# A New Approach to the Discretization of Continuous-Time Controllers

Jürg P. Keller and Brian D. O. Anderson, *Fellow, IEEE*

*Abstract*—**A new method for discretizing continuous-time controllers is derived, using principles of controller approximation [1]. It focuses on the closed-loop use of the discrete-time controller. The resulting approximation criterion is a measure for stability of the control system, provides an upper bound on the sampling time for which stability can be guaranteed, and because it is based on continuous-time controller approximation it indicates the cost of discretization in terms of performance degradation. The discrete-time controller is obtained through minimization of this criterion, which can be performed with standard software used in $H_\infty$ controller design.**

## I. Introduction

Most controller design procedures allow direct discrete-time controller design. Nevertheless there is some reason to design first a continuous-time controller, which is afterwards converted into a discrete-time one. Direct design of discrete-time controllers requires a predetermination of the sampling time in order to model plant, noise, process disturbances, etc. Upper bounds for the sampling time can only be obtained after considering closed-loop bandwidth ([5, ch. 10]) and, depending on specifications, this may only be available *after* controller design. Because this obvious dilemma does not occur in continuous-time controller design, it may be reasonable to design first a continuous-time controller, which is in a second step approximated by a discrete-time one. A second reason for first executing a continuous-time design is that physical insight is better retained.

There are many methods available for discretizing continuous-time controllers, for instance, bilinear transformation, hold input approximation, and signal invariant transformations, but only a few of them take account of the closed-loop use of the controller. Yet discretization of a continuous-time controller is a type of approximation, and, as argued in [1], it is logical to take closed-loop performance into account, whenever a controller is approximated. Methods for discrete-time implementation of continuous-time linear state-feedback laws are proposed by [11], [20], and [10]. A closed-loop redesign method is proposed by Rattan and Yeh [15] and Rattan [16], but none of these approximation methods can guarantee the elementary property of closed-loop stability after controller discretization. Stability of sampled data systems was investigated by Thompson *et al.* [18], [19] using conic sector analysis. In [19] the conservativeness of the conic sector proposed in [18] was reduced. Based on this analysis tool, Niemann [14] proposed a controller discretization, which minimizes the distance (induced norm) between the continuous-time controller and the center of the conic sector associated with the discrete-time controller. The consequences of this optimization are not clear, since also either the cone radius [18] or the cone weighting factor [19] are a function of the discrete-time controller. As a consequence, closed-loop stability has to be checked after controller determination.

In the following, a new discretization method is presented, which takes account of closed-loop issues. Using ideas from perturbation theory, stability of the closed-loop system is related to the approximation error, resulting from controller discretization, much as in some of the order reduction methods of Anderson and Liu [1]. The resulting stability criterion is derived in Section II. The discrete-time controller is chosen to maximize the stability margin. To evaluate this margin, the value of an operator norm is required and this is hard to calculate. In order to get a tractable problem, the original criterion is approximated (arbitrarily closely) by an auxiliary criterion. After some transformations an $H_\infty$ characterization of a stabilizing discrete-time controller, which maximizes the stability margin and approximates the continuous-time controller, is obtained. This is shown in Sections III and IV. In Section V possible applications of the discrete-time controller characterization are proposed. In Section VI examples are presented and Section VII contains concluding remarks.

## II. Stability of Discretized Control Systems

Let $P(s)$ be the plant transfer function of a possibly multivariable plant, $C(s)$ the known continuous-time controller, $C_d(z)$ the unknown discrete-time controller with given antialiasing filter $F_a(s)$ and let $H(s)$ represent a zero-order hold. Fig. 1 shows a rearrangement of the system with the discrete-time controller. We shall assume here that $C(s)$ is open-loop stable (as in common)—removal of this assumption would presumably be possible, as discussed briefly in the conclusions. It is natural to seek to approximate $C(s)$ using a $C_d(z)$ which is also stable, and we make this assumption. In addition we assume that $C(s)$ is a stabilizing controller for $P(s)$, and for convenience, that $P(s)$ is strictly proper. Last, we assume that $P(s)$, $C(s)$, and $F_a(s)$ are rational in $s$.

The symbol $\Delta(s, z = e^{sT}, t)$ is used for the approximation-error operator, which is linear and periodically time variant; it cannot be represented by a real-rational transfer function. To investigate the influence of $\Delta$ on closed-loop

Manuscript received November 24, 1989; revised July 10, 1991. Paper recommended by Associate Editor, D. F. Delchamps.

The authors are with the Department of Systems Engineering, Research School of Physical Sciences and Engineering, Australian National University, Canberra, A.C.T., Australia.

IEEE Log Number 9105124.

0018-9286/92$03.00 

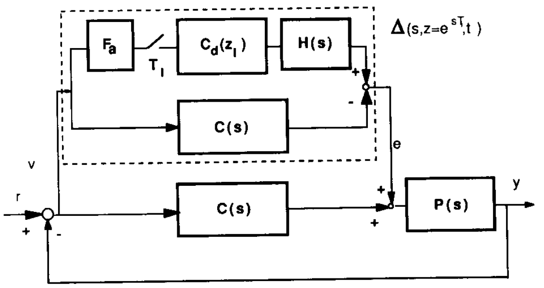


Fig. 1.

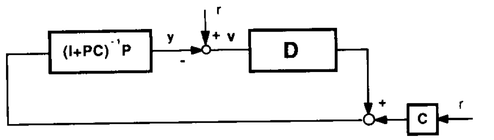


Fig. 2.

stability an equivalent representation to Fig. 1 is drawn in Fig. 2.

Let

$$J_c \triangleq \Delta (I + PC)^{-1} P \tag{2.1}$$

$$\| J_c \| = \max_{u \in L_2[0, \infty)} \frac{\| J_c u \|_2}{\| u \|_2}. \tag{2.2}$$

Notice that in $J_c$ both $\|\Delta\|$ and $\|(I + PC)^{-1}P\|$ are bounded. This is obvious for the second term, i.e., the closed-loop formed from $P$ and $C$. In the first term the antialiasing filter $F_a$ ensures that the mapping of $F_a$ followed by the sampler has a finite gain. Since the controllers $C(s)$ and $C_d(z_l)$ are both stable and $H(s)$ has a finite gain, there results boundedness of $\|\Delta\|$. Therefore, also $\| J_c\|$ is finite.

By a variant on the small-gain theorem, see [2, p. 46], the loop in Fig. 2 is bounded-input bounded-output stable and has a mapping from $r \in L_2[0, \infty)$ to $y \in L_2[0, \infty)$ if

$$\| J_c \| < 1. \tag{2.3}$$

Using the definition of $J_c$, it becomes clear how controller approximation relates to stability, and how controller approximation might be pursued. To guarantee stability, the approximation must ensure that (2.3) holds, i.e., that $\Delta$, when weighted by $(I + PC)^{-1}P$, is suitably small in norm. It is reasonable to seek $C_d$ not merely to ensure that (2.3) holds, but to ensure that $\| J_c\|$ is minimized.

