acta physica slovaca vol. 45 No. 2, 93 – 102 April 1995

SELF–ORGANIZATION OF PINWHEELS IN THE VISUAL CORTEX BY STOCHASTIC HEBB DYNAMICS: EQUIVALENCE TO KOSTERLITZ–THOULESS MODEL

Hans - Otto Carmesin[1]
Institut für Theoretische Physik, Universität Bremen, 28334 Bremen, Germany

Received 22 November 1994, accepted 30 November 1994

A neuron of area V1 of the human visual cortex typically responds preferentially to a specific stimulus orientation; thus an orientation is assigned to a neuron of area V1. These orientations establish a global order: they form circles at so called pinwheel singularities. In the present paper, the self- organization of that global order is explained as follows: The couplings from neurons of the retina to neurons of area V1 are modeled with a neurostatistical Hebb dynamics. As a result, neurons of area V1 exhibit orientation preference and neighbouring neurons exhibit an effective quadrupolar orientation interaction. So the global orientation emerges according to a quadrupolar model, this is mapped to an xy- model, the latter is eqivalent to the Kosterlitz–Thouless model, exhibits so called pinwheel singularities at any temperature T and the Kosterlitz–Thouless phase transition at a critical temperature T_c.

1. Introduction

Phenomena. The human brain can be regarded as a highly complex physical system that exhibits a variety of observable and measurable phenomena. In the present paper, the self- organization of orientation preferences of neurons in the visual area V1 is modeled. A neuron of that area responds preferentially to a specific orientation; so an orientation is assigned to each neuron of area V1. These orientations exhibit a global order: the orientations form circles around so called pinwheel singularities [2, 16, 20] (see Fig. 1.). Fred Wolf et al. showed that the experimentally measured orientations can be reproduced by assuming electrical charges at the singularities and by drawing the orientations parallel to the lines of constant electrostatic potential [19]. How does such a specific order emerge?

Dynamics. The neurons in the retina stimulate (indirectly) the neurons in area V1, the latter neurons fire according to that stimulation and to stochastic fluctuations.

[1] E-mail address: Carmesin@theo.physik.uni-bremen.de

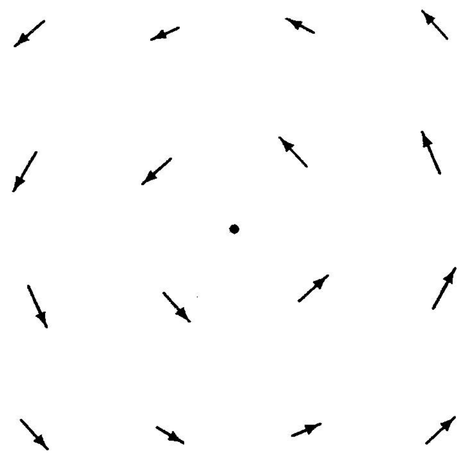

Fig. 1. **Pinwheel singularity.** Arrows: indicators for orientation preferences of neurons in area V1; directions of arrows are ignored here. •: singularity. The orientations are shown in the vicinity of one singularity. Formally similar singularites occur for electric charges in two dimensions (Kosterlitz- Thouless model) and for the xy- model of planar vector spins on the square lattice.

As a result of correlated neuronal activity, the synapses change according to the physiologically reasonable Hebb- rule [9, 10].

Analysis of dynamical properties. The whole dynamics exhibits ergodicity properties [6]. So the averaged coupling change can be determined. It turns out that the averaged coupling change is the *gradient of a potential* V. So the gradient of the potential can be interpreted as an *effective force* acting on the couplings in coupling space. Moreover, it turns out that local minima of the potential correspond to neurons with orientation preference and that neighbouring neurons of area V1 exhibit an *effective two orientation force* tending to make their orientations parallel. This effective force can be expressed as a *quadrupolar interaction* in leading order of a multipole expansion. The resulting quadrupolar model can be mapped to an *xy- model* [3]. The xy- model is equivalent to the *Kosterlitz- Thouless model* [12], that is a model for electrical charges in the plane; it exhibits the so called *pinwheel singularities* with orientations parallel to equipotential lines around electrical charges at singularities. In addition, a Kosterlitz-Thouless *phase transition* might occur at low temperature; it will be interesting to search for it in the physiological data in the future; in particular, one may investigate normal as well as pathological states of area V1.

