A Function Space Approach to Sampled Data Control Systems and Tracking Problems

Yutaka Yamamoto, *Senior Member, IEEE*

***Abstract*— This paper presents a new framework for hybrid sampled-data control systems. Instead of considering the state only at sampling instants, this paper introduces a function piece during the sampling period as the state and gives an infinite-dimensional model with such a state space. This gives the advantage that sampled-data systems with built-in intersample behavior can be regarded as linear, time-invariant, discrete-time systems. As a result, the approach makes it possible to introduce such time-invariant concepts as transfer functions, poles, and zeros to the sampled-data systems even with the presence of the intersample behavior. In particular, tracking problems can be studied in this setting in a simple and unified way, and ripples are completely characterized as a mismatch between the intersample reference signal and transmission zero directions. This leads to the internal model principle for sampled-data systems.**

I. Introduction

It is well recognized that digital control provides various advantages over the usual time-invariant feedback controls. Deadbeat control, for example, makes a type of stabilization that cannot be achieved with continuous-time linear time-invariant feedback. We can also implement a much more complex logic in control actions making use of the recent advances in computer technology. For example, it is recognized that multirate sampling and/or generalized hold functions provide much greater capability in control; see, for example, [8], [10].

On the other hand, sampled-data systems give rise to a difficulty not encountered in the classical situation. In designing sampled-data systems, one usually designs the continuous-time system and then designs a discrete-time controller over the discrete-time domain. For example, the fundamental work of [12] shows that as far as the regulation of initial states is concerned, one needs only consider the discrete-time system

$$\begin{aligned} w_{k+1} &= e^{Ah}w_k + \int_0^h e^{A\tau}Bu_k\,d\tau, \\ y_k &= Cw_k \end{aligned} \tag{1}$$

where h is the sampling period and the matrices (A, B, C) are the system matrices of the linear time-invariant plant

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t). \end{aligned} \tag{2}$$

Manuscript received September 13, 1991; revised April 23, 1993. Recommended by Past Associate Editor D. F. Delchamps.

The author is with the Division of Applied Systems Science, Faculty of Engineering, Kyoto University, Kyoto 606-01 Japan.

IEEE Log Number 9215718.

Certainly the stability of a sampled-data control system for the plant (2) is determined by the behavior of (1) (and the discrete-time controller), and this can be determined based only on the information at the sampled instants. A drawback in discretizing the continuous-time plant as above, however, is that it loses the intersample information (or makes it implicit, at least). Since the plant is continuous time, the overall performance must be evaluated in continuous time, but this is difficult once we discretize the plant and make it discrete time. A typical problem is that of ripples in servo control systems. In such a situation, the system is subject to continuous excitation by exogenous signals, and there can remain stationary ripples, even though the error tends to zero at sampled instants. One can compute the intersample behavior *(after the control system design)* via the modified Z-transform, but this is often very annoying due to the complexity of formulas [17].

In view of this, recent investigations of sampled-data systems place more emphasis upon intersample behavior, and there are now a number of investigations along this line: [5] and [3] for stability analysis; [3] and [14] for H^2-optimization; [9], [20], and [1] for H^∞-problem; and [7], [22], [19], [15], and [24] for tracking.

The study of sampled-data systems with built-in intersample behavior induces another technical difficulty. Even though the (continuous-time) plant and the (discrete-time) controller are both time-invariant, the underlying time sets are different, so when combined together to form a hybrid sampled-data control system, the resulting system is not time-invariant and only periodically time-varying. This mixture of two types of time sets makes the analysis of sampled-data systems technically very awkward. In particular, it makes the classical state space formalism ineffective.

To remedy this, we introduce a new framework for sampled-data systems. Since the fundamental issue lies in incorporating the intersample behavior, we take the basic idea of taking the full histories of each sampling period of input/state/output functions as input/state/output vectors and then derive linear, time-invariant, discrete-time transition rules. The difference here is that since we take functions as states, inputs, etc., the system becomes *infinite-dimensional*, but on the other hand it has the advantage that continuous-time systems can be viewed as linear, time-invariant, discrete-time systems. Thus it is particularly suitable for the study of sampled-data systems in that both digital and analog components can be placed into the unified framework of linear, time-invariant, discrete-time framework. This technique is a generalization to continuous-time systems of *lifting* employed by [13] for discrete-time periodic systems. It "lifts" the original system to that considered in a discrete-time yet with larger input/state/output spaces. Unlike the discrete-time case [13], however, the present case involves infinite-dimensional function spaces due to the continuous-time nature of the plant.

This new framework of lifting continuous-time systems has been introduced to control theory by [20], [24] and by [2], [1], independently. While we use infinite-dimensional state space, [20], [2] and [1] use a finite-dimensional state space. This difference actually induces some difference in formulas in the lifted system, and it is to be investigated in the future as to which is more advantageous. While [20] and [1] studies H^∞-type problem for sampled-data systems, we are here concerned with tracking problems, in particular, characterization of ripples.

The paper is organized as follows: Section II introduces a function-space valued sequence space and Z-transforms there. We then define the lifting of the continuous-time linear system in Section III. After giving some facts on stability, stabilizability, and transfer functions in Section IV, we discuss ripple-free tracking conditions in Sections V and VI. Section V deals with the case of tracking signals generated by a simple pole. The key feature here is the notion of transmission zeros and their associated zero directions. This directional vector (naturally a function in the present setting) gives a proper intersample tracking signal, and roughly speaking, ripple-free tracking occurs when and only when this direction coincides with the intersample behavior of the reference signal. Section VI deals with the general case: Here the tracking problem for sampled-data systems becomes more delicate and interesting in that combination of a digital compensator and a hold element yields a continuous-time internal model, so that one can often split the internal model of the exogenous signal to the digital and analog parts. A necessity and also a sufficiency (under a mild assumption) of the internal model principle is obtained.

This paper is the full journal version of the conference paper [24], in which the basic framework and preliminary tracking results were presented.

II. Sequence Space and Z-Transforms

Let h be a fixed sampling period. Our basic idea is to regard a trajectory $x(t)$ as the sequence of functions $\{x_k(\theta)\}_{k=1}^{\infty}$ defined by

$$x_k(\theta) := x((k-1)h+\theta) \tag{3}$$

and give a discrete-time state transition rule of x_k. To rigorously introduce this, we need a few preliminaries.

Let $\tilde{C}[0, \infty)$ denote the space of all piecewise continuous functions on $[0, \infty)$. Similarly, let $\tilde{C}(0, h]$ denote the space of all such functions $\varphi(t)$ on $(0, h]$ with finite right limit at 0. The latter space will be denoted $\mathbf{X}$ in the sequel. $\mathbf{X}$ is clearly a Banach space with supremum norm

$$\|\varphi\| := \sup_{0<\theta\leq h}|\varphi(\theta)|.$$

Defining the mapping $\tilde{\mathcal{S}}$ as

$$\tilde{\mathcal{S}}: \tilde{C}[0, \infty) \to \mathbf{X}^\infty : \varphi \mapsto \{\varphi_k\}_{k=1}^{\infty} : \quad \varphi_k(\theta) := \varphi((k-1)h+\theta),\ 0<\theta\leq h \tag{4}$$

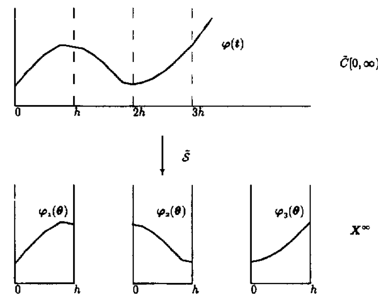


Fig. 1. Mapping $\tilde{\mathcal{S}}$.

we see that $\tilde{C}[0, \infty)$ and the sequence space $\mathbf{X}^\infty$ (the countably many direct product of $\mathbf{X}$) are placed in one-to-one correspondence. This mapping $\tilde{\mathcal{S}}$ "lifts" the continuous-time signal $\varphi(t)$ to an element $\{\varphi_k\}$ residing in the sequence space $\mathbf{X}^\infty$. The space $\mathbf{X}^\infty$ can also be identified with the space of formal power series $\mathbf{X}[[z^{-1}]]$ with coefficients in $\mathbf{X}$ via the *Z-transform*

$$Z: \mathbf{X}^\infty \to \mathbf{X}[[z^{-1}]]$$

$$\{\varphi_k\}_{k=1}^{\infty} \mapsto \sum_{k=1}^{\infty} \varphi_k z^{-k}. \tag{5}$$

With slight abuse of notation, we will also denote $Z(\varphi)$ for $Z(\tilde{\mathcal{S}}(\varphi))$ when $\varphi \in \tilde{C}[0, \infty)$; it may also be denoted by $\hat{\varphi}$. Note that it is a function of θ.

