In: beim Graben, P., Zhou, C., Thiel, M. & Kurths, J.(Eds.)
Lectures in Supercomputational Neuroscience:
Dynamics in Complex Brain Networks. Berlin: Springer, 2008.

7

Neural and Cognitive Modeling with Networks of Leaky Integrator Units

Peter beim Graben[1,3], Thomas Liebscher[2] and Jürgen Kurths[3]

[1] School of Psychology and Clinical Language Sciences,
University of Reading, United Kingdom
p.r.beimgraben@reading.ac.uk
[2] Bundeskriminalamt, Wiesbaden, Germany
[3] Institute of Physics, Nonlinear Dynamics Group, Universität Potsdam, Germany

Summary. After reviewing several physiological findings on oscillations in the electroencephalogram (EEG) and their possible explanations by dynamical modeling, we present neural networks consisting of leaky integrator units as a universal paradigm for neural and cognitive modeling. In contrast to standard recurrent neural networks, leaky integrator units are described by ordinary differential equations living in continuous time. We present an algorithm to train the temporal behavior of leaky integrator networks by generalized back-propagation and discuss their physiological relevance. Eventually, we show how leaky integrator units can be used to build oscillators that may serve as models of brain oscillations and cognitive processes.

7.1 Introduction

The electroencephalogram (EEG) measures the electric fields of the brain generated by large formations of certain neurons, the *pyramidal cells*. These nerve cells roughly possess an axial symmetry and they are aligned in parallel perpendicular to the surface of the cortex [1–4]. They receive excitatory input at the superficial apical dendrites from thalamic relay neurons and inhibitory input at the basal dendrites and at their somata from local interneurons [1,3–5]. Excitatory and inhibitory synapses cause different ion currents through the cell membranes thus leading to either depolarization or hyperpolarization, respectively. When these synapses are activated, a single pyramidal cell behaves as a microscopic electric dipole surrounded by its characteristic electric field [1,6].

According to the inhomogeneity of the cortical gray matter, a mass of approximately 10,000 synchronized pyramidal cells form a dipole layer whose fields sum up to the *local field potentials* that polarize the outer tissues of the scalp, which acts thereby as a low pass filter [1, 3, 5, 6]. These filtered sum potentials are macroscopically measurable as the EEG at the surface of a subject's head (cf. Chap. 1).

Some of the most obvious features of the EEG are oscillations in certain frequency bands. The *alpha waves* are sinusoidal-like oscillations between 8–14 Hz, strongly pronounced over parietal and occipital recording sites that reflect a state of relaxation during wakefulness, with no or only low visual attention. Figure 7.1 shows a characteristic power spectrum for the alpha rhythm: There is one distinguished peak superimposed to the $1/f$ background EEG. When a subject starts paying attention, the powerful slow alpha waves disappear, while smaller oscillations with higher frequencies around 14–30 Hz (the *beta waves*) arise [2, 7, 8]. We will refer to this finding, sometimes called *desynchronization* of the EEG, as to the *alpha blocking* [7].[4] Alpha waves are assumed to be related to awareness and cognitive processes [11–14]. Experimental findings suggest that thalamocortical feed-back loops are involved in the origin of the alpha EEG [1, 2, 4, 8, 15, 16].

The $1/f$-behavior and the existence of distinguished oscillations in the EEG such as the alpha waves are cornerstones in the evaluation of computational models of the EEG. Indeed, modeling these brain rhythms has a long tradition. Wilson and Cowan [17] were the first to use populations of excitatory and inhibitory neurons innervating each other (see Sect. 7.3.2). They introduced a two-dimensional state vector whose components describe the proportion of firing McCulloch-Pitts neurons [18] within a unit volume of neural tissue at an instance of time. This kind of ensemble statistics leads to the

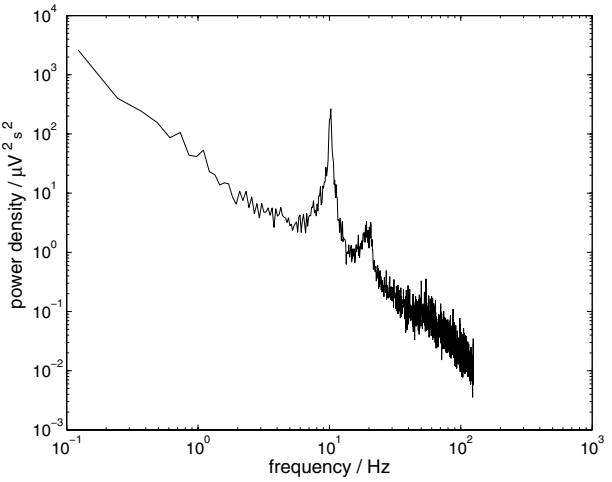


Fig. 7.1. Power spectrum of the alpha EEG at one parietal electrode site (PZ)

[4] The term "desynchronization" is misleading since it has no direct relation to synchronization in the sense of, for example, [9, 10]. From the viewpoint of data analysis it simply means: decreasing power in the alpha band of the spectrum. However, biophysical theories of the EEG explain the loss of spectral power by a loss of coherence of neuron activity, i.e. a reduction of synchronization [7, 8, 11].

well-known sigmoidal activation functions for neural networks [19] through the probability distributions of either synapses or activation thresholds (see also [5]). The model further includes the refractory time in which a neuron that has been activated just before cannot be activated again, and the time-course of postsynaptic potentials as *impulse response functions*. This model has been strongly simplified by Wilson [20], leading to a network of only two recurrently coupled *leaky integrator units* (see Sect. 7.2). Wilson reported limit cycle dynamics of this system for a certain range of the excitatory input, playing the role of a control parameter. However, this network does not exhibit an equivalent to the alpha blocking, because the frequency of the oscillations becomes slower for increasing input.

Lopez da Silva et al. [21] pursued two different approaches: a distributed model of the thalamus where relay cells and interneurons are considered individually, and a "lumped" model analogous to the one of Wilson and Cowan [17] but without refractory time and with even more complicated postsynaptic potentials. In order to determine the sum membrane potential of each population as a model EEG, one has to compute the convolution integral of the postsynaptic impulse response functions with the spike rate per unit of volume. Linearizing the activation functions allows a system-theoretic treatment of the model by means of the Laplace transform, thus allowing the analytical computation of the power spectrum. Lopez da Silva et al. [21,22] have shown that their model of thalamical or cortical feedback loops actually exhibits a peak around 10 Hz, i.e. alpha waves, in the spectrum, although they were not able to demonstrate alpha blocking. This population model [21] has been further developed by Freeman [23], Jansen et al. [24,25], Wendling et al. [26,27], and researchers from the Friston group [28–30] in order to model the EEG of the olfactory system, epileptic EEGs, and event-related potentials (ERP), respectively.

A further generalization of the Lopez da Silva et al. model [21] led Rotterdam et al. [31] to a description of spatio-temporal dynamics by considering a chain of coupled cortical oscillators. A similar approach has been pursued by Wright and Liley [32,33] who discussed a spatial lattice of coupled unit volumes of excitatory and inhibitory elements obeying cortical connectivity statistics. The convolution integrals of the postsynaptic potentials with the spike rates were substituted by convolution sums over discrete time. The most important result for us is that the power spectrum shows the alpha peak, and, that there is a "shift to the right" (towards the beta band) of this peak with increasing input describing arousal, i.e. actually alpha blocking.

Additionally, Liley et al. [34] also suggested a distributed model of cortical alpha activity using a compartmental description of membrane potentials [35]. In such an approach, nerve cells are thought to be built up of cylindrical compartments that are governed by generalized Hodgkin-Huxley equations [36] (see also Chap. 1). Liley et al. [34] reported two oscillatory regimes of this dynamics: one having a broad-band spectrum with a peak in the beta range and the other narrowly banded with a peak around the alpha frequency.

There are also field theoretic models of neural activity [37–41] (see Chap. 8). In these theories, the unit volumes of cortical tissue are considered to be infinitesimally small. Thus, the systems of coupled ordinary differential equations are substituted by nonlinear partial differential equations. Robinson et al. [41] have proposed such a theory in order to describe thalamocortical interactions and hence the alpha EEG.

Another approach that could lead to the explanation of the EEG is Hebb's concept of a *cell assembly* [42], where *reverberatory circles* form neural oscillators. We shall see in Sect. 7.4.3 how such circles may emerge in an evolving neural network.

On the other hand, Kaplan et al. [43], van der Velde and de Kamps [44], Wennekers et al. [45], and Smolensky and Legendre [46] argue how neural networks could bridge the gap between the sub-symbolic representation of single neurons and "a symbol-like unit of thought" in models of cognitive processes. Kaplan et al. proposed that the cell assembly be an assembly of neural units that are recurrently connected to exhibit reverberatory circles, in which information needs to cycle around until the symbolic meaning is fully established. They presented a series of experiments in which they made use of physiological principles that should be present in the functioning of cell assemblies: temporally structured input, dependency on prior experience, competition between assemblies and control of its activation. A main result is that after a cell assembly is provided with input, its activation gradually increases until an asymptotic activation is reached or the input is removed. After removal of the input, the activation gradually decreases until it comes back to its resting level.

7.2 Leaky Integrator Networks

7.2.1 Description of Leaky Integrator Units

When neural signals are exchanged between different cell assemblies that are typically involved in brain functions, oscillations caused by recurrent connections between the neurons should become visible. A possible way to model this behavior is by describing each cell assembly by a leaky integrator unit [47], which integrates input over time while the internal activation is continuously decreased by a dampening leakage term. We shall present the relationship between cell assemblies and leaky integrator units in Sect. 7.3.2. However, also single neurons can be described by a leaky integrator unit, though with quite different leakage constants, as we shall see in Sect. 7.3.1. In terms of standard units (as e.g. used by Rumelhard et al. [48]), a leaky integrator unit looks like the one depicted in Fig. 7.2.

