Design Concepts for Process Control

A. Kestenbaum,[1] R. Shinnar,*[1] and F. E. Thau[2]

Departments of Chemical Engineering and Electrical Engineering, The City College of The City University of New York, New York, New York 10031

Six criteria are offered for the evaluation of the performance of process controllers. Classical techniques for the design of PID controllers and modern control theoretic techniques for controller design are reviewed. It is shown that modern optimal control algorithms often do not lead to useful controller designs. They are useful only as tools in heuristic design methods. Their main problem is sensitivity to the nature of the process transfer function and the lack of conservative design methods. The problems faced in developing better algorithms are outlined.

Downloaded via 63.209.178.11 on February 27, 2023 at 14:06:07 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

1. Introduction

In recent years there has been a growing literature on the applications of optimal control techniques in the process industries (Bryson, 1967; Cegla, 1969; Denn, 1972; Douglass, 1972; Gould, 1969; Lim, 1970; Lapidus, 1967; Lueck, 1968). However, it is generally felt that these methods have not made the expected impact on the industrial practice of process control, a problem which was the subject of panel discussions at several recent AIChE meetings (National Science Foundation Workshop, 1973; Foss, 1973).

Often the hypothesis is raised that the lack of applications is due to the lack of sophistication of the practitioner and might be the result of the normal time lag between the conception of modern control system design techniques in the academic community and their industrial application. Some are less generous and claim that the problems dealt with in the academic world are too far removed from reality to lead to useful results. One of the authors (R.S.) has been involved in industry for many years and has often tried to apply some of these optimal control techniques with rather limited success and reluctantly has come to the conclusion that in the present state of the art it is quite difficult to apply optimal control theory to an industrial problem. It is not that the techniques are not useful. In fact, they often contribute significantly to our understanding of the problem. It is rather that they are often not in a state where their application is straightforward enough to allow their use in a reasonable amount of time without extensive study and research, and where straightforward application might even lead to serious troubles.

We therefore need to rethink our approach to the problem, and this paper is intended to be a first start in this direction. We will try to present a critical review of the application of modern control theory techniques to process control, and, although the paper contains considerable amounts of heretofore unpublished work of our own, it is not a regular research paper, but rather a special type of review.

We do not intend to cover the whole spectrum of process control, but we choose to concentrate on a rather simple and almost trivial problem: the design of analog controllers for single-input single-output systems, and specially overdamped systems. An example of such a system is the transfer function

$$G_p(s) = \frac{e^{-\theta s}}{(\tau s + 1)} \tag{1.1}$$

[1] Department of Chemical Engineering.
[2] Department of Electrical Engineering.

It is trivial in the sense that classical frequency response methods lead to satisfactory designs. Interestingly enough, a considerable part of our optimal control literature deals with examples of this type (Lim, 1970; Lueck, 1968; Cegla, 1969; Koppel, 1968; Kalman, 1960).

We started our own research with the same type of problem for the reason that a well understood example is the best test for any algorithm. If the algorithm fails to produce reasonable process performance we can thoroughly analyze the example system and hopefully understand why.

We are fully aware that the interesting problems are far more complex and that optimal control has been applied to multiple-input, multiple-output problems. We will show, however, that our approach and conclusions also apply to those cases.

An important feature of a good control system design algorithm is that it provides the practicing engineer with a framework within which to cast his problem and provides a systematic design procedure which can be applied to a larger number of similar problems. If, however, an algorithm gives unsatisfactory results, we have to understand why, and put proper safeguards into the design algorithm. Hence, in section 2 below we state our interpretation of the principal desirable characteristics that should be achieved by a process controller. In section 3 we describe and compare various design methods for the classic PID controller in terms of their achieving these goals. In section 4 we describe some approaches to the design of process controllers using frequency-domain and time-domain optimization procedures. Section 5 contains an evaluation of three competing process controller designs for a simplified problem. Our conclusions are summarized in section 6.

2. Design Criteria for Process Control

Before discussing in detail the advantages or disadvantages of specific controllers, let us restate the aims and goals of process control and the specifications that such a controller must fulfill. We will list them first without any intention of ranking them. Again, we are referring here to a simple continuous feedback controller, as discussed in the optimal control literature (Horowitz, 1963; Hougen, 1964; Ziegler, 1942; Cohen, 1952), which may be part of a more complex system but which can be designed separately.

It is important to remember that in process control we seldom totally rely on such controllers, and that they normally are part of a more complex scheme which is managed by an operator who achieves the desired control of the total process by adjusting the set-points of individual controllers. Interestingly, contrary to quality control in mass production by machine tools, there has been little systematic research on the way an operator should or does control a

Petitioners Lenovo (United States) Inc. and Motorola Mobility LLC - Ex. 1016

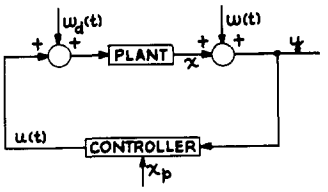

Figure 1. Process control configuration; $G_p(s)$ = plant transfer function; $G_c(s)$ = controller transfer function; x_p = set point.

complex unit. But we still have a pretty good idea on what demands are put on the controller. In the following an attempt is made to summarize these into six criteria.

(1) Ability to Maintain the Controlled Variable at a Given Set-Point. The first demand seems rather trivial, as this is the most obvious goal of process control. But this most essential requirement of process control is often the most difficult to fulfill as it creates mathematical difficulties for most of the optimization algorithms proposed thus far in the process control literature (Koppel, 1968) and we therefore would like to define it rather precisely.

The fact that the process of Figure 1 can be controlled by a controller measuring a single variable $y(t)$ and manipulating another u does not mean that the process has a single input. In fact it may have several feed-streams as well as other uncontrolled inputs such as environmental temperatures, moisture, etc. When the operator makes a set-point change he may also simultaneously change the set-points of other units which may change some of the input to the process under control. The value of the controller is that it is able to maintain the controlled variable x at a given set-point x_p and compensate for all the other unknown and changing inputs to the plant. Obviously, in a real system it can only accomplish this for a given range of inputs, and when one of the inputs exceeds that range the operator must step in. But within its operating range this controller must be able to handle the process to be controlled, without any specific knowledge about the value of the different inputs.

Therefore, when changing a set-point an operator has only very imperfect knowledge as to what the proper steady-state control should be and the controller must be able to estimate this input, which in conventional control is achieved by the integral control mode.

(2) Set-Point Changes Should be Fast and Smooth. As the overall system may be slow and complex, it is important for the operator to be able to perform individual set-point changes as fast as possible. However, minimum time response often leads to large excursions in the system transient response, and smooth response (or low overshoot) has a significant advantage. The operator normally does not know what the proper set-point is. If there is severe overshoot he will normally try to counteract and thereby often aggravate the problem. While this problem of minimum-time response has attracted a great deal of attention in the optimal control literature (Bryson and Ho, 1967) and while it is a very important one, fast time response cannot be treated in isolation and is only one part of a process controller specification.

(3) Asymptotic Stability and Satisfactory Performance for a Wide Range of Frequencies. The total system (not necessarily the controller) should obviously be asymptotically stable to be suitable for operator control. This asymptotic stability should be achieved even though the process parameters may change within a reasonable range of system parameter values.

Furthermore, the closed-loop transfer function frequency response should not have peaks indicating strong amplification of certain input signals. This means that the maximum amplification in the transfer function from disturbance input to process output should be low. This should be true for disturbances such as w_d in Figure 1, which are filtered through the entire plant as well as measurement noise, disturbances which appear in an "unfiltered" form such as $w(t)$ in Figure 1. Most of the inputs to a plant are filtered through the plant, and, though the transfer function from each disturbance to the output may be different, they all may have the same denominator. Some disturbances, however, may be created in the process itself or may be entering the process in a later stage. The first type are normally the more important as in most cases they determine the steady-state level of the manipulated inputs as well as of the uncontrolled state variables.

(4) The Controller Should be Designable with a Minimum of Information with Respect to the Nature of the Input and the Structure of the System. In many cases in process control we have a rather imprecise knowledge of the nature of the disturbances and their variation with time. If the payoff is sufficient, we can try to obtain that knowledge. But even if we can obtain this knowledge we would still like the controller to protect us against contingencies, sudden changes in its inputs, etc.

The second requirement of obtaining a controller design despite the lack of knowledge of the system structure is more stringent. In most situations the process to be controlled is only inaccurately known, or is so complex that we try to get away with using a design based upon a simple approximate model of the actual process. We have to be careful that the control action achieved in theory is not strongly dependent on that part of our model which is inaccurate. A short example will illustrate this. Consider, for example, a distributed-parameter system (as for example, a heat exchanger), which features both mixing process and transport delays. For mixing studies we might successfully model it as a series of three stirred tanks. However, if we design an optimal controller for three stirred tanks we might obtain a controller which combines derivative action with a very high gain. While this would function well in three stirred tanks, it will lead to instability in the real system due to the finite timelags involved. It is this need for structural insensitivity which is the hardest one to evaluate in practical applications of optimal control and will be discussed in detail later.