Let us introduce the sampling and hold operations. Let δ_h be the delta function placed at h, i.e., $\langle \delta_h, \varphi\rangle = \varphi(h)$, or $\delta_h \varphi = \varphi(h)$ for $\varphi \in \mathbf{X}$. The sampling operator $\mathcal{S}$ is identified with δ_h acting on $\mathbf{X}$. It has a natural extension: $\mathbf{X}^\infty \to \mathbb{R}^\infty$ (or $\mathcal{S}: \tilde{C}[0, \infty) \to \mathbb{R}^\infty$ with identification $\tilde{\mathcal{S}}: \tilde{C}[0, \infty) \cong \mathbf{X}^\infty$) by

$$\mathcal{S}(\{\varphi_k\}) := \{\delta_h \varphi\}_{k=1}^{\infty} = \{\varphi_k(h)\}_{k=1}^{\infty}. \tag{6}$$

The *hold operator* $\mathcal{H}: \mathbb{R}^\infty \to \mathbf{X}^\infty$ is defined by

$$\mathcal{H}(\{x_k\}) := \{\varphi_k(\theta)\}, \quad \varphi_{k+1}(\theta) \equiv H(\theta)x_k, \quad 0<\theta\leq h \tag{7}$$

where $H(\theta)$ is a fixed hold function. We remark that $Z(\mathcal{S}(\varphi)) = \sum_{k=1}^{\infty} \varphi(kh) z^{-k}$, so it agrees with the usual Z-transform of the sampled data $\{\varphi(kh)\}$. We also remark that

$$(\mathcal{H}\mathcal{S})(\{\varphi_k\})(\theta) = \begin{cases} 0, & k=1 \\ \varphi_{k-1}(h), & k\geq 2 \end{cases} \tag{8}$$

so that the sample-hold operation induces a unit-time delay (forward shift) operator.

We note that the space $\mathbf{X}$ admits an algebra structure with respect to convolution

$$(\varphi * \psi)(\theta) := \int_0^\theta \varphi(\theta-\tau)\psi(\tau)\, d\tau. \tag{9}$$

This naturally induces an algebra structure to $\mathbf{X}[[z^{-1}]]$, which we freely use in sequel. We remark that the multiplication of z and z^{-1} to an element in $\mathbf{X}[[z^{-1}]]$ acts as the one-step backward and forward shift operators, respectively.

III. Function Space Model

Using the lifting $\tilde{S}: \varphi(t) \mapsto \{\varphi_k\}$, we can derive a discrete-time, time-invariant system equation for a continuous-time system. This has the obvious advantage that the underlying time set can be made the same for digital and analog elements.

We start by giving the discrete-time, time-invariant model for a continuous-time system. Let (A_c, B_c, C_c) be a given continuous-time system

$$\begin{aligned} \dot{x}(t) &= A_c x(t) + B_c u(t), \\ y(t) &= C_c x(t), \quad u \in \mathbf{R}^m, \quad x \in \mathbf{R}^n, \quad y \in \mathbf{R}^p. \end{aligned} \tag{10}$$

Let $x_k(\theta) \in \mathbf{X}^n$, $y_k(\theta) \in \mathbf{X}^p$, $u_k(\theta) \in \mathbf{X}^m$ be the functions obtained by applying $\tilde{S}$ to x, y, u, respectively, i.e., $x_k(\theta) = x((k-1)h + \theta)$, (cf. (4)), etc. Suppose that system (10) is at time $t = kh$. Then the past h second history of the state is $x_k(\theta) \in \mathbf{X}^n$. If input $u_{k+1}(\theta) \in \mathbf{X}^m$ is applied on the interval $(kh, (k+1)/h]$, then the state trajectory $x_{k+1}(\theta)$ and the output trajectory $y_k(\theta)$ follows the transition rules

$$\begin{aligned} x_{k+1}(\theta) &= e^{A_c\theta} x_k(h) + \int_0^\theta e^{A_c(\theta-\tau)} B_c u_{k+1}(\tau)\, d\tau, \\ y_k(\theta) &= C_c x_k(\theta), \quad 0 < \theta \le h. \end{aligned} \tag{11}$$

Introducing the operators

$$\begin{aligned} &F: \mathbf{X}^n \to \mathbf{X}^n : x(\theta) \mapsto e^{A_c\theta} x(h), \\ &G: \mathbf{X}^m \to \mathbf{X}^n : u(\theta) \mapsto \int_0^\theta e^{A_c(\theta-\tau)} B_c u(\tau)\, d\tau, \\ &H: \mathbf{X}^n \to \mathbf{X}^p : x(\theta) \mapsto C_c x(\theta) \end{aligned} \tag{12}$$

(11) can be written simply as

$$\begin{aligned} x_{k+1} &= F x_k + G u_{k+1}, \\ y_k &= H x_k. \end{aligned} \tag{13}$$

It is clear that (10) and (11) or (13) give precisely the same input-state/state-output correspondence. There is, however, a point that calls for more attention. On the right-hand side of (13), the input term is u_{k+1}, not u_k. Therefore, in the strict sense, the quantity $x_{k+1}(\theta)$ does not satisfy strict causality. To remedy this, one can introduce a new state variable

$$\xi_k := x_k - G u_k.$$

With respect to this new state, (13) can be rewritten as

$$\begin{aligned} \xi_{k+1} &= F\xi_k + FGu_k \\ y_k &= H\xi_k + HGu_k \end{aligned} \tag{14}$$

and satisfies causality. In what follows, however, we will use mainly (13) because the unit-time delay induced by the hold operation just cancels this effect of u_{k+1}. *Also, when a continuous-time system Σ_c is viewed as a discrete-time system under the lifting here, we will denote it by $\tilde{S}(\Sigma_c)$ or by $\tilde{\Sigma}_c$.*

Consider the hybrid control system depicted in Fig. 2. Here Σ_d and Σ_c denote discrete-time and continuous-time systems (A_d, B_d, C_d) and (A_c, B_c, C_c), respectively. Under the identification $\tilde{S}$ introduced in the previous section, we note that at $t = kh$, the sampled values of signals e_k, x_k, y_k, r_k, etc., are $e_k(h)$, $x_k(h)$, $y_k(h)$, $r_k(h)$, respectively. It follows

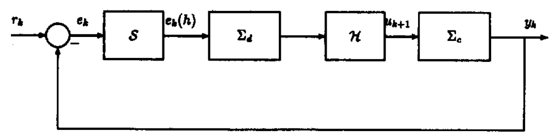


Fig. 2. Hybrid closed-loop system Σ_{cl}.

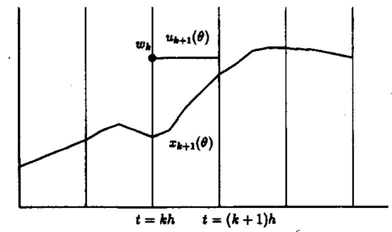


Fig. 3. Signal timing.

that the closed-loop system, denoted Σ_{cl} in the sequel, obeys the equations:

$$\begin{aligned} w_{k+1} &= A_d w_k + B_d e_k(h), \\ u_{k+1}(\theta) &\equiv H(\theta) C_d w_k. \end{aligned} \tag{15}$$

$$\begin{aligned} x_{k+1}(\theta) &= e^{A_c\theta} x_k(h) + \int_0^\theta e^{A_c(\theta-\tau)} B_c u_{k+1}(\tau)\, d\tau, \\ y_k(\theta) &= C_c x_k(\theta), \quad 0 < \theta \le h. \end{aligned} \tag{16}$$

and

$$e_k(\theta) = r_k(\theta) - y_k(\theta).$$

The one-unit time shift in the output equation in (15) is a result of hold operation (see Fig. 3 for the precise correspondence of timing of each signal). We then obtain the following model for Σ_{cl}

$$\begin{bmatrix} w_{k+1} \\ x_{k+1}(\theta) \end{bmatrix} = \begin{bmatrix} A_d & -B_d C_c \delta_h \\ B(\theta) C_d & e^{A_c\theta} \delta_h \end{bmatrix} \begin{bmatrix} w_k \\ x_k(\theta) \end{bmatrix} + \begin{bmatrix} B_d \delta_h \\ 0 \end{bmatrix} r_k(\theta) \tag{17}$$

$$e_k(\theta) = [0 \ \ -C_c] \begin{bmatrix} w_k \\ x_k(\theta) \end{bmatrix} + r_k(\theta) \tag{18}$$

$$y_k(\theta) = [0 \ \ C_c] \begin{bmatrix} w_k \\ x_k(\theta) \end{bmatrix} \tag{19}$$

where $B(\theta)$ is given by

$$B(\theta) := \int_0^\theta e^{A_c(\theta-\tau)} B_c H(\tau)\, d\tau. \tag{20}$$

We will denote the first operator on the right-hand side of (17) by $\mathbf{A}$.

