

## Relating Neural Dynamics to Neural Coding

G. Bard Ermentrout,<sup>1,3</sup> Roberto F. Galán,<sup>2,3</sup> and Nathaniel N. Urban<sup>2,3</sup>

<sup>1</sup>University of Pittsburgh, Department of Mathematics, Thackeray Hall, Pittsburgh, Pennsylvania 15260, USA

<sup>2</sup>Carnegie Mellon University, Department of Biological Sciences, Mellon Institute, Pittsburgh, Pennsylvania 15213, USA

<sup>3</sup>Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania 15213, USA

(Received 2 July 2007; published 14 December 2007)

We demonstrate that two key theoretical objects used widely in computational neuroscience, the phase-resetting curve (PRC) from dynamics and the spike triggered average (STA) from statistical analysis, are closely related when neurons fire in a nearly regular manner and the stimulus is sufficiently small. We prove that the STA due to injected noisy current is proportional to the derivative of the PRC. We compare these analytic results with numerical calculations for the Hodgkin-Huxley neuron and we apply the method to neurons in the olfactory bulb of mice. This observation allows us to relate the stimulus-response properties of a neuron to its dynamics, bridging the gap between dynamical and information theoretic approaches to understanding brain computations and facilitating the interpretation of changes in channels and other cellular properties as influencing the representation of stimuli.

DOI: [10.1103/PhysRevLett.99.248103](https://doi.org/10.1103/PhysRevLett.99.248103)

PACS numbers: 87.19.La, 05.45.Xt, 87.18.Sn

Dynamical systems models like the Hodgkin-Huxley model of the squid giant axon are very effective at replicating the firing properties of individual neurons. Such models have been extremely useful tools for understanding the mechanisms of neural excitability and for simulating neural circuits. However, these models have been less successful in aiding the understanding of what neurons compute. Given a model that perfectly describes the behavior of a neuron or a network, we are in most cases still at a loss to say what computation this neuron or circuit is performing. Rather, the natural concepts and objects of the theory of computation (e.g., stimuli, features, coding) seem only distantly related to those of the theory of dynamical systems (e.g., differential equations and attractors). Here we relate two mathematical objects central to these two approaches: the phase-resetting curve (PRC) and the spike triggered average (STA). We start by introducing these objects and then prove that the STA is proportional to the derivative of the PRC in the weak stimulus limit. We show that this approach allows us to efficiently and accurately compute the STA from the PRC and vice versa, in the case of numerical simulations as well as in the case of real neurons. The ability to compute these functions from each other allows us to make some progress in relating dynamics of neurons to their ability to code features.

The PRC describes how the spiking of a regularly firing neuron is altered by incoming input, that is, how the time of the next spike is shifted as a function of the stimulus time relative to the previous spike:  $\Delta(t) \equiv [T - \hat{T}(t)]/K$  where  $T$  is the natural period and  $\hat{T}(t)$  is the time of the spike given a stimulus at time  $t$  after the last spike. The constant  $K$  is the magnitude of the stimulus, e.g., 1 mV since the PRC depends on the amplitude of the perturbation. [1,2] show that for small stimuli,  $x(t)$ , any stable limit cycle oscillator can be reduced to a scalar model for its phase,  $\theta$ :

$$\frac{d\theta}{dt} = 1 + \Delta(\theta)x(t), \quad (1)$$

where  $\theta \in [0, T)$  and  $\Delta(\theta)$  is the PRC. For neurons, the stimulus has dimensions of millivolts per millisecond. We take  $\theta = 0$  to be the time of spiking. The PRC is valid for neurons which are repetitively firing (that is, on a stable limit cycle), but the concept of the PRC can be applied even when the neurons are quite noisy, often permitting a PRC to be measured experimentally from neurons *in vitro* [3,4].  $\Delta(\theta)$  is readily computed for differential equation models, and can also be computed for neural models either via direct perturbation [5–7] or indirectly [3,4]. Equation (1) shows that for a regularly firing neuron, the PRC provides, an answer to the question of “how will a stimulus influence when the next spike will come?”