Any design method must therefore be related to a method of systems identification and take into account the difference between the model and the real system.

(5) The Controller Should be Insensitive to Change in System Parameters. In a real control situation the parameters of the system and of noise parameters are not accurately known and in addition often change with time. The controller must be able to handle reasonable changes, with a sufficient stability margin. The reason for this need is twofold. First, the throughput through process equipment changes due to varying overall needs of the plant. That means in a process with a time lag the controller must be able to perform while the actual time constants of the system change and these changes are in no way negligible (a factor of 2 is a reasonable range for many processes). The second reason is that as mentioned above, linear system equations are often a linearization around a steady state and when the steady-state set-point is changed these linearized system parameters may change significantly.

(6) Excessive Control Actions Should be Avoided. There are two main reasons for limiting the control effort. The first is mathematical. When dealing with a linear problem we neglect one important nonlinearity, the finite limits on the magnitude of allowed control signals. To avoid errors we must put reasonable limits on magnitude of control

or we have to take this nonlinearity into account in our design. Strong control action might also be costly and there is an economic reason for limiting control action as, for example, in space flight. In process control reducing the control actions is seldom of economic significance as there are very few cases when a minimization of control effort can really be justified on cost considerations alone.

If one looks at any of the optimum control strategies as to the way they fulfill these requirements it becomes obvious that in the present state of the art it is impossible to incorporate all of them simultaneously into an algorithm. All existing algorithms are written for one or two of the above points and the rest must be tested for in a rather pragmatic way. After several such tests in practical situations it soon becomes evident that some of these requirements are contradictory, such as 2 and 4, and that a compromise is needed. And here we come to a basic problem in the application of optimal design methods for problems which basically demand a compromise solution. It is not clear a priori that a method that searches for an optimum for one specification or maybe for two together (2 and 6 are quite accessible together) leads even to the proper specification of the structure of a controller which will give a sensible compromise for all of them. One of the advantages of the PID controller is that we have for its design, methods which will almost always lead to a reasonable, workable compromise. Hence, before we can find wide general applications to optimal control, we have to work out similar methods which either guarantee a sensible compromise or tell the designer more clearly the type of problems for which the method will lead to a good overall controller.

To illustrate the problems involved we will discuss in more detail the properties of PID controllers and of controllers designed using the framework of optimization procedures. PID controllers have been selected since a large literature on PID controller design exists although for many processes simpler PI, or even proportional controllers, would often be quite adequate. The optimal controllers to be examined are designed using the theory outlined in section 4.

3. The PID Controller

Several conventional design methods have been proposed in the literature for the PID Controller (Horowitz, 1963; Ziegler, 1942; Cohen and Coon, 1953; Hougen, 1964; Gould, 1969) and some, such as the Cohen and Coon method, deal specifically with the case which is the base for our comparison, an overdamped system with delay. It is interesting to note that all of these design techniques lead to controllers which are almost identical (Koppel, 1968), despite the fact that they are based on different design procedures. We will indicate below one possible reason for this similarity in controller structures.

Any useful design method must have an identification or modelling procedure as an integral part (Seinfeld, 1974). Let us first define the design problem. We start with a process, and we will consider here a case of single manipulated variable u and a single measured variable y, keeping in mind that the set point of y is adjusted by other measurements. If we linearize the system then we can write a linear transfer function

$$Y(s) = G_p(u(s))$$

Our first problem is to get an estimate of $G_p[u(s)]$. This is the initial step in any controller design. One way is to set up a simplified theoretical model of the system and identify the parameters of the model experimentally. Or we can base the model on identification experiments, manipulating u and measuring y. $G_p(s)$ can be directly measured by using sinusoidal inputs. Another way is to use a step input or a pulse in u and fit a simplified model to the resulting output. Or $G_p(s)$ can also be computed numerically from the response to a step or a pulse (Hougen, 1964).

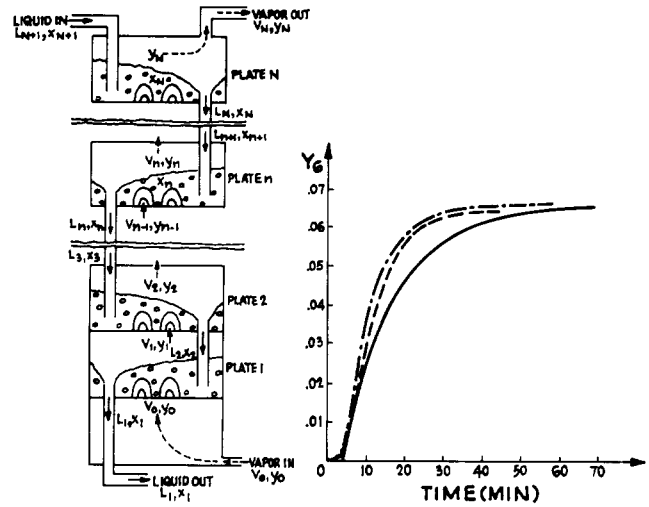

Figure 2. Step response of plate gas absorption unit and two approximations (J. M. Douglas): - - - -, exact response; -, first-order approximation; $G_p(s) = 0.221e^{-3.65s}/(14.15s + 1)$; – · —, second-order approximation; $G_p(s) = 0.221e^{-1.4s}/[(7.08s + 1)\cdot(3.4s + 1)]$.

In practice such experiments have often limited accuracy due to the presence of disturbances and measurement noise, hence requiring sophisticated statistical identification procedures. Furthermore, $G_p(s)$ changes with throughput, and also with the set point y_{ss}, since $G_p(s)$ results from a linearization.

One often used method (Ziegler, 1942) avoids the problem of identifying $G_p(s)$ by just identifying two critical parameters associated with $G_p(s)$ the critical frequency ω_{cr}, at which the phase lag is 180° and the critical gain which is the inverse of the amplitude ratio at this frequency. (All our gains are normalized such that $\Delta y/\Delta u$ at steady state is unity.) These parameters can be measured by using proportional control and increasing the gain until the system starts to cycle. The settings of the PID controller are

$$U(s) = -K_c\left(1 + \frac{1}{\tau_I s} + \tau_D s\right) \tag{3.1}$$

when $K_c = 0.6K_{cr}$ and $\tau = \pi/\omega_{cr}$ and $\tau_D = \pi/\omega_{cr}$.

In practice identification by cycling is sometimes difficult and one can get similar results by estimating K_{cr} and ω_{cr} from either a step or a pulse response. Ziegler (1942) suggested using the step response and fitting it using the model given in eq (1.1). Parameters are obtained by drawing a tangent through the inflection point, the slope being $1/\tau$, and θ being the intercept with the time base. This procedure is only good for overdamped systems. Other models have been proposed (Murril, 1967; Koppel, 1968). An example of such an approximation is given in Figure 2, which gives the step response for an extractor (Douglass, 1972). Murril uses the same model only with a smaller value of τ ($(\tau + \theta)$ equal to the time when the step reaches 0.63 of the final value), and Koppel uses a lag in conjunction with a second-order system). Let us remind ourselves that neither the complex theoretical response nor any of the models really represents an extraction column.

Before discussing other modelling and design methods, let us first look at the performance of a typical PID controller. Let us for a moment assume that the system really is given by eq 1.1 and that $\theta = 0.5$ and $\tau = 1.0$.

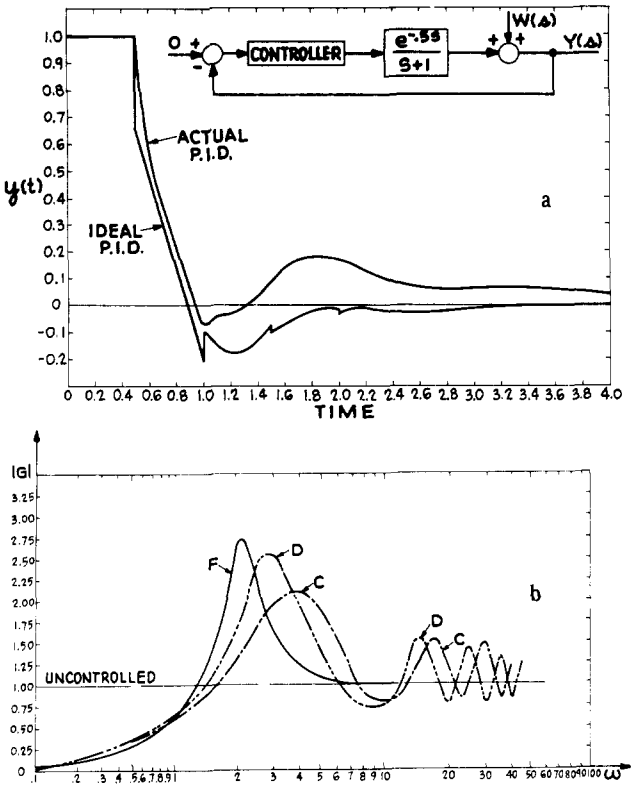

Figure 3. (a) Time response of PID controller to output unit step disturbance. (b) Magnitude of transfer function from output disturbance to output response: C, PID control, exact parameters; D, PID control, disturbed parameters ($\beta = 0.2$); F, PID control, different structure, $G_p(s) = 1/(1 + 0.375S)^4$.