Remark 3.1: There are now several different frameworks incorporating intersample behavior to sampled-data systems. Francis and Georgiou [5] and Chen and Francis [3], [4] gave a finite-dimensional, periodically time-varying model. Hara and Kabamba [9], [11] used a mixed discrete-time/continuous-time model. A similar lifting as presented here is also employed by [20], [2] and [1] with a finite-dimensional state space. In this model, $x_k := x(kh)$ is taken as the state, and as a result one needs an infinite-dimensional direct transmission term from u to y to describe intersample behavior. On the other hand, this is built into $x_k(\theta)$ in our framework. While finite-dimensionality of the state has obvious merit—typically the input operator is of finite rank—our framework also has the (dual) advantage that the output equation is simply expressed as $y_k(\theta) = C_c x_k(\theta)$, whereas an infinite-dimensional operator is needed in the former approach. The present framework also naturally exhibits the nature of sampling: it is an evaluation operator δ_h (delta function). Of course, these two realizations yield entirely the same input/output behavior and hence the transfer function. Actually, in the definition of zeros (Definition 5.2), they yield almost the same relations. Again the difference lies in whether we allow infinite-dimensional operator either in the state or in the output equation.

IV. Spectrum, Stability, Stabilizability and Transfer Matrices

Since our lifted model (17)–(19) is, in principle, infinite-dimensional, we need to investigate some basic system properties. Let us start by characterizing the spectrum of system Σ_{cl}.

Theorem 4.1: The spectrum of system Σ_{cl} (17)–(19) is the union of $\{0\}$ and the spectrum of the matrix

$$A_0 := \begin{bmatrix} A_d & -B_d C_c \\ B(h)C_d & e^{A_c h} \end{bmatrix}. \tag{21}$$

Furthermore, except 0, they are all eigenvalues. In other words, the spectrum of the hybrid system is entirely determined by the evaluation at $\theta = h$.

Proof: Let us note that the operator $\mathbf{A}$ in (17) is a bounded operator. It is clearly a finite-rank operator, so it is also compact [18, Theorem 4.18]. Then $\mathbf{A}$ is not continuously invertible because $\mathbf{X}$ is infinite-dimensional so that zero belongs to the spectrum $\sigma(\mathbf{A})$, and any other point in the spectrum is an eigenvalue ([18, Theorems 4.18, 4.25]). To compute the nonzero eigenvalues, suppose that $\lambda \neq 0$ and $(w', x(\theta)')'$ satisfy

$$(\lambda I - \mathbf{A})\begin{bmatrix} w \\ x(\theta) \end{bmatrix} = \begin{bmatrix} \lambda I - A_d & B_d C_c \delta_h \\ -B(\theta)C_d & \lambda I - e^{A_c\theta}\delta_h \end{bmatrix}\begin{bmatrix} w \\ x(\theta) \end{bmatrix} = 0. \tag{22}$$

To solve these equations, first set $\theta = h$

$$(\lambda I - A_d)w + B_d C_c x(h) = 0$$
$$-B(h)C_d w + (\lambda I - e^{A_c h})x(h) = 0.$$

These have a nonzero solution if and only if λ is an eigenvalue of (21) and $(w', x(h)')'$ is a corresponding eigenvector. Once these vectors are found, (22) can be trivially solved as

$$x(\theta) = \frac{1}{\lambda}\{e^{A_c\theta}x(h) + B(\theta)C_d w\}.$$

This completes the proof. □

We now study the internal stability of Σ_{cl}. We say that Σ_{cl} is *(asymptotically) stable* if

$$\|\mathbf{A}^k\| \to 0 \text{ as } k \to \infty.$$

The norm $\|\cdot\|$ is the one induced from that of $\mathbf{X}^n$. This condition holds if and only if $\mathbf{A}^k x$ decay to zero uniformly for all x in the unit ball of $\mathbf{X}^n$.

The following theorem is well known (e.g., [3]), stating that the internal stability of Σ_{cl} is determined by the behavior at the sampled instants, but is given in the time-invariant operator theoretic setting here.

Theorem 4.2: The closed-loop system Σ_{cl} (17)–(19) is internally stable if and only if the matrix A_0 in (21) is a stable matrix.

Proof: Since every nonzero spectrum is an eigenvalue, the condition is clearly necessary. Conversely, suppose A_0 is a stable matrix. Theorem 4.1 implies that the spectral radius of $r_\sigma(\mathbf{A})$ of $\mathbf{A}$ coincides with the magnitude of the largest eigenvalue of (21), which is less than one by hypothesis. Recall the well-known formula for spectral radius ([18, page 235])

$$r_\sigma(\mathbf{A}) = \lim_{k\to\infty}\|\mathbf{A}^k\|^{1/k}.$$

Take any $\epsilon > 0$ such that $r_\sigma(\mathbf{A}) + \epsilon < 1$. Then

$$\|\mathbf{A}^k\| < (r_\sigma(\mathbf{A}) + \epsilon)^k$$

for all sufficiently large k, and hence the conclusion follows. □

Theorem 4.2 shows that if the sampled digital system $(e^{A_c h}, B(h), C_c)$ is stabilized by a digital controller (A_d, B_d, C_d) as a finite-dimensional discrete-time system, then the closed-loop hybrid system Σ_{cl} (17)–(19) is also internally stabilized. This agrees with the classical knowledge, and it shows the consistency of the present framework.

We now discuss the preservation of stabilizability. In doing so, we make the following standard assumption:

Assumption A (Spectrum Nondegeneracy Assumption): None of the eigenvalues of A_c with nonnegative real parts differ by $2k\pi j/h$, $k \neq 0$.

Now consider the sample-holded state feedback

$$u_{k+1}(\theta) = H(\theta)Kx_k(h). \tag{23}$$

This yields the feedback system

$$x_{k+1}(\theta) = (e^{A_c\theta} + B(\theta)K)x_k(h) \tag{24}$$

and as in Theorem 4.2, system (24) is stable if and only if $e^{A_c h} + B(h)K$ is a stable matrix. Therefore, we readily obtain the following theorem.

Theorem 4.3: The system (11) is stabilizable by sampled state feedback (23) if the original continuous-time system (A_c, B_c) is stabilizable and Assumption A is satisfied.

Before closing this section, we give some facts on transfer matrices. Let (F, G, H) be a given system over $\mathbf{X}$:

$$\begin{aligned} x_{k+1} &= Fx_k + Gu_k, \\ y_k &= Hx_k. \end{aligned} \tag{25}$$

Using the algebra structure of $\mathbf{X}[[z^{-1}]]$, especially $\{zx\}_k = x_{k+1}$, we can compute the Z-transform of both sides of (25) to obtain

$$\hat{y} = H(zI - F)^{-1}G\hat{u} \tag{26}$$

where $\hat{}$ denotes Z-transform. Allowing formal division for $(zI - F)^{-1}$, $H(zI - F)^{-1}G$ agrees with the Z-transform of the sequence

$$HG, HFG, \cdots, HF^{k-1}G, \cdots$$

and multiplication of $H(zI - F)^{-1}G$ with u gives the output in the Z-transform domain. On the other hand, by the standard technique of considering characteristic solutions of type $x_k = \lambda^{k-1}x_0(\theta)$, one can consider substituting a complex number λ to (26), and using the fact that $zx_k = x_{k+1} = \lambda x_k$, it is obvious that the solution is obtained by solving

$$\begin{aligned} \lambda x_k &= Fx_k + Gu_k, \\ y_k &= Hx_k. \end{aligned}$$

In this case $(\lambda I - F)^{-1}$ appearing in the solution is the resolvent operator. Therefore, by this formal agreement, the transfer function (26) may be computed by formally substituting z into $H(\lambda I - F)^{-1}G$ and vice versa.

