Modification of Learning Optimal Gait Generation Method in Considering Discontinuous Velocity Transitions

Satoshi Satoh
Hiroshima University, 1-4-1, Kagamiyama,
Higashi-Hiroshima, 739-8527, JAPAN
Email: satoh@hiroshima-u.ac.jp

Masahito Ikeda, Kenji Fujimoto
and Yoshikazu Hayakawa
Nagoya University, Furo-cho, Chikusa-ku,
Nagoya-shi, Aichi, 464-8603, JAPAN
E-mail: ikeda@haya.nuem.nagoya-u.ac.jp,
E-mail: fujimoto@nagoya-u.jp,
E-mail: hayakawa@nuem.nagoya-u.ac.jp

***Abstract*—In this paper, we propose a modification of our previous learning gait generation method. Our framework can generate an optimal feedforward control input and the corresponding periodic trajectory minimizing the L_2 norm of the control input by iteration of laboratory experiments. In order to generate a periodic gait, the previous result imposed a constraint that the initial state of the robot is equivalent to the state just after the collision between the foot and the ground, by equipping a reference trajectory defined as a time-reversal version of a pair of the output signal and its time derivative. However, it occasionally happens that the reference velocity at the terminal time conflicts a desired terminal velocity for a periodic trajectory. This paper proposes a modified learning algorithm with another reference trajectory whose velocity coincides with the desired one. Although calculation of such reference trajectory generally requires information of the transition mapping, this method estimates the mapping by the least-squares with the stored experimental data. We also propose a technique to generate an optimal gait not only energy-efficient but also avoiding the foot-scuffing problem.**

Keywords*—*nonlinear control; Hamiltonian systems; iterative learning control; gait generation; passive dynamic walking.

I. Introduction

Recently, control of walking robots has become an active research area. As the technology for walking robots evolves, an optimization problem of gaits with respect to the energy consumption becomes increasingly important. Most of walking pattern generation and control methods have been based on the zero moment point (ZMP) criterion, e.g., [1], [2], [3]. Passive dynamic walker studied by McGeer [4] also attracts attention. Behavior analysis of passive walkers were studied by, e.g., [5], [6]. Walking control methods based on passive dynamic walking have been proposed by many researchers, e.g. [7], [8], [9], [10], [11], as is antithetical to the ZMP based control with respect to the energy consumption.

On the other hand, we consider that physical property and learning control are useful tools to tackle the optimal gait generation problem. In [12], [13], [14], we have proposed optimal gait generation framework via iterative learning control (ILC) proposed in [15], which utilizes a property of Hamiltonian systems called variational symmetry. Hamiltonian systems have been introduced to represent physical systems and they explicitly possess good properties for the control design such as passivity, symmetry and so on. Our technique can generate an optimal periodic gait which minimizes a cost function by iteration of laboratory experiments. The cost function consists of two terms: one attempts to minimize the L_2 norm of the control input, and the other attempts to make a trajectory periodic, which is a constraint term for a periodic gait. By taking advantage of variational symmetry of Hamiltonian systems, the proposed method does not require precise knowledge of the plant system. So far, in numerical simulations, we have generated an optimal running gait of a planer one-legged hopping robot [12] and optimal walking gaits of a planer compass-like biped robot [13] and one with a torso [14], respectively.

In this paper, we consider a kneed biped robot with torso and propose a modification of our previous learning gait generation method in considering discontinuous velocity transitions. The following are some reasons why we introduce knees. Firstly, we can investigate more general and human-like walking motions. Secondly, controlling knees so that the robot achieves foot clearance properly, enables us to avoid the foot scuffing problem [4], [16]. This problem is that, in the case of compass-like robot, the swing leg scuffs the ground when it passes the stance leg, and this phenomenon causes the robot to fall down. We consider a necessary condition for a periodic trajectory as that the state just after the collision between the swing leg and the ground coincides with the initial state. However, our previous works [13], [14] require information of the inertia matrix of the robot to deal with such a state constraint. Since it is sometimes difficult to obtain the accurate inertia information of robots, we assigned another constraint which has a similar effect and is available without the inertia information, by introducing a reference trajectory. The reference trajectory is defined on a time interval $[t^1 - \Delta t,\ t^1]$ as the time reversal of the output signal y and its time derivative $\dot{y}$ on $[t^0,\ t^0 + \Delta t]$, where t^0 and t^1 represent the initial time and the terminal time (just before the collision), respectively and Δt denotes a sufficiently small positive constant. The output y corresponds to the set of joint angles of the robot. However, it occasionally happens that the reference velocity at t^1, i.e. $\lim \Delta t \to 0(y(t^1) - y(t^1 - \Delta t))/\Delta t$, conflicts a desired terminal velocity for a periodic trajectory. In this paper, we propose a modified learning algorithm with another reference trajectory whose velocity coincides with the desired one for a periodic trajectory. Although calculation of such reference trajectory generally requires information of the transition mapping which maps from the velocity just before the collision to that just after the transition, the proposed method estimates the mapping by the least-squares with the stored experimental data. We also propose a technique to generate an optimal gait not only energy-efficient but also avoiding the foot-scuffing problem. Finally, numerical simulations demonstrate the validity of the proposed method.

II. Preliminaries

This section briefly refers to preliminary backgrounds.

