

The Theory of Weakly Coupled Oscillators

Michael A. Schwemmer and Timothy J. Lewis

Department of Mathematics, One Shields Ave, University of California

Davis, CA 95616

1 Introduction

A phase response curve (PRC) (Winfree, 1980) of an oscillating neuron measures the phase shifts in response to stimuli delivered at different times in its cycle. PRCs are often used to predict the phase-locking behavior in networks of neurons and to understand the mechanisms that underlie this behavior. There are two main techniques for doing this. Each of these techniques requires a different kind of PRC, and each is valid in a different limiting case. One approach uses PRCs to reduce neuronal dynamics to firing time maps, e.g (Ermentrout and Kopell, 1998; Guevara et al., 1986; Goel and Ermentrout, 2002; Mirollo and Strogatz, 1990; Netoff et al., 2005b; Oprisan et al., 2004). The second approach uses PRCs to obtain a set of differential equations for the phases of each neuron in the network.

For the derivation of the firing time maps, the stimuli used to generate the PRC should be similar to the input that the neuron actually receives in the network, i.e. a facsimile of a synaptic current or conductance. The firing time map technique can allow one to predict phase-locking for moderately strong coupling, but it has the limitation that the neuron must quickly return to its normal firing cycle before subsequent input arrives. Typically, this implies that input to a neuron must be sufficiently brief and that there is only a single input to a neuron each cycle. The derivation and applications of these firing time maps are discussed in Chapter ZZ.

This chapter focuses on the second technique, which is often referred to as the theory of weakly coupled oscillators (Ermentrout and Kopell, 1984; Kuramoto, 1984; Neu, 1979). The theory of weakly coupled oscillators can be used to predict phase-locking in neuronal networks with any form of coupling, but as the name suggests, the coupling between cells must be sufficiently “weak” for these predictions to be quantitatively accurate. This implies that the coupling can only have small effects on neuronal dynamics over any given period. However, these small effects can accumulate over time and lead to phase-locking in the neuronal network. The theory of weak coupling allows one to reduce the dynamics

of each neuron, which could be of very high dimension, to a single differential equation describing the phase of the neuron. These “phase equations” take the form of a convolution of the input to the neuron via coupling and the neuron’s infinitesimal PRC (iPRC). The iPRC measures the response to a small brief (δ -function-like) perturbation and acts like an impulse response function or Green’s function for the oscillating neurons. Through the dimension reduction and exploiting the form of the phase equations, the theory of weakly coupled oscillators provides a way to identify phase-locked states and understand the mechanisms that underlie them.

The main goal of this chapter is to explain how a weakly coupled neuronal network is reduced to its phase model description. Three different ways to ‘derive’ the phase equations are presented, each providing different insight into the underlying dynamics of phase response properties and phase-locking dynamics. The first derivation (the “Seat-of-the-Pants” derivation in section 3) is the most accessible. It captures the essence of the theory of weak coupling and only requires the reader to know some basic concepts from dynamical system theory, and have a good understanding of what it means for a system to behave linearly. The second derivation (The Geometric Approach in section 4) is a little more mathematically sophisticated and provides deeper insight into the phase response dynamics of neurons. To make this second derivation more accessible, we tie all concepts back to the explanations in the first derivation. The third derivation (The Singular Perturbation Approach in section 5) is the most mathematically abstract but it provides the cleanest derivation of the phase equations. It also explicitly shows that the iPRC can be computed as a solution of the “adjoint” equations.

During these three explanations of the theory of weak coupling, the phase model is derived for a pair of coupled neurons to illustrate the reduction technique. The later sections (section 6 and 7) briefly discuss extensions of the phase model to include heterogeneity, noise, and large networks of neurons.

For more mathematically detailed discussions of the theory of weakly coupled oscillators, we direct the reader to (Ermentrout and Kopell, 1984; Hoppensteadt and Izhikevich, 1997; Kuramoto, 1984; Neu, 1979).

2 Neuronal Models and Reduction to a Phase Model

2.1 General Form of Neuronal Network Models

The general form of a single or multi-compartmental Hodgkin-Huxley-type neuronal model (Hodgkin and Huxley, 1952) is

$$\frac{dX}{dt} = F(X), \quad (1)$$

where X is a N -dimensional state variable vector of containing the membrane potential(s) and gating

variables¹, and $F(X)$ is a vector function describing the rate of change of the variables in time. For the Hodgkin-Huxley (HH) model (Hodgkin and Huxley, 1952), $X = [V, m, h, n]^T$ and

$$F(X) = \begin{bmatrix} \frac{1}{C} (-g_{Na}m^3h(V - E_{Na}) - g_Kn^4(V - E_K) - g_L(V - E_L) + I) \\ \frac{m_\infty(V) - m}{\tau_m(V)} \\ \frac{h_\infty(V) - h}{\tau_h(V)} \\ \frac{n_\infty(V) - n}{\tau_n(V)} \end{bmatrix}, \quad (2)$$

In this chapter, we assume that the isolated model neuron (equation (1)) exhibits stable T -periodic firing (e.g. top trace of Figure 2). In the language of dynamical systems, we assume that the model has an asymptotically stable T -periodic *limit cycle*. These oscillations could be either due to intrinsic conductances or induced by applied current.

A pair of coupled model neurons is described by

$$\frac{dX_1}{dt} = F(X_1) + \varepsilon I(X_1, X_2) \quad (3)$$

$$\frac{dX_2}{dt} = F(X_2) + \varepsilon I(X_2, X_1), \quad (4)$$

where $I(X_1, X_2)$ is a vector function describing the coupling between the two neurons, and ε scales the magnitude of the coupling term. Typically, in models of neuronal networks, cells are only coupled through the voltage (V) equation. For example, a pair of electrically coupled HH neurons would have the coupling term

$$I(X_1, X_2) = \begin{bmatrix} \frac{1}{C} (g_C(V_2 - V_1)) \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (5)$$

where g_C is the coupling conductance of the electrical synapse.

¹The gating variables could be for ionic membrane conductances in the neuron, as well as those describing the output of chemical synapses.

2.2 Phase Models, the G -Function and Phase-Locking

The power of the theory of weakly coupled oscillators is that it reduces the dynamics of each neuronal oscillator in a network to single phase equation that describes the rate of change of its relative phase, ϕ_j . The phase model corresponding to the pair of coupled neurons (3-4) is of the form

$$\frac{d\phi_1}{dt} = \varepsilon H(\phi_2 - \phi_1) \quad (6)$$

$$\frac{d\phi_2}{dt} = \varepsilon H(-(\phi_2 - \phi_1)). \quad (7)$$

The following sections present three different ways of deriving the function H , which is often called the interaction function.

Subtracting the phase equation for cell 1 from that of cell 2, the dynamics can be further reduced to a single equation that governs the evolution of the phase difference between the cells, $\phi = \phi_2 - \phi_1$

$$\frac{d\phi}{dt} = \varepsilon(H(-\phi) - H(\phi)) = \varepsilon G(\phi). \quad (8)$$

In the case of a pair of coupled Hodgkin-Huxley neurons (as described above), the number of equations in the system is reduced from the original 8 describing the dynamics of the voltage and gating variables to a single equation. The reduction method can also be readily applied to multi-compartment model neurons, e.g. (Lewis and Rinzel, 2004; Zahid and Skinner, 2009), which can render a significantly larger dimension reduction. In fact, the method has been applied to real neurons as well, e.g. (Mancilla et al., 2007).

Note that the function $G(\phi)$ or “ G -function” can be used to easily determine the phase-locking behavior of the coupled neurons. The zeros of the G -function, ϕ^* , are the steady state phase differences between the two cells. For example, if $G(0) = 0$, this implies that the synchronous solution is a steady state of the system. To determine the stability of the steady state note that when $G(\phi) > 0$, ϕ will increase and when $G(\phi) < 0$, ϕ will decrease. Therefore, if the derivative of G is positive at a steady state ($G'(\phi^*) > 0$) then the steady state is unstable. Similarly, if the derivative of G is negative at a steady state ($G'(\phi^*) < 0$) then the steady state is stable. Figure 1 shows an example G -function for two coupled identical cells. Note that this system has 4 steady states corresponding to $\phi = 0, T$ (synchrony), $\phi = T/2$ (anti-phase), and two other non-synchronous states. It is also clearly seen that $\phi = 0, T$ and $\phi = T/2$ are stable steady states and the other non-synchronous states are unstable. Thus, the two cells in this system exhibit bistability, and they will either synchronize their firing or fire in anti-phase depending upon the initial conditions.

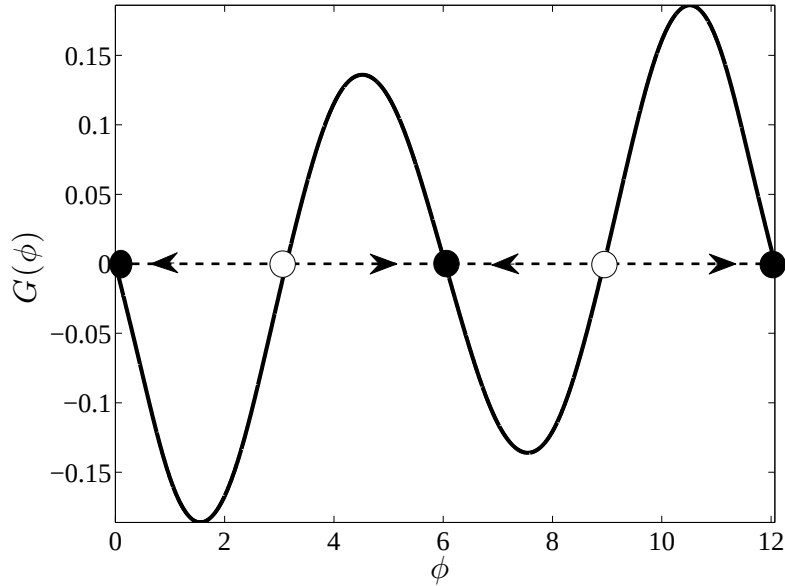


Figure 1: **Example G function.** The G function for two model Fast-Spiking (FS) interneurons (Erisir et al., 1999) coupled with gap junctions on the distal ends of their passive dendrites is plotted. The arrows show the direction of the trajectories for the system. This system has four steady state solutions $\phi_S = 0, T$ (synchrony), $\phi_{AP} = T/2$ (anti-phase), and two non-synchronous states. One can see that synchrony and anti-phase are stable steady-states for this system (filled in circles) while the non-synchronous solutions are unstable (open circles). Thus, depending on the initial conditions, the two neurons will fire synchronously or in anti-phase.