We now compute the transfer function of Σ_{cl} (17)–(18). This will be given by solving

$$\begin{bmatrix} \lambda I - A_d & B_d C_c \delta_h \\ -B(\theta)C_d & \lambda I - e^{A_c\theta}\delta_h \end{bmatrix} \begin{bmatrix} w \\ x(\theta) \end{bmatrix} = \begin{bmatrix} B_d r(h) \\ 0 \end{bmatrix} \tag{27}$$

and

$$e(\theta) = r(\theta) - C_c x(\theta).$$

Suppose λ belongs to the resolvent set of matrix A_0 (21). Let $(w_0', x_0')'$ be the solution of $(\lambda I - A_0)(w_0', x_0')' = ((B_d r(h))', 0)'$:

$$\begin{bmatrix} w_0 \\ x_0 \end{bmatrix} = \begin{bmatrix} \lambda I - A_d & B_d C_c \\ -B(h)C_d & \lambda I - e^{A_c h} \end{bmatrix}^{-1} \begin{bmatrix} B_d r(h) \\ 0 \end{bmatrix}. \tag{28}$$

The second row of (27) yields

$$x(\theta) = \frac{1}{\lambda}\{B(\theta)w_0 + e^{A_c\theta}x_0\}.$$

Hence

$$\hat{e}(\theta) = \hat{r}(\theta) - \frac{1}{\lambda}C_c[B(\theta)C_d\, e^{A_c\theta}] \cdot \begin{bmatrix} \lambda I - A_d & B_d C_c \\ -B(h)C_d & \lambda I - e^{A_c h} \end{bmatrix}^{-1} \begin{bmatrix} B_d \hat{r}(h) \\ 0 \end{bmatrix}. \tag{29}$$

This correspondence $\hat{r}(z) \mapsto \hat{e}(z)$ (or $\hat{y}(z)$ with suitable changes) will be referred to as the *closed-loop transfer matrix operator* of system Σ_{cl}. Combining (27) with (28), we see $x(h) = (1/\lambda)\{B(h)w_0 + e^{A_c h}x_0\} = x_0$. Therefore, we have the sampled-transfer matrix

$$\hat{e}(h) = \hat{r}(h) - [0\ \ C_c] \cdot \begin{bmatrix} \lambda I - A_d & B_d C_c \\ -B(h)C_d & \lambda I - e^{A_c h} \end{bmatrix}^{-1} \begin{bmatrix} B_d \hat{r}(h) \\ 0 \end{bmatrix}. \tag{30}$$

This shows that, at sampled instants, the closed-loop transfer function is obtained by composing the discrete-time (finite-dimensional) transfer functions in the usual way, and this guarantees consistency of the present method with the classical approach.

V. Zeros and Tracking

One of the problems in digital control is that it may induce ripples. If we focus our attention only on the discrete-time model describing the behavior at the sampled instants, it is difficult to analyze ripples, since during the intersample periods the system works as an open-loop system. This makes it difficult to obtain the internal model principle for servo control problems as obtained by [6] (see [7], [22] for some extensions to the case of ripple-free tracking). Taking advantage of the present framework, however, we can give a much clearer view on the ripples. In this section, we first study the case where the reference signals are generated by a simple pole and give a ripple-free tracking condition for this case. The general case where the reference signals are generated by systems with multiple poles is studied in the next section.

As it turns out, making use of the fact that the present approach allows transfer matrices, we can naturally talk about zeros of the transfer matrix, and then the ripple behavior can be clearly understood as a result of the mismatch between a directional function intrinsically associated with a transmission zero and the intersample function. To see the idea, let $W(z)$ be a transfer matrix from the reference signal r to the error e of a finite-dimensional, discrete-time closed-loop system. Suppose that the tracking signal r is $v/(z - \lambda)$ where v is a vector giving the direction into the input channels. It is well known that the error e tends to zero if and only if

$$W(\lambda)v = 0.$$

That is, not only $W(z)$ has λ as a transmission zero, but also should admit v as a directional vector associated to this zero λ.

Precisely the same result holds in the present situation, only that vector v must be replaced by a function $v(\theta)$. We start with the following lemma.

Theorem 5.1: Let $W_{er}(z)$ be the closed-loop transfer matrix operator of Σ_{cl} from the reference r to the error e given by (29). Suppose it is stable. Then its response to the input $\{\lambda^{k-1}v(\theta)\}$, $|\lambda| \geq 1$, asymptotically approaches

$$\lambda^{k-1}W_{er}(\lambda)v(\theta). \tag{31}$$

In particular,

1) there is no stationary ripple if and only if $W_{er}(\lambda)v(\theta) \equiv 0$;
2) when $\lambda = 1$, the stationary ripple is given by $W_{er}(\lambda)v(\theta)$;

3) if $|\lambda| > 1$ and if $W_{er}(\lambda)v(\theta) \not\equiv 0$, then the response will always diverge.

Proof: Since $W_{er}(z)$ is stable, we can disregard the initial state response. Expand $W_{er}(z)\{v(\theta)/(z-\lambda)\}$ as

$$W_{er}(z)\frac{v(\theta)}{z-\lambda} = \frac{W_{er}(\lambda)v(\theta)}{z-\lambda} + R(z). \tag{32}$$

According to Theorem 4.2 and formula (29), the second term $R(z)$ is analytic outside the unit disk $\{z;\ |z|<1\}$. Therefore, the response (32) asymptotically approaches $\lambda^{k-1}W_{er}(\lambda)v(\theta)$. This readily yields the conclusions. □

In summary,

1) asymptotic tracking to $\lambda^{k-1}v(\theta)$ is possible only when $W_{er}(\lambda)v(\theta) \equiv 0$;
2) when this condition does not hold, and if $|\lambda| > 1$, ripples always diverge.

The second property is suggestive. In practice, what we can implement precisely are almost exclusively digital devices. Then, unless the continuous-time plant $\tilde{\Sigma}_c$ possesses λ as a pole, the tracking condition $W_{er}(\lambda)v(\theta) \equiv 0$ is hardly satisfied, so that the intersample ripples always diverge; see also [16]. (Tracking at sampled points is still possible by a digital compensator.) This suggests that from the engineering point of view, tracking to $e^{\mu t}$, $Re\,\mu > 0$ is impractical, and the only practical cases are tracking to steps, ramps, and sinusoids.

In what follows we specialize Theorem 5.1 to the particular form of (29). We start with the definition of zeros. We first make the following assumption throughout.

Assumption B: Consider the closed-loop system Σ_{cl} in Fig. 2. Let $\mathcal{C}(z)$ be the transfer function of Σ_d, and $\mathcal{P}(s)$ the transfer function of Σ_c, with lifted (discrete-time) transfer function $\tilde{\mathcal{P}}(z)$ of the lifted system $\tilde{\Sigma}_c$ of Σ_c. We assume that

- for any unstable pole λ of $\mathcal{C}(z)$, no $v \in \mathbf{X}$ satisfies $\tilde{\mathcal{P}}(\lambda)v = 0$, and
- for any unstable pole μ of $\mathcal{P}(s)$, $e^{\mu h}$ is not a transmission zero of $\mathcal{C}(z)$.

Since the poles of $\tilde{\mathcal{P}}(z)$ are precisely $e^{\mu h}$ with μ being poles of $\mathcal{P}(s)$, this assumption simply requires that there be no unstable pole-zero cancellation between Σ_d and $\tilde{\Sigma}_c$.

Definition 5.2: A complex number λ is an *invariant zero* and a function $v(\theta) \in \mathbf{X}^m$ is an *associated zero direction* of the closed-loop system (17)–(18) if there exists $(w',\ x(\theta)')'$ such that

$$\begin{bmatrix} \lambda I - A_d & B_d C_c \delta_h & B_d \delta_h \\ -B(\theta)C_d & \lambda I - e^{A_c\theta}\delta_h & 0 \\ 0 & C_c & I \end{bmatrix} \begin{bmatrix} w \\ x(\theta) \\ v(\theta) \end{bmatrix} = 0. \tag{33}$$

The operator on the left-hand side is a clear analog of the system matrix in the finite-dimensional case. The difference here is that the matrix depends on the intersample parameter θ. We call this operator the *closed-loop system operator.*

Under Assumption B and closed-loop stability, unstable λ (i.e., $|\lambda| \geq 1$) is an invariant zero with associated zero direction $v(\theta)$ if and only if

$$W_{er}(\lambda)v(\theta) \equiv 0$$

i.e., it is also a transmission zero. Therefore, we can identify these two notions for tracking problems. Let us start by solving (33). The last row of (33) yields

$$B_d C_c x(h) + B_d v(h) = 0$$

so that we must have

$$(\lambda I - A_d)w = 0, \tag{34}$$
$$-B(\theta)C_d w + \lambda x(\theta) - e^{A_c\theta}x(h) = 0. \tag{35}$$

Existence of a nonzero solution to (34) and (35) is a necessary and sufficient condition for λ to be an transmission zero of $W_{er}(z)$ with $v(\theta) = -C_c x(\theta)$ an associated zero direction function. Now (34) and (35) can admit a solution in the following way:

1) there exists a solution $(w',\ x(\theta)')'$ with nonzero w, or
2) a solution of type $(0,\ x(\theta)')'$ with $x(\theta) \not\equiv 0$ exists.