A. Hamiltonian systems and variational symmetry

Our plant is a Hamiltonian system with dissipation $\Sigma^{x_{t^0}} : L_2^m[t^0, t^1] \to L_2^m[t^0, t^1] : u \mapsto y$ with a controlled Hamiltonian $H(x, u, t)$ as

$$\Sigma^{x_{t^0}} : \begin{cases} \dot{x} = (J - R)\dfrac{\partial H(x,u,t)}{\partial x}^\top, x(t^0) = x_{t^0} \\ y = -\dfrac{\partial H(x,u,t)}{\partial u}^\top \end{cases} . \quad (1)$$

Here $x(t) \in \mathbb{R}^n$, $u(t), y(t) \in \mathbb{R}^m$ describe the state, the input and the output, respectively. The structure matrix $J \in \mathbb{R}^{n\times n}$ and the dissipation matrix $R \in \mathbb{R}^{n\times n}$ are skew-symmetric and symmetric positive semi-definite, respectively. The variational system $\delta\Sigma^{x_{t^0}}$ of the system $\Sigma^{x_{t^0}}$ represents the Fréchet derivative of $\Sigma^{x_{t^0}}$. According to [15], under certain conditions the adjoint of the variational system $(\delta\Sigma^{x_{t^0}})^*$ has the following relationship with the variational system $\delta\Sigma^{x_{t^0}}$

$$\begin{aligned}(\delta\Sigma^{x_{t^0}}(u))^*(v) &= \mathcal{R}(\delta\Sigma^{\phi_{t^0}}(w))\mathcal{R}(v) \\ &\approx \frac{1}{\epsilon}\mathcal{R} \circ (\Sigma^{\phi_{t^0}}(w + \epsilon\mathcal{R}(v)) - \Sigma^{\phi_{t^0}}(w)), \quad (2)\end{aligned}$$

where ϕ_{t^0} and w denote appropriate initial condition and input, respectively, ϵ represents sufficiently small positive constant and $\mathcal{R}$ is a time-reversal operator defined by $\mathcal{R}(u)(t - t^0) = u(t^1 - t)$ for $\forall t \in [t^0, t^1]$. This property is called **variational symmetry** of Hamiltonian systems. Equation (2) implies that one can calculate the input-output mapping of the adjoint by only using the input-output data of the original system. For how to concretely apply the property to iterative learning control, see [15].

B. Previously proposed learning gait generation method

In our previous works [13], [14], we proposed the following cost function to generate an optimal periodic gait:

$$\begin{aligned}\hat{\Gamma}(y, \dot{y}, \bar{u}) := \frac{1}{2}\int_{t^0}^{t^1} \Big((\mathcal{R}(y)(\tau) - Cy(\tau))^\top \nu_1(\tau)\Lambda_y \\ \times(\mathcal{R}(y)(\tau) - Cy(\tau)) + (\mathcal{R}(\dot{y})(\tau) - f_\Pi(y, \dot{y})(\tau))^\top \quad (3) \\ \times\nu_1(\tau)\Lambda_{\dot{y}}(\mathcal{R}(\dot{y})(\tau) - f_\Pi(y, \dot{y})(\tau)) + \bar{u}(\tau)^\top\Lambda_{\bar{u}}\bar{u}(\tau)\Big)\, \mathrm{d}\tau,\end{aligned}$$

where appropriate positive definite matrices $\Lambda_y, \Lambda_{\dot{y}}, \Lambda_{\bar{u}} \in \mathbb{R}^{m\times m}$ represent weight matrices for the configuration and phase coordinates restrictions and the input minimization, respectively. Since the support and swing legs change each other after the collision, a matrix $C \in \mathbb{R}^{m\times m}$ represents the leg exchange. $\nu_1(t) \in \mathbb{R}$ represents a filter function with respect to the time t defined as

$$\nu_1(t) := \begin{cases} 0 & (t^0 \le t < t^1 - \Delta t) \\ \dfrac{1}{2}\left(1 - \cos\left(\dfrac{t - (t^1 - \Delta t)}{\Delta t}\pi\right)\right) & (t^1 - \Delta t \le t \le t^1) \end{cases} \quad (4)$$

where Δt should be chosen as a sufficiently small positive constant. Figure 2 illustrates $\nu_1(t)$. $f_\Pi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m : (q_-, \dot{q}_-) \mapsto \dot{q}_+$ represents the transition mapping which maps from the pair of the configuration and phase coordinates just before the collision to the phase coordinate just after the transition. In what follows, the subscripts $(\cdot)_-$ and $(\cdot)_+$ denote just before and after the transition, respectively. Note that the configuration coordinate is assumed not to change under the collision, i.e., $q_+ \equiv q_-$. Here we also assume the following

Assumption 1: For any q_-, a mapping $f_\Pi^{q_-} := f_\Pi(q_-, \cdot) : \mathbb{R}^m \to \mathbb{R}^m : \dot{q}_- \mapsto \dot{q}_+$ is diffeomorphic.

The previously proposed method minimizes the cost function (3) by combining iterative learning control proposed in [15] and an estimation method of the transition mapping f_Π via the least-squares and eventually generates an optimal gait. See [13], [14] for the details.

III. Main results

This section proposes a modification of learning optimal gait generation method.