In Figure 3a we give the response of the system to a unit-step disturbance in the load $w(t)$ for both the ideal PID controller and for a practical PID controller. The main difference between the two is that the practical PID controller has a high and low frequency filter in addition to the terms given by (3.1). The transfer function is

$$G_c(s) = \frac{u(s)}{y(s)} = \frac{K_c\left(1 + \tau_D s + \dfrac{1}{\tau_I s}\right)}{\dfrac{\tau_D}{\gamma}s + 1 + \dfrac{1}{\alpha\tau_I s}} \tag{3.2}$$

where

$$\gamma \gg \tau_D,\ \alpha \gg \frac{1}{\tau_I}$$

In general, the effect of the low-frequency filter is very small and for all practical applications we may neglect it. $G_c(s)$ then becomes

$$G_c(s) = \frac{K_c\left(1 + \tau_D s + \dfrac{1}{\tau_I s}\right)}{\left(\dfrac{\tau_D}{\gamma}S + 1\right)} \tag{3.3}$$

The controller parameters used to generate Figures 3a and 3b are: $\tau_D = 0.17$, $\tau_I = 1$, $K_c = 2$, $\alpha = 8$. [The actual Ziegler–Nichols settings would be $K_c = 2.4$, $\tau_I = 1$, $\tau_D = 0.25$. Our settings are slightly more conservative.] From the step response we note that the performance of the actual PID controller is slightly worse than that of the ideal PID controller. However, controller (3.3) eliminates one of the main problems of derivative control action, which is the high control effort due to disturbances with appreciable high-frequency content. For completeness we give also the frequency response of system (1.1) controlled by a PID controller (Figure 3b) and the Nyquist plot (Figures 4 and 5).

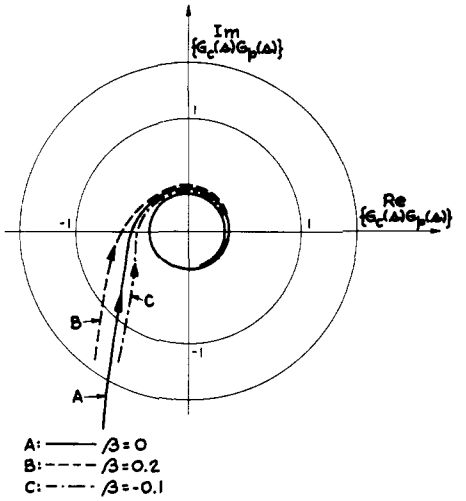

Figure 4. The effect of parameter changes on Nyquist plot of system using ideal PID controller. System configuration as in Figure 3a. $G_p(s) = e^{-s\theta}/(\tau s + 1)$; $\tau = 1$; $\beta = \theta/\theta_c - 1$; $\theta_c = 0.5$; $G_c(s) = 2(1 + (1/s) + 0.17s)$.

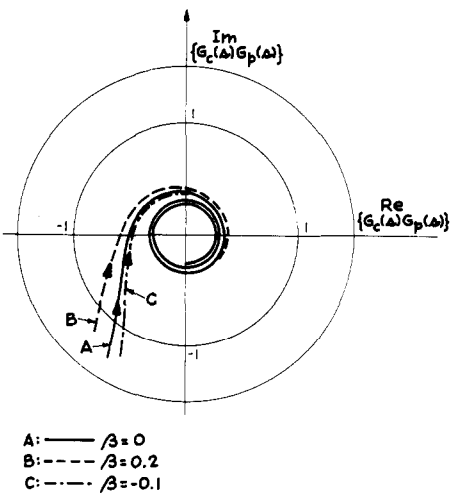

Figure 5. The effect of parameter changes on Nyquist plot of system using actual PID controller. System configuration as in Figure 3a. $G_p(s) = e^{-s\theta}/\tau s + 1$; $\tau = 1$; $\beta = \theta/\theta_c - 1$; $\theta_c = 0.5$; $G_c(s) = 2(1 + (1/s) + 0.17s)/(0.17/8)(s + 1)$.

The Ziegler–Nichols settings on a PID controller give a reasonable compromise among all of the requirements we defined in the previous section. In addition to the ability to maintain a desired set point (requirement 1), it has a good step response (requirement 2) and reasonable low amplification (requirements 3 and 6); it is quite insensitive to parameter changes and not very sensitive to the structure of the plant (requirement 5). This will be demonstrated in section 5 on evaluation.

One of the frequently encountered process changes in chemical processes is an increase or reduction of throughput. This causes a change in the residence time of the plant, thus changing the characteristic times in the plant's model. However, the ratio between these characteristic times may not change. Obviously, we would take care of extreme changes by changing the controller parameter settings but the controller should not be too sensitive to such changes even at fixed settings. In Figure 3b we give the frequency response of the output y to output disturbances w when no change in throughput has occurred (curve c) and when a change in throughput has occurred such that 20% error in the previously assumed time characteristics have

Table I. Controller Settings for System with Transfer Function $[1/(1 + (s/n))]^n$ Using Different Identification Methods and Ziegler—Nichols Controller Settings

Identification method	Controller settings method	n	2	3	4	5
(1) Ziegler—Nichols ultimate (closed loop) method	Ziegler—Nichols loop turnings method	$k_c$max	∞	8	4	2.89
		ω_{cr}	—	5.2	4	3.63
		K_c		4.80	2.40	1.73
		τ_I	—	0.60	0.79	0.89
		τ_D	—	0.15	0.20	0.22
(2) Ziegler—Nichols process reaction curve method	Ziegler–Nichols process reaction curve method	θ	0.146	0.27	0.36	0.42
		τ	1.364	1.23	1.116	1.024
		K_c	11.21	5.47	3.72	2.93
		τ_I	0.292	0.54	0.72	0.84
		τ_D	0.073	0.14	0.18	0.21
(3) P. W. Murril method	Ziegler–Nichols	θ	0.146	0.27	0.36	0.42
		τ^a	0.92	0.81	0.72	0.66
		K_c	7.4	3.6	2.47	1.88
		τ_I	0.292	0.54	0.72	0.84
		τ_D	0.073	0.14	0.18	0.21

[a] $\tau = t_{0.632} - \theta$.

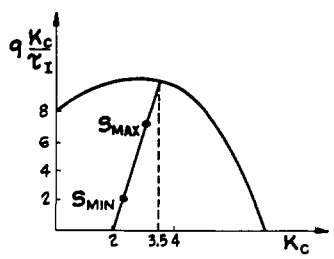

Figure 6. Parameter plane for PI controller.

resulted (curve D). We observe that the difference between the two curves is not significant and the controller has a very reasonable frequency response. We also plot the response of the controller to a different but similar transfer function model. It is also insensitive to the exact form of the plant model assumed since the only important parameters of the model are the maximum allowable gain $K_{cr\,max}$ (K_{cr} is the maximum gain that can be used (still maintain stability) when only proportional control is used.) and the critical frequency K_{cr} (τ_D and τ_I will depend on ω_{cr}). (ω_{cr} is the frequency of oscillation of the system when only proportional control of gain K_{cr} is used.) In any case, approximate knowledge of those features (K_{cr} and ω_{cr}) is a minimum requirement for any controller design.

Other settings for K_c, τ_I, and τ_D have been proposed but they are all rather close to the Ziegler setting. Cohen and Coon dealt with a method based on the model given in (1.1). They first found an optimal controller by using settings which minimize the integral square error of a step response and as this controller is on the verge of instability, they modify the settings on the basis of stability considerations.

There is a third way devised by the authors whose value for this case is merely didactic as it leads to similar results. In other cases this point of view might be useful since it may lead to a more direct design procedure. We try to combine the first three requirements mentioned in section 2, in terms of requirements on the frequency response of the system. Denoting the magnitude of the amplitude ratio of the frequency response of the system by $G(\omega)$ we may improve conditions that will satisfy the actual physical requirements. The first requirement dictates that $G(0) = 0$. To ensure fast response $G(\omega)$ should stay as low as possible at low frequencies. This can be specified by

$$\left.\frac{dG(\omega)}{d\omega}\right|_{\omega\to 0} \leq \epsilon$$

Requirement 3 can be expressed quantitatively by limiting the maximum value of $G(\omega)$. However, the last two specifications may be contradicting indicating the need for a compromise. If we specify

$$\left.\frac{dG(\omega)}{d\omega}\right|_{\omega\to 0} = \epsilon$$

then the settings that minimize $G_{max}(\omega)$ (where $G_{max}(\omega) = \max_\omega G(\omega)$) might be considered to be the best, or when $G_{max}(\omega)$ is specified one might try to minimize

$$\left.\frac{dG(\omega)}{d\omega}\right|_{\omega\to 0}$$

Hence, what is needed is not a controller that is optimum in a mathematical sense but a "mini-max" controller.