In the first case, w is an eigenvector of Σ_d and λ is a pole of Σ_d. In the second case, λ is a pole of $\tilde{\Sigma}_c$ and $x(\theta)$ is an eigenfunction. This argument also shows that a pole of, Σ_d or $\tilde{\Sigma}_c$ always yields a transmission zero of $W_{er}(z)$ under Assumption B. Therefore, we have the following proposition.

Proposition 5.3: Let $W_{er}(z)$ be the closed-loop transfer matrix operator from r to e of Σ_{cl}, and let Σ_d and $\tilde{\Sigma}_c$ be its digital and continuous parts, respectively. Assume closed-loop stability. Then, under Assumption B, unstable poles of Σ_d or $\tilde{\Sigma}_c$ induce a transmission zero of $W_{er}(z)$ and vice versa.

This proposition says nothing about the zero directions. To elaborate more upon this, we now confine ourselves to the case of zero-order hold $H(t) \equiv 1$ and $r(t) = e^{\mu t}v_0$, $Re\,\mu \geq 0$. The following theorems may be regarded as an internal model principle for the special case of tracking to signals generated by a single pole μ.

We first consider the tracking to steps: $\mu = 0$.

Theorem 5.4: Consider the closed-loop system Σ_{cl} in Fig. 2. Assume that $H(t) \equiv 1$, and Σ_d and Σ_c are stabilizable and detectable. Consider the tracking to the step reference signal (represented in the continuous-time).

$$r(t) = v_0.$$

Assume the closed-loop stability. Then tracking without ripples is possible either by incorporating 0 into Σ_c or by incorporating one into Σ_d as a pole.

Proof: The Z-transform of $r(t)$ is given by $v_0/(z-1)$. Put $v(\theta) \equiv v_0$. Let us first consider the case where zero is a pole of Σ_c. From (33), $v(\theta)$ can be a transmission zero direction if and only if

$$C_c x(\theta) + v(\theta) = 0 \tag{36}$$
$$(I - A_d)w = 0 \tag{37}$$
$$-B(\theta)C_d w + x(\theta) - e^{A_c\theta}x(h) = 0 \tag{38}$$

for some w, $x(\theta)$. Let x_0 be an eigenvector of A_c corresponding to 0. Then (36)–(38) can be easily solved as

$$v(\theta) = -C_c x_0, \quad w = 0, \quad x(\theta) = x_0.$$

In this case $v_0 = -C_c x_0$ gives an allowable trackable direction. (Note that $C_c x_0 \neq 0$ by the detectability of $(A_c,\ B_c)$.)

We next consider the case where one is a pole of Σ_d but not of $\tilde{\Sigma}_c$. As shown above, $v(\theta)$ is a transmission zero direction if and only if (36)–(38) are satisfied. Since one is not a pole of $\tilde{\Sigma}_c$, w must be an eigenvector corresponding to one, and this in turn yields a unique solution

$$x(\theta) = \{e^{A_c\theta}(I - e^{A_c h})^{-1}B(h) + B(\theta)\}C_d w$$

and $v(\theta) = -C_c x(\theta)$ is a corresponding trackable direction. This $x(\theta)$ is indeed a constant because

$$\begin{aligned}&\frac{d}{d\theta}(e^{A_c\theta}(I - e^{A_c h})^{-1}B(h) + B(\theta))\\ &= e^{A_c\theta}(I - e^{A_c h})^{-1}A_c B(h) + e^{A_c\theta}B_c\\ &= e^{A_c\theta}(I - e^{A_c h})^{-1}(e^{A_c h} - I)B_c + e^{A_c\theta}B_c\\ &= 0.\end{aligned}$$

Finally, if zero is not a pole of Σ_c and one is not a pole of Σ_d, then one cannot be a transmission zero of the sampling time closed-loop transfer matrix (30). Hence by the discrete-time internal model principle, the closed-loop system does not track the step reference as above. □

We next consider the tracking to $r(t) = e^{\mu t}v_0$, $Re\,\mu \neq 0$. Since the hold element is a zero-order hold, ripple-free tracking is possible only by incorporating μ into Σ_c as a pole and not so by incorporating a digital internal model $1/(z - e^{\mu h})$ into Σ_d.

Theorem 5.5: Consider the tracking to the reference signal (represented in the continuous time)

$$r(t) = e^{\mu t}v_0, \quad Re\,\mu \geq 0$$

under the same hypotheses on Σ_{cl} and $H(t)$ as in Theorem 5.4. Assume $\mu \neq 0$. Then tracking without stationary ripples is possible only by incorporating μ into Σ_c as a pole.

Proof: The signal $r(t)$ is represented as $\tilde{S}r = \{\lambda^{k-1}e^{\mu\theta}v_0\}_{k=1}^{\infty}$ in the notation in Section II, and its Z-transform is given by

$$\frac{e^{\mu\theta}}{z - \lambda}v_0.$$

Put $v(0) := e^{\mu\theta}v_0$. To prove that ripple-free tracking is not possible by incorporating $\lambda := e^{\mu h}$ into Σ_d, suppose that λ is a pole of Σ_d but not of $\tilde{\Sigma}_c$. As shown above, $v(\theta)$ is a transmission zero direction if and only if

$$C_c x(\theta) + v(\theta) = 0 \quad (39)$$

$$(\lambda I - A_d)w = 0 \quad (40)$$

$$-B(\theta)C_d w + \lambda x(\theta) - e^{A_c\theta}x(h) = 0 \quad (41)$$

for some w, $x(\theta)$. Since λ is not a pole of $\tilde{\Sigma}_c$, this implies that w must be a (nonzero) eigenvector of A_d corresponding to λ, and we have a unique solution

$$x(\theta) = \frac{1}{\lambda}\{e^{A_c\theta}(\lambda I - e^{A_c h})^{-1}B(h) + B(\theta)\}C_d w. \quad (42)$$

Since λ is not an eigenvalue of $e^{A_c h}$, μ is not an eigenvalue of A_c, so that $e^{A_c\theta}$ does not contain the mode of $e^{\mu\theta}$. Since $H(t) \equiv 1$, we have

$$B(\theta) = \int_0^\theta e^{A_c\tau}B_c\,d\tau$$

so that $x(\theta)$ above cannot contain the mode of $e^{\mu\theta}$. Therefore, $v(\theta) = C_c x(\theta)$ cannot be a zero direction associated to λ, so stationary ripples exist.

It remains to show that tracking without ripples is possible by implementing λ as a pole of $\tilde{\Sigma}_c$. In this case we can solve (39)–(41) as

$$v(\theta) = -C_c x(\theta), \quad w = 0, \quad x(\theta) = \frac{1}{\lambda}e^{A_c\theta}x_0$$

where x_0 is an eigenvector of $e^{A_c h}$. Now in view of Assumption A, x_0 is also an eigenvector of A_c with respect to μ. This implies $e^{A_c\theta}x_0 = e^{\mu\theta}x_0$, and hence

$$v(\theta) = -\frac{1}{\lambda}e^{\mu\theta}C_c x_0$$

and such $v(\theta)$ gives a trackable direction. (In particular, $v_0 = -(1/\lambda)C_0 x_0$.) Note also that $C_c x_0 \neq 0$ by the detectability of (A_c, B_c). □

Remark 5.6: Related results have been studied by [7], [22], etc. We note, however, that the present characterization in terms of poles and zeros has been made possible via the notion of transfer matrices resulting from lifting.

VI. Tracking and The Internal Model Principle

The results in the previous section show that tracking to signals generated by a single pole can be well described by poles and zeros. In particular, when the tracking signal is $v(\theta)/(z - 1)$ and if the internal model is in the digital part, then the only trackable intersample signal $v(\theta)$ is necessarily a constant function for the case of zero-order hold; conversely, it is possible to track step functions by incorporating the pole one into the digital compensator Σ_d. This situation occurs because the cascade connection of $1/(z - 1)$ and the hold element somehow works as a continuous-time internal model $1/s$. This is not, however, well described in the previous section.

