

Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Structure and dynamics of recurrent
neuromodules

by

Frank Pasemann

Preprint-Nr.: 38

1997



Structure and Dynamics of Recurrent Neuromodules *

Frank Pasemann
Max-Planck-Institute for Mathematics in the Sciences
D-04103 Leipzig, Germany
email: f.pasemann@mis.mpg.de

December 12, 1997

Abstract

The article calls attention to complex dynamical phenomena in artificial neural systems, which are - as is asserted - of relevance also for understanding biological brain functions. Examples of various dynamical effects (hysteresis, oscillations, deterministic chaos, synchronization and coherence) are discussed in terms of the discrete dynamics of small recurrent networks. The relevance of a dynamical systems approach for understanding the emergence of higher level information processing or cognitive abilities of biological and artificial neural systems is discussed.

*submitted for publication.

1 Introduction

Biological brains have a massively recurrent connectivity, i.e. there are very many directed closed signal processing loops on different scales of the brain: between single neurons as well as between hypercolumns and different brain areas which are functionally discernible. They involve excitatory as well as inhibitory connections. From the dynamical systems point of view, this is one reason to expect complex dynamics in biological brains. In fact oscillations of various periodicities are observed almost everywhere in the brain, as well as synchronization effects, wave patterns of activity, and even chaotic dynamics (Duke and Pritchard, 1991; Krüger, 1991; Freeman, 1992; Buzsáki et al. 1994; Elbert et al. 1994; Pantev et al., 1995). Therefore, it seems natural to assume that complex dynamics is the substantial basis for the cognitive abilities of biological brains (Kelso, 1995). Nevertheless, its significances and functional principles are still obscure. The intention of this article is to demonstrate with a few simple examples the abundance of dynamical features one should be aware of, when discussing the functional role of biological brain dynamics. Our line of arguments originates in analysis and simulations of discrete nonlinear dynamics of small artificial neural networks with recurrent architecture. We call them *neuromodules* because we are thinking of them as being interacting parts of a larger (cognitive) system.

As was pointed out already in the days of cybernetics (e.g. von Foerster 1969), the recurrent connectivity structure of neural systems will open up a new paradigm for discussing, modeling and understanding phenomena like perception and cognition. Recurrences will not just result in a simple input-output, i.e. reactive, behavior of a system, but instead, they will *create* certain processes, activity patterns or “images” for given stationary or dynamical input signals (see e.g. Johannesma et al. 1986, Harth 1993). Perception under this respect is therefore not an uni-directional process - like a mapping from the external world to some brain states. But it generates local dynamical activities already at lower levels of signal processing. By virtue of cooperative or competing interactions, these finally result in an appropriate behavior. And from this point of view cognition can not be understood any longer as pure rule based symbol processing (Port and van Gelder 1995). Thus, we believe that development and analysis of artificial recurrent neural networks can generate enough background material for controversial debates about the realization of perception, existence or non-existence of internal representations of the external world, static versus dynamical memories, and the like. Furthermore, artificial recurrent neural structures can also provide an extensive test-ground for hypotheses about neural processing principles.

Why is it so difficult to describe the possible role of complex dynamics for cognitive or higher-level information processing in more concrete terms? This is perhaps due to the fact, that we are still just beginning to understand the complex properties of interacting nonlinear dynamical systems. From a theoretical point

of view, there is the need to derive complexity measures on the basis of dynamical network properties, isomorphisms of network dynamics, more concrete definitions of concepts like self-organization and emergent properties. To understand the role of biological brain dynamics, one should have an idea about what kind of signals to look for, and how they can be correlated with observable motor action, thinking or general behavior.

With respect to artificial neural systems, one difficulty arises from the fact that there are in general no learning rules that can implement appropriate dynamic attractors into the phase space of recurrent neural networks. Furthermore, most of the time it is not at all clear what kind of recurrent architecture to use for a specific cognitive task. Even worse, in systems acting and learning in a given environment, knowledge of appropriate internal dynamics responding to particular (time-varying) sensor inputs is commonly not at hand; so any kind of supervised learning will not be applicable.

Therefore, today mainly *convergent* (Hirsch 1989) neural networks are examined and utilized for technical purposes. Typical representatives are the feedforward networks, Hopfield and Kohonen networks (Hertz et al. 1991). Convergent networks are most effectively applied, if for given stationary inputs a definite stationary output signal is desired, as is the case e.g. for pattern classification tasks, or for image recognition or completion. More complex, i.e. higher information processing or cognitive tasks, like generation of temporal sequences, recognition, storage and reproduction of temporal patterns, or the control of systems which requires memory to compute derivatives or integrals, are in general not easily achieved with convergent networks. This is the typical domain of applications for recurrent *non-convergent* networks, which are themselves capable of having complex internal dynamics.

Besides recurrent connectivity, modularity is a basic structural property of biological brains - and also an effective design principle for artificial systems. So, when studying the the dynamics of neuromodules, as we do in the following section, one should keep in mind that a cognitive system will be composed of many such modules interacting in a cooperative or competing way to produce a desired behavior. Although the dynamical effects outlined in section 2 are found for the discrete dynamics of neuromodules with simple formal neurons (additive units with sigmoidal transfer functions), it is argued that a comparable complexity should be expected also for systems with continuous time dynamics, and also for neural networks with more biology-like spiking neurons. In section 3 we discuss the examples of section 2 with respect to a dynamical behavior-oriented approach (compare e.g. Mallot 1997) to cognitive systems.

2 Discrete dynamics of recurrent neuromodules

In terms of dynamical systems theory (Guckenheimer and Holmes 1983; Jackson 1991; Abraham and Shaw 1992; Ott 1993) the behavior of a (time-continuous) system under consideration is described by the “flow” of its states in time. This is determined usually by a system of coupled first order differential equations. For every *initial condition*, i.e. a state $x(0)$ at time $t = 0$, there exists a unique “stream line” of the flow passing through $x(0)$: the corresponding solution of the differential equations. The dynamics of dissipative systems, like neural networks, is characterized by its *attractors*. These are sets of states which are permanently revisited, either periodically (*periodic attractors*) or apparently “chaotic” (*chaotic attractors*). There are also attracting stationary states, called *fixed point attractors*. As the naming suggests, the states in the neighborhood of attractors are flowing towards these sets, i.e. for $t \rightarrow \infty$ they approach them arbitrarily close. The *basin* of an attractor consists of all states, which flow to the attractor for $t \rightarrow \infty$. If there exists only one attractor for the system, it is called a *global attractor*. If there are co-existing attractors, the state space is partitioned into their basins. For chaotic attractors the basin boundaries can be fractal. If a system has only fixed point attractors we call it *convergent*, otherwise *non-convergent*.

