Digital Controllers for SISO Systems: A Review and a New Algorithm

bу

E. Zafiriou and M. Morari

# Digital controllers for SISO systems: a review and a new algorithm

#### EVANGHELOS ZAFIRIOU† and MANFRED MORARI†

Several digital control algorithms for linear single-input single-output systems are examined and the effect of the sampling period on their performance is analysed in terms of rippling, overshoot and settling time. The problem is addressed in the frequency domain (z-transform) and it is shown that each controller works for some classes of systems but that none works for all. The similarities and differences of these controllers are established and an explanation of their deficiencies is given based on the location of the zeros of the discrete system. The insight gained leads to a simple new rule for the design of a controller which combines the advantages of the different algorithms but at the same time is free of their problems. A single tuning parameter is included which directly affects the closed-loop speed of response and bandwidth. The parameter can be used to detune the controller in the event that the real system differs from the model on which the controller design is based. No tuning is necessary when the available model is exact, unless smaller values for the manipulated variable, at the cost of a slower response, are preferred.

#### 1. Introduction

In the literature one can find a very large number of digital control algorithms for single-loop systems. Numerous design criteria are used and often they are formulated so as to satisfy the special requirements of a specific process. The synthesis of many of those controllers is quite complicated and it is usually very difficult to know when and why problems may occur.

A common property of the control algorithms examined in this paper is the computational simplicity of their synthesis; it so happens that all of them can in fact be described as pole-zero placement controllers, though that may not be the way they were initially designed. Each of them tries to satisfy some commonly accepted criteria, such as:

- (i) A performance criterion for the discrete output of the system (for example, minimization of the sum of squared errors).
- (ii) Behaviour between the sampling points: hidden oscillations in the system output can be caused by ringing of the manipulated variable or by unobservable oscillatory open-loop modes.
- (iii) Settling time (for example, the requirement to reach the set-point in a finite number of time steps).
- (iv) Overshoot and/or undershoot.
- (v) Manipulated variable: large values must be avoided because they usually cannot be implemented due to saturation of the manipulated variable.

In this paper we compare the set-point step responses obtained with various controllers. In the first part, an analysis of the problems is given for the case of no

Received 20 November 1984.

<sup>†</sup> Department of Chemical Engineering, California Institute of Technology, Pasadena, California 91125, U.S.A.

modelling error and a new algorithm which avoids the disadvantages of the examined controllers is proposed. The second part deals with stability issues of the new algorithm when a mismatch between the model and the plant exists.

## 2. Evaluation of the control algorithms

In the case of no modelling error, the classical feedback (Fig. 1 (a)) is equivalent, for command following, to the open-loop structure shown in Fig. 1 (b), where  $G_c(z)$  may be considered as the controller to be designed.

There is a simple relation between  $G_c(z)$  and the controller of the classical feedback structure C(z):

$$G_{c}(z) = \frac{C(z)}{1 + C(z)G(z)} \tag{1}$$

where G(z) is the transfer function of the plant. We can similarly obtain C(z) from  $G_c(z)$  by

$$C(z) = \frac{G_c(z)}{1 - G_c(z)G(z)} \tag{2}$$

Designing  $G_c(z)$  and then obtaining C(z) by eqn. (2) is an established technique (see, for example, Jury and Schroeder (1956)). This technique will be used throughout the paper because it allows presentation of the results in a transparent manner.

For the pulse transfer function of the plant we shall use the expression

$$G(z) = K \frac{(z - a_1^-) \dots (z - a_k^-)(z - a_{k+1}^+) \dots (z - a_{n-1}^+)}{(z - p_1) \dots (z - p_n)} z^{-N}$$
(3)

where N is the largest integer such that NT is less than or equal to the dead-time (T being the sampling time). The superscripts '—' and '+' denote the zeros inside and outside the unit circle respectively and k is the number of zeros inside the unit circle. G(z) is assumed to be open-loop stable.

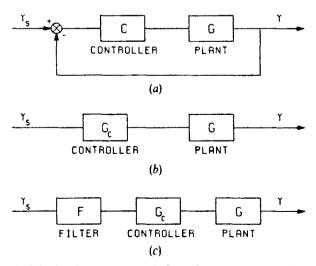


Figure 1. (a) Classical feedback structure. (b) Open-loop structure equivalent to the feedback for the exact model. (c) The open-loop structure with the filter included.

# 2.1. Digital control algorithms

#### (i) Dahlin's controller

This is a well known controller included in current textbooks like Smith (1972) and Kuo (1977). It was introduced by Dahlin (1968), who proposed a method for designing the controller by specifying the closed-loop transfer function V(z) to be first order with a dead-time equal to that of the plant and a steady-state gain of 1.0, thus:

$$V(z) = \frac{[1 - \exp(-T/\lambda)]z^{-1-N}}{1 - \exp(-T/\lambda)z^{-1}}$$
(4)

where  $\lambda$  is the time constant of the closed-loop response and T is the sampling interval. From V(z) and G(z), the classical feedback controller is found to be

$$C_{DC}(z) = \frac{[1 - \exp(-T/\lambda)]z^{-1-N}}{1 - \exp(-T/\lambda)z^{-1} - [1 - \exp(-T/\lambda)]z^{-1-N}} \frac{1}{G(z)}$$
(5)

The time constant of the closed-loop response serves as a tuning parameter for adjusting the speed of the response.

If we take the equivalent structure of Fig. 1 (b) the controller is found from eqn. (1) to be

$$G_{\rm DC}(z) = \frac{1 - \exp(-T/\lambda)}{1 - \exp(-T/\lambda)z^{-1}} \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{z(z - a_1) \dots (z - a_{n-1})}$$
(6)

where the expression for G(z) given by (3) is used. Hence we can rearrange to get the structure of Fig. 1 (c) with controller

$$G_{C}(z) = \frac{1}{K} \frac{(z - p_{1}) \dots (z - p_{n})}{z(z - a_{1}) \dots (z - a_{n-1})}$$
(7)

and a filter

$$F(z) = \frac{1 - \exp(-T/\lambda)}{1 - \exp(-T/\lambda)z^{-1}}$$
 (8)

The reason for separating out F(z) and calling it a 'filter' will become apparent in § 4 of the paper dealing with modelling errors. Note that for  $\lambda = 0$  (no filter), Dahlin's controller becomes a deadbeat controller which brings the output to the set-point value after one sampling period. The stability problems are discussed in § 2.2.

Dahlin also proposed a modification to be used when there are problems with the initial algorithm; we will examine this modified controller in detail in § 2.2, after the reasons for the problems are explained.