When the tracking signal is generated by a repeated pole, the situation is more complicated and cannot be easily described by simple pole-zero arguments. For example, take the ramp signal $r(t) = t$. The Z-transform (defined in Section II) of this signal is

$$\frac{h}{(z-1)^2} + \frac{\theta}{(z-1)}.$$

Since this is an output of a continuous-time plant $1/s^2$, however, neither $1/(z-1)^2$ nor $\theta/(z-1)$ can appear independently as the output of this plant. Therefore, we cannot separate the treatment of simple and double poles in such a case. This situation is quite different from the standard situation in the internal model principle for the usual finite-dimensional systems and requires a more elaborate treatment.

Obtaining a tracking condition for the general case in the sampled-data systems has attracted recent research interest: Using a mixed continuous-time/discrete-time model, Hara and Sung [19] derived a state-space necessary and sufficient condition for tracking. A geometric approach was taken by Kawano *et al.* [15] to get a necessary and sufficient condition. Somewhat earlier than these, Franklin and Emami-Naeini [7] gave an internal model principle in a more classical setting, and Urikura and Nagata [22] gave a geometric condition for the case of deadbeat tracking. While they are more of state-space nature, we here look for a condition in the frequency domain.

To see why the standard machinery does not work, let us write the loop transfer function of Fig. 2 as $P(z, \theta)/q(z)$. Here the denominator $q(z)$ is independent of θ by (29). This means that if we consider intersample tracking, its continuous-time behavior is reflected only upon the *numerator* $P(z, \theta)$ and not the denominator $q(z)$. On the other hand, the usual treatment of the internal model principle for finite-dimensional systems expresses it in terms of divisibility between the two denominators of the loop transfer function and the exogenous signal generator. This is clearly not suitable for the present case where the intersample information is represented by the numerator.

To remedy this situation, it is desirable to recover the continuous-time transfer function to prove the internal model principle. To this end, let us first introduce the notion of finite Laplace transforms.

Definition 6.1: Let φ be any function or distribution on the interval $[0, h]$. The *finite Laplace transform*, denoted $\mathcal{L}_h[\varphi]$ is defined by

$$\mathcal{L}_h[\varphi](s) := \int_0^h e^{-s\theta}\varphi(\theta)\, d\theta. \tag{43}$$

The integral must be understood in the sense of distributions if φ is a distribution.

Note that, in view of the well-known Paley–Wiener theorem [21], $\mathcal{L}_h[\varphi](s)$ is always an entire function of exponential type.

The Z-transform of a function ϕ on $[0, \infty)$ and its Laplace transform can be related by the following lemma.

Lemma 6.2: Suppose that ϕ satisfies the estimate

$$\|\phi\|_{[kh,\,(k+1)h]} \le Ce^{\beta k} \tag{44}$$

for some $C, \beta > 0$, where $\|\phi\|_{[kh,\,(k+1)h]}$ is the L^2-norm on $[kh, (k+1)h]$. Then the Laplace transform $\mathcal{L}[\phi]$ exists for $Re\, s > \beta$, and

$$\mathcal{L}_h[\mathcal{Z}[\phi][z]]|_{z=e^{hs}} = e^{hs}L[\phi](s). \tag{45}$$

A remark on (45) is in order. According to the definition of (5), $\mathcal{Z}[\phi][z] = \Sigma\phi_k z^{-k}$, where $\tilde{S}(\phi) = \{\phi_k\}_{k=1}^{\infty}$. Here the multiplication operator by z acts as the one-step left shift: $\{\phi_k\} \mapsto \{\phi_{k+1}\}$. In the time domain, this clearly corresponds to $\phi(t) \mapsto \phi(t+h)$, so that its Laplace transform is e^{hs}. Identity (45) claims that taking the finite Laplace transform of each piece ϕ_k and expanding the sum via the substitution $z \mapsto e^{hs}$ actually yield the Laplace transform. The multiplication by e^{hs} on the right-hand side becomes necessary to account for our definition of Z-transform starting with z^{-1} rather than z^0.

Proof: That ϕ is Laplace transformable and converges absolutely for $Re\, s > \beta$ is obvious from (44). To show (45), observe

$$\begin{aligned}\mathcal{L}[\phi](s) &= \sum_{k=1}^{\infty}\int_0^h \phi_k(\theta)e^{-((k-1)h+\theta)s}\, d\theta \\ &= \sum_{k=1}^{\infty} e^{-(k-1)hs}\mathcal{L}_h[\phi_k](s) \\ &= \mathcal{L}_h\left[z\left(\sum_{k=1}^{\infty} z^{-k}\phi_k\right)\right]\Bigg|_{z=e^{hs}} \\ &= e^{-hs}\cdot\mathcal{L}_h[\mathcal{Z}[\phi][z]]|_{z=e^{hs}}.\end{aligned}$$

The interchange of integral and summation in the second line is justified by the absolute convergence of the Laplace transform. □

Let us now see the finite Laplace transforms of the transfer functions of digital and analog components in Fig. 2. Let $\mathcal{C}(z) = C_d(zI - A_d)^{-1}B_d$ and $\mathcal{P}(s) = C_c(sI - A_c)^{-1}B_c$ be the discrete-time and continuous-time transfer matrices of Σ_d and Σ_c, respectively. By Lemma 6.2, the transfer function of the continuous-time plant in the sense above is $e^{hs}\mathcal{P}(s)$. The finite Laplace transform $\mathcal{L}_h[H](s)$ of the hold function is an entire function of s by the Paley–Wiener theorem [21]. The digital part $\mathcal{C}(z)$ becomes $\mathcal{C}(e^{hs})$ by the finite Laplace transform. Therefore, the loop transfer function (in continuous-time) becomes the product of these three: $G(s) := e^{hs}\mathcal{C}(e^{hs})\mathcal{L}_h[H](s)\mathcal{P}(s)$. This falls into the category of pseudorational transfer functions studied in [23]. Roughly speaking, a transfer function $W(s)$ is said to be *pseudorational* if $W(s)$ admits a fractional representation $W(s) = p(s)/q(s)$ such that $p(s)$ and $q(s)$ are Laplace transforms of distributions with compact support in $(-\infty, 0]$. (There is one more technical condition which does not concern us here [23].) In view of the facts $e^{hs} = \mathcal{L}[\delta_{-h}](s)$, $s = \mathcal{L}[\delta']$, we see immediately that our loop transfer matrix G is pseudorational.

Since $\mathcal{C}(z)$ and $\mathcal{P}(s)$ are both rational, they admit coprime factorizations over polynomials (in z and s, respectively). Also, $\mathcal{L}_h[H](s)$ is itself the Laplace transform of a distribution with compact support in $[0, h]$, and it admits a coprime factorization $e^{hs}\mathcal{L}_h[H](s)/e^{hs}$ (this looks odd, but e^{hs} does not appear as a common factor after multiplication with $\mathcal{L}_h[H](s)$). Furthermore, since $\mathcal{P}(s)$ has only finitely many poles, the product of these three functions admits a coprime factorization $G(s) = Q^{-1}(s)P(s)$ [23]. Now define

$$X^Q := \{x \in L^2_{loc}[0, \infty) : \check{Q} * x|_{(0,\infty)} = 0\}$$

where $\check{Q}$ is the inverse Laplace transform of $Q(s)$, and $*$ denotes convolution. Then the following fact is known.

Fact 6.3 [23]: If $W = Q^{-1}(s)P(s)$ is a coprime factorization, then the closure of the set of all input-free outputs generated by $W(s)$ is precisely X^Q. Furthermore, for another such $D(s)$, $X^D \subset X^Q$ holds if and only if $\check{Q} = \Pi * \check{D}$ for some matrix Π over the ring of distributions with compact support in $(-\infty, 0]$. When this holds, we write $D(s)|Q(s)$.

Let us now consider the tracking problem. Suppose that the reference signal generator is represented by a left coprime factorization $D^{-1}(s)N(s)$ over polynomials. By the fact above, the input-free outputs of this system are precisely X^D which is the same as those generated by $D^{-1}(s)$. Therefore, we may as well take $D^{-1}(s)$ as the reference signal generator. The following theorem gives the necessity of the internal model for the hybrid sampled-data system Fig. 2.

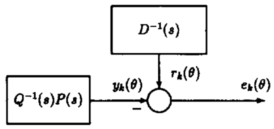


Fig. 4. Steady-state mode.