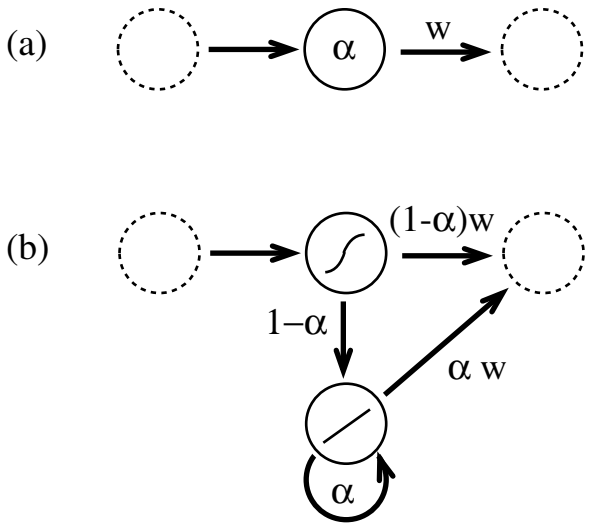


Fig. 7.2. Simulation of a leaky integrator unit (**a**) and a recurrent combination of two standard units (**b**). The function of the leakage rate α is mimicked by two parallel standard units with a logistic and a linear activation function, respectively. The synaptic weights to subsequent units are denoted by w (cf. (7.3))

The activation of this leaky integrator unit is described by

$$\begin{aligned} \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} &= -x_i(t) + (1-\alpha_i)\, x_i(t) + \alpha_i f(y_i(t)) \\ &= -\alpha_i x_i(t) + \alpha_i f(y_i(t))\,, \end{aligned} \tag{7.1}$$

or, in another form:

$$\tau_i \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} + x_i(t) = f(y_i(t))\,. \tag{7.2}$$

The symbols have the following meanings:

$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t}$	change of activation of unit i at time t
$x_i(t)$	activation of unit i at time t
$y_i(t)$	net input of unit i at time t
α_i	leakage rate of unit i
$\tau_i = \alpha_i^{-1}$	time constant of unit i
f	activation function of each unit; usually sigmoidal (e.g. logistic as in (7.5)) or linear.

The leakage rate α tells how much a unit depends on the actual net input. Its value is between 0 and 1. The lower the value of α, the stronger the influence of the previous level of activation and the less the influence of the actual net input. If $\alpha = 1$, the previous activation does not have any influence and the new activation is only determined by the net input (this is the case e.g. for the

standard units used by the PDP group [48]). By contrast, $\alpha = 0$ means that the actual net input does not have any influence and the activation remains constant. ($1 - \alpha$ could also be regarded as the strength of its *memory* with respect to earlier activations.)

The net input of unit i is given as the sum of all incoming signals:

$$y_i(t) = \sum_j w_{ij} x_j(t) + b_i + I_i^{\text{ext}}(t)\,. \tag{7.3}$$

With

$y_i(t)$	net input of unit i at time t
w_{ij}	weight of connection from unit j to unit i
b_i	bias of unit i
I_i^{ext}	external input to unit i

Equation (7.1) is very similar to the general form of neural networks equations for continuous-valued units (described, for example, in [19]). The difference lies in the presence of the leakage term α that makes the current activation dependent on its previous activation. We motivate (7.1) by the equivalent recurrent network of Fig. 7.2 and we shall use it in Sect. 7.2.2 subsequently to derive a generalized back-propagation algorithm as a learning rule for temporal patterns. On the other hand, (7.2) is well-known from the theory of ordinary differential equations. Its associated homogeneous form

$$\tau_i \frac{\mathrm{d}x_i}{\mathrm{d}t} + x_i = 0$$

simply describes an exponential decay process. Therefore, the inhomogeneous (7.2) can be seen as a forced decay process integrating its input on the right hand side.

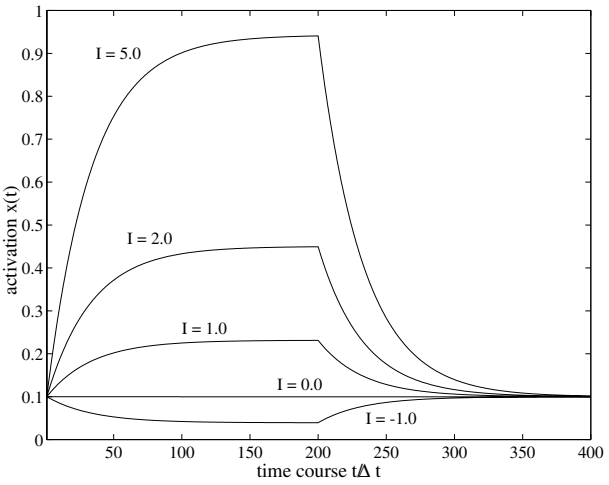
Hertz et al. [19, p. 54] discuss a Hopfield network of leaky integrator units which is characterized by (7.2) with symmetric synaptic weights w_{ij}. Such a network is a dynamical system whose attractors are the patterns which are to be learned. Moreover, Hertz et al. [19, p. 55] consider another dynamical system

$$\tau_i \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} + x_i(t) = \sum_j w_{ij} f(x_j(t)) + b_i + I_i^{\text{ext}}(t) \tag{7.4}$$

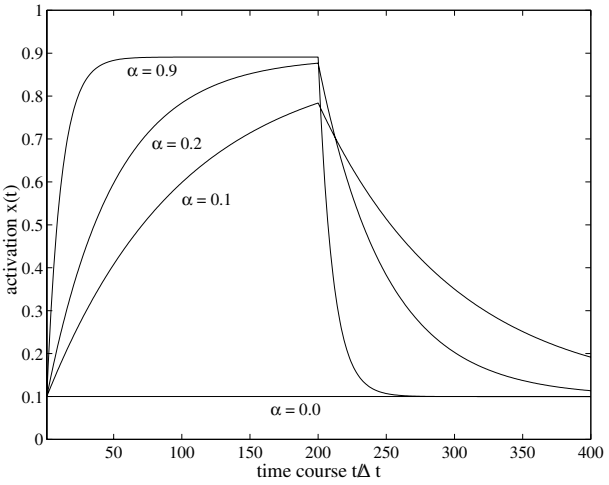
having the same equilibrium solutions as (7.2). As we shall see in Sect. 7.3.1, (7.4) appropriately models small networks of single neurons. The time-course of activation for a leaky integrator unit using a logistic activation function

$$f(x) = \frac{1}{1 + \mathrm{e}^{-\beta x}} \tag{7.5}$$

with respect to input and leakage rate is shown in Figs. 7.3(a) and (b).



(a)



(b)

Fig. 7.3. Time-course of activation (7.1); (**a**) for different input with $\Delta t = 0.1$, $\alpha = 0.3$ and $b = -2.2$; (**b**) Time-course of activation (7.1) for different leakage rates with $\Delta t = 0.1$, $I^{\text{ext}} = 4.3$ and $b \approx -2.2$

7.2.2 Training Leaky Integrator Networks

In order to use leaky integrator units to create network models for simulation experiments, a learning rule that works in continuous time is needed. The following formulation is motivated by [49, 50] and describes how a back-propagation algorithm for leaky integrator units can be derived.

In a first step, Euler's algorithm is used to change the differential equations into difference equations:[5]

$$\begin{aligned} x_i(t+\Delta t) &\approx x_i(t) + \frac{\mathrm{d}x_i(t)}{\mathrm{d}t}\Delta t \\ \Rightarrow \qquad \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} &\approx \frac{x_i(t+\Delta t) - x_i(t)}{\Delta t}\,. \end{aligned} \tag{7.6}$$

Combining (7.1) and (7.6) yields

$$\begin{aligned} \widetilde{x}_i(t+\Delta t) &= (1-\Delta t)\widetilde{x}_i(t) + \Delta t\left\{(1-\alpha_i)\widetilde{x}_i(t) + \alpha_i f(\widetilde{y}_i(t))\right\} \\ &= (1-\Delta t\alpha_i)\,\widetilde{x}_i(t) + \Delta t\alpha_i f(\widetilde{y}_i(t))\,, \end{aligned} \tag{7.7}$$

where tildes above variables (e.g. $\widetilde{x}$) denote continuous functions that have been discretized.

Figures 7.3(a) and (b) show the time-course of activation for a leaky integrator unit with different values of external input I and leakage parameters α with $I \neq 0$ for $t \in [0, 20]$. In order to train a network, one needs to define an error function

$$E = \int_{t_0}^{t_1} f_{\mathrm{err}}\left[\boldsymbol{x}(t), t\right] \mathrm{d}t\,. \tag{7.8}$$

Here, we choose the least mean square function

$$E = \frac{1}{2}\sum_i \int_{t_0}^{t_1} s_i \left[x_i(t) - d_i(t)\right]^2 \mathrm{d}t\,, \tag{7.9}$$

where $d_i(t)$ is the desired activation of unit i at time t and s_i is the relative importance of this activation: $s = 0$ means unimportant and $s = 1$ means most important.

If one changes the activation of unit i at time t for a small amount, one gets a measure of how much this change influences the error function:

$$e_i(t) = \frac{\partial f_{\mathrm{err}}\left[\boldsymbol{x}(t), t\right]}{\partial x_i(t)} \tag{7.10}$$

with

$$f_{\mathrm{err}} = \frac{1}{2}\sum_i s_i \left[x_i(t) - d_i(t)\right]^2\,.$$

With (7.9) as error function, we get

$$e_i(t) = s_i \left[x_i(t) - d_i(t)\right]\,. \tag{7.11}$$

[5] Note that the following derivation could also be achieved using the variational calculus well-known from Hamilton's principle in analytical mechanics [49]. We leave this as an exercise for the reader.

