Bettering Operation of Robots by Learning

Suguru Arimoto, Sadao Kawamura, and Fumio Miyazaki
Faculty of Engineering Science, Osaka University, Toyonaka, Osaka, 560 Japan
Received January 26, 1984; accepted March 12, 1984

This article proposes a betterment process for the operation of a mechanical robot in a sense that it betters the next operation of a robot by using the previous operation's data. The process has an iterative learning structure such that the $(k+1)$th input to joint actuators consists of the kth input plus an error increment composed of the derivative difference between the kth motion trajectory and the given desired motion trajectory. The convergence of the process to the desired motion trajectory is assured under some reasonable conditions. Numerical results by computer simulation are presented to show the effectiveness of the proposed learning scheme.

前回の作動データを用い次回の作動を改善する方式を用いた、ロボット機械系の作動の改善について述べる。この方式は反復学習構造を持ち、関節アクチュエーターへのＫ＋１番目の入力は、Ｋ番目の入力及び軌跡作動の要求値と実値との差の両者でもって決定される。ある妥当なる条件の下でのこの方式の収束性についての確認が出来た。ここで提案された方式の有効性を実証するため、シュミレーションの結果を示す。

I. INTRODUCTION

It is human to make mistakes, but it is also human to learn much from experience. Athletes have improved their form of body motion by learning through repeated training, and skilled hands have mastered the operation of machines or plants by acquiring skill in practice and gaining knowledge from experience. Upon reflection, can machines or robots learn autonomously (without the help of human beings) from measurement data of previous operations and make better their performance of future operations? Is it possible to think of a way to implement such a learning ability in the automatic operation of dynamic systems? If there is a way, it must be applicable to affording autonomy and intelligence to industrial robots.

Motivated by this consideration, we propose a practical approach to the problem of bettering the present operation of mechanical robots by using the data of

 CCC 0471-2223/84/020123-18$04.00

previous operations. We construct an iterative betterment process for the dynamics of robots so that the trajectory of their motion approaches asymptotically a given desired trajectory as the number of operation trials increases.

In relation to this problem, we should point out that a learning machine called the "perception"[1] was investigated almost twenty years ago in the field of pattern recognition and some related methods of learning were proposed in the field of control engineering,[2] stimulated by the introduction of the perceptron. However, most of the past literature made nothing of the underlying dynamical structures of treated systems, or proposed learning processes had nothing to do with dynamics, since the "perceptron" was a static machine. As opposed to those static learning processes, our proposed learning scheme uses directly the underlying dynamics of object systems.

In the next section we first introduce a betterment process with iterative learning structure for a linear servomechanism system with single degree of freedom. To see the convergence of the iterative process and its speed, a computer simulation result is given. In Section III the proposed scheme of the betterment process is extended to a class of general linear time-varying dynamical systems, in order to cope with nonlinear dynamics of many degrees of freedom. A mathematically rigorous proof of the convergence for this general class of the iterative betterment process will be given in the Appendix. In Section IV we discuss the dynamics of serial-link-type robot manipulators and formulate a linear time-varying differential equation that governs the motion of manipulators around a given desired motion trajectory. In Section V we discuss the problem of applying the betterment process to the learning control of manipulators and use a few numerical examples that show the possibility of the application.

Some partial results on practical applications of the proposed learning scheme for actual robot manipulators, together with some numerical simulation results, will be given in a subsequent paper.[3]

II. PROPOSAL OF BETTERMENT PROCESS

First we consider a linear servomechanism system with a single degree of freedom, which consists of a feedback control (as shown in Fig. 1) for the speed control of a voltage-controlled dc servomotor (as shown in Fig. 2). It is well known that if the armature inductance is sufficiently small and mechanical fric-

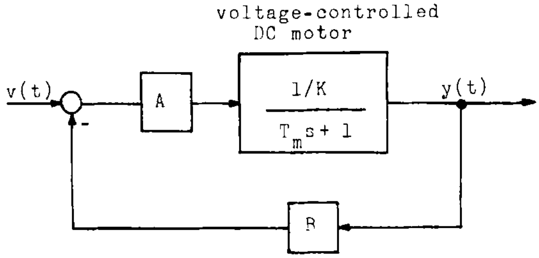

Figure 1. Feedback control system for speed control of a dc motor.

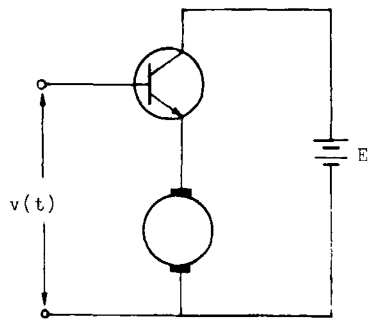

Figure 2. Voltage-controlled dc servomotor.

tion is ignored, the dynamics of the motor in Figure 2 is expressed by a linear differential equation

$$T_m \frac{d}{dt} y(t) + y(t) = v(t)/K, \tag{1}$$

where y and v denote the angular velocity of the motor and the input voltage, respectively, and K and T_m denote the torque constant and the time constant of the motor, respectively. Then, the closed-loop system shown in Figure 1 is subject to the following differential equation:

$$T_m \dot{y} + y = A(v - By)/K, \tag{2}$$

where $\dot{y} = dy/dt$. This can be represented in general form

$$\dot{y} + ay = bv, \tag{3}$$

where we set

$$a = (1 + AB/K)/T_m, \qquad b = A/KT_m.$$

It should be noted that the solution of Eq. (3) is expressed in integral form:

$$y(t) = e^{-at} y(0) + \int_0^t b e^{-a(t-\tau)} v(\tau) d\tau. \tag{4}$$

Now we suppose that a desired time evolution $y_d(t)$ of angular velocity is given for a fixed finite time interval $t \in [0, T]$. Throughout this article, we implicitly assume that the description of the system dynamics is unknown, that is, exact values of a and b in Eq. (3) or (4) need not be known. It is also assumed implicitly that $y_d(t)$ is continuously differentiable and sufficiently smooth so that every value of $y_d(t)$ can be reproduced approximately with high accuracy by sampled and digitized data $\{y_d(k\Delta t), k = 0, 1, \ldots, N = T/\Delta t\}$, which could be stored on a VLSI RAM memory. We consider the problem of finding an input voltage $v(t)$ whose response is coincident with given $y_d(t)$ over the interval $t \in [0, T]$. To find such an ideal input, we first choose an arbitrary voltage function $v_0(t)$ and try to supply the motor with this input voltage. An error btween the desired output $y_d(t)$ and the first response $y_0(t)$ to supplied input $v_0(t)$ may arise. This error is expressed as

$$e_0(t) = y_d(t) - y_0(t), \tag{5}$$

where

$$\dot{y}_0(t) + ay_0(t) = bv_0(t). \tag{6}$$

If sufficiently densely sampled data on $e_0(t)$ over $t \in [0, T]$ are stored on another RAM memory, it is possible to produce a next input voltage of the form

$$v_1(t) = v_0(t) + b^{-1}\dot{e}_0(t). \tag{7}$$

Next we apply this input to the motor in the second trial of operation, store the second resultant error $e_1(t)$, and so forth, which results in the following iterative process:

$$\begin{aligned} v_{k+1}(t) &= v_k(t) + b^{-1}\dot{e}_k(t), \\ e_k(t) &= y_d(t) - y_k(t), \\ y_k(t) &= e^{-at}y_k(0) + \int_0^t be^{-a(t-\tau)}v_k(\tau)d\tau. \end{aligned} \tag{8}$$