For example, when a system consisting of three stirred tanks is considered (the one stirred tank and delay is mathematically more tedious) with a PI controller it is found that all mini-max settings lie approximately on a line in the K_c–K_c/τ_I plane as indicated by the solid line in Figure 6. Note that both the Ziegler–Nichols and the Cohen and Coon controller settings are close to the mini-max line. It is also interesting to note that there is a minimum controller parameter setting S_{min} on the line below which smaller values of K_c have very little effect on $G_{max}(\omega)$. Similarly there is an upper limiting parameter setting S_{max} which is determined by the need to ensure a reasonable safety margin with respect to closed-loop system stability. Clearly, this mini-max procedure does not take into account stability in a direct way. Stability is taken into consideration indirectly by limiting the maximum allowable value of $G(\omega)$.

More complex methods (Gould, 1969; Horowitz, 1963; Hougen, 1964) deal directly with the closed-loop frequency response $G(s)$, but, while the criteria differ, the net results sought are similar: to allow good control at low frequencies without amplifying the resonance frequency. Frequency response methods allow one, therefore, to make a compromise between criteria 1–4 in a direct way, and are preferable whenever this is possible or justifiable.

Criterion 5 is harder and is often overlooked. If K_{cr} and ω_{cr} are measured experimentally, this is taken care of automatically. If models are used, this is more problematic. The original procedure of Ziegler and Cohen and Coon does not always guarantee good results as can be shown by a simple example. In Table I we give the results of this method for the transfer function $1/(1 + (s/n))^n$ and we note that for $n = 5$ the Z–N reaction curve identifying method gives an unstable controller. The method proposed by Murril is more conservative and gives here good results. Methods based on

$G_p(s)$ itself will not have this problem, and if $G_p(s)$ is inaccurately measured we can always be conservative. In practice an experienced control engineer would be able to adjust K_c and τ_I to ensure stability even in this case. We brought it to show that this modelling procedure has a far greater effect on performance than the controller design algorithm. Let us, for example, consider a cascade of three identical stirred tanks, for which $K_{c\,max} = 8$ when only proportional controller is used. If we would really believe that the three stirred tank model accurately represents the process to be controlled, we could, by appropriately designing a controller (for example, using a strong derivative control with a large proportional gain) actually achieve nearly perfect control.

However, this may be a fiction. A real system has a finite delay in it and a high gain K_c would cause that system to be unstable. Thus, if we want to stabilize the system by derivative action we need much more detailed information about the system itself to guarantee the stability.

4. Optimization Techniques

Optimization procedures that have been offered for the design of process controllers usually involve a two-step design procedure: (1) the mathematical minimization of a scalar functional performance criterion

$$V[\mathbf{x}(t), \mathbf{u}(t)] \tag{4.1}$$

where $\mathbf{x}$ denotes the process dynamic state vector, $\mathbf{u}$ denotes the control vector which is to be determined, and where the control interval may be finite or infinite; (2) the physical realization of a feedback control law

$$\mathbf{u} = \phi(\mathbf{x}, t) \tag{4.2}$$

to effect the minimization of (4.1).

The optimization procedures that have received the greatest attention for process plant regulation are those that involve a quadratic performance criterion. The problems are stated either in a deterministic formulation where a controller is sought that minimizes an integral involving quadratic forms in the state and control variables for a given initial deviation in state, or in a stochastic formulation where the controller is sought to minimize the expected value of an integral of a weighted sum of quadratic forms in the state and control variables, conditioned on knowledge of the random processes perturbing the process to be controlled and on the available measurements.

4.1 Deterministic Formulation. One class of deterministic optimum control problems can be stated as follows: given the linear system of first-order differential equations

$$\mathbf{x} = A\mathbf{x} + \mathbf{b}u(t),\ \mathbf{x}(0) = \mathbf{x}_0 \tag{4.3}$$

where $\mathbf{x}$ is an n vector and u is a scalar input, find the feedback control law

$$u = f(\mathbf{x}(t)) \tag{4.4}$$

that will minimize the quadratic performance criterion

$$J = \frac{1}{2}\int_0^\infty (\mathbf{x}'Q\mathbf{x} + ru^2)\,dt \tag{4.5}$$

where $(\ )'$ denotes the transpose of a matrix or vector. Under the conditions that (a) the pair (A, B) is completely controllable (Kalman) (b) matrices Q and R are positive semi-definite and positive definite, respectively, the optimum control law for this problem is a linear feedback control law

$$u = -\mathbf{k}'\mathbf{x} \tag{4.6}$$

where

$$\mathbf{k}' = r^{-1}\mathbf{b}'M \tag{4.7}$$

and M is the positive-definite solution of the following nonlinear algebraic equation

$$MA + A'M - r^{-1}M\mathbf{b}\mathbf{b}'M + Q = 0 \tag{4.8}$$

When conditions (a) and (b) above are satisfied the closed-loop system that results from use of control law (4.6) is guaranteed to be asymptotically stable. Note, however, that control law (4.6) requires measurement of all process state variables. This may be costly or physically impossible in certain practical applications. It is also important to realize that care must be taken in properly translating a physical process control problem into the above mathematical framework in order to ensure that the resulting controller will yield a feedback control system that meets the performance requirements mentioned in section 2 above.

When a change in process steady-state operation is desired, one standard procedure for employing the above formulation is as follows: the dynamics of the nonlinear process are expanded in a Taylor series about the desired steady state to yield a system of linear differential equations

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}u \tag{4.9}$$

which approximately describes the dynamics of state deviations $\mathbf{x}(t)$ away from equilibrium as a function of control deviations $u(t)$ away from the controller setting required to maintain equilibrium.

However, an important characteristic of control problems for chemical processes is that the required control u_s to maintain a desired steady state $\mathbf{x}_s$ will not generally be known so that the parameters of the linear system will not be known precisely. Hence, a designer will use a nominal linearized model which we will denote by

$$\dot{\mathbf{x}}_c = A_c\mathbf{x}_c + \mathbf{b}_c u_c(t) \tag{4.10}$$

Johnson (1971) has shown that parameter uncertainty of this kind can be taken into account by including "step-disturbance" terms in the formulation of the linearized deviation system which is to be controlled.

In early attempts to apply optimum control theory (Koppel, 1968) to chemical processes, parameter uncertainties and external disturbances were often ignored and linear models of the form (4.10) were employed. When these parameter uncertainties or, equivalently, step disturbances are ignored in the formulation of the first-order state equation for the plant of (1.1) it is found that the controller transfer function is (Koppel, 1968)

$$G_c(s) = \frac{Ke^{-\theta_c}(\tau_c s + 1)}{(\tau_c s + 1) + K(1 - e^{-\theta_c(1+s)})} \tag{4.11}$$

where θ_c is the nominal value of the plant transport delay and τ_c is the nominal plant time constant. K is the gain of the controller and depends on the weight given to the control effort in (4.5). When no weight is given to control effort, the transfer function of the feedback controller becomes

$$G_c(s) = \frac{e^{-\theta_c}(\tau_c s + 1)}{1 - e^{-\theta_c(1+s)}} \tag{4.12}$$

An undesirable feature of the controllers of (4.11) and (4.12) is that steady-state offsets will occur in some state variables. Section 5 contains a more detailed evaluation of these controllers with reference to the design criteria of section 2.

If the designer desires to account for parameter uncertainties or external step disturbances, a class of compensa-

tors, called observer-based controllers after the work of Luenberger and others (Luenberger, 1960; Luenberger, 1971), might be employed. These controllers act on the available measurements to provide estimates of unmeasured state variables and unmeasurable disturbances. The design of this type of controller is based on Johnson (1971) and Luenberger (1960): the step disturbance resulting from applying an inaccurate steady-state control action is modeled as a first-order system and an additional state variable augmented to the system equations. The constant disturbance is then approximately reconstructed using a Luenberger observer. Following the work of Johnson (1975), constant disturbances can be counteracted by the control law, which is designed to be a sum of two terms, one that acts to cancel disturbances, and the other which acts to satisfy the design requirements as if the constant disturbance did not exist. Although the resulting controller will no longer be strictly optimum for performance index (4.5), this kind of procedure will result in an integral control action, thus eliminating off-sets of some states.

For a single manipulated-input, single measured-output system the construction of the observer eliminates the need for taking derivatives in order to reconstruct the state vector. Indeed, for a second-order system the compensator achieved by the above method resembles a PID controller where the actual system "derivative" action is an approximation to a pure derivative, thus tending to eliminate frequent saturation of the controller due to the effect of measurement noise.

When applying this procedure to the base comparison system the resulting feedback controller becomes

$$G_c(s) = \frac{[c_1(s-f)+c_2](s+1)}{s^2-(f-1-k)s+se^{-\theta s}[-fe^{-\theta}(f-1)+f(1-k)]\cdot f(1+k)[1-e^{-\theta s}]} \tag{4.13}$$

where

$$c_1 = ke^{-\theta} - f[1+k(1-e^{-\theta})] \qquad (f = \text{observer pole}, <0)$$

$$c_2 = -(1+f)[1+k(1-e^{-\theta})] \qquad (k = \text{controller gain})$$

We see that this controller has a pole at zero which indicates that integral control action will be applied in order to eliminate offsets.

