Internal Model Control. 4. PID Controller Design

Daniel E. Rivera, Manfred Morari,* and Sigurd Skogestad

Chemical Engineering, 206-41, California Institute of Technology, Pasadena, California 91125

For a large number of single input-single output (SISO) models typically used in the process industries, the Internal Model Control (IMC) design procedure is shown to lead to PID controllers, occasionally augmented with a first-order lag. These PID controllers have as their only tuning parameter the closed-loop time constant or, equivalently, the closed-loop bandwidth. On-line adjustments are therefore much simpler than for general PID controllers. As a special case, PI- and PID-tuning rules for systems modeled by a first-order lag with dead time are derived analytically. The superiority of these rules in terms of both closed-loop performance and robustness is demonstrated.

I. Introduction

Synthesis and tuning of control structures for SISO systems comprises the bulk of process control problems. In the past, hardware considerations dictated the use of the PID controller, but through the use of computers, controllers have now advanced to the stage where virtually any conceivable control policy can be implemented. Despite these advances, the most widely used controller is still of the PID type. Finding design methods which lead to the optimal operation of PID controllers is therefore of significant interest.

For controller tuning, simplicity, as well as optimality, is important. The three modes of the ordinary PID controller, k_c, τ_I, and τ_D, do not readily translate into the desired performance and robustness characteristics which the control system designer has in mind. The presence of simple rules which relate model parameters and/or experimental data to controller parameters serves to simplify the task of the designer.

The literature contains a number of these "tuning rules"; possibly the best known are the Ziegler-Nichols rules proposed in 1942. Given the wide use of the first-order lag/dead time model for chemical processes, tuning rules for PID control of this structure have received wide attention in the literature. Most common are the rules proposed by Cohen and Coon 1953). Smith (1972) contains a good summary of efforts in this area.

Our intention is to present a clearer and more logical framework for PID controller design which is simple to understand and implement while possessing a sound fundamental basis. Instead of fixing a control structure and then attempting to "extract" optimality from this controller (as is usually the case with classical methods), our approach will be to postulate a model, state desirable control objectives, and, from these, proceed in a straightforward manner to obtain both the appropriate controller structure and parameters.

The Internal Model Control (IMC) structure provides a suitable framework for satisfying these objectives. IMC was introduced by Garcia and Morari (1982), but a similar concept has been used previously and independently by a number of other researchers. Using the IMC design procedure, controller complexity depends exclusively on two factors: the complexity of the model and the performance requirements stated by the designer. The goal of this article is to show that for the objectives and simple models common to chemical process control, the IMC design procedure leads naturally to PID-type controllers, occasionally augmented by a first-order lag. Furthermore, the proposed procedure provides valuable insight regarding controller tuning effects on both performance and robustness.

II. Performance and Robustness Measures

Probably the best indicator of performance is the *sensitivity function*

$$S = \frac{1}{1 + gc} = \frac{e}{y_s - d} \tag{1}$$

(The nomenclature should be apparent from Figure 1.) It is desirable to keep the sensitivity function small over as wide a frequency range as possible. For any proper system, $|S|$ will approach unity as the frequency becomes large. Instead of the sensitivity function, the closed-loop *bandwidth* can be used as a simple performance measure; it is the frequency ω_b at which $|S|$ first reaches $1/2^{1/2}$

$$|S| < \frac{1}{2^{1/2}} \quad \forall\ \omega < \omega_b \tag{2}$$

Increasing the bandwidth implies less attenuation of the reference signal, more effective disturbance rejection, and a faster response. For a phase margin (PM) less than or equal to $\pi/2$ (the most common situation), the bandwidth is less than or equal to the (gain) crossover frequency ω_c, defined as the frequency at which the open-loop gain first drops to unity

$$|gc| > 1 \quad \forall\ \omega < \omega_c \tag{3}$$

Occasionally, we will also refer to the Integral Square Error (ISE) and to the Integral Absolute Error (IAE) for a specified set point or disturbance change to compare the performance of different controllers:

$$J = \text{ISE} = \int_0^\infty (y - y_s)^2 \, dt \tag{4}$$

$$J' = \text{IAE} = \int_0^\infty |y - y_s| \, dt \tag{5}$$

It is crucial in control system design to ensure the stability and performance of the closed-loop system in the presence of plant/model mismatch, i.e., to guarantee *robustness*. We will use a superscript $(\tilde{\ })$ to distinguish the (known) model and its properties from the (generally unknown) "real" plant. Plant/model mismatch can be caused, for example, by model reduction (the representation of a high-order system by a low-order approximate model) or by system parameters which depend on the operating conditions. Though we do not know the real plant g, it is often reasonable to assume it to be a member

* To whom all correspondence should be addressed.

0196-4305/86/1125-0252$01.50/0 © 1985 American Chemical Society

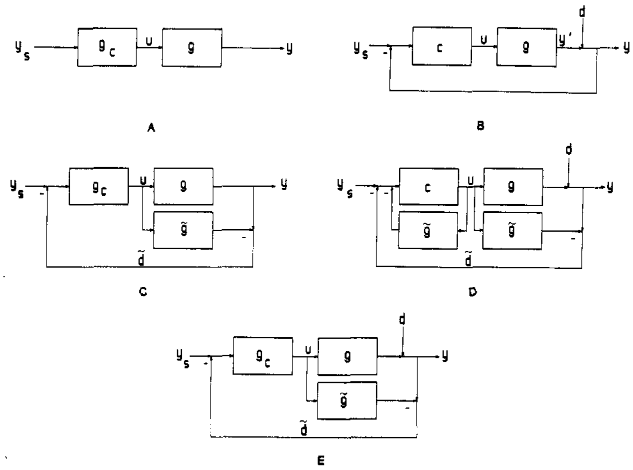

Figure 1. Evolution of the IMC structure.

of a family Π of linear plants defined by a norm-bounded multiplicative error e_m

$$\Pi = \{g: |e_m| \leq l_m\} \tag{6}$$

where

$$e_m = \frac{g - \tilde{g}}{\tilde{g}} \tag{7}$$

Usually $|e_m|$ approaches a value equal to or greater than 1 for high frequencies.

We will also establish in the following that the *complementary sensitivity function* $\tilde{H}$

$$\tilde{H} = \frac{\tilde{g}c}{1 + \tilde{g}c} \tag{8}$$

is a good robustness measure. The name "complementary sensitivity" follows from the equality

$$\tilde{H} + \tilde{S} = 1 \tag{9}$$

Let us assume that g, $\tilde{g}$, and c have no poles in the open right-half plane (RHP) and that the closed-loop system with the "nominal" plant $\tilde{g}$ and the controller c is stable. Then Doyle and Stein (1981) have shown that the closed-loop system is stable for all plants in the family Π if and only if

$$|\tilde{H}| < \frac{1}{l_m} \quad \forall \omega \tag{10}$$

Because l_m increases with frequency and eventually exceeds 1, $|\tilde{H}|$ has to drop below 1 at some frequency. Because of (9), $|\tilde{S}|$ has to be close to 1 in this frequency range. Thus, the achievable closed-loop bandwidth is limited by the bandwidth over which the process model is good. The smallest uncertainty $l_m(\omega)$ is allowed at the frequency where $|\tilde{H}(j\omega)|$ has its maximum peak. As a consequence, the M value defined by (11) (e.g.: Rosenbrock, 1974) is a suitable robustness indicator.

$$M = \max_{\omega} |\tilde{H}| \tag{11}$$

M is convenient and widely accepted as more useful than gain margins (GM) or phase margins (PM). Gain and phase margins only measure robustness with respect to model uncertainties which are independent of ω and thus tend to be overly optimistic. The following relationships indicate how M establishes lower bound on GM and PM:

$$\text{GM} \geq 1 + \frac{1}{M} \tag{12}$$

$$\text{PM} \geq 2 \sin^{-1} \left(\frac{1}{2M}\right) \cong \frac{1}{M} \tag{13}$$

For the special case of $M = 1$, (12) and (13) become

$$\text{GM} \geq 2 \tag{14}$$

$$\text{PM} \geq 60° \tag{15}$$

One must note that M by itself yields only a qualitative indication of robustness. The allowable uncertainty in specific model parameters can be deduced from M only when the bandwidth ω_b is known. Consider, for example, an analysis of the allowable dead-time error in a closed-loop system (the plant dead time exceeds that of the model by the quantity δ):

$$g = \tilde{g}e^{-s\delta} \tag{16}$$

Because the dead-time error introduces a phase lag of $\omega\delta$ at frequency ω, the system will remain stable for the dead-time error δ if

$$\delta < \frac{\text{PM}}{\omega_c} \tag{17}$$

When (13) is substituted, this exact condition can be replaced by the more conservative condition

$$\delta < \frac{1}{\omega_c M} \tag{18}$$

For PM = 90°, $\omega_b = \omega_c$ and (18) becomes

$$\delta < \frac{1}{\omega_b M} \tag{19}$$

Equation 19 clearly illustrates the trade-off between performance and robustness. Good performance (high ω_b) is obtained only at the expense of robustness (small allowed dead-time error).

Our study is aimed at systems of "type 1" and "type 2" (Wiberg, 1971):

$$\text{type 1:} \ \lim_{s \to 0} sgc \neq 0 \tag{20}$$

$$\text{type 2:} \ \lim_{s \to 0} s^2gc \neq 0 \tag{21}$$

Type 1 and type 2 systems exhibit no offset to step and ramp changes on $(y_s - d)$, respectively. Furthermore, the following limits hold:

$$\lim_{s \to 0} H(s) = 1 \tag{22}$$

$$\lim_{s \to 0} S(s) = 0 \tag{23}$$

III. Internal Model Control (IMC)

III.1. Fundamentals. The goal of control system design is fast and accurate set-point tracking

$$y \cong y_s \quad \forall\ t, \quad \forall\ d \tag{24}$$

This implies that the effect of external disturbances should be corrected as efficiently as possible (good regulatory behavior)

$$y' \cong y_s - d \quad \forall\ t, \quad \forall\ d \tag{25}$$

Furthermore, the control system designer wishes to obtain (24) and (25), while also being assured of insensitivity to modeling error.

It is well-known that an open-loop (feedforward) arrangement (Figure 1A) represents the optimal way to satisfy (24). For the open-loop scheme, the stability question is trivial (the system is stable when both the controller and the system are stable); also the controller is easy to design ($g_c = \tilde{g}^{-1}$). The disadvantages are the sensitivity of the performance to plant/model mismatch and the inability to cope with unmeasured disturbances.

With the feedback arrangement (Figure 1B), the situation is reversed. Plant/model mismatch and unmeasured disturbances can be dealt with effectively, but tuning is complicated by the closed-loop stability problem.