In sections 3, 4 and 5, we present three different ways of derive the interaction function H and therefore the G -function. These derivations make several approximations that require the coupling between neurons to be sufficiently weak. “Sufficiently weak” implies that the neurons’ intrinsic dynamics dominate the effects due to coupling at each point in the periodic cycle, i.e. during the periodic oscillations, $|F(X_j(t))|$ should be an order of magnitude greater than $|\varepsilon I(X_1(t), X_2(t))|$. However, it is important to point out that, even though the phase models quantitatively capture the dynamics of the full system for sufficiently small ε , it is often the case that they can also capture the qualitative behavior for moderate coupling strengths (Lewis and Rinzel, 2003; Netoff et al., 2005a).

3 A “Seat-of-the-Pants” Approach

This section will describe perhaps the most intuitive way of deriving the phase model for a pair of coupled neurons (Lewis and Rinzel, 2003). The approach highlights the key aspect of the theory of weakly coupled oscillators, which is that neurons behave linearly in response to small perturbations and therefore obey the principle of superposition.

3.1 Defining Phase

T -periodic firing of a model neuronal oscillator (equation (1)) corresponds to repeated circulation around an asymptotically stable T -periodic limit cycle, i.e. a closed orbit in state space X . We will denote this T -periodic limit cycle solution as $X_{LC}(t)$. The phase of a neuron is a measure of the time that has elapsed as the neuron's moves around its periodic orbit, starting from an arbitrary reference point in the cycle. We define the phase of the periodically firing neuron j at time t to be

$$\theta_j(t) = (t + \phi_j) \mod T, \quad (9)$$

where $\theta_j = 0$ is set to be at the peak of the neurons' spike (Figure 2).² The constant ϕ_j , which is referred to as the *relative phase* of the j^{th} neuron, is determined by the position of the neuron on the limit cycle at time $t = 0$. Note that each phase of the neuron corresponds to a unique position on the cell's T -periodic limit cycle, and any solution of the uncoupled neuron model that is on the limit cycle can be expressed as

$$X_j(t) = X_{LC}(\theta_j(t)) = X_{LC}(t + \phi_j). \quad (10)$$

When a neuron is perturbed by coupling current from other neurons or by any other external stimulus, its dynamics no longer exactly adhere to the limit cycle, and the exact correspondence of time to phase (equation (9)) is no longer valid. However, when perturbations are sufficiently weak, the neuron's intrinsic dynamics are dominant. This ensures that the perturbed system remains close to the limit cycle and the inter-spike intervals are close to the intrinsic period T . Therefore, we can approximate the solution of neuron j by $X_j(t) \simeq X_{LC}(t + \phi_j(t))$, where the relative phase ϕ_j is now a function of time t . Over each cycle of the oscillations, the weak perturbations to the neurons produce only small changes in ϕ_j . These changes are negligible over a single cycle, but they can slowly accumulate over many cycles and produce substantial effects on the relative firing times of the neurons.

The goal now is to understand how the relative phase $\phi_j(t)$ of the coupled neurons evolves slowly in time. To do this, we first consider the response of a neuron to small abrupt current pulses.

3.2 The Infinitesimal Phase Response Curve

Suppose that a small brief square current pulse of amplitude εI_0 and duration Δt is delivered to a neuron when it is at phase θ^* . This small, brief current pulse causes the membrane potential to abruptly increase

²Phase is often normalized by the period T or by $T/2\pi$, so that $0 \leq \theta < 1$ or $0 \leq \theta < 2\pi$ respectively. Here, we do not normalize phase and take $0 \leq \theta < T$.

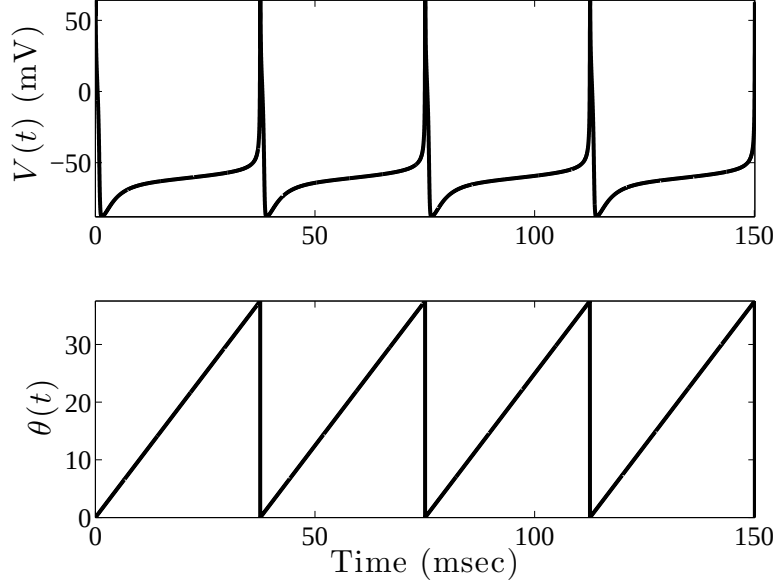


Figure 2: **Phase.** a) Voltage trace for the Fast-Spiking interneuron model from Erisir et al. (Erisir et al., 1999) with $I_{appl} = 35 \mu A/cm^2$ showing T-periodic firing. b) The phase, $\theta(t)$ of these oscillations increases linearly from 0 to T , and we have assumed that zero phase occurs at the peak of the voltage spike.

by $\delta V \simeq \varepsilon I_0 \Delta t / C$, i.e. the change in voltage will approximately equal the total charge delivered to the cell by the stimulus, $\varepsilon I_0 \Delta t$, divided by the capacitance of the neuron, C . In general, this perturbation can cause the cell to fire sooner (phase advance) or later (phase delay) than it would have fired without the perturbation. The magnitude and sign of this *phase shift* depends on the amplitude and duration of the stimulus, as well as the phase in the oscillation at which the stimulus was delivered. This relationship is quantified by the Phase Response Curve (PRC), which gives the phase shift $\Delta\phi$ as a function of the phase θ^* for a fixed $\varepsilon I_0 \Delta t$ (Figure 3).

For sufficiently small and brief stimuli, the neuron will respond in a linear fashion, and the PRC will scale linearly with the magnitude of the current stimulus

$$\Delta\phi(\theta^*) \simeq Z_V(\theta^*) \delta V = Z_V(\theta^*) \left(\frac{1}{C} \varepsilon I_0 \Delta t \right), \quad 0 \leq \theta^* < T, \quad (11)$$

where $Z_V(\theta^*)$ describes the proportional phase shift as a function of the phase of the stimulus. The function $Z_V(\theta)$ is known as the infinitesimal phase response curve (iPRC) or the phase-dependent sensitivity function for voltage perturbations. The iPRC $Z_V(\theta)$ quantifies the normalized phase shift due to an infinitesimally small δ -function-like voltage-perturbation delivered at any given phase on the limit cycle.

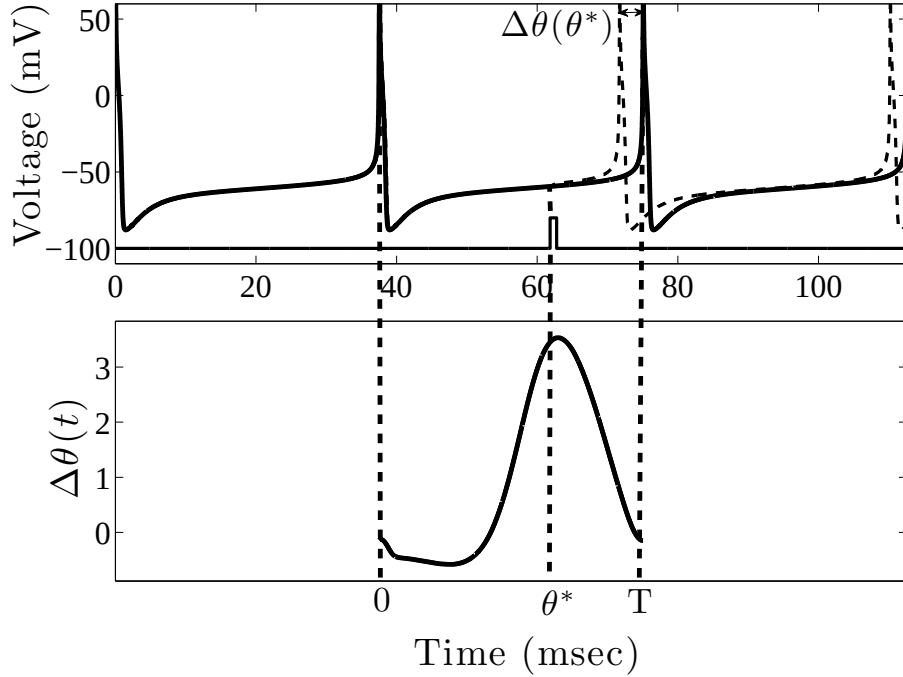


Figure 3: **Measuring the Phase Response Curve from Neurons.** The voltage trace and corresponding PRC is shown for the same FS model neuron from Figure 2. The PRC is measured from a periodically firing neuron by delivering small current pulses at every point, θ^* , along its cycle and measuring the subsequent change in period, $\Delta\theta$, caused by the current pulse.

3.3 The Phase Model for a Pair of Weakly Coupled Cells

Now we can reconsider the pair of weakly coupled neuronal oscillators (equations (3-4)). Recall that, because the coupling is weak, the neurons' intrinsic dynamics dominate the dynamics of the coupled-cell system, and $X_j(t) \simeq X_{LC}(\theta_j(t)) = X_{LC}(t + \phi_j(t))$ for $j = 1, 2$. This assumes that the coupling current can only affect the speed at which cells move around their limit cycle and does not affect the amplitude of the oscillations. Thus, the effects of the coupling are entirely captured in the slow time dynamics of the relative phases of the cells $\phi_j(t)$.

The assumption of weak coupling also ensures that the perturbations to the neurons are sufficiently small so that the neurons respond linearly to the coupling current. That is, (i) the small phase shifts of the neurons due to the presence of the coupling current for a brief time Δt can be approximated using the iPRC (equation (11)), and (ii) these small phase shifts in response to the coupling current sum linearly (i.e. the principle of superposition holds). Therefore, by equation (11), the phase shift due to the coupling current from t to $t + \Delta t$ is

$$\begin{aligned}
\Delta\phi_j(t) &= \phi_j(t + \Delta t) - \phi_j(t) \\
&\simeq Z_V(\theta_j(t)) (\varepsilon I(X_j(t), X_k(t))) \Delta t. \\
&= Z_V(t + \phi_j(t)) (\varepsilon I(X_{LC}(t + \phi_j(t)), X_{LC}(t + \phi_k(t)))) \Delta t.
\end{aligned} \tag{12}$$

Furthermore, by dividing the above equation by Δt and taking the limit as $\Delta t \rightarrow 0$, we obtain a system of differential equations that govern the evolution of the relative phases of the two neurons

$$\frac{d\phi_j}{dt} = \varepsilon Z_V(t + \phi_j) I(X_{LC}(t + \phi_j), X_{LC}(t + \phi_k)), \quad j, k = 1, 2; j \neq k. \tag{13}$$

Note that, by integrating this system of differential equations to find the solution $\phi_j(t)$, we are assuming that phase shifts in response to the coupling current sum linearly.