The dynamics of a system usually depends on *control parameters*, that is, on variables that vary much more slowly than the states of the system. The behavior of a system then may change qualitatively, i.e. the type of attractors will change at certain critical parameter values, called *bifurcation points*. A sequence of such transitions is then displayed by a corresponding *bifurcation diagram*.

In the following we will discuss the discrete activation dynamics of neural networks with n units corresponding to a map $f : \mathbf{R}^n \rightarrow \mathbf{R}^n$ given by

$$a_i(t+1) = \theta_i + \sum_{j=1}^n w_{ij} \cdot \sigma(a_j(t)), \quad i, j = 1, \dots, n, \quad (1)$$

where a_i denotes the *activity* of neuron i , w_{ij} the synaptic strength or *weight* of the connection from neuron j to neuron i , and the term $\theta_i = \theta_{i0} + I_i$ will be considered here as the sum of a fixed internal bias θ_{i0} and a stationary external input I_i of the i th neuron. The output o_i of neuron i is given by a sigmoidal (S-shaped) transfer function, and we use the strictly positive sigmoide given by $\sigma(x) = (1 + e^{-x})^{-1}$. Since it is assumed, that the bias terms θ_i and the weights w_{ij} vary only slowly compared with the activity of the neurons, they are treated here as parameters for the activation dynamics.

One should keep in mind that using stationary external inputs for the modules is primarily done for the *classification* of dynamical module properties. Modules in biological brains (or advanced artificial systems like autonomous robots) will never get stationary inputs, neither from the sensors nor from other parts of the system. So their dynamics in general will be of *transient* type, which is still classified by a corresponding attractor. It is on this background, that we believe

“cognitive entities” to be best represented by *basins of attraction* (Pasemann 1996).

2.1 “One-loop” networks

The simplest type of recurrent networks with n neurons have only one closed directed loop of connections. They are called n -ring networks (Pasemann 1995a), and their discrete dynamics is given by the map

$$a_i(t+1) := \theta_i + w_{ii-1}\sigma(a_{i-1}(t)), \quad a_0 := a_n, \quad w_{10} := w_{1n}, \quad i = 1, \dots, n.$$

For the special case $n = 1$ we have a single neuron with self-interaction (Pasemann 1993). We classify n -ring networks and their dynamical properties corresponding to their loop property: They are called *even* (*odd*), if the number of inhibitory connections in the signal loop is even (odd). Besides having global fixed point attractors for large parameter domains, n -rings display the following dynamics: For even networks there exists a parameter domain (*hysteresis domain*), where two fixed point attractors co-exist with roughly $(2^n - 2)/n$ periodic attractors, mainly with period n ; for odd networks there exists a parameter domain (*oscillation domain*), where we find roughly $2^n/2n$ co-existing periodic attractors, mainly with period $2n$. For example an even 16-ring can have 4080 period-16, 30 period-8, three period-4 and one period-2 attractors, all co-existing (for fixed parameters) with two fixed point attractors. An odd 15-ring will have 1096 co-existing periodic attractors: 1091 with period 30, three with period 10, one with period 6 and one period-2 attractor (for details see Pasemann 1995a). The “firing pattern” corresponding to one of the 1091 period-30 attractors is shown in figure 1. Note that although the pattern is periodic in time, at every time slice of the period different neurons are active.

Not only the large number of co-existing periodic attractors is interesting, but also a specific role of the inhibitory connections should be noted: They arrange the neurons of a ring into groups with identical but phase shifted activity patterns. If the outputs of the ring neurons are projected on an additional neuron then every attractor generates a specific oscillatory wave form at its output (Pasemann 1995a). Thus for certain fixed parameters n -ring networks are able to generate a large variety of different activity patterns or wave forms that are selected by initial conditions. If an n -ring network is in a specific oscillatory mode, slight changes of parameters (inside the hysteresis or oscillatory domain) will not change this mode. On the other hand, moving the parameters - external inputs, bias terms or weights - slowly across the bifurcation sets, ring networks may be used as switchable multi-frequency oscillators.

For a self-excitatory single neuron (and also for even n -rings), in addition a hysteresis effect can be observed: if parameters are varied back and forth across the hysteresis domain, on which the system is bi-stable (plus oscillatory modes

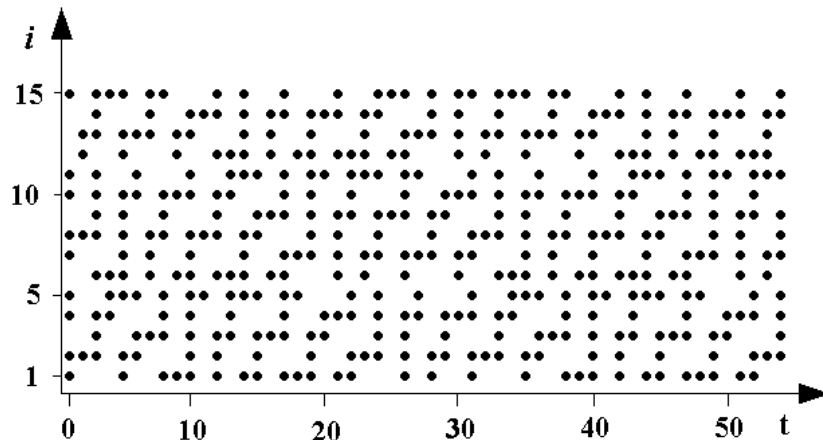


Figure 1: The firing pattern on a period-30 attractor of an odd 15-ring module: high activity of neuron i is marked by black dots.

in the case of n -rings), the system “jumps” from one stable state into the other at different bifurcation points (Pasemann 1993, Pasemann 1995a).

