On the Dynamics of Small Continuous-Time Recurrent Neural Networks

Randall D. Beer*
Case Western Reserve University

Dynamical neural networks are being increasingly employed in a variety of contexts, including as simple model nervous systems for autonomous agents. For this reason, there is a growing need for a comprehensive understanding of their dynamical properties. Using a combination of elementary analysis and numerical studies, this article begins a systematic examination of the dynamics of continuous-time recurrent neural networks. Specifically, a fairly complete description of the possible dynamical behavior and bifurcations of one- and two-neuron circuits is given, along with a few specific results for larger networks. This analysis provides both qualitative insight and, in many cases, quantitative formulas for predicting the dynamical behavior of particular circuits and how that behavior changes as network parameters are varied. These results demonstrate that even small circuits are capable of a rich variety of dynamical behavior (including chaotic dynamics). An approach to understanding the dynamics of circuits with time-varying inputs is also presented. Finally, based on this analysis, several strategies for focusing evolutionary searches into fruitful regions of network parameter space are suggested.

Key Words: dynamical neural networks; computational neuroethology; evolutionary search; nonlinear dynamics

Introduction

Continuous-time recurrent neural networks (CTRNNs) are networks of model neurons of the following general form:

$$\dot{y}_i = f_i(y_1, \ldots, y_N) \equiv \frac{1}{\tau_i}\left(-y_i + \sum_{j=1}^{N} w_{ji}\sigma(y_j + \theta_j) + I_i\right) \quad i = 1, 2, \ldots, N \quad (1)$$

where y is the state of each neuron, τ is its time constant ($\tau > 0$), w_{ji} is the strength of the connection from the j^{th} to the i^{th} neuron, θ is a bias term, $\sigma(x) =$

* Department of Computer Engineering and Science and Department of Biology, Case Western Reserve University, Cleveland, Ohio; E-mail: beer@alpha.ces.cwru.edu

$1/(1 + e^{-x})$ is the standard logistic activation function, and I represents a constant external input. It will also sometimes be convenient to express Equation 1 in vector notation: $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}; \mathbf{W}, \boldsymbol{\theta}, \mathbf{I}, \boldsymbol{\tau})$, where $\dot{\mathbf{y}}$, $\mathbf{f}$, $\mathbf{y}$, $\boldsymbol{\theta}$, $\mathbf{I}$, $\boldsymbol{\tau}$ are length N vectors of the corresponding scalar quantities and $\mathbf{W}$ is the $N \times N$ matrix of connection weights. Although variations of this basic model neuron were studied much earlier (e.g., Grossberg, 1969), a restricted form of Equation 1 was popularized by Hopfield (1984) in his work on associative memories.

Within the autonomous agents community, there is a growing interest in the use of dynamical neural networks for controlling the behavior of agents (Beer, 1990) and in the evolution of such networks (Collins & Jefferson, 1991; Werner & Dyer, 1991; Beer & Gallagher, 1992; de Garis, 1992; Spiessens & Torreele, 1992; Cliff, Harvey, & Husbands, 1993; Yamauchi & Beer, 1994; Floreano & Mondada, 1994; Miller & Cliff, 1994). In contrast to purely static feedforward networks, which support only reactive behaviors, dynamical neural networks allow an agent to initiate action independent of its immediate situation and to organize its behavior in anticipation of future events (Beer, 1995a).

CTRNNs are an obvious choice for this work because (1) they are arguably the simplest nonlinear, continuous dynamical neural network model; (2) despite their simplicity, they are universal dynamics approximators in the sense that, for any finite interval of time, CTRNNs can approximate the trajectories of any smooth dynamical system on a compact subset of $\mathbb{R}^n$ (Funahashi and Nakamura, 1993); and (3) they have a plausible neurobiological interpretation, where the state y often is associated with a nerve cell's mean membrane potential and the output $\sigma(y)$ is associated with its short-term average firing frequency. CTRNNs are also being applied to a wide variety of other problems, including associative memories (Hopfield, 1984), optimization (Hopfield & Tank, 1985), biological modeling (Lockery, Fang, & Sejnowski, 1990), and many others (Erlanson & Abu-Mostafa, 1991; Goudreau & Giles, 1992; Simard & Cun, 1992). In addition, a variety of learning algorithms for CTRNNs have been developed (for review, see Pearlmutter, 1990).

With the growing use of CTRNNs, there is a growing need for a comprehensive understanding of their dynamical properties. What kinds of dynamical behavior can a given network exhibit? How does a network's dynamical behavior depend on its parameters? Even partial answers to these questions would contribute significantly to our understanding of existing networks and could help to guide the synthesis of new networks for solving particular problems. This is especially true for evolved networks, for which it is common to achieve a perfectly competent network whose operation appears to be completely incomprehensible. Although a few attempts have been made to understand the operation of evolved dynamical neural networks (Gallagher and Beer, 1993; Cliff, Husbands, & Harvey, 1993; Yamauchi & Beer, 1994; Beer,

1995a,b; Husbands, Harvey, & Cliff, in press), the results thus far obtained are highly specific to the particular circuits considered. In addition, progress on these questions may have implications for understanding the dynamics of more biologically realistic networks.

The goal of this article is to illustrate how the mathematical tools of dynamical systems theory can be used to gain significant insight into the dynamics of CTRNNs. While a great deal of work has been done on these and closely related models (e.g., Wilson & Cowan, 1972; Cohen & Grossberg, 1983; Hopfield, 1984; Hirsch, 1989), no comprehensive description of the dynamics of even small unconstrained CTRNNs is currently available. Because of the interest in associative memory applications, a great deal of previous work has focused on the question of how to constrain CTRNNs so that they exhibit only equilibrium points (for review, see Grossberg, 1988). However, oscillatory (Atiya & Baldi, 1989) and chaotic (Sompolinsky & Crisanti, 1988; Das, Schieve, & Zeng, 1991) dynamics are also possible and are potentially very important in applications (e.g., rhythmical or evasive behavior in an autonomous agent). Although the particular results obtained here are mostly specific to Equation 1, the general techniques can be applied to any CTRNN, and analogous techniques for iterated maps can be applied to discrete-time recurrent neural networks (Blum & Wang, 1992; Tino, Horne, & Giles, 1995).

This article is organized as follows. It begins with a fairly complete theory of one- and two-neuron circuits in sections 2 and 3, respectively. Section 4 presents a few specific results concerning the dynamics of CTRNNs of arbitrary size, using three-neuron circuits to illustrate particular results. An approach to understanding the dynamics of CTRNNs with time-varying inputs is then described in section 5. Finally, in section 6, I discuss some possible implications of these results for the analysis and synthesis of CTRNNs produced by evolutionary algorithms. For those unfamiliar with the formalism of dynamical systems theory, I have tried to augment the mathematics with text and pictures throughout. Good tutorial introductions to dynamical systems theory can be found in the books by Abraham and Shaw (1992), Strogatz (1994), and Hale and Koçak (1991).

A Single Neuron with Self-Connection

The obvious place to begin the study is with a single neuron with a self-connection, as it forms the basic building block for any larger network. For convenience, I will drop the subscripts in this case and write Equation 1 as:

$$\dot{y} = f(y; w, \theta, I, \tau) \equiv \frac{1}{\tau}(-y + w\sigma(y + \theta) + I) \qquad \textbf{(2)}$$

In the absence of a self-connection, Equation 2 reduces to a linear differential equation whose solution is obviously just an exponential decay to I with time constant τ. However, as we shall see, the presence of a simple nonlinear self-connection endows an individual neuron with significantly more complicated behavior. Neurobiologically, this self-connection can be interpreted as either a very simple type of active conductance or as a literal self-collateral. Furthermore, under certain conditions, a self-connection can be interpreted as representing a larger circuit whose dynamics have been adiabatically eliminated (Schieve, Bulsara, & Davis, 1991).

2.1 Qualitative theory

Ideally, we would like an explicit expression for the general solution $\phi(t, y_0, w, \theta, I, \tau)$ describing how the trajectories of Equation 2 evolve in time from a given initial state y_0. Unfortunately, no elementary expression for the solution of Equation 2 exists. Therefore, we must settle for a qualitative description of its dynamics. Specifically, I will describe the limit sets of Equation 2, including their stability and their dependence on the parameters, as well as the bifurcations that can occur as the parameters are varied.

For a scalar dynamical system such as that presented in Equation 2, the only possible limit sets are equilibrium points, which may be either stable or unstable. Equilibrium points represent constant solutions of Equation 2 and correspond to zeroes of f. The zeroes of f are clearly independent of τ, and so we will henceforth assume that $\tau = 1$. The stability of an equilibrium point of Equation 2 is given by the sign of $f'(y, w, \theta) = w\sigma'(y + \theta) - 1$, where f' denotes $\partial f/\partial y$ and $\sigma'(x)$ is often written as $\sigma(x)(1 - \sigma(x))$. A given equilibrium point $\bar{y}$ of Equation 2 is stable if $f'(\bar{y}, w, \theta) < 0$ and unstable if $f'(\bar{y}, w, \theta) > 0$ (Hale & Koçak, 1991). Because $0 \leq \sigma'(x) \leq 1/4, f' < 0$ for $w < 4$ and so only stable equilibria are possible in this case. On the other hand, when $w > 4$, the sign of f' depends on y, w and θ. Thus, a qualitative change in behavior can occur as w passes through 4, leading us to expect at least two qualitatively different phase portraits in this system.

I will use the notation $\bar{y}(I, w, \theta)$ to represent the surface of equilibrium points of Equation 2 as a function of the parameters. By solving $f(y, w, \theta, I) = 0$ for I, it is easy to see that this surface is implicitly defined by the expression $I = y - w\sigma(y + \theta)$. Unfortunately, this expression cannot be algebraically solved for y. Assuming for the moment that $\theta = 0$, Figure 1 shows the numerically computed equilibria of Equation 2 as a function of I at $w = -20$ (Fig. 1A) and at $w = 20$ (Fig. 1B). Note that, while Equation 2 exhibits a single stable equilibrium point for all I when $w = -20$, it exhibits three equilibria for a range of I values when $w = 20$. In the latter case, the outer two equilibria are stable and the inner one is unstable. In this case, the unstable equilibrium point separates the basins of attraction of the

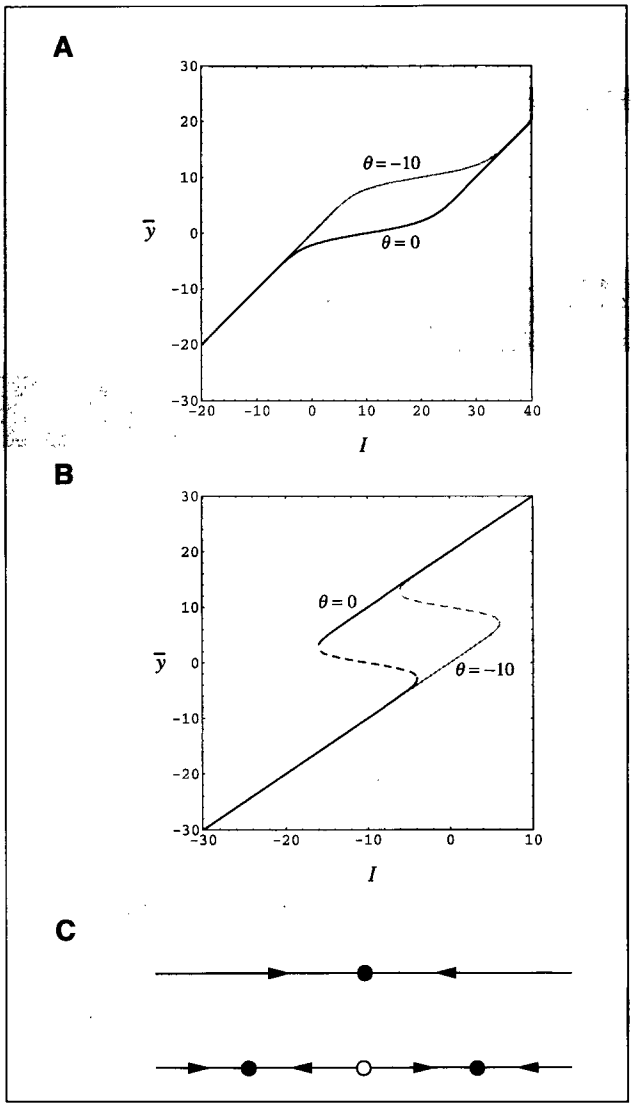


Figure 1
The equilibria of Equation 2 as a function of I when (A) $w = -20$ and (B) $w = 20$. Solid lines represent stable equilibria, and dashed lines represent unstable equilibria. Dark curves are for $\theta = 0$, and gray curves are for $\theta = -10$. The two possible phase portraits of Equation 2 are shown in C. Here filled circles represent stable equilibrium points, and open circles represent unstable equilibrium points. The bottom phase portrait occurs when the input is within the fold shown in B, whereas the top phase portrait occurs everywhere else.

two stable equilibrium points. This corresponds to a neuron whose self-excitation is sufficiently large that it is capable of remaining active through positive feedback alone even in the absence of external input. These two distinct phase portraits are shown in Fig. 1C.

Still assuming that $\theta = 0$, the three-dimensional surface of equilibria $\bar{y}(I, w, 0)$ is

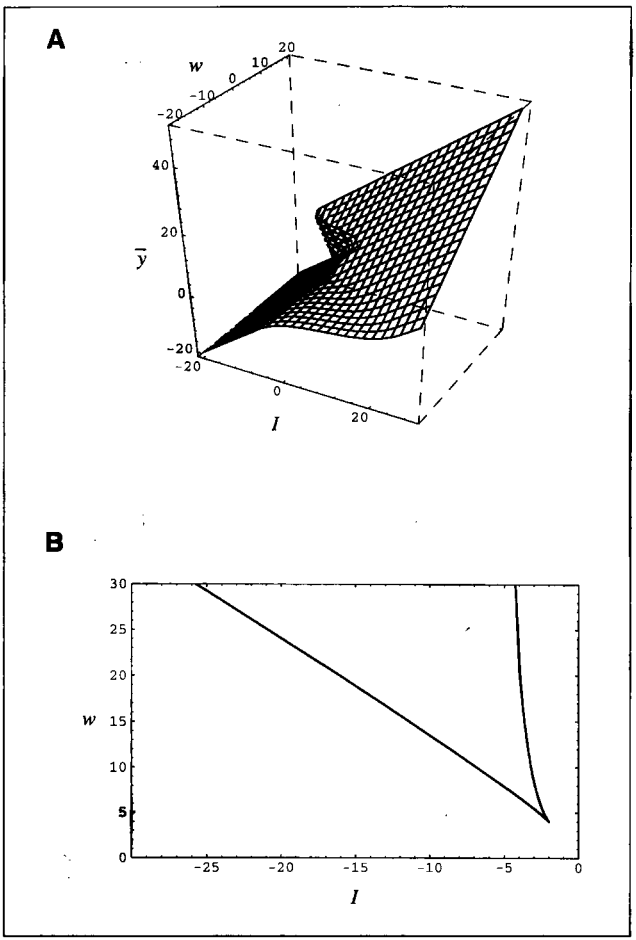


Figure 2
The (A) equilibrium surface and (B) bifurcation set of Equation 2 when $\theta = 0$. The curves shown in Figure 1A and B represent slices through this surface corresponding to particular values of w and θ. Geometrically, B is the projection of the edge of the fold in A. Analytically, it is the curve along which $f(y, w, \theta, I) = f'(y, w, \theta) = 0$. For parameter values outside the cusp, the system exhibits a single stable equilibrium point (top phase portrait in Fig. 1C). For parameter values within the cusp, the system exhibits a pair of stable equilibrium points separated by an unstable equilibrium point (bottom phase portrait in Fig. 1C).

shown in Figure 2A. We can see that, as w increases through 4, this surface develops a fold, the middle branch of which is unstable. The width of this fold increases with increasing w. Whenever the values of I and w cross into or out of this fold, Equation 2 undergoes a bifurcation—that is, its dynamical behavior switches between the two qualitatively different phase portraits shown in Figure 1C. Thus, the edges of the fold form the bifurcation set of Equation 2. This set is a cusp (Cowan & Ermentrout, 1978) and is shown in Figure 2B. The point of the cusp occurs at $(I, w) = (-2, 4)$. When I and w cross the cusp at the cusp point, Equation 2 undergoes a pitchfork bifurcation, while varying these parameters across any other point on the cusp results in a saddle-node bifurcation (Hale & Koçak, 1991).