A. Kneed biped with torso

We consider a fully actuated planar kneed biped with torso depicted in Fig. 1. Assumptions on this robot conforms [16] and they are omitted here. The configuration coordinate is defined as $q := (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)^\top$. The leg exchange matrix C is given by

$$C := \begin{pmatrix} 0&0&0&1&0 \\ 0&0&1&0&0 \\ 0&1&0&0&0 \\ 1&0&0&0&0 \\ 0&0&0&0&1 \end{pmatrix}. \quad (5)$$

Here is a new input v defined as $v := (u_1 + u_2, -u_2 + u_3, -u_4 - u_5, u_5, -u_3 + u_4)^\top$. Then, the dynamics of this robot is described by a Hamiltonian system of the form (1) with the output $y = q$ and a fixed initial state $x_{t^0} = (q_{t^0}^\top, p_{t^0}^\top)^\top$.

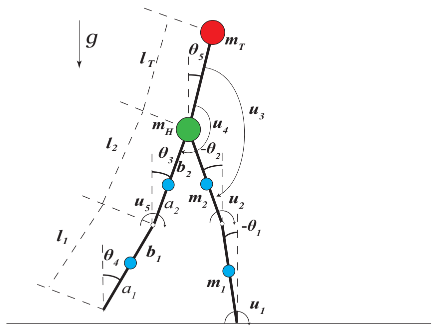

Fig. 1. Kneed biped with torso

Before the iterative learning control method is applied, feedback controllers are typically employed to the control system in order to render the system asymptotically stable. In [17], a generalized canonical transformation, which is a pair of feedback and coordinate transformations preserving the Hamiltonian structure in (1), is proposed. It is known that in the case of a typical mechanical system, a simple PD feedback preserves the structure of the Hamiltonian system [17], [15]. Let us consider a PD controller

$$v = -K_P q - K_D \dot{q} + \bar{u}, \tag{6}$$

where the input $\bar{u}$ is for iterative learning control and $K_P, K_D \in \mathbb{R}^{5\times 5}$ are symmetric positive definite matrices.

B. Modification of the constraint term for periodic trajectories

Due to the time-reversal operator $\mathcal{R}$ and the filter function ν_1, the first term of the cost function (3) attempts to make $y(t^0) = Cy(t^1)$ hold and the second one attempts to make $\dot{y}(t^0) = f_\Pi(y(t^1), \dot{y}(t^1))$ hold, respectively. However, the following conflict occasionally happens that $\lim \Delta t \to 0\ C(y(t^1) - y(t^1 - \Delta t))/\Delta t$ does not coincide with a desired terminal velocity for a periodic trajectory, i.e., $f_\Pi(y(t^1), \dot{y}(t^1))$.

In this paper, firstly, we propose another constraint term for a periodic trajectory instead of the first and second terms of the previous cost function (3). We consider conditions for a desired periodic trajectory y_d. At the terminal time t^1, y_d has to satisfy $y_d(t^1) = Cy_{t^0}$, where y_{t^0} represents a fixed initial value of the output generated by a fixed initial state of the system x_{t^0}, i.e., $y_{t^0} = q_{t^0}$. Here we define $g_\Pi(q_+, \dot{q}_+) := (f_\Pi^{q_+})^{-1}(\dot{q}_+)$ (note that $q_- \equiv q_+$). Then the time derivative of y_d at $t = t^1$ has to satisfy $\dot{y}_d(t^1) = \dot{q}_{d-} = g_\Pi(Cy_{t^0}, \dot{y}_{t^0})$. Thus we obtain the following desired trajectory around the terminal time as

$$\begin{aligned} y_d(t) &:= \dot{q}_{d-}(t - t^1) + Cy_{t^0} \\ &= g_\Pi(Cy_{t^0}, \dot{y}_{t^0})(t - t^1) + Cy_{t^0}, \end{aligned} \tag{7}$$

which satisfies the above two conditions. We utilize y_d in Eq. (7) as a reference trajectory.

Secondly, we propose a constraint term in order to avoid the foot-scuffing problem. It attempt to lift the toe of the swing leg to a certain height around the middle time of the period, i.e., $(t^1 - t^0)/2$. Due to Fig. 1, the height of the toe at the time t is given by

$$h(y)(t) = l_1(\cos y_1(t) - \cos y_4(t)) + l_2(\cos y_2(t) - \cos y_3(t)),$$

where y_i denotes the i-th element of y. We introduce the following two filter functions: one is with respect to the height and the other is with respect to the time. They are defined as

$$F(h) = \begin{cases} h - h_d & (0 \le h \le h_d) \\ 0 & (h > h_d) \end{cases},$$

where a design parameter h_d denotes the reference height, and

$$\nu_2(t) := \begin{cases} \frac{1}{2}\left(1 - \cos\left(\frac{t^0+t^1-2t}{2\Delta\bar{t}}\pi\right)\right) & \left(\frac{t^0+t^1-2\Delta\bar{t}}{2} \le t \le \frac{t^0+t^1+2\Delta\bar{t}}{2}\right) \\ 0 & \left(t^0 \le t < \frac{t^0+t^1-2\Delta\bar{t}}{2}, \frac{t^0+t^1+2\Delta\bar{t}}{2} < t \le t^1\right) \end{cases},$$

where a design parameter $\Delta\bar{t}$ represents an appropriate positive constant. Figure 2 illustrates $\nu_2(t)$.

