Biol Cybern (2006) 94: 245–261
DOI 10.1007/s00422-005-0047-3

ORIGINAL PAPER

Biological Cybernetics

T. Iwasaki · M. Zheng

Sensory feedback mechanism underlying entrainment of central pattern generator to mechanical resonance

Received: 12 May 2005 / Accepted: 21 November 2005 / Published online: 10 February 2006

Abstract Rhythmic body motions observed in animal locomotion are known to be controlled by neuronal circuits called central pattern generators (CPGs). It appears that CPGs are energy efficient controllers that cooperate with biomechanical and environmental constraints through sensory feedback. In particular, the CPGs tend to induce rhythmic motion of the body at a natural frequency, i.e., the CPGs are entrained to a mechanical resonance by sensory feedback. The objective of this paper is to uncover the mechanism of entrainment resulting from the dynamic interaction of the CPG and mechanical system. We first develop multiple CPG models for the reciprocal inhibition oscillator (RIO) and examine through numerical experiments whether they can be entrained to a simple pendulum. This comparative study identifies the neuronal properties essential for the entrainment. We then analyze the simplest model that captures the essential dynamics via the method of harmonic balance. It is shown that robust entrainment results from a *strong, positive-feedback coupling* of a *lightly damped* mechanical system and the RIO consisting of neurons with the *complete adaptation property.*

1 Introduction

The control system for animal locomotion is formed by a group of neurons interconnected in a specific manner, called the central pattern generator (CPG). The CPG can be viewed as a nonlinear dynamical system capable of generating patterns, i.e., phase-locked oscillations, in the time courses of cell membrane potentials. Such patterns are used as activation signals for muscle contractions to achieve coordinated rhythmic body motions observed in swimming, walking, flying, etc. The sensory information is not essential for pattern generation itself, but rather modifies the oscillation profile to achieve robustness against, and adaptation to, environmental changes.

T. Iwasaki (✉) · M. Zheng
Department of Mechanical and Aerospace Engineering
University of Virginia, Charlottesville, VA 22904-4746, USA
E-mail: iwasaki@virginia.edu

The neuronal circuits of CPGs have been extensively studied for a wide variety of animals, and their mathematical models have been developed and validated through comparisons of model predictions with experimental observations (Orlovsky et al., 1999; Cohen et al., 1992; Ermentrout and Chow, 2002). The simplest and perhaps most fundamental CPG consists of two (groups of) neurons with mutually inhibitory synaptic connections (Brown, 1911), and is called the half-center model or the reciprocal inhibition oscillator (RIO). The inhibitory connections in the RIO generate out-of-phase oscillations in the activities of the two neurons, which drive a pair of muscles (flexor/extensor, dorsal/ventral etc.) alternately to generate periodic motions of limbs or limbless bodies. Reciprocal inhibition is considered to be an essential mechanism underlying the neuronal control of animal locomotion (Friesen, 1994).

The RIOs and their extensions (e.g., recurrent cyclic inhibition oscillators) have been studied in terms of realistic neuron models with supporting biological data (Friesen and Stent, 1978; Friesen, 1994; Getting, 1989) as well as simplified models with mathematical analysis (Matsuoka, 1985, 1987). While the former approach focuses on accurately reproducing the biological behavior, the latter aims to rigorously analyze the dynamical behavior using simple models at the expense of some reality. There is, however, a common principle in neuronal dynamics found in both lines of research. The property that diminishes the neuronal response to a tonic stimulus over time, called *adaptation*, has been shown crucial for generation of stable out-of-phase oscillations in RIOs (Matsuoka, 1985; Friesen, 1994).

Neuronal control circuits for locomotion have been studied in the context of driving the mechanical body, possibly with sensory feedback (Taga, 1991; Ekeberg, 1993; Wadden and Ekeberg, 1998; Williamson, 1998; Ijspeert, 2001; Lewis et al., 2003; Fukuoka et al., 2003). These references consider rhythmic motions generated by coupling of CPGs with various mechanical systems, including lamprey swimming, salamander swimming and walking, and human-like legged locomotion. Stable locomotions were generated using the specific control architectures that were developed

from biological knowledge on neuronal CPG circuits. Successful parameter tuning in each work indicated that the nonlinear control architecture employed was dynamically rich enough to generate the expected behaviors. However, a fundamental question still remains: Which dynamical properties and mechanisms are essential for achieving natural motions?

One of the intriguing properties of neuronal control systems is the ability to achieve natural motions by cooperating with the physical constraints imposed by the dynamics of the body and the surrounding environment. In particular, the frequency of rhythmic body motion seems to be set close to the natural frequency of the mechanical body (possibly modified through the interaction with the environment). For instance, bees flap much faster than hawks, and elephants walk with a much slower pace than ants. From the neuronal control viewpoint, such phenomena occur as a result of the entrainment of the CPG to a mechanical resonance through sensory feedback. The entrainment to resonance is an energy efficient control strategy for generating natural periodic motions, and hence its underlying biological mechanism, if uncovered, will be useful in engineering applications. Indeed, methods for achieving this phenomenon have been studied in the engineering literature with different terminologies; "self excitation" (Ono, 1998; Ono et al., 2001), "feedback resonance" (Fradkov, 1999a,b), and "resonance tuning" (Raney and Slominski, 2004).

This paper attempts to identify the principles underlying the natural rhythmic motion achieved by the entrainment of the CPG to a mechanical resonance. In contrast to the previous works on the neuro-mechanical interactions that studied the entire locomotion system of particular animals, we try to focus on the simplest relevant dynamics that captures the essential mechanism of the resonance entrainment. In particular, we will study the motion of a simple pendulum coupled with an RIO, driven by the following hypothesis: *The RIO is capable of achieving robust entrainment to the mechanical resonance.* This means that an RIO, with intrinsic frequency ω_{RIO}, tends to drive the body (represented by the pendulum in our study) at its natural frequency ω_o even if $\omega_{\text{RIO}} \neq \omega_o$. The term "robust" is added to emphasize the ability to maintain this tendency for a wide range of ω_o. In other words, an RIO is expected to change its oscillation frequency in accordance with the mechanical property of the body to which the RIO is coupled (e.g. slower oscillation for a straight arm and faster oscillation for an arm folded at the elbow). The objective of this paper is to provide evidence for the correctness of this hypothesis, and to uncover the underlying mechanism for the robust entrainment.

Our approach is based on mathematical modeling of RIOs, followed by numerical experiments on, and theoretical analyses of, an RIO-driven pendulum system. We consider several neuron models with different complexity and structure, and investigate which neuronal properties are crucial for achieving the robust entrainment. In particular, for each of the neuron models, an RIO is formed and conditions for existence of oscillatory trajectory are obtained. Each RIO is then used to drive a simple pendulum and the entrainment property is examined and compared with the results for other RIO models to identify the crucial neuronal properties. Finally, the method of harmonic balance (see e.g., Khalil 1996; Glad and Ljung 2000) is applied to the RIO-pendulum system to reveal how combinations of various dynamical properties determine the frequency of resulting oscillations. Through this process, we were able to identify a few properties essential for achieving the robust entrainment to resonance.

2 Neuron models

In this section, we will present several models of neuronal dynamics with different complexity and structure. Later, each model will be used as a basic unit to form an RIO and tested against capability of robust entrainment. Comparisons of the results for different models reveal which neuronal properties are (or are not) important for achieving robust entrainment.

2.1 Dynamic axon with feedforward adaptation

Many neurons consist of three components: dendrite, soma, and axon. Within a neuron, the information flows from the dendrites to the soma, and then to the axon, in the form of an electrical signal. The neuronal dynamics can be modeled as an input–output system from the current injection input u to the membrane potential output v. We consider the following structure for the neuron model

$$v = \mathcal{A}(q), \quad q = \mathcal{F}(u)$$

where $\mathcal{A}$ and $\mathcal{F}$ represent the dynamical mappings of the axon and (soma, dendrites), respectively, and q is the current flowing from the soma to the axon.

The dynamics of the soma and dendrites have often been modeled by a linear low-pass filter (Hunt et al., 1992; Hadeler, 1974). However, such models do not capture a neuronal property important for generating oscillations in RIOs, namely, the adaptation (Matsuoka, 1985; Friesen, 1994). The adaptation property may be characterized by the existence of a peak in the output q in response to a step (constant) input u; this means that q increases immediately after the input stimulus, but soon decreases as it adapts to the stimulus. A low-pass filter typically has a monotonically increasing step response and hence does not possess the adaptation property. To embed this important property in $\mathcal{F}$, we consider a linear band-pass filter with the transfer function from u to q given by

$$F(s) = \frac{ks}{(1+\tau_1 s)(1+\tau_2 s)} \tag{1}$$

where k, τ_1, and τ_2 are positive constants. The poles of $F(s)$ capture the processing time lag, and the zero at the origin achieves the adaptation by completely blocking the constant input in the steady state. We shall explain the behavior of this model later in comparison with other models.

The mapping $\mathcal{A}$ is responsible for generation of action potentials (spike trains). The essential dynamics of the nerve impulse can be captured by the so-called two-variable models (Hindmarsh and Rose, 1982; Rinzel, 1985; Av-Ron et al., 1993) described by the following general form:

$$\begin{aligned} \dot{v} &= \psi_1(v) - w + q \\ \dot{w} &= \rho(\psi_2(v) - w) \end{aligned} \tag{2}$$

where v is the membrane potential, w represents the dynamics of voltage gated ion channels, and $\rho > 0$ is the rate of channels opening. The nonlinear functions $\psi_k(v)$ $(k = 1, 2)$ are such that the slope of $\psi_2(v)$ is always nonnegative, and the slope of $\psi_1(v)$ varies from negative, positive, to negative as v increases. A typical example of such functions is shown in Fig. 1 (Iwasaki and Zheng, 2002). The shapes and relative positioning of the curves are important for capturing certain neuronal properties, but their sizes are immaterial as their effects can be captured by scaling the input and output by a constant.

For the oscillation analysis and simulation studies in this paper, we consider the dynamic axon model (2) with functions described by

$$\psi_1(v) := \phi_1(v) - bv, \quad \psi_2(v) := \phi_2(v) \tag{3}$$

where b is a positive scalar and $\phi_k \in \mathbf{S}$ where $\mathbf{S}$ is the class of monotonically increasing, bounded, Lipschitz functions. As shown later, this class of functions simplifies the oscillation analysis. The particular functions and parameter values used in our simulation studies are taken from Iwasaki and Zheng (2002) and are summarized in the Appendix. The functions ψ_k are chosen so that the equilibrium point is close to the origin and the operating regions of v and w are roughly between 0 and 1 (see Fig. 1). The time constant τ_1 determines the initial response speed and is roughly equal to a typical duration of the nerve impulses ($\cong$10 ms), while the other time constant $\tau_2 (= 100$ ms) determines the rate of adaptation and is much larger than τ_1.

