RESEARCH ARTICLE

Towards a unifying model of neural net activity in the visual cortex

Hermann Haken

Received: 20 February 2006 / Accepted: 11 August 2006 / Published online: 25 November 2006 © Springer Science+Business Media B.V. 2006

Abstract A neural net model describing the nonlinear interactions between axonal spikes is presented. It reconciles aspects of pattern recognition (as action of an associative memory) with those of spike synchronization and phase locking. The stability of the synchronized state is studied in detail.

Keywords Neural net · Spike synchronization · Phase locking · Pattern recognition

Introduction

Today there is a considerable body of experimental facts on the visual cortex of cats, monkeys and, to a somewhat lesser extent, of humans (see, for instance Kandel et al. 2000). As we know, the visual cortex is composed of specific layers (called areas). The function of the neurons of lower layers, such as area 17 etc. of cats, is well studied and shows that neurons respond by their firing in a specific manner to edges, bars and corners in their receptive fields. Furthermore, neurons are arranged in columns or hypercolumns of orientation specific cells. One might speculate that, when going up the layers, neurons can be found that are specific to more complex patterns-and eventually are responsive, say to faces, which leads to the concept of the "grandmother" cell. However, such cells have not been found and there are theoretical reasons to dismiss their concept. For instance, if recognition is based on such

H. Haken (🖂)

cells, the visual system will become very vulnerable. If such a specific cell is destroyed, the corresponding face can no more be recognized. Thus neuroscientists are-at least generally-convinced that pattern recognition is achieved by the action of a whole net of neurons. But then the questions arises how the visual cortex can combine the various features so that, say, a specific face or scene is recognized. Actually, the whole problem is far more general. As we know, the different features of an object, e.g. a lemon, such as shape, colour, smell, weight, denotation, are processed in quite different parts of the brain. Nevertheless, we perceive the lemon as a whole. What binds all these features together? On the experimental side, an interesting and perhaps relevant effect was found, first in anesthetized cats. For instance, when two bars with the same orientation are moving in the same direction and are lying in the receptive fields of two different groups of neurons, these neurons fire synchronously (Gray and Singer 1989; Eckhorn et al. 1988. For a more complete list of references cf. Haken 2002). Is synchronization the key to the solution of the "binding problem" (see for example Singer 1999a, b)?

In the wake of these experimental findings a number of models of synchronizing neural nets were developed (Mirollo and Strogatz 1990; Keener et al. 1981; van Vreeswijk 1996; Bressloff and Coombes 1998, 2000; Gerstner 1995. For a more comprehensive list of references cf. Haken 2002)—but to the best of my knowledge—none has been concerned with the problem of pattern recognition. On the other hand, leaving aside the fact that neurons interact via axonal spikes (pulses), models on pattern recognition have been developed. They are based on statistical mechanics, dynamic systems theory and more specifically on synergetics

Institute for Theoretical Physics I, Center of Synergetics, University of Stuttgart, Stuttgart 70550, Germany e-mail: hermann.haken@itp1.uni-stuttgart.de

and use the concept of attractors (Hopfield 1982; Haken 2004, where many further references can be found). Each attractor represents a prototype pattern. When an incomplete (test) pattern is offered to the recognition system, this latter acts as an associative memory which is achieved by a dynamics that pulls the test pattern into an attractor state closest to it. In these models there is no place for any synchronization effects, however. From these considerations the obvious question arises: Is there a unifying model which allows us to deal with both pattern recognition and spike synchronization?

Above that, I wish to develop a model that captures essential properties of realistic neurons and that can be studied analytically. In the present paper, I present such a unifying model and study its properties. In this model, I start from a typical set-up of a neuron with its coupling to other neurons. The signals from other neurons are transformed via synapses into electric currents of the dendrites of the considered neuron. The dynamics of the dendritic currents is taken care of by means of driven differential equations, the driving force being the axonal pulses from other neurons. The dendritic currents are summed up in a weighted fashion by the neuron's soma whereupon pulses are emitted through the axon which splits and makes contact with the dendrites of other neurons. The dynamics of pulse generation is described by differential equations referring to a phase angle ϕ_i of neuron *j* whose "motion" (rotation speed) is determined by the incoming dendritic currents and limited by saturation. The model is insofar a minimal model as the saturation which is conventionally considered is taken care of by a quadratic term of the input. In addition to conventional models, bilinear couplings of the dendritic currents stemming from other neurons are taken into account. As it turns out, the dendritic variables can be exactly eliminated so that coupled differential equations for the phase angles ϕ_i of the whole neural net result. (Actually, in a neurophysiological interpretation, the phase angles are proportional to the corresponding action potentials.) The corresponding equations are the starting point of our present paper (cf. 2.1). The properties of these equations with respect to pattern recognition are studied in my articles (Haken 2006a, b). As is shown there, pattern recognition requires a specific choice of the coefficients $\lambda_{ii'}, \lambda_{ii'i''}$ of the Eq. 2.1 below. These coefficients are specific combinations of so called prototype vectors representing the stored or learned patterns. In the present paper, I want to study in how far synchronization of pulses or, more generally, phase locking is compatible with pattern recognition. The quite interesting answer to this question will be presented in the conclusions at the end of this paper.

Also the indispensable concept of quasi-attractors will be explained there.

The synchronized state

Our starting point is the set of equations

$$\dot{\phi}_{j}(t) + \gamma' \phi_{j}(t) \mod 2\pi = \sum_{j'} \tilde{\lambda}_{jj'} X_{j'} + \sum_{j'j''} \tilde{\lambda}_{jj'j''} X_{j'} X_{j'} X_{j''} + p_{j}$$

$$(2.1)$$

that we have derived in a previous paper (Haken 2006b). For convenience we briefly recapitulate the meaning of the individual parts. ϕ_j is the phase of the axonal pulse of neuron *j*. γ' is a damping constant. Modulo 2π means that ϕ_j must be reduced by integer multiples of 2π until it runs in-between 0 and 2π . The quantity $\phi_j \mod 2\pi$ is proportional to the action potential of neuron *j*. The λ s on the right-hand side are coupling coefficients between neurons and contain synaptic strengths, whereas X_j is defined by