It is not surprising that $\| J_c\|$ is hard to calculate, however, in the next two sections, we develop a convenient approximation to $\| J_c\|$ which leads to a useful characterization of stabilizing discrete-time controllers. Not only can we calculate for a given $C_d(z)$ a convenient approximation to $\| J_c\|$ but we can then choose $C_d$ to minimize this approximation error index.

## III. Approximation of $J_c$

The main idea of this section is to approximate the continuous-time operators involved in forming $J_c$, i.e., $(I + PC)^{-1}P$, $F_a$, $H$, and $C$, by discrete-time operators with arbitrarily small sampling time which is a submultiple of the sampling time of $C_d$. It turns out that the operator $J_d$, obtained by replacing the continuous-time operators with their discrete-time hold-input approximation, converges to $J_c$ as their sampling time (but not that for $C_d$) tends to zero. The approximation idea is illustrated in Fig. 3, which depicts $J_c$, and Fig. 4, which depicts a multirate sampled data system $J_d$, which is an approximation of $J_c$.

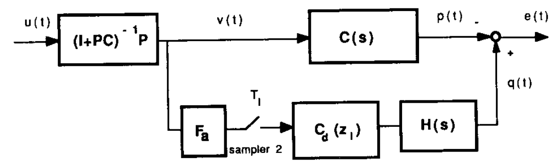


Fig. 3.

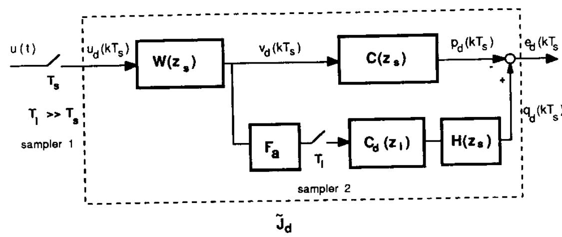


Fig. 4.

Two different sampling times, integrally related, appear now in $J_d$. Discrete-time transfer functions mapping signals sampled in intervals $T_s$, respectively, $T_l$ will be denoted by the variable $z_s$, respectively, $z_l$.

In the next section, we shall show how a further step reduces the calculation of $\| J_d\|$ to a standard problem. We now turn to justifying the claim that $\| J_d\|$ approximates $\| J_c\|$. Let $W(s)$ denote $(I + PC)^{-1}P$ and suppose $(A_i, B_i, C_i, D_i)$, $i = \{W, C, F_a\}$ are minimal state-space realizations of $W(s)$, $C(s)$, and $F_a(s)$. By assumption $D_W$ and $D_{F_a}$ are zero. Further, let

$$C_d(z_l) = E_d + H_d(z_l I - F_d)^{-1} G_d \tag{3.1}$$

be a minimal realization of $C_d(z_l)$. The sampling rate for $C_d(z_l)$ is $T_l$ and that for the fast sampled data approximations of $W(s)$, $C(s)$, and $F_a$ is assumed to be $T_s$, where

$$T_l = NT_s \tag{3.2}$$

for some integer $N$, which will be chosen large. The discrete-time replacements of the continuous-time transfer functions are their hold-input discretization with minimal realization $(F_i, G_i, H_i, E_i)$, $i = \{\hat{W}, \hat{C}, \hat{F}_a\}$:

$$F_i = e^{A_i T_s}, \quad G_i = \int_0^{T_s} e^{A_i \tau} B_i \, d\tau,$$
$$H_i = C_i, \quad E_i = D_i. \tag{3.3}$$

The function of the block labeled $H(z_s)$ (see Fig. 4) is easily described. $H$ receives a discrete-time input at time intervals $T_l$ s apart. Each discrete-time input pulse generates $N$ discrete-time output pulses, equal to the input pulse, starting at the time the input pulse is received and spaced apart by $T_s$ s.

*Proposition 3.1:* Consider the operator $J_c$ mapping $u(\cdot) \in L_2[0,\infty)$ to $e(\cdot) \in L_2[0,\infty)$ depicted in Fig. 3 and the operator $J_d$ mapping $u_d(kT_s)$, $k = 0,1,2,\cdots, \in l_2[0,\infty)$ to $e(kT_s)$, $k = 0,1,2,\cdots, \in l_2[0,\infty)$ depicted in Fig. 4, with the relation between the blocks making up $J_d$ and $J_c$ as described above. Then with $T_l$ fixed, there holds

$$\lim_{N\to\infty} \|J_d\| = \|J_c\|, \qquad NT_s = T_l. \tag{3.4}$$

For the proof of this result, see Appendix A.

The intuition is clear; if we sample fast enough and use conventional discretization of the continuous-time blocks, we obtain a multirate discrete-time system with the same gain (in the limit) as the hybrid system.

## IV. Evaluation of $J_d$ as a Standard Problem

Two different sampling times appear in the operator $J_d$. As such, it is not immediately clear how $J_d$ should be represented by a transfer function (matrix). However, a simple idea allows $J_d$ to be represented by a norm-equivalent transfer function with respect to the larger sampling time. The idea has been used in the design of periodically time-varying controllers [9], and is an old one in signal processing, where it is termed "blocking" (see [13]).

Consider for example the transfer function

$$\hat{C}(z_s) = E_c + H_c(z_sI - F_c)^{-1}G_c. \tag{4.1}$$

The processing achieved by $\hat{C}(z_s)$ can be viewed in an alternative way. One groups $N$ successive inputs into an input block and similarly for the output. Input and output blocks occur at intervals of $NT_s = T_l$, and the processing is viewed as mapping input blocks into output blocks. These inputs and output blocks are vectors of dimension $N$ in case $\hat{C}(z_s)$ is a scalar transfer function, and are vectors of vectors in case $\hat{C}(z_s)$ is a transfer function matrix. A state-space representation is easily obtained as

$$x_c[(k+1)NT_s] = F_c^N x_c(kNT_s) + \begin{bmatrix} F_c^{N-1}G_c & F_c^{N-2}G_c & \cdots & G_c \end{bmatrix} \cdot \begin{bmatrix} v_d(kNT_s) \\ v_d[(kN+1)T_s] \\ \vdots \\ v_d[(kN+N-1)T_s] \end{bmatrix}$$

$$\begin{bmatrix} y_d(kNT_s) \\ y_d[(kN+1)T_s] \\ \vdots \\ y_d[(kN+N-1)T_s] \end{bmatrix} = \begin{bmatrix} H_c \\ H_cF_c \\ \vdots \\ H_cF_c^{N-1} \end{bmatrix} x_c(kNT_s) + \begin{bmatrix} E_c & 0 & \cdots \\ H_cG_c & E_c & \ddots \\ H_cF_cG_c & H_cG_c & \ddots \\ \vdots & & \\ H_cF_c^{N-2}G_c & H_cF_c^{N-3}G_c & \cdots \end{bmatrix} \cdot \begin{bmatrix} v_d(kNT_s) \\ v_d[(kN+1)T_s] \\ \vdots \\ v_d[(kN+N-1)T_s] \end{bmatrix}. \tag{4.2}$$

A similar representation can be obtained for $\hat{W}(z_s)$ and $\hat{F}_a(z_s)$. Evidently $\hat{C}(z_s)$ can be replaced by transfer function matrices in the variable $z_l$, e.g.,

$$\tilde{C}(z_l) = \tilde{E}_c + \tilde{H}_c(z_lI - \tilde{F}_c)^{-1}\tilde{G}_c \tag{4.3}$$

where $\tilde{F}_c$, $\tilde{G}_c$, $\tilde{H}_c$, and $\tilde{E}_c$ are the matrices appearing in (4.2) with $\tilde{W}(z_l)$ and $\tilde{F}_a(z_l)$ being constructed similarly. We denote the input and output vectors in (4.2) by $\tilde{v}_d(kT_l)$ and $\tilde{y}_d(kT_l)$.