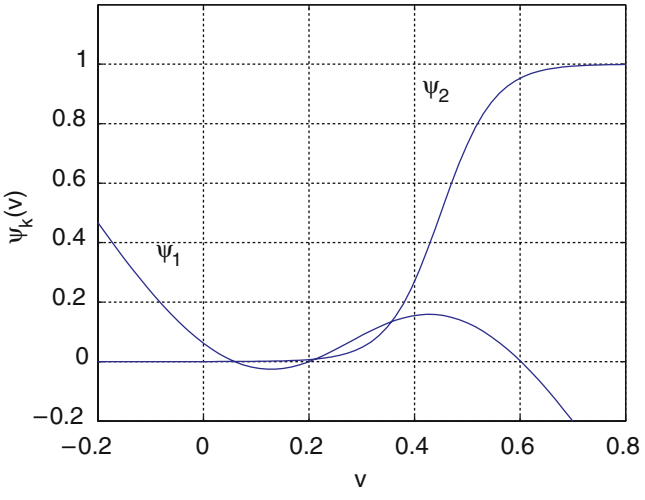

Fig. 1 Typical shapes of $\psi_1(v)$ and $\psi_2(v)$. The equations are given in the Appendix

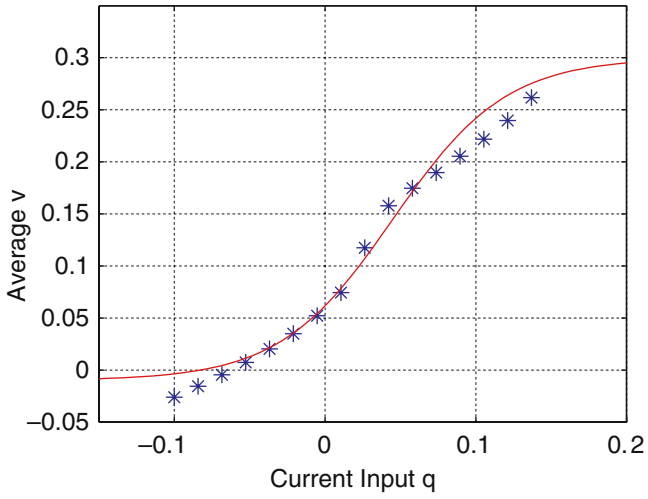

Fig. 2 Input/output data of the dynamic axon model (*asterisks*: data from dynamics axom with feedforward adaptation (DAFF);*solid line* curve fit)

The dynamic axon model described by (2) and (3) has been shown to possess various neuronal properties including threshold, refractory period, rate coding, and bistability (Iwasaki and Zheng, 2002). We shall refer to the neuron model described here as the Dynamic Axon with FeedForward adaptation (DAFF) model.

2.2 Static axon with feedforward adaptation

In this section, we develop another neuron model by simplifying the DAFF model. We keep the (soma, dendrite) model $F(s)$ as is, but replace the dynamic axon model (2) by a static nonlinearity $v = \varphi(q)$ with $\varphi \in \mathbf{S}$. This class of functions will capture the threshold property of the axon, and in this case, v may be interpreted as the firing rate of nerve impulses rather than the membrane potential. The neuronal dynamics is thus modeled by the following Static Axon with Feed Forward adaptation (SAFF) structure:

$$v = \varphi(q), \quad q = F(s)u. \tag{4}$$

An SAFF model that roughly approximates a given DAFF model has been developed through numerical experiments by the following procedure.[1] Apply a constant input q to the dynamic axon model (2) and measure the steady state output v. If q is less than a certain threshold level, the output v approaches a constant. Let this constant value be the data y. If q is above the threshold, the output v will be a periodic signal (spike trains). In this case, let the average value of v over a period be the data y. Repeat the numerical experiment for different values of q to collect the data points (q_k, y_k), $(k = 1, \ldots, n)$. Find a function $\varphi \in \mathbf{S}$ that approximates the data points in the sense that $|y_k - \varphi(q_k)|$ is small for all k.

[1] We do not claim or expect that the procedure described here is optimal in any sense, although it seems reasonable. The accuracy of the approximation is not so important in our analysis, as the robust entrainment property seems insensitive to it.

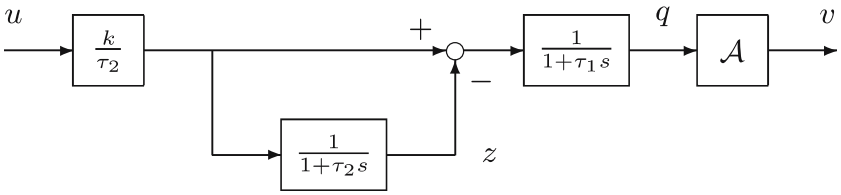

Fig. 3 Feedforward adaptation mechanism. The axon model $\mathcal{A}$ is given by φ for static axon with feedback adaptation (SAFF) and (2) for DAFF.

For the dynamic axon model with the parameters given in the Appendix, the numerical experiments were performed to obtain the data points indicated by "asterisk" in Fig. 2. As the input q increases, the axon model starts to generate spike trains when q becomes greater than $q_{\text{th}} \cong 0.03$. The spike trains are observed with increasing frequency until q becomes greater than 0.15. The data points in this operating region are approximated by the following function to capture the threshold and saturation effects:

$$\varphi(q) := a_s \tanh\left(\frac{q - c_s}{b_s}\right) + d_s.$$

The parameters may be optimized to achieve the least-square error, but a rough manual tuning yielded the parameter values in the Appendix with the solid curve in Fig. 2. The behavior of the DAFF model, on average, is expected to be captured by the SAFF model thus obtained.

2.3 Static axon with feedback adaptation

This section introduces another neuron model proposed by Matsuoka (1985) for comparison purposes, and shows how to choose the model parameters to roughly match its dynamic behavior with that of the SAFF model. The Matsuoka model may be interpreted as a static axon with feedback adaptation (SAFB).

The input–output mapping ($u \to v$) of the SAFB model is defined by

$$\begin{aligned} &\tau_r \dot{q} + q = \gamma u - b_a z \\ &\tau_a \dot{z} + z = v \\ &v = \max(q + \beta, 0) \end{aligned}$$

where $\max(\cdot, \cdot)$ takes the larger value of the two arguments, v is the firing rate of the membrane potential, z is the adaptation variable, u is the (current) input, and the parameters γ, β, b_a, τ_r, and τ_a are all taken to be positive.

The behavior of the SAFB model would be similar to a given SAFF model if the parameters are chosen as follows. In the active operating region, where $q + \beta \geq 0$, the SAFB model behaves linearly, with the transfer function

$$\frac{v(s)}{u(s)} = \frac{\gamma(1 + \tau_a s)}{(1 + \tau_r s)(1 + \tau_a s) + b_a}.$$

On the other hand, linearization of the SAFF model around the equilibrium with any constant input u is given by

$$\frac{v(s)}{u(s)} = \frac{ks}{(1 + \tau_1 s)(1 + \tau_2 s)} \cdot \varphi'(0)$$

where φ' is the derivative of φ. The two transfer functions are close to each other if

$$\tau_a \gg 1, \quad \frac{\gamma \tau_a}{1 + b_a} = k\varphi'(0),$$

$$\frac{\tau_r + \tau_a}{1 + b_a} = \tau_1 + \tau_2, \quad \frac{\tau_r \tau_a}{1 + b_a} = \tau_1 \tau_2.$$

Thus τ_r and τ_a are obtained as the solutions x to

$$x^2 - (1 + b_a)(\tau_1 + \tau_2)x + (1 + b_a)\tau_1 \tau_2 = 0$$

where $\tau_r < \tau_a$, and γ is given by

$$\gamma = \frac{k(1 + b_a)}{\tau_a} \cdot \varphi'(0).$$

Note that once b_a is fixed, the values of τ_r, τ_a, and γ can be calculated by the above equations. To ensure $\tau_a \gg 1$, we should choose sufficient large $b_a > 0$.

To choose the value of the bias β, consider the equilibrium of the SAFB model with $u = 0$. The equilibrium-output value is $v_{\text{e}} := \beta/(1 + b_a)$. It may seem reasonable to match v_{e} with the equilibrium output of the SAFF model, i.e., $v_{\text{e}} = \varphi(0)$. However, if β is chosen to achieve this, it is observed that the resulting SAFB model produces an RIO oscillation with a very small amplitude when compared with the case with the SAFF model. Since the oscillation amplitude of an SAFB-based RIO is proportional to β (Matsuoka, 1985), we can choose β so that the amplitude is matched with that of an SAFF-based RIO. When doing so, the condition $v_{\text{e}} = \varphi(0)$ would give a rough initial estimate of β from which its value can be adjusted. In particular, we may let

$$v_{\text{e}} = \delta\varphi(0), \text{ or } \beta = \delta(1 + b_a)\varphi(0)$$

and tune the parameter $\delta \cong 1$ to achieve amplitude matching.

Given an SAFF model and $b_a \gg 1$, the procedure described above generates the parameters of the SAFB model such that the linearized dynamics of the SAFF and SAFB models are close to each other. Note that b_a can be thought of as the strength of adaptation. We consider two SAFB models with small and large values of b_a to compare the effects of adaptation on the robust entrainment. The parameters thus fixed are summarized in the Appendix.

2.4 Adaptation effects

The neuron models presented in the previous subsections capture the adaptation property. We considered two basic mechanisms of adaptation – feedforward and feedback. These

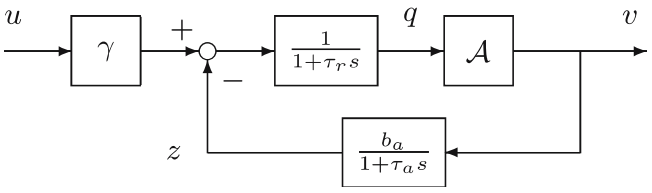

Fig. 4 Feedback adaptation mechanism. The axon model $\mathcal{A}$ is given by $\max(q+\beta, 0)$ for SAFB

mechanisms are illustrated by the block diagrams in Figs. 3 and 4. In each figure, z is the adaptation variable that develops with some time lag to turn off the effect of the input. Note that z is generated from either the feedforward of the input u (Fig. 3) or the feedback of the output v (Fig. 4).

Figure 5 shows the responses of the four neuron models (DAFF, SAFF, SAFB with $b_a = 5$, and SAFB with $b_a = 10^4$) when the step input of magnitude $u = 1$ is applied at time $t = 0.1\,s$. Each neuron model responds to the input right after it is applied, but the model output v tends to return to the resting value after a while. In particular, the outputs of SAFF and DAFF models approach the resting values because the zero of $F(s)$ at the origin blocks the constant input in the steady state, indicating complete adaptation. The SAFB model would also achieve the complete adaptation in the limit where b_a goes to infinity. For finite values of b_a, the adaptation is incomplete and the distance between the steady state and resting values is negatively correlated with b_a.

3 Reciprocal inhibition oscillator

3.1 Basics

The reciprocal inhibition oscillator is the simplest CPG, consisting of two neurons with mutually inhibitory synaptic connections (Friesen, 1994). The block diagram of the RIO is shown in Fig. 6 where two neurons $\mathcal{N}$ are connected via inhibitory synapses with strength σ. Note that the RIO has

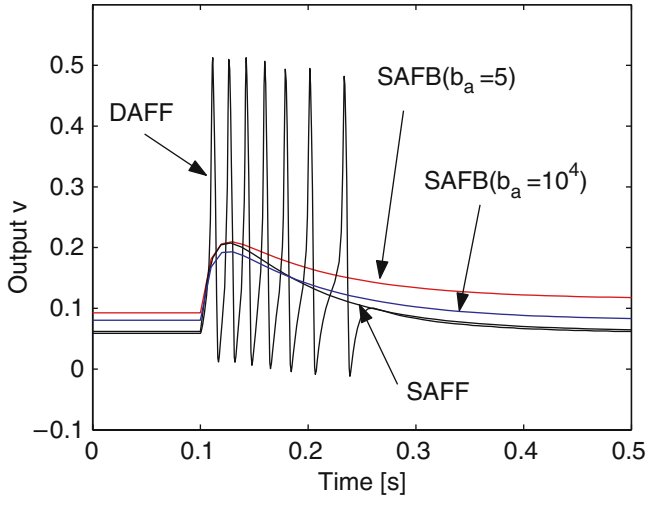

Fig. 5 Step responses of the neuron models

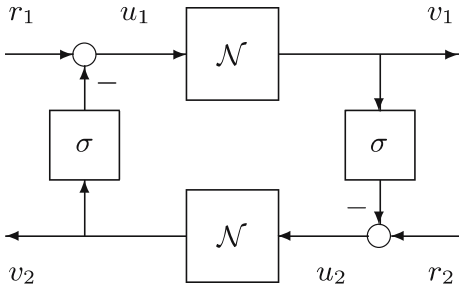

Fig. 6 Reciprocal inhibition oscillator

inputs r_i and outputs v_i $(i = 1, 2)$. Below, we analyze the RIOs with different neuron models to obtain conditions for the existence of oscillation.