A contrasting approach to understanding neural computation has focused on neural coding, by which we mean determining what features of stimuli are represented by single spikes and spike trains [8]. Such analysis of the stimulus dependence of spike trains often includes the calculation of the STA, which is related to the reverse correlation [8,9]. The STA is defined as the average stimulus with a given prior statistics preceding an action potential in a neuron. For our purposes, by stimulus we refer to the current injected into a neuron (divided by the capacitance). In other experimental protocols, the stimulus may be a sensory stimulus presented to an animal while a neuron is recorded. If  $x(t)$  is the stimulus

$$\text{STA}(t) = \langle x(\tau_j - t) \rangle, \quad (2)$$

where the average is taken over all spike times,  $\tau_j$ . In the present context, the STA is an answer to the question of “on average, what temporal features of the stimulus lead to spiking?”

*Theory.*—To derive the main result which applies when the stimulus is *small*, we write  $x(t) = \sigma \xi(t)$ , where  $\sigma$  is the magnitude of the signal and  $\xi(t)$  is a zero-mean stationary stochastic process with unit variance. (Note that for white noise applied to the voltage equation for a neuron model,  $\sigma^2$  has dimensions of  $\text{mV}^2/\text{msec}$ .) We will use the smallness of  $\sigma$  to estimate the time of a spike in Eq. (1) and from this obtain the STA from (2). With  $\theta(0) = 0$  as the initial condition (we assume the neuron has just spiked) we write

$$\theta(t) = \theta_0(t) + \sigma \theta_1(t) + \dots$$

(as, for example, [10]) and substitute into (1). Clearly  $\theta_0(t) = t$ , and from this we find that

$$\frac{d\theta_1}{dt} = \Delta(t)\xi(t),$$

which upon integration yields

$$\theta_1(t) = \int_0^t \Delta(s)\xi(s)ds.$$

We want to determine the time  $\tau$  at which the oscillator spikes again, that is  $\theta(\tau) = T$ . As with  $\theta$ ,  $\tau$  depends on  $\sigma$ , so we write  $\tau = \tau_0 + \sigma \tau_1 + \dots$ , and find  $\tau_0 = T$  and

$$\tau_1 = - \int_0^T \Delta(s)\xi(s)ds.$$

Thus, we find that to order  $\sigma$ ,

$$\tau = T - \sigma \int_0^T \Delta(s)\xi(s)ds. \quad (3)$$

Substituting Eq. (3) into (2) and taking expectations, we get

$$\begin{aligned} \text{STA}(t) &= \sigma \left\langle \xi \left( T - \sigma \int_0^T \Delta(s)\xi(s)ds - t \right) \right\rangle \\ &= \sigma \langle \xi(T-t) \rangle - \sigma^2 \langle \xi'(T-t) \int_0^T \Delta(s)\xi(s)ds \rangle + \dots \\ &= -\sigma^2 \int_0^T \Delta(s) \langle \xi'(T-t)\xi(s) \rangle ds + \dots \\ &= -\sigma^2 \int_0^T \Delta(s) C'(T-t-s) ds + \dots \\ &= -\sigma^2 \int_0^T \Delta'(s) C(T-t-s) ds + \dots, \end{aligned} \quad (4)$$

where  $C(t) = \langle \xi(t)\xi(0) \rangle$  is the auto correlation of the stimulus. To get the fourth line, we note that the expected value of  $\xi(t)$  is zero; to get the last line, we have integrated by parts. Dropping the higher-order terms and assuming the stimulus is white noise, we obtain the main result

$$\text{STA}(t) = -\sigma^2 \Delta'(-t), \quad (5)$$

where we use the  $T$  periodicity of  $\Delta(t)$  to drop the  $T$ . This result shows that the dynamics of a neuron, as captured by the PRC, can be used to predict the STA and conversely, given the STA, we can estimate the PRC of a repetitively firing neuron. Furthermore, (4) is valid for any small, zero-mean, stationary, stochastic stimulus.