$$X_j(t) = \int_0^t K(t-\sigma) f\Big(\phi_j(\sigma-\tau)\Big) \mathrm{d}\sigma.$$
(2.2)

 p_j represents the external signals that in this section will be assumed to be independent of *j*. The quantity *K* is a response function of the dendrites. In this paper, we assume it in the form

$$K(t - \sigma) = e^{-\gamma(t - \sigma)}.$$
(2.3)

for positive $(t - \sigma)$ and equal to zero for negative $(t - \sigma)$. γ is the decay constant of the dendritic electric currents. The function *f* is defined by means of

$$f(\phi(t)) = \sum_{n} \delta(t - t_n), \mathbf{n} : \text{integer}$$
(2.4)

where t_n are the arrival times of axonal pulses (spikes). We specialize (2.1) to the case of phase locking. In this case, we assume that the phase angles ϕ_j are independent of j

$$\phi_j = \phi. \tag{2.5}$$

The times t_n in (2.4) are defined by means of

$$\phi(t_n) = 2\pi n \tag{2.6}$$

and choosing an appropriate zero of time we may put

$$t_n = n\Delta. \tag{2.7}$$

Clearly we also have

$$t_{n+1} - t_n = \Delta. \tag{2.8}$$

In the following we will allow for a small time delay τ

$$\tau : \text{time delay} > 0 \tag{2.9}$$

which will help us to make occasionally the evaluation of integrals over δ -functions unique. In the first step of our analysis we evaluate (2.2) assuming that X is independent of j. Using (2.4) and the properties of the δ -functions we may evaluate (2.2) in the form

$$X(t) = \int_0^t e^{-\gamma(t-\sigma)} \sum_n \delta(\sigma - \tau - t_n) d\sigma = \sum_{n=0}^N e^{-\gamma(t-n\Delta - \tau)},$$
(2.10)

or still more explicitly (with $\tau \rightarrow 0$)

$$X(t) = e^{-\gamma (t \mod \Delta)} W - \frac{e^{-\gamma t}}{e^{\gamma \Delta} - 1}, \quad W = \frac{1}{1 - e^{-\gamma \Delta}}.$$
(2.11)

The steady state is obtained by dropping the term of X(t) that decays exponentially. In the following we thus will use the expression

$$X(t) = e^{-\gamma (t \mod \Delta)} \frac{1}{1 - e^{-\gamma \Delta}}.$$
(2.12)

Because in the following some of the formulas will become rather involved, occasionally we will treat special cases, e.g.

$$\gamma \Delta \gg 1; \quad X(t) = e^{-\gamma (t \mod \Delta)}$$
 (2.13)

which means that the pulse intervals are rather large. But we will also consider the opposite case in which the pulse intervals are small. After these preliminary steps we may return to the solution of (2.1). In the special case (2.5) and under the assumption that all signals p_j are the same, (2.1) reduces to

$$\dot{\phi}(t) + \gamma' \phi(t) \operatorname{mod} 2\pi = AX + BX^2 + p \tag{2.14}$$

where we use the abbreviations

$$\sum_{j} \tilde{\lambda}_{jj'} = A \tag{2.15}$$

and

$$\sum_{j'j''} \tilde{\lambda}_{jj'j''} = B. \tag{2.16}$$

Note that implicitly in (2.15) and (2.16) an important assumption enters, namely that the corresponding sums in (2.15) and (2.16) are independent of the index *j* which is a requirement for the network connectivities. We shall not dwell on the consequence of this requirement here which we discussed elsewhere (Haken 2002). For later purposes we evaluate $\dot{\phi}(t_0)$. Using the result (2.12) we readily find

$$\dot{\phi}(t_0) = \frac{1}{1 - e^{-\gamma \Delta}} A + B \frac{1}{(1 - e^{-\gamma \Delta})^2} + p, \quad t_0 = 0.$$
 (2.17)

The only free parameter that occurs in (2.11), (2.12) and (2.17) that still must be determined is the pulseinterval Δ . In order to fix it we choose a time interval

$$t_n \le t \le t_{n+1} \tag{2.18}$$

and put

$$\phi(t) = \phi(t_n) + x(t) \tag{2.19}$$

with

$$x(t_n) = 0, \quad \dot{x} \ge 0.$$
 (2.20)

Having an eye on (2.14) we assume that x(t) obeys an equation of the form

$$\dot{x}(t) + \gamma' x(t) = G(t). \tag{2.21}$$

Its solution can be written as

$$\begin{aligned} x(t) &\equiv \phi(t) - \phi(t_n) \\ &= \int_{t_n}^t e^{-\gamma'(t-\sigma')} \Big(AX(\sigma') + BX(\sigma')^2 \Big) d\sigma' \\ &+ \int_{t_n}^t e^{-\gamma'(t-\sigma')} p d\sigma'. \end{aligned}$$
(2.22)

Putting the upper time equal to t_{n+1} and using the explicit form of X (2.12) we can cast (2.22) into

$$\phi(t_{n+1}) - \phi(t_n)$$

$$= \int_{t_n}^{t_{n+1}} e^{-\gamma'(t_{n+1} - \sigma')} \left(A e^{-\gamma(\sigma' - t_n)} W + B e^{-2\gamma(\sigma' - t_n)} W^2 \right)$$

$$\times d\sigma' + \frac{p}{\gamma'} (1 - e^{-\gamma' \Delta}). \qquad (2.23)$$

The integrals can be easily evaluated. Thus we can cast (2.23) using the requirement (2.6) into the form

$$2\pi = A \frac{1}{\gamma - \gamma'} \left(e^{-\gamma'\Delta} - e^{-\gamma\Delta} \right) W + B \frac{1}{2\gamma - \gamma'} \times \left(e^{-\gamma'\Delta} - e^{-2\gamma\Delta} \right) W^2 + \frac{p}{\gamma'} \left(1 - e^{-\gamma'\Delta} \right).$$
(2.24)

This is an implicit equation for Δ . In order to discuss it we consider two limiting cases, namely: (1) that the pulse interval is small so that

$$\Delta \gamma \ll 1$$
, $\Delta \gamma' \ll 1$, (which implies *p* large enough).
(2.25)

In this case (2.24) reduces to

$$2\pi = A\frac{1}{\gamma} + B\frac{1}{\gamma^2\Delta} + p\Delta.$$
(2.26)

This is a quadratic equation for Δ . Choosing the positive root and assuming $2\pi\gamma - A$ big enough we obtain

$$\Delta = \frac{1}{p} (2\pi - A/\gamma) + |B|/\gamma \cdot \frac{1}{2\pi\gamma - A}.$$
(2.27)

Let us discuss the individual terms on the right-hand side of (2.27). $\frac{2\pi}{p}$ just means that the pulse interval is proportional to the size of the incoming signal, or in other words, that the pulse rate is proportional to the signal size. If there is a positive feedback A, the pulse interval is shortened, otherwise lengthened. Furthermore, if there is a nonlinear saturation effect, i.e. B < 0, the pulse interval is enlarged by this feedback.