When the SAFB neuron model is substituted for each of $\mathcal{N}$, the RIO is known to oscillate if

$$\left(1 + \frac{\tau_r}{\tau_a}\right) < \sigma\gamma < (1 + b_a) \tag{5}$$

with the amplitude proportional to β, provided $\beta > 0$ and $r_1 = r_2 = 0$ (Matsuoka, 1985, 1987). This condition is in fact necessary and sufficient for the nonexistence of any stable equilibrium points, which constitutes a sufficient condition for existence of stable oscillation with a proof that every solution is bounded.

When the DAFF or the SAFF neuron model is substituted for $\mathcal{N}$, the RIO can be represented through the separation of linear and nonlinear parts (Doyle et al., 1991) by the feedback system in Fig. 7, where $G(s)$ is a transfer function and Φ is a static nonlinear mapping given by either $\mathrm{diag}(\phi_1, \phi_1, \phi_2, \phi_2)$ or $\mathrm{diag}(\varphi, \varphi)$. In general, a sufficient condition for the existence of oscillation can be given as follows.

Lemma 1 *Consider the system in Fig. 7 where r is a constant input vector, $G(s)$ is a transfer function, and Φ is a diagonal mapping with each entry belonging to* **S**. *Then the system oscillates for a generic initial condition*[2] *if $G(s)$ is stable and every equilibrium point is unstable.*

Proof Since Φ is bounded, stability of $G(s)$ guarantees boundedness of every trajectory. A generic trajectory does not converge to any equilibrium point due to the absence of stable equilibrium points. Since it does not diverge to infinity nor converge to a point, it has to oscillate. □

The main idea of this result is that boundedness of every trajectory and instability of every equilibrium point force the states to oscillate (but not necessarily in a periodic manner). This idea has been used by Matsuoka (1985) and Efimov and Fradkov (2004); the latter provide a mathematically rigorous exposition of the idea starting from a formal definition of an "oscillatory system." We shall specialize Lemma 1 for the RIOs based on the SAFF and the DAFF neuron models to derive conditions on the synaptic parameter σ under which the existence of oscillation is guaranteed.

[2] In other words, for any initial state except for equilibria and points on the trajectory converging to an unstable equilibrium (e.g. separatrix).

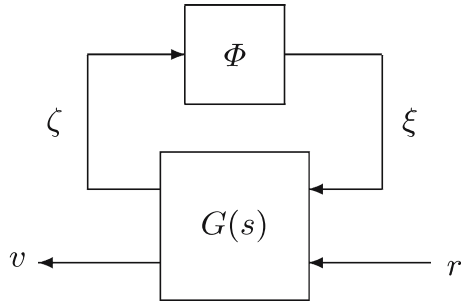

Fig. 7 General framework for analysis

3.2 Conditions for existence of oscillation

Let us first consider the SAFF model. In this case, the boundedness of the trajectory is always guaranteed regardless of the value of σ.

Lemma 2 *Consider the RIO with the SAFF neuron model. Then every signal in the RIO is bounded for any bounded input r and for any value of* σ.

Proof Since $\varphi \in \mathbf{S}$, its outputs v_i (i=1, 2) are always bounded. When r is bounded, the input to each neuron $\mathcal{N}$ is also bounded. Since $F(s)$ is stable, its output must be bounded as well. □

The stability property of an equilibrium point can be analyzed by examining the linearized system around the equilibrium (Khalil, 1996).

Lemma 3 *For the RIO with the SAFF neuron model, there is a unique equilibrium point for each constant input r. Moreover, the RIO system linearized around the equilibrium is exponentially unstable if and only if*

$$|\sigma k| > \frac{\tau_1 + \tau_2}{\varphi'(0)}$$

where φ' *is the derivative of* φ.

Combining the boundedness result (Lemma 2) and the instability result (Lemma 3) we have the following.

Theorem 1 *The RIO with the SAFF neuron model oscillates for a generic initial condition if*

$$|\sigma k| > \frac{\tau_1 + \tau_2}{\varphi'(0)}.$$

This result indicates that the synaptic connection must be strong enough to induce oscillations. If the parameters of the SAFB model are chosen as described in Sect. 2.3 to approximate the dynamics of the SAFF model, then the lower bound in (5) becomes

$$\sigma > \frac{\tau_1 + \tau_2}{k\varphi'(0)}$$

which is exactly the same as the lower bound for the SAFF case given in Theorem 1.

Next, we consider the DAFF model. In this case, the boundedness of trajectories is not always guaranteed, and the following result gives a boundedness condition. See the Appendix for a proof.

Lemma 4 *Consider the RIO with the DAFF neuron model. Then every signal in the RIO is bounded for any bounded input r if*

$$|\sigma k| < \frac{(1 + b\tau_1)(1 + b\tau_2)}{1 + b\tau_o}, \quad \tau_o := \frac{\tau_1\tau_2}{\tau_1 + \tau_2}. \tag{6}$$

Thus, boundedness of every trajectory is guaranteed if the synaptic connections are weak enough. The following result provides the condition for instability. A proof is given in the Appendix.

Lemma 5 *Consider the system obtained by linearizing the RIO with DAFF around an equilibrium point. Let* ψ_1' *and* ψ_2' *be the derivatives of* ψ_1 *and* ψ_2 *evaluated at the equilibrium. If*

$$\psi_2' < \psi_1' \quad or \quad \tau_+ + \tau_\times(\rho - \psi_1') < 0 \tag{7}$$

where $\tau_+ := \tau_1 + \tau_2$, *and* $\tau_\times := \tau_1\tau_2$, *then the system is unstable. If*

$$\psi_1' < \psi_2' \quad and \quad \psi_1' < \rho \tag{8}$$

then the system is unstable if and only if

$$|\sigma k| > \eta$$

where η *is defined by the root of a quadratic equation with the smaller magnitude:*

$$\eta := \min_{\lambda}\{\, |\lambda| : \; f(\lambda) = 0 \,\}, \tag{9}$$

$$f(\lambda) := \rho c_2\lambda^2 + (\rho c_1 + a_1 c_2)\lambda + a_1 c_1 - a_0 a_3^2$$
$$\alpha := \rho - \psi_1', \quad \beta := \rho(\psi_2' - \psi_1'),$$
$$a_0 := \beta, \quad a_1 := \alpha + \tau_+\beta, \quad a_3 := \tau_+ + \tau_\times\alpha,$$
$$c_1 := \tau_+(1 + \tau_1\alpha)(1 + \tau_2\alpha) + \tau_\times^2\alpha\beta, \quad c_2 := \tau_+ - \psi_1'\tau_\times.$$

For a typical DAFF model, every equilibrium point satisfies either (7) or (8), with at least one satisfying (8), and hence all the linearized systems are unstable if and only if $|\sigma k| > \eta$. Combining Lemmas 4 and 5, we have the following.

Theorem 2 *Consider the RIO with the DAFF neuron model. Suppose that every equilibrium point satisfies (7) or (8). Then its trajectory oscillates for a generic initial condition if*

$$\eta < |\sigma k| < \frac{(1 + b\tau_1)(1 + b\tau_2)}{1 + b\tau_o}$$

where η *and* τ_o *are defined in (9) and (6), respectively.*

For the DAFF-based RIO, the oscillation condition involves both lower and upper bounds on the synaptic strength. The lower bound comes from the instability condition, while the upper bound guarantees boundedness of every trajectory.

Table 1 Period and σ-interval for oscillation

	$\sigma_{\ell b}$	σ_{ub}	Period [s]
DAFF	8.368	33.00	0.187
SAFF	7.480	∞	0.199
SAFB ($b_a = 10^4$)	7.480	74807	0.209
SAFB ($b_a = 5$)	7.480	44.25	0.231

DAFF dynamic axon with feedforward adaption, *SAFF* static axon with feedforward adaption; *SAFB* static axon with feed back adaption

3.3 Basic oscillation profiles

The RIOs with different neuron models were numerically simulated for the model parameter values given in the Appendix. The synaptic parameter was chosen as $\sigma = 8$, which is near the threshold values for oscillations; see Table 1 for the lower and upper bounds on σ, for which the existence of oscillation is guaranteed, calculated from (5) and Theorems 1 and 2. A short pulse was applied to r_1 to initiate the oscillation but both r_1 and r_2 were zero for the rest of each simulation. The oscillation behavior of v_1 and v_2 in the steady state for each RIO is plotted in Fig. 8. The cycle periods of the oscillations were measured and are summarized in Table 1.

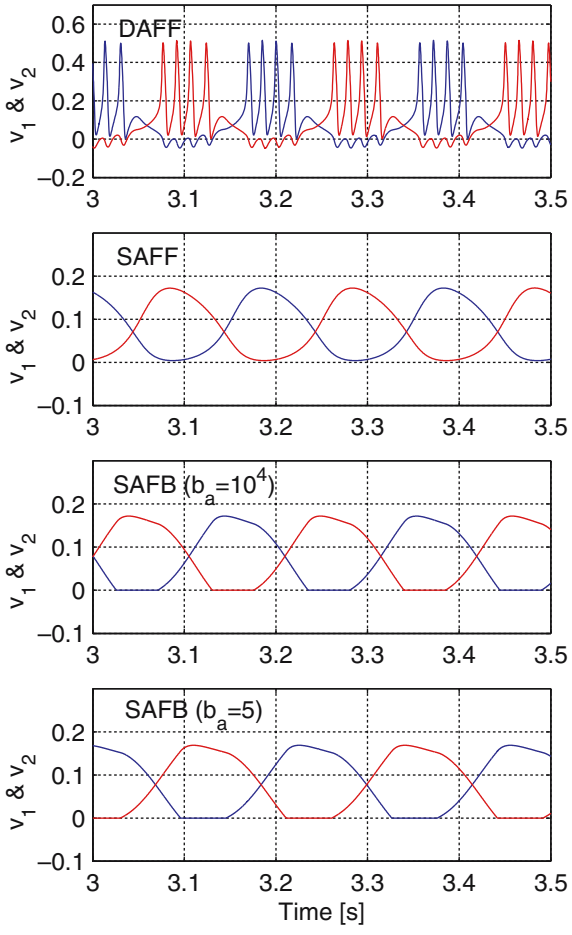

Fig. 8 Oscillation profiles of RIOs

We see that the oscillation of the DAFF-based RIO is well approximated by that of the SAFF-based RIO in terms of the average amplitude and the period. The SAFB-based RIOs also reasonably approximate the SAFF-based RIO for both values of b_a. However, it is observed that the deviation of the period of the SAFB-based RIO from that of the SAFF-based RIO becomes significant if b_a is further decreased.

Numerical simulations with other values of σ (not shown) indicate that the predicted σ-ranges of oscillations are quite accurate. All the upper and lower bounds of σ are in agreement with the simulation results within at most 1% error except for the lower bound for the DAFF case. The DAFF-based RIO exhibits bistability when σ is just below $\sigma_{\ell b}$. In particular, $\sigma < \sigma_{\ell b}$ implies that there is a stable equilibrium point, but at the same time, a stable oscillatory trajectory is also observed as shown in Fig. 8 for the case $\sigma = 8$. Thus, the RIO has both a stable equilibrium and a stable oscillatory trajectory. The bistability property is observed for the range $6.5 \leq \sigma < \sigma_{\ell b}$. The other three RIOs do not seem to possess the bistability property.