Theorem 6.4: Consider the unity feedback system given in Fig. 2. Suppose that the exogenous signal is generated by $D^{-1}(s)$ with poles only in the closed right-half plane. Suppose also that the closed-loop system is internally stable, and let $G(s) = Q^{-1}(s)P(s)$ be a coprime factorization of the loop transfer function. If the closed-loop system asymptotically tracks any signal generated by $D^{-1}(s)$, then $D(s)$ divides $Q(s)$.

Proof: Let $r(t)$ be any signal generated by $D^{-1}(s)$, and $y(t)$ the corresponding output tracking $r(t)$. Let us also write r_k and y_k in accordance with the notation in Section II. We first consider the system in the discrete-time mode. If the asymptotic tracking occurs, then it follows that the error $e(t)$ must also converge to zero at the sampling instants, i.e., $e(kh) = e_k(h) \to 0,\ k \to \infty$. By the linearity of the system we can decompose the response as

$$y_k(\theta) = y_k^1(\theta) + y_k^2(\theta)$$

where $y_k^1(\theta)$ is the initial-state response (corresponding to the state at time kh), and $y_k^2(\theta)$ is the response corresponding to the input $e(kh)$. Since the input $e_k(h) \to 0$, $y_k^2(\theta)$ approaches 0 as $k \to \infty$. This means that

$$y_k^1(\theta) - r_k(\theta) \to 0 \text{ as } k \to \infty. \tag{46}$$

(See Fig. 4 for the steady-state interpretation.) Since r is an output generated by $D^{-1}(s)$ and it is shift invariant and arbitrary, (46) and Fact 6.3 imply that there exists a sequence of states x_k in X^Q whose corresponding outputs approach $r(t)$. But since $Q^{-1}P$ is coprime, the closure of the set of outputs generated by this system is precisely X^Q by Fact 6.3. Since X^Q is closed, it follows that

$$X^D \subset X^Q.$$

Again by Fact 6.3, this implies $D(s)|Q(s)$. □

The necessity theorem above is still unsatisfactory because the delay element $\mathcal{C}(e^{hs})$ contains too many poles and zeros. For example, $1/(e^{hs} - 1)$ has $2n\pi j/h,\ n = 0, \pm 1, \cdots$ as poles, but the corresponding digital compensator is just $1/(z - 1)$, and it need not have the capability of producing continuous-time signals that can be produced with $1/(e^{hs} - 1)$. Only in combination with the hold device can the digital part produce continuous-time signals. Therefore, we make the following assumption.

Assumption C: The finite Laplace transform $\psi(s)$ of the hold function $H(t)$ is expressible as the ratio $\psi(s) = \Gamma(z)\Omega^{-1}(z)\Psi(s)^{-1}|_{z=e^{hs}}$ where $\Gamma(z)$ and $\Omega(z)$ are polynomial matrices in z and $\Psi(s)$ is a polynomial matrix in s.

Since $\psi(s)$ must be entire, all zeros in $\Psi(s)$ must be cancelled by $\Gamma(e^{hs})$. This assumption is satisfied for most of the practical cases; for example,

$$\psi(s) = \frac{1 - e^{-hs}}{s} = \frac{z-1}{z} \cdot \frac{1}{s}.$$

Remark 6.5: A similar hypothesis is made in [15]; see also Example 6.8.

Let us now state and prove the sufficiency of our internal model principle.

Theorem 6.6: Let $\mathcal{C}(z)$ and $\mathcal{P}(s)$ be the transfer matrices of the digital and analog parts in Fig. 2, and $\mathcal{C}(z) = P_C(z)Q_C(z)^{-1}$ and $\mathcal{P}(s) = Q_P(s)^{-1}P_P(s)$ be their coprime factorizations over polynomials. Let $\psi(s) = \Gamma(z)\Omega^{-1}(z)\Psi(s)^{-1}|_{z=e^{hs}}$ be the transfer function of the hold element satisfying Assumption C. Suppose that the exogenous signal is generated by $D^{-1}(s)$ with polynomial matrix $D(s)$. Assume the following five conditions:

1) The closed-loop system Σ_{cl} is stable.
2) $Q_C(z) = \Gamma(z)\Pi(z)$ for some polynomial matrix $\Pi(z)$.
3) $Q(s) := Q_P(s)\Psi(s) = N(s)D(s)$ for some polynomial matrix $N(s)$, i.e., $Q(s)$ contains $D(s)$ as an internal model.
4) The minimal realizations of $D^{-1}(s)$ and $Q^{-1}(s)$ both satisfy the spectrum nondegeneracy Assumption A.
5) For any pole μ of $D^{-1}(s)$, $e^{\mu h}$ is not a pole of $\Pi^{-1}(z)\Omega^{-1}(z)$.

Then the closed-loop system asymptotically tracks any signal generated by $D^{-1}(s)$.

Proof: By property 2, the numerator $\Gamma(z)$ of ψ is cancelled by the denominator $Q_C(z)$. This means that by suitably setting an initial state in $\mathcal{C}(z)$, one can get a continuous-time output generated by $\Psi(s)^{-1}$ by combination of the digital element $\mathcal{C}(z)$ with the hold element. Therefore, we can regard the loop transfer function as the cascade connection of $P_C(z)\Pi^{-1}(z)\Omega^{-1}(z)$ and $Q^{-1}(s)P_P(s)$.

Since $D(s)|Q(s)$, if we look at only the sampled instants, the closed-loop system contains the discrete-time internal model. Therefore, by the internal stability, at least tracking at sampled instants occurs. Therefore, the error input to the loop transfer function tends to zero, and the sampled output $y(kh)$ approaches the sampled reference signal $r(kh)$. In the same way as in the proof of Theorem 6.4, we see that there must exist a sequence x_j of initial states in the forward-loop system such that their corresponding responses approach the sampled reference signal $r(kh)$. Therefore, there exists an initial state x in the forward-loop system such that its corresponding output agrees with $r(t)$ at sampled instants $kh,\ k = 1, 2, \cdots$. Now by property 5, the digital part $P_C(z)\Pi^{-1}(z)\Omega^{-1}(z)$ cannot contribute to this output. Therefore, this initial state must lie in the continuous-time part $Q^{-1}(s)P_P(s)$. Thus we return to the steady-state diagram Fig. 4. In this steady-state mode, the output $r(t)$ must be cancelled by $y(t)$ at sampled instants. Since both the exogenous signal generator $D^{-1}(s)$ and the continuous-time part $Q^{-1}(s)$ satisfy Assumption A, however, and since $D(s)$ divides $Q(s)$, if $y(t)$ agrees with $r(t)$ at sampled instants, it must agree with $r(t)$ on the intersample intervals. (Otherwise, there would be $e^{\mu t} \in X^D$ that agrees with some output in X^Q at sampled instants but not on the

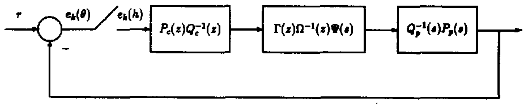


Fig. 5. Closed-loop system.

intersample intervals. But this is clearly impossible under Assumption A, by the linear independence of exponentials.)□

Remark 6.7: The spectrum nondegeneracy condition on both the exogenous signal generator and the augmented plant $Q^{-1}(s)$ cannot be removed in Theorem 6.6. For example, if we have $s = 0$, $2\pi j/h$ both as poles in the continuous-time part, then it may well happen that the closed-loop system may track to the sine wave even if the exogenous signal is a step, because the closed-loop system does not have the capability of knowing the intersample error. The same can be said of the exogenous signal generator.

To summarize: Theorems 6.4 and 6.6 give variants of the continuous-time internal model principle for sampled-data systems. The z and s domains are linked by the finite Laplace transform and the substitution $z = e^{hs}$. With this, the combination of some part of the digital compensator with the hold element can work as a continuous-time system and hence an internal model (cf. Fig. 5). Typical examples now follow.

Example 6.8: The conditions of the theorem are satisfied for the following system with $D(s) = s^2 + \omega^2$:

$$\begin{aligned} \mathcal{C}(z) &= \frac{h}{z-1} \\ H(t) &= \sin(\omega t)|_{[0,\,h]}, \quad \omega = 2\pi/h \\ \psi(s) &= \frac{1-e^{-hs}}{s^2+\omega^2} = \frac{z-1}{z} \cdot \frac{1}{s^2+\omega^2}\bigg|_{z=e^{hs}} \\ \mathcal{P}(s) &= \frac{1}{s+1}. \end{aligned}$$

Here $Q(s) = (s+1)(s^2+\omega^2)$ and contains the internal model $s^2 + \omega^2$. Hence the closed-loop system Fig. 5 can track a sinusoidal wave with $\omega = 2\pi/h$. The internal model is made up with combination of $1/(z-1)$ and the hold element $(1 - e^{-hs})/(s^2 + \omega^2)$.