(ii) The controller that minimizes the sum of the squared errors of the output This controller is designed so as to minimize the objective function

$$\Phi = \sum_{j=0}^{\infty} (y_j - r)^2$$

where  $y_i$  is the value of the system output at t = jT and r the desired set-point. The

3× =

problem is solved in the z-domain by Chang (1961) and in the state space of Kucera (1972).

The controller for the structure of Fig. 1 (b) is (Kucera 1972):

$$G_{SE}(z) = \frac{1}{K} \frac{\left(1 - \frac{1}{a_{k+1}^+}\right) \dots \left(1 - \frac{1}{a_{n-1}^+}\right)}{(1 - a_{k+1}^+) \dots (1 - a_{n-1}^+)} \frac{(z - p_1) \dots (z - p_n)}{z(z - a_1^-) \dots (z - a_k^-) \left(z - \frac{1}{a_{k+1}^+}\right) \dots \left(z - \frac{1}{a_{n-1}^+}\right)}$$
(9)

# (iii) Output and state deadbeat controllers

The output deadbeat controller is defined as the controller that drives the discrete output of the system to the set-point value and keeps it there, after a minimum number of time steps. The earliest discussion of this kind of controller was given by Bergen and Ragazzini (1954). Kalman (1954) raised the question of the behaviour of the continuous output between the sampling points. Jury and Schroeder (1956) and Jury (1958) studied the problem by using the modified z-transform and Kalman and Bertram (1959) in the state-space.

The design of the state deadbeat controller is based on the idea of having the states of the system at equilibrium when the discrete output reaches the set-point so that the continuous output will remain there as well. The minimum number of time steps required for all the states to be at equilibrium is at most equal to the number of the states, which is equal to the order of the system. For SISO systems it is easier to design this controller using the z-transform. As a design criterion, the manipulated variables are required to remain constant after the output of the system has reached the set-point (minimal prototype controllers; see, Ragazzini and Franklin (1958) and Luyben (1973)).

Increasing the specified settling time by one or more sampling intervals introduces additional degrees of freedom which allow one to prescribe a value for the manipulated variable (Isermann 1981) or to optimize an objective function or to prescribe the system output (Janiszowski 1983).

Output deadbeat

The controller for the open-loop structure of Fig. 1 (b) is (Kucera 1972)

$$G_{\text{OD}}(z) = \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{(z - a_1^-) \dots (z - a_k^-) z^{n-k} (1 - a_{k+1}^+) \dots (1 - a_{n-1}^+)}$$
(10)

State deadbeat

The controller for the open-loop structure is

$$G_{SD}(z) = \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{(1 - a_1) \dots (1 - a_{n-1}) z^n}$$
(11)

Note that for state deadbeat control the closed-loop characteristic equation has (n + N) roots at the origin (Aström and Wittenmark 1984).

## (iv) Vogel-Edgar controller

Vogel and Edgar (Vogel 1982) derived a control algorithm by using the general pole-zero placement controller design procedure presented by Aström and Witten-

mark (1980). For the conventional feedback structure, the Vogel-Edgar controller is  $C_{VC}(z) =$ 

$$\frac{1}{K} \frac{[1 - \exp(-T/\lambda)](z - p_1) \dots (z - p_n)}{z^n [1 - \exp(-T/\lambda)z^{-1}](1 - a_1) \dots (1 - a_{n-1}) - [1 - \exp(-T/\lambda)](z - a_1) \dots (z - a_{n-1})z^{-N}}$$
(12)

For the equivalent structure of Fig. 1 (b), this controller becomes

$$G_{VE}(z) = \frac{1 - \exp(-T/\lambda)}{1 - \exp(-T/\lambda)z^{-1}} \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{(1 - a_1) \dots (1 - a_{n-1})z^n}$$
(13)

Comparing (13) with (8) and (11) we note that the Vogel-Edgar controller is equivalent to the state deadbeat controller (11) with an added filter (8). The filter parameter  $\lambda$  is used for tuning, as in the case of Dahlin's controller. By comparing this controller with Dahlin's, Vogel and Edgar (Vogel 1982) found their's to be superior for second-order systems.

The expressions for the controller  $G_c(z)$  of the open-loop structure (Fig. 1 (b)) are summarized for the discussed control algorithms in Table 1.

Dahlin  $G_{DC}(z) = F(z) \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{z(z - q_n) \dots (z - q_{n-1})}$ 

Output deadbeat  $G_{\text{OD}}(z) = \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{(1 - a_{k+1}^+) \dots (1 - a_{n-1}^+)(z - a_1^-) \dots (z - a_k^-) z^{n-k}}$ 

Minimum sum of squared errors  $G_{SE}(z) = \frac{1}{K} \frac{\left(1 - \frac{1}{a_{k+1}^+}\right) \dots \left(1 - \frac{1}{a_{n-1}^+}\right)}{(1 - a_{k+1}^+) \dots (1 - a_{n-1}^+)}$ 

$$\times \frac{(z-p_1)...(z-p_n)}{z(z-a_1^-)...(z-a_k^-)\left(z-\frac{1}{a_{k+1}^+}\right)...\left(z-\frac{1}{a_{n-1}^+}\right)}$$

State deadbeat

$$G_{SD}(z) = \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{(1 - a_1) \dots (1 - a_{n-1})z^n}$$

Vogel-Edgar

$$G_{VE}(z) = F(z) \frac{1}{K} \frac{(z - p_1) \dots (z - p_n)}{(1 - a_1) \dots (1 - a_{n-1}) z^n}$$

Where

$$F(z) = \frac{1 - \exp(-T/\lambda)}{1 - \exp(-T/\lambda)z^{-1}}$$

the plant is

$$G(z) = K \frac{(z - a_1^-) \dots (z - a_k^+)(z - a_{k+1}^+) \dots (z - a_{k-1}^+)}{(z - p_1) \dots (z - p_k)} z^{-N}$$

and the superscripts '-' and '+' denote 'inside' and 'outside' the unit circle, respectively.

Table 1. Controller  $G_c(z)$  of the open-loop structure of Fig. 1 (b).

### 2.2. Analysis of the deficiencies of the algorithms

From Table 1, it is apparent that all these control algorithms are in fact pole-zero placement controllers and that the poles of the controller and the closed-loop transfer function are explicitly related to the zeros of the plant. Hence, the performance of these controllers is strongly related to the location of the zeros of G(z) on the complex plane.

Aström et al. (1984) have proved the following theorems regarding the effect of the sampling period on the location of the zeros of sampled systems.

For the system shown in Fig. 2 the pulse transfer function G(z) between y and u is

$$G(z) = (1 - z^{-1}) \mathcal{Z} \mathcal{L}^{-1} \left\{ \frac{A(s)}{s} \right\}$$

$$(14)$$

Figure 2. Pulse transfer function:  $G(z) = (1 - z^{-1}) \cdot \mathcal{L} \mathcal{L}^{-1} \{A(s)/s\}$ .