The explicit time-dependence on the righthand side of equation (13) can be eliminated by “averaging” over the period T . Note that $Z_V(t)$ and $X_{LC}(t)$ are T -periodic functions, and the scaling of the righthand side of equation (13) by the small parameter ε indicates that changes in the relative phases ϕ_j occur on a much slower time scale than T . Therefore, we can integrate the righthand side over the full period T holding the values of ϕ_j constant to find the average rate of change of the ϕ_j over a cycle. Thus, we obtain equations that approximate the slow time evolution of the relative phases ϕ_j

$$\begin{aligned}
\frac{d\phi_j}{dt} &= \varepsilon \frac{1}{T} \int_0^T Z_V(\tilde{t}) (I(X_{LC}(\tilde{t}), X_{LC}(\tilde{t} + \phi_k - \phi_j))) d\tilde{t} \\
&= \varepsilon H(\phi_k - \phi_j), \quad j, k = 1, 2; j \neq k,
\end{aligned} \tag{14}$$

i.e. the relative phases ϕ_j are assumed to be constant with respect to the integral over T in $\tilde{t}$, but they vary in t . This averaging process is made rigorous by averaging theory (Ermentrout and Kopell, 1991; Guckenheimer and Holmes, 1983).

We have reduced the dynamics of a pair of weakly coupled neuronal oscillators to an autonomous system of two differential equations describing the phases of the neurons and therefore finished the first derivation of the phase equations for a pair of weakly coupled neurons.³ Note that the above derivation can be easily altered to obtain the phase model of a neuronal oscillator subjected to T -periodic external forcing as well. The crux of the derivation was identifying the iPRC and exploiting the approximately

³Note that this reduction is not valid when T is of the same order of magnitude as the time scale for the changes due to the weak coupling interactions (e.g. close to a SNIC bifurcation), however an alternative reduction can be performed in this case (Ermentrout, 1996).

linear behavior of the system in response to weak inputs. In fact, it is useful to note that the interaction function H takes the form of a convolution of the iPRC and the coupling current, i.e. the input to the neuron. Therefore, one can think of the iPRC as the oscillator acting like an impulse response function or Green's function.

3.3.1 Averaging theory

Averaging theory (Ermentrout and Kopell, 1991; Guckenheimer and Holmes, 1983) states that there is a change of variables that maps solutions of

$$\frac{d\phi}{d\tilde{t}} = \varepsilon g(\phi, \tilde{t}), \quad (15)$$

where $g(\phi, \tilde{t})$ is a T -periodic function in ϕ and $\tilde{t}$, to solutions of

$$\frac{d\varphi}{dt} = \varepsilon \bar{g}(\varphi) + \mathcal{O}(\varepsilon^2), \quad (16)$$

where

$$\bar{g}(\varphi) = \frac{1}{T} \int_0^T g(\varphi, \tilde{t}) d\tilde{t}, \quad (17)$$

and $\mathcal{O}(\varepsilon^2)$ is Landau's "Big O" notation which represents terms that either have a scaling factor of ε^2 or go to zero at the same rate as ε^2 goes to zero as ε goes to zero.

4 A Geometric Approach

In this section, we describe a geometric approach to the theory of weakly coupled oscillators originally introduced by Yoshiki Kuramoto (Kuramoto, 1984). The main asset of this approach is that it gives a beautiful geometric interpretation of the iPRC and deepens our understanding of the underlying mechanisms of the phase response properties of neurons.

4.1 The One-to-One Map Between Points on the Limit Cycle and Phase

Consider again a model neuron (1) that has a stable T -periodic limit cycle solution, $X_{LC}(t)$ such that the neuron exhibits a T -periodic firing pattern (e.g. top trace of Figure 2). Recall that the phase of the oscillator along its limit cycle is defined as $\theta(t) = (t + \phi) \bmod T$, where the relative phase ϕ is a constant that is determined by the initial conditions. Note that there is a one-to-one correspondence between phase and each point on the limit cycle. That is, the limit cycle solution takes phase to a unique

point on the cycle, $X = X_{LC}(\theta)$, and its inverse maps each point on the limit cycle to a unique phase, $\theta = X_{LC}^{-1}(X) = \Phi(X)$.

Note that it follows immediately from the definition of phase (9) that the rate of change of phase in time along the limit cycle is equal to 1, i.e. $\frac{d\theta}{dt} = 1$. Therefore, if we differentiate the map $\Phi(X)$ with respect to time using the chain rule for vector functions, we obtain the following useful relationship

$$\frac{d\theta}{dt} = \nabla_X \Phi(X_{LC}(t)) \cdot \frac{dX_{LC}}{dt} = \nabla_X \Phi(X_{LC}(t)) \cdot F(X_{LC}(t)) = 1, \quad (18)$$

where $\nabla_X \Phi$ is the gradient of the map $\Phi(X)$ with respect to the vector of the neuron's state variables $X = (x_1, x_2, \dots, x_N)$

$$\nabla_X \Phi(X) = \left[\left(\frac{\partial \Phi}{\partial x_1}, \frac{\partial \Phi}{\partial x_2}, \dots, \frac{\partial \Phi}{\partial x_N} \right) \Big|_X \right]^T. \quad (19)$$

(We have defined the gradient as a column vector for notational reasons).

4.2 Asymptotic Phase and the Infinitesimal Phase Response Curve

The map $\theta = \Phi(X)$ is well-defined for all points X on the limit cycle. We can extend the domain of $\Phi(X)$ to points off the limit cycle by defining the concept of *asymptotic phase*. If X_0 is a point on the limit cycle and Y_0 is a point in a neighborhood of the limit cycle⁴, then we say that Y_0 has the same asymptotic phase as X_0 if $\|X(t; X_0) - X(t; Y_0)\| \rightarrow 0$ as $t \rightarrow \infty$. This means that the solution starting at the initial point Y_0 off the limit cycle converges to the solution starting at the point X_0 on the limit cycle as time goes to infinity. Therefore, $\Phi(Y_0) = \Phi(X_0)$. The set of all points off the limit cycle that have the same asymptotic phase as the point X_0 on the limit cycle is known as the *isochron* (Winfree, 1980) for phase $\theta = \Phi(X_0)$. Figure 4 shows some isochrons around the limit cycle for the Morris-Lecar neuron (Morris and Lecar, 1981). It is important to note that the figure only plots isochrons for a few phases and that *every* point on the limit cycle has a corresponding isochron.

Equipped with the concept of asymptotic phase, we can now show that the iPRC is in fact the gradient of the phase map $\nabla_X \Phi(X_{LC}(t))$ by considering the following phase resetting “experiment”. Suppose that, at time t , the neuron is on the limit cycle in state $X(t) = X_{LC}(\theta^*)$ with corresponding phase $\theta^* = \Phi(X(t))$. At this time, it receives a small abrupt external perturbation εU , where ε is the magnitude of the perturbation and U is the unit vector in the direction of the perturbation in state space. Immediately after the perturbation, the neuron is in the state $X_{LC}(\theta^*) + \varepsilon U$, and its new asymptotic phase is $\tilde{\theta}^* = \Phi(X_{LC}(\theta^*) + \varepsilon U)$. Using Taylor series,

⁴In fact, the point Y_0 can be anywhere in the basin of attraction of the limit cycle.

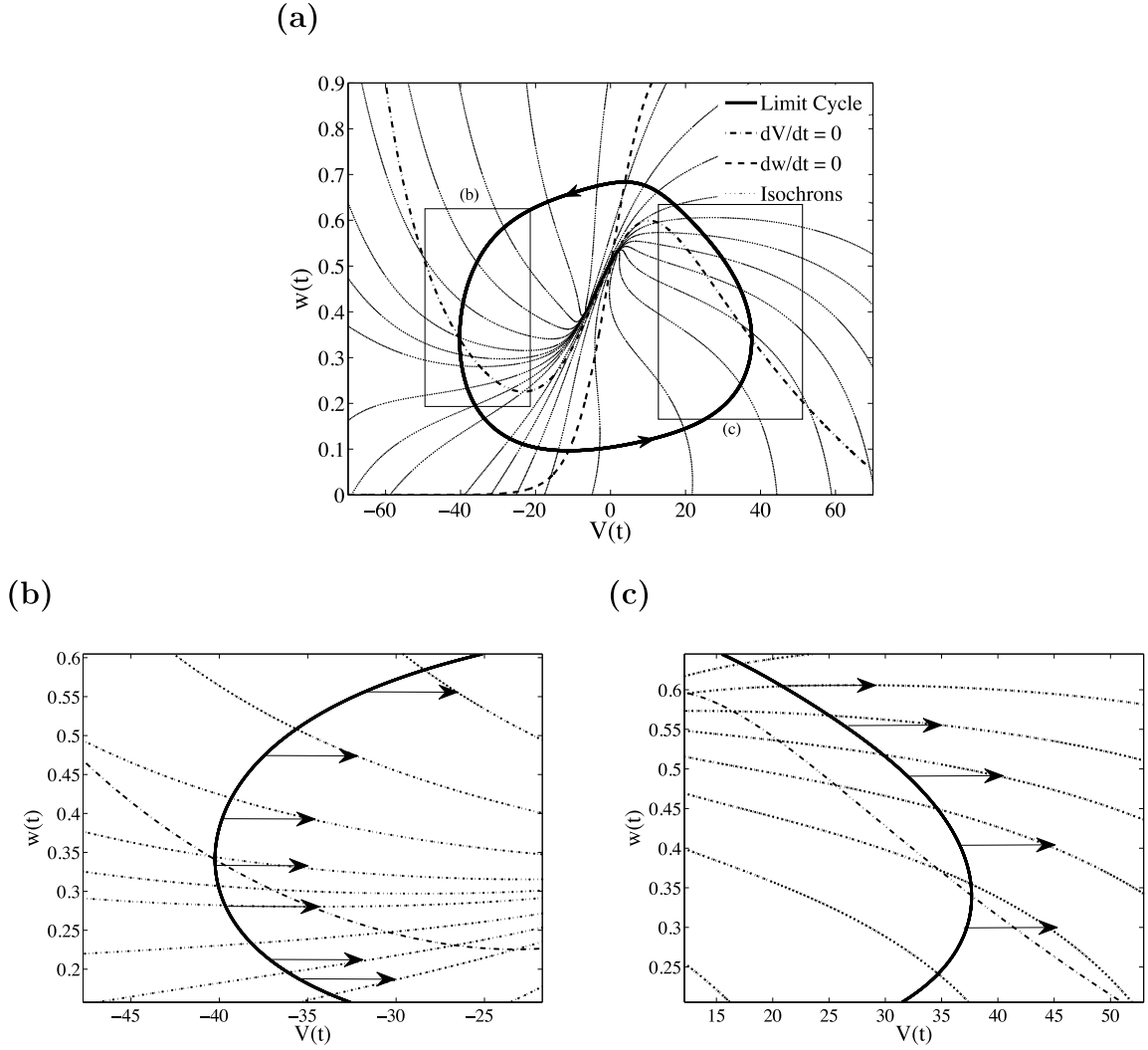


Figure 4: **Example Isochron Structure.** a) The limit cycle and isochron structure for the Morris-Lecar neuron (Morris and Lecar, 1981) is plotted along with the nullclines for the system. b) Blow up of a region on the left side of the limit cycle showing how the same strength perturbation in the voltage direction can cause different size phase delays and even a phase advance. c) Blow up of a region on the right side of the limit cycle showing also that the same size voltage perturbation can cause different size phase advances.

