

Mathematical neuroscience: from neurons to circuits to systems

Boris Gutkin, David Pinto ^{*}, Bard Ermentrout

Division of Biology and Medicine, Department of Neuroscience, Brown University, Providence, RI, USA

Abstract

Applications of mathematics and computational techniques to our understanding of neuronal systems are provided. Reduction of membrane models to simplified canonical models demonstrates how neuronal spike-time statistics follow from simple properties of neurons. Averaging over space allows one to derive a simple model for the whisker barrel circuit and use this to explain and suggest several experiments. Spatio-temporal pattern formation methods are applied to explain the patterns seen in the early stages of drug-induced visual hallucinations.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Mathematical models; Non-linear dynamics; Barrel cortex; Spike-time statistics; Hallucinations

1. Introduction

For many years, mathematics and computational methods have played an important role in our understanding of the nervous system. The goal of this chapter is to present some examples of how mathematical techniques can be applied at a variety of levels to increase our understanding of neural systems. We begin with a description of the biophysical principles underlying the generation of action potential in single neurons. The crucial modeling idea is to represent the electrical properties of biological membranes using an equivalent circuit consisting of capacitors and resistors in parallel. We then use phaseplane methods to study a simplified single neuron model and show how the dynamics on the plane can be further reduced to a scalar dynamical system on a circle. Simulations of the reduced model are used to explain the spiking statistics of single neurons driven by noisy stimuli. We then turn our attention to simple neuronal circuits involving networks of excitatory and inhibitory neurons. A mean field approximation reduces such networks to a planar system, and phaseplane analysis enables us to explain experimental results from the somatosensory (touch) system of the rat. Finally, we examine large spatially organized networks. We apply bifurcation theory to

these networks and use the results to explain the patterns seen during visual hallucinations.

2. Equivalent circuit models

The equivalent circuit model has become standard for representing the dynamics of electrical activity observed in single neurons. It is based on the idea that neuronal activity can be completely described by the flow of different currents associated with the neuron's membrane. Currents are divided into those that can be represented by linear circuit elements (passive currents) and those that are voltage and/or time dependent and require more complex dynamics (active currents). Both sets can be understood in terms of the experimental observations shown in Fig. 1a (see [13] for review).

2.1. Passive currents

The first observation is that neurons maintain a constant voltage differential across their membrane called the resting potential ($V_m \approx -65$ mV). This potential arises from the fact that (1) neuronal membranes are semi-permeable, allowing only certain ions to pass from one side to the other (mostly sodium (Na^+), Potassium (K^+), and calcium (Ca^{++})), (2) neurons actively maintain a concentration gradient across the membrane of those same ions, and (3) the ions involved

^{*} Corresponding author.

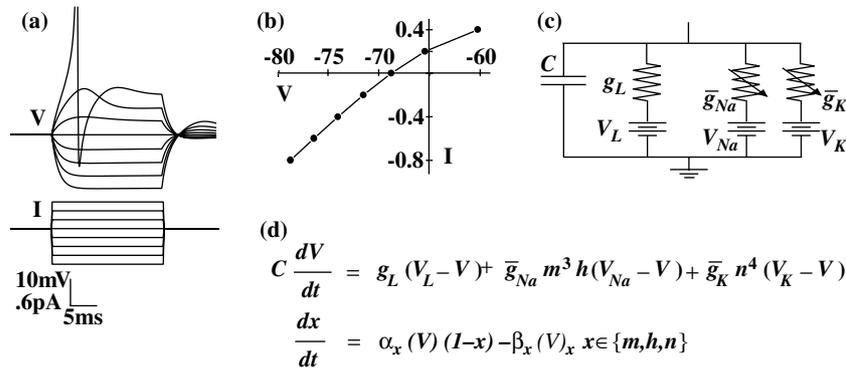


Fig. 1. Equivalent circuit model of neuronal activity. (a) Changes in neuronal membrane potential (V) in response to different levels of injected current (I). (b) $I-V$ relationship of neuronal responses. (c) Circuit diagram representing capacitive and both passive and active resistive currents. (d) Hodgkin Huxley current-balance equations and gating kinetics.

carry an electrical charge. Potassium (K^+), for instance, has a higher concentration inside the cell and carries a positive charge. Diffusive forces drive K^+ out of the cell. The subsequent loss of positive ions leads to a net negative charge inside the membrane. The resulting electrical force attracts positive ions, including those attached to K^+ , back into the cell. The resting potential, also called the equilibrium potential, is the voltage level at which the diffusive and electrical forces arising from all of the permeable ions are in balance. The equivalent circuit element used to represent the resting potential is an electromotive force (EMF) or battery.

The second observation is that direct injection of current into a neuron induces an incremental change in voltage. Within a range of current intensities, the induced change in voltage is linear (Fig. 1b). According to Ohm's law ($V = IR$), this suggests that the membrane is functioning in part as a linear resistor. (Non-linearities in the relationship are discussed below.) Experimentally, the conductance of the membrane is measured from the slope of the line obtained by plotting the change in voltage induced by different intensities of injected current. Even though many different ions contribute to the change in voltage, a single *leak* conductance, g_L , is often used to account for all of the passive ionic currents.

The third observation is that voltage changes due to current injection follow an exponential time course. This suggests that neuronal membranes operate as an RC-circuit, i.e. a linear resistor in parallel with a capacitor. Neuronal membranes are composed of a double layer of lipid molecules peppered with proteins. Proteins have low specific resistivities ($\sim 1 \Omega \text{cm}$) and comprise the protein channels through which charged ions flow across the membrane. Lipids, on the other hand, have high specific resistivities ($\sim 10^9 \Omega \text{cm}$). They provide an effective insulation between the highly electrolytic internal and external cellular solutions, exactly the arrangement required for a parallel plate capacitor. A current injected into the membrane is divided between a capacitive cur-

rent that charges the lipid bilayer, and an ionic current that flows through the protein channels. For an RC-circuit, the time constant of voltage change is $\tau = RC$. Experimentally, therefore, the capacitance of a neuron is obtained by measuring the observed time constant of the membrane potential's exponential rate of change and dividing it by the resistance measured as described above.

2.2. Active currents

Most observed currents exhibit a constant conductance only within a range of voltages. In fact, some non-linearity in the current-to-voltage relationship is expected even for completely passive currents. For instance, when the voltage is such that both diffusive and electrical forces are driving ions in the same direction, the observed conductance can be larger than a simple linear relationship would predict. In many cases, however, the non-linearity, or *rectification*, in a channel's conductance cannot be explained by passive properties alone. Moreover, as seen in Fig. 1a, conductance is often observed to not only depend on voltage but also to change over time.