In the block representation, the output of the sampler 2 (see Fig. 4) is the first element of the vectorized signal $\tilde{v}_d$. The transfer matrix of the sampler is therefore a row vector of length $N$, having 1 as first element and being zero elsewhere. The hold element $H(z_l)$ produces a sequence of $N$ pulses equal to the input pulse which corresponds to a block vector with equal entries. It is therefore a mapping of the output of $C_d(z_l)$ by a column vector of ones.

Then the block rearrangement of $\tilde{J}_d$ is the time-invariant operator with transfer function matrix

$$\tilde{J}_d(z_l) = \left(\tilde{C}(z_l) - \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} C_d(z_l)[1\,0\cdots 0]\tilde{F}_a(z_l)\right)\tilde{W}(z_l). \tag{4.4}$$

For multivariable systems the 1 and zero entries in the column and row vectors are replaced by identity and zero matrices.

Observe also that

$$\sum_k u_d'(kT_s)u_d(kT_s) = \sum_k \tilde{u}_d'(kT_l)\tilde{u}_d(kT_l) \tag{4.5}$$

$$\sum_k e_d'(kT_s)e_d(kT_s) = \sum_k \tilde{e}_d'(kT_l)\tilde{e}_d(kT_l). \tag{4.6}$$

Hence the connection between $J_d$, mapping $u_d(\cdot)$ to $e_d(\cdot)$ and $\tilde{J}_d$, mapping $\tilde{u}_d(\cdot)$ to $\tilde{e}_d(\cdot)$ ensures

$$\|J_d\| = \|\tilde{J}_d\| \tag{4.7}$$

where the norms are the operator-induced norms on $l_2^p[0,\infty)$ and $l_2^{pN}[0,\infty)$. However since $\tilde{J}_d$ has a transfer matrix representation, this operator norm is precisely the $H_\infty$-norm of the transfer function matrix.

This result and that of the last section have shown how, through approximation and rearrangement, a mathematically tractable (i.e., readily computable) criterion for controller discretization is obtained, at least approximately.

*Theorem 4.1:* With quantities as defined earlier

$$\| J_c \| = \lim_{T_s \to 0} \| J_d \| = \lim_{T_s \to 0} \| \tilde{J}_d \| = \lim_{T_s \to 0} \| \tilde{J}_d(e^{j\omega T_l}) \|_\infty. \tag{4.8}$$

In the above analysis, we have assumed that the output of sampler 2 is the first element of a block signal. We could however have chosen to define the block signal associated with $v_d(\cdot)$ by

$$\tilde{v}_d(kT_l) = \begin{bmatrix} v_d[(kN+\alpha)T_s] \\ v_d[(kN+1+\alpha)T_s] \\ \vdots \\ v_d[(k+1)NT_s] \\ \vdots \\ v_d[((k+1)N-1+\alpha)T_s] \end{bmatrix} \tag{4.9}$$

and similarly for $\tilde{y}_d(kT_l)$. Thus the assembling of the fast signals into blocks, while still occurring every $T_l$ s, occurs at times displaced from the sampling instants of the sampler 2 and the time of generation of outputs from the slow compensator. The effect is to vary $\tilde{J}_d$ but, as it turns out, not to vary $\| \tilde{J}_d \|$.

## V. Extremizing the Stability Measure

Different measures for the determination of a discrete-time controller have been proposed [4]–[8], but most of them, except in Rattan [16], consider only performance degradation at sampling instants. Because $\tilde{J}_d$ approximates the continuous-time approximation error arbitrarily closely, intersampling behavior of the system is also reflected in the error measure used for controller discretization. The resulting advantages are that the occurrence of intersampling ripple is minimized and that therefore larger sampling periods may be used while still achieving acceptable intersampling behavior of the continuous-time signals (e.g., output $y$).

By means of the representation (4.4), several discretization problems can be solved:

- Given a sampling time $T_l$ an optimal controller $C_d(z_l)$ can be obtained through minimization of $\| \tilde{J}_d \|_\infty$. The resulting controller approximates the continuous-time one and is stabilizing if $\| \tilde{J}_d \|_\infty < 1$ (for $T_s$ reasonably small).
- Using the same approach, there is also the possibility of approximating a discrete-time controller by another stabilizing discrete-time controller with larger sampling time.
- A measure for the impact of the sampling time on controller discretization is the value of $\| \tilde{J}_d \|_\infty$ for an optimal controller $C_d(z_l)$ and the effect of varying sampling time can be easily examined. If for an optimal controller $\| \tilde{J}_d \|_\infty = 1$ an upper bound for the sampling period $T_l$ based on a sufficient condition for stability is reached.

The minimization of $\| \tilde{J}_d \|_\infty$ is *a standard $H_\infty$ problem*, when $C(s)$ is stable (which it is by assumption). Controllers $C_d$ were determined using the minimization procedure described in [17]. The $H_\infty$-problem generally has no unique solution. One reason becomes evident from a closer look at the underlying 4-block problem of achieving $\inf \| \tilde{J}_d \|_\infty$

$$\inf_{C_d} \| \tilde{J}_d \|_\infty = \inf_{C_d} \left\| \begin{bmatrix} A - C_d M & B \\ C & D \end{bmatrix} \right\|_\infty. \tag{5.1}$$

A lower bound $l_b$ for $\| \tilde{J}_d \|_\infty$ is given by

$$l_b = \max \left( \left\| \begin{bmatrix} B \\ D \end{bmatrix} \right\|_\infty, \| [C \quad D] \|_\infty \right). \tag{5.2}$$

The examples show that for small sampling times $T_l$, this lower bound is usually reached. Safonov's solution involves the determination of an antistable approximation of a stable transfer function with approximation error norm smaller or equal than a given value, i.e., 1. If $\inf \| J_d \|_\infty = l_b$, the antistable approximation achieving the specified error bound is usually not unique.[1] This results in a set of solutions with equal $\| J_d \|_\infty$. Within that set $C_d$ can be chosen for example either to be of minimal order or to shape $\bar{\sigma}(\tilde{J}_d(j\omega))$ in a desirable way, without affecting $\| \tilde{J}_d \|_\infty$. An alternative way for the determination of an unique solution (more suitable for MIMO systems) may be the application of super optimal controller design proposed by Foo and Postlethwaite [4], but the advantages are not clear.

