DIGITAL CONTROL

The revolutionary advances in computer technology today have made it possible to replace conventional controllers with digital computers. *Digital control* thus refers to the control scheme in which the controller is a digital device, generally a digital computer. This means that we can make use of a much more advanced control logic and versatility than those made possible with conventional analog controllers. On the other hand, we also need an interface that connects a computer with real plants. In particular,

- Measurement is made at discrete instants in time
- Data must be spatially discretized to allow digital data handling

In other words, digital controllers can handle data that are discretized both in time and space. The former discretization is usually referred to as *sampling* and the latter *quantization.* These two features place digital control systems outside the scope of the usual linear, time-invariant control systems. (There is also the problem of saturation effect when controllers have a fixed word length. But this problem is much less studied in the context of digital control.)

To see the situation more concretely, consider the unity-feedback digital control system shown in Fig. 1. Here r is the reference signal, y the system output, and e the error signal. These are continuous-time signals. The error $e(t)$ goes through the *sampler* (or an *A/D converter*) $\mathcal{S}$. This sampler

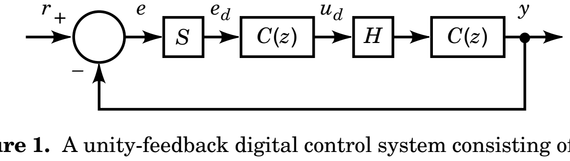

Figure 1. A unity-feedback digital control system consisting of a continuous-time plant $P(s)$, discrete-time controller $C(z)$, sampler $\mathcal{S}$ and a hold device $\mathcal{H}$.

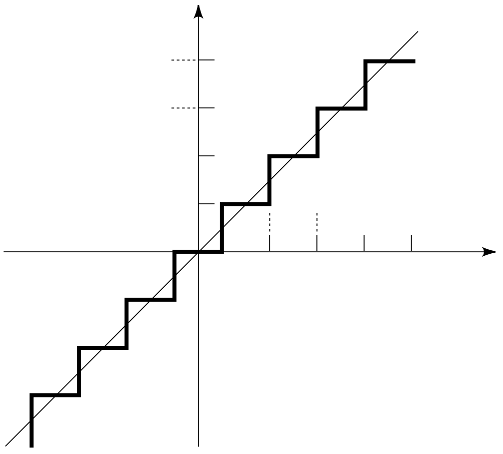

Figure 2. Quantization converts the slanted straight line (thin) to the piecewise step zigzag function (thick).

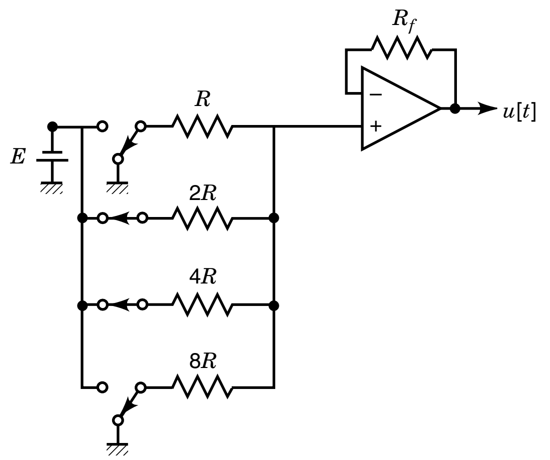

Figure 4. A D/A converter is constructed with an operational amplifier, switching, and resistors.

reads out the values of $e(t)$ at every time step h called the *sampling period,* and produces a discrete-time signal $e_d[k]$, $k = 0, 1, 2, \ldots$. In this process, a quantization error (due to round-off) as shown in Fig. 2 occurs. The sampling operator $\mathscr{S}$ acts on a continuous-time signal $w(t)$, $t \geq 0$ as

$$\mathscr{S}(w)[k] := w(kh), \quad k = 0, 1, 2, \ldots$$

(The quantization effect is omitted here.) The discretized signal is then processed by a discrete-time controller $C(z)$ and becomes a control input u_d. This signal then goes through another interface $\mathscr{H}$ called a *hold device* or a *D/A converter* to become a continuous-time signal. A typical example is the *zero-order hold,* where $\mathscr{H}$ simply keeps the value of a discrete-time signal $w[k]$ as a constant until the next sampling time:

$$(\mathscr{H}(w[k]))(t) := w[k], \quad \text{for } kh \leq t < (k+1)h$$

A typical sample-hold action [with $C(z)$ = identity] is shown in Fig. 3. A simple D/A converter can be constructed with operational amplifiers, resistors, and switching devices as depicted in Fig. 4. Because this construction requires high precision in resistors, more elaborate circuitry is adopted in practice.

There are other types of hold devices, for example, a first-order hold for various reasons. In this article, however, we confine ourselves to the zero-order hold above.

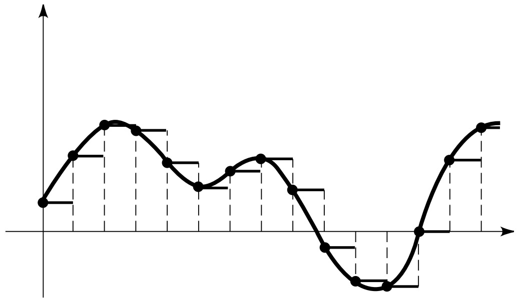

Figure 3. A simple sample-hold combination maps a continuous-time signal to a piecewise step function.

In the process above, a quantization error occurs in the A/D conversion. This is a round-off error that occurs when we convert analog values to digital data (often with a fixed wordlength), as shown in Fig. 2. This introduces a nonlinearity into the system although other system components may be linear. A possible effect is that the closed-loop system may exhibit typical nonlinear behavior, such as limit cycles. Such phenomena are, however, much less studied compared to the effect arising from data sampling in time, and one usually assumes that sufficient spatial resolution is guaranteed so that the effect of quantization is negligible.

The term digital control is thus used almost synonymously with sampled-data control (that is, the control scheme where measurement and control actions occur intermittently with a fixed period) and quantization effects are ignored. Usually one considers *single-rate* sampled-data control systems where sampling and hold actions occur periodically with a fixed period in a synchronized way. In practice, however, there are varied situations in which different sampling rates are employed at distributed control stations. Such a situation leads to *multirate* sampled-data control systems. However, for the sake of simplicity this article deas with single-rate systems.

z-TRANSFORM

We start with a fundamental description of systems and sequences. Let $\{w[k]\}_{k=0}^{\infty}$ denote a sequence with values in some vector space X. Typically, X is the n-dimensional vector space $\mathbb{R}^n$, but we will later encounter an example where X is not finite-dimensional. The *z-transform* of $w = \{w[k]\}_{k=0}^{\infty}$ is defined to be the formal sum (mathematically, this is called a *formal power series*):

$$\mathscr{Z}[w](z) := \sum_{k=0}^{\infty} w[k]z^{-k}$$

with indeterminate z. It is also denoted as $\hat{w}(z)$. The negative powers of z is in accord with the usual convention. Here z is just a formal variable, and z-transform at this stage simply gives a convenient way of coding sequences via the correspondence $t \leftrightarrow z^{-t}$.

It can be readily verified that the z-transform $\mathscr{Z}[w * u]$ of the *discrete convolution*

$$(w * u)[k] := \sum_{j=0}^{k} w[k-j]u[j]$$

is given by the product $\mathscr{Z}[w]\mathscr{Z}[u]$ of the z-transforms of the two sequences, i.e.,

$$\mathscr{Z}[w * u] = \mathscr{Z}[w]\mathscr{Z}[u]$$

As a special case, the multiplication by z^{-1} yields the time-shift (delay): $\{w[k]\}_{k=0}^{\infty} \mapsto \{w[k-1]\}_{k=1}^{\infty}$. Similarly, the multiplication by z yields the time-advance operator: $\{w[k]\}_{k=0}^{\infty} \mapsto \{w[k+1]\}_{k=0}^{\infty}$.

The z-transformation plays the role of the Laplace transformation in the continuous-time case. As with Laplace transforms, it is useful to consider the substitution of a complex number to the variable z. For example, the geometric sequence $\{\lambda^k v\}_{k=0}^{\infty}$ has the z-transform

$$\sum_{k=0}^{\infty} \lambda^k z^{-k} v = \frac{zv}{z-\lambda} \tag{1}$$

We can consider this as a function with complex variable z. The sequence $\{\lambda^k v\}$ tends to zero if and only if $|\lambda| < 1$; this is equivalent to its z-transform being analytic in $\{z : |z| \geq 1\}$.