Equations (7.10) and (7.11) describe the influence of a change of activation only for t. In a neural net that is described by (7.1) and (7.3), each change of activation at t also influences the activation at later times t' ($t < t'$). The amount of this influence can be described by using time-ordered derivatives [51, 52]:

$$\begin{aligned} \tilde{z}_i(t) &= \frac{\partial^+ E}{\partial \widetilde{x}_i(t)} \\ &:= \frac{\partial E}{\partial \widetilde{x}_i(t)} + \sum_{t'>t} \sum_j \frac{\partial^+ E}{\partial \widetilde{x}_j(t')} \frac{\partial \widetilde{x}_j(t')}{\partial \widetilde{x}_i(t)} \end{aligned} \tag{7.12}$$

with $j = 1, 2, \ldots, n$ n number of units
$t' = t + \Delta t, t + 2\Delta t, \ldots, t_1$ t_1 last defined time

$\tilde{z}_i(t)$ measures how much a change of activation of unit i at time t influences the error function at all times.

Performing the derivations in (7.12) with (7.9), (7.11), (7.7) and (7.3) and setting $t' = t + \Delta t$ gives:

$$\frac{\partial E}{\partial \widetilde{x}_j(t)} = \Delta t e_i \tag{7.13}$$

$$\frac{\partial \widetilde{x}_i(t + \Delta t)}{\partial \widetilde{x}_i(t)} = (1 - \Delta t \alpha_i) + \Delta t \alpha_i w_{ii} f'(\widetilde{y}_i(t)) \tag{7.14}$$

$$\frac{\partial \widetilde{x}_j(t + \Delta t)}{\partial \widetilde{x}_i(t)} = \Delta t \alpha_j w_{ji} f'(\widetilde{y}_j(t)) \tag{7.15}$$

for all units j that are connected with unit i.

All other derivatives are zero. With this, one gets

$$\begin{aligned} \widetilde{z}_i(t) = \Delta t e_i &+ (1 - \Delta t \alpha_i)\, \widetilde{z}_i(t + \Delta t) \\ &+ \sum_j \Delta t \alpha_j w_{ji} f'(\widetilde{y}_j(t)) \widetilde{z}_j(t + \Delta t)\,. \end{aligned} \tag{7.16}$$

The back-propagated error signal $\boldsymbol{z}(t)$ is equivalent to the $\boldsymbol{\delta}$ in standard back-propagation. After the last defined activation $d_i(t_1)$, there is no further change of E, so $z_i(t_1 + \Delta t) = 0$.

Making use of Euler's method in the opposite direction, we find that the back-propagated error signal can be described by the following differential equation:

$$\frac{\mathrm{d}z_i(t)}{\mathrm{d}t} = \alpha_i z_i(t) - e_i - \sum_j \alpha_j w_{ji} f'(y_j(t)) z_j(t)\,. \tag{7.17}$$

With (7.16), it is possible to calculate how the error function changes if one changes the parameters α_i and w_{ij}. Each variation also changes the activation x_i. The influence of this activation on E can be calculated using the chain rule of derivatives.

If w_{ij} changes over Δt by ∂w_{ij}, then the influence of this change on the error function can be described by

$$\left.\frac{\partial E}{\partial w_{ij}}\right|_t^{t+\Delta t} := \frac{\partial^+ E}{\partial x_i(t+\Delta t)} \frac{\partial x_i(t+\Delta t)}{\partial w_{ij}} = z_i(t+\Delta t)\alpha_i x_j(t) f'(y_i(t))\Delta t\,. \tag{7.18}$$

A change of ∂w_{ij} during the *whole* time $t_0 \le t \le t_1$ produces:

$$\frac{\partial E}{\partial w_{ij}} = \alpha_i \int_{t_0}^{t_1} z_i(t) x_j(t) f'(y_i(t)) \mathrm{d}t\,. \tag{7.19}$$

For the influence of a change in α_i on E one finds

$$\left.\frac{\partial^+ E}{\partial \alpha_i}\right|_t^{t+\Delta t} = \frac{\partial E}{\partial x_i(t+\Delta t)} \frac{\partial x_i(t+\Delta t)}{\partial \alpha_i} = z_i(t+\Delta t)\left\{f(y_i(t)) - x_i(t)\right\}\Delta t\,. \tag{7.20}$$

For the whole time:

$$\frac{\partial E}{\partial \alpha_i} = \int_{t_0}^{t_1} z_i(t)\left\{f(y_i(t)) - x_i(t)\right\} \mathrm{d}t\,. \tag{7.21}$$

Now, we have nearly all the equations that are needed to train a neural network of leaky integrator units. Finally, we must keep in mind the fact that the leakage term α must be between 0 and 1. This can be done by using

$$\alpha = \frac{1}{1+\mathrm{e}^{-\bar{\alpha}}} \tag{7.22}$$

and learning $\bar{\alpha}$ instead of α. With this replacement we set

$$\frac{\partial E}{\partial \bar{\alpha}_i} = \frac{1}{1+\mathrm{e}^{-\bar{\alpha}_i}}\left(1 - \frac{1}{1+\mathrm{e}^{-\bar{\alpha}_i}}\right) \times \times \int_{t_0}^{t_1} z_i(t)\left\{f(y_i(t)) - x_i(t)\right\} \mathrm{d}t\,. \tag{7.23}$$

7.2.3 Overview of the Learning Procedure

To start the training, one needs to have the following information:

(i) topology of the net with number of units (n) and connections
(ii) values of the parameters $\boldsymbol{W}(0) = (w_{ij}(0))$ and $\bar{\boldsymbol{\alpha}}(0)$ at $t = 0$
(iii) activations $\boldsymbol{x}(t_0)$ at $t = 0$

(iv) time-course of the input $\boldsymbol{I}^{\text{ext}}(t), t_0 \leq t \leq t_1$
(v) time-course of the desired output $\boldsymbol{d}(t)$
(vi) activation function f for each unit
(vii) error function E
(viii) time-step size Δt that resembles the required resolution of the time-course ($\Delta t = 0.1$ turned out to be a good default value).

After having fixed these parameters according to the desired learning schedule, the goal is then to find a combination of $\boldsymbol{W}$ and $\bar{\boldsymbol{\alpha}}(0)$ that gives a minimum for E. This can be achieved by the following algorithm:

(i) At first one has to calculate the net input (7.3) for each unit successively and for each time-step *forward* in time. Simultaneously, the activations are calculated with (7.7).
(ii) With (7.9), one calculates the main error E and the error vector $\boldsymbol{e}(t)$ using (7.11).
(iii) Then, the error signals are propagated *backwards* through time with (7.16), making use of the condition $\tilde{z}_i(t_1 + \Delta t) = 0$.
(iv) Now, one calculates the gradient of each free parameter with respect to the error function E with the discrete versions of (7.19) and (7.23):

$$\frac{\partial E}{\partial w_{ij}} = \frac{1}{1 + \mathrm{e}^{-\bar{\alpha}_i}} \sum_{t=t_0}^{t_1} \widetilde{z}_i(t + \Delta t)\widetilde{x}_j(t) f'(\widetilde{y}_i(t))\Delta t \tag{7.24}$$

$$\frac{\partial E}{\partial \bar{\alpha}_i} = \frac{1}{1 + \mathrm{e}^{-\bar{\alpha}_i}} \left(1 - \frac{1}{1 + \mathrm{e}^{-\bar{\alpha}_i}}\right) \sum_{t=t_0}^{t_1} \widetilde{z}_i(t) \left\{ f(\widetilde{y}_i(t)) - \widetilde{x}_i(t) \right\} \Delta t \,. \tag{7.25}$$

(v) After this, the parameters are changed along the negative gradient (*gradient descent*):

$$w_{ij} = w_{ij} - \eta_w \frac{\partial E}{\partial w_{ij}} \tag{7.26}$$

$$\bar{\alpha}_i = \bar{\alpha}_i - \eta_{\bar{\alpha}} \frac{\partial E}{\partial \bar{\alpha}_i} \,, \tag{7.27}$$

with η_w and $\eta_{\bar{\alpha}}$ as learning rates. ($\eta = 0.1$ is commonly a suitable starting value.) The gradient can be used for *steepest descent*, *conjugate gradient* or other numeric approximations (see e.g. [53]).
(vi) Having obtained the new values $\boldsymbol{W}$ and $\bar{\boldsymbol{\alpha}}$, the procedure goes back to step (i) and is followed until the main error falls below a certain value in step (ii) or this criterion is not reached after a maximal number of iterations.

(For a model that uses this type of learning algorithm with leaky integrator units, see [54]). In the context of modeling oscillating brain activity, recurrent networks of leaky integrator units become interesting. Section 7.4 will describe three typical examples.