## VI. Example

The new discretization method was tested with two examples, the first from Rattan [16] and the second from Katz [7]. For both examples, it was reported that common discretization methods produce either nonstabilizing controllers or systems with very poor closed-loop performance. Using the proposed discretization method of this paper, controllers were calculated for various sampling times and the resulting control systems were compared to the results of Rattan [16] and Kennedy and Evans [8], who further investigated the example of Katz [7]. Since the authors are not using antialiasing filters, $F_a$ was set to the identity which is an approximation of $F_a = a/s + a$ for $a$ arbitrarily large.

The following particular aspects were investigated:

- step response of the closed-loop system
- the frequency response to input signals $\boldsymbol{r}$
- gain and phase margin of the discrete-time system
- possible increase of the sampling time above the values, proposed in [16] and [7].

For calculating a discrete-time controller, i.e., minimization of $\| \tilde{J}_d \|_\infty$, the small sampling time $T_s$ must be chosen. As a rule of thumb, the small sampling period $T_s$ needs to be about 20 times smaller than a sampling time determined from closed-loop bandwidth considerations. (See [5]; the sampling frequency should be more than twice the closed-loop bandwidth of the continuous-time system.) Smaller $T_s$ alter $\| \tilde{J}_d \|_\infty$

[1] Since the optimal antistable approximation leads to an error norm smaller than 1.

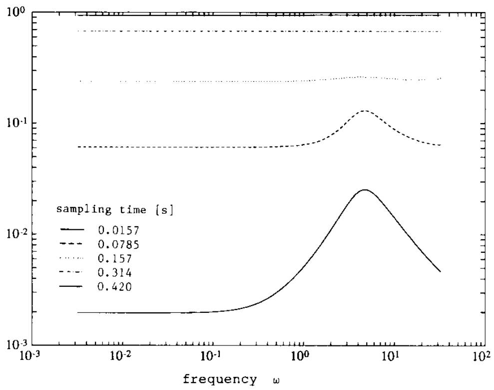



Fig. 5.

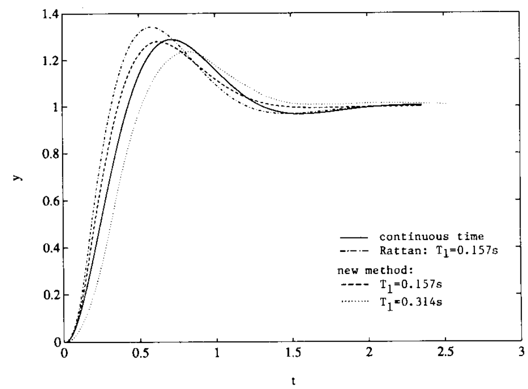



Fig. 6.

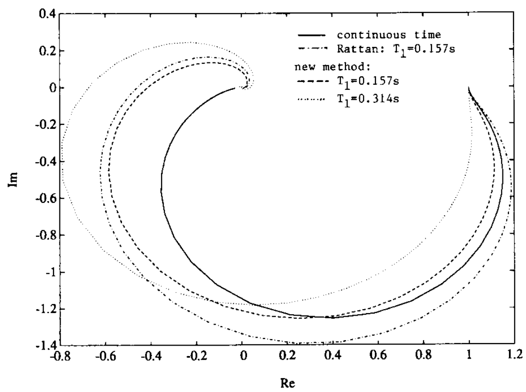



Fig. 7.

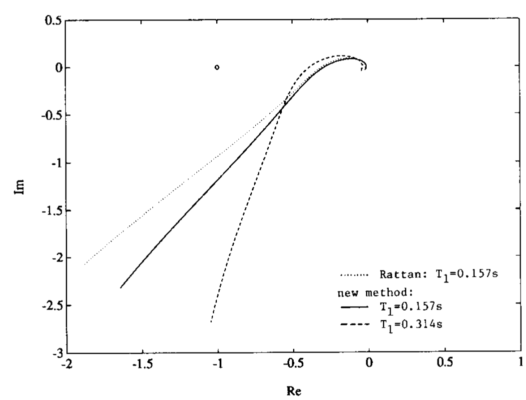



Fig. 8.

only negligibly, but increase computational burden.

The transfer functions of the plants and controllers are given in Appendix B.

### *A. Rattan's Example*

In Fig. 5 the approximation error $\bar{\sigma}(\tilde{J}_d(j\omega))$ is plotted for optimal controllers designed for different sampling periods $T_l$. The plot shows, that $T_l = 0.420s$ is close to the proposed upper bound. A further increase in sampling time leads to $\|\tilde{J}_d\|_\infty > 1$ and for a controller with $\|\tilde{J}_d\|_\infty = 1.1$ an unstable closed-loop system was obtained. For the controllers with sampling time $T_l \le 0.157s$, $\|J_d\|_\infty$ is determined by (5.2) and a relatively large flexibility is available to determine optimal reduced-order controllers. The order reduction leads to the characteristic that the optimal $\tilde{J}_d$ is not all-pass, i.e., $\bar{\sigma}(\tilde{J}_d(j\omega)) \neq$ const.. For $T_l > 0.157s$ a negligible suboptimality was allowed to obtain a reasonable controller order and also to circumvent the problem of finding a minimal controller realization. The sampling time proposed by Rattan is 0.157 s. In Figs. 6 and 7, the step response and the frequency response are plotted. The continuous-time system and Rattan's optimization result with $T_l = 0.157s$ are compared to our results for $T_l = 0.157s$ and $T_l = 0.314s$. The plots reveal that in this example performance is maintained which is comparable to Rattan's optimization, even for a sampling time which is twice Rattan's sampling time. Gain and phase margin of the discrete-time system are analyzed using a Nyquist diagram. From Fig. 8 it can be concluded that both gain and phase margin for $T_l = 0.157s$ are similar to Rattan's result, whereas for $T_l = 0.314s$ gain margin is slightly reduced.

The continuous-time controller has degree 1, the discrete-time controller of Rattan has degree 1, and those generated by our method have degree 2.

### *B. Katz's Example*

According to the discussion in Katz [7], of 8 commonly used discretization methods only one, *viz.* prewarped bilinear transformation yields a stabilizing controller at $T_l = 0.030s$ and even then there are very poor closed-loop properties. Kennedy and Evans [8] also investigated this example and achieved good controller discretization. They proposed a procedure in which the continuous-time feedback controller is approximated by a discrete-time model following controller, i.e., one having a feedforward and a feedback part. Our controller-discretization method will be compared to this result.