In order to explain the *anomalous rectification* observed in many neuronal currents, Hodgkin and Huxley [11] proposed the gating model of channel conductance. The idea is that some of the protein channels that allow ions to pass through the membrane also have charged gating particles that open or close the channel in a voltage and time dependent manner, thus changing the conductance. For instance, the conductance of *active* K^+ currents are represented by a maximal conductance g_K times n^4 , where n is the probability that a gating particle is open, and the exponent signifies the number of gating particles on the channel.

The dynamics of the gating particles are modeled using first order kinetics as shown in Fig. 1d. Here $\alpha(V)$ is the voltage-dependent rate at which open channels

close and $\beta(V)$ is the rate at which closed channels open. In the case of active Na^+ currents, experiments suggest that two types of independent gating particles are involved, three activation gates, m , and one inactivation gate, h . The voltage and time dependence of each gate can be determined experimentally by examining conductance changes associated with each individual current. The equivalent circuit element used to represent active neuronal conductance is a variable resistor.

2.3. Current-balance equations

Taken together, the passive and active currents described above comprise the equivalent circuit or parallel conductance model of neural activity shown in Fig. 1c. The equation describing the dynamics of the circuit is a simple instantiation of Kirchoff's first law: the sum of all currents flowing toward a junction is zero. In particular, an applied current divides into a capacitive current that charges the membrane and resistive currents that pass through active and passive channels. Customarily, the equation is written as in Fig. 1d, with the capacitive current (CdV/dt) set equal to the sum of the various ionic currents.

A major success of the Hodgkin and Huxley model was its ability to explain much more than the voltage and time sensitivity of the various neuronal conductances. Once each of the currents had been properly characterized, it was found that the same non-linear gating mechanisms were sufficient to explain the generation of action potentials, i.e. explosive all or none voltage spikes that initiate synaptic signaling between neurons. Using the same set of current-balance equations, Hodgkin and Huxley demonstrated how action potentials result from a stereotyped sequence of conductance changes among the various ionic currents.

Besides Na^+ and K^+ , a number of other voltage and time-dependent conductances have been described. Ca^{++} channels are thought to underlie the generation of action potential bursts that characterize activity in some neurons of the thalamus and cerebral cortex [21]. Spike adaptation, a decrease in the rate of action potential generation over time, depends on K^+ channels that are sensitive not only to voltage, but also to the concentration of Ca^{++} inside the neuron [16]. Na^+ channels that function on a slower time scale and lack an inactivation gate may be responsible for enhancing signals that arrive from other neurons [31]. Each of these currents can be incorporated into the current-balance framework with the addition of appropriate conductances and kinetics to the set of equations.

2.4. Dynamic clamp

A recent and exciting application of the current-balance equations involve using them to modify the func-

tion of real neurons. This is done by using computer models in real time, and the method is known as the “dynamic clamp” or “hybrid model” [29]. The idea is to couple real neurons to simulated ionic currents, or even entire simulated membranes. An electrode is used to measure the membrane voltage of a real neuron. The measured voltage is used to calculate simulated ionic currents. The simulated currents are then injected back into the neuron through the same electrode. The most common applications of the technique are to “insert” new electrical channels into a neuron's membrane or to “subtract” an existing channel and study the effect on the neuron's behavior. The method was first applied to study neurons in the lobster digestive system [29], while more recent studies have examined single neuron behavior in the mammalian brain, [12]. Another developing application is to use the hybrid model method to construct circuits of simulated neurons modeled on programmable analogue microchips [17]. Recent collaborations have explored this technique to study how neuronal circuits control brain rhythms during sleep [18].

3. Phaseplanes and spiking

The Hodgkin-Huxley current-balance model described above is a four-dimensional dynamical system. This makes its rigorous mathematical analysis difficult. Morris and Lecar (see [26]) devised a very simple model neuron based on only three conductances, a fast calcium channel, a slow potassium channel, and a passive leak channel. Using the equivalent circuit formulation described above, the equations are:

$$C \frac{dV}{dt} = I + g_L(V_L - V) + g_{\text{Ca}} m_\infty(V)(V_{\text{Ca}} - V) + g_{\text{K}} w(V_{\text{K}} - V),$$

$$\frac{dw}{dt} = \frac{w_\infty(V) - w}{\tau_w(V)}.$$

The gating kinetics of the calcium channel ($m_\infty(V)$) are fast enough to be considered as acting instantaneously. Thus, the system requires only two dynamic variables and can be rigorously and completely analyzed using phaseplane methods.

Phaseplane analysis is one of the most useful methods that comes to theoretical neuroscience from the qualitative theory of differential equations. The idea is to understand the dynamics of two variables by plotting one as a function of the other. The phaseplane gets its name from the idea that by plotting two dynamical variables, particularly ones that undergo some sort of periodic, or stereotyped behavior, we can study the relative “phase” of one variable against the other. For example, consider a neuron generating an action

potential. The membrane voltage makes an excursion from some resting level, to the peak voltage level, back through an undershoot or *after-hyperpolarization* and then returns to rest. The recovery variable, w , follows a corresponding trajectory in response to the membrane voltage. Plotting the voltage against the recovery variable (see Fig. 2a and b) results in a closed loop, or *trajectory*, on the phaseplane.

The phaseplane is characterized by several landmarks that help to define and visualize the dynamics. For instance, the nullclines of the system are defined as the curves along which one or the other variable remains constant. The curve along which $dV/dt = 0$ defines the V -nullcline, while $dw/dt = 0$ defines the w -nullcline. At the points where the two nullclines intersect neither variable changes; the intersection points define the steady states of the system. One such point is the neuron's resting potential, marked R . This is an attracting fixed point, small transient disturbances of the membrane voltage, by means of an externally applied current or activation of voltage-dependent conductances for instance, will in time relax and the potential will return to rest. Another steady state point is the neuron's threshold potential, marked T . This is a repelling fixed point, so that trajectories that approach threshold either return to rest or continue past resulting in the generation of an action potential.