Two key issues arise in this analysis. First, we must translate the periodic PRC to the generally aperiodic STA. This conversion is most straightforward when neuron is firing at a nearly constant rate and the PRC is well defined. While the STA is in principle aperiodic, in reality it is only sensible to define the STA over the time interval prior to a spike in which there are no other spikes. Thus the time over which both the STA and the PRC can be clearly defined is the interval between spikes, i.e., the average period. Second, we must note that because the PRC is the integral of the STA, it is defined only up to an additive constant term. However, if we assume that the PRC vanishes at  $t = 0, T$  (as is common in neurons), then we can determine the integration constant.

*Examples.*—To test the theory, we compute the STA for the Hodgkin-Huxley (HH) equations and then use Eq. (5) to compute the PRC subject to the constraints that it vanishes at  $t = 0$  and at the average period,  $t = T$ . We drive the four variable biophysical Hodgkin-Huxley model with a constant bias current ( $I = 10 \mu\text{A}/\text{cm}^2$ ) to make it fire at 70 Hz, inject noise and compute the spike triggered average. (7000 spikes were used.) We then numerically integrate the STA and time-reverse it to reconstruct an approximate PRC. We compute the exact PRC using the method in [11] and compare the two methods for two different amplitudes of noise. Figure 1 shows that in both the low and high noise case, the PRCs calculated from the STAs are almost identical to the actual PRC. Later (see Fig. 3), we will systematically quantify the dependence of the reconstruction on the statistics of the spike trains for both the HH and phase models.

We next tested this transformation on real neurons. We performed whole cell recordings from olfactory bulb mitral cells. We injected these cells with dc current to cause them

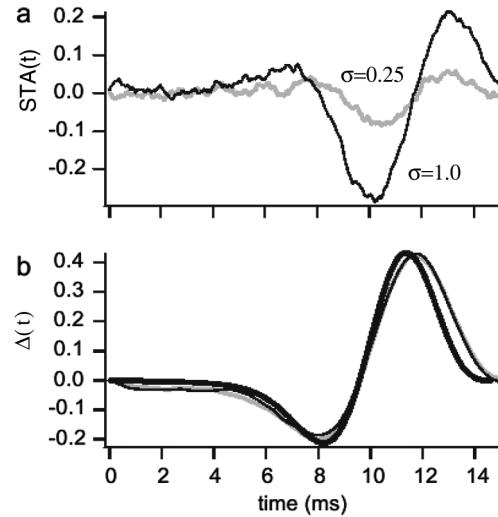


FIG. 1. Estimation of the PRC from the STA for the Hodgkin-Huxley model. (a) STA for two levels of noise (gray:  $\sigma^2 = 0.0625 \text{ mV}^2/\text{msec}$ ; black:  $\sigma^2 = 1.0 \text{ mV}^2/\text{msec}$ ) (b) PRCs reconstructed from the STA (thick line, true PRC; thin lines, reconstructed PRCs).

to fire repetitively at  $50 \pm 6$  Hz, added noisy current (with an amplitude 10% of the dc) and recorded spike times over intervals of 2–2.5 seconds, repeated 100–120 times. From the recorded spike trains we first eliminated the initial period (250–600 ms) of spike frequency accommodation and then calculated the STA. We calculated the PRC from the STA as per Eq. (5). For comparison, we also calculated the PRC using a method based on injection of aperiodic perturbing pulses [4]. Figure 2 shows the estimated PRC obtained from these two methods from the an olfactory bulb mitral cell. In the PRCs estimated by both methods, there is a clear negative region after the spike followed by a larger positive region, consistent with our earlier estimation of the PRC for olfactory bulb mitral cells. We also were interested in the fact that the STA-based method was possible despite the fact that cells were not firing in a precisely oscillatory manner. We measured the standard deviation of the interspike intervals in our recordings and found it to be approximately 10% of the firing rate. Thus our method is robust for at least this level of variation in the firing rate. We also applied these methods to recordings of neurons in the mouse somatosensory cortex (data not shown) with similar results.

