PHYSICAL REVIEW E **68**, 020901(R) (2003)

Balancing at the border of instability

Luc Moreau
SYSTeMS-EESA, Ghent University, Technologiepark 914, 9052 Zwijnaarde, Belgium

Eduardo Sontag
Department of Mathematics, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, USA
(Received 4 December 2002; published 22 August 2003)

Some biological systems operate at the critical point between stability and instability, and this requires a fine tuning of parameters. We bring together two examples from the literature that illustrate this: neural integration in the nervous system and hair cell oscillations in the auditory system. In both examples the question arises as to how the required fine tuning may be achieved and maintained in a robust and reliable way. We study this question using tools from nonlinear and adaptive control theory. We illustrate our approach on a simple model which captures some of the essential features of neural integration. As a result, we propose a large class of feedback adaptation rules that may be responsible for the experimentally observed robustness of neural integration. We mention extensions of our approach to the case of hair cell oscillations in the ear.

DOI: 10.1103/PhysRevE.68.020901 PACS number(s): 87.19.La, 64.60.My, 87.19.St

Persistent neural activity is prevalent throughout the nervous system. Numerous experiments have demonstrated that persistent neural activity is correlated with short-term memory. A prominent example concerns the oculomotor system—see Refs. [1,2] for a review and experimental facts. The brain moves the eyes with quick saccadic movements. Between saccades, it keeps the eyes still by generating a continuous and constant contraction of the eye muscles; thus requiring a constant level of neural activity in the motor neurons controlling the eye muscles. This constant neural activity level serves as a short-term memory for the desired eye position. During a saccade, a brief burst of neural activity in premotor command neurons induces a persistent change in the neural activity of the motor neurons, via a mechanism equivalent to integration in the sense of calculus. Neural activity of an individual neuron, however, has a natural tendency to decay with a relaxation time of the order of milliseconds. Therefore the question arises as to how a transient stimulus can cause persistent changes in neural activity. According to a long-standing hypothesis, persistent neural activity is maintained by synaptic feedback loops. Positive feedback can oppose the tendency of a pattern of neural activity to decay. If the feedback is weak, then the natural tendency to decay dominates and neural activity decays. As the feedback strength is increased, the neural dynamics undergo a bifurcation and become unstable. When the feedback is tuned to exactly balance the decay, then neural activity neither increases nor decreases but persists without change. This, however, requires a fine tuning of the synaptic feedback strength, and the question arises as to how a biological system can achieve and maintain this fine tuning [3–6]. Some gradient descent and function approximation algorithms performing this fine tuning have been proposed [3,5] and a feedback learning mechanism based on differential anti-Hebbian synaptic plasticity has been studied in Ref. [7]. Nevertheless, it is still unclear how the required fine tuning is physiologically feasible. For this reason, a different model for neural integration based upon bistability has recently been proposed in Ref. [8]. In the present paper, we do not follow the line of research based upon bistability. Instead, we pursue the hypothesis of precisely tuned synaptic feedback. The present paper proposes an adaptation mechanism that may be responsible for the fine tuning of neural integrators and that may explain the experimentally observed robustness of neural integrators with respect to perturbations. Before we present this adaptation mechanism in detail, we first discuss a similar phenomenon in the auditory system.

In order to detect the sounds of the outside world, hair cells in the cochlea operate as nanosensors which transform acoustic stimuli into electric signals. In Refs. [9–12] these hair cells are described as active systems capable of generating spontaneous oscillations. Ions such as Ca^{2+} are believed to contribute to the hair cell's tendency to self-oscillate. For low concentrations of the ions, damping forces dominate and the hair cell oscillations are damped. As the concentration increases, the system undergoes a Hopf bifurcation, the dynamics become unstable, and the hair cells exhibit spontaneous oscillations. In Refs. [9–12] the hair cells are postulated to operate near the critical point, where the activity of the ions exactly compensates for the damping effects. As before, this requires a fine tuning of parameters (the ion concentrations), and again the question arises as to how this fine tuning can be achieved and maintained. In Refs. [9,11], a feedback mechanism has been proposed, which could be responsible for maintaining this fine tuning.