Using the phaseplane we can visualize the effect of injecting currents into a neuron. Mathematically, injecting a steady positive current corresponds with adding a positive constant to the right hand side of the voltage equation. On the phaseplane, the voltage nullcline is raised. This alters the location of the fixed points. When the voltage nullcline rises sufficiently, the rest and threshold steady states converge, then vanish. This reveals a periodic or *limit cycle* solution on the phaseplane corresponding to the repetitive generation of action potentials or “firing” by the neuron. Like steady states, limit cycle can be either attractive or repelling. In physiological terms, the two cases correspond to sustained firing that is robust to small deflections in the membrane parameters (e.g. noise in V_m or g_L) or repetitive activity that is exquisitely sensitive to perturbations. The robustness or sensitivity can be in either the amplitude, or frequency of the firing events. Note also that neurons can fire at extremely low frequencies, as well as extremely high frequencies, depending on the amplitude of the input.

An important and rather surprising detail is that the trajectory of the limit cycle remains constant in shape no matter how fast the cell fires (or traverses the loop). Experimentally, this corresponds to the invariant shape of generated action potentials. Mathematically it allows us to consider the loop as an invariant structure that

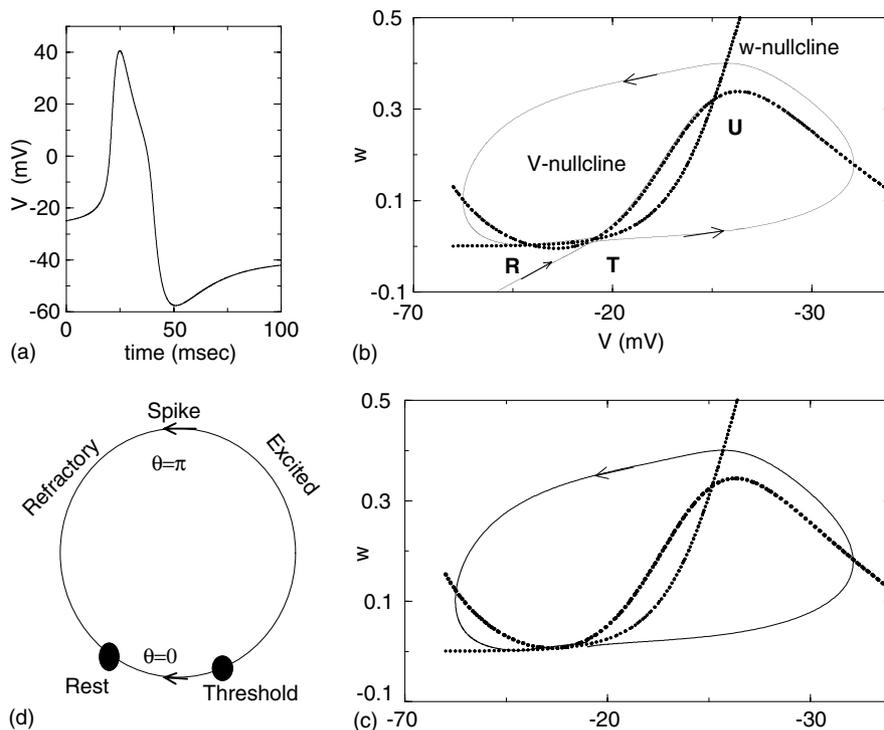


Fig. 2. (a) An action potential, or *spike*, in the membrane voltage plotted in time. (b) Phase diagram of the same spike. The V - and w -nullclines are labeled. Their intersections correspond to fixed points: R is the attracting rest point, T is the threshold, which is a saddle point (attracting from one direction and repelling from others), U is a repelling point. (c) After increasing the injected current, the V -nullcline rises and there are no stable rest points but rather an attracting limit cycle corresponding to a state of repetitive firing. (d) Equivalent state diagram for the theta-neuron. Filled circles mark the critical points achieved by the model—rest and threshold. The model also includes the spiking and refractory behavior of the full model.

does not alter with the systems dynamics. Thus if we can find a way to track only the quantities that change we should be able to simplify the model. Formally this is done by the method of normal forms that comes to us from bifurcation theory. Here we will talk about the heuristic idea (see [8] for details).

As the steady states converge and the limit cycle emerges, the system approaches a point of bifurcation in the dynamic solutions. When we examine the eigenvalues of the linearized model near this bifurcation, we see that there is one real eigenvalue that goes through zero. This implies that the bifurcation is of a *saddle-node on an invariant circle*, which gives us the mathematical basis for treating the trajectory loop as an invariant. To find the mathematical “center” of this loop, we describe the changing voltage around the loop using an angle variable, θ . Let us set the zero angle, $\theta = 0$ at the resting voltage level and let θ increase in a counterclockwise direction around the loop. The action potential is then a revolution of the w/v relative angle or “phase” from 0 through π and back to 2π . Now the trick is to find a mathematical transformation of variables from the original equivalent circuit model such that one of the new variables lies along the invariant limit cycle and the remaining variables are orthogonal to it. For a dynamical system having a saddle-node bifurcation such a transformation can always be found, and yields the equation:

$$\frac{dx}{dt} = qx^2 + pI, \quad (1)$$

where the variable x describes the dynamics, p and I are parameters that includes such things as external inputs, and $q > 0$ is a parameter that depends on the details of the original full model. This equation is a bit cumbersome since it goes to infinity in finite time (in terms of the neural model the action potential happens at infinity). We deal with the singularity by wrapping x onto a circle using a tangent change of variables: $x = \tan \frac{\theta}{2}$, thus arriving mathematically at the phase variable described above. The final equation for the theta-model becomes:

$$\frac{d\theta}{dt} = q(1 - \cos \theta) + (1 + \cos \theta)pI. \quad (2)$$

The important thing to notice is that this model is very simple, and yet reproduces the behavior of the more complex equivalent circuit model. In addition, the theta-neuron is a canonical model for a whole class of conductance based models exhibiting saddle-node bifurcation dynamics. More specifically, any conductance based model of a neuron that exhibits a saddle-node bifurcation can be described by the theta-neuron. Thus if we can understand the behavior of the theta-neuron, we will also understand more about the function of many different types of neurons. For example here are some of the properties of the theta-neuron that

reflect the general behavior of pyramidal neurons found in the neocortex:

- Action potentials are all-or-none events. That is, when stimulated subthreshold activity returns to rest. But when the firing threshold is passed, a complete spike occurs.
- The model, and the neurons, generate continuous trains of action potentials in response to constant current injections.
- Repetitive firing appears with arbitrarily low frequencies.
- The input-current/output-frequency (IF) curve can be readily fit with a square root (instantaneous IF) or linear (steady state IF) function, both of which have been observed in real neurons.