For a fixed sampling period h, the z-transform of a continuous-time signal $x(t)$ is understood to be the z-transform of its sampled sequence:

$$\mathscr{Z}[x](z) := \sum_{k=0}^{\infty} x(kh) z^{-k}$$

There is a way to compute $\mathscr{Z}[x](z)$ from its Laplace transform (see Theorem below). Note also that the z-transform of an exponential function $e^{\mu t}$ is $z/(z - e^{\mu h})$:

$$\mathscr{Z}[e^{\mu t}](z) = \frac{z}{z - e^{\mu h}} \tag{2}$$

Let us now give a system description. Suppose that a discrete-time system

$$\begin{aligned} x[k+1] &= Ax[k] + Bu[k] \\ y[k] &= Cx[k] + Du[k] \end{aligned} \tag{3}$$

is given. Taking z-transforms of sequences $\{x[k]\}$, $\{u[k]\}$, $\{y[k]\}$ and using the fact that the multiplication by z induces the time-advance operator, we see that

$$\begin{aligned} z\hat{\mathbf{x}} &= A\hat{\mathbf{x}} + B\hat{\mathbf{u}} \\ \hat{\mathbf{y}} &= C\hat{\mathbf{y}} + D\hat{\mathbf{u}} \end{aligned}$$

Solving this, we have

$$\hat{\mathbf{y}} = C(zI - A)^{-1}x_0 + [D + C(zI - A)^{-1}B]\hat{\mathbf{u}}$$

where x_0 is the initial state at time 0. The second term $D + C(zI - A)^{-1}B$ is the *transfer function* of this system. It is (asymptotically) stable if and only if $zI - A$ has no eigenvalues on or outside the unit circle $\{z : |z| = 1\}$. This is equivalent to the transfer function being analytic in $\{z : |z| \geq 1\}$, provided that there are no hidden poles of $(zI - A)^{-1}$ cancelled by the numerator.

Let us now give a sample-point description of the continuous-time plant $P(s)$. Let (A, B, C) be its (minimal) realization. For brevity, we assume that the direct feedthrough term of $P(s)$ is zero. This means

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{4}$$

The first objective is to give a description of this plant at sampling instants $t = kh$, $k = 0, 1, 2, \ldots$. By the zero-order hold, the input to the plant for $kh \leq t < (k+1)h$ is the constant $u_d[k]$. Suppose that the state of the plant at $t = kh$ is $x[k] = x(kh)$. Then by integrating Eq. (4) from kh to $(k+1)h$, we obtain

$$\begin{aligned} x[k+1] = x((k+1)h) &= e^{Ah}x[k] + \int_0^h e^{A(h-\tau)}Bu_d[k]\,d\tau \\ &= e^{Ah}x[k] + \int_0^h e^{A(\tau)}Bd\tau u_d[k] =: A_d x[k] + B_d u[k] \\ y[k] = y(kh) &= Cx[k] =: C_d x[k] \end{aligned} \tag{5}$$

In other words, the behavior of $P(s)$ at the sampling instants can be described by a time-invariant *discrete-time system* (A_d, B_d, C_d). This is the formula due to Kalman and Bertram (1). Its transfer function

$$P_d(z) := C(zI - e^{Ah})^{-1} \int_0^h e^{A\tau} Bd\tau$$

is called the *pulse transfer function* of $P(s)$. Then the composite transfer function of the closed-loop system, when confined to sampled instants, is given by $P_d(z)C(z)(I + P_d(z)C(z))^{-1}$. We then have the following result.

Theorem 1. The behavior of the sampled-data system in Fig. 1 at sampled instants can be described by a time-invariant, discrete-time equation. To be more precise, let (A_0, B_0, C_0, D_0) and (A, B, C) be the minimal realizations of $C(z)$ and $P(s)$, respectively. Then the behavior of the closed-loop system at sampled instants can be represented by system matrices

$$\left(\begin{bmatrix} A_0 & -B_0C_d \\ B_dC_0 & A_d - B_dD_0C_d \end{bmatrix}, \begin{bmatrix} B_0 \\ B_dDD_0 \end{bmatrix}, [0 \quad C_d] \right)$$

Example 1. Let $P(s) = 1/s^2$. This has a realization

$$\begin{aligned} \frac{dx}{dt} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y &= [1 \quad 0]x \end{aligned}$$

Equation (5) is then computed easily as

$$\begin{aligned} x[k+1] &= \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} x[k] + \begin{bmatrix} h^2/2 \\ h \end{bmatrix} u_d[k] \\ y[k] &= [1 \quad 0]x[k] \end{aligned}$$

Much of the classical theory for sampled-data control is devoted to the computation of the pulse transfer function from a given continuous-time transfer function. Note that this precedes the advent of the modern state space theory, and elaborate formulas in the z and Laplace transform domains have been found. For example, the following theorem is well known (2,3):

Theorem 2. Let $P(s)$ be a rational function such that it is analytic for $\{s : |s| > R\}$ for some $R > 0$ and $sP(s) \to \alpha$ as $|s| \to \infty$ for some real α. Then

$$\begin{aligned}\mathscr{Z}[P](z) &= \frac{1}{2\pi j}\oint_{\gamma} \frac{P(s)}{1 - e^{hs}z^{-1}}\,ds \\ &= \sum_{\text{poles of } P(s)} \text{Res}\left[\frac{P(s)}{1 - e^{hs}z^{-1}}\right]\end{aligned} \tag{6}$$

where γ denotes a contour that travels from $c - j\infty$ to $c + j\infty$ [c: abscissa of convergence of $P(s)$; the coordinate to the right of which the Laplace integral defining $P(s)$ converges] and goes around a semcircle on the left half-plane that encircles all poles of $P(s)$.

Once the sample-point behavior is specified by the procedure above, it is easy to give a description of the intersample behavior of the output or the state. Suppose, for example, that the state of the plant $P(s)$ takes values $x[k]$ at $t = kh$, $k = 0, 1, 2, \ldots$. By integrating Eq. (4) from $t = kh$ to $t = kh + \theta$ $(0 \le \theta \le h)$, we get

$$x(kh + \theta) = e^{A\theta}x[k] + \int_0^{\theta} e^{A(\theta-\tau)}Bu_d[k]\,d\tau \tag{7}$$

and

$$y(kh + \theta) = Cx(kh + \theta) \tag{8}$$

This shows that if $x[k]$ and $u_d[k]$ tend to zero as $k \to \infty$, then the intersampling behavior $x(kh + \theta)$, $0 \le \theta < h$ also tends to zero uniformly for θ as $k \to \infty$. This is because the right-hand side of Eq. (7) is just the multiplication and convolution of known continuous functions with $x[k]$ and $u_d[k]$ over a finite interval. Therefore, the stability of a sampled-data system can be determined solely by its sample-point behavior. We summarize this observation in the form of a theorem:

Theorem 3. The closed-loop system in Fig. 1 is stable if the discrete-time closed-loop system consisting of $C(z)$ and $P_d(z)$ is stable. Therefore, to stabilize the plant $P(s)$ in the sampled-data setting Fig. 1, it is necessary and sufficient that $P_d(z)$ be stabilized by $C(z)$.

This result gives a foundation for the classical treatment of sampled-data systems. To design (or at least to stabilize) a sampled-data system, one can equivalently stabilize the pulse transfer function $P_d(z)$ derived from $P(s)$. This led to the classical design procedure based on pulse transfer functions.

Equations (7) and (8) are closely related to the notion of the *modified z-transform.* For a continuous-time function $x(t)$, its modified z-transform $\mathscr{Z}[x](z, m)$, $0 \le m \le 1$ is defined as

$$\mathscr{Z}[x](z, m) := \sum_{k=0}^{\infty} x(kh + mh)z^{-k} \tag{9}$$

As in Theorem 2, the following result holds.

Theorem 4. Assuming the same conditions as in Theorem 2, the following formulas for the modified z-transform holds:

$$\begin{aligned}\mathscr{Z}[P](z, m) &= \frac{1}{2\pi j}\oint_{C} \frac{P(s)e^{mhs}}{1 - e^{mhs}z^{-1}}\,ds \\ &= \sum_{\text{poles of } P(s)} \text{Res}\left[\frac{P(s)e^{mhs}}{1 - e^{mhs}z^{-1}}\right]\end{aligned}$$

The modified z-transform has a close connection with lifted transfer functions in the modern approach (see the section entitled Modern Approach).

CLASSICAL DESIGN METHODS AND THEIR LIMITATIONS

We briefly review the classical design methods and their limitations. The first is a design method based on the continuous-time design.

Continuous-Time Design

A simple, straightforward method is to employ a continuous-time design, obtain a continuous-time controller, and then convert the controller to a discrete-time system via some kind of discretization. Let $C_c(s)$ be a continuous-time controller. Typical discretization methods are the following:

- Use the Tustin (bilinear) transformation:

$$C(z) = C_c\left(\frac{2}{h} \cdot \frac{z-1}{z+1}\right)$$

- Employ the backward difference $(z - 1)/hz$ for approximating the differential operator s.
- Approximate $C_c(s)$ by using the sample/hold equivalent Eq. (5).

Among these, the first method is most commonly used.

It is well known that the Tustin transformation preserves stability: if $C_c(s)$ is a stable transfer function (in the sense of continuous-time systems), then the transformed function $C_c(2(z - 1)/h(z + 1))$ gives a stable (in the sense of discrete-time systems) discrete-time transfer function. Although this is a great advantage in signal processing, care must be exercised in control system design, because this property does not guarantee the closed-loop stability. In fact, as h becomes larger, there is even a case in which the closed-loop stability is violated (see the example in the section entitled "H^{∞} Design"). This is because the original continuous-time design does not usually take account of the sampling period. To take care of this, one has to pay more attention to various robustness properties, such as gain and phase margins, and so on. To discuss such properties, frequency domain considerations are highly desirable.