It is observed (not shown here) that, for each RIO model, the oscillations in v_1 and v_2 are out-of-phase if the connection is inhibitory ($\sigma > 0$) and are in-phase (or synchronized) if excitatory ($\sigma < 0$). For the RIOs with SAFF or SAFB neuron models, the main difference between the inhibitory/excitatory cases is the phase, and the wave forms can both be described as a distorted sinusoid. However, for the DAFF-based RIO model, there is a significant difference in the wave forms as well as the phase. Specifically, while alternate bursting of nerve impulses are observed when $\sigma > 0$, synchronized plateaux with no spikes are observed when $\sigma < 0$.

4 Robust entrainment of RIOs to pendulum

Observations of animal locomotion systems suggest that the neuronal controller is able to entrain to the natural motion of the body. Therefore, each of our RIO models is expected to achieve entrainment to a pendulum if it is properly coupled with the pendulum and captures essential dynamical properties of neuronal circuits. The objective of this section is to verify this expectation by numerical experiments and to identify the neuronal properties essential for entrainment through comparative studies of the four RIO models.

4.1 System description

Consider a simple pendulum whose normalized equation of motion is given by

$$\ddot{\theta} + 2\zeta\omega_o\dot{\theta} + \omega_o^2 \sin\theta = \omega_o^2\tau. \tag{10}$$

where $\omega_o > 0$ and $\zeta > 0$ are the natural frequency and the damping ratio of the system linearized around the stable equilibrium without input, θ is the angular displacement, and τ is the torque input applied at the pivot. The period of the linearized pendulum is $T_o := 2\pi/\omega_o$. We shall call T_o the

fundamental period of the pendulum. The natural period of the pendulum (when $\zeta = 0$ and $\tau = 0$) is close to T_o if the amplitude is small, and becomes larger as the amplitude gets larger.

The objective is to design a feedback controller that generates the torque input τ, using the information on the pendulum angle θ, to achieve a maintained oscillation in the presence of the energy loss due to the positive damping. Intuitively, the most energy efficient solution would be to achieve entrainment to the natural motion of the pendulum, i.e., to make the pendulum oscillate at its (undamped) natural frequency, provided the damping is sufficiently small.

Consider the feedback system of the pendulum and an RIO depicted in Fig. 9. The pendulum angle θ is measured and conditioned by a saturation function

$$\chi(\theta) := \varrho \tanh(\lambda\theta)$$

before entering the RIO. Both neurons in the RIO are directly driven by the sensory signal, but with different signs. The torque τ is generated by rectifying the RIO outputs through $\pi^+(v) := \max(v, 0)$, taking the difference, and multiplying the feedback gain μ. This rectifying nonlinearity is introduced to reflect the fact that the muscle can only contract but cannot actively extend. Finally, a pulse signal is supplied at r_2 to initiate the oscillation.

4.2 Numerical experiments

In this section, we shall first verify through numerical simulations that our RIO models are able to achieve entrainment to the pendulum with some choice of parameters. We will then examine robustness of the entrainment property with respect to perturbations in the fundamental period T_o of the pendulum and the feedback gain μ. Based on the comparison of different RIO models, a conclusion will be drawn regarding which neuronal properties appear to be crucial for achieving robust entrainment.

For the numerical experiments described below, we set some of the parameters as

$$\zeta = 0.1, \quad \varrho = 0.4, \quad \lambda = 2$$

and consider various values for the remaining parameters (T_o, μ). Note that the pendulum is lightly damped. If the sensor gains are too small, then the RIOs insist on oscillating at their intrinsic frequencies and act like open-loop control systems. On the other hand, if the input to the DAFF-based RIO is too large, then the spike trains tend to be lost and the system states go out of the intended operation range. The above parameter values have been roughly chosen with these considerations in mind.

For each of the RIOs, we have manually searched for the values of parameters (T_o, μ) so that the entrainment to the pendulum occurs. The procedure is as follows. For each (T_o, μ), run a simulation of the system in Fig. 9 and, if the pendulum oscillates, measure its period T and amplitude ϑ. Simulate the free motion of the undamped pendulum ($\zeta = 0$, $\tau = 0$) with the initial condition $\theta(0) = \vartheta$ and $\dot\theta(0) = 0$ to calculate the natural period T_n at this amplitude. If the period error $e := (T - T_n)/T_n$ is small, we declare that the entrainment is achieved.

It turned out to be relatively easy to find the parameter values by fixing T_o to be a few times larger than the intrinsic periods of the RIOs ($\cong$ 200 ms) and gradually increasing the feedback gain μ. When μ is small, each RIO oscillates with its intrinsic period. If μ becomes larger than a threshold value, the period of the RIO suddenly becomes close to the natural period of the pendulum. Entrainment is observed for a range of μ where the oscillation amplitude increases but the period is kept close to the natural period as μ gets larger. The upper bound of the range is typically defined by the event that the amplitude goes beyond 180°. We have verified that SAFB-based RIO with $b_a = 5$ is able to achieve entrainment with less than $e = 3\%$ error from the natural period of the pendulum, and the other three RIOs with less than 1% error.

Figure 10 shows a sample of the entrainment behavior for the DAFF-based RIO with T_o=400 ms and μ=1. Initiated by a small pulse into the RIO, the system starts to oscillate, and

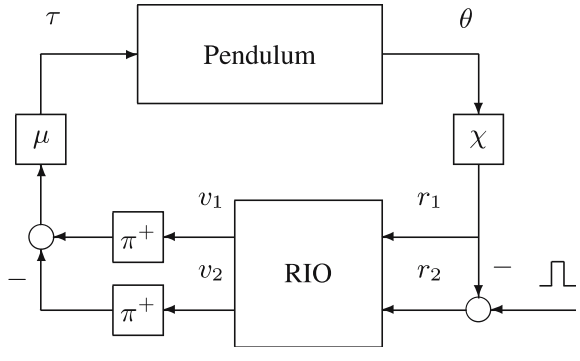

Fig. 9 Pendulum driven by RIO

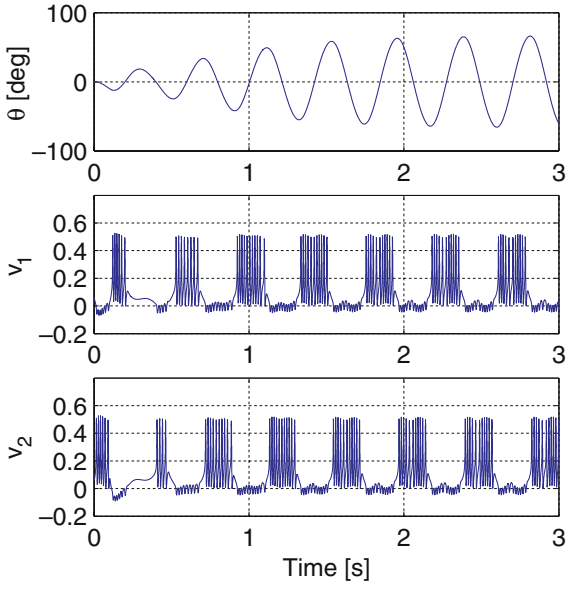

Fig. 10 Entrainment of the DAFF-based RIO to the pendulum; T_o = 0.4 s, μ=1

the amplitude of the pendulum oscillation gradually increases to reach a steady state. The θ-amplitude is ϑ=67° and the corresponding natural period is T_n=438 ms. The period of the oscillation achieved by the RIO is 432 ms, yielding the entrainment error of $e = -1.4\%$.

We now examine the robustness. We grid the (T_o, μ) parameter plane, run a simulation for each grid point, and calculate the period error e. The result is summarized by the contour plots of e as a function of (T_o, μ) as shown in Fig. 11. The contours are plotted at levels e=5, 6, . . . , 10%. Also shown are the boundaries for the region with nearly natural oscillation. As mentioned earlier, oscillations occur on the left of the region but the frequencies are close to the intrinsic frequency of the RIO. On the right of the region, the pendulum angle goes beyond 180° and the simulations are terminated. The dark region and the lightly-shaded region correspond to the sets of points such that $e < 1\%$ and $1\% < e < 3\%$, respectively. The results are found insensitive to initial conditions. Some of the complex borders appear to be an artifact resulting from the coarse grid and the simulation run time shorter than required to reach the steady state. However, such "noises" are small enough to make the following important observations.

We see that both RIOs with DAFF and SAFF achieve entrainment robustly against the perturbations in (T_o, μ). This means that *the dynamics for generating nerve impulses are not essential for robust entrainment*, and that *the basic mechanism for robust entrainment is effective even under the influence of the additional spike dynamics.* The SAFB-based RIOs also exhibit some capability of robust entrainment, although the robustness is diminished especially for small b_a, when compared with the RIOs with DAFF and SAFF. Recall that the DAFF and SAFF models have a band-pass filter with a zero at the origin. The linearized SAFB model also has a zero near the origin when b_a is large, but the zero moves away as b_a gets smaller. In view of these facts, it seems that having zero at the origin in the neuron model is crucial for achieving robust entrainment. In other words, *the "complete adaptation" property of neuronal dynamics appears to be essential.* Moreover, *the detailed mechanism for embedding the adaptation property (i.e., feedforward or feedback) appears nonessential.*

5 Mechanism of robust entrainment

In the previous section, we have verified through numerical simulations that our RIO models are capable of achieving robust entrainment to the pendulum. Complete adaptation has been identified as an essential property of the neuronal dynamics, while the properties that were deemed unimportant include the dynamics for spike generation and the choice of the feedforward or feedback mechanism for adaptation. Based on these findings, we now proceed to the theoretical analysis of the entrainment phenomenon for the feedback system of the pendulum and SAFF-based RIO. Our aim here is to uncover the dynamical mechanism for the robust entrainment and to explain why the adaptation property is important.

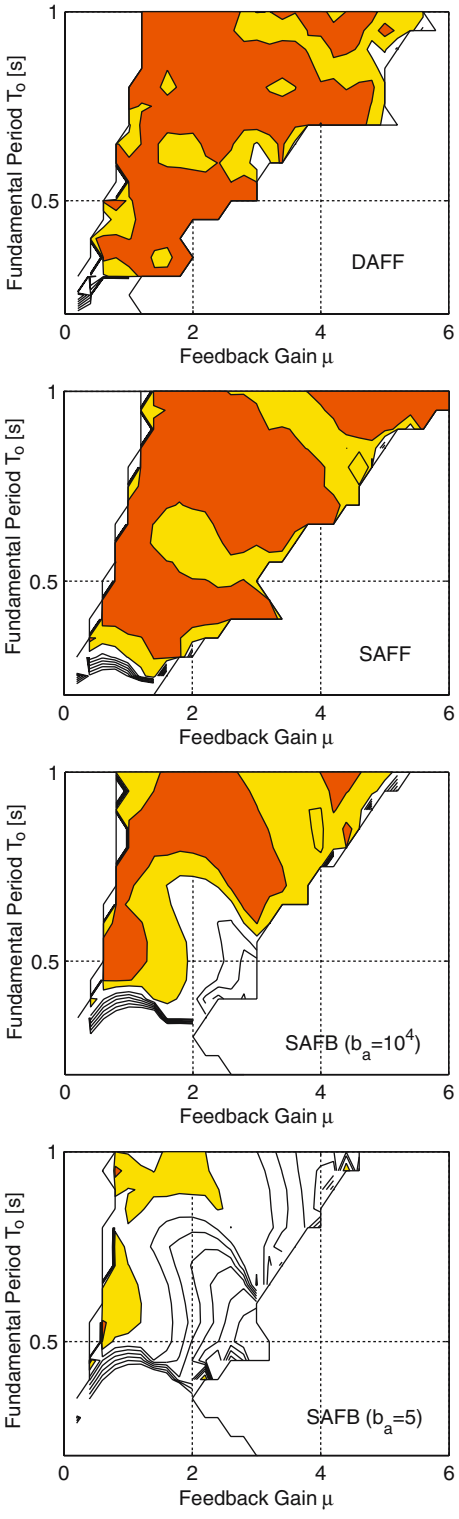

Fig. 11 Contour plots of period error $e(T_o, \mu)$. Contours at levels e=5, 6, . . . , 10%. *Dark region*: e<1%; *Lightly shaded region*: $1\% < e < 3\%$

We hypothesize that the basic mechanism of entrainment is embedded in the phasic interaction between the oscillatory dynamics of the pendulum and RIO, and that the nonlinearities

of the pendulum, sensor, and actuator have secondary effects only. Based on this hypothesis, we proceed to remove those nonlinearities and analyze the resulting simpler feedback system. In particular, we shall replace the pendulum dynamics by its linearization

$$P(s) := \frac{\omega_o^2}{s^2 + 2\zeta\omega_o s + \omega_o^2},$$

remove the actuator nonlinearity π^+ in Fig. 9, and assume linear sensor dynamics $\chi(\theta) = \eta\theta$ with a constant gain η. In the following subsections, we will analyze the entrainment property of the feedback system consisting of $P(s)$ and SAFF-based RIO with the linear gains μ and η.