To explore further how this relationship between the PRC and the STA depends on the regularity of the periodic firing, we drive both the HH model and simple phase models of the form of (1) with larger and larger amplitude noise to make their firing more irregular. We examined the effects of noise on phase models with two commonly used PRCs [ $\Delta(t) = \sin t$ , Figure 3(a1);  $\Delta(t) =$

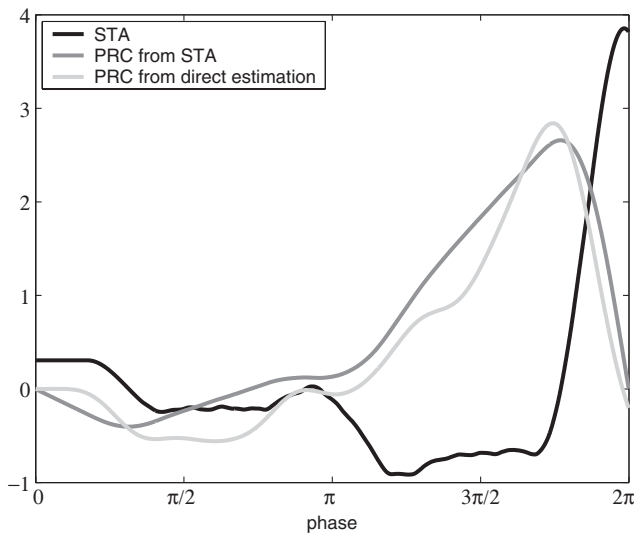


FIG. 2. Estimation of the PRC of a neuron recorded *in vitro*. Graph shows the STA (black) and the PRC estimated using two different methods from recordings of an olfactory bulb mitral cell. The light gray line shows an estimate from the method described previously [4]. The dark gray line shows average PRC for the same cell calculated from the STA, as described above. (Note the STA has been plotted in time rescaled by  $2\pi/T$ , the mean period.)

$1 - \cos t$ , Fig. 3(a2)]. Calculating the STA and the PRC for these higher noise simulations (after correcting for the change in average firing rate) resulted in PRCs that less and less closely resembled the actual PRC for these models although the general shape of the PRC was somewhat maintained [Figs. 3(a1) and 3(a2)]. We injected increasingly larger noise ( $\sigma = 0.25 - 16$ ) into the HH model. While the shape of the PRC is degraded, the zero crossing is preserved remarkably well. We quantified this degradation in the quality of the estimated PRC by calculating the correlation coefficient ( $R$ ) between the actual PRC and the PRC estimated from the STA and plotted this against the CV of the ISI distribution. We observed that for both phase oscillator models the estimate provided a good approximation ( $R > 0.75$ ) of the PRC up to  $CV = 0.4$ . For CVs above 0.4, the correlation between the actual and estimated PRC declined rapidly for both models. For the HH model, it was difficult (even with the large noise applied) to increase the CV beyond about 0.3, when in the oscillatory mode. Trying larger noise values, led to numerical difficulties. We remark that in Fig. 3, we have normalized the magnitude of the PRC to a maximum of 1 and that the un-normalized reconstructed PRCs for the HH model had amplitudes that were smaller at large noise values.