It thus seems that operating in the vicinity of a bifurcation is a recurrent theme in biology. And the question as to how proximity to the bifurcation point may be achieved and maintained in a noisy environment may be of considerable general interest. We view the two presented examples as special instances of the following general problem. Consider a forced dynamical system, described by a differential equation $\dot{x}=f_\mu(x,u(t))$. The right-hand side of this equation depends on a parameter μ, and the unforced dynamics $\dot{x}=f_\mu(x,0)$ are assumed to exhibit a bifurcation when μ equals a critical value μ_0. The problem consists of finding a feedback adaptation rule for the parameter μ, which guaran-

tees proximity to the bifurcation point; that is, which steers μ toward its critical value μ_0. This adaptation law may depend on μ and x but should be independent of μ_0, since this critical value is not known precisely. This abstract formulation captures common features of both biological examples and suggests some unexpected links with the literature. Questions very similar to the present one have been studied extensively in the literature on adaptive control [13] and stabilization [14]; and the general problem is closely related to extremum seeking [16] and to instability detection [15], where an operating parameter is adapted on line in order to experimentally locate bifurcations.

Although the above general formulation is convenient, there is little hope that a complete and satisfactory theory can be developed, which applies to all possible instances of the problem. Introducing simplifying assumptions makes it more tractable. In this paper, we study in detail what is probably the most simple but nontrivial instance of the general problem. We consider the one-dimensional system

$$\dot{x} = -\mu_0 x + \mu x + u(t) \quad (\dot{x} = dx/dt), \tag{1}$$

which captures some of the essential features of neural integration and is, in fact, closely related to the autapse model from Ref. [6]. With this interpretation, x is a strictly positive variable representing neural activity in the integrator network and $u(t)$ represents the signal generated by the premotor command neurons. The term $-\mu_0 x$ corresponds to the natural decay of neural activity and μx represents a positive, synaptic feedback loop. Of course, when studying neural integration, questions can be investigated at varying levels of detail. It is clear that a simple model such as that described by Eq. (1) has several limitations. Because of its one-dimensional nature, the present model is, for example, unable of reproducing the distributed nature of persistent activity patterns observed in the brain. Nevertheless, Eq. (1) captures a key feature of neural integration: when the feedback is tuned to exactly balance the decay, Eq. (1) behaves as an integrator and produces persistent neural activity. Equation (1) is therefore a valuable model when studying fine tuning of neural integrator networks [1,17].

We are interested in the fine tuning of Eq. (1) and study this question using tools from nonlinear and adaptive control theory. First, we ignore the presence of the input $u(t)$ and consider the simpler equation

$$\dot{x} = -\mu_0 x + \mu x. \tag{2}$$

We present a large class of feedback adaptation laws for Eq. (2), which steer μ to its critical value μ_0; thus enabling the automatic self-tuning of parameters and the spontaneous generation of persistent neural activity. We consider adaptation laws [21] of the form

$$\dot{\mu} = f(x) - g(\mu). \tag{3}$$

We show that, under three very mild conditions, this adaptation rule guarantees convergence to the bifurcation point for Eq. (2). The first condition requires that g is a strictly increasing function. This means that the term $-g(\mu)$ in Eq. (3) acts as a negative feedback. As a consequence, if the neural activity x were constant in Eq. (3), then the synaptic feedback gain μ would naturally relax to a rest value depending on x via the equation $f(x) = g(\mu)$. The second condition states that there exists x^* such that $f(x^*) = g(\mu_0)$. This condition implies that, if the neural activity would be constant and equal to x^* in Eq. (3), then the synaptic feedback gain μ would naturally relax to its critical, desired value μ_0. Of course there is no guarantee that the neural activity would be equal to, or even converge to, this special value x^*. Instead, the level of neural activity is governed by Eq. (2). Therefore, in order for the adaptation law (3) to work, we need to impose a last condition that f is a decreasing function. This means that the level of neural activity negatively regulates the synaptic feedback strength.