2.2 “Two loop” networks

Networks with only two closed directed connection loops can already display the full spectrum of complex dynamical behavior. This can be demonstrated already for the two simplest architectures shown in figure 2. They have two neurons (5 parameters) and three neurons (7 parameters), respectively. In both cases there is an inhibitory neuron involved. The excitatory units may represent for instance pyramidal cells and the inhibitory ones inter-neurons like stellar cells, a standard combination found in the columns of the cortex. The discrete dynamics of the 2-neuron network is e.g. given by

$$\begin{aligned} a_1(t+1) &:= \theta_1 + w_{12} \sigma(a_2(t)) \quad , \\ a_2(t+1) &:= \theta_2 + w_{21} \sigma(a_1(t)) + w_{22} \sigma(a_2(t)) \quad ; \end{aligned} \quad (2)$$

and the dynamics of the 3-module by

$$\begin{aligned} a_1(t+1) &:= \theta_1 + w_{12} \sigma(a_2(t)) + w_{13} \sigma(a_3(t)) \quad , \\ a_2(t+1) &:= \theta_2 + w_{21} \sigma(a_1(t)) \quad , \\ a_3(t+1) &:= \theta_3 + w_{31} \sigma(a_1(t)) \quad , \end{aligned} \quad (3)$$

For these modules we observe (e.g. Pasemann 1995b) global fixed point attractors as well as various periodic and chaotic attractors. In fact, there are period doubling routes to chaos. But also quasi-periodic attractors can be found for certain parameter values. Figure 3, for example, shows a typical bifurcation diagram

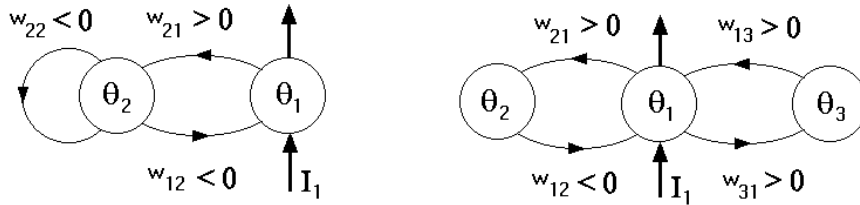


Figure 2: Two configurations with interesting dynamical features: Stationary states, oscillations and chaos, co-existing attractors and hysteresis effects.

for the 2-module with θ_1 as control parameter and $\bar{o} := 0.5 \cdot (o_1 + o_2)$ denoting the averaged output of the module. Starting with period-2 attractors we observe quasi-periodic attractors (around α), larger windows with periodic attractors (periods 5 and 8 around β) and chaotic attractors reached via period-doubling (around γ). The distinction between attractors of higher periods, quasi-periodic and chaotic attractors can be verified also by calculating the corresponding Lyapunov exponents (see for example Ott 1993).

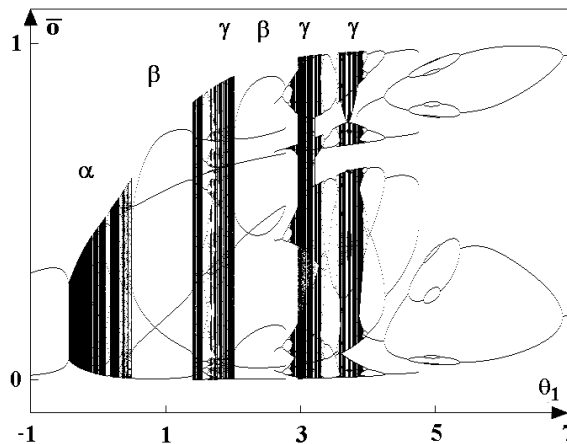


Figure 3: Bifurcation diagram for the 2-module with fixed parameters $\theta_2 = -0.44$, $-w_{12} = w_{21} = 8$, $w_{22} = -16$ and varied input to the excitatory neuron.

Again, in both configurations (2), (3) co-existing attractors can be found. For example, a periodic and a chaotic attractor co-exist for the same parameter values, as is shown for the 2-neuron network (2) in figure 4a. But even two chaotic attractors may co-exist, as is the case for the 3-module (3). If a chaotic attractor co-exists with other attractors fractal basin boundaries are frequently observed. They partition the phase space in an irregular way. This is demonstrated in figure 4b for the 2-module (2). The essential effect of this irregular decomposition of phase space is, that different attractors can be reached from one and the same

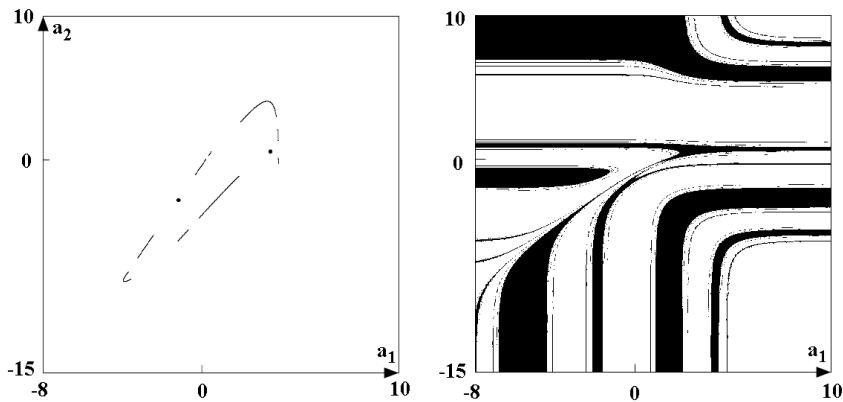


Figure 4: Coexisting attractors for the 2-module: a.) period-2 and chaotic attractor for $\theta_1 = 3.9$, $\theta_2 = -0.44$, $-w_{12} = w_{21} = 8$, $w_{22} = -16$, b.) corresponding basins with fractal boundary (black: period-2).

phase space domain. This is of course quite different from the regular “decision regions” obtained for instance by fixed point attractor networks like Hopfield networks. Keeping in mind that real neurodynamics is living on transients, the advantage of irregular phase space partition is that switching between different basins of attraction or represented “cognitive entities”) can be caused by small disturbances. These can be generated, for instance, by short (or pulsed) external signals.