$$\tilde{\theta} = \Phi(X_{LC}(\theta^*) + \varepsilon U) = \Phi(X_{LC}(\theta^*)) + \nabla_X \Phi(X_{LC}(\theta^*)) \cdot (\varepsilon U) + \mathcal{O}(\varepsilon^2). \quad (20)$$

223 Keeping only the linear term (i.e. $\mathcal{O}(\varepsilon)$ term), the phase shift of the neuron as a function of the phase
 224 θ^* at which it received the εU perturbation is given by

$$\Delta\phi(\theta^*) = \tilde{\theta} - \theta^* \simeq \nabla_X \Phi(X_{LC}(\theta^*)) \cdot (\varepsilon U). \quad (21)$$

As was done in section 3.2, we normalize the phase shift by the magnitude of the stimulus,

$$\frac{\Delta\phi(\theta^*)}{\varepsilon} \simeq \nabla_X \Phi(X_{LC}(\theta^*)) \cdot U = Z(\theta^*) \cdot U. \quad (22)$$

Note that $Z(\theta) = \nabla_X \Phi(X_{LC}(\theta))$ is the iPRC. It quantifies the normalized phase shift due to a small delta-function-like perturbation delivered at any given on the limit cycle. As was the case for the iPRC Z_V derived in the previous section (see equation (11)), $\nabla_X \Phi(X_{LC}(\theta))$ captures only the linear response of the neuron and is quantitatively accurate only for sufficiently small perturbations. However, unlike Z_V , $\nabla_X \Phi(X_{LC}(\theta))$ captures the response to perturbations in any direction in state space and not only in one variable (e.g. the membrane potential). That is, $\nabla_X \Phi(X_{LC}(\theta))$ is the vector iPRC; its components are the iPRCs for every variable in the system (see Figure 5).

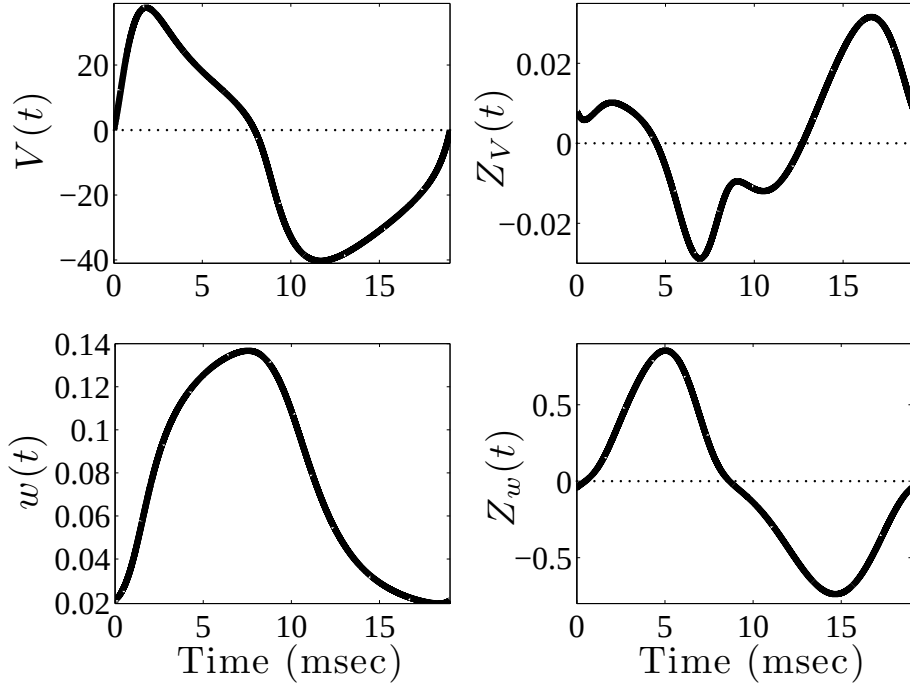


Figure 5: iPRCs for the Morris-Lecar Neuron. The voltage, $V(t)$ and channel, $w(t)$, components of the limit cycle for the same Morris-Lecar neuron as in Figure 4 are plotted along with their corresponding iPRCs. Note that the shape of voltage iPRC can be inferred from the insets of Figure 4. For example, the isochronal structure in Figure 4 (c) reveals that perturbations in the voltage component will cause phase advances when the voltage is increasing from roughly 30 to 38 mV.

In the typical case of a single-compartment HH model neuron subject to an applied current pulse (which perturbs only the membrane potential), the perturbation would be of the form $\varepsilon U = (u, 0, 0, \dots, 0)$ where x_1 is the membrane potential V . By equation (20), the phase shift is

$$\Delta\phi(\theta) = \frac{\partial\Phi}{\partial V}(X_{LC}(\theta)) u = Z_V(\theta) u, \quad (23)$$

which is the same as the equation (11) derived in the previous section.

With the understanding that $\nabla_X\Phi(X_{LC}(t))$ is the vector iPRC, we now derive the phase model for two weakly coupled neurons.

4.3 A Pair of Weakly Coupled Oscillators

Now consider the system of weakly coupled neurons (3-4). We can use the map Φ to take the variables $X_1(t)$ and $X_2(t)$ to their corresponding asymptotic phase, i.e. $\theta_j(t) = \Phi(X_j(t))$ for $j = 1, 2$. By the chain rule, we obtain the change in phase with respect to time

$$\begin{aligned} \frac{d\theta_j}{dt} &= \nabla_X\Phi(X_j(t)) \cdot \frac{dX_j}{dt} \\ &= \nabla_X\Phi(X_j(t)) \cdot [F(X_j(t)) + \varepsilon I(X_j(t), X_k(t))] \\ &= \nabla_X\Phi(X_j(t)) \cdot F(X_j(t)) + \nabla_X\Phi(X_j(t)) \cdot [\varepsilon I(X_j(t), X_k(t))] \\ &= 1 + \varepsilon \nabla_X\Phi(X_j(t)) \cdot I(X_j(t), X_k(t)), \end{aligned} \quad (24)$$

where we have used the “useful” relation (18). Note that the above equations are exact. However, in order to solve the equations for $\theta_j(t)$, we would already have to know the full solutions $X_1(t)$ and $X_2(t)$, in which case you wouldn’t need to reduce the system to a phase model. Therefore, we exploit that fact that ε is small and make the approximation $X_j(t) \sim X_{LC}(\theta_j(t)) = X_{LC}(t + \phi_j(t))$, i.e. the coupling is assumed to be weak enough so that it does not affect the amplitude of the limit cycle, but it can affect the rate at which the neuron moves around its limit cycle. By making this approximation in equation (24) and making the change of variables $\theta_j(t) = t + \phi_j(t)$, we obtain the equations for the evolution of the relative phases of the two neurons

$$\frac{d\phi_j}{dt} = \varepsilon \nabla_X\Phi(X_{LC}(t + \phi_j(t))) \cdot I(X_{LC}(t + \phi_j(t)), X_{LC}(t + \phi_k(t))). \quad (25)$$

Note that these equations are the vector versions of the equations (13) with the iPRC written as $\nabla_X\Phi(X_{LC}(t))$. As described in the previous section, we can average these equations over the period T to eliminate the explicit time dependence and obtain the phase model for the pair of coupled neurons

$$\frac{d\phi_j}{dt} = \varepsilon \frac{1}{T} \int_0^T \nabla_X\Phi(X_{LC}(\tilde{t})) \cdot I(X_{LC}(\tilde{t}), X_{LC}(\tilde{t} + (\phi_k - \phi_j))) d\tilde{t} = \varepsilon H(\phi_k - \phi_j). \quad (26)$$

Note that, while the above approach to deriving the phase equations provides substantial insight into the geometry of the neuronal phase response dynamics, it does not provide a computational method to compute the iPRC for model neurons, i.e. we still must directly measure the iPRC using extensive numerical simulations as described in the previous section.

5 A Singular Perturbation Approach

In this section, we describe the singular perturbation approach to derive the theory of weakly coupled oscillators. This systematic approach was developed independently by Malkin (Malkin, 1949; Malkin, 1956), Neu (Neu, 1979), and Ermentrout and Kopell (Ermentrout and Kopell, 1984). The major practical asset of this approach is that it provides a simple method to compute iPRCs for model neurons.

Consider again the system of weakly coupled neurons (3-4). We assume that the isolated neurons have asymptotically stable T -periodic limit cycle solutions $X_{LC}(t)$ and that coupling is weak (i.e. ε is small). As previously stated, the weak coupling has small effects on the dynamics of the neurons. On the time-scale of a single cycle, these effects are negligible. However, the effects can slowly accumulate on a much slower time-scale and have a substantial influence on the relative firing times of the neurons. We can exploit the differences in these two time-scales and use the method of multiple scales to derive the phase model.