We now show that, under these three conditions, the feedback adaptation law (3) indeed tunes the synaptic feedback gain μ to exactly balance the natural decay rate μ_0. We begin with noticing that the combined system of equations (2) and (3) has a unique rest point. This equilibrium is determined by setting the right-hand sides of Eqs. (2) and (3) equal to zero, yielding $x = x^*$ and $\mu = \mu_0$. Although the precise value of μ_0 is unknown, if we are able to prove that all trajectories of Eqs. (2) and (3) converge to this (unknown) fixed point, then it follows that μ indeed converges to its desired critical value μ_0. In order to prove this, we introduce a coordinate transformation $q = \ln(x) - \ln(x^*)$ and $p = \mu - \mu_0$. This transforms Eqs. (2) and (3) into $\dot{q} = p$ and $\dot{p} = f(\exp(q)x^*) - g(p + \mu_0)$. In these new coordinates, the dynamics take the form of a nonlinear mass-spring-damper system [with unit mass, nonlinear spring characteristic $f(\exp(\cdot)x^*)$, and nonlinear damping function $g(\cdot + \mu_0)$]. It follows from physical energy considerations that this system exhibits damped oscillations [20]. This shows that all trajectories of Eqs. (2) and (3) indeed converge to the unique fixed point, where $\mu = \mu_0$.

The above coordinate transformation reveals a subtle relationship between self-tuning of bifurcations and the internal model principle (IMP) from robust control theory (see Refs. [18,19] for a discussion of the IMP from a systems biology perspective). This relation is made explicit by the equation $\dot{q} = p$, which represents an integrator and corresponds to integral action studied in robust control theory. One regards the constant μ_0 as an unknown perturbation acting on the system. The IMP implies that, in order to track this constant perturbation, the system dynamics should contain integral action. The integral action is generated by the biological system itself, and not by the feedback adaptation law.

We have so far ignored the presence of the signal $u(t)$. We showed that the adaptation law (3) tunes the synaptic feedback gain to exactly compensate for the natural decay rate, resulting in the spontaneous generation of persistent neural activity. At these equilibrium conditions, the action potential firing rate equals x^*, which is related to μ_0 by $f(x^*) = g(\mu_0)$. In the next paragraphs, we take into account the effect of the input $u(t)$. In this case, the value x^* will

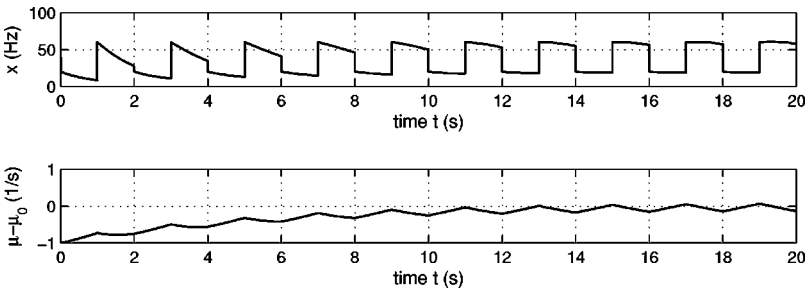

FIG. 1. Tuning of a neural integrator. Simulation of Eq. (1) with $\dot{\mu}=\varepsilon(-ax-b\mu+c)$ and $u(t)=\Sigma_{t_i}\delta(t-t_i)\{x_{\text{desired}}(t_i+)-x(t_i-)\}$, where $\delta(\cdot)$ represents the Dirac impulse and where the sum goes over all saccade times t_i. The constants are $\mu_0=200\ \text{s}^{-1}$, $a=1\ \text{s}^{-1}$, $b=0.01\ \text{s}^{-1}$, $c=42\ \text{s}^{-2}$, and $\varepsilon=0.01$. The adaptation law satisfies the compatibility condition $a(20+60\ \text{Hz})/2+b\mu_0=c$.

play the role of a parameter that influences the accuracy with which the feedback adaptation law guarantees proximity to the bifurcation point.