However, the notion of frequency response is not readily available. To see the situation, let $C(z)$ be a discrete-time transfer function. Suppose that a sinusoid $e^{j\omega t}$ is applied after sampling. Then the actual input to $C(z)$ is $\{e^{kj\omega h}\}_{k=0}^{\infty}$ with z-transform $z/(z - e^{j\omega h})$ given by Eq. (2). The steady-state response of $C(z)$ against this input is then given by $\{e^{kj\omega h}C(e^{j\omega h})\}_{k=0}^{\infty}$. It appears that we can discuss the frequency domain properties via $C(e^{j\omega h})$. For example, one might attempt to employ phase lead/lag compensation based on this quantity. However, due to sampling, this frequency response does not fully represent the nature of continuous-time inputs. For example, not only $e^{j\omega t}$ but also $e^{j(\omega+2n\pi/h)t}$, $n = \pm1, \pm2, \ldots$ give exactly the same sampled values $\{e^{kj\omega h}\}_{k=0}^{\infty}$. The response is then governed by the same $C(e^{j\omega h})$. This means that sampling does not have enough resolution to distinguish all these sinusoids, and the notion of phase, which depends on the response against sinusoids, is unclear in such a sampled-data control setup. Another way of seeing this is to note that $e^{-j\omega h} = \overline{e^{j\omega h}}$ and hence $C(e^{j(2\pi/h-\omega)h}) = \overline{C(e^{j\omega h})}$. This means that beyond the frequency π/h, the same gain characteristic repeats periodically and $C(e^{j\omega h})$ cannot be treated as the same frequency response concept as in the continuous-time case. This is related to the notion of aliasing, which we examine in more detail in the section "Modern Approach."

It may be still possible to execute an elaborate continuous-time design that also works well in the sampled-data setting by looking more closely into the nature of the Tustin transformation. However, in such a method, a systematic design method such as H^{∞} design theory is difficult to apply. Furthermore, one needs a more concrete understanding of the phenomena above, and this is much better done in the modern approach treated in the subsequent sections.

Discrete-Time Design

Yet another classical approach is based on the pulse transfer function $P_d(z)$. As far as stability is concerned, one can deal only with $P_d(z)$. It was also recognized that sampled-data control can achieve performance that is not possible with linear, time-invariant, continuous-time controller. For example, the so-called deadbeat control achieves the property that the output (or state) settles exactly to zero after a finite time period. This is done by placing all poles of the closed-loop system to zero; the output, at least at sampled instants, then becomes zero after a finite time—a performance not possible with continuous-time controllers.

It should be, however, noted that such a classical treatment also shares the weakness of the classical transfer function approach. Namely, it did not take account of hidden pole–zero configurations. In particular, it was observed that merely settling the output might sometimes induce very large intersample ripples.

It was Kalman and Bertram (1) who introduced the state space approach for sampled-data systems. As we have already seen, the sample-time input–output relation is described by

$$x[k+1] = A_d x[k] + B_d u_d[k] = e^{Ah}x[k] + \int_0^h e^{A\tau}\, Bd\tau\, u_d[k]$$

Let us consider the stabilization by state feedback. If we employ a sampled state feedback

$$u_d[k] = -Kx[k]$$

then by Theorem 3 the stability of the closed-loop system is determined by the spectrum of

$$A_d - B_d K \tag{10}$$

Thus this is a purely discrete-time pole-placement problem. Furthermore, if we can set the eigenvalues of Eq. (10) all to zero, then $x[k]$ becomes zero after a finitely number of steps if there is no external input. Also with Eqs. (7) and (8), not only $x[k]$ but also the intersampling trajectory will settle to zero. This clearly shows the advantage of the state space theory, which was introduced around that time.

The problem is thus reduced to the pole allocation for the discrete-time system (A_d, B_d) and the feasibility of this is reduced to the problem of determining the controllability and stabilizability of this pair. Naturally, we may as well assume that the continuous-time plant (A, B, C) is stabilizable or controllable. Otherwise, it is not possible to stabilize the plant even with continuous-time controllers.

For brevity, let us consider controllability. The following result is well known (4):

Theorem 5. Let $\lambda_1, \ldots, \lambda_n$ be eigenvalues of A. Suppose that for no pair λ_i, λ_j $(i \neq j)$, $\lambda_i - \lambda_j$ is an integer multiple of $2\pi/h$. Then (A_d, B_d) is controllable if and only if (A, B) is controllable.

The proof is an easy consequence of the fact that the eigenvalues of $A_d = e^{Ah}$ are $\{e^{\lambda_1}, \ldots, e^{\lambda_n}\}$. This is a variant of the spectral mapping theorem, but we skip the details here (4,5).

By the discussions above, it may appear that sampled-data control systems can be safely designed via discrete-time design methods. Note that, at least for the deadbeat control via state space, we can also settle the intersample behavior identically to zero after a finite number of steps. However, this is valid only for regulation problems, and the issue of the intersample behavior for tracking (servo control) problems, where exogenous signals are present, is quite different. To see this, consider the example depicted in Fig. 5. Here the continuous-time plant is $1/(s^2 + 1)$ whose natural frequency is 1 rad/s. On the other hand, the tracking signal is $\sin\,(1 + 2\pi/0.1)t$ where the sampling period h is 0.1 s. It so happens that, at sampling instants $t = kh$, $k = 0, 1, 2, \ldots$, this signal is identical with $\sin t$ because $(2\pi/0.1)kh = 2k\pi$. Therefore, for the discrete-time controller the tracking signal is no different

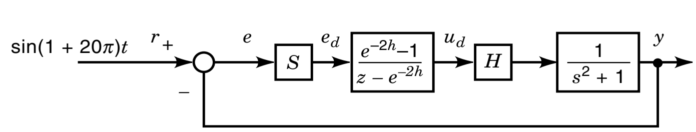

Figure 5. A unity feedback system with tracking reference signal $\sin(1 + 20\pi)t$.

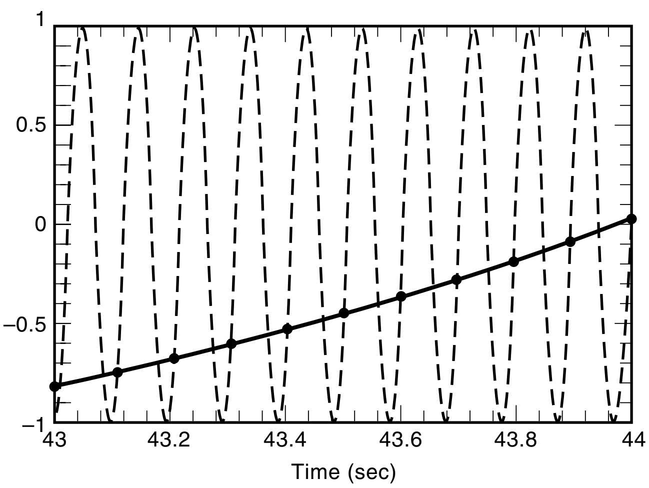

Figure 6. The simulation of Figure 5 shows that the input $\sin(1 + 20\pi)t$ does not yield a sinusoid at the same frequency, and large intersample ripples result.

from sin t. The simulation result is shown in Fig. 6. The plant output is shown by the solid curve while the reference input is shown by the dashed curve. The output tracks sin t rather than $\sin(1 + 2\pi/0.1)t$, and there is a large amount of intersample ripples due to the difference between $\sin(1 + 2\pi/0.1)t$ and sin t.

This example shows the following:

- There can be large intersample ripples for sampled-data systems.
- Such ripples are difficult to characterize via the discrete-time framework as described above.
- The ripples do not appear to be stationary.

The observations above indicate that the discrete-time model Eq. (5) is generally not appropriate for describing sampled-data systems when there are nontrivial intersample ripples. What is indicated here is that we need a framework that can give a description for the continuous-time behavior of a sampled-data system.

Suppose that we wish to describe a frequency response. Let sin ωt be an input applied to the sampled-data system shown in Fig. 1. For linear, time-invariant, stable continuous-time systems, it is well known that a single sinusoid yields another sinusoid in the steady-state output, with exactly the same frequency, possibly with gain and phase shifts. To be more precise, let $G(s)$ be the transfer function of such a system. It is well known that the steady-state output is

$$G(j\omega)\sin\omega t$$

That is, as $t \to \infty$, the output asymptotically approaches $G(j\omega)\sin\omega t$.

Such a separation principle does not hold for sampled-data systems. In fact, the example shown in Fig. 6 shows a counterexample: the steady-state output against $\sin(1 + 2\pi/0.1)t$ is sin t—sinusoid, but with a different frequency.

One of the reasons for such a phenomenon is that sampled-data systems are no longer time-invariant systems in a very strict sense, if we take the intersample behavior into account. This is closely related to the issue of the notion of aliasing effects and Shannon's sampling theorem. We briefly review these in the next section.

SAMPLING THEOREM

Let $f(t)$ be a given continuous-time signal on $(-\infty, \infty)$. To make the sampling well defined, we assume that f is a continuous function. The sampled sequence is $\{f(kh)\}_{k=-\infty}^{\infty}$. As it is, this is just a sequence defined on the set of integers. The question here is how we should represent this sequence in the continuous-time domain.