Theorem 1 (Åström et al. 1984)

Let A(s) be a strictly proper (m < n) rational function:

$$A(s) = K \frac{(s - v_1) \dots (s - v_m)}{(s - w_1) \dots (s - w_n)}$$
(15)

Then as  $T \to 0$ , m zeros of G(z) go to 1 as  $\exp(v_i T)$  and the remaining (n - m - 1) zeros of G(z) go to the zeros of  $B_{n-m}(z)$ , where

$$B_k(z) = b_1^k z^{k-1} + b_2^k z^{k-2} + \dots + b_k^k$$
 (16)

and

$$b_i^k = \sum_{l=1}^i (-1)^{i-l} l^k \binom{k+1}{i-l}, \quad i = 1, 2, ..., k$$
 (17)

Some of the zeros of the polynomials  $B_k$  are outside or on the unit circle. The 'unstable' zeros for a few values of k are listed in Table 2.

k	Unstable zero of $B_k$
2	-1
3	-3.732
4	-1, -9.899
5	-2.322, -23.20
6	-1, -4.542, -51.22
7	-1.868, $-8.160$ , $-109.3$
8	-1, -3.138, -13.96, -228.5
9	-1.645, $-4.957$ , $-23.14$ , $-471.4$

Table 2. Unstable zeros of the polynomial  $B_k(z)$  of eqn. (16).

Theorem 2 (Åström et al. 1984)

Let A(s) be a strictly proper rational transfer function with  $A(0) \neq 0$  and Re  $(w_i) < 0$ . Then all zeros of the pulse transfer function G(z) given by eqn. (14) go to zero as the sampling period T goes to infinity.

# (i) Dahlin's controller

There are two problems associated with this controller:

(a) When there exist zeros of G(z) outside the unit circle then it follows from (5) or (6) that the transfer function between the set-point and the input to the plant is unstable. As for the transfer function between the output and the set-point, it will be unstable unless exact cancellation occurs between the unstable poles of  $G_{\rm DC}(z)$  and the unstable zeros of G(z).

It is important to notice that G(z) may have unstable zeros not only when there are RHP zeros in the Laplace transfer function of the plant, but even for systems with only LHP zeros or no zeros at all. As we can see from Theorem 1 and Table 2, a system with three poles and no zeros in the Laplace transfer function, for example, will have a zero outside the unit circle for 'small' sampling periods.

(b) Even when all the zeros are inside the unit circle, there will still be problems if there are zeros close to -1. That would cause ringing of the controller output (oscillations with period twice the sampling period) and its effect on the system output would be rippling between the sampling instants. In addition, the ringing of the manipulated variable may cause unnecessary equipment wear. This problem is illustrated in Example 1.

Example 1

For T=3,

$$G(s) = \frac{1}{(10s+1)(25s+1)} \Rightarrow G(z) = 0.0157 \frac{(z+0.869)}{(z-0.887)(z-0.741)}$$

For  $\lambda = 0$  the response is shown in Fig. 3. For  $\lambda = 3$  the response is shown in Fig. 4. As  $\lambda$  increases, the response becomes slower but the oscillations remain, though their amplitude is reduced.

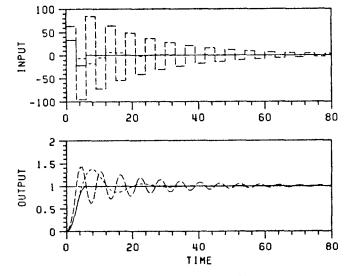


Figure 3. G(s) = 1/(10s + 1)(25s + 1), T = 3,  $\lambda = 0$ : --- Dahlin's controller (also the one that minimizes the sum of squared errors); ··· modified Dahlin; --- the new algorithm; --- set-point.

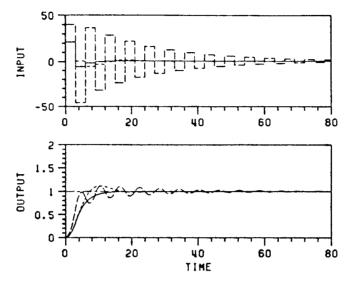


Figure 4. G(s) = 1/(10s + 1)(25s + 1), T = 3,  $\lambda = 3$ : --- Dahlin's controller; ··· modified Dahlin; --- the new algorithm; --- set-point.

Note that zeros near -1 may occur, depending on G(s), not only for small Ts but also for intermediate Ts. This will be the case, for example, for a third-order system with no zeros in the Laplace transfer function (see Table 2 and Theorems 1 and 2).

# Modified Dahlin controller

The modification introduced by Dahlin (1968) is to substitute zeros at the origin for the zeros of G(z) (in (5)) which are unstable or cause ringing, while keeping the same steady-state gain for  $C_{\rm DC}(z)$ .

Assume that there are m zeros of G(z) that have to be discarded (unstable or those that cause ringing). The expression for the modified Dahlin controller can be derived from eqn. (5).

$$C'_{DC}(z) = \frac{\left[1 - \exp\left(-T/\lambda\right)\right]z^{-1}}{1 - \exp\left(-T/\lambda\right)z^{-1} - \left[1 - \exp\left(-T/\lambda\right)\right]z^{-1-N}} \frac{1}{K}$$

$$\times \frac{(z - p_1) \dots (z - p_n)}{z^m (1 - a_1) \dots (1 - a_m)(z - a_{m+1}) \dots (z - a_{n-1})}$$
(18)

For the system of Example 1, the modified controller will give the response shown in Fig. 3, which is clearly better than that obtained with the unmodified algorithm. By increasing  $\lambda$ , the response can be improved significantly (Fig. 4).

As we will see in § 3 of this paper, the modification would be much more effective if the substitution of the 'ringing' zeros of G(z) with zeros at the origin were applied to the controller  $G_{\rm DC}(z)$  (eqn. (6)) of the equivalent structure of Fig. 1 (b). When the modification is made on  $C_{\rm DC}(z)$ , the problem may remain, as Example 2 illustrates. An additional disadvantage is that we do not know beforehand for which systems problems will appear.

Example 2

For 
$$T=1$$
,

$$G(s) = \frac{\exp(-s)}{(2s+1)(5s+1)} \Rightarrow G(z) = 0.0398 \frac{(z+0.792)}{(z-0.819)(z-0.607)} z^{-1}$$

The responses with both the unmodified and the modified Dahlin controllers are shown in Fig. 5. Although the modification is supposed to eliminate ringing, the time during which the oscillation of the manipulated variable persists, increases significantly when the modification is made.

Let us now examine the performance of the modified Dahlin controller when unstable zeros are present.