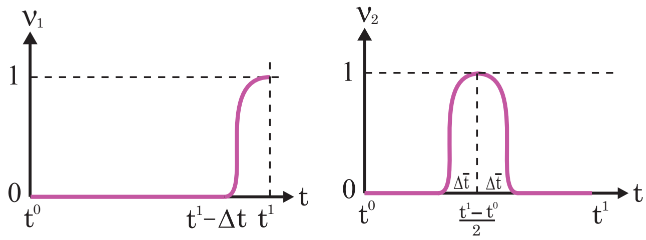

Fig. 2. Filter functions ν_1 and ν_2

Finally, we propose the following cost function:

$$\begin{aligned} \Gamma(y, \bar{u}) := \frac{1}{2}\int_{t^0}^{t^1} \Big(&(y(\tau) - y_d(\tau))^\top \nu_1(\tau)\Lambda_y(y(\tau) - y_d(\tau)) \\ &+ \lambda_h \nu_2(\tau)(F(h(y(\tau))))^2 + \bar{u}(\tau)^\top \Lambda_{\bar{u}} \bar{u}(\tau)\Big)\, \mathrm{d}\tau, \end{aligned} \tag{8}$$

where an appropriate positive constant λ_h represents a weighting coefficient and y_d is defined in Eq. (7). Although information of the transition mapping g_Π is required to derive the iteration law for learning, we estimate the mapping via the least-squares with the stored experimental data.

The iteration law for the control input is derived based on the steepest descent method [15], [13], [14]. We calculate the Fréchet derivative of the cost function (8) as

$$\begin{aligned} \delta\Gamma(y, \bar{u})(\delta y, \delta\bar{u}) &= \langle \nu_1 \Lambda_y (y - y_d), \delta y\rangle \\ &\quad + \langle \lambda_h \nu_2 F(h(y)), \frac{\mathrm{d}F}{\mathrm{d}h}\frac{\partial h}{\partial y}\delta y\rangle + \langle \Lambda_{\bar{u}}\bar{u}, \delta\bar{u}\rangle \\ &= \langle \nu_1\Lambda_y(y - y_d) + \lambda_h \nu_2 F(h(y)) \frac{\mathrm{d}F}{\mathrm{d}h}\frac{\partial h}{\partial y}^\top, \delta y\rangle + \langle \Lambda_{\bar{u}}\bar{u}, \delta\bar{u}\rangle \\ &=: \langle \nabla_y\Gamma(y, \bar{u}), \delta y\rangle + \langle \nabla_{\bar{u}}\Gamma(y, \bar{u}), \delta\bar{u}\rangle \\ &= \langle \nabla_{\bar{u}}\Gamma(y, \bar{u}) + \delta\Sigma^{x_{t^0}}(\bar{u})^*(\nabla_y\Gamma(y, \bar{u})), \delta\bar{u}\rangle, \end{aligned} \tag{9}$$

where $\nabla_y\Gamma(y, \bar{u})$ and $\nabla_{\bar{u}}\Gamma(y, \bar{u})$ denote the partial gradients of the cost function with respect to the output and the input, respectively, and we have

$$\frac{\mathrm{d}F}{\mathrm{d}h} = \begin{cases} 1 & (0 \le h \le h_d) \\ 0 & (h > h_d) \end{cases},$$
$$\frac{\partial h}{\partial y} = (-l_1 \sin y_1, -l_2 \sin y_2, l_2 \sin y_3, l_1 \sin y_4, 0). \qquad (10)$$

From Eq. (9), the gradient of the cost function

$$\Gamma^{\bar{u}}(\bar{u}) := \Gamma(\Sigma(\bar{u}), \bar{u}),$$

denoted by $\nabla_{\bar{u}}\Gamma^{\bar{u}}(\bar{u})$ is given by

$$\nabla_{\bar{u}}\Gamma^{\bar{u}}(\bar{u}) = \nabla_{\bar{u}}\Gamma(\Sigma(\bar{u}), \bar{u}) + \delta\Sigma^{x_{t^0}}(\bar{u})^*(\nabla_y \Gamma(\Sigma(\bar{u}), \bar{u})). \qquad (11)$$

Then the steepest descent method implies that the input should be updated as follows in order to minimize the cost function:

$$u_{(i+1)} = u_{(i)} - K_{(i)} \nabla_{\bar{u}}\Gamma^{\bar{u}}(\bar{u}_{(i)}), \qquad (12)$$

where the subscript i denotes the i-th iteration in a laboratory experiment and a positive definite matrix $K_{(i)}$ represents a step parameter. Variational symmetry in Eq. (2) enables one to calculate the gradient by only using the input-output data of the original system, that is experimental data, as

$$\nabla_{\bar{u}}\Gamma^{\bar{u}}(\bar{u}) \approx$$
$$\nabla_{\bar{u}}\Gamma(\Sigma(\bar{u}), \bar{u}) + \frac{1}{\epsilon}\mathcal{R}(\Sigma^{\phi_{t^0}}(w + \epsilon\mathcal{R}(\nabla_y \Gamma(\Sigma(\bar{u}), \bar{u}))) - \Sigma^{\phi_{t^0}}(w)) \qquad (13)$$