5.1 Oscillation analysis by harmonic balance

Consider the feedback system in Fig. 9 where "Pendulum" and χ are given by $P(s)$ and $\chi(\theta) = \eta\theta$, respectively, and the RIO is given by Fig. 6 with $\mathcal{N}$ being the SAFF neuron model (4). In the analysis below, we start with arbitrary transfer functions $F(s)$ and $P(s)$ and arbitrary nonlinear function $\varphi(q)$, and later assume that φ is monotonically increasing and $F(0) = 0$. Ignoring the actuator nonlinearity π^+, its dynamical behavior is described by

$$q_1 = F(s)(r_1 - \sigma\varphi(q_2)), \quad q_2 = F(s)(r_2 - \sigma\varphi(q_1)), \tag{11}$$

$$r_1 = \eta\theta, \quad r_2 = -\eta\theta, \quad \theta = P(s)\tau, \tag{12}$$

$$\tau = \mu(\varphi(q_1) - \varphi(q_2)). \tag{13}$$

We are interested in the periodic solutions of these equations. In particular, we would like to predict the frequency and amplitude of such oscillations. To this end, let us assume that the above equations admit T-periodic solutions q_i, r_i, τ, and θ. We will develop the condition for harmonic balance (see e.g., Khalil 1996) that the solutions must satisfy, and use the condition to estimate the frequency and amplitude.

Recall that a generic T-periodic signal $x(t)$ can be expressed as a Fourier series

$$x(t) = \sum_{k=0}^{\infty} [\alpha_k \sin(k\omega t) + \beta_k \cos(k\omega t)]$$

for appropriately chosen parameters α_k and β_k where $\omega := 2\pi/T$. Let us denote by $\hat{x} := \alpha_1 + j\beta_1$ the phasor representation of the fundamental frequency component. Define

$$v_i := \varphi(q_i), \quad \kappa_i := \hat{v}_i/\hat{q}_i \tag{14}$$

for $i = 1, 2$. By the symmetry of the feedback system structure, we may assume that q_1 and q_2 have the same time-course with a possible time-shift. In this case, we have $\kappa_1 = \kappa_2 =: \kappa$.

Now, substitution of (14) into (11) yields

$$\begin{bmatrix} 1 & \sigma\kappa F(j\omega) \\ \sigma\kappa F(j\omega) & 1 \end{bmatrix} \begin{bmatrix} \hat{q}_1 \\ \hat{q}_2 \end{bmatrix} = F(j\omega) \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \end{bmatrix}. \tag{15}$$

Solving the equation for $\hat{q}_i$ and substituting the results into (12) and (13), we have

$$\hat{\tau} = \mu\kappa(\hat{q}_1 - \hat{q}_2) = \frac{2\mu\kappa\eta P(j\omega)F(j\omega)}{1 - \sigma\kappa F(j\omega)}\hat{\tau}.$$

Thus the oscillation frequency ω must satisfy the following harmonic balance equation:

$$\kappa F(j\omega)H(j\omega) = 1, \quad H(j\omega) := \sigma + 2\mu\eta P(j\omega). \tag{16}$$

A solution ω to (16) provides an estimate for the oscillation frequency. However, it is difficult in general to solve it because κ depends on the unknown time-course of q_i. As in the standard describing function method (see e.g., Khalil 1996), we now introduce an approximation where the second and higher harmonic terms in q_i are ignored. This approximation would be reasonable when $P(s)$ and/or $F(s)$ have low-pass filter characteristics with sufficient attenuation at twice the oscillation frequency. In this case, we have

$$\kappa = \frac{c_1 + jd_1}{a}, \tag{17}$$

$$\varphi(a\sin\omega t + b) = \sum_{k=0}^{\infty} [c_k \sin(k\omega t) + d_k \cos(k\omega t)]. \tag{18}$$

By Lemma 6 in the Appendix, κ is a real number. The harmonic balance equation (16) then implies that $F(j\omega)H(j\omega)$ must be real-valued. Moreover, if the function φ is monotonically increasing as is the case for our SAFF model, κ is positive. In this case, comparing the phase components on the both sides of (16), we have

$$\angle[F(j\omega)H(j\omega)] = 0 \tag{19}$$

where $\angle[\cdot]$ denotes the phase angle of the complex number in the argument. This condition gives candidate(s) for possible oscillation frequencies.

The harmonic balance equation provides an estimate for the oscillation amplitude as well for each estimated oscillation frequency ω_e. To explain the idea, let us assume $F(0)=0$, which is the case for the SAFF neuron model.[3] In this case, the bias term b in $q_i(t)$ is zero, and κ can be considered as a function of a. The amplitude of q_i oscillations can then be estimated to be the value of a consistent with (16) for $\omega = \omega_e$. In particular, (16) gives an estimate for κ as $\kappa_e := 1/(F(j\omega_e)H(j\omega_e))$, and an estimate for the amplitude is given by a_e satisfying $\kappa(a_e) = \kappa_e$. If there is no such a_e, then a possible reason is that there is no oscillation at frequency ω_e.

The stability of oscillation at frequency ω_e may be predicted by following the argument in Glad and Ljung (2000). Consider the roots of characteristic equation $\kappa F(s)H(s) = 1$ and let $\lambda(\kappa)$ be the maximum value of their real parts. Then, the oscillation at frequency ω_e is expected to be stable if

$$\lambda(\kappa_e) = 0, \quad \kappa'(a_e)\lambda'(\kappa_e) < 0, \tag{20}$$

[3] In principle, the amplitude can still be estimated without this assumption by determining the values of κ and its zero-frequency version, and then solving for a and b.

where the prime denotes the derivative. We explain the underlying idea for the case $\lambda'(\kappa_e)>0$. The other case $\lambda'(\kappa_e)<0$ can be explained similarly. Note that the approximated closed-loop system is stable if and only if $\lambda(\kappa)<0$. When $\kappa = \kappa_e$, the characteristic equation has some root(s) on the imaginary axis. The first condition $\lambda(\kappa_e)=0$ implies that all the other roots are in the open left half complex plane. The property $\lambda'(\kappa_e) > 0$ then implies that a small positive (negative) perturbation of κ leads to instability (stability). Now, if the oscillation is perturbed to have a slightly larger amplitude, then κ decreases due to $\kappa'(a_e)<0$. It may be reasonable to assume that the stability property then tends to decrease the oscillation amplitude. If it is perturbed to have a slightly smaller amplitude, then instability may lead to an increasing amplitude. Thus, the oscillation amplitude always tends to return to the original value after a perturbation, indicating stability of the oscillation.

In summary, the frequency of a possible oscillation is estimated to be ω_e such that the phase angle of $F(j\omega_e)H(j\omega_e)$ is zero. The amplitude of q_i-oscillation is estimated to be the value of a_e for which $\kappa(a_e) = 1/(F(j\omega_e)H(j\omega_e))$ holds, where $\kappa(a)$ is the describing function defined by (17) and (18). The oscillation is expected to be stable if (20) holds where $\kappa_e := \kappa(a_e)$ and $\lambda(\kappa)$ is the maximum real part of the roots of $1 = \kappa F(s)H(s)$.

5.2 Harmonic balance predicts entrainment

We now apply the harmonic balance method in the previous section to a particular case of parameters and verify that the method predicts entrainment. Robustness of the entrainment property against parameter perturbations will be addressed in the next section. Below, we set μ=2, η=1, and T_o=1 s (or $\omega_o=2\pi$ rad/s), and the other parameters are specified as before.

Figure 12 shows the Bode plots of $F(s)$, $P(s)$, and $H(s)$. The harmonic phase balance (19) is achieved at the three frequencies $\omega_e = 6.11$, 7.93, and 31.6 rad/s, where the phase plots of $H(s)$ and $F(s)^{-1}$ intersect so that $F(j\omega_e)H(j\omega_e)$ is a positive real number. These are the candidates for the estimated oscillation frequency. Figure 13 shows the Nyquist plot of $F(s)H(s)$. The harmonic phase balance occurs when the curve intersects with the positive real axis (marked by points A, B, and C). The amplitude of q_i-oscillation, a_e, is estimated from the harmonic gain balance and the describing function for φ. Figure 14 shows the describing function $\kappa(a)$, numerically calculated by the Fast Fourier Transform. The oscillation amplitude of q_i is estimated to be the value of a that gives $\kappa(a) = 1/(F(j\omega_e)H(j\omega_e)) =: \kappa_e$. For instance, the oscillation at point A of Fig. 13 with ω_e=6.11 rad/s is estimated to have q_i-amplitude of $a_e = 0.22$ because $\kappa_e \cong 1/1.2 \cong 0.83$ and $\kappa(a_e) = \kappa_e$. Figure 15 shows the simulation of the RIO-pendulum system and confirms that these estimates are close to the actual values (frequency 6.11 rad/s and q_1-amplitude 0.21).