We can now augment the open-loop and closed-loop systems as indicated in Figure 1C and 1D without affecting performance: In Figure 1C, $\tilde{d} = 0$, and therefore the system is still open-loop; in Figure 1D, the two blocks $\tilde{g}$ cancel each other. Relating Figure 1C and 1D through the definitions

$$g_c = \frac{c}{1 + c\tilde{g}} \tag{26}$$

$$c = \frac{g_c}{1 - \tilde{g}g_c} \tag{27}$$

we arrive at the general structure in Figure 1E which has the advantages of both the open-loop and closed-loop structures: When the model of the plant is perfect ($g = \tilde{g}$) and there are no disturbances ($d = 0$), feedback is not needed and structure E behaves identically to structure A. Because the plant model $\tilde{g}$ appears explicitly in E, this structure is referred to as the Internal Model Control (IMC) structure. As a simplification, we can say that the controller in E can be designed with the ease of an open-loop controller while retaining the benefits of a feedback system. It is our goal to describe, in detail, such a design procedure.

From the block diagram for the IMC structure (Figure 1E), follow the relationships

$$u = \frac{g_c}{1 + g_c(g - \tilde{g})}(y_s - d) \tag{28}$$

$$y = \frac{gg_c}{1 + g_c(g - \tilde{g})}(y_s - d) + d = H(y_s - d) + d \tag{29}$$

$$e = y_s - d = \frac{1 - \tilde{g}g_c}{1 + g_c(g - \tilde{g})}(y_s - d) = S(y_s - d) \tag{30}$$

Four properties can be shown which suggest the advantages of this structure.

P1: Dual Stability. Assume $g = \tilde{g}$. Then the system is effectively open-loop and "closed-loop stability" is implied by the stability of g and g_c:

$$y = gg_c(y_s - d) + d \tag{31}$$

While for the classical structure (Figure 1B) it is not at all clear what type of controller c and what parameter choices lead to closed-loop stable systems, the IMC structure guarantees closed-loop stability for all stable controllers g_c.

P2: Perfect Control. Assume that the controller is equal to the model inverse ($g_c = \tilde{g}^{-1}$) and that the closed-loop system in Figure 1E is stable. Then $y = y_s$ for all $t > 0$ and all disturbances d(t).

P3: Type-1 System. Assume that the controller steady-state gain is equal to the inverse of the model gain

$$g_c(0) = \tilde{g}(0)^{-1} \tag{32}$$

and that the closed-loop system in Figure 1E is stable. Then the system is of type 1 and the control error vanishes asymptotically for all asymptotically constant inputs y_s and d. This property implies no offset at steady state and follows from (30) via the final value theorem.

P4: Type-2 System. Select g_c to staisfy P3 and

$$\frac{\mathrm{d}}{\mathrm{d}s}(\tilde{g}g_c)|_{s=0} = 0 \tag{33}$$

Then the system is of type 2 and the control error vanishes asymptotically for all asymptotically ramp-shaped inputs y_s and d (Brosilow, 1983). (P4 also follows from (30) via the final value theorem.)

P1 simply expresses the fact that in the absence of plant/model mismatch, the stability issue is trivial, as long as the open-loop system is stable. P2 asserts that the ideal open-loop controller leads to perfect closed-loop performance when the IMC structure is employed. P3 and P4 state that inherent integral action can be achieved without the need for introducing additional tuning parameters. P2, however, represents an idealized situation. We know intuitively that P2 requires an infinite controller gain; this is confirmed by substituting $g_c = \tilde{g}^{-1}$ in (27). By setting $g_c(0) = \tilde{g}(0)^{-1}$ as postulated for P3, we find $c(0) = \infty$, which implies integral control action, as expected.

There are several reasons why the "perfect controller" implied by P2 cannot be realized in practice.

1. **Right-Half Plane (RHP) Zeros**: If the model has a RHP zero, the controller $g_c = \tilde{g}^{-1}$ has a RHP pole, and if $\tilde{g} = g$, the closed-loop system will be unstable according to P1.

2. **Time Delay.** If the model contains a time delay, the controller $g_c = \tilde{g}^{-1}$ is predictive and cannot be realized.

3. **Constraints on the Manipulated Variables.** If the model is strictly proper, then the perfect controller $g_c = \tilde{g}^{-1}$ is improper, which implies $\lim_{\omega\to\infty} |g_c| = \infty$. Thus, infinitely small high-frequency disturbances would give rise to infinitely large excursions of the manipulative variables which are physically unrealizable.

4. **Modeling Error.** If $g \neq \tilde{g}$, P1 does not hold and the closed-loop system will generally be unstable for the controller $g_c = \tilde{g}^{-1}$.

In resolving these four issues, the ideal of perfect control must be abandoned. The IMC design procedure handles this in two steps; first, performance is addressed with no regard to robustness or input constraints. Second, a filter is introduced and designed for properness (input constraints) and robustness without looking at how this affects the performance. Though there obviously does not exist any separation principle which makes this approach "optimal", the design procedure is very simple and direct. Also, there seem to be very few cases where other more complicated and indirect procedures (e.g., LQG) give better results. The freedom which the designer is given to choose the filter makes it possible to take into account considerations which may be difficult to pin down mathematically.

Step 1. Factor the model

$$\tilde{g} = \tilde{g}_+\tilde{g}_- \tag{34}$$

such that $\tilde{g}_+$ contains all the time delays and RHP zeros; consequently $\tilde{g}_-^{-1}$ is stable and does not involve predictors.

Step 2. Define the IMC controller by

$$g_c = \tilde{g}_-^{-1}f \tag{35}$$

where f, a low-pass filter, must be selected such that g_c is proper or, if "derivative" action is allowed (as in the ideal PID controller), such that g_c has a zero excess of at most 1. By definition of the factorization in (34), g_c is realizable.

Having introduced these definitions, the closed-loop relationships (29) and (30) become

$$y = \frac{\tilde{g}_+f(1 + e_m)}{1 + \tilde{g}_+fe_m}(y_s - d) + d = \frac{\tilde{H}(1 + e_m)}{1 + \tilde{H}e_m}(y_s - d) + d \tag{36}$$

$$e = y_s - y = \frac{1 - \tilde{g}_+f}{1 + \tilde{g}_+fe_m}(y_s - d) = \frac{1 - \tilde{H}}{1 + \tilde{H}e_m}(y_s - d) \tag{37}$$

For the special case of a perfect model ($e_m = 0$), (36) and (37) reduce to

$$y = \tilde{g}_+ f(y_s - d) + d = \tilde{H}(y_s - d) + d \qquad (38)$$

$$e = (1 - \tilde{g}_+ f)(y_s - d) = \tilde{S}(y_s - d) \qquad (39)$$

Equations 38 and 39 demonstrate clearly that for the case of no plant/model mismatch, the nominal closed-loop transfer function $\tilde{H} = \tilde{g}_+ f$ is at the designer's discretion except that (1) $\tilde{g}_+$ must contain all the delays and RHP zeros and (2) f must be of sufficiently high order to avoid physically unrealizable control action. Thus, the closed-loop transfer function can be designed directly and not ambiguously via c as in the classic controller design procedure (Figure 1B).

Our treatment is not complete without indicating how to select $\tilde{g}_+$ and f.

III.2. Factorization of $\tilde{g}$. Assume $\tilde{g} = g$. For step inputs in y_s and d, selecting $\tilde{g}_+$ and f such that $|\tilde{g}_+ f| = 1$ $\forall\ \omega$ minimizes the ISE (Holt and Morari, 1984, 1985). This implies that f must be unity and that $\tilde{g}_+$ has the form of an all-pass

$$\tilde{g}_+ = e^{-\theta s}\prod_i \frac{-\beta_i s + 1}{\beta_i s + 1} \qquad \text{Re}\ (\beta_i) > 0 \qquad (40)$$

where β_i^{-1} are all the RHP zeros and θ is the time delay present in $\tilde{g}$. As a consequence of this factorization, poles corresponding to the LHP image of the RHP zeroes have been added to the closed-loop response.

For step inputs in y_s and d, selecting f to be unity and $\tilde{g}_+$ as

$$\tilde{g}_+ = e^{-\theta s}\prod_i\ (-\beta_i s + 1) \qquad \text{Re}\ (\beta_i) > 0 \qquad (41)$$

minimizes the IAE (Holt and Morari, 1984, 1985).

When $\tilde{g}$ is a minimum-phase model, $\tilde{g}_+ = 1$.

III.3. Filter Selection. In order to satisfy P3 (zero offset to step inputs), we adopt the following convention for $\tilde{g}_+(s)$ and $f(s) = p(s)/q(s)$

$$\tilde{g}_+(0) = p(0) = q(0) = 1 \qquad (42)$$

The simplest filter f satisfying (42) is of the form

$$f(s) = \frac{1}{(\epsilon s + 1)^r} \qquad (43)$$

where r is sufficiently large to guarantee that the IMC controller g_c is proper. If $g = \tilde{g}$ and $\tilde{g}_+ = 1$ (i.e., the model is minimum-phase), then $y/y_s = \tilde{H} = f$. The parameter ϵ, which can be adjusted by the operator, determines the speed of the response. For a minimum-phase system, the bandwidth is proportional to $1/\epsilon$

$$\omega_b = \omega_c = 1/\epsilon \text{ for } r = 1 \qquad (44)$$

$$1/\epsilon > \omega_b > 1/r\epsilon > \omega_c \text{ for } r > 1 \qquad (45)$$

For nonminimum-phase systems, the achievable bandwidth is inherently limited by the plant. For example, consider the following representative factorizations with $\epsilon = 0$

$$\tilde{g}_+ = e^{-s\theta}: \quad \omega_c = \frac{\pi}{3\theta} \quad \omega_b = \frac{0.724}{\theta} = 0.69\omega_c \qquad (46)$$

$$\tilde{g}_+ = (-\beta s + 1)/(\beta s + 1): \quad \omega_c = \frac{1}{\beta(3)^{1/2}} \quad \omega_b = \frac{1}{\beta(7)^{1/2}} = 0.65\omega_c \qquad (47)$$

$$\tilde{g}_+ = -\beta s + 1: \quad \omega_c = \infty \quad \omega_b = \frac{1}{\beta(2)^{1/2}} \qquad (48)$$

For $\epsilon > 0$, ω_c and ω_b decrease from the bounds established through (46)–(48). For $r = 1$, exact formulas for the bandwidth and crossover are included in Appendix A; these are effectively approximated by

$$\tilde{g}_+ f = (e^{-s\theta})/(\epsilon s + 1): \quad \omega_c \cong \omega_b \cong \frac{1}{\theta + \epsilon} \qquad (49)$$

$$\tilde{g}_+ f = [(-\beta s + 1)/(\beta s + 1)][1/(\epsilon s + 1)]: \quad \omega_c \cong \omega_b \cong \frac{1}{2\beta + \epsilon} \qquad (50)$$

$$\tilde{g}_+ f = (-\beta s + 1)/(\epsilon s + 1): \quad \omega_c \cong \omega_b \cong \frac{1}{\beta + \epsilon} \qquad (51)$$

One notices from these expressions that until $1/\epsilon$ is of an order of magnitude comparable to θ or β, respectively, ω_c and ω_b are virtually unaffected by the presence of the filter. Thus, making ϵ very small for nonminimum-phase systems has little effect on the bandwidth and performance but is very detrimental to the robustness, as we will see later. For ϵ large compared to θ or β approximately, the same proportionality holds as for MP systems (eq 44).