For a given pair of the state $x = (q^\top, p^\top)^\top$ and the input $\bar{u}$, the literature [18] (see also [19]) gives a way to produce a pair of the initial condition ϕ_{t^0} and the input w satisfying conditions for variational symmetry. The initial condition in the configuration and phase coordinates corresponding to ϕ_{t^0}, denoted by $Q^{\phi}_{t^0}$, and w are given by

$$Q^{\phi}_{t^0} = (q(t^1)^\top, -\dot{q}(t^1)^\top)^\top,$$
$$w = K_P \mathcal{R}(q) - K_D \mathcal{R}(\dot{q}). \qquad (14)$$

Although it is necessary to calculate y_d in Eq. (7) in order to obtain the partial gradient $\nabla_y \Gamma(\Sigma(\bar{u}), \bar{u})$, the unknown desired velocity just before touchdown $\dot{q}_{d-}$ which mapped from the nonlinear mapping g_Π is generally required. Here we propose a technique to estimate $\dot{q}_{d-}$ by the least-squares with the stored experimental data. Since the following relation holds:

$$\mathrm{d}\dot{q}_- = \frac{\partial g_\Pi(q_+, \dot{q}_+)}{\partial(q_+, \dot{q}_+)} \begin{pmatrix} \mathrm{d}q_+ \\ \mathrm{d}\dot{q}_+ \end{pmatrix}, \qquad (15)$$

we approximate $\mathrm{d}q_+, \mathrm{d}\dot{q}_+$ and $\mathrm{d}\dot{q}_-$ in (15) by differences between the desired state and the stored data by the i-th experiment. We define the estimate value of $\dot{q}_{d-}$ at the i-th iteration as

$$\widetilde{\dot{q}_{d-(i)}} := \dot{q}_{-(i)} + \widetilde{\frac{\partial g_\Pi(q_+, \dot{q}_+)}{\partial(q_+, \dot{q}_+)}}\Big|_{\substack{q_+ = q_{+(i)} \\ \dot{q}_+ = \dot{q}_{+(i)}}} \begin{pmatrix} q_{d+} - q_{+(i)} \\ \dot{q}_{d+} - \dot{q}_{+(i)} \end{pmatrix} \qquad (16)$$

and we calculate the estimate value of the Jacobian $\widetilde{\frac{\partial g_\Pi}{\partial(q_+, \dot{q}_+)}}$ by the least-squares. It should be recalled that the rest desired state q_{d+} and $\dot{q}_{d+}$ are already given as $q_{d+} = Cq_{t^0}$ and $\dot{q}_{d+} = \dot{q}_{t^0}$, respectively. Let us define the following data sets

$$\Delta Y_{-(i)} := \left[\dot{q}_{-(1)} - \dot{q}_{-(i)}, \cdots, \dot{q}_{-(i-1)} - \dot{q}_{-(i)}\right],$$
$$\Delta Y_{+(i)} := \begin{bmatrix} q_{+(1)} - q_{+(i)}, \cdots, q_{+(i-1)} - q_{+(i)} \\ \dot{q}_{+(1)} - \dot{q}_{+(i)}, \cdots, \dot{q}_{+(i-1)} - \dot{q}_{+(i)} \end{bmatrix}. \qquad (17)$$

The size of $\Delta Y_{-(i)}$ is $5 \times (i-1)$ and that of $\Delta\dot{Y}_{+(i)}$ is $10 \times (i-1)$. From Eqs. (15) and (17), we can estimate the Jacobian as

$$\widetilde{\frac{\partial g_\Pi(q_+, \dot{q}_+)}{\partial(q_+, \dot{q}_+)}}\Big|_{\substack{q_+ = q_{+(i)} \\ \dot{q}_+ = \dot{q}_{+(i)}}} := \Delta Y_{-(i)} \Delta\dot{Y}^\dagger_{+(i)}, \qquad (18)$$

where $(\cdot)^\dagger$ represents the pseudo inverse matrix of $(\cdot)$. We can also utilize MATLAB's arithmetic operator of the matrix left division to solve Eq. (18) easily. Consequently, from Eqs. (16) and (18) we obtain

$$\widetilde{\dot{q}_{d-(i)}} = \dot{q}_{-(i)} + \Delta Y_{-(i)} \Delta Y^\dagger_{+(i)} \begin{pmatrix} q_{d+} - q_{+(i)} \\ \dot{q}_{d+} - \dot{q}_{+(i)} \end{pmatrix}.$$

Finally, let us summarize the proposed learning algorithm.

Step 0 : Set appropriate positive definite matrices Λ_y and $\Lambda_{\bar{u}}$, an appropriate positive constant λ_h as weighting parameters, positive constants Δt, $\Delta\bar{t}$ and h_d as design parameters, a sufficiently small positive constant ϵ_{stop} as a convergence parameter, appropriate positive definite matrices K_P and K_D for PD feedback (6) and an appropriate initial condition $Q_{t^0} := (q^\top_{t^0}, \dot{q}^\top_{t^0})^\top$. Set $i = 1$. Then, go to Step 1.

Step i $(1 \le i \le 10)$: Execute preliminary experiment with an appropriate initial condition around Q_{t^0} and zero control input (or an appropriate initial input) in order to obtain data sets ΔY_- and ΔY_+ defined in Eq. (19), which are slight modifications of those in Eq. (17), in order to calculate $\widetilde{y_d}$ which is an estimate value of y_d in Eq. (7).
Set $i = i + 1$. Then, go to Step i.