Recall that the z-transform of $\{f(kh)\}_{k=-\infty}^{\infty}$ is

$$\sum_{k=-\infty}^{\infty} f(kh)z^{-k}$$

We have extended the definition in a natural way to the negative ks. We also recall that the multiplication by z^{-1} is the right shift operator. Since the kth signal value $f(kh)$ is placed at $t = kh$, this right shift corresponds to the right shift by time length h in the continuous-time domain. It is also well known that in the Laplace transform domain the right shift operator by h is represented by the multiplication by e^{-hs}. Therefore, it is natural to represent the Laplace transform of the *sequence* $\{f(kh)\}_{k=-\infty}^{\infty}$ by

$$\sum_{k=-\infty}^{\infty} f(kh)e^{-khs}$$

The inverse Laplace transform of this is the train of impulses (Delta functions) multiplied by $f(kh)$ at the kth step:

$$\sum_{k=-\infty}^{\infty} f(kh)\delta(t-kh) \tag{11}$$

Observe that this is formally a multiplication of $f(t)$ with the train of impulses

$$\sum_{k=-\infty}^{\infty} \delta(t-kh) \tag{12}$$

and thus it is called the *impulse modulation* of $f(t)$.

The question that concerns us here is the following: Suppose that we are given a sampled sequence $\{f(kh)\}_{k=-\infty}^{\infty}$, or $\sum_{k=-\infty}^{\infty} f(kh)e^{-khs}$ just as well. How much can we recover the original signal $f(t)$ out of this piece of data?

If we impose no condition on $f(t)$, then the solution is clearly nonunique. There is infinite freedom in the intersampling periods while passing through $f(kh)$, $k \in \mathbb{Z}$. A typical solution is obtained by assuming that $f(t)$ is *band-limited;* that is, its Fourier transform is zero outside a bounded interval. This is the content of the following Shannon sampling theorem.

Theorem 6. Let f be a continuous function that is Fourier transformable. Suppose that its Fourier transform is identically zero outside the interval $(-\pi/h + \epsilon, \pi/h - \epsilon)$ for some $\epsilon > 0$. Then

$$f(t) = \sum_{n=-\infty}^{\infty} f(nh)\frac{\sin\pi(t/h-n)}{\pi(t/h-n)} \tag{13}$$

We now briefly indicate the outline of a proof.

As noted above, Eq. (11) is obtained by multiplying $f(t)$ to the train of impulses [Eq. (12)]. Hence its Fourier transform is just the convolution of the respective Fourier transforms (6). For the Fourier transform of Eq. (12), the following Poisson summation formula is well known (6):

$$\mathscr{F}\left(\sum_{n=-\infty}^{\infty}\delta(t-nh)\right) = \frac{2\pi}{h}\sum_{n=-\infty}^{\infty}\delta\left(t-\frac{2n\pi}{h}\right) \tag{14}$$

It follows that

$$\begin{aligned}\mathscr{F}\left(\sum_{n=-\infty}^{\infty} f(nh)\delta(t-nh)\right) &= \mathscr{F}\left(f(t)\sum_{n=-\infty}^{\infty}\delta(t-nh)\right)\\ &= \frac{1}{2\pi}\hat{\mathbf{f}}(\omega) * \left(\frac{2\pi}{h}\sum_{n=-\infty}^{\infty}\delta\left(t-\frac{2n\pi}{h}\right)\right)\\ &= \frac{1}{h}\hat{\mathbf{f}}(\omega) * \left(\sum_{n=-\infty}^{\infty}\delta\left(t-\frac{2n\pi}{h}\right)\right)\\ &= \frac{1}{h}\sum_{n=-\infty}^{\infty}\hat{\mathbf{f}} * \delta\left(t-\frac{2n\pi}{h}\right)\\ &= \frac{1}{h}\sum_{n=-\infty}^{\infty}\hat{\mathbf{f}}\left(\omega-\frac{2n\pi}{h}\right)\end{aligned}$$

In other words, we get infinitely many copies of the shifted image of $\hat{\mathbf{f}}(\omega)$ as shown in Fig. 7. This is because a sinusoid $\sin \omega t$ behaves precisely the same as $\sin(\omega + 2m\pi/h)t$ at sampled points $t = nh$, $n = 0, \pm 1, \pm 2, \ldots$. Such higher frequency signals that arise from sampling are called *alias components*. It is clearly not possible to recover the original signal $f(t)$ from such data contaminated by *aliasing*. In particular, there is in general an overlapping of $\hat{\mathbf{f}}(\omega - 2n\pi/h)$. (The period $\omega_s := 2\pi/h$ of these spectra is called the *sampling frequency* and its half π/h the *Nyquist frequency*.)

However, it is possible to recover $f(t)$ if such an overlapping does not occur. Indeed, it will be clear from Fig. 7 that if the original spectrum $\hat{\mathbf{f}}$ is zero outside the interval $(-\pi/h, \pi/h)$, then there is no overlapping among those copies. The band-limited hypothesis that $\hat{\mathbf{f}}$ is zero outside $(-\pi/h + \epsilon, \pi/h - \epsilon)$ guarantees this. To eliminate all unnecessary alias components, multiply the function

$$\alpha(\omega) := \begin{cases}1 & (|\omega| \le \pi/h)\\ 0 & (|\omega| > \pi/h)\end{cases} \tag{15}$$

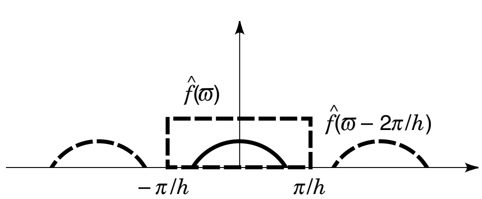

Figure 7. The spectrum of f repeats periodically with period $2\pi/h$.

to $\sum_{n=-\infty}^{\infty}\hat{\mathbf{f}}(\omega - 2n\pi/h)$. Then only the spectrum $\hat{\mathbf{f}}(\omega)$ in the fundamental frequency range $(-\pi/h, \pi/h)$ remains. Applying the inverse Fourier transform, we obtain

$$\begin{aligned}f(t) = \overline{\mathscr{F}}[\hat{\mathbf{f}}(\omega)] &= \overline{\mathscr{F}}\left[\alpha(\omega)\sum_{n=-\infty}^{\infty}\hat{\mathbf{f}}(\omega-2n\pi/h)\right]\\ &= h\overline{\mathscr{F}}[\alpha] * \left(\sum_{n=-\infty}^{\infty} f(nh)\delta(t-nh)\right)\\ &= \sum_{n=-\infty}^{\infty} f(nh)\mathrm{sinc}(t-nh)\end{aligned}$$

where $\mathrm{sinc}\, t := h\overline{\mathscr{F}}[\alpha]$. This function is easily computed as

$$\mathrm{sinc}\, t = \frac{\sin\pi t/h}{\pi t/h}$$

This readily implies Eq. (13).

This result also clarifes the meaning of Theorem 5. When there is a pair of eigenvalues that differ only by an integer multiple of $2\pi/h$, the corresponding two modes cannot be distinguished because they yield the same eigenvalues when discretized.

Some remarks are in order. Although Eq. (13) certainly gives a well-defined reconstruction formula, it is crucially based on the assumption that the original signal $f(t)$ is band limited. This assumption, which appears quite innocent, is seldom satisfied in practice. In fact, if f is band limited, it must be an entire function; that is, it is analytic on the whole complex plane. We can hardly expect real signals to be analytic functions. Therefore, the assumption for Eq. (13) can be satisfied only in an approximate sense. The second drawback is that Eq. (13) is not causal. In other words, it makes use of future sampled values $f(nh)$ to reconstruct the current value $f(t)$. It is therefore not physically realizable. To remedy this, one should be content with approximation, and a large portion of digital signal processing is devoted to the various solutions of this problem.

Theorem 6 also yields the following observations:

- By sampling, intersampling information is generally lost.
- In particular, sinusoids $\sin(\omega + 2n\pi/h)t$, $n = 0, \pm 1, \pm 2, \ldots$ cannot be mutually distinguished.
- However, this is about the maximum uncertainty introduced by sampling. After sampling, all the components that arise in the output are combinations of all such alias components $\{\sin(\omega + 2n\pi/h)t\}_{n=-\infty}^{\infty}$.

The last statement still needs to be clarified. The basic idea is the following: When a sinusoid $\sin \omega t$ is sampled, it is converted to a modulated train of impulses as shown in Eq. (12). In other words, infinitely many alias components $\{\sin(\omega + 2n\pi/h)t\}$ are excited by sampling. To avoid an undesirable effect arising from such aliased components, it is generally necessary to place an analog low-pass filter (usually called an *anti-aliasing filter*) in front of the sampler. Since this cannot cancel the alias components completely, how much such alias components affect the overall performance is a concern. Such a question has been studied in the literature (2,3,7). However, its general structure is better understood in the scope of the modern approach, which we describe in the subsequent sections.

MODERN APPROACH

We now turn our attention to the foundation of the modern treatment of sampled-data systems. From what we have presented up to this section, it is clear that the fundamental difficulty in sampled-data control systems lies in the fact that they involve two different time sets: one is discrete (arising from the digital controller) and the other is continuous (arising from the continuous-time plant). This difficulty has been successfully circumvented in the modern approach.

A New Model with Intersample Behavior—Lifting

While it is possible to recover intersample behavior via the modified z-transform, it implicitly assumes sampling inputs in its formulation. It is therefore not adequate for describing correspondence from the exogenous continuous-time inputs to continuous-time outputs.

A new solution was introduced in 1990–1991 (8–12). The new idea, currently called *lifting,* makes it possible to describe sampled-data systems via a time-invariant, discrete-time model while maintaining the intersample behavior.