ϵ is directly related to important closed-loop characteristics, unlike the parameters available in the general lead/lag network c of the classical structure (e.g., PID controllers). The larger ϵ is the slower the response and the smaller the actions of the manipulated variable. With (43), the maximum peak for $|f|$ is 1; i.e., the robustness characteristics are good.

For $r > 1$, filter forms other than (43) can lead to faster response. For example, for $r = 2$, the filter

$$f = \frac{1}{\epsilon^2 s^2 + 2\zeta\epsilon s + 1} \qquad (52)$$

with damping factor $\zeta = 0.5$ minimizes the ISE (Frank, 1974). However, with this filter $|f|_{\max} = 1.15$; thus, performance improvement occurs at the expense of a reduced robustness margin. In practice, choosing filters with structures more general than (43) is usually not worthwhile.

Additional conditions on f are necessary in order to satisfy P4 (zero offset to ramp inputs). With the adopted conventions (42), (33) becomes

$$\tilde{g}_+'(0) = q'(0) - p'(0) \qquad (53)$$

where the prime denotes differentiation with respect to s. An example of a filter satisfying (53) is

$$f = \frac{(2\epsilon - \tilde{g}_+'(0))s + 1}{(\epsilon s + 1)^2} \qquad (54)$$

where, as before, the adjustable parameter ϵ is, for minimum-phase systems, the closed-loop time constant and $1/\epsilon$ is proportional to the closed-loop bandwidth. Values of $\tilde{g}_+'(0)$ for representative factorizations are

$$\frac{\mathrm{d}}{\mathrm{d}s}\,(e^{-s\theta})|_{s=0} = -\theta \qquad (55)$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\,(-\beta s + 1)|_{s=0} = -\beta \qquad (56)$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{-\beta s + 1}{\beta s + 1}\right)\bigg|_{s=0} = -2\beta \qquad (57)$$

Because in general $\tilde{g}_+'(0) < 0$, one obtains

$$M = \max_{\omega}\ |\tilde{H}| > 1 \qquad (58)$$

i.e., M is strictly greater than unity for all filters satisfying (53). Again, the tighter performance specification (no offset for ramps) is paid for with decreased robustness margins.

III.4. Accounting for Modeling Error. Thus far, all the discussion on filter selection has assumed a perfect model, in which case ϵ can be selected freely; this is not the case in practice where plant/model mismatch exists. It follows from (10) that for robust stability of the closed-loop system

$$|\tilde{g}_+ f| < \frac{1}{l_{\mathrm{m}}} \quad \forall\ \omega \tag{59}$$

Assuming for simplicity $|\tilde{g}_+| = 1$, it becomes clear that the filter magnitude $|f|$ must be small wherever the plant/model mismatch e_{m} is large. Because l_{m} approaches or exceeds 1 for high frequencies in all practical situations, we find again that *the allowable range for ϵ is limited by the degree of plant/model mismatch.* As stated previously, the closed-loop bandwidth can never be larger than the bandwidth over which the process model is valid. The models used in process control are usually good enough to set $1/\epsilon$ at least equal to the open-loop bandwidth.

In the presence of plant/model mismatch, the structure of f fails to automatically guarantee the shape of the response. However, for the suggested IMC design procedure ($\tilde{g}_+(0)f(0) = 1$ and $f = (\epsilon s + 1)^{-r}$) and using the Triangle Inequality

$$|e| \le \frac{|1 - \tilde{g}_+ f|}{1 - |\tilde{g}_+ f e_{\mathrm{m}}|}\, |y_{\mathrm{s}} - d| \tag{60}$$

one can discern general frequency intervals for which (60) and the ideal error function

$$|e| = |1 - \tilde{g}_+ f||y_{\mathrm{s}} - d| = |\tilde{S}||y_{\mathrm{s}} - d| \tag{61}$$

are very similar. At low frequency ($\omega \ll 1/\epsilon$), $\tilde{g}_+ f \cong 1$ and $e \cong 0$. For $\omega \gg 1/\epsilon$, $|f|$ is exceedingly small, $|\tilde{g}_+ f e_{\mathrm{m}}| \cong 0$, and (60) and (61) become close to each other. For $\omega \cong 1/\epsilon$, the situation is uncertain. We conclude that for ϵ sufficiently large, the closed-loop response to high-frequency or low-frequency inputs (e.g., steps) will become similar to the response of the nominal system $\tilde{g}_+ f$.

In summary, the key advantage of the IMC design procedure is that all controller parameters are related in a unique, straightforward manner to the model parameters. There is only one adjustable parameter ϵ which has intuitive appeal because it determines the speed of response of the system. Furthermore, ϵ is approximately proportional to the closed-loop bandwidth which must always be smaller than the bandwidth over which the process model is valid. This leads to a good initial estimate of ϵ, which can be adjusted on-line if necessary.

IV. IMC in the Context of Classical Control

For linear systems, the IMC controller g_{c} represents an alternate parametrization of the classic controller c, albeit with very useful properties. Through the transformation

$$c = \frac{g_{\mathrm{c}}}{1 - \tilde{g} g_{\mathrm{c}}} = \frac{\tilde{g}_-^{-1}}{f^{-1} - \tilde{g}_+} \tag{62}$$

Figure 1B becomes equivalent with 1E. If there is no delay in g, c is rational and can be implemented as a lead/lag network. Indeed, for minimum-phase systems ($\tilde{g}_+ = 1$) and a first-order filter ($f = \epsilon s + 1)^{-1}$), c becomes

$$c = \frac{1}{\epsilon}\frac{\tilde{g}^{-1}}{s} \tag{63}$$

IV.1. IMC Implemented as a PID Controller (Table I). Naturally, one would expect that for certain process models, the lead/lag network c obtained from (62) via the IMC design procedure is a PID controller. Indeed, we find that IMC leads to PID controllers for virtually all models common in industrial practice (Table I). Note that Table I includes systems with pure integrators and RHP zeros. Occasionally, the PID controllers are augmented by a first-order lag with time constant τ_{F}. A few remarks regarding Table I ar appropriate:

Remark 1. When the PID controller of the specified form is applied to the model $\tilde{g}$, the closed-loop system is stable for all values of $\epsilon > 0$.

Remark 2. For about one-third of the cases studied, ϵ appears only in the expression for the controller gain k_{c}. For cases A–C, H, and J, the controller gain is inversely proportional to ϵ, thus demonstrating that on-line PID controller adjustment is effectively achieved by simply manipulating k_{c}. These are minimum-phase models, for which $\omega_{\mathrm{b}} = \omega_{\mathrm{c}} = 1/\epsilon$, and the model itself imposes no limitations on the bandwidth. For cases D, F, L, and P, the controller gain k_{c} is the only parameter dependent on ϵ, but because of the presence of a RHP zero, there is a maximum gain which cannot be surpassed no matter how small ϵ is. Cases D, F, L, and P correspond to systems factored according to (41), in which case ω_{b} is limited by approximately $1/(\beta + \epsilon)$ (recall eq 51); the IMC design procedure recognizes naturally that increasing the gain beyond a certain value leads to performance deterioration and eventually stability.

For a significant number of the considered models, ϵ appears in *all* the parameters of the classic feedback controller, e.g., K and R. It is not surprising then that for such processes, trial and error tuning of PID controllers is notoriously difficult. However, the IMC parametrization shows how all the controller parameters may be adjusted *simultaneously* in an effective manner.

Remark 3. In all cases, there is no offset for set-point and/or disturbance step changes. If the process has an integrator, a step disturbance entering through the integrator becomes a ramp, thus requiring that there should be no offset for ramp changes. This performance specification is met in cases I, K, N, O, R, and S by selecting the filter f to be of the form shown in (54).

Remark 4. For systems with RHP zeros, two options for choosing $\tilde{g}_+ f$ are available.

Option 1. $\tilde{g}_+ f$ follows (40):

$$\tilde{g}_+ f = \left(\frac{-\beta s + 1}{\beta s + 1}\right)\frac{1}{\epsilon s + 1} \tag{64}$$

(64) is optimal in the ISE sense to step changes when $\epsilon = 0$ (Holt and Morari, 1985). For no offset to ramps, f has to be selected in accordance with (54):

$$\tilde{g}_+ f = \left(\frac{-\beta s + 1}{\beta s + 1}\right)\frac{(2(\beta + \epsilon)s + 1)}{(\epsilon s + 1)^2} \tag{65}$$

(64) and (65) require augmenting the PID controller with a lag term ($\tau_{\mathrm{F}} s + 1$). The filter constant ϵ may, in principle, be chosen freely. However, as already pointed out in (47), a single RHP zero factored according to (40) limits the bandwidth to $1/\beta(7)^{1/2}$. Therefore, selecting $\epsilon \ll \beta(7)^{1/2}$ has very little effect on the response. We recommend that to improve robustness, $\epsilon > \beta/2$.

Option 2. $\tilde{g}_+ f$ follows (41)

$$\tilde{g}_+ f = \frac{-\beta s + 1}{\epsilon s + 1} \tag{66}$$

For step inputs, (66) is IAE optimal when $\epsilon = 0$ and ISE optimal when $\epsilon = \beta$ (Holt and Morari, 1985). For no offset to ramps,

$$\tilde{g}_+ f = \frac{(-\beta s + 1)((\beta + 2\epsilon)s + 1)}{(\epsilon s + 1)^2} \tag{67}$$

Option 2 gives a simpler controller and is favorable for situations where $\epsilon > \beta$ is acceptable. It results in a PID controller without the need for an additional lag (as shown in cases D, F, L. N, P and R). However, noting that the closed-loop transfer function is not strictly proper, one must require that

$$\lim_{\omega\to\infty} |\tilde{g}_+ f| < 1 \qquad (68)$$

or (59) will be violated for high frequencies (where $l_m \geq 1$) and instability is bound to occur in all practical situations. This explains why $\epsilon > \beta$ is required for D, F, L, N, P, and R. The effect of this practical recommendation is that the RHP zero is pushed outside the bandwidth of the closed-loop system.

