2d}{|d|}} - 1 \right] \tag{17}$$

This difference is negative for

$$a + d < 0 \tag{18}$$

which is one of the Routh-Hurwitz stability criteria. The other Routh-Hurwitz criterion is automatically satisfied by the previous restriction of Equation (10). The quadratic form of Equation (17) suggests a Lyapunov function with a first derivative of the form

$$\dot{V} = M(x^2 + y^2) \tag{19}$$

where

$$M = \left[\left(\frac{a^2}{bc} - 1\right)^{\frac{2a}{|a|}} \left(\frac{d^2}{bc} - 1\right)^{\frac{2d}{|d|}} - 1 \right] \tag{20}$$

The corresponding Lyapunov function would be

$$V = \underline{x}' \underline{A} \underline{x} \tag{21}$$

where

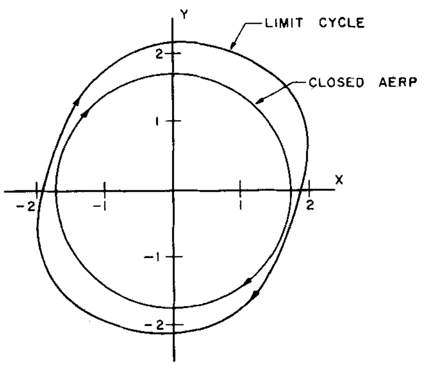

Fig. 3. The van der Pol equation
$\epsilon = 1.0$.

$$\underline{x} = \begin{bmatrix} x \\ y \end{bmatrix}; \; x' = [x\, y] = \text{transpose of } \underline{x}$$

$$\underline{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \text{ (symmetric, positive definite)}$$

The elements of $\underline{A}$ are

$$a_{11} = -\frac{M}{2a}\left[1 + \frac{c(ac + bd)}{(ad - bc)(a + d)}\right] \tag{22}$$

$$a_{12} = \frac{M(ac + bd)}{2(ad - bc)(a + d)} \tag{23}$$

$$a_{22} = -\frac{M}{2d}\left[1 + \frac{b(ac + bd)}{(ad - bc)(a + d)}\right] \tag{24}$$

One may easily assure the positive definite nature of $\underline{A}$ by observing that

$$a_{11} > 0 \tag{25}$$

$$a_{11}\, a_{22} - a_{12}^2 > 0 \tag{26}$$

The van der Pol Equation

Among the simplest of the nonlinear systems which one might consider is the van der Pol equation

$$\ddot{x} + \epsilon(x^2 - 1)\,\dot{x} + x = 0 \tag{27}$$

which may be transformed to

$$\dot{x} = y - \epsilon\left(\frac{x^3}{3} - x\right) \tag{28}$$

$$\dot{y} = -x$$

This is a particularly interesting system because the integral exhibits limit-cycle behavior (*1*). The AERP cycles have been constructed in Figure 2, and a closed alternating extreme path also exists (Figure 3). This is the first indication from a nonlinear system that the stability of the AERP is qualitatively the same as the stability of the true integral. From Equation (28), it is obvious that this closed AERP will exist for all ϵ. It is also known that limit cycle exists for all ϵ (*11*). This suggests a definite correlation between the existence of a closed AERP and the existence of a limit cycle. If this correlation proves valid, the first fundamental problem of stability analysis will be solved for second-order systems. A sufficient condition for the existence of limit cycles will have been found.

STABILITY AND CONTROL OF A CONTINUOUS STIRRED-TANK CHEMICAL REACTOR

The analysis of chemical process control systems has usually been accomplished through the use of convenient linear approximations. In many cases this approach has been very satisfactory. Fortunately, the nonlinear phenomena may even have beneficial effects in many chemical and staged operations. One may also resort to large holdups in chemical systems to minimize the effects of disturbances. It is not necessary therefore to consider the effects of nonlinear phenomena on the stability and control of such systems. One important exception to this statement was first pointed out by Frank-Kamenetskii (*6*). In the study of heterogeneous combustion reactions he found multiple equilibrium states. Each equilibrium state represented a balance of the heat-generation rate with the heat removal-rate. Bilous and Amundson (*4*) and others have also discussed similar phenomena in continuous stirred reactors with exothermic reactions.

The local stability of each equilibrium state could be determined by the relative magnitudes of the heat-generation rate and heat-removal rate near equilibrium. Aris and Amundson considered the global stability problem and the control quality problem for this continuous stirred-tank reactor (2). Their results were obtained by numerically integrating the dynamic equations in the form of the second-order autonomous system $\dot{x} = P(x, y)$, $\dot{y} = Q(x, y)$. They found that limit cycles do exist in such a system for certain values of the control parameters. In what follows here the Aris and Amundson problem will be analyzed with the use of the alternating extreme radius path.