7.3 From Physiology to Leaky Integrator Models

7.3.1 Leaky Integrator Model of Single Neurons

Let us consider the somatic membrane of a neuron i in the vicinity of its trigger zone. For the sake of simplicity, we shall assume that the membrane behaves only passively at this site. For further simplification, we do not describe the trigger zone by the complete Hodgkin-Huxley equations [36], but instead as a McCulloch-Pitts neuron [18], i.e. as a threshold device: the neuron fires if its membrane potential $U_i(t)$ exceeds the activation threshold $\theta \approx -50\,\mathrm{mV}$ from below due to the law of "all-or-nothing" [55, 56]. Because of this, the membrane potential $U_i(t)$ becomes translated into a spike train which can be modeled by a sum of delta functions

$$R_i(t) = \sum_{\substack{k:U_i(t_k)=\theta \\ \dot{U}_i(t_k)>0}} \delta(t - t_k)\,. \tag{7.28}$$

Now, we can determine the number of spikes in a time interval $[0, t]$ [35], which is given by

$$N_i(t) = \int_0^t R_i(t')\mathrm{d}t'\,.$$

Thus, from the *spike rate* per unit time, we regain the original signal

$$\frac{\mathrm{d}}{\mathrm{d}t} N_i(t) = R_i(t)\,. \tag{7.29}$$

In the next step, we consider the membrane potential U_i in the vicinity of the trigger zone which obeys Kirchhoff's First Law (see Fig. 7.4), i.e.

$$\sum_j I_{ij} = \frac{U_i - E_m}{r_m} + c_i \frac{\mathrm{d}U_i}{\mathrm{d}t}\,, \tag{7.30}$$

here, E_m is the Nernst equilibrium potential of the leakage channels with resistance r_m. c_i is the capacitance of the membrane of neuron i and I_{ij} is the current through the membrane at the chosen site coming from the synapse formed by the jth neuron with neuron i.

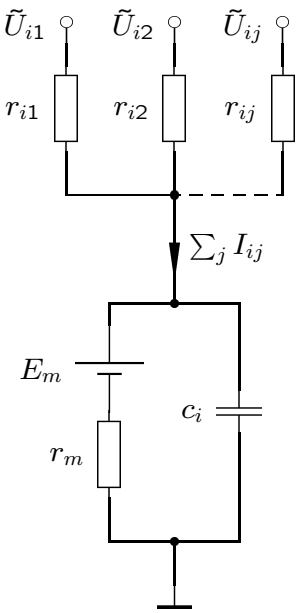


Fig. 7.4. Equivalent circuit for the leaky integrator neuron

These synaptic currents depend upon both the potential difference $\tilde{U}_{ij} - U_i$ between the postsynaptic potential $\tilde{U}_{ij}$ at the synapse connecting neuron j to i and the potential U_i at the trigger zone of i, and the intracellular resistance along the current's path r_{ij}. Therefore

$$I_{ij} = \frac{\tilde{U}_{ij} - U_i}{r_{ij}} \tag{7.31}$$

applies. Inserting (7.31) into (7.30) yields

$$\sum_j \frac{\tilde{U}_{ij} - U_i}{r_{ij}} = \frac{U_i - E_m}{r_m} + c_i \frac{\mathrm{d}U_i}{\mathrm{d}t} ,$$

and after some rearrangements

$$r_m c_i \frac{\mathrm{d}U_i}{\mathrm{d}t} + U_i \left(1 + \sum_j \frac{r_m}{r_{ij}} \right) = E_m + \sum_j \frac{r_m}{r_{ij}} \tilde{U}_{ij} . \tag{7.32}$$

After letting $E_m = 0$, without loss of generality and introducing the *time constants*

$$\tau_i = \frac{r_m c_i}{1 + \sum_j \frac{r_m}{r_{ij}}} \tag{7.33}$$

and provisory *synaptic weights*

$$\tilde{w}_{ij} = \frac{\frac{r_m}{r_{ij}}}{1 + \sum_j \frac{r_m}{r_{ij}}}\,, \tag{7.34}$$

we eventually obtain

$$\tau_i \frac{\mathrm{d}U_i}{\mathrm{d}t} + U_i = \sum_j \tilde{w}_{ij}\tilde{U}_{ij}\,. \tag{7.35}$$

Next, the postsynaptic potentials $\tilde{U}_{ij}$ require our attention. We assume that an action potential arriving at the presynaptic terminal of the neuron j releases, on average, one transmitter vesicle.[6] The content of the vesicle diffuses through the synaptic cleft and reacts with receptor molecules embedded in the postsynaptic membrane. After chemical reactions described by kinetic differential equations (cf. Chap. 1, [35]), opened ion channels give rise to a postsynaptic impulse response potential $G_{ij}(t)$. Because we characterize the dendro-somatic membranes as linear systems here, the postsynaptic potential elicited by a spike train $R_j(t)$ is given by the convolution

$$\tilde{U}_{ij}(t) = G_{ij}(t) * R_j(t)\,. \tag{7.36}$$

Let us make a rather crude approximation here by setting the postsynaptic impulse response function proportional to a delta function:

$$G_{ij}(t) = g_{ij}\delta(t)\,, \tag{7.37}$$

where g_{ij} is the *gain* of the synapse $j \to i$. Then, the postsynaptic potential is given by the product of the gain with the spike rate of the presynaptic neuron j.

Finally, we must take the stochasticity of the neuron into account as thoroughly described in Chap. 1. This is achieved by replacing the membrane potential U_j at the trigger zone by its average obtained from the distribution function, which leads to the characteristic sigmoidal activation functions [57], e.g. the logistic function (see (7.5))

$$R_j(t) = f(U_j(t)) = \frac{1}{1 + \mathrm{e}^{-\beta[U_j(t)-\theta]}}\,. \tag{7.38}$$

Collecting (7.35, 7.36) and (7.38) together and introducing the proper *synaptic weights*

$$w_{ij} = g_{ij}\tilde{w}_{ij} \tag{7.39}$$

yields the leaky integrator model of a network of distributed single neurons

$$\tau_i \frac{\mathrm{d}U_i}{\mathrm{d}t} + U_i = \sum_j w_{ij}\, f(U_j(t)) \tag{7.40}$$

which is analogous to (7.4).

[6] The release of transmitter is a stochastic process that can be approximately described by a binomial distribution [55], and hence, due to the limit theorem of de Moivre and Laplace, is normally distributed (see Chap. 1).

7.3.2 Leaky Integrator Model of Neural Populations

According to Freeman [58] (see also [59]), a neuronal *population* ("KI" set) consists of many reciprocally connected neurons of one kind, either excitatory or inhibitory. Let us consider such a set of McCulloch-Pitts neurons [18] distributed over a unit volume i of neural tissue. We introduce the proportions of firing cells (either excitatory or inhibitory, in contrast to [17]) in volume i at the instance of time t, $Q_i(t)$, as the state variables [17, 32].

A neuron belonging to volume i will fire if its net input U_i (analogous to the membrane potential at the trigger zone, see Sect. 7.3.1) crosses the threshold θ. But now, we have to deal with an ensemble of neurons possessing randomly distributed thresholds within the unit volume i. We therefore obtain an ensemble activation function [5] (cf. Chap. 14) by integrating the corresponding probability distribution density $D(\theta)$ of thresholds [17],

$$f(U_i) = \int_0^{U_i} D(\theta) \mathrm{d}\theta \,. \tag{7.41}$$

Depending upon the modality of the distribution $D(\theta)$, the activation function could be sigmoidal or even more complicated. For unimodal distributions such as Gaussian or Poissonian distributions, $f(U_i)$ might be approximated by the logistic function (7.38). As for the single neuron model, the net input is obtained by a convolution

$$U_i(t) = \int_{-\infty}^{t} G(t - t') \sum_j w_{ij} \, Q_j(t') \mathrm{d}t' \,, \tag{7.42}$$

with "synaptic weights" w_{ij} characterizing the neural connectivity and whether the population is excitatory or inhibitory.

In the following, we shall simplify the model of Wilson and Cowan [17] by neglecting the refractory time. The model equations are then

$$Q_i(t + \tau_i) = f \left(\int_{-\infty}^{t+\tau_i} G(t - t') \sum_j w_{ij} \, Q_j(t') \mathrm{d}t' \right) \,, \tag{7.43}$$

such that $Q_i(t + \tau_i)$ is the proportion of cells being above threshold in the time interval $[t, t + \tau_i]$. Expanding the left hand side into a Taylor series at t and assuming again that $G(t - t') = \delta(t - t')$, we obtain

$$\tau_i \, \frac{\mathrm{d}Q_i(t)}{\mathrm{d}t} + Q_i(t) = f \left(\sum_j w_{ij} \, Q_j(t) \right) \,, \tag{7.44}$$

a leaky integrator model again, yet characterized by (7.2).

7.4 Oscillators from Leaky Integrator Units

7.4.1 Linear Model

In this section, we demonstrate that a damped harmonic oscillator can be obtained from a simple model of two recurrently coupled leaky integrator units with linear activation functions [49]. Figure 7.5 shows the architecture of this model.

The network of Fig. 7.5 is governed by (7.4):

$$\tau_1 \frac{\mathrm{d}x_1}{\mathrm{d}t} + x_1 = w_{11}x_1 + w_{12}x_2 + p \tag{7.45}$$

$$\tau_2 \frac{\mathrm{d}x_2}{\mathrm{d}t} + x_2 = w_{21}x_1 + w_{22}x_2\,, \tag{7.46}$$

where x_1 denotes the activity of unit 1 and x_2 that of unit 2. Correspondingly, τ_1 and τ_2 are the time constants of the units 1 and 2, respectively. The synaptic weights w_{ij} are indicated in Fig. 7.5. Note that the weights w_{11} and w_{22} describe autapses [60]. The quantity p refers to excitatory synaptic input that might be a periodic forcing or any other function of time.