(2) Let us now consider the other case, i.e. large pulse intervals so that

$$\gamma \Delta \gg 1, \quad \gamma > \gamma'$$
 (2.28)

holds. In this case (2.24) is reduced to

$$2\pi = \frac{1}{\gamma} e^{-\gamma' \Delta} \left(A + \frac{B}{2} \right) + \frac{P}{\gamma'} \left(1 - e^{-\gamma' \Delta} \right)$$
(2.29)

or after resolving (2.29) for $e^{-\gamma\Delta}$ we obtain

$$e^{-\gamma'\Delta} = \left(\frac{A}{\gamma} + \frac{B}{2\gamma} - \frac{p}{\gamma'}\right) \left(2\pi - \frac{p}{\gamma'}\right)^{-1}.$$
 (2.30)

Because of the exponential function, the right-hand side must be positive. This requires small enough signals, i.e.

$$p < 2\pi/\gamma'. \tag{2.31}$$

Our results can easily be generalized if not only linear and quadratic terms in X in equation (2.14) are taken into account but terms up to arbitrary order. In other words, the first two terms on the right-hand side of (2.14) are replaced by

$$\sum_{m=1}^{\infty} A_m X^m. \tag{2.32}$$

Stability of the synchronized state

In the preceding section, we have seen that it is rather simple to obtain the synchronized state and to determine the pulse interval. There remains, however, a basic question, namely "Is the mathematically constructed state realizable?" i.e. in particular "Is it stable?" Therefore in this section, we have to study the stability, which we will do by linear stability analysis. Though the idea is basically simple, the individual steps are slightly complicated so that we proceed in several steps, where we want to take care of the interests of the speedy reader. Thus we present the basic stability analysis equations and discuss their solutions whereas we skip the purely technical-mathematical derivation of the stability equations. Linear stability analysis means that we start from the ideal solution $\phi(t)$ and superpose on it small deviations that we call ξ_j . These deviations are of course time-dependent. If they decay in the course of time, the system is stable, otherwise unstable. According to these remarks we introduce the decomposition

$$\phi_i(t) = \phi(t) + \xi_i(t), \tag{3.1}$$

which we insert in (2.1), from which we subtract the equation (2.14) of the foregoing section. In order to write down the right-hand side Z of the thus resulting equation

$$\dot{\xi}_{j}(t) + \gamma' \big(\phi(t) + \xi_{j}(t)\big) \operatorname{mod} 2\pi - \phi(t) \operatorname{mod} 2\pi = Z$$
(3.2)

we introduce

$$\hat{X}_k(t) = \int_0^t K(t,\sigma) f(\phi(\sigma') + \xi_k(\sigma')) \mathrm{d}\sigma'$$
(3.3)

with the variable

$$\sigma' = \sigma - \tau \tag{3.4}$$

where τ is an infinitesimally small quantity. We further introduce

$$\hat{X}_{k}(t) - X_{k}(t)$$

$$= \int_{0}^{t} K(t,\sigma) \{ f(\phi(\sigma') + \xi_{k}(\sigma')) - f(\phi(\sigma')) \} \mathrm{d}\sigma. \quad (3.5)$$

With these abbreviations we may write Z in the form

$$Z = \sum_{k} \tilde{\lambda}_{jk} \left(\hat{X}_{k} - X_{k} \right) + \sum_{kl} \tilde{\lambda}_{jkl} \left(\hat{X}_{k} \hat{X}_{l} - X_{k} X_{l} \right). \quad (3.6)$$

For further evalutation we put

$$\hat{X}_k = X_k + \varepsilon_k \tag{3.7}$$

where we assume that ε_k is a small quantity. This allows us to transform (3.6) into

$$I = \hat{\gamma} e^{-\gamma' t} \sum_{n \ge 1}^{N-1} e^{\gamma' t_n} \xi_j(t_n)$$
(3.12)

where we use the abbreviation

$$\hat{\gamma} = \gamma' 2\pi \dot{\phi}(t_0)^{-1}. \tag{3.13}$$

Note that $\dot{\phi}$ has been calculated in the previous section (cf. Eq. 2.17). We further introduce the abbreviations

$$\hat{\lambda}_{jk} = \dot{\phi}(t_0)^{-1} \tilde{\lambda}_{jk}; \quad \hat{\lambda}_{jkl} = \dot{\phi}(t_0)^{-1} \tilde{\lambda}_{jkl}$$
(3.14)

which allows us to write II in the form

$$II = \sum_{k} \hat{\lambda}_{jk} \sum_{n=0}^{N-1} \xi_{k}(t_{n}) \frac{1}{\gamma' - \gamma} \Big(\gamma' \mathrm{e}^{-\gamma'(t_{N}-t_{n})} - \gamma \mathrm{e}^{-\gamma(t_{N}-t_{n})} \Big).$$
(3.15)

Furthermore *III* is given by

$$III = 2\sum_{kl} \hat{\lambda}_{jkl} \sum_{n=0}^{N-1} \left\{ e^{-\gamma'(t_N - t_n)} \left(1 + \frac{\gamma}{\gamma' - 2\gamma} \frac{e^{(\gamma' - 2\gamma)\Delta} - 1}{e^{(\gamma' - \gamma)\Delta} - 1} \right) - e^{-\gamma(t_N - t_n)} \frac{\gamma}{\gamma' - 2\gamma} \frac{e^{(\gamma' - 2\gamma)\Delta} - 1}{e^{(\gamma' - \gamma)\Delta} - 1} \right\} W \xi_k(t_n).$$

$$(3.16)$$

$$Z = \sum_{k} \tilde{\lambda}_{jk} \varepsilon_k + \sum_{kl} \tilde{\lambda}_{jkl} (X_k \varepsilon_l + X_l \varepsilon_k).$$
(3.8)