The signal $u(t)$ will in general result in a time-varying action potential firing rate $x(t)$. The mechanism with which this happens is determined by the neural integrator equation (1) and the adaptation law (3). For the purpose of analysis, we make two simplifying assumptions, both of which seem to be natural and physically relevant for neural integration. First, we assume that, over any sufficiently large time interval $[t_0, t_0+T]$, the time spent by $x(t)$ in any interval $[x_1, x_2]$ is approximately independent of t_0. In more mathematical terms, we assume the existence of a function $P(x)$ such that for every test function $\alpha(x)$, the time average $(1/T)\int_{t_0}^{t_0+T}\alpha[x(t)]dt$ converges to $\int_0^\infty P(x)\alpha(x)dx$ as $T\to\infty$, uniformly with respect to t_0. Second, we assume that the adaptation law acts on a much slower time scale than the time variations in $x(t)$. Under these assumptions, the effect of the action potential firing rate $x(t)$ on the adaptation law (3) may be approximated by the average effect $\dot{\mu}=\int_0^\infty P(x)f(x)dx-g(\mu)$. It is now clear when the adaptation law guarantees proximity to the bifurcation point: if the compatibility condition $\int_0^\infty P(x)f(x)dx=f(x^*)$ is satisfied, then time scale separation arguments suggest that μ will converge approximately to μ_0 and the neural integrator will approximately behave as a perfect integrator. The compatibility condition may by interpreted as follows [20]. When the premotor command signal $u(t)$ has zero time average and the adaptation law acts on a slow time scale, then Eq. (1) behaves as a good integrator and the firing rate $x(t)$ equals the time integral of $u(t)$ plus an integration constant. The compatibility condition ensures that this integration constant is compatible with the desired range for the firing rate $x(t)$.

We illustrate this result on a particular example representative for saccadic eye movements. We consider the case of periodic saccadic eye movements asking for an action potential firing rate in the motor neurons alternating between 20 Hz and 60 Hz every second. At each saccade, a brief burst of neural activity in premotor command neurons changes the actual firing rate. We assume that this change is such that immediately after each saccade, the actual firing rate equals the desired firing rate. Between saccades, we assume that no input is applied [22]. If the neural integrator is perfectly tuned, then the actual firing rate will remain constant between saccades and equal to the desired firing rate (eyes are fixed). If the neural integrator is not perfectly tuned, then the actual firing rate will deviate from the desired firing rate (eyes drift) until a new saccade occurs, which brings the actual firing rate to its new desired value. Figure 1 shows the results of a simulation where the adaptation law satisfies the compatibility condition of the previous paragraph. In the beginning of the simulation, we have mistuned the neural integrator. Clearly, after a short transient, the adaptation law achieves excellent tuning and the drift between two successive saccades becomes negligible.

We have thus shown that an adaptation law can tune a neural integrator with great accuracy to its bifurcation point. In order to achieve perfect tuning, however, the adaptation law itself needs to satisfy a compatibility condition. It seems that we have merely moved the problem of fine tuning from the neural integrator to the adaptation law. The crucial observation and one of the main contributions of the present paper, however, is that this results in a significant decrease in sensitivity. *The adaptation law is robust with respect to perturbations in its parameters.*

In order to illustrate this significant increase in robustness, let us first summarize the well-known [5] sensitivity proper-