It remains to understand the effect of θ on the equilibrium points of Equation 2. As can be seen for $w = -20$ and $w = 20$ in Figure 1, a nonzero θ merely shifts the location of the fold (or "kink" if $w < 4$) both horizontally and vertically on the equilibrium surface by $-\theta$. Thus, with respect to the bifurcation set of Equation 2, the effects of varying θ or I are interchangeable, implying that the bifurcation set of Equation 2 in the (θ, w) plane at $I = a$ is identical to the bifurcation set in the (I, w) plane at $\theta = a$. Of course, varying θ or I does affect the state space locations of the equilibrium points in different ways.

2.2 Quantitative theory

2.2.1 Equilibrium surface There are obviously several limiting cases where approximate expressions for the equilibrium surface $\bar{y}(I, w, \theta)$ can easily be derived. For example, when either $w \approx 0$ or $I \ll -(w/2 + \theta)$,

$$\bar{y}(I, w, \theta) \approx I \tag{3}$$

On the other hand, when $I \gg -(w/2 + \theta)$, then

$$\bar{y}(I, w, \theta) \approx w + I \tag{4}$$

We can also derive an approximate expression for $\bar{y}(I, w, \theta)$ when $\gamma + \theta \approx 0$, which occurs along the line $\hat{I} = -(w/2 + \theta)$. When $\gamma + \theta \approx 0$, $\sigma(\gamma + \theta)$ can be replaced by its Taylor expansion: $\sigma(\gamma + \theta) \approx (\gamma + \theta)/4 + 1/2$. Substituting this expression into $f(\gamma, w, \theta, I) = 0$ and solving for γ, we obtain the following expression for $\bar{y}(I, w, \theta)$ in the neighborhood of the line along which $\gamma = -\theta$:

$$\bar{y}(I, w, \theta) \approx \frac{4}{4 - w}(I - \hat{I}) - \theta \tag{5}$$

Note that the slope of $\bar{y}(I, w, \theta)$ along this line tends to 0 as $w \to \pm\infty$, implying that the central portion of the equilibrium surface becomes insensitive to I as w grows very large either negatively or positively. Note also that, taken together, Equations 3, 4, and 5 give approximate expressions for each of the three "pieces" of a neuron's steady-state I/γ relation $\bar{y}(I; w, \theta)$ (see Fig. 1).

2.2.2 Bifurcation set We can also derive expressions for the cusp bifurcation set of Equation 2 (Fig. 2B). By allowing us to calculate the location and size of the fold in input space as a function of the other parameters, such expressions tell us where in parameter space a single neuron with a self-connection will switch between unistable and bistable behavior.

Analytically, the cusp is defined by the simultaneous zeroes of $f(\gamma, w, \theta, I)$ and $f'(\gamma, w, \theta)$ (Hale & Koçak, 1991). Because $\sigma(x)$ can be written as $\left(\tanh(x/2) + 1\right)/2$, f' can be written as:

$$f'(\gamma, w, \theta) = \frac{w}{4}\operatorname{sech}^2\left(\frac{\gamma + \theta}{2}\right) - 1$$

If $f'(\gamma, w, \theta) = 0$, then $\gamma = \pm 2\operatorname{sech}^{-1}\left(2/\sqrt{w}\right) - \theta$ (retaining only the real solutions). Substituting this into $f(\gamma, w, \theta, I) = 0$ and solving for I gives:

$$I = \pm 2\operatorname{sech}^{-1}\left(\frac{2}{\sqrt{w}}\right) - w\sigma\left(\pm 2\operatorname{sech}^{-1}\left(\frac{2}{\sqrt{w}}\right)\right) - \theta$$

This can be rewritten in terms of elementary functions as follows to give exact expressions for the left (*lb*) and right (*rb*) branches of the cusp valid for $w \geq 4$:

$$\begin{aligned} lb(w, \theta) &\equiv 2\ln\left(\frac{\sqrt{w} + \sqrt{w-4}}{2}\right) - \frac{w + \sqrt{w(w-4)}}{2} - \theta \\ rb(w, \theta) &\equiv -2\ln\left(\frac{\sqrt{w} + \sqrt{w-4}}{2}\right) - \frac{w - \sqrt{w(w-4)}}{2} - \theta \end{aligned} \quad \textbf{(6)}$$

It will also be convenient to define an expression for the width of the fold (which is independent of θ) valid for $w \geq 4$ as:

$$d(w) \equiv rb(w, \theta) - lb(w, \theta) = \sqrt{w(w-4)} - 4\ln\left(\frac{\sqrt{w} + \sqrt{w-4}}{2}\right) \quad (7)$$

For $w \gg 4$, these three expressions approach the following simpler forms:

$$\begin{aligned} lb(w, \theta) &\approx \ln(w) - w + 1 - \theta \\ rb(w, \theta) &\approx -\ln(w) - 1 - \theta \\ d(w) &\approx w - 2\ln(w) - 2 \end{aligned}$$

2.3 Generalizations

The single-neuron theory described here can be straightforwardly translated to any model neuron related to Equation 2 by a linear affine change of variables. For example, the model neuron $\dot{y} = (-\gamma + \sigma(w\gamma + \theta + I))/\tau$ can be transformed into Equation 2 using the substitution $\gamma \mapsto (\gamma - I)/w$.

In addition, following the same basic derivation used earlier for $\sigma(x)$, it can be shown that the bifurcation set of $\dot{\gamma} = -\gamma + w\psi(\gamma + \theta) + I$ for an arbitrary smooth

activation function $\psi(x)$ will be given by the expression

$$I = \psi'^{-1}\left(\frac{1}{w}\right) - w\psi\left(\psi'^{-1}\left(\frac{1}{w}\right)\right) - \theta$$

where $\psi'^{-1}(x)$ will, in general, be a relation rather than a function, with restrictions on w and multiple branches. In the typical case where $\psi(x)$ is a bounded sigmoidal (S-shaped) function, w must be greater than or equal to the reciprocal of the maximum slope of $\psi(x)$, and $\psi'^{-1}(x)$ will have two branches. For example, when $\psi(x) = \tan^{-1}(x)$, $w \geq 1$ and $\psi'^{-1}(x) = \pm\sqrt{(1-x)/x}$.

Two-Neuron Circuits

We now turn our attention to the dynamics of two-neuron circuits:

$$\begin{aligned}\dot{y}_1 &= f_1(y_1, y_2) \equiv \frac{1}{\tau_1}(-y_1 + w_{11}\sigma(y_1 + \theta_1) + w_{21}\sigma(y_2 + \theta_2) + I_1)\\ \dot{y}_2 &= f_2(y_1, y_2) \equiv \frac{1}{\tau_2}(-y_2 + w_{12}\sigma(y_1 + \theta_1) + w_{22}\sigma(y_2 + \theta_2) + I_2)\end{aligned} \quad (8)$$

As we did for the single-neuron theory, we will attempt to characterize the range of dynamical behavior that is possible in two-neuron circuits and to understand how that behavior changes as parameters are varied. First, however, a way to relate the dynamics of two-neuron circuits to the single-neuron theory described in the previous section will be presented.

3.1 Synaptic input space diagrams

The analysis of a planar dynamical system is typically performed by studying the intersections of the nullclines (the curves along which either $\dot{y}_1 = 0$ or $\dot{y}_2 = 0$) in the state space (y_1, y_2). By setting $\dot{y}_1 = 0$ and $\dot{y}_2 = 0$ in Equation 8 and solving for y_2 and y_1, respectively, we can derive the following expressions for the nullclines:

$$\text{The } \dot{y}_1 \text{ nullcline: } y_2 = \sigma^{-1}\left(\frac{y_1 - w_{11}\sigma(y_1 + \theta_1) - I_1}{w_{21}}\right) - \theta_2$$

$$\text{The } \dot{y}_2 \text{ nullcline: } y_1 = \sigma^{-1}\left(\frac{y_2 - w_{22}\sigma(y_2 + \theta_2) - I_2}{w_{12}}\right) - \theta_1$$

where $\sigma^{-1}(x) = \ln\left(\frac{x}{1-x}\right)$.

However, these nullcline expressions in state space suffer from some important disadvantages for our purposes here. Because $\sigma^{-1}(x)$ is defined only for $0 < x < 1$, the preceding expression for the $\dot{y}_1$ nullcline, for example, is defined only when $0 < y_1 - w_{11}\sigma(y_1 + \theta_1) - I_1 < w_{21}$ (if $w_{21} > 0$) or $w_{21} < y_1 - w_{11}\sigma(y_1 + \theta_1) - I_1 < 0$

(if $w_{21} < 0$). When $w_{11} > 4$, these inequalities may be satisfied by as many as three disjoint ranges of y_1 values (Fig. 3C), complicating the phase plane analysis. It is also not obvious how these multiple branches are related to the single-neuron theory described earlier (Fig. 3A, B).

For these reasons, it will be more convenient to study the dynamics of Equation 8 in the synaptic input space (J_1, J_2) of the neurons, defined as $J_1 \equiv w_{21}\sigma(y_2 + \theta_2)$ and $J_2 \equiv w_{12}\sigma(y_1 + \theta_1)$. That is, the synaptic input to a neuron in a two-neuron circuit is just the input that it receives from the other neuron. Assuming that both w_{12} and w_{21} are nonzero, Equation 8 can be reformulated in synaptic input space as:

$$\dot{J}_1 = \frac{1}{\tau_2}\left(J_1 - \frac{J_1^2}{w_{21}}\right)\left(\frac{w_{22}}{w_{21}}J_1 - \ln\left(\frac{J_1}{w_{21} - J_1}\right) + J_2 + I_2 + \theta_2\right)$$

$$\dot{J}_2 = \frac{1}{\tau_1}\left(J_2 - \frac{J_2^2}{w_{12}}\right)\left(\frac{w_{11}}{w_{12}}J_2 - \ln\left(\frac{J_2}{w_{12} - J_2}\right) + J_1 + I_1 + \theta_1\right)$$

By setting these expressions equal to 0 and solving for J_2 and J_1, respectively, we obtain the following expressions for the nullclines in synaptic input space:

$$\text{The } \dot{J}_1(\dot{y}_2) \text{ nullcline:} \quad J_2 = \ln\left(\frac{J_1}{w_{21} - J_1}\right) - \frac{w_{22}}{w_{21}}J_1 - I_2 - \theta_2$$

$$\text{The } \dot{J}_2(\dot{y}_1) \text{ nullcline:} \quad J_1 = \ln\left(\frac{J_2}{w_{12} - J_2}\right) - \frac{w_{11}}{w_{12}}J_2 - I_1 - \theta_1$$

How are these expressions related to the steady-state I/y relations of the individual neurons? Because, according to the single-neuron theory, the state of neuron 1 in steady-state is given by $\bar{y}_1(J_1 + I_1, w_{11}, \theta_1)$, its steady-state synaptic input to neuron 2 is just $w_{12}\sigma(\bar{y}_1(J_1 + I_1, w_{11}, \theta_1) + \theta_1)$. Similar considerations apply to the steady-state synaptic input from neuron 2 to neuron 1. Using the linear nature of the effects of θ to simplify these expressions, the nullclines in synaptic input space can thus also be expressed as

$$\text{The } \dot{J}_1(\dot{y}_2) \text{ nullcline:} \quad \bar{J}_1(J_2; w_{22}, \theta_2, I_2, w_{21}) \equiv w_{21}\sigma(\bar{y}_2(J_2 + I_2 + \theta_2, w_{22}))$$

$$\text{The } \dot{J}_2(\dot{y}_1) \text{ nullcline:} \quad \bar{J}_2(J_1; w_{11}, \theta_1, I_1, w_{12}) \equiv w_{12}\sigma(\bar{y}_1(J_1 + I_1 + \theta_1, w_{11}))$$

Thus, in synaptic input space, the nullclines (Fig. 3D) are just shifted, squashed, and scaled versions of the steady-state I/y relations of each neuron (Fig. 3A,B), of whose properties we have a very good qualitative and quantitative understanding from the single-neuron theory. For example, we know the following about the $\dot{y}_1$ nullcline $\bar{J}_2(J_1)$:

- w_{11} controls the width and location of the fold (or "kink" if $w_{11} < 4$) of $\bar{J}_2(J_1)$ as described by Equation 6.