All of the above are rather basic properties that show us that the reduced model does contain dynamical “information” about real neurons. Next we use this reduced model to explore the dynamics underlying more complicated behaviors.

4. Statistics of cortical neural activity in vivo

A long standing debate in neuroscience concerns whether neurons code information about the world in terms of the average frequency of spike generation, or through the precise timing of individual spikes. The second hypothesis has been criticized as unlikely since the firing of neurons in the living brain is very irregular (e.g. see [28]). On the other hand, several experiments have shown that neurons are capable of producing reliably timed action potentials. In one particular experiment, it was demonstrated that, with constant levels of current injection, neurons generate spikes at a set rate but with irregular timing. With repeated noise-like current injections, however, the particular firing times were much more precise with regard to the input [20]. Modeling results demonstrate that the theta-neuron displays similar responses to constant vs. noisy input (Fig. 3) [8]. Here, we examine how the model can be used to capture and explain these rather non-intuitive findings.

The theta-neuron with noisy V_m but constant current input is essentially an intrinsic oscillator perturbed by random voltage noise or, in the language of mathematics, a non-linear renewal process. Thus, the start time of each successive spiking cycle depends on the time of the previous spike. Since the spike times are perturbed by the noise in the “voltage”, uncertainty builds with each successive spike. In fact, it can be shown analytically [8] and numerically (Fig. 3), that this uncertainty, measured as the trial-to-trial variance of spike times, builds as the square root of the spike number multiplied by the variance of the first spike time [6].

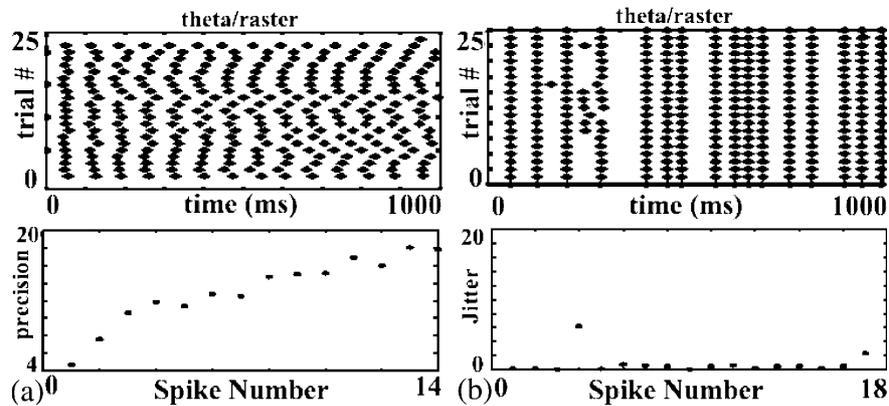


Fig. 3. (a) Response of the theta-neuron model (with noisy V_m) when stimulated by a constant current. The upper graph shows spike timings over repeated trials presenting the same constant stimulus. Early spikes come with the same timing, while later spikes shift randomly. The bottom graph shows plot of the uncertainty in spike times, calculated as the standard deviation of the spike-times across trials. As we can see, the uncertainty builds with each successive spike. (b) Response of the same theta-neuron model when repeatedly stimulated by a noisy current. The spikes remain consistent throughout each trial (upper graph), and the uncertainty remains low from spike to spike (bottom graph).

On the other hand, when the same theta-neuron is driven by jagged, repeated noise-like currents, the spike times depend not on each other but on the dynamics of the stimulus. If the stimulus is such that it rapidly pushes the voltage beyond threshold, then an action potential is generated, and it is generated at the same point of the input across all trials. Thus the uncertainty is low, and does not build with each spike.

This analysis implies that for neurons to encode information with precise spike-timing they should spike in response to sharp shock-like inputs. At the same time, slow inputs would be encoded in a neurons average firing rate. Ongoing analysis of the theta-neuron suggests that these two encoding schemes may not be mutually exclusive, but that the two modes can multiplex in the spike output of a single neuron [7]. Furthermore the canonical model approach has also clarified some possible biophysical mechanisms for synchronous oscillations in circuits of neurons coupled with excitatory synapses and how these related to the underlying bifurcation structure of membrane excitability [35].

5. Activity models

The previous sections showed how to reduce the current-balance equations to a simple scalar model to understand the dynamics of single neurons. In many experimentally relevant cases, it is desirable to simplify the dynamics still further by using so-called activity or firing rate models. In these models the relevant quantity is not the spike time or potential of an individual neuron, but rather the generalized activity level or firing rate in single neurons or neuronal populations.

In contrast to the biophysical representation provided by the current-balance equations, activity models are

based on a more functional description of neuronal activity. They are most often used to study interactions between large neuronal populations or, more generally, in cases where a biophysical model would be too impractical or too complex. In this section, we present a functional derivation of the activity model and then describe an example in which it is used to examine the processing of sensory input in the rodent whisker system.

5.1. Deriving the equations

The central equation of activity models describe the relationship between three distinct measures of neuronal activity: voltage, firing rate, and synaptic drive (described below). The generation of an action potential, or the “firing” of a neuron activates synaptic connections made onto other neurons. Activated synapses evoke voltage changes, or post-synaptic potentials (PSPs), on the receiving neurons. PSPs are often represented as following a characteristic time course described using an ‘alpha’ function, $\alpha(t)$ [25]. PSPs can be either positive or negative depending on whether the synapse is excitatory (EPSPs) or inhibitory (IPSPs). For a series of action potentials, alpha functions sum and the voltage, V , in the post-synaptic neuron is given by

$$V(t) = \sum_i \alpha(t - t_i),$$

where i spans the number of action potentials and t_i is the arrival time for each. In a population of neurons, the distribution of action potentials over time is often described using a continuous function that represents the average firing rate of the population, $F(t)$. Moreover, the average firing rate depends on the average voltage level so that $F(t) = F(V(t))$. Splitting the population into excitatory vs. inhibitory neurons, the equation for the average excitatory population voltage is given by

$$V_e(t) = ee \int_{-\infty}^t \alpha_e(t-t') F_e(V_e(t')) dt' - ie \int_{-\infty}^t \alpha_i(t-t') F_i(V_i(t')) dt'. \quad (3)$$

Here, the two populations have distinct alpha and firing rate functions and ie (for instance) represents the relative strength of synapses from the inhibitory to excitatory population. The average voltage of the inhibitory population is defined similarly.