First, we define a “fast time” $t_f = t$, which is on the time-scale of the period of the isolated neuronal oscillator, and a “slow time” $t_s = \varepsilon t$, which is on the time-scale that the coupling affects the dynamics of the neurons. Time, t , is thus a function of both the fast and slow times, i.e. $t = f(t_f, t_s)$. By the chain rule, $\frac{d}{dt} = \frac{\partial}{\partial t_f} + \varepsilon \frac{\partial}{\partial t_s}$. We then assume that solutions $X_1(t)$ and $X_2(t)$ can be expressed as power-series in ε that are dependent both on t_f and t_s ,

$$X_j(t) = X_j^0(t_f, t_s) + \varepsilon X_j^1(t_f, t_s) + \mathcal{O}(\varepsilon^2), \quad j = 1, 2.$$

Substituting these expansions into equations (3-4) yields

$$\begin{aligned} \frac{\partial X_j^0}{\partial t_f} + \varepsilon \frac{\partial X_j^0}{\partial t_s} + \varepsilon \frac{\partial X_j^1}{\partial t_f} + \mathcal{O}(\varepsilon^2) &= F(X_j^0 + \varepsilon X_j^1 + \mathcal{O}(\varepsilon^2)) \\ &+ \varepsilon I(X_j^0 + \varepsilon X_j^1 + \mathcal{O}(\varepsilon^2), X_k^0 + \varepsilon X_k^1 + \mathcal{O}(\varepsilon^2)), \quad j, k = 1, 2; j \neq k. \end{aligned} \tag{27}$$

Using Taylor series to expand the vector functions F and I in terms of ε , we obtain

$$F(X_j^0 + \varepsilon X_j^1 + \mathcal{O}(\varepsilon^2)) = F(X_j^0) + \varepsilon DF(X_j^0)X_j^1 + \mathcal{O}(\varepsilon^2) \quad (28)$$

$$\varepsilon I(X_j^0 + \varepsilon X_j^1 + \mathcal{O}(\varepsilon^2), X_k^0 + \varepsilon X_k^1 + \mathcal{O}(\varepsilon^2)) = \varepsilon I(X_j^0, X_k^0) + \mathcal{O}(\varepsilon^2), \quad (29)$$

where $DF(X_j^0)$ is the Jacobian, i.e. matrix of partial derivatives, of the vector function $F(X_j)$ evaluated at X_j^0 . We then plug these expressions into equations (27), collect like terms of ε , and equate the coefficients of like terms.⁵

The leading order ($\mathcal{O}(1)$) terms yield

$$\frac{\partial X_j^0}{\partial t_f} = F(X_j^0), \quad j = 1, 2. \quad (30)$$

These are the equations that describe the dynamics of the uncoupled cells. Thus, to leading order, each cell exhibits the T -periodic limit cycle solution $X_j^0(t_f, t_s) = X_{LC}(t_f + \phi_j(t_s))$. Note that equation (30) implies that the relative phase ϕ_j is constant in t_f , but it can still evolve on the slow time-scale t_s .

Substituting the solutions for the leading order equations (and shifting t_f appropriately), the $\mathcal{O}(\varepsilon)$ terms of equations (27) yield

$$\mathcal{L}X_j^1 \equiv \frac{\partial X_j^1}{\partial t_f} - DF(X_{LC}(t_f))X_j^1 = I(X_{LC}(t_f), X_{LC}(t_f - (\phi_j(t_s) - \phi_k(t_s)))) - X_{LC}'(t_f) \frac{d\phi_j}{dt_s}. \quad (31)$$

To simplify notation, we have defined the linear operator $\mathcal{L}X \equiv \frac{\partial X}{\partial t_f} - DF(X_{LC}(t_f))X$, which acts on a T -periodic domain and is therefore bounded. Note that equation (31) is a linear differential equation with T -periodic coefficients. In order for our power series solutions for $X_1(t)$ and $X_2(t)$ to exist, a solution to equation (31) must exist. Therefore, we need to find conditions that guarantee the existence of a solution to equation (31), i.e. conditions that ensure that the righthand side of equation (31) is in the range of the operator $\mathcal{L}$. The Fredholm Alternative explicitly provides us with these conditions.

Theorem (Fredholm's Alternative). *Consider the following equation*

$$(*) \quad \mathcal{L}x = \frac{dx}{dt} + A(t)x = f(t); \quad x \in \mathbb{R}^N,$$

where $A(t)$ and $f(t)$ are continuous and T -periodic. Then, there is a continuous T -periodic solution $x(t)$

⁵Because the equation should hold for arbitrary ε , coefficients of like terms must be equal.

to (*) if and only if

$$(**) \quad \frac{1}{T} \int_0^T Z(t) \cdot f(t) dt = 0,$$

for each continuous T -periodic solution, $Z(t)$, to the adjoint problem

$$\mathcal{L}^* x = -\frac{dZ}{dt} + \{A(t)\}^T Z = 0.$$

In the language of the above theorem,

$$A(t) = -DF(X_{LC}(t_f)) \quad \text{and} \quad f(t) = I(X_{LC}(t_f), X_{LC}(t_f - (\phi_j(t_s) - \phi_k(t_s)))) - X'_{LC}(t_f) \frac{d\phi_j}{dt_s}.$$

292 Thus, the solvability condition (**) requires that

$$\frac{1}{T} \int_0^T Z(t_f) \cdot \left[I(X_{LC}(t_f), X_{LC}(t_f - (\phi_j(t_s) - \phi_k(t_s)))) - X'_{LC}(t_f) \frac{d\phi_j}{dt_s} \right] dt_f = 0 \quad (32)$$

293 where Z is a T -periodic solution of the adjoint equation

$$\mathcal{L}^* Z = -\frac{\partial Z}{\partial t_f} - DF(X_{LC}(t_f))^T Z = 0. \quad (33)$$

294 Rearranging equation (32),

$$\frac{d\phi_j}{dt_s} = \frac{1}{T} \int_0^T Z(t_f) \cdot [I(X_{LC}(t_f), X_{LC}(t_f - (\phi_j(t_s) - \phi_k(t_s))))] dt_f \quad (34)$$

295 where we have normalized $Z(t_f)$ by

$$\frac{1}{T} \int_0^T Z(t_f) \cdot [X'_{LC}(t_f)] dt_f = \frac{1}{T} \int_0^T Z(t_f) \cdot F(X_{LC}(t_f)) dt_f = 1. \quad (35)$$

296 This normalization of $Z(t_f)$ is equivalent to setting $Z(0) \cdot X'_{LC}(0) = Z(0) \cdot F(X'_{LC}(0)) = 1$, because

297 $Z(t) \cdot X'_{LC}(t)$ is a constant (see below).

298 Finally, recalling that $t_s = \varepsilon t$ and $t_f = t$, we obtain the phase model for the pair of coupled neurons

$$\frac{d\phi_j}{dt} = \varepsilon \frac{1}{T} \int_0^T Z(\tilde{t}) \cdot [I(X_{LC}(\tilde{t}), X_{LC}(\tilde{t} - (\phi_j - \phi_k)))] d\tilde{t} = \varepsilon H(\phi_k - \phi_j), \quad (36)$$

299 By comparing these phase equations with those derived in the previous sections, it is clear that the

300 appropriately normalized solution to the adjoint equations $Z(t)$ is the iPRC of the neuronal oscillator.

5.0.1 A Note on the Normalization of $Z(t)$

$$\begin{aligned}
\frac{d}{dt} [Z(t) \cdot F(X_{LC}(t))] &= \frac{dZ}{dt} \cdot F(X_{LC}(t)) + Z(t) \cdot \frac{d}{dt} [F(X_{LC}(t))] \\
&= (-DF(X_{LC}(t))^T Z) \cdot F(X_{LC}(t)) + Z(t) \cdot (DF(X_{LC}(t))X'_{LC}(t)) \\
&= -Z(t) \cdot (DF(X_{LC}(t))F(X_{LC}(t))) + Z(t) \cdot (DF(X_{LC}(t))F(X_{LC}(t))) \\
&= 0.
\end{aligned}$$

This implies that $Z(t) \cdot F(X_{LC}(t))$ is a constant. The integral form of the normalization of $Z(t)$ (equation (35)) implies that this constant is 1. Thus, $Z(t) \cdot F(X_{LC}(t)) = Z(t) \cdot X'_{LC}(t) = 1$ for all t , including $t = 0$.

5.1 Adjoints and Gradients

The intrepid reader who has trudged their way courageously through the preceding three sections may be wondering if there is direct way to relate the gradient of the phase map $\nabla_X \Phi(X_{LC}(t))$ to solution of the adjoint equation $Z(t)$. Here, we present Brown et. al's (Brown et al., 2004) direct proof that $\nabla_X \Phi(X_{LC}(t))$ satisfies the adjoint equation (33) and the normalization condition (35).

Consider again the system of differential equations for an isolated neuronal oscillator (1) that has an asymptotically stable T -periodic limit cycle solution $X_{LC}(t)$. Suppose that $X(t) = X_{LC}(t + \phi)$ is a solution of this system that is on the limit cycle, which starts at point $X(0) = X_{LC}(\phi)$. Further suppose that $Y(t) = X_{LC}(t + \phi) + p(t)$ is a solution that starts at from the initial condition $Y(0) = X_{LC}(\phi) + p(0)$, where $p(0)$ is small in magnitude. Because this initial perturbation $p(0)$ is small and the limit cycle is stable, (i) $p(t)$ remains small and, to $\mathcal{O}(|p|)$, $p(t)$ satisfies the linearized system

$$\frac{dp}{dt} = DF(X_{LC}(t + \phi))p, \quad (37)$$

and (ii) the phase difference between the two solutions is

$$\Delta\phi = \Phi(X_{LC}(t + \phi) + p(t)) - \Phi(X_{LC}(t + \phi)) = \nabla_X \Phi(X_{LC}(t + \phi)) \cdot p(t) + \mathcal{O}(|p|^2). \quad (38)$$

Furthermore, while the asymptotic phases of the solutions evolve in time, the phase difference between the solutions $\Delta\phi$ remains constant. Therefore, by differentiating equation (38), we see that to $\mathcal{O}(|p|)$

$$\begin{aligned}
0 &= \frac{d}{dt} [\nabla_X \Phi(X_{LC}(t + \phi)) \cdot p(t)] \\
&= \frac{d}{dt} [\nabla_X \Phi(X_{LC}(t + \phi))] \cdot p(t) + \nabla_X \Phi(X_{LC}(t + \phi)) \cdot \frac{dp}{dt} \\
&= \frac{d}{dt} [\nabla_X \Phi(X_{LC}(t + \phi))] \cdot p(t) + \nabla_X \Phi(X_{LC}(t + \phi)) \cdot (DF(X_{LC}(t + \phi))p(t)) \\
&= \frac{d}{dt} [\nabla_X \Phi(X_{LC}(t + \phi))] \cdot p(t) + (DF(X_{LC}(t + \phi))^T \nabla_X \Phi(X_{LC}(t + \phi))) \cdot p(t) \\
&= \left\{ \frac{d}{dt} [\nabla_X \Phi(X_{LC}(t + \phi))] + DF(X_{LC}(t + \phi))^T (\nabla_X \Phi(X_{LC}(t + \phi))) \right\} \cdot p(t).
\end{aligned}$$

Because p is arbitrary, the above argument implies that $\nabla_X \Phi(X_{LC}(t))$ solves the adjoint equation (33).

