Studies on Iterative Learning Control for Linear Systems

Takuya Sogo

February 1999

Acknowledgment

It is a great pleasure to express the author's gratitude to many people for helping him to produce this thesis. First and foremost, the author wishes to express sincere thanks to his supervisor Professor Norihiko Adachi for guiding constantly during the course of this research and for giving his time generously to discuss the ideas in this thesis. Appreciation is extended to Professor Yutaka Yamamoto at Kyoto University, Professor Toshiharu Sugie at Kyoto University, Associate Professor Yoji Iiguni at Osaka University and Dr. Yoichi Hirashima at Okayama University for their valuable comments and advice. The author also thank Dr. Katsuya Ogino, Mr. Takanori Fukao and students of Professor Adachi's laboratory for discussions.

Takuya Sogo
Kyoto University
February 1999

Abstract

This dissertation is concerned with iterative learning control, which was originally developed as a trial-based control method for robot manipulators to track a given trajectory defined on a short time interval precisely and repeatedly. One of the most effective method of iterative learning control is a method utilizing adjoint systems, which is based on the gradient method. Some experimental results has shown its advantage. In this dissertation, convergence of this type of iterative learning control is theoretically discussed for linear systems.

First, iterative learning control using adjoint system for linear continuous-time systems is discussed. We demonstrate its convergence and express the convergence conditions as conditions of system matrices or transfer matrices for applications. Moreover, convergence rate of the iterative learning control is presented. We also discuss relationship between convergence rates and robustness of the iterative learning control against measurement noise, perturbation caused by initialization errors etc. A modification of the iterative learning control is proposed for the robustness.

Second, iterative learning control for sampled-data system is discussed. Since iterative learning control is a method to obtain the desired input from the output, its performance is determined by the inverse system. On the other hand, it is not trivial to discuss stability of the inverse system or zeros of sampled-data system on the short-fixed continuous-time interval because both the zeros and the number of sample points are determined by the sampling period. We examine limiting properties of the inverse sampled-data system on the finite continuous-time interval to determine whether it is possible or not to formulate iterative learning control as a minimization problem of the output error at sample points. We present a affirmative result which is independent of stability of zeros.

Finally, iterative learning control for linear discrete-time systems is discussed based on the results of the preceding chapters. It is formulated as a minimization problem of the Euclidean norm of the output errors. As is done for the continuous-time systems, the gradient method is applied to iterative learning control and a method using adjoint systems is developed. From an application viewpoint, convergence conditions are discussed for discrete-time systems with structured uncertainty or with parameters given as intervals.

Contents

List of Figures

Chapter 1

Introduction

1.1 Iterative Learning Control and Its Brief Review

Industrial robot manipulators are often required to track a given trajectory defined on a short time interval precisely and repeatedly. However, it is difficult for conventional servo systems to satisfy this requirement because the effect of transient response cannot be ignored over the short time interval if feedback controllers are employed. On the other hand, if feedforward controllers are introduced, effects of uncertainty of the system model emerge. In order to overcome these difficulties, a control method called iterative learning control has been developed by robotics researchers[1]. The basic idea of the method is described as follows. Consider a robot manipulator, which can be set to a fixed initial condition and is required to yield a fixed trajectory y_d defined on a finite time interval $[0, t_f]$, repeatedly. Let Σ be its input-output mapping with the same initial condition. Then the output trajectory $y \in \mathcal{Y}$ for an input function $u \in \mathcal{U}$ can be denoted by $y = \Sigma u$ where $\mathcal{Y}$ and $\mathcal{U}$ are classes of functions defined on $[0, t_f]$. Iterative learning control uses recursive iteration, including trial operation on the real system, so that the output function y converges to the desired trajectory. (see Fig.1.1)

Step 1 Give an input function $u \in \mathcal{U}$ to Σ.

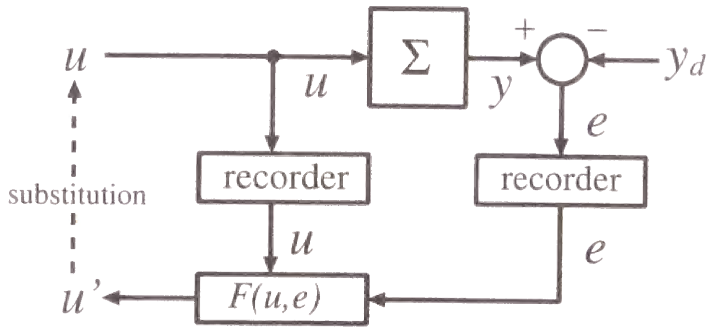


Figure 1.1: Iterative learning control algorithm

Step 2 Measure the output trajectory $y = \Sigma u$.

Step 3 Substitute u by $F(u, e)$, where $e = y - y_d$ and $F{:}\mathcal{U} \times \mathcal{Y} \to \mathcal{U}$.

Step 4 Reset the initial condition. Go to **Step 1**.

Assume that $F(u, e)$ is chosen so that

$$\|e_k(t)\| = \|y_k(t) - y_d(t)\| \to 0 \quad \text{or the minimum value} \tag{1.1}$$

as $k \to \infty$ where $\|\cdot\|$ indicates a functional norm on $[0, t_f]$ and k denotes an index of the iteration. Then, naturally, such iteration is equivalent to numerical algorithms for solving the functional equation $y_d = \Sigma u$ or to methods for minimizing the norm of $y - y_d$. However, it is possible to apply those algorithms or methods only if Σ is completely known. Therefore, the problem of this control method is to determine the revision mapping $F(u, e)$ at **Step 3** when only partial information on Σ is available.

As stated above, iterative learning control has been developed as an approach to the problem of fast and precise servo systems for robotic manipulators. One of the first ideas of iterative learning control was presented by Uchiyama [2]. He studied updating algorithm

$$u_{k+1}(t) = u_k(t) + Ae_k(t) \quad t \in [0, t_f] \tag{1.2}$$

where A is a matrix and $e_k(t) = [\Sigma u_k](t) - y_d(t)$. Convergence of the algorithm was discussed in the frequency domain by using approximations. Advantage of introduction of the method was demonstrated experimentally for a mechanical arm. Subsequently, Arimoto and his research group investigated more general class of iterative learning control algorithms both theoretically and experimentally. They examined so-called PID-type algorithms

$$u_{k+1}(t) = u_k(t) + (A_1 + A_2 \int_0^t d\tau + A_3 \frac{d}{dt}) e_k(t) \quad t \in [0, t_f] \tag{1.3}$$

applied to linear systems or nonlinear systems where A_i $(i = 1, 2, 3)$ is a matrix [3, 4, 5]. Bondi et al.[6, 7] also investigated similar algorithms for robotic manipulators. Convergence of those algorithms were proved theoretically for a class of nonlinear systems including robotic manipulators [8, 9, 10, 11]. However it was pointed out that Γ in the algorithm (1.3) must include a differential operator in order to apply the algorithm to wider class of plants [12, 13]. Complete differential operation cannot be implemented on physical systems; it can be realized only approximately. Approaches to improvement of applicability to a larger class are examined also from a different viewpoint. Moore et al. [14] or Mita et al. [15, 16] discussed relaxation of the requirement (1.1) to the algorithm. They showed that the application class of the algorithm can be improved by replacing the limit 0 or the minimum value in (1.1) with a neighborhood of them. It was also demonstrated by Heinzinger et al. [17, 9] and Arimoto [8] that such relaxation improves robustness of the algorithm against measurement noises, initialization errors and fluctuations of system dynamics.

On the other hand, Inoue et al.[18] proposed repetitive control as a precise tracking control method for a proton synchrotron magnet power supply, whose desired trajectories are periodic. Since repetitive control records the output errors and delay them to utilize for adjusting the input, it seems to be based on the same idea as iterative learning control. However, repetitive control is essentially different from iterative learning control since the former include no reset operation of the initial condition at each t_f period. Stability and tracking performance of repetitive control were discussed on the infinite time horizon or the frequency domain independently of iterative learning control [19, 20, 21, 22]; its application was also developed [23].

One of the most distinctive feature of iterative learning control is reset operation of the initial condition at each t_f period, which is not included by repetitive control. This makes it possible to choose a non-causal operator as the Γ in the algorithm (1.3) since the input function can be updated off-line; some researchers noticed the fact. Togai et al.[24] and Saab [25] studied algorithms for linear discrete-time systems, which utilize the output error at the one-step forward time. Kurek et al.[26] discussed similar algorithms in the context of 2-D system theory. Togai et al.[24] applied the steepest descent method, Newton-Raphson method and Gauss-Newton method to determination of the optimum Γ in the algorithm. Furuta et al.[27, 28] examined one of the most effective non-causal methods based on the gradient method in functional spaces, which utilizes adjoint systems. Since the algorithm is a generalized version of numerical optimization methods, it can be applied to any system provided that the system model is completely known. It has been shown experimentally that even if only partial information on the system model is available, the algorithm can be applied to a more general class of systems than the PID-type, and moreover that the scheme is easily realized with digital controllers[28]. However, convergence of the algorithm was discussed only in the frequency domain, which is useless when one deals with non-causal systems defined on the finite time domain.

1.2 An Overview of the Dissertation

In this dissertation, we discuss iterative learning control using adjoint systems for linear systems, convergence of which has not shown theoretically. We present its convergence conditions in order to give design methodologies of

the algorithm for linear systems with uncertainty. Moreover, we discuss iterative learning control for sampled-data systems, especially non-minimum phase systems. Design methods of iterative learning control is also presented for linear discrete-time systems.

Chapter 2: Iterative Learning Control for Linear Continuous-Time Systems

In this chapter, we discuss iterative learning control for linear continuous-time systems. As stated in the preceding section, iterative learning control is regarded as iterative methods to solve functional equations or minimize functional norms. However, conventional numerical methods cannot be applied to iterative learning control because those methods require the precise system models. The main topic of iterative learning control is how to design the method when there is uncertainty of the systems model. In this chapter, based on the gradient method in the functional space, we propose an algorithm of iterative learning control for linear continuous-time system with uncertainty. The algorithm utilizes the adjoint system of the nominal system for updating the input function. We present a convergence condition of the algorithm, which is expressed as strictly coerciveness of the operator that represents uncertainty of the system. Moreover, for convenience of the application, the condition is transformed into the strictly positive real (SPR) condition of the transfer function; a design methodology of iterative learning control for linear time-invariant systems is given based on the SPR condition. We also discuss the convergence condition for linear time-invariant systems with structured uncertainty, namely the system parameters are given as intervals where the parameter exist. Next, from an application viewpoint, we discuss relaxation of the convergence conditions, especially the condition imposed on the desired trajectory. We demonstrate that the relaxation preserves uniform convergence of the output to the desired trajectory. Finally, some numerical examples are presented to illustrate the iterative learning control and the design methodology proposed in this chapter.

Chapter 3: Convergence Rate and Robustness of Iterative Learning Control

In this chapter, we discuss convergence rate of iterative learning control and its relation with robustness against measurement noises, initialization errors, fluctuations of system dynamics etc. First, we demonstrate that convergence speed of iterative learning control should be of exponential functions from robustness point of view. It is shown, however, that it is impossible to give such algorithms for linear continuous-time systems provided that we consider the minimization problem formulated in the previous chapter as the problem of iterative learning control. In order to overcome this difficulty, we modify the functional to be minimized by introducing a so-called regularization term. We present a design method of iterative learning control with the regularization term. From an application viewpoint of the iterative learning control, we also discuss properties of the minimizer that the input function converges to.

Chapter 4: Iterative Learning Control for Sampled-data Systems and the Inverse Systems

The continuous-time system discussed in the preceding chapters was defined on the finite time interval. Since the time interval is short for most applications of iterative learning control, stability of poles and zeros of the system transfer functions is not considered. On the other hand, for implementation of iterative learning control, it is necessary to record input functions or the measured output functions and process those functions repetitively. Therefore, it is convenient to implement the iterative learning control with a sampler, a hold and digital computers. In this case, however, we have to to consider stability of poles and zeros of the discrete-time system because

the number of the sample points or the discrete-time interval increases on the fixed continuous-time interval as the sampling period approaches 0. In this chapter, we consider iterative learning control for sampled-data systems with a 0-order hold and a sampler which have the same sampling period. It is known that zeros of the transfer function of sampled-data systems have no simple relationship with zeros of the original continuous-time systems; unstable zeros of the sampled-data systems are common even if there is no unstable zero of the continuous-time system. We discuss effect of those zeros to inverse systems and iterative learning control for the sampled-data system. It is demonstrated that, contrary to intuitive expectation, unstable zeros do not cause divergence of the inverse system when the sampling period goes to 0 as far as relative degree of the transfer function of the continuous-time system is 0, 1 or 2. It is shown that this property implies one can define iterative learning control for sampled-data systems as a minimization problem only on the sample points; the ripple on the inter-sample points can be reduced by shrinking the sampling period. Those results are illustrated by numerical examples.

Chapter 5: Iterative Learning Control for Linear Discrete-time Systems

In this chapter, in order to develop iterative learning control for linear sampled-data system, we discuss iterative learning control for linear discrete-time systems as minimization problems in finite dimension vector spaces. This is supported by the result given in the preceding chapter. First, we present iterative learning control using adjoint systems, which is based on the gradient method in the vector space. As is given for the case of continuous-time systems in Chapter 2, convergence conditions of the method are presented as strictly positive real condition of a system which represent uncertainty of the system. Moreover, we also discuss the convergence condition for linear time-invariant systems with uncertainty when the system parameters are given as intervals where the parameter exist. Finally, the iterative learning control and its design are illustrated by examples.

Chapter 2

Iterative Learning Control for Linear Continuous-Time Systems

2.1 Introduction

In this chapter, we discuss iterative learning control using adjoint systems for linear continuous-time systems theoretically and demonstrate its convergence. As stated in chapter 1, iterative learning control can be regarded as iterative methods to solve functional equations or minimize functional norms. However, conventional numerical methods cannot be applied to iterative learning control because those methods require the precise system models.

In the following sections, first, we formalize iterative learning control using adjoint systems and then present convergence conditions for linear systems with uncertainty. Second, those conditions are transferred to conditions on system matrices for convenience of application. We also give convergence conditions for single-input single-output linear time-invariant systems with structured uncertainty. Moreover, relaxation of the convergence conditions is discussed, again from an applications viewpoint. Finally, some numerical examples are presented.

2.2 Convergence of Iterative Learning Control Using Adjoint Systems

Consider a linear time-invariant system

$$\begin{aligned} \frac{d}{dt}x(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned} \tag{2.1}$$

defined on the finite time interval $[0, t_f]$ where $u \in R^m$, $x \in R^n$ and $y \in R^p$. $A(t)$, $B(t)$, $C(t)$ and $D(t)$ are compatible matrices consisting of continuous functions of $t \in [0, t_f]$. If an initial condition $x(0)$ and a continuous function $u(t)$ are given, then $y(t)$ is uniquely determined as

$$y(t) = h(t) + [Su](t)$$

where

$$\begin{aligned} h(t) &= C(t)R(t,0)x(0), \\ [Su](t) &= \int_0^t C(t)R(t,\sigma)B(\sigma)u(\sigma)d\sigma + D(t)u(t) \end{aligned} \tag{2.2}$$

and $R(t, \sigma)$ is a $n \times n$ matrix of continuous functions of t and σ. Consider L_2 space of R^m-valued functions,

$$L_2^m[0, t_f] = \{u(t) | \int_0^{t_f} u^T(t)u(t)dt < +\infty\}$$

where T denotes the transpose. Then S defines a linear operator mapping $L_2^m[0, t_f]$ into $L_2^p[0, t_f]$. The inner product and the norm in $L_2^m[0, t_f]$ are defined as

$$\langle u_1, u_2 \rangle = \int_0^{t_f} u_1^T(t)u_2(t)dt$$

and

$$\|u\|_2 = \sqrt{\langle u, u \rangle},$$

respectively. When there is no possibility of confusion about the dimension, the superscript of $L_2^m[0, t_f]$ is omitted. The induced norm, the range and

the null space of an operator K: $L_2[0, t_f] \to L_2[0, t_f]$ are expressed as $\|K\|$, $\mathcal{R}(K)$ and $\mathcal{N}(K)$, respectively; the adjoint operator is denoted by K^*.

Let us consider the problem of making the output of the system (2.1) track a desired output trajectory $y_d \in L_2[0, t_f]$ precisely. Then an ideal input function is $u^* \in L_2[0, t_f]$ satisfying $y_d = h + Su^*$. However, u^* can be known only if the system model (2.1) is completely known. In this paper, we will develop an iterative method to obtain u^* with the following assumptions:

1. The initial value is always the same $x(0)$.

2. One can measure the response $y = h + Su$ for any input function u.

Since these assumptions imply that $y_d - h$ can be substituted for y_d, without loss of generality, it is assumed that $x(0) = 0$ or $h \equiv 0$ in the following discussion. Therefore, we will discuss the functional equation

$$y_d = Su \quad u \in L_2^m[0, t_f]$$

Suppose that (2.1) is a partly unknown system and can be decomposed into an unknown system and a known system as follows:

$$\begin{aligned} \frac{d}{dt}\xi(t) &= E(t)\xi(t) + F(t)u(t) \\ \eta(t) &= G(t)\xi(t) + H(t)u(t) \end{aligned} \tag{2.3}$$

and

$$\begin{aligned} \frac{d}{dt}x(t) &= \hat{A}(t)x(t) + \hat{B}(t)\eta(t) \\ y(t) &= \hat{C}(t)x(t) + \hat{D}(t)\eta(t) \end{aligned} \tag{2.4}$$

with $\xi(0) = 0$ and $x(0) = 0$ where $\eta \in R^m$ and elements of matrices are continuous functions of $t \in [0, t_f]$. Then

$$S = \hat{S}U \tag{2.5}$$

where $\hat{S}$ and U are defined in the same way as S in order to represent input-output mappings of (2.3) and (2.4), namely

$$\begin{aligned} \eta &= Uu \quad \eta, u \in L_2^m[0, t_f] \\ y &= \hat{S}\eta \quad y \in L_2^p[0, t_f], \eta \in L_2^m[0, t_f], \end{aligned}$$

respectively. (see Figure 2.1) The next theorem presents an iterative algorithm using $\hat{S}$ instead of the unavailable S.

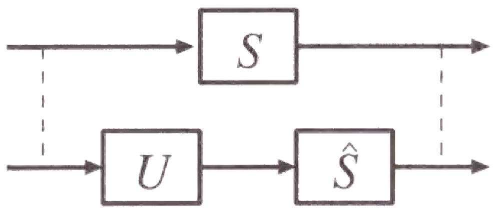


Figure 2.1: Decomposition of S

Theorem 1 *Suppose that there exists $v_d \in L_2[0, t_f]$ such that*

$$y_d = S\hat{S}^* v_d \tag{2.6}$$

and U is strictly coercive, namely

$$\langle U\eta, \eta \rangle \geq \beta \|\eta\|_2^2 \tag{2.7}$$

for any $\eta \in L_2^p[0, t_f]$ where β is a positive constant. Then the sequence $\{u_k; k = 0, 1, \cdots\}$ generated by

$$u_{k+1} = u_k - \alpha \hat{S}^*(Su_k - y_d) \tag{2.8}$$

with

$$u_0 \in \mathcal{R}(\hat{S}^*) \tag{2.9}$$

satisfies

$$\|u_k - \hat{S}^* v_d\|_2 \to 0$$

as $k \to \infty$ where the constant α is chosen as $0 < \alpha < 2\beta / \|\hat{S}U\|^2$.

Proof: Since (2.8) with (2.9) implies $u_k \in \mathcal{R}(\hat{S}^*)$ for $k = 0, 1, \cdots$, there exists a unique $v_k \in \mathcal{N}(\hat{S}^*)^\perp$ such that $u_k = \hat{S}^* v_k$ where $\mathcal{N}(\hat{S}^*)^\perp$ denotes the orthogonal complement of $\mathcal{N}(\hat{S}^*)$. Furthermore, (2.6) implies that there exists a unique $\bar{v}_d \in \mathcal{N}(\hat{S}^*)^\perp$ such that $\hat{S}^* \bar{v}_d = \hat{S}^* v_d$. Therefore, (2.8) with (2.5) implies

$$\hat{S}^*(v_{k+1} - \bar{v}_d) = \hat{S}^*\{v_k - \bar{v}_d - \alpha \hat{S}U\hat{S}^*(v_k - \bar{v}_d)\} \tag{2.10}$$

Since $R(\hat{S}) \subseteq \overline{R(\hat{S})} = N(\hat{S}^*)^\perp$ where $\overline{R(\hat{S})}$ indicates the closure of $R(\hat{S})$, we have

$$v_{k+1} - \bar{v}_d, v_k - \bar{v}_d \text{ and } \hat{S}U\hat{S}^*(v_k - \bar{v}_d) \in N(\hat{S}^*)^\perp$$

Hence, (2.10) implies

$$v_{k+1} - \bar{v}_d = v_k - \bar{v}_d - \alpha \hat{S}U\hat{S}^*(v_k - \bar{v}_d) \tag{2.11}$$

From this equality and (2.7), we have

$$\begin{aligned} \|v_{k+1} - \bar{v}_d\|_2^2 &\leq \|v_k - \bar{v}_d\|_2^2 - 2\alpha\beta\|\hat{S}^*(v_k - \bar{v}_d)\|_2^2 + \alpha^2 \|\hat{S}U\hat{S}^*(v_k - \bar{v}_d)\|_2^2 \\ &\leq \|v_k - \bar{v}_d\|_2^2 - \alpha(2\beta - \alpha\|\hat{S}U\|^2)\|\hat{S}^*(v_k - \bar{v}_d)\|_2^2 \end{aligned}$$

Since $\alpha(2\beta - \alpha\|\hat{S}U\|^2) > 0$ and $\{\|v_k - \bar{v}_d\|_2^2; k = 0, 1, \cdots\}$ are bounded from below, we establish

$$\|\hat{S}^*(v_k - \bar{v}_d)\|_2 = \|u_k - \hat{S}^* v_d\|_2 \to 0 \tag{2.12}$$

as $k \to \infty$. ■

Remark 1 *If $S = \hat{S}$ or $U = I$ (the identity operator) then the algorithm (2.8) is equivalent to the Landweber-Fridman method[29] or one of the gradient methods such as the steepest descent method[30]. The condition (2.7) gives a margin of the convergence of the iteration for the system (2.1) with uncertainty.*

The operator $\hat{S}^*$ corresponds to an input-output mapping of the system

$$\begin{aligned} \frac{d}{dt}p(t) &= -\hat{A}^T(t)p(t) - \hat{C}^T(t)y(t) \\ \eta(t) &= \hat{B}^T(t)p(t) + \hat{D}^T(t)y(t) \end{aligned} \quad (2.13)$$

with the initial condition $p(t_f) = 0$, as is shown from an identical equation

$$x^T(t_f)p(t_f) - x^T(0)p(0) = \left\langle \frac{d}{dt}x(t), p(t) \right\rangle + \left\langle x(t), \frac{d}{dt}p(t) \right\rangle$$

Therefore, the function $\hat{S}^*(Su_k - y_d)$ in the algorithm can be obtained by numerical calculation of the response for each error function $Su_k - y_d$. Note that (2.9) can be satisfied by letting $u_0 \equiv 0$. Then the algorithm (2.8) can easily be realized for a recursive process of iterative learning control as shown in Figure 2.2.

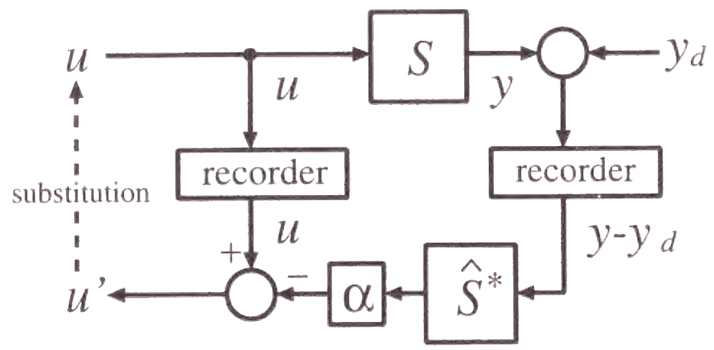


Figure 2.2: Iterative learning control using the adjoint system

2.3 Convergence Conditions for Linear Time-Variant Systems and Linear Time-Invariant Systems

Convergence of the algorithm is guaranteed for a partly unknown system S, if S can be decomposed into a known $\hat{S}$ and a strictly coercive U, which is unknown. The next theorem gives a method to determine whether U

is strictly coercive based on the system matrices of (2.3). In the following discussions, $H(t) > 0$ means $u^T H(t) u > 0$ for all $u \in R^m$ and the other inequalities of a matrix are defined likewise.

Theorem 2 *If*

$$H(t) > 0$$

for all $t \in [0, t_f]$ and there exist matrices $K(t)$, $L(t)$, $P(t)$ and $Q(t)$ such that

$$\begin{aligned} P(t)E(t) + E^T(t)P(t) &= -L^T(t)L(t) - Q(t) \\ F^T(t)P(t) + K^T(t)L(t) &= G(t) \\ K^T(t)K(t) &= H(t) + H^T(t) \end{aligned}$$

where $P(t)$ and $Q(t)$ are symmetric and satisfy

$$P(t_f) \geq 0 \quad (2.14)$$

$$\frac{d}{dt}P(t) \leq 0 \quad (2.15)$$

$$Q(t) > 0 \quad (2.16)$$

for all $t \in [0, t_f]$. Then U is strictly coercive.

Proof: Since there exists a constant $c \in (0, 1)$ such that

$$Q(t) - \left(\frac{1}{c} - 1\right) L^T(t)L(t) > 0 \quad (2.17)$$

for all $t \in [0, t_f]$, we have

$$P(t)E(t) + E^T(t)P(t) = -N^T(t)N(t) - R(t) \quad (2.18)$$

$$F^T(t)P(t) + M^T(t)N(t) = G(t) \quad (2.19)$$

$$M^T(t)M(t) = c(H(t) + H^T(t)) \quad (2.20)$$

where $M(t) = \sqrt{c}K(t)$, $N(t) = L(t)/\sqrt{c}$ and $R(t) = Q(t) - (1/c - 1)L^T(t)L(t)$. Moreover, from (2.3), (2.15), (2.18), (2.19) and (2.20), we have

$$\frac{d}{dt}(\xi^T P \xi) = (E\xi + Fu)^T P\xi + \xi^T P(E\xi + Fu) + \xi^T \left(\frac{d}{dt}P\right)\xi$$

$$
\begin{aligned}
&= \xi^T(-N^TN - R)\xi + 2u^T(G - M^TN)\xi + \xi^T\left(\frac{d}{dt}P\right)\xi \\
&= \xi^T(-N^TN - R)\xi + 2u^T\eta - u^T(H + H^T)u \\
&\quad -2u^TM^TN\xi + \xi^T\left(\frac{d}{dt}P\right)\xi \\
&= -(N\xi + Mu)^T(N\xi + Mu) - \xi^TR\xi + 2u^T\eta \\
&\quad -(1-c)u^T(H + H^T)u + \xi^T\left(\frac{d}{dt}P\right)\xi
\end{aligned}
$$

Integration on $[0, t_f]$ of both sides of the equalities yields

$$
\begin{aligned}
\xi(t_f)^TP(t_f)\xi(t_f) &= -\|N\xi + Mu\|_2^2 - \langle \xi, R\xi\rangle + 2\langle u, \eta\rangle - (1-c)\langle u, (H + H^T)u\rangle \\
&\quad + \left\langle \xi, \left(\frac{d}{dt}P\right)\xi \right\rangle
\end{aligned}
$$

From (2.14), (2.15), (2.17) and this equality, we establish

$$\langle u, \eta\rangle \geq \frac{1-c}{2}\lambda_{[0,t_f]}\|u\|_2^2$$

where $\lambda_{[0,t_f]}(t)$ indicates the minimum value of the smallest eigenvalue of $H(t) + H^T(t)$ on $[0, t_f]$. This completes the proof. ■

The following example illustrates the design of the algorithm presented in Theorem 1 based on Theorem 2.