Step i $(i = 11)$: Execute laboratory experiment with x_{t^0} and zero control input (or an appropriate initial input). Let $Q_{(i)} := (q^\top_{(i)}, \dot{q}^\top_{(i)})^\top$, $y_{(i)}$ and $\bar{u}_{(i)}$ denote the corresponding data obtained by the experiment, respectively. Set $k = 4$. Then, go to Step $3k$.

Step $3k$: Execute the $3k$-th laboratory experiment via the following iteration law

$$\begin{cases} Q_{t^0(3k)} = (q(t^1)^\top_{(3k-1)}, -\dot{q}(t^1)^\top_{(3k-1)})^\top \\ \bar{u}_{(3k)} = K_P \mathcal{R}(q_{(3k-1)}) - K_D \mathcal{R}(\dot{q}_{(3k-1)}) \end{cases}.$$

Go to Step $3k$+1.

Step $3k + 1$: Execute the $(3k$+1)-th laboratory experiment via the following iteration law with a sufficiently small positive constant $\epsilon_{(k)}$

$$\begin{cases} Q_{t^0(3k+1)} = Q_{t^0(3k)} \\ \bar{u}_{(3k+1)} = \bar{u}_{(3k)} + \epsilon_{(k)}\mathcal{R}(\nu_1\Lambda_y(y_{(3k-1)} - \widetilde{y_{d(k)}}) \\ \qquad + \lambda_h\nu_2 F(h(y_{(3k-1)}))\frac{\mathrm{d}F}{\mathrm{d}h}\frac{\partial h}{\partial y}^\top) \end{cases}.$$

Here $\widetilde{y_{d(k)}}$ is calculated by

$$\widetilde{y_{d(k)}} = \left(\dot{q}_{(3k-1)}(t^1) + \Delta Y_{-(k)}\Delta Y_{+(k)}^\dagger \begin{pmatrix} Cq_{t^0} - q_{+(3k-1)} \\ \dot{q}_{t^0} - q_{+(3k-1)} \end{pmatrix}\right) \times (t - t^1) + Cq_{t^0}$$

with data sets $\Delta Y_{-(k)}$ and $\Delta Y_{+(k)}$ as

$$\Delta Y_{-(k)} := \begin{cases} \left[\dot{q}_{-(1)} - \dot{q}_{-(3k-1)}, \cdots, \dot{q}_{-(10)} - \dot{q}_{-(3k-1)}\right] & (k=4) \\ \left[\dot{q}_{-(1)} - \dot{q}_{-(3k-1)}, \cdots, \dot{q}_{-(10)} - \dot{q}_{-(3k-1)}, \right. \\ \left. \quad \dot{q}_{-(14)} - \dot{q}_{-(3k-1)}, \cdots, \dot{q}_{-(3k-4)} - \dot{q}_{-(3k-1)}\right] & (k>4) \end{cases},$$

$$\Delta Y_{+(k)} := \begin{cases} \begin{bmatrix} q_{+(1)} - q_{+(3k-1)}, \cdots, q_{+(10)} - q_{+(3k-1)} \\ \dot{q}_{+(1)} - \dot{q}_{+(3k-1)}, \cdots, \dot{q}_{+(10)} - \dot{q}_{+(3k-1)} \end{bmatrix} & (k=4) \\ \begin{bmatrix} q_{+(1)} - q_{+(3k-1)}, \cdots, q_{+(10)} - q_{+(3k-1)}, \\ \dot{q}_{+(1)} - \dot{q}_{+(3k-1)}, \cdots, \dot{q}_{+(10)} - \dot{q}_{+(3k-1)}, \\ \quad q_{+(14)} - q_{+(3k-1)}, \cdots, q_{+(3k-4)} - q_{+(3k-1)}, \\ \quad \dot{q}_{+(14)} - \dot{q}_{+(3k-1)}, \cdots, \dot{q}_{+(3k-4)} - \dot{q}_{+(3k-1)}, \end{bmatrix} & (k>4) \end{cases} \quad (19)$$

Go to Step $3k$+2.

Step $3k+2$: Execute the ($3k$+2)-th laboratory experiment via the following iteration law with an appropriate positive definite matrix $K_{(k)}$

$$\begin{cases} Q_{t^0(3k+2)} = Q_{t^0(3k-1)} \\ \bar{u}_{(3i+2)} = \bar{u}_{(3k-1)} - K_{(i)}\big(\Lambda_{\bar{u}}\bar{u}_{(3k-1)} \\ \qquad + \dfrac{1}{\epsilon_{(k)}}\mathcal{R}(y_{(3k+1)} - y_{(3k)})\big) \end{cases}.$$

If $\Gamma(y_{(3k-1)}, \bar{u}_{(3k-1)}) - \Gamma(y_{(3k+2)}, \bar{u}_{(3k+2)}) < \epsilon_{stop}$, the procedure terminates. Otherwise, set $k = k+1$ and go to Step $3k$.