The normalization condition simply follows from the definition of the phase map (see equation(18)), i.e.

$$\frac{d\theta}{dt} = \nabla_X \Phi(X_{LC}(t)) \cdot X'_{LC}(t) = 1. \quad (39)$$

5.2 Computing the PRC Using the Adjoint method

As stated in this beginning of this section, the major practical asset of the singular perturbation approach is that it provides a simple method to compute the iPRC for model neurons. Specifically, the iPRC is a T -period solution to

$$\frac{dZ}{dt} = -DF(X_{LC}(t))^T Z \quad (40)$$

subject to the normalization constraint

$$Z(0) \cdot X'_{LC}(0) = 1. \quad (41)$$

This equation is the adjoint equation for the isolated model neuron (equation (1)) linearized around the limit cycle solution $X_{LC}(t)$.

In practice, the solution to equation (40) is found by integrating the equation backwards in time (Williams and Bowtell, 1997). The adjoint system has the opposite stability of the original system (equation (1)), which has an asymptotically stable T -periodic limit cycle solution. Thus, we integrate backwards in time from an arbitrary initial condition so as to dampen out the transients and arrive at the (unstable) periodic solution of equation (40). To obtain the iPRC, we normalize the periodic solution using (41). This algorithm is automated in the software package XPPAUT (Ermentrout, 2002), which is available for free on Bard Ermentrout's webpage www.math.pitt.edu/~bard/bardware/.

6 Extensions of Phase Models for Pairs of Coupled Cells

Up to this point, we have been dealing solely with pairs of identical oscillators that are weakly coupled. In this section, we show how the phase reduction technique can be extended to incorporate weak heterogeneity and weak noise.

6.1 Weak Heterogeneity

Suppose that the following system

$$\frac{dX_j}{dt} = F_j(X_j) + \varepsilon I(X_k, X_j) = F(X_j) + \varepsilon [f_j(X_j) + I(X_k, X_j)] \quad (42)$$

describes two weakly coupled neuronal oscillators (note that the vector functions $F_j(X_j)$ are now specific to the neuron). If the two neurons are weakly heterogeneous, then their underlying limit cycles are equivalent up to an $\mathcal{O}(\varepsilon)$ difference. That is, $F_j(X_j) = F(X_j) + \varepsilon f_j(X_j)$, where $f_j(X_j)$ is a vector function that captures the $\mathcal{O}(\varepsilon)$ differences in the dynamics of cell 1 and cell 2 from the function $F(X_j)$. These differences may occur in various places such as the value of the neurons' leakage conductances, the applied currents, or the leakage reversal potentials, to name a few.

As in the previous sections, equation (42) can be reduced to the phase model

$$\begin{aligned} \frac{d\phi_j}{dt} &= \varepsilon \left(\frac{1}{T} \int_0^T Z(\tilde{t}) \cdot [f_j(X_{LC}(\tilde{t})) + I(X_{LC}(\tilde{t}), X_{LC}(\tilde{t} + \phi_k - \phi_j))] d\tilde{t} \right) \\ &= \varepsilon \omega_j + \varepsilon H(\phi_k - \phi_j), \end{aligned} \quad (43)$$

where $\omega_j = \frac{1}{T} \int_0^T Z(\tilde{t}) \cdot f_j(X_{LC}(\tilde{t})) d\tilde{t}$ represents the difference in the intrinsic frequencies of the two neurons caused by the presence of the weak heterogeneity. If we now let $\phi = \phi_2 - \phi_1$, we obtain

$$\begin{aligned} \frac{d\phi}{dt} &= \varepsilon (H(-\phi) - H(\phi) + \Delta\omega) \\ &= \varepsilon (G(\phi) + \Delta\omega), \end{aligned} \quad (44)$$

where $\Delta\omega = \omega_2 - \omega_1$. The fixed points of (44) are given by $G(\phi) = -\Delta\omega$. The addition of the heterogeneity changes the phase-locking properties of the neurons. For example, suppose that in the absence of heterogeneity ($\Delta\omega = 0$) our G function is the same as in Figure 1, in which the synchronous solution, $\phi_S = 0$, and the anti-phase solution, ϕ_{AP} , are stable. Once heterogeneity is added, the effect will

be to move the neurons away from either firing in synchrony or anti-phase to a constant non-synchronous phase shift, as in Figure 6. For example, if neuron 1 is faster than neuron 2, then $\Delta\omega < 0$ and the stable steady-state phase-locked values of ϕ will be shifted to left of synchrony and to the left of anti-phase, as is seen in Figure 6 when $\Delta\omega = -0.5$. Thus, the neurons will still be phase-locked, but in a non-synchronous state that will either be to the left of synchronous state or to the left of the anti-phase state depending on the initial conditions. Furthermore, if $\Delta\omega$ is decreased further, saddle node bifurcations occur in which a stable and unstable fixed point collide and annihilate each other.

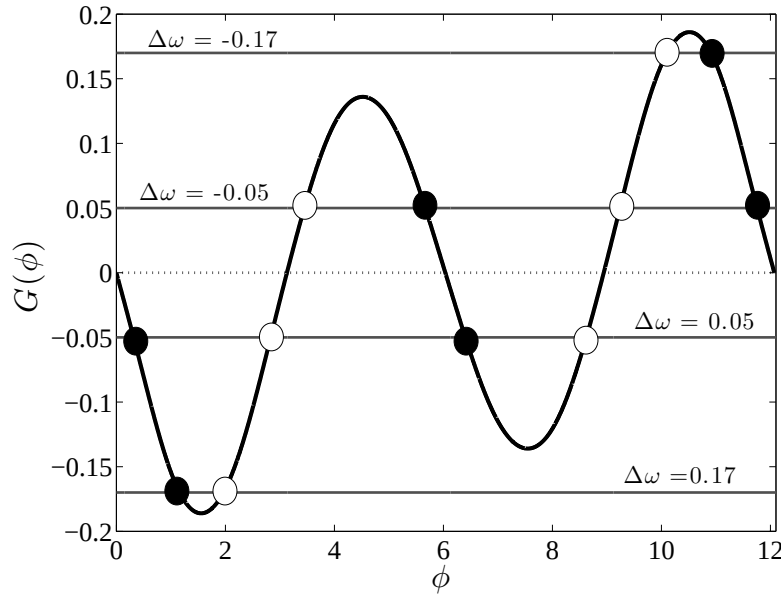


Figure 6: **Example G Function with Varying Heterogeneity.** Example of varying levels of heterogeneity with the same G function as in Figure 1. One can see that the addition of any level of heterogeneity will cause the stable steady-state phase-locked states to move to away from the synchronous and anti-phase states to non-synchronous phase-locked states. Furthermore, if the heterogeneity is large enough, the stable steady-state phase-locked states will disappear completely through saddle node bifurcations.

6.2 Weakly Coupled Neurons with Noise

In this section, we show how two weakly coupled neurons with additive white noise in the voltage component can be analyzed using a probability density approach (Kuramoto, 1984; Pfeuty et al., 2005).

The following set of differential equations represent two weakly heterogeneous neurons being perturbed with additive noise

$$\frac{dX_j}{dt} = F_j(X_j) + \varepsilon I(X_k, X_j) + \delta N_j(t), \quad i, j = 1, 2; i \neq j, \quad (45)$$

where δ scales the noise term to ensure that it is $\mathcal{O}(\varepsilon)$. The term $N_j(t)$ is a vector with Gaussian white noise, $\xi_j(t)$, with zero mean and unit variance (i.e. $\langle \xi_j(t) \rangle = 0$ and $\langle \xi_j(t) \xi_j(t') \rangle = \delta(t - t')$) in the voltage component, and zeros in the other variable components. In this case, the system can be mapped to the phase model

$$\frac{d\phi_j}{dt} = \varepsilon(\omega_j + H(\phi_k - \phi_j)) + \delta\sigma_\phi \xi_j(t), \quad (46)$$

where the term $\sigma_\phi = \left(\frac{1}{T} \int_0^T [Z(\tilde{t})]^2 d\tilde{t} \right)^{1/2}$ comes from averaging the noisy phase equations (Kuramoto, 1984). If we now let $\phi = \phi_2 - \phi_1$ we arrive at

$$\frac{d\phi}{dt} = \varepsilon(\Delta\omega + (H(-\phi) - H(\phi))) + \delta\sigma_\phi \sqrt{2}\eta(t), \quad (47)$$

where $\Delta\omega = \omega_2 - \omega_1$ and $\sqrt{2}\eta(t) = \xi_2(t) - \xi_1(t)$ where $\eta(t)$ is Gaussian white noise with zero mean and unit variance.

The non-linear Langevin equation (47) corresponds to the Fokker-Planck equation (Risken, 1989; Stratonovich, 1967; Van Kampen, 1981)

$$\frac{\partial \rho}{\partial t}(\phi, t) = -\frac{\partial}{\partial \phi}[\varepsilon(\Delta\omega + G(\phi))\rho(\phi, t)] + (\delta\sigma_\phi)^2 \frac{\partial^2 \rho}{\partial \phi^2}(\phi, t), \quad (48)$$

where $\rho(\phi, t)$ is the probability that the neurons have a phase difference of ϕ at time t . The steady-state $\left(\frac{\partial \rho}{\partial t} = 0 \right)$ solution of equation (48) is

$$\rho(\phi) = \frac{1}{N} e^{M(\phi)} \left[\frac{e^{-\alpha T \Delta\omega} - 1}{\int_0^T e^{-M(\bar{\phi})} d\bar{\phi}} \int_0^\phi e^{-M(\bar{\phi})} d\bar{\phi} + 1 \right], \quad (49)$$

where

$$M(\phi) = \alpha \int_0^\phi (\Delta\omega + G(\bar{\phi})) d\bar{\phi}, \quad (50)$$

N is a normalization factor so that $\int_0^T \rho(\phi) d\phi = 1$, and $\alpha = \frac{\varepsilon}{\delta^2 \sigma_\phi^2}$ represents the ratio of the strength of the coupling to the variance of the noise.

The steady-state distribution $\rho(\phi)$ tells us the probability that the two neurons will have a phase difference of ϕ as time goes to infinity. Furthermore, Pfeuty et al. (Pfeuty et al., 2005) showed that spike-train cross-correlogram of the two neurons is equivalent to the steady state distribution (49) for small ε . Figure 7 (a) shows the cross-correlogram for two identical neurons ($\Delta\omega = 0$) using the G function from Figure 1. One can see that there is a large peak in the distribution around the synchronous solution ($\phi_S = 0$), and a smaller peak around the anti-phase solution ($\phi_{AP} = T/2$). Thus, the presence of the

noise works to smear out the probability distribution around the stable steady-states of the noiseless system.