The process of this scheme is depicted in Figure 3. We call this iterative pocess a betterment process because the error $e_k(t)$ vanishes rapidly as the number of trials increases and, moreover, for a norm $\|e_k\|$ it becomes less than for the previous value $\|e_{k-1}\|$. In fact, if $y_0(0) = y_d(0) = y_k(0)$ for any k, then, in view of Eq. (8), the error is described by

$$\begin{aligned} \dot{e}_k(t) &= \dot{y}_d(t) - \dot{y}_k(t) \\ &= \dot{y}_d(t) - \left[-ae^{-at}y_k(0) + bv_k(t) - \int_0^t abe^{-a(t-\tau)}v_k(\tau)d\tau\right] \\ &= \dot{y}_d(t) - \left[-ae^{-at}y_{k-1}(0) + bv_{k-1}(t) - \int_0^t abe^{-a(t-\tau)}v_{k-1}(\tau)d\tau\right] \\ &\qquad - \left[\dot{e}_{k-1}(t) - \int_0^t ae^{-a(t-\tau)}\dot{e}_{k-1}(\tau)d\tau\right] \\ &= \dot{y}_d(t) - \dot{y}_{k-1}(t) - \dot{e}_{k-1}(t) + \int_0^t ae^{-a(t-\tau)}\dot{e}_{k-1}(\tau)d\tau \\ &= a\int_0^t e^{-a(t-\tau)}\dot{e}_{k-1}(\tau)d\tau \end{aligned} \tag{9}$$

which, in general, implies

$$\begin{aligned} |\dot{e}_k(t)| &\le |a|^2\int_0^t dt_1\int_0^{t_1} e^{-a(t-t_2)}|\dot{e}_{k-2}(t_2)|\,dt_2 \\ &\le |a|^k\int_0^t\int_0^{t_1}\cdots\int_0^{t_{k-1}}|\dot{e}_0(t_k)|\,dt_k\,dt_{k-1}\cdots dt_1 \\ &\le \frac{K|a|^kT^k}{k!} \to 0 \qquad \text{as } k \to \infty, \end{aligned} \tag{10}$$

where we set $K = \max_{0\le t\le T}|\dot{e}_0(t)|$.

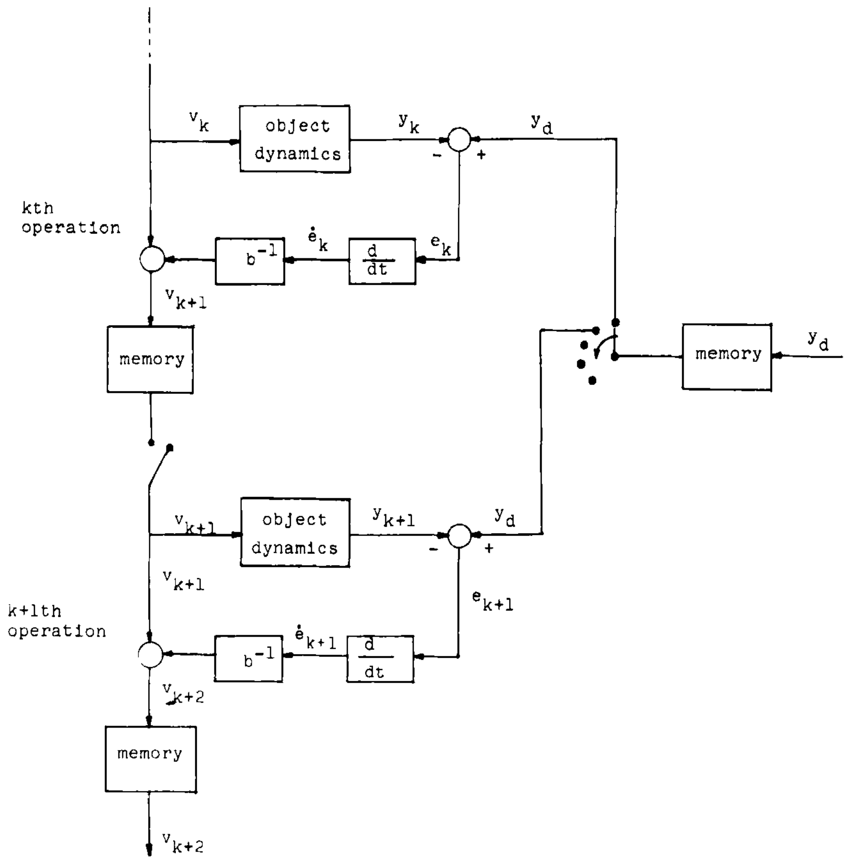

Figure 3. Iterative betterment process.

To see the speed of convergence for this betterment process a numerical example is shown in Figure 4, in which $a = b = 1$, $y_d(t) \equiv t(1-t)$ over $t \in [0, 1]$, $u_0(t) \equiv 0$, and $y_d(0) = y_k(0)$.

It should be noted that, instead of the exact value of b used in the betterment process defined by Eq. (8), its arbitrary approximate value γ can be used, provided the following condition is satisfied:

$$|1 - b\gamma^{-1}| < 1. \tag{11}$$

This generalization and further extended treatment will be discussed in the next section.