*Discussion.*—Relating neural dynamics to neural coding has been termed a grand challenge for neuroscience and we believe that our work describes an important step towards relating these two subjects, albeit for a restricted set of stimuli and neural operating range. Strengthening this connection will provide valuable means of relating the wealth of data on biophysical properties of neurons (as captured in dynamical systems models of neuronal properties) to questions about the properties of stimuli that are being computed by neurons. There are two limitations to our approach. However, our methods can still be applied to many situations. First, the PRC, as useful as it is, has limits to its applicability; specifically, in this context, it is defined only for nearly periodically firing neurons where only the timing of spikes (and not the rate) are altered. Nonetheless, neurons in many brain networks are active spontaneously and the strength of any one input is weak. Thus, the assumption that inputs modulate the timing of spikes rather than adding more spikes may hold. For example, [12] demonstrated that realistic synaptic conductances in the *aplysia* satisfy the mathematical criteria of “weak coupling” in the sense that the notion of phase still makes sense.

Second, traditional STA analysis refers to sensory stimuli rather than directly injected neurons. However, a number of authors have used current injections to understand the “neural code.” The authors of [13] were among the first to try to extract statistical information such as the STA and the spike triggered covariance (STC) from a biophysical model for a neuron. Using simulations of the Hodgkin-Huxley equations, they compute both the STA and the STC in a situation where the stimulus is sufficiently strong to elicit spikes. They show that there are important features of

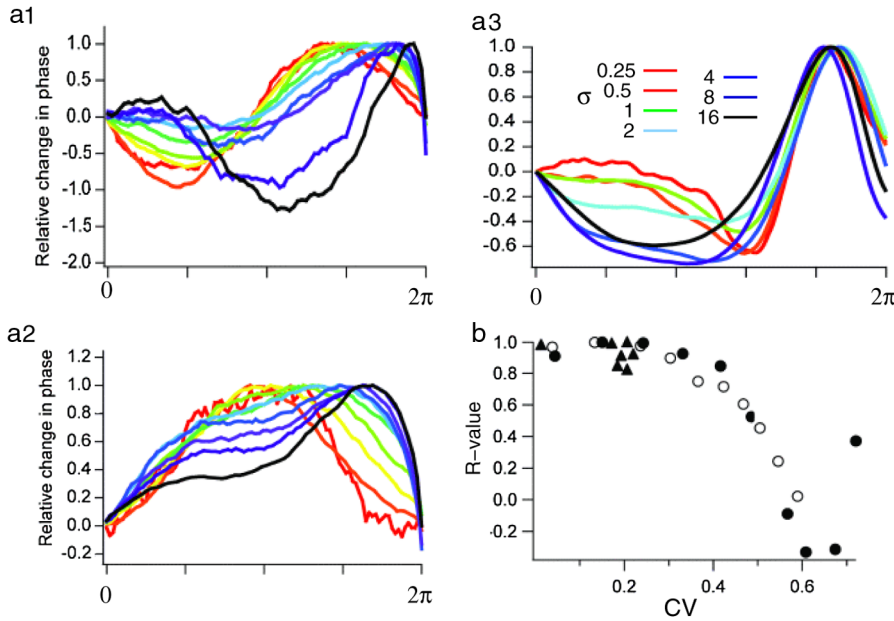


FIG. 3 (color online). Effect of irregular firing on the estimate of the PRC from the STA for phase models and the HH model. Results of simulations of the phase model with  $\Delta(t) = \sin t$  (a1) and  $\Delta(t) = 1 - \cos t$  (a2) with increasing noise amplitude from 0.2 (red)-2.8 (black) in steps of 0.3 (in the order of colors in the rainbow). For both PRCs increasing noise amplitude causes an increase in firing rate. In both cases low noise levels (orange traces) produce most precise estimates of the PRCs. (a3) Reconstruction of the HH PRC with increasing noise values ( $\sigma$  shown in figure) (b) Plot of correlation between actual and estimated PRCs vs CV. (Hollow circles,  $1 - \cos(t)$ , solid circles  $\sin(t)$  and triangles, HH.)