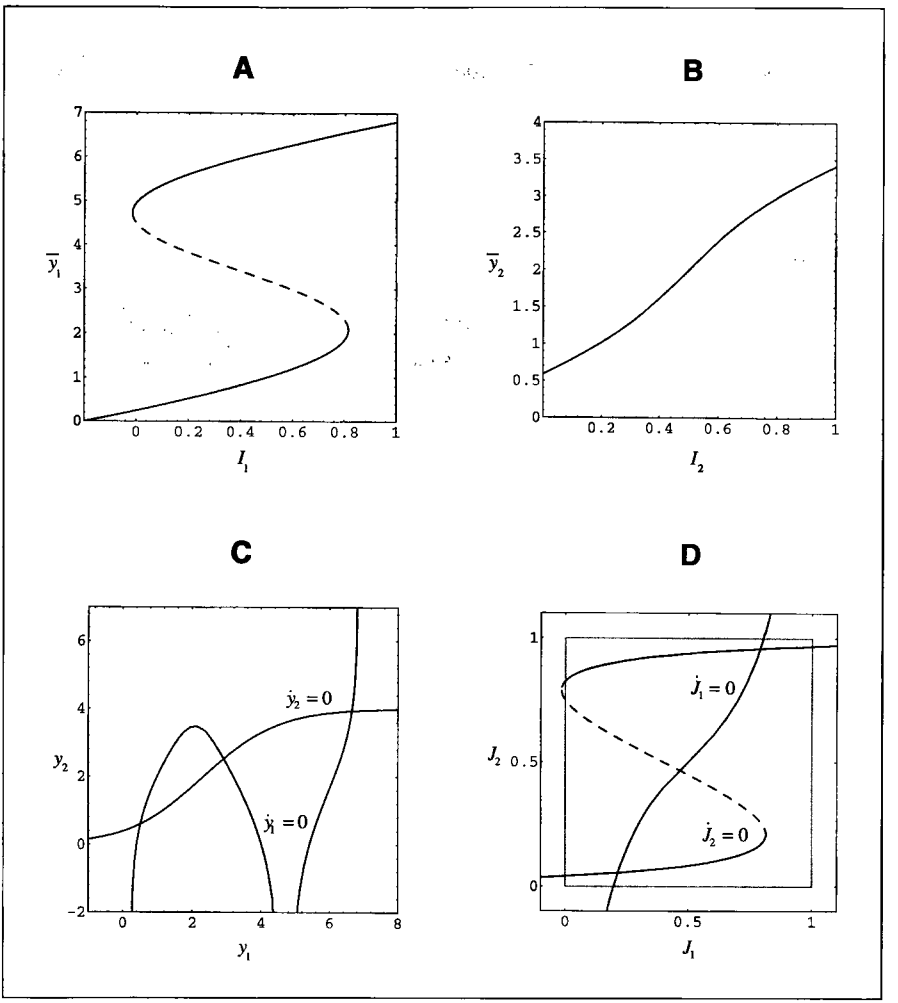


Figure 3

Synaptic input space diagrams. The steady-state $1/\gamma$ relations for neurons with (A) $w = 6$, $\theta = -3.4$ and (B) $w = 3$, $\theta = -2$. (C) The state space nullclines of a circuit formed by connecting the neurons in A (neuron 1) and B (neuron 2) with $w_{12} = w_{21} = 1$. Note that the nullcline of neuron 1 has two distinct branches in state space. (D) The synaptic input space nullclines of the same circuit shown in C. The gray box delimits the range of possible synaptic inputs that can occur in this circuit. Note that both nullclines are continuous curves in synaptic input space.

- θ_1 and I_1 shift the location of the fold of $\bar{J}_2(J_1)$ horizontally by $-\theta_1$ and $-I_1$, respectively.
- w_{12} controls the vertical scale and orientation of $\bar{J}_2(J_1)$.

Similar considerations apply to $\bar{J}_1(J_2)$. Of course, the intersections of $\bar{J}_2(J_1)$ and $\bar{J}_1(J_2)$ give the equilibrium points of Equation 8. If desired, the results of this analysis can be mapped back into state space by $\bar{y}_1 = \sigma^{-1}\left(\bar{J}_2/w_{12}\right) - \theta_1$ and $\bar{y}_2 = \sigma^{-1}\left(\bar{J}_1/w_{21}\right) - \theta_2$.

3.2 Phase portraits

What sorts of dynamical behavior are possible in a two-neuron circuit? It is relatively easy to see that, under the transformations just described, the nullclines of a two-neuron circuit can generically intersect only an odd number of times totaling between one and nine inclusive. While two, four, six, or eight intersections are possible, they are not structurally stable because they can be destroyed by infinitesimal perturbations to the parameters. Thus, Equation 8 can generically exhibit only phase portraits with one, three, five, seven, or nine equilibria (Ermentrout, 1995). The purpose of this section is to characterize these generic phase portraits.

To simplify this task, we can distinguish between the "global" or overall qualitative form of a phase portrait and the local behavior around each of the stable and unstable equilibrium points (i.e., whether they are nodes or spirals). Taking into account the various symmetries of the system, I have found only eleven qualitatively distinct phase portraits in a two-neuron CTRNN. These are listed in Table 1 and have been labeled *1, 1lc, 3a, 3b, 3lc, 5a, 5b, 5c, 5lc, 7,* and *9*. The first number in these names simply reflects the number of equilibrium points in the corresponding phase portrait. In the case of three and five equilibrium points, an additional letter is used to distinguish between distinct phase portraits having the same number of equilibrium points. The suffix *lc* signals the presence of a limit cycle.

Phase portraits *3lc* and *5lc* are particularly interesting. For example, in *5lc*, a stable limit cycle coexists with two stable equilibrium points, making it possible to switch this system between rhythmical behavior and two different quiescent states with appropriate input pulses. It is not known whether limit cycles can also occur in conjunction with more than two stable equilibrium points, or whether more than one limit cycle can occur simultaneously, though I have never found any examples of such behavior.

In contrast to the small number of qualitatively distinct phase portraits, there are many phase portraits with different local behavior. These local variations are denoted by appending a number to the global name. For example, *3a.3* has the same global layout as *3a.1*, but the two stable nodes in *3a.1* are stable spirals in *3a.3*. Table 1

Table 1 Names and configurations of the two-neuron phase portraits

Global name	Local name	Stable nodes	Stable spirals	Unstable nodes	Unstable spirals	Saddle nodes	Limit cycles
1	1.1	1					
	1.2		1				
1lc	1lc.1				1		1
	1lc.2			1			1
3a	3a.1	2				**1**	
	3a.2	1	1			**1**	
	3a.3		2			**1**	
3b	3b.1	1			1	1	
	3b.2		1		1	1	
3lc	3lc	1			1	1	1
5a	5a.1	3				2	
	5a.2	2	1			**2**	
5b	5b.1	2		1		2	
	5b.2	2			**1**	**2**	
	5b.3		2		**1**	**2**	
	5b.4	1	1		**1**	**2**	
5c	5c	2			1	2	
5lc	5lc	2			1	2	1
7	7.1	3		**1**		**3**	
	7.2	3			1	3	
9	9.1	4		1		4	
	9.2	4			**1**	**4**	
	9.3	**2**	2	1		**4**	
	9.4	2	2		1	4	

does not show all possible local variations. Rather, only those phase portraits that actually appear in the bifurcation maps described in the next section, as well as a few additional phase portraits that illustrate the range of possibilities, are listed in the table.

Examples of each of the eleven qualitatively distinct phase portraits, along with their corresponding input space diagrams, are shown in Figure 4. Each intersection

between the two nullclines in synaptic input space corresponds to an equilibrium point in the associated phase portrait. Note how the stable and unstable manifolds of the saddle points serve to separate the basins of attraction of the stable limit sets. Of course, in general there are many different input space diagrams that give rise to the same qualitative global behavior. However, it is clear that our understanding of how the various parameters transform the input space diagrams gives us a great deal of information. For example, in phase portrait *3a*, shifting the **N**-shaped nullcline down and the **S**-shaped nullcline to the left by modifying the biases appropriately will lead to phase portrait *1lc*.

Reasoning about the geometry of synaptic input space diagrams can also lead to quantitative conditions for the occurrence of various phase portraits. For example, it is obvious from the example of phase portrait 9 in Figure 4 that a necessary condition for 9 equilibria is that the folds of both nullclines be sufficiently wide. However, even setting $d(w_{11}) > |w_{21}|$ and $d(w_{22}) > |w_{12}|$ (using Equation 7 for $d(w)$) is not sufficient because the biases must also be set appropriately so as to compensate for the shift in the location of the fold caused by the self-weights. Assuming that $w_{12}, w_{21} > 0$, sufficient conditions for 9 equilibria are that $lb(w_{11}, \theta_1) < 0$, $rb(w_{11}, \theta_1) > w_{21}$, $lb(w_{22}, \theta_2) < 0$, and $rb(w_{22}, \theta_2) > w_{12}$. These conditions will be generalized to networks of arbitrary size in section 4.2.

While synaptic input space diagrams allow us to understand the overall number and location of equilibrium points, it is also important to know the local behavior around each equilibrium point (i.e., do nearby trajectories approach or recede and in what manner?). In general, the local behavior in the neighborhood of an equilibrium point $\bar{\mathbf{y}} = (\bar{y}_1, \bar{y}_2)$ of Equation 8 is given by the eigenvalues of $D\mathbf{f}(\bar{\mathbf{y}})$, the Jacobian matrix of $\mathbf{f}$ evaluated at $\bar{\mathbf{y}}$ (Hale & Koçak, 1991):

$$D\mathbf{f}(\bar{\mathbf{y}}) = \begin{bmatrix} \frac{\partial f_1}{\partial y_1}(\bar{y}_1, \bar{y}_2) & \frac{\partial f_1}{\partial y_2}(\bar{y}_1, \bar{y}_2) \\ \frac{\partial f_2}{\partial y_1}(\bar{y}_1, \bar{y}_2) & \frac{\partial f_2}{\partial y_2}(\bar{y}_1, \bar{y}_2) \end{bmatrix} = \begin{bmatrix} \frac{w_{11}\sigma'(\bar{y}_1+\theta_1)-1}{\tau_1} & \frac{w_{21}\sigma'(\bar{y}_2+\theta_2)}{\tau_1} \\ \frac{w_{12}\sigma'(\bar{y}_1+\theta_1)}{\tau_2} & \frac{w_{22}\sigma'(\bar{y}_2+\theta_2)-1}{\tau_2} \end{bmatrix}$$

The eigenvalues of this matrix (that is, the λ that satisfy the characteristic equation $\det(D\mathbf{f}(\bar{\mathbf{y}}) - \lambda\mathbf{I}) = 0$, where $\mathbf{I}$ is the identity matrix) can be written as:

$$\lambda_1, \lambda_2 = \frac{w_{11}\sigma'(\bar{y}_1 + \theta_1) - 1}{2\tau_1} + \frac{w_{22}\sigma'(\bar{y}_2 + \theta_2) - 1}{2\tau_2} \pm \frac{1}{2}\sqrt{\left(\frac{w_{11}\sigma'(\bar{y}_1 + \theta_1) - 1}{\tau_1} - \frac{w_{22}\sigma'(\bar{y}_2 + \theta_2) - 1}{\tau_2}\right)^2 + \frac{4w_{12}w_{21}\sigma'(\bar{y}_1 + \theta_1)\sigma'(\bar{y}_2 + \theta_2)}{\tau_1\tau_2}} \quad (9)$$

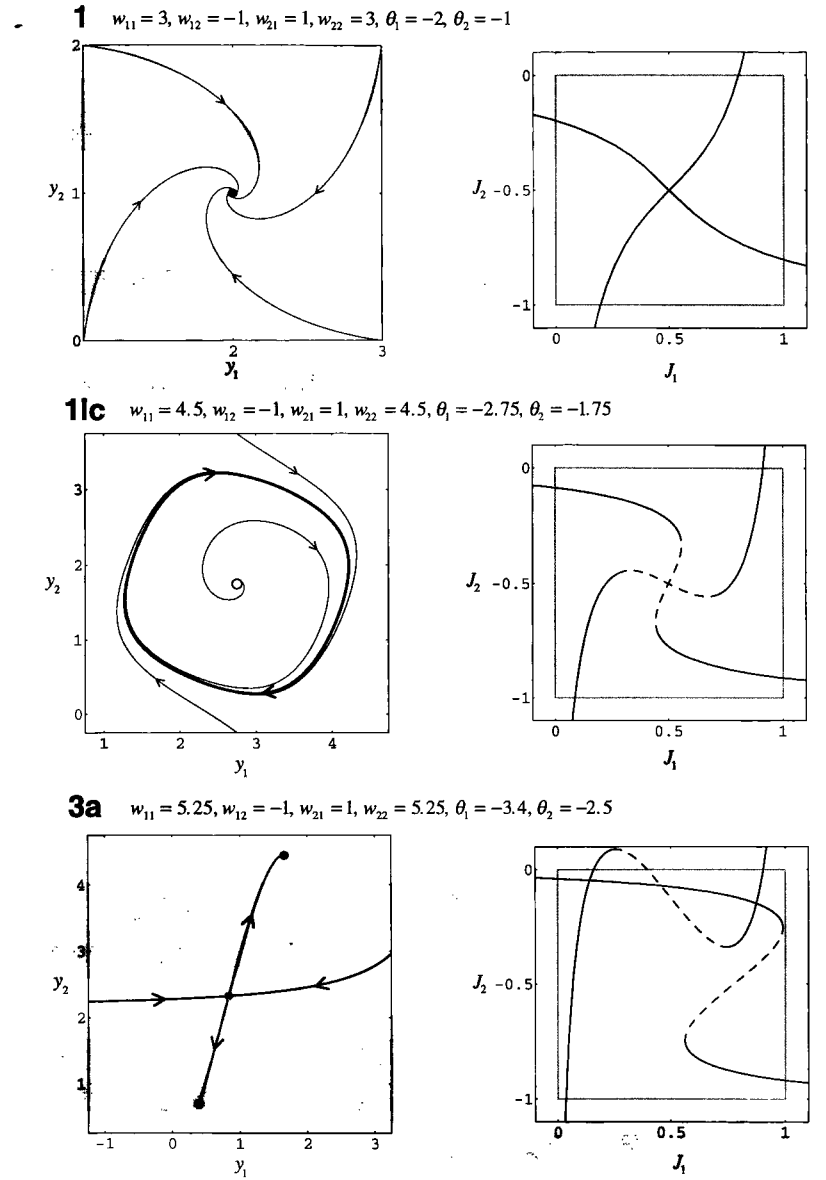


Figure 4a Examples of 11 generic, qualitatively distinct phase portraits that can occur in a two-neuron circuit. For each example, the phase portrait is shown on the left and the corresponding synaptic input space diagram is shown on the right. Stable equilibrium points are shown as solid disks, unstable equilibrium points are shown as circles, and saddle points are shown as gray disks. Special invariant trajectories (limit cycles and the stable and unstable manifolds of saddle points) are shown as thick lines. In the case of phase portraits *1* and *1lc* only, sample trajectories are shown as thin lines. The gray boxes in the synaptic input space diagrams delimit the ranges of possible synaptic inputs that can occur. All examples have $\tau_1 = \tau_2 = 1$.