Example 1 *Consider a linear time-variant system*

$$\frac{d}{dt}x(t) = (d - t)x(t) + u(t) \tag{2.21}$$

$$y(t) = x(t) \tag{2.22}$$

as S where $t \in [0, 1]$ and the parameter d is unknown but the range is known; $d \in [2, 3]$. Let $\hat{S}$ for the algorithm of Theorem 1 be

$$
\begin{aligned}
\frac{d}{dt}x(t) &= (1 - t)x(t) + \eta(t) \\
y(t) &= x(t)
\end{aligned} \tag{2.23}
$$

then the system U is

$$
\begin{aligned}
\frac{d}{dt}\xi(t) &= (d - t)\xi(t) + (d - 1)u(t) \\
\eta(t) &= \xi(t) + u(t)
\end{aligned}
$$

Since there exist K, L, P and Q satisfying the conditions of Theorem 2, namely

$$
\begin{aligned}
K &= \sqrt{2} \\
L(t) &= \sqrt{2}\{1 + (1 - d)P(t)\}/2 \\
P(t) &= \{-2t - 1 + 3d + \sqrt{2(2t^2 + 2t - 6dt - 1 + 3d^2)}\}/(d - 1)^2 \\
Q &= 1
\end{aligned}
$$

U is strictly coercive. Note that the system $\hat{S}^*$ is

$$
\begin{aligned}
\frac{d}{dt}p(t) &= -(1 - t)p(t) - v(t) \\
u(t) &= p(t)
\end{aligned} \tag{2.24}
$$

with $p(t_f) = 0$. Then a sufficient condition for (2.6) is

$$y_d \in C^2[0, t_f]$$

and

$$y_d(0) = y_d(t_f) = \frac{d}{dt}(t_f) = 0$$

where $C^n[0, t_f]$ denotes the class of n-times continuously differentiable functions on $[0, t_f]$. Hence, we can guarantee convergence of the algorithm with a sufficiently small α based on (2.24) for (2.22).

Suppose that the system (2.3) is time-invariant. Then by the Popov-Kalman-Yakubovic Lemma[31] we can rewrite the presuppositions of Theorem 2 as follows:

1. $H > 0$
2. The transfer matrix $U(s) = G(sI - E)^{-1}F + H$ is strictly positive real (SPR) i.e., there exists a positive λ such that $U(s - \lambda) + U(\bar{s} - \lambda)^T \geq 0$ for all s on the right-half plane and the imaginary axis where $\bar{s}$ denotes the conjugate complex number of s.

If the linear systems (2.1), (2.4) and (2.3) are time-invariant, to choose $\hat{S}$ so that U is strictly coercive for the algorithm of Theorem 1 is to extract the transfer function

$$\hat{S}(s) = \hat{C}(sI - \hat{A})\hat{B} + \hat{C}$$

from

$$S(s) = C(sI - A)B + C$$

so that $U(s)$ satisfying $S(s) = \hat{S}(s)U(s)$ has the above two properties.

The process to design the algorithm of Theorem 1 for linear time-invariant systems is summarized as follows[32, 33, 34].

step 1 Choose a transfer function matrix $\hat{S}(s)$ such that $S(s) = C(sI - A)^{-1}B + D$ is expressed as

$$S(s) = \hat{S}(s)U(s)$$

where $U(s) = G(sI - E)^{-1}F + H$ is SPR and

$$H > 0. \tag{2.25}$$

step 2 Let a realization of the transfer function matrix $\hat{S}^T(s)$ be $(\hat{A}^T, \hat{C}^T, \hat{B}^T, \hat{D}^T)$

step 3 Let the mapping $\eta = \hat{S}^*(y - y_d)$ of the algorithm (2.8) be the input-output mapping of the linear system

$$\begin{aligned} \frac{d}{dt}p(\tau) &= \hat{A}^T p(\tau) + \hat{C}^T(y(t) - y_d(t)) \\ \eta(t) &= \hat{B}^T p(\tau) + \hat{D}^T(y(t) - y_d(t)) \end{aligned} \tag{2.26}$$

with the initial condition $p(0) = 0$ where $\tau = t_f - t$.

If the matrix $S(s)$ is invertible, step 1 can be replaced with:

step 1 Choose $\hat{S}^{-1}(s)$ as an approximate inverse of $S(s)$ so that $U(s) = S(s)\hat{S}^{-1}(s)$ satisfies strictly-positive-real conditions and (2.25).

2.4 Convergence Conditions for Linear Time-Invariant Systems with Structured Uncertainty

In this section, we will develop a method to check whether $\hat{S}(s)$ is chosen so that $U(s)$ is SPR, provided $S(s)$ is single-input single-output and has structured uncertainty, namely

$$S(s) = \frac{N(s)}{D(s)}$$

where

$$\begin{aligned} N(s) &= \sum_{k=0}^{m} q_{m-k}s^k \\ D(s) &= \sum_{k=0}^{n} p_{n-k}s^k \\ p_0 &= q_0 = 1 \\ p_i &\in [\underline{p_i}, \overline{p_i}] \quad (i = 1, \cdots, n) \qquad (2.27) \\ q_i &\in [\underline{q_i}, \overline{q_i}] \quad (i = 1, \cdots, m) \qquad (2.28) \end{aligned}$$

If we choose $\hat{S}(s)$ as

$$\hat{S}(s) = \frac{\hat{N}(s)}{\hat{D}(s)}$$

where $\hat{N}(s) = \sum_{k=0}^{\hat{m}} \hat{q}_{\hat{m}-k}s^k$ and $\hat{D}(s) = \sum_{k=0}^{\hat{n}} \hat{p}_{\hat{n}-k}s^k$ then the first condition $H > 0$ imposes $\hat{n} - \hat{m} = n - m$ and $\hat{p}_0\hat{q}_0 > 0$. Without loss of generality, it is assumed that $\hat{p}_0 = \hat{q}_0 = 1$ in the following discussion. The second condition requires that for all parameters satisfying (2.27) and (2.28)

$$U(s) = \frac{\hat{D}(s)N(s)}{\hat{N}(s)D(s)} \tag{2.29}$$

is SPR, equivalently in this case[35]

1. $U(s)$ has no pole on the right-half plane or the imaginary axis, i.e., $\hat{N}(s)D(s)$ is Hurwitz.

2. $\mathrm{Re}U(j\omega) > 0$ for all $\omega \in (-\infty, \infty)$.

It is already known that a transfer function is SPR for all parameters from the convex set as long as the transfer function is SPR for only a finite number of parameters. These results were developed to make the verification of the SPR condition computationally feasible in terms of adaptive control applications. Some results have been published for transfer functions with such forms as $N(s)/\hat{N}(s)$[36], $N(s)(1+s)^{\hat{n}-m}/\hat{D}(s)$[37] and $N(s)/D(s)$[38]. However, these authors have not considered the function (2.29) with which we are concerned. A similar result for (2.29) is presented as follows.

Theorem 3 *The transfer function $\hat{D}(s)N(s)/(\hat{N}(s)D(s))$ is SPR for all parameters satisfying (2.27) and (2.28) if the transfer function*

$$U_{uv}(s) = \frac{\hat{D}(s)N_u(s)}{\hat{N}(s)D_v(s)} \tag{2.30}$$

is SPR for $u = 1,2,3,4$ and $v = 1,2,3,4$ where

$$\begin{aligned} D_1(s) &= \sum_{k=0}^{n/4]} \overline{p_{n-4k}} s^{4k} + \sum_{k=0}^{n/4]} \overline{p_{n-4k-1}} s^{4k+1} \\ &\quad + \sum_{k=0}^{n/4]} \underline{p_{n-4k-2}} s^{4k+2} + \sum_{k=0}^{n/4]} \underline{p_{n-4k-3}} s^{4k+3} \end{aligned} \tag{2.31}$$

$$\begin{aligned} D_2(s) &= \sum_{k=0}^{n/4]} \underline{p_{n-4k}} s^{4k} + \sum_{k=0}^{n/4]} \underline{p_{n-4k-1}} s^{4k+1} \\ &\quad + \sum_{k=0}^{n/4]} \overline{p_{n-4k-2}} s^{4k+2} + \sum_{k=0}^{n/4]} \overline{p_{n-4k-3}} s^{4k+3} \end{aligned} \tag{2.32}$$

$$\begin{aligned} D_3(s) &= \sum_{k=0}^{n/4]} \underline{p_{n-4k}} s^{4k} + \sum_{k=0}^{n/4]} \overline{p_{n-4k-1}} s^{4k+1} \\ &\quad + \sum_{k=0}^{n/4]} \overline{p_{n-4k-2}} s^{4k+2} + \sum_{k=0}^{n/4]} \underline{p_{n-4k-3}} s^{4k+3} \end{aligned} \tag{2.33}$$

$$\begin{aligned} D_4(s) &= \sum_{k=0}^{n/4]} \overline{p_{n-4k}} s^{4k} + \sum_{k=0}^{n/4]} \underline{p_{n-4k-1}} s^{4k+1} \\ &\quad + \sum_{k=0}^{n/4]} \underline{p_{n-4k-2}} s^{4k+2} + \sum_{k=0}^{n/4]} \overline{p_{n-4k-3}} s^{4k+3} \end{aligned} \tag{2.34}$$

where $\underline{p_0} = \overline{p_0} = 1$ and $\underline{p_i} = \overline{p_i} = 0$ for $i < 0$; $n/4]$ denotes the largest integer that is less than or equal to $n/4$. Polynomials $N_1(s)$, $N_2(s)$, $N_3(s)$ and $N_4(s)$ are defined likewise.

To prove this theorem , we utilize the next theorem, given by Kharitonov[39].

Theorem 4 *A polynomial*

$$D(s) = \sum_{k=0}^{n} p_{n-k} s^k$$

is Hurwitz for all parameters satisfying (2.27) where $p_0 = 1$ if the four polynomials (2.31), (2.32), (2.33) and (2.34) are Hurwitz.

Proof of Theorem 3. First, since (2.30) is SPR for $v = 1,2,3,4$, the denominator $\hat{N}(s)D_v(s)$ is Hurwitz for $v = 1,2,3,4$, and, hence, by Theorem 4 $\hat{N}(s)D(s)$ is Hurwitz for all parameters satisfying (2.27).

Second, we show $\mathrm{Re}U(j\omega) > 0$ for $\omega \in (-\infty, \infty)$. Since

$$\begin{aligned} \mathrm{Re}D(j\omega) &= \sum_{k=0}^{n/4]} p_{n-4k}\omega^{4k} + \sum_{k=0}^{n/4]} (-p_{n-4k-2})\omega^{4k+2} \\ \mathrm{Im}D(j\omega) &= \sum_{k=0}^{n/4]} p_{n-4k-1}\omega^{4k+1} + \sum_{k=0}^{n/4]} (-p_{n-4k-3})\omega^{4k+3} \end{aligned}$$

etc., we have

$$\mathrm{Re}D_2(j\omega) = \mathrm{Re}D_3(j\omega) \le \mathrm{Re}D(j\omega) \le \mathrm{Re}D_1(j\omega) = \mathrm{Re}D_4(j\omega) \tag{2.35}$$

$$\mathrm{Im}D_2(j\omega) = \mathrm{Im}D_4(j\omega) \le \mathrm{Im}D(j\omega) \le \mathrm{Im}D_1(j\omega) = \mathrm{Im}D_3(j\omega) \tag{2.36}$$

for $\omega \in [0, \infty)$ and

$$\mathrm{Re}D_2(j\omega) = \mathrm{Re}D_3(j\omega) \le \mathrm{Re}D(j\omega) \le \mathrm{Re}D_1(j\omega) = \mathrm{Re}D_4(j\omega) \tag{2.37}$$

$$\mathrm{Im}D_1(j\omega) = \mathrm{Im}D_3(j\omega) \le \mathrm{Im}D(j\omega) \le \mathrm{Im}D_2(j\omega) = \mathrm{Im}D_4(j\omega) \tag{2.38}$$

for $\omega \in (-\infty, 0]$. Inequalities (2.35), (2.36), (2.37) and (2.38) where Ns are substituted for Ds also hold. Note that $\mathrm{Re}U(j\omega) > 0$ for $\omega \in (-\infty, \infty)$ is equivalent to

$$(\mathrm{Re}D(j\omega)\mathrm{Re}N(j\omega)+\mathrm{Im}D(j\omega)\mathrm{Im}N(j\omega))\mathrm{Re}\left\{\frac{\hat{D}(j\omega)}{\hat{N}(j\omega)}\right\}$$
$$+(\mathrm{Im}D(j\omega)\mathrm{Re}N(j\omega)-\mathrm{Re}D(j\omega)\mathrm{Im}N(j\omega))\mathrm{Im}\left\{\frac{\hat{D}(j\omega)}{\hat{N}(j\omega)}\right\}>0 \quad (2.39)$$

for $\omega\in(-\infty,\infty)$; the left-hand side of (2.39) is linear for $\mathrm{Re}D(j\omega)$, $\mathrm{Im}D(j\omega)$, $\mathrm{Re}N(j\omega)$ and $\mathrm{Im}N(j\omega)$. Then $\mathrm{Re}U_{uv}(j\omega)>0$ for $u=1,2,3,4$, $v=1,2,3,4$ implies that $\mathrm{Re}U(j\omega)>0$ for all parameters satisfying (2.27) and (2.28). This completes the proof. ∎

The following example illustrates the design of the algorithm of Theorem 1 based on Theorem 3.

Example 2 *Consider a linear time-invariant system*

$$S(s)=\frac{N(s)}{D(s)}=\frac{s+q_1}{s^2+p_1s+p_2} \quad (2.40)$$

with uncertain parameters

$$p_i \in [\underline{p_i},\overline{p_i}] \quad (i=1,2)$$
$$q_1 \in [\underline{q_1},\overline{q_1}]$$

Let

$$\hat{S}(s)=\frac{\hat{N}(s)}{\hat{D}(s)}=\frac{s+\hat{q}_1}{s^2+\hat{p}_1s+\hat{p}_2}$$

Then $(\hat{p}_1,\hat{p}_2,\hat{q}_1)$ *must be chosen so that*

$$\frac{N(s)\hat{D}(s)}{D(s)\hat{N}(s)}=\frac{(s+q_1)(s^2+\hat{p}_1s+\hat{p}_2)}{(s^2+p_1s+p_2)(s+\hat{q}_1)}$$

is SPR for $(p_1,p_2,q_1)\in\mathcal{P}_1\times\mathcal{P}_2\times\mathcal{Q}_1$ *where*

$$\mathcal{P}_i = \{\underline{p_i},\overline{p_i}\} \quad (i=1,2)$$
$$\mathcal{Q}_1 = \{\underline{q_1},\overline{q_1}\}$$

specifically,

$$(s^2+p_1s+p_2)(s+\hat{q}_1) : \textit{Hurwitz}$$

for $(p_1,p_2)\in\mathcal{P}_1\times\mathcal{P}_2$ *and*

$$\{(p_2-\omega^2)q_1+p_1\omega^2\}\{(\hat{p}_2-\omega^2)\hat{q}_1+\hat{p}_1\omega^2\}$$
$$+\{p_1q_1\omega-(p_2-\omega^2)\omega^2\}\{\hat{p}_1\hat{q}_1\omega-(\hat{p}_2-\omega^2)\omega^2\}>0$$

for $\omega\in(-\infty,\infty)$ *and* $(p_1,p_2,q_1)\in\mathcal{P}_1\times\mathcal{P}_2\times\mathcal{Q}_1$.

2.5 Relaxation of the Convergence Condition

The example in Section 2 shows that the existence condition (2.6) imposes restrictions upon the value of the desired trajectory $y_d(t)$ or its derivatives at $t=t_f$ as well as at $t=0$. This can be an inconvenience for some applications. On the other hand, the aim of iterative learning control is to make the output of the system track the desired trajectory rather than to estimate the ideal input function $\hat{S}^*v_d$ precisely. The next theorem shows that even if the existence condition is relaxed, we can guarantee at least convergence of the output to the desired trajectory and boundedness of the input.

Theorem 5 *Suppose that there exists* $u_d\in L_2[0,t_f]$ *such that*

$$y_d=Su_d \quad (2.41)$$

and U *which satisfies* $S=U\hat{S}$ *is strictly coercive, namely*

$$\langle U\eta,\eta\rangle\geq\beta\|\eta\|_2^2 \quad (2.42)$$

for any $\eta\in L_2^p[0,t_f]$ *where* β *is a positive constant. Then the sequence* $\{u_k;k=0,1,\cdots\}$ *generated by*

$$u_{k+1}=u_k-\alpha\hat{S}^*(Su_k-y_d) \quad (2.43)$$

with

$$u_0\in L_2[0,t_f] \quad (2.44)$$

satisfies

$$\|u_k\|_2\leq M \quad (2.45)$$

$$\|Su_k - y_d\|_2 \to 0 \tag{2.46}$$

as $k \to \infty$ where the constant α is chosen as $0 < \alpha < 2\beta/\|\hat{S}^* U\|^2$; M is a positive constant.

Proof: From (2.41), (2.42) and (2.43), we have

$$\|u_{k+1} - u_d\|_2^2 \le \|u_k - u_d\|_2^2 - \alpha(2\beta - \|\hat{S}^* U\|^2\alpha)\|S(u_k - u_d)\|_2^2 \tag{2.47}$$

for $k = 0, 1, \cdots$. Since $\alpha(2\beta - \|\hat{S}^* U\|^2\alpha)$ is positive, this inequality implies

$$\|u_{k+1} - u_d\|_2^2 \le \|u_k - u_d\|_2^2$$

and

$$\|S(u_k - u_d)\|_2^2 \to 0$$

as $k \to \infty$. This completes the proof. ■

Remark 2 *If there is no direct term from input to output in (2.1), i.e., $D(t) = 0$, and every element of $C(t)$ is in $C^1[0, t_f]$, then (2.46) with (2.45) implies uniform convergence of the output, which is desirable for the aim of iterative learning control, namely*

$$\|Su_k - y_d\|_\infty \to 0 \tag{2.48}$$

as $k \to \infty$ where $\|Su_k - y_d\|_\infty = \sup\{|[Su_k](t) - y_d(t)|; t \in [0, t_f]\}$ and $|\cdot|$ denotes Euclidean norm. This is demonstrated as follows. From (2.2), (2.41) and (2.45), we have

$$\left\| \frac{d}{dt}[S(u_k - u_d)] \right\|_2^2 \le M'$$

where M' is a positive constant. On the other hand,

$$[S(u_k - u_d)]^j(t_2) - [S(u_k - u_d)]^j(t_1) = \int_{t_1}^{t_2} \frac{d}{dt}[S(u_k - u_d)]^j(\tau)d\tau \tag{2.49}$$

for $t_1, t_2 \in [0, t_f]$, $j = 1, \cdots, p$ and $k = 0, 1, \cdots$ where $[S(u_k - u_d)]^j(t)$ denotes the j-th element of $[S(u_k - u_d)](t)$. Hence, by the Cauchy-Schwarz inequality

$$([S(u_k - u_d)]^j(t_2) - [S(u_k - u_d)]^j(t_1))^2 \le |t_2 - t_1| M'$$

which implies that $\{[S(u_k - u_d)]^j(t); j = 1, \cdots, p; k = 0, 1, \cdots\}$ is equicontinuous. By this property, (2.46) implies pointwise convergence

$$|[S(u_k - u_d)]^j(t)| \to 0 \quad (j = 1, \cdots, p; t \in [0, t_f])$$

as $k \to \infty$ whence, by the Ascoli-Arzelà theorem[40], (2.48) follows.

Example 3 *Consider the system (2.22) as S. If $y_d \in C^1[0, t_f]$ and $y_d(0) = 0$ then there exists $u_d \in C^0[0, T]$ such that $y_d = Su_d$, which is sufficient for (2.41).*

Example 4 *Consider the system (2.40) as S. If $y_d \in C^1[0, t_f]$ and $y_d(0) = 0$ then*

$$x(t) = e^{-q_1 t} \int_0^t e^{q_1 \tau} y_d(\tau)d\tau$$

satisfies $y_d(t) = \frac{d}{dt}x(t) + q_1 x(t)$ with $\frac{d}{dt}x(0) = x(0) = 0$, hence, there exists

$$u_d(t) = (q_1^2 - p_1 q_1 + p_2)e^{-q_1 t} \int_0^t e^{q_1 \tau} y_d(\tau)d\tau + (p_1 - q_1)y_d(t) + \frac{d}{dt}y_d(t)$$

such that $y_d = Su_d$, which is sufficient for (2.41).

2.6 Numerical Examples

In this section, we illustrate design of the iterative learning control presented in the preceding sections. Numerical examples are also given.

Example 5 *Let's consider a time-invariant system*

$$\frac{d^2}{dt^2}y(t) + a\frac{d}{dt}y(t) + by(t) = cu(t) \tag{2.50}$$

$$y(0) = 0$$

and

$$\frac{d}{dt}y(0) = 0$$

where a,b and c are positive constants with uncertainty. Laplace transformation yields

$$\mathcal{L}[y] = \frac{y(0)s + ay(0) + \frac{d}{dt}y(0)}{s^2+as+b} + \frac{c}{s^2+as+b}\mathcal{L}[u]$$

where $\mathcal{L}$ denotes the Laplace-transformed function. Let $y \equiv y_d$ then

$$\begin{aligned}\mathcal{L}\left[\frac{d^2}{dt^2}y_d + a\frac{d}{dt}y_d + by_d\right] &= (y(0)-y_d(0))s + (ay(0) + \frac{d}{dt}y(0)) \\ &\quad - \left(ay_d(0) + \frac{d}{dt}y_d(0)\right) + c\mathcal{L}[u]\end{aligned}$$

Sufficient conditions for (2.41) are

$$y_d \in C^2[0, t_f] \tag{2.51}$$

$$y_d(0) = y(0)$$

and

$$\frac{d}{dt}y_d(0) = \frac{d}{dt}y(0). \tag{2.52}$$

The design process of the iterative learning control is illustrated as follows.

step 1 *Let the transfer function be*

$$S(s) = \frac{c}{s^2+as+b}$$

and choose $\hat{S}(s)$ as

$$\hat{S}(s) = \frac{\hat{c}}{s^2+\hat{a}s+\hat{b}}$$

then

$$U(s) = S(s)\hat{S}^{-1}(s) = \frac{c(s^2+\hat{a}s+\hat{b})}{\hat{c}(s^2+as+b)}$$

The SPR condition of $U(s)$ is equivalent to

$$U(j\omega) + U(-j\omega) = \frac{2c\{\omega^4 + (a\hat{a} - b - \hat{b})\omega^2 + \hat{b}b\}}{\hat{c}\{(b-\omega^2)^2 + (a\omega)^2\}} > 0 \tag{2.53}$$

for all $\omega \in (-\infty, \infty)$ [41].

step 2 & 3 *The mapping $\eta = \hat{S}^*(y - y_d)$ of the algorithm (2.43) is*

$$\begin{aligned}\frac{d}{dt}p(\tau) &= \begin{pmatrix} 0 & 1 \\ -\hat{b} & -\hat{a}\end{pmatrix} p(\tau) + \begin{pmatrix} 0 \\ \hat{c}\end{pmatrix}(y(t) - y_d(t)) \\ \eta(t) &= \begin{pmatrix} 1 & 0 \end{pmatrix} p(\tau)\end{aligned}$$

with the initial condition $p(0) = (0\ 0)^T$, where $t = t_f - \tau$ and $(\hat{a}, \hat{b}, \hat{c})$ satisfy (2.53)

Figure 2.3 and 2.4 show results of numerical simulations for Example 5, where the parameters, $u_0(t)$, and the desired trajectory $y_d(t)$ on $[0, 20]$ are chosen as follows.

$$(a, b, c) = (1.0, 2.0, 3.0)$$

$$(\hat{a}, \hat{b}, \hat{c}) = (20.0, 5.0, 5.0)$$

$$\alpha = 1.0$$

$$u_0(t) \equiv 0 \quad t \in [0, 20]$$

$$y_d(t) = \begin{cases} f(t) & \text{if } t \in [0, 10] \\ f(20-t) & \text{if } t \in [10, 20] \end{cases} \tag{2.54}$$

where

$$f(t) = \frac{-2t^3 + 30t^2}{100}$$

Note that there exists $u_d(t)$ satisfying (2.41) for (2.54):

$$u_d(t) = \begin{cases} \dfrac{-4t^3 + 54t^2 + 48t + 60}{300} & \text{if } 0 \le t \le 10 \\ \dfrac{-4(20-t)^3 + 66(20-t)^2 - 72(20-t) + 60}{300} & \text{if } 10 < t \le 20 \end{cases}$$

Dotted lines in Figure 2.3 show the output functions and input functions for $k = 0, 1, 2, 3, 5, 7$. Points in Figure 2.4 are values of norms for $k = 0, 1, \cdots, 19$.

Example 5 implies that our method presented in Theorem 5 can be applied to systems which many conventional methods failed to be applied to. The reason is that our method needs only (2.53) on

$$A = \begin{pmatrix} 0 & -b \\ 1 & -a \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \end{pmatrix}^T, C = \begin{pmatrix} 0 & c \end{pmatrix}$$

while the conventional methods require conditions which are not satisfied by this system, for example,

$$CB \neq 0$$

in [4, 5, 10],

$$B^T = PC$$

where $P > 0$ in [11], or

$$\int_0^{t_f} u(t) \int_0^t Ce^{A(t-\tau)} Bu(\tau) d\tau dt \geq 0$$

for any $u(t) \in L_2[0, t_f]$ [5, 42].

Example 6 *Consider a multi-input multi-output linear system*

$$\begin{pmatrix} y^1(s) \\ y^2(s) \end{pmatrix} = S(s) \begin{pmatrix} u^1(s) \\ u^2(s) \end{pmatrix}$$

$$S(s) = \begin{pmatrix} as^2 + 4s + 2 & -4s - 2 \\ -4s - 2 & s^2 + 9s + 5 \end{pmatrix}^{-1}$$

where a is a positive constant. Only step 2 of the design process will be examined.

step 2 *Let*

$$\hat{S}(s) = \begin{pmatrix} 2s^2 + 4s + 2 & -4s - 2 \\ -4s - 2 & s^2 + 9s + 5 \end{pmatrix}^{-1}$$

then

$$\begin{aligned} U(s) &= S(s)\hat{S}^{-1}(s) \\ &= \begin{pmatrix} \frac{d(s,2)}{d(s,a)} & 0 \\ \frac{(2-a)s^2(4s+2)}{d(s,a)} & 1 \end{pmatrix} \end{aligned}$$

where

$$d(s, a) = as^4 + (9a + 4)s^3 + (5a + 22)s^2 + 22s + 6$$

Since

- *If $a > 0$ then $U(s)$ is analytic for $Re\{s\} > 0$*
- *If $4.4 > a > 0$ then $U^T(-j\omega) + U(j\omega) > 0$ for $\omega \in R$*
- *If $a > 0$ then $H = 2/a > 0$*

$\hat{S}(s)$ with

$$4.4 > a > 0$$

satisfies the convergence condition.

Figure 2.5-2.7 show results of numerical simulations for the example. The parameters, $u_0(t)$, and the desired trajectory $y_d(t)$ on $[0, 10]$ are chosen as follows.

$$a = 1.0$$

$$y^1(0) = \dot{y}^1(0) = 0$$

$$y^2(0) = \dot{y}^2(0) = 0$$

$$y_d^1(t) = 1 - \cos\frac{\pi t}{5} \tag{2.55}$$

$$y_d^2(t) = \cos\frac{\pi t}{5} - 1 \tag{2.56}$$

$$\alpha = 2.0$$

$$u_0^1(t) \equiv 0 \quad (t \in [0, 10])$$

$$u_0^2(t) \equiv 0 \quad (t \in [0, 10])$$

Note that there exists $u_d(t)$ satisfying (2.41) for (2.55) and (2.56):

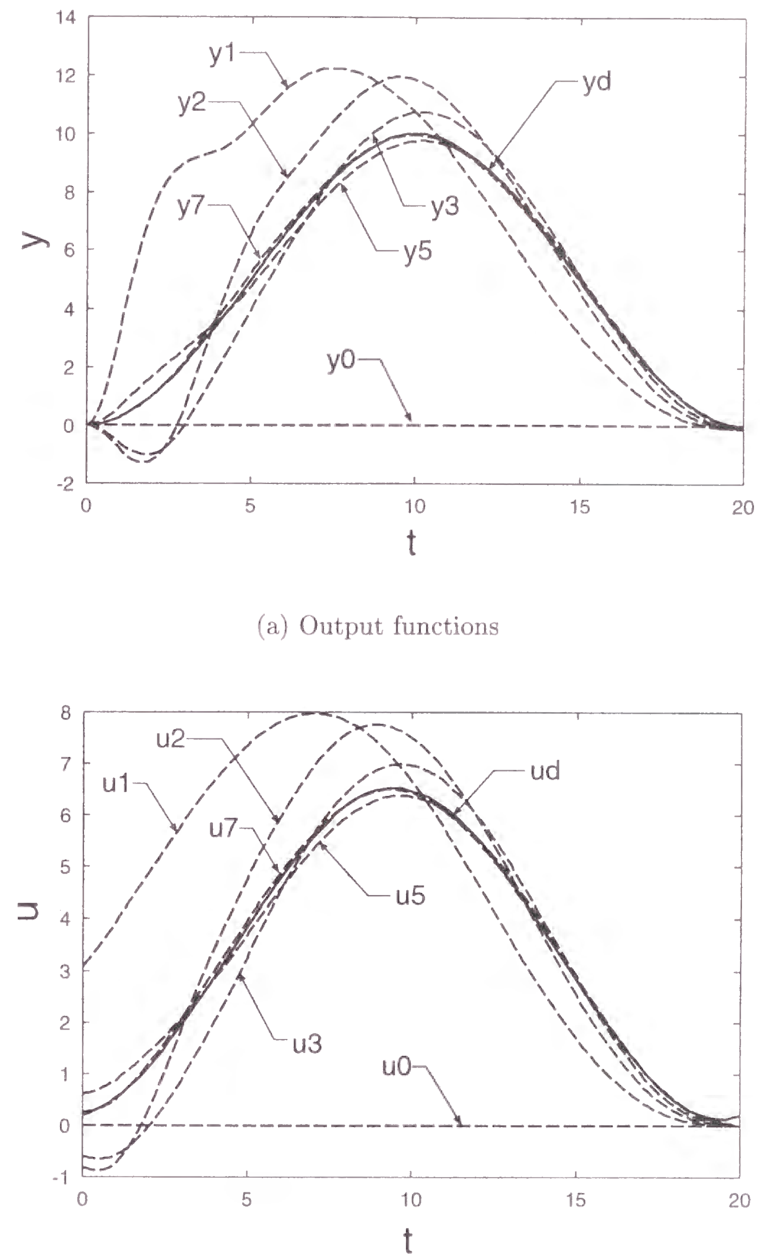
$$u_d^1(t) = \frac{1}{25}\{(\pi^2 - 100)\cos\frac{\pi t}{5} + 40\pi \sin\frac{\pi t}{5} + 100$$

$$u_d^2(t) = \frac{1}{25}\{(175 - \pi^2)\cos\frac{\pi t}{5} - 65\pi \sin\frac{\pi t}{5} - 175$$

Dotted lines in Figure 2.5 and Figure 2.6 show the output functions and input functions for $k = 0, 1, 2, 3, 4, 49, 99$. Plots in Figure 2.7 show increasing values of norms for $k = 0, 1, \cdots, 499$.