In this design the value k is to be determined on the basis of the relative weights given to the state and control effort in the performance index and the value of f is arbitrary but negative. However, in order for the observer to accurately reconstruct the state of the system, the value of f should be larger (one order of magnitude or more) than the value of $1+k$, since only then will the eigenvalue of the observer be much more negative than the eigenvalue of the system it observes, thus causing the observation error to rapidly approach zero.

The effect of modelling errors on observer controllers has been investigated by two of the authors (Thau and Kestenbaum, 1973) and it is shown that caution must be taken when designing an observer, both in regard to measurement noise amplification and in regard to sensitivity with respect to parameter variations.

Another technique for eliminating offsets (O'Connor and Denn, 1972) is based on an idea developed by a number of investigators in which a performance criterion of the form

$$V = \frac{1}{2}\int_0^\infty [C^2x^2(t) + \dot{u}^2(t)]\,dt \tag{4.14}$$

is minimized for a system described by (4.9). In order to apply the approach to processes with transport delay the exponential term in the transfer function of the system must be approximated by a rational transfer function. If the approximating transfer function has a denominator of first order, the resulting optimum controller will have the same form as a PID controller. If the dynamics of the process or of the approximating transfer function is of higher order, then measurement of additional state variables or a controller with higher-order derivatives will be required.

4.2 Stochastic Formulation. The above approach of considering a linear process model and a quadratic performance criterion can be extended to provide a design procedure which explicitly accounts for stochastic disturbances to the process dynamics and for measurement noise.

(a) First consider the case where a white noise random process perturbs the process dynamics

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}u + \mathbf{w}(t) \tag{4.15}$$

where $\mathbf{w}(t)$ is white Gaussian noise with zero mean, $E\{\mathbf{w}(t)\} = 0$, and with covariance matrix

$$E\{\mathbf{w}(t)\mathbf{w}'(t_1)\} = W\delta(t-t_1)$$

The performance criterion to be minimized is

$$J = E\left\{\frac{1}{2}\int_0^\infty (\mathbf{x}'Q\mathbf{x} + ru^2)\,dt\right\} \tag{4.16}$$

It can be shown that the control law that minimize (4.16) is (4.6), the same control law that minimizes (4.5) for the deterministic control problem. The optimum value of (4.16), of course, differs from the optimum value of (4.5) because of the presence of the disturbance $\mathbf{w}(t)$ in (4.15).

(b) Next, assume that the process to be controlled is given by (4.15) but that the measurements available to controller are perturbed by random noise and are given by

$$\mathbf{y} = C\mathbf{x} + \mathbf{v}(t) \tag{4.17}$$

where the measurement vector $\mathbf{y}$ is of dimension $m \le n$ and $\mathbf{v}(t)$ is zero-mean white Gaussian noise with

$$E\{\mathbf{v}(t)\mathbf{v}'(t_1)\} = V\delta(t-t_1)$$

The performance criterion is (4.16) but now the expectation is conditioned on the available measurements.

The solution to this problem is expressed by the Separation Theorem (Bryson and Ho, 1967): the optimum controller comprises a Kalman–Bucy filter which provides a minimum-variance estimate $\hat{\mathbf{x}}(t)$ of the process state and a control law

$$u = -\frac{1}{r}\mathbf{b}'M\hat{\mathbf{x}}(t) \tag{4.18}$$

that has the same form as the deterministic optimum control law (4.6). The optimum estimate $\hat{\mathbf{x}}(t)$ is generated by a dynamic system

$$\dot{\hat{\mathbf{x}}} = A\hat{\mathbf{x}} + \mathbf{b}u(t) + S(t)[\mathbf{y} - C\hat{\mathbf{x}}],\ \mathbf{x}(0) = 0 \tag{4.19}$$

where time-varying gain matrix $S(t)$ is given by

$$S = P(t)C'V^{-1} \tag{4.20}$$

and $P(t)$ is the positive-definite solution of the matrix Riccati equation

$$\dot{P} = AP + PA' + W - PC'V^{-1}CP \tag{4.21}$$

If it is assumed that the measurements available to the controller have been observed over the entire interval $-\infty \le \tau \le t$, where t denotes the current time, then the control-

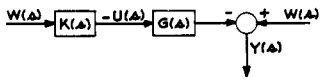

Figure 7. Open loop configuration for Wiener design. $G_c(s) = -[k(s)/(1 - k(s)G_p(s))]$.

ler has the time-invariant structure specified by (4.18) and (4.19) with a constant gain matrix S given by (4.20) where the matrix P is the solution to the nonlinear algebraic equation

$$0 = AP + PA' + W - PC'V^{-1}CP \tag{4.22}$$

Extension of this separation theorem to systems with time delays yields a controller identical in form to (4.11). One may observe the resemblance between the gain matrix S of the Kalman filter eq 4.19 and the gain matrix G of the Luenberger observer.

While the Kalman filter gain (S) is calculated on the basis of the knowledge of the measurement noise coveriance matrix V the observer gain is arbitrary as long as it satisfies the requirement that the observer error dynamics by asymptotically stable (Luenberger, 1971).

Earlier work (Cegla, 1969) on the control of processes subject to random disturbances was based on techniques developed by Wiener (Wiener, 1949; Newton, 1957) for the filtering and prediction of stationary random processes. Using this approach the performance criterion to be minimized is given by

$$J = E\{y^2(t)\} + \lambda E\{u^2(t)\} \tag{4.23}$$

where λ is a positive number denoting the relative weight of control effort. Following the development described by Cegla (1969), it is assumed that the output disturbance w has zero mean, that the spectral density $\Phi_{ww}(s)$ is known, and that the plant transfer function $G_p(s)$ is also known. Minimization of (4.23) is performed on an equivalent open-loop system, Figure 7, consisting of the plant and a cascaded compensator, $K(s)$, from which the feedback controller is then constructed. After $K(s)$ is found via spectral factorization, the feedback controller $G_c(s)$ is given by

$$G_c(s) = -\frac{K(s)}{1 - K(s)G_p(s)} \tag{4.24}$$

It is important to note that the stability of the cascaded compensator, $K(s)$, is guaranteed by the procedure, not that of the feedback controller derived from it (Lim and Bankhoff, 1970; Leuck and McGuire, 1968; Cegla, 1969). While $K(s)$ has all its poles in the left-hand plane it is not guaranteed that $G_c(s)$ will have all its poles in the left-hand plane (Cegla, 1969). The problem of the stability of the controller has been treated by Bankhoff and Cegla. The former solves the problem by constraining the structure of the cascaded compensator, but does not consider the stability of the controller itself. The latter stabilizes the controller by imposing a larger weight on the control effort (λ), which consequently changes the parameters of the controller and thus affects its stability. A method of how to choose a suitable weight, λ, that meets some constraints on the variance of the control effort is described in Lim and Bankhoff (1970).

When applied to the plant that is the base for the comparison of controllers to be discussed in section 5, we use in the design

$$G_p(s) = \frac{e^{-\theta_c s}}{1 + \tau_c s}$$

and find the controller transfer function that minimizes (4.23) to be identical with (4.11) and, of course, to be identical with the asymptotic form of the controller that employs a Kalman filter to provide an estimate of the process state for the time-delayed system with no measurement noise.

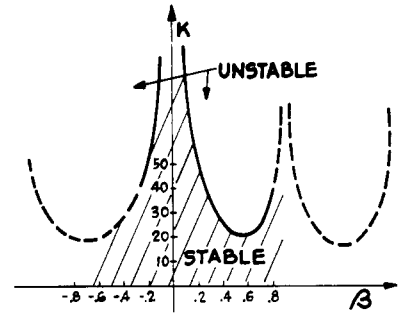

Figure 8. Stability limits of optimal controller as a function of estimation error. $G_p(s) = [e^{-\theta s}/(s + 1)]$; $G_c(s) = [ke^{-\theta_c}(s + 1)/(s + 1) + k(1 - e^{-\theta_c(1+s)})]$; $\beta = \theta/\theta_c - 1$; $\theta_c = 0.5$.

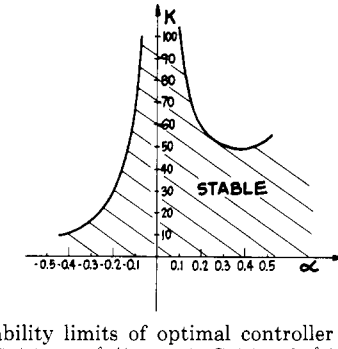

Figure 9. Stability limits of optimal controller as a function of load change. $G_p(s) = e^{-\theta s}/(\tau s + 1)$; $G_c(s) = ke^{\theta_c}(S = 1)/[(s + 1) + k(1 - e^{-\theta_c(1+s)})]$; $\theta = \theta_c(1 + \alpha)$; $\tau = \tau_c(1 + \alpha)$; $\tau_c = 1$; $\theta_c = 0.5$.

In section 5 below we examine the performance of a number of the above controllers designed using the framework of optimum control theory and compare the resulting performance with that of PID controllers.