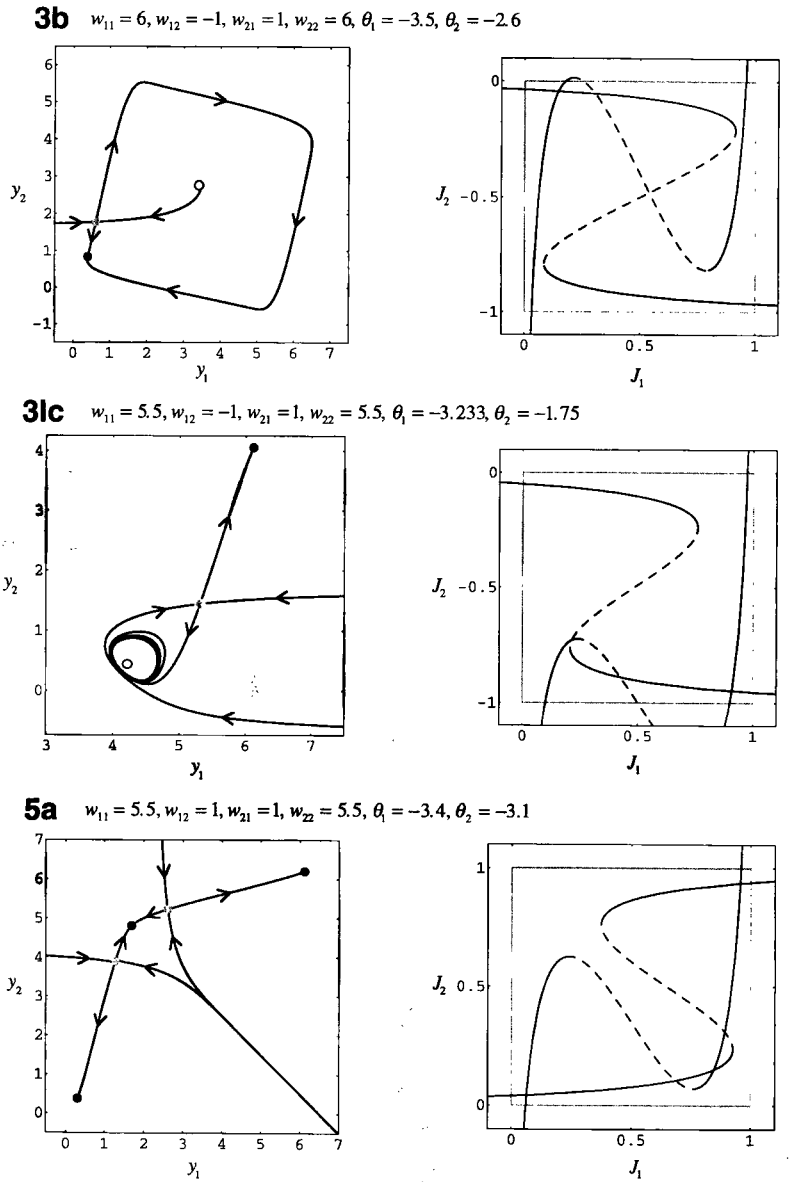
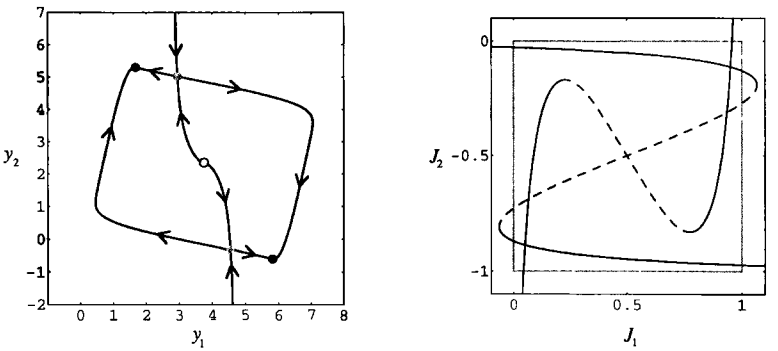
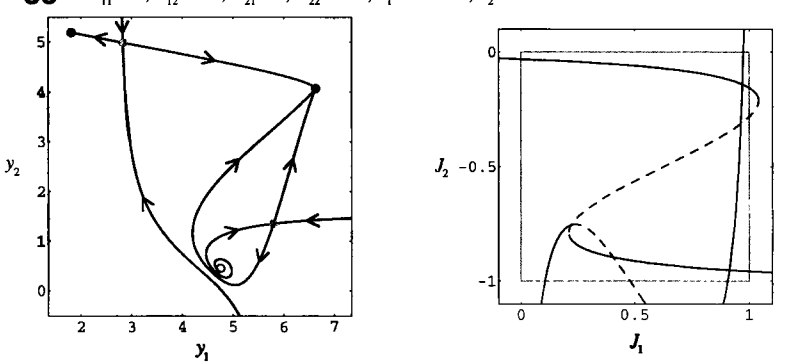


Figure 4b Continued.

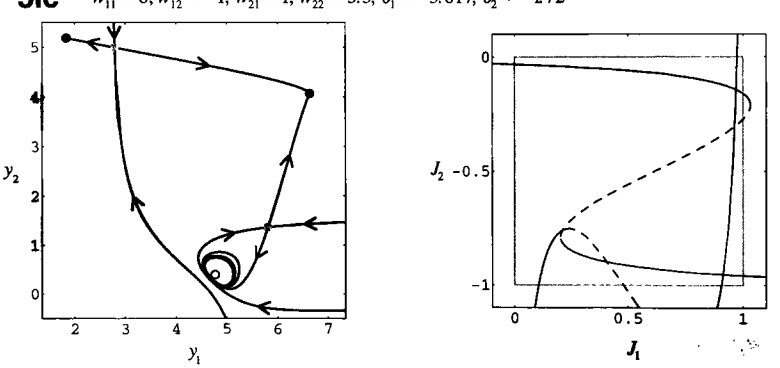
5b $w_{11} = 6.5, w_{12} = -1, w_{21} = 1, w_{22} = 5.7, \theta_1 = -3.75, \theta_2 = -2.35$



5c $w_{11} = 6, w_{12} = -1, w_{21} = 1, w_{22} = 5.5, \theta_1 = -3.625, \theta_2 = -1.72$



5lc $w_{11} = 6, w_{12} = -1, w_{21} = 1, w_{22} = 5.5, \theta_1 = -3.617, \theta_2 = -1.72$



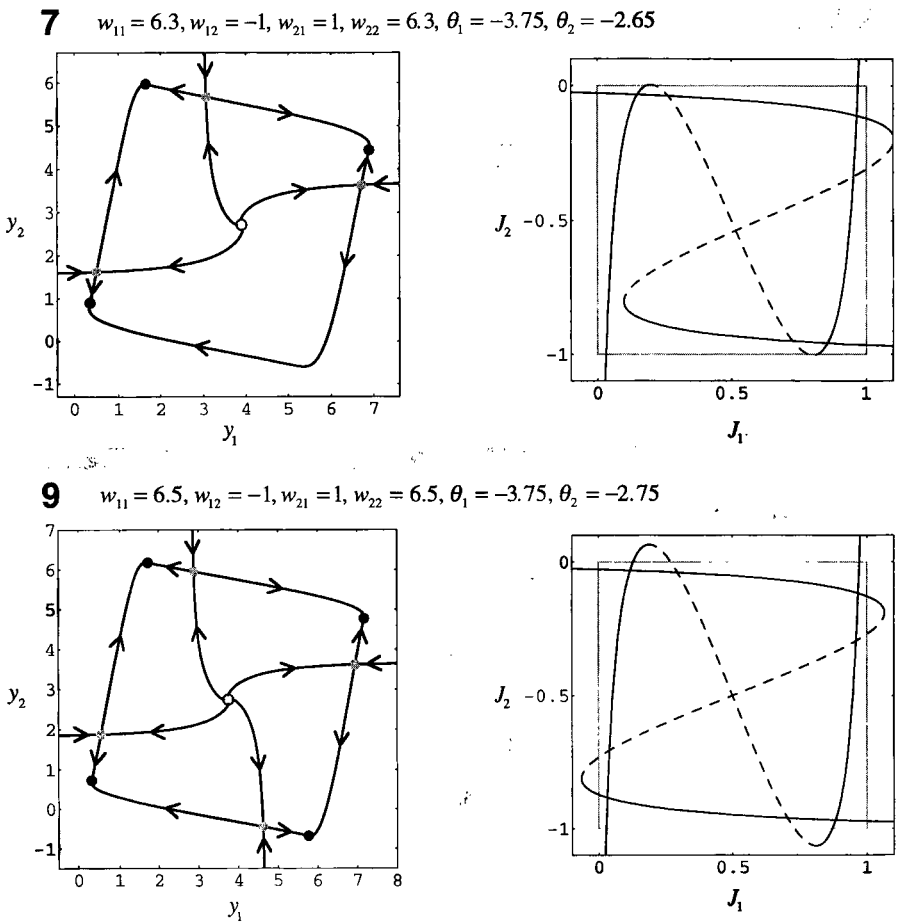


Figure 4d Continued.

An equilibrium point $\bar{\mathbf{y}}$ of Equation 8 is *stable* if $\mathrm{Re}[\lambda_1], \mathrm{Re}[\lambda_2] < 0$, *unstable* if $\mathrm{Re}[\lambda_1], \mathrm{Re}[\lambda_2] > 0$, and a *saddle* if $\mathrm{Re}[\lambda_1]$ and $\mathrm{Re}[\lambda_2]$ have opposite signs (Hale & Koçak, 1991). Furthermore, $\bar{\mathbf{y}}$ is a *node* if $\mathrm{Im}[\lambda_1] = \mathrm{Im}[\lambda_2] = 0$ and a *spiral* if $\mathrm{Im}[\lambda_1] \neq 0$ and $\mathrm{Im}[\lambda_2] \neq 0$. From Equation 9, it can be seen that $\bar{\mathbf{y}}$ is a node if the expression under the radical is positive and a spiral if this expression is negative. Because $\sigma'(x), \tau_1, \tau_2 > 0$, it is therefore obvious that spiral behavior can occur only when w_{12} and w_{21} are of opposite sign.

3.3 Bifurcations

Now that we understand the range of phase portraits that are possible in two-neuron circuits, in this section we will study the layout of these phase portraits in parameter space and the ways in which one phase portrait can change into another as network parameters are varied. Because a two-neuron CTRNN has ten parameters, the best that we can hope to do is examine selected slices through this ten-dimensional parameter space and try to develop a general understanding of the effects of the different parameters.

3.3.1 Connection weights Let us begin with the roles of the connection weights. Recall that the self-weights affect the extent and location of the nullcline folds, whereas the cross-weights merely scale the nullclines in input space.

To allow some analysis of the effects of these important parameters, we will study networks in which the nullclines intersect at their exact centers. This can be achieved by setting $\theta_1^* = -(w_{11}+w_{21}+I_1)/2$ and $\theta_2^* = -(w_{22}+w_{12}+I_2)/2$, corresponding to networks in which the sigmoidal activation function of each neuron is centered over the range of net inputs that it receives. In this case, the central equilibrium point will always occur at $\overline{\mathbf{y}}^* = (-\theta_1^*, -\theta_2^*)$. Because $\sigma'(\overline{y}_1^* + \theta_1^*) = \sigma'(\overline{y}_2^* + \theta_2^*) = \sigma'(0) = 1/4$ (and assuming for now that $\tau_1 = \tau_2 = 1$), Equation 9 for the eigenvalues of $\overline{\mathbf{y}}^*$ simplifies to:

$$\lambda_1^*, \lambda_2^* = \frac{w_{11}}{8} + \frac{w_{22}}{8} - 1 \pm \sqrt{\left(\frac{w_{11}}{8} - \frac{w_{22}}{8}\right)^2 + \frac{w_{12}w_{21}}{16}} \tag{10}$$

Bifurcation maps showing the effects of varying w_{11} and w_{22} in center-crossing networks with $w_{12}w_{21} = 1$ and $w_{12}w_{21} = -1$ are shown in Figures 5A and B, respectively. The curves in these figures represent boundaries between regions with qualitatively distinct phase portraits. Not surprisingly, the map for the $w_{12}w_{21} = 1$ network is much simpler than that for the $w_{12}w_{21} = -1$ network, because the many local variations listed in Table 1 involve spirals that can only occur in the latter case. Generally speaking, however, it is clear in both cases that increasing the self-weights leads to phase portraits with larger numbers of equilibrium points in center-crossing networks because nullclines with wide, deep folds can intersect more frequently than can those with narrow, shallow folds or no folds at all.

The advantage of center-crossing networks is that we can fully analyze the stability of the central equilibrium point $\overline{\mathbf{y}}^*$. When $w_{12}w_{21} > 0$, $\overline{\mathbf{y}}^*$ must be a node because the eigenvalues are always real. By setting expression (10) equal to 0, we find that the critical surface across which $\overline{\mathbf{y}}^*$ changes stability is given by the following expression:

$$4w_{11} + 4w_{22} - w_{11}w_{22} + w_{12}w_{21} = 16 \tag{11}$$

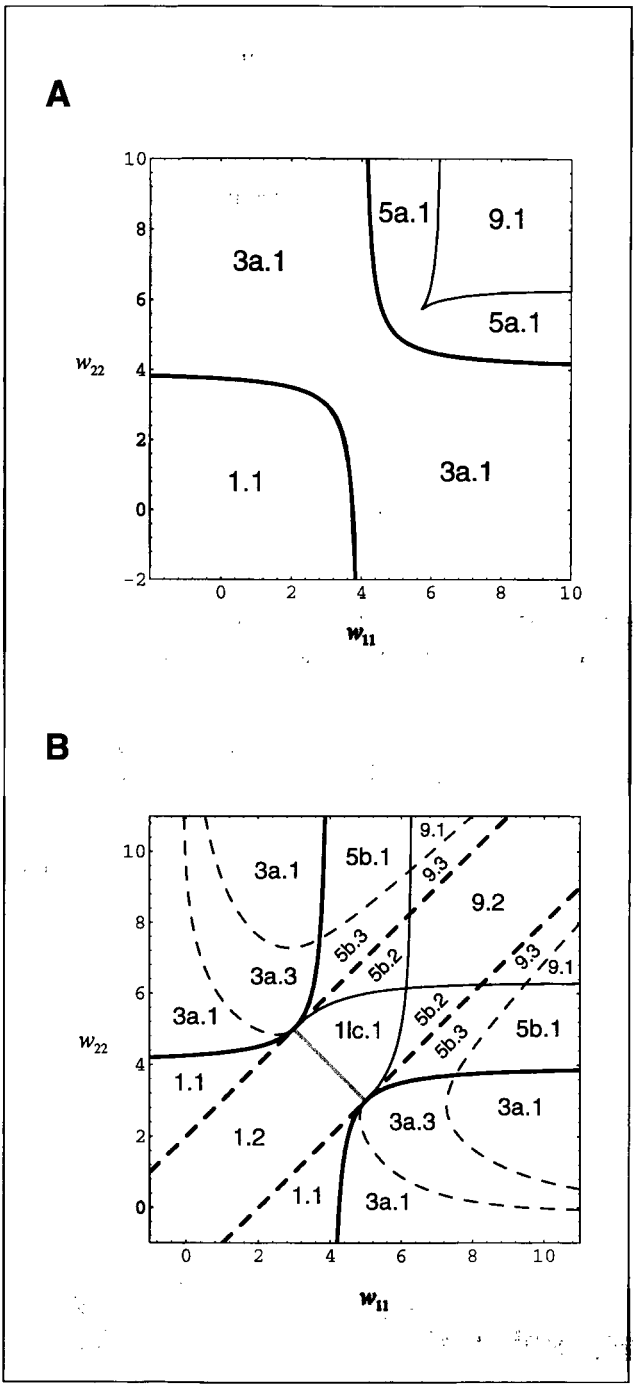


Figure 5
Self-weight bifurcation maps of center-crossing circuits with (A) symmetric ($w_{12} = w_{21} = 1$) and (B) antisymmetric ($w_{12} = -1$, $w_{21} = 1$) cross-connections. Thick lines denote analytically computed bifurcation curves (see Section 3.3.1), while thin lines denote numerically computed bifurcation curves. Solid curves represent a single zero eigenvalue (i.e., a saddle-node bifurcation). Gray lines represent purely imaginary eigenvalues (i.e., a Hopf bifurcation). Dashed lines represent critically real eigenvalues (i.e., a change in local behavior around an equilibrium point). The labels refer to phase portraits listed in Table 1, examples of which can be found in Figure 4.