First-Order Exothermic Reaction

Consider an irreversible, exothermic reaction in the vessel shown in Figure 4. The reaction is a simple transformation of A to B

$$A \rightarrow B \tag{29}$$

where the rate of reaction is given by

$$r = k_0 C_A^n \tag{30}$$

This rate of reaction must be coupled with the other transport rates to complete the formulation of the dynamic equations

$$V \frac{dC_A}{dt^*} = F(C_{A0} - C_A) - kVe^{-E/R'T^*} C_A^n$$

$$\rho V C_p \frac{dT^*}{dt^*} = \rho F C_p (T_o - T^*) + \Delta H k V e^{-E/R'T^*} C_A^n - (hA + g_c)(T^* - T_c) \tag{31}$$

Proportional Control

For proportional control on temperature, g_c has the form

$$g_c = K^*(T - T_s) \tag{32}$$

With this form for the control function, the problem of offset will be eliminated if it is assumed that T_s is also an equilibrium temperature in the absence of control. The following definitions are now made in order to make the origin an equilibrium state:

$$T = \frac{T^* - T_s}{T_s}; \quad c = \frac{C_A - C_{As}}{C_{As}}; \quad t = t^*/\tau$$

Equation (31) will now have the form

$$\frac{dc}{dt} = -c - \beta\left[(1 + c)^n e^{\theta \frac{T}{T+1}} - 1\right] = P(c, T)$$

$$\frac{dT}{dt} = -\gamma T + \epsilon\left[(1 + C)^n e^{\theta \frac{T}{T+1}} - 1\right] - KT[T + (1 - \alpha_c)] = Q(c, T) \tag{33}$$

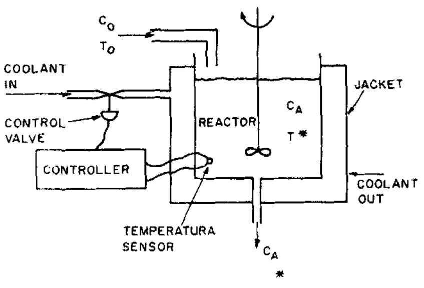

Fig. 4. Continuous stirred-tank reactor with controlled jacket cooling.

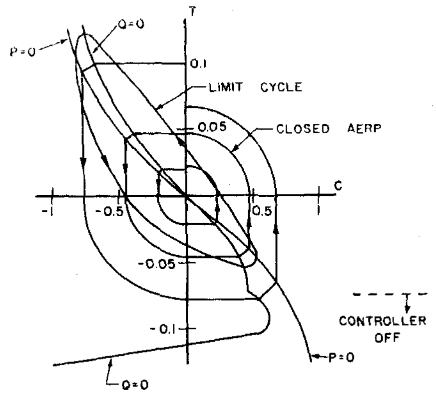

Fig. 5. First-order reaction with proportional control
$K = 16$
$(hA)_{max} = \infty$

where

$$P(0, 0) = Q(0, 0) = 0 \qquad \epsilon = \frac{\Delta H \tau C_{As}^n \kappa e^{-\theta}}{\rho C_p T_s}$$

$$\alpha_c = T_c/T_s$$

$$\beta = \tau k e^{-\theta} C_{As}^{n-1} \qquad \theta = E/R'T_s$$

$$\gamma = 1 + \frac{\tau k A}{\rho V C_p} \qquad K = \frac{\tau K^* T_s}{\rho V C_p}$$

The formulation of the problem is not complete without considering the constraints on the controller. The first and obvious constraint is that a negative flow rate to the jacket is not allowed:

$$hA + g_c \geq 0 \tag{34}$$

Secondly, there must exist some maximum flow rate for any real cooling system:

$$(hA + g_c) \leq (hA)^*_{max} \tag{35}$$

In dimensionless form this constraint becomes

$$(hA)_{max} = \max \text{ of } \left(\gamma - 1 + \frac{\tau g_c}{\rho V C_p}\right) \tag{36}$$

From the linear parts of Equation (33) one may determine the scale factor for constructing the alternating extreme radius paths (AERP):

$$m^2 = \frac{\beta\theta}{n\epsilon} \tag{37}$$

Examples of AERP constructions are shown in Figures 5 and 6. Where limit cycles exist they are shown superimposed upon the AERP. The results of these constructions