5. Controller Evaluation

In section 3 we discussed the performance and stability of the PID controller for our design example. Here, we will try to compare it to some of the optimal controller designs given in section 4.

We will start with stability (condition 3). The methods of optimal controller design discussed above guarantee stability of the total system. But this is not sufficient. We need a sufficient stability margin to take care of inaccuracy in the model as well as the fact that parameters may change with changes in throughput, set-point, and operating conditions. For minimum phase transfer functions the unconstrained controller has an infinite gain and would become unstable if we added a small time delay to the system. This sensitivity is typical of a number of unconstrained controllers. The deterministic unconstrained controller for our model is also unstable for small changes in the lag θ_c. It can be stabilized by introducing a constraint such as the quadratic constraint on the control effort. Other ways will be discussed later.

In Figures 8 and 9 we give the stability limits of the controller for perturbations in system parameters by plotting the value of K in eq 4.11 at which the controller becomes unstable as a function of two system parameters. In Figure 8 we study the effect of changing just the lag θ keeping τ constant using as a parameter $\beta = \theta/\theta_c - 1$.

A PID controller under these circumstances is stable for any value of β less than 0.775 (or θ for our example less

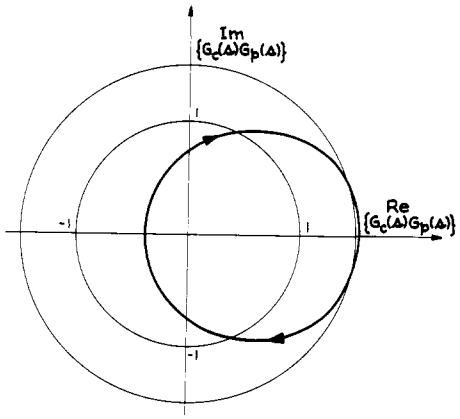

Figure 10. Nyquist plot of systems using optimal controller. $G_p(s) = [e^{-0.5s}/(s + 1)]$; $G_c(s) = [e^{-0.5}(s + 1)/1 - e^{-0.5(s+1)}]$.

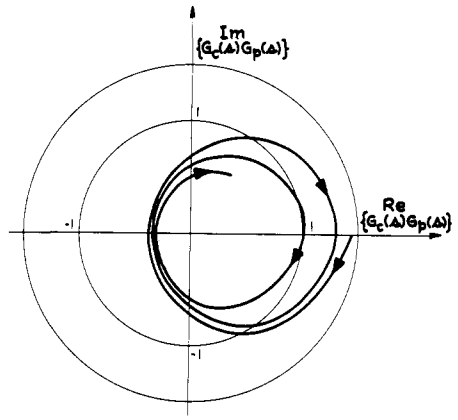

Figure 12. Nyquist plot of system using constrained optimal controller. $G_p(s) = [e^{-0.5s}/(s + 1)]$; $G_c(s) = [ke^{0.5(s+1)}/((s + 1) + k(1 - e^{-0.5(1+s)}))]$; $k = 40$.

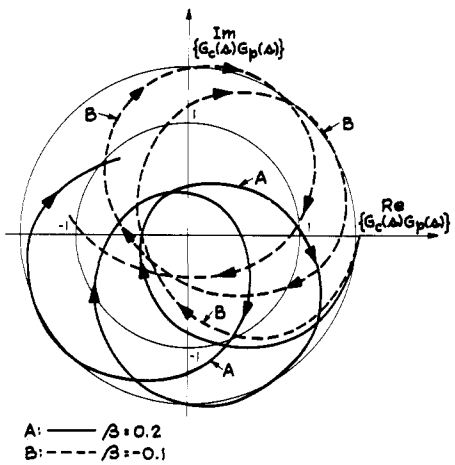

Figure 11. Effect of parameter variations on Nyquist plot of systems using unconstrained optimal controller. $G_c(s) = [e^{-0.5}(s + 1)/1 - e^{-0.5(s+1)}]$; $G_p(s) = [e^{-\theta s}/(s + 1)]$; $\beta = \theta/\theta_c - 1$; $\theta_c = 0.5$.

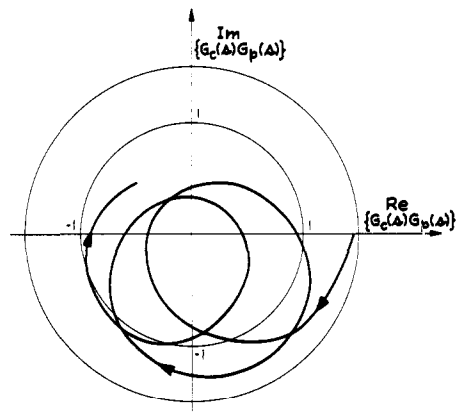

Figure 13. Nyquist plot of system with parameter variation using constrained controller. $G_p(s) = [e^{-\theta s}/(s + 1)]$; $G_c(s) = 40e^{-0.5}(s + 1)/[(s + 1) + 40(1 - e^{-0.5(1+s)})]$; $\beta = \theta/\theta_c - 1 = 0.2$; $\theta_c = 0.5$.

than 0.89). That means we can always ensure stability by slightly over-estimating θ. The optimal controller (eq 4.11) shows much more complex behavior as it is unstable for both increases and decreases in θ. The only way we can achieve a large stability margin is to choose a very small K.

In Figure 9 a similar plot is given for changes in throughputs $\alpha = ((\theta + \tau)/(\theta_c + \tau_c)) - 1$, keeping θ/τ constant. A PID controller is stable for $-0.65 < \alpha < 1.5$, or in other words, we can change throughput by a factor of 2 in both directions while still staying far from the stability limit. The unconstrained optimal controller is sensitive to changes in α in both directions, though the plot is less symmetric. Again, a low value of K is needed to ensure reasonable stability.

For another view of the stability problem we give the Nyquist diagram for some points taken from Figures 8 and 9. We note that for exact parameters the stability is excellent but for small changes in α or β the Nyquist diagram changes its form, leading to instability. (See Figures 10–13.)

In many systems, changes in throughput or set point will change α and β simultaneously. An efficient design technique must guarantee stability over the whole range of reasonable changes. It must also be insensitive to the exact nature of the transfer function. None of the present optimal design algorithms has an inbuilt mechanism to guarantee this. The same problem exists in classical control, especially if the open loop system is unstable. For the single input case discussed here, classical control solves the problem by avoiding complex stabilization.

Any controller must be able to keep the system at the desired steady state (criterion 1) and must be able to function over a wide frequency range. The easiest test for this is to look at the closed loop frequency response. Criterion 1 is fulfilled if the amplitude approaches zero as $\omega \rightarrow$ zero and reaches a low value for reasonable large values of ω. The faster it reaches zero the better is the low-frequency response. However, as shown in section 2, this will always cause amplification in the resonant range, therefore requiring a compromise. The frequency response of the controller in 4.11 for $K = 40$ is given in Figures 14–17 and compared with the PID controller. The optimal controller shows no amplification for input disturbances (Figure 14), but has unsatisfactory low-frequency behavior, due to the lack of integral action. We will discuss this later. However, the satisfactory response in the resonant region is rather illusory. In Figure 15 we give the same plot for unfiltered (or output) disturbances. We note that the optimal controller for exact parameters is only slightly better, shifting the maximum to higher frequencies. However, for a slight perturbation in β, hardly noticed by the PID controller, the system shows high amplification for some higher resonance frequencies, which is expected as $\beta = 0.2$ for $K = 40$ is close to instability.

In Figure 16 we give the control effort. We note that in the low-frequency range the PID controller has a larger control effort. This is desirable as any controller should be

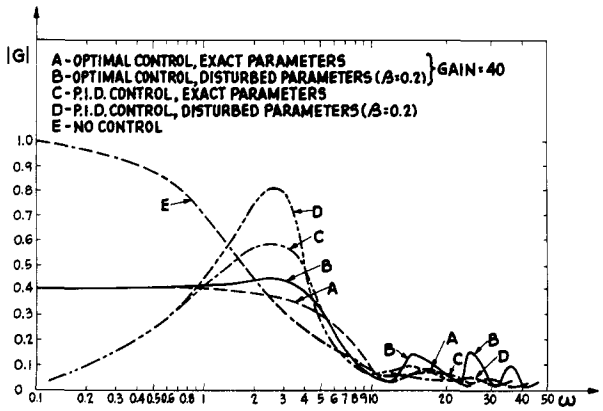

Figure 14. Magnitude of transfer function from input disturbance to output response for closed loop system.

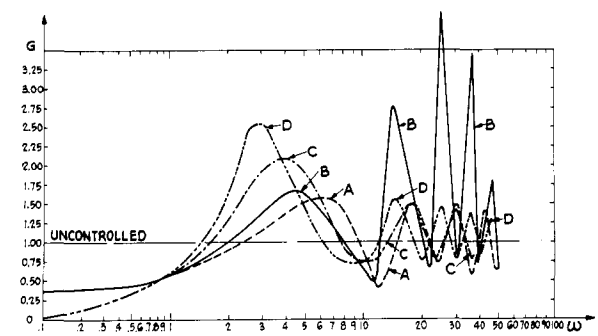

Figure 15. Magnitude of transfer function from output disturbance to output response: A, optimal control, exact parameters; B, optimal control, disturbed parameters (β = 0.2), gain = 40; C, PID control, exact parameters; D, PID control, disturbed parameters (β = 0.2).