As can be seen in Figure 5A for fixed coupling weights $w_{12} = w_{21} = 1$, this expression defines a hyperbola in (w_{11}, w_{22}) space. When the self-weights are below the lower branch of this hyperbola, $\bar{\mathbf{y}}^*$ is stable. Between the two branches, $\bar{\mathbf{y}}^*$ is a saddle. Above the upper branch, $\bar{\mathbf{y}}^*$ is unstable. The remaining bifurcation boundary shown in Figure 5A, which separates regions containing five and nine equilibrium points, was computed numerically because it does not involve any change in $\bar{\mathbf{y}}^*$. Such saddle-node bifurcations, which occur whenever the edge of the fold of one nullcline crosses the other nullcline, are a common feature of this system. A sequence of synaptic input space diagrams illustrating the changes that occur along the diagonal of this bifurcation map is shown in Figure 6.

When $w_{12}w_{21} < 0$, there are several cases to consider. By setting the portion of Equation 10 under the radical to 0, we find that $\bar{\mathbf{y}}^*$ will switch type across the critical surface defined by

$$(w_{11} - w_{22})^2 = 4|w_{12}w_{21}|$$

This condition gives rise to the two dashed diagonal lines in the bifurcation map shown in Figure 5B. When the self-weights fall between these two lines, $\bar{\mathbf{y}}^*$ will be a spiral. Outside, it will be a node.

When $\bar{\mathbf{y}}^*$ is a spiral, its critical stability surface can be found by setting the real part of Equation 10 to 0:

$$w_{11} + w_{22} = 8$$

This condition gives rise to the short gray line connecting the two dashed diagonals in Figure 5B. As the self-weights cross from below this line to above it, $\bar{\mathbf{y}}^*$ undergoes a Hopf bifurcation (Hale & Koçak, 1991), losing stability and giving birth to a limit cycle.

On the other hand, when $\bar{\mathbf{y}}^*$ is a node (i.e., when $w_{12}w_{21} < 0$ but $(w_{11} - w_{22})^2 > 4|w_{12}w_{21}|$), the critical stability surface is once again given by Equation 11, derived earlier, except in this case $w_{12}w_{21} = -1$. This expression defines the thick black hyperbola branches shown in Figure 5B. The remaining bifurcation curves in this map were computed numerically. Note that, as the product of the cross-weights passes from -1 through 0 to 1, many of the bifurcation curves in Figure 5B merge to produce the curves shown in Figure 5A.

3.3.2 Biases Another important set of parameters are those that shift the nullclines in input space. Because the effects of the biases (θ_1 and θ_2) and the external inputs (I_1 and I_2) on the locations of the nullclines in synaptic input space are identical, we will study only the biases here. Bifurcation maps for the cases $w_{12}w_{21} = 1$ and $w_{12}w_{21} = -1$ are shown in Figures 7A and B, respectively. In Figure 7A, the

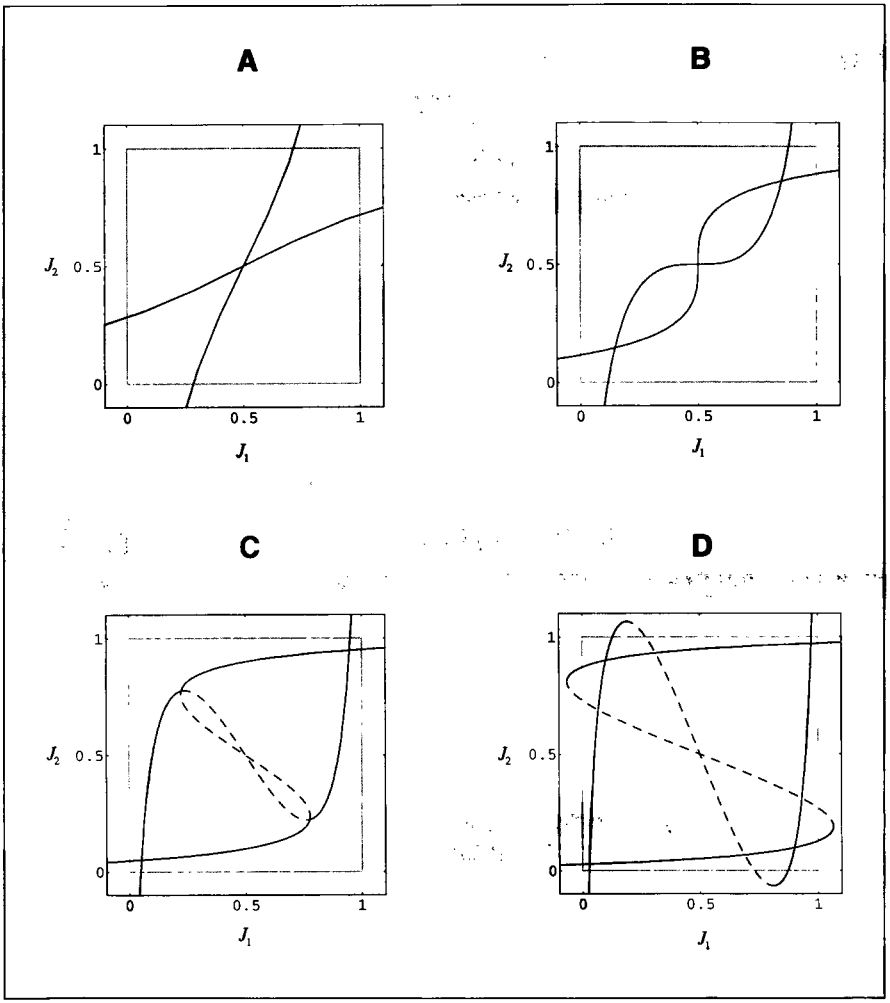


Figure 6
A sequence of synaptic input diagrams taken from points along the diagonal of the self-weight bifurcation map shown in Figure 5A. These diagrams exhibit (A) one equilibrium point ($w_{11} = w_{22} = 2$), (B) three equilibrium points ($w_{11} = w_{22} = 4$), (C) five equilibrium points ($w_{11} = w_{22} = 5.5$), and (D) nine equilibrium points ($w_{11} = w_{22} = 6.5$).

self-weights were chosen sufficiently large that nine equilibria were possible. In Figure 7B, on the other hand, the self-weights were chosen sufficiently small that limit cycles were possible (so that the folds of each nullcline "fit inside" one another). Otherwise, regions with seven and nine equilibria would appear instead of limit cycles. All bifurcation curves shown here were computed numerically. Once again,

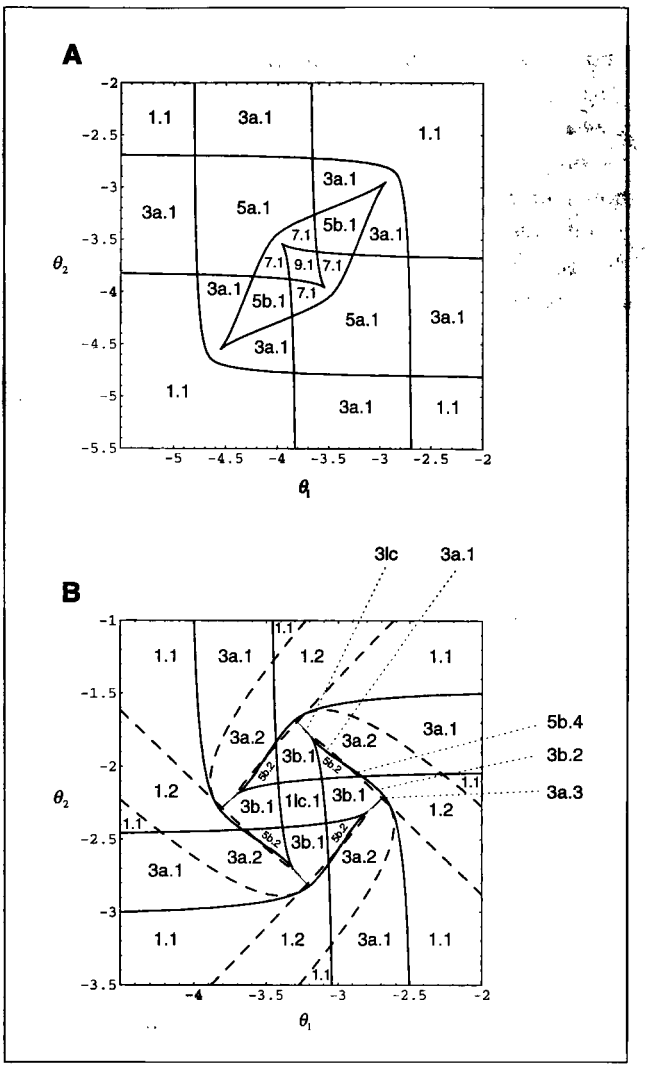


Figure 7
Bias bifurcation maps. (A) Symmetric circuit with $w_{11} = w_{22} = 6.5$ and $w_{12} = w_{21} = 1$. (B) Antisymmetric circuit with $w_{11} = w_{22} = 5.5$, $w_{12} = -1$ and $w_{21} = 1$. Line style conventions are the same as for Figure 5 except that all bifurcation curves shown here were numerically computed. As described in the text, narrow regions of different local behavior surround the central diamond in (B). Although these regions are labeled with dotted lines for one edge only, the other edges should be identically labeled.

the $w_{12}w_{21} = 1$ bifurcation map is simpler than the $w_{12}w_{21} = -1$ map. Generally speaking, however, it is clear how the number of possible intersections increases and decreases as the two nullclines slide past one another in input space.

Many interesting phenomena surround the central diamond in Figure 7B. For example, there are narrow regions of different local behavior. The neighborhoods of the gray bifurcation boundaries are especially noteworthy. Outside these boundaries,

phase portrait *3a.2* appears. When this boundary is crossed, the stable spiral of *3a.2* undergoes a Hopf bifurcation, giving rise to a limit cycle and an unstable spiral (phase portrait *3lc*). The region of existence of *3lc* is smaller than the width of the gray line in this figure and is therefore not shown. At the inner boundary of this region, the limit cycle in *3lc* swells until it touches the saddle point, undergoing a homoclinic loop bifurcation (Hale & Koçak, 1991) (a global bifurcation in which an unstable and stable manifold of a saddle point coincide without any change in the local properties of any equilibrium points), leaving phase portrait *3b.1*. Finally, near the center of the map, the stable and saddle equilibrium points in *3b.1* merge in a saddle-node bifurcation on a loop (Hale & Koçak, 1991), leaving the stable limit cycle and unstable spiral of phase portrait *1lc.1*.

3.3.3 Time constants The final parameters that need to be considered are the time constants τ_1 and τ_2. I will not make any detailed study of the effects of these parameters here. However, a few general observations can easily be made. First, it is clear that only the ratio of these time constants matters because we can always eliminate one of the time constants by rescaling time appropriately. Second, because the nullclines are independent of the time constants, varying the time constants cannot change the number of equilibria. Finally, varying the ratio of the time constants can affect the stability and type of equilibria. For example, changing this ratio can cause transitions between phase portraits *1* and *1lc* (i.e., a Hopf bifurcation), as well as transitions between local variations of a given phase portrait (e.g., *3b.1* and *3b.2*).

3.3.4 Asymptotic forms of bifurcation curves Although the exact form of many of the bifurcation curves described previously must be computed numerically, the geometry of synaptic input space diagrams allows us to derive asymptotic expressions for some of these curves. For example, consider the right-hand boundary between regions of five and nine equilibria in both Figures 5A and 5B. In both of these cases, this boundary is defined geometrically by the condition that the edges of the fold of the $\dot{y}_2$ nullcline are tangent to the lower and upper branches of the $\dot{y}_1$ nullcline. As w_{11} increases, the location of these lower and upper branches in synaptic input space approaches 0 and w_{12}. Thus, for a center-crossing network, the asymptotic form of this bifurcation curve is defined simply by the condition that the width of the fold in the $\dot{y}_2$ nullcline be equal to the magnitude of w_{12}—that is, $d(w_{22}) = |w_{12}|$. Numerically solving this equation for w_{22} with $|w_{12}| = 1$ gives an approximate asymptotic value for this boundary of $w_{22} \approx 6.28755$. Likewise, the other numerically computed bifurcation curve shown in Figures 5A and 5B is defined by the equation $d(w_{11}) = |w_{21}|$. Of course, the advantage of these equations is that

we can compute the asymptotic location of these boundaries for arbitrary values of the cross-weights.

Using similar reasoning, the following asymptotic equations can be derived for the eight numerically computed saddle-node bifurcation curves that extend beyond the edges of Figures 7A and 7B: $lb(w_{11}, \theta_1) = 0$, $rb(w_{11}, \theta_1) = 0$, $lb(w_{11}, \theta_1) = w_{12}$, $rb(w_{11}, \theta_1) = w_{12}$, $lb(w_{22}, \theta_2) = 0$, $rb(w_{22}, \theta_2) = 0$, $lb(w_{22}, \theta_2) = w_{21}$, and $rb(w_{22}, \theta_2) = w_{21}$. By substituting the particular weight values used in Figures 7A and 7B and solving for the biases in these equations (which can be done algebraically in this case as the biases enter only linearly), we can obtain exact asymptotic values for these eight saddle-node bifurcation curves.

Circuits of Arbitrary Size

Because three-neuron circuits already raise most of the difficulties of the general case, let us now turn to a consideration of Equation 1 for arbitrary N. Before we proceed, it is important to understand that general results are notoriously difficult to achieve in nonlinear dynamical systems of dimension greater than two. Whereas many of the local techniques of dynamical systems theory (e.g., the stability analysis of equilibrium points through linearization) can be straightforwardly applied in higher-dimensional systems, the range of phase portraits and bifurcations that can occur in higher dimensions is bewildering. Consequently, this section presents only a brief sampling of interesting results, and there is much room for further research. Where appropriate, these results are illustrated with examples drawn from three-neuron circuits.

4.1 Nullsurfaces

Assuming that the weight matrix is zero-diagonal nonsingular, we can always reformulate Equation 1 in synaptic input space. Unfortunately, for $N > 2$ the corresponding nullsurfaces in synaptic input space do not bear the straightforward relationship to the individual neuron I/γ relations that they did for $N = 2$, and thus the synaptic input space diagrams we used there cannot be generalized to larger networks. In state space, the $\dot{\gamma}_1$ nullsurface of an N-neuron circuit can be written as:

$$\gamma_2 = \sigma^{-1}\left(\frac{\gamma_1 - w_{11}\sigma(\gamma_1 + \theta_1) - \sum_{j=3}^{N} w_{j1}\sigma(\gamma_j + \theta_j) - I_1}{w_{21}}\right) - \theta_2$$

where this expression could have just as easily been solved for any of the other state variables except γ_1. Analogous expressions can be derived for the $N - 1$ other nullsurfaces. No detailed study of the properties of these surfaces is made here except to note that, as for two-neuron circuits, these surfaces can exhibit distinct branches.