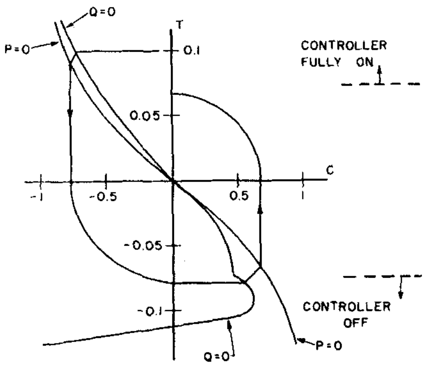

Fig. 6. First-order reaction with proportional control
$K = 16$
$(hA)_{max} = 2$

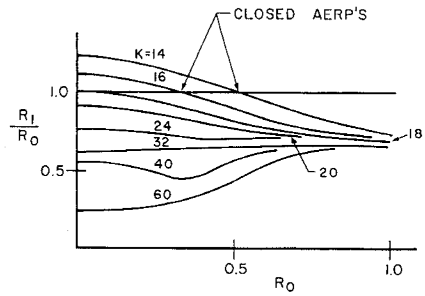

Fig. 7. Alternating extreme radius paths for first-order exothermic chemical reaction in a continuous stirred tank $(hA)_{max} = \infty$

are then summarized in Figures 7 and 8. In every case the existence of a closed AERP was sufficient to predict the existence of a limit cycle. These results confirm those of Aris and Amundson which were obtained by numerical integration. The numerical integrations were also repeated in order to determine the location of the limit cycles very accurately and to compare the times required for the AERP calculations and the numerical integration. With an IBM-650 computer, the pertinent loci for the AERP construction could be found in approximately 2 min. This information is sufficient to determine the stability of any disturbance in the (c, t) plane. An average time of 30 min. for each particular initial condition was required for the numerical integration using the Edelman method. This numerical integration will produce stability information only for that set of initial conditions along one solution path.

In each case the stability of the actual integral found by numerical integration was compared with the stability of the alternating extreme radius path. This AERP never failed to be a valid stability criterion. The existence of a closed AERP was always sufficient to predict a closed integral (that is, a limit cycle). It would appear that the first problem of control for this chemical reactor system has been solved. A stability criterion for large disturbances has been found.

The second problem of control, the problem of control quality, may now be considered. One might expect maximum quality from that set of control parameters which minimize the area under the curves of Figure 7. For this problem of a first-order reaction with proportional control, the optimum value of K would appear to be its largest possible value. In any real system there will be time delays in the control system, however, which would make very large proportional gains undesirable. In addition, any other lags or nonlinearities in the controller would make very large gains unattainable. The principal question here is one of desirability. The presence of time delays places an upper bound on the values of proportional gain which assure local stability. One might resort then to conventional practice in control system synthesis where one chooses the proportional gain at about twice the minimum allowable value for stability. In this case, $K = 40$ might be a reasonable choice. Now the AERP can be used to assure that this choice gives a stable system for all disturbances regardless of size.

Proportional Plus Derivative Control

Now consider the continuous stirred reactor of Figure 4 with both proportional and derivative control. The control function is now

$$g_c = K^* \left[(T^* - T_s) + \tau_d^* \frac{dT^*}{dt^*} \right] \qquad (38)$$

The dynamic equations in the form of Equation (33) now become

$$\frac{dc}{dt} = -c - \beta \left[(1 + c)^n e^{\theta \frac{T}{T+1}} - 1 \right] = P(c, T) \qquad (39)$$

$$\frac{dT}{dt} = -\gamma T + \epsilon \left[(1 + c)^n e^{\theta \frac{T}{T+1}} - 1 \right] - K \left[T + \tau_d \frac{dT}{dt} \right] [T + (1 - \alpha_c)] \qquad (40)$$

In the latter of these the derivative may be solved for explicitly:

$$\frac{dT}{dt} = \frac{1}{1 + K\tau_d [T + (1 - \alpha_c)]} \left\{ -\gamma T + \epsilon \left[(1 + c)^n e^{\theta \frac{T}{T+1}} - 1 \right] - KT [T + (1 - \alpha_c) \right\} = Q(c, T) \qquad (41)$$

It can be seen from Equations (39) and (41) that the two loci of $P = 0$ and $Q = 0$ are the same as with proportional control alone. The effect of derivative action is to change the scale factor, m, where now

$$m^2 = \frac{\beta\theta}{n\epsilon} [1 + K\tau_d (1 - \alpha_c)] \qquad (42)$$

The constraints on the controller are also distorted by the derivative action

$$0 \leq \gamma - 1 + KT + \frac{K\tau_d}{1 + K\tau_d [T + (1 - \alpha_c)]} - \gamma T + \epsilon \left[(1 + c)^n e^{\theta \frac{T}{T+1}} - 1 \right] - KT [T + (1 - \alpha_c)] \leq (hA)_{max} \qquad (43)$$

When these constraints are exceeded, the system behaves as though no controller existed, and the scale factor reverts to that given by Equation (37):

$$m^2 = \frac{\beta\theta}{n\epsilon} \qquad (37)$$

The construction of the alternating extreme radius path is thus complicated by the use of two scale factors. The pertinent loci for the construction are shown in Figures 9 and 10 for $K = 40$. The constructions have been carried out and the results are summarized in Figure 11. From these results one observes an improvement in control quality with the use of derivative control. The im-