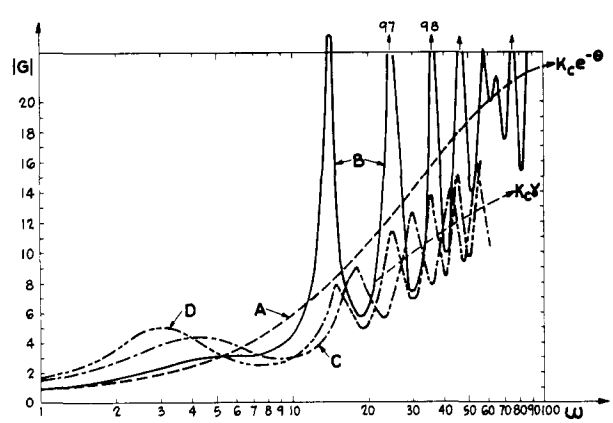

Figure 16. Magnitude of transfer function from output disturbance to controller output; A, optimal control, exact parameters; B, optimal control, disturbed parameters (β = 0.2); gain = 40; C, PID control, exact parameters; D, PID control, disturbed parameters (β = 0.2).

able to compensate for disturbances and values of $U(j\omega)/W_i(j\omega)$ less than 4 are perfectly acceptable. Even the unconstrained optimal controller (4.12) has a too low control effort (less than unity, which is the minimum needed for perfect compensation). At higher frequencies the optimum controller has higher control effort. For practical PID controllers this region is unimportant due to their filtering action.

To get good stability and frequency response for perturbed parameters we need a K around six. In Figure 17 we give the frequency response for an optimal controller (K =

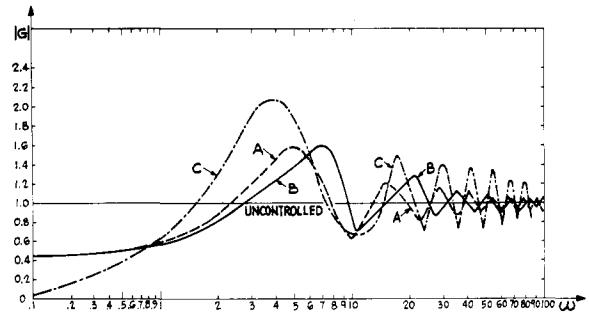

Figure 17. Magnitude of frequency response from output disturbance to process output for systems using optimal controller (gain = 6.2); A, optimal control, exact parameters; B, optimal control, disturbed parameters (β = 0.2); gain = 40; C, PID control, exact parameters.

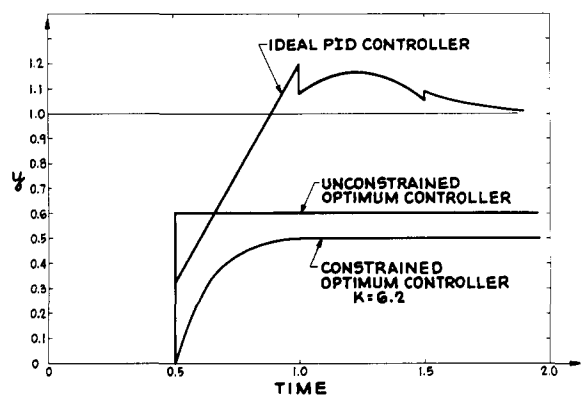

Figure 18. Response of closed loop system to step change in set point.

6.2), which shows good response in the higher frequencies for β = 0.2, K = 6.2 corresponds to a strong constraint giving hardly any control. In Figure 18 we give the response of this controller to a step change in set point and we note that the speed of the response is about the same as for our PID controller.

Good frequency response is not the sole criterion for a good controller and it is perfectly legitimate to put other criteria, such as fast step response or good response for some typical process disturbances, into our design criteria, but such a controller must, in addition, also have satisfactory stability and frequency response. By using a K low enough to guarantee this, we wound up with a controller which in the criterion chosen is hardly better than a standard controller, while being worse in any other criterion. It is "optimal" in only one sense, having a low control effort for a single perturbation, which is of no interest whatsoever to the design of process control systems. This criterion is only chosen because it is mathematically convenient. Many of the so-called optimization procedures in engineering suffer from this fallacy, as they often optimize mathematically convenient criteria with little relevance to the problem. Using relevant constraints can often give good results, but we have to understand the underlying reasons.

We want to stress that this is not a fair judgment of optimal control in general, which has many useful applications. We just brought this example as it helps to understand the limitations of optimal control.

There are other methods in optimal controller design which result in much better overall performance than the example given. But none of them in its present state is a straightforward design algorithm like those available in chemical control. Rather, these are useful tools in heuristic

controller design. Their most important application is to establish bounds as to what controllers can achieve. No feedback controller will have a faster response than an unconstrained optimal controller. Thus, we get a useful bound for fast response. Regretfully, optimal control does not give a similar bound for the frequency response, but a modified method is useful. We can compute an optimum response for a Gaussian input with a spectral density

$$\Phi(\omega) = \frac{2\nu}{\nu^2 + \omega^2}$$

and compute the ratio of the variance of y to the variance of the uncontrolled system, optimizing the controller for this input (see Lueck and McGuire, 1968; Bankhoff, 1970; and Cegla, 1969). We can then plot the optimum ratio for each and compare it to that for suggested controller. This gives useful limits as to what can be achieved.

To directly get a useful controller we have to incorporate our different criteria as constraints into the algorithm. There are various ways in which this can be done. Criterion 1, or satisfactory low-frequency behavior, can be achieved in several ways.

O'Connor and Denn (1969) suggested the use of a constraint based on the derivative of the control effort (see eq 4.14). This adds integral control but otherwise gives controllers similar to the deterministic controller. For a second-order system the unconstrained optimal controller is a proportional derivative controller. Using a constraint on $(\dot{u})^2$ one gets a PID controller. O'Connor shows that if instead of eq 2.1 a first-order Pade approximation (which is a second order system) is used, one gets a PID controller with settings very close to the Ziegler–Nichols setting. The latter therefore can also be considered an optimal controller. The procedure is ingenious but hard to generalize as in this case it strongly depended on the model chosen. It might be of wider applicability but this has not yet been demonstrated.

For our model a similar result can be achieved by adding an observer to the system (see Thau and Kestenbaum, 1973). This is a trial and error procedure, which can give good results, but the structural sensitivity and stability of the system has to be checked empirically. One example of such a controller is given in Figure 19 and it can be seen that the controller while having good overall frequency response for exact parameters, is sensitive to changes in β. Aoki shows how the system parameters can be treated as stochastic variables and included in the optimization. Unless a proper model is chosen, this does not protect against sensitivity to the structure of the model assumed. It is, however, a promising approach (see also Massaghic, 1973; Kreiselman, 1972.)

In section 4 we also mentioned stochastic techniques. A detailed evaluation is outside the scope of this paper but we checked and found them to suffer from basically the same deficiencies. Good performance for a specific disturbance does not guarantee either good overall stability nor fulfillment of the six criteria mentioned in section 2. There is, however, another use of these methods. The characteristics of the inputs provide a constraint for excessive control actions and by proper formulation of the nature of the input one can obtain good overall performance. One also can obtain integral control action by specifying a nonstationary random input. Box and Jenkins (1970) have successfully developed this approach for sampled data controllers and the method could be extended to continuous controllers. Lueck and McGuire (1968) suggested using a Gaussian noise with a correlation time close to the time constant of the system. This has poor low-frequency response. Cegla showed that one can improve this by using an input with the spectral density

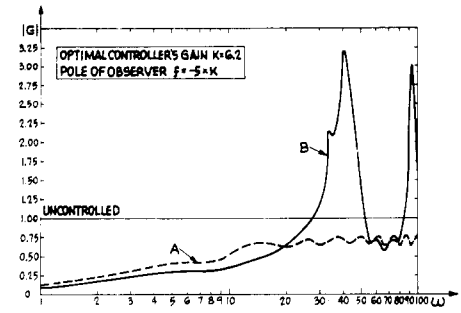

Figure 19. Magnitude of frequency response from output disturbance to process output for system using observer controller.

$$\Phi = \frac{2\nu_1}{\nu_1^2 + \omega^2} + k\frac{2\nu_2}{\nu_2^2 + \omega^2} \qquad (0 < k < 1)$$

where ν_2 is the same as used by Lueck and ν_1 is 10 times larger. No data exist on sensitivity or margin of stability. A preliminary check showed it to depend on k and the values of ν chosen. Further work is needed to develop reliable design algorithm based on this rather promising approach.

We knew from the beginning that in the example given here, the potential gains compared to simple controllers are limited. We have chosen our simple example to examine the methods, and the fact that optimal control methods fail in a trivial example teaches us much more than a complex example would.