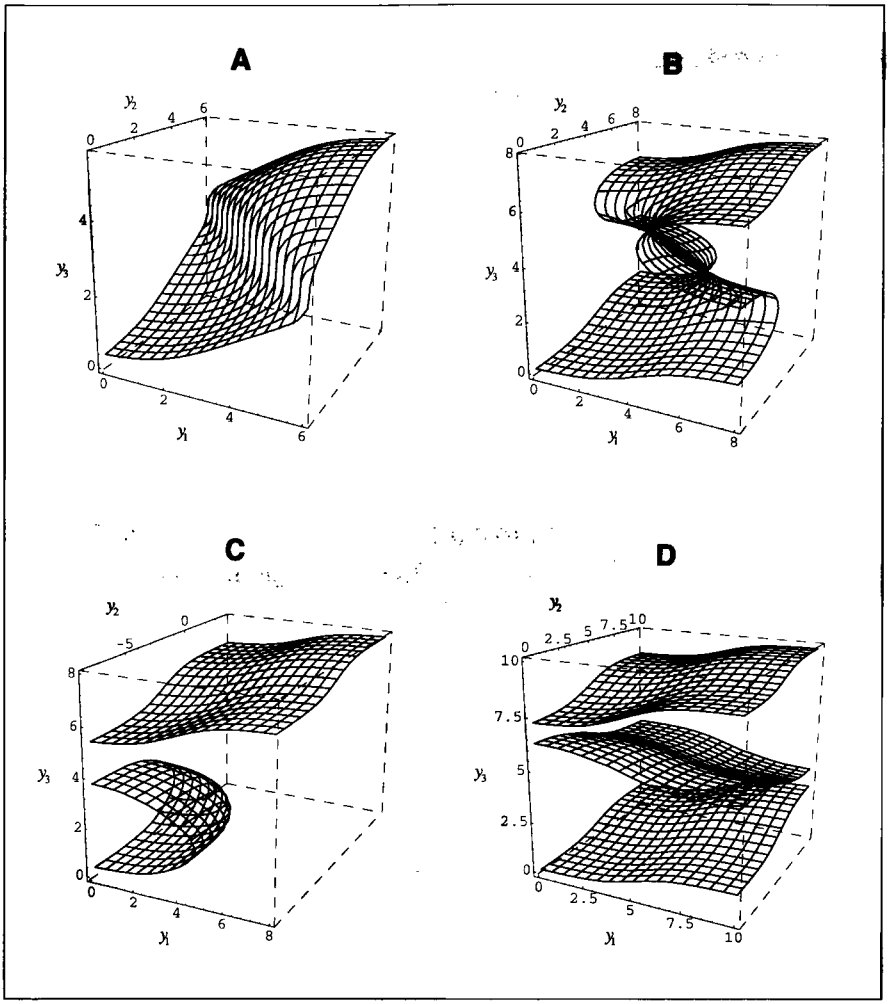


Figure 8
Example $\dot{y}_3$ nullsurfaces for a three-neuron circuit. In all cases, $w_{13} = w_{23} = 1$. (A) $w_{33} = 4$, $\theta_1 = \theta_2 = \theta_3 = -3$. (B) $w_{33} = 6$, $\theta_1 = \theta_2 = \theta_3 = -4$. (C) $w_{33} = 6$, $\theta_1 = -3$, $\theta_2 = 2$, $\theta_3 = -4$. (D) $w_{33} = 8$, $\theta_1 = \theta_2 = \theta_3 = -5$. The $\dot{y}_1$ and $\dot{y}_2$ nullsurfaces are similar but orthogonally oriented. The intersections of these three sets of nullsurfaces correspond to the equilibrium points of a three-neuron circuit.

Several examples of the nullsurfaces for one neuron of a three-neuron circuit are shown in Figure 8.

4.2 Phase portraits

Already in three-neuron circuits there are too many different phase portraits to enumerate explicitly, even if we restrict ourselves to qualitatively distinct ones and take the various symmetries into account. Figure 9 shows some of the many interesting phase portraits that can occur in three-neuron circuits, including a circuit with 27 equilibrium points (8 stable, 18 saddles, and 1 unstable [Fig. 9A]), a circuit with two distinct periodic orbits (Fig. 9B), a circuit with a doubly periodic orbit (Fig. 9C), and a circuit with chaotic dynamics (Fig. 9D). The latter three phase portraits were obtained merely by varying a single time constant in the same circuit.

The largest Lyapunov exponent of the chaotic attractor is 0.010, and its Lyapunov dimension is 2.021 (Parker & Chua, 1989). Roughly speaking, these numbers imply that this circuit is mildly chaotic. To my knowledge, this is the smallest autonomous CTRNN in which chaotic dynamics have been observed. It is also the smallest possible, since a smooth dynamical system must be at least three-dimensional for chaos to occur (Hale & Koçak, 1991). The chaotic circuit described by Das et al. (1991) had four neurons, and the chaotic dynamics described by Sompolinsky and Crisanti (1988) occurred only when the number of neurons was large.

The maximum number of equilibria that an N-neuron circuit can generically exhibit is 3^N. A sufficient (but not necessary) condition for 3^N equilibria occurs when the fold of each neuron in the circuit stretches beyond the range of synaptic input that it receives from all the other neurons (i.e., for each neuron i, $lb(w_{ii}, \theta_i)$ must be less than the minimum synaptic input that it can receive and $rb(w_{ii}, \theta_i)$ must be greater than the maximum synaptic input that it can receive). In this case, the nullsurface of each neuron consists of three distinct $(N-1)$-dimensional manifolds (see Fig. 8D for the $N = 3$ case) that intersect a total of 3^N times.

A particular family of N-neuron circuits with 3^N equilibria can be designed as follows. Suppose that the weight matrix takes the highly symmetric form $w_{ii} = w_s$ (with $w_s > 4$ so that a fold exists) and that $w_{ij,j\neq i} = 1$, $\theta_i = -(w_s + N - 1)/2$, $I_i = 0$, and $\tau_i = 1$. Given these constraints, the synaptic input received by each neuron is in the range $(0, N-1)$ and each neuron's fold is exactly centered over this range. Thus, the sufficient condition mentioned earlier can be stated as $d(w_s) > N - 1$. When this condition is satisfied, the network will exhibit 3^N equilibria. The central equilibrium point will be unstable, 2^N equilibria will be stable, and the remaining $3^N - 2^N - 1$ equilibria will be saddle points (see Fig. 9A for the $N = 3$ case).

4.3 Bifurcations

The number and types of bifurcations that can occur in larger networks also defies enumeration. Consequently, I will present no detailed bifurcation maps of the parameter space of larger circuits here. However, it is worth illustrating briefly how

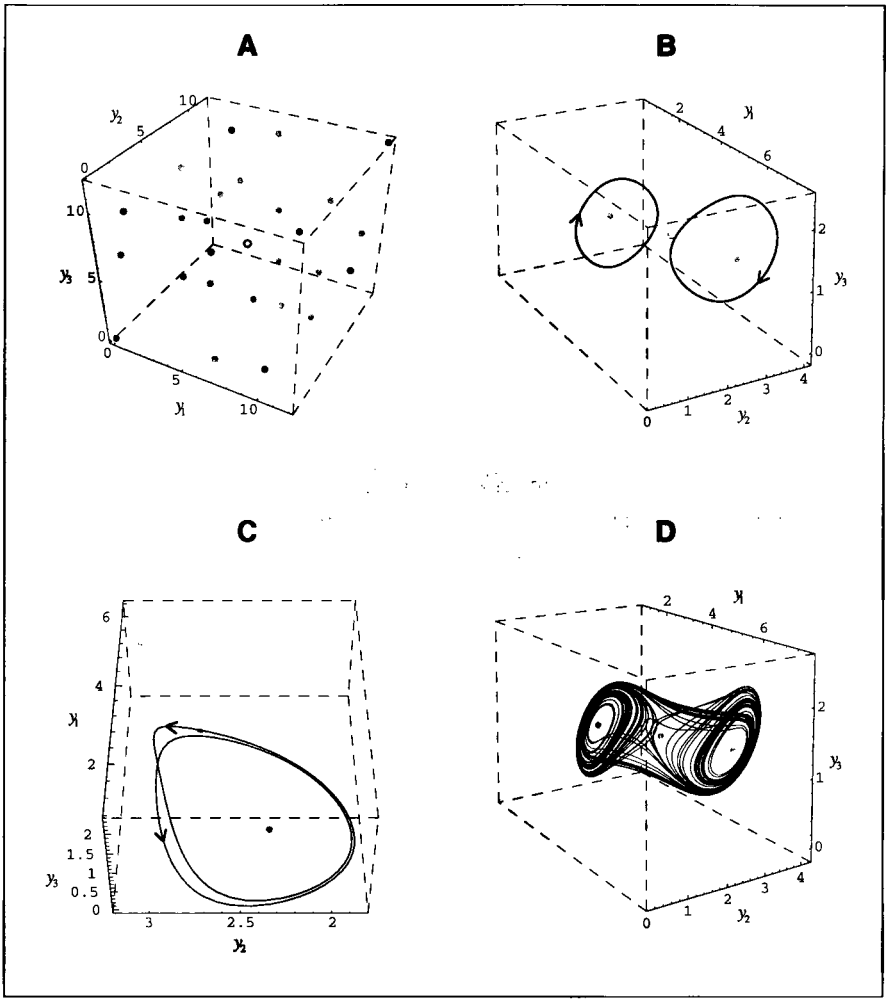


Figure 9
Example phase portraits of three-neuron circuits. The shading conventions for equilibrium points are identical to those of Figure 4. Stable and unstable manifolds of saddle points are not shown here because some of them are two-dimentional. Only stable nonequilibrium limit sets are shown. (A) A circuit with 27 equilibrium points. Here $w_{ii} = 10$, $w_{ij,j\neq i} = 1$, $\theta_i = -6$, $\tau_i = 1$. (B) A circuit with three saddle points and two stable limit cycles. Here $w_{11} = 5.422$, $w_{12} = -0.24$, $w_{13} = 0.535$, $w_{21} = -0.018$, $w_{22} = 4.59$, $w_{23} = -2.25$, $w_{31} = 2.75$, $w_{32} = 1.21$, $w_{33} = 3.885$, $\theta_1 = -4.108$, $\theta_2 = -2.787$, $\theta_3 = -1.114$, $\tau_1 = \tau_2 = \tau_3 = 1$. (C) A circuit with a doubly periodic limit cycle derived from (B) by setting $\tau_2 = 1.92$. (D) A circuit with chaotic dynamics derived from (B) by setting $\tau_2 = 2.5$.

the approach presented in section 3.3.1 to analyzing the central equilibrium point of a two-neuron center-crossing network can be extended to larger circuits in certain special cases.

A general center-crossing network is defined by the restriction that each neuron's activation function is exactly centered over the range of net input that it receives:

$$\theta_i^* = -\frac{I_i + \sum_{j=1}^{N} w_{ji}}{2} \tag{12}$$

The Jacobian matrix evaluated at an equilibrium point $\bar{\mathbf{y}}$ of an N-neuron circuit has the form:

$$D\mathbf{f}(\bar{\mathbf{y}}) = \begin{bmatrix} \frac{w_{11}\sigma'(\bar{y}_1+\theta_1)-1}{\tau_1} & \cdots & \frac{w_{N1}\sigma'(\bar{y}_N+\theta_N)}{\tau_1} \\ \vdots & \ddots & \vdots \\ \frac{w_{1N}\sigma'(\bar{y}_1+\theta_1)}{\tau_N} & \cdots & \frac{w_{NN}\sigma'(\bar{y}_N+\theta_N)-1}{\tau_N} \end{bmatrix}$$

For the central equilibrium point $\bar{\mathbf{y}}^* = (-\theta_1^*, \cdots, -\theta_N^*)$ of a center-crossing network with $\tau_i = 1$, this Jacobian matrix becomes

$$D\mathbf{f}(\bar{\mathbf{y}}^*) = \begin{bmatrix} \frac{w_{11}}{4} - 1 & \cdots & \frac{w_{N1}}{4} \\ \vdots & \ddots & \vdots \\ \frac{w_{1N}}{4} & \cdots & \frac{w_{NN}}{4} - 1 \end{bmatrix}$$

Even for a center-crossing network, it is only possible to analyze fully the stability of $\bar{\mathbf{y}}^*$ in highly symmetric cases. We will consider the two simplest cases here.

First, suppose that $w_{ii} = w_s$ and $w_{ij,j\neq i} = w_c$. Then it can be shown that the N eigenvalues of $D\mathbf{f}(\bar{\mathbf{y}}^*)$ are:

$$\lambda_1^*, \ldots, \lambda_N^* = \frac{w_s + (N-1)w_c - 4}{4}, \frac{w_s - w_c - 4}{4}, \ldots, \frac{w_s - w_c - 4}{4}$$

Note that, because of the high degree of symmetry, only two of the eigenvalues of $D\mathbf{f}(\bar{\mathbf{y}}^*)$ are distinct. Because all the eigenvalues are real, $\bar{\mathbf{y}}^*$ is always a node. When $w_c > 0$, $\bar{\mathbf{y}}^*$ will be stable when $w_s < 4 - (N-1)w_c$, a saddle with a one-dimensional unstable manifold and an $(N-1)$-dimensional stable manifold when $4 - (N-1)w_c < w_s < w_c + 4$, and unstable when $w_s > w_c + 4$. On the other hand, when $w_c < 0$, $\bar{\mathbf{y}}^*$ will be stable when $w_s < w_c + 4$, a saddle with an $(N-1)$-dimensional unstable manifold and a one-dimensional stable manifold when $w_c + 4 < w_s < 4 - (N-1)w_c$, and unstable when $w_s > 4 - (N-1)w_c$.

Now suppose that $w_{ii} = w_s$ and $w_{ij,i\neq j} = w_c$ but $w_{ji,i\neq j} = -w_c$. Then it can be shown that the eigenvalues of $D\mathbf{f}(\bar{\mathbf{y}}^*)$ have the following form:

$$N \text{ odd:} \quad \lambda_1^*, \ldots, \lambda_N^* = \frac{w_s - 4}{4}, \frac{w_s - 4}{4} \pm C_1 i, \cdots, \frac{w_s - 4}{4} \pm C_{\frac{N-1}{2}} i$$

$$N \text{ even:} \quad \lambda_1^*, \ldots, \lambda_N^* = \frac{w_s - 4}{4} \pm C_1 i, \cdots, \frac{w_s - 4}{4} \pm C_{\frac{N}{2}} i$$

where the coefficients C_k are (not necessarily distinct) functions of w_c. Because $\mathrm{Re}[\lambda_i^*] = (w_s - 4)/4$, $\bar{\mathbf{y}}^*$ will be stable when $w_s < 4$ and unstable when $w_s > 4$. In addition, because all of the complex conjugate pairs of eigenvalues will be purely imaginary at $w_s = 4$, $\bar{\mathbf{y}}^*$ will undergo a Hopf bifurcation there. Indeed, as multiple eigenvalues are purely imaginary at $w_s = 4$ (and the single real eigenvalue is also zero at this point when N is odd), this Hopf bifurcation is highly degenerate in larger networks, raising the possibility of complex dynamics (e.g., invariant tori and chaos) nearby in parameter space (Guckenheimer & Holmes, 1983).