An alternative formulation of the activity model can be expressed if we define the *synaptic drive* of each population as

$$S_e = \int_{-\infty}^t \alpha_e(t-t') F_e(V_e(t')) dt',$$

$$S_i = \int_{-\infty}^t \alpha_i(t-t') F_i(V_i(t')) dt'$$

so that, again taking the excitatory population as an example,

$$S_e = \int_{-\infty}^t \alpha_e(t-t') F_e(eeS_e - ieS_i) dt'. \quad (4)$$

Moreover, if we assume that the alpha functions are described by simple exponential decay (e.g. $\alpha_e = e^{-t/\tau_e}$), then we may differentiate to obtain,

$$\tau_e \frac{dS_e}{dt} + S_e = F_e(eeS_e - ieS_i).$$

Eq. (3) is called the voltage formulation while (4) is the activity formulation [5]. Intuitively, the term *synaptic drive* derives from the units of eeS_e (for instance), describing the voltage change induced by excitatory synapses. In most models, ee has units of volts · synapse, the connection strength times the number of synapses. Thus S_e has units of 1/synapse, a dimensionless quantity per synapse, i.e. synaptic drive [22]. As described below, the activity formulation in particular allows for straightforward incorporation of biological data into the model system.

5.2. Whiskers and Barrels

Deflecting a single whisker on a rodent's face activates a chain of neurons starting from the periphery, to the brain stem, to the thalamus, and finally to an anatomically defined cluster of neurons in neocortex collectively referred to as a whisker barrel [34]. Each barrel contains two principal cell populations, excitatory neurons and inhibitory neurons. Both populations receive excitatory synaptic inputs from the thalamus. In addition, both populations send reciprocal connections to neurons of the other type and recurrent connections to neurons of the same type (Fig. 4a) [30]. Excitatory barrel neurons also send synaptic connections beyond

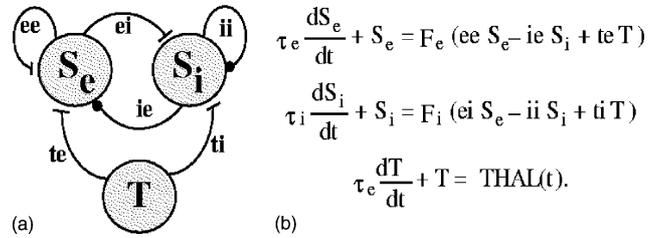


Fig. 4. Activity model of rodent barrel cortex. (a) Schematic diagram of synaptic connections between excitatory (S_e) and inhibitory (S_i) barrel populations and from the thalamus (T). Excitatory synapses are represented by lines, inhibitory synapses by dots. (b) Activity formulation of barrel activity incorporating thalamic input.

the barrel onto neurons in other brain regions [33]. Therefore, using both the activity model and experimental data, the goal is to understand how synaptic interactions within a barrel operate on population input signals arriving from the thalamus and transform them into output signals emanating from the excitatory population.

In the laboratory, experiments measure the number and timing of action potentials generated by thalamic and barrel neurons in response to a whisker deflection. More generally, data from many neurons is pooled into an average firing rate for each population. The usefulness of the activity model derives from the fact that it represents neuronal activity in a form compatible with this experimental data. Since data from both thalamic and barrel neuron populations is measured in terms of average firing rate, thalamic data can be used directly as input to the model ($THAL(t)$), and excitatory barrel population data can be used directly to evaluate the excitatory output of the model (F_e , Fig. 4b).

In practice, the precise shape of the firing rate functions (F_e and F_i) are determined experimentally by examining the firing rates of thalamic and barrel neurons in response to a well-chosen battery of whisker deflections [15,22]. Once simulated responses match the experimental data, the model can then be used both to predict responses to novel stimuli and to understand the mechanisms by which those responses are generated.

One prediction of the barrel activity model is that the size of the excitatory response is sensitive to the distribution of thalamic population input. To understand why this is so, we examine the phaseplane obtained from the model system in response to two thalamic population inputs [24]. On the phaseplane, activity levels of the excitatory and inhibitory populations are plotted on the x - and y -axis, respectively. Over the course of an input signal, changing activity levels are tracked by the response trajectory over successive planes at time points indicated by the small roman numerals (Fig. 5a and b). The thin curved (straight) line is the excitatory (inhibitory) nullcline, i.e. points at which $dS_e/dt = 0$ ($dS_i/dt = 0$).

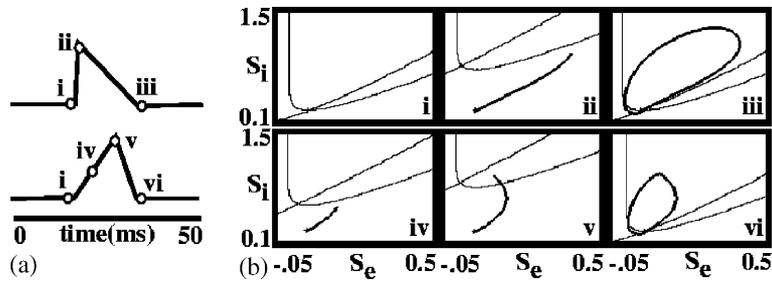


Fig. 5. Phaseplane analysis of simulated barrel response to thalamic input distributions. (a) Two simulated thalamic input distributions differing in time course but with the same total activity. (b) Phaseplane representation of barrel population responses. Successive planes show ongoing response at the time points indicated by small roman numerals.

For the rapidly rising thalamic input, the nullclines rise quickly as does the network's response (Fig. 5b, i to ii). For the slowly rising input, the nullclines rise more slowly. The response trajectory remains close to the nullcline where changes in activity are relatively smaller (Fig. 5b, i to iv to v). Thus, despite the fact that both inputs consist of the same total amount of thalamic activity (Fig. 5a), the rapidly rising input distribution generates a much larger response from the network than the slowly rising input distribution (Fig. 5b iii vs. vi). This prediction of the model has been verified experimentally by examining the response of real barrel neurons to different distributions of thalamic input [23].