III. BETTERMENT PROCESS FOR GENERAL TIME-VARYING SYSTEMS

Consider a general linear time-varying dynamical system whose input–output relation is described by

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_0^t H(t, \tau)\mathbf{u}(\tau)d\tau, \tag{12}$$

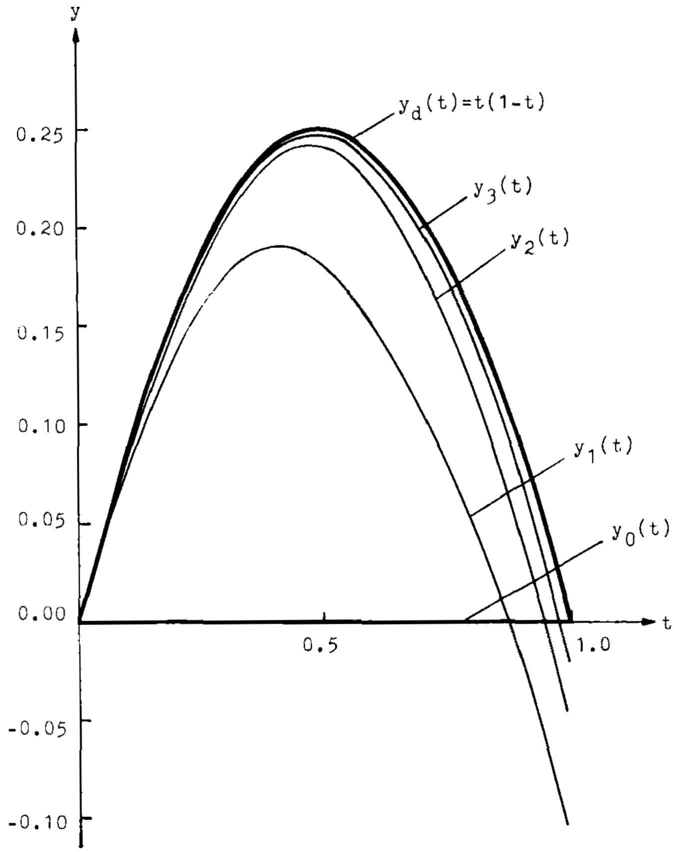

Figure 4. Convergence of betterment process.

where $\mathbf{u}$ denotes the input vector, $\mathbf{y}$ the output vector, and $\mathbf{g}(t)$ a given fixed vector such that $\mathbf{u}, \mathbf{y}, \mathbf{g} \in R^r$ and hence $H(t, \tau) \in R^{r \times r}$.

Suppose that a desired trajectory of system output $\mathbf{y}_d(t)$ is given over a fixed finite interval $t \in [0, T]$ and an initial input function $\mathbf{u}_0(t)$ is arbitrarily chosen. We assume that $\mathbf{y}_d(t)$, $\mathbf{u}(t)$, $\mathbf{g}(t)$, and $H(t, \tau)$ are continuously differentiable in t and τ.

Now we construct an iterative betterment process for system Eq. (12) as follows:

$$\begin{aligned} \mathbf{y}_k(t) &= \mathbf{g}(t) + \int_0^t H(t, \tau)\mathbf{u}_k(\tau)d\tau, \\ \mathbf{e}_k(t) &= \mathbf{y}_d(t) - \mathbf{y}_k(t), \\ \mathbf{u}_{k+1}(t) &= \mathbf{u}_k(t) + \Gamma(t)\dot{\mathbf{e}}_k(t), \end{aligned} \tag{13}$$

where $\Gamma(t)$ is a given $r \times r$ matrix function that is easily programmable or reproducible from a RAM memory. This process is depicted in Figure 5.

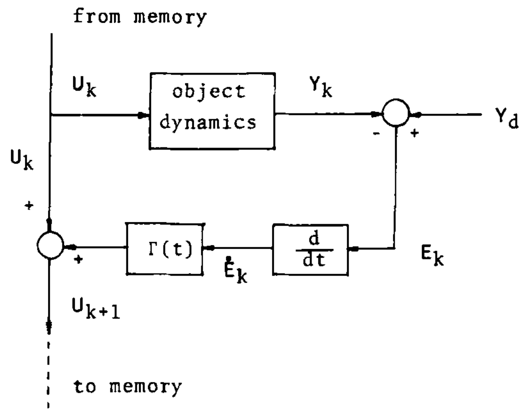

Figure 5. Betterment process for a general time-varying dynamical system.

To see the effect of betterment for the iteration (13), it is necessary to introduce a vector norm for r-vector-valued functions $\mathbf{e}(t)$ defined on $[0, T]$ as follows:

$$\|\mathbf{e}\|_\lambda = \sup_{0\le t\le T} \{e^{-\lambda t} \max_{1\le i\le r} |e_i(t)|\}, \tag{14}$$

where λ is a positive constant. Then, for the betterment process defined by Eq. (13), the following theorem holds:

THEOREM 1: If $\mathbf{y}_d(0) = \mathbf{g}(0)$, $\|I_r - H(t, t)\Gamma(t)\|_\infty < 1$ for all $t \in [0, T]$, and a given initial input $\mathbf{u}_0(t)$ is continuous on $[0, T]$, then there exist positive constants λ and ρ_0 such that

$$\|\dot{\mathbf{e}}_{k+1}\|_\lambda \le \rho_0 \|\dot{\mathbf{e}}_k\|_\lambda \quad \text{and} \quad 0 \le \rho_0 < 1 \tag{15}$$

for $k = 0, 1, 2, \ldots$, where the matrix norm $\|G\|_\infty$ for $r \times r$ matrix $G = (g_{ij})$ stands for

$$\|G\|_\infty = \max_{1\le i\le r} \left\{ \sum_{j=1}^{r} |g_{ij}| \right\}.$$

The proof will be given in the Appendix. It follows from this theorem that

$$\|\dot{\mathbf{e}}_k(t)\|_\lambda \to 0 \quad \text{as } k \to \infty, \tag{16}$$

and hence

$$\dot{\mathbf{y}}_k(t) \to \dot{\mathbf{y}}_d(t) \quad \text{as } k \to \infty. \tag{17}$$

It is clear by definition of the norm $\|\cdot\|_\lambda$ that these convergences are uniform in $t \in [0, T]$. Moreover, on account of the fact that

$$\mathbf{y}_k(0) = \mathbf{g}(0) = \mathbf{y}_d(0),$$

Eq. (17) implies that

$$\mathbf{y}_k(t) \to \mathbf{y}_d(t) \quad \text{as } k \to \infty \tag{18}$$

uniformly in $t \in [0, T]$.