In Fig. 9 $\bar{\sigma}(\tilde{J}_d(j\omega))$ is plotted. The controller order was determined in the same way as in the preceding example. The sampling-time $T_l = 0.030s$ is close to the proposed upper bound, which is near to 0.039 s. The relatively large value of $\min\|\tilde{J}_d\|_\infty$ for $T_l = 0.030s$ predicts therefore the discretization problems reported in [7]. The step response and the

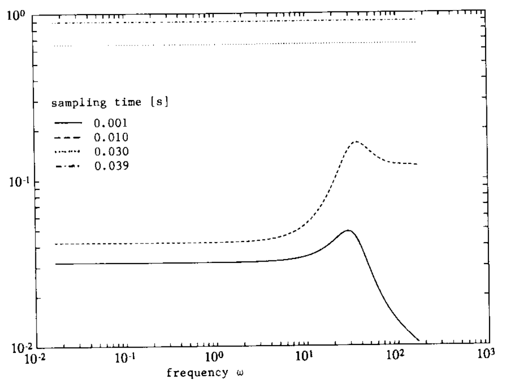



Fig. 9.

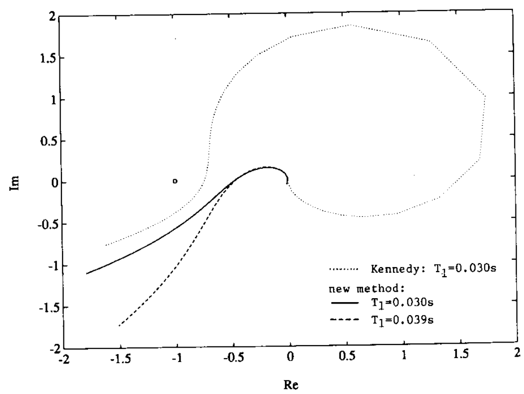



Fig. 12.

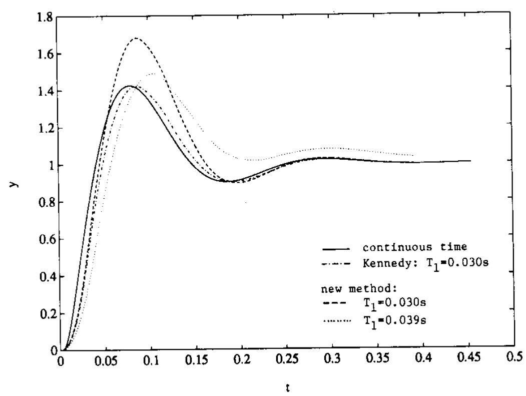



Fig. 10.

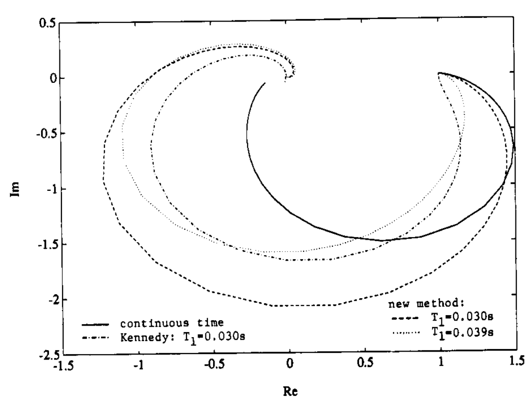



Fig. 11.

frequency response are shown in Figs. 10 and 11. The design of Kennedy for $T_l = 0.030s$ shows better agreement with the continuous-time step response. This is not surprising because he makes use of the additional freedom of a feedforward controller. In the frequency-response plot, it can be seen that the design of Kennedy is better at high frequencies. It is remarkable that our controller with $T_l = 0.039s$ still gives acceptable closed-loop performance. A more important difference between our design and the design of Kennedy can be seen in Fig. 12. The Nyquist plot of the discrete-time loop-transfer function shows that the gain and phase margin of our design are considerably better than those of Kennedy even for the sampling-time $T_l = 0.039s$. The continuous-time controller, the controller of Kennedy and our controller for $T_l = 0.030s$ are second-order while our controller for $T_l = 0.039s$ is first-order.

## VII. Conclusion

Our proposed method for controller discretization offers the advantage that it focuses on closed-loop behavior. It is based on principles of continuous-time controller approximation and includes therefore also intersampling behavior of the system. The examples show that the value assumed by our approximation criterion can be interpreted as a possible measure for the practical cost of discretization in terms of performance degradation. Furthermore, the approximation criterion provides an upper bound on the sampling time, based on a sufficient stability criterion. The method works well with the selected examples, which were reported to be difficult to discretize. Even for relatively large sampling time, the results are comparable to earlier proposed discretizations operating at a smaller sampling time. The determination of the discrete-time controller is straightforward and can be performed with standard software used in $H_\infty$ controller design.

Several aspects of the proposed discretization method need to be investigated. When the optimal $J_d$ is all-pass it may be possible to derive an exact value for the order of $C_d(z_l)$ in terms of the orders of $C(s)$, $P(s)$, and $F_a$. This value is an upper bound in the case when the controller is determined as described in Appendix B. Another interesting point is how the resulting controllers relate to controllers obtained with conventional discretization methods, especially for extreme values of the sampling time. Two questions arise in this context: are the conventionally discretized controllers in the set of optimal controllers? and if not: how far away are they from optimality with respect to the criterion $\| J_d \|_\infty$?

In the examples presented, nonuniqueness of controllers was used to minimize controller order. It is reasonable to investigate, if some other auxiliary criterion, possibly based on performance considerations, may be preferable to force a unique controller.

Using the proposed formalism, it may be possible to formulate a controller approximation problem, including other performance oriented weighting functions than stability margin as proposed for controller order reduction in [1]. Also, fractional representations of controllers may be approximated [12] allowing the discretization of unstable controllers.

The approximation of continuous-time controllers by multirate discrete-time controllers (e.g., multivariable systems with samplers, operating with different sampling times) can be easily formalized using the block representation $\tilde{J}_d$, but for controller determination, boundary conditions ensuring causality and time-invariance must be respected. Neglecting the last mentioned constraint, time-varying multirate controllers could be obtained.

It is desirable to gain further insight into block representation optimization problems like $\min\|\tilde{J}_d\|$. As a result, sampling time could probably be related in a quantitative way to performance degradation resulting from system discretization and therefore provide a clear procedure for sampling time determination. Also the efficiency of more complicated control systems, for example multirate systems could be examined.

It is also possible to impose a constraint that there be some processing delay in the control $C_d(z_l)$; this requires some development and will be explained elsewhere.