Comparisons. Other models of orientational ordering in area V1 use other dynamical models (mainly Kohonen models containing an artificial minimum rule) instead of a neurostatistical Hebb- dynamics [11, 16, 7, 18, 15]. As a result, other studies of area

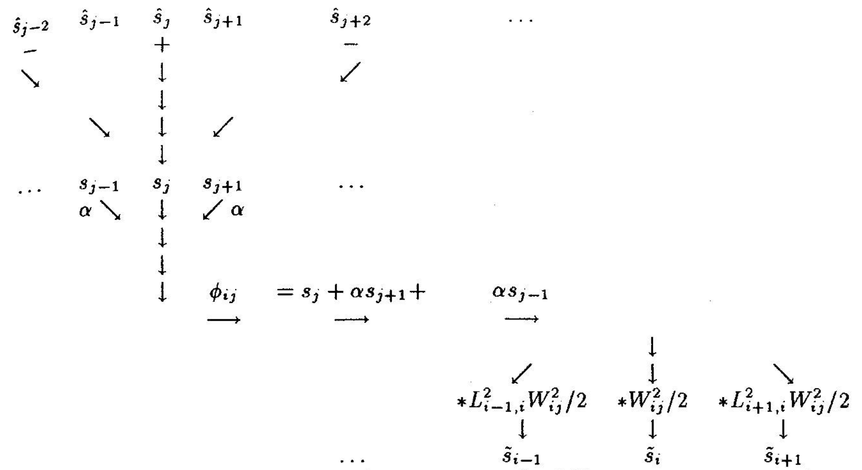

Fig. 2. **Network architecture with presynaptic ON- center neurons.** The network is illustrated in one dimension. Top: Sensor neurons $\hat{s}_j$ of the retina. Middle: Presynaptic ON- center neurons s_j of the retina. Bottom: postsynaptic neurons $\tilde{s}_i$ of area V1. Arrows: Transfer of electrical potentials ϕ_{ij}: Membrane potential in the nerve fibre from s_j to $\tilde{s}_i$. $W^2_{ij}/2$: afferent synaptic efficiency. $L^2_{i-1,i}/2$ and $L^2_{i+1,i}/2$: lateral synaptic efficiency. $L^2_{i-1,i}W^2_{ij}/2$ and $L^2_{i-1,i}W^2_{ij}/2$: synaptic efficiencies that transfer the membrane potential ϕ_{ij} to a neighbouring postsynaptic neuron.

V1 do neither identify a coupling dynamics according to a gradient of a potential, nor an equivalent xy- model, nor a Kosterlitz- Thouless type pinwheel singularity, nor a Kosterlitz- Thouless phase transition.

Altogether, the self- organization of area V1 is modeled straight forwardly with the simple and physiologically plausible neurostatistical Hebb- dynamics [6]. This yields the following results: Neurons in area V1 exhibit orientation preference. The ordering of orientation preferences takes place according to an xy- model, equivalent to the Kosterlitz- Thouless model. The experimentally observed [2, 19, 20] pinwheel structure is explained quantitatively by the xy- model. A Kosterlitz- Thouless phase transition is predicted for areas V1 with an effective temperature below the Kosterlitz- Thouless transition temperature.

Organization of the paper. The complete network model is specified in section 2. The complete analysis of the network model is presented in [6] and it is beyond the scope of the present paper. In section 3, the used results about the above mentioned potential V are cited from [6]; a similar potential is derived in [5], so the reader may refer to that paper for a rough description of the potential derivation. Moreover, the pinwheel structure is derived from the potential V and the conditions for the Kosterlitz-Thouless phase transition in the visual area V1 are derived in section 3. The explained phenomena, future perspectives and conclusions are summarized in section 4.

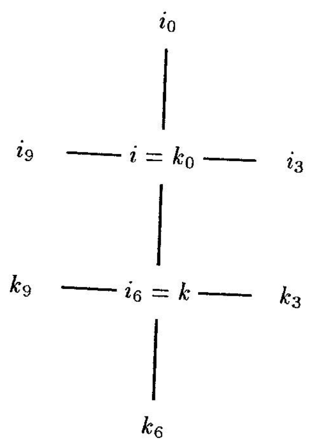

Fig. 3. **Neighbourship of neurons indicated according to the clock.** A neuron s_i has the neighbours s_{i_0} (top), s_{i_3} (right), s_{i_6} (bottom), s_{i_9} (left). Thus, if s_i is the neuron s_{k_0}, then s_k is the neuron s_{i_6}.

2. Network Model

Network architecture. Each neuron takes the value +1 or −1 at discrete time steps $t = 1, 2, 3, \ldots$. The network consists of a model retina and a model area V1 (see Fig. 2.), both are modeled as square lattices for simplicity. The retina receives uniformly distributed stimuli at the sensor neurons $\hat{s}_j$. The model retina (as well as the retina [10]) contains so called ON- center neurons s_j, each of which is stimulated by the corresponding sensor neuron $\hat{s}_j$ and inhibited by the four (on the modeled square lattice) next nearest (that is diagonal) neighbours of $\hat{s}_j$. Each ON- center neuron transfers its electrical potential via its axon to its specific neuron $\tilde{s}_i$ of the area V1; Thereby the axon transfers additional small potentials proportional to α and to $L^2 \approx \beta^2$ as indicated in Fig. 2. These small additional potentials are necessary and sufficient for the orientation self- organization