The triple iteration laws imply that this learning procedure needs three experiments to execute a single update in (12). In the $3k$-th iteration, we produce a trajectory ϕ corresponds to $x_{(3k-1)}$ by Eq. (14), in order to utilize variational symmetry (2). In the ($3k$+1)-th iteration, we calculate the output $\Sigma^{\phi_{t^0}}(w + \epsilon\mathcal{R}(\nabla_y\Gamma(\Sigma(\bar{u}), \bar{u})))$ in Eq. (13). Then the input and output signals of $\delta\Sigma^{x_{t^0}}(\bar{u})^*(\nabla_y\Gamma(\Sigma(\bar{u}), \bar{u}))$ in Eq. (11) can be calculated from Eq. (13). With this information, the gradient of the cost function with respect to the input $\nabla_{\bar{u}}\Gamma^{\bar{u}}(\bar{u}_{(3k-1)})$ (see Eq. (11)) is obtained. Finally, the input for the ($3k$+2)-th iteration is given by Eq. (12) with these signals.

Remark 1: If the learning procedure is executed around a symmetric trajectory, one can iterate the following procedure after executing Step 11 in the above algorithm (in this case, the initial value of k is 6 not 4)

$$\begin{cases} Q_{t^0(2k)} = Q_{t^0(2k-1)} \\ \bar{u}_{(2k)} = \bar{u}_{(2k-1)} + \epsilon_{(k)}\mathcal{R}(\nu_1\Lambda_y(y_{(2k-1)} - \widetilde{y_{d(k)}}) \\ \qquad + \lambda_h\nu_2 F(h(y_{(2k-1)}))\frac{\mathrm{d}F}{\mathrm{d}h}\frac{\partial h}{\partial y}^\top) \end{cases}$$

$$\begin{cases} Q_{t^0(2k+1)} = Q_{t^0(2k-1)} \\ \bar{u}_{(2k+1)} = \bar{u}_{(2k)} - K_{(k)}\big(\Lambda_{\bar{u}}\bar{u}_{(2k)} + \dfrac{1}{\epsilon_{(k)}}\mathcal{R}(y_{(2k+1)} - y_{(2k)})\big) \end{cases} \quad (20)$$

with appropriately modified data sets $\Delta Y_{-(k)}$ and $\Delta Y_{+(k)}$.

IV. NUMERICAL EXAMPLE

We apply the proposed algorithm in the previous section to the kneed biped with torso depicted in Fig. 1 to generate an optimal gait trajectory on the level ground. The physical parameters of the robot are chosen as $m_T = 10.0, m_H = 13.0, m_1 = m_2 = 2.25$ [kg] and $a_1 = b_1 = a_2 = b_2 = 0.23, l_T = 0.30$ [m]. Gains $K_P = \mathrm{diag}(4,4,4,4,4)$ and $K_D = \mathrm{diag}(2,2,2,2,2)$ are chosen for the PD feedback (6). We utilize the following design parameters with respect to weighting functions for the cost function (8) as $\Lambda_y = \mathrm{diag}(2,0,2,2,0)$, $\Lambda_{\bar{u}} = \mathrm{diag}(1\times10^{-9}, 1\times10^{-9}, 1\times10^{-9}, 1\times10^{-9}, 1\times10^{-9})$ and $\lambda_h = 1$, those with respect to the filter functions ν_1 and ν_2 as $\Delta t = \Delta\bar{t} = 5\times10^{-3}[s]$ and those with respect to ILC algorithm as $K_{(\cdot)} = \mathrm{diag}(200,0,100,100,0)$ and $\epsilon_{(\cdot)} = 1$. In this section, we show the results by 2 Step learning scheme in Remark 1. We proceed 800 steps of the learning procedure which means 1610 simulations including 10 preliminary experiments with the following initial condition:

$$q_{t^0} = (-0.15, -0.15, 0.15, 0.15, 0.0),$$
$$\dot{q}_{t^0} = (0.65, 0.65, 0.24, 1.1, 0.0),$$

which is heuristically determined.

Figure 3 shows the history of the cost function (8) along the iteration decreasing monotonically. It implies that the output trajectory converges to an (at least locally) optimal one smoothly. Figure 4 represents the animations of the robot before and after learning, respectively and it implies that a walking motion seems to be generated eventually. While many conventional results equipped heuristically designed reference trajectories for knee joints in order to control kneed biped robots, the proposed method can generate an optimal trajectory which achieves foot clearance properly. Figure 5 exhibits the phase portrait of q-$\dot{q}$. The fact that an almost periodic trajectory is generated follows from that the phase portraits in these figures form almost closed orbits.

V. CONCLUSION

In this paper, we have proposed a modification of our previous learning gait generation method by equipping a reference trajectory considering discontinuous velocity transitions. This method can generate an optimal feedforward control input and the corresponding periodic trajectory minimizing