Finally, we remark that linear time-varying systems described by

$$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t) + B(t)\mathbf{u}(t), \tag{19a}$$

$$\mathbf{y}(t) = C(t)\mathbf{x}(t), \qquad \mathbf{x} \in R^n,\ \mathbf{u}, \mathbf{y} \in R^r \tag{19b}$$

are included in the class of dynamical systems described by Eq. (12). In fact, it is well known that Eq. (19) yields

$$\mathbf{y}(t) = C(t)X(t)\mathbf{x}(0) + \int_0^t C(t)X(t)X^{-1}(\tau)B(\tau)\mathbf{u}(\tau)d\tau, \tag{20}$$

where $X(t)$ is a unique matrix solution to the homogeneous matrix differential equation

$$\dot{X}(t) = A(t)X(t), \qquad X(0) = I_n. \tag{21}$$

Evidently, Eq. (20) is reduced to Eq. (12) by setting $\mathbf{g}(t) = C(t)X(t)\mathbf{x}(0)$ and $H(t, \tau) = C(t)X(t)X^{-1}(\tau)B(\tau)$. Hence, if

$$B(t) \equiv B, \qquad C(t) \equiv C,$$

and CB is nonsingular, then there exists a constant matrix such that

$$\| I_r - H(t, t)\Gamma \|_\infty = \| I_r - BC\Gamma \|_\infty < 1. \tag{22}$$

In conclusion, the betterment process for the system (19) with above conditions is convergent in a sense that $\mathbf{y}_k(t) \to \mathbf{y}_d(t)$ uniformly in $t \in [0, T]$ as $k \to \infty$.

IV. DYNAMICS OF ROBOT MANIPULATORS

Now we consider the dynamics of a serial-link manipulator with n degrees of freedom like the one shown in Figure 6. It is well known (see Ref. 4 or 5) that its dynamics can be expressed generally by the form

$$R(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + \mathbf{f}(\dot{\boldsymbol{\theta}}, \boldsymbol{\theta}) + \mathbf{g}(\boldsymbol{\theta}) = \boldsymbol{\tau}, \tag{23}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ denotes the joint angle coordinates as shown in Figure 6, and $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)$ denotes a generalized force vector. $R(\boldsymbol{\theta})$ is called an inertia matrix and is usually positive definite. The term $\mathbf{g}(\boldsymbol{\theta})$ comes from the potential energy of the manipulator system and has a simpler form than the inertia matrix. The term $\mathbf{f}(\dot{\boldsymbol{\theta}}, \boldsymbol{\theta})$ consists of centrifugal and Coriolis forces and other nonlinear characteristics such as frictional forces.

The task is usually described in terms of Cartesian coordinates or other task-oriented coordinates. However, we assume in this article that a desired motion of the manipulator is determined from a description of the task as a time function $\boldsymbol{\theta}_d(t)$ over $t \in [0, T]$ for joint angle coordinates $\boldsymbol{\theta}$. An extension to the treatment based on the task-oriented coordinates will be discussed in our subsequent paper.[3]

Now we apply a linear PD feedback control law to the manipulator, which we proposed in our previous papers[6-8] where we proved its effectiveness as a positioning control of manipulators. The control input is given by

$$\boldsymbol{\tau} = \mathbf{g}(\boldsymbol{\theta}) + A(\boldsymbol{\theta}_d - \boldsymbol{\theta}) + B(\dot{\boldsymbol{\theta}}_d - \dot{\boldsymbol{\theta}}), \tag{24}$$

where A and B are constant matrices. By using this control method with appropriate gain matrices, the motion of the manipulator usually follows in the neighborhood of a given desired trajectory $\boldsymbol{\theta}_d(t)$ in $\boldsymbol{\theta}$ space. However, there is a small difference between these two trajectories, which we denote

$$\mathbf{x}(t) = \boldsymbol{\theta}(t) - \boldsymbol{\theta}_d(t). \tag{25}$$

Since $\mathbf{x}(t)$ is small, it is possible to rewrite Eq. (23) in $\mathbf{x}$ as follows:

$$R[\boldsymbol{\theta}_d(t)]\ddot{\mathbf{x}}(t) + [B + C(t)]\dot{\mathbf{x}}(t) + [A + D(t)]\mathbf{x}(t) = \mathbf{h}(t) + \mathbf{u}(t), \tag{26}$$

where $\mathbf{u}(t)$ is an additional control input to further improve the motion of the manipulator. Since $R[\boldsymbol{\theta}_d(t)]$ is always positive definite, it is possible to multiply both sides of Eq. (26) from the left by $R^{-1}[\boldsymbol{\theta}_d(t)]$, which results in

$$\ddot{\mathbf{x}}(t) + C_1(t)\dot{\mathbf{x}}(t) + C_2(t)\mathbf{x}(t) = \mathbf{g}_1(t) + B_1(t)\mathbf{u}(t). \tag{27}$$

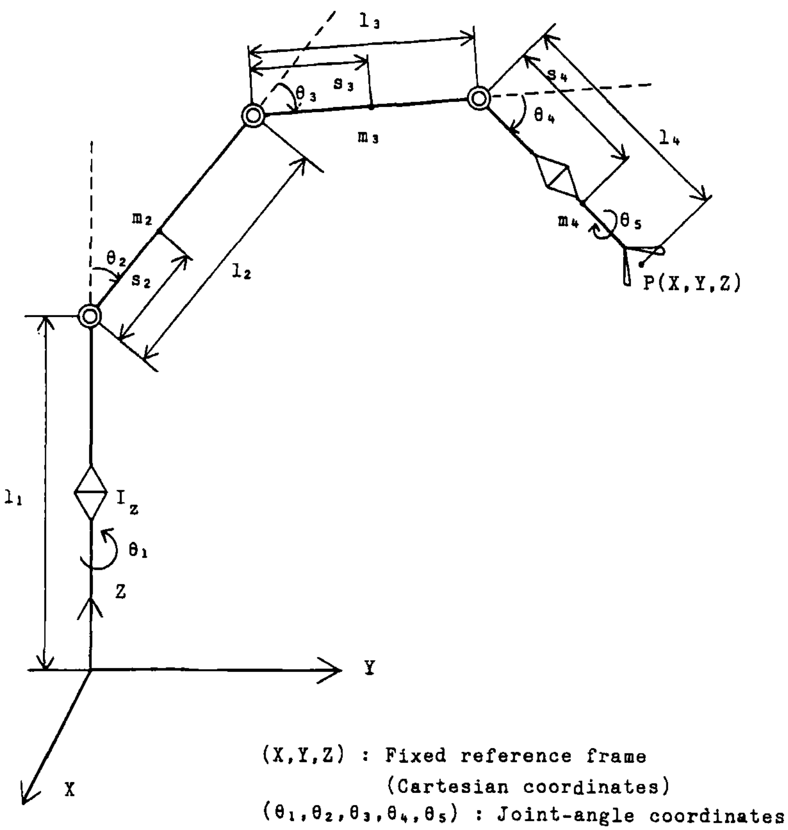

Figure 6. Serial-link manipulator with 5 degrees of freedom.