2.7 Concluding Remarks

In this chapter, convergence of functions generated by iterative learning control using adjoint systems was demonstrated for linear continuous-time systems. Several convergence conditions were presented and discussed from an applications viewpoint. Design method of the iterative learning control were presented for time-variant or time-invariant linear systems. Relaxation of the convergence conditions and L_∞ convergence were discussed, again from an applications viewpoint.



(a) Output functions

(b) Input functions

Figure 2.3: Simulation results of Example 5

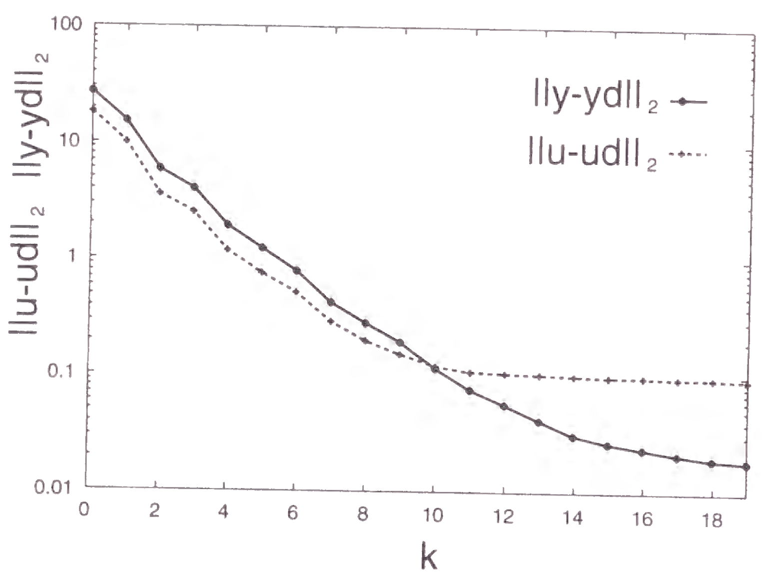
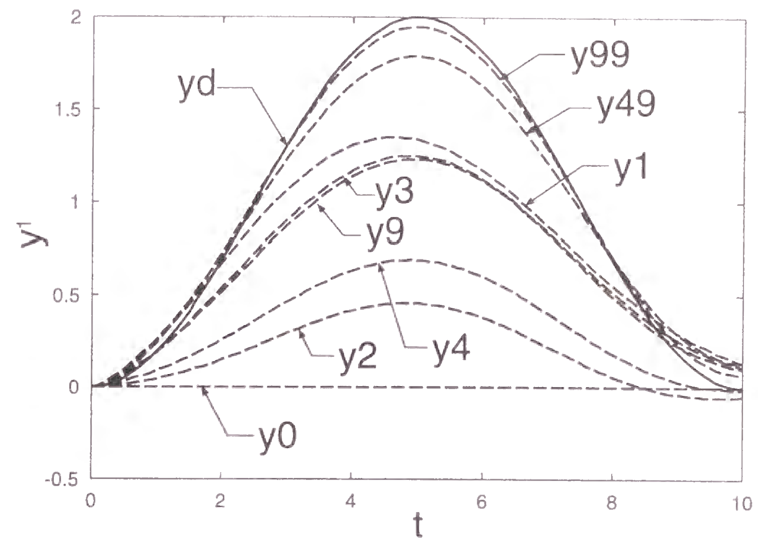
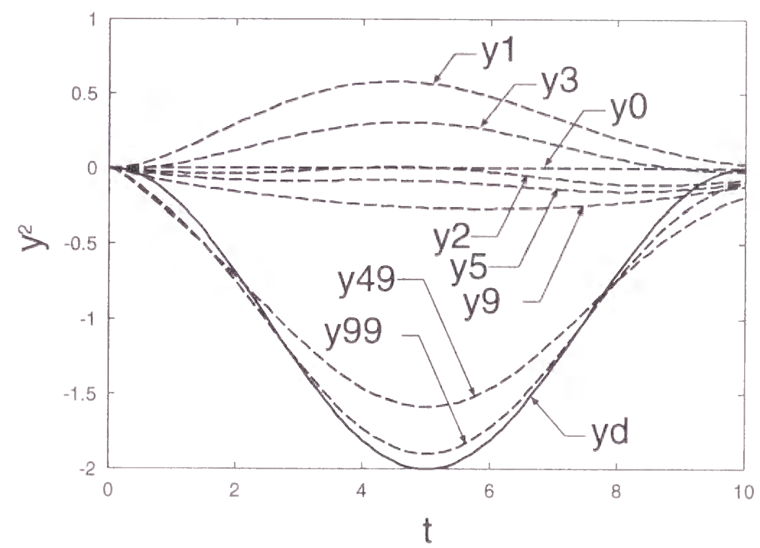


Figure 2.4: Errors vs. the number of the iteration of Example 5

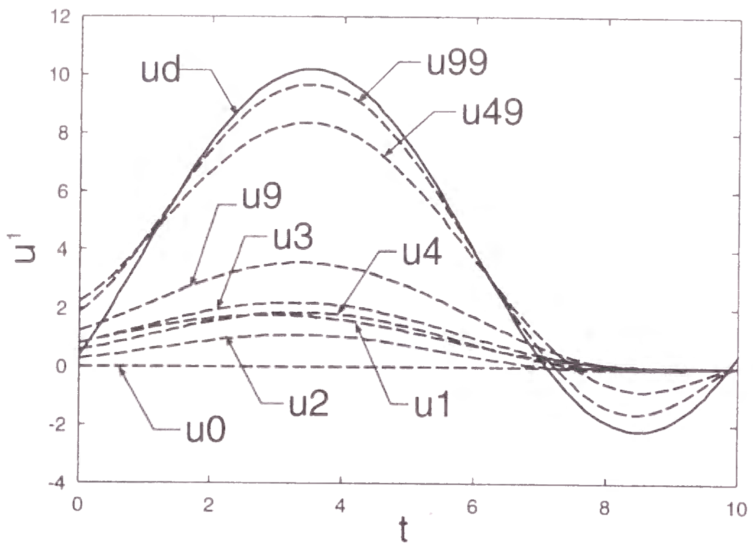


(a) Plots of $y^1(t)$

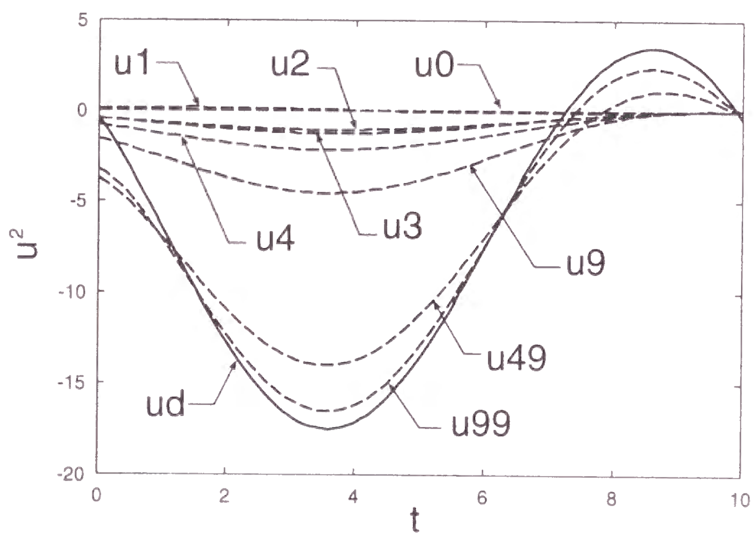


(b) Plots of $y^2(t)$

Figure 2.5: Output functions of Example 6



(a) Plots of $u^1(t)$



(b) Plots of $u^2(t)$

Figure 2.6: Input functions of Example 6

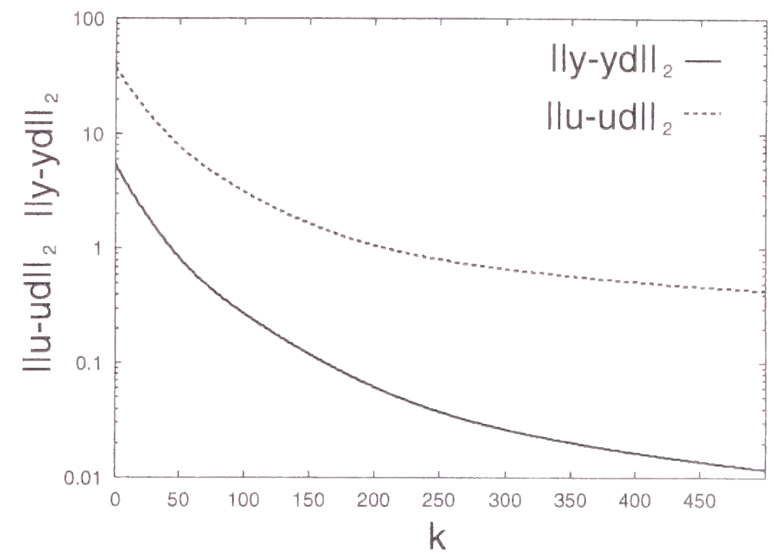


Figure 2.7: Errors vs. the number of the iteration of the example 6

Chapter 3

Convergence Rate and Robustness of Iterative Learning Control

3.1 Introduction

In the last chapter, we discussed design of the iterative algorithms to generate input functions, which is summarized as follows.

Problem 1 *Determine a mapping* $K : L_2[0, t_f] \to L_2[0, t_f]$ *of the algorithm*

$$u_{k+1} = u_k + K(Su_k - y_d) \quad (k = 0, 1, \cdots) \tag{3.1}$$

so that

$$\|Su_k - y_d\|_2 \to 0 \tag{3.2}$$

as $k \to \infty$ *where* S *is defined by (2.2).*

Many kinds of mappings other than ones discussed in this dissertation have been proposed as K, e.g. PID-type operators [43], approximate adjoint operators[27] etc. However, convergence rate of (3.2) has not been studied very much. On the other hand, it is important to investigate robustness of the algorithms against measurement noise, perturbation caused by initialization errors etc. The robustness means boundedness of the error sequence

$\{Su_k - y_d\}$ generated by (3.1) with bounded noise, i.e.

$$u_{k+1} = u_k + K(Su_k - y_d + d_k) \quad (k = 0, 1, \cdots) \tag{3.3}$$

and

$$\|d_k\|_2 \leq M \quad (k = 0, 1, \cdots)$$

where M is a positive constant which represents the noise level. One of the sufficient conditions to guarantee the robustness is

$$r = \sup \left\{ \frac{\|(I + SK)(Su - y_d)\|_2}{\|Su - y_d\|_2}; u \in L_2[0, t_f] \right\} < 1 \tag{3.4}$$

provided that it is assumed that there exists $u_d \in L_2[0, t_f$ such that $y_d = Su_d$ where I denotes the identical operator. The reason is that (3.3) with (3.4) leads to

$$Su_k - y_d = (I + SK)(Su_{k-1} - y_d) + SKd_{k-1}$$

and hence

$$\begin{aligned} \|Su_k - y_d\|_2 &\leq \|(I + SK)(Su_{k-1} - y_d)\|_2 + \|SKd_{k-1}\|_2 \\ &\leq r\|Su_{k-1} - y_d\|_2 + \|SKd_{k-1}\|_2 \\ &\leq r^k\|Su_0 - y_d\|_2 + \sum_{n=0}^{k-1} r^n \|SKd_{n-1}\|_2 \\ &\leq r^k\|Su_0 - y_d\|_2 + \frac{\|SK\|M}{1 - r} \end{aligned}$$

This inequality means that the limit of $\|Su_k - y_d\|_2$ is bounded by a value proportional to the noise level M; accumulation of the noise at iterations does not cause divergence of $\|Su_k - y_d\|_2$. Therefore, iterative learning control algorithms should satisfy (3.4) when a large number of iterations are needed or the noise is not small.

In this chapter, we demonstrate that there exists no bounded operator K of the algorithm (3.1) such that (3.4) is satisfied. Therefore, we modify the formulation of iterative learning control by introducing regularization ideas for the case that countermeasures against noise are needed.

3.2 Convergence Rate of Iterative Learning Control

Consider a linear system (2.1) with $D(t) = 0$, i.e.

$$\begin{aligned} \frac{d}{dt}x(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{3.5}$$

that is defined on the finite time interval $[0, t_f]$. Let y_d be a desired trajectory. Note that for most of applications such as robot manipulators, one can assume that $D(t) = 0$, i.e. there is no term which directly transform the input signal to the output signal without dynamics. Suppose that (3.5) is a single-input-single-output system for simplicity. Then the output $y(t)$ is denoted as

$$y(t) = h(t) + [Su](t)$$
$$h(t) = C(t)R(t, 0)x(0)$$

and

$$[Su](t) = \int_0^t C(t)R(t, \sigma)B(\sigma)u(\sigma)d\sigma \tag{3.6}$$

where $R(t, \sigma)$ is the transition matrix. Note that we can assume the initial condition $x(0) = 0$ or $h(t) \equiv 0$ by considering $y_d + h \in R(S)$ as desired trajectories instead of y_d where $R(S)$ denotes range of the operator S. In the following discussions, we consider

$$y = Su \quad u \in L_2[0, t_f]$$

and

$$y_d \in R(S)$$

as input-output mapping of the system and a desired trajectory, respectively.

If $U = I$ i.e. $S = \hat{S}$ in Theorem 1 or Theorem 5, an estimation of convergence speed of the algorithm is known [30]. The next theorem gives a similar estimation of convergence speed of the algorithm given in Theorem 5 for the general U.

Theorem 6 *Assume that the algorithm of Theorem 5 is applied to the SISO system (3.5) and the constant α satisfies*

$$0 < \alpha < \frac{2\beta}{\|S\|^2} \tag{3.7}$$

Then

$$\|Su_k - y_d\|_2^2 \leq \frac{\|Su_0 - y_d\|_2^2}{1 + kC\|Su_0 - y_d\|_2^2} \tag{3.8}$$

for $k = 0, 1, \cdots$ where

$$C = \frac{\alpha(2\beta - \alpha\|S\|^2)}{\|U\|^2\|u_0 - u_d\|_2^2}$$

Proof: From the assumption, we can easily see

$$S = \hat{S}U = U\hat{S} \tag{3.9}$$

From this equality and the Schwarz inequality, we have

$$\begin{aligned}\|S(u_k - u_d)\|_2^2 &= \langle S^*S(u_k - u_d), u_k - u_d\rangle \\ &\leq \|U^*\hat{S}^*S(u_k - u_d)\|_2\|u_k - u_d\|_2\end{aligned}$$

Moreover from (2.5), we obtain

$$\|S(u_k - u_d)\|_2^2 \leq \|U^*\|\|\hat{S}^*S(u_k - u_d)\|_2\|u_0 - u_d\|_2$$

and hence

$$\frac{\|S(u_k - u_d)\|_2^4}{\|U^*\|^2\|u_0 - u_d\|_2^2} \leq \|\hat{S}^*S(u_k - u_d)\|_2^2 \tag{3.10}$$

On the other hand, the algorithm (2.43) with (3.9) leads to

$$\begin{aligned}\|S(u_{k+1} - u_d)\|_2^2 &= \|S(u_k - u_d)\|_2^2 - \alpha\langle S(u_k - u_d), \hat{S}U\hat{S}^*S(u_k - u_d)\rangle \\ &\quad +\alpha^2\|S\hat{S}^*S(u_k - u_d)\|_2^2 \qquad (3.11)\\ &\leq \|S(u_k - u_d)\|_2^2 \\ &\quad -\alpha(2\beta - \alpha\|S\|^2)\|\hat{S}^*S(u_k - u_d)\|_2^2 \qquad (3.12)\end{aligned}$$

Since (3.7) implies $\alpha(2\beta - \alpha\|S\|^2) > 0$, substitution of (3.10) to (3.12) yields

$$\|S(u_{k+1} - u_d)\|_2^2 \leq \|S(u_k - u_d)\|_2^2(1 - C\|S(u_k - u_d)\|_2^2)$$

From this inequality we can obtain (3.8) since $\{\|S(u_k - u_d)\|_2^2; k = 0, 1, \cdots\}$ is a sequence of nonnegative numbers[30]. ■

The convergence speed of the right hand side in (3.8) is much slower than exponential functions such as r^k ($|r| < 1$) which play important roles for robustness against noise as stated in the section 3.1. On the other hand, since the integral operator S defined by (3.6) is a compact linear operator and has the infinite dimensional range, $\mathcal{R}(S)$ is non-closed, i.e.

$$\mathcal{R}(S) \neq \overline{\mathcal{R}(S)}$$

where $\overline{\mathcal{R}(S)}$ denotes the closure of $\mathcal{R}(S)$[29]. In this case, the equation

$$y_d = Su$$

for a given y_d is called an ill-posed problem because a small perturbation of y_d causes a large error of the solution u_d[29][44].

As stated in 2.5 of Chapter 2 convergence of the input function sequence of iterative learning control

$$u_k \to u_d$$

is not indispensable; convergence of the output function sequences

$$Su_k \to y_d, \tag{3.13}$$

is required. However, we can demonstrate that (3.13) cannot be exponential from the following proposition.

Proposition 1 *[45] If $\bar{y} \in \overline{R(S)}$ and $\bar{y} \notin R(S)$, then any sequence $\{u_k; u_k \in L_2[0, t_f]\}$ such that*

$$\lim_{n\to\infty} \|Su_n - \bar{y}\|_2 = 0$$

is unbounded.

Theorem 7 *There exists no bounded operator K which satisfies (3.4).*

Proof: Suppose that there exists a bounded operator K such that (3.4). Then the sequence $\{u_k; k = 0, 1, \cdots\}$ defined by (3.1) with $u_0 = 0$ satisfies

$$\begin{aligned} \|Su_k - y_d\|_2 &= \|(I + SK)(Su_k - y_d)\|_2 \\ &\leq r^k \|Su_0 - y_d\|_2 \end{aligned}$$

and

$$\begin{aligned} \|u_k\|_2 &= \left\| \sum_{n=0}^{k-1} K(Su_n - y_d) \right\|_2 \\ &\leq \sum_{n=0}^{k-1} \|K(Su_n - y_d)\|_2 \\ &\leq \|K\| \|Su_0 - y_d\|_2 \frac{1 - r^{k-1}}{1 - r} \end{aligned}$$

where r is a constant satisfying $0 < r < 1$. Therefore we have

$$\|u_k\|_2 \leq \frac{\|K\| \|y_d\|_2}{1 - r} \quad (k = 0, 1, \cdots) \tag{3.14}$$

which implies that $\{u_k; k = 0, 1, \cdots\}$ includes a weakly converging subsequence $\{u_{k_i}; i = 0, 1, \cdots\}$. Let u_d be a weak limit of the sequence. Then

$$\langle u_{k_i} - u_d, S^* g \rangle \to 0$$

equivalently,

$$\langle Su_{k_i} - y_d, g \rangle - \langle Su_d - y_d, g \rangle \to 0 \tag{3.15}$$

as $i \to \infty$ for any $g \in L_2[0, t_f]$. From $\|Su_{k_i} - y_d\|_2 \to 0$, we obtain

$$\langle Su_d - y_d, g \rangle = 0$$

for any $g \in L_2[0, t_f]$, and hence

$$y_d = Su_d \tag{3.16}$$

Note that (3.14) holds for $u_k = u_d$ [46], i.e.

$$\|u_d\|_2 \leq \frac{\|K\| \|y_d\|_2}{1 - r} \tag{3.17}$$

Consider $\bar{y}$ defined in Proposition 1 and a sequence $\{y_k; y_k \in R(S)\}$ such that

$$\|y_k - \bar{y}\|_2 \to 0 \tag{3.18}$$

as $k \to \infty$. Since (3.16) and (3.17) hold for each y_k, there exists a sequence $\{u_k\}$ such that

$$y_k = Su_k \tag{3.19}$$

and

$$\|u_k\|_2 \leq \frac{\|K\| \|y_k\|_2}{1 - r} \tag{3.20}$$

Substituting (3.19) for (3.18) yields

$$\|Su_k - \bar{y}\|_2 \to 0$$

as $k \to \infty$. On the other hand, since $\{y_k\}$ is a converging sequence, (3.20) leads to

$$\|u_k\|_2 \leq M$$

where M is a positive constant. This contradicts Proposition 1. ■

Since we can implement only bounded operators as physical systems, the theorem claims that it is impossible to choose an operator as K defined in Problem 1 so that the algorithm has the exponentially decreasing property (3.4).

3.3 Iterative Learning Control with a Regularization Term

As stated in section 3.1, iterative learning control should be designed so that it converges exponentially, from a robustness viewpoint. On the other hand, this is impossible for the algorithm of Problem 1. Therefore, we have to modify the formulation of the problem when countermeasures against noise are needed, e.g. when the noise is not small and a large number of iterations are needed. In this section, we modify the formulation of the problem of iterative learning control by introducing regularization ideas in order to design exponentially converging algorithms.

In Problem 1, the functional to be minimized is

$$\|Su - y_d\|_2^2 \tag{3.21}$$

The minimizers of (3.21) do not always exist. The fact is one of the reasons why we have to check existence of u_d such that $y_d = Su_d$ in order to apply the algorithms of Theorem 1 and 5. Moreover, non-closedness of S implies that the convergence (3.21) cannot be exponential. Therefore, we will replace the functional (3.21) to be minimized with

$$J_Q(u) = \|Su - y_d\|_2^2 + \langle u, Qu \rangle \tag{3.22}$$

where $Q : L_2^m[0, t_f] \to L_2^m[0, t_f]$ is a positive self-adjoint operator, i.e.

$$\langle u, Qu \rangle > 0 \quad (u \neq 0)$$

and

$$Q = Q^*.$$

Since

$$J_Q(u + \eta) = J_Q(u) + 2\langle \eta, (S^*S + Q)u - S^*y_d \rangle + \langle \eta, (S^*S + Q)\eta \rangle$$

where u and $\eta \in L_2[0, t_f]$, a necessary condition for the minimizer u is

$$(S^*S + Q)u - S^*y_d = 0 \tag{3.23}$$

On the other hand, the operator $S^*S + Q$ has the bounded inverse because $S^*S + Q$ is positive self-adjoint. Therefore $J_Q(u)$ always has the unique minimizer[47]

$$u^* = (S^*S + Q)^{-1}S^*y_d. \tag{3.24}$$

Although u_d is not equivalent to u^*, at the final part of this section we will show that it is possible to choose Q of $J_Q(u)$ so that

$$y^* = Su^*$$

is arbitrarily close to $y_d = Su_d$ with $\|u^*\|_2$ kept bounded. By adding $\langle u, Qu \rangle$ to (3.21), the minimizing problem is replaced with a problem which has the unique minimizer for any $y_d \in L_2[0, t_f]$, and hence there exists exponentially converging algorithms as is shown in the following discussions. The added term $\langle u, Qu \rangle$ of (3.22) is called regularization term[44]. Note that the left hand side of (3.23) corresponds to the gradient function of $J_Q(u)$.

As stated in Chapter 2, conventional gradient methods cannot be applied to generating a sequence of input functions which minimize $J_Q(u)$ because the adjoint operator S^* is unavailable. We will examine a method utilizing $\hat{S}^*$ which is different from S^*. Let S be

$$S = \hat{S}U$$

or

$$S^* = U^*\hat{S}^*$$

where U is a positive self-adjoint operator which represents uncertainty included in the system S. Consider

$$J_{\epsilon U}(u) = \|Su - y_d\|_2^2 + \langle u, \epsilon Uu \rangle$$

where ϵ is a positive constant. Then

$$J_{\epsilon U}(u + \eta) = J_{\epsilon U}(u) + 2\langle \eta, (U^*\hat{S}^*S + \epsilon U)u - U^*\hat{S}^*y_d \rangle + \langle \eta, (S^*S + \epsilon U)\eta \rangle$$

and hence the unique minimizer of $J_{\epsilon U}(u)$ is the solution of

$$U^*\{\hat{S}^*(Su - y_d) + \epsilon u\} = 0. \tag{3.25}$$

The next theorem gives an algorithm to minimize $J_{\epsilon U}(u)$ by utilizing

$$\hat{S}^*(h + Su - y_d) + \epsilon u$$

instead of the unavailable $U^*\{\hat{S}^*(Su - y_d) + \epsilon u\}$. (See Figure 3.1)

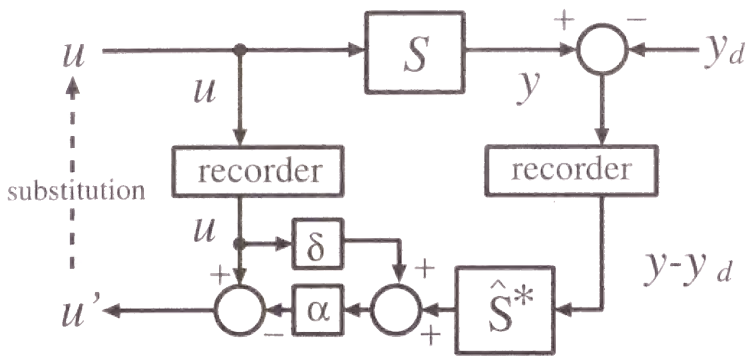


Figure 3.1: Iterative learning control for $J_{\epsilon U}(u)$

Theorem 8 *Assume that U is positive self-adjoint and $\{u_k\}$ is generated by*

$$u_{k+1} = u_k - \alpha\{\hat{S}^*(Su_k - y_d) + \epsilon u_k\} \tag{3.26}$$

with $u_0 \in L_2[0, t_f]$ where

$$0 < \alpha < \frac{2}{M_1 M_2} \tag{3.27}$$

$$M_1 = \sup\left\{\frac{\|U^{-1}u\|}{\|u\|_2}; u \in L_2[0, t_f], u \neq 0\right\}$$

and

$$M_2 = \sup\left\{\frac{\|(S^*S + \epsilon U)u\|}{\|u\|_2}; u \in L_2[0, t_f], u \neq 0\right\}$$

Then

$$\|u_k - u^*\|_2 \leq r^k M \tag{3.28}$$

where u^* is the minimizer of $J_{\epsilon U}(u)$; r and M are constants satisfying $0 < r < 1$ and $0 < M$, respectively.