It follows from eqn. (18) that as  $\exp(-T/\lambda) \to 1$   $(\lambda \to \infty)$ , (m+N) of the roots of the closed-loop characteristic polynomial go to the origin and one goes to +1 (from inside the unit circle). Hence, there is no doubt that a sufficiently large  $\lambda$  can be found to stabilize the system. The problem is that this tuning is necessary even when the model is exact and we do not know beforehand where to look for exp  $(-T/\lambda)$ .

A system with an unstable zero is examined in Example 3.

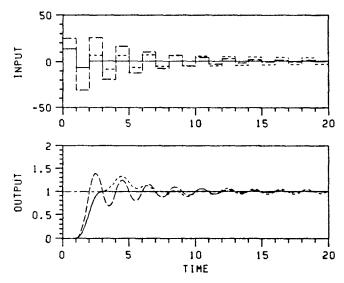


Figure 5.  $G(s) = \exp(-s)/(2s+1)(5s+1)$ , T=1,  $\lambda = 0$ : --- Dahlin's controller; ··· modified Dahlin; --- the new algorithm; --- set-point.

Example 3

For 
$$T = 0.1$$
.

For 
$$T = 0.1$$
,  

$$G(s) = 3.333 \frac{(-s+1.5)}{(s+1)(s+2)(s+2.5)} \Rightarrow G(z) = -0.01316 \frac{(z-1.162)(z+0.792)}{(z-0.905)(z-0.819)(z-0.779)}$$

Both zeros (one is unstable, and one causes ringing) have to be substituted by zeros at the origin when applying eqn. (18). Then

$$\lambda = 0$$
 unstable.

$$\lambda = 0.718 \, (\exp{(-T/\lambda)} = 0.87)$$

 $\lambda = 0.718 \, (\exp{(-T/\lambda)} = 0.87)$  stable (at the edge of instability, i.e. undamped oscillations of the discrete, as well as of the continuous output).

$$\lambda = 1.95 \left( \exp \left( - T/\lambda \right) = 0.95 \right)$$

response shown in Fig. 6.

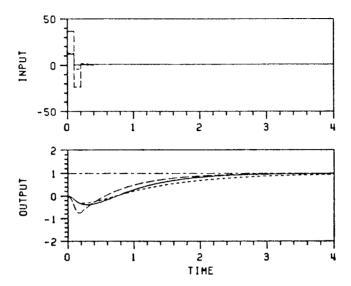


Figure 6. G(s) = 3.333(-s+1.5)/(s+1)(s+2)(s+2.5), T = 0.1: ... modified Dahlin,  $\lambda = 1.95$ ; ... the new algorithm,  $\lambda = 0$ : ... the new algorithm,  $\lambda = 0.25$ ; ... set-point.

For  $\lambda = 1.95$ , the roots of the closed-loop characteristic polynomial are at 0.30, 0.56 and 0.91. Hence, one of them is close to 1/1.162 (i.e. the inverse of the unstable zero). The other two are near enough to the origin for their effect on the response to be close to the effect that two roots at the origin would have. These observations form part of the basis of the new algorithm introduced in § 3.

# (ii) Output deadbeat

There are two problems associated with this controller.

- (a) Potential ringing of the controller output. This will happen when there are zeros with negative real part *inside* the unit circle; the closer they are to -1 the worse the rippling.
- (b) The second problem is associated with the way the control system handles the unstable zeros. Since those zeros cannot be used as poles of  $G_c(z)$ , in order to have steady-state gain 1 for the closed-loop transfer function, i.e. V(1) = 1, the terms  $(1 a_{k+1}^+), ..., (1 a_{n-1}^+)$  have to be included in the denominator of the controller and they remain in the closed-loop transfer function V(z), too. We have

$$V(z) = \frac{(z - a_{k+1}^+) \dots (z - a_{n-1}^+)}{z^{n-k} (1 - a_{k+1}^+) \dots (1 - a_{n-1}^+)} z^{-N}$$
(19)

From eqn. (19) we find that for a set-point step change the first non-zero value of the system output will occur at time t = (N + 1)T and will be of magnitude

$$y_1 = \frac{1}{(1 - a_{k+1}^+) \dots (1 - a_{n-1}^+)}$$
 (20)

Equation (20) implies that if some of those zeros are near +1,  $y_1$  will be large and significant undershoot or overshoot will occur. From Theorem 1, we see that this

happens for small sampling times when the zeros outside the unit circle are introduced by RHP zeros of the Laplace transfer function.

The response for the system of Example 3 is shown in Fig. 7, where oscillations, as well as large undershoot and overshoot, occur.

# (iii) The controller that minimizes the sum of the squared errors of the output

There is only one problem with this controller, but it is nevertheless important, i.e. ringing of the manipulated variable and, as a result, rippling of the system output between the samples when there are zeros with negative real part near the unit circle.

On the other hand, this controller does not suffer the second problem of the output deadbeat controller, illustrated in Fig. 7. The first non-zero value of the system output is

$$y_1 = \frac{\left(1 - \frac{1}{a_{k+1}^+}\right) \dots \left(1 - \frac{1}{a_{n-1}^+}\right)}{\left(1 - a_{k+1}^+\right) \dots \left(1 - a_{n-1}^+\right)} \tag{21}$$

Hence when one of those  $a^+$ s is near + 1, then its inverse,  $1/a^+$ , is also near + 1, in fact it is closer to +1 so that  $|y_1| < 1$ .

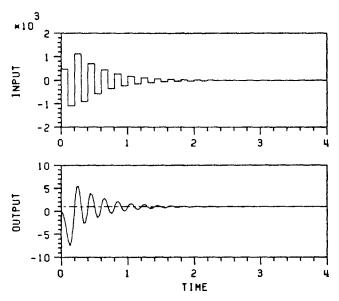


Figure 7. G(s) = 3.333(-s+1.5)/(s+1)(s+2)(s+2.5), T = 0.1: ——output deadbeat controller; —set-point.

For the case of Example 1, this controller is identical to Dahlin's (its response can be seen in Fig. 3). Another case is examined in Example 4.

Example 4

For T = 1.65,

$$G(s) = \frac{2}{(s^2 + 1.2s + 1)(s + 2)} \Rightarrow G(z) = 0.4168 \frac{(z + 0.0708)(z + 1.058)}{(z^2 - 0.184z + 0.138)(z - 0.0369)}$$

By using the inverse of -1.058 as a pole for the controller  $G_c(z)$  (according to eqn. (9)), we get the oscillatory response shown in Fig. 8.

In both Figs. 3 and 8 we see that this controller clearly minimizes the sum of the squared errors of the *discrete* output, but at the same time no attention is paid to the system output between the samples.