6. From cave paintings to the structure of the cortex

For years, people have been fascinated by Paleolithic rock paintings such as are found in the famous Lascaux caves in France and on the sandstone walls in the American Southwest. These pictures often depict animals and human-like images. However, there are also more abstract designs such as spirals, sunbursts and mandalas which occur across all cultures in this art. One of the questions that anthropologists have asked is what are the possible meanings of these abstract symbols. A number of researchers have suggested that they are paintings drawn by shamans (“medicine men”) while in trance states and that the drawings represent visual imagery that is a consequence of an altered state of awareness [9,10,19]. In particular, anthropologists draw an analogy between the iconic forms of cave art and the patterns reported in the early stages of hallucination [14], as well as the patterns perceived during certain visual stimuli such as eyeball pressure and flickering light [32]. The latter patterns, called phosphenes, and their drug-induced analogues were classified by Kluver into four different types he called “form constants”:

- tunnels and funnels,
- spirals,
- lattices such as honeycombs and checkerboards,
- cobwebs.

Representatives of these are shown in Fig. 6. Thus, anthropologists propose that the abstract geometric designs found in cave and rock paintings are depictions of internal imagery imagined by shamans who are in drug or trance-induced mental states. In fact, as we will see, many of the more common designs can be predicted by examining the structure of neural activity patterns in the visual cortex.

We have seen how to go from detailed biophysics to reduced models that can be used to study local sensory circuits. What happens when many such circuits are coupled together in a spatially structured manner? The striate visual cortex is notable due to its almost crystalline organization of cells that respond to orientation, color, and position in the visual field. That is, the visual cortex is organized into columns (not unlike the barrels) and within each column are neurons that respond to (i) spatial positions on the retina; (ii) inputs from the left

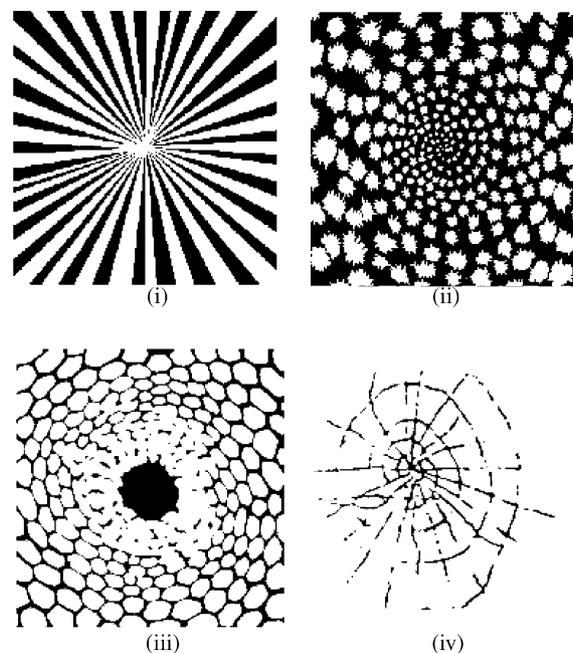


Fig. 6. Form constants (i) funnel, (ii) spiral, (iii) lattice and (iv) cobweb.

and right eyes; (iii) oriented lines; and many other visual modalities. The connections between cells are not random. For example, cells in one column responding to vertical lines are connected more strongly to cells of the same orientation in neighboring columns. The consequences of this highly structured connectivity have been explored in a number of papers. Ermentrout and Cowan [3] showed that spontaneous symmetry breaking in spatially organized cortical networks could account for coarse spatial activity patterns perceived during hallucinations. More recently, Bressloff et al. [1] have incorporated coupling between orientation-sensitive cells to explain other types of more complex hallucinatory patterns. In this final section, we will use these ideas to understand a possible context for the ancient cave drawings described above.

Let us first consider a very general neural network architecture:

$$\tau_i \frac{du_i}{dt} = -u_i + F_i \left(\sum_j W_{ij} u_j \right) \quad (5)$$

which represents the activities of many coupled neurons. W_{ij} represents the strength of connections between neurons so that Eq. (5) is a generalization of Eq. (4). This network can be separated into as many layers as desired; we include no intrinsic organization. Suppose that we subtract off the background activity of the network so that $u_i(t)$ represents the deviation from background and $F_i(0) = 0$. For simplicity, we set $\tau_i = 1$. We can ask whether or not this background state is stable. To determine this, we need to linearize about the rest state, $u = 0$. The stability is determined by the eigenvalues of the matrix:

$$M = -I + Q,$$

where $Q_{ij} = F_i'(0)W_{ij}$. If Q is small enough (that is, there is only weak coupling between units or $F_i'(0)$ is small) then the eigenvalues of M have negative real parts and any perturbations decay to the background state. However, suppose some external influence (e.g. flickering light, hallucinogens, etc) “excites” the network by, for example, making the slopes of F_i large. Then, it is possible that some eigenvalues of M can become positive and this leads to instabilities. The “shape” of the activities that arise from this instability is, to lowest order, a combination of all the eigenvectors of Q whose eigenvalues have simultaneously become positive. Generically for a random symmetric network, we expect either a transcritical or pitchfork bifurcation leading to a single excited mode reflecting the organization of the random network. If the effect of hallucinogens or other external stimuli is to produce an overall increase in excitability of the network, then is it possible that the background state can become unstable leading to a spontaneous pattern of activity reflecting the associative

organization of the recurrent network. If there are many nearly identical eigenvalues, then the resulting patterns can be complex mixtures of the eigenvectors and the dynamics governing the patterns.

Consider, for example, an associative weight matrix,

$$W_{ij} = \sum_{k=1}^m \xi_i^k \xi_j^k,$$

where $\xi^1, \xi^2, \dots, \xi^m$ are m vectors to be memorized. If each of these has unit magnitude and are nearly orthogonal, then the weight matrix, W has an m -fold eigenvalue $\lambda = 1$ with these m vectors as eigenvectors. Standard bifurcation methods can be applied to derive a set of amplitude equations for the m eigenvectors. These have the form:

$$A_i' = vA_i + p_2(A_1, \dots, A_m) + p_3(A_1, \dots, A_m), \quad (6)$$

where v is a parameter which governs the stability of the background state and p_2, p_3 are quadratic and cubic terms respectively. The activity pattern that arises past the bifurcation point has the form:

$$\mathbf{u}(t) \approx A_1(t)\xi^1 + \dots + A_m(t)\xi^m.$$

Thus, for an associatively connected network, the resulting patterns of activity which occur when the background state becomes unstable are dynamically varying mixtures of various “learned memories.”