This can be rewritten to yield the following state equation:

$$\frac{d}{dt}\begin{bmatrix}\mathbf{x}(t)\\ \dot{\mathbf{x}}(t)\end{bmatrix}=\begin{bmatrix}0 & \mathrm{I}\\ -C_2(t) & -C_1(t)\end{bmatrix}\begin{bmatrix}\mathbf{x}(t)\\ \dot{\mathbf{x}}(t)\end{bmatrix}+\begin{bmatrix}0\\ \mathbf{g}_1(t)\end{bmatrix}+\begin{bmatrix}0\\ B_1(t)\end{bmatrix}\mathbf{u}(t). \tag{28}$$

At this stage, we must bear in mind that Eq. (28) itself is a linear time-varying dynamical system, like Eq. (19a) [except for the forcing term $\mathbf{g}_1(t)$ in Eq. (28)], though all coefficient time-varying functions together with $\mathbf{g}_1(t)$ cannot be evaluated exactly. However, if we choose an adequate output vector $\mathbf{y}$ for system (28) and a gain matrix $\Gamma(t)$ in order to satisfy the assumption of Theorem 1, then it is possible to apply the betterment process proposed in the previous section. It should be noted again that in the application of such a betterment process all time-varying functions of the coefficients in the linear dynamical system [Eqs. (19a) and (19b)] need not be known in principle, but only the output $\mathbf{y}_k(t)$ for each operation trial must be known.

V. APPLICATION OF THE BETTERMENT PROCESS TO THE CONTROL OF MANIPULATORS

To construct a betterment process for system (28), we choose the output vector as

$$\mathbf{y}(t)=[0\ I_n]\begin{bmatrix}\mathbf{x}(t)\\ \dot{\mathbf{x}}(t)\end{bmatrix}=\dot{\mathbf{x}}(t) \tag{29}$$

and introduce an iterative process

$$\mathbf{u}_{k+1}(t)=\mathbf{u}_k(t)+\Gamma(t)[\dot{\mathbf{y}}_d(t)-\dot{\mathbf{y}}_k(t)], \tag{30}$$

where $\Gamma(t)$ is an $n\times n$ time-varying or constant matrix and $\mathbf{y}_k(t)$ is the output response to the control input $\mathbf{u}_k(t)$. Next we note that the solution of Eq. (28) is expressed by

$$\begin{aligned}\mathbf{x}(t)&=\bar{\mathbf{g}}_1(t)+\int_0^t H_{12}(t,\tau)B_1(\tau)\mathbf{u}(\tau)d\tau,\\ \dot{\mathbf{x}}(t)&=\bar{\mathbf{g}}_2(t)+\int_0^t H_{22}(t,\tau)B_1(\tau)\mathbf{u}(\tau)d\tau,\end{aligned} \tag{31}$$

where we put

$$X(t)X^{-1}(\tau)=\begin{array}{c}\begin{array}{cc} n & n\end{array}\\ \begin{bmatrix}H_{11}(t,\tau) & H_{12}(t,\tau)\\ H_{21}(t,\tau) & H_{22}(t,\tau)\end{bmatrix}\end{array}\begin{array}{c} \\ n\\ n\end{array}.$$

Hence, $X(t)$ is a fundamental solution to a homogeneous $2n\times 2n$ matrix differential equation

$$\frac{d}{dt}X(t)=\begin{bmatrix}0 & I_n\\ -C_2(t) & -C_1(t)\end{bmatrix}X(t),$$

where

$$X(t)=\begin{bmatrix} X_{11}(t) & X_{12}(t) \\ X_{21}(t) & X_{22}(t) \end{bmatrix}, \qquad X(0)=\begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix}.$$

It should be noted that

$$\begin{aligned} \tilde{\mathbf{g}}_1(t) &= X_{11}(t)\mathbf{x}(0) + X_{12}(t)\dot{\mathbf{x}}(0) + \int_0^t H_{12}(t,\tau)\mathbf{g}_1(\tau)d\tau, \\ \tilde{\mathbf{g}}_2(t) &= X_{21}(t)\mathbf{x}(0) + X_{22}(t)\dot{\mathbf{x}}(0) + \int_0^t H_{22}(t,\tau)\mathbf{g}_1(\tau)d\tau. \end{aligned} \tag{32}$$

This means that Eq. (28) reduces to a form of Eq. (12) even if a forcing term $g_1(t)$ exists in Eq. (28).

Now we check the validity of the most important assumption in Theorem 1, which is described in this case by

$$\begin{aligned} \| I_r - H(t,t)\Gamma(t)\|_\infty &= \| I_n - [0 \; I_n]X(t)X^{-1}(t)\begin{bmatrix} 0 \\ B_1(t)\end{bmatrix}\Gamma(t)\|_\infty \\ &= \| I_n - B_1(t)\Gamma(t)\|_\infty . \end{aligned} \tag{33}$$

As seen in the derivation of Eq. (27) from Eq. (26), $B_1(t)=R^{-1}[\boldsymbol{\theta}_d(t)]$. Hence, if we set $\Gamma(t)=R[\boldsymbol{\theta}_d(t)]$, the assumption of Theorem 1 is clearly satisfied with the condition

$$\begin{aligned} \| I_n - H(t,t)\Gamma(t)\|_\infty &= \| I_n - R^{-1}[\boldsymbol{\theta}_d(t)]R[\boldsymbol{\theta}_d(t)]\|_\infty \\ &= \| I_n - I_n\|_\infty = 0 < 1. \end{aligned}$$

In general, most entries of the inertia matrix $R(\boldsymbol{\theta})$ are complicated functions of $\boldsymbol{\theta}$ and therefore it is difficult to compute exact values of $R[\boldsymbol{\theta}_d(t)]$ in real time. However, it is true that $R(\boldsymbol{\theta})$ remains positive definite for any $\boldsymbol{\theta}$. This suggests that in most configurations of the manipulator there is the possibility of selecting an appropriate constant matrix Γ such that

$$\| I_n - B_1(t)\Gamma\|_\infty = \| I_n - R^{-1}(\boldsymbol{\theta}_d)\Gamma\|_\infty < 1.$$

Finally, we remark that, from the original problem of controlling the manipulator posed in the previous section, a desired trajectory $\mathbf{y}_d(t)$ for the variational system should be given as $\mathbf{y}_d(t) \equiv 0$ because $\mathbf{x}(t)$ represents the small difference of $\boldsymbol{\theta}(t)$ with the original desired motion $\boldsymbol{\theta}_d(t)$, that is,

$$\mathbf{x}(t) = \boldsymbol{\theta}(t) - \boldsymbol{\theta}_d(t).$$

Hence, if it holds that $\dot{\mathbf{x}}_k(0) = \mathbf{y}_d(0) = 0$ for all k, then all assumptions of Theorem 1 are satisfied and thereby the betterment process converges. Fortunately, the initial condition $\dot{\mathbf{x}}(0) = \mathbf{0}$ is usually satisfied if the manipulator is still at the beginning of every operation trial, that is, $\dot{\boldsymbol{\theta}}_k(0) = \mathbf{0}$, and $\boldsymbol{\theta}_d(t)$ is chosen to satisfy the initial condition $\dot{\boldsymbol{\theta}}_d(0) = \mathbf{0}$.