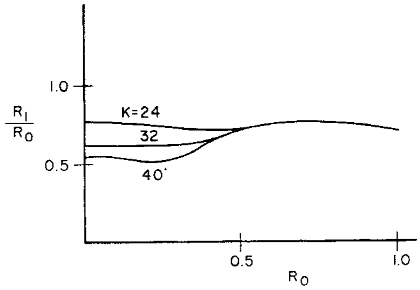

Fig. 8. Alternating extreme radius paths for a first-order exothermic chemical reaction in a continuous stirred tank $(hA)_{max} = 2$

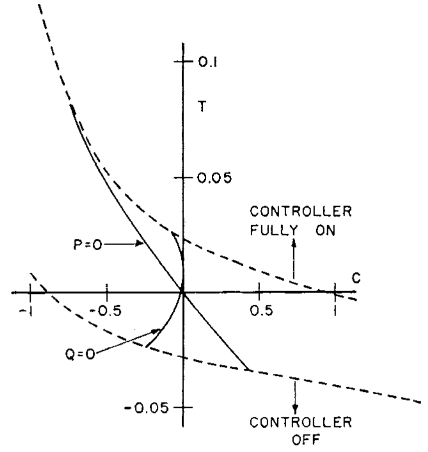

Fig. 9. First-order chemical reaction with proportional and derivative control
$K = 40$
$\tau_d = 0.25$
$(hA)_{\max} = 2$

provement is greatest for small disturbances where the controller action is within the constraints. For large disturbances the distortion of the constraints by derivative action may cause the quality of control to diminish with increasing τ_d.

A Control Quality Problem with Two Controllers

So far the geometric stability criterion which uses the alternating extreme radius path has been used to assure stability for large disturbances in control systems. It has been applied to systems designed by conventional linear methods to assure that the nonlinearities do not cause a control system failure. Refinements on these designs will now be considered with the goal of improving the control quality for large disturbances.

Consider the chemical reactor system of Figure 12 where the heat produced in the reactor is used to preheat the feed to the reactor. Heat will also be removed from the reactor by a cooling jacket. The temperature in the reactor will be controlled by controlling the jacket coolant flow rate and the bypass of the heat exchanger. This will be done with two independent proportional controllers, each with constraints on their actions. The delay in the feed line between the heat exchanger and the reactor will be ignored. The dynamic equations in the form of Equation (33) become

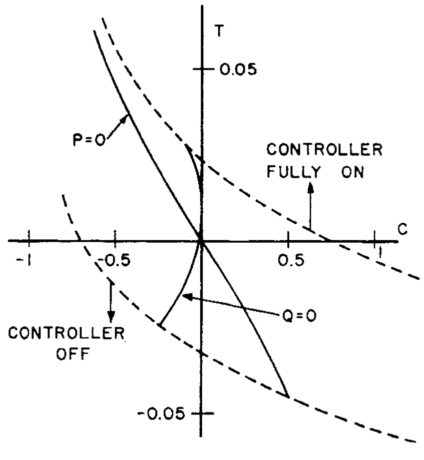

Fig. 10. First-order chemical reaction with proportional and derivative control
$K = 40$
$\tau_d = 0.60$
$(hA)_{\max} = 2$

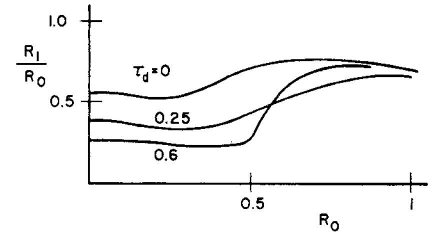

Fig. 11. Alternating extreme radius paths for a first-order exothermic chemical reaction in a continuous stirred tank with proportional and derivative control
$K = 40$
$(hA)_{\max} = 2$

$$\frac{dc}{dt} = -c - \beta\left[(1+c)^n e^{\theta\frac{T}{T+1}} - 1\right] = P(c, T)$$

$$\frac{dT}{dt} = -\frac{T}{1+f(\gamma_2-1)} - (\gamma_1-1)\,T + \epsilon\left[(1+c)^n e^{\theta\frac{T}{T+1}} - 1\right] - K_1T\,[T+(1-\alpha_c)] - \frac{K_2T(\gamma_2-1)(1-\alpha_c)}{\gamma_2[1-f(\gamma_2-1)]} = Q(c, T) \quad (44)$$

In this formulation the ΔT in the heat exchanger is assumed constant for its entire length although it may change with time. The heat capacity of the heat exchanger is also neglected.