The next steps should consist in the evaluation of

$$\varepsilon_k(t) = \hat{X}_k - X_k \tag{3.9}$$

and

$$y = \left(\phi(t) + \xi_j(t)\right) \mod 2\pi - \phi(t) \mod 2\pi \tag{3.10}$$

both of which occur in (3.2) with (3.8). We skip this explicit evaluation and immediately proceed to the basic equations for ξ_j . As it turns out the differential equations (3.2) can be converted into algebraic equations for ξ_j taken at discrete times t_n which correspond to the spiking times in accordance with the previous section. The algebraic equations are given by

$$\xi_j(t) = I + II + III, \quad t = t_N \tag{3.11}$$

where the quantities *I*, *II*, *III* are defined as follows. *I* stems from the damping terms containing γ' and reads

Quite clearly (3.15) stems from terms of Eq. 2.1 that are linear in X, whereas (3.16) stems from terms that are bilinear in X. The relation (3.16) refers to the general case where γ , γ' , Δ are still arbitrary. In order to find formulas that are more handy we consider special cases, namely

$$\gamma' \Delta \ll 1, \quad \gamma \Delta \ll 1,$$
 (3.17)

which corresponds to short spike intervals. In this case *III* reduces to

$$III = 2\sum_{kl} \hat{\lambda}_{jkl} \sum_{n=0}^{N-1} \frac{1}{\gamma' - \gamma} \left(e^{-\gamma'(t_N - t_n)} \gamma' - e^{-\gamma(t_N - t_n)} \gamma \right) W \xi_k(t_n).$$
(3.18)

The other limiting case of comparatively large spike intervals characterized by

$$\gamma \Delta \gg 1, \quad \gamma > \gamma'$$
 (3.19)

allows us to reduce (3.16) to the form

$$III = 2 \sum_{kl} \hat{\lambda}_{jkl} \sum_{n=0}^{N-1} \frac{1}{\gamma' - 2\gamma} \\ \times \left((\gamma' - \gamma) e^{-\gamma'(t_N - t_n)} - \gamma e^{-\gamma(t_N^- t_n)} \right) \xi_k(t_n) W.$$
(3.20)

The equations (3.11) with (3.15) and (3.16) or (3.18), (3.20) are linear equations in the unknown variables ξ_j taken at discrete times t_n . Leaving aside some more technical questions we may rather simply solve these equations which we will do in the next section.

Stability analysis continued: solution of the stability equations

In order to solve the stability equations we make the hypothesis

$$\xi_j(t_n) = \beta^n \xi_j(t_0) \tag{4.1}$$

with a still unknown constant β . Inserting (4.1) into (3.12) we obtain

$$I = \hat{\gamma} e^{-\gamma' t_N} \sum_{n \ge 1}^{N-1} e^{\gamma' n \Delta} \beta^n \xi_j(t_0)$$
(4.2)

which can immediately be evaluated to yield

$$I = \hat{\gamma}\xi_j(t_0)\frac{\beta^N}{\mathrm{e}^{\gamma'\Delta}\beta - 1}.$$
(4.3)

An additive term has been neglected as will be done similarly in the following formulas. (For a detailed justification cf. (Haken 2002).) The expressions *II* and *III* can be evaluated similarly. We obtain for *II*

$$II = \beta^{N} \left(\frac{\gamma'}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} - \frac{\gamma}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma \Delta} - 1} \right) \sum_{k} \hat{\lambda}_{jk} \xi_{k}(t_{0})$$

$$(4.4)$$

For the evaluation of *III* it is advisable to distinguish between the limiting cases of short and long pulse intervals. In the case

$$\gamma' \Delta \ll 1, \quad \gamma \Delta \ll 1$$
 (4.5)

we obtain

$$III = \beta^{N} \left(\frac{\gamma'}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} - \frac{\gamma}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma \Delta} - 1} \right) \\ \times \frac{2}{\gamma \Delta} \sum_{kl} \hat{\lambda}_{jkl} \xi_{k}(t_{0}),$$
(4.6)

whereas in the case of

$$\gamma \Delta \gg 1, \quad \gamma > \gamma'$$

$$(4.7)$$

we obtain

$$III = \beta^{N} \left(\frac{\gamma' - \gamma}{\gamma' - 2\gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} - \frac{\gamma}{\gamma' - 2\gamma} \frac{1}{\beta e^{\gamma \Delta} - 1} \right) \\ \times 2 \sum_{kl} \hat{\lambda}_{jkl} \xi_{k}(t_{0}) W.$$
(4.8)

In order to illustrate our further procedure we consider the case (4.7). It is convenient to treat the variables ξ_j as components of a vectors ξ . Inserting (4.1), (4.4) and (4.8) into (3.11) and dividing both sides by β^N we obtain the following vector equation

$$\xi(t_0) = \left\{ \hat{\gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} + \left(\frac{\gamma'}{\gamma' - \gamma \beta e^{\gamma' \Delta} - 1} - \frac{\gamma}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma \Delta} - 1} \right) \Lambda^{(1)} + \left(\frac{\gamma' - \gamma}{\gamma' - 2\gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} - \frac{\gamma}{\gamma' - 2\gamma} \frac{1}{\beta e^{\gamma \Delta} - 1} \right) \Lambda^{(2)} W \right\} \xi(t_0).$$

$$(4.9)$$

where $\Lambda^{(1)}$, $\Lambda^{(2)}$ are matrices corresponding to their components $\hat{\lambda}_{jk}$ and $2\sum \hat{\lambda}_{jkl}$, respectively. In order to facilitate our analysis we assume that the matrices $\Lambda^{(1)}$, $\Lambda^{(2)}$ commute

$$\left[\Lambda^{(1)},\Lambda^{(2)}\right] = 0 \tag{4.10}$$

so that they can be simultaneously diagonalized. We assume that the eigenvalues are given by