# (iv) State deadbeat

Equation (11) shows that the ringing problem of the controller output will not occur. Nevertheless, the state deadbeat controller frequently suffers from large overshoot or undershoot of the discrete output itself before equilibrium is reached. The reason is similar to that discussed for the output deadbeat controller, but now the problem arises not only for zeros of G(z) outside the unit circle, but also for zeros inside the unit circle.

The closed-loop transfer function is

$$V(z) = \frac{(z - a_1) \dots (z - a_{n-1})}{(1 - a_1) \dots (1 - a_{n-1}) z^n} z^{-N}$$
 (22)

The first non-zero value of the discrete output for a set-point step change is

$$y_1 = \frac{1}{(1 - a_1) \dots (1 - a_{n-1})}$$
 (23)

Equation (23) implies that when there are zeros near +1,  $y_1$  will be large, which results in significant overshoot or undershoot or both.

From Theorem 1 we know that this will happen for small Ts whenever the Laplace transfer function of the system has zeros either in the LHP or the RHP. Also note that for systems with more than one zero in the Laplace transform, all the zeros contribute to the large value of  $y_1$ .

The problem is illustrated in Example 5.

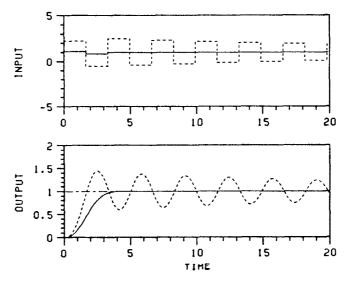


Figure 8.  $G(s) = 2/(s^2 + 1 \cdot 2s + 1)(s + 2)$ ,  $T = 1 \cdot 65$ ,  $\lambda = 0$ : ... the controller that minimizes the sum of squared errors; —— the new algorithm; --- set-point.

Example 5

For T = 0.1.

$$G(s) = 1.5 \frac{(s+2)}{(s+1)(s+3)} \Rightarrow G(z) = 0.136 \frac{(z-0.819)}{(z-0.905)(z-0.741)}$$

The response is given in Fig. 9; the overshoot is about 450%.

A case with a RHP zero is the following.

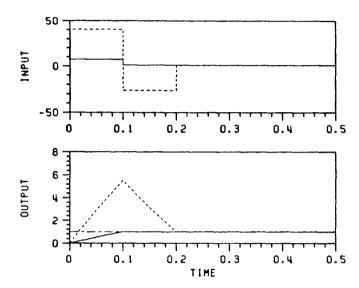


Figure 9. G(s) = 1.5(s+2)/(s+1)(s+3), T = 0.1; ... state deadbeat (also Vogel-Edgar,  $\lambda = 0$ ); ... the new algorithm,  $\lambda = 0$  (also Vogel-Edgar,  $\lambda = 0.5$ ); ... set-point.

Example 6

For T = 0.1,

$$G(s) = 0.5 \frac{(-s+2)}{(s^2+1.5s+1)} \Rightarrow G(z) = -0.0416 \frac{(z-1.223)}{(z^2-1.851z+0.861)}$$

The response with an extremely large undershoot is shown in Fig. 10.

# (v) Vogel-Edgar controller

This controller is identical to the state deadbeat controller when the tuning parameter  $\lambda$  is equal to zero. The tuning parameter can only help in certain cases to eliminate the overshoot/undershoot problem of the state deadbeat controller.

The closed-loop transfer function is

$$V(z) = \frac{1 - \exp(-T/\lambda)}{1 - \exp(-T/\lambda)z^{-1}} \frac{(z - a_1) \dots (z - a_{n-1})}{(1 - a_1) \dots (1 - a_{n-1})z^n} z^{-N}$$
(24)

and the first non-zero value of the discrete output is

$$y_1 = \frac{1 - \exp(-T/\lambda)}{(1 - a_1) \dots (1 - a_{n-1})}$$
 (25)

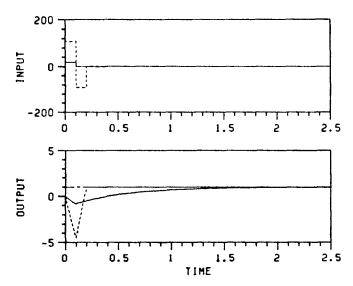


Figure 10.  $G(s) = 0.5(-s+2)/(s^2+1.5s+1)$ , T = 0.1: ... state deadbeat (also Vogel-Edgar,  $\lambda = 0$ ); ... the new algorithm,  $\lambda = 0$  (also Vogel-Edgar,  $\lambda = 0.497$ ); ... set-point.

The tuning parameter can always be made sufficiently large so that  $y_1$  is as small as desired. However for large  $\lambda$  the term  $1 - \exp(-T/\lambda)z^{-1}$  in the denominator of the closed-loop transfer function will slow down the response significantly, unless it is cancelled by one of the zeros of the system. Hence, if there is only one zero near +1 and inside the unit circle, by selecting  $\lambda$  such that  $\exp(-T/\lambda)$  is equal to that zero, not only the overshoot is eliminated, but, at the same time, the response becomes faster. This is illustrated in Fig. 9 for the system of Example 5, where by selecting  $\lambda = 0.5$ ,  $\exp(-T/\lambda)$  becomes equal to the zero at 0.819.

In the case where there is only one zero near + 1 but it is outside the unit circle, the best selection for  $\lambda$  would be such that  $\exp(-T/\lambda)$  is equal to the inverse of that zero. We can see the result for the system of Example 6, in Fig. 10  $(\lambda = 0.497 \Rightarrow \exp(-T/\lambda) = 1/1.223)$ .

Vogel and Edgar did not provide any guidelines for the selection of  $\lambda$ . They suggest gradually increasing  $\lambda$  until the response becomes satisfactory.

In the case where there are two or more zeros near +1,  $\lambda$  has to be made very large in order to bring the overshoot or undershoot down to acceptable values (cancellation with only one of the zeros is not sufficient). The result is that the response becomes too slow to be acceptable. This will happen for systems with two or more zeros in the Laplace transfer function, either in the LHP or the RHP. The problem is illustrated in Example 7.

Example 7

For T = 0.1,

$$G(s) = 2.25 \frac{(s+1)(s+2)}{(s+0.5)(s+1.5)(s^2+2.5s+6)}$$
  
$$\Rightarrow G(z) = 0.01067 \frac{(z-0.905)(z-0.819)(z+0.951)}{(z-0.951)(z-0.861)(z^2-1.726z+0.779)}$$

Two of the zeros are close enough to +1 to produce the extremely large overshoot shown in Fig. 11 (for  $\lambda = 0$ ). By selecting  $\lambda$  such that  $\exp(-T/\lambda)$  is equal to one of the zeros ( $\lambda = 0.5$ ,  $\exp(-T/\lambda) = 0.819$ ), the response shown in Fig. 11 is obtained. For  $\exp(-T/\lambda)$  equal to the other zero ( $\lambda = 1$ ,  $\exp(-T/\lambda) = 0.905$ ), the response is given in Fig. 12. The overshoot is still clearly unacceptable. In order to reduce it to 75%,  $\lambda$  has to be increased to 2.5 and extremely sluggish behaviour results (Fig. 12).