In conclusion, it is shown that the betterment process of Eq. (30) for the system described by Eq. (28) and (29) becomes convergent if an appropriate constant $n \times n$ matrix Γ is chosen and the initial condition $\dot{\boldsymbol{\theta}}_d(0) = \dot{\boldsymbol{\theta}}_k(0) = \mathbf{0}$ is satisfied in every operation trial.

To see the applicability of this scheme, we show numerical simulation results in Figures 7(a), 7(b), and 7(c), and Figures 8(a), 8(b), and 8(c), which are obtained on the basis of the following system:

$$\ddot{x} + 2\dot{x} + x = u \tag{34}$$

or

$$\frac{d}{dt}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \tag{35}$$

with output process

$$y = [0 \quad 1]\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \dot{x} \tag{36}$$

and betterment process

$$u_{k+1} = u_k + \Gamma(\dot{y}_d - \dot{y}_k). \tag{37}$$

In Figures 7(a), 7(b), and 7(c), we set

$$\Gamma = 1.0, \qquad y_d(t) = 12t^2(1-t), \tag{38}$$

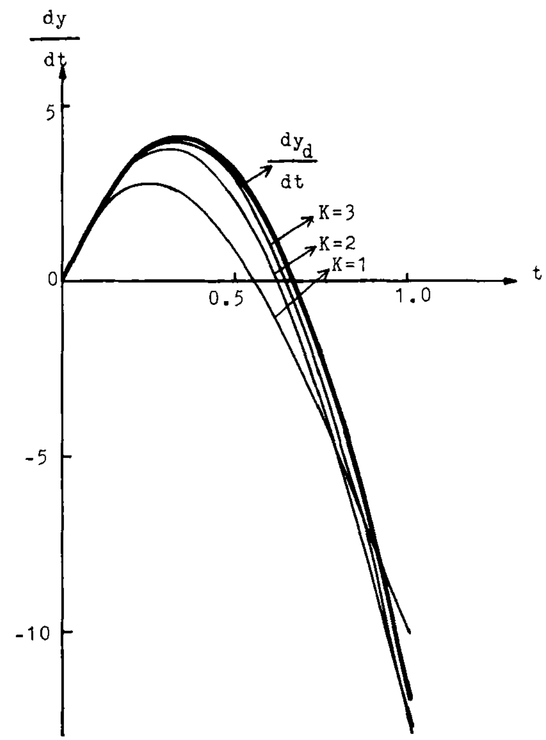

Figure 7(a). Plot of $\dot{y}_k(t) = \ddot{x}_k(t)$ (acceleration) for betterment process with $\Gamma = 1.0$.

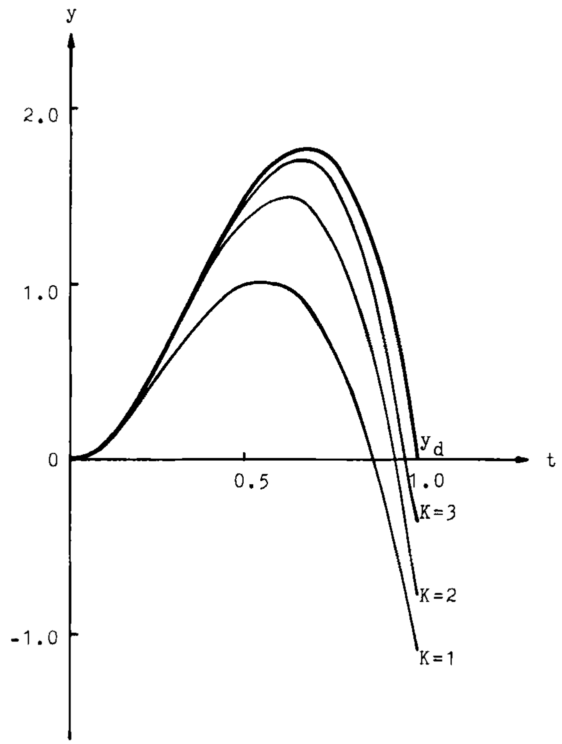

Figure 7(b). Plot of $y_k(t) = \dot{x}_k(t)$ (velocity) for betterment process with $\Gamma = 1.0$.

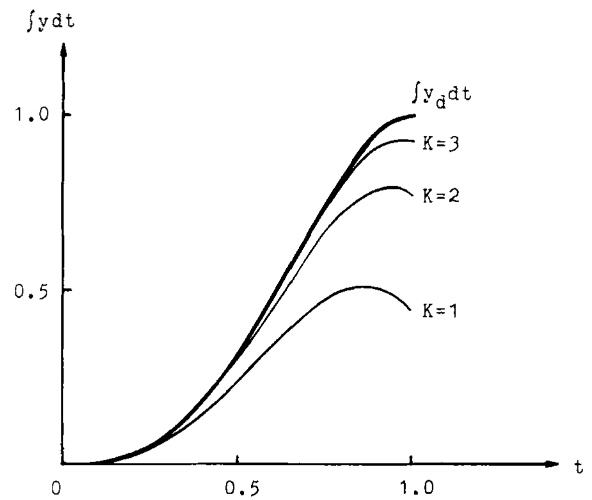

Figure 7(c). Plot of $\int_0^t y_k(\tau)d\tau = x_k(t)$ (position) for betterment process with $\Gamma = 1.0$.

and in Figures 8(a), 8(b), and 8(c),

$$\Gamma = 0.5, \qquad y_d(t) = 12t^2(1-t). \tag{39}$$

Comparing the three parts of Figure 7 with the corresponding parts of Figure 8, we see that the convergence for the case where $\Gamma = 1.0$ is faster than the convergence when $\Gamma = 0.5$, as could be theoretically expressed by comparing the proof of Theorem 1 in the Appendix with the proof of convergency for the system described by Eq. (8). In fact, the key assumption of Theorem 1 turns out to be

$$\| I_r - H(t,t)\Gamma(t) \|_\infty = 0$$

for the case of Eq. (38) and

$$\| I_r - H(t,t)\Gamma(t) \|_\infty = 0.5 < 1$$

for the case of Eq. (39). In either case, however, the speed of convergence is relatively high and the output approaches the desired trajectory with sufficient accuracy within several trials.