The constraints on the two controllers become

$$0 \leq \gamma - 1 + K_1T \leq (hA)_{\max} \quad \text{(jacket)} \quad (45)$$

$$0 \leq f \leq f_{\max} \quad \text{(heat exchanger)} \quad (46)$$

Once again the loci of $P = 0$ and $Q = 0$ may be computed and the alternating extreme radius paths may be constructed. The scale factor for a first-order reaction is given by

$$m^2 = \frac{\beta\theta}{\epsilon}$$

The problem to be solved with the AERP's is to decide how to distribute the control action between the two controllers. Conventional linear methods would give a value of d, where

$$d = \epsilon\theta - (\gamma_1 - 1) - \frac{1}{\gamma_2} - K_1(1-\alpha_c) - \frac{K_2(\gamma_2-1)(1-\alpha_R)}{\gamma_2^2} \quad (47)$$

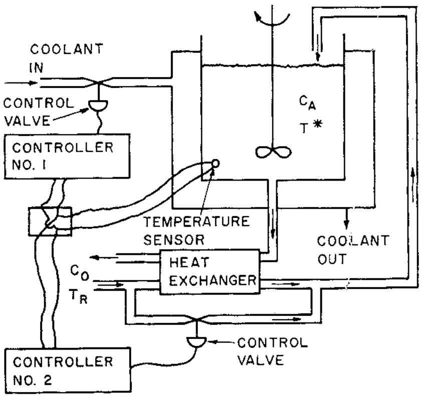

Fig. 12. Continuous stirred-tank reactor with two controllers.

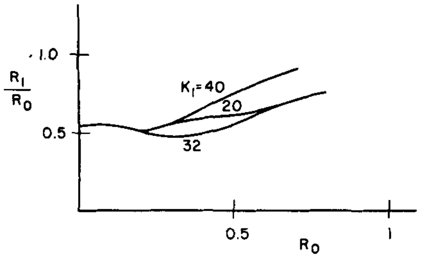

Fig. 13. Alternating extreme radius paths for a first-order exothermic chemical reaction in a continuous stirred tank with two controllers.

Once a suitable d is chosen, linear theory offers no help in deciding how to choose K_1 and K_2. In the numerical example which follows, d will be held constant, and the AERP's will be used to compare the control quality for various K_1 and K_2 combinations.

The results of the AERP constructions are summarized in Figure 13. It is found that better control quality is obtained for $K_1 = 32$ than for $K_1 = 20$ or 40. This value of $K_1 = 32$ also better isolates the origin from the other equilibrium points. The AERP method has thus given a solution to the control quality problem. This solution must now be compared with the actual solution obtained by numerical integration. The control quality will be expressed in terms of a performance index, P':

$$P' = \int_0^t (c^2 + \rho^2 T^2)\, dt \qquad (48)$$

Equation (45) was integrated numerically and the performance index, P', was evaluated at each step in the integration. The lowest value of P' would indicate the best control quality. The results of the numerical integration are given in Table 1, where ρ is chosen such that $\rho = m$. The smallest performance index was obtained for $K_1 = 32$, thus the AERP would have given the correct answer to the control quality question. The best controller arrangement for handling large disturbances appears to be one which distributes the control action between the two controllers. The use of either controller alone does not produce a good control quality. No distinction can be made in the way the controllers handle small disturbances. (See Figure 13 for small R_0.)

The alternating extreme radius paths have been used to solve the two fundamental problems of control for a first-order exothermic reaction in a continuous stirred tank. First, stability was assured and confirmed by numerical integration. Then, a control quality problem was solved and an optimum controller was designed. The validity of this optimum design was confirmed by numerical integration.

Second-Order Exothermic Reaction

Consider the same chemical reactor system as shown in Figure 4, where the reaction is now second-order. The

TABLE 1. FIRST-ORDER EXOTHERMIC REACTION WITH TWO CONTROLLERS: PERFORMANCE INDEX FROM NUMERICAL INTEGRATIONS

$\beta = 1, \gamma_1 = 2, \epsilon = .25, \theta = 25, 1 - \alpha_c = .125$
$1 - \alpha_R = .25, \gamma_2 = 2, f_{max} = 2, (hA)_{max} = 2$

Initial conditions: $c = 0$, $T = 0.1$
Total dimensionless time = 3.9325

K_1	K_2	P'
20	48	2.197
32	24	2.159
40	8	2.194

dynamic equations assume the same form as Equation (33) where now $n = 2$:

$$\frac{dc}{dt} = -c - \beta\left[(1+c)^2 e^{\theta \frac{T}{T+1}} - 1\right] = P(c, T)$$

$$\frac{dT}{dt} = -\gamma T + \epsilon\left[(1+c)^2 e^{\theta \frac{T}{T+1}} - 1\right] - KT\left[T + (1-\alpha_c)\right] = Q(c, T) \qquad (49)$$

The loci of $P = 0$ and $Q = 0$ may be computed by solving quadratics in c. The scale factor for constructing the alternating extreme radius path is given by

$$m^2 = \frac{\beta\theta}{2\epsilon}$$

A number of examples have been worked out with one result shown in Figure 14. The results are what would be expected from the previous problem. The alternating extreme radius path is once again useful in predicting the existence of limit cycles.