The idea is very simple. Let $f(t)$ be a given continuous-time signal. Sampling surely results in a loss of intersample information. Then, instead of sampling $f(t)$, we will represent it as a *sequence* of *functions.* Namely, we set up the correspondence (Fig. 8):

$$\mathcal{L} : f \mapsto \{f[k](\theta)\}_{k=0}^{\infty}, \quad f[k](\theta) = f(kh+\theta), \quad 0 \le \theta < h$$

This idea makes it possible to view (time-invariant or even periodically time-varying) continuous-time systems as linear, time-invariant discrete-time systems. (Basically the same idea that converts periodically time-varying discrete-time systems to time-invariant systems were encountered and rediscovered many times in the literature. It appears to date back at least to Ref. 13. Such a discrete-time lifting is also frequently used in signal processing, especially in multirate signal processing, and is called *blocking.*)

The basic idea is the following: Let

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{16}$$

be a given continuous-time plant. Then we lift the input $u(t)$ to obtain $u[k](\cdot)$. We consider that this lifted input is applied at the timing $t = kh$ (h is a prespecified sampling rate), and observe how it affects the system. Let $x[k]$ be the state at time $t = kh$. As in Eq. (5), the state $x[k+1]$ at time $(k+1)h$ is given by

$$x[k+1] = e^{Ah}x[k] + \int_0^h e^{A(h-\tau)}Bu[k](\tau)\,d\tau \tag{17}$$

The difference here is that the lifted input $u[k](\cdot)$ need not be a constant on $(kh, (k+1)h)$, and the right-hand side integral gives an operator

$$L^2[0, h) \to \mathbb{R}^n : u(\cdot) \mapsto \int_0^h e^{A(h-\tau)}Bu(\tau)\,d\tau$$

While state-transition is described in the discrete timing as above, the system keeps producing an output. If we consider lifting of $x(t)$, it is easily seen to be described by

$$x[k](\theta) = e^{A\theta}x[k] + \int_0^\theta e^{A(\theta-\tau)}Bu[k](\tau)\,d\tau$$

Then the lifted output $y[k](\cdot)$ is given by

$$y[k](\theta) = Ce^{A\theta}x[k] + \int_0^\theta Ce^{A(\theta-\tau)}Bu[k](\tau)\,d\tau \tag{18}$$

Observe that Eqs. (17) and (18) take the form

$$\begin{aligned} x[k+1] &= \mathcal{A}x[k] + \mathcal{B}u[k] \\ y[k] &= \mathcal{C}x[k] + \mathcal{D}u[k] \end{aligned}$$

and the operators $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ do not depend on k. In other words, it is possible to describe this continuous-time system with discrete-timing, once we adopt the lifting point of view. To be more precise, the operators $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$ are defined as follows:

$$\begin{aligned} \mathcal{A} &: \quad \mathbb{R}^n \to \mathbb{R}^n : x \mapsto e^{Ah}x \\ \mathcal{B} &: \quad L^2[0, h) \to \mathbb{R}^n : u \mapsto \int_0^h e^{A(h-\tau)}Bu(\tau)\,d\tau \\ \mathcal{C} &: \quad \mathbb{R}^n \to L^2[0, h) : x \mapsto Ce^{A\theta}x \\ \mathcal{D} &: \quad L^2[0, h) \to L^2[0, h) : u \mapsto \int_0^\theta Ce^{A(\theta-\tau)}Bu(\tau)\,d\tau \end{aligned} \tag{19}$$

Thus the continuous-time plant in Eq. (16) can be described by a time-invariant discrete-time model. Once this is done, it is entirely routine to connect this expression with a discrete-time controller, and hence sampled-data systems (for example, see Fig. 1) can be fully described by time-invariant discrete-time equations, this time without sacrificing the intersampling information. We will also denote this overall equation abstractly as

$$\begin{aligned} x[k+1] &= \mathcal{A}x[k] + \mathcal{B}u[k] \\ y[k] &= \mathcal{C}x[k] + \mathcal{D}u[k] \end{aligned} \tag{20}$$

Since this is purely a discrete-time system except that the input–output spaces are infinite-dimensional (L^2 spaces), all the formal developments presented earlier carries over to the

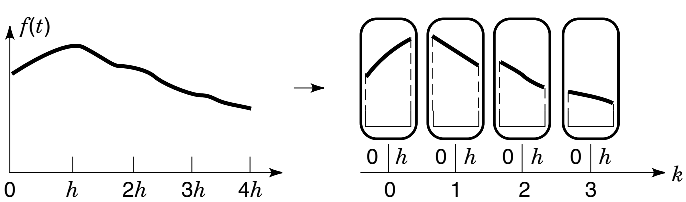

Figure 8. Lifting maps a continuous-time signal to a discrete-time signal with function components.

present situation without any change. For example, the transfer function is defined as

$$G(z) := \mathscr{D} + \mathscr{C}(z\mathscr{I} - \mathscr{A})^{-1}\mathscr{B}$$

When an input $u(z)$ is applied to system in Eq. (20), its zero-initial state output is given by $G(z)u(z)$. Note also that operator $\mathscr{A}$ in Eq. (19) is a matrix, and hence $\mathscr{A}$ in Eq. (20) is also a matrix. This means that the stability of Eq. (20) can be tested by the poles of $G(z)$. It is stable if $G(z)$ is analytic for $\{z : |z| \geq 1\}$.

Steady State and Frequency Response

The time-invariance of the lifted model Eq. (20) naturally yields a definition of stead-state and frequency responses.

Let $G(z)$ be a stable transfer function as defined above. Take a sinusoid $e^{j\omega t}$ as an input. Its lifted image is

$$\{e^{j\omega kh}e^{j\omega\theta}\}_{k=0}^{\infty}$$

According to Eq. (1), the z-transform of this sequence is

$$\frac{ze^{j\omega\theta}}{z - e^{j\omega h}}$$

Since $G(z)$ is stable, expand it in a neighborhood of $z = e^{j\omega h}$:

$$G(z) = G(e^{j\omega h}) + (z - e^{j\omega h})\tilde{G}(z)$$

with some $\tilde{G}(z)$ that is also analytic in $|z| \geq 1$ (by the stability of G). It follows that

$$G(z)\frac{ze^{j\omega\theta}}{z - e^{j\omega h}} = \frac{zG(e^{j\omega h})e^{j\omega\theta}}{z - e^{j\omega h}} + \tilde{G}(z)e^{j\omega\theta}$$

The second term on the right tends to zero as $k \to \infty$ by the analyticity of $\tilde{G}$, and hence the output approaches $zG(e^{j\omega h})e^{j\omega\theta}/(z - e^{j\omega h})$. Therefore, the lifted output $y[k](\cdot)$ asymptotically approaches

$$y[k](\theta) = (e^{j\omega h})^k G(e^{j\omega h})[e^{j\omega\theta}](\theta) \tag{21}$$

as $k \to \infty$.

Unless $e^{j\omega h} = 1$, the asymptotic response above is really not in steady state. However, its *modulus* $|G(e^{j\omega h})e^{j\omega\theta}|$ remains invariant at each sampling time. The change at each step is a phase shift induced by the multiplication by $e^{j\omega h}$. This explains why the ripples in Fig. 6 look similar but not really the same in different sampling periods.

This observation motivates the following definition:

Definition 1 Let $G(z)$ be the transfer function of the lifted system as above. The *frequency response operator* is the operator

$$G(e^{j\omega h}) : L^2[0, h) \to L^2[0, h) \tag{22}$$

regarded as a function of $\omega \in [0, \omega_s)$ ($\omega_s := 2\pi/h$). Its *gain* at ω is defined to be

$$\|G(e^{j\omega h})\| = \sup_{v \in L^2[0,h)} \frac{\|G(e^{j\omega h})v\|}{\|v\|} \tag{23}$$

The maximum $\|G(e^{j\omega h})\|$ over $[0, \omega_s)$ is the H^∞ norm of $G(z)$.

Since $e^{j(\omega+n\omega_s)t}$ is lifted to be $\{e^{j\omega kh}e^{j(\omega+n\omega_s)\theta}\}_{k=0}^{\infty}$, such aliasing high-frequency signals are in the form $e^{j\omega kh}v(\theta)$. Thus the definition above takes all the aliasing components into account, and takes the largest magnitude of enlargement among them.

Frequency Response via Sequence Spaces

The observation above clearly shows that the gain of the frequency response takes all aliasing effects into account. It is however unclear that aliasing exhausts all the freedom in $v \in L^2[0, h)$. In other words, is it true that if we consider the largest gain imposed by considering all aliased components, does this give the same gain as Eq. (23)?

This is indeed true; its proof is based on the following lemma which guarantees that the family $\{e^{j\omega_n\theta}/\sqrt{h}\}_{n=-\infty}^{\infty}$ ($\omega_n = \omega + n\omega_s$) forms an orthonormal basis of $L^2[0, h)$, and hence any $v \in L^2[0, h)$ can be expanded into a series of aliased signals $e^{j\omega_n\theta}$, $n = 0, \pm1, \pm2, \ldots$.