## Appendix A
## Proof of Proposition 3.1

*Step 1:* We shall show that

$$\lim_{N\to\infty}\|J_d\| \geq \|J_c\|. \tag{A.1}$$

*Lemma A.1:* For arbitrarily small $\epsilon > 0$, there exists a piecewise constant function $u(\cdot)$, with points of discontinuity at $kT_o$, $k = 1, 2, \cdots$, where $T_0$ is a submultiple of $T_l$, such that with $e_u(\cdot)$ the response of $J_c$ to $u(\cdot)$

$$\|u\|_2 = 1 \quad \|e_u\|_2 \geq \|J_c\| - \epsilon. \tag{A.2}$$

*Proof:* Define

$$\epsilon_1 \hat{=} \frac{\epsilon}{5 - \dfrac{2\epsilon}{\|J_c\|}} > 0. \tag{A.3}$$

Then there exists $v \in L_2[0, \infty)$ with $\|v\|_2 = 1$ such that

$$\|e_v(\cdot)\|_2 \geq \|J_c\| - \epsilon_1. \tag{A.4}$$

By the properties of Lebesgue integration, there exists a piecewise constant function, call it $w(\cdot)$, such that

$$\|w - v\|_2 \leq \epsilon_1 / \|J_c\|. \tag{A.5}$$

Evidently

$$e_w = J_c w = J_c v + J_c(w - v) = \epsilon_v + J_c(w - v) \tag{A.6}$$

so that

$$\begin{aligned}\|e_w\|_2 &\geq \|e_v\|_2 - \|J_c\|\,\|w - v\|_2 \\ &\geq \|J_c\| - 2\epsilon_1.\end{aligned} \tag{A.7}$$

Now let $x(\cdot)$ be a second piecewise constant function obtained by shifting the points of discontinuity of $w(\cdot)$ to numbers rationally related to $T_l$.

Let $t_1$ be such that

$$\left[\int_{t_1}^{\infty} w'(t)w(t)\,dt\right]^{1/2} \leq \frac{\epsilon_1}{2\|J_c\|} \tag{A.8}$$

and clearly, the points of discontinuity of $x(\cdot)$ in the interval $[0, t_1]$ can be chosen so that

$$\int_0^{t_1}[x(t) - w(t)]'[x(t) - w(t)]\,dt]^{1/2} \leq \frac{\epsilon_1}{2\|J_c\|}. \tag{A.9}$$

Then with $x(t) = 0$ for $t > t_1$

$$\|x - w\|_2 \leq \frac{\epsilon_1}{\|J_c\|}. \tag{A.10}$$

From (A.8) to (A.10)

$$\|e_x(\cdot)\|_2 \geq \|J_c\| - 3\epsilon_1. \tag{A.11}$$

Now there exists a $T_0$ which is a submultiple of $T_l$ such that $x(\cdot)$ is constant on half open intervals $[kT_0, (k+1)T_0)$ for any integer $k$. Define

$$u(\cdot) = \frac{x(\cdot)}{\|x(\cdot)\|_2}. \tag{A.12}$$

Clearly, $\|u\|_2 = 1$ and $u(\cdot)$ has the piecewise constancy property of the lemma statement. Also

$$\|x(\cdot)\|_2 \leq \|x - w\|_2 + \|w - v\|_2 + \|v\|_2 \leq 1 + \frac{2\epsilon_1}{\|J_c\|}.$$

So then

$$\|e_u(\cdot)\|_2 \geq \frac{\|e_x(\cdot)\|_2}{1 + \dfrac{2\epsilon_1}{\|J_c\|}} \geq \frac{\|J_c\| - 3\epsilon_1}{1 + \dfrac{2\epsilon_1}{\|J_c\|}} = \|J_c\| - \epsilon$$

[by simple algebra, using (A.3)]. This proves the lemma.

Let us now drop the subscript $u$ on $e_u(\cdot)$. In order to establish (A.1), we shall select a discrete-time signal $u_d(kT_s)$, $k = 0, 1, \cdots$, for $J_d$ of unit norm and show that the resulting output $e_d(kT_s)$ has norm exceeding $\|J_c\| - (L+1)\epsilon$ for some integer $L$ and arbitrarily small $\epsilon > 0$. We shall later identify $T_s$ with a submultiple of $T_0$, the latter having been identified in Lemma A.1. Obviously, $u(\cdot)$ identified in the Lemma is constant on half open intervals $[kT_s, (k+1)T_s)$, no matter what this interval is. We shall chose $u_d(\cdot)$ by requiring

$$u_d(kT_s) = u(kT_s) \qquad \forall k. \tag{A.13}$$

We shall also define $l_2[0, \infty]$-norms by introducing a scaling factor, thus

$$\|u_d(\cdot)\|_2 = \sqrt{T_s}\left[\sum_{k=0}^{\infty} u_d'(kT_s)u_d(kT_s)\right]^{1/2}. \tag{A.14}$$

Observe then that

$$\|u_d(\cdot)\|_2 = \|u(\cdot)\|_2 = 1. \tag{A.15}$$

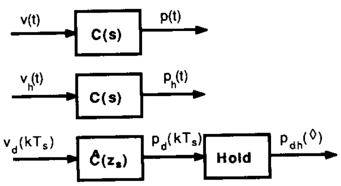



Fig. 13.

Now consider Figs. 3 and 4 and observe that, with $v(\cdot)$ the response of $W(s) = (I + PC)^{-1}P$ to $u(\cdot)$ and $v_d(\cdot)$ the response of the block $\hat{W}(z_s)$ to $u_d(\cdot)$, assuming zero initial conditions in both cases, the relation between the discrete- and continuous-time blocks ensures that for all $T_s$ which are submultiples of $T_0$

$$v_d(kT_s) = v(kT_s). \tag{A.16}$$

Now in preparation for the next lemma, let us define, with reference to Fig. 13, three new continuous-time signals

$$v_h(t) = v(kT_s) \qquad \text{for } t \in [kT_s, (k+1)T_s) \tag{A.17}$$

$$p_h(\cdot) = \text{response of } C(s) \text{ to } v_h(\cdot), \text{ with zero initial conditions} \tag{A.18}$$

$$p_{dh}(\cdot) = \text{result of applying a hold operation to } p_d(kT_s). \tag{A.19}$$

The subscript $h$ denotes "hybrid." Notice that, because of the way $\hat{C}(z_s)$ is defined, there holds

$$p_h(kT_s) = p_d(kT_s) = p_{dh}(kT_s). \tag{A.20}$$

The signals $q_h(\cdot)$ and $q_{dh}(\cdot)$ are defined in the same with

$$q_h(kT_s) = q_d(kT_s) = q_{dh}(kT_s). \tag{A.21}$$

Our aim in the next Lemma is to show closeness of $(p(\cdot) - q(\cdot))$ and $(p_{dh}(\cdot) - q_{dh}(\cdot))$. This will follow because: i) closeness of $v(\cdot)$ and $v_h(\cdot)$ ensures closeness of $(p(\cdot) - q(\cdot))$ and $(p_h(\cdot) - q_h(\cdot))$ and ii) the relation between the continuous-time transfer functions and their discretizations and their inputs ensure closeness of $(p_h(\cdot) - q_h(\cdot))$ and $(p_{dh}(\cdot) - q_{dh}(\cdot))$.