2.3 More loops

It seems plausible from these observations that systems with more than two closed signal loops can display even more complex dynamical features. We will demonstrate this again for the simple example by *bi-directional n-rings* or *n-chains*. Their discrete dynamics is given by

$$a_i = \theta_i + w_{ii+1}\sigma(a_{i+1}) + w_{ii-1}\sigma(a_{i-1}), \quad w_{nn+1} = w_{n1} \quad w_{10} = w_{1n}, \quad (4)$$

where $w_{1n} = w_{n1} = 0$ for *n-chains*.

As simulations show, purely excitatory or inhibitory bi-directional *n-rings*, have, besides parameter domains for global fixed point attractors, only domains for which there is a (large) number of co-existing period-2 and fixed point attractors. At least for the symmetric case (i.e. $w_{ij} = w_{ji}$) this can be proven analytically by applying the technique of Lyapunov functions as used for instance in (Herz 1996). Furthermore, it is observed in simulations that these period-2 attractors appear as groups of neighboring neurons with synchronously oscillating activity. This is quite different from the situation we observed for *n-rings* as in figure 1. It can be read from the typical firing pattern corresponding to an

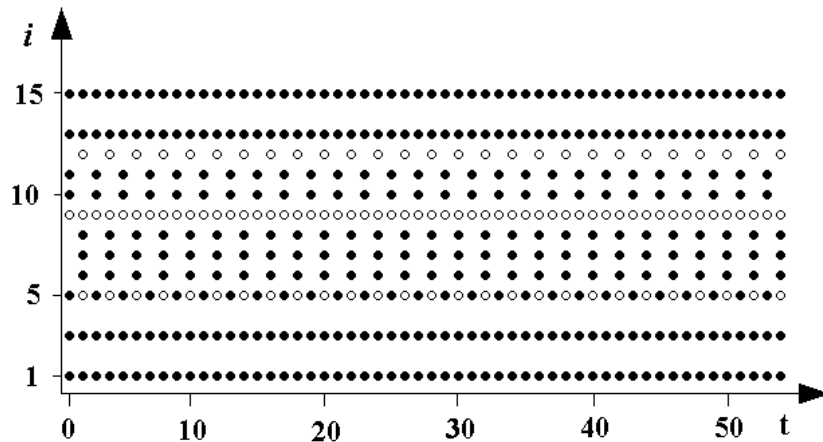


Figure 5: The firing pattern on a period-2 attractor of a bidirectional 15-ring module: high activity of neuron i is marked by black dots, medium activity by circles.

attractor of an (even) bi-directional 15-chain shown in figure 5. There neurons 1, 3, 13 and 15 have constant high, neurons 2, 4 and 14 constant low activity, and neuron 9 constant medium activity. The synchronously active groups of neurons (6,7,8) and (10,11) are alternately “firing”. Neurons 5 and 12 display periodic activity.

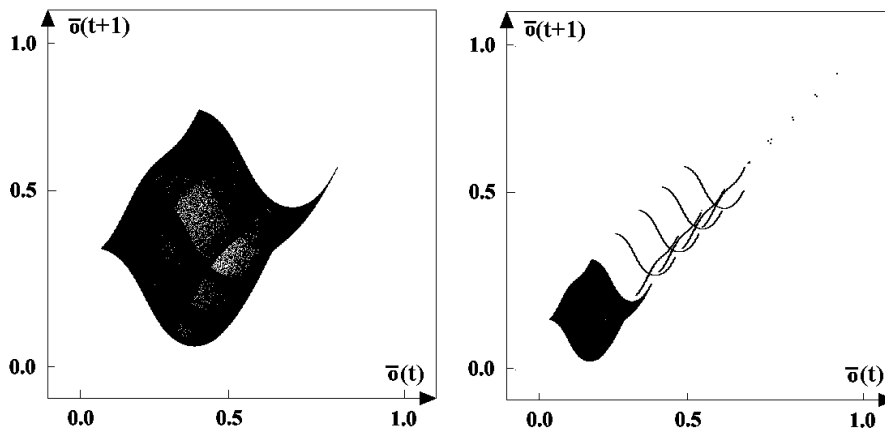


Figure 6: a.) A chaotic attractor for the 3-module (see text), b.) Coexisting chaotic, period-2 and fixed point attractors for a bi-directional 7-chain (see text).

Introducing only one inhibitory connection in an otherwise excitatory bi-directional chain or ring has already a dramatic effect. This can be seen by the following construction: Take the two loop network (3) with parameters corresponding to a chaotic attractor; for example $-w_{12} = w_{21} = w_{31} = w_{13} = 8.0$ and

$\theta_1 = 4.5$, $\theta_2 = -2.7$, $\theta_3 = -8.0$. The corresponding chaotic attractor is displayed in figure 6a by the first return map of the averaged output $\bar{o}(t) := \frac{1}{n} \sum_i^n o_i(t)$ of an n -module. Then add an excitatory neuron 4 connected bi-directionally with the excitatory end-neuron 3 according to $\theta_4 = -8.0$ and $w_{34} = w_{43} = 8.0$. In the resulting 4-dimensional phase space we now observe two chaotic attractors co-existing with a fixed point attractor. Adding more excitatory neurons in the same way, we observe the systematic appearance of $n - 2$ chaotic attractors co-existing with several fixed point and period-2 attractors. This is shown for a bi-directional 7-chain in figure 6b, where any crossed structure corresponds to an additional chaotic attractor. Also some fixed points (on the main diagonal) and period-2 attractors can be seen. Technically, there is a characteristic difference observed in bifurcation diagrams between the “original” chaotic attractor (seen for the 3-module) and the additional ones (in n -chains): The original one appears in a period doubling bifurcation sequence starting from a fixed point with a strong resonance of period four; the additional chaotic attractors come from (co-existing) period-doubling bifurcations starting with a period-2 orbit.