Lemma 1. Fix any $\omega \in [0, \omega_s)$. Then every $\varphi \in L^2[0, h)$ can be expanded in terms of $\{e^{j\omega_n\theta}\}_{n=-\infty}^{\infty}$ as

$$\varphi(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{j\omega_n\theta} \tag{24}$$

with

$$a_n = \frac{1}{h}\int_0^h e^{-j\omega_n\tau}\varphi(\tau)\,d\tau = \frac{1}{h}\hat{\varphi}(j\omega_n) \tag{25}$$

where $\hat{\varphi}$ denotes the Laplace transform of φ when extended to $L^2[0, \infty)$ as 0 outside $[0, h)$. Furthermore, the L^2 norm $\|\varphi\|$ is given by

$$\|\varphi\|^2 = h\sum_{n=-\infty}^{\infty} |a_n|^2 \tag{26}$$

Let us apply this result to the frequency response defined by Eq. (23). Expand $v \in L^2[0, h)$ as $v(\theta) = \sum_{\ell=-\infty}^{\infty} v_\ell e^{j\omega_\ell\theta}$. Note that $\{e^{j\omega kh}e^{j\omega_n\theta}\}_{k=0}^{\infty}$ is the lifted image of $e^{j\omega_n t}$, and hence $\{e^{j\omega kh}v(\theta)\}_{k=0}^{\infty}$ is $\sum_{\ell=-\infty}^{\infty} v_\ell e^{j\omega_\ell t}$. By Eq. (21), the asymptotic response of $G(z)$ against this input is given by

$$\begin{aligned} e^{j\omega kh}G(e^{j\omega h})[v] &= e^{j\omega kh}G(e^{j\omega h})\left[\sum_{\ell=-\infty}^{\infty} v_\ell e^{j\omega_\ell\theta}\right] \\ &= \sum_{\ell=-\infty}^{\infty} e^{j\omega kh}G(e^{j\omega h})[e^{j\omega_\ell\theta}]v_\ell \end{aligned} \tag{27}$$

Expand $G(e^{j\omega h})[e^{j\omega_\ell\theta}]$ in terms of $\{e^{j\omega_n\theta}\}$ to get

$$G(e^{j\omega h})[e^{j\omega_\ell\theta}] = \sum_{n=-\infty}^{\infty} g_n^\ell e^{j\omega_n\theta}$$

Substituting this into Eq. (27), we obtain

$$e^{j\omega kh}G(e^{j\omega h})[v] = e^{j\omega kh}\sum_{\ell=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} g_n^\ell e^{j\omega_n\theta}v_\ell$$

Since $e^{j(\omega+n\omega_s)h} = e^{j\omega h}$, this is the kth step response of

$$\sum_{\ell=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} g_n^{\ell} e^{j\omega_n t} v_{\ell}$$

where $t = kh + \theta$. Interchanging the order of summation, this is equal to

$$\sum_{n=-\infty}^{\infty}\left(\sum_{\ell=-\infty}^{\infty} g_n^{\ell}(\omega) v_{\ell}\right) e^{j\omega_n t} \tag{28}$$

This means that the response against $\sum_{\ell=-\infty}^{\infty} v_{\ell} e^{j\omega_{\ell} t}$ is again expressible as an infinite sum of all such aliased signals. It should be intuitively clear that the largest gain among them again gives the gain of the frequency response, when such signals are equipped with norm $(\sum_n |v_{\ell}|^2)^{1/2}$. This isometric correspondence is guaranteed by the Parseval identity Eq. (26). This is the viewpoint adopted in Refs. 14 and 15 to discuss the frequency response of sampled-data systems; see also Ref. 16. It is also closer to the classical treatment based on the impulse modulation (3,7).

Gain Computation

The gain function $G(e^{j\omega h})$ is given as the operator norm at each frequency, and its computation is primarily an infinite-dimensional problem. However, for most of the practical purposes, it can be computed as the maximal singular value (17).

Our problem is thus reduced to that of solving the singular value equation

$$[\gamma^2 I - G^* G(e^{j\omega h})]w = 0 \tag{29}$$

This is still an infinite-dimensional equation. However, since $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$ are finite-rank operators, we can reduce this to a finite-dimensional rank condition. Note that, by lifting, a realization of $G(z)$ can be written in the form

$$\begin{aligned} x[k+1] &= \mathcal{A}x[k] + \mathcal{B}w[k] \\ y[k] &= \mathcal{C}x[k] + \mathcal{D}w[k] \end{aligned}$$

Its adjoint can then be easily derived as

$$\begin{aligned} p[k] &= \mathcal{A}^* p[k+1] + \mathcal{C}^* v[k] \\ e[k] &= \mathcal{B}^* p[k+1] + \mathcal{D}^* v[k] \end{aligned}$$

Taking the z transforms of both sides, setting $z = e^{j\omega h}$, and substituting $v = y$ and $e = \gamma^2 w$, we get

$$\begin{aligned} e^{j\omega h} x &= \mathcal{A}x + \mathcal{B}w \\ p &= e^{j\omega h}\mathcal{A}^* p + \mathcal{C}^*(\mathcal{C}x + \mathcal{D}w) \\ (\gamma^2 - \mathcal{D}^*\mathcal{D})w &= e^{j\omega h}\mathcal{B}^* p + \mathcal{D}^*\mathcal{C}x \end{aligned}$$

Solving these, we obtain

$$\left(e^{j\omega h}\begin{bmatrix} I & \mathcal{B}R_\gamma^{-1}\mathcal{B}^* \\ 0 & \mathcal{A}^* + \mathcal{C}^*\mathcal{D}R_\gamma^{-1}\mathcal{B}^* \end{bmatrix} - \begin{bmatrix} \mathcal{A} + \mathcal{B}R_\gamma^{-1}\mathcal{D}^*\mathcal{C} & 0 \\ \mathcal{C}^*(I + \mathcal{D}R_\gamma^{-1}\mathcal{D}^*)\mathcal{C} & I \end{bmatrix}\right)\begin{bmatrix} x \\ p \end{bmatrix} = 0 \tag{30}$$

where $R_\gamma = (\gamma I - \mathcal{D}^*\mathcal{D})$. The important point to be noted here is that all the operators appearing here are actually matrices. For example, by checking the domain and range spaces, we easily see that $\mathcal{B}R_\gamma^{-1}\mathcal{B}^*$ is a linear operator from $\mathbb{R}^n$ into itself, i.e., a matrix. Therefore, in principle, one can solve the singular value Eq. (29) by finding a nontrivial solution for Eq. (30) (provided R_γ is invertible) (17,18).

H^∞/H^2 CONTROL PROBLEMS

A significant consequence of the modern approach to sampled-data control is that various robust control problems such as H^∞/H^2 control problems are now completely solved. The problem was initiated by Chen and Francis (19) and later solved in Refs. 9, 10, and 20–22 in more complete forms; see Ref. 5 for the pertinent historical accounts.

To state the problem more precisely, let us introduce the notion of generalized plants. Suppose that a continuous-time plant is given in the following form:

$$\begin{aligned} \dot{x}_c(t) &= Ax_c(t) + B_1 w(t) + B_2 u(t) \\ z(t) &= C_1 x_c(t) + D_{11} w(t) + D_{12} u(t) \\ y(t) &= C_2 x_c(t) \end{aligned}$$

Here w is the exogenous input, $u(t)$ is the control input, $y(t)$ is the measured output, and $z(t)$ is the controlled output. The controller is of the following form:

$$\begin{aligned} x_d[k+1] &= A_d x_d[k] + B_d \mathcal{S} y[k] \\ v[k] &= C_d x_d[k] + D_d \mathcal{S} y[k] \\ u[k](\theta) &= H(\theta) v[k] \end{aligned}$$

where $H(\theta)$ is a suitable hold function. This is shown in Fig. 9. The objective here is to design or characterize a controller that achieves a prescribed performance level $\gamma >> 0$ in such a way that

$$\|T_{zw}\|_\infty < \gamma \tag{31}$$

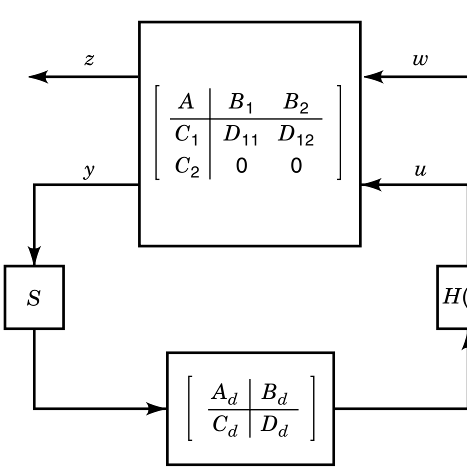

Figure 9. Generalized plant construction of a sampled-feedback system where z denotes the controlled output, y is the measured output, w is the exogenous input, and u is the control input.

where T_{zw} denotes the closed-loop transfer function from w to z. This is the H^∞ control problem for sampled-data systems. The H^2 control problem is obtained by replacing the H^∞ norm above by the H^2 norm.

The difficulty here is that both w and z are continuous-time variables, and hence their lifted variables are infinite-dimensional. A remarkable fact here is that the H^∞ problem (and H^2 problem as well) [Eq. (31)] can be equivalently transformed to the H^∞ problem for a finite-dimensional discrete-time system. While we skip the details here [see the references above and (5)], we remark that this norm-equivalent discrete-time system is entirely different from the one given in the section on "Discrete-Time Design" in that it fully takes intersampling behavior into account. The difference will be exhibited by the design examples in the next section.

SOME EXAMPLES

To see the power of the modern design methods, let us consider two design examples. We start with the H^∞ design.