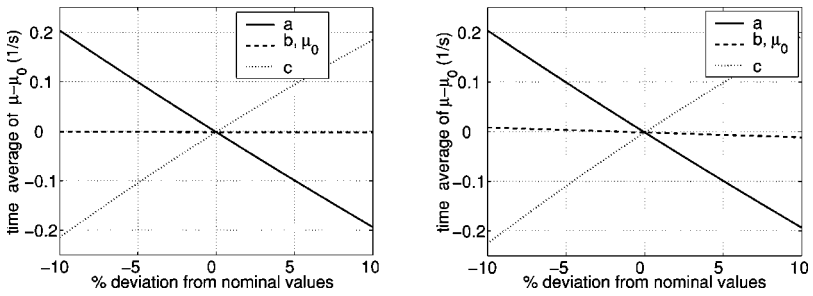

FIG. 2. Robustness of tuning with respect to parameter perturbations in the adaptation law. Simulation results for Eq. (1) with $\dot{\mu}=0.001(-ax-b\mu+c)$ and $u(t)$ as in Fig. 1. The plots show the average value of $\mu-\mu_0$ in periodic regime for different values of the parameters a, b, c, and μ_0. The nominal values of the parameters satisfy the compatibility condition $a(20+60\ \text{Hz})/2+b\mu_0=c$, and are given by $a=1\ \text{s}^{-1}$, $b=0.01\ \text{s}^{-1}$, $c=40.1\ \text{s}^{-2}$, $1/\mu_0=100$ ms (left) and $a=1\ \text{s}^{-1}$, $b=0.01\ \text{s}^{-1}$, $c=42\ \text{s}^{-2}$, $1/\mu_0=5$ ms (right).

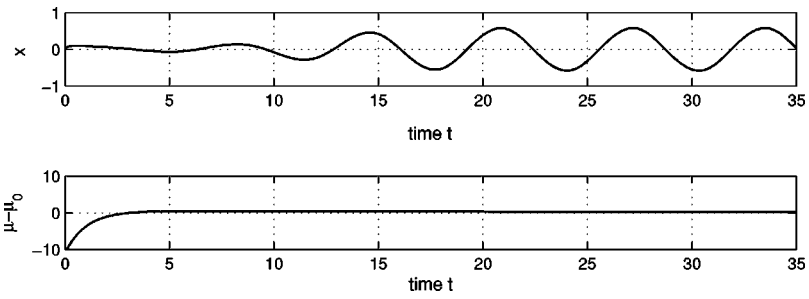

FIG. 3. Tuning of a nonlinear oscillator. Simulation of equations $\ddot{x}+(\mu_0-\mu)\dot{x}+\lambda\dot{x}^3+\omega^2 x=0$ and $\dot{\mu}=f(r)-g(\mu)$ with $\mu_0=1$, $\lambda=1$, $\omega=1$ and $f(r)-g(\mu)=1/(1+r^2)-\mu+1/2$. The variable r is determined by $r^2=x^2+(\dot{x}/\omega)^2$.

ties of neural integration. Experiments suggest that the actual time constant obtained in a tuned neural integrator circuit is typically greater than 10 s; that is, $|\mu-\mu_0|\leqslant 0.1\ \mathrm{s}^{-1}$. This requires for the fine tuning of μ a relative precision $\Delta\mu/\mu$ ranging from 1/100 to 1/2000, depending on whether the intrinsic time constant $1/\mu_0$ equals 100 ms or 5 ms (typical values suggested in the literature). The required precision for μ should be contrasted with the required precision for the parameters of the adaptation law proposed in the present paper. The simulations of Fig. 2 show that, in order to have $|\mu-\mu_0|\leqslant 0.1\ \mathrm{s}^{-1}$ as observed in experiments, the parameters of the adaptation law need to be tuned with a precision of 1/20, independently of the intrinsic time constant $1/\mu_0$. Comparing this with the originally required precision for the synaptic feedback strength μ, we conclude that *the proposed adaptation mechanism could improve the robustness of neural integration with a factor ranging from 5 to 100*.