On the other hand, oscillations at other frequencies corresponding to points B and C in Fig. 13 were not observed in simulations with various initial state/pulse conditions we examined. This is consistent with the prediction from the harmonic balance equation and the describing function. Recall that the characteristic equation of the approximated closed-loop system is given by $1=\kappa F(s)H(s)$. Hence, by the Nyquist criterion (see e.g., Ogata 1996), the system is stable if and only if the Nyquist plot of $F(s)H(s)$ does not encircle the point $1/\kappa + j0$, provided $F(s)H(s)$ is stable. It then follows that point A in Fig. 13 divides the positive real axis into stable and unstable segments: the system is stable (unstable) if $1/\kappa + j0$ is on the right (left) of point A. Hence $\lambda(\kappa_e)=0$ and $\lambda'(\kappa_e)>0$ hold at this point. Moreover, Fig. 14 shows that the slope of the describing function is negative at $a = a_e$, i.e., $\kappa'(a_e)<0$. Therefore, the oscillation at point A is expected to be stable. However, points B and C do not satisfy these requirements. In particular, there is no transition between stability and instability at point B, violating (20). For point

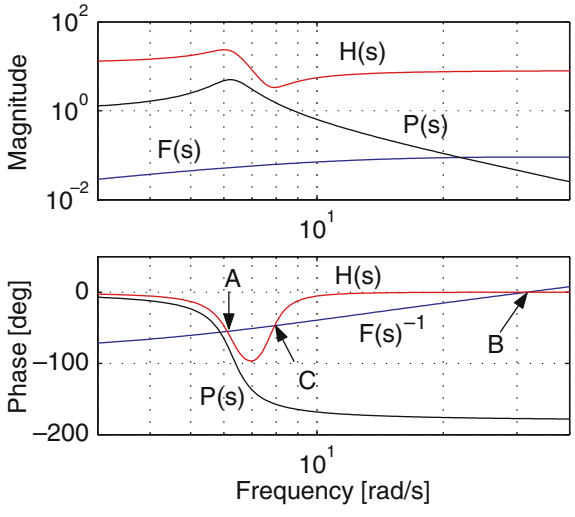

Fig. 12 Bode plots of $F(s)$, $P(s)$, and $H(s)$

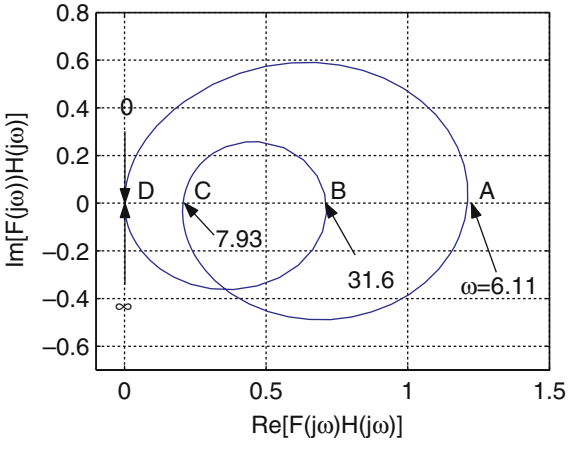

Fig. 13 Nyquist plot of $F(s)H(s)$

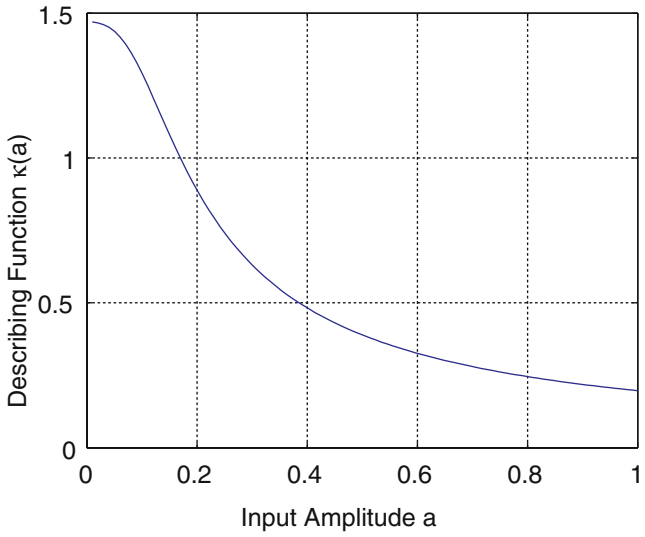

Fig. 14 Describing function $\kappa = c/a$ for $\varphi(a\sin\omega t) \cong c\sin\omega t + d$

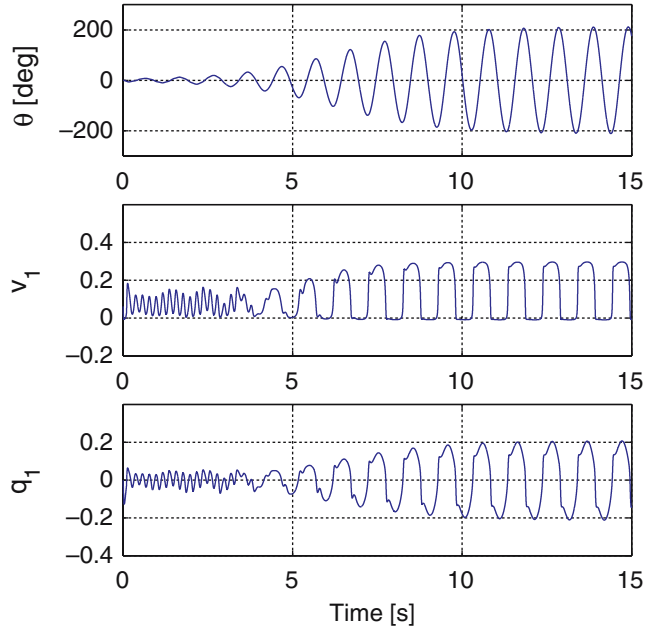

Fig. 15 Entrainment of RIO/SAFF to linear pendulum

C, κ_e=1/0.21=4.7 and no value of a_e satisfies $\kappa(a_e)$=κ_e. Hence, stable oscillations are not expected at points B and C.

In summary, the harmonic balance condition predicts that a stable oscillation occurs at point A of Figs. 12 and 13 where the phase angle of $F(j\omega)H(j\omega)$ becomes zero for the first time as ω increases from zero. The frequency at which this occurs, ω_e, provides an estimate for the oscillation frequency. It turns out that this is close to the natural frequency of $P(s)$, predicting the entrainment, because the first phase balance occurs when the phase of $P(s)$ changes abruptly. This estimate also turns out to be close to the actual oscillation frequency. Indeed, for the particular case considered here, the undamped natural frequency of $P(s)$ is 2π rad/s, and the estimated and actual frequencies are both 6.11 rad/s.

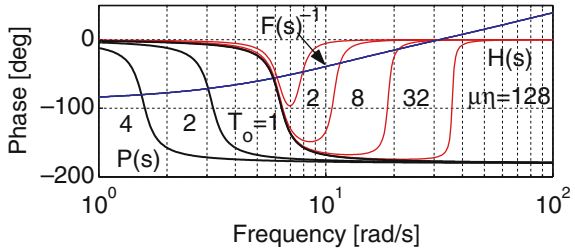

Fig. 16 Effects of T_o and $\mu\eta$ on the phase

5.3 Essential mechanism for robust entrainment

Let us now address the mechanism for robust entrainment. Our first claim is that the entrainment occurs regardless of the natural frequency ω_o of the pendulum *if the loop gain $\mu\eta$ is sufficiently large*. This can be explained by Fig. 16 where the phase plots are shown for various values of $\mu\eta$ and ω_o (or equivalently $T_o := 2\pi/\omega_o$). The phase of $H(s)$ is plotted for $\mu\eta$=2, 8, 32, and 128 with fixed T_o=1 s, while that of $P(s)$ is for T_o=1, 2, and 4 s.

If the loop gain is small, the harmonic phase balance occurs only at ω_e=31.6 rad/s which predicts a stable oscillation with period T=199 ms. Note that this coincides with the intrinsic period of the RIO (see Table 1). This is not a coincidence. The harmonic equation (15) with nonzero oscillations ($\hat{q}_i \neq 0$) and zero input to the RIO ($\hat{r}_i$=0) implies that $F(j\omega)$ is real. Hence the intrinsic frequency of the RIO, ω_{RIO}, can be estimated from Im$[F(j\omega_{\text{RIO}})]$=0 as $\omega_{\text{RIO}} \cong 1/\sqrt{\tau_1\tau_2}$ = 31.6 rad/s.

On the other hand, as the loop gain gets larger, the phase of the transfer function $H(s)$ approaches that of $P(s)$, and the harmonic phase balance occurs around $\omega_e \cong 6$ rad/s, predicting a stable oscillation around the natural frequency ω_o=2π of $P(s)$. Note that the intersection of the phase curves for $P(s)$ and $F(s)^{-1}$, defining ω_e, will always be around ω_o where the phase of $P(s)$ changes abruptly from 0^o to -180^o, regardless of the value of ω_o. Thus, entrainment ($\omega_e \cong \omega_o$) is expected to occur robustly against perturbations in ω_o under a high loop-gain condition. Numerical simulations confirm that this is indeed the case (not shown).

An essential prerequisite to the above argument is that $P(s)$ *is lightly damped.* If the damping ratio ζ is sufficiently small, then the phase of $P(s)$, and hence $H(s)$, will change from 0°—180° very quickly at the natural frequency ω_o, so that the harmonic phase balance occurs at this frequency. On the other hand, if ζ is not small, then the phase of $H(s)$ would not change abruptly, and the intersection with the phase of $F(s)^{-1}$ may occur at a frequency ω_e that is away from ω_o. The situation is illustrated in Fig. 17 where the phase response of $H(s)$ is plotted for various values of ζ with fixed loop-gain $\mu\eta$=10. The vertical solid line indicates the natural frequency ω_o. The estimated oscillation period $T_e := 2\pi/\omega_e$ is plotted as a function of ζ in Fig. 18. Clearly, both the estimated period T_e and actual (simulated) period T are close to the natural period T_o if ζ is sufficiently small. However, the difference between T_e, T and T_o can be large if ζ is not small. The fig-

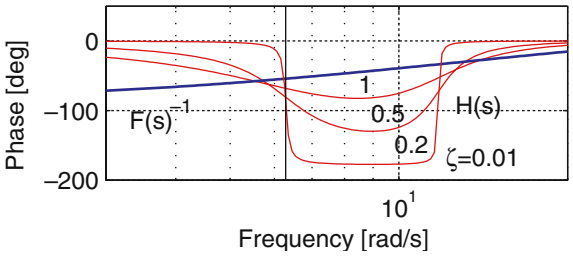

Fig. 17 Phase responses for various damping ζ where $\mu\eta = 10, z_o = 0, T_o = 1$

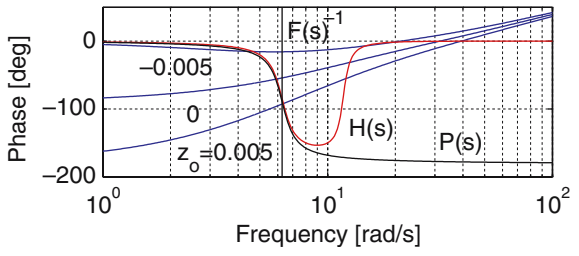

Fig. 19 Phase responses for various zero location z_o where $\mu\eta$=10, ζ=0.1, $T_o = 1$

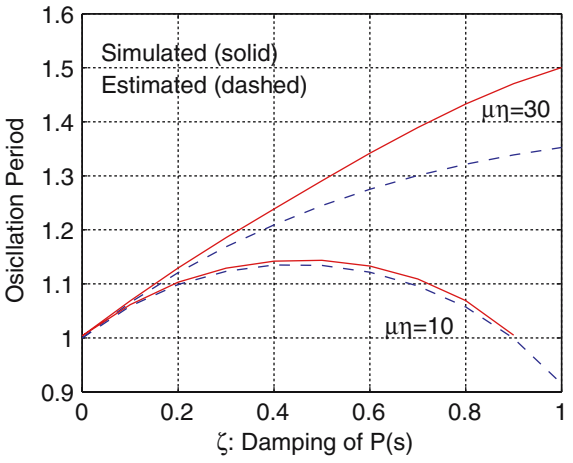

Fig. 18 Estimated and actual periods as functions of ζ where $\mu\eta = 10, 30, z_o = 0, T_o = 1$

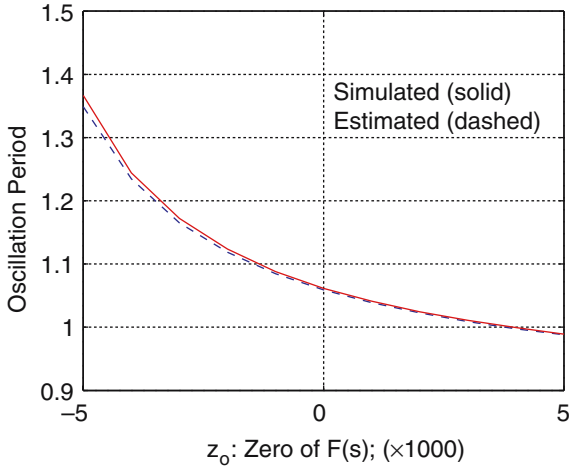

Fig. 20 Estimated and actual periods as functions of z_o where $\mu\eta$=10, ζ=0.1, $T_o = 1$

ure also indicates that the discrepancy is larger with a higher loop-gain.