What does this have to do with form constants? The analysis sketched above assumes that the weight matrix, W_{ij} , has no structure or that it has structure arising from learned memories. In associative areas of the brain, this is likely to be true. However, as we mentioned above, visual cortex has a great deal of structure. The connections are far from random and have numerous intrinsic symmetries. For example, if the weight matrix is that associated with a two-dimensional spatial grid with coupling that depends only on distance, then the eigenvectors associated with W have doubly periodic structure (e.g. stripes or hexagons or spots). The resulting patterns that arise in the two-dimensional network consist of combinations of striped and spotted patterns.

How would a striped pattern of activity in cortex be perceived? The answer is the key to understanding Kluver’s form constants. Consider a pattern of light imposed on the retina. Since there is a topographic map from the retina to the visual cortex, this results in a pattern on the cortex. Conversely, each spatial pattern of activation on the cortex is associated with a virtual pattern on the retina. In particular, Schwartz [27] showed that each point (r, θ) in polar retinal coordinates centered at the fovea is associated with a point (x, y) in the cortex. The transformation is:

$$x = \frac{\alpha}{\epsilon} \ln \left(1 + \frac{\epsilon}{w_0} r \right), \quad y = \frac{\beta r \theta}{w_0 + \epsilon r},$$

and the inverse of this map,

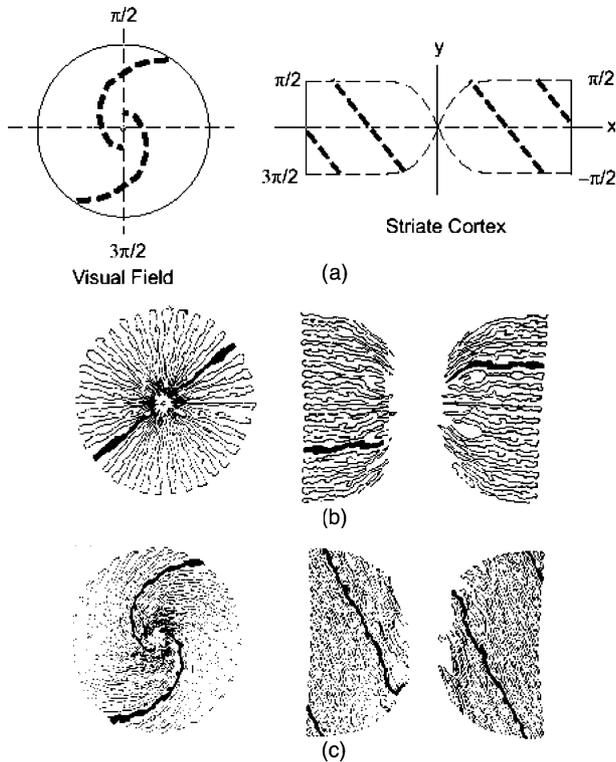


Fig. 7. (a) The retinal-cortical transformation and its action the (b) funnel and (c) spiral form constants (after Bressloff, et al. [1]).

$$r = \frac{w_0}{\epsilon} (e^{\frac{x}{\epsilon}} - 1), \quad \theta = \frac{\epsilon y}{\beta(1 - e^{-\frac{x}{\epsilon}})},$$

allows us to transform a cortical pattern to its retinal analogue. Near the fovea, the map is the scaled identity. Far from the fovea, the map is the complex log and its inverse is the complex exponential. Fig. 7 shows the action of this map along with its effect on a funnel and spiral form constant.

This map coupled with the idea that visual hallucinations arise as a consequence of stability loss, first in striate cortex and later in higher association areas, allows us to explain the form constants of Kluver and indirectly explain the abstract designs of shamanistic paintings. Ermentrout and Cowan [3] exploited the symmetry of connections based on spatial position and used group theory to show that the spatial patterns which bifurcate from the background state are nothing more than stripes and spots. The normal form Eqs. (6) have a vastly simplified structure. For example, the following form arises:

$$A_1' = \nu A_1 - A_1(aA_1^2 + bA_2^2),$$

$$A_2' = \nu A_2 - A_2(aA_2^2 + bA_1^2),$$

where a, b are model-dependent constants and A_1 is the amplitude corresponding to horizontal stripes of activity while A_2 is the amplitude corresponding to vertical stripes. If $A_1 \neq 0, A_2 = 0$ ($A_1 = 0, A_2 \neq 0$) then horizontal (vertical) stripes occur while if $A_1 = A_2$ then spots occur. The stability of these bifurcating patterns is dependent only on the constants a, b . Fig. 8 shows examples of horizontal stripes and spots of activity in cortical coordinates (top left) as well as the perceived image in retinal coordinates (top right).

If the structure of the connectivity matrix incorporates the orientation specificity of cortical neurons, then the associated eigenvectors are more complicated but are interpreted by the visual system as periodically arrayed line segments. Bressloff et al. [1] have analyzed this case and derived the normal form equations. The patterns corresponding to the fixed points of the normal form equations are more complicated than just stripes and spots. However, they complete the enumeration of the form constants. Two examples are shown in the lower panels of Fig. 8 and correspond to lattice and cobweb form constants.

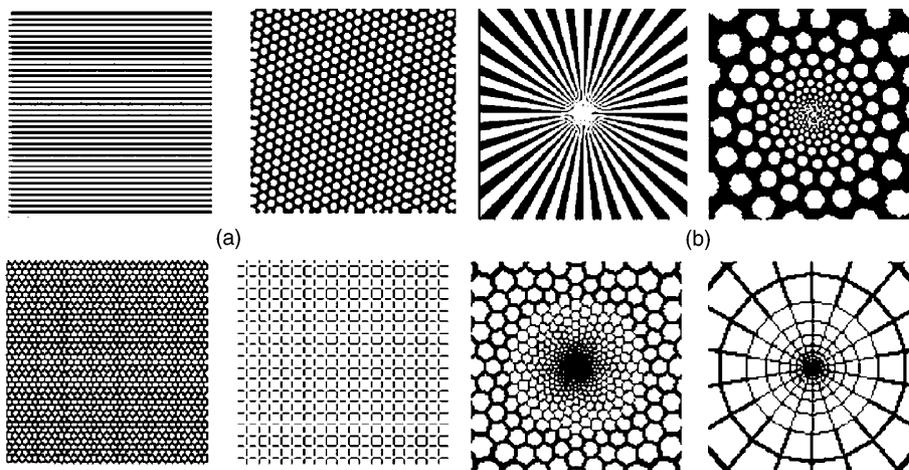


Fig. 8. Bifurcating patterns in cortical (left) and retinal (right) coordinates. Upper panels correspond to patterns arising from distance-dependent connections and lower panels to distance- and orientation-dependent connections (after Bressloff, et al. [1]).