$$\lambda_{\mu}^{(1)}, \lambda_{\mu}^{(2)},$$
 (4.11)

respectively. Choosing the eigenvectors ξ correspondingly the Eq. 4.9 can be reduced to

$$1 = \frac{a}{\beta e^{\gamma' \Delta} - 1} + \frac{b}{\beta e^{\gamma \Delta} - 1}$$
(4.12)

where we introduce the abbreviations

$$a = \hat{\gamma} + \frac{\gamma'}{\gamma' - \gamma} \lambda_{\mu}^{(1)} + \frac{\gamma' - \gamma}{\gamma' - 2\gamma} \lambda_{\mu}^{(2)} W, \quad W = \frac{1}{1 - e^{-\gamma \Delta}}$$
(4.13)

and

$$b = -\frac{\gamma}{\gamma' - \gamma} \lambda_{\mu}^{(1)} - \frac{\gamma}{\gamma' - 2\gamma} \lambda_{\mu}^{(2)} W.$$
(4.14)

Equation 4.12 can be considered as an eigenvalue equation for β which is of second order and can rather easily be solved. To get an insight into the structure of the eigenvalues we capitalize on the specific form of

the Eq. 4.12, which means that there are specific singularities when one of the numerators vanishes. Note that the following approach works provided the roots $\beta = \beta_1$ and $\beta = \beta_2$ are sufficiently separated. Thus in lowest order we may put

$$1 = \frac{a}{\beta_1 e^{\gamma' \Delta} - 1} \tag{4.15}$$

with the solution

$$\beta_1 = \mathrm{e}^{-\gamma'\Delta}(1+a). \tag{4.16}$$

Similarly we may put

$$1 = \frac{b}{\beta_2 e^{\gamma \Delta} - 1} \tag{4.17}$$

with the solution

$$\beta_2 = \mathrm{e}^{-\gamma \Delta} (1+b). \tag{4.18}$$

Making the further assumption

$$\gamma \gg \gamma' \tag{4.19}$$

the expressions for the eigenvalues can be further reduced by means of

$$a = \hat{\gamma} - \frac{\gamma'}{\gamma} \lambda_{\mu}^{(1)} + \frac{1}{2} \lambda_{\mu}^{(2)}, \qquad (4.20)$$

$$b = \lambda_{\mu}^{(1)} + \frac{1}{2}\lambda_{\mu}^{(2)}.$$
(4.21)

In order to obtain stability the absolute values of the eigenvalues β_1,β_2 must be smaller than unity. This is secured if a and b are smaller than 0. The case (4.20) is rather simple, when we focus our attention on the conventional procedure in pattern recognition (Haken 2004, 2006b). In this case the eigenvalues $\lambda^{(1)}$ must be positive (for pattern amplification), whereas the eigenvalues $\lambda^{(2)}$ must be negative for pattern saturation. If the corresponding eigenvalues $\lambda_{\mu}^{(1,2)}$ have sufficiently large absolute values, a is surely negative. The case (4.21) is somewhat more intricate because the eigenvalue $\lambda^{(1)}$ is positive. In this case the eigenvalue $\lambda^{(2)}$ that corresponds to the nonlinear term must be sufficiently negative. The situations is somewhat facilitated, however, by the prefactor $e^{-\gamma\Delta}$ because we are operating under the assumption that $\gamma \Delta \gg 1$ so that $\beta_2 < 1$ even if b > 0.

In conclusion of this section we consider the case $\gamma \Delta \ll 1$, $\gamma' \Delta \ll 1$, where we may proceed in close analogy to the previous case. We obtain the vector equation

$$\xi(t_0) = \left(\hat{\gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} + \left(\frac{\gamma'}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma' \Delta} - 1} - \frac{\gamma}{\gamma' - \gamma} \frac{1}{\beta e^{\gamma \Delta} - 1}\right) \\ \times \left(\Lambda^{(1)} + \frac{1}{\gamma \Delta} \Lambda^{(2)}\right) \xi(t_0).$$
(4.22)

Under the assumptions (4.10) and (4.11) we then find

$$1 = \frac{a'}{\beta e^{\gamma' \Delta} - 1} + \frac{b'}{\beta e^{\gamma \Delta} - 1}$$
(4.23)

The coefficients a', b' are given by

$$a' = \hat{\gamma} + \frac{\gamma'}{\gamma' - \gamma} \lambda_{\mu}, \qquad (4.24)$$

and

$$b' = \frac{\gamma}{\gamma - \gamma'} \lambda_{\mu}, \tag{4.25}$$

where

$$\lambda \mu = \left(\lambda_{\mu}^{(1)} + \frac{1}{\gamma \Delta} \lambda_{\mu}^{(2)}\right) \tag{4.26}$$

and

$$b' = \frac{\gamma}{\gamma - \gamma'} \left(\lambda_{\mu}^{(1)} + \frac{1}{\gamma \Delta} \lambda_{\mu}^{(2)} \right).$$
(4.27)

Again under the assumption that the eigenvalues are sufficiently separated from each other, in lowest approximation they read

$$\beta'_1 = e^{-\gamma' \Delta} (1 + a')$$
 (4.28)

and

$$\beta'_2 = e^{-\gamma \Delta} (1 + b').$$
 (4.29)

For a further discussion, we first consider the case $\gamma' > \gamma$. Then the condition a' < 0 requires $\lambda_{\mu} < 0$, from which follows b' > 0, i.e. instability (provided we ignore the exponential factors in (4.28), (4.29) that are close to unity). In the case $\gamma' > \gamma$ we need $\lambda_{\mu} > 0$ for a' < 0 which leads to b' > 0, i.e. again instability. Let us therefore consider the case $\gamma' = \gamma$ starting from (4.23). There is only one root, namely

$$\beta = \mathrm{e}^{-\gamma \Delta} (1 + a' + b') = \hat{\gamma} + \lambda_{\mu}. \tag{4.30}$$

Provided λ_{μ} is sufficiently negative, the pulse sequence is stable. Since in pattern recognition $\lambda_{\mu}^{(2)}$ is negative and $\gamma \Delta \ll 1$ (in the present case), $\lambda_{\mu} < 0$ can surely be fulfilled. Clearly, from here we may extrapolate that for $\gamma' \approx \gamma$ still stability is possible. So far, it appears as if the process of pattern recognition is compatible with the formation of pulse trains. There is, however, still the conditions (2.15), (2.16) and $p_j = p$ to be fulfilled for pulse train formation. In the next section we therefore want to study, in how far these conditions can be relaxed.