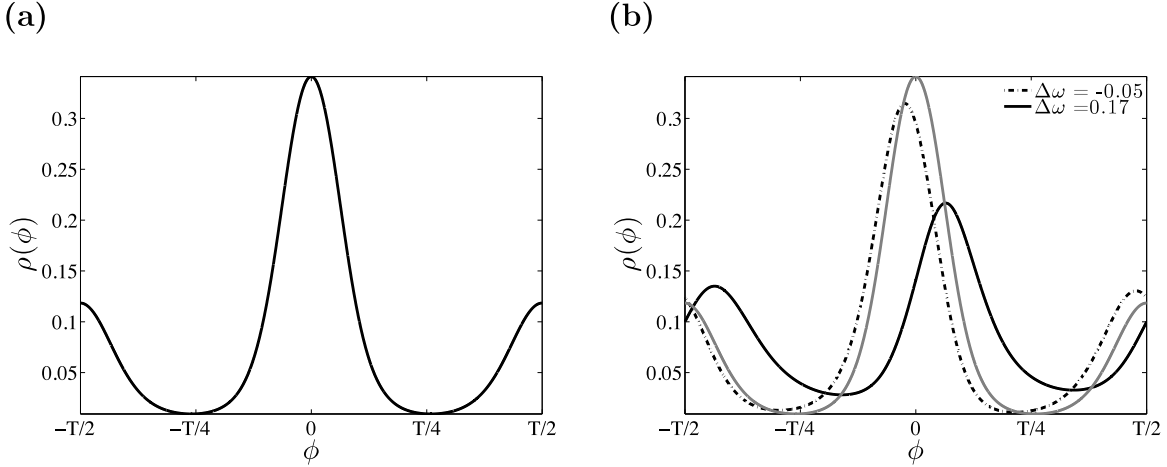


Figure 7: **The Steady-State Phase Difference Distribution $\rho(\phi)$ is the Cross-Correlogram for the Two Neurons.** (a) Cross-correlogram for the G function given in Figure 1 with $\alpha = 10$. Note that we have changed the x -axis so that ϕ now ranges from $-T/2$ to $T/2$. The cross-correlogram has two peaks corresponding to the synchronous and anti-phase phase-locked states. This is due to the fact that in the noiseless system, synchrony and anti-phase were the only stable fixed points. (b) Cross-correlograms for two levels of heterogeneity from Figure 6. The cross-correlogram from (a) is plotted as the light solid line for comparison. The peaks in the cross-correlogram have shifted to correspond with the stable non-synchronous steady-states in Figure 6.

If heterogeneity is added to the G function as in Figure 6, one would expect that the peaks of the cross-correlogram would shift accordingly so as to correspond to the stable steady-states of the noiseless system. Figure 7 (b) shows that this is indeed the case. If $\Delta\omega < 0$ ($\Delta\omega > 0$), the stable steady-states of the noiseless system shift to the left (right) of synchrony and to the left (right) of anti-phase, thus causing the peaks of the cross-correlogram to shift left (right) as well. If we were to increase (decrease) the noise, i.e. decrease (increase) α , then we would see that the variance of the peaks around the stable steady-states becomes larger (smaller), according to equation (49).

7 Networks of Weakly Coupled Neurons

In this section, we extend the phase model description to examine networks of weakly coupled neuronal oscillators.

Suppose we have a one spatial dimension network of M weakly coupled and weakly heterogeneous neurons

$$\frac{dX_i}{dt} = F_i(X_i) + \frac{\varepsilon}{M_0} \sum_{j=1}^M s_{ij} I(X_j, X_i), \quad i = 1, \dots, M; \quad (51)$$

where $S = \{s_{ij}\}$ is the connectivity matrix of the network, M_0 is the maximum number of cells that any neuron is connected to and the factor of $\frac{1}{M_0}$ ensures that the perturbation from the coupling is $\mathcal{O}(\varepsilon)$. As before, this system can be reduced to the phase model

$$\frac{d\phi_i}{dt} = \omega_i + \frac{\varepsilon}{M_0} \sum_{j=1}^M s_{ij} H(\phi_j - \phi_i), \quad i = 1, \dots, M. \quad (52)$$

The connectivity matrix, S , can be utilized to examine the effects of network topology on the phase-locking behavior of the network. For example, if we wanted to examine the activity of a network in which each neuron is connected to every other neuron, i.e. all-to-all coupling, then

$$s_{ij} = 1, \quad i, j = 1, \dots, M. \quad (53)$$

Because of the non-linear nature of equation (52), analytic solutions normally cannot be found. Furthermore, it can be quite difficult to analyze for large numbers of neurons. Fortunately, there exist two approaches to simplifying equation (52) so that mathematical analysis can be utilized, which is not to say that simulating the system equation (52) is not useful. Depending upon the type of interaction function that is used, various types of interesting phase-locking behavior can be seen, such as total synchrony, traveling oscillatory waves, or, in two spatial dimensional networks, spiral waves and target patterns, e.g. (Ermentrout and Kleinfeld, 2001; Kuramoto, 1984).

A useful method of determining the level of synchrony for the network (52) is the so-called Kuramoto synchronization index (Kuramoto, 1984)

$$r e^{2\pi i \psi / T} = \frac{1}{M} \sum_{j=1}^M e^{2\pi i \phi_j / T}, \quad (54)$$

where $i = \sqrt{-1}$, ψ is the average phase of the network, and r is the level of synchrony of the network. This index maps the phases, ϕ_j , to vectors in the complex plane and then averages them. Thus, if the neurons are in synchrony, the corresponding vectors will all be pointing in the same direction and r will be equal to one. The less synchronous the network is, the smaller the value of r .

In the following two sections, we briefly outline two different mathematical techniques for analyzing these phase oscillator networks in the limit as M goes to infinity.

7.1 Population Density Method

A powerful method to analyze large networks of all-to-all coupled phase oscillators was introduced by Strogatz and Mirollo (Strogatz and Mirollo, 1991) where they considered the so-called Kuramoto model with additive white noise

$$\frac{d\phi_i}{dt} = \omega_i + \frac{\varepsilon}{M} \sum_{j=1}^M H(\phi_j - \phi_i) + \sigma \xi(t), \quad (55)$$

where the interaction function is a simple sine function, i.e. $H(\phi) = \sin(\phi)$. A large body of work has been focused on analyzing the Kuramoto model as it is the simplest model for describing the onset of synchronization in populations of coupled oscillators (Acebrón et al., 2005; Strogatz, 2000). However, in this section, we will examine the case where $H(\phi)$ is a general T -periodic function.

The idea behind the approach of (Strogatz and Mirollo, 1991) is to derive the Fokker-Planck equation for (55) in the limit as $M \rightarrow \infty$, i.e. the number of neurons in the network is infinite. As a first step, note that by equating real and imaginary parts in equation (54) we arrive at the following useful relations

$$r \cos(2\pi(\psi - \phi_i)/T) = \frac{1}{M} \sum_{j=1}^M \cos(2\pi(\phi_j - \phi_i)/T) \quad (56)$$

$$r \sin(2\pi(\psi - \phi_i)/T) = \frac{1}{M} \sum_{j=1}^M \sin(2\pi(\phi_j - \phi_i)/T). \quad (57)$$

Next, we note that since $H(\phi)$ is T -periodic, we can represent it as a Fourier series

$$H(\phi_j - \phi_i) = \frac{1}{T} \sum_{n=0}^{\infty} a_n \cos(2\pi n(\phi_j - \phi_i)/T) + b_n \sin(2\pi n(\phi_j - \phi_i)/T). \quad (58)$$

Recognizing that equations (56) and (57) are averages of the functions cosine and sine, respectively, over the phases of the other oscillators, we see that, in the limit as M goes to infinity (Neltner et al., 2000; Strogatz and Mirollo, 1991)

$$r a_n \cos(2\pi n(\psi_n - \phi)/T) = a_n \int_{-\infty}^{\infty} \int_0^T g(\omega) \rho(\tilde{\phi}, \omega, t) \cos(2\pi n(\tilde{\phi} - \phi)/T) d\tilde{\phi} d\omega \quad (59)$$

$$r b_n \sin(2\pi n(\psi_n - \phi)/T) = b_n \int_{-\infty}^{\infty} \int_0^T g(\omega) \rho(\tilde{\phi}, \omega, t) \sin(2\pi n(\tilde{\phi} - \phi)/T) d\tilde{\phi} d\omega, \quad (60)$$

where we have used the Fourier coefficients of $H(\phi_j - \phi_i)$. $\rho(\phi, \omega, t)$ is the density of oscillators with uncoupled frequency ω and phase ϕ at time t , and $g(\omega)$ is the density function for the distribution of the

frequencies of the oscillators. Furthermore, $g(\omega)$ also satisfies $\int_{-\infty}^{\infty} g(\omega) d\omega = 1$. With all this in mind, we can now rewrite the infinite M approximation of equation (55)

$$\frac{d\phi}{dt} = \omega + \varepsilon \frac{1}{2\pi} \sum_{n=0}^{\infty} [ra_n \cos(2\pi n(\psi_n - \phi)/T) + rb_n \sin(2\pi n(\psi_n - \phi)/T)] + \sigma \xi(t). \quad (61)$$

The above nonlinear Langevin equation corresponds to the Fokker-Planck equation

$$\frac{\partial \rho}{\partial t}(\phi, \omega, t) = -\frac{\partial}{\partial \phi} [J(\phi, t) \rho(\phi, \omega, t)] + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial \phi^2}(\phi, \omega, t), \quad (62)$$

with

$$J(\phi, t) = \omega + \varepsilon \frac{1}{T} \sum_{n=0}^{\infty} [ra_n \cos(2\pi n(\psi_n - \phi)/T) + rb_n \sin(2\pi n(\psi_n - \phi)/T)], \quad (63)$$

and $\int_0^T \rho(\phi, \omega, t) d\phi = 1$ and $\rho(\phi, \omega, t) = \rho(\phi + T, \omega, t)$. Equation (62) tells us how the fraction of oscillators with phase ϕ and frequency ω evolves with time. Note that equation (62) has the trivial solution $\rho_0(\phi, \omega, t) = \frac{1}{T}$, which corresponds to the incoherent state in which the phases of the neurons are uniformly distributed between 0 and T .