In practice, there exists no ideal PID controller as required in option 2. An additional lag is always present in the controller to provide roll-off at high frequencies. Option 1 suggests a "practical" PID controller with an "optimal" roll-off element $(\tau_F s + 1)^{-1}$.

Remark 5. No systems with LHP zeros are listed in Table I. As seen from (62), LHP zeros translate into lags in the feedback controller structure when the IMC design procedure is used. Therefore, for models with LHP zeros, the PID controller from Table I should be augmented with the corresponding lags.

Remark 6. Controller complexity, as stated in the Introduction, depends on the model and the control system objectives. Consider the cases H (a pure integrator), A (a first-order model) and B (a second-order noninteracting model), for which the desired closed-loop response is that of a first-order lag. Only a proportional controller is necessary for H; a PI controller must be used for A, while a PID controller is needed for B. Likewise, consider cases P–S, where the process model is the same: as the demands on the control system increase (as in requiring no offset to ramp changes), so does the complexity of the controller.

Remark 7. Table I can also be used for systems with delays by approximating the dead time with a Padé element; the entry for the rational approximate model then provides the controller parameters. This procedure is illustrated with two examples.

Example 1.

$$g(s) = \frac{1 - k_1 e^{-\theta s}}{s} \qquad (69)$$

Using a first-order Padé approximation

$$\tilde{g}(s) = \frac{1 - k_1 + \dfrac{\theta}{2}(1 + k_1)s}{s\left(\dfrac{\theta}{2}s + 1\right)} \qquad (70)$$

If $k_1 > 1$, then (70) has a RHP zero, and a controller from entries P–S can be selected. If $k_1 < 1$, the resulting LHP zero should be removed by a simple lag, as explained in remark 5. PID parameters can then be obtained from entries J or K.

Example 2.

$$g(s) = \frac{k e^{-\theta s}}{\tau s + 1} \qquad (71)$$

A "zeroth-order" Padé approximation ($e^{-\theta s} \cong 1$) yields

$$\tilde{g}(s) = \frac{k}{\tau s + 1} \qquad (72)$$

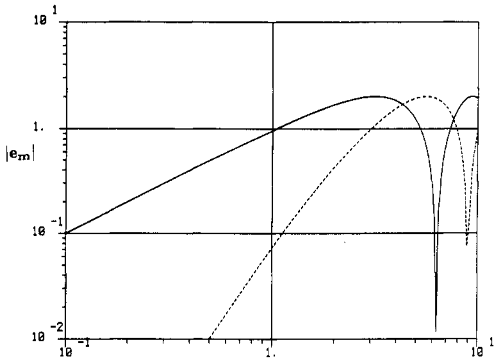

Figure 2. Multiplicative uncertainty $|e_m|$ for the zeroth (—) and first-order (···) Padé approximation. Zeroth-order: $|e_m| = |1 - e^{-\theta s}|$. First-order: $|e_m| = |[1 - (\theta/2)s/(1 + (\theta/2)s)] - e^{-\theta s}|$. $|e_m|$ for the zeroth-order approximation is equal to $|S|$ for $H = e^{-\theta s}$.

Entry A in Table I provides a PI controller for this structure. The "zeroth-order" Padé approximation is equivalent to designing a controller with no information on the dead time.

A first-order Padé approximation yields

$$\tilde{g}(s) = \frac{k\left(-\dfrac{\theta}{2}s + 1\right)}{(\tau s + 1)\left(\dfrac{\theta}{2}s + 1\right)} \qquad (73)$$

Entries F and G (PID controller and PID controller with first-order lag) are applicable to this problem. This problem is discussed in more detail in section V.

IV.2. Effects of the Padé Approximation. Examples 1 and 2 are interesting because they indicate circumstances under which three-term lead–lag controllers can be used to control processes with dead time. The use of the Padé approximation, however, introduces modeling error, which consequently limits the achievable bandwidth ω_b and the minimum value for ϵ. From (59), one obtains a good guess on the smallest value of ϵ which still maintains a stable control system.

$|e_m|$ for the zeroth- and first-order Padé approximations are shown in Figure 2. For the first-order Padé approximation, $|e_m| = 1$ at $\omega \cong 3/\theta$, and thus a sufficient condition for stability is to choose $\epsilon > \theta/3$. For the zeroth-order approximation, $|e_m| = 1$ at $\omega \cong 1/\theta$ and therefore $\epsilon > \theta$ is required. Because ω_b is inherently limited by $0.724/\theta$ (recall (46)), one can expect that using the first-order Padé approximation will yield designs very close to optimal (i.e., if no approximation were present). The zeroth-order approximation will be adequate, however, when small bandwidths and low-frequency inputs are involved.

V. IMC-Based PID Control for a First-Order Lag with Dead Time

The important role of the first-order lag/dead-time model (71) in process control mandates a more detailed discussion of example 2. Our attention is directed to a further understanding of the PI and PID rules generated by cases A and F; the advantages of the augmented PID controller (case G) are also indicated.

V.1. Tuning Procedures. The IMC-based controllers obtained by using first- and zeroth-order Padé approximations for the time delay are (cases A and F in Table I)

$$\text{PID} \qquad c = \frac{(1 + \tau s)\left(1 + \dfrac{\theta}{2}s\right)}{k\left(\dfrac{\theta}{2} + \epsilon\right)s} \qquad (74)$$

$$\text{PI} \qquad c = \frac{(1 + \tau s)}{k\epsilon s} \qquad (75)$$

Table I. IMC-Based PID Controller Parameters[a]

	model	$y/y_s = \tilde{g}_+ f$	controller	$k_c k$	τ_I	τ_D	τ_F	comments
A	$\frac{k}{\tau s + 1}$	$\frac{1}{\epsilon s + 1}$	$\frac{1}{k}\frac{\tau s + 1}{\epsilon s}$	$\frac{\tau}{\epsilon}$	τ	–	–	–
B	$\frac{k}{(\tau_1 s + 1)(\tau_2 s + 1)}$	$\frac{1}{\epsilon s + 1}$	$\frac{(\tau_1 s + 1)(\tau_2 s + 1)}{k\epsilon s}$	$\frac{\tau_1 + \tau_2}{\epsilon}$	$\tau_1 + \tau_2$	$\frac{\tau_1 \tau_2}{\tau_1 + \tau_2}$	–	–
C	$\frac{k}{\tau^2 s^2 + 2\zeta\tau s + 1}$	$\frac{1}{\epsilon s + 1}$	$\frac{\tau^2 s^2 + 2\zeta\tau s + 1}{k\epsilon s}$	$\frac{2\zeta\tau}{\epsilon}$	$2\zeta\tau$	$\frac{\tau}{2\zeta}$	–	–
D	$k\frac{-\beta s + 1}{\tau s + 1}$	$\frac{-\beta s + 1}{\epsilon s + 1}$	$\frac{\tau s + 1}{k(\beta + \epsilon)s}$	$\frac{\tau}{\beta + \epsilon}$	τ	–	–	(2, 3, 5)
E	$k\frac{-\beta s + 1}{\tau s + 1}$	$\frac{-\beta s + 1}{(\beta s + 1)(\epsilon s + 1)}$	$\frac{\tau s + 1}{ks(\beta\epsilon s + 2\beta + \epsilon)}$	$\frac{\tau}{2\beta + \epsilon}$	τ	–	$\frac{\beta\epsilon}{2\beta + \epsilon}$	(1, 4)
F	$k\frac{-\beta s + 1}{\tau^2 s^2 + 2\zeta\tau + 1}$	$\frac{-\beta s + 1}{\epsilon s + 1}$	$\frac{\tau^2 s^2 + 2\zeta\tau s + 1}{k(\beta + \epsilon)s}$	$\frac{2\zeta\tau}{\beta + \epsilon}$	$2\zeta\tau$	$\frac{\tau}{2\zeta}$	–	(2, 3, 5)
G	$k\frac{-\beta s + 1}{\tau^2 s^2 + 2\zeta\tau s + 1}$	$\frac{-\beta s + 1}{(\beta s + 1)(\epsilon s + 1)}$	$\frac{\tau^2 s^2 + 2\zeta\tau s + 1}{k(\beta\epsilon s + 2\beta + \epsilon)s}$	$\frac{2\zeta\tau}{2\beta + \epsilon}$	$2\zeta\tau$	$\frac{\tau}{2\zeta}$	$\frac{\beta\epsilon}{2\beta + \epsilon}$	(1, 4)
H	$\frac{k}{s}$	$\frac{1}{\epsilon s + 1}$	$\frac{1}{k\epsilon}$	$\frac{1}{\epsilon}$	–	–	–	–
I	$\frac{k}{s}$	$\frac{2\epsilon + 1}{(\epsilon s + 1)^2}$	$\frac{2\epsilon s + 1}{k\epsilon^2 s}$	$\frac{2}{\epsilon}$	2ϵ	–	–	(6)
J	$\frac{k}{s(\tau s + 1)}$	$\frac{1}{\epsilon s + 1}$	$\frac{\tau s + 1}{k\epsilon}$	$\frac{1}{\epsilon}$	–	τ	–	–
K	$\frac{k}{s(\tau s + 1)}$	$\frac{2\epsilon s + 1}{(\epsilon s + 1)^2}$	$\frac{(\tau s + 1)(2\epsilon s + 1)}{k\epsilon^2 s}$	$\frac{2\epsilon + \tau}{\epsilon^2}$	$2\epsilon + \tau$	$\frac{2\epsilon\tau}{2\epsilon + \tau}$	–	(6)
L	$k\frac{-\beta s + 1}{s}$	$\frac{-\beta s + 1}{\epsilon s + 1}$	$\frac{1}{k(\beta + \epsilon)}$	$\frac{1}{\beta + \epsilon}$	–	–	–	(2, 3, 5)
M	$k\frac{-\beta s + 1}{s}$	$\frac{-\beta s + 1}{(\beta s + 1)(\epsilon s + 1)}$	$\frac{1}{k(\beta\epsilon s + 2\beta + \epsilon)}$	$\frac{1}{2\beta + \epsilon}$	–	–	$\frac{\beta\epsilon}{2\beta + \epsilon}$	(1, 4)
N	$k\frac{-\beta s + 1}{s}$	$\frac{(-\beta s + 1)((\beta + 2\epsilon)s + 1)}{(\epsilon s + 1)^2}$	$\frac{(\beta + 2\epsilon)s + 1}{ks(\beta + \epsilon)^2}$	$\frac{\beta + 2\epsilon}{(\beta + \epsilon)^2}$	$\beta + 2\epsilon$	–	–	(5, 6)