As with n -rings, again we have a neural structure where the total number of co-existing attractors $N(n)$ grows faster than n . This can be observed in simulations, but the systematics of periodic attractor appearance may be deduced also analytically. For example for $n = 5$ we have three chaotic attractors co-existing with two fixed point and one period-2 attractor, i.e. $N(5) = 6$. Different from n -rings, since chaotic attractors are involved, the basin structure will be fractal and the phase space will be partitioned in complex “decision domains”. It is an interesting feature that some of these attractors can be brought to extinction by slightly varying the external inputs of corresponding added neurons. Thus, the appearance and disappearance of specific attractors may be controlled by slow external input signals.

2.4 Synchronization and coherence in coupled modules

Stimulus induced coherent firing of neurons in biological brains is nowadays a well established experimental fact. In particular synchronous neuron activity is often discussed as a fundamental temporal mechanism for feature binding and integration of distributed brain processes (cf. among many others Eckhorn et al., 1988; Gray et al., 1989; Singer, 1994; Engel et al., 1997). Since we are thinking in terms of modular systems, synchronization is a result of coupling modules. We want to point out that again the situation may be even more complex than discussed in the standard literature on synchronization effects in biological brains. We refer to observations made in simulations with two coupled identical chaotic 2-modules of the type described above (figure 2a), here with parameters $\theta_2 = \theta'_2 = -1$ and $w_{12} = w'_{12} = -8$, $w_{21} = w'_{21} = 8$; θ_1 and θ'_1 correspond to the varying external inputs. We consider only the (inhibitory) coupling from the inhibitory neuron of a module to the excitatory neuron of the

other module as shown in figure 7, where coupling weights are set to $w = w' = -4$. This type of coupling seems to be most efficient with respect to synchronization *and* de-synchronization of modules.

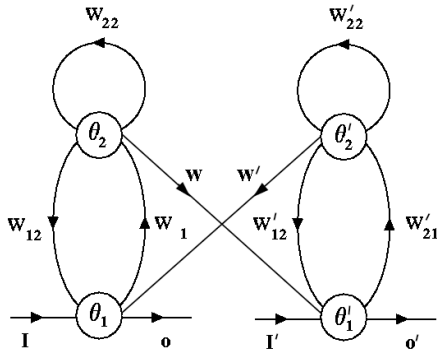


Figure 7: Coupling of two chaotic neuromodules.

The synchronization of the 2-modules appears over a large range of (approximately equal) stationary external inputs to the excitatory units. This effect is robust, that is, it appears also if the internal parameters, for instance the module weights, are slightly varied. Modules de-synchronize immediately after the external input signals diverge. But the following observation is more noteworthy: For certain (equal) input signals there co-exist different attractors in the coupled system; for instance two periodic attractors, but also periodic attractors with quasi-periodic or chaotic attractors. On one of these attractors the dynamics of the modules synchronizes, on the other it is coherent in the sense, that there is a fixed (though sometimes complex) phase relation. Furthermore, synchronization can occur as well on periodic attractors as on quasi-periodic or chaotic attractors. This is demonstrated in figures 8, 9, and 10) where the projections of co-existing attractors onto the phase space of one module (figures a.) and onto the module output space (figures b.) is depicted. Synchronized outputs will appear as states lying on the main diagonal in (o_2, o'_2) -space.

The essential point is, that with fixed parameters and stationary (equal) input signals synchronization of module dynamics can depend on initial conditions, that is, on the “history” of the coupled system. Furthermore, if the coupled system starts in a synchronized mode, then it persists in this mode even if the external inputs $I := I_1 = I'_1$ are varying slowly. It appears, for example, that the system starts in a synchronous mode at $I = 3.0$ on the (then global) period-3 attractor and, with increasing inputs I , switches to the synchronous chaotic mode around $I_1 = I'_1 = 7.0$; and back again to the period-3 mode with I decreasing. So synchronization of these coupled modules is really a sign for time-varying signals with *fixed ratio* (recall, that the inputs may correspond to the weighted outputs of other neurons of the system).

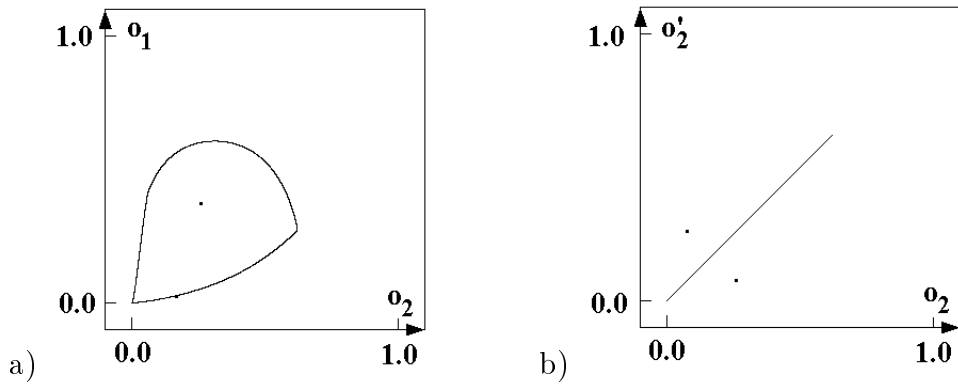


Figure 8: Coexistent attractors in coupled 2-modules: a.) Projection onto the phase space of one module. b.) Outputs of the excitatory units (synchronization corresponds to points on the main diagonal). Here: synchronization on quasi-periodic attractor, coherence on period-2 attractor.

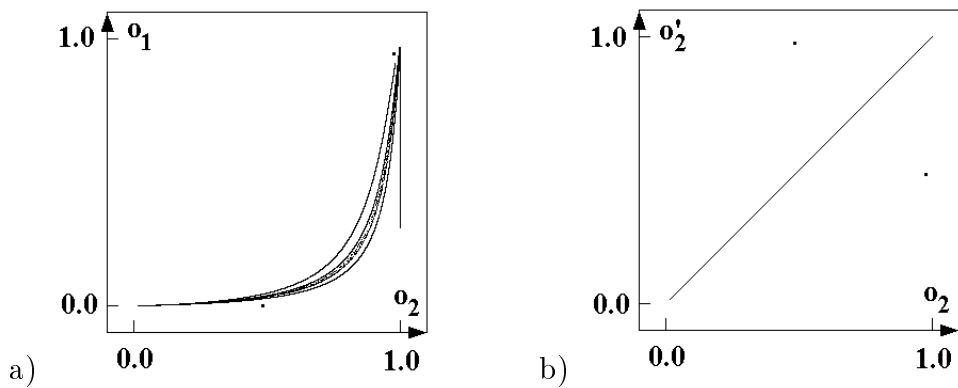


Figure 9: Coexistent attractors in coupled 2-modules: synchronization on chaotic attractor, coherence on period-2 attractor.