Equations (7.45) and (7.46) can be converted into two second-order ordinary differential equations

$$\frac{\mathrm{d}^2x_1}{\mathrm{d}t^2} + \gamma\frac{\mathrm{d}x_1}{\mathrm{d}t} + \omega_0^2 = p_1 \tag{7.47}$$

$$\frac{\mathrm{d}^2x_2}{\mathrm{d}t^2} + \gamma\frac{\mathrm{d}x_2}{\mathrm{d}t} + \omega_0^2 = p_2\,, \tag{7.48}$$

where we have introduced the following simplifying parameters:

$$\gamma = \frac{\tau_1(1-w_{22}) + \tau_2(1-w_{11})}{\tau_1\tau_2}$$

$$\omega_0^2 = \frac{w_{11}w_{22} - w_{12}w_{21} - w_{11} - w_{22} + 1}{\tau_1\tau_2}$$

$$p_1 = \frac{1}{\tau_1}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{1-w_{22}}{\tau_1\tau_2}$$

$$p_2 = \frac{w_{21}}{\tau_1\tau_2}p$$

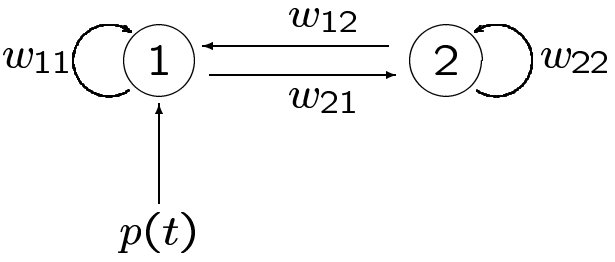


Fig. 7.5. Architecture of an oscillator formed by leaky integrator units

Now, (7.47) and (7.48) describe two damped, decoupled, harmonic oscillators with external forcing when $\gamma \geq 0$ and $\omega_0^2 > 0$, i.e. one unit must be excitatory and the other inhibitory.

7.4.2 Simple Nonlinear Model

Next, we discuss a simple nonlinear system, consisting of three coupled leaky integrator units, which provides a model of the thalamocortical loop. Figure 7.6 displays its architecture.

According to Fig. 7.6, the model (7.4) are

$$\tau_1 \frac{\mathrm{d}x_1}{\mathrm{d}t} + x_1 = -\alpha f(x_3(t)) \tag{7.49}$$

$$\tau_2 \frac{\mathrm{d}x_2}{\mathrm{d}t} + x_2 = \beta f(x_1(t)) \tag{7.50}$$

$$\tau_3 \frac{\mathrm{d}x_3}{\mathrm{d}t} + x_3 = \gamma f(x_2(t))\,. \tag{7.51}$$

Setting all $\tau_i = 1$ and rearranging, we get

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}t} &= -x_1 - \alpha f(x_3(t)) \\
\frac{\mathrm{d}x_2}{\mathrm{d}t} &= -x_2 + \beta f(x_1(t)) \\
\frac{\mathrm{d}x_3}{\mathrm{d}t} &= -x_3 + \gamma f(x_2(t))\,.
\end{aligned}$$

These equations define a vector field $\boldsymbol{F}$ with the Jacobian matrix

$$\mathrm{D}\boldsymbol{F} = \begin{pmatrix} -1 & 0 & -\alpha f'(x_3(t)) \\ \beta f'(x_1(t)) & -1 & 0 \\ 0 & \gamma f'(x_2(t)) & -1 \end{pmatrix}.$$

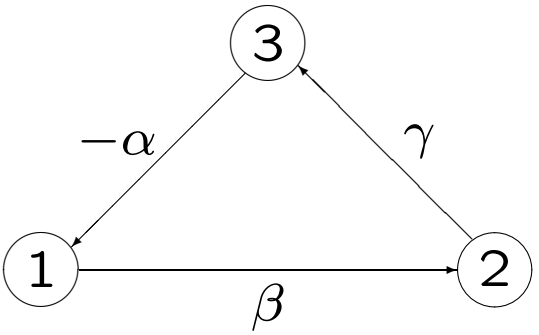


Fig. 7.6. Thalamocortical oscillator of three leaky integrator units: (*1*) pyramidal cell; (*2*) thalamus cell; (*3*) cortical interneuron (star cell)

For the activation function, we chose $f(x) = \tanh x$, which can be obtained by a coordinate transformation from the logistic function in (7.38). Therefore, $\boldsymbol{F}(x_1, x_2, x_3) = 0$ and we can look at whether the center manifold theorem [61] can be applied. The Jacobian at $(0, 0, 0)$ is

$$\mathrm{D}\boldsymbol{F}(0) = \begin{pmatrix} -1 & 0 & -\alpha \\ \beta & -1 & 0 \\ 0 & \gamma & -1 \end{pmatrix},$$

having eigenvalues

$$\begin{aligned} \lambda_1 &= -1 - \sqrt[3]{\alpha\beta\gamma} \\ \lambda_2 &= -1 + \frac{1}{2}(1 - \mathrm{i}\sqrt{3})\sqrt[3]{\alpha\beta\gamma} \\ \lambda_3 &= -1 + \frac{1}{2}(1 + \mathrm{i}\sqrt{3})\sqrt[3]{\alpha\beta\gamma}\,. \end{aligned}$$

Since $\lambda_1 < 0$ for $\alpha, \beta, \gamma \geq 0$, we seek for the weight parameters making $\mathrm{Re}(\lambda_{2|3}) = 0$. This leads to the condition

$$\alpha\beta\gamma = 8\,, \tag{7.52}$$

which can be easily fulfilled, for example, by setting

$$\alpha = 4, \qquad \beta = 1, \qquad \gamma = 2\,.$$

In this case, the center manifold theorem applies: the dynamics stabilizes along the eigenvector corresponding to λ_1, exhibiting a limit cycle in the center manifold spanned by the eigenvectors of λ_2 and λ_3. Figure 7.7 shows a numerical simulation of this oscillator. It is also possible to train a leaky integrator network using the algorithm described in Sect. 7.2.2 in order to replicate a limit cycle dynamics [49].

7.4.3 Random Neural Networks

In this last subsection, we describe a network model that is closely related to those presented in Chaps. 3, 5, 12, 13, and 14, namely a random graph carrying leaky integrator units described by (7.4) or, equivalently, (7.40), at its nodes:

$$\tau_i \frac{\mathrm{d}x_i(t)}{\mathrm{d}t} + x_i(t) = \sum_j w_{ij} f(x_j(t))\,.$$

We shall see that the onset of oscillatory behavior is correlated with the emergence of super-cycles in the topology of the network provided by an evolving directed and weighted Erdős-Rényi graph of N nodes where all connections between two nodes are equally likely with increasing probability [62–64].

As explained in Chap. 3, a directed Erdős-Rényi graph consists of a set of vertices V that are randomly connected by arrows taken from an edge set

$E \subset V \times V$ with equal probability q. The topology of the graph is completely described by its *adjacency matrix* $\boldsymbol{A} = (a_{ij})$ where $a_{ij} = 1$, if there is an arrow connecting the vertex j with the vertex i (i.e. $(j,i) \in E$ for $i,j \in V$) while $a_{ij} = 0$ otherwise. A directed and weighted Erdős-Rényi graph is then described by the *weight matrix* $\boldsymbol{W} = (w_{ij})$ which is obtained by element-wise multiplication of the adjacency matrix with constants g_{ij}: $w_{ij} = g_{ij}\, a_{ij}$. Biologically plausible models must satisfy Dale's law, which says that excitatory neurons only have excitatory synapses while inhibitory neurons only possess inhibitory synapses [56]. Therefore, the column vectors of the weight matrix are constrained to have a unique sign. We achieve this requirement by randomly choosing a proportion p of the vertices to be excitatory and the remainder to be inhibitory.

In our model, the weights become time-dependent due to the following evolution algorithm:

(i) Initialization: $\boldsymbol{W}(0) = 0$.
(ii) At evolution time t, select a random pair of nodes i, j.
(iii) If they are not connected, create a synapse with weight $w_{ij}(t+1) = +\delta$ if j is excitatory, and $w_{ij}(t+1) = -\delta$ if j is inhibitory. If they are already connected, enhance the weight $w_{ij}(t+1) = w_{ij}(t) + \delta$ if $w_{ij}(t) > 0$ and $w_{ij}(t+1) = w_{ij}(t) - \delta$ if $w_{ij}(t) < 0$. All other weights remain unchanged.
(iv) Repeat from (ii) for a fixed number of iterations L.

As the "learning rate", we choose $\delta = 1$, while the connectivity increases for L time steps. In order to simplify the simulations, we further set $\tau_i = 1$ for all $1 \le i \le N$.

Since (7.40) describes the membrane potential of the ith neuron, we can estimate its dendritic field potential by the inhomogeneity of (7.40),

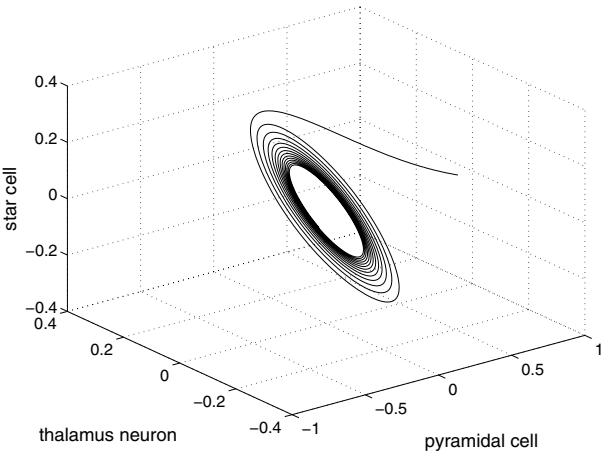


Fig. 7.7. Limit cycle of the thalamocortical oscillator in its center manifold plane

$$F_i(t) = \sum_j w_{ij}\, f(x_j)\,. \tag{7.53}$$

Then, the model EEG[7] is given by the sum of the dendritic field potentials of all excitatory nodes

$$E(t) = \sum_{i^+} F_i(t)\,. \tag{7.54}$$

The indices i^+ indicate that the neuron i belongs to the population of excitatory neurons, namely the EEG generating pyramidal cells.