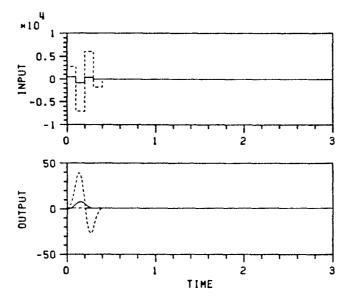


Figure 11.  $G(s) = 2.25(s+1)(s+2)/(s+0.5)(s+1.5)(s^2+2.5s+6)$ , T = 0.1: ··· state deadbeat (also Vogel-Edgar,  $\lambda = 0$ ); ··· Vogel-Edgar,  $\lambda = 0.5$ ; ··- set-point.

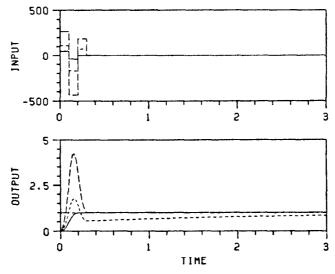


Figure 12.  $G(s) = 2.25(s+1)(s+2)/(s+0.5)(s+1.5)(s^2+2.5s+6)$ , T = 0.1: --- Vogel-Edgar,  $\lambda = 1$ ; ··· Vogel-Edgar,  $\lambda = 2.5$ ; —- the new algorithm,  $\lambda = 0$ ; --- set-point.

#### 2.3. Discussion

There are two basic problems associated with the performance of the controllers examined. One is rippling of the system output between the samples due to ringing of the manipulated variable and the other is large overshoot and/or undershoot of the discrete output itself. Dahlin's controller will also be unstable when the system has zeros outside the unit circle and the modified Dahlin controller has to be stabilized in this case by appropriate selection of the tuning parameter.

The transfer functions of the controller  $G_c(z)$  of the open-loop structure (Fig. 1 (b)) are given in Table 1. In § 2.2 we explained how the problems are related to the location of the zeros of the model and how the location depends on the continuous system Laplace transfer function and on the sampling period T (Theorems 1 and 2).

The relations between the continuous system, the discrete controller, the sampling period and the associated problems are given in Table 3. The entries in the first

Continuous system	Control algorithm	Small T	$T_{ m cr} - \Delta T$	$T_{\rm cr} + \Delta T$
	DC		<del>-</del>	Unstable†
	OD	_	<del>+</del>	Overshoot†
LHP zero	SE		+ +	†
	SD	Overshoot	Overshoot†	Overshoot†
	VE	One zero:		
		More: overshoot	Overshoot†	Overshoot†
	DC	Unstable	Unstable	Rippling
	OD	Undershoot		Rippling
RHP zero	SE	<del></del>	Rippling	Rippling
	SD	Undershoot	_	
	VE	One zero: —	-	
		More: undershoot		
	DC	n-m=2: rippling		
		$n-m \ge 4$ : unstable	Unstable	Rippling
Excess zeros	OD	Rippling	******	Rippling
n-m even <sup>+</sup>	SE	Rippling	Rippling	Rippling
	SD			<del></del>
	VE		_	<del></del>
	DC	Unstable	Unstable	Rippling
Excess zeros	OD			Rippling
n-m odd	SE		Rippling	Rippling
	SD	manuscon.	——————————————————————————————————————	
** *	VE	-		
(≥ 3)	SD		— — —	— — — —

<sup>†</sup> Two or more zeros (see § 2.3).

Table 3. Summary, of the problems of the examined control algorithms.

<sup>‡</sup> For n - m = 2 only the column for small T applies.

DC Dahlin's controller

OD Output deadbeat controller

SE Minimizes sum of squared errors of the output

SD State deadbeat controller

VE Vogel-Edgar controller

Tcr Sampling period at which the zero of the discrete system crosses the unit circle

 $<sup>\</sup>Delta T$  Range around  $T_{cr}$  for which the corresponding problem occurs

m, n Degrees of numerator and denominator of G(s)

column refer to the Laplace transfer function of the system; hence the table shows what effect a LHP or a RHP zero or a pole excess (n-m) of 2 or more, can have on the performance of each controller for various ranges of the sampling period T.

How small T has to be in order for the problems in Table 3 to arise, depends on the continuous system and the type of problem. An overshoot of 100% caused by a LHP zero at small Ts can occur for  $T \le 0.75\tau$  for a real zero and for  $T \le 0.9\tau$  for a pair of complex-conjugate zeros, where  $\tau$  is the inverse of the distance of the zero from the origin. Similarly for a RHP zero, we may get a significant undershoot for  $T \le 0.75\tau$  for a real zero and for  $T \le 0.9\tau$  for complex zeros. The rippling caused by excess zeros at small Ts (when (n-m) is even) can occur for T up to  $0.75\tau_D$ , where  $\tau_D$  is the dominant time constant of the system, but usually the problem appears for smaller Ts which can start as low as  $0.2\tau_D$ .

 $T_{\rm cr}$  denotes the critical sampling period at which a zero crosses the unit circle. Rippling may occur around  $T_{\rm cr}$  for a range from  $T_{\rm cr} - \Delta T$  to  $T_{\rm cr} + \Delta T$ . The value of  $T_{\rm cr}$  for each zero of G(z) has to be calculated from the Laplace transfer function G(s).  $\Delta T$  appears to be a fraction of  $T_{\rm cr}$ : it is about  $0.3T_{\rm cr}$  when we have a zero of G(z) corresponding to a RHP zero of G(s) and about  $0.7T_{\rm cr}$  when we have an excess zero crossing the unit circle. The reason for the difference is that a zero corresponding to a RHP zero covers a greater distance on the complex plane than an excess zero, as T goes from zero to infinity, and therefore it moves faster with T.

When G(s) has more than one zero in the LHP and when these zeros are much smaller than the poles (by an order of magnitude or more), then the corresponding discretized system G(z) can have a zero either inside or outside the unit circle, depending on the sampling time. From Theorems 1 and 2, it follows that if such a zero crosses the unit circle, then there are two  $T_{\rm cr}$ s for which a crossing occurs. For the smaller  $T_{\rm cr}$  at which the zero leaves the unit circle, the problems of the examined algorithms are listed in Table 3. For the larger  $T_{\rm cr}$  (at which the zero returns to the unit circle), all algorithms will generate significant overshoot before t=T and, in addition, have all the other problems particular to each algorithm. The reason is that the open-loop step response of systems of this type shows large overshoot and the second  $T_{\rm cr}$  is larger than the time t at which this overshoot occurs. This problem is not of practical importance since such a large T would result in a folding frequency  $\pi/T$  smaller than the bandwidth of G(s) and would therefore not be selected.