Proof: Since U is a positive self-adjoint operator, there exists the bounded inverse U^{-1} [46]. From (3.26), we obtain

$$u_{k+1} = u_k - \alpha U^{-1}\{S^*(h + Su_k - y_d) + \epsilon U u_k\} \tag{3.29}$$

Since $S^*S + \epsilon U$ is also a positive self-adjoint operator, (3.29) leads to

$$u_{k+1} = u_k - \alpha U^{-1}(S^*S + \epsilon U)\{u_k - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\}$$

and hence

$$\begin{aligned} & u_{k+1} - (S^*S + \epsilon U)^{-1}S^*(y_d - h) \\ &= \{I - \alpha U^{-1}(S^*S + \epsilon U)\}\{u_k - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\} \end{aligned}$$

On the other hand,

$$0 < \|I - \alpha U^{-1}(S^*S + \epsilon U)\|_2 < \left|1 - \alpha\|U^{-1}\|\|S^*S + \epsilon U\|\right|$$

for $\alpha > 0$. For α chosen as (3.27), we have

$$\|u_{k+1} - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2 \leq r\|u_k - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2$$

where $0 < r < 1$, therefore

$$\|u_k - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2 \leq r^k\|u_0 - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2$$

This completes the proof. ■

It is easy to see robustness of the algorithm. Consider the algorithm (3.26) with additional term d_k which stems from measurement noise, perturbation caused by initialization error etc,

$$u_{k+1} = u_k - \alpha\{\hat{S}^*(h + Su_k - y_d + d_k) + \epsilon u_k\}$$

where

$$\|d_k\| \leq L$$

for $k = 0, 1, \cdots$; L is a positive constant. Then reasoning which is similar to the proof of Theorem 8 yields

$$\begin{aligned} & \|u_{k+1} - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2 \\ \leq\ & r\|u_k - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2 - \alpha\hat{S}^*d_k \end{aligned}$$

and hence

$$\begin{aligned} & \|u_k - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2 \\ \leq\ & r^k\|u_0 - (S^*S + \epsilon U)^{-1}S^*(y_d - h)\|_2 + \frac{\alpha\|\hat{S}^*\|L}{1 - r} \end{aligned}$$

This implies that the limit of u_k is within neighborhood of the minimizer by the radius proportional to the noise level L.

3.4 Interpretation of the Minimizer of the Modified Problem

The input functions $\{u_k\}$ generated by the algorithm (3.26) converges to

$$u^* = (S^*S + \epsilon U)^{-1} S^* y_d$$

and hence the output functions converges to

$$y^* = S(S^*S + \epsilon U)^{-1} S^* y_d$$

that is different from the original desired trajectory y_d. However, we can show the norm of the difference is bounded by a positive constant which is proportional to ϵ.

Theorem 9 *Assume that there exists $u_d \in L_2[0, t_f]$ such that*

$$y_d = Su_d$$

Then

$$\|Su^* - y_d\|_2^2 \leq \epsilon M_U \|u_d\|_2^2 \tag{3.30}$$

and

$$\|u^*\| \leq \frac{M_U}{m_U} \tag{3.31}$$

where

$$M_U = \sup \left\{ \frac{\langle u, Uu \rangle}{\|u\|_2^2}; u \in L_2[0, t_f], u \neq 0 \right\}$$

$$m_U = \inf \left\{ \frac{\langle u, Uu \rangle}{\|u\|_2^2}; u \in L_2[0, t_f], u \neq 0 \right\}$$

and u^* is the minimizer of $J_{\epsilon U}(u)$.

Proof: From positiveness of U, we have

$$\|Su^* - y_d\| + \epsilon m_U \|u^*\|_2^2 \leq J_{\epsilon U}(u^*)$$

This inequality leads to

$$\|Su^* - y_d\|_2^2 \leq \|Su - y_d\|_2^2 + \epsilon M_U \|u\|_2^2 \tag{3.32}$$

and

$$\|u^*\|_2^2 \leq \frac{1}{\epsilon m_U} \|Su - y_d\|_2^2 + \frac{M_U}{m_U} \|u\|_2^2 \tag{3.33}$$

for any $u \in L_2[0, t_f]$ because

$$J_{\epsilon U}(u^*) \leq J_{\epsilon U}(u)$$

and

$$J_{\epsilon U}(u) \leq \|Su - y_d\|_2^2 + \epsilon M_U \|u\|_2^2$$

Substitution of u_d for u of (3.32) and (3.33) yield (3.30) and (3.31), respectively. ■

This theorem means that $\|Su^* - y_d\|_2$ can be arbitrarily close to 0 by making ϵ small whereas the least upper bound of $\|u^*\|_2$ is independent of ϵ.

3.5 Concluding Remarks

In this chapter, first, we discussed speed of the convergence of iterative learning control from a robustness viewpoint; we demonstrated that speed of the convergence should be of exponential functions. Second, we pointed out that there is no exponentially converging algorithm for the problem given in the last chapter. Finally, in order to deal with the case that the noise level is not small, we modified the formulation of the problem of iterative learning control by introducing the regularization term. Based on the idea given in the last chapter, we also presented an iterative learning control algorithm which converges exponentially.

Chapter 4

Iterative Learning Control for Sampled-Data Systems and the Inverse Systems

4.1 Introduction

In the last two chapters, we discussed iterative learning control for continuous-time systems. The dynamical systems to be controlled was defined on the finite time interval $[0, t_f]$. Since the time interval is short for most applications of iterative learning control, stability of poles and zeros of the system transfer functions was not considered. On the other hand, for implementation of iterative learning control, it is necessary to record input functions or the measured output functions and process those functions repetitively. Therefore, it is convenient to implement the iterative learning control with a sampler, a hold and digital computers. In this case, however, we have to consider stability of poles and zeros of the discrete-time system because the number of the sample points or the discrete-time interval increases as the sampling period Δ goes to 0 even if the continuous-time interval $[0, t_f]$ is fixed.

In this chapter, we consider iterative learning control for sampled-data systems with a 0-order hold and a sampler which have the same sampling period. First, we discuss unstable zeros which emerge when we consider sampled-data systems and iterative learning control for the systems. Second, we demonstrate that it is not necessary to consider stability of zeros of the sampled-data systems for any small sampling period when relative degree of the transfer function of the continuous-time system is 0, 1 or 2 even if it has unstable zeros. Finally, numerical examples are given to illustrate the results.

4.2 Mathematical Preliminaries and Motivation of the Study

Consider a linear continuous-time SISO system defined on $[0, t_f]$

$$\begin{aligned} \frac{d}{dt}x(t) &= A_c x(t) + b_c u(t) \\ y(t) &= cx(t) + du(t) \end{aligned} \tag{4.1}$$

with the initial condition $x(0) = 0$ where $x \in R^n$, $u \in R$ and $y \in R$. Then the input-output mapping defined by (4.1) on $[0, t_f]$ is expressed as

$$y = Su \quad (u, y \in L_2[0, t_f])$$

where

$$Su = \int_0^t ce^{A_c(t-\tau)} b_c u(\tau) d\tau + du(t) \tag{4.2}$$

The transfer function of (4.1) is

$$\begin{aligned} G(s) &= c(sI - A_c)^{-1} b_c + d \\ &= \frac{c\,\mathrm{adj}(sI - A_c) b_c + d \det(sI - A_c)}{\det(sI - A_c)} \\ &= \frac{K(s - \gamma_1) \cdots (s - \gamma_m)}{(s - p_1) \cdots (s - p_n)} \end{aligned} \tag{4.3}$$

where $\mathrm{adj}(sI - A_c)$ indicates the adjoint matrix of $sI - A_c$.

Consider the continuous-time system (4.1) with a sampler of the sampling period Δ and a 0-order hold at the same cycle which generates the input.

Then the relationship of y, x and u at sample points $\{k\Delta; k = 0, 1, \cdots\}$ is expressed as the discrete-time system

$$\begin{aligned} x((k+1)\Delta) &= A_\Delta(k\Delta) + b_\Delta u(k\Delta) \\ y(k\Delta) &= cx(k\Delta) + du(k\Delta) \end{aligned} \tag{4.4}$$

where

$$A_\Delta = \exp A_c \Delta$$

$$b_\Delta = \int_0^\Delta \exp(A_c \tau) b_c d\tau \tag{4.5}$$

The pulse transfer function is

$$\begin{aligned} H(z) &= c(zI - A_\Delta)^{-1} b_\Delta + d \\ &= \frac{N(z)}{D(z)} + d \end{aligned} \tag{4.6}$$

where

$$N(z) = \det \begin{bmatrix} zI - A_\Delta & -b_\Delta \\ c & 0 \end{bmatrix} \tag{4.7}$$

$$D(z) = \det(zI - A_\Delta) \tag{4.8}$$

Relationship between poles and zeros of sampled-data systems and those of continuous-time systems has been studied. By considering eigenvalues of the matrix A_c and ones of the matrix $A_\Delta = \exp(A_c\Delta)$, we can see that p_i of (4.3) corresponds to $\exp(p_i\Delta)$ of (4.6). It should be noted that as $\Delta \to 0$, all poles of (4.6) converges to 1 that is on the boundary of the stable area and the unstable one. However, there is no such simple relationship between zeros of (4.3) and ones of (4.6). Unstable zeros can emerge in (4.6) even if there is no unstable ones in (4.3). When $cb_\Delta \neq 0$, $H(z)$ is expressed as

$$H(z) = \frac{cb_\Delta(z - q_1(\Delta)) \cdots (z - q_{n-1}(\Delta))}{(z - \exp(p_1\Delta)) \cdots (z - \exp(p_n\Delta))} + d. \tag{4.9}$$

The next lemma should be noted for the following discussions.

Lemma 1 *If $n-m \geq 1$ then there exist $\epsilon_0 > 0$ such that*

$$\left|\frac{cb_\Delta}{\Delta^{n-m}}\right| > 0 \tag{4.10}$$

for any $\Delta \in (0, \epsilon_0)$.

Proof: See Appendix A. ■

Remark 3 *(4.10) implies*

$$|cb_\Delta| > 0 \tag{4.11}$$

for any $\Delta \in (0, \epsilon_0)$.

Suppose

$$\Delta = \frac{t_f}{N}$$

where N is a natural number. Then the input-output mapping of (4.4) on $\{0, \Delta, 2\Delta, \cdots, t_f - \Delta, t_f\}$ is

$$w_\Delta = \Gamma_\Delta v_\Delta$$

where

$$v_\Delta = \begin{bmatrix} u(0) & u(\Delta) & \cdots & u(t_f) \end{bmatrix}^T$$
$$w_\Delta = \begin{bmatrix} y(0) & y(\Delta) & \cdots & y(t_f) \end{bmatrix}^T$$
$$\Gamma_\Delta = \begin{bmatrix} d & 0 & \cdots & & 0 \\ cb_\Delta & d & & & \\ cA_\Delta b_\Delta & cb_\Delta & \ddots & & \vdots \\ \vdots & \ddots & \ddots & d & 0 \\ cA_\Delta^{N-1} b_\Delta & \cdots & cA_\Delta b_\Delta & cb_\Delta & d \end{bmatrix} \tag{4.12}$$

Let y^* be a desired trajectory defined on $[0, t_f]$ Then, among previous studies on iterative learning control for discrete-time systems [25][24], one of the most common formulation of the problem is

$$\text{minimize} \quad |\Gamma_\Delta v - \sigma_\Delta y^*| \quad v \in R^{N+1} \tag{4.13}$$

where $|\cdot|$ indicates the Euclidean norm; the operator

$$\sigma_\Delta : L_2[0, t_f] \to R^{N+1}$$

is the sampling operator defined as

$$\sigma_\Delta y^* = [y^*(0)\ y^*(\Delta)\ \cdots\ y^*((N-1)\Delta)\ y^*(t_f)]^T .$$

Since the minimizer of (4.13) with the minimum norm (It should be noted that the minimizers are generally not unique.) is

$$v_\Delta^* = \Gamma_\Delta^+ \sigma_\Delta y^*$$

where Γ_Δ^+ indicates the Moore-Penrose pseudo-inverse matrix of Γ_Δ, one can define the iterative learning control as an iterative algorithm which generates $\{v_n; v_n \in R^{N+1}, n = 0, 1, \cdots\}$ such that

$$v_n \to \Gamma_\Delta^+ \sigma_\Delta y^* \tag{4.14}$$

as $n \to \infty$. It should be noted that we can obtain the minimizer $\Gamma_\Delta^+ \sigma_\Delta y^*$ as the output of the inverse system of (4.4). If $d \neq 0$ then $\Gamma_\Delta^+ = \Gamma_\Delta^{-1}$ that corresponds to

$$\begin{aligned} x((k+1)\Delta) &= \left(A_\Delta - \frac{b_\Delta c}{d}\right) x(k\Delta) + \frac{b_\Delta}{d} y(k\Delta) \\ u(k\Delta) &= -\frac{c}{d} x(k\Delta) + \frac{1}{d} y(k\Delta); \end{aligned}$$

there exists a positive δ such that $cb_\Delta \neq 0$ for any $\Delta \in (0, \delta)$ and if $d = 0$ and $cb_\Delta \neq 0$ then

$$\Gamma_\Delta^+ = \begin{bmatrix} 0 & \begin{bmatrix} cb_\Delta & & 0 \\ \vdots & \ddots & \\ cA_\Delta^{N-1} b_\Delta & \cdots & cb_\Delta \end{bmatrix}^{-1} \\ \vdots & \\ 0 & \\ 0 & 0 \quad \cdots \quad 0 \end{bmatrix} \tag{4.15}$$

that corresponds to

$$\begin{aligned} x((k+1)\Delta) &= \left(A_\Delta - \frac{b_\Delta c A_\Delta}{cb_\Delta}\right) x(k\Delta) + \frac{b_\Delta}{cb_\Delta} y((k+1)\Delta) \\ u(k\Delta) &= -\frac{cA_\Delta}{cb_\Delta} x(k\Delta) + \frac{1}{cb_\Delta} y((k+1)\Delta). \end{aligned}$$

Most of previous studies on iterative learning control focused only on how to design algorithms to update v_n so that (4.14) holds. However, behavior of the minimizer $\Gamma_\Delta^+\sigma_\Delta y^*$ for varying Δ is not studied very much. Since the dimension $N+1 = t_f/\Delta + 1 \to \infty$ as $\Delta \to 0$, undesirable effect of unstable poles or zeros of (4.6) may possibly occur for sampled-data systems with a small Δ.

Åström et. al.[48] showed the next result on limiting behavior of zeros in $H(z)$.

Theorem 10 *[48] Assume that $G(s)$ is strictly proper, i.e. $m < n$. Then for almost all $\Delta > 0$ there exist $n-1$ zeros in $H(z)$ and $H(z)$ approaches*

$$K\frac{\Delta^{n-m}}{(n-m)!}\frac{(z-1)^m B_{n-m}(z)}{(z-1)^n}$$

as $\Delta \to 0$ where

$$B_p(z) = \beta_{p1}z^{p-1} + \beta_{p2}z^{p-2} + \cdots + \beta_{pp}$$

$$\beta_{pk} = \sum_{i=1}^{k}(-1)^{k-i}i^p\begin{pmatrix} p+1 \\ k-i \end{pmatrix}$$

This theorem shows that $n-1$ zeros in $H(z)$ can be classified into m zeros which converge to 1 and $n-m-1$ ones which converge to zeros of $B_{n-m}(z)$. The next result is known about the former zeros.

Theorem 11 *[49] The zero of $H(z)$ that converges to 1 is Taylor-expanded as*

$$q_i(\Delta) \;=\; 1 + \gamma_i\Delta + \frac{\gamma_i^2}{2}\Delta^2 + \left(\frac{\gamma_i^3}{6} + \frac{\gamma_i c b_c}{12G'(\gamma_i)}\right)\Delta^3 + O(\Delta^4)$$

where $i = 1, 2, \cdots, m$.

If the relative degree $n-m=2$, then zeros of $H(z)$ converge to 1 or -1 because $B_2(z) = z+1$. We can show the next result about the zero that converges to -1.

Lemma 2 *Assume that $n-m=2$ and zeros $\{q_i(\Delta); i = 1, 2, \cdots, m\}$ of $H(z)$ is expanded as (4.16) when $m \geq 1$. Then, the other zero is expanded as*

$$q_{n-1}(\Delta) = \begin{cases} -1 - \frac{4cA_c^2b_c - \mathrm{trace}(A_c)cA_cb_c}{3cA_cb_c}\Delta \\ +O(\Delta^2) \quad (if \;\; n=2) \\ -1 - \frac{cA_c^2b_c}{3cA_cb_c}\Delta + O(\Delta^2) \\ \quad (if \;\; n \geq 3) \end{cases}$$

Proof: See Appendix B ■

In the following sections, we discuss the difference

$$\theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^* - u^* \tag{4.16}$$

and its relationship with behavior of the zeros of $H(z)$ for small Δ where u^* is a function defined on $[0, t_f]$ satisfying $y^* = Su^*$; $\theta_\Delta : R^{N+1} \to L_2[0, t_f]$ is a 0-order hold operator defined as

$$[\theta_\Delta v](t) = \begin{cases} v(k) & \text{if} \quad t \in [(k-1)\Delta, k\Delta) \\ & \quad\;\;\; (k = 1, 2, \cdots, \frac{t_f}{\Delta}) \\ v\left(\frac{t_f}{\Delta}+1\right) & \text{if} \quad t = t_f \end{cases}$$

It should be note that if (4.16) is small, the output error including the inter-sample error or the ripple

$$S\theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^* - y^*$$

is also small.

4.3 A Limiting Property of the Inverse of the Sampled-Data System

In the following sections, the following notations are used.

$$\begin{aligned} C^k[0, t_f] \;\; &: \;\; \text{the class of } k\text{-times continuously differentiable} \\ & \quad\;\; \text{functions on } [0, t_f] \\ \|u\|_\infty \;\; &= \;\; \sup\{|u(t)|; t \in [0, t_f]\} \\ \|u\|'_\infty \;\; &= \;\; \sup\{|u(t)|; t \in [0, t_f)\} \end{aligned}$$

We present the next proposition for sake of the following discussions about the difference (4.16).

Proposition 2 *If $u^* \in C^0[0,t_f]$ then*

$$\left\| \frac{d^i}{dt^i} S\theta_\Delta \sigma_\Delta u^* - \frac{d^i}{dt^i} Su^* \right\|_\infty \to 0 \tag{4.17}$$

as $\Delta \to 0$ for $i = 0, 1, \cdots, n-m$.

Proof: From (4.2), we have

$$\begin{aligned} \left\| \frac{d^i}{dt^i} S \right\| &= \sup \left\{ \frac{\left\| \frac{d^i}{dt^i} Su \right\|_\infty}{\|u\|_\infty} ; u(t) : \text{piece-wise continuous function on } [0, t_f] \right\} \\ &< +\infty \end{aligned} \tag{4.18}$$

Since $u^* \in C^0[0,t_f]$, we have

$$\lim_{\Delta \to 0} \|\theta[\sigma[u^*]] - u^*\|_\infty = 0 \tag{4.19}$$

(4.18) and (4.19) leads to (4.17). ■

Remark 4 *The convergence (4.17) is independent of stability of poles of $H(z)$ which is equivalent to stability of poles of $G(s)$.*

4.3.1 The Case of Relative Degree 0 or 1

Intuitively, convergence of the difference (4.16) as $\Delta \to 0$ depends on stability of zeros of $H(z)$, behavior of which is complicated as stated in Section 4.2. However, the next theorem shows that the convergence is independent of the stability.

Theorem 12 *Assume that y^* satisfies $y^* = Su^*$ where $u^* \in C[0,t_f]$. If $n - m = 0$ then*

$$\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*\|_\infty \to 0 \tag{4.20}$$

as $\Delta \to 0$. If $n - m = 1$ then

$$\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*\|'_\infty \to 0 \tag{4.21}$$

as $\Delta \to 0$.

Remark 5 *If the relative degree $n - m = 1$ equivalently $d = 0$ then from (4.15) we have $[\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*](N\Delta) \equiv 0$ equivalently $[\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*](t_f) \equiv 0$. In this case, the convergence (4.20) does not generally hold; the convergence (4.21) that is convergence except $t = t_f$ is the best result.*

The following lemmas are prepared to prove Theorem 12.

Lemma 3 *If $u^* \in C[0,t_f]$ then there exists $\epsilon_0 > 0$ such that*

$$\Gamma_\Delta \Gamma_\Delta^+ \sigma_\Delta Su^* = \sigma_\Delta Su^* \tag{4.22}$$

for any $\Delta \in (0, \epsilon_0)$.

Proof: See Appendix C ■

Lemma 4 *Assume that $n - m = 0$ and*

$$\begin{aligned} v &= [v(0)\ v(1)\ \cdots\ v(N)]^T \in R^{N+1} \\ w &= [w(0)\ w(1)\ \cdots\ w(N)]^T \in R^{N+1} \end{aligned}$$

satisfy $w = \Gamma_\Delta v$. Then

$$\begin{aligned} &\max \{|v(k)|; k = 0, 1, \cdots, N\} \\ &\quad \le L \max \{|w(k)|; k = 0, 1, \cdots, N\} \end{aligned} \tag{4.23}$$

for any $\Delta = t_f/N \in (0, \epsilon_1)$ where ϵ_1 and L are positive constants.

Proof: See Appendix D ■

Lemma 5 *Assume that $n - m = 1$ and*

$$\begin{aligned} v &= [v(0)\ v(1)\ \cdots\ v(N)]^T \in R^{N+1} \\ w &= [w(0)\ w(1)\ \cdots\ w(N)]^T \in R^{N+1} \end{aligned}$$

satisfy $w = \Gamma_\Delta v$. Then

$$\begin{aligned} &\max \{|v(k)|; k = 0, 1, \cdots, N-1\} \\ &\quad \le M \max \{|[\delta_\Delta w](k)|; k = 0, 1, \cdots, N-1\} \end{aligned} \tag{4.24}$$

for any $\Delta = t_f/N \in (0, \epsilon_0)$ *where* M *is a positive constant;* ϵ_0 *was given in Lemma 3. The operator* $\delta_\Delta : R^{N+1} \to R^N$ *is defined as*

$$[\delta_\Delta w](k) = \frac{|w(k+1) - w(k)|}{\Delta}$$

Proof: See Appendix E ■

Proof of Theorem 12: Note that

$$\begin{aligned} &\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*\|_\infty \\ &\leq \|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - \theta_\Delta \sigma_\Delta u^*\|_\infty \\ &\quad + \|\theta_\Delta \sigma_\Delta u^* - u^*\|_\infty \end{aligned} \tag{4.25}$$

and since $u^* \in C[0, t_f]$,

$$\|\theta_\Delta \sigma_\Delta u^* - u^*\|_\infty \to 0$$

as $\Delta \to 0$. Those properties hold with the norm $\|\cdot\|'_\infty$. Therefore, what we have demonstrate in order to prove (4.20) or (4.21) is $\|f_\Delta\|_\infty \to 0$ or $\|f_\Delta\|'_\infty \to 0$, respectively, where

$$f_\Delta = \theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - \theta_\Delta \sigma_\Delta u^* \tag{4.26}$$

(1) The case of $n - m = 0$

From $\sigma_\Delta S \theta_\Delta = \Gamma_\Delta$, we have

$$\begin{aligned} \sigma_\Delta S f_\Delta &= \Gamma_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - \Gamma_\Delta \sigma_\Delta u^* \\ &= \Gamma_\Delta \sigma_\Delta \theta_\Delta (\Gamma_\Delta^+ \sigma_\Delta y^* - \sigma_\Delta u^*) \\ &= \Gamma_\Delta \sigma_\Delta f_\Delta \end{aligned}$$

Moreover, by Lemma 4, we obtain

$$\begin{aligned} &\max\{|[\sigma_\Delta f_\Delta](k)|; k = 1, 2, \cdots, N+1\} \\ &\leq L \max\{|[\sigma_\Delta S f_\Delta](k)|; k = 1, 2, \cdots, N+1\} \end{aligned} \tag{4.27}$$

for $\Delta \in (0, \epsilon_1)$. On the other hand, Proposition 2 leads to

$$\|S\theta_\Delta \sigma_\Delta u^* - y^*\|_\infty \to 0 \tag{4.28}$$

as $\Delta \to 0$. From Lemma 3, we have

$$\begin{aligned} \sigma_\Delta S f_\Delta &= \Gamma_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - \sigma_\Delta S \theta_\Delta \sigma_\Delta u^* \\ &= \sigma_\Delta y^* - \sigma_\Delta S \theta_\Delta \sigma_\Delta u^* \end{aligned} \tag{4.29}$$

(4.28) and (4.29) implies

$$\max\{|[\sigma_\Delta S f_\Delta](k)|; k = 1, 2, \cdots, N+1\} \to 0 \tag{4.30}$$

and moreover

$$\max\{|[\sigma_\Delta f_\Delta](k)|; k = 1, 2, \cdots, N+1\} \to 0$$

because of (4.27). This leads to (4.20) since f_Δ was defined by (4.26).

(2) The case of $n - m = 1$

Since we have

$$\sigma_\Delta S f_\Delta = \Gamma_\Delta \sigma_\Delta f_\Delta$$

as the case of $n - m = 0$, Lemma 5 leads to

$$\begin{aligned} &\max\{|[\sigma_\Delta f_\Delta](k)|; k = 1, 2, \cdots, N\} \\ &\leq M \max\{|[\delta_\Delta \sigma_\Delta S f_\Delta](k)|; k = 1, \cdots, N\} \end{aligned} \tag{4.31}$$

for $\Delta \in (0, \epsilon_0)$. On the other hand, since (4.29) is satisfied for $n - m = 1$, we have

$$\delta_\Delta \sigma_\Delta S f_\Delta = \delta_\Delta \sigma_\Delta y^* - \delta_\Delta \sigma_\Delta S \theta_\Delta \sigma_\Delta u^* \tag{4.32}$$

for $\Delta \in (0, \epsilon_0)$. Since $n - m = 1$, $G(s)$ is expressed as

$$\begin{aligned} G(s) &= K \frac{s^{n-1} + b_1 s^{n-2} + \cdots + b_{n-1}}{s^n + a_1 s^{n-1} + \cdots + a_n} \\ &= \frac{K}{s} + G_1(s) \end{aligned}$$

where

$$G_1(s) \;=\; K\frac{(b_1 - a_1)s^{n-1} + \cdots + (b_{n-1} - a_{n-1})s - a_n}{s(s^n + a_1 s^{n-1} + \cdots + a_n)}$$

that is the system of relative degree 2. Let $\overline{S}$ be the input-output mapping of $G_1(s)$ on $[0, t_f]$ as is done for S. Then $Su = K\int_0^t u(\tau)d\tau + \overline{S}u$ and hence (4.32) leads to

$$\begin{aligned}
&[\delta_\Delta\sigma_\Delta S f_\Delta](k)\\
&= K\left\{\left[\delta_\Delta\sigma_\Delta\int_0^t u^*(\tau)d\tau\right](k)\right.\\
&\quad \left. - \left[\delta_\Delta\sigma_\Delta\int_0^t [\theta_\Delta\sigma_\Delta u^*](\tau)d\tau\right](k)\right\}\\
&\quad + [\delta_\Delta\sigma_\Delta\overline{S}u^*](k) - [\delta_\Delta\sigma_\Delta\overline{S}\theta_\Delta\sigma_\Delta u^*](k)\\
&= K\left\{\frac{1}{\Delta}\int_{k\Delta}^{(k+1)\Delta} u^*(\tau)d\tau - u^*(k\Delta)\right\}\\
&\quad + \left[\frac{d}{dt}\overline{S}u^*\right]((k+\alpha_k^\Delta)\Delta)\\
&\quad - \left[\frac{d}{dt}\overline{S}\theta_\Delta\sigma_\Delta u^*\right]((k+\beta_k^\Delta)\Delta)
\end{aligned}$$

for $k = 0, 1, \cdots, N-1$, where the mean value theorem was applied since we have $\overline{S}u^* \in C^1[0, t_f]$ and $\overline{S}\theta_\Delta\sigma_\Delta u^* \in C^1[0, t_f]$ from $u^* \in C[0, t_f]$ and the relative degree of $G_1(s)$ is 2. Moreover, we obtain

$$\begin{aligned}
&|[\delta_\Delta\sigma_\Delta S f_\Delta](k)|\\
&\leq \left|K\left\{\frac{1}{\Delta}\int_{k\Delta}^{(k+1)\Delta} u^*(\tau)d\tau - u^*(k\Delta)\right\}\right|\\
&\quad + \left|\left[\frac{d}{dt}\overline{S}u^*\right]((k+\alpha_k^\Delta)\Delta) - \left[\frac{d}{dt}\overline{S}u^*\right]((k+\beta_k^\Delta)\Delta)\right|\\
&\quad + \left|\left[\frac{d}{dt}\overline{S}u^*\right]((k+\beta_k^\Delta)\Delta)\right.\\
&\quad \left. - \left[\frac{d}{dt}\overline{S}\theta_\Delta\sigma_\Delta u^*\right]((k+\beta_k^\Delta)\Delta)\right| \qquad (4.33)
\end{aligned}$$

for $k = 0, 1, \cdots, N-1$. On the other hand, from $u^* \in C[0, t_f]$, we have

$$\max\left\{\left|K\left\{\frac{1}{\Delta}\int_{k\Delta}^{(k+1)\Delta} u^*(\tau)d\tau - u^*(k\Delta)\right\}\right|\right.$$

$$\left.; k = 0, \cdots, N-1\right\} \to 0 \qquad (4.34)$$

as $\Delta \to 0$. Note that $\frac{d}{dt}\overline{S}u^* \in C^1[0, t_f]$ because the relative degree of $G_1(s)$ is 2. Then, from $\max\{|(k+\alpha_k^\Delta)\Delta - (k+\beta_k^\Delta)\Delta|; k = 0, 1, \cdots, N-1\} \to 0$ (as $\Delta \to 0$), we have

$$\begin{aligned}
&\max\left\{\left|\left[\frac{d}{dt}\overline{S}u^*\right]((k+\alpha_k^\Delta)\Delta) - \left[\frac{d}{dt}\overline{S}u^*\right]((k+\beta_k^\Delta)\Delta)\right|;\right.\\
&\left. k = 0, 1, \cdots, N-1\right\} \to 0 \qquad (4.35)
\end{aligned}$$

as $\Delta \to 0$. Note that Proposition 2 holds for $\overline{S}$. Then, from $\left\|\frac{d}{dt}Su^* - \frac{d}{dt}S\theta\sigma u^*\right\|_\infty \to 0$ (as $\Delta \to 0$), we have

$$\begin{aligned}
&\max\left\{\left|\left[\frac{d}{dt}\overline{S}u^*\right]((k+\beta_k^\Delta)\Delta)\right.\right.\\
&\quad \left.\left. - \left[\frac{d}{dt}\overline{S}\theta_\Delta\sigma_\Delta u^*\right]((k+\beta_k^\Delta)\Delta)\right|; k = 0, 1, \cdots, N-1\right\}\\
&\to 0 \qquad (4.36)
\end{aligned}$$

as $\Delta \to 0$.