Reversible Reaction

Unfortunately all process control systems cannot be described by equations of the form of $\dot{x} = P(x, y)$ $\dot{y} = Q(x, y)$. Higher order differential equations are many times necessary for a complete mathematical description of the system. One such system would be obtained if the chemical reaction in the previous problem was made reversible:

$$A \rightleftharpoons B$$

If both reactions are considered first-order, the rate of reaction becomes

$$r = k_{10} C_A - k_{20} C_B \qquad (50)$$

where

$$k_{10} = k_1 e^{-E_1/R'T^*} \quad \text{(forward)}$$
$$k_{20} = k_2 e^{-E_2/R'T^*} \quad \text{(reverse)}$$

This rate may now be coupled with the other transport rates to give the dynamic equations for a continuous stirred tank:

$$V \frac{dc_A}{dt^*} = F(c_{A0} - c_A) - k_1 V e^{-E_1/R'T^*} c_A + k_2 V e^{-E_2/R'T^*} c_B \qquad (51)$$

$$V \frac{dc_B}{dt^*} = F(c_{B0} - c_B) - k_2 V e^{-E_2/R'T^*} c_B + k_1 V e^{-E_1/R'T^*} c_A \qquad (52)$$

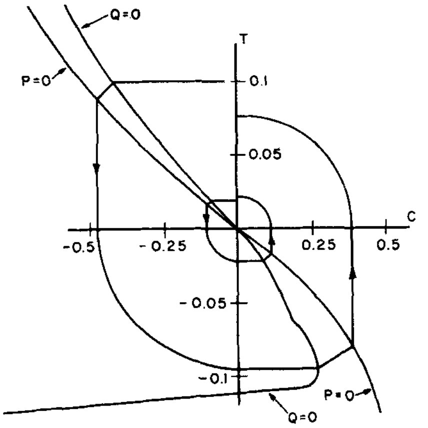

Fig. 14. Second-order exothermic reaction with proportional control.

$$\rho V C_p \frac{dT^*}{dt^*} = \rho F C_p (T_o - T^*) + \Delta H[k_1 V e^{-E_1/R'T^*} c_A - k_2 V e^{-E_2/R'T^*} c_B] - (hA + g_c)(T^* - T_c) \quad (53)$$

Equations (51) and (53) may be added and a new variable defined by

$$c^* = c_A + c_B \quad (54)$$

With proportional control, the dynamic equations in dimensionless form become

$$\frac{da}{dt} = -a - \beta_1 \left[(1+a) e^{\theta_1 \frac{T}{T+1}} - 1 \right] + \beta_2 \left[\left\{ \frac{c_s}{a_s}(c+1) - (a+1) \right\} e^{\theta_2 \frac{T}{T+1}} + \left(1 - \frac{c_s}{a_s} \right) \right] \quad (55)$$

$$\frac{dc}{dt} = -c \quad (56)$$

$$\frac{dT}{dt} = -\gamma T + \epsilon_1 \left[(1+a) e^{\theta_1 \frac{T}{T+1}} - 1 \right] + \epsilon_2 \left[\left\{ \frac{c_s}{a_s}(c+1) - (a+1) \right\} e^{\theta_2 \frac{T}{T+1}} + \left(1 - \frac{c_s}{a_s} \right) \right] - KT\,[T + (1 - \alpha_c)] \quad (57)$$

$$a = \frac{C_A - C_{As}}{C_{As}} \qquad \epsilon_1 = \frac{\Delta H_\tau C_{As} k_1}{\rho C_p T_s} e^{-\theta_1}$$

$$T = \frac{T^* - T_s}{T_s} \qquad \epsilon_2 = \frac{\Delta H_\tau C_{As} k_2}{\rho C_p T_s} e^{-\theta_1}$$

$$\beta_1 = \tau k_1 e^{-\theta_1} \qquad \theta_1 = E_1/R'T_s$$

$$\beta_2 = \tau k_2 e^{-\theta_2} \qquad \theta_2 = E_2/R'T_s$$

The resulting system of differential equations is third-order. For purposes of stability analysis, however, one observes immediately from Equation (56) that the variable, c, is always stable. Since the variable, c, approaches zero as $t \to \infty$, it need not be considered in the other two equations. The resulting system for stability analysis becomes