From an inspection of the open-loop step response, it can easily be determined if any of the zeros of G(z) arising from LHP zeros of G(s) cross the unit circle as follows. Assume that G(s) has been scaled so that G(0) > 0. If the unit step response is negative for some t' and if T = t' then G(z) has an odd number of zeros on the real axis between +1 and  $+\infty$ . The converse is also true. The reason is that K in (3) is the value of the open-loop step response at t = T. For all the systems tested with such crossings there was always a range of T for which the number of zeros between +1 and  $+\infty$  was odd. Hence, the above condition on the step response seems to be both necessary and sufficient for the existence of such crossings. However no proof is available.

### 3. A new algorithm

In § 2 it became clear that the controller  $G_c(z)$  of the structure of Fig. 1 (b) should not have poles with negative real part close to the unit circle. These poles cause oscillation of the controller output (with period equal to twice the sampling period if

the poles are negative reals) and as a result the process output exhibits intersample rippling.

Hence, when there are zeros of G(z) with negative real part, one should not use those zeros, or their inverses, as poles of  $G_c(z)$ , but use poles at the origin instead.

When there are unstable zeros of G(z) with positive real part, one should use their inverses as poles of  $G_c(z)$ . If zeros close to +1 were substituted with poles at the origin, then, in order to avoid steady-state offset (i.e. in order to have  $G_c(1) = G(1)^{-1}$ ), large undershoot or overshoot would have to be accepted.

In addition, for all the control algorithms examined in § 2, the zeros of  $G_c(z)$  are chosen to be equal to the poles of G(z).

Hence a simple rule to design the controller  $G_c(z)$  is the following.

- (1) Use as zeros of  $G_c$ , the poles of G(z) (they are assumed to be stable.
- (2) Use as poles of  $G_c$ , the zeros of G(z) with positive real part which are inside the unit circle, the inverses of those with positive real part which are outside the unit circle and as many at the origin as there are zeros with negative real part.
- (3) An additional pole of  $G_c$  at the origin must be present because of the inherent time delay of a discrete system.
- (4) The steady-state gain of  $G_c$  should be:

$$G_{\rm c}(1) = \frac{1}{G(1)}$$

Let the superscripts '-' and '+' denote 'inside' and 'outside' the unit circle, respectively, and the subscripts '+' and '-' denote 'positive' and 'negative' real part. We have

$$G_{c}(z) = \frac{1}{K} \frac{\left(1 - \frac{1}{a_{+,k+1}^{+}}\right) \dots \left(1 - \frac{1}{a_{+,l}^{+}}\right)}{(1 - a_{+,k+1}^{+}) \dots (1 - a_{+,l}^{+})(1 - a_{-,l+1}) \dots (1 - a_{-,n-1})} \times \frac{(z - p_{1}) \dots (z - p_{n})}{z^{n-l}(z - a_{+,1}^{-}) \dots (z - a_{+,k}^{-})\left(z - \frac{1}{a_{+,k+1}^{+}}\right) \dots \left(z - \frac{1}{a_{+,l}^{+}}\right)}$$
(26)

where k zeros  $(a_{+,1}^-, ..., a_{+,k}^-)$  have positive real part and are inside the unit circle, l-k  $(a_{+,k+1}^+, ..., a_{+,l}^+)$  have positive real part and are outside the unit circle and n-1-l  $(a_{-,l+1}, ..., a_{-,n-1})$  have negative real part.

A filter of the form  $F(z) = (1 - \alpha)/(1 - \alpha z^{-1})$  (where  $0 \le \alpha < 1$ ) should also be included. Theoretical justification for the inclusion of the filter will be given in § 4. At this point one can, nevertheless, note that the tuning parameter can be used to lower the values of the manipulated variable. This of course will be accomplished at the cost of a slower response.

The classical feedback controller C(z) for the structure of Fig. 1 (a) can be obtained from

$$C(z) = \frac{F(z)G_{c}(z)}{1 - F(z)G_{c}(z)G(z)}$$
(27)

The behaviour of the new algorithm was examined for the examples of § 2 and, as the corresponding plots show, it performs very well (Figs. 3-12).

The algorithm is expected to work well all the time, unless of course there are unobservable open-loop oscillatory modes (Åström and Wittenmark 1984). This will happen if the sampling period matches the frequency of any oscillatory mode of the continuous open-loop system. The problem can be met by appropriate selection of the sampling period (for example, less than half the period of oscillations of the continuous system impulse response (Jury 1957)). The controller  $G_c(z)$  can then be designed from G(z) as above.

It should also be noted that this design procedure can be extended to external system inputs other than steps. One might already have observed that one can obtain the new algorithm of eqn. (26) by writing the controller transfer function  $G_{SE}(z)$  (structure of Fig. 1 (b)) that minimizes the sum of the squared errors and substituting the poles with negative real part with poles at the origin, while keeping the same steady-state gain. The same procedure can be used for inputs other than steps to yield a problem-free controller. One should first obtain the transfer function of the controller (structure of Fig. 1 (b)) that minimizes the sum of the squared errors for the specified input and then make the correction mentioned above.

# 4. Stability in the presence of modelling error

For no modelling error, the classical feedback structure of Fig. 1 (a) is equivalent to the open-loop structure of Fig. 1 (b), where eqns. (1) and (2) relate the controllers C(z) and  $G_c(z)$ . In the case where the model is not exact, let us denote the model by  $\tilde{G}(z)$  and the actual plant by G(z). Then the feedback structure is equivalent to that of Fig. 13 (a) where the two added  $\tilde{G}(z)$  blocks cancel each other. This structure can be simplified to that in Fig. 13 (b).

The relationships between the controller  $G_c(z)$  in Fig. 13 (b) and the classical feedback controller C(z) are

$$G_{c}(z) = \frac{C(z)}{1 + C(z)\tilde{G}(z)}$$
(28)

$$C(z) = \frac{G_{c}(z)}{1 - G_{c}(z)\widetilde{G}(z)}$$
(29)

For an exact model, the structure of Fig. 13 (b) simplifies for command-following to that shown in Fig. 1 (b), i.e. it becomes open-loop, which makes the design of  $G_c$  transparent. For example, if Dahlin had substituted the zeros which cause ringing in the corresponding  $G_c(z)$  instead of C(z) he would have got the desired result regarding ringing, without the complications and problems of the modified Dahlin controller.