We create such random neural networks with size $N = 100$, 200, 500, and 1000 nodes. Since about 80% of cortical neurons are excitatory pyramidal cells, $p = 80\%$ of the network's nodes are chosen to be excitatory [66]. For each iteration of the network's evolution, the dynamics of its nodes is calculated. After preparing them with normally distributed initial conditions ($\mu = 0, \sigma = 1$), (7.40) is solved numerically with the activation functions $f_i(x) = \tanh x$ for an ensemble of $K = 10$ time series of length $T = 100$ with a step-width of $\Delta t = 0.0244$. The dendritic field potential and EEG are computed according to (7.53) and (7.54).

From the simulated EEGs, the power spectra are computed and averaged over all K realizations of the network's dynamics. In order to monitor sudden changes in the topologies of the networks, three characteristic statistics are calculated:

(1) The *mean degree* (the average number of vertices attached to the nodes) $\langle k \rangle$ of the associated undirected graphs, described by the symmetrized adjacency matrix $\boldsymbol{A}^s = \Theta(\boldsymbol{A} + \boldsymbol{A}^T)$ (Θ denotes Heaviside's step function),

(2) the *total distribution*

$$d(l) = \frac{\mathrm{tr}(\boldsymbol{A}^l)}{l\mathcal{N}} \tag{7.55}$$

of cycles of the exact length l [62–64, 67–70]. In (7.55), $\mathrm{tr}(\boldsymbol{A}^l)$ provides the total number of (not necessarily self-avoiding) closed paths of length l through the network. Since any node at such a path may serve as the starting point and there are l nodes, the correct number of cycles is obtained by dividing by l. Finally, $\mathcal{N} = \sum_l \mathrm{tr}(\boldsymbol{A}^l)/l$ is a normalization constant. From the cycle distribution (7.55), we derive

(3) an *order parameter* s for topological transitions from the averaged slopes of the envelope of $d(l)$, where the envelopes are estimated by connecting the local maxima of $d(l)$. The above mentioned procedure is repeated for each network size $M = 10$ times where we have chosen $L_{100} = 150, L_{200} = 400, L_{500} = 800$, and $L_{1000} = 1700$ iterations of network evolution in order to ensure sufficiently dense connectivities.

[7] In fact, (7.54) describes better the local field potential (LFP) rather than the EEG. Considering (7.4) as a model of coupled neural populations, instead, seems to be more appropriate for describing the EEG [65].

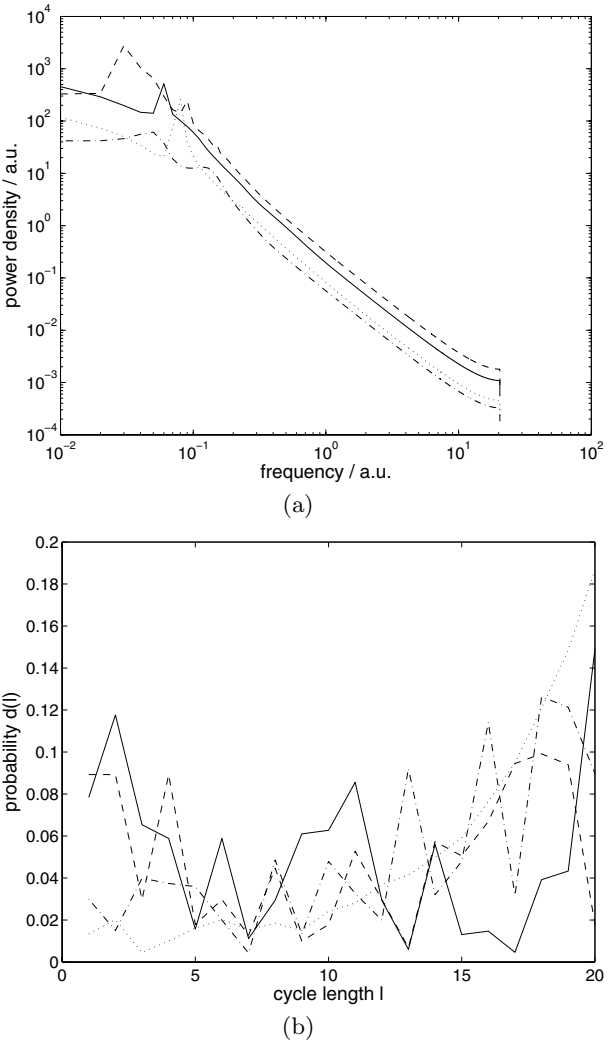


Fig. 7.8. (**a**) Power spectra of representative simulated time series during the oscillatory transition (critical phase) for four different network sizes: $N = 100$ (*dotted*), $N = 200$ (*dashed-dotted*), $N = 500$ (*solid*), and $N = 1000$ (*dashed*); (**b**) Total distributions of cycles (7.55) for the same networks

Figure 7.8 shows four representative networks in the critical phase characterized by the smallest positive value of the cycle order parameter s, averaged over the $M = 10$ network simulations, when sudden oscillations occur in the dynamics of the units, as is visible by the peaks in the power spectra [Fig. 7.8(a)]. The cycle distributions $d(l)$ [Fig. 7.8(b)] for network sizes $N = 200, 500,$ and 1000 display a transition from geometrically decaying to

exponentially growing functions while this transition has already taken place for $N = 100$. As Fig. 7.8(a) reveals, the power spectra display a broad $1/f$ continuum. Superimposed to this continuum are distinguished peaks that can be regarded as the "alpha waves" of the model.

According to random graph theory, Erdős-Rényi networks exhibit a percolation transition when a giant cluster suddenly occurs for $\langle k \rangle = 1$ [62–64]. A second transition takes place for $\langle k \rangle = 2$, indicating the appearance of mainly isolated cycles in the graph. Isolated cycles are characterized by a geometrically decaying envelope of the total cycle distribution. Our simulations suggest the existence of a third transition when super-cycles are composed from merging smaller ones. This is reflected by a transition of the total cycle distribution $d(l)$ from a geometrically decaying to an exponentially growing behavior due to a "combinatorial explosion" of possible self-intersecting paths through the network (super-cycles are common in regular lattices with $\langle k \rangle \geq 3$). We detect this transition by means of a suitably chosen order parameter s derived from $d(l)$ as the averaged slope of its envelope. For decaying $d(l)$, $s < 0$ and for growing $d(l)$, $s > 0$. The appearance of super-cycles is associated with $s \approx 0$ if $d(l)$ is approximately symmetric in the range of l. In this case, sustained oscillations emerge in the network's dynamics due to the presence of reverberatory circles. For further details, see [65].

7.5 Cognitive Modeling

In this chapter we have reviewed neurophysiological findings on oscillations in the electroencephalogram as well as certain approaches to model these through coupled differential equations. We have introduced the theory of networks of *leaky integrator units* and presented a general learning rule to train these networks in such a way that they are able to reproduce temporal patterns in continuous time. This learning rule is a generalized back-propagation algorithm that has been applied for the first time to model reaction times from a psychological experiment [54]. Therefore, leaky integrator networks provide a unique and physiologically plausible paradigm for neural and cognitive modeling.

Mathematically, leaky integrator models are described by systems of coupled ordinary differential equations that become nonlinear dynamical systems by using sigmoidal activation functions. Networks of leaky integrator units may display a variety of complex behaviors: limit cycles, multistability, bifurcations and hysteresis [17]. They could therefore act as models of perceptional instability [71] or cognitive conflicts [72], as has already been demonstrated by Haken [73, 74] using *synergetic computers*. As Haken [74, p. 246] pointed out, the order parameter equations of synergetic computers are analogous to neural networks whose activation function is expanded into a power series. However, these computers are actually leaky integrator networks as we will see subsequently.

Basically, synergetic computers are time-continuous Hopfield nets [19] governed by a differential equation

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} - \sum_{k=1}^{K} \eta_k \boldsymbol{v}_k (\boldsymbol{v}_k^+ \boldsymbol{x}) = -B \sum_{k' \neq k}^{K} (\boldsymbol{v}_{k'}^+ \boldsymbol{x})^2 (\boldsymbol{v}_k^+ \boldsymbol{x}) \boldsymbol{x} - C(\boldsymbol{x}^+ \boldsymbol{x}) \boldsymbol{x} \tag{7.56}$$

where $\boldsymbol{x}(t)$ denotes the activation vector of the network; the K vectors $\boldsymbol{v}_k$ are training patterns with adjoints $\boldsymbol{v}_k^+$ such that the orthonormality relations $\boldsymbol{v}_k^+ \boldsymbol{v}_l = \delta_{kl}$ hold. In this notation, $\boldsymbol{x}^+ \boldsymbol{y} = \sum_i x_i y_i$ means the *inner product* of the row vector $\boldsymbol{x}^+$ with a column vector $\boldsymbol{y}$ yielding a scalar. On the other hand, the outer product $\boldsymbol{y}\boldsymbol{x}^+$ of a column vector $\boldsymbol{y}$ with a row vector $\boldsymbol{x}^+$ is a matrix with elements $y_i x_j$.

Therefore, the second term of the left hand side of (7.56) can be rewritten as

$$\sum_{k=1}^{K} \eta_k \boldsymbol{v}_k (\boldsymbol{v}_k^+ \boldsymbol{x}) = \sum_{k=1}^{K} \eta_k (\boldsymbol{v}_k \boldsymbol{v}_k^+) \boldsymbol{x} = \left(\sum_{k=1}^{K} \eta_k \boldsymbol{v}_k \boldsymbol{v}_k^+ \right) \boldsymbol{x} = \boldsymbol{W} \boldsymbol{x}$$

where

$$\boldsymbol{W} = \sum_{k=1}^{K} \eta_k \boldsymbol{v}_k \boldsymbol{v}_k^+$$

is the synaptic weight matrix obtained by Hebbian learning of the patterns $\boldsymbol{v}_k$ with learning rates η_k.