H^∞ Design

Consider the unstable second-order plant

$$P(s) := C_p(sI - A_p)^{-1}B_p = \frac{1}{s^2 - 0.1s + 1}$$

with weight matrices

$$Q^{1/2} = 1 \quad R^{1/2} = 0.01 \quad E = 0.01 \quad N = 0.01$$

and the antialiasing filter

$$\hat{F}_{aa}(s) := \frac{1}{hs + 1}$$

depicted in Fig. 10.

We here compare the following two design results:

The direct sampled-data H^∞ design

The continuous-time H^∞ design with Tustin transformation

In the continuous-time design, the antialiasing filter is bypassed. On the other hand, it is inserted in the sampled-data design to make the total design well posed.

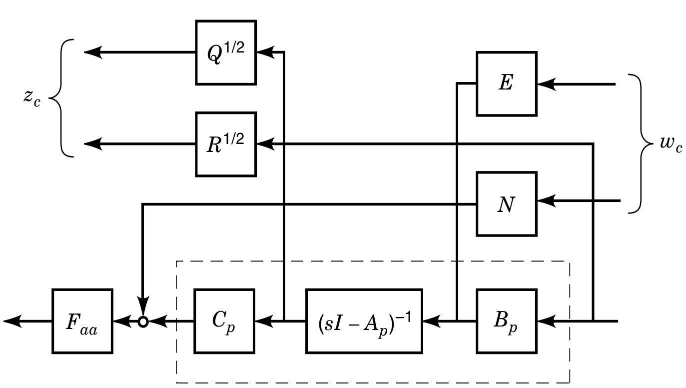

Figure 10. Generalized plant for sampled-data and continuous-time H^∞ design.

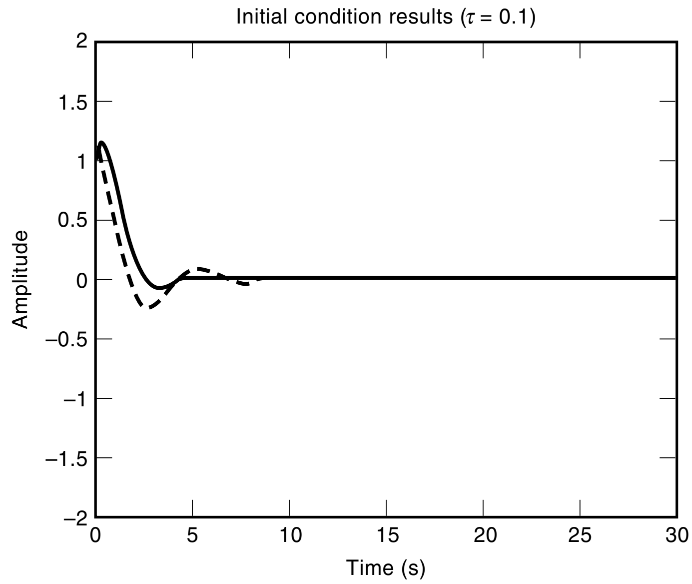

Figure 11. Time responses for $h = 0.1$ by sampled-data (solid) and continuous-time (dash) H^∞ designs do not show much difference.

Figure 11 shows the impulse responses of the designed closed loop for the sampling period $h = 0.1$. The solid curve represents the response for the sampled-data design and the dashed curve shows that for the continuous-time design with Tustin transformation. They do not present much difference at this stage. However, when we increase the sampling period (i.e., decrease the sampling rate) to $h = 0.55$ (Fig. 12), the continuous-time design is already very close to the stability margin. In the conventional design, one may conclude that this sampling period is already too long, and the whole configuration is not feasible for sampled-data implementation. But quite contrary to such an intuition, the sampled-data H^∞ design can tolerate such a long sampling period. The crucial

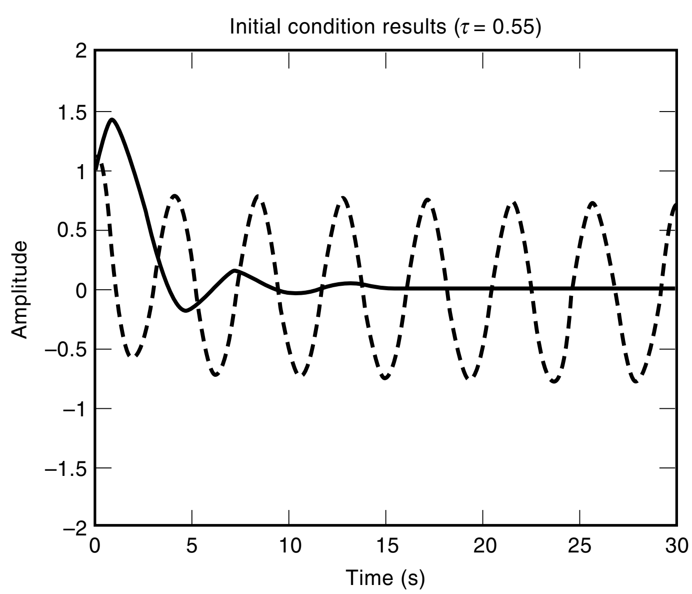

Figure 12. Time responses for $h = 0.55$ exhibit a clear difference between sampled-data (solid) and continuous-time (dash) H^∞ designs.

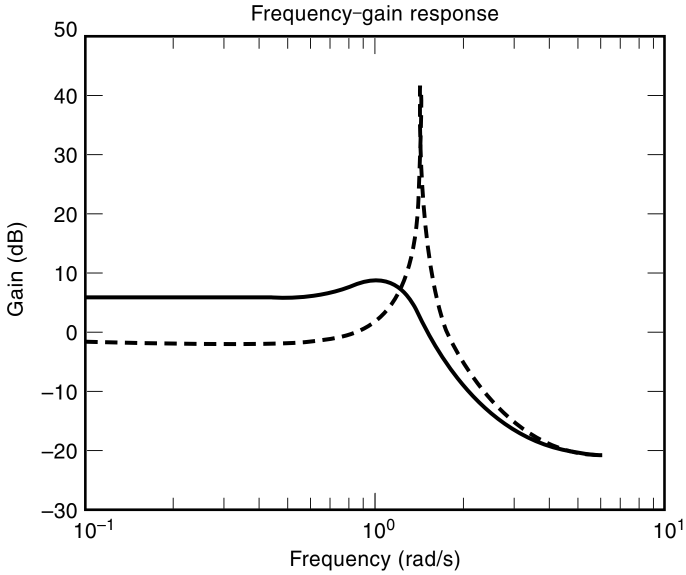

Figure 13. Frequency response plots for $h = 0.55$ support the observation in Fig. 11.

difference here is that the sampled-data design incorporates the sampling period in the design procedure, whereas the continuous-time design does not. This gap becomes even clearer when we compare two designs via their frequency responses (Fig. 13). Whereas the sampled-data design exhibits a rather mild curve, the continuous-time design shows a very sharp peak at around 1.5 rad/s. Observe also that this frequency agrees precisely with the period of oscillation in the impulse response (Fig. 13).

H^2 Design

In the case of continuous-time design, slower sampling rates yield problems. For the sample-point discretization, fast sampling rates can induce very wild responses.

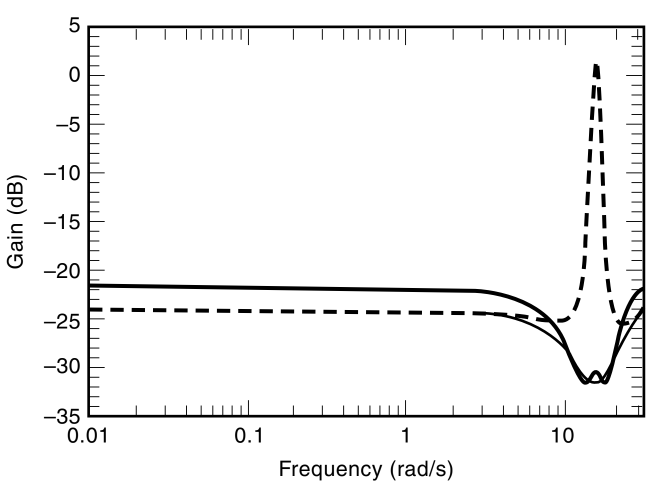

Figure 14. Frequency response plots show the difference between the sampled-data (solid), discrete-time (dashed) difference; the thin dotted curve shows the frequency response with intersample behavior ignored.

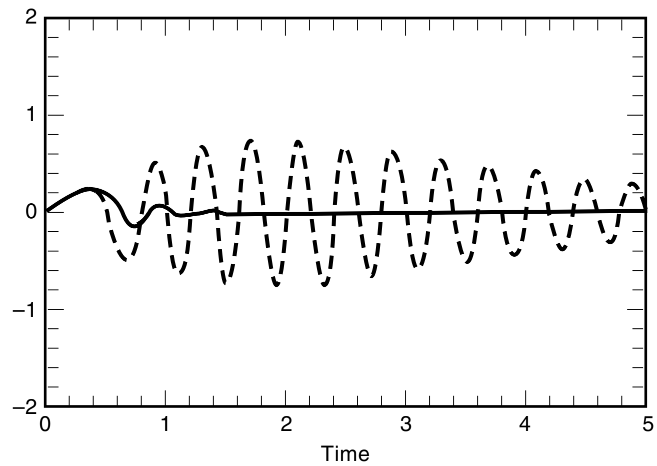

Figure 15. Time responses for sampled-data (solid) and discrete-time (dash) design show a clear difference.