*Lemma A.2:* With quantities defined as above, there exists a suitably small submultiple $T_s$ of $T_0$ such that

$$\|(p(\cdot) - q(\cdot)) - (p_{dh}(\cdot) - q_{dh}(\cdot))\|_2 \le \epsilon. \tag{A.22}$$

*Proof:* Since $v(\cdot)$ is the output of the block $W(s)$ driven by $u(\cdot)$, there holds, for $t \in [kT_s, (k+1)T_s)$

$$\begin{aligned} v(t) - v_h(t) &= v(t) - v(kT_s) \\ &= C_w(e^{A_w(t-kT_s)} - I)x(kT_s) \\ &\quad + C_w \int_0^{t-kT_s} e^{A_w\tau} B_w u(kT_s + \tau)\, d\tau. \end{aligned} \tag{A.23}$$

Since $\|u_d(\cdot)\|_2$ and $\|u_d(\cdot)\|_\infty$ are bounded, an easy calculation shows that there exists a $T$, so that for all $T_s \le T$, there holds

$$\|v(\cdot) - v_h(\cdot)\|_2 \le \epsilon/3K \tag{A.24}$$

where $K$ is $\max\{\|C\|, \|HCSF_a\|\}$, the $L_2$-gain of $C$ or $HCSF_a$ (which were shown to be finite).

Now it follows that for the same $T_s$

$$\|p(\cdot) - p_h(\cdot)\|_2 \le \epsilon/3. \tag{A.25}$$

Now $v_d(kT_s)$ is the set of sampled values of the piecewise constant signal $v_h(\cdot)$, and $C(z_s)$ is constructed from $C(s)$ by a standard discretization operation. Hence the same argument that led to (A.24) also show that $T_s$ suitably small

$$\|p_h(\cdot) - p_{dh}(\cdot)\|_2 \le \epsilon/3 \tag{A.26}$$

and therefore

$$\|p(\cdot) - p_{dh}(\cdot)\|_2 \le \epsilon 2/3. \tag{A.27}$$

Since $q_h(\cdot)$ is the output of the hold element $H(s)$, there is

$$q_h(\cdot) = q_{dh}(\cdot) \tag{A.28}$$

which results in

$$\|q(\cdot) - q_{dh}(\cdot)\|_2 \le \epsilon 1/3. \tag{A.29}$$

Combining the last two inequalities, the inequality (A.22) is immediate.

Finally, we can pin down a lower bound for $\|J_d\|$. Using (A.22), we can put a lower bound on the norm of the response $e_d(\cdot)$ of $J_d$ to $u_d(\cdot)$, which has been chosen to have unit norm.

*Lemma A.3:* With quantities as defined earlier, and with $T_s$ a sufficiently small submultiple of $T_o$ and thus of $T_l$, there holds

$$\|e_d(kT_s)\|_2 \ge \|J_c\| - 2\epsilon \tag{A.30}$$

and, since $\|u_d(kT_s)\|_2 = 1$

$$\|J_d\| \ge \|J_c\| - 2\epsilon. \tag{A.31}$$

*Proof:* Observe from Figs. 3 and 4

$$e(t) = q(t) - p(t)$$
$$e_d(kT_s) = q_d(kT_s) - p_d(kT_s).$$

Defining

$$e_{dh}(\cdot) \triangleq q_{dh}(\cdot) - p_{dh}(\cdot).$$

Then

$$e(\cdot) - e_{dh}(\cdot) = (q(\cdot) - q_{dh}(\cdot)) - (p(\cdot) - p_{dh}(\cdot)). \tag{A.32}$$

Moreover

$$\begin{aligned} \|e_d(kT_s)\|_2 &= \|e_{dh}(\cdot)\|_2 \\ &\ge \|e(\cdot)\|_2 - \|e(\cdot) - e_{dh}(\cdot)\|_2. \end{aligned}$$

Now use Lemmas A.1 and A.2 and (A.32). There follows

$$\begin{aligned} \|e_d(kT_s)\|_2 &\ge \|J_c\| - \epsilon - \epsilon \\ &= \|J_c\| - 2\epsilon. \end{aligned}$$

Since $\epsilon$ is arbitrary, and noting that $T_s$ depends on $\epsilon$, with $NT_s = T_l$, the result (A.1) is immediate.

*Step 2:* We shall show that

$$\|J_c\| \geq \lim_{N\to\infty} \|J_d\|. \tag{A.33}$$

Choose an arbitrarily small $\epsilon > 0$. Then we may choose $T_s$ sufficiently small such that, for *all* signals $u(\cdot)$ which are piecewise constant on intervals of length $T_s$ and have $\|u(\cdot)\|_2 = 1$, there holds (by the same argument as in the proof of Lemma A.2) the inequalities (A.24), (A.25), (A.27), and (A.29). Hence (A.22) holds for all such $u(\cdot)$.

Now we use a modified form of the argument in the proof of Lemma A.3. As before (A.32) holds. Hence

$$\begin{aligned}\|e(\cdot)\|_2 &\geq \|e_{dh}(\cdot)\|_2 - \|e(\cdot) - e_{dh}(\cdot)\|_2 \\ &\geq \|e_d(kT_s)\|_2 - \epsilon.\end{aligned} \tag{A.34}$$

Now choose $u(\cdot)$ by choosing first $u_d(kT_s)$ to have unit norm and to produce $e_d(kT_s)$ with

$$\|e_d(kT_s)\| \geq \|J_d\| - \epsilon \tag{A.35}$$

(which can always be done for suitably small $T_s$); of course, as before we ensure that

$$u(t) = u_d(kT_s) \qquad kT_s \leq t < (k+1)T_s.$$

Then $\|u(\cdot)\|_2 = 1$, $\|e(\cdot)\|_2 \geq \|J_d\| - 2\epsilon$, from which A.33 follows. Proposition 3.1 is an immediate consequence of A.33 and A.1.

## Appendix B
## Controller Data

*Transfer Functions of Rattan's Example*

*Continuous-time:*

*Plant:*

$$G(s) = \frac{10}{s(s+1)}.$$

*Controller:*

$$C(s) = \frac{0.416s + 1}{0.139s + 1}.$$

*Discrete-time controllers:*

| $T_l[s]$ | $N$ | $\|\tilde{J}_d\|_\infty$ | zeros: $z_1$ | $z_2$ | $z_3$ | poles: $p_1$ | $p_2$ | $p_3$ | gain $k$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.0157 | 5 | 0.026 | −0.1657 | 0.2464 | 0.9633 | 0.0001 | −0.2776 | 0.8932 | 4.2173 |
| 0.0785 | 10 | 0.135 | −0.1450 | 0.3080 | 0.8332 | 0.0137 | −0.2682 | 0.5801 | 3.7321 |
| 0.157 | 20 | 0.265 | −0.1681 | 0.7088 | | −0.0173 | −0.2710 | | 2.8926 |
| 0.314 | 40 | 0.680 | 0.1710 | 0.6499 | | 0.1034 | −0.2057 | | 1.1931 |
| 0.420 | 50 | 0.950 | 0.2379 | 0.8412 | | 0.1414 | −0.2399 | | 0.5266 |

for $C_d(z_l) = k\dfrac{(z-z_1)(z-z_2)(z-z_3)}{(z-p_1)(z-p_2)(z-p_3)}$.