the stimulus encoded in the STC which are not evident in the STA. The calculations in section II could be extended to incorporate higher-order effects and thus relate the STC to other aspects of the PRC [14]. [13] also consider the effects of the amplitude of the stimulus on the STA; such effects would be reflected in changing the shape of the reconstructed PRC [as in Fig. 3(a3)]. While our theory is linear (with respect to the effects of the stimulus), it is known that the shape of the PRC is also affected by stimuli which are sufficiently large [2], so that some of the shape effects in Fig. 3 may be due to pushing the simulations beyond the linear range. More recently, [15] derived the STA for the integrate-and-fire model and produced formulas in the limit of small noise. [16] characterized the spike triggered average voltage (a related quantity which requires intracellular recording of the membrane voltage fluctuations) for several different models and related it to the underlying dynamics while [17] studied the relationship between the STA and STC to subthreshold dynamics. [18] computes the PSTH and STA of a general class of spike-response models, thus providing some insight into the relationships between dynamics of spiking and the neural code. This work is the closest to ours in its generality. They apply their results to the data in [19]. Our methods, while in a more restricted situation (regularly firing) are very general in that they apply to any model of an experimental system for which a PRC is defined. In particular, PRCs encode dynamic information about subthreshold behavior [3,20], which is not possible in spike-response models. The PRC can be directly computed from any biophysical model for periodic neuronal firing.

The authors were supported by NSF Grants, No. DC005798, No. MH079504, and No. DMS0513500 and they would like to thank Brent Doiron and Nicolas Fourcaud-Trocmé for their comments on this

manuscript.

- 
- [1] Y. Kuramoto, *Chemical Oscillations Waves and Turbulence* (Dover, Mineola NY, 2003);
  - [2] A. Winfree *The Geometry of Biological Time* (Springer-Verlag, New York, 2000), 2nd ed.
  - [3] B. S. Gutkin *et al.*, *J. Neurophysiol.* **94**, 1623 (2005).
  - [4] R. F. Galán *et al.*, *Phys. Rev. Lett.* **94**, 158101 (2005).
  - [5] A. D. Reyes and E. E. Fetz, *J. Neurophysiol.* **69**, 1673 (1993).
  - [6] T. Tateno and H. P. Robinson, *Biophys. J.* **92**, 683 (2007).
  - [7] T. I. Netoff *et al.*, *J. Neurophysiol.* **93**, 1197 (2005).
  - [8] F. Rieke *et al.*, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1999); H. L. Bryant and J. P. Segundo, *J. Physiol. (Paris)* **260**, 279 (1976).
  - [9] P. Dayan and L. F. Abbott, *Theoretical Neuroscience Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, MA, 2001).
  - [10] H. Nakao *et al.*, *Phys. Rev. Lett.* **98**, 184101 (2007).
  - [11] B. Ermentrout, *Simulating, Analyzing, and Animating Dynamical Systems: A guide to XPPAUT for Researchers and Students* (SIAM, Philadelphia, PA, 2002), p. 223.
  - [12] A. J. Preyer and R. J. Butera, *Phys. Rev. Lett.* **95**, 138103 (2005).
  - [13] B. Aguera y Arcas *et al.*, *Neural Comput.* **15**, 1715 (2003).
  - [14] B. Ermentrout and B. Doiron (to be published).
  - [15] L. Paninski, *Neural Comput.* **18**, 2592 (2006); L. Paninski *et al.*, *J. Comput. Neurosci.* **21**, 71 (2006).
  - [16] L. Badel *et al.*, *Neurocomputing; Variable Star Bulletin* **69**, 1062 (2006).
  - [17] S. Hong *et al.*, arXiv:q-bio/0612025.
  - [18] A. Herrmann and W. Gerstner, *J. Comput. Neurosci.* **11**, 135 (2001); A. Herrmann and W. Gerstner, *J. Comput. Neurosci.* **12**, 83 (2002).
  - [19] A. V. Poliakov *et al.*, *J. Physiol.* **504**, 401 (1997).
  - [20] B. Ermentrout, *Neural Comput.* **8**, 979 (1996); E. Brown *et al.*, *ibid.* **16**, 673 (2004).