References

- [1] P.C. Bressloff, J.D. Cowan, M. Golubitsky, P.J. Thomas, M. Wiener, Geometric visual hallucinations, Euclidean symmetry, and the functional architecture of striate cortex, *Philos. T Roy. Soc. B* 356 (2001) 299–330.
- [3] G.B. Ermentrout, J.D. Cowan, A mathematical theory of visual hallucinations, *Kybernetik* 34 (1979) 137–150.
- [5] G.B. Ermentrout, Neural Nets as spatio-temporal pattern forming systems, *Rep. Prog. Phys.* 61 (1998) 353–430.
- [6] B.S. Gutkin, G.B. Ermentrout, M. Rudolph, Spike generating dynamics and the conditions for spike time precision in cortical neurons, *J. Comput. Neurosci.* 15 (2003) 91–103.
- [7] B.S. Gutkin, G.B. Ermentrout, How neurons multiplex coding of fast and slow stimuli, in preparation.
- [8] B.S. Gutkin, G.B. Ermentrout, Dynamics of membrane excitability determine interspike interval variability: a link between spike generation mechanisms and cortical spike train statistics, *Neural Comput.* 10 (5) (1998) 1047–1065.
- [9] K. Hedges, Phosphores in the context of Native American rock art, in: F.G. Bock (Ed.), *American Indian Rock Art*, vols. VII, VIII, 1981.
- [10] T. Hudson, G. Lee, Function and symbolism in Chumash rock art, *J. New World Archaeol.* 6 (1984) 26–47.
- [11] A.L. Hodgkin, A.F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Phys. (London)* 117 (1952) 500–544.
- [12] D. Jaeger, J.M. Bower, Synaptic control of spiking in cerebellar Purkinje cells: dynamic current clamp based on model conductances, *J. Neurosci.* 19 (14) (1999) 6090–6101.
- [13] D. Johnston, S. Wu, *Foundations of Cellular Neurophysiology*, MIT Press, 1997.
- [14] H. Kluver, *Mescal and the Mechanisms of Hallucination*, University of Chicago Press, Chicago, 1966.
- [15] H.T. Kyriazi, D.J. Simons, Thalamocortical response transformations in simulated whisker barrels, *J. Neurosci.* 13 (1993) 1601–1615.
- [16] B. Lancaster, P.R. Adams, Calcium-dependent current generating the after-hyperpolarization of hippocampal neurons, *J. Neurophys.* 55 (6) (1986) 1268–1282.
- [17] G. Le Masson, S. Le Masson, M. Moulins, From conductances to neural network properties: analysis of simple circuits using the hybrid network method, *Prog. Biophys. Mol. Biol.* 64 (2–3) (1995) 201–220.
- [18] G. Le Masson, S. Renaud-Le Masson, D. Debay, T. Bal, Feedback inhibition controls spike transfer in hybrid thalamic circuits, *Nature* 417 (2002) 854–858.
- [19] J.D. Lewis-Williams, T.A. Dowson, The signs of all times entoptic phenomena in upper Paleolithic art, *Current Anthropol.* 29 (1988) 201–245.
- [20] Z.F. Mainen, T.S. Sejnowski, Reliability of spike timing in neocortical neurons, *Science* 268 (1995) 1503–1506.
- [21] D.A. McCormick, J.R. Huguenard, A model of the electrophysiological properties of thalamocortical relay neurons, *J. Neurophys.* 68 (1992) 1384–1400.
- [22] D.J. Pinto, J.C. Brumberg, D.J. Simons, G.B. Ermentrout, A quantitative population model of whisker barrels: re-examining the Wilson–Cowan equations, *J. Comput. Neurosci.* 3 (1996) 247–264.
- [23] D.J. Pinto, J.C. Brumberg, D.J. Simons, Circuit dynamics and coding strategies in rodent somatosensory cortex, *J. Neurophysiol.* 83 (2000) 1158–1166.
- [24] D.J. Pinto, J.D. Hartings, J.C. Brumberg, D.J. Simons, Cortical damping: analysis of thalamocortical response transformations in rodent barrel cortex, *Cerebral Cortex* 13 (2003) 33–44.
- [25] W. Rall, A. Agmon-Snir, Cable theory for dendritic neurons, in: C. Koch, I. Segev (Eds.), *Methods in Neuronal Modeling*, MIT Press, Cambridge, MA, 1998, pp. 27–93.
- [26] J. Rinzel, G.B. Ermentrout, Analysis of neural excitability and oscillations, in: C. Koch, I. Segev (Eds.), *Methods in Neuronal Modeling*, MIT Press, Cambridge, MA, 1998.
- [27] E. Schwartz, Spatial mapping in the primary sensory projection: analytic structure and relevance to projection, *Biol. Cybernet.* 25 (1977) 181–194.
- [28] W.M. Shadlen, W.T. Newsome, The variable discharge of cortical neurons: implications for connectivity, computation and information coding, *J. Neurosci.* 18 (1998) 3870–3896.
- [29] A.A. Sharp, M.B. O’Neil, L.F. Abbott, E. Marder, Dynamic clamp: computer-generated conductances in real neurons, *J. Neurophysiol.* 69 (3) (1993) 992–995.
- [30] D.J. Simons, Response properties of vibrissa units in rat SI somatosensory neocortex, *J. Neurophysiol.* 41 (1978) 798–820.
- [31] C.E. Stafstrom, P.C. Schwindt, W.E. Crill, Negative slope conductance due to a persistent subthreshold sodium current in cat neocortical neurons in vitro, *Brain Res.* 236 (1982) 221–226.
- [32] C. Tyler, Some new entoptic phenomena, *Vis. Res.* 181 (1978) 1633–1639.
- [33] E.L. White, *Cortical Circuits: Synaptic Organization of the Cerebral Cortex*, Birkhauser, Boston, MA, 1989.
- [34] T.A. Woolsey, H. van der Loos, The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex, *Brain Res.* 17 (1970) 205–242.
- [35] G.B. Ermentrout, M. Pascal, B.S. Gurkin, The effects of spike frequency adaptation and negative feedback on the synchronization of neural oscillators, *Neural Comput.* 13 (2001) 1285–1310.