O	$k\frac{-\beta s + 1}{s}$	$\frac{(-\beta s + 1)(2(\beta + \epsilon)s + 1)}{(\beta s + 1)(\epsilon s + 1)^2}$	$\frac{2(\beta + \epsilon)s + 1}{ks(\beta\epsilon^2 s + \epsilon^2 + 4\beta\epsilon + 2\beta^2)}$	$\frac{2(\beta + \epsilon)}{2\beta^2 + 4\beta\epsilon + \epsilon^2}$	$2(\beta + \epsilon)$	–	$\frac{\beta\epsilon^2}{2\beta^2 + 4\beta\epsilon + \epsilon^2}$	(4, 6)
P	$k\frac{-\beta s + 1}{s(\tau s + 1)}$	$\frac{-\beta s + 1}{\epsilon s + 1}$	$\frac{\tau s + 1}{k(\beta + \epsilon)}$	$\frac{1}{\beta + \epsilon}$	–	τ	–	(2, 3, 5)
Q	$k\frac{-\beta s + 1}{s(\tau s + 1)}$	$\frac{-\beta s + 1}{(\beta s + 1)(\epsilon s + 1)}$	$\frac{\tau s + 1}{k(\beta\epsilon s + 2\beta + \epsilon)}$	$\frac{1}{2\beta + \epsilon}$	–	τ	$\frac{\beta\epsilon}{2\beta + \epsilon}$	(1, 4)
R	$k\frac{-\beta s + 1}{s(\tau s + 1)}$	$\frac{(-\beta s + 1)((\beta + 2\epsilon)s + 1)}{(\epsilon s + 1)^2}$	$\frac{(\tau s + 1)((\beta + 2\epsilon)s + 1)}{ks(\beta + \epsilon)^2}$	$\frac{\beta + 2\epsilon + \tau}{(\beta + \epsilon)^2}$	$\beta + 2\epsilon + \tau$	$\frac{\tau(\beta + 2\epsilon)}{\beta + 2\epsilon + \tau}$	–	(5, 6)
S	$k\frac{-\beta s + 1}{s(\tau s + 1)}$	$\frac{(-\beta s + 1)(2(\beta + \epsilon)s + 1)}{(\beta s + 1)(\epsilon s + 1)^2}$	$\frac{(\tau s + 1)(2(\beta + \epsilon)s + 1)}{ks(\beta\epsilon^2 s + \epsilon^2 + 4\beta\epsilon + 2\beta^2)}$	$\frac{2(\beta + \epsilon) + \tau}{2\beta^2 + 4\beta\epsilon + \epsilon^2}$	$2(\beta + \epsilon) + \tau$	$\frac{2\tau(\beta + \epsilon)}{2(\beta + \epsilon) + \tau}$	$\frac{\beta\epsilon^2}{2\beta^2 + 4\beta\epsilon + \epsilon^2}$	(4, 6)

[a] Controller form: $c = [k_c/(\tau_F s + 1)](1 + [1/(\tau_I s)] + \tau_D s)$. ϵ is the only adjustable parameter; for most cases ϵ is equivalent to the closed-loop time constant and $1/\epsilon$ is approximately the closed-loop bandwidth. In all cases, there exists no offset for step set-point/disturbance changes. Comments: 1. ISE optimal for step set-point changes when $\epsilon = 0$. 2. IAE optimal for step set-point changes when $\epsilon = 0$. 3. ISE optimal for step set-point changes when $\epsilon = \beta$. 4. Filter/factorization option 1 (64). Practical recommendation $\epsilon > \beta/2$. 5. Filter/factorization option 2 (66). Practical recommendation $\epsilon > \beta$. 6. No offset for ramp set-point/disturbance changes.

Table II. IMC-Based PID Parameters for $g(s) = ke^{-\theta s}/(\tau s + 1)$ and Practical Recommendations for ϵ/θ

controller	kk_c	τ_I	τ_D	recommended ϵ/θ (> $0.1\tau/\theta$ *always*)
PID	$(2\tau + \theta)/(2\epsilon + \theta)$	$\tau + (\theta/2)$	$\tau\theta/(2\tau + \theta)$	>0.8
PI	$\theta/\tau = 0.1$	1.54		>1.7
improved PI	$(2\tau + \theta)/2\epsilon$	$\tau + (\theta/2)$		>1.7

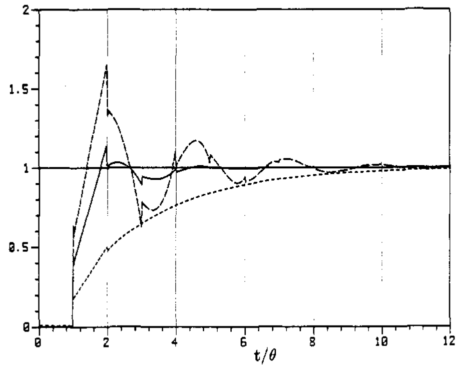

Figure 3. IMC-PID tuning rule. Effect of ϵ/θ on the closed-loop response to a unit step set-point change. $g(s) = ke^{-\theta s}/(\tau s + 1)$. (—) $\epsilon/\theta = 0.8$; (- - -) $\epsilon/\theta = 0.4$; (···) $\epsilon/\theta = 2.5$.

Option 2 (eq 66) was chosen for the filter for the first-order Padé approximation in order to get a PID controller without an additional lag term. These controllers are represented compactly in Table II. The closed-loop transfer functions for system (71) with these controllers indicate a number of advantages:

PID
$$y = \frac{e^{-\theta s}}{\dfrac{\left(\dfrac{\epsilon}{\theta} + \dfrac{1}{2}\right)\theta s}{1 + \dfrac{1}{2}\theta s} + e^{-\theta s}}(y_s - d) + d \qquad (76)$$

PI
$$y = \frac{e^{-\theta s}}{\left(\dfrac{\epsilon}{\theta}\right)\theta s + e^{-\theta s}}(y_s - d) + d \qquad (77)$$

The closed-loop response is independent of the system time constant τ. (The process lag $(1 + \tau s)$ is cancelled by the controller.) The time is scaled by θ. The shape of the response depends on ϵ/θ only.

In other words, specifying one value of ϵ/θ for any first-order lag with the dead-time model results in an identical response when the time is caled by θ, regardless of k, θ, and τ. For instance, if the dead time in system I is twice as long as the dead time in system II, then for a specific ϵ/θ, the response characteristics will be identical except that it will take the response of system I exactly twice as long to reach the same point as system II. The choice of the "best" ratio ϵ/θ must be based on performance and robustness considerations.

For the PID controller, Figure 3 demonstrates the dependence of the step response on ϵ/θ. $\epsilon/\theta = 0.4$ is fairly close to the value where instability occurs ($\epsilon/\theta = 0.145$), and the large overshoot and poorly damped oscillations are therefore not surprising. Note that $\epsilon/\theta = 0.5$ is the lower value recommended in Table I for models with a RHP zero factored according to (66). For $\epsilon/\theta = 0.8$, the response looks very good: the rise time is about 1.5θ and the settling time is 4.5θ; the overshoot is about 10%, and the decay

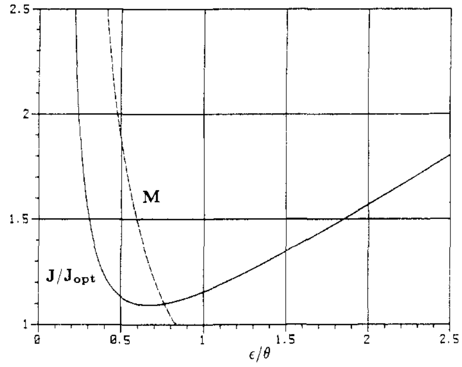

Figure 4. IMC-PID tuning rule (74). Effect of ϵ/θ on M and ISE (J) for step changes. $g(s) = ke^{-\theta s}/(\tau s + 1)$.

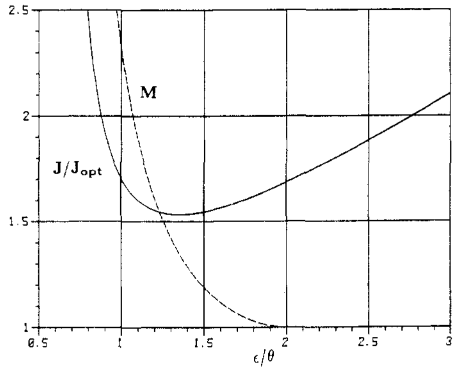

Figure 5. IMC-PI tuning rule (75). Effect of ϵ/θ on M and ISE (J) for step changes. $g(s) = ke^{-\theta s}/(\tau s + 1)$.

ratio is quite good. For $\epsilon/\theta = 2.5$, the response becomes highly overdamped and almost identical with that of a first-order system with time constant ϵ and delay θ.

The scaled form of the closed-loop transfer functions (76) and (77) allows convenient design plots to be made (Figures 4 and 5). The performance measure J, the integral square error to a step disturbance/set-point change, and the robustness measure M have been plotted as a function of ϵ/θ. In Figures 4 and 5, J is normalized by J_{opt}, the error corresponding to the optimum response $y/y_s = e^{-\theta s}$. In theory, a Smith predictor with infinite gain ($k_c = \infty$) accomplishes this response.

For PID control (Figure 4), J/J_{opt} reaches a minimum of 1.092 for $\epsilon/\theta = 0.68$. At this point, $M = 1.3$. For practical purposes, a better compromise between performance and robustness is attained for $\epsilon/\theta = 0.8$; here, the ISE is almost minimum but M has dropped to 1. For PI control (Figure 5), $\epsilon/\theta = 1.4$ results in the minimum J/J_{opt} value 1.55. M for this setting is approximately 1.3. $M = 1$ first occurs at $\epsilon/\theta \cong 2$, where $J/J_{opt} \cong 1.7$.

Figure 4 also confirms that the first-order Padé approximation leads to relatively little performance deterioration. For $\epsilon/\theta = 0.8$, the result is a PID controller that performs with only 10% greater ISE than the *optimal* Smith predictor, while retaining favorable robustness characteristics. Compared to the PI controller, however, the Smith predictor provides significant performance improvement; one must realize that the PI rule originates from a reduced model with no dependence on the process dead time. An alternate rule is described in section V.3 which takes into account this deficiency.

Figures 4 and 5 have been obtained under the assumption of no plant uncertainty; only the model error induced by the Padé approximation is considered. Significant plant uncertainty *within the bandwidth of the controller* will require the designer to select a larger value of ϵ. This consideration is of particular concern when $\epsilon/\theta \ll 1$. Because for the process industries the closed-loop bandwidth can rarely exceed 10 times the open-loop bandwidth ($10/\tau$), a practical requirement is to always select $\epsilon > \tau/10$. For the IMC-PID parameters, this inequality is dominant for $\theta/\tau < 1/7$; for the PI parameters, it will become important for $\theta/\tau < 1/14$.