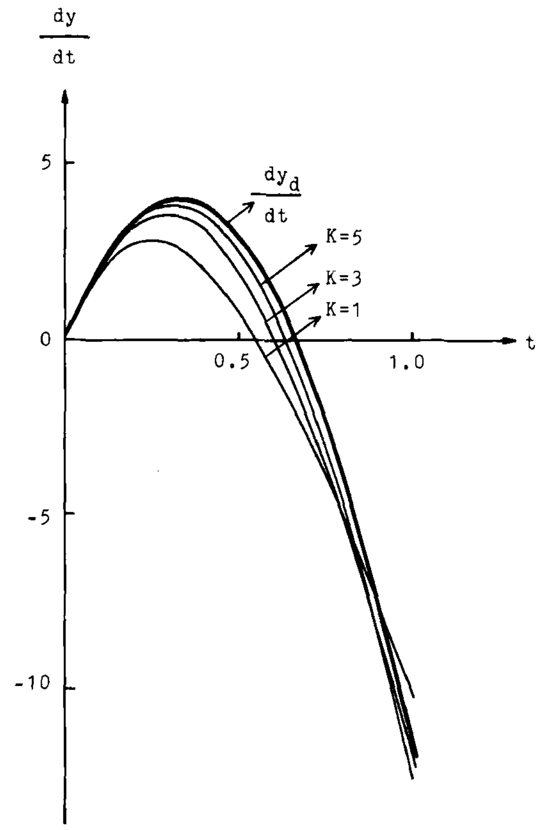

Figure 8(a). Plot of $\dot{y}_k(t) = \ddot{x}_k(t)$ (acceleration) for betterment process with $\Gamma = 0.5$.

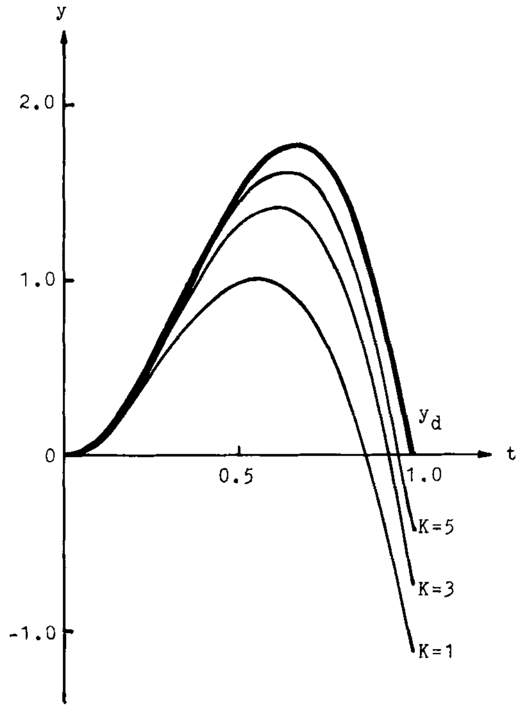

Figure 8(b). Plot of $y_k(t) = \dot{x}_k(t)$ (velocity) for betterment process with $\Gamma = 0.5$.

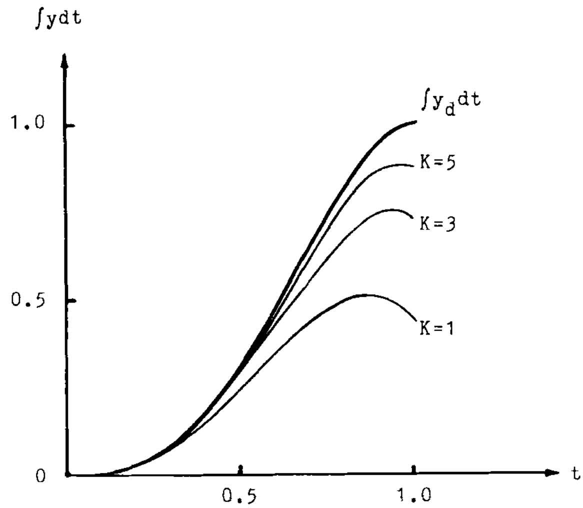

Figure 8(c). Plot of $\int_0^t y_k(\tau)d\tau = x_k(t)$ (position) for betterment process with $\Gamma = 0.5$.

Finally, we must remark that it is impossible to choose the positional variable $x(t)$ as an output. In this case, we write

$$y = [1\ 0]\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = x$$

which, together with system Eq. (35), yields

$$H(t, t) = [1\ 0]\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$

Hence, the key assumption of Theorem 1 is not satisfied for any choice of Γ.

VI. CONCLUSION

With the goal of applying "learning" control to robot manipulators, we proposed an iterative betterment process for mechanical systems that improves their present performance of motion by use of updated data of the previous operation. It was shown theoretically, together with numerical examples, that the betterment process can be applied to the motion control of manipulators if a desired motion trajectory is given.

In this article, the discussion is restricted to the case of linear dynamical systems and therefore the equation of motion of the manipulator had to be represented in the neighborhood of the desired trajectory. However, this representation was necessary only for the theoretical proof of the applicability of the betterment process to the control of robots. In practice, this iterative scheme could be associated directly with the original nonlinear dynamics of robots, that is, the actual betterment process is constructed as shown in Figure 9.

Finally, we point out that the betterment process for linear dynamical systems with constant coefficients has many interesting aspects in relation to linear system theory. It should also be pointed out that it is possible to prove the convergence of the iterative betterment process constructed for a class of nonlinear dynamical systems like robot manipulators. These discussions are given in subsequent papers.[9,10]

APPENDIX

Proof of Theorem 1

First we note that the norm $\| G \|_\infty$ is the natural matrix norm induced by the vector norm