From (4.33), (4.34), (4.35) and (4.36), we establish

$$\max\{|[\delta_\Delta\sigma_\Delta S f_\Delta](k)|; k = 0, 1, \cdots, N-1\} \to 0 \quad (\text{as}\Delta \to 0)$$

that implies

$$\max\{|[\sigma_\Delta f_\Delta](k)|; k = 0, 1, \cdots, N\} \to 0 \quad (\text{as}\Delta \to 0)$$

because of (4.31). This completes the proof. ■

Remark 6 *As stated in Section 4.2, when $G(s)$ has an unstable zero, $H(z)$ have the corresponding unstable one for the sufficiently small Δ. However, Theorem 12 implies that when $\Delta \to 0$ on the finite time interval $[0, t_f]$, divergence doesn't occur even if $H(z)$ has unstable zeros*

4.3.2 The Case of Relative Degree 2

In this section, we discuss the case of relative degree 2. We show the next theorem that is similar to the case of relative degree 0 or 1.

Theorem 13 *Assume that y^* satisfies $y^* = Su^*$ where $u^* \in C^1[0, t_f]$. If the relative degree $n - m = 2$, then*

$$\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*\|'_\infty \to 0 \tag{4.37}$$

as $\Delta \to 0$

Remark 7 *Theorem 13 is independent of stability of zeros of $H(z)$. The convergence (4.37) with respect to the norm $\|\cdot\|'_\infty$ is the best result in this case because of the same reason as stated in Remark 5.*

In the following discussion, it is assumed that $n - m = 2$ and Δ is chosen sufficiently small that $cb_\Delta \neq 0$. (See Lemma 4.10 and Remark 4.11.) From (4.9), we can express $H^{-1}(z)$ as

$$H^{-1}(z) = H_1^{-1}(z) H_2^{-1}(z) H_3^{-1}(z)$$

where

$$H_1^{-1}(z) = \frac{\Delta^2}{cb_\Delta} \frac{(z+1)(z - \exp(p_1\Delta)) \cdots (z - \exp(p_n\Delta))}{(z-1)^2 (z - q_1(\Delta)) \cdots (z - q_{n-1}(\Delta))},$$

$$H_2^{-1}(z) = \frac{1}{z+1}$$

and

$$H_3^{-1}(z) = \frac{(z-1)^2}{\Delta^2}.$$

$\Gamma_\Delta^+(N, N+1)$, $N \times (N+1)$ matrix that is made of Γ_Δ^+ except the $N+1$-th row, is expressed as the multiplication of three matrices which correspond to $H_1^{-1}(z)$, $H_2^{-1}(z)$ and $H_3^{-1}(z)$, respectively.

$$\Gamma_\Delta^+(N, N+1) = \Lambda_{\Delta 1} \Lambda_{\Delta 2} \Lambda_{\Delta 3} \tag{4.38}$$

where $\Lambda_{\Delta 1}$, $\Lambda_{\Delta 2}$ and $\Lambda_{\Delta 3}$ are $N \times N$, $N \times (N-1)$ and $(N-1) \times (N+1)$, respectively. The elements of these matrices are expressed as follows.

$$\Lambda_{\Delta 1} = \begin{bmatrix} \frac{\Delta^2}{cb_\Delta} & & & & 0 \\ \bar{c}_\Delta \bar{b}_\Delta & \frac{\Delta^2}{cb_\Delta} & & & \\ \bar{c}_\Delta \bar{A}_\Delta \bar{b}_\Delta & \bar{c}_\Delta \bar{b}_\Delta & \ddots & & \\ \vdots & \ddots & \ddots & \frac{\Delta^2}{cb_\Delta} & \\ \bar{c}_\Delta \bar{A}_\Delta^{N-2} \bar{b}_\Delta & \cdots & \bar{c}_\Delta \bar{A}_\Delta \bar{b}_\Delta & \bar{c}_\Delta \bar{b}_\Delta & \frac{\Delta^2}{cb_\Delta} \end{bmatrix} \tag{4.39}$$

where $\left(\bar{A}_\Delta, \bar{b}_\Delta, \bar{c}_\Delta, \frac{\Delta^2}{cb_\Delta}\right)$ is the realization of $H_1^{-1}(z)$ in the controllable canonical form.

$$\Lambda_{\Delta 2} = \begin{bmatrix} 0 & & & & & 0 \\ 1 & 0 & & & & \\ -1 & 1 & 0 & & & \\ 1 & -1 & 1 & 0 & & \\ & & \ddots & \ddots & & \\ \vdots & & & & 1 & 0 \\ (-1)^N & \cdots & & 1 & -1 & 1 \end{bmatrix} \tag{4.40}$$

$$\Lambda_{\Delta 3} = \frac{1}{\Delta^2} \begin{bmatrix} 1 & -2 & 1 & & & 0 \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ 0 & & & 1 & -2 & 1 \end{bmatrix} \tag{4.41}$$

For the sake of the proof of Theorem 13, we present the following lemmas about $H_1^{-1}(z)$ and $H_2^{-1}(z)$.

Lemma 6 *Consider a sequence $\{\eta(k); k = 0, 1, \cdots, N-1(= t_f/\Delta - 1)\}$ and let $\{u(k)\}$ be*

$$\begin{aligned} &\begin{bmatrix} u(0) & u(1) & \cdots & u(N-1) \end{bmatrix}^T \\ &= \Lambda_{\Delta 1} \begin{bmatrix} \eta(0) & \eta(1) & \cdots & \eta(N-1) \end{bmatrix}^T \end{aligned}$$

Then, for any $\Delta \in (0, \epsilon)$ and any sequence $\{\eta(k)\}$, we have

$$\begin{aligned} &\max\{|u(k)|; k = 0, 1, \cdots N-1(= t_f/\Delta - 1)\} \\ &\leq M_1 \max\{|\eta(k)|; k = 0, 1, \cdots N-1(= t_f/\Delta - 1)\} \end{aligned} \tag{4.42}$$

where ϵ and M_1 are positive constants.

Proof: See appendix F ■

Lemma 7 *Consider a sequence $\{\zeta(k); k = 0, 1, \cdots N-2\}$ which is defined by one of the following equations.*

1. $\zeta(k) = \Delta^p f_\Delta(t_k)$ *($p \geq 1$) where* $t_k \in [0, t_f]$ *and* $\|f_\Delta\|_\infty \to 0$ *as* $\Delta \to 0$.
2. $\zeta(k) = \Delta^p f(t_k)$ *($p \geq 1$) where* $t_k \in [0, t_f]$, $|t_{k+1} - t_k| \leq M_2\Delta$ *(M_2: a positive constant) and* $f \in C^0[0, t_f]$.
3. $\zeta(k) = f_\Delta(k\Delta) - f(k\Delta)$ *where* $f_\Delta, f \in C^1[0, t_f]$ *and* $\left\|\frac{d}{dt}f_\Delta - \frac{d}{dt}f\right\|_\infty \to 0$ *as* $\Delta \to 0$.

Let $\{\eta(k)\}$ be

$$\begin{aligned} &\begin{bmatrix} \eta(0) & \eta(1) & \cdots & \eta(N-1) \end{bmatrix}^T \\ &= \Lambda_{\Delta 2}\begin{bmatrix} \zeta(0) & \zeta(1) & \cdots & \zeta(N-2) \end{bmatrix}^T \end{aligned}$$

Then

$$\max\{|\eta(k)|; k = 0, 1, \cdots, N-1 (= t_f/\Delta - 1)\} \to 0 \tag{4.43}$$

as $\Delta \to 0$.

Proof: See appendix G ■

Proof of Theorem 13: Since

$$\begin{aligned} &\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*\|'_\infty \\ &\leq \|\theta_\Delta \sigma_\Delta u^* - u^*\|'_\infty \\ &\quad + \|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - \theta_\Delta \sigma_\Delta u^*\|'_\infty \\ &\leq \|\theta_\Delta \sigma_\Delta u^* - u^*\|'_\infty \\ &\quad + \max\{|[\Gamma_\Delta^+ \sigma_\Delta y^* - \sigma_\Delta u^*](k)|; \\ &\quad k = 0, 1, \cdots, N-1\} \end{aligned}$$

and $\lim_{\Delta \to 0} \|\theta_\Delta \sigma_\Delta u^* - u^*\|'_\infty = 0$, what we have to prove is

$$\begin{aligned} &\max\{|[\Gamma_\Delta^+ \sigma_\Delta y^* - \sigma_\Delta u^*](k)|; \\ &\quad k = 0, 1, \cdots, N-1\} \to 0 \end{aligned} \tag{4.44}$$

as $\Delta \to 0$. Moreover, from the definition of Γ_Δ^+, $y^* = Su^*$ and $\Gamma_\Delta = \sigma_\Delta S \theta_\Delta$ we have

$$\begin{aligned} &\Gamma_\Delta^+ \sigma_\Delta y^* - \sigma_\Delta u^* \\ &= \Gamma_\Delta^+ (\sigma_\Delta y^* - \Gamma_\Delta \sigma_\Delta u^*) \\ &= \Gamma_\Delta^+ (\sigma_\Delta S u^* - \sigma_\Delta S \theta_\Delta \sigma_\Delta u^*) \end{aligned}$$

This equality with (4.38) implies that what we have to prove is

$$\begin{aligned} &\max\{|[\Lambda_{\Delta 1}\Lambda_{\Delta 2}\Lambda_{\Delta 3}(\sigma_\Delta S u^* - \sigma_\Delta S \theta_\Delta \sigma_\Delta u^*)](k)|; \\ &\quad k = 0, 1, \cdots, N-1\} \to 0 \end{aligned} \tag{4.45}$$

First, we discuss $\Lambda_{\Delta 3}(\sigma_\Delta S u^* - \sigma_\Delta S \theta_\Delta \sigma_\Delta u^*)$. Since the relative degree of $G(s)$ is 2, we have

$$G(s) = K\frac{1}{s^2} + L\frac{1}{s^3} + \frac{b_1 s^{n-1} + \cdots + b_n}{s^3(s^n + a_1 s^{n-1} + \cdots + a_n)} \tag{4.46}$$

and hence

$$y = Su = S_1 u + S_2 u + S_3 u$$

where

$$\begin{aligned} S_1 u &= K\int_0^t \int_0^\tau u(\sigma) d\sigma d\tau \\ S_2 u &= L\int_0^t \int_0^\tau \int_0^\sigma u(x) dx d\sigma d\tau, \\ S_3 u &= \int_0^t \hat{c}\exp(\hat{A}(t-\tau))\hat{b}u(\tau)d\tau, \end{aligned}$$

and $(\hat{A}, \hat{b}, \hat{c})$ is realization of the third term of (4.46). Note that

$$\begin{aligned} &\Lambda_{\Delta 3}(\sigma_\Delta S u^* - \sigma_\Delta S \theta_\Delta \sigma_\Delta u^*) \\ &= \Lambda_{\Delta 3}\sigma_\Delta S_1 u^* + \Lambda_{\Delta 3}\sigma_\Delta S_2 u^* + \Lambda_{\Delta 3}\sigma_\Delta S_3 u^* - \\ &\quad (\Lambda_{\Delta 3}\sigma_\Delta S_1 \theta_\Delta \sigma_\Delta u^* + \Lambda_{\Delta 3}\sigma_\Delta S_2 \theta_\Delta \sigma_\Delta u^* \\ &\quad + \Lambda_{\Delta 3}\sigma_\Delta S_3 \theta_\Delta \sigma_\Delta u^*) \end{aligned} \tag{4.47}$$

and

$$[\Lambda_{\Delta 3}Y](k) = \frac{y(k+2) - 2y(k+1) + y(k)}{\Delta^2}$$

$(k = 0, \cdots, N-2)$ where $Y = [y(0) \cdots y(N)]^T$. Then we obtain the following expression of each term of (4.47) by the mean value theorem.

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_1 u^*](k) \\ &= K\left\{u^*(k\Delta) + \Delta\frac{4}{3}\frac{d}{dt}u^*((k+2\phi_1(k,\Delta))\Delta)\right. \\ &\quad \left. -\Delta\frac{1}{3}\frac{d}{dt}u^*((k+2\phi_2(k,\Delta))\Delta)\right\} \end{aligned} \tag{4.48}$$

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_2 u^*](k) \\ &= L\left\{\int_0^{k\Delta} u^*(\tau)d\tau + \Delta\frac{4}{3}u^*((k+2\phi_3(k,\Delta))\Delta)\right. \\ &\quad \left. -\Delta\frac{1}{3}u^*((k+2\phi_4(k,\Delta))\Delta)\right\} \end{aligned} \tag{4.49}$$

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_3 u^*](k) \\ &= K\left\{\frac{d^2}{dt^2}S_3 u^*(k\Delta) + \Delta\frac{4}{3}\frac{d^3}{dt^3}S_3 u^*((k+2\phi_5(k,\Delta))\Delta)\right. \\ &\quad \left. -\Delta\frac{1}{3}\frac{d^3}{dt^3}S_3 u^*((k+2\phi_6(k,\Delta))\Delta)\right\} \end{aligned} \tag{4.50}$$

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_1\theta_\Delta\sigma_\Delta u^*](k) \\ &= K\left(u^*(k\Delta) + \frac{\Delta}{2}\frac{d}{dt}u^*((k+\phi_7(k,\Delta))\Delta)\right) \end{aligned} \tag{4.51}$$

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_2\theta_\Delta\sigma_\Delta u^*](k) \\ &= L\left(\int_0^{k\Delta}[\theta_\Delta\sigma_\Delta u^*]d\tau\right. \\ &\quad \left. +\Delta u^*(k\Delta) + \frac{\Delta^2}{6}\frac{d}{dt}u^*((k+\phi_8(k,\Delta))\Delta)\right) \end{aligned} \tag{4.52}$$

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_3\theta_\Delta\sigma_\Delta u^*](k) \\ &= \frac{d^2}{dt^2}[S_3\theta_\Delta\sigma_\Delta u^*](k\Delta) \\ &\quad +\Delta\frac{4}{3}\frac{d^3}{dt^3}[S_3\theta_\Delta\sigma_\Delta u^*]((k+2\phi_9(k,\Delta))\Delta) \\ &\quad -\Delta\frac{1}{3}\frac{d^3}{dt^3}[S_3\theta_\Delta\sigma_\Delta u^*]((k+2\phi_{10}(k,\Delta))\Delta) \end{aligned} \tag{4.53}$$

where $0 \le \phi_i(k,\Delta) \le 1$ $(i = 1, 2, \cdots, 10)$. Moreover, (4.53) leads to

$$\begin{aligned} &\Lambda_{\Delta 3}\sigma_\Delta[S_3\theta_\Delta\sigma_\Delta u^*](k) \\ &= \frac{d^2}{dt^2}[S_3\theta_\Delta\sigma_\Delta u^*](k\Delta) \\ &\quad +\Delta\frac{4}{3}\left\{\frac{d^3}{dt^3}[S_3\theta_\Delta\sigma_\Delta u^*]((k+2\phi_9(k,\Delta))\Delta)\right. \\ &\quad \left. -\frac{d^3}{dt^3}[S_3 u^*]((k+2\phi_9(k,\Delta))\Delta)\right\} \\ &\quad -\Delta\frac{1}{3}\left\{\frac{d^3}{dt^3}[S_3\theta_\Delta\sigma_\Delta u^*]((k+2\phi_{10}(k,\Delta))\Delta)\right. \\ &\quad \left. -\frac{d^3}{dt^3}[S_3 u^*]((k+2\phi_{10}(k,\Delta))\Delta)\right\} \\ &\quad +\Delta\frac{4}{3}\frac{d^3}{dt^3}[S_3 u^*]((k+2\phi_9(k,\Delta))\Delta) \\ &\quad -\Delta\frac{1}{3}\frac{d^3}{dt^3}[S_3 u^*]((k+2\phi_{10}(k,\Delta))\Delta) \end{aligned} \tag{4.54}$$

Therefore, from Proposition 2 , we can see

$$[\Lambda_{\Delta 3}(\sigma_\Delta S u^* - \sigma_\Delta S\theta_\Delta\sigma_\Delta u^*)](k) \tag{4.55}$$

$(k = 0, 1, \cdots, N-2)$ is the linear combination of 1., 2. or 3. of Lemma 7. By Lemma 7, we have

$$\max\{|[\Lambda_{\Delta 2}\Lambda_{\Delta 3}(\sigma_\Delta S u^* - \sigma_\Delta S\theta_\Delta\sigma_\Delta u^*)](k)|; \\ k = 0, 1, \cdots, N-1\} \to 0$$

furthermore, by Lemma 6, we obtain (4.45). This completes the proof. ■

4.4 Implication of the Result and Numerical Examples

If the desired trajectory $y^*(t)$ $(t \in [0, t_f])$ is defined as in Theorem 12 or Theorem 13, one can obtain u^* satisfying $y^* = Su^*$ by using the following inverse system of (4.1).

$$\begin{aligned}\frac{d}{dt}x(t) &= (A_c - d^{-1}c)x(t) + d^{-1}b_c y^*(t)\\ u^*(t) &= -d^{-1}cx(t) + d^{-1}y^*(t)\end{aligned} \tag{4.56}$$

when $d \neq 0$;

$$\begin{aligned}\frac{d}{dt}x(t) &= \{A_c + b_c F^{-1}cA_c\}x(t) + b_c F^{-1}\left(\frac{d}{dt}\right)^{n-m} y^*(t)\\ u^*(t) &= F^{-1}cA_c^{n-m}x(t) + F^{-1}\left(\frac{d}{dt}\right)^{n-m} y^*(t)\end{aligned} \tag{4.57}$$

when $d = 0$ where $F = cA_c^{n-m-1}b_c$ and $x(0) = 0$. Let S^{-1} be the input-output mapping of (4.56) or (4.57) on $[0, t_f]$. Then the conclusion of Theorem 12 or Theorem 13 is

$$\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - S^{-1}y^*\|_\infty \to 0 \tag{4.58}$$

or

$$\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - S^{-1}y^*\|'_\infty \to 0 \tag{4.59}$$

These convergences imply that the inverse system of the sampled-data system approximates the inverse system of the continuous-time system provided that $n - m = 0$, 1 or 2.

As stated in the section 4.2, one of the most common problem formulation of iterative learning control is the minimization problem (4.13). Theorem 12 or Theorem 13 supports this problem formulation because (4.58) or (4.59) guarantees that, by shrinking the sampling period Δ, we can make the minimizer $\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ of (4.13) arbitrarily close to the ideal input $S^{-1}y^*$. It should be noted that the ripple between sample points is reduced independently of stability of the zeros because

$$\|S\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - y^*\|_\infty \leq \|S\|\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - S^{-1}y^*\|_\infty$$

or

$$\|S\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - y^*\|'_\infty \leq \|S\|\|\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - S^{-1}y^*\|'_\infty$$

Numerical examples are presented to illustrate the results of Theorem 12 or Theorem 13.

Example 7 *Consider a system with relative degree 1,*

$$S(s) = \frac{s-1}{(s+5)(s+6)}$$

on $[0, t_f]$. Then the pulse transfer function is

$$H(z) = \frac{K(z - \frac{35e^{-5\Delta} - 36e^{-6\Delta} + e^{-11\Delta}}{36e^{-5\Delta} - 35e^{-6\Delta} - 1})}{30(z - e^{-5\Delta})(z - e^{-6\Delta})}$$

that has an unstable zero for small Δ where

$$K = 36e^{-5\Delta} - 35e^{-6\Delta} - 1.$$

Let a desired trajectory be

$$y^*(t) = -t^2(2t - 3)$$

Then we can see that a u^* satisfying $y^* = Su^*$ is defined by the inverse system

$$\begin{aligned}\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\frac{d}{dt}y^*\\ u^* &= \begin{bmatrix} 30 & 12 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \frac{d}{dt}y^*\end{aligned}$$

It is easily checked that there exists $u^*(t) \in C[0, t_f]$ which satisfies the assumption of Theorem 12. In Figure 4.1 (a), $u = \theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ is plotted for $\Delta = 0.2$ and 0.05; the dashed line refers to u^*. In Figure 4.1 (b), $e = S(\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*)$ is plotted for $\Delta = 0.2$ and 0.05. We can see that, as the sampling period Δ goes to 0, $\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ approaches u^* and hence the residual $S\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - y^*$ is reduced.

Example 8 *Consider a system with relative degree 2,*

$$G(s) = \frac{s-1}{s^3}$$

Then

$$H(z) = \frac{(3-\Delta)\Delta^2}{6}\frac{(z-q_4(\Delta))(z-q_5(\Delta)}{(z-1)^3}$$

where $q_4(\Delta) = 1+\Delta+O(\Delta^2)$ *and* $q_5(\Delta) = -1+\frac{\Delta}{3}+O(\Delta^2)$. *Let* u^* *be*

$$u^*(t) = t+1 \quad t \in [0, 10] \tag{4.60}$$

and $y^* = Su^*$. *Then the assumption of Theorem 13 is satisfied. In Figure 4.2 (a),* $u = \theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^*$ *is plotted for* $\Delta = 1$ *and* 0.5*; the dashed line refers to* u^*. *In Figure 4.2 (b),* $e = S(\theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^* - u^*)$ *is plotted for* $\Delta = 1$ *and* 0.5. *We can see effect of the zero* $q_1(\Delta)$ *is reduced even though* $q_1(\Delta)$ *approaches to* 1 *from inside the unstable area.*

Example 9 *Consider a system with relative degree 2,*

$$G(s) = \frac{2}{2s^2-3s+1} \tag{4.61}$$

Then

$$H(z) = 2(e^{\Delta/2}-1)^2\frac{z-q_1(\Delta)}{(z-e^{\Delta})(z-e^{\Delta/2})} \tag{4.62}$$

where $q_1(\Delta) = -1-\frac{\Delta}{2}+O(\Delta^2)$. *The function* u^* *is chosen as (4.60) and* $y^* = Su^*$. *In Figure 4.3,* $u = \theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^*$ *for* $\Delta = 1$ *and* 0.25 *and* $e = S(\theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^* - u^*)$ *for* $\Delta = 1$ *and* 0.25 *are plotted, respectively.*

Example 10 *Consider a system with relative degree 3,*

$$G(s) = \frac{1}{(s+1)^3} \tag{4.63}$$

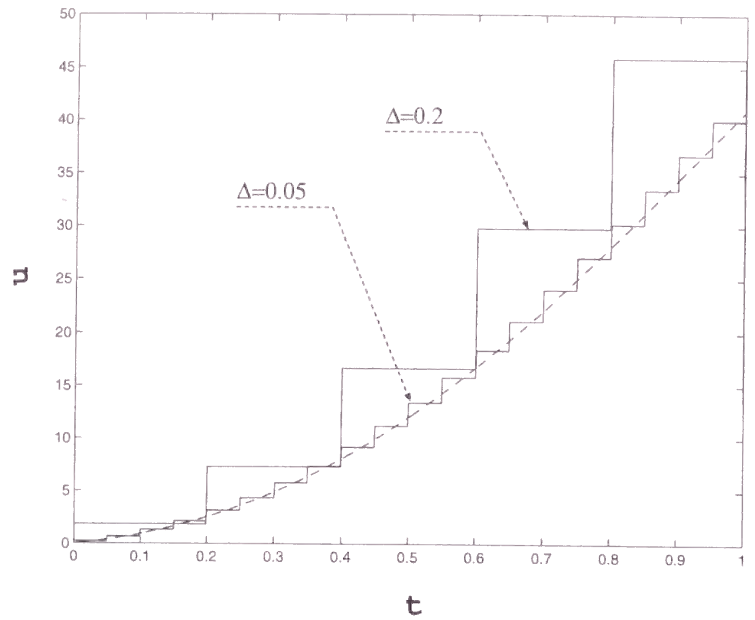
Then

$$H(z) = \frac{2-(\Delta^2+2\Delta+2)e^{-\Delta}}{2} \frac{(z-q_2(\Delta))(z-q_3(\Delta))}{(z-e^{-\Delta})^3} \tag{4.64}$$

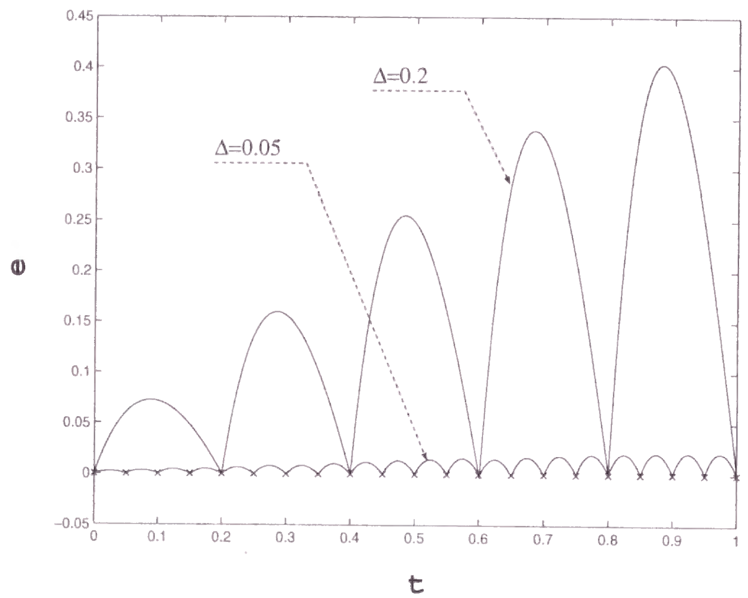
where $q_2(\Delta) \to -2-\sqrt{3}$ *and* $q_3(\Delta) \to -2+\sqrt{3}$ *as* $\Delta \to 0$. *The function* u^* *is chosen as (4.60) and* $y^* = Su^*$. *In Figure 4.4,* $u = \theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^*$ *for* $\Delta = 2.5$ *and* 1 *and* $e = S(\theta_\Delta\Gamma_\Delta^+\sigma_\Delta y^* - u^*)$ *for* $\Delta = 2.5$ *and* 1 *are plotted, respectively. We can see the norms of these functions increase as* Δ *goes to* 0 *This is the effect of the zero* $q_2(\Delta)$ *which converges to the unstable* $z = -2-\sqrt{3}$.

4.5 Concluding Remarks

In this chapter, we discussed implementation of iterative learning control with a 0-order hold, a sampler and digital computers. It was proved that if relative degree of transfer function of the continuous-time system is 0, 1 or 2 then it is not necessary to consider stability of zeros of the sampled-data system for any small sampling period when sampled-data systems are considered on the fixed continuous time interval; this property holds even if the transfer function of the continuous time system has unstable zeros. It was guaranteed that one can simply define problems of iterative learning control for sampled-data systems as minimization problems of output errors on the sample points. We also presented some numerical examples to illustrate the main result and demonstrated that the result is uncommon for systems with general relative degree of the transfer functions. This implies that when we consider iterative learning control for sampled-data systems and the relative degree of the original continuous-time system is greater than 2, we must pay much attention to unstable zeros or length of the sampling period in order to avoid undesirable effect on the inter sample points.