5 Nonautonomous Circuits

Thus far, we have considered only the dynamics of autonomous circuits (i.e., circuits whose inputs are fixed to constant values). However, we must also consider circuits with time-varying inputs (i.e., those in which $I_i \mapsto I_i(t)$ in Equation 1). Such nonautonomous circuits are important for at least two reasons. First, any agent will have sensors that provide time-varying signals from its environment. Second, when trying to understand the operation of larger networks, it is sometimes useful to decompose them into sets of smaller circuits that provide time-varying inputs to one another. Unfortunately, there is little that can be said in general about the response of nonlinear dynamical systems to arbitrary time-varying inputs. However, in this section, I will illustrate one simple strategy for achieving at least some qualitative understanding of the dynamics of CTRNNs with time-varying inputs. I will also present one example of the sort of new phenomena that can arise when time-varying inputs are introduced.

5.1 Time-varying inputs

A simple strategy for understanding the behavior of nonautonomous CTRNNs is to think of the time-varying inputs to a circuit as parameters that take on fixed values at any given point in time. Thus, the dynamics of a nonautonomous circuit can be instantaneously decomposed into a sequence of autonomous dynamics. At any given instant, the state of the nonautonomous circuit will be moving toward the attractor in whose basin it finds itself along a trajectory of the corresponding autonomous circuit. Of course, because the inputs are changing in time, the instantaneous autonomous dynamics will also be changing, with limit sets and boundaries between distinct basins of attraction deforming and moving around, possibly even undergoing bifurcations.

When the inputs change relatively slowly compared to the timescale of the autonomous dynamics, then the system state will always be found near an attractor of

these dynamics. In the special case where this attractor is an equilibrium point and the inputs are changing very slowly, the relationship between the nonautonomous system and the corresponding parameterized autonomous system can be given a precise analytical formulation using so-called singular perturbation or quasistatic methods (Hoppensteadt, 1993). However, even when the attractors are not equilibrium points or the inputs change very quickly (causing the system state to lag behind the deforming and moving attractors by a considerable amount), it is often still useful to attempt to relate the behavior of a nonautonomous system to the changing structure of the underlying autonomous dynamics. Thus, by studying how the autonomous dynamics of a circuit vary as a function of its input parameters and then decomposing a trajectory of the nonautonomous circuit in these terms, we can sometimes gain significant qualitative insight into the response of a nonautonomous circuit to a given family of input signals.

As a very simple concrete example of this strategy, consider a single neuron with $w = 5.5$ and $\theta = -2.75$. We know from our analysis in section 2 that the steady-state I/γ relation of the neuron will exhibit a fold of width $d(5.5) \approx 0.555$ centered over $I = 0$. Now suppose that we sinusoidally drive this neuron back and forth across its fold. If $I(t)$ varies sufficiently slowly, then the state of the neuron will always track the location of one of the two autonomous equilibrium points, switching between the upper and lower attractors whenever the attractor it is currently following disappears in a saddle-node bifurcation as $I(t)$ crosses the edges of the fold at $I = lb(5.5, -2.75) \approx -0.277$ and $I = rb(5.5, -2.75) \approx 0.227$. This gives rise to a classic hysteresis diagram (Fig. 10A). The instantaneous autonomous phase portraits of the neuron at selected points along this hysteresis curve are shown in Figure 10B.

When $I(t)$ varies more quickly, the state of the neuron always lags behind the moving attractors, but we can still understand the resulting trajectory qualitatively in terms of the changing autonomous dynamics (Fig. 10C, D). For example, in Figure 10D, $I(t)$ is decreasing so quickly after it reaches its peak that the neuron state never has a chance to move very far toward the upper attractor before it crosses the unstable equilibrium point and falls back toward the lower attractor.

Because nothing in this decomposition strategy depends in any essential way on the periodicity of the input, we can also use it to understand the response of the neuron to nonperiodic signals. For example, the relationship between the magnitude and duration of input pulse necessary to switch the neuron between its upper and lower attractors can be directly related to the location of the unstable equilibrium point that separates the two stable equilibria. Examples of the use of this strategy to understand the response of larger nonautonomous CTRNNs with more complicated autonomous dynamics to more complex input signals can be found in Gallagher and Beer (1993), Yamauchi and Beer (1994) and Beer (1995a,b).

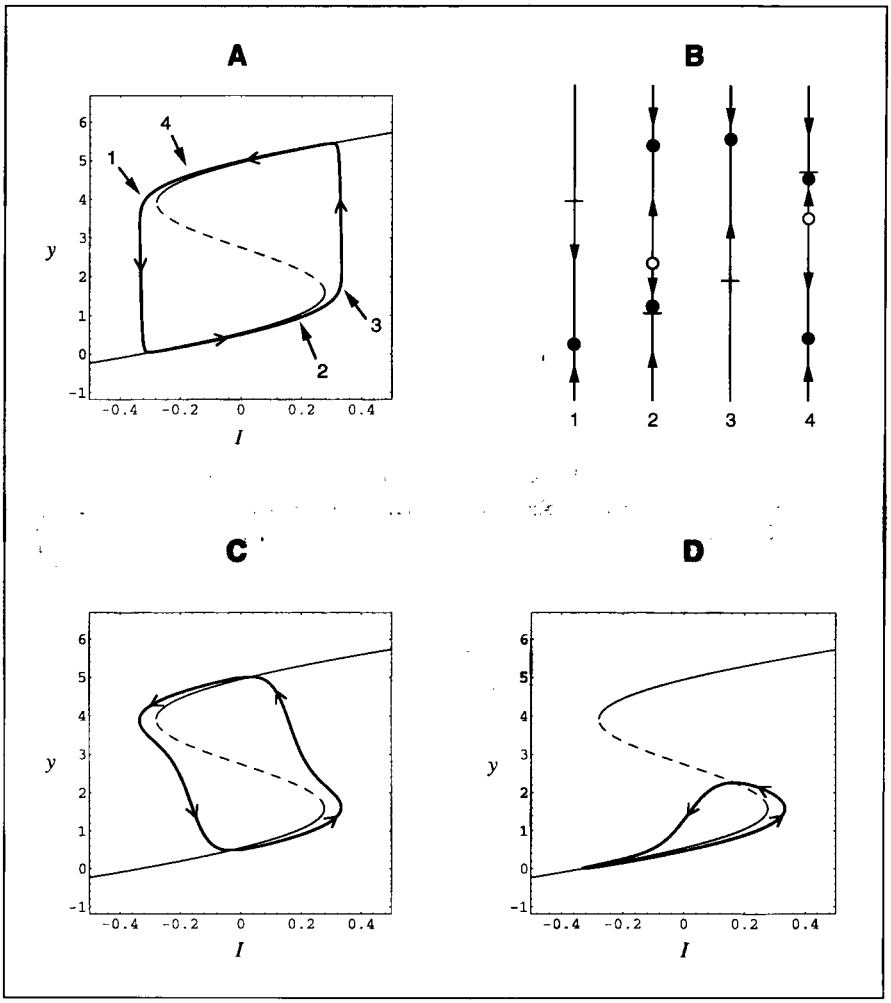


Figure 10
Hysteresis in a single nonautonomous neuron ($w = 5.5, \theta = -2.75$) driven by an external input of the form $I(t) = A\sin(\omega t)$. Here the amplitude A was set to 0.33 so that the extrema of the sine wave extend just beyond the edges of the folds in the neuron's I/y relation. (A) $\omega = 0.033$. The system's trajectory (thick line) is shown superimposed on the neuron's steady-state I/y relation. (B) The instantaneous autonomous dynamics of the neuron at the four numbered points shown in A, with a short horizontal line marking the instantaneous state of the neuron at each point. (C) $\omega = 0.067$. (D) $\omega = 0.071$.

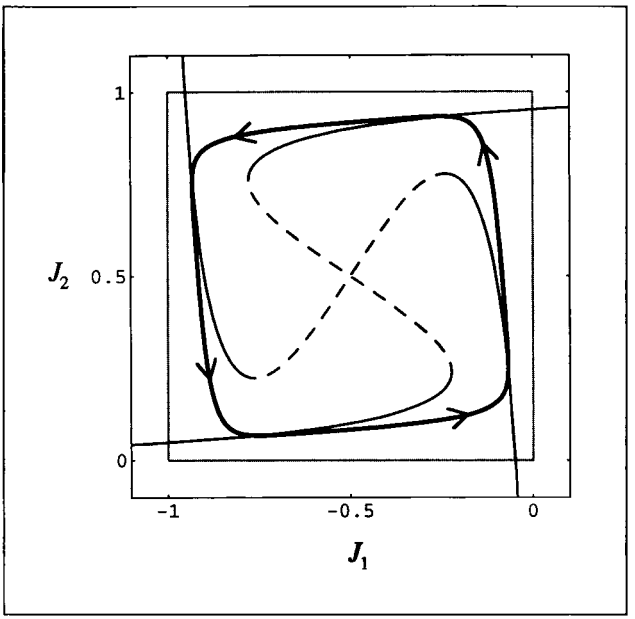


Figure 11
Limit cycles can sometimes be understood using the same approach as that illustrated in Figure 10 for hysteresis. Here two copies of the neuron examined in Figure 10 are reciprocally interconnected in a center-crossing circuit, with $w_{11} = w_{22} = 5.5$, $w_{12} = 1$, $w_{21} = -1$, $\theta_1 = -2.25$, $\theta_2 = -3.25$. The system trajectory (thick line) is shown superimposed on the I/γ relations of the individual neurons in synaptic input space.

5.2 Decomposing larger circuits

This same general strategy can sometimes also be used to decompose the dynamics of a larger autonomous circuit into smaller interacting subcircuits. As a simple example of this type of application, we will briefly examine how a limit cycle arises from a reciprocal interaction between the autonomous dynamics of two neurons. As usual for two-neuron circuits, it will be most convenient to work in the synaptic input space (J_1, J_2) (Fig. 11). Here the horizontal motion of the system is governed by the N-shaped nullcline (the $\dot{J}_2$ nullcline) and the vertical motion is governed by the S-shaped nullcline (the $\dot{J}_1$ nullcline).

Now notice how the two nullclines are arranged. When the state of the system is below the lower branch of the S-shaped nullcline, the N-shaped nullcline pulls it to the right because this is the only stable branch of the I/γ relation of neuron 1. As the state moves toward the right, it "falls off" the edge of the right fold of the S-shaped nullcline (because the synaptic input from neuron 1 takes neuron 2 through a saddle-node bifurcation that destroys this equilibrium point) and begins moving upward toward its upper stable branch. This upward movement eventually causes the state to fall off the edge of the upper fold in the N-shaped nullcline and move toward its left stable branch, which in turn causes the state to fall off the left edge of the fold of the S-shaped nullcline and move toward its lower branch. Finally,

this downward motion eventually causes the state to fall off the lower edge of the fold in the N-shaped nullcline, after which the entire process repeats.

Thus, we can see how a reciprocal interaction between the autonomous dynamics of the individual neurons (which is fully described by the single-neuron theory) generates the observed limit cycle by a mechanism very similar to the hysteresis illustrated in the previous section. In this relatively simple case, this decomposition can be given a more precise analytical formulation using relaxation oscillator techniques (Arnol'd, 1994). However, if a qualitative understanding is sufficient, then the basic idea sketched here can be applied more generally.

5.3 Complex dynamics in nonautonomous circuits

When time-varying inputs to CTRNNs are allowed, qualitatively new dynamical phenomena can arise. For example, though it is well-known that smooth two-dimensional autonomous dynamical systems cannot exhibit chaotic dynamics (Hale & Koçak, 1991), this limitation no longer exists in the presence of time-varying input signals. One scenario for chaotic dynamics involves periodically forcing a planar dynamical system that possesses a homoclinic loop (Wiggins, 1990). To illustrate this behavior in the context of CTRNNs, consider the example *3lc* phase portrait shown in Figure 4. Although this phase portrait does not exhibit a homoclinic orbit, a small variation of it does. Periodically driving such a circuit back and forth across this homoclinic loop using a sinusoidal external input of the proper amplitude and frequency produces the chaotic dynamics shown in Figure 12. The largest Lyapunov exponent of this chaotic attractor is 0.023, and its Lyapunov dimension is 1.157 (Parker & Chua, 1989).

6 Discussion

Dynamical neural networks are being used increasingly as simple model nervous systems in autonomous agents research. As their use grows, a thorough understanding of the dynamical capabilities of such networks will become essential. In this article, I have illustrated how the mathematical tools of dynamical systems theory can be used to gain significant insight into the operation of small continuous-time recurrent neural networks. Using a combination of elementary analysis and numerical studies, I have given a fairly complete description of the possible dynamical behavior and bifurcations of one- and two-neuron circuits, along with a few specific results for larger networks. These results provide both qualitative insight and, in many cases, quantitative formulas for predicting the dynamical behavior of particular circuits and how that behavior changes as network parameters are varied. The synaptic input diagrams described in section 3.1 are especially useful for understanding the dynamics

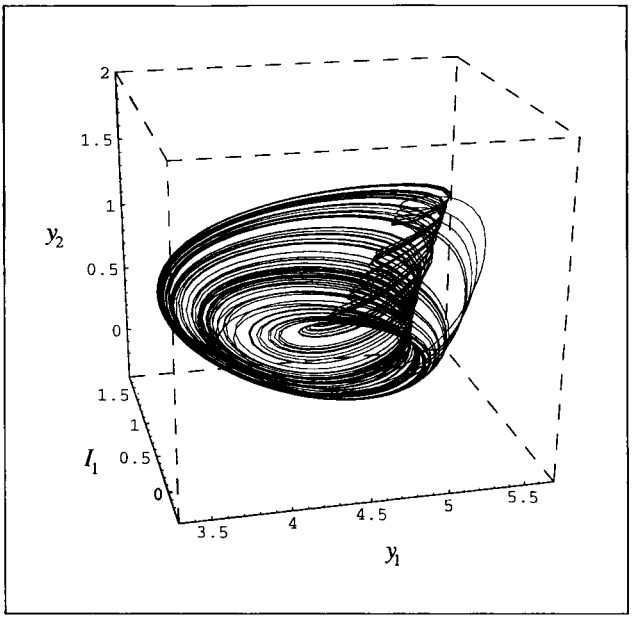


Figure 12
Chaotic dynamics in a sinusoidally driven two-neuron circuit possessing a homoclinic loop. This circuit is very close to the *3lc* example shown in Figure 4 except that here $\theta_1 = -3.23596$. Neuron 1 is driven with a time-varying external input of the form $I_1(t) = 0.25\sin(t/2.45) + 0.06205$.

of two-neuron circuits. In addition, I have illustrated one simple method for gaining a qualitative understanding of CTRNNs with time-varying inputs and presented an example of the complicated dynamics that can arise in such cases.