Consider a simple second-order plant $P(s) = 1/(s^2 + 2s + 1)$. For $h = 0.2$, we execute

Sampled-data (continuous-time based) H^2 design
Discrete-time H^2 design

Figures 14 and 15 show the frequency and time responses of the closed-loop systems, respectively. In Fig. 14, the solid (thick) curve shows the response of the sampled-design, whereas the dotted (thin) curve shows the discrete-time frequency response when the designed controller K is connected with the discretized plant G_d (i.e., a purely discrete-time frequency response). At a first glance, it appears that the discretized design performs better, but actually it performs poorer when we compute the real (continuous-time) frequency response of G connected with K_d. The dashed curve shows this frequency response; it is similar to the discrete-time frequency response in the low-frequency range but exhibits a very sharp peak at the Nyquist frequency ($\pi/h \sim 15.7$ rad/s, i.e., $1/2h = 2.5$ Hz).

In fact, the impulse responses in Fig. 15 exhibit a clear difference between them. The solid curve shows the sampled-data design, and the dashed curve the discrete-time one. The latter shows an oscillatory response. Also, both responses decay to zero very rapidly at sampled instants. The difference is that the latter exhibits very large ripples, with periods of approximately 0.4 s. This corresponds to 1/0.4 Hz, which is the same as $(2\pi)/0.4 = \pi/h$ rad/s, i.e., the Nyquist frequency. This is precisely captured in the modern (lifted) frequency response in Fig. 14.

It is worth noting that when h is smaller, the response for the discrete-time design becomes even more oscillatory, and shows a very high peak in the frequency response. The details may be found in Ref. 23.

BIBLIOGRAPHICAL NOTES

For classical treatments of sampled-data control, it is instructive to consult Refs. 2, 3, and 7. The textbooks (24,25) cover both classical and modern aspects of digital control. For discrete-time design methods and other related topics, the reader is also referred to the handbook (26). For Shannon's sampling theorem, consult Ref. 27 for various extensions and some historical accounts.

As noted in the main text, discrete-time lifting has been introduced and rediscovered by several authors (see, for example, Refs. 13, 28, and 29). Aside from the H^∞ design, the H^2 control problem has also been studied extensively (30–33). L^1-norm problems are studied in Refs. 34–36. In relation to the H^∞ control problem, various robust stability problems have been studied (see, for example, Refs. 15 and 37).

The treatment of frequency response given here is based on Refs. 16 and 17. Some computational aspects are discussed in Ref. 38. For other approaches not dealt with here, see also Refs. 39 and 40.

BIBLIOGRAPHY

1. R. E. Kalman and J. E. Bertram, A unified approach to the theory of sampling systems, *J. Franklin Inst.,* **267**: 405–436, 1959.
2. E. I. Jury, *Sampled-Data Control Systems,* New York: Wiley, 1958.
3. J. R. Ragazzini and G. F. Franklin, *Sampled-Data Control Systems,* New York: McGraw-Hill, 1958.
4. R. E. Kalman, Y. C. Ho, and K. Narendra, Controllability of linear dynamical systems, *Contrib. Differ. Equations,* **1**: 189–213, 1963.
5. T. Chen and B. A. Francis, *Optimal Sampled-Data Control Systems,* New York: Springer, 1995.
6. A. H. Zemanian, *Distribution Theory and Transform Analysis,* New York: Dover, 1987.
7. J. T. Tou, *Digital and Sampled-Data Control Systems,* New York: McGraw-Hill, 1959.
8. B. Bamieh et al., A lifting technique for linear periodic systems with applications to sampled-data control systems, *Syst. Control Lett.,* **17**: 79–88, 1991.
9. G. Tadmor, Optimal $\mathcal{H}_\infty$ sampled-data control in continuous time systems, *Proc. ACC,* 1991, pp. 1658–1663.
10. H. T. Toivonen, Sampled-data control of continuous-time systems with an $\mathcal{H}_\infty$ optimality criterion, *Automatica,* **28**: 45–54, 1992.
11. Y. Yamamoto, New approach to sampled-data systems: A function space method, *Proc. 29th CDC,* 1990, pp. 1882–1887.
12. Y. Yamamoto, A function space approach to sampled-data control systems and tracking problems, *IEEE Trans. Autom. Control,* **AC-39**: 703–712, 1994.
13. B. Friedland, Sampled-data control systems containing periodically time varying members, *Proc. 1st IFAC Congr.,* 1961, pp. 361–367.
14. M. Araki, Y. Ito, and T. Hagiwara, Frequency response of sampled-data systems, *Automatica,* **32**: 483–497, 1996.
15. G. Dullerud and K. Glover, Robust stabilization of sampled-data systems to structured LTI perturbations, *IEEE Trans. Autom. Control,* **AC-38**: 1497–1508, 1993.
16. Y. Yamamoto and M. Araki, Frequency responses for sampled-data systems—their equivalence and relationships, *Linear Algebra Its Appl.,* **205–206**: 1319–1339, 1994.
17. Y. Yamamoto and P. P. Khargonekar, Frequency response of sampled-data systems, *IEEE Trans. Autom. Control,* **AC-41**: 166–176, 1996.
18. Y. Yamamoto, On the state space and frequency domain characterization of H^∞-norm of sampled-data systems, *Syst. Control Lett.,* **21**: 163–172, 1993.
19. T. Chen and B. A. Francis, On the $\mathcal{L}_2$-induced norm of a sampled-data system, *Syst. Control Lett.,* **15**: 211–219, 1990.
20. B. Bamieh and J. B. Pearson, A general framework for linear periodic systems with applications to H_∞ sampled-data control, *IEEE Trans. Autom. Control,* **AC-37**: 418–435, 1992.
21. P. T. Kabamba and S. Hara, Worst case analysis and design of sampled data control systems, *IEEE Trans. Auto. Control,* **AC-38**: 1337–1357, 1993.
22. N. Sivashankar and P. P. Khargonekar, Characterization and computation of the $\mathcal{L}_2$-induced norm of sampled-data systems, *SIAM J. Control. Optim.,* **32**: 1128–1150, 1994.
23. S. Hara, Y. Yamamoto, and H. Fujioka, Modern and classical analysis/synthesis methods in sampled-data control—a brief overview with numerical examples, *Proc. 35th IEEE CDC,* 1996, pp. 1251–1256.
24. K. J. Åström and B. Wittenmark, *Computer Controlled Systems—Theory and Design,* Upper Saddle River, NJ: Prentice-Hall, 1996, 3rd ed.
25. D. Williamson, *Digital Control and Implementation,* New York: Prentice-Hall, 1991.
26. W. S. Levine (ed.), *The Control Handbook,* Boca Raton, FL: CRC Press, 1996.
27. A. I. Zayed, *Advances in Shannon's Sampling Theory,* Boca Raton, FL: CRC Press, 1993.
28. J. H. Davis, Stability conditions deried from spectral theory: Discrete systems with periodic feedback, *SIAM J. Control,* **10**: 1–13, 1972.
29. P. P. Khargonekar, K. Poolla, and A. Tannenbaum, Robust control of linear time-invariant plants using periodic compensation, *IEEE Trans. Autom. Control,* **AC-30**: 1088–1096, 1985.
30. T. Chen and B. A. Francis, $\mathcal{H}_2$-optimal sampled-data control, *IEEE Trans. Autom. Control,* **AC-36**: 387–397, 1991.
31. P. P. Khargonekar and N. Sivashankar, $\mathcal{H}_2$ optimal control for sampled-data systems, *Syst. Control Lett.,* **17**: 425–436, 1991.
32. B. Bamieh and J. B. Pearson, The $\mathcal{H}_2$ problem for sampled-data systems, *Syst. Control Lett.,* **19**: 1–12, 1992.
33. S. Hara, H. Fujioka, and P. T. Kabamba, A hybrid state-space approach to sampled-data feedback control, *Linear Algebra Its Appl.,* **205–206**: 675–712, 1994.
34. G. Dullerud and B. A. Francis, $\mathcal{L}_1$ performance in sampled-data systems, *IEEE Trans. Autom. Control,* **AC-37**: 436–446, 1992.
35. N. Sivashankar and P. P. Khargonekar, Induced norms for sampled-data systems, *Automatica,* **28**: 1267–1272, 1992.
36. B. Bamieh, M. A. Dahleh, and J. B. Pearson, Minimization of the L^∞-induced norm for sampled-data systems, *IEEE Trans. Autom. Control,* **AC-38**: 717–732, 1993.
37. N. Sivashankar and P. P. Khargonekar, Robust stability and performance analysis of sampled-data systems, *IEEE Trans. Autom. Control,* **AC-38**: 58–69, 1993.
38. S. Hara et al., Computational aspects of gain-frequency response for sampled-data systems, *Proc. 34th IEEE CDC,* 1995, pp. 1784–1789.
39. A. Feuer and G. C. Goodwin, Generalized sample hold functions—frequency domain analysis of robustness, sensitivity and intersample difficulties, *IEEE Trans. Autom. Control,* **AC-39**: 1042–1047, 1994.
40. J. S. Freudenberg, R. H. Middleton, and J. H. Braslavsky, Inherent design limitations for linear sampled-data feedback systems, *Int. J. Control,* **61**: 1387–1421, 1995.

YUTAKA YAMAMOTO
Kyoto University

DIGITAL CONTROL SYSTEMS DESIGN. See DISCRETE TIME SYSTEMS DESIGN METHODS.