One of the consequences of the above theorem is that we can incorporate an internal model for sampled-data systems in such a way that we can split poles to analog and digital parts. A typical example is the case of tracking to ramp $r(t) = t$, where we can incorporate one pole $\lambda = 1$ into the digital compensator and the other to the analog plant $\mathcal{P}(s)$:

Example 6.9:

Consider the unity feedback sampled-data system Fig. 5 with

$$\begin{aligned} \mathcal{C}(z) &= \frac{h}{z-1} \\ \psi(s) &= \frac{1-e^{-hs}}{s} = \frac{z-1}{z} \cdot \frac{1}{s}\bigg|_{z=e^{hs}} \\ \mathcal{P}(s) &= \frac{4s+1}{s(s^2+4s+6)}. \end{aligned}$$

It is easy to see that they satisfy the conditions of Theorem 6.6. The Laplace transform of t is $1/s^2$, and the loop transfer function is

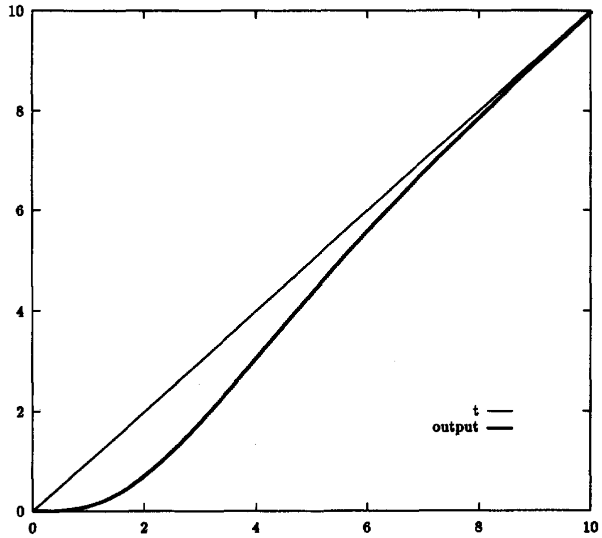


Fig. 6. Tracking to ramp.

$$\frac{h(4s+1)}{e^{hs}s^2(s^2+4s+6)}$$

so that it surely has the internal model $1/s^2$. Here one pole $s = 0$ is in the plant $\mathcal{P}(s)$ and the other is made by the combination of the hold $(1-e^{-hs})/s$ and the digital compensator $h/(z-1)$ through the cancellation $(1 - e^{-hs})/(z-1)|_{e^{hs}=z} = 1/z$. Asymptotic tracking is achieved as seen in Fig. 6.

VII. Concluding Remarks

We have given a function space approach to sampled-data control systems. The introduction of a time-invariant, though infinite-dimensional, model made it possible to discuss tracking and ripples in terms of the familiar notions of poles and zeros and associated zero directions. A new element here is that this theory enables us to regard intersample ripples as the problem of transmission zero directions just as in the case of the finite-dimensional multivariable systems where zero directional vector also governs tracking properties.

The general case allowing multiple poles is more interesting in that it exhibits a nonclassical internal model splitting into digital and analog parts. Although we have made some restrictive assumptions on the type of a hold device for the sufficiency part of the internal model principle, this assumption is satisfied for most practically encountered cases. Generalization to the case of more general hold devices is a topic for future investigation.

Acknowledgment

The author wishes to thank Profs. M. Ikeda and S. Hara for their valuable comments and discussions. He also wishes to thank an anonymous referee for calling his attention to the references [7] and [16].

REFERENCES

[1] B. Bamieh and J. B. Pearson, "A general framework for linear periodic systems with applications to H_∞ sampled-data control," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 418–435, 1992.

[2] B. Bamieh, J. B. Pearson, B. A. Francis, and A. Tannenbaum, "A lifting technique for linear periodic systems with applications to sampled-data control systems," *Syst. Contr. Lett.*, vol. 17, pp. 79–88, 1991.

[3] T. Chen and B. A. Francis, "Input–output stability of sampled-data systems," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 50–58, 1991.

[4] ——, "$\mathcal{H}_2$-optimal sampled-data control," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 387–397, 1991.

[5] B. A. Francis and T. T. Georgiou, "Stability theory for linear time-invariant plants with periodic digital controllers," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 820–832, 1988.

[6] B. A. Francis and W. M. Wonham, "The internal model principle for linear multivariable regulators," *J. Appl. Math. Optim.*, vol. 2, pp. 170–194, 1975.

[7] G. F. Franklin and A. Emami-Naeini, "Design of ripple-free multivariable robust servomechanism," *IEEE Trans. Automat. Contr.*, vol. 31, pp. 661–664, 1986.

[8] T. Hagiwara and M. Araki, "Design of a stable feedback controller based on the multirate sampling of the plant output," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 812–819, 1988.

[9] S. Hara and P. T. Kabamba, "Worst case analysis and design of sampled data control systems," in *Proc. 29th CDC*, 1990, pp. 202–203.

[10] P. T. Kabamba, "Control of linear systems using generalized sampled-data hold functions," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 772–783, 1987.

[11] P. T. Kabamba and S. Hara, "Worst case analysis and design of sampled data control systems,"*IEEE Trans. Automat. Contr.*, vol. 38, pp. 1337–1357, 1993.

[12] R. E. Kalman and J. E. Bertram, "General synthesis procedure for computer control of single and multiloop linear systems," *AIEE Trans.*, vol. 77, pp. 602–609, 1959.

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

[14] P. P. Khargonekar and N. Sivashankar, "$\mathcal{H}_2$ optimal control for sampled-data systems," preprint, Department of Electrical Engineering, Univ. Michigan, 1991, to appear in SIAM J. Control and Optimiz.

[15] A. Kawano, T. Hagiwara, and M. Araki, "Robust servo condition for sampled-data systems," (in Japanese), in *Proc. SICE Contr. Theory Symp.*, 1990, pp. 35–40.

[16] J. B. Lewis and Y.-S. Jan, "On linear discrete-time deadbeat control," in *Proc. 20th CDC*, 1980, pp. 494–499.

[17] J. R. Ragazzini and G. F. Franklin, *Sampled-Data Control Systems.* New York: McGraw-Hill, 1958.

[18] W. Rudin, *Functional Analysis.* New York: McGraw-Hill, 1973.

[19] S. Hara and H.-K. Sung, "Ripple-free conditions in sampled-data control systems," to be presented in *30th CDC Brighton*, 1991.

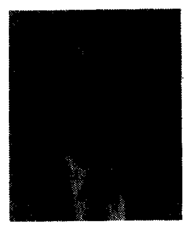
[20] H. T. Toivonen, "Sampled-data control of continuous-time systems with an $\mathcal{H}_\infty$ optimality criterion," *Automatica*, vol. 28, pp. 45–54, 1992.

[21] F. Treves, *Topological Vector Spaces, Distributions and Kernels.* New York: Academic, 1967.

[22] S. Urikura and A. Nagata, "Ripple-free deadbeat control for sampled-data systems," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 474–482, 1987.

[23] Y. Yamamoto, "Pseudo-rational input/output maps and their realizations: A fractional representation approach to infinite-dimensional systems," *SIAM J. Control Optim.*, vol. 26, pp. 1415–1430, 1988.

[24] ——, "New approach to sampled-data systems: A function space method," in *Proc. 29th CDC*, 1990, pp. 1882–1887.



Yutaka Yamamoto received his Ph.D. degree in mathematics from the University of Florida in 1978 under the guidance of R. E. Kalman.

From 1978 to 1987 he was with Department of Applied Mathematics and Physics, Kyoto University. In 1987 he joined the Department of Applied Systems Science as an Associate Professor.

Dr. Yamamoto is currently an Associate Editor of *Automatica* and *Mathematics of Control, Signals and Systems*. He is a member of the Society of Instrument and Control Engineers (SICE) and the Institute of Systems, Control and Information Engineers. He received Sawaragi memorial paper award in 1985, outstanding paper award of SICE in 1987 (joint with S. Hara), and best author award of SICE in 1990. His current research interests are in realization and robust control of distributed parameter systems, learning control, and sampled-data systems.