$$\frac{da}{dt} = -a - \beta_1 \left[(1+a) e^{\theta_1 \frac{T}{T+1}} - 1 \right] - \beta_2 \left[(1+a) e^{\theta_2 \frac{T}{T+1}} - 1 \right] + \beta_2 \frac{c_s}{a_s} \left(e^{\theta_2 \frac{T}{T+1}} - 1 \right) \quad (58)$$

$$\frac{dT}{dt} = -\gamma T + \epsilon_1 \left[(1+a) e^{\theta_1 \frac{T}{T+1}} - 1 \right] + \epsilon_2 \left[(1+a) e^{\theta_2 \frac{T}{T+1}} - 1 \right] - \epsilon_2 \frac{c_s}{a_s} \left(e^{\theta_2 \frac{T}{T+1}} - 1 \right) - KT\,[T + (1 - \alpha_c)] \quad (59)$$

It should be noted that the steady state value of c (that is, c_s) still enters the dynamic equations as a parameter.

The differential equations are now second-order, and the state space is once again two-dimensional. The geometric stability criterion using the alternating extreme radius paths may once again be applied. In this one example a change of variable reduced a third-order system to the more convenient second-order form.

HIGHER ORDER SYSTEMS

The Search for Stable Variables

The method used in a previous problem for stability analysis of a third-order system will now be applied to higher order systems. In that method it was found that one of the components of the state vector was always stable. The system could therefore be reduced to second-order where the alternating extreme radius paths could be constructed. Consider now a set of chemical reactors in series where the output of each reactor depends on the conditions in that reactor and the feed from the reactor immediately preceding it. If each reactor was designed as in Figure 4 and the reaction was first-order and irreversible, then $2N$ simultaneous ordinary differential equations would be necesary to describe the dynamic behavior of the complete system. For purposes of stability analysis, however, each reactor may be considered individually but in order from 1 to N. Equation (33) would be sufficient to describe reactor Number 1. If this first reactor is stable, then as $t \to \infty$ the feed to the second reactor becomes steady. Since stability pertains to very large times, the stability of reactor Number 2 may be analyzed with respect to this steady feed condition. Likewise the analysis may proceed step-by-step until an unstable reactor is found. Once an unstable reactor is found, all reactors thereafter will surely be unstable. The important point to be made is that at each step the system is only second-order. The methods employed previously are thus still applicable to this very high order system. In the limit for very large N, this chain of reactors approximates a tubular reactor with restricted backmixing. Thus one more application of the alternating extreme radius path is now at hand.

The control quality problem for a set of reactors in series is quite far removed from the stability problem. Here one is concerned with finite times. The only general result that can be obtained by step-by-step considerations may be found in the following: If the entire set of stages is stable, if the disturbance in question is acting on only one of the stages, and if each of the stages acting separately and individually gives an acceptable level of control quality for this disturbance and all smaller disturbances, then the control quality of each of the stages in the series-connected set is acceptable.

Extensions of the Geometric Stability Criterion

Many higher order systems do not permit the type of analysis described in the previous section. It is not always possible to find variables which are independently stable. The next possible approach would be to extend the geometric stability criterion described previously to higher order systems. While this has not been achieved as yet, some of the important considerations will be discussed here.

As a result of the uniqueness properties of the solutions of Equation (1), one knows that the solution path is always one-dimensional regardless of the dimensionality of the state space. One would also expect that the alternating extreme radius path would remain one-dimensional for higher order systems. The geometric construction of the AERP requires a very clear definition of the radius in terms of the linear coefficients [Equation (8)]. If the radius were defined by the usual sum of squares,

$$R = \sqrt{\sum_{1}^{N} p_i^2 x_i^2} \quad (60)$$

where $p_1 = 1$, then $N-1$ scale factors, p_i, would be necessary. (N = order of the system). There are $N!/2(N-2)!$ scale factors available, however, from the off diagonal elements of the matrix of linear coefficients, $\bar{A}$. This suggests that $N!/2(N-2)!$ radii should be defined by considering only two variables at a time:

$$R_{ij} = \sqrt{x_i^2 + p_{ij}^2 x_j^2} \qquad i < j \tag{61}$$

$$p_{ij}^2 = - a_{ij}/a_{ji} \tag{62}$$

If one focuses attention upon only two variables at a time, three fundamental questions must be answered.

1. Which two variables receive the attention?
2. What happens to the other variables?
3. What happens to the origin of coordinates?

It is hoped that future work will produce criteria for answering these questions based upon the signs of the x_i, f_i and $\dot{R}_{ij}$.

CONCLUSIONS

The qualitative theory of differential equations as developed by Poincaré and Lyapunov has been employed to solve some very practical control problems. The results have come from the use of a geometric stability criterion which involves a very simple construction in two dimensions. This stability criterion gave necessary and sufficient conditions for the existence of limit cycles in second-order systems. It was also employed successfully in a control quality problem at a great time saving compared to conventional numerical methods.