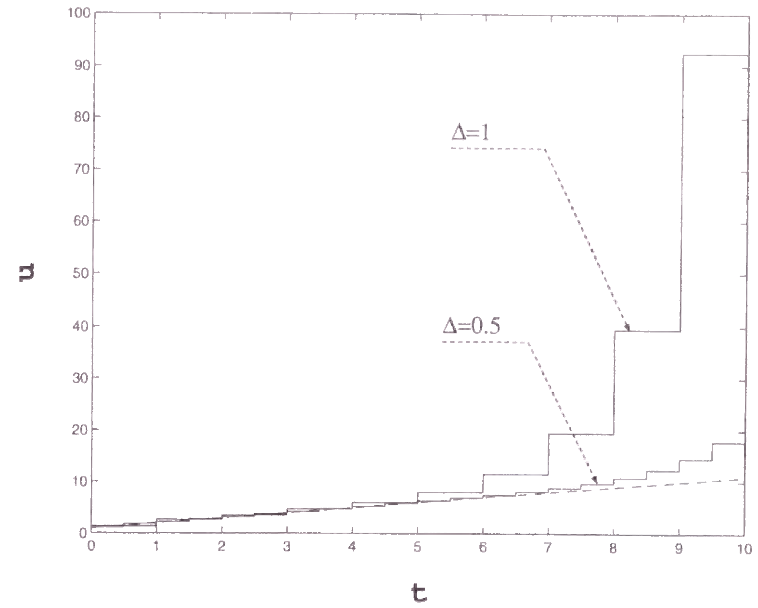


(a) $\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ and u^*

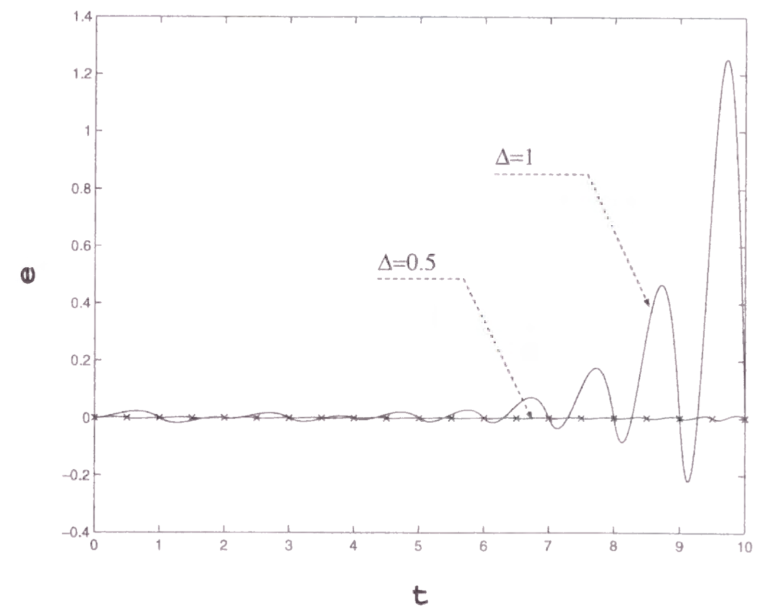


(b) $S(\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*)$

Figure 4.1: Simulation Results of Example 7.

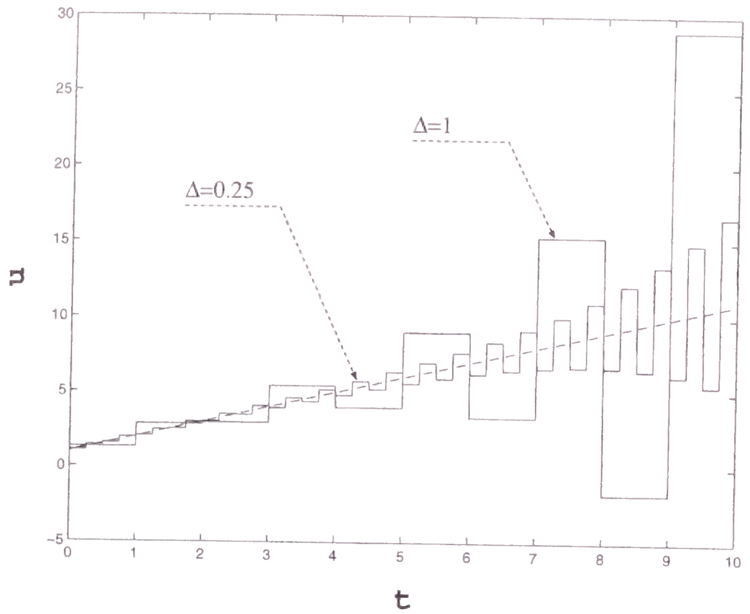


(a) $\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ and u^*

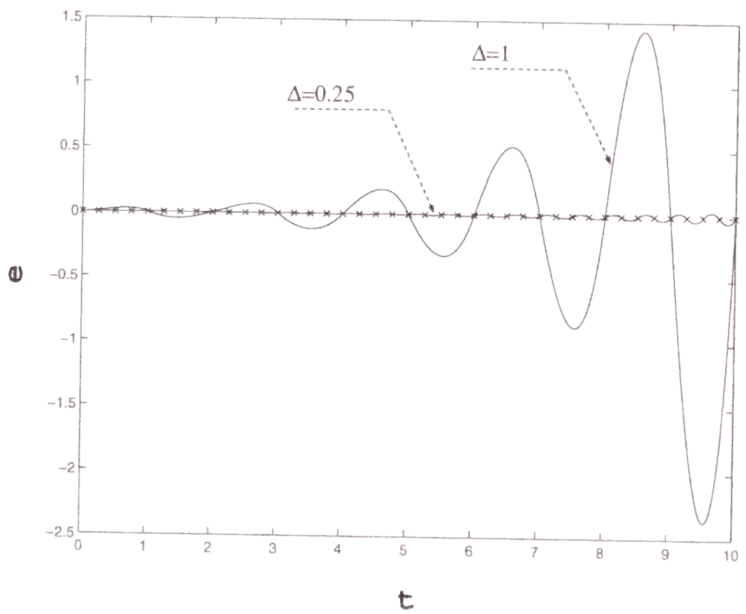


(b) $S(\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*)$

Figure 4.2: Simulation Results of Example 8.

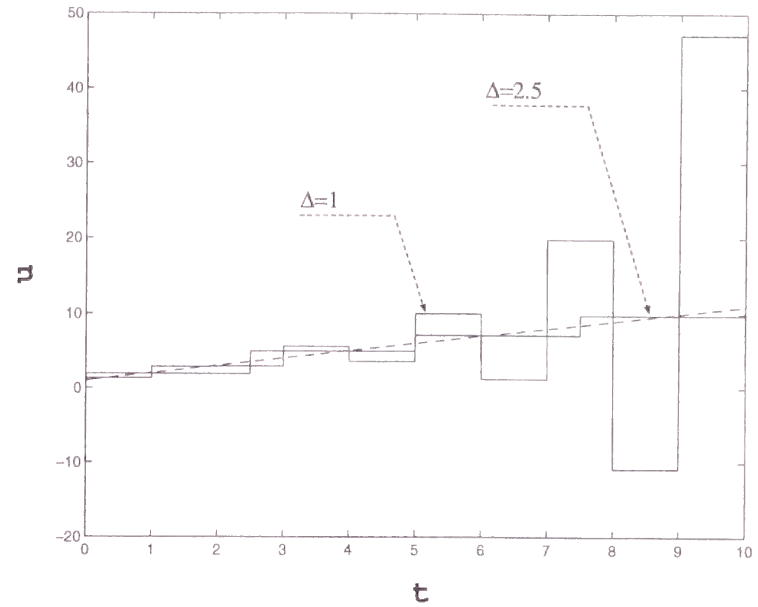


(a) $\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ and u^*

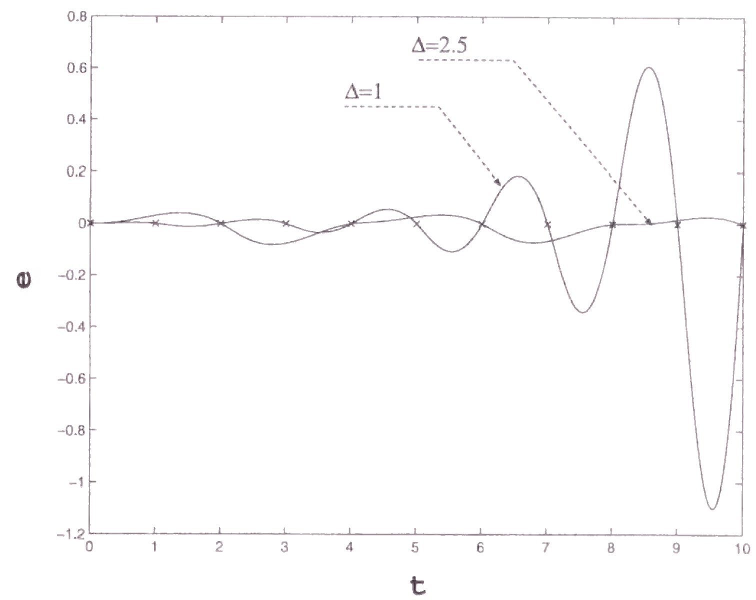


(b) $S(\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*)$

Figure 4.3: Simulation Results of Example 9.



(a) $\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^*$ and u^*



(b) $S(\theta_\Delta \Gamma_\Delta^+ \sigma_\Delta y^* - u^*)$

Figure 4.4: Simulation Results of Example 10.

Chapter 5

Iterative Learning Control for Linear Discrete-Time Systems

5.1 Introduction

In the last chapter, we discussed iterative learning control for sampled-data systems. It was shown that one can define iterative learning control as an iterative algorithm of minimizing output errors on the sample points when relative degree of the transfer function of the continuous-time system is 0, 1 or 2. Even if there are unstable zeros of the transfer function of the sampled-data system, effect of the unstable zeros is reduced by shrinking the sample period; the ripple on the inter-sample points is small when the sample period is sufficiently small. In this chapter, we develop iterative learning control for linear discrete-time systems which refer to sampled-data systems with the limiting property stated above.

Many researchers studied iterative learning control for linear discrete-time systems and proposed the design methods [50, 51, 52, 53, 54, 26, 55, 25, 56]. However, no specific design method of iterative learning control using adjoint systems was presented for discrete-time systems with uncertainty. On the other hand, we demonstrated advantage of the iterative learning control using adjoint systems for linear continuous-time systems. In the following sections, we discuss this kind of iterative learning control for linear discrete-time systems. First, we formulate iterative learning control for linear discrete-time systems as an iterative algorithm of minimizing the output error in the vector space. Second, we present iterative learning control using adjoint systems and demonstrate its convergence. The convergence condition is given as strictly positive realness of the transfer function that represents uncertainty of the system. Moreover, we present a convergence condition when the system has structured uncertainty, i.e. when parameters of the system is given as intervals. Finally, numerical examples is presented to illustrate the results.

5.2 Mathematical Preliminaries

Consider a linear discrete-time system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \quad (k = 0, 1, \cdots) \end{aligned} \tag{5.1}$$

with $x(0) = 0$ where $x \in R^n$, $u \in R^m$ and $y \in R^p$. Then the input-output mapping of (5.1) for times $k = 0, 1, \cdots, N-1$ is expressed as

$$w = \Gamma v$$

where

$$v = \begin{bmatrix} u(0)^T & u(1)^T & \cdots & u(N-1)^T \end{bmatrix}^T$$

$$w = \begin{bmatrix} y(0)^T & y(1)^T & \cdots & y(N-1)^T \end{bmatrix}^T$$

$$\Gamma = \begin{bmatrix} D & 0 & \cdots & & 0 \\ CB & D & \ddots & & \vdots \\ CAB & CB & \ddots & & \\ \vdots & \ddots & \ddots & D & 0 \\ CA^{N-2}B & \cdots & CAB & CB & D \end{bmatrix}$$

Let a desired trajectory at $k = 0, 1, \cdots, N-1$ be

$$w_d = \begin{bmatrix} y_d(0)^T & y_d(1)^T & \cdots & y_d(N-1)^T \end{bmatrix}^T \tag{5.2}$$

then the best input v^* is the solution of the equation $\Gamma v = w_d$. More generally, v^* is defined as the least-squares solution with the minimum norm, namely

$$v^* = \Gamma^\dagger w_d$$

where $\Gamma^\dagger$ indicates the Moore-Penrose generalized inverse matrix. If Γ is expressed as

$$\Gamma = FG$$

where F and G are matrices with the maximum rank, then

$$\Gamma^+ = G^T(GG^T)^{-1}(F^TF)^{-1}F^T$$

In the following discussion, we consider

$$\begin{aligned} v_{n+1} &= v_n - \Phi e_n \quad (n = 0, 1, \cdots) \\ e_n &= \Gamma v_n - w_d + d_n \end{aligned} \tag{5.3}$$

as algorithms that update the input v where d_n represents measurement error or disturbance; it is assumed that

$$|d_n| \le M$$

where M is a positive constant. We discuss a problem to determine Φ so that the input sequence $\{v_n; n = 0, 1, \cdots\}$ generated by (5.3) satisfies

$$v_n \to \Gamma^+ w_d \tag{5.4}$$

as $n \to \infty$ and $M \to 0$. If Γ^T is available and Φ is chosen as

$$\Phi = \alpha \Gamma^T \tag{5.5}$$

where α is a sufficiently small positive constant, then (5.3) is coincident with the gradient method. The main topic of iterative learning control is how to choose Φ when there is uncertainty in the system (5.1) and hence Γ^T is unavailable.

In the following discussions, $R(\Gamma)$ and $N(\Gamma)$ indicate the range and the null space of Γ, respectively.

5.3 Convergence of Iterative Learning Control

Assume that the parameters of the system (5.1) have uncertainty and let the nominal model be

$$x(k+1) = \hat{A}x(k) + \hat{B}u(k)$$
$$y(k) = \hat{C}x(k) + \hat{D}u(k) \tag{5.6}$$

Then the counterpart of Γ for (5.6) is

$$\hat{\Gamma} = \begin{bmatrix} \hat{D} & 0 & \cdots & & 0 \\ \hat{C}\hat{B} & \hat{D} & \ddots & & \vdots \\ \hat{C}\hat{A}\hat{B} & \hat{C}\hat{B} & \ddots & & \\ \vdots & \ddots & \ddots & \hat{D} & 0 \\ \hat{C}\hat{A}^{N-2}\hat{B} & \cdots & \hat{C}\hat{A}\hat{B} & \hat{C}\hat{B} & \hat{D} \end{bmatrix}.$$

In the following discussion, we consider the algorithm (5.3) with

$$\Phi = \alpha\hat{\Gamma}^T$$

instead of unavailable Γ^T.

Suppose

$$\hat{\Gamma} = \Lambda\Gamma \tag{5.7}$$

where $\Lambda \in R^{pN \times pN}$. Then we have the next theorem.

Theorem 14 *If*

$$(\Lambda w)^T w \geq \mu|w|^2 \quad (\mu\text{: a positive constant}) \tag{5.8}$$

for any $w \in R^{pN}$ *and*

$$w_d \in R(\Gamma) \tag{5.9}$$

then the sequence $\{v_n; n = 0, 1, \cdots\}$ *defined by*

$$v_0 \in N(\Gamma)^{\perp} \tag{5.10}$$

$$v_{n+1} = v_n - \alpha\hat{\Gamma}^T e_n \tag{5.11}$$
$$e_n = \Gamma v_n - w_d + d_n \tag{5.12}$$

$$0 < \alpha < \frac{2\mu}{|\hat{\Gamma}^T|^2} \tag{5.13}$$

satisfies

$$|v_n - \Gamma^+ w_d|$$
$$\leq r^n|v_0 - \Gamma^+ w_d| + \frac{1}{1-r}|\alpha\hat{\Gamma}^T|M \tag{5.14}$$

where $|d_n| \leq M$ *and* r *is a positive constant less than 1;* $|\cdot|$ *indicates the Euclidean norm or the induced norm.*

Proof: Since (5.9) implies

$$\Gamma\Gamma^+ w_d = w_d$$

(5.11) with (5.12) and (5.7) leads to

$$v_{n+1} - \Gamma^+ w_d$$
$$= v_n - \Gamma^+ w_d - \alpha\Gamma^T\Lambda^T\Gamma(v_n - \Gamma^+ w_d)$$
$$+\alpha\Gamma^T\Lambda^T d_n \tag{5.15}$$

From this equation, we have

$$|v_{n+1} - \Gamma^+ w_d|$$
$$\leq |v_n - \Gamma^+ w_d - \alpha\Gamma^T\Lambda^T\Gamma(v_n - \Gamma^+ w_d)|$$
$$+|\alpha\hat{\Gamma}^T|M \tag{5.16}$$

moreover, from (5.8),

$$|v_n - \Gamma^+ w_d - \alpha\Gamma^T\Lambda^T\Gamma(v_n - \Gamma^+ w_d)|^2$$
$$\leq |v_n - \Gamma^+ w_d|^2$$
$$-\alpha(2\mu - \alpha|\hat{\Gamma}^T|^2)|\Gamma(v_n - \Gamma^+ w_d)|^2 \tag{5.17}$$

Since $\Gamma^+ w_d \in N(\Gamma)^\perp$ and $R(\Gamma^T) = N(\Gamma)^\perp$, (5.15) with (5.10) implies that

$$v_n \in N(\Gamma)^\perp \quad (n = 0, 1, \cdots)$$

On the other hand, since $\Gamma | N(\Gamma)^\perp : N(\Gamma)^\perp \to R(\Gamma)$ is bijection, there exists a positive constant

$$\sigma = \min \left\{ \frac{|\Gamma v|}{|v|}; v \in N(\Gamma)^\perp, v \neq 0 \right\} \tag{5.18}$$

and hence (5.17) leads to

$$\begin{aligned} & |v_n - \Gamma^+ w_d - \alpha \Gamma^T \Lambda^T \Gamma (v_n - \Gamma^+ w_d)|^2 \\ \leq \; & r^2 |v_n - \Gamma^+ w_d|^2 \end{aligned} \tag{5.19}$$

where

$$r^2 = 1 - \alpha(2\mu - \alpha|\hat{\Gamma}^T|^2)\sigma^2$$

which satisfies $0 \leq r^2 < 1$ because (5.13) implies

$$0 \; < \; \alpha(2\mu - \alpha|\hat{\Gamma}^T|^2)\sigma^2$$

Therefore, from (5.16) and (5.19), we obtain

$$\begin{aligned} & |v_n - \Gamma^+ w_d| \\ \leq \; & r|v_{n-1} - \Gamma^+ w_d| + |\alpha \hat{\Gamma}^T| M \\ \leq \; & r^n |v_0 - \Gamma^+ w_d| + \frac{1 - r^n}{1 - r} |\alpha \hat{\Gamma}^T| M \end{aligned}$$

(5.14) follows this inequality. ■

Remark 8 *The conclusion of Theorem 14 means that if the iteration $n \to \infty$ and the noise level $M \to 0$ then the input v_n tends to the ideal input $\Gamma^+ w_d$.*

Remark 9 *Let*

$$e_n = \left[\; \epsilon(0)^T \quad \epsilon(1)^T \quad \cdots \quad \epsilon(N-1)^T \; \right]^T$$

and

$$\hat{\Gamma}^T e_n = \left[\; \eta(0)^T \quad \eta(1)^T \quad \cdots \quad \eta(N-1)^T \; \right]^T$$

then $\epsilon(k)$ and $\eta(k)$ are related by the adjoint system

$$\begin{aligned} p(k-1) &= \hat{A}^T p(k) + \hat{C}^T \epsilon(k) \\ \eta(k) &= \hat{B}^T p(k) + \hat{D}^T \epsilon(k) \end{aligned} \tag{5.20}$$

with $p(N-1) = 0$. $\hat{\Gamma}^T e_n$ is easily calculated by using (5.20). The condition (5.10) is satisfied by $v_0 = 0$.

The next theorem show that adding a condition about $\hat{\Gamma}$ eliminates the restriction (5.9) on the desired trajectory w_d.

Theorem 15 *Suppose that $\hat{\Gamma}$ is expressed as*

$$\hat{\Gamma} = \Lambda_1 \Gamma \tag{5.21}$$

and

$$\hat{\Gamma} = \Gamma \Lambda_2 \tag{5.22}$$

where Λ_1 and Λ_2 satisfy

$$(\Lambda_1 w)^T w \geq \mu_1 |w|^2 \quad (\mu_1\text{: a positive constant}) \tag{5.23}$$

for any $w \in R^{pN}$ and

$$(\Lambda_2 v)^T v \geq \mu_2 |v|^2 \quad (\mu_2\text{: a positive constant}) \tag{5.24}$$

for any $v \in R^{mN}$, respectively. Then the sequence $\{v_n; n = 0, 1, \cdots\}$ defined by

$$v_0 \in N(\Gamma)^\perp \tag{5.25}$$

$$v_{n+1} = v_n - \alpha \hat{\Gamma}^T e_n \tag{5.26}$$

$$e_n = \Gamma v_n - w_d + d_n \tag{5.27}$$

$$0 < \alpha < \frac{2\mu_1}{|\hat{\Gamma}^T|^2} \tag{5.28}$$

satisfies

$$\begin{aligned} &|v_n - \Gamma^+ w_d| \\ &\leq \; r^n |v_0 - \Gamma^+ w_d| + \frac{1}{1-r} |\alpha \hat{\Gamma}^T| M \end{aligned} \tag{5.29}$$

where $|d_n| \leq M$ *and* r *is a positive constant less than* 1.

Proof: From (5.26) and (5.27), we have

$$\begin{aligned} v_{n+1} &= v_n - \alpha \hat{\Gamma}^T (\Gamma v_n - P w_d) \\ &\quad + \alpha \hat{\Gamma}^T (I - P) w_d + \alpha \hat{\Gamma}^T d_n \end{aligned}$$

where $P : R^{pN} \to R(\Gamma)$ is the orthogonal projection.

Since (5.24) implies $R(\Lambda_2)^\perp = \{0\}$ and $|\Lambda_2 v| \geq \mu_2 |v|$ because of the Schwarz inequality, Λ_2 is bijective and hence there exist Λ_2^{-1}. Since (5.22) leads to $\hat{\Gamma}^T = \Lambda_2^T \Gamma^T$ and $\Lambda_2^{-T} \hat{\Gamma}^T = \Gamma^T$,

$$N(\hat{\Gamma}^T) = N(\Gamma^T)$$

moreover

$$N(\hat{\Gamma}^T) = N(\Gamma^T) = R(\Gamma)^\perp \tag{5.30}$$

By making use of this relationship in (5.30), we obtain

$$v_{n+1} = v_n - \alpha \hat{\Gamma}^T (\Gamma v_n - P w_d) + \alpha \hat{\Gamma}^T d_n$$

furthermore, by $P w_d = \Gamma \Gamma^+ w_d$,

$$\begin{aligned} &v_{n+1} - \Gamma^+ w_d \\ &= \; v_n - \Gamma^+ w_d - \alpha \Gamma^T \Lambda^T \Gamma (v_n - \Gamma^+ w_d) \\ &\quad + \alpha \Gamma^T \Lambda^T d_n \end{aligned}$$

From this equation, (5.29) is established by reasoning which is similar to Theorem 14. ■

Assume that the systems (5.1) and (5.6) are single-input single-output and let the transfer functions be

$$H(z) = \frac{b_0 z^m + b_1 z^{m-1} + \cdots + b_m}{z^n + a_1 z^{n-1} + \cdots + a_n} \tag{5.31}$$

$$\hat{H}(z) = \frac{\hat{b}_0 z^{\hat{m}} + \hat{b}_1 z^{\hat{m}-1} + \cdots + \hat{b}_{\hat{m}}}{z^{\hat{n}} + \hat{a}_1 z^{\hat{n}-1} + \cdots + \hat{a}_{\hat{n}}} \tag{5.32}$$

where

$$m + \hat{n} \geq \hat{m} + n$$

Then

$$J(z) := \hat{H}(z) H(z)^{-1} \tag{5.33}$$

satisfies $\hat{H}(z) = J(z) H(z)$ and hence the matrix Λ in (5.7) is

$$\Lambda = \begin{bmatrix} \tilde{d} & 0 & \cdots & & 0 \\ \tilde{c}\tilde{b} & \tilde{d} & \ddots & & \vdots \\ \tilde{c}\tilde{A}\tilde{b} & \tilde{c}\tilde{b} & \ddots & & \\ \vdots & \ddots & \ddots & \tilde{d} & 0 \\ \tilde{c}\tilde{A}^{N-2}\tilde{b} & \cdots & \tilde{c}\tilde{A}\tilde{b} & \tilde{c}\tilde{b} & \tilde{d} \end{bmatrix} \tag{5.34}$$

where $(\tilde{A}, \tilde{b}, \tilde{c}, \tilde{d})$ is a realization of $J(z)$. In this case, $\Gamma\Lambda = \Lambda\Gamma$ because

$$[\Gamma\Lambda]_{ij} = \sum_{k=1}^{N} \gamma_{i-k} \lambda_{k-j} = \sum_{k=0}^{i-j} \gamma_{i-j-k} \lambda_k$$

and

$$[\Lambda\Gamma]_{ij} = \sum_{k=1}^{N} \lambda_{i-k} \gamma_{k-j} = \sum_{k=0}^{i-j} \lambda_{i-j-k} \gamma_k$$

where

$$\gamma_{i-j} = \begin{cases} 0 & i < j \\ D & i = j \\ CA^{i-j+1}B & i > j \end{cases}$$

and

$$\lambda_{i-j} = \begin{cases} 0 & i < j \\ \tilde{d} & i = j \\ \tilde{c}\tilde{A}^{i-j+1}\tilde{b} & i > j \end{cases}$$

Therefore, in Theorem 15, the condition (5.23) is equivalent to the condition (5.24).

The next lemma replace those conditions with a property of the transfer function $J(z)$.

Lemma 8 *If the transfer function $J(z) = \hat{H}(z)H(z)^{-1}$ is strictly positive real, equivalently,*

1. *(Schur stability)*
   $J(z)$ and $J(z)^{-1}$ has no pole outside or on the unit circle.

2. *(positiveness)*

$$Re[J(e^{j\omega})] > 0 \tag{5.35}$$

for any $\omega \in R$

then Λ satisfies

$$(\Lambda w)^T w \geq \mu |w|^2$$

for any $w \in R^N$ where μ is a positive constant.

Proof: See Appendix H. ■

Even if there is uncertainty in $H(z)$, choosing $\hat{H}(z)$ so that $J(z)$ is strictly positive real makes it possible to guarantee the convergence of the algorithm given in Theorem 14 or 15.

5.4 Convergence of the Iterative Learning Control for Systems with Uncertain Parameters

In section 5.3, convergence conditions for iterative learning control are expressed by a strictly-positive-real (SPR) condition on the transfer function

$J(z) = \hat{H}(z)H(z)^{-1}$. In this section, moreover, the SPR condition will be replaced with inequalities in the parameters of $H(z)$ for practical applications. Suppose that the orders m and n in $H(z)$ are known. Let $H(z)$ and $\hat{H}(z)$ be

$$H(z) = \frac{N(z)}{D(z)} \quad \text{and} \quad \hat{H}(z) = \frac{\hat{N}(z)}{\hat{D}(z)}$$

where

$$N(z) = b_0 z^m + b_1 z^{m-1} + \cdots + b_m$$
$$D(z) = z^n + a_1 z^{n-1} + \cdots + a_n$$
$$\hat{N}(z) = \hat{b}_0 z^{\hat{m}} + \hat{b}_1 z^{\hat{m}-1} + \cdots + \hat{b}_{\hat{m}}$$

and

$$\hat{D}(z) = z^{\hat{n}} + \hat{a}_1 z^{\hat{n}-1} + \cdots + \hat{a}_{\hat{n}}$$

If $\hat{m}$ and $\hat{n}$ are chosen as $\hat{m} = m$ and $\hat{n} = n$, it is possible to make $J(z)$ be SPR, because choosing $\hat{H}(z)$ as $\hat{H}(z) = H(z)$ yields $J(z) = 1$. The next theorem gives margins of corresponding differences between $(a_1, \cdots, a_n, b_0, \cdots, b_m)$ and $(\hat{a}_1, \cdots, \hat{a}_n, \hat{b}_0, \cdots, \hat{b}_m)$.

Theorem 16 *Suppose that $\hat{D}(z)$ and $\hat{N}(z)$ are Schur stable. If $\hat{n} = n$, $\hat{m} = m$ and parameters $(a_1, \cdots, a_n, b_0, \cdots, b_m)$ satisfy*

$$|a_1 - \hat{a}_1| + \cdots + |a_n - \hat{a}_n| < \frac{1}{\sqrt{2}} \min_{\omega \in R} |\hat{D}(e^{j\omega})| \tag{5.36}$$

and

$$|b_0 - \hat{b}_0| + \cdots + |b_m - \hat{b}_m| < \frac{1}{\sqrt{2}} \min_{\omega \in R} |\hat{N}(e^{j\omega})| \tag{5.37}$$

Then $J(z)$ is SPR.