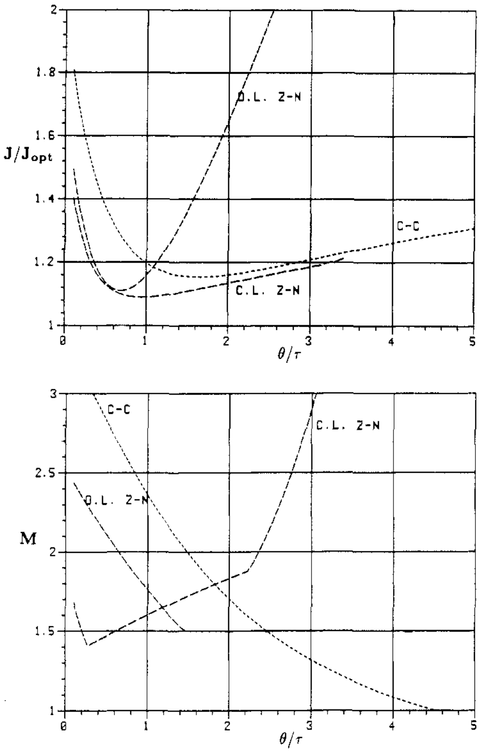

Figure 6. PID controllers for $g(s) = e^{-\theta s}/(\tau s + 1)$. Performance (top) and robustness (bottom) properties of the Cohen–Coon, open-loop Ziegler–Nichols, and closed-loop Ziegler–Nichols tuning rules.

V.2. Comparisons with Other Methods. Next we compare the IMC-PID parameters with the classic Ziegler–Nichols and Cohen–Coon tuning rules (Figure 6). The first notable difference between these rules and those from IMC is that J and M depend strongly on θ/τ, while for the IMC rules the performance and robustness measures are independent of this ratio. The Cohen–Coon rules give reasonable performance ($J/J_{opt} < 1.3$) for $0.6 < \theta/\tau < 4.5$. In this range, M varies between 2.7 and 1.0; i.e., robustness is quite poor, especially for small ratios of θ/τ. The performance obtained with the closed-loop Ziegler–Nichols parameters is good for the range $0.2 < \theta/\tau < 3.5$, but again the robustness is poor except for $\theta/\tau \cong 0.3$. Indeed, for $\theta/\tau > 4$, the closed-loop system is *unstable* with the $c - l$ Ziegler–Nichols parameters. In terms of performance, the open-loop Ziegler–Nichols parameters are only useful in the range $0.2 < \theta/\tau < 1.4$. The advantages of the IMC tuning rules are further demonstrated through simulations (Figures 7–9).

It should be emphasized, however, that by themselves, the higher M values for the Ziegler–Nichols and Cohen–Coon settings do not imply that these control systems can tolerate less plant/model mismatch than IMC before becoming unstable. As was explained in section II, model error tolerance depends on both M and the closed-loop bandwidth ω_b. Thus, only for a *particular* bandwidth/performance specification is IMC more robust than Ziegler–Nichols and Cohen–Coon. Comparing Ziegler–Nichols and Cohen–Coon with a small bandwidth/poor performance and IMC with a larger bandwidth/better performance can demonstrate a larger robustness of the former despite larger M values.

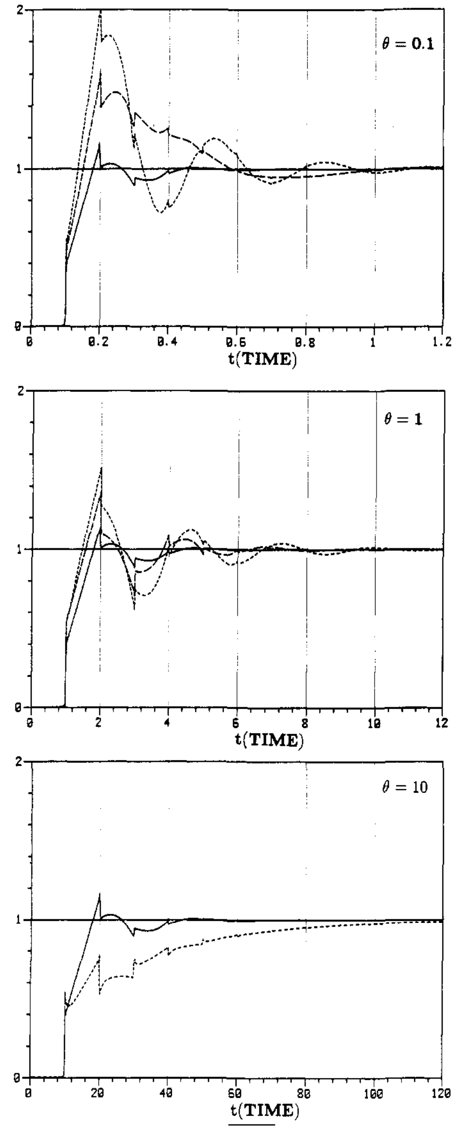

Figure 7. PID tuning rules for process $e^{-\theta s}/(\tau s + 1)$. Closed-loop responses to a unit step set-point change for $\theta = 0.1$, 1, and 10. (—) IMC ($\epsilon/\theta = 0.8$); (- - -) closed-loop Zielger–Nichols (unstable for $\theta = 10$); (···) Cohen–Coon.

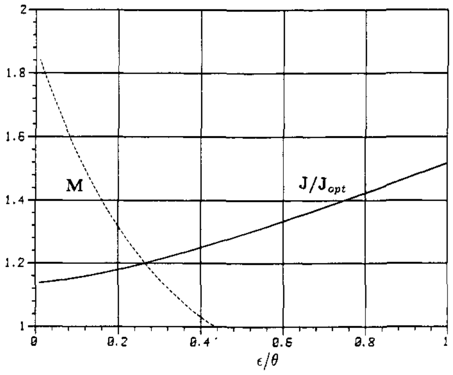

Figure 8. IMC-PID controller with exponential filter (Table I, entry G). Effect of ϵ/θ on M and the ISE (J) for step changes. $g(s) = ke^{-\theta s}/(\tau s + 1)$.

In Figure 7, the presence of pure derivative action (which is physically unrealizable) leads to a somewhat jerky response and to even more violent moves in the manipulated variable. If the proper IMC controller implied by entry G is used, the consequences are an increase in the ISE and a slower speed of response (Figure 8); the response, however, is smooth and looks more attractive (Figure 9). Here $\epsilon/\theta = 0.45$ was chosen to obtain good robustness characteristics ($M = 1$).

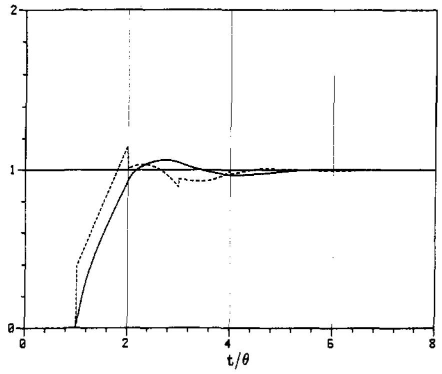

Figure 9. IMC-PID controllers. Closed-loop responses to a unit step set-point change. $g(s) = ke^{-\theta s}/(\tau s + 1)$. (—) IMC-PID with exponential filter (Table I, entry G), $\epsilon/\theta = 0.45$. (···) IMC-PID (Table I, entry F), $\epsilon/\theta = 0.8$.

V.3. Development of an Improved PI Rule. The IMC-PI rule (from entry A, Table I), despite its compactness and simplicity (as evidenced by Table II and Figure 5), has at best 55% greater performance cost than the optimal Smith predictor and is not overall superior to the Z-N and C-C expressions. This is a consequence of the zeroth-order Padê approximation and can be remedied by incorporating the dead time in the internal model through other means.

The IMC design procedure prescribes that first a process model $\tilde{g}$ should be established which closely approximates the real process; the controller structure and parameters follow directly from $\tilde{g}$. As a second step, the filter parameters are adjusted to compensate for the plant/model mismatch. In the context of the present example, it is clear that a zeroth-order Padê approximation is inadequate. If, in order to obtain a PI controller, a first-order lag is used to approximate a first-order lag with dead time, it appears reasonable to increase the model's lag over that of the process in order to account for the presence of the delay. Thus, we postulate the model

$$\tilde{g} = \frac{k}{(\tau\lambda)s + 1} \tag{78}$$

where λ depends on the process time delay. λ must be chosen such that "best" approximates the first-order lag/delay process. Rivera (1984) has established

$$\lambda = 1 + 0.5(\theta/\tau) \tag{79}$$

as suitable.

The PI rules obtained by using (79) appear in Table II. Comparing the "improved" PI controller

$$c = \frac{1 + (\tau + \theta/2)s}{k\epsilon s} \tag{80}$$

with the PID controller (74) and the PI-controller (75) based on the zeroth-order Padê approximation, the following becomes clear: For small time delays ($\tau/\theta \gg 1$) the improved PI rule and the original one are the same. For very large time delays ($\tau/\theta \ll 1$) the term $(1 + \tau s)$ in (74) will be outside the closed-loop bandwidth, and the improved PI controller and PID controller are equivalent when the following relationship is used:

$$\left(\frac{\epsilon}{\theta}\right)_{\mathrm{PI}} = \left(\frac{\epsilon}{\theta}\right)_{\mathrm{PID}} + 0.5 \tag{81}$$

Thus, for very large time delays, the PID controller approaches the PI controller with some gain correction according to (81); i.e., derivative action becomes ineffective.

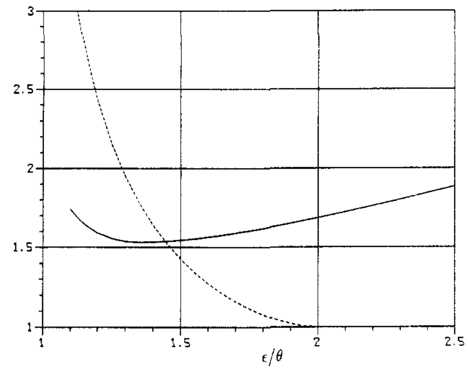

Figure 10. Improved IMC-PI tuning rule. Lower bound on performance and robustness for all θ/τ. $g(s) = ke^{-\theta s}/(\tau s + 1)$. (—) $(J/J_{\text{opt}})_{\max}$; (···) $M_{\max}$.