On the other hand, when the model  $\tilde{G}(z)$  is not exact, the structure of Fig. 13 (b) is not effectively open-loop and stability problems can arise.

A very useful property of this structure is that any controller  $G_c(z)$  which satisfies  $G_c(1) = \tilde{G}(1)^{-1}$  yields zero offset (Garcia and Morari 1982). From Fig. 13 (b), we find

$$u(z) = \frac{G_{c}(z)}{1 + G_{c}(z)[G(z) - \tilde{G}(z)]} y_{s}(z)$$
(30)

$$y(z) = \frac{G(z)G_{c}(z)}{1 + G_{c}(z)[G(z) - \tilde{G}(z)]}y_{s}(z)$$
(31)

where u is the manipulated variable.

For stability, it is necessary and sufficient that both of the following characteristic equations have their roots strictly inside the unit circle:

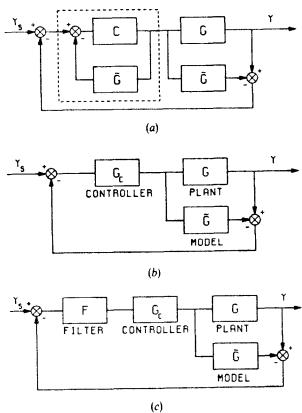


Figure 13. (a) Structure equivalent to the classical feedback. (b) Simplification of the structure shown in (a). (c) The equivalent structure with the filter included.

$$\frac{1}{G_{c}(z)} + [G(z) - \tilde{G}(z)] = 0$$
 (32)

$$\frac{1}{G(z)G_{c}(z)} + \frac{1}{G(z)}[G(z) - \tilde{G}(z)] = 0$$
 (33)

By adding a filter F(z) to the controller  $G_c(z)$  we get the structure in Fig. 13 (c). The importance of the filter is shown by the following theorem (Garcia and Morari 1985).

## Theorem 3

Let G(z) and  $\widetilde{G}(z)$  be stable,  $G_c(1) = 1/\widetilde{G}(1)$  and

$$F(z) = \frac{1 - \alpha}{1 - \alpha z^{-1}}, \quad 0 \leqslant \alpha < 1 \tag{34}$$

Then there exists an  $\alpha^*$  ( $0 \le \alpha^* < 1$ ) such that the system is closed-loop stable for all  $\alpha$  in the range  $\alpha^* \le \alpha < 1$  if and only if G(z) and  $\tilde{G}(z)$  satisfy

$$G(1)\tilde{G}(1) > 0 \tag{35}$$

Hence, by using the structure of Fig. 13 (c) and a  $G_c(z)$  such that  $G_c(1) = 1/\tilde{G}(1)$ ,

stability can always be achieved for any mismatch of model and plant, as long as the steady-state gain of the model has the correct sign.

This structure does not have to be actually implemented, since it is entirely equivalent to the classical feedback. The conventional feedback controller C(z) can be obtained from

$$C(z) = \frac{F(z)G_c(z)}{1 - F(z)G_c(z)\widetilde{G}(z)}$$
(36)

An advantage of using the structure of Fig. 13 (c) for design purposes is that it provides a clear physical meaning to the tuning parameter  $\alpha$ .

#### 5. Conclusions

The purpose of this paper was to present and explain, in a transparent manner, the problems of some well known digital controllers for SISO systems. Based on the results of this study, a simple rule was derived for the design of a controller which will always perform well. The reason for trying to establish a simple rule is to make sure that this design procedure is of use to the practicing engineer.

The algorithm developed is free of the basic problems of the controllers examined, i.e. intersample rippling and overshoot or undershoot. In the case where all the unstable zeros of the pulse transfer function of the system G(z) have negative real part, it yields a deadbeat controller which drives the discrete output of the system to the setpoint in a finite number of time steps. When G(z) has unstable zeros with positive real part, the controller drives the output to the set-point asymptotically in order to avoid large overshoot or undershoot. When all the zeros, stable or unstable, have positive real part, it minimizes the sum of the squared errors of the output.

An advantage of the proposed structure is that a tuning parameter is included (filter of Fig. 13 (c)) whose physical meaning and effect is clear to the designer. If the model is exact, no tuning is necessary in contrast with the Dahlin and Vogel-Edgar controllers. In the case of a mismatch between the system and the model, stability can always be guaranteed when some simple conditions hold. It is also important to note that the same tuning parameter can be used to reduce the values of the manipulated variable at the cost of a slower response.

#### ACKNOWLEDGEMENT

Support from the National Science Foundation and the Department of Energy is gratefully acknowledged.

# REFERENCES

ASTRÖM, K. J., HAGANDER, P., and STERNBY, J., 1984, Automatica, 20, 31.

ÅSTRÖM, K. J., and WITTENMARK, B., 1980, Proc. Instn elect. Engrs., Pt. D, 127, 120; 1984, Computer Controlled Systems (Englewood Cliffs: Prentice-Hall).

BERGEN, A. R., and RAGAZZINI, J. R., 1954, A.I.E.E. Trans., 73, 236.

CHANG, S. S. L., 1961, Synthesis of Optimum Control Systems (New York: McGraw-Hill).

DAHLIN, E. B., 1968, Instrum. Control Syst., 41, 77.

GARCIA, C. E., and MORARI, M., 1982, Ind. Engng chem. Proc. Des. Dev., 21, 308; 1985, Ibid., 24, 472.

ISERMANN, R., 1981, Digital Control Systems (Springer-Verlag).

JANISZOWSKI, K., 1983, Int. J. Control, 37, 159.

Jury, E. I., 1957, A.I.E.E. Trans., 75, 391; 1958, Sampled-Data Control Systems (New York: John

JURY, E. I., and SCHROEDER, W., 1956, A.I.E.E. Trans., 75, 317.

KALMAN, R. E., 1954, A.I.E.E. Trans., 73, 245.

KALMAN, R. E., and BERTRAM, J. E., 1958, A.I.E.E. Trans., 78, 602.

KUCERA, V., 1972, Kybernetica, 8, 233.

Kuo, B. C., 1977, Digital Control Systems (Champaign, Illinois: SRI Publishing Co.). Luyben, W. L., 1973, Process Modelling, Simulation and Control for Chemical Engineers (New York: McGraw-Hill).

RAGAZZINI, J. R., and FRANKLIN, G. F., 1958, Sampled-Data Control Systems (New York: McGraw-Hill).

SMITH, C. L., 1972, Digital Computer Process Control (Scranton, Pennsylvania: Intext Educational Publishers).

VOGEL, E. F., 1982, Adaptive control of chemical processes with variable dead time. Ph.D. dissertation, University of Texas, Austin.