Proof: Note that (5.36) and (5.37) imply $|D(e^{j\omega}) - \hat{D}(e^{j\omega})| < |\hat{D}(e^{j\omega})|$ and $|N(e^{j\omega}) - \hat{N}(e^{j\omega})| < |\hat{N}(e^{j\omega})|$, respectively. Then by Rouché's theorem[57] $D(z) = \hat{D}(z) + D(z) - \hat{D}(z)$ and $N(z) = \hat{N}(z) + N(z) - \hat{N}(z)$ are Schur stable if $\hat{D}(z)$ and $\hat{N}(z)$ are Schur stable. Since

$$J(z) = \frac{1 + \{N(z) - \hat{N}(z)\}/\hat{N}(z)}{1 + \{D(z) - \hat{D}(z)\}/\hat{D}(z)}$$

we have

$$|\arg J(e^{j\omega})| \leq \left|\arg\left(1+\frac{N(e^{j\omega})-\hat{N}(e^{j\omega})}{\hat{N}(e^{j\omega})}\right)\right| + \left|\arg\left(1+\frac{D(e^{j\omega})-\hat{D}(e^{j\omega})}{\hat{D}(e^{j\omega})}\right)\right| \quad (5.38)$$

On the other hand, note that $\hat{N}(z)$ and $\hat{D}(z)$ have no zero on the unit circle; then from (5.36) and (5.37) we obtain

$$\left|\frac{D(e^{j\omega})-\hat{D}(e^{j\omega})}{\hat{D}(e^{j\omega})}\right| \leq \frac{|a_1-\hat{a}_1|+\cdots+|a_n-\hat{a}_n|}{|\hat{D}(e^{j\omega})|} < \frac{1}{\sqrt{2}}$$

and

$$\left|\frac{N(e^{j\omega})-\hat{N}(e^{j\omega})}{\hat{N}(e^{j\omega})}\right| \leq \frac{|b_0-\hat{b}_0|+\cdots+|b_m-\hat{b}_m|}{|\hat{N}(e^{j\omega})|} < \frac{1}{\sqrt{2}}$$

Those inequalities lead to

$$\left|\arg\left(1+\frac{D(e^{j\omega})-\hat{D}(e^{j\omega})}{\hat{D}(e^{j\omega})}\right)\right| < \frac{\pi}{4} \quad (5.39)$$

for any $\omega \in R$ and

$$\left|\arg\left(1+\frac{N(e^{j\omega})-\hat{N}(e^{j\omega})}{\hat{N}(e^{j\omega})}\right)\right| < \frac{\pi}{4} \quad (5.40)$$

for any $\omega \in R$, respectively. From (5.38), (5.39) and (5.40), we establish $\mathrm{Re}[J(e^{j\omega})] > 0$ for any $\omega \in R$. This completes the proof. ■

The assumptions (5.36) and (5.37) in Theorem 16 are simple and convenient for applications. However they could be conservative estimates of the parameter margins, because they are sufficient conditions for $J(z)$ to be SPR. Next, a necessary and sufficient condition will be discussed. Suppose that parameters $(a_1, \cdots, a_n, b_0, \cdots, b_m)$ satisfy

$$b_k \in [\underline{b}_k, \overline{b}_k] \quad (k = 0, \cdots, m) \quad (5.41)$$

and

$$a_k \in [\underline{a}_k, \overline{a}_k] \quad (k = 1, \cdots, n) \quad (5.42)$$

It should be noted that $n = \hat{n}$ or $m = \hat{m}$ are not necessarily assumed in the following discussions. A condition for

$$H(z) = \frac{b_0 z^m + b_1 z^{m-1} + \cdots + b_m}{z^n + a_1 z^{n-1} + \cdots + a_n}$$

with (5.41) and (5.42) to be SPR is already known[58]. Although their result cannot be applied directly to

$$J(z) = \frac{(\hat{b}_0 z^{\hat{m}} + \hat{b}_1 z^{\hat{m}-1} + \cdots + \hat{b}_{\hat{m}})}{(z^{\hat{n}} + \hat{a}_1 z^{\hat{n}-1} + \cdots + \hat{a}_{\hat{n}})} \frac{(z^n + a_1 z^{n-1} + \cdots + a_n)}{(b_0 z^m + b_1 z^{m-1} + \cdots + b_m)}$$

which we are interested in, we can show a similar result as follows.

Theorem 17 *$J(z)$ is SPR in a parameter space (hyper cube)*

$$\{(a_1, \cdots, a_n, b_0, \cdots, b_m); \; b_k \in [\underline{b}_k, \overline{b}_k] \; (k = 0, \cdots, m), \; a_k \in [\underline{a}_k, \overline{a}_k] \; (k = 1, \cdots, n)\}$$

if and only if

1. *$J(z)$ and $J(z)^{-1}$ are Schur stable on $m/2] + n/2] + 1$ line segments, $\Phi_0, \cdots, \Phi_{m/2]}$ and $\Psi_1, \cdots, \Psi_{n/2]}$ where*

$$\Phi_k = \{(b_0, \cdots, b_k, \cdots, b_m) | b_k \in [\underline{b}_k, \overline{b}_k], \; b_i \in \{\underline{b}_i, \overline{b}_i\} (i = 0, \cdots, m; \text{ with the exception that } i = k)\}$$

$$\Psi_k = \{(a_1, \cdots, a_k, \cdots, a_n) | a_k \in [\underline{a}_k, \overline{a}_k], \; b_i \in \{\underline{a}_i, \overline{a}_i\} (i = 1, \cdots, n; \text{ with the exception that } i = k)\}$$

and $m/2]$ indicates the largest integer less than $m/2$.

2. $J(z)$ satisfies

$$Re[J(e^{j\omega})] > 0 \tag{5.43}$$

for any $\omega \in R$ on 2^{m+n+1} corner points

$$\begin{aligned}&\{(a_1, \cdots, a_n, b_0, \cdots, b_m);\\ &\quad b_k \in \{\underline{b}_k, \overline{b}_k\}\ (k = 0, \cdots, m),\\ &\quad a_k \in \{\underline{a}_k, \overline{a}_k\}\ (k = 1, \cdots, n)\}\end{aligned}$$

Proof: First, we can see that $J(z)$ and $J(z)^{-1}$ are Schur stable if $N(z)$, $D(z)$, $\hat{N}(z)$ and $\hat{D}(z)$ are Schur stable. Direct application of the known result[58] to the polynomials $N(z)$ and $D(z)$ establish condition No. *1*.

Second, since (5.35) is equivalent to

$$\mathrm{Re}[\hat{N}(e^{j\omega})D(e^{j\omega})N(e^{-j\omega})\hat{D}(e^{-j\omega})] > 0$$

for any $\omega \in R$, we will discuss the positiveness of

$$\begin{aligned}&f(\omega, a_1, \cdots, a_n, b_0, \cdots, b_m)\\ &= \mathrm{Re}[\hat{N}(e^{j\omega})D(e^{j\omega})N(e^{-j\omega})\hat{D}(e^{-j\ omega})]\end{aligned}$$

From

$$\hat{N}(z)D(z) = \sum_{k_1=0}^{n+\hat{m}} \sum_{i_1=0}^{\hat{m}} \hat{b}_{\hat{m}-i_1} a_{n-k_1+i_1} z^{k_1}$$

and

$$N(z)\hat{D}(z) = \sum_{k_2=0}^{\hat{n}+m} \sum_{i_2=0}^{m} b_{m-i_2} \hat{a}_{\hat{n}-k_2+i_2} z^{k_2}$$

where $a_0 = \hat{a}_0 = 1$, $a_i = 0$ $(i = -\hat{m}, \cdots, -1, n+1, \cdots, n+\hat{m})$ and $\hat{a}_i = 0$ $(i = -m, \cdots, -1, \hat{n}+1, \cdots, \hat{n}+m)$, we have

$$\begin{aligned}&f(\omega, a_1, \cdots, a_n, b_0, \cdots, b_m)\\ &= \left(\sum_{k_1=0}^{n+\hat{m}} \sum_{i_1=0}^{\hat{m}} \hat{b}_{\hat{m}-i_1} a_{n-k_1+i_1} \cos k_1\omega\right)\\ &\quad \left(\sum_{k_2=0}^{\hat{n}+m} \sum_{i_2=0}^{m} b_{m-i_2} \hat{a}_{\hat{n}-k_2+i_2} \cos k_2\omega\right)\\ &\quad + \left(\sum_{k_1=0}^{n+\hat{m}} \sum_{i_1=0}^{\hat{m}} \hat{b}_{\hat{m}-i_1} a_{n-k_1+i_1} \sin k_1\omega\right)\\ &\quad \left(\sum_{k_2=0}^{\hat{n}+m} \sum_{i_2=0}^{m} b_{m-i_2} \hat{a}_{\hat{n}-k_2+i_2} \sin k_2\omega\right)\end{aligned}$$

Since the right hand side of the equation shows that $f(\omega, a_1, \cdots, a_n, b_0, \cdots, b_m)$ is a linear function of parameters $a_1, \cdots, a_n, b_0 \cdots, b_m$, we obtain

$$\begin{aligned}&f(\omega, a_1, \cdots, a_n, b_0, \cdots, b_m)\\ &= \lambda f(\omega, a_1, \cdots, \underline{a}_k, \cdots, a_n, b_0, \cdots, b_m)\\ &\quad +(1-\lambda) f(\omega, a_1, \cdots, \overline{a}_k, \cdots, a_n, b_0, \cdots, b_m)\end{aligned}$$

etc. where $\lambda \in [0, 1]$. This implies that if $f(\omega, a_1, \cdots, a_n, b_0, \cdots, b_m) > 0$ for

$$\begin{aligned}&\omega \in R\\ &a_i \in \{\underline{a}_i, \overline{a}_i\}\quad (i = 1, \cdots, k)\\ &a_i \in [\underline{a}_i, \overline{a}_i]\quad (i = k+1, \cdots, n)\\ &b_i \in [\underline{b}_i, \overline{b}_i]\quad (i = 0, \cdots, m)\end{aligned}$$

then $f(\omega, a_1, \cdots, a_n, b_0, \cdots, b_m) > 0$ for

$$\begin{aligned}&\omega \in R\\ &a_i \in \{\underline{a}_i, \overline{a}_i\}\quad (i = 1, \cdots, k-1)\\ &a_i \in [\underline{a}_i, \overline{a}_i]\quad (i = k, \cdots, n)\\ &b_i \in [\underline{b}_i, \overline{b}_i]\quad (i = 0, \cdots, m)\end{aligned}$$

Iteration of such induction for a_i and b_i establishes condition No. *2*. ■

5.5 Numerical Examples

In this section, we present numerical examples to illustrate the iterative learning control and its convergence conditions given in the preceding section.

Example 11 *Consider a system*

$$H(z) = \frac{b_0 z + b_1}{z^2 + a_1 z + a_2}$$

Let $\hat{H}(z)$ *be*

$$\hat{H}(z) = \frac{0.15z + 0.09}{z^2 - 1.0z + 0.2} \tag{5.44}$$

Then $\min_{\omega\in R}|\hat{D}(e^{j\omega})| = 0.2$ *and* $\min_{\omega\in R}|\hat{N}(e^{j\omega})| = 0.06$*; by Theorem 16,* $J(z) = \hat{H}(z)H^{-1}(z)$ *is SPR if*

$$\begin{aligned} |a_1 + 1.0| + |a_2 - 0.2| &\leq 0.14 \\ |b_0 - 0.15| + |b_1 - 0.09| &\leq 0.042 \end{aligned}$$

The algorithm that is derived from (5.44) based on Theorem 15 is

$$u_{n+1}(k) = u_n(k) - \alpha\eta_n(k) \quad (k = 0, \cdots, N-1)$$

$$\begin{aligned} p_n(k-1) &= \begin{bmatrix} 0 & 1 \\ -0.2 & 1.0 \end{bmatrix} p_n(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \epsilon_n(k) \\ \eta_n(k) &= \begin{bmatrix} 0.09 & 0.15 \end{bmatrix} p_n(k) \end{aligned}$$

$$\epsilon_n(k) = y_n(k) - y_d(k)$$

where $\{y_n(k)|k = 0, \cdots, N-1\}$ *is the output of (5.1) for* $\{u_n(k)|k = 0, \cdots, N-1\}$*. Figure 5.1 shows* $|e_n|$ *for* $n = 0, 1, \cdots, 10$ *when*

$$\begin{aligned} (a_1, a_2) &= (-1.1, 0.18) \\ (b_0, b_1) &= (0.16, 0.12) \\ y_d(k) &= \sin\frac{2k}{100}\pi \quad (k = 0, 1, \cdots, 100) \\ u_0(k) &= 0 \quad (k = 0, 1, \cdots, 100) \\ \alpha &= 0.2 \\ M &= 0 \end{aligned}$$

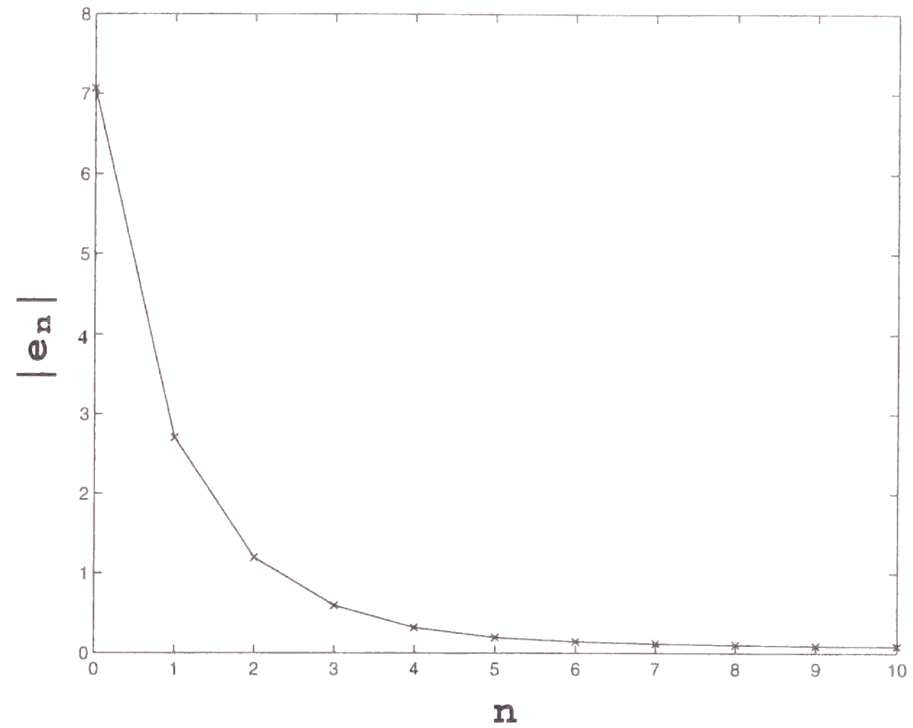


Figure 5.1: $|e_n|$ of Example 11.

Example 12 *Consider a system*

$$H(z) = \frac{b_0 z + b_1}{z^2 + a_1 z + a_2}$$

with interval parameters

$$\begin{aligned} a_1 &\in [-0.55, -0.45] && (5.45) \\ a_2 &\in [0.04, 0.06] && (5.46) \\ b_0 &\in [2.0, 3.0] && (5.47) \\ b_1 &\in [-1.5, -0.5] && (5.48) \end{aligned}$$

Let $\hat{H}(z)$ *be*

$$\hat{H}(z) = \frac{1}{z - 0.5} \tag{5.49}$$

Then $b_0 z + b_1$ *and* $z^2 + a_1 z + a_2$ *are Schur stable on line elements*

$$\Phi_0 = \{(b_0, b_1); b_0 \in [2.0, 3.0],$$

$$b_1 \in \{-1.5, -0.5\}\}$$

and

$$\begin{aligned}\Psi_1 &= \{(a_1, a_2); a_1 \in [-0.55, -0.45], \\ &\quad a_2 \in \{0.04, 0.06\}\}\end{aligned}$$

$Re[J(e^{j\omega})] > 0$ for any $\omega \in R$ on points

$$\begin{aligned}&\{(a_1, a_2, b_0, b_1); \\ &\quad a_1 \in \{-0.55, -0.45\}, a_2 \in \{0.04, 0.06\}, \\ &\quad b_0 \in \{2.0, 3.0\}, b_1 \in \{-1.5, -0.5\}\}\end{aligned}$$

Therefore, $J(z) = \hat{H}(z)H(z)^{-1}$ is SPR for any (5.45), (5.46), (5.47) and (5.48).

5.6 Concluding Remarks

In this chapter, iterative learning control based on the gradient method for linear discrete-time systems was presented. When there is uncertainty in the system, a convergence condition is given by matrix inequalities or the SPR condition on a transfer function which represents the uncertainty of the system. Furthermore, the SPR condition on the transfer function was replaced with inequalities of parameters that are convenient for practical applications.

It should be noted again that the iterative learning control presented in this paper can be directly applied to sampled-data systems with 0-order hold if the relative degree of the transfer function of the original continuous-time system is 0, 1 or 2. As demonstrated in chapter 4, the inter-sample error between the output and the desired trajectory converges to 0 as the sampling period goes to 0 independently of stability of zeros in the transfer function of the sampled-data system.

Chapter 6

Conclusion

In this thesis, we discussed iterative learning control using adjoint systems for linear continuous-time systems and its extension to sampled-data systems.

In Chapter 2, we demonstrated convergence of the iterative learning control applied to linear continuous-time systems with uncertainty. The main convergence condition is strictly coerciveness or strictly positiveness of the unknown part of the system. Next, in Chapter 3, we estimated convergence rate of the iterative learning control and proved that it cannot be exponential one, which yields robustness against noise. In order to achieve exponential convergence, we introduce a regularization method into the iterative learning control in exchange for tracking performance.

In Chapter 4 and 5, we discussed extension of the iterative learning control presented in the preceding chapters to linear sampled-data systems. In Chapter 4, we treated ripples at inter-sample points which are effects of unstable zeros of sampled-data systems. We proved that if the relative degree of transfer function of the continuous-time system is 0, 1 or 2, shrinking the sampling period reduces the ripples independently of stability zeros. In Chapter 5, we developed iterative learning control for linear sampled-data systems based on the same idea as in Chapter 2. Since the results in 4 presents how to decrease ripples at inter-sample points, we dealt with linear discrete-time systems as object systems of the iterative learning control.

Appendix A

Proof of Lemma 1

Note that $cA_c^k b_c = 0$ $(k = 0, 1, \cdots, n-m-2)$ if $n - m \geq 2$ and $cA_c^{n-m-1} b_c \neq 0$ Then we have

$$cb = \sum_{p=n-m}^{\infty} \frac{cA_c^{p-1} b_c \Delta^p}{p!}$$

and hence

$$\frac{cb}{\Delta^{n-m}} = \frac{cA_c^{n-m-1} b_c}{(n-m)!} + \sum_{p=n-m+1}^{\infty} \frac{cA_c^{p-1} b_c \Delta^{p-n+m}}{p!} \tag{A.1}$$

On the other hand, since

$$\left| \sum_{p=n-m+1}^{\infty} \frac{cA_c^{p-1} b_c \Delta^{p-n+m}}{p!} \right| \leq \sum_{p=n-m+1}^{\infty} \frac{|c||A_c|^{p-1}|b_c|\Delta^{p-n+m}}{p!}$$

the right hand side of which converges to a continuous function for $\Delta \in (-\infty, \infty)$, we have

$$\lim_{\Delta \to 0} \left| \sum_{p=n-m+1}^{\infty} \frac{cA_c^{p-1} b_c \Delta^{p-n+m}}{p!} \right| = 0 \tag{A.2}$$

Therefore, there exists $\epsilon_0 > 0$ such that

$$\left| \frac{cb}{\Delta^{n-m}} \right| > 0.$$

for any $\Delta \in (0, \epsilon_0)$ ∎

Appendix B

Proof of Lemma 2

By applying a differentiation formula for matrix determinant to (4.7), we obtain

$$\left.\frac{d}{d\Delta}N(z)\right|_{\Delta=0} = (z-1)^{n-1}cb_c$$
$$\left.\frac{d^2}{d\Delta^2}N(z)\right|_{\Delta=0} = (z+1)(z-1)^{n-2}cA_cb_c$$

$$\left.\frac{d^3}{d\Delta^3}N(z)\right|_{\Delta=0}$$
$$= \begin{cases} (z+5)(z-1)^{n-2}cA_c^2b_c \\ -3\text{trace}(A_c)cA_cb_c(z-1)^{n-2} \quad \text{if} \quad n=2 \\ \\ (z^2+4z+1)(z-1)^{n-3}cA_c^2b_c \\ -3\text{trace}(A_c)cA_cb_c(z+1)(z-1)^{n-3} \\ \quad \text{if} \quad n \geq 3 \end{cases}$$

Note that $cb_c = 0$ and $cA_c^kb_c \neq 0 (k = 1, 2, \cdots)$ when $n - m = 2$. Then we have

$$\begin{aligned} N(z) &= (z+1)(z-1)^{n-2}cA_cb_c\frac{\Delta^2}{2} \\ &+ \left.\frac{d^3}{d\Delta^3}N(z)\right|_{\Delta=0} cA_c^2b_c\frac{\Delta^3}{6} \\ &+O(\Delta^4) \end{aligned}$$

moreover

$$\frac{N(-1)}{\Delta^3} = \begin{cases} (-2)^{n-2}\frac{4cA_c^2b_c - 3\mathrm{trace}(A_c)cA_cb_c}{6} \\ +O(\Delta) \quad \text{if } n = 2 \\ \\ (-2)^{n-2}\frac{cA_c^2b_c}{6} + O(\Delta) \quad \text{if } n \geq 3 \end{cases}$$

On the other hand, from (4.9), we have

$$\frac{N(-1)}{\Delta^3} = \frac{cb_\Delta}{\Delta^2}\frac{(-1 - q_1(\Delta))\cdots(-1 - q_{n-1}(\Delta))}{\Delta}$$

moreover

$$\begin{aligned} &\frac{q_{n-1}(\Delta)+1}{\Delta} \\ &= -\left\{(-2)^{n-2}\frac{M_n}{6} + O(\Delta)\right\} \\ &\quad \left\{\frac{cb_\Delta}{\Delta^2}(-1 - q_1(\Delta))\cdots(-1 - q_{n-2}(\Delta))\right\}^{-1} \end{aligned}$$

where

$$M_n = \begin{cases} 4cA_c^2b_c - 3\mathrm{trace}(A_c)cA_cb_c & \text{if } n = 2 \\ cA_c^2b_c & \text{if } n \geq 3 \end{cases}$$

Since we obtain $\frac{cb_\Delta}{\Delta^2} = \frac{cA_cb_c}{2} + O(\Delta)$ and

$$(-1 - q_1(\Delta))\cdots(-1 - q_{n-2}(\Delta)) = (-2)^{n-2} + O(\Delta)$$

from (4.16), we establish

$$\begin{aligned} &\frac{q_{n-1}(\Delta)+1}{\Delta} \\ &= -\left\{(-2)^{n-2}\frac{M_n}{6} + O(\Delta)\right\} \\ &\quad \left\{(-2)^{n-2}\frac{cA_cb_c}{2} + O(\Delta)\right\}^{-1} \\ &= -\frac{M_n}{3cA_cb_c} + O(\Delta). \end{aligned}$$

■

Appendix C

Proof of Lemma 3

If $d \neq 0$ then Γ is nonsingular, equivalently $\Gamma^+ = \Gamma^{-1}$. We obtain immediately (4.22). We will discuss the case of $d = 0$, i.e. $n - m \geq 1$. From (4.11) and (4.15), we have

$$\Gamma_\Delta\Gamma_\Delta^+\sigma_\Delta y^* = [0\ y^*(\Delta)\ y^*(2\Delta)\ \cdots\ y^*(N\Delta)]^T$$

for $\Delta \in (0, \epsilon_0)$. On the other hand, $u^* \in C[0, t_f]$ and $y^* = Su^*$ imply $y^*(0) = 0$. This completes the proof. ■

Appendix D

Proof of Lemma 4

Since $d \neq 0$, we can rewrite (4.4) as

$$\begin{aligned} x(k+1) &= (A - d^{-1}bc)x(k) + d^{-1}bw(k) \\ v(k) &= -d^{-1}cx(k) + d^{-1}w(k) \end{aligned}$$

These equation with $x(0) = 0$ leads to

$$\begin{aligned} v(0) &= d^{-1}w(0) \\ v(k) &= -d^{-1}c\sum_{p=0}^{k-1}(A - d^{-1}bc)^p d^{-1}bw(k-1-p) \\ &\quad +d^{-1}w(k) \quad (k = 1, 2, \cdots, N) \end{aligned}$$

and hence

$$\begin{aligned} |v(0)| &\leq |d^{-1}||w(0)| \\ |v(k)| &\leq |d^{-1}c|\sum_{p=0}^{k-1}|A - d^{-1}bc|^p|d^{-1}||b||w(k-1-p)| \\ &\quad +|d^{-1}||w(k)| \\ &\leq |d^{-1}c||d^{-1}| \max\{|w(p)|; p = 0, 1, \cdots, k-1\} \cdot \\ &\quad |b|\sum_{p=0}^{k-1}|A - d^{-1}bc|^p + |d^{-1}||w(k)| \end{aligned}$$

Moreover

$$\max\{|v(k)|; k = 1, 2, \cdots, N\}$$

$$\leq \; |d^{-1}c||d^{-1}| \max\{|w(p)|; p = 0, 1, \cdots, N-1\} \cdot$$
$$\max\{|b| \sum_{p=0}^{k-1} |A - d^{-1}bc|^p; k = 1, 2, \cdots, N\}$$
$$+|d^{-1}| \max\{|w(k)|; k = 1, 2, \cdots, N\} \qquad \text{(D.1)}$$

On the other hand, since

$$\begin{aligned} |A| &= \left| \sum_{k=0}^{\infty} \frac{A_c^k \Delta^k}{k!} \right| \\ &\leq \sum_{k=0}^{\infty} \frac{|A_c|^k \Delta^k}{k!} = \exp(|A_c|\Delta) \end{aligned} \qquad \text{(D.2)}$$

and

$$\begin{aligned} |b| &= \left| |b_c| \sum_{k=1}^{\infty} \frac{|A_c|^{k-1} \Delta^k}{k!} \right| \\ &\leq |b_c| \sum_{k=1}^{\infty} \frac{|A_c|^{k-1} \Delta^k}{k!} \\ &= |b_c| \frac{\exp(|A_c|\Delta) - 1}{|A_c|} \end{aligned} \qquad \text{(D.3)}$$

we can estimate

$$|A - d^{-1}bc| \leq \exp(|A_c|\Delta) + L_1\{\exp(|A_c|\Delta) - 1\}$$

where L_1 is a positive constant. This implies

$$\begin{aligned} &\sum_{p=0}^{k-1} |A - d^{-1}bc|^p \\ &\quad \leq \frac{[\exp(|A_c|\Delta) + L_1\{\exp(|A_c|\Delta) - 1\}]^k - 1}{\{\exp(|A_c|\Delta) - 1\}(1 + L_1)} \end{aligned}$$

furthermore

$$\begin{aligned} &\max\{|b| \sum_{p=0}^{k-1} |A - d^{-1}bc|^p; k = 1, 2, \cdots, N\} \\ &\leq \frac{|b_c|}{|A_c|} \frac{\{\exp(|A_c|\Delta) + L_1(\exp(|A_c|\Delta) - 1)\}^{T/\Delta} - 1}{1 + L_1} \\ &\leq L_2 \end{aligned}$$

for $\Delta = t_f/N \in (0, \epsilon_0)$ where L_2 is a positive constant. Therefore, from (D.1) we obtain

$$\begin{aligned} &\max\{|v(k)|; k = 1, 2, \cdots, N\} \\ &\quad \leq (|d^{-1}c||d^{-1}|L_2 + |d^{-1}|) \max\{|w(p)|; p = 0, 1, \cdots, N\} \end{aligned}$$