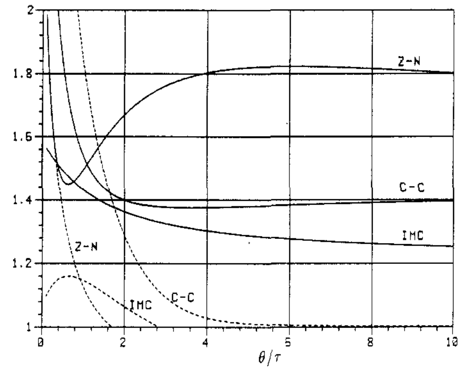

Figure 11. PI tuning rules for $g(s) = ke^{-\theta s}/(\tau s + 1)$. Performance and robustness properties for the improved IMC-PI tuning rule ($\epsilon/\theta = 1.7$), the closed-loop Ziegler–Nichols rule, and the Cohen–Coon rule. (—) J/J_{opt}; (···) M.

So far, no rules have been given on how to select ϵ for the improved IMC-PI controller. To provide an idea of the performance and robustness properties, Figure 10 was constructed. This plot indicates the maximum values of J/J_{opt} and M (over the entire θ/τ range) as a function of ϵ/θ. From Figures 4, 5, and 10, one finds that $\epsilon/\theta = 1.7$ provides a reasonable compromise between performance and robustness ($J/J_{\text{opt}} = 1.58$, $M = 1.15$). Note that this is slightly higher than $\epsilon/\theta = 1.3$ which is suggested from (81) based on the PID rule ($(\epsilon/\theta)_{\text{PID}} = 0.8$). Figure 11 demonstrates that with this choice, one obtains performance and robustness properties equal or superior to those of the Ziegler–Nichols and Cohen–Coon PI rules. Simulation results (Figure 12) confirm this.

Not only do the improved IMC-PI parameters lead to better performance and robustness than the traditional methods, the IMC design procedure also makes the search for the appropriate parameters simpler. In the IMC context, the PI controller is reparametrized with the parameters λ and ϵ. While in general it is necessary to search over k_c and τ_I simultaneously, IMC allows to search first for λ to obtain a good model fit and then for ϵ to obtain good performance and robustness.

V.4. Robustness to Dead-Time Errors. The following criterion which is sufficient for stability in the face of dead-time error was derived in section II:

$$\delta < \frac{1}{\omega_c M} \tag{82}$$

For the IMC-PID controller and $\epsilon/\theta \geq 0.8$ ($M = 1$), the following crossover frequency approximation holds (see Appendix B)

$$\frac{1}{\omega_c} \cong \frac{\theta}{2} + \epsilon \tag{83}$$

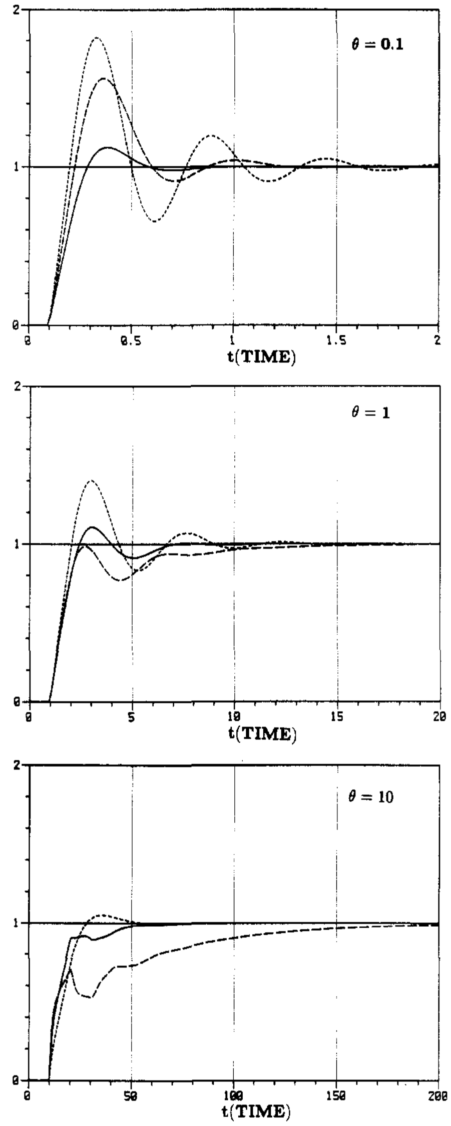

Figure 12. PI tuning rules for $g(s) = e^{-\theta s}/(\tau s + 1)$. Closed-loop responses to a unit step set-point change for $\theta = 0.1$, 1, and 10. (—) IMC-improved ($\epsilon/\theta = 1.7$); (---) closed-loop Ziegler–Nichols; (···) Cohen–Coon.

Likewise, for the improved PI controller ($\epsilon/\theta \geq = 1.7$, M varies with τ, as shown in Figure 11), one can approximate ω_c as

$$\frac{1}{\omega_c} \cong \epsilon \tag{84}$$

(83) and (84) lead to convenient expressions for the allowable dead-time error in terms of the design parameters ϵ, θ, and M

$$\text{PID} \qquad \delta < \epsilon + \frac{\theta}{2} \tag{85}$$

$$\text{improved PI} \qquad \delta < \frac{\epsilon}{M} \tag{86}$$

Table III confirms that (85) and (86) provide extremely accurate predictions for the allowable dead-time error. Furthermore, it shows that the suggested PI and PID settings provide dead-time error robustness in excess of 100%.

V.5. Tuning Based on Crossover Information. The classic Ziegler–Nichols identification procedure (1942) and its modern counterpart by Åström and co-workers (1983) do not provide a parametric model directly but only the

Table III. Allowable Dead-Time Error δ/θ for IMC-Based PID and PI Parameters[a]

controller	dead-time range	allowable dead-time error δ/θ	
		approx bound	exact bound
PID $\epsilon/\theta = 0.8$	$0 < \theta/\tau < \infty$	1.30	1.36
improved PI;	$\theta/\tau = 0.1$	1.54	1.55
$\epsilon/\theta = 1.7$	$\theta/\tau = 0.6$	1.46	1.46
	$\theta/\tau = 1.0$	1.50	1.52
	$\theta/\tau = 10.0$	1.70	1.97

[a] The exact bound is obtained from (17) while the approximate bounds are obtained from (85) and (86) for the PID and improved PI rules, respectively.

ultimate gain K_u and the ultimate period P_u. P_u is related to the phase crossover frequency ω_ϕ by

$$P_u = \frac{2\pi}{\omega_\phi} \tag{87}$$

K_u is the inverse of the process gain at ω_ϕ

$$K_u = |g(\omega_\phi)|^{-1} \tag{88}$$

Generally, a good estimate of the process time delay θ is also available from step tests. As Shinskey (1979) argues, it is reasonable to model processes with

$$2\theta < P_u < 4\theta \tag{89}$$

as a first-order lag with dead time as denoted by (71). If θ is known, k and τ can be estimated from K_u and P_u through the formulas

$$\tau = \frac{P_u \tan\left[\pi\left(1 - \frac{2\theta}{P_u}\right)\right]}{2\pi} \tag{90}$$

$$k = \frac{1}{K_u}\left(1 + \left[\frac{2\pi\tau}{P_u}\right]^2\right)^{1/2} \tag{91}$$

With the aid of (90) and (91), the PI and PID tuning rules from Table II can be applied when K_u, P_u, and θ are available, and a model of the form (71) is assumed.

VI. Summary

IMC-derived PID controller parameters are listed in Table I for most models commonly used in process control. In some cases, the IMC design procedure leads naturally to the need for a first-order lag to augment the PID controller structure. The single adjustable parameter ϵ is directly related to the speed of response, with $1/\epsilon$ approximately equal to the bandwidth of the closed-loop system. When the PID controller with the specified structure is applied to the *model*, the closed-loop system is stable for all $\epsilon > 0$. In practice, $1/\epsilon$ should be chosen to be smaller than the bandwidth over which the process model is valid. A good generally conservative initial guess is to set ϵ equal to the dominant time constant of the open-loop system.

If the model includes LHP zeros, these zeros should be cancelled first by an appropriate lag and then the entries from Table I can be used. If the system includes dead time, Padé approximations may be used to simplify the model; the result is a simple, lead-lag type controller. The zeroth-order approximation requires $\epsilon > \theta$, for the first-order approximation $\epsilon > \theta/3$.

When a first-order Padé approximation is used for a first-order lag with dead time, the IMC design technique yields the PID parameters listed in Table II. The ISE to a step change is minimal for $\epsilon/\theta = 0.68$. A better trade-off between performance and robustness is reached for $\epsilon/\theta = 0.8$. For small dead times ($\theta/\tau < 1$), robustness considerations (unmodeled dynamics) will dictate a larger ϵ.

By approximating a first-order lag with the dead-time model by a first-order lag without dead time, the IMC procedure leads to the "improved" PI parameters in Table II. For a choice of $\epsilon/\theta = 1.7$, the lower bound on performance, over the entire θ/τ range, is $J/J_{\mathrm{opt}} = 1.58$ and $M = 1.15$.

Furthermore, simple expressions for the allowable dead-time error δ are available in terms of the parameter ϵ. For the IMC-PID rule, this expression is

$$\delta < \left(\epsilon + \frac{\theta}{2}\right) \tag{85}$$

and for the improved PI rule

$$\delta < \frac{\epsilon}{M} \tag{86}$$

VII. Conclusions

We have shown that for most of the models used to describe the dynamics of chemical process systems, the PID controller is the natural choice. In the absence of nonlinearities, constraints, or multivariate interactions, it is infeasible to improve the performance with more complex controllers unless higher order, more accurate process models are available.

Furthermore, by substituting Padê approximations, these PID rules have been extended to models with dead time. For the particular case of a first-order lag with dead-time process, the improvement of the ISE for a step set point/disturbance by the Smith predictor over a PID controller is at most 10% *regardless* of θ/τ. For small values of θ/τ, this 10% improvement is generally not attainable because of model uncertainties. For large values of θ/τ, some improvement is possible if the process model is valid over a large enough bandwidth.

Although we show that PID-type controllers are adequate for most common process models, we find that the classical feedback structure is inadequate for a clear understanding of control system design. IMC formed the basis of all the rules in Tables I and II. If one were to use IMC directly and not insist on the traditional PID parameters, no rules and no involved tables would be needed. The IMC design procedure is generally applicable regardless of the system involved. No special provisions are required to deal with very single type of system. The complexity of the rules in Tables I and II demonstrates that the PID parameters k_C, τ_I, and τ_D are the consequences of a long hardware tradition rather than because they represent the most practical tuning tools. The unfortunate parametrization of the PID controller might also explain why some modern control methods (possessing structures that fall under that of IMC) have claimed improvements in control quality over PID for simple systems where a properly tuned PID controller would have yielded an equally good result. The results presented here also clearly point out the limitations of PID controllers. The practical occurrences of systems where no nonlinearities, constraints, or multivariate interactions are present are very rare. In all other situations, the PID controller must be "patched up" with antireset windup, dead-time compensators, and decouplers, while the IMC technique allows a unified treatment of all cases.