*Rattan's controller ($T_l$ = 0.157s):*

$$C_d(z_l) = \frac{3.436z - 2.191}{z + 0.2390}.$$

*Transfer Functions of Katz/Kennedy's Example*

*Continuous-time:*

*Plant:*

$$G(s) = \frac{863.3}{s^2}.$$

*Controller:*

$$C(s) = \frac{2940s^2 + 86436}{(s + 86436)^2}.$$

*Discrete-time controllers:*

| $T_l[s]$ | $N$ | $\|\tilde{J}_d\|_\infty$ | zeros: $z_1$ | $z_2$ | $z_3$ | poles: $p_1$ | $p_2$ | $p_3$ | gain $k$ |
|---|---|---|---|---|---|---|---|---|---|
| 0.001 | 5 | 0.030 | 0.9699 | −1.2570 | −5.4273 | −0.2244 | −0.0445 | 0.8338 | 0.5019 |
| 0.010 | 10 | 0.166 | 0.7450 | 0.2482 | −0.3426 | −0.2876 | 0.0145 | 0.0845 | 4.3237 |
| 0.030 | 20 | 0.652 | 0.7063 | −0.0959 | | −0.1810 | −0.1027 | | 1.4107 |
| 0.039 | 40 | 0.892 | 0.8476 | | | −0.2089 | | | 0.8513 |

*Kennedy's controller ($T_l$ = 0.030s):*

*Feedforward:*

$$F(z) = \frac{t(z)}{r(z)}.$$

*Feedback:*

$$C(z) = \frac{s(z)}{r(z)}.$$

with

$$\begin{aligned} r(z) &= z^2 + 1.1773z + 0.7418 \\ s(z) &= 2.6403z - 1.8714 \\ t(z) &= 1.3084z^2 - 0.5390z - 0.0007. \end{aligned} \tag{B.1}$$

## References


[1] B. D. O. Anderson and Y. Liu, "Controller reduction: Concepts and approaches," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 802–812, 1989.

[2] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic, 1975.

[3] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State space solutions to the standard $H_2$ and $H_\infty$ control problem," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 831–848, 1989.

[4] K. Foo and I. Postlethwaite, "An $H_\infty$-minimax approach to the

design of robust control systems—Part II: All solutions, all-pass form solutions, and the 'best' solution," *Syst. Contr. Lett.*, vol. 7, pp. 261-268, 1986.

[5] G. F. Franklin and J. D. Powell, *Digital Control of Dynamic Systems.* Reading, MA: Addison-Wesley, 1980.

[6] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems with their $L^\infty$-error norm," *Int. J. Contr.*, vol. 39, pp. 1105-1193, 1984.

[7] P. Katz, *Digital Control Using Microprocessors.* Englewood Cliffs, NJ: Prentice-Hall, 1981.

[8] R. A. Kennedy and R. J. Evans "Digital redesign of a continuous controller based on closed-loop performance," in *Proc. 29th Conf. Decision Contr.*, HI, 1990, pp. 1898-1901.

[9] P. P. Khargonekar, K. Poolla, and A. Tannenbaum, "Robust control of linear time-invariant plants using periodic compensation," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1088-1096, 1985.

[10] D. L. Kleinman and P. K. Rao, "Continuous-discrete gain transformation methods for linear feedback control," *Automatica*, vol. 13, pp. 425-428, 1977.

[11] B. C. Kuo and D. W. Peterson, "Optimal discretization of continuous data control systems," *Automatica*, vol. 9, pp. 125-129, 1973.

[12] Y. Liu, B. D. O. Anderson, and U.-L. Ly, "Coprime factorization controller reduction with Bezout identity induced frequency weighting," *Automatica*, vol. 26, pp. 233-249, 1990.

[13] S. K. Mitra and R. Gnanasekaran, "Block implementation of recursive digital filters—New structures and properties," *IEEE Trans. Circuits and Syst.*, vol. CAS-25, pp. 200-207, 1978.

[14] H. H. Niemann, "A frequency domain design method of sampled-data compensators," presented at the Amer. Contr. Conf., San Diego, CA, 1990.

[15] K. S. Rattan and H. H. Yeh, "Discretizing continuous-data control systems," *Computer-Aided Design*, vol. 10, pp. 299-306, 1978.

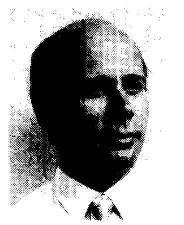
[16] K. S. Rattan, "Digitalization of existing continuous control systems," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 282-285, 1984.

[17] M. G. Safonov, E. A. Jonckheere, M. Verma, and D. J. N. Limebeer, "Synthesis of positive real multivariable feedback systems," *Int. J. Contr.*, vol. 45, pp. 817-842, 1987.

[18] P. M. Thompson, G. Stein, and M. Athans, "Conic sectors for sampled-data feedback systems," *Syst. Contr. Lett.*, vol. 3, pp. 88-82, 1983.

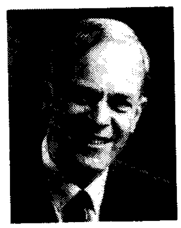
[19] P. M. Thompson, R. L. Dailey, and J. C. Doyle, "New conic sectors for sampled-data feedback systems," *Syst. Contr. Lett.*, vol. 7, pp. 395-404, 1986.

[20] R. A. Yackel, B. C. Kuo, and G. Singh, "Digital redesign of continuous systems by matching states at multiple sampling periods," *Automatica*, vol. 10, pp. 105-111, 1974.



**Jürg P. Keller** was born in Samedan, Switzerland on November 6, 1958. He received the diploma in chemical engineering in 1983, the diploma in control theory in 1989, and the Doctor of technical sciences in 1989 from the Eidgenössische Technische Hochschule ETH, Zürich, Switzerland.

From April 1989 to May 1990 he was Visiting Fellow at the Department of Systems Engineering, Australian National University, Canberra, Australia. Since 1990 he has been with Hoffman-La Roche AG, Basel, Switzerland, where he has been working as Control Engineer. His interests include model reduction, control structure selection, and optimization of sampled data systems.



**Brian D. O. Anderson** (S'62-M'66-SM'74-F'75) was born in Sydney, Australia, and received his education at the University of Sydney and Stanford University, Stanford, CA.

He holds a doctorate (*honoris causa*) from the Université Catholique de Louvain, Belgium. Following completion of his education, he worked in industry in Silicon Valley and served as an Assistant Professor in the Department of Electrical Engineering at Stanford. He was Foundation Professor of Electrical Engineering at the University of Newcastle, Australia from 1967 till 1981 and is now Professor of Systems Engineering at the Australian National University, Canberra. His interests are in control and signal processing.

Dr. Anderson is a Fellow of the Royal Society, the Australian Academy of Science, Australian Academy of Technological Sciences and Engineering, and the Institute of Electrical and Electronic Engineers, and an Honorary Fellow of the Institution of Engineers, Australia. He is serving a term as President of the International Federation of Automatic Control from 1990 to 1993.