We now turn our attention to the "complete adaptation" property that has been deemed essential for robust entrainment via numerical experiments. As noted before, this property is a result of the zero of $F(s)$ at the origin. *The role of this zero is to place the phase of* $F(s)^{-1}$ *near* $-90°$ *in the low frequency range.* Noting that the phase of $P(s)$ is $-90°$ at the natural frequency, the intersection of the phase plots of $F(s)^{-1}$ and $P(s)$ would be found near the natural frequency if ω_o is sufficiently small when compared with the poles of $F(s)$. In this case, the entrainment is predicted and actually occurs. On the other hand, if $F(s)$ had a zero away from the origin or had no zeros, the phase of $F(s)^{-1}$ may not be close to $-90°$, resulting in a phase intersection away from the natural frequency, or no intersection. It then follows that the entrainment is not predicted, and in fact, does not occur.

Figure 19 shows the phase responses of $F(s) := k(s - z_o)/((1+\tau_1 s)(1+\tau_2 s))$ for several values of the zero z_o. We see that the intersection of $H(s)$ and $F(s)^{-1}$ curves move to the left, and hence ω_e decreases, as z_o decreases. Figure 20 clearly shows this tendency of estimated period T_e together with the actual period T obtained by simulations. Note that the entrainment ($T \cong T_o = 1$) does not occur when the zero of $F(s)$ is away from the origin in the negative direction. This may explain why the SAFB model does not achieve robust entrainment when b_a=5 but it does when b_a=10^4. While positive values of z_o in the range $0<z_o<0.005$ lead to the entrainment for the case T_o=1 (Fig. 20), the property can be lost after a perturbation of T_o.[4] On the other hand, if $z_o = 0$, then the robust entrainment is achieved for any natural frequency $\omega_o \ll \omega_{\text{RIO}} = 31.6$ rad/s because the phase of $F(s)^{-1}$ is near $-90°$ in this frequency range. Thus, our analysis by harmonic balance explains why it is essential for $F(s)$ to have a zero at (or near) the origin to achieve robust entrainment over a wide range of natural frequencies.

Finally, it should be noted that *the positivity of* $\mu\eta$ *is crucial but the sign of* σ *is not important for entrainment.* If $\mu\eta$ were negative, then the phase angle of $H(j\omega)$ would vary between 0 and $+180°$ and approaches that of $-P(j\omega)$ as $|\mu\eta|$ gets larger. Consequently, the harmonic phase balance (19) does not occur near the natural frequency, and indeed, the entrainment is not observed in simulations. On the other hand, under the high loop-gain $\mu\eta>0$ relative to the neuronal coupling σ, the phase of the transfer function $H(s)$ is close to that of $P(s)$ regardless of the sign of σ so that the entrainment mechanism discussed above applies. This means that not only a reciprocal inhibition oscillator ($\sigma>0$) but also a reciprocal excitation oscillator ($\sigma<0$) can achieve entrainment. Moreover, even a pair of uncoupled neurons

[4] For instance, if $\omega_o = 2$ rad/s and $z_o = 0.005$, then the phase plot of $H(s)$ in Fig. 19 moves to the left to give $\omega_e = 2.33$ rad/s, which is 16% off from ω_o.

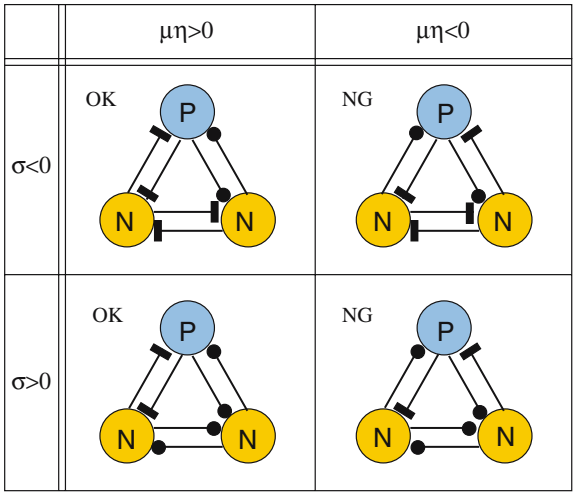

Fig. 21 Connectivity architectures of pendulum-CPG system for different signs of $\mu\eta$ and σ. "OK" and "NG" mean that the entrainment is and is not achieved, respectively. The *circles* with P and N stand for the pendulum and neuron, respectively. The *lines* terminating with a *circle* and *bar* indicate inhibitory and excitatory connections, respectively

(σ=0) can induce natural oscillation of the pendulum. These predictions from the harmonic balance analysis have been validated by numerical simulations (not shown). The point of this discussion may be summarized in Fig. 21 that shows which connectivity architecture does (or does not) lead to entrainment.

6 Conclusion

We have considered several RIOs with different neuron models and derived sufficient conditions for the existence of oscillations. The ability of each RIO to achieve robust entrainment was examined through numerical experiments of driving a simple pendulum. Comparative studies of the results for various RIOs indicated the following.

- The complete adaptation property is essential for achieving robust entrainment.
- The neuronal dynamics for generating nerve impulses are not important.
- The specific mechanism of adaptation (i.e., feedforward or feedback) is not important.

Based on these findings, we have chosen the SAFF-based RIO model for further theoretical analysis. The method of harmonic balance has been employed to reveal the fundamental mechanism of robust entrainment. Specifically, we have found the following.

- The pendulum-RIO oscillation occurs at the lowest frequency at which the harmonic phase balance (19) is satisfied.
- This frequency is close to the pendulum natural frequency ω_o if:
  - the pendulum is lightly damped;
  - the coupling within the RIO ($|\sigma|$) is sufficiently weak, and the coupling between the RIO and pendulum ($\mu\eta > 0$) is sufficiently strong;
  - the neuronal dynamics have the phase of $F(j\omega_o)$ near 90^o.
- The phase condition $\angle[F(j\omega_o)] \cong 90^o$ is satisfied when the $F(s)$ has a zero at the origin, or equivalently, when the neuron model has the complete adaptation property.
- The coupling between the pendulum and each neuron should be a positive feedback loop ($\mu\eta > 0$), while the coupling between the two neurons can be positive or negative, or even zero.

It is interesting to note that, under the extreme condition $|\sigma| \ll \mu\eta$, the mechanism of robust entrainment is essentially the same as the one used in engineering applications (Raney and Slominski, 2004; Ono, 1998) – positive rate feedback with saturation. The idea is to destabilize the equilibrium point by adding a locally negative damping through feedback and to guarantee the boundedness of every trajectory by the saturation nonlinearity. Since the feedback does not affect the stiffness, the resulting oscillation is expected to occur at the natural frequency. This type of control strategy also resulted from the speed gradient algorithm (Fradkov, 1979) to achieve the feedback resonance (Fradkov, 1999b). Within the framework of our analysis, this corresponds to having $F(s) = s$ and σ=0 so that the harmonic phase balance occurs at the natural frequency: $\angle[F(j\omega_o)H(j\omega_o)] = \angle[jP(j\omega_o)] = 0$. Thus the complete adaptation property may be viewed as a result of the neuronal dynamics $F(s)$ acting as an approximate differentiator in the neighborhood of ω_o.

We have gained some insights into the biological mechanism of robust entrainment by focusing on the simplest possible yet relevant setting (i.e., RIO-pendulum system). The principle uncovered in this work would be useful for engineering designs to achieve natural rhythmic motions. For instance, the RIO-based control architecture explored here may provide a way to improve the engineering approach of positive-rate feedback with saturation for robust entrainment. Possible improvements include incorporation of additional functionalities such as the speed (or frequency) control and the phase coordination among multiple limbs (or segments) through the coupled-oscillator framework (Ermentrout and Kopell, 1984; Cohen et al., 1992). The problem of how to generate natural rhythmic motion for general locomotion systems through feedback control is still largely open and awaits further studies in the intersection of biology and engineering.

Acknowledgements This material is based on work supported by the National Science Foundation under Grants No. 0201386 and No. 0237708, and by the National Institutes of Health NINDS R01 NS46057-01 as part of the NSF/NIH Collaborative Research in Computational Neuroscience Program.

Appendix A: Neuron model parameters

The parameter values for the neuron models used in our numerical experiments are summarized below. The time unit is millisecond for all of the following models.

The (soma, dendrites) model $\mathcal{F}$ (Iwasaki and Zheng, 2002):

$k = 10, \quad \tau_1 = 10, \quad \tau_2 = 100.$

The dynamic axon model $\mathcal{A}$ (Iwasaki and Zheng, 2002):

$$\begin{matrix} \psi_1(v) = c\phi(av) + q_o - bv \\ \psi_2(v) = \phi(d(v + v_o)) \end{matrix}, \quad \phi(x) := \frac{1}{1 + e^{2-4x}}$$

$q_o = -0.2, \quad v_o = -0.35, \quad \rho = 0.3,$
$a = 1.8, \quad b = 3, \quad c = 2.2, \quad d = 5.$

The static axon model :

$a_s = 0.155, \quad b_s = 0.075, \quad c_s = 0.045, \quad d_s = 0.145.$

The Matsuoka model :

Case 1 $\begin{cases} b_a = 10^4, \quad \tau_r = 9.09, \quad \tau_a = 1.1 \times 10^6, \\ \gamma = 0.134, \quad \delta = 1.3, \quad \beta = 803. \end{cases}$

Case 2 $\begin{cases} b_a = 5, \quad \tau_r = 9.22, \quad \tau_a = 651, \\ \gamma = 0.136, \quad \delta = 1.5, \quad \beta = 0.556. \end{cases}$

Appendix B: Proofs

Proof of Lemma 3 Noting that q_i must be zero at equilibrium because $F(0) = 0$, the unique equilibrium point is characterized by

$v_i = \varphi(0), \quad q_i = 0, \quad (i = 1, 2)$

regardless of the value of r. The linearized system around the equilibrium is described by

$q_i = F(s)u_i, \quad v_i = \varphi'(0)q_i, \quad (i = 1, 2),$

from which we obtain

$(1 - \sigma\varphi'(0)F(s))(1 + \sigma\varphi'(0)F(s))v_i = 0.$

Thus, the characteristic polynomial of the system is given by the numerator of the transfer function multiplying v_i. Then we see that the system is stable if and only if

$p(s) := \tau_1\tau_2 s^2 + (\tau_1 + \tau_2 + \lambda\varphi'(0))s + 1$

are Hurwitz polynomials for $\lambda := \pm\sigma k$, which is the case when

$$|\sigma k| < \frac{\tau_1 + \tau_2}{\varphi'(0)}.$$

Proof of Lemma 4 The relations among the signals in the RIO can be written as

$$\begin{bmatrix} 1 & VF\sigma \\ VF\sigma & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = V \begin{bmatrix} Fr_1 + \phi_1(v_1) - W\phi_2(v_1) \\ Fr_2 + \phi_1(v_2) - W\phi_2(v_2) \end{bmatrix}$$

where

$$V(s) := \frac{1}{s + b}, \quad W(s) := \frac{\rho}{s + \rho}.$$

The transfer function $G(s)$ in the framework of Fig. 7 is the mapping from $(\phi_1(v_1), \phi_1(v_2), \phi_2(v_1), \phi_2(v_2), r_1, r_2)$ to $(v_1, v_2, v_1, v_2, v_1, v_2)$. Since $V(s)$, $W(s)$, and $F(s)$ are all stable, we then see that $G(s)$ is stable if and only if $(1 - \sigma^2 V(s)^2 F(s)^2)^{-1}$ is stable. This is the case when the numerator polynomials of $1 \pm \sigma V(s)F(s)$,

$\tau_1\tau_2 s^3 + (b\tau_1\tau_2 + \tau_1 + \tau_2)s^2 + (1 + b(\tau_1 + \tau_2) \pm \sigma k)s + b$

is Hurwitz. By the Routh criterion, we obtain

$$\pm\sigma k > -\frac{(1 + b\tau_1)(1 + b\tau_2)}{1 + b\tau_o}$$

which is equivalent to the claimed condition.