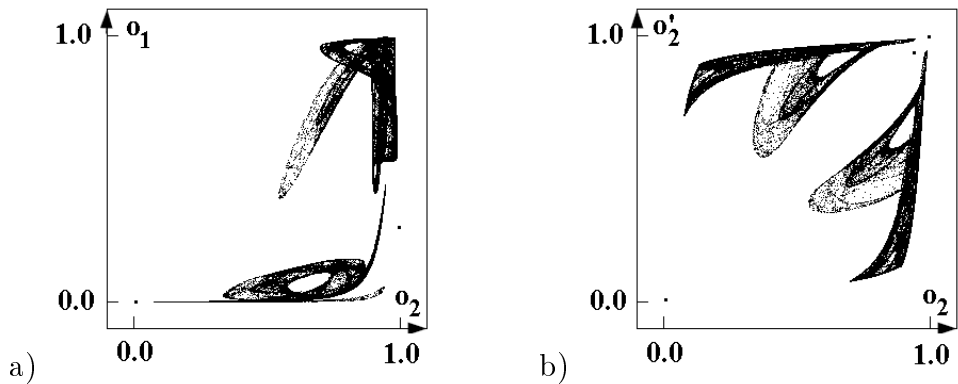


Figure 10: Coexistent attractors in coupled 2-modules: synchronization on period-3 attractor, “coherence” on chaotic attractor.

Similar observations are made with a slightly different setup, where self-inhibition is replaced by self-excitation. This corresponds to biologically motivated configurations, which were studied for instance by Nagashino and Kelso (1991) and Cohen et.al. (1993). They also reported the co-existence of in-phase and anti-phase modes and the switching between them, although in a quite different dynamical context. In this second setup, and for the corresponding discrete dynamics, the isolated modules will have no chaotic attractors. For these coupled modules synchronous modes are therefore observed only for periodic and quasiperiodic attractors.

3 Why and what dynamics?

Is there something to learn from the simple examples presented above; can we learn something for a theoretical description of biological brain function? The main message is of course, that the complex dynamical properties of nonlinear systems - and biological brains are asserted to be of this kind - provide an immense reservoir for possible signal processing capabilities, a domain of intricate procedures or methods one can hardly think of today. Therefore at least the interest of neuro-scientists in complex neuro-dynamical phenomena is justified. Let us first summarize the main effects reported in this paper.

For all the simple neural structures of the above examples we observed dynamical attractors for specific parameter domains. As was mentioned earlier, various kinds of oscillations, and perhaps also chaotic activity, are also characteristic for biological brains. They are observed on different functional levels and anatomical scales. From the last section it becomes obvious that complex dynamical properties are associated with the presence of both, excitatory as well as inhibitory connections in recurrent neuromodules. And we can specify complex dynamics of a neuromodule to be represented, for instance, by the existence of bifurcation sequences involving not only periodic but also quasi-periodic and chaotic attractors (like the one in figure 3). One may infer that from the dynamical systems point of view the essential role of inhibitory connections is to keep the system near critical parameter values, where its behavioral flexibility is largest. This conception is close to the idea of “computing at the edge of chaos” (Langton 1990).

All of the above described modules allowed the co-existence of dynamic attractors, and often the number of of co-existing attractors grows faster than the number of neurons in corresponding architectures. This becomes exceptionally manifest already for the n -ring networks of section 2.1. Thinking about dynamic attractors - or better their basins - as representing particular “cognitive entities”, can lead to new models for dynamic memories, attention mechanisms, and the like. The hypothesis is that dynamic attractors are a much more flexible instrument for higher level information or cognitive signal processing than the fixed

point attractors of classical neural networks. They only code for signal classes, specific stationary patterns or images, or for probability distributions in signal space. But dynamic attractors may store whole behavior patterns, or “world” models for predicting external input signals. Dynamic recurrent modules may be effectively used, for example, in situations where short-time memory is required to “compute” derivatives or integrals, or where temporal activity patterns have to be recognized, stored or reproduced. Moreover, there is also the technical argument that in a phase space of given dimension (= number of neurons) the set of possible fixed point attractors is finite dimensional, where as the set of possible dynamic attractors is an infinite dimensional space.

Very often hysteresis effects are associated with co-existing attractors. In fact, they do not only appear for the well known example of two co-existing fixed point attractors (Pasemann 1993), where the systems jumps from one stable state to the other at different critical parameter values when crossing the hysteresis domain from different sides. They can also be observed for periodic, and even chaotic attractors; for example, in the 3-module given by equation (3). This kind of a more general hysteresis phenomenon may also underly the behavior observed on various levels of sensorimotor signal processing. Classical examples the perspective ambiguity, the figure-ground ambiguity, and the ambiguity of meaning in visual perception. Co-existing attractors can shed also new light on synchronization phenomena observed in biological brains. Usually it is assumed that these cooperative brain states acquire “meaning” because they are observed as a stimulus driven effect or as associated with meaningful behavioral events. From the example of coupled chaotic modules in section 2.5 one can learn that synchronization of modules is a very general effect, existing for a broad range of external input values and internal module parameters. These modules also desynchronize immediately after input signals diverge. But we also saw that for different values of equal inputs synchronization takes place on different periodic or even chaotic attractors. Furthermore, for one and the same set of specific input signals different coherent states can appear, as can be read from figure 9: for example synchronous chaotic dynamics and asynchronous periodic dynamics.