We have studied a simple model for neural integration and proposed a class of feedback adaptation rules that could explain the experimentally observed robustness of neural integration with respect to perturbations. The analysis tools that we have introduced extend to the study of fine tuning involved in other systems such as hair cell oscillations in the ear [20]. Consider the nonlinear oscillator equation $\ddot{x}+(\mu_0-\mu)\dot{x}+\lambda\dot{x}^3+\omega^2 x=u(t)$ which captures some of the essential features of hair cell oscillations [11]. Inspired by our previous analysis, we consider a feedback adaptation law for the parameter μ of the form $\dot{\mu}=f(r)-g(\mu)$, with r a positive variable characterizing the magnitude of oscillations and related to x and $\dot{x}$ via the expression $r^2=x^2+(\dot{x}/\omega)^2$. Figure 3 shows that, in the absence of the stimulus $u(t)$, this type of adaptation law is indeed able to bring and keep the bifurcation parameter close to its critical value, resulting in the spontaneous generation of oscillations.

L.M. was financially supported by the Fund for Scientific Research—Flanders. He was also supported by the Belgian American Educational Foundation, while visiting the Princeton University Mechanical and Aerospace Engineering Department. This paper presents research results of the Belgian Program on Inter-University Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors. E.S. was supported in part by the USAF Grant No. F49620-01-1-0063 and by the NSF Grant No. CCR-0206789.

[1] D. Robinson, Annu. Rev. Neurosci. **12**, 33 (1989).
[2] E. Aksay, G. Gamkrelidze, H.S. Seung, R. Baker, and D.W. Tank, Nat. Neurosci. **4**, 184 (2001).
[3] D.B. Arnold and D.A. Robinson, Biol. Cybern. **64**, 447 (1991).
[4] H.S. Seung, Proc. Natl. Acad. Sci. U.S.A. **93**, 13 339 (1996).
[5] H.S. Seung, D.D. Lee, B.Y. Reis, and D.W. Tank, Neuron **26**, 259 (2000).
[6] H.S. Seung, D.D. Lee, B.Y. Reis, and D.W. Tank, J. Comput. Neurosci. **9**, 171 (2000).
[7] X. Xie and H.S. Seung, Adv. Neural Inf. Process. Syst. **12**, 199 (2000).
[8] A. Koulakov, S. Raghavachari, A. Kepecs, and J. Lisman, Nat. Neurosci. **5**, 709 (2002).
[9] S. Camelet, T. Duke, F. Jülicher, and J. Prost, Proc. Natl. Acad. Sci. U.S.A. **97**, 3183 (2000).
[10] V. Eguíluz, M. Ospeck, Y. Choe, A. Hudspeth, and M. Magnasco, Phys. Rev. Lett. **84**, 5232 (2000).
[11] F. Jülicher, Acad. Sci., Paris, C. R. **2**(6), 849 (2001).
[12] F. Jülicher, D. Andor, and T. Duke, Proc. Natl. Acad. Sci. U.S.A. **98**, 9080 (2001).
[13] M. Kristic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design* (Wiley, New York, 1995).
[14] E. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, 2nd ed. (Springer, New York, 1998).
[15] J.S. Anderson, S.Y. Shvartsman, G. Flätgen, I.G. Kevrekidis, R. Rico-Martínez, and K. Krischer, Phys. Rev. Lett. **82**, 532 (1999).
[16] M. Krstić and H.-H. Wang, Automatica **36**, 595 (2000).
[17] L. Shen, Biol. Cybern. **61**, 319 (1989).
[18] T.-M. Yi, Y. Huang, M.I. Simon, and J. Doyle, Proc. Natl. Acad. Sci. U.S.A. **97**, 4649 (2000).
[19] E. Sontag, Systems Control Lett. (to be published).
[20] L. Moreau, E. Sontag, and M. Arcak, Systems Control Lett. (to be published).
[21] The dynamics for μ could come from synaptic plasticity. In particular, the term $f(x)$ might be related to types of synaptic plasticity, which depend on the temporal ordering of presynaptic and postsynaptic spiking, as in Ref. [7].
[22] We assume that no feedback is applied to keep x at its desired level between saccades, which is consistent with experimental observations.