The analysis described in this article demonstrates that even small CTRNNs are capable of complicated dynamical behavior. Indeed, it appears that an N-neuron CTRNN can exhibit all the qualitatively different sorts of behavior that are possible in smooth N-dimensional dynamical systems. Although this certainly makes their analysis more difficult, this demonstration is important because, although Funahashi and Nakamura (1993) proved that *sufficiently large* CTRNNs can approximate the dynamics of arbitrary smooth dynamical systems for finite time, their result tells us little about the capabilities of the relatively small CTRNNs that are currently evolved in autonomous agent research. In addition, the fact that even small and simple dynamical neural networks can exhibit dynamics of significant complexity may have important neurobiological implications.

What implications do these results have for the researcher faced with understanding a particular CTRNN (e.g., an evolved neural controller for an autonomous agent)? If one wishes to understand the dynamics of particular single neurons or two-neuron circuits, then the results presented in this article are obviously directly applicable. Of course, it will be more typical that one is presented with the problem of understanding

the dynamics of larger networks. In this case, the results presented here should be viewed as illustrating how the mathematical tools of dynamical systems theory can be applied to particular CTRNNs. In addition, one important question in tackling the dynamics of larger systems is whether each individual neuron in the network has a fold in its steady-state I/γ relation and, if so, where that fold lies relative to the range of synaptic inputs that neuron receives. This is exactly the kind of information provided by the single-neuron theory presented in section 2.

The results presented here can also be viewed as providing a first step toward a more systematic, comprehensive theory of the dynamics of CTRNNs. It is likely that significant further progress can be made on small to moderate-sized CTRNNs. In addition, if a larger network has a natural decomposition into subcircuits of one or two neurons, then the one- and two-neuron theories combined with the strategy presented in section 5 may provide some insight into the dynamics of the larger circuit. Finally, progress can sometimes be made if a circuit is highly symmetric (see also Collins & Stewart, 1994, for examples of the use of group theoretical techniques to analyze the dynamics of symmetric circuits).

What are the implications of the results presented here for attempts to synthesize CTRNNs with desired properties? Once again, if the desired dynamical behavior can be generated by one- or two-neuron circuits, then the results presented here are directly applicable. Furthermore, two-neuron circuits may serve as powerful building blocks for the modular construction of larger networks with desired dynamics. Thus, an evolutionary algorithm could evolve larger networks by evolving the coupling between a set of canonical two-neuron circuits exhibiting each of the eleven distinct phase portraits shown in Figure 4. As our understanding of three- and four-neuron CTRNNs improves, this canonical set could be expanded. Given the dynamical richness of coupled oscillators, another possibility would be to start with canonical two-neuron circuits exhibiting oscillations of different frequencies and amplitudes and evolving their coupling to achieve some overall dynamics of interest.

More generally, the results of the analysis presented here can be used to focus an evolutionary search into regions of parameter space most likely to yield desired behavior. From the structure of an individual neuron's steady-state I/γ relation, it is clear that unless a neuron has a fold that intersects the range of synaptic inputs that it receives, its input-output behavior will be monotonic. Indeed, the overwhelmingly most likely phase portrait of a randomly generated CTRNN is a single stable equilibrium point. Even in two-neuron circuits, all the other phase portraits shown in Figure 4 require the coordination of the self-weights, the cross-weights, and the biases (note the way that most of the interesting behavior that a two-neuron circuit can exhibit occurs in a neighborhood of the center of the bifurcation maps in Figures 5 and 7, which provides a graphical illustration of this point). In larger networks,

this coordination only becomes more difficult to achieve and therefore less likely to be found by a stochastic search technique. For this reason, it would be extremely fruitful to focus the search into regions of parameter space where most, if not all, of the qualitatively different dynamics that a network can possibly exhibit are relatively nearby. Such regions are called *organizing centers* (Golubitsky & Schaefer, 1985).

One possibility would be to search in the neighborhood of center-crossing networks as specified by Equation 12. This can be accomplished in one of two ways: (1) Seed an initial population with small perturbations of randomly generated center-crossing networks and then allow the search to proceed normally (i.e., search the parameter vector $\mathbf{W}|\boldsymbol{\theta}|\boldsymbol{\tau}$). (2) "Anchor" the search on center-crossing networks by searching not the biases themselves but rather perturbations from the center-crossing values (i.e., search the parameter vector $\mathbf{W}|\Delta\boldsymbol{\theta}^*|\boldsymbol{\tau}$, where $\boldsymbol{\theta}^*$ depends on $\mathbf{W}$ as described by Equation 12). Preliminary experiments using the latter approach to evolve oscillatory circuits suggest that significant improvement over an unbiased and unconstrained search can be obtained, both in terms of the overall success of the search and in terms of the number of generations required to achieve a given level of performance. However, further experiments are required to determine the general utility of this approach.

Another possibility would be to focus on networks with the following structure: $\mathbf{W}^*$: $w_{ii} = 4$, $w_{ij,j\neq i} = 0$, $\boldsymbol{\theta}^*$: $\theta_i = -2$, $\boldsymbol{\tau}^*$: $\tau_i = 1$. Such networks exhibit a single equilibrium point $\bar{\mathbf{y}}^* = (2, \ldots, 2)$. Because this equilibrium point is highly degenerate (the N eigenvalues of $D\mathbf{f}(\bar{\mathbf{y}}^*)$ are all 0 and critically real), relatively small changes in the weight and bias parameters will cause this phase portrait to bifurcate into more complex dynamical behavior. Indeed, one would expect many of the qualitatively different dynamics that a network can exhibit to occur in the neighborhood of this point in parameter space. Once again, we can either seed an initial population with random perturbations of these highly degenerate networks and allow the search to proceed normally or search perturbations from the degenerate values (i.e., search the parameter vector $\Delta\mathbf{W}^*|\Delta\boldsymbol{\theta}^*|\Delta\boldsymbol{\tau}^*$).

To date, attempts to evolve dynamical neural controllers for autonomous agents have been exclusively empirical in nature. It is common to place the entire burden of finding circuits with useful dynamics on an evolutionary algorithm, with the result that these algorithms often fail on even simple tasks unless the fitness function, parameter encoding, population size, mutation rate, and so on are carefully chosen, and they do not scale well to more difficult tasks. It is also common to treat the networks that do evolve as mysterious black boxes. However, this article has shown how the mathematical tools of dynamical systems theory can be used to gain significant insight into the dynamics of CTRNNs. Although fairly exhaustive analysis of the sort presented here for one- and two-neuron circuits is not generally possible

for larger networks, these and other techniques can be applied to particular circuits or highly symmetric classes of circuits. I have also suggested ways in which the sort of analysis presented here can be used to focus evolutionary searches into fruitful regions of parameter space and thereby improve the performance and yield of such searches.

Acknowledgments

I am extremely grateful for many helpful discussions with Ken Loparo, Nessan Fitzmaurice, Hillel Chiel, and Alan Calvitti, all of whom also read and commented on an earlier draft of this article. This work was supported in part by grant no. N00014-90-J-1545 from the Office of Naval Research.

References

Abraham, R. H., & Shaw, C. D. (1992). *Dynamics—the geometry of behavior.* Redwood City, CA: Addison-Wesley.

Arnol'd, V. I. (Ed.). (1994). *Dynamical systems V: Bifurcation theory and catastrophe theory.* New York: Springer-Verlag.

Atiya, A., & Baldi, P. (1989). Oscillations and synchronizations in neural networks: An exploration of the labeling hypothesis. *International Journal of Neural Systems, 1*(2), 103–124.

Beer, R. D. (1990). *Intelligence as adaptive behavior: An experiment in computational neuroethology.* San Diego: Academic Press.

Beer, R. D. (1995a). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence, 72,* 173–215.

Beer, R. D. (1995b). Computational and dynamical languages for autonomous agents. In R. Port and T. van Gelder (Eds.), *Mind as motion: Explorations in the dynamics of cognition.* Cambridge, MA: MIT Press.

Beer, R. D., & Gallagher, J. C. (1992). Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior, 1,* 91–122.

Blum, E. K., & Wang, X. (1992). Stability of fixed points and periodic orbits and bifurcations in analog neural networks. *Neural Networks, 5,* 577–587.

Cliff, D., Harvey, I., & Husbands, P. (1993). Explorations in evolutionary robotics. *Adaptive Behavior, 2*(1), 73–110.

Cliff, D., Husbands, P., & Harvey, I. (1993). Analysis of evolved sensory-motor controllers. In *Proceedings of the Second European Conference on Artificial Life.*

Cohen, M. A., & Grossberg, S. (1983). Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. *IEEE Transactions on Systems, Man and Cybernetics, 13,* 813–825.

Collins, J. J., & Stewart, I. (1994). A group-theoretic approach to rings of coupled biological oscillators. *Biological Cybernetics, 71,* 95–103.

Collins, R. J., and Jefferson, D. R. (1991). Representations for artificial organisms. In J.-A. Meyer, H. Roitblat, & S. Wilson (Eds.), *From animals to animats 2: Pro-*

ceedings of the Second International Conference on the Simulation of Adaptive Behavior. Cambridge, MA: MIT Press.

Cowan, J. D., & Ermentrout, G. B. (1978). Some aspects of the eigenbehavior of neural nets. In S. A. Levin (Ed.), *Studies in mathematical biology 1: Cellular behavior and the development of pattern.* Washington, DC: The Mathematical Association of America.

Das, P. K., Schieve, W. C. & Zeng, Z. (1991). Chaos in an effective four-neuron neural network. *Physics Letters* [A], *161*, 60–66.

de Garis, H. (1992). Steerable GenNets: The genetic programming of steerable behaviors in GenNets. In F. J. Varela and P. Bourgine (Eds.), *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life.* Cambridge, MA: MIT Press.

Erlanson, R., & Abu-Mostafa, Y. (1991). Analog neural networks as decoders. In R. P. Lippman, J. E. Moody, & D. S. Touretzky (Eds.), *Advances in neural information processing systems 3.* San Mateo, CA: Morgan Kaufmann.

Ermentrout, B. (1995). Phase plane analysis of neural activity. In M.A. Arbib (Ed.), *The handbook of brain theory and neural networks.* Cambridge, MA: MIT Press.

Floreano, D., & Mondada, F. (1994). Automatic creation of an autonomous agent: Genetic evolution of a neural-network driven robot. In D. Cliff, P. Husbands, J. Meyer, & S. Wilson (Eds.), *From animals to animats 3: Proceedings of the Third International Conference on the Simulation of Adaptive Behavior.* Cambridge, MA: MIT Press.

Funahashi, K., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks, 6,* 801–806.

Gallagher, J. C., & Beer, R. D. (1993). A qualitative dynamical analysis of evolved locomotion controllers. In J.-A. Meyer, H. Roitblat, & S. Wilson (Eds.), *From animals to animats 2: Proceedings of the Second International Conference on the Simulation of Adaptive Behavior.* Cambridge, MA: MIT Press.

Golubitsky, M., & Schaefer, D. G. (1985). *Singularities and groups in bifurcation theory,* Vol. 1. New York: Springer-Verlag.

Goudreau, M. W., & Giles, C. L. (1992). Neural network routing for random multistage interconnection networks. In J. E. Moody, S. J. Hanson & R. P. Lippmann (Eds.), *Advances in neural information processing systems 4.* San Mateo, CA: Morgan Kaufmann.

Grossberg, S. (1969). On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics, 1,* 319–350.

Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms and architectures. *Neural Networks, 1,* 17–61.

Guckenheimer, J., & Holmes, P. (1983). *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields.* New York: Springer-Verlag.

Hale, J. K., & Koçak, H. (1991). *Dynamics and bifurcations.* New York: Springer-Verlag.

Hirsch, M. (1989). Convergent activation dynamics in continuous time networks. *Neural Networks, 2,* 331–349.

Hopfield, J. J. (1984). Neurons with graded response properties have collective com-

putational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences, 81*, 3088–3092.

Hopfield, J. J., & Tank, D. W. (1985). Neural computation of decisions in optimization problems. *Biological Cybernetics, 52*, 141–152.

Hoppensteadt, F. C. (1993). *Analysis and simulation of chaotic systems*. New York: Springer-Verlag.

Husbands, P., Harvey, I., & Cliff, D. (in press). Circle in the round: State space attractors for evolved sighted robots. *Robotics and Autonomous Systems*.

Lockery, S. R., Fang, Y., & Sejnowski, T. J. (1990). A dynamic neural network model of sensorimotor transformations in the leech. *Neural Computation, 2*(3), 274–282.

Miller, G. F., & Cliff, D. (1994). Protean behavior in dynamic games: Arguments for the co-evolution of pursuit-evasion tactics. In D. Cliff, P. Husbands, J. Meyer, & S. Wilson (Eds.), *From animals to animats 3: Proceedings of the Third International Conference on the Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press.

Parker, T. S., & Chua, L. O. (1989). *Practical numerical algorithms for chaotic systems*. New York: Springer-Verlag.

Pearlmutter, B. A. (1990). *Dynamic recurrent neural networks* (Tech. Rep. CMU-CS-90-196). Pittsburgh: Carnegie Mellon University, School of Computer Science.

Schieve, W. C., Bulsara, A. R., & Davis, G. M. (1991). Single effective neuron. *Physical Review* [A], *43*(6), 2613–2623.

Simard, P., & Cun, Y. L. (1992). Reverse TDNN: An architecture for trajectory generation. In J. E. Moody, S. J. Hanson, & R. P. Lippman (Eds.), *Advances in neural information processing systems 4*. San Mateo, CA: Morgan Kaufmann.

Sompolinsky, H., & Crisanti, A. (1988). Chaos in random neural networks. *Physical Review Letters, 61*(3), 259–262.

Spiessens, P., & Torreele, J. (1992). Massively parallel evolution of recurrent networks: An approach to temporal processing. In F. J. Varela & P. Bourgine (Eds.), *Toward a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life*. Cambridge, MA: MIT Press.

Strogatz, S. H. (1994). *Nonlinear dynamics and chaos*. Reading, MA: Addison-Wesley.

Tino, P., Horne, B. G., & Giles, C. L. (1995). Fixed points in two-neuron discrete time recurrent networks: Stability and bifurcation considerations (Tech. Rep. UMIACS-TR-95-51 & CS-TR-3461). College Park, MD: University of Maryland, Institute for Advanced Computer Studies.

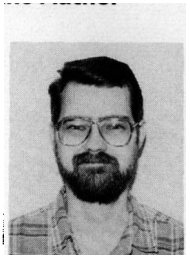
Werner, G. M., & Dyer, M. G. (1991). Evolution of communication in artificial organisms. In C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen (Eds.), *Artificial life II*. Reading, MA: Addison-Wesley.

Wiggins, S. (1990). *Introduction to applied nonlinear dynamical systems and chaos*. New York: Springer-Verlag.

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal, 12*, 1–24.

Yamauchi, B., & Beer, R. D. (1994). Sequential behavior and learning in evolved dynamical neural networks. *Adaptive Behavior, 2*(3), 219–246.

bout the Author



Randall D. Beer

Randall D. Beer received his PhD in computer science from Case Western Reserve University in 1989. He is currently an associate professor at CWRU in the department of computer engineering and science, with a joint appointment in the department of biology. He is interested in the evolution and analysis of dynamical neural networks for autonomous agents, biologically inspired robotics, neuroethological modeling, and theoretical issues in adaptive behavior.