From synchronization to phase locking

In the previous sections from "The synchronized state" to "Stability analysis continued: solution of the stability equations" we have dealt with synchronization in which case the phases ϕ_j of the neurons with indices (j)are given by $\phi_j = \phi$. This solution has become possible under the assumption that all external signals are equal and that the coupling coefficients λ obey specific conditions (cf. (5.4), (5.6)). In this section, we want to extend our former results to the case in which these conditions are violated, at least to some degree. We still will be concerned with the case in which the frequencies of all axonal pulses are the same but in which phase shifts according to

 $\phi_i(t) = \phi + \delta \phi_i,$

or equivalently,

$$t_{j,n} = t_n + \tau_j \tag{5.1}$$

occur. Our starting point is the equations

$$\dot{\phi}_{j} + \gamma' \phi_{j} \operatorname{mod} 2\pi = \sum_{j'} \lambda_{jj'} X_{j'} + \sum_{j'j''} \lambda_{jj'j''} X_{j} X_{j'} + p_{j} \equiv R(t)$$
(5.2)

where we make the decompositions

$$\lambda_{jj'} = \lambda_{jj'}^o + \kappa_{jj'} \tag{5.3}$$

with the constraints

$$\sum_{j'} \lambda_{jj'}^0 = \lambda^0 \tag{5.4}$$

and

$$\lambda_{jj'j''} = \lambda_{jj'j''}^{0} + \kappa_{jj'j''}$$
(5.5)

with the constraints

$$\sum_{j'j''} \lambda_{jj'j''}^0 = \lambda_2^0.$$
(5.6)

The constraints (5.4) and (5.6) are the ones we had to introduce previously in order to achieve synchronization. We further will use the decomposition

$$p_j = p + \pi_j. \tag{5.7}$$

In the following we will assume that the quantities κ and π are small enough so that we will be able to use perturbation theory. X_i is given by

$$X_{j}(t) = \int^{t} K(t,\sigma) \sum_{n} \delta(\sigma - t_{j,n}) \mathrm{d}\sigma$$
(5.8)

or because of (5.1) by

$$X_j(t) = \int^t K(t,\sigma) \sum_n \delta(\sigma - t_n - \tau_j) \mathrm{d}\sigma.$$
(5.9)

In complete analogy to our previous procedure of section "The synchronized state" we integrate the Eq. 5.2 from one pulse to the next and use the usual Green's function formalism. Thus we obtain

$$\phi_j(t_{j,n+1}) - \phi_j(t_{j,n}) = \int_{t_{j,n}}^{t_{j,n+1}} e^{-\gamma'(t_{j,n+1}-\sigma)} R(\sigma) d\sigma.$$
(5.10)

Making the substitution

$$\sigma \to \sigma + \tau_i \tag{5.11}$$

transforms (5.10) into

$$\phi_j(t_{j,n+1}) - \phi_j(t_{j,n}) = \int_{t_n}^{t_{n+1}} e^{-\gamma'(t_{n+1}-\sigma)} R(\sigma+\tau_j) d\sigma.$$
(5.12)

To proceed further we compare (5.9) with our former expression (2.12). This allows us to write (5.9) as

$$X_j(t) = e^{-\gamma((t-\tau_j) \mod \Delta)} W$$
(5.13)

or equally well

$$X_j(t) = X(t - \tau_j), \tag{5.14}$$

i.e. the index *j* appears only via τ_j . Because of the modulo function in (5.13) we may use

)
$$X(t - \tau_i) = X(t + t_n - \tau_i)$$
 (5.15)

because of

$$t_n = n\Delta. \tag{5.16}$$

We further recall that the phase difference between two pulses is just 2π . Using these intermediate steps as well as the explicit form of *R* we can transform (5.10) into

$$2\pi = \int_{t_n}^{t_{n+1}} e^{-\gamma'(t_{n+1}-\sigma)} \Biggl\{ \sum_{j'} \lambda_{jj'} X(\sigma + \tau_j - \tau_{j'}) + \sum_{j'j''} \lambda_{jj'j''} X(\sigma + \tau_j - \tau_{j'}) X(\sigma + \tau_j - \tau_{j''}) + p_j(\sigma + \tau_j) \Biggr\} d\sigma.$$
(5.17)

Finally we introduce the transformation

$$\sigma \to \sigma + t_n \tag{5.18}$$

which allows us to cast (8) into

$$2\pi = \int_{0}^{\Delta} e^{-\gamma'(\Delta-\sigma)} \left\{ \sum_{j'} \lambda_{jj'} X(\sigma + \tau_{j} - \tau_{j'}). + \sum_{j'j''} \lambda_{jj'j''} X(\sigma + \tau_{j} - \tau_{j'}) X(\sigma + \tau_{j} - \tau_{j''}) + p_{j}(\sigma + t_{n} + \tau_{j}) \right\} d\sigma.$$
(5.19)

Already at this moment we can observe that the Eqs. 5.19 are invariant against the simultaneous replacement of τ_j by $\tau_j + \tau$ provided p_j is time-independent which we will assume from now on. In order to treat the Eq. 5.19 further, we use the abbreviations

$$\int_0^{\Delta} e^{-\gamma'(\Delta-\sigma)} X(\sigma+\tau_j-\tau_j) \mathrm{d}\sigma = G^{(1)}(\Delta,\tau_j-\tau_{j'}), \quad (5.20)$$

$$\int_0^{\Delta} e^{-\gamma(\Delta-\sigma)} X(\sigma+\tau_j-\tau_{j'}) X(\sigma+\tau_j-\tau_{j''}) d\sigma$$

= $G^{(2)}(\Delta,\tau_j-\tau_{j'},\tau_j-\tau_{j''})$ (5.21)

and

$$\int_0^{\Delta} e^{-\gamma'(\Delta-\sigma)} p_j(\sigma+t_n+\tau_j) d\sigma = P_j(\Delta, t_n+\tau_j).$$
 (5.22)