The notion "synergetic computer" refers to the possibility of describing the network (7.56) by the evolution of *order parameters*, which are appropriately chosen as the "loads" of the training patterns $\boldsymbol{v}_k$ in a kind of principal component analysis. We therefore separate activation space and time by the ansatz

$$\boldsymbol{x}(t) = \sum_k \xi_k(t) \boldsymbol{v}_k + \boldsymbol{w}(t)\,,$$

where $\xi_k(t) = \boldsymbol{v}_k^+ \boldsymbol{x}(t)$ and $\boldsymbol{w}(t)$ is a fast decaying remainder. Multiplying (7.56) from the left with $\boldsymbol{v}_l^+$ and exploiting the orthonormality relations, we eventually obtain

$$\frac{\mathrm{d}\xi_l}{\mathrm{d}t} - \eta_l \xi_l = -B \sum_{k \neq l}^{K} \xi_k^2 \xi_l - C \left(\sum_k \xi_k^2 \right) \xi_l\,. \tag{7.57}$$

Division by $-\eta_l = 1/\tau_l$ then yields the leaky integrator equations for the order parameters with rescaled constants B', C' and a cubic activation function

$$\tau_l \frac{\mathrm{d}\xi_l}{\mathrm{d}t} + \xi_l = B' \sum_{k \neq l}^{K} \xi_k^2 \xi_l + C' \left(\sum_k \xi_k^2 \right) \xi_l\,. \tag{7.58}$$

The "time constants" play then the role of *attention parameters* describing the amount of attention devoted to a particular pattern. These parameters might also depend on time, e.g. for modeling habituation.

From a formal point of view, the attention model of Lourenço [75], the *cellular neural networks* (CNN) of Chua [76] (see also [77–79]) and the disease model of Huber et al. [80] can also be regarded as leaky integrator networks.

Also, higher cognitive functions such as language processing and their neural correlates such as event-related brain potentials (ERPs) [72,81,82] can be modeled with leaky integrator networks. Kawamoto [83] used a Hopfield net with exponentially decaying activation and habituating synaptic weights to modeling lexical ambiguity resolution. The activations of the units in his model are governed by the equations

$$x_i(t+1) = f\left(\delta x_i(t) + \sum_j w_{ij} x_j(t)\right) . \tag{7.59}$$

Setting $\delta = 1-\alpha = 1-\tau^{-1}$ and approximating $f'(x) \approx 1$ for typical activation values yields, after a Taylor expansion of the activation function,

$$f\left(\delta x_i(t) + \sum_j w_{ij} x_j(t)\right) \approx f\left(\sum_j w_{ij} x_j(t)\right) + f'\left(\sum_j w_{ij} x_j(t)\right) \delta x_i(t) ,$$

the leaky integrator equation (7.2).

Smolensky and Legendre [46] consider Hopfield nets of leaky integrator units that can be described by a Lyapunov function E. They call the function $H = -E$ the *harmony* of the network and argue that cognitive computations maximize this harmony function at the sub-symbolic level. Additionally, the harmony value can also be computed at the symbolic level of linguistic representations in the framework of harmonic grammars or optimality theory [84]. By regarding the harmony as an order parameter of the network, one could also model neural correlates of cognitive processes, e.g., ERPs.

This has recently been attempted by Wennekers et al. [83], who built a six-layer model of the perisylvian language cortex by randomly connecting leaky integrator units within each layer (similar to our exposition in Sect. 7.4.3). The network was trained with a Hebbian correlation learning rule to memorize "words" (co-activated auditory and motor areas) and "pseudowords" (activation of the auditory layer only). After training, cell assemblies of synchronously oscillating units across all six layers emerged. Averaging their event-related oscillations in the recall phase then yielded a larger amplitude for the "words" than for the "pseudowords", thus emulating the mismatch negativity (MMN) ERP known from word recognition experiments [45].

Acknowledgements

We gratefully acknowledge helpful discussions with W. Ehm, P. Franaszcuk, M. Garagnani, A. Hutt, V. K. Jirsa, and J. J. Wright. This work has been supported by the Deutsche Forschungsgemeinschaft (research group "Conflicting

Rules in Cognitive Systems"), and by the Helmholtz Institute for Supercomputational Physics at the University of Potsdam.

References

1. O. Creutzfeld and J. Houchin. Neuronal basis of EEG-waves. In *Handbook of Electroencephalography and Clinical Neurophysiology*, Vol. 2, Part C, pp. 2C-5–2C-55. Elsevier, Amsterdam, 1974.
2. F. H. Lopes da Silva. Neural mechanisms underlying brain waves: from neural membranes to networks. *Electroencephalography and Clinical Neurophysiology*, 79: 81–93, 1991.
3. E.-J. Speckmann and C. E. Elger. Introduction to the neurophysiological basis of the EEG and DC potentials. In E. Niedermeyer and F. Lopez da Silva, editors, *Electroencephalography. Basic Principles, Clinical Applications, and Related Fields*, Chap. 2, pp. 15–27. Lippincott Williams and Wilkins, Baltimore, 1999.
4. S. Zschocke. *Klinische Elektroenzephalographie*. Springer, Berlin, 1995.
5. W. J. Freeman. Tutorial on neurobiology: from single neurons to brain chaos. *International Journal of Bifurcation and Chaos*, 2(3): 451–482, 1992.
6. P. L. Nunez and R. Srinivasan. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, New York, 2006.
7. E. Başar. *EEG-Brain Dynamics. Relations between EEG and Brain Evoked Potentials*. Elsevier/North Holland Biomedical Press, Amsterdam, 1980.
8. M. Steriade, P. Gloor, R. R. Llinás, F. H. Lopes da Silva, and M.-M. Mesulam. Basic mechanisms of cerebral rhythmic activities. *Electroencephalography and Clinical Neurophysiology*, 76: 481–508, 1990.
9. C. Allefeld and J. Kurths. Testing for phase synchronization. *International Journal of Bifurcation and Chaos*, 14(2): 405–416, 2004.
10. C. Allefeld and J. Kurths. An approach to multivariate phase synchronization analysis and its application to event-related potentials. *International Journal of Bifurcation and Chaos*, 14(2): 417–426, 2004.
11. R. Srinivasan. Internal and external neural synchronization during conscious perception. *International Journal of Bifurcation and Chaos*, 14(2): 825–842, 2004.
12. G. Pfurtscheller. EEG rhythms — event related desynchronization and synchronization. In H. Haken and H. P. Koepchen, editors, *Rhythms in Physiological Systems*, pp. 289–296, Springer, Berlin, 1991.
13. E. Başar, M. Özgören, S. Karakaş, and C. Başar-Eroğlu. Super-synergy in brain oscillations and the grandmother percept is manifested by multiple oscillations. *International Journal of Bifurcation and Chaos*, 14(2): 453–491, 2004.
14. W. Klimesch, M. Schabus, M. Doppelmayr, W. Gruber, and P. Sauseng. Evoked oscillations and early components of event-related potentials: an analysis. *International Journal of Bifurcation and Chaos*, 14(2): 705–718, 2004.
15. N. Birbaumer and R. F. Schmidt. *Biologische Psychologie*. Springer, Berlin, 1996.
16. M. Steriade. Cellular substrates of brain rhythms. In E. Niedermeyer and F. Lopez da Silva, editors, *Electroencephalography. Basic Principles, Clinical Applications, and Related Fields*, Chap. 3, pp. 28–75. Lippincott Williams and Wilkins, Baltimore, 1999.

17. H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12: 1–24, 1972.
18. W. S. McCulloch and W. Pitts. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5: 115–33, 1943. Reprinted in J. A. Anderson and E. Rosenfeld (1988), pp. 83ff.
19. J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the Theory of Neural Computation*. Perseus Books, Cambridge (MA), 1991.
20. H. R. Wilson. *Spikes, Decisions and Actions. Dynamical Foundations of Neuroscience*. Oxford University Press, New York (NY), 1999.
21. F. H. Lopes da Silva, A. Hoecks, H. Smits, and L. H. Zetterberg. Model of brain rhythmic activity: the alpha-rhythm of the thalamus. *Kybernetik*, 15: 27–37, 1974.
22. F. H. Lopes da Silva, A. van Rotterdam, P. Bartels, E. van Heusden, and W. Burr. Models of neuronal populations: the basic mechanisms of rhythmicity. In M. A. Corner and D. F. Swaab, editors, *Perspectives of Brain Research*, Vol. 45 of *Progressive Brain Research*, pp. 281–308. 1976.
23. W. J. Freeman. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biological Cybernetics*, 56: 139–150, 1987.
24. B. H. Jansen and V. G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73: 357–366, 1995.
25. B. H. Jansen, G. Zouridakis, and M. E. Brandt. A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biological Cybernetics*, 68: 275–283, 1993.
26. F. Wendling, F. Bartolomei, J. J. Bellanger, and P. Chauvel. Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *European Journal of Neuroscience*, 15: 1499–1508, 2002.
27. F. Wendling, J. J. Bellanger, F. Bartolomei, and P. Chauvel. Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. *Biological Cybernetics*, 83: 367–378, 2000.
28. O. David, D. Cosmelli, and K. J. Friston. Evaluation of different measures of functional connectivity using a neural mass model. *NeuroImage*, 21: 659–673, 2004.
29. O. David and K. J. Friston. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage*, 20: 1743–1755, 2003.
30. O. David, L. Harrison, and K. J. Friston. Modelling event-related respones in the brain. *NeuroImage*, 25: 756–770, 2005.
31. A. van Rotterdam, F. H. Lopes da Silva, J. van den Ende, M. A. Viergever, and A. J. Hermans. A model of the spatial-temporal characteristics of the alpha rhythm. *Bulletin of Mathematical Biology*, 44(2): 283–305, 1982.
32. J. J. Wright and D. T. L. Liley. Simulation of electrocortical waves. *Biological Cybernetics*, 72: 347–356, 1995.
33. J. J. Wright and D. T. J. Liley. Dynamics of the brain at global and microscopic scales: neural networks and the EEG. *Behavioral and Brain Sciences*, 19: 285–320, 1996.
34. D. T. J. Liley, D. M. Alexander, J. J. Wright, and M. D. Aldous. Alpha rhythm emerges from large-scale networks of realistically coupled multicompartmental model cortical neurons. *Network: Computational. Neural Systems*, 10: 79–92, 1999.