Proof of Lemma 5 At an equilibrium, the input to the filter $F(s)$ is constant and hence its output is zero due to the zero at the origin. Thus the equilibrium is characterized by the equations for the dynamic axon with zero input and zero derivatives:

$w_i = \psi_1(v_i) = \psi_2(v_i), \quad q_i = 0, \quad (i = 1, 2).$

Note that the equilibrium point is independent of the input r. When linearized around an equilibrium, the dynamic axon model becomes

$\dot{v} = \psi_1' v - w + q, \quad \dot{w} = \rho(\psi_2' v - w)$

whose transfer function from q to v is given by

$$H(s) := \frac{s + \rho}{(s + \rho)(s - \psi_1') + \rho\psi_2'}.$$

In this case, the perturbed signals in the RIO satisfy

$v_i = H(s)F(s)u_i, \quad u_1 = -\sigma v_2, \quad u_2 = -\sigma v_1$

where $i = 1, 2$, from which we obtain

$(1 + \sigma HF)(1 - \sigma HF)v_i = 0.$

Thus the closed-loop poles are characterized by the roots of $1 \pm \sigma HF = 0$.

The numerator polynomial of $1 \pm \sigma HF = 0$ is given by

$n(s) := a_4 s^4 + a_3 s^2 + (a_2 + \lambda)s^2 + (a_1 + \rho\lambda)s + a_0$

where $\lambda := \pm\sigma k$ and

$a_2 := 1 + \tau_+\alpha + \tau_\times\beta, \quad a_4 := \tau_\times.$

From the Routh criterion, it can readily be shown that this polynomial has a root with positive real part if and only if at least one of the following conditions holds:

$a_3 < 0, \quad a_0 < 0, \quad c_1 + c_2\lambda < 0, \quad f(\lambda) < 0.$

If (7) holds, then $a_3 < 0$ or $a_0 < 0$ holds and hence the system is unstable. If (8) holds, then we have $a_3 > 0$ and $a_0 > 0$ and it suffices to show that either $c_1 + c_2\lambda < 0$ or $f(\lambda) < 0$ holds if and only if $|\sigma k| > \eta$. If $c_2 = 0$, the condition $c_1 + c_2\lambda < 0$ is violated due to $c_1 > 0$, and the condition $f(\lambda) < 0$ becomes $\lambda < -f(0)/(\rho c_1)$ which is equivalent to $|\sigma k| > \eta$ since

$$f(0) = \alpha\tau_+[(1+\tau_1^2\beta)(1+\tau_2^2\beta) + \alpha(\tau_+ + \tau_\times\alpha + \tau_+\tau_\times\beta)] > 0.$$

Consider the case $c_2 \neq 0$. Noting that

$$f(\lambda_o) = -a_0 a_3^2 < 0, \quad \lambda_o := -c_1/c_2$$

and $f(0) > 0$, we see that $f(\lambda) = 0$ has two real roots. Let λ and $\lambda+$ be the two roots such that $\lambda_< \lambda_+$. Now, if $c_2 > 0$, the instability condition is given by

$$\lambda < \lambda_o \quad \text{or} \quad \lambda_< \lambda < \lambda_+.$$

Since

$$\lambda_< \lambda_o < \lambda_+ < 0,$$

the condition is equivalent to $\lambda < \lambda_+$. This holds for $\lambda = \sigma k$ or $\lambda = -\sigma k$ if and only if $|\sigma k| > \eta$. If $c_2 < 0$, the instability condition is

$$\lambda > \lambda_o \quad \text{or} \quad \lambda < \lambda \quad \text{or} \quad \lambda+ < \lambda$$

with

$$\lambda_< 0 < \lambda_+ < \lambda_o,$$

and hence we have $\eta < |\lambda| = |\sigma k|$.

Proof of Theorem 2 The result follows from Lemmas 1, 4, and 5.

Lemma 6 *Let a continuous function φ and a T-periodic function q be given. Suppose $q(t)$ can be expressed as*

$$q(t) = \sum_{k=0}^{\infty} [a_k \sin(k\omega t) + b_k \cos(k\omega t)],$$

where $a_1 \neq 0$ and $\omega := 2\pi/T$. Consider the Fourier series expansion

$$\varphi(q(t)) = \sum_{k=0}^{\infty} [c_k \sin(k\omega t) + d_k \cos(k\omega t)]$$

and define

$$\kappa := \frac{c_1 + jd_1}{a_1 + jb_1}.$$

Then the following hold.

(i) If $a_i = b_j = 0$ for even i and odd j, then κ is real.
(ii) If $a_i = b_j = 0$ for $i \neq 1$ and $j \neq 0$, and φ is monotonically increasing, then κ is positive.

Proof Define

$$f(t) := \varphi(q(t)), \quad F(\theta) := f(\theta/\omega), \quad G(\theta) := F(\theta) + F(-\theta).$$

Then

$$\begin{aligned} d_1 &= \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos(\omega t)dt \\ &= \frac{1}{\pi}\int_0^{\pi} [F(\theta) + F(-\theta)]\cos(\theta)d\theta \\ &= \frac{1}{\pi}\int_0^{\pi/2} [G(\theta) - G(\pi-\theta)]\cos(\theta)d\theta \end{aligned}$$

Noting that

$$\sin k(\pi - \theta) = (-1)^{k+1}\sin k\theta,$$

$$\cos k(\pi - \theta) = (-1)^k \cos k\theta,$$

it can be verified that $G(\theta) = G(\pi - \theta)$ holds when $q(t)$ can be expressed as a linear combination of $\sin(k\omega t)$ with odd k and $\cos(k\omega t)$ with even k. Hence statement (i) holds.

To prove statement (ii), note that

$$\begin{aligned} c_1 &= \frac{2}{T}\int_{-T/2}^{T/2} f(t)\sin(\omega t)dt \\ &= \frac{1}{\pi}\int_0^{\pi} [F(\theta) - F(-\theta)]\sin(\theta)d\theta \end{aligned}$$

and

$$\begin{aligned} &a_1(F(\theta) - F(-\theta)) \\ &= a_1\left(\varphi(b_0 + a_1\sin\theta) - \varphi(b_0 - a_1\sin\theta)\right) \end{aligned}$$

is positive when $0 < \theta < \pi$ and φ is monotonically increasing. Hence we have $\kappa = c_1/a_1 > 0$.

References

Av-Ron E, Parnas H, Segel L (1993) A basic biophysical model for bursting neurons. Biol Cybern 69(1):87–95

Brown T (1911) The intrinsic factors in the act of progression in the mammal. Proc Roy Soc Lond B Biol Sci 84:308–319

Cohen A, Ermentrout G, Kiemel T, Kopell N, Sigvardt K, Williams T (1992) Modelling of intersegmental coordination in the lamprey central pattern generator for locomotion. Trends Neurosci 15(11):434–438

Doyle J, Packard A, Zhou K (1991) Review of LFTs, LMIs, and μ. In: Proceedings of IEEE conference on decision and control pp 1227–1232

Efimov D, Fradkov A (2004) Excitation of oscillations in nonlinear systems under static feedback. In: Proceedings of IEEE conference on decision and control pp 2521–2526

Ekeberg O (1993) A combined neuronal and mechanical model of fish swimming. Biol Cybern 69(5/6):363–374

Ermentrout G, Chow C (2002) Modeling neural oscillations. Physiol Behav 77:629–633

Ermentrout G, Kopell N (1984) Frequency plateaus in a chain of weakly coupled oscillators. I SIAM J Math Anal 15(2):215–237
Fradkov A (1979) Speed-gradient scheme and its application in adaptive control problems. Autom Remote Contr 40(9):1333–1342 (Translated from Avtomatika i Telemekhanika 9:90–101, 1979.)
Fradkov A (1999a) Exploiting nonlinearity by feedback. Physica D 128:159–168
Fradkov A (1999b) Feedback resonance in nonlinear oscillators. Proc Eur Contr Conf
Friesen W (1994) Reciprocal inhibition:A mechanism underlying oscillatory animal movements. Neurosci Biobehav Rev 18(4):547–553
Friesen W, Stent G (1978) Neural circuits for generating rhythmic movements. Annu Rev Biophys Bioeng 7:37–61
Fukuoka Y, Kimura H, Cohen A (2003) Adaptive dynamic walking of a quadruped robot on irregular terrain based on biological concepts. Int J Robot Res 22(3-4):187–202
Getting P (1989) Reconstruction of small neural networks. In: Koch C, Segev I (eds) Methods in Neuronal Modeling. MIT Press, Cambridge, pp 171–194
Glad T and Ljung L (2000) Control theory – multivariable and nonlinear methods. Taylor & Francis, Boca Raton, London
Hadeler K (1974) On the theory of lateral inhibition. Kybernetik 14:161–165
Hindmarsh J, Rose R (1982) A model of the nerve impulse using two first-order differential equations. Nature 296(5853):162–164
Hunt K, Sbarbaro D, Zbikowski R, Gawthrop P (1992) Neural networks for control systems – a survey. Automatica 28(6):1083–1112
Ijspeert A (2001) A connectionist central pattern generator for the aquatic and terrestrial gaits of a simulated salamander. Biol Cybern 84:331–348
Iwasaki T, Zheng M (2002) The Lur'e model for neuronal dynamics. In: Proceedings of IFAC world congress July 21–26, 2002, Barcelona, Spain
Khalil H (1996) Nonlinear systems. Prentice Hall, Englewood Cliffs
Lewis M, Etiennne-Cummings R, Hartmann M, Xu Z, Cohen A (2003) An in silico pattern generator:silicon oscillator, coupling, entrainment, and physical computation. Biol Cybern 88:137–151
Matsuoka K (1985) Sustained oscillations generated by mutually inhibiting neurons with adaptation. Biol Cybern 52:367–376
Matsuoka K (1987) Mechanisms of frequency and pattern control in the neural rhythm generators. Biol Cybern 56:345–353
Ogata K (1996) Modern control engineering. Prentice Hall, Englewood Cliffs
Ono K (1998) Self-excited natural periodical motion mechanism for efficient and robust locomotion. In: Proceedings of TITech COE/Super-Mechano Systems Workshop pp 32–43
Ono K, Takahashi R, Shimada T (2001) Self-excited walking of a biped mechanism. Int J Robot Res 20(12):953–966
Orlovsky G, Deliagina T, Grillner S (1999) Neuronal Control of Locomotion:From Mollusc to Man. Oxford University Press, Oxford
Raney D, Slominski E (2004) Mechanization and control concepts for biologically inspired micro air vehicles. J Aircr 41(6):1257–1265
Rinzel J (1985) Excitation dynamics:insights from simplified membrane models. Fed Proc 44(15):2944–2946
Taga G (1991) Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment. Biol Cybern 65:147–159
Wadden T, Ekeberg O (1998) A neuro-mechanical model of legged locomotion:single leg control. Biol Cybern 79:161–173
Williamson M (1998) Neural control of rhythmic arm movements. Neural Netw 11:1379–1394