Thus the Eqs. 5.17 eventually acquire the form

$$2\pi = \sum_{j'} \lambda_{jj'} G^{(1)}(\Delta, \tau_j - \tau_{j'}) + \sum_{j'j''} \lambda_{jj'j''} G^{(2)} \times (\Delta, \tau_j - \tau_{j'}, \tau_j - \tau_{j'''}) + P_j(\Delta, t_n + \tau_j).$$
(5.23)

In order to make contact with our previous results in which the strict synchronization conditions were fulfilled, we introduce the decompositions

$$G^{(1)}(\Delta,\tau_j-\tau_{j'}) = G^{(1)}(\Delta_0,0) + h_{jj'} \equiv G^{(1)}_0 + h_{jj'}, \quad (5.24)$$

$$G^{(2)}(\Delta, \tau_j - \tau_{j'}, \tau_j - \tau_{j''}) = G^{(2)}(\Delta_0, 0) + h_{jj'j''} \equiv G_0^{(2)} + h_{jj'j''},$$
(5.25)

$$P_j(\Delta, t_n + \tau_j) = p \frac{1}{\gamma'} \left(1 - e^{-\gamma' \Delta_0} \right) + \pi_j(\Delta_0, \Delta')$$
(5.26)

as well as

$$\Delta = \Delta_0 + \Delta'. \tag{5.27}$$

We assume that in accordance with the results of section "The synchronized state" the equations

$$2\pi = \sum_{j'} \lambda_{jj'}^0 G^{(1)}(\Delta_0, 0) + \sum_{j''} \lambda_{jj'j''}^0 G^{(2)}(\Delta_0, 0, 0) + p \frac{1}{\gamma'} \left(1 - e^{-\gamma' \Delta_0} \right)$$
(5.28)

are fulfilled where Δ_0 plays the same role as Δ in section "The synchronized state". So far our equations are exact. From now on we assume that the quantities κ,h,Δ' are small. Inserting the expressions (5.3), (5.5), (5.24), (5.25), (5.26), (5.27) into (5.23) we arrive at

$$2\pi = \sum_{j'} \left(\lambda_{jj'}^{0} + \kappa_{jj'} \right) \left(G_0^{(1)} + h_{jj'} \right) + \sum_{j'j''} \left(\lambda_{jj'j''}^{0} + \kappa_{jj'j''} \right) \left(G_0^{(2)} + h_{jj'j''} \right) + p \frac{1}{\gamma'} (1 - e^{-\gamma' \Delta_0}) + \pi_j (\Delta_0, \Delta').$$
(5.29)

Using (5.28) and keeping the small quantities up to first order we eventually find

$$0 = \sum_{j'} \lambda_{jj'}^{0} h_{jj'} + \sum_{j'} \kappa_{jj'} G_{0}^{(1)} + \sum_{j'j''} \lambda_{jj'j''}^{0} h_{jj'j''} + \sum_{j'j''} \kappa_{jj'j''} G_{0}^{(2)} + \pi_{j}(\Delta_{0}, \Delta).$$
(5.30)

In order to make the smallness assumptions explicit, we introduce the decompositions

$$h_{jj'} = a(\tau_j - \tau_{j'}) + b\Delta'$$
(5.31)

and a corresponding one for $h_{jj'j''}$ as well as

$$\pi_i(\Delta_0, \Delta) \approx \pi_i(\Delta_0, 0) \tag{5.32}$$

Note that π_j is a small quantity. From what follows it is not difficult to convince oneself that it is sufficient to treat only the example (5.31) because the additional terms stemming from $h_{jj'j''}$ play a similar role. Thus we shall drop these terms. Then the Eq. 5.30 can be cast into the form

$$\sum_{j'} \lambda_{jj'}^0 h_{jj'} = f_j \tag{5.33}$$

where we use the abbreviation

$$f_j = -G^{(1)}(\Delta_0, 0) \sum_{j'} \kappa_{jj'} + \pi_j(\Delta_0, 0).$$
(5.34)

Making use of (5.31) we cast (5.33) into the form

$$a\lambda^0\tau_j - \sum_{j'}\lambda^0_{jj'}a\tau_{j'} + b\lambda^0\Delta' = f_j.$$
(5.35)

(Remember that $\sum_{jj'} \lambda_{jj'}^0 = \lambda^0$.) Putting the components τ_i together to form' a vector τ we find

$$L\tau = F \tag{5.36}$$

where

$$L = (L_{jj'}) = (a\lambda^0 \delta_{jj'} - \lambda_{jj'}a)$$
(5.37)

and

$$\boldsymbol{F} = \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} - b\lambda^0 \Delta' \boldsymbol{E}.$$
 (5.38)

E is the unity matrix. We are now able to make contact with well known rules from linear algebra and introduce the eigenvalue equations

$$L\boldsymbol{\tau}^k = \lambda_k \boldsymbol{\tau}^k. \tag{5.39}$$

One can readily identify the null vector which reads

$$\boldsymbol{\tau}^0 = \boldsymbol{E} \tag{5.40}$$

and possesses a vanishing eigenvalue

 $\lambda_0 = 0. \tag{5.41}$

To verify this just recall that

$$a\lambda^0 - \sum_{j'} \lambda_{jj'} a = 0 \tag{5.42}$$

holds because of the definition of λ^0 . Again using linear algebra we can write the solution to (5.36) in the form

$$\boldsymbol{\tau} = \boldsymbol{M}\boldsymbol{F} + \alpha \boldsymbol{\tau}^{\boldsymbol{\theta}} \tag{5.43}$$

where *M* is the Green's function. The solvability of (3) requires that *F* is orthogonal on the adjoint null vector τ^{0+} , i.e.