$$\|\mathbf{f}\|_\infty = \max_{1 \le i \le r} |f_i|$$

for vector $\mathbf{f} = (f_1, \ldots, f_r)$. Hence, it is evident that

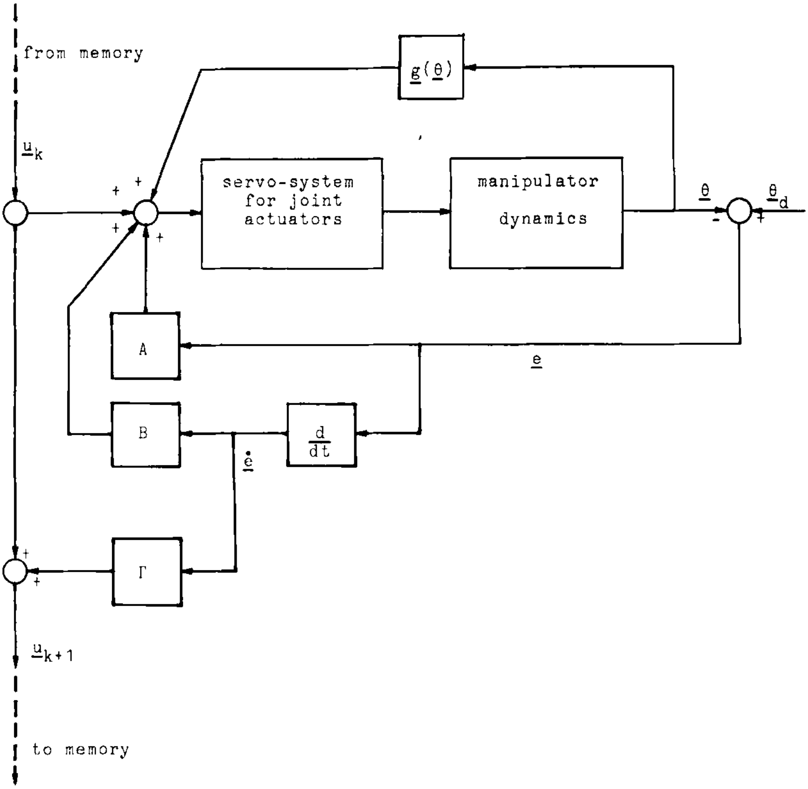

Figure 9. Learning control scheme by betterment process for robot manipulators.

$$\|G\mathbf{f}\|_\infty \leq \| G \|_\infty \cdot \| \mathbf{f} \|_\infty . \tag{A1}$$

Now it follows from Eq. (13) that

$$\begin{aligned}\dot{\mathbf{e}}_k(t) &= \dot{\mathbf{y}}_d(t) - \dot{\mathbf{y}}_k(t)\\ &= \dot{\mathbf{y}}_d(t) - [\dot{\mathbf{g}}(t) + H(t, t)\mathbf{u}_k(t) + \int_0^t \frac{\partial}{\partial t} H(t, \tau)\mathbf{u}_k(\tau)d\tau]\\ &= \dot{\mathbf{y}}_d(t) - [\dot{\mathbf{g}}(t) + H(t, t)\mathbf{u}_{k-1}(t) + \int_0^t \frac{\partial}{\partial t} H(t, \tau)\mathbf{u}_{k-1}(\tau)d\tau]\\ &\quad - [H(t, t)\Gamma(t)\dot{\mathbf{e}}_{k-1}(t) + \int_0^t \frac{\partial}{\partial t} H(t, \tau)\Gamma(\tau)\dot{\mathbf{e}}_{k-1}(\tau)d\tau]\\ &= [I_r - H(t, t)\Gamma(t)]\dot{\mathbf{e}}_{k-1}(t) - \int_0^t \frac{\partial}{\partial t} H(t, \tau)\Gamma(\tau)\dot{\mathbf{e}}_{k-1}(\tau)d\tau. \end{aligned} \tag{A2}$$

Multiplying both sides of this equation by $e^{-\lambda t}$, taking the norm $\|\cdot\|_\lambda$, and using the property of Eq. (A1), we have

$$\begin{aligned}
\|\dot{\mathbf{e}}_k\|_\lambda &= \max_{0\le t\le T}\{e^{-\lambda t}\|\dot{\mathbf{e}}_k(t)\|_\infty\} \\
&\le \max_{0\le t\le T}\|I_r - H(t,t)\Gamma(t)\|_\infty \cdot \max_{0\le t\le T}\{e^{-\lambda t}\|\dot{\mathbf{e}}_{k-1}(t)\|_\infty\} \\
&\quad + h_0 \cdot \max_{0\le t\le T}\int_0^t e^{-\lambda(t-\tau)} \max_{0\le \tau\le T}\{e^{-\lambda\tau}\|\dot{\mathbf{e}}_{k-1}(\tau)\|_\infty\} d\tau \\
&= \rho\|\dot{\mathbf{e}}_{k-1}\|_\lambda + h_0\|\dot{\mathbf{e}}_{k-1}\|_\lambda \cdot \max_{0\le t\le T}\int_0^t e^{-\lambda(t-\tau)} d\tau \\
&= \left(\rho + \frac{h_0(1-e^{-\lambda T})}{\lambda}\right)\|\dot{\mathbf{e}}_{k-1}\|_\lambda,
\end{aligned} \tag{A3}$$

where

$$\rho = \max_{0\le t\le T}\|I_r - H(t,t)\Gamma(t)\|_\infty,$$

$$h_0 = \max_{0\le t,\tau\le T}\|\frac{\partial}{\partial t}H(t,\tau)\Gamma(\tau)\|_\infty.$$

Since $0\le\rho<1$ by assumption, it is possible to choose λ sufficiently large so that

$$\rho_0 = \rho + \frac{h_0(1-e^{-\lambda T})}{\lambda} < 1. \tag{A4}$$

Substituting this into Eq. (A3), we obtain Eq. (15). Thus, Theorem 1 has been proved.

References

1. F. Rosenblatt, *Principles of Neurodynamics, Perceptrons and the Theory of Brain Mechanisms,* Spartan Books Co., Washington, DC, 1961.
2. Ya. Z. Tsypkin, *Adaption and Learning in Automatic Systems,* Academic, New York, 1971.
3. S. Arimoto, S. Kawamura, and F. Miyazaki, "Can mechanical robots learn by themselves?" *Proc. 2nd Int. Symp. Robotics Res.,* Kyoto, Japan, August 1984.
4. R. P. Paul, *Robot Manipulators,* The MIT Press, Cambridge, 1981.
5. P. Vukobratovic, *Scientific Fundamentals of Robotics 1 and 2,* Springer-Verlag, Berlin, 1982.
6. M. Takegaki and S. Arimoto, "A new feedback method for dynamic control of manipulators," *Trans. ASME J Dyn. Syst. Meas. Control,* **103**, 119–125 (1982).
7. S. Arimoto and F. Miyazaki, "Stability and robustness of PID feedback control for robot manipulators of sensory capability," *Proc. 1st Int. Symp. Robotics Res.,* The MIT Press, Cambridge, MA, 1983.
8. F. Miyazaki, S. Arimoto, M. Takegaki, and Y. Maeda, "Sensory feedback based on the artificial potential for robot manipulators," *Proc. 9th IFAC,* Budapest, Hungary, July 1984.
9. S. Kawamura, F. Miyazaki, and S. Arimoto, "Iterative learning control for robotic systems," paper presented at IECON '84, Tokyo, Japan, 1984.
10. S. Arimoto, S. Kawamura, F. Miyazaki, "Bettering operation of dynamic systems by learning; A new control theory for servomechanism or mechatronics systems," paper presented at 23rd IEEE Conf. Decision and Control, Las Vegas, NV, 1984.