Dynamic attractors of recurrent networks can be characterized and made “visible” by varying spatial and temporal activity patterns of corresponding groups of neurons. Although the system may live on a periodic attractor, that is, spatial patterns of activity will reappear periodically, there must not be a clustering or “grouping” of neurons in the sense of always synchronously active neurons. Compare, for example, the firing patterns of n -rings in figure 1). Furthermore, attractors may “live” effectively in lower dimensional subspaces. This means that some of the neurons have constant activity although the module dynamics is on a periodic attractor. A typical example is that of bidirectional n -rings: Here groups of neurons “fire” synchronously with period two, and there are also neurons having constant low or high output (compare figure 1).

One more remark with respect to the interpretation of dynamic attractors as

representing “cognitive entities” should be made here. It concerns the problem of switching between different attractors, or “cognitive entities”, respectively. We realized already with the examples of the last section that there are different ways of achieving this. 1.) If there are several co-existing attractors for a fixed set of weights and bias terms then an attractor is selected by “pulsed” inputs setting the initial conditions in one of the corresponding basins of attraction. This corresponds, for example, to the situation given by n -ring networks (section 2.1), and it is an analog to the case of Hopfield networks where a pattern is given as initial condition and then the network dynamics relaxes to a fixed point attractor. 2.) An attractor may be selected for a given (stationary) input by slowly varying a parameter like the synaptic strength of a connection along a bifurcation sequence; or, with weights and bias terms fixed, by slowly varying the external inputs (compare figure 3). But there exists an even more intricate mechanism: 3.) The infinitely many unstable periodic orbits spanning a chaotic attractor may be stabilized by a so called chaos control (see e.g. Ott 1993). It was shown (Stollenwerk and Pasemann 1996) that this technique can be implemented consistently in the neural network framework: a chaotic 2-module like the one given by equation (2) is controlled by neuromodules, each of which stabilizes a given periodic orbit. Both, deterministic and stochastic switching between different periodic attractors is possible.

The described dynamical phenomena are of course not specific for the case of discrete time dynamics considered here. In general they reflect the (perhaps higher dimensional) continuous-time dynamics of corresponding neural systems. In general they will be given in terms of ordinary differential equations with *delays*. Furthermore, the described phenomena are not limited to the very small modules of our examples. In fact, they describe also dynamical properties of very large systems in the following sense: Given a large network consisting of randomly coupled neuron pools, then in the limit of infinite time and infinite network size, the network dynamics corresponds to the module dynamics, where the activity of module units corresponds to the (synchronized) activity of pool neurons (for details see Wennekers and Pasemann 1996). In these formal neural systems the neuron model is of course not very biological. But in simulations with corresponding small modules of spiking neurons (of the leaky-integrate-and-fire type) we found at least hysteresis effects (bi-stability) and oscillations. In this case state variables were represented by inter-spike intervals, and co-existing periodic attractors and deterministic chaos are of course not so easily identified.

One essential point we want to re-emphasize here is that the dynamics of recurrent modules is parametrized by quantities like synaptic strength of connections, bias terms, and the like. Recall that specific variables are signified as parameters if their time-like changes are much slower than the activity changes of neurons. Because it is well known that in biological systems corresponding quantities can be manipulated on various (slow) time scales by complicated electrical and/or biochemical processes, for instance learning by changing synaptic

strength, parametrized dynamics seems to be the real basis for understanding cognitive brain functions.

It is also due to their parametrized dynamics that we understand modules as multi-functional units of signal processing systems: during their interaction with other modules or larger parts of the system their parameters can be changed so that a context-dependent processing of “peripheral” signals becomes possible; and, on a slower time-scale, “learning” of parameter domains for which the module responds with an effective dynamics to input signals becomes possible. Following this point of view, we can further deduce that isolated modules are functionally not determined. They receive their specific functional properties only during cooperative or competitive interactions with other modules of the system. This also leads us to the assumption that chaotic modules are good candidates for basic building blocks of modular cognitive systems: not because chaotic dynamics must be ever realized during interaction in an assemblage of modules, but because chaotic modules are endowed with the full spectrum of dynamical behavior. Almost every type of periodic, quasi-periodic and chaotic dynamics can be selected by appropriate tuning of parameters (compare the bifurcation diagram of figure 3 as an example).

Assuming that biological brains can be modeled as modular systems, the dynamical systems approach will assert that cognitive abilities, or cognitive processes, of these systems will emerge as global processes through the many cooperative or competitive local interactions of modules. With respect to this view, it possibly does not make sense to localize any of the various cognitive abilities in the form of functional neural assemblies or a local connectivity structure, although specific modules or ensembles of modules may be involved in processing subtasks.

In the context of artificial neural networks dynamical properties have not been considered as being essential up to now. This is mainly because one follows in general a constructive approach to artificial systems, that is, one examines convergent networks with pre-designed architectures and a given learning rule. In fact, most of the known learning rules work effectively only on convergent networks.

To learn something about the principles of neural signal processing we favor an emergence-oriented approach to artificial neural networks. For this type of investigations the dynamical systems framework is fundamental. To single out dynamical properties that relate to “cognitive” behavior we have to consider systems acting in a sensorimotor loop. We suggest that using evolutionary algorithms to generate appropriate recurrent dynamic networks (Pasemann and Dieckmann 1997, Pasemann 1997) is a reasonable way to follow this behavior-oriented approach.

4 Conclusion

The aim of this article was to call attention to the large variety of dynamical phenomena inherent in formal neural systems with recurrent connectivity. The experimental exploration of biological brain dynamics is still in its beginning, as well as work on the conceptual and theoretical foundations for this field. We claimed that experiences with complex dynamical properties of artificial neural networks can provide guiding principles for modeling biological brain functions.