$$F \perp \tau^{0+} \tag{5.44}$$

or explicitly

$$\left(\boldsymbol{\tau}^{0+}\boldsymbol{F}\right) = 0. \tag{5.45}$$

We remember that the adjoint null vector has to obey the equations

$$(\tau^{0+}\tau^{0}) = 1 \tag{5.46}$$

and

$$\left(\boldsymbol{\tau}^{0+}\boldsymbol{\tau}^{k}\right) = \delta_{kk'} \tag{5.47}$$

where the right-hand side of (5.47) is the Kronecker symbol. Inserting (5.38) into (5.45) yields

$$\left(\boldsymbol{\tau}^{0+}\boldsymbol{E}\right)\left(-b\lambda^{0}\Delta'\right)+\left(\boldsymbol{\tau}^{0+}\boldsymbol{f}\right)=0$$
(5.48)

which allows us to calculate Δ' explicitly by

$$\Delta' = \left((\boldsymbol{\tau}^{0+} \boldsymbol{E}) (\boldsymbol{b} \boldsymbol{\lambda}^{\boldsymbol{\theta}}) \right)^{-1} (\boldsymbol{\tau}^{0+} \boldsymbol{f}).$$
(5.49)

This result can be interpreted as follows. When there are unequal signals, i.e. $\pi_j \neq 0$ and/or the condition (5.4) is violated the oscillation frequency of the phase locked state is changed and the change of the corresponding time interval is explicitly given (5.49). Denoting the eigenvectors of the equation (5.39) as usual v^k (instead of τ^k) the Green's function is explicitly given which allows us to write the solutions to (5.36) in the form

$$\boldsymbol{\tau} = \sum_{k \neq 0} \boldsymbol{v}^k \circ \boldsymbol{v}^{k+} \frac{1}{a(\lambda^0 - \lambda_k)} \boldsymbol{F} + \alpha \boldsymbol{\tau}^{\boldsymbol{\theta}}$$
(5.50)

where \circ denotes the direct product.

At this moment we are able to make contact with pattern recognition by the very same model. To this end, we proceed in several steps.

(1) In (5.50) the constant α can be chosen arbitrarily which is a consequence of the time translation invariance of (5.19). Without loss of generality we may put it equal to zero.

(2) The vector components of F(5.38) are of smallness of first order. This allows us to make an important step: In (5.50), in accordance with (5.35), we may replace v^k, v^{k+} , λ_k which stem from $\lambda_{jj'}^{\circ}$, by the corresponding quantities when the original matrix $(\lambda_{jj'})$ in (2.1) is decomposed into its eigenvectors. The total error made in (5.50) is then of second order. But now, the eigenvectors of $(\lambda_{jj'})$ are just the stored prototype vectors in the case of pattern recognition.

(3) According to (5.38), \mathbf{F} consists of two parts. For our subsequent discussion, we ignore the overall frequency shift Δ' and focus our attention on the role of π_j , i.e. the derivations of the true signal p_j from the constant signal p. This means that all the signatures of the offered pattern ("test pattern") are carried by π_j . Because of the scalar product $\mathbf{v}^{k+}\mathbf{F}$ those configurations π_j give the essential contributions to the phase shifts τ_j that correspond to the stored prototype patterns. We may suspect that this effect will be enhanced (in the sense of selection) if the nonlinear terms $\lambda_{jjj'}$ are taken into account. The role of the contributions $\sum_{j'} \kappa_{jj'}$ is hard to discuss. We assume that due to compensation effects they only play a minor role as compared to π_j .

(4) As a detailed study of Eq. 5.19 reveals, their solution becomes impossible if the π_j s become too large so that frequency locking breaks down (even if we go beyond the smallness approximation (5.30)).

Conclusion

The results of the present paper, jointly with those I have been publishing elsewhere (Haken 2006a, b) lead me to the following conclusions. Hereby we proceed in two steps. In the first step, in accordance with the present paper, we ignore the dynamics of the attention parameters λ_k , i.e. we assume that they are constant. Then the Eq. 2.1 provide us with a model that, depending on the stored (or learned) prototype patterns as well as on the externally offered patterns, describes two kinds of operations. If signals (i.e. pattern features) are similar (i.e. π_j in (5.26) are small) and the conditions (2.15), (2.16) approximately fulfilled, the responding neuronal populations synchronize, whereby phase-shifts τ_i depending on pattern variations may

occur. If the signals (features) vary too strongly, the (model) system acts like an attractor network (Haken 2006a, b). In this case, a pattern is encoded by a whole set of limit cycle oscillators. Each such set is governed by a specific order parameter in the sense of synergetics (Haken 2004).

In a second step, we must take into account that both the phase-locked states as well as the attractor states exist only for a limited time. Thus e.g. attractors must be replaced by what I called quasi-attractors. As the mechanism of this fading away of phase-locked and attractor states I suggested the saturation of attention, i.e. the attention parameters λ_k , vanish after a short time. Then new attractor states open etc. Thus, we have to deal with a complicated dynamics that still has to be explored in more detail.

References

- Bressloff PC, Coombes S (1998) Spike train dynamics underlying pattern formation in integrate-and-fire oscillator networks. Phys Rev Lett 81:2384–2387
- Bressloff PC, Coombes S (2000) Dynamics of strongly coupled spiking neurons. Neural Comput 12:91–129
- Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988) Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. Biol Cybern 60:121–130
- Gerstner W (1995) Time structure of the activity in neural network models. Phys Rev E 51:738–758
- Gray CM, Singer W (1989) Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc Natl Acad Sci USA 86:1698–1702
- Haken H (2002) Brain Dynamics, 2nd edn. Springer, Berlin
- Haken H (2004) Synergetic computers and cognition, 2nd edn. Springer, Berlin
- Haken H (2006a) Synergetics of brain function. Int J Psychophysiol 60:110–124
- Haken H. (2006b) Beyond attractor neural networks. Nonlinear Phenomena in Complex Systems. (to be published)
- Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. PNAS 72:2554– 2558
- Kandel, ER, Schwartz JH, Jessel TM (2000) Principles of neural sciences, 4th edn. Mc Graw Hill, New York
- Keener JP, Hoppensteadt FC, Rinzel J (1981) Integrate and fire models of nerve membrane response to oscillatory input. SIAM J Appl Math 41:503–517
- Mirollo RE, Strogatz SH (1990) Synchronization of pulsecoupled biological oscillators. SIAM J Appl Math 50:1645– 1662
- Singer W (1999a) Time as coding space. Curr Opin Neurobiol 9:189–194
- Singer W (1999b) Neural synchrony: a versatile code for the definition of relations? Neuron 24:49–65
- van Vreeswijk C (1996) Partial synchronization in populations of pulse-coupled oscillators. Phys Rev E 54:5522–5537