This inequality with (D.1) leads to (4.23). ■

Appendix E

Proof of Lemma 5

Since $\Delta \in (0, \epsilon_0)$ implies $\frac{cb_\Delta}{\Delta} \neq 0$, we can rewrite (4.4) as

$$\begin{aligned} x(k+1) &= \left\{ A - b\left(\frac{cb}{\Delta}\right)^{-1} c\frac{A-I}{\Delta} \right\} x(k) \\ &\quad + b\left(\frac{cb}{\Delta}\right)^{-1} [\delta_\Delta w](k) \end{aligned} \tag{E.1}$$

$$\begin{aligned} v(k) &= -\left(\frac{cb}{\Delta}\right)^{-1} c\frac{A-I}{\Delta} x(k) \\ &\quad + \left(\frac{cb}{\Delta}\right)^{-1} [\delta_\Delta w](k) \end{aligned} \tag{E.2}$$

for $k = 0, 1, \cdots, N-1$. Those equations with $x(0) = 0$ lead to

$$\begin{aligned} v(k) &= -\left(\frac{cb}{\Delta}\right)^{-1} c\frac{A-I}{\Delta} \sum_{p=0}^{k-1} F_\Delta^p g_\Delta [\delta_\Delta w](k-1-p) \\ &\quad + \left(\frac{cb}{\Delta}\right)^{-1} [\delta_\Delta w](k) \end{aligned}$$

for $k = 0, 1, \cdots, N-1$ where

$$F_\Delta = A - b\left(\frac{cb}{\Delta}\right)^{-1} c\frac{A-I}{\Delta} \tag{E.3}$$

$$g_\Delta = b\left(\frac{cb}{\Delta}\right)^{-1} \tag{E.4}$$

This implies

$$
\begin{aligned}
|v(k)| \leq \; & \left|\left(\frac{cb}{\Delta}\right)^{-1}\right|\left|c\frac{A-I}{\Delta}\right| \cdot \\
& \max\{|[\delta_\Delta w](p)|; p=0,1,\cdots,k-1\}\cdot \\
& |g_\Delta|\sum_{p=0}^{k-1}|F_\Delta|^p \\
& +\left|\left(\frac{cb}{\Delta}\right)^{-1}\right||[\delta_\Delta w](k)|
\end{aligned}
$$

and hence

$$
\begin{aligned}
& \max\{|v(k)|; k=1,2,\cdots,N-1\} \\
\leq \; & \left|\left(\frac{cb}{\Delta}\right)^{-1}\right|\left|c\frac{A-I}{\Delta}\right| \cdot \\
& \max\{|[\delta_\Delta w](p)|; p=0,1,\cdots,N-2\}\cdot \\
& \max\{|g_\Delta|\sum_{p=0}^{k-1}|F_\Delta|^p; k=1,2,\cdots,N\} \\
& +\left|\left(\frac{cb}{\Delta}\right)^{-1}\right| \cdot \\
& \max\{|[\delta_\Delta w](p)|; p=1,2,\cdots,N-1\} \qquad \text{(E.5)}
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
\left|\frac{A-I}{\Delta}\right| &= \left|\sum_{k=1}^{\infty}\frac{A_c^k\Delta^{k-1}}{k!}\right| \\
&\leq \sum_{k=1}^{\infty}\frac{|A_c|^k\Delta^{k-1}}{k!} = \frac{\exp(|A_c|\Delta)-1}{\Delta}
\end{aligned}
$$

and, from (A.1) and (A.2),

$$
0 < \left|\left(\frac{cb}{\Delta}\right)^{-1}\right| < M_1 \qquad \text{(E.6)}
$$

for any $\Delta \in (0, \epsilon_0)$ where M_1 is a positive constant. Moreover, from (E.3), (E.4), (D.2), (E.6), (D.3), and (E.6), we obtain

$$
|g_\Delta| \leq |b_c|\frac{\exp(|A_c|\Delta)-1}{|A_c|}M_1
$$

$$
\begin{aligned}
|F_\Delta| \leq \; & \exp(|A_c|\Delta) \\
& +|b_c|\frac{\exp(|A_c|\Delta)-1}{|A_c|}M_1|c|\frac{\exp(|A_c|\Delta)-1}{\Delta} \\
\leq \; & \exp(|A_c|\Delta)+M_2(\exp(|A_c|\Delta)-1)
\end{aligned}
$$

for $\Delta \in (0, \epsilon_0)$ where

$$
M_2 = \frac{|b_c|M_1|c|(\exp(|A_c|\epsilon_0)-1)}{|A_c|\epsilon_0}
$$

By these inequalities, we can estimate

$$
\begin{aligned}
& |g_\Delta|\sum_{p=0}^{k-1}|F_\Delta|^p \\
= \; & |b_c|\frac{\exp(|A_c|\Delta)-1}{|A_c|}M_1 \cdot \\
& \frac{\{\exp(|A_c|\Delta)+M_2(\exp(|A_c|\Delta)-1)\}^k-1}{\{\exp(|A_c|\Delta)+M_2(\exp(|A_c|\Delta)-1)-1} \\
= \; & \frac{|b_c|M_1}{|A_c|}\frac{\{\exp(|A_c|\Delta)+M_2(\exp(|A_c|\Delta)-1)\}^k-1}{1+M_2}
\end{aligned}
$$

and hence

$$
\begin{aligned}
& \max\left\{|g_\Delta|\sum_{p=0}^{k-1}|F_\Delta|^p; k=1,2,\cdots,N\right\} \\
= \; & \frac{|b_c|M_1}{|A_c|}\frac{\{\exp(|A_c|\Delta)+M_2(\exp(|A_c|\Delta)-1)\}^{t_f/\Delta}-1}{1+M_2} \\
\leq \; & M_3 \qquad \text{(E.7)}
\end{aligned}
$$

for $\Delta = t_f/N \in (0, \epsilon_0)$ where M_3 is a positive constant. The inequality (E.5) with (E.6), (E.6) and (E.7) leads to

$$
\begin{aligned}
& \max\{|v(k)|; k=1,2,\cdots,N-1\} \\
\leq \; & M_1\left\{|c|\frac{\exp(|A_c|\epsilon_0)-1}{\epsilon_0}M_3-1\right\}\cdot \\
& \max\{|[\delta_\Delta w](p)|; p=0,1,\cdots,N-1\}
\end{aligned}
$$

and the equation (E.2) with $x(0)=0$ and (E.6) implies

$$
|v(0)| \leq M_1|[\delta_\Delta w](0)|
$$

for $\Delta \in (0, \epsilon_0)$. Combination of those inequalities leads to (4.24). ■

Appendix F

Proof of Lemma 6

From Theorem 11 and Lemma 2, we have

$$
\begin{aligned}
& H_1^{-1}(z) \\
= & \frac{\Delta^2}{cb_\Delta}\{1+ \\
& \left.\frac{r_0(\Delta)z^n + r_1(\Delta)z^{n-1} + \cdots + r_n(\Delta)}{(z-1)^2(z-q_1(\Delta))\cdots(z-q_{n-1}(\Delta))}\right\}
\end{aligned}
$$

where

$$\frac{|r_j(\Delta)|}{\Delta} \leq +\infty \quad (j = 0, 1, \cdots n) \tag{F.1}$$

Consider the realization of $H_1^{-1}(z)$ in the controllable canonical form and

$$
\begin{aligned}
x(k+1) &= \bar{A}_\Delta x(k) + \begin{bmatrix} r_0(\Delta) \\ r_0(\Delta) \\ \vdots \\ r_n(\Delta) \end{bmatrix} \eta(k) \\
u(k) &= \begin{bmatrix} 0 & \cdots & 0 & \frac{\Delta^2}{cb_\Delta} \end{bmatrix} x(k) + \frac{\Delta^2}{cb_\Delta}\eta(k)
\end{aligned} \tag{F.2}
$$

with the initial value $x(0) = 0$. Then we have

$$
\begin{aligned}
& \left|\frac{cb_\Delta}{\Delta^2}u(k)\right| \\
\leq & \sum_{j=0}^{k-1} |\bar{A}_\Delta|^j \left|\begin{bmatrix} r_0(\Delta) \\ r_0(\Delta) \\ \vdots \\ r_n(\Delta) \end{bmatrix}\right| |\eta(k-j-1)| \\
& +|\eta(k)|
\end{aligned}
$$

for $k \geq 1$ where $|\cdot|$ indicates Euclidean norm or the induced norm. Note that by (F.1)

$$\left\| \begin{bmatrix} r_0(\Delta) \\ r_0(\Delta) \\ \vdots \\ r_n(\Delta) \end{bmatrix} \right\| \leq \bar{M}_1 \Delta$$

Since Theorem 11 and Lemma 2 imply

$$\begin{aligned} & ||\bar{A}_\Delta| - 1| \\ = & \ |\max\{1, |q_1(\Delta)|, \cdots, |q_{n-1}(\Delta)|\} - 1| \\ \leq & \ \bar{M}_2 \Delta \end{aligned}$$

where $\bar{M}_2$ is a positive constant, we have

$$|\bar{A}_\Delta| \leq 1 + \bar{M}_2\Delta \leq e^{\bar{M}_2\Delta}$$

Therefore, we obtain

$$\begin{aligned} & \left| \frac{cb_\Delta}{\Delta^2} u(k) \right| \\ \leq & \ \frac{e^{k\bar{M}_2\Delta} - 1}{e^{\bar{M}_2\Delta} - 1} \bar{M}_1 \Delta \max\{|\eta(i)|; i = 0, 1, \\ & \cdots, k-1\} + |\eta(k)| \qquad \text{(F.3)} \\ \leq & \ \frac{e^{\frac{t_f}{\Delta}\bar{M}_2\Delta} - 1}{\bar{M}_2} \bar{M}_1 \max\{|\eta(i)|; i = 0, 1, \\ & \cdots, k-1\} + |\eta(k)| \qquad \text{(F.4)} \end{aligned}$$

for $k = 1, 2, \cdots, N - 1 (= t_f/\Delta - 1)$. On the other hand, since

$$\frac{cb_\Delta}{\Delta^2} = \frac{cA_c b_c}{2} + O(\Delta)$$

there exists $\epsilon > 0$ such that

$$\left| \frac{\Delta^2}{cb_\Delta} \right| < \bar{M}_3 \qquad \text{(F.5)}$$

for any $\Delta \in (0, \epsilon)$ where $\bar{M}_3$ is a positive constant. The inequality (F.4) with (F.5) establishes

$$\begin{aligned} & |u(k)| \\ \leq & \ \bar{M}_3 \frac{e^{\frac{T}{\Delta}\bar{M}_2\Delta} - 1}{\bar{M}_2} \bar{M}_1 \\ & \max\{|\eta(i)|; i = 0, 1, \cdots, k-1\} \\ & + |\eta(k)| \qquad \text{(F.6)} \end{aligned}$$

for $k = 1, \cdots, N (= T/\Delta)$. From (F.2), we have $u(0) = \frac{\Delta^2}{cb_\Delta}\eta(0)$ and hence

$$|u(0)| \leq \bar{M}_3 |\eta(0)| \qquad \text{(F.7)}$$

From (F.6) and (F.7), we obtain (4.42). ∎

Appendix G

Proof of Lemma 7

The case of *1.*:

Since $\eta(k) = \sum_{j=1}^{k}(-1)^{j-1}\zeta(k-j)$, we have

$$\begin{aligned}&\max\{|\eta(k)|; k = 0, 1, \cdots, N-1\}\\ \leq\ & (N-1)\max\{|\zeta(k)|; k = 0, 1, \cdots, N-2\}\end{aligned}$$

and hence

$$\begin{aligned}&\max\{|\eta(k)|; k = 0, 1, \cdots, N-1\}\\ \leq\ & (t_f/\Delta - 1)\Delta^p\|f_\Delta\|_\infty\end{aligned}$$

This implies (4.43).

The case of *2.*:

Note that

$$\eta(k) = \begin{cases} \sum_{j=1}^{(k-1)/2}\{\zeta(k-2j+1) - \zeta(k-2j)\} \\ +\zeta(0) \text{ if } k\text{: odd} \\ \\ \sum_{j=1}^{k/2}\{\zeta(k-2j+1) - \zeta(k-2j)\} \\ \text{ if } k\text{: even}\end{cases}$$

and hence

$$|\eta(k)|$$

$$\le \begin{cases} \frac{k-1}{2}\max\{|\zeta(k-2j+1)-\zeta(k-2j)|; \\ j=1,\cdots,\frac{k-1}{2}\}+|\zeta(0)| \quad \text{if } k\text{: odd} \\ \\ \frac{k}{2}\max\{|\zeta(k-2j+1)-\zeta(k-2j)|; \\ j=1,\cdots,\frac{k}{2}\} \quad \text{if } k\text{: even} \end{cases}$$

$$\le \begin{cases} \frac{N-2}{2}\max\{|\zeta(k-2j+1)-\zeta(k-2j)|; \\ j=1,\cdots,\frac{N-2}{2}\} \\ +|\zeta(0)| \quad \text{if } N-1\text{: odd} \\ \\ \frac{N-1}{2}\max\{|\zeta(k-2j+1)-\zeta(k-2j)|; \\ j=1,\cdots,\frac{N-1}{2}\} \quad \text{if } N-1\text{: even} \end{cases}$$

Then

$$|\eta(k)| \le \begin{cases} \frac{T-2\Delta}{2}\Delta^{p-1}\max\{|f(t_{k-2j+1})-f(t_{k-2j})|; \\ j=1,\cdots,\frac{N-2}{2}\}+\Delta^p|f(t_0)| \\ \quad \text{if } N-1\text{: odd} \\ \\ \frac{T-\Delta}{2}\Delta^{p-1}\max\{|f(t_{k-2j+1})-f(t_{k-2j})|; \\ j=1,\cdots,\frac{N-1}{2}\} \quad \text{if } N-1\text{: even} \end{cases}$$

On the other hand, from $f\in C^0[0,t_f]$ and $|t_{k-2j+1}-t_{k-2j}|\le M_2\Delta$, we have

$$\max\{|f(t_{k-2j+1})-f(t_{k-2j})|; \quad j=1,\cdots,\frac{N-2}{2}\left(\text{or } \frac{N-1}{2}\right)\}\to 0$$

as $\Delta\to 0$. This implies (4.43).

The case of *3.*:

Note that, when $\Delta\to 0$, $\zeta(0)=f_\Delta(0)-f(0)\to 0$ and

$$\begin{aligned} &\zeta(k-2j+1)-\zeta(k-2j) \\ &= \Delta\left\{\frac{f_\Delta((k-2j+1)\Delta)-f_\Delta((k-2j)\Delta)}{\Delta} - \frac{f((k-2j+1)\Delta)-f((k-2j)\Delta)}{\Delta}\right\} \end{aligned}$$

$$\begin{aligned} &= \Delta\left\{\frac{d}{dt}f_\Delta((k-2j)\Delta+\alpha(k,\Delta)\Delta) - \frac{d}{dt}f((k-2j)\Delta+\alpha(k,\Delta)\Delta)\right\} \\ &= \Delta\left\{\frac{d}{dt}f_\Delta((k-2j)\Delta+\alpha(k,\Delta)\Delta) - \frac{d}{dt}f((k-2j)\Delta+\alpha(k,\Delta)\Delta)\right\} \\ &\quad +\Delta\left\{\frac{d}{dt}f((k-2j)\Delta+\alpha(k,\Delta)\Delta) - \frac{d}{dt}f((k-2j)\Delta+\beta(k,\Delta)\Delta)\right\} \end{aligned}$$

where $0\le\alpha(k,\Delta)\le 1$ and $0\le\beta(k,\Delta)\le 1$. Since $\frac{d}{dt}f\in C^0[0,T]$ and $\left\|\frac{d}{dt}f_\Delta-\frac{d}{dt}f\right\|_\infty\to 0$, (4.43) is established by the same reasoning as the case of *2.* ■

Appendix H

Proof of Lemma 8

Given $\{u(k); k = 0, 1, \cdots, N-1\}$, $\{x(k); k = 0, 1, \cdots, N\}$ and $\{y(k); k = 0, 1, \cdots, N-1\}$ defined by

$$\begin{aligned} x(k+1) &= \tilde{A}x(k) + \tilde{b}u(k) \\ y(k) &= \tilde{c}x(k) + \tilde{d}u(k) \end{aligned} \tag{H.1}$$

with $x(0) = 0$, by the discrete strictly positive real lemma[59] there exist matrices K and L and positive-definite symmetric matrices P and Q satisfying

$$\begin{aligned} 2\sum_{k=0}^{N-1} u(k)y(k) &= x(N)^T P x(N) + \sum_{k=0}^{N-1} x(k)^T Q x(k) \\ &\quad + \sum_{k=0}^{N-1} |Lx(k) + Ku(k)|^2 \end{aligned} \tag{H.2}$$

This leads to $w^T \Lambda_1 w \geq 0$ where

$$w = \begin{bmatrix} u(0) & u(1) & \cdots & u(N-1) \end{bmatrix}^T$$

Suppose that $w^T \Lambda_1 w = 0$, then from (H.2) we have $x(k) = 0$ $(k = 0, 1, \cdots, N)$ because P and Q are positive definite. This leads to $\tilde{b}u(k) = 0$ $(k = 0, 1, \cdots, N-1)$ by (H.1). Suppose, without loss of generality, that $(\tilde{A}, \tilde{b}, \tilde{c}, \tilde{d})$ is in controllable canonical form, then we can obtain $u(k) = 0$ $(k = 0, 1, \cdots, N-1)$ namely $w = 0$. Therefore since there exists the smallest positive eigenvalue μ of $(\Lambda + \Lambda^T)/2$, we obtain $w^T \Lambda w \geq \mu |w|^2$ ■

Bibliography

[1] K. L. Moore. *Iterative Learning Control for Deterministic Systems.* Springer-Verlag, 1993.

[2] M. Uchiyama. Formation of high-speed motion pattern of a mechanical arm by trial. *Transaction of the Society of Instrument and Control Engineers*, 14(6):706–712, 1978. (In Japanese).

[3] S. Kawamura, F. Miyazaki, and S. Arimoto. Realization of robot motion based on a learning method. *IEEE Transactions on systems, man, and cybernetics*, 18(1):126–134, 1988.

[4] S. Arimoto, S. Kawamura, and F. Miyazaki. Bettering operation of dynamic systems by learning: A new control theory for servomechanism or mechatronics systems. *Proceedings of 23rd IEEE Conference on Decision and Control*, pages 1064–1069, 1984.

[5] S. Arimoto, S. Kawamura, F. Miyazaki, and S. Tamaki. Learning control theory for dynamical systems. *Proceedings of 24th IEEE Conference on Decision and Control*, pages 1375–1380, 1985.

[6] P. Bondi, G. Casalino, and L. Gambardella. On the iterative learning control theory for robotic manipulators. *IEEE Journal of Robotics and Automation*, 4(1):14–22, 1988.

[7] G. Casalino and L. Gambardella. Learning of movements in robotic manipulators. *Proceedings of 1986 IEEE Conference on Robotics and Automation*, pages 572–578, 1986.

[8] S. Arimoto. Learning control theory for robotic motion. *International Journal of Adaptive Control and Signal Processing*, 4(6):543–564, 1990.

[9] G. Heinzinger, D.Fenwick, B. Paden, and F. Miyazaki. Stability of learning control with disturbances and uncertain initial conditions. *IEEE Transactions on Automatic Control*, 37(1):110–114, 1992.

[10] J. E. Hauser. Learning control for a class of nonlinear systems. *Proceedings of the 26th IEEE Conference on Decision and Control*, pages 859–860, 1987.

[11] S. S. Saab. On the P-type learning control. *IEEE Transactions on Automatic Control*, 39(11):2298–2302, 1994.

[12] T. Sugie and T. Ono. On a learning control law. *System & Control*, 31(2):129–135, 1987. (In Japanese).

[13] T. Sugie and T. Ono. An iterative learning control law for dynamical systems. *Automatica*, 27(4):729–732, 1991.

[14] K. L. Moore, M Dahleh, and S. P. Bhattacharyya. Iterative learning control: A survey and new results. *Journal of Robotic Systems*, 9(5):563–594, 1992.

[15] T. Mita and E. Kato. Iterative control and its application to motion control of robot arm - a direct approach to servo-problems -. *Proceedings of 24th IEEE Conference on Decision and Control*, pages 1393–1398, 1985.

[16] T.Mita and E.Kato. Iterative control of robot manipulators. *Proceedings of 15th ISIR*, pages 665–672, 1985.

[17] G. Heinzinger, D. Fenwick, B. Paden, and F. Miyazaki. Robust learning control. *Proceedings of the 28th IEEE Conference on Decision and Control*, pages 436–440, December 1989. Tampa, Florida.

[18] T. Inoue, M. Nakano, T. Kubo, S. Matsumoto, and H. Baba. High accuracy control of a proton synchrotron magnet power supply. *Proceedings of the Eighth Triennial World Congress of the International Federation of Automatic Control*, 6:3137–3142, 1982.

[19] T. Omata, M. Nakano, and T. Inoue. Application of repetitive control method to multivariable systems. *Transactions of the Society of Instrument and Control Engineers*, 20(9):795–800, 1984.

[20] Y. Yamamoto and S. Hara. The internal model principle and stabilizability of repetitive control systems. *Transactions of the Society of Instrument and Control Engineers*, 22(8):830–834, 1986.

[21] M. Ikeda and M. Takano. Repetitive control for systems with nonzero relative degree. *Transactions of the Society of Instrument and Control Engineers*, 24(6):575–582, 1988.

[22] S. Hara, Y. Yamamoto, T. Omata, and M. Nakano. Repetitive control system: A new type servo system for periodic exogenous signals. *IEEE Transactions on Automatic Control*, 33(7):659–668, 1988.

[23] M. Tomizuka, T. Tsao, and K. Chew. Analysis and synthesis of discrete-time repetitive controllers. *ASME Journal of Dynamic Systems, Measurement, and Control*, 111(9):353–358, 1989.

[24] M. Togai and O. Yamano. Analysis and design of an optimal learning control scheme for industrial robots: a discrete system approach. *Proceedings of 24th IEEE Conference on Decision and Control*, pages 1399–1404, 1985.

[25] S. S. Saab. A discrete-time learning control algorithm for a class of linear time-invariant systems. *IEEE Transactions on Automatic Control*, 40(6):1138 1142, 1995.

[26] J. E. Kurek and M. B. Zaremba. Iterative learning control synthesis based on 2-D system theory. *IEEE Transactions on Automatic Control*, 38(1):121 125, 1993.

[27] K. Furuta and M. Yamakita. Iterative generation of optimal input of a manipulator. *Proceedings of IEEE Robotics and Automation*, pages 579–583, 1986.

[28] M. Yamakita and K. Furuta. Iterative generation of virtual reference for a manipulator. *Robotica*, 9:71–80, 1991.

[29] C. W. Groetsch. *The theory of Tikhonov regularization for Fredholm equations of the first kind*. Pitman Publishing, 1984.

[30] W. J. Kammerer and M. Z. Nashed. Steepest descent for singular linear operators with nonclosed range. *Applicable Analysis*, 1:143–159, 1971.

[31] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle. *Stability of Adaptive Systems: Passivity and Averaging Analysis*. The MIT Press, 1986.

[32] T. Sogo and N. Adachi. A design method of the learning control algorithm for linear systems. *Proc. 1994 Japan-U.S.A. Symposium on Flexible Automation*, 1994.

[33] T. Sogo and N. Adachi. On the learning control scheme for linear systems. *Transactions of the Institute of Systems, Control and Information Engineers*, 7(9):339–346, 1994. (In Japanese).

[34] T. Sogo and N. Adachi. A method to design iterative learning control algorithm for linear systems. *Transactions of the Society of Instrument and Control Engineers*, 31(10):1601–1607, 1995. (In Japanese).

[35] P. Ioannou and G. Tao. Frequency domain conditions for strictly positive real functions. *IEEE Transactions on Automatic Control*, 32(1):53–54, 1987.

[36] S. Dasgupta and A. S. Bhagwat. Conditions for designing strictly positive real transfer functions for adaptive output error identification. *IEEE Transactions on Circuits and Systems*, 34(7):731–736, July 1987.

[37] B. D. O. Anderson, S. Dasgupta, P. Khargonekar, F. J. Kraus, and M. Mansour. Robust strict positive realness: Characterization and construction. *IEEE Transactions on Circuits and Systems*, 37(7):869–876, July 1990.

[38] T. Mori and H. Kokame. Extended Kharitonov's theorem and their application. *Proceedings of American Control Conference*, pages 627–632, 1989.

[39] V. L. Kharitonov. Asymptotic stability of an equilibrium position of a family of systems of linear differential equations. *Differential Equations*, pages 1483–1485, Nov. 1979.

[40] K. Yosida. *Functional Analysis*. Springer-Verlag, sixth edition, 1980.

[41] P. Ioannou and G. Tao. Frequency domain conditions for strictly positive real functions. *IEEE Transactions on Automatic Control*, AC-32(1):53 54, 1987.

[42] A. Hać. Learning control in the presence of measurement noise. *Proceedings of American Control Conference*, pages 2846–2851, 1990.

[43] S. Arimoto, S. Kawamura, and F. Miyazaki. Convergence, stability, and robustness of learning control schemes for robot manipulators. *Recent Trends in Robotics: Modeling, Control, and Education*, pages 307–316, 1986.

[44] C. W. Groetsch. *Inverse Problems in the Mathematical Sciences.* Vieweg, 1993.

[45] F. J. Beutler and W. L. Root. The operator pseudoinverse in control and systems identification. *Generalized Inverses and Applications*, pages 397–494, 1976.

[46] F.Riesz and B.Sz.-Nagy. *Functional Analysis.* Ungar, 1955.

[47] D. G. Luenberger. *Optimization by Vector Space Methods.* John Wiley & Sons,Inc., 1969.

[48] K. J. Åström, P. Hagander, and J.Sternby. Zeros of sampled systems. *Automatica*, 20(1):31–38, 1984.

[49] T. Hagiwara. Analytic study on the intrinsic zeros of sampled-data systems. *IEEE Transactions on Automatic Control*, 41(2):261–263, 1996.

[50] M. Togai and O. Yamano. Analysis and design of an optimal learning control scheme for industrial robots: a discrete system approach. *Proceedings of 24th IEEE Conference on Decision and Control*, pages 1399–1404, 1985.

[51] Y. Nishida, T. Sugie, and T. Ono. A digital learning control considering input saturation and disturbances. *Transactions of the Institute of Systems, Control and Information Engineers*, 2(3):80–87, 1989. (In Japanese).

[52] M. Yamakita and K. Furuta. Generation of virtual reference for discrete system by learning. *Transactions of the Society of Instrument and Control Engineers*, 25(8):867–873, 1989. (In Japanese).

[53] B. Porter and S. S. Mohamed. Digital iterative learning control of linear multivariable plants. *INT. J. SYSTEMS. SCI.*, 23(9):1393–1401, 1992.

[54] B. Porter and S. S. Mohamed. Digital iterative learning control of compensated linear multivariable plants. *INT. J. SYSTEMS. SCI.*, 23(11):1793–1804, 1992.

[55] S. S. Saab. A discrete-time learning control algorithm. *Proceedings of the American Control Conference*, pages 749–753, 1994.

[56] M. Q. Phan and J. A. Frueh. Learning control for trajectory tracking using basis functions. *Proceedings of the 35th IEEE Conference on Decision and Control*, pages 2490–2492, 1996.

[57] Walter Rudin. *Real and Complex Analysis.* McGraw–Hill, 1966.

[58] A. Katbab and E. I. Jury. On the strictly positive realness of schur interval functions. *IEEE Transactions on Automatic Control*, 35(12):1382–1385, 1990.

[59] Yoan D. Landau. *Adaptive Control.* Marcel Dekker,Inc., 1979.

Published Papers by the Author

Chapter 2:

- T. Sogo and N. Adachi. A design method of the learning control algorithm for linear systems. *Proc. 1994 Japan-U.S.A. Symposium on Flexible Automation*, pp. 681–684, July 1994

- T. Sogo and N. Adachi. On the learning control scheme for linear systems. *Transactions of the Institute of Systems, Control and Information Engineers*, Vol. 7, No. 9, pp. 339–346, 1994. (In Japanese).

- T. Sogo and N. Adachi. A method to design iterative learning control algorithm for linear systems. *Transactions of Instrument and Control Engineers*, Vol. 31, No. 10, pp. 1601–1607, 1995. (In Japanese).

Chapter 3:

- T. Sogo and N. Adachi. A gradient-type learning control algorithm for linear systems. *Proceedings of 1994 Asian Control Conference*, pp. 227–230, July 1994

Chapter 4:

- T. Sogo and N. Adachi. Convergence rates and robustness of iterative learning control. *Proceedings of the 35th IEEE Conference on Decision and Control*, pp. 3050–3055, 1996

- T. Sogo and N. Adachi. On properties of limiting zeros of sampled-data systems on finite time domains and the inverse systems. *Transactions of Instrument and Control Engineers*, Vol. 34. No. 10, pp. 1395–1403, 1998 (In Japanese).

- T. Sogo and N. Adachi, A limiting property of the inverse of sampled-data systems on a finite time interval, *Proceedings of the 37th IEEE Conference on Decision and Control*, pp. 835–836, 1998

Chapter 5:

- T. Sogo and N. Adachi. Iterative Learning Control Based on the Gradient Method for Linear Discrete-Time Systems *Transactions of the Institute of Systems, Control and Information Engineers*, Vol. 12, No. 4, 1999. (to appear; in Japanese).

- T. Sogo and N. Adachi, Iterative Learning Control Based on the Gradient Method for Linear Discrete-time Systems, *Proceedings of the 14th IFAC World Congress*, July 1999 (to appear).