In each problem considered, the geometric stability criterion was compared with the results of numerical integration of the same problem. In each case, the geometric stability criterion was valid. Among the problems considered were problems from applied mechanics and chemical process control problems. Extensions of the method to higher order systems were then pointed out.

A great many higher order systems remain, however, to be analyzed. The path of extending the geometric stability criterion to such systems remains very unclear. Two approaches to this problem suggest themselves for the future. First, one should seek a rigorous mathematical proof of the validity of the geometric stability criterion for second-order systems. Such a proof might indicate conclusively the possibilities of using the method on higher order systems. Secondly, the idea of a geometric stability criterion might be generalized directly without the aid of a mathematical proof. Based upon the experience gained here with the second approach, it seems that the first is to be favored.

The successful application of any qualitative method of designing control systems is only possible if the system itself is well defined by a set of dynamic equations. Now that nonlinear equations can be treated, there is a great need for more data on the nonlinear characteristics of process systems.

ACKNOWLEDGMENT

The authors wish to express their gratitude to the National Science Foundation, the International Business Machines Corporation, and the American Oil Company whose financial assistance made this work possible.

NOTATION

A_o = heat transfer area
C_A = concentration of component A
C_{Ao} = input concentration of A
C_{As} = steady state concentration of component A
C_p = heat capacity of the reaction mixture
E = activation energy
f = fraction of steady state flow of feed through the heat exchanger which is actually flowing at a time (that is, $f = 1 - K_2T$)
F = volumetric flow rate to the reactor
g_c = control function for the jacket
h = heat transfer coefficient at steady state
ΔH = heat of reaction
K_1 = proportional gain of the jacket controller
K_2 = proportional gain of the heat exchanger controller
k_o = specific reaction rate constant = $ke^{-E/R'T^*}$
k = preexponential factor
m = scale factor
n = order of reaction
p = right-hand side of differential equation
Q = right-hand side of differential equation
R = radius
R' = gas constant
t^* = time
T^* = absolute temperature
T_o = input temperature
T_c = average coolant temperature
T_R = input temperature of cold feed to the heat exchanger
T_s = steady state temperature
V = volume of the reactor
x_i = ith dependent variable
$\bar{x}$ = vector of dependent variables
α_R = T_R/T_s
γ_1 = $1 + \left(\frac{hA_T}{\rho V C_p}\right)$ jacket
γ_2 = $1 + \left(\frac{hA_T}{\rho V C_p}\right)$ heat exchanger
ρ = density of the reaction mixture
τ = time constant
τ_d^* = rate time for derivative control
τ_d = dimensionless rate constant for derivative control

LITERATURE CITED

1. Andronov, A. A., and S. E. Chaikin, "Theory of Oscillations," Princeton University Press, Princeton, N. J. (1949).
2. Aris, R., and N. R. Amundson, *Chem. Eng. Sci.*, **7**, 121, 132, 148 (1958).
3. Bertram, J. E., and R. E. Kalman, *Trans. Am. Soc. Mech. Engrs.*, **820**, 370 (1960).
4. Bilous, O., and N. R. Amundson, *A.I.Ch.E. Journal*, **1**, 513 (1955).
5. Chang, S. S. L., *Trans. Am. Soc. Mech. Engrs.*, **830**, 91 (1961).
6. Frank-Kamenetskii, D. A., "Diffusion and Heat Exchange in Chemical Kinetics," Princeton University Press, Princeton, N. J. (1955).
7. Higgins, T. J., *Am. Inst. Elec. Engrs. Nonlinear Control Committee Trans.* (1960).
8. Ingwerson, D. R., *Inst. Radio Engrs. Trans. on Automatic Control*, **AC-6**, 199 (1961).
9. Kovach, L. D., "Life Can Be So Nonlinear," *Amer. Scientist*, **48**, 218, (1960).
10. Krasovskii, N. N., *Prikladnay Matematika i Mekkanika*, **21**, 309 (1957).
11. LaSalle, J. P., and S. Lefschetz, "Stability by Lyapunov's Direct Method," Academic Press, New York (1961).
12. Letov, A. M., "Stability in Nonlinear Control Systems," Princeton University Press, Princeton, N. J. (1961).
13. Lyapunov, A. M., *Ann. Math. Study*, **17**, Princeton University Press, Princeton, N. J. (1949).
14. Minorsky, N., "Nonlinear Oscillations," Van Nostrand, New York (1962).
15. Nemytskii, V. V., and V. V. Stepanov, "Qualitative Theory of Differential Equations," Princeton University Press, Princeton, N. J., (1960).
16. Poincaré, H., *Jr. de Math. Pures*, Ser. 3, **7**, 375 (1881).

Manuscript received February 6, 1963; revision received April 15, 1963; paper accepted April 17, 1963.