References

- [1] Abraham, R.H.; Shaw, C.D. (1992) *Dynamics - The Theory of Behavior*, 2nd. ed., Addison-Wesley, Redwood City.
- [2] Buzsáki, G.; Llinás, R.; Singer, W.; Berthoz, A.; Christen, Y. (eds.) (1994) *Temporal Coding in the Brain*, Springer, Berlin.
- [3] Cohen, M. A.; Grossberg, S.; Pribe, C. A. (1993) Quadruped gait transitions from a neural pattern generator with arousal modulated interactions. In: *Proceedings of the World Congress on Neural Networks*, Portland, Oregon July, 1993, Vol. II. pp. 610-613.
- [4] Duke, W.; Pritchard, W.S. (eds.) (1991) *Proceedings of the Conference on Measuring Chaos in the Human Brain*, World Scientific, Singapore.
- [5] Eckhorn, R.; Bauer, R.; Jordan, W.; Brosch, M.; Kruse, W.; Munk, M.; Reitboeck H. J. (1988) Coherent oscillations: a mechanism for feature linking in the visual cortex. *Biol. Cybern.* 60: 121-130.
- [6] Elbert, T.; Ray, W.J.; Kowalik, Z.J.; Skinner, J.E.; Graf, K.E.; Birbaumer, N. (1994) Chaos and physiology: Deterministic chaos in excitable cell assemblies. *Physiological Review* 74: 1-47.
- [7] Engel, A. K.; Roelfsema, P. R.; Fries, P.; Brecht, M.; Singer, W. (1997) Binding and response selection in the temporal domain - a new paradigm for neurobiological research? *Theory in Biosciences* 116: 241-266.
- [8] Freeman, W. J. (1992) Tutorial on neurobiology: From single neurons to brain chaos. *Int. J. Bif. and Chaos* 2: 451-482.
- [9] Gray, C. M.; König, P.; Engel, A. K.; Singer, W. (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338: 334-337.
- [10] Guckenheimer, J.; Holmes, P. (1983) *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer Verlag, New York.

- [11] Harth, E. (1993) *The Creative Loop - How the Brain makes a Mind*. Addison-Wesley, Reading.
- [12] Hertz, J. A.; Krogh, A.; Palmer, R.G. (1991) *Introduction to the Theory of Neural Computation*, Addison-Wesley, Redwood City.
- [13] Herz, A. V. M. (1996) Global analysis of recurrent neural networks. In: Doman, E.; van Hemmen, J.L.; Schulten, K. (eds.) *Models of Neural Networks*, Vol. III. Springer Verlag, New York.
- [14] Hirsch, M. (1989) Convergent activation dynamics in continuous time networks. *Neural Networks* 2: 331-350.
- [15] Jackson, E.A. (1991) *Perspectives of Nonlinear Dynamics*, Cambridge University Press, Cambridge.
- [16] Johannesma, P.; Aertsen, A.; van den Boogaard, H.; Eggermont, J.; Epping, W. (1986) From synchrony to harmony: ideas on the function of neural assemblies and on the interpretation of neural synchrony. In: Palm, G.; and Aertsen, A. (eds.) *Brain Theory*. Springer Verlag, Berlin.
- [17] Kelso, J. A. S. (1995) *Dynamic Patterns - The Self-Organization of Brain and Behavior*, MIT Press, Cambridge.
- [18] Krüger, J. (ed.) (1991) *Neural Cooperativity*, Springer Verlag, Berlin.
- [19] Langton, C. G. (1990) Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D* 42: 12-37.
- [20] Mallot, H. A. (1997) Behavior-oriented approaches to cognition: Theoretical perspectives. *Theory in Biosciences* 116: 196-220.
- [21] Nagashino, H; Kelso, J. A. S. (1992) Phase transitions in oscillatory neural networks. *Science of Artificial Neural Networks*, SPIE 1710: pp.279-287.
- [22] Ott, E. (1993) *Chaos in Dynamical Systems*, Cambridge University Press, Cambridge.
- [23] Pantev, C.; Elbert, T.; Lutkenhöner B. (eds.) (1995) *Oscillatory event-related brain dynamics*, Plenum Press, London.
- [24] Pasemann, F. (1993) Dynamics of a single model neuron. *Int. J. Bif. and Chaos* 3: 271-278.
- [25] Pasemann, F. (1995a) Characteristics of periodic attractors in neural ring networks. *Neural Networks* 8: 421-429.

- [26] Pasemann, F. (1995b) Neuromodules: A dynamical systems approach to brain modeling. In: Herrmann, H. J.; Wolf, D. E.; Pöppel, E. (eds.) Supercomputing in Brain Research: From Tomography to Neural Networks. World Scientific, Singapore. pp. 331-348.
- [27] Pasemann, F. (1996), Repräsentation ohne Repräsentation - Überlegungen zu einer Neurodynamik modularer kognitiver Systeme. In: Rusch, G.; Schmidt, S. J.; Breidbach, O. (eds.) Interne Repräsentationen - Neue Konzepte der Hirnforschung. Suhrkamp Verlag, Frankfurt.
- [28] Pasemann, F.; Dieckmann, U. (1997) Evolved neurocontrollers for pole-balancing. In: Mira, J.; Moreno-Diaz, R.; Cabestany, J. (eds.) Biological and Artificial Computation: From Neuroscience to Technology. Proceedings IWANN'97, Lanzarote, Canary Islands, Spain, June 1997. Springer Verlag. pp. 1279 - 1287.
- [29] Pasemann, F. (1997) Pole-balancing with different evolved neurocontrollers. In: Gerstner, W.; Germond, A.; Hasler, M.; Nicoud, J.-D. (eds.) Artificial Neural Networks - ICANN'97. October 7-10, Lausanne, Switzerland, Proceedings. LNCS 1327. Springer-Verlag. pp. 823 - 829.
- [30] Port, R. F.; van Gelder, T. (eds.) (1995) Mind as Motion - Explorations in the Dynamics of Cognition, MIT Press, Cambridge.
- [31] Singer, W. (1994) Time as coding space in neocortical processing. In: Buzsáki, G.; Llinás, R.; Singer, W.; Berthoz, A.; Christen, Y. (eds.) Temporal Coding in the Brain. Springer, Berlin. pp. 51-80.
- [32] von Foerster, H. (1969) What is memory that it may have hindsight and foresight as well? In: Bogoch, S. (ed.) The Future of the Brain Sciences. Plenum Press, New York; in german: Was ist Gedächtnis - das es Rückschau und Vorschau ermöglicht?. In: Schmidt, S. J. (ed.) Heinz von Foerster: Wissen und Gewissen. Versuch einer Brücke, Suhrkamp, Frankfurt.
- [33] Wennekers, T.; Pasemann, F. (1996) Synchronous chaos in high-dimensional modular neural networks. International Journal of Bifurcation and Chaos 6: 2055-2067.