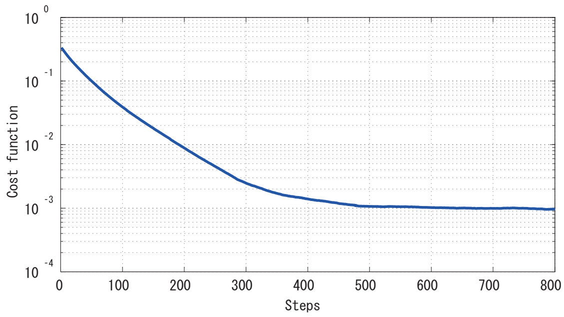

Fig. 3. Cost function

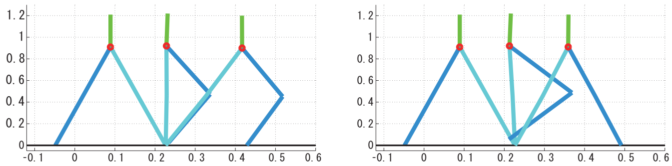

Fig. 4. Stick diagrams before and after learning

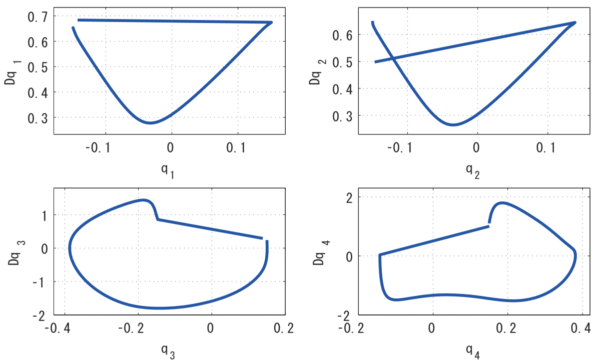

Fig. 5. Phase portraits

the L_2 norm of the control input. Although calculation of such reference trajectory generally requires information of the transition mapping, the proposed method estimates the mapping by the least-squares with the stored experimental data. Thus, it does not require the precise knowledge of the plant system nor the discontinuous state transition model. We also have proposed a technique to generate an optimal gait not only energy-efficient but also avoiding the foot-scuffing problem. Finally, numerical simulations of a kneed biped with torso have demonstrated the validity of the proposed method.

REFERENCES

[1] M. Vukobratović and J. Stepanenko, "On the stability of anthropomorphic systems," *Mathematical Biosciences*, vol. 15, pp. 1–37, 1972.

[2] A. Takanishi, M. Ishida, Y. Yamazaki, and I. Kato, "The realization of dynamic walking by the biped walking robot WL-10RD," in *Proc. Int. Conf. Advanced Robotics*, 1985, pp. 459–466.

[3] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, "The development of honda humanoid robot," in *Proc. IEEE Int. Conf. Robotics and Automation*, 1998, pp. 1321–1326.

[4] T. McGeer, "Passive dynamic walking," *Int. J. Robotics Research*, vol. 9, no. 2, pp. 62–82, 1990.

[5] K. Osuka and K. Kirihara, "Motion analysis and experiments of passive walking robot quartet II," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2000, pp. 3052–3056.

[6] A. Sano, Y. Ikemata, and H. Fujimoto, "Analysis of dynamics of passive walking from storage energy and supply rate," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2003, pp. 2478–2483.

[7] A. Goswami, B. Espiau, and A. Keramane, "Limit cycles in a passive compass gait biped and passivity-mimicking control laws," *Autonomous Robots*, vol. 4, no. 3, pp. 273–286, 1997.

[8] M. W. Spong, "Passivity-base control of the compass gait biped," in *Proc. of IFAC World Congress*, 1999, pp. 19–23.

[9] F. Asano, M. Yamakita, N. Kamamichi, and Z. W. Luo, "A novel gait generation for biped walking robots based on mechanical energy constraint," *IEEE Trans. Robotics and Automation*, vol. 20, no. 3, pp. 565–573, 2004.

[10] S. Hyon and T. Emura, "Symmetric walking control: Invariance and global stability," in *Proc. IEEE ICRA*, 2005, pp. 1455–1462.

[11] F. Asano and Z. W. Luo, "Energy-efficient and high-speed dynamic biped locomotion based on principle of parametric excitation," *IEEE Trans. Robotics*, vol. 24, no. 6, pp. 1289–1301, 2008.

[12] S. Satoh, K. Fujimoto, and S. Hyon, "Gait generation for passive running via iterative learning control," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2006, pp. 5907–5912.

[13] ——, "Biped gait generation via iterative learning control including discrete state transitions," in *Proc. 17th IFAC World Congress*, 2008, pp. 1729–1734.

[14] ——, "A framework for optimal gait generation via learning optimal control using virtual constraint," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2008, pp. 3426–3432.

[15] K. Fujimoto and T. Sugie, "Iterative learning control of Hamiltonian systems: I/O based optimal control approach," *IEEE Trans. Autom. Contr.*, vol. 48, no. 10, pp. 1756–1761, 2003.

[16] A. Goswami, B. Thuilot, and B. Espiau, "Compass-like biped robot part i: Stability and bifurcation of passive gaits," *INRIA Research Report*, no. 2996, 1996.

[17] K. Fujimoto and T. Sugie, "Canonical transformation and stabilization of generalized Hamiltonian systems," *Systems & Control Letters*, vol. 42, no. 3, pp. 217–227, 2001.

[18] K. Fujimoto, T. Horiuchi, and T. Sugie, "Iterative learning optimal control of hamiltonian systems based on variational symmetry," *Trans. the Institute of Systems Control and Information Engineers*, vol. 21, no. 1, pp. 10–17, 2008, (in Japanese).

[19] K. Fujimoto and I. Koyama, "Iterative feedback tuning for Hamiltonian systems," in *Proc. 17th IFAC World Congress*, 2008, pp. 15 678–15 683.