35. C. Koch and I. Segev, editors. *Methods in Neuronal Modelling. From Ions to Networks*. Computational Neuroscience. MIT Press, Cambridge (MA), 1998.
36. A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal Physiology*, 117: 500–544, 1952.
37. V. K. Jirsa and H. Haken. Field theory of electromagnetic brain activity. *Physical Review Letters*, 77(5): 960–963, 1996.
38. V. K. Jirsa. Information processing in brain and behavior displayed in large-scale scalp topographies such as EEG and MEG. *International Journal of Bifurcation and Chaos*, 14(2): 679–692, 2004.
39. J. J. Wright, C. J. Rennie, G. J. Lees, P. A. Robinson, P. D. Bourke, C. L. Chapman, E. Gordon, and D. L. Rowe. Simulated electrocortical activity at microscopic, mesoscopic, and global scales. *Neuropsychopharmacology*, 28: S80–S93, 2003.
40. J. J. Wright, C. J. Rennie, G. J. Lees, P. A. Robinson, P. D. Bourke, C. L. Chapman, E. Gordon, and D. L. Rowe. Simulated electrocortical activity at microscopic, mesoscopic and global scales. *International Journal of Bifurcation and Chaos*, 14(2): 853–872, 2004.
41. P. A. Robinson, C. J. Rennie, J. J. Wright, H. Bahramali, E. Gordon, and D. L. Rowe. Prediction of electroencephalic spectra from neurophysiology. *Physical Reviews E*, 63, 021903, 2001.
42. D. O. Hebb. *The Organization of Behavior*. Wiley, New York (NY), 1949.
43. S. Kaplan, M. Sonntag, and E. Chown. Tracing recurrent activity in cognitive elements (TRACE): a model of temporal dynamics in a cell assembly. *Connection Science*, 3: 179–206, 1991.
44. F. van der Velde and M. de Kamps. Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences*, 29:37–108, 2006.
45. T. Wennekers, M. Garagnani, and F. Pulvermüller. Language models based on hebbian cell assemblies. *Journal of Physiology - Paris*, 100: 16–30, 2006.
46. P. Smolensky and G. Legendre. *The Harmonic Mind. From Neural Computation to Optimality-Theoretic Grammar*, Vol. 1: Cognitive Architecture. MIT Press, Cambridge (MA), 2006.
47. R. B. Stein, K. V. Leung, M. N. Oğuztöreli, and D. W. Williams. Properties of small neural networks. *Kybernetik*, 14: 223–230, 1974.
48. D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, editors. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. I. MIT Press, Cambridge (MA), 1986.
49. B. A. Pearlmutter. Learning state space trajectories in recurrent neural networks. *Neural Computation*, 1(2): 263–269, 1989.
50. B. A. Pearlmutter. Gradient calculations for dynamic recurrant neural networks: A survey. *IEEE Transaction on Neural Networks*, 6(5): 1212–1228, 1995.
51. P. Werbos. Back-propagation through time: What it does and how to do it. Vol. 78 of *Proc. IEEE*, pp. 1550–1560, 1990.
52. P. Werbos. Maximizing long-term gas industry profits in two minutes in Lotus using neural network models. *IEEE Transaction on Systems, Man, and Cybernetics*, 19: 315–333, 1989.
53. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipies in C*. Cambridge University Press, New York, 1996.

54. T. Liebscher. Modeling reaction times with neural networks using leaky integrator units. In K. Jokinen, D. Heylen, and A. Nijholt, editors, *Proc. 18th Twente Workshop on Language Technology*, Vol. 18 of *TWLT*, pp. 81–94, Twente (NL), 2000. Univ. Twente.
55. E. R. Kandel, J. H. Schwartz, and T. M. Jessel, editors. *Principles of Neural Science*. Appleton & Lange, East Norwalk, Connecticut, 1991.
56. P. Dayan and L. F. Abbott. *Theoretical Neuroscience*. Computational Neuroscience. MIT Press, Cambridge (MA), 2001.
57. D. J. Amit. *Modeling Brain Function. The World of Attractor Neural Networks*. Cambridge University Press, Cambridge (MA), 1989.
58. W. J. Freeman. *Mass Action in the Nervous System*. Academic Press, New York (NY), 1975.
59. W. J. Freeman. How and why brains create meaning from sensory information. *International Journal of Bifurcation and Chaos*, 14(2): 515–530, 2004.
60. C. S. Herrmann and A. Klaus. Autapse turns neuron into oscillator. *International Journal of Bifurcation and Chaos*, 14(2): 623–633, 2004.
61. J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Vol. 42 of *Springer Series of Appl. Math. Sciences*. Springer, New York, 1983.
62. R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1): 47–97, 2002.
63. S. Bornholdt and H. G. Schuster, editors. *Handbook of Graphs and Networks. From the Genome to the Internet*. Wiley-VCH, Weinheim, 2003.
64. B. Bollobás. *Random Graphs*. Cambridge University Press, Cambridge (UK), 2001.
65. P. beim Graben and J. Kurths. Simulating global properties of electroencephalograms with minimal random neural networks. *Neurocomputing*, doi:10.1016/j.neucom.2007.02.007, 2007.
66. A. J. Trevelyan and O. Watkinson. Does inhibition balance excitation in neocortex? *Progress in Biophysics and Molecular Biology*, 87: 109–143, 2005.
67. S. Itzkovitz, R. Milo, N. Kashtan, G. Ziv, and U. Alon. Subgraphs in random networks. *Physical Reviews E*, 68, 026127, 2003.
68. G. Bianconi and A. Capocci. Number of loops of size h in growing scale-free networks. *Physical Review Letters*, 90(7), 2003.
69. H. D. Rozenfeld, J. E. Kirk, E. M. Bollt, and D. ben Avraham. Statistics of cycles: how loopy is your network? *Journal of physics A: Mathematical General*, 38: 4589–4595, 2005.
70. O. Sporns, G. Tononi, and G. M. Edelman. Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex*, 10(2): 127–141, 2000.
71. J. Kornmeier, M. Bach, and H. Atmanspacher. Correlates of perceptive instabilities in visually evoked potentials. *International Journal of Bifurcation and Chaos*, 14(2): 727–736, 2004.
72. S. Frisch, P. beim Graben, and M. Schlesewsky. Parallelizing grammatical functions: P600 and P345 reflect different cost of reanalysis. *International Journal of Bifurcation and Chaos*, 14(2): 531–549, 2004.
73. H. Haken. *Synergetic Computers and Cognition. A top-down Approach to Neural Nets*. Springer, Berlin, 1991.
74. H. Haken. *Principles of Brain Functioning*. Springer, Berlin, 1996.

75. C. Lourenço. Attention-locked computation with chaotic neural nets. *International Journal of Bifurcation and Chaos*, 14(2): 737–760, 2004.
76. L. O. Chua. *CNN: A paradigm for complexity.* World Scientific, Singapore, 1998.
77. D. Bálya, I. Petrás, T. Roska, R. Carmona, and A. R. Vázquez. Implementing the multi-layer retinal model on the complex-cell cnn-um chip prototype. *International Journal of Bifurcation and Chaos*, 14(2): 427–451, 2004.
78. V. Gál, J. Hámori, T. Roska, D. Bálya, Zs. Borostyánköi, M. Brendel, K. Lotz, L. Négyessy, L. Orzó, I. Petrás, Cs. Rekeczky, J. Takács, P. Venetiáner, Z. Vidnyánszky, and Á Zarándy. Receptive field atlas and related CNN models. *International Journal of Bifurcation and Chaos*, 14(2): 551–584, 2004.
79. F. S. Werblin and B. M. Roska. Parallel visual processing: A tutorial of retinal function. *International Journal of Bifurcation and Chaos*, 14(2): 843–852, 2004.
80. M. T. Huber, H. A. Braun, and J.-C. Krieg. Recurrent affective disorders: nonlinear and stochastic models of disease dynamics. *International Journal of Bifurcation and Chaos*, 14(2): 635–652, 2004.
81. P. beim Graben, S. Frisch, A. Fink, D. Saddy, and J. Kurths. Topographic voltage and coherence mapping of brain potentials by means of the symbolic resonance analysis. *Physical Reviews E*, 72: 051916, 2005.
82. H. Drenhaus, P. beim Graben, D. Saddy, and S. Frisch. Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language*, 96(3): 255–268, 2006.
83. A. H. Kawamoto. Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language*, 32: 474–516, 1993.
84. A. Prince and P. Smolensky. Optimality: from neural networks to universal grammar. *Science*, 275: 1604–1610, 1997.