The existence of a satisfactory controller illuminates the shortcomings of other methods and we have to conclude that the main problem in applying optimal control methods to this case is the lack of methods which take into account the inaccuracies inherent in modelling and identification procedures and guarantee good overall performance and stability over changing process conditions.

We can ask ourselves, why does the optimization procedure fail here when it has contributed so much to aerospace applications? An airplane or missile designer hopefully knows the properties of his servomechanisms. Many of his controls for which optimization methods are normally used deal with short time actions closely followed. A process plant must work for years and the process control problem is to keep it at a steady state.

Maybe adaptive methods (Bellman, 1961) will be in the future able to overcome this difficulty but again much more work needs to be done before they can be straightforwardly applied by the process engineer. The beauty of the classical method is that one can get satisfactory controllers for a wide variety of systems with a very limited knowledge of the transfer function (see Horowitz and Sidi, 1972). For more complex systems those methods do not work well. We cannot use them, for example, to stabilize an unstable reactor or dampen a limit cycle. There are many situations in which we need a better design method.

But modern control methods have not reached the stage where they offer a practical alternative; the main problem is that by their nature these algorithms do not lead to conservative design methods. What we would need to use them is experience in the formulation of an algorithm which lead to stable conservative designs. Such methods would formulate either the inputs or the constraints. Based on our experience it is not necessary that the constraints (or input) formulate a really important goal. Rather they should heuristically assume overall good performance.

The discussion in this paper has centered till now on a very simple problem: continuous controllers for over-

damped simple in-put–single out-put systems. The present challenges of process control are in other areas, multi-variable systems, direct computer controls etc. But the principles illustrated here apply to those cases as well. Let's briefly consider multi-variable control, (see Rosenbrock, 1971) which will be discussed in detail in a future paper.

One of the methods advanced is non-interacting controls (Gould, 1969). It is an inversion of the transfer function and depends again strongly on our knowledge of this transfer function. While it is not theoretically clear that a noninteracting control scheme is optimum in any sense, it suffers from sensitivity problems similar to those of the optimal control scheme discussed here.

Another similar problem in multi-variable control is that the formulation which is mathematically most convenient for optimal controller design, namely fast return to steady state, does not always lead to sensible compromises. Weekman (1974) discusses an interesting case (Kurihara, 1967) in which an optimum design led to a controller which is not usable in a plant. On the other hand it gave useful information for improving controller design. The problem here is that exclusive emphasis on fast return neglects the fact that the inputs are not at steady state and the operator must be able to fix the steady-state conditions. Measured variables which are useful for fast control are not necessarily those for which we want to maintain a given value.

Here modern control methods have some very important applications. They often give insight into what variables should be chosen for measurement and manipulation and how to connect them. Thus they provide a useful tool in heuristic design methods. Much more work is needed before they can be transplanted into reliable design algorithms that can be used in a straightforward way.

6. Discussion and Conclusion

It should be pointed out that the foregoing discussion dealt mainly with continuous controllers. Sampled data controllers with infrequent sampling involve other problems which will be discussed in a future paper.

We presented a comparison between a PID controller and different optimum control schemes. We purposely chose an extremely simple case in which the PID controller gives rather satisfactory performance. As mentioned earlier none of the optimum controllers has an overall performance better than that of the classical controller and most are definitely inferior. The main fault of all the optimum schemes that were examined above was their sensitivity to the exact transfer function of the designed system. We would expect that a proper optimum design algorithm should use the available information to the largest extent possible, and it is quite difficult to build into the algorithm the fact that it should not really "believe" our process model.

Although it is possible to include parameter uncertainties in the formulation of the optimum control problem, it is the structural sensitivity which is more difficult to include and which is so important. In most cases we design controllers for processes which are not only very imperfectly known but are also so complex that any manageable process model is only a very crude description of the process.

Optimum control methods have one important application. They give an exact limit as to what can be achieved for any given criterion. Any practical controller will be a compromise between the different criteria, and here a knowledge of what could be achieved is an excellent guideline as to how good a compromise the real controller is.

For the example given above it is quite evident that a classical PID controller is a good compromise. In more complex cases there will be much more room for improvement, but at present the best we can do is to use systematic trial and error methods.

In complex control problems some of the algorithms proposed also offer important insight as to the choice of measured and manipulated variables and how to connect them. But in the present state these are not design methods but rather helpful tools in heuristic design and here modern control methods have much to offer.

There is a considerable challenge in trying to integrate the lessons of modern control theory into efficient usable design procedures for the practicing engineer. To achieve this the academic researcher should first realize that design of process controllers is not really an optimization problem but rather a minimax problem or a compromise between a considerable number of conflicting requirements. Fulfilling all these requirements puts some strong constraints on using design. At present it is not clear that this leaves much place for optimizing one of them. There is a need for better design procedures for complex control cases and modern control techniques may provide a lead on how to achieve them. But such design procedure will need to be tied into specific modelling or identification techniques and has to take into account their inherent limitations.

Literature Cited

Aoki, M., "Optimization of Stochastic Systems," Academic Press, New York, N.Y., 1967.

Bellman, R., "Adaptive Control Processes," Princeton University Press, Princeton, N.J., 1961.

Box, G. E. P., Jenkins, G. M., "Time Series Analysis, Forecasting and Control," Holden-Day, San Francisco, Calif., 1970.

Bryson, A. E., Jr., Ho, Y. C., "Applied Optimal Control," Blaisdell, 1967.

Cegla, U., Ph.D. Dissertation, The City University of New York, 1969.

Cohen, G. H., Coon, G. A., *Trans. ASME,* (July 1953).

Denn, M., *Chem. Eng. Sci.,* **27,** 121 (1972).

Douglass, "Process Dynamics and Control," Prentice-Hall, 1972.

Foss, A. A., *IEEE Trans. Autom. Control,* 646–652 (Dec 1973).

Gould, L. A., "Chemical Process Control," Addison-Wesley, 1969.

Horowitz, I. M., "Synthesis of Feedback Systems," Academic Press, New York, N.Y., 1963.

Horowitz, I. M., Sidi, M., *Int. J. Cont.,* **16,** 287 (1972).

Hougen, J. O., *Chem. Eng. Progr. Monograph Ser. No. 4,* 60 (1964).

Johnson, C. D., *IEEE Trans. Autom. Control,* **16,** 635 (Dec 1971).

Kalman, R. E., *Trans. ASME, J. Basic Eng.,* **82,** 34 (1960).

Kalman, R. E., "On the General Theory of Control Systems," Proceedings of the First International Congress on Automatic Control, V1, p 481, Butterworth, London, 1961.

Koppel, L. B., "Introduction to Control Theory," Prentice-Hall, 1968.

Kreisselmeir, G., Grubel, G., "The Design of Optimally Parameter Insensitive Control Systems," Proceedings of the IFAC 5th World Congress, Paris, Paper 31.1, 1972.

Kurihara, H., PH.D. Thesis, M.I.T., 1967.

Lapidus, L., Luus, R., "Optimal Control of Engineering Processes," Blaisdell 1967.

Lees, S., Hougen, J. O., *Ind. Eng. Chem.,* **48,** 1064 (1956).

Lee, W., Weekman, V. W., "Advanced Control Practice in the Process Industry: A View from Industry," Plenary Lecture at the 1974 JACC, Austin, Texas, 1974.

Lim, H. C., Bankhoff, S. G., *AIChE J.,* **16,** 233 (1970).

Lueck, R. H., McGuire, N. L., *AIChE J.,* **14,** 173, 181 (1968).

Luenberger, D. G., *IEEE Trans. Autom. COntrol,* **11,** 190 (April 1960).

Luenberger, D. G., *IEEE Trans. Autom. Control,* **16,** 596 (Dec 1971).

Missaghic-Mamaghnic, M., Fairman, F. W., *Int. J. Syst. Sci.,* **4,** 859–864 (1973).

Murrill, P. W., "Automatic Control of Processes," International Textbook Co., 1967.

Newton, G. C., Jr., et al., "Analytical Design of Linear Feedback Controls," Wiley, New York, N.Y., 1957.

O'Connor, G. E., Denn, M. M., *Chem. Eng. Sci.,* **27,** 121 (1972).

Rosenbrock, H. M., "On the Design of Linear Multivariable Control Systems," Trans. IFAC, London 1966; also Rosenbrock and McMorran, IFEE Trans. on Automatic Control, Vol. AC 16, No. 552 (1971).

Seinfeld, J. H., Lapidus, L., "Process Modelling Estimation and Identification," Prentice-Hall, 1974.

Thau, F. E., Kestenbaum, A., "The Effect of Modelling Errors on Linear State Reconstruction and Regulators," presented at ASME Winter Annual Meeting, 1973.

"Priorities in Process Control Research," Report of National Science Foundation Workshop, Tulane University, Mar 11, 1973.

Wiener, N., "The Extrapolation Interpolation and Smoothing of Stationary Time Series," M.I.T. Press, 1949.

Received for review June 27, 1974
Accepted August 14, 1975