To study the onset of synchronization in these networks, Strogatz and Mirollo (Strogatz and Mirollo, 1991) and others, e.g. (Neltner et al., 2000), linearized equation (62) around the incoherent state, ρ_0 , in order to determine its stability. They were able to prove that below a certain value of ε , the incoherent state is neutrally stable and then loses stability at some critical value $\varepsilon = \varepsilon_C$. After this point, the network becomes more and more synchronous as ε is increased.

7.2 Continuum Limit

Although the population density approach is powerful method for analyzing the phase-locking dynamics of neuronal networks, it is limited by the fact that it does not take into account spatial effects of neuronal networks. An alternative approach to analyzing (52) in the large M limit that takes into account spatial effects is to assume that the network of neuronal oscillators forms a spatial continuum (Bressloff and Coombes, 1997; Crook et al., 1997; Ermentrout, 1985).

Suppose that we have a one-dimensional array of neurons in which the j^{th} neuron occupies the position $x_j = j\Delta x$ where Δx is the spacing between the neurons. Further suppose that the connectivity matrix is defined by $S = \{s_{ij}\} = W(|x_j - x_i|)$, where $W(|x|) \rightarrow 0$ as $|x| \rightarrow \infty$ and $\sum_{j=-\infty}^{\infty} W(x_j) \Delta x = 1$. For example, the spatial connectivity matrix could correspond to a Gaussian function, $W(|x_j - x_i|) = e^{-\frac{|x_j - x_i|^2}{2\sigma^2}}$, so that closer neurons have more strongly coupled to each other than to neurons that are

further apart. We can now rewrite equation (52) as

$$\frac{d\phi}{dt}(x_i, t) = \omega(x_i) + \varepsilon \sum_{j=-\infty}^{\infty} [W(|x_j - x_i|) \Delta x H(\phi(x_j, t) - \phi(x_i, t))], \quad (64)$$

where $\phi(x_i, t) = \phi_i(t)$, $\omega(x_i) = \omega_i$ and we have taken $1/M = \Delta x$. By taking the limit of $\Delta x \rightarrow 0$ ($M \rightarrow \infty$) in equation (64), we arrive at the continuum phase model

$$\frac{\partial \phi}{\partial t}(x, t) = \omega(x) + \varepsilon \int_{-\infty}^{\infty} W(|x - \bar{x}|) H(\phi(\bar{x}, t) - \phi(x, t)) d\bar{x}, \quad (65)$$

where $\phi(x, t)$ is the phase of the oscillator at position x and time t . Note that this continuum phase model can be modified to account for finite spatial domains (Ermentrout, 1992) and to include multiple spatial dimensions.

Various authors have utilized this continuum approach to prove results about the stability of the synchrony and traveling wave solutions of equation (65) (Bressloff and Coombes, 1997; Crook et al., 1997; Ermentrout, 1985; Ermentrout, 1992). For example, Crook et al. (Crook et al., 1997) were able to prove that presence of axonal delay in synaptic transmission between neurons can cause the onset of traveling wave solutions. This is due to the presence of axonal delay which encourages larger phase shifts between neurons that are further apart in space. Similarly, Bressloff and Coombes (Bressloff and Coombes, 1997) derived the continuum phase model for a network of integrate-and-fire neurons coupled with excitatory synapses on their passive dendrites. Using this model, they were able to show that long range excitatory coupling can cause the system to undergo a bifurcation from the synchronous state to traveling oscillatory waves. For a rigorous mathematical treatment of the existence and stability results for general continuum and discrete phase model neuronal networks, we direct the reader to (Ermentrout, 1992).

8 Summary

- The infinitesimal PRC (iPRC) of a neuron measures its sensitivity to infinitesimally small perturbations at every point along its cycle.
- The theory of weak coupling utilizes the iPRC to reduce the complexity of neuronal network to consideration of a single phase variable for every neuron.
- The theory is valid only when the perturbations to the neuron, from coupling or an external source, is sufficiently “weak” so that the neuron’s intrinsic dynamics dominate the influence of the

coupling. This implies that coupling does not cause the neuron's firing period to differ greatly from its unperturbed cycle.

- For two weakly coupled neurons, the theory allows one to reduce the dynamics to consideration of a single equation describing how the phase difference of the two oscillators changes in time. This allows for the prediction of the phase-locking behavior of the cell-pair through simple analysis of the phase difference equation.
- The theory of weak coupling can be extended to incorporate effects from weak heterogeneity and weak noise.

Acknowledgements. This work was supported by the National Science Foundation under grants DMS-09211039 and DMS-0518022.

References

- Acebrón, J., Bonilla, L., Vicénte, C., Ritort, F., and Spigler, R. (2005). The kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.*, 77:137–185.
- Bressloff, P. and Coombes, S. (1997). Synchrony in an array of integrate-and-fire neurons with dendritic structure. *Phys. Rev. Lett.*, 78:4665–4668.
- Brown, E., Moehlis, J., and Holmes, P. (2004). On the phase reduction and response dynamics of neural oscillator populations. *Neural Comp.*, 16:673–715.
- Crook, S., Ermentrout, G., Vanier, M., and Bower, J. (1997). The role of axonal delay in the synchronization of networks of coupled cortical oscillators. *J. Comp. Neurosci.*, 4:161–172.
- Erisir, A., Lau, D., Rudy, B., and Leonard, C. (1999). Function of specific k^+ channels in sustained high-frequency firing of fast-spiking neocortical interneurons. *J. Neurophysiol.*, 82:2476–2489.
- Ermentrout, B. (2002). *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students*. SIAM.
- Ermentrout, G. (1985). The behavior of rings of coupled oscillators. *J. Math. Biology*, 23:55–74.
- Ermentrout, G. (1992). Stable periodic solutions to discrete and continuum arrays of weakly coupled nonlinear oscillators. *SIAM J. Appl. Math.*, 52(6):1665–1687.

Ermentrout, G. (1996). Type 1 membranes, phase resetting curves, and synchrony. *Neural Computation*, 8:1979–1001.

Ermentrout, G. and Kleinfeld, D. (2001). Traveling electrical waves in cortex: Insights from phase dynamics and speculation on a computational role. *Neuron*, 29:33–44.

Ermentrout, G. and Kopell, N. (1984). Frequency plateaus in a chain of weakly coupled oscillators, i. *SIAM J. Math. Anal.*, 15(2):215–237.

Ermentrout, G. and Kopell, N. (1991). Multiple pulse interactions and averaging in systems of coupled neural oscillators. *J. Math. Bio.*, 29:33–44.

Ermentrout, G. and Kopell, N. (1998). Fine structure of neural spiking and synchronization in the presence of conduction delays. *Proc. Nat. Acad. of Sci.*, 95(3):1259–1264.

Goel, P. and Ermentrout, G. (2002). Synchrony, stability, and firing patterns in pulse-coupled oscillators. *Physica D*, 163:191–216.

Guckenheimer, J. and Holmes, P. (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag, NY.

Guevara, M. R., Shrier, A., and Glass, L. (1986). Phase resetting of spontaneously beating embryonic ventricular heart cell aggregates. *Am J Physiol Heart Circ Physiol*, 251(6):H1298–1305.

Hodgkin, A. and Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500–544.

Hoppensteadt, F. C. and Izhikevich, E. M. (1997). *Weakly Connected Neural Networks*. Springer-Verlag, New York.

Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag, Berlin.

Lewis, T. and Rinzel, J. (2003). Dynamics of spiking neurons connected by both inhibitory and electrical coupling. *J. Comp. Neurosci.*, 14:283–309.

Lewis, T. and Rinzel, J. (2004). Dendritic effects in networks of electrically coupled fast-spiking interneurons. *Neurocomputing*, 58-60:145–150.

Malkin, I. (1949). *Methods of Poincare and Liapunov in Theory of Non-Linear Oscillations*. Gostexizdat, Moscow.

Malkin, I. (1956). *Some Problems in Nonlinear Oscillation Theory*. Gostexizdat, Moscow.

542 Mancilla, J., Lewis, T., Pinto, D., Rinzel, J., and Connors, B. (2007). Synchronization of electrically
543 coupled pairs of inhibitory interneurons in neocortex. *J. Neurosci.*, 27(8):2058–2073.

544 Mirollo, R. and Strogatz, S. (1990). Synchronization of pulse-coupled biological oscillators. *SIAM J.*
545 *Applied Math.*, 50(6):1645–1662.

546 Morris, C. and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys J*,
547 35:193–213.

548 Neltner, L., Hansel, D., Mato, G., and Meunier, C. (2000). Synchrony in heterogeneous networks of
549 spiking neurons. *Neural Comp.*, 12:1607–1641.

550 Netoff, T., Acker, C., Bettencourt, J., and White, J. (2005a). Beyond two-cell networks: experimental
551 measurement of neuronal responses to multiple synaptic inputs. *J. Comput. Neurosci.*, 18:287–295.

552 Netoff, T., Banks, M., Dorval, A., Acker, C., Haas, J., Kopell, N., and White, J. (2005b). Synchronization
553 of hybrid neuronal networks of the hippocampal formation strongly coupled. *J. Neurophysiol.*,
554 93:1197–1208.

555 Neu, J. (1979). Coupled chemical oscillators. *SIAM J. Appl. Math.*, 37(2):307–315.

556 Oprisan, S., Prinz, A., and Canavier, C. (2004). Phase resetting and phase locking in hybrid circuits of
557 one model and one biological neuron. *Biophys. J.*, 87:2283–2298.

558 Pfeuty, B., Mato, G., Golomb, D., and Hansel, D. (2005). The combined effects of inhibitory and electrical
559 synapses in synchrony. *Neural Computation*, 17:633–670.

560 Risken, H. (1989). *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer-Verlag,
561 NY.

562 Stratonovich, R. (1967). *Topics in the Theory of Random Noise*. Gordon and Breach, NY.

563 Strogatz, S. (2000). From kuramoto to crawford: Exploring the onset of synchronization in populations
564 of coupled oscillators. *Physica D*, 143:1–20.

565 Strogatz, S. and Mirollo, R. (1991). Stability of incoherence in a population of coupled oscillators. *J.*
566 *Stat. Physics*, 63:613–635.

567 Van Kampen, N. (1981). *Stochastic Processes in Physics and Chemistry*. Amsterdam: Elsevier Science.

568 Williams, T. and Bowtell, G. (1997). The calculation of frequency-shift functions for chains of coupled
569 oscillators, with application to a network model of the lamprey locomotor pattern generator. *J.*
570 *Comput. Neurosci.*, 4:47–55.

- 571 Winfree, A. T. (1980). *The Geometry of Biological Time*. Springer, NY.
- 572 Zahid, T. and Skinner, F. (2009). Predicting synchronous and asynchronous network groupings of hip-
573 pocampal interneurons coupled with dendritic gap junctions. *Brain Research*, 1262:115 – 129.