Finally, we must acknowledge (Lau and Balhoff, 1984) that the discrete form of IMC (Garcia and Morari, 1982), because of the increased number of tuning parameters and the added flexibility allowed by the discrete representation in formulating control objectives, can lead to performance and robustness improvements not possible with the PID parameters suggested here.

Acknowledgment

The idea for a paper of this type arose from discussion with J. P. Shunta and J. Richards from Du Pont Corp., Wilmington, DE. We are thankful for criticism from C. Economou and K. Levien. Financial support from the National Science Foundation (CPE-8115022) and the Department of Energy (DOE Contract DE-AC02-80ER10645) is gratefully acknowledged.

Appendix

A. Analytical forms for ω_b and ω_c can be found for the following:

Option 1.

$$\tilde{g}_+f = \frac{-\beta s + 1}{\beta s + 1}\frac{1}{(\epsilon s + 1)} \tag{64}$$

The sensitivity operator is

$$S = 1 - \tilde{g}_+f = \left(\frac{(2\beta + \epsilon)s\left(\dfrac{\epsilon\beta}{(2\beta + \epsilon)}s + 1\right)}{(\beta s + 1)(\epsilon s + 1)}\right) \tag{A.1}$$

which yields

$$\omega_b = \frac{(-(7\beta^2 + 8\beta\epsilon + \epsilon^2) + [(7\beta^2 + 8\beta\epsilon + \epsilon^2)^2 + 4(\beta\epsilon)^2]^{1/2})^{1/2}}{2^{1/2}\beta\epsilon} \tag{A.2}$$

Using the asymptote approximation for the amplitude of (A.1), one obtains

$$\omega_b \cong \frac{1}{2\beta + \epsilon}$$

To obtain the crossover frequency, we have cg

$$cg = \frac{\tilde{g}_+}{f^{-1} - \tilde{g}_+} = \frac{-\beta s + 1}{\epsilon\beta s^2 + (\epsilon + 2\beta)s} \tag{A.3}$$

from which one obtains

$$\omega_c = \frac{(-(\epsilon^2 + 4\beta\epsilon + 3\beta^2) + [(\epsilon^2 + 4\beta\epsilon + 3\beta^2)^2 + 4(\beta\epsilon)^2]^{1/2})^{1/2}}{2^{1/2}\epsilon\beta} \tag{A.4}$$

Again, using the asymptote approximation for the amplitude of (A.3), one obtains the simpler expression

$$\omega_c \cong \frac{1}{2\beta + \epsilon}$$

Option 2.

$$\tilde{g}_+f = \frac{-\beta s + 1}{\epsilon s + 1} \tag{66}$$

The sensitivity operator is

$$S = \frac{(\epsilon + \beta)s}{(\epsilon s + 1)} \tag{A.5}$$

which leads to

$$\omega_b = \frac{1}{[2(\epsilon + \beta)^2 - \epsilon^2]^{1/2}} \tag{A.6}$$

or, from the asymptote amplitude approximation to (A.5),

$$\omega_b \cong \frac{1}{\epsilon + \beta}$$

For the crossover frequency, the expressions are

$$cg = \frac{-\beta s + 1}{(\epsilon + \beta)s} \tag{A.7}$$

from which one obtains

$$\omega_c = \frac{1}{[(\epsilon + \beta)^2 - \beta^2]^{1/2}} \tag{A.8}$$

and from the asymptote amplitude ratio

$$\omega_c \cong \frac{1}{\epsilon + \beta}$$

Option 3.

$$\tilde{g}_+f = \frac{e^{-\theta s}}{\epsilon s + 1}$$

For this case, it is not possible to write the bandwidth or crossover expressions in explicit form. One can obtain an approximate expression by representing the dead time as a Padê approximation

$$\tilde{g}_+f \cong \frac{\left(1 - \dfrac{\theta}{2}s\right)}{\left(1 + \dfrac{\theta}{2}s\right)}\frac{1}{(\epsilon s + 1)}$$

and then proceeding according to option 1. The resulting expressions are

$$\omega_b \cong \omega_c \cong \frac{1}{\theta + \epsilon}$$

B. Crossover Approximation for PI and PID Rules. The expression for cg arising from the use of the IMC-PID rule to a first-order lag with the dead-time process is

$$cg = \frac{\left(1 + \dfrac{\theta}{2}s\right)e^{-\theta s}}{\left(\dfrac{\theta}{2} + \epsilon\right)s} \tag{B.1}$$

One can solve explicitly for the crossover frequency to obtain

$$\omega_c = \frac{1}{\theta\left[\left(\dfrac{\epsilon}{\theta} + \dfrac{1}{2}\right)^2 - \left(\dfrac{1}{2}\right)^2\right]^{1/2}} \tag{B.2}$$

For $\epsilon/\theta > 0.8$, it is reasonable to neglect the latter $(1/2)^2$ term and thus approximate ω_c as

$$\omega_c \cong \frac{1}{\theta\left[\left(\dfrac{\epsilon}{\theta} + \dfrac{1}{2}\right)^2\right]^{1/2}} = \frac{1}{\epsilon + \dfrac{\theta}{2}}$$

For the IMC-improved PI rule, cg is

$$cg = \frac{\left(1 + \left(\tau + \dfrac{\theta}{2}\right)s\right)e^{-\theta s}}{\epsilon s(1 + \tau s)} \tag{B.3}$$

As $\theta \rightarrow 0$, it is clear that the crossover frequency reaches the value

$$\omega_c \rightarrow \frac{1}{\epsilon} \tag{B.4}$$

which implies

$$\omega_c \cong \frac{1}{\epsilon} \qquad \text{(B.5)}$$

Assuming, however, that $\tau = 0$ (the worst case), we see that (B.5) is still a good approximation. Consider that for $\tau = 0$, the crossover frequency is determined explicitly by

$$\omega_c = \frac{1}{\theta\left[\left(\frac{\epsilon}{\theta}\right)^2 - \left(\frac{1}{2}\right)^2\right]^{1/2}}$$

For $\epsilon/\theta > 1.7$, $(\epsilon/\theta)^2 \gg (1/2)^2$ and therefore

$$\omega_c \cong \frac{1}{\epsilon}$$

Literature Cited

Åström, K.; Hagglund, T. IFAC Workshop on Adaptive Control, San Francisco, 1983.

Brosilow, C., personal communication, Case Western Reserve University, Cleveland, OH, 1983.

Cohen, G. H.; Coon, G. A. *Trans. ASME* **1953**, *75*, 827.

Doyle, J.; Stein, G. *IEEE Trans. Autom. Control* **1981**, *AC-26*, 4.

Frank, P. M. "Entwurf von Regelkreisen mit vorgeschriebenem Verhalten"; G. Braun Verlag: Karlsruhe, 1974.

Garcia, C. E.; Morari, M. *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 308.

Holt, B. R.; Morari, M. *Chem. Eng. Sci.*, in press.

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

Lau, H. K.; Balhoff, R. A. personal communication, Shell Development, Houston, TX, 1984.

Rivera, D. M.S. Thesis, University of Wisconsin, Madison, 1984.

Rosenbrock, H. H. "Computer-Aided Control System Design"; Academic Press: New York, 1974.

Shinskey, F. G. "Process Control Systems"; McGraw-Hill: New York, 1979.

Smith, C. L. "Digital Computer Process Control"; Intext Educational Publishers: Scranton, PA, 1972.

Wiberg, D. M. "State Space and Linear Systems"; Schaum, T., Ed.; McGraw-Hill: New York, 1971.

Ziegler, J. G.; Nichols, N. B. *Trans. ASME* **1942**, *64*, 759.

Received for review July 10, 1984
Revised manuscript received July 22, 1985
Accepted August 1, 1985

Flash Pyrolysis of Sewage Sludge

Jan Piskorz, Donald S. Scott,* and Ian B. Westerberg

Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

A dried mixture of raw and activated sewage sludge was pyrolyzed in a bench-scale fluidized-bed reactor at residence times of less than 1 s over the temperature range 400–700 °C. Organic liquid yields of up to 52% maf were obtained at the optimum temperature of 450 °C and 0.55-s residence time. Shorter reaction times gave slightly higher liquid yields. The effects of reaction atmosphere, char accumulation, and deashing on liquid yields were also investigated. A test in a larger unit at a feed rate of 2 kg/h duplicated the bench-scale result. From elemental analyses, the char appears to be aromatic in character, while the liquid tar product has a H/C atomic ratio of about 1.77. A certain amount of water-soluble alcohols, ketones, etc., appear to be present also.

Treatment of municipal and industrial wastewaters generates significant quantities of sludge, particularly if biological processes are used. The cost of disposal of these sludges, whether treated or untreated, has increased greatly in recent years, and it has been estimated that the disposal cost is now 50% of the total wastewater treatment cost. Some recently used disposal methods may become more restricted in future years, for example, landfill, water, or ocean disposal and agricultural use. In 1980, sludge disposal costs were estimated to be \$208/ton in Canada, \$182/ton in the United States, and \$162/ton in the European Economic Community (Bridle, 1982). There is an urgent need for alternate solutions to the sludge disposal problem, and one of these may be the conversion of sludge to a liquid fuel. Sludge combustion has been practiced for a long time in a variety of processes but is normally a net consumer of fuel and often encounters severe problems with ash clinkering or slagging and with air pollution.

The possibility of converting municipal waste to a useful fuel oil was demonstrated several years ago by the development of the Occidental pyrolysis process, which used a hot circulating flow of ash or char to rapidly decompose organic material (Prober and Bauer, 1977). More recently, German work has demonstrated that a synthetic crude oil could be produced from sewage sludge by heating at 300 °C for 3 h under an inert atmosphere (Bayer and Kutubbudin, 1982). Workers at Battelle Northwest Laboratories have developed a similar process in which sludge is heated with sodium carbonate for 1 h under an argon atmosphere (as quoted in "Chemical Engineering", 1981). The German process is claimed to be a net producer of energy if the sludge is dewatered.

Bridle et al. (1983a, 1983b) in Canada have recently produced a good quality oil from a dried mixture of raw and waste activated sludge using a retort type of reactor at atmospheric pressure. Yields of up to 28% (dry basis) were obtained at 425 °C. A high proportion of aliphatic hydrocarbon was found in the pyrolytic oil produced.

Another possible pyrolysis process which has been used successfully with various biomass materials, and with coal, for the production of liquids is the short residence time fluidized bed. The use of a fluidized bed of sand permits very high heating rates of the solid and the rapid removal of vapors formed by thermal decomposition of the feed. This process has been called flash pyrolysis and has been applied at atmospheric pressure to a wide variety of biomass materials by Scott and Piskorz (1981, 1982a) with good yields of liquid products. The application of this process to dried sewage sludge will be described in the following sections.

0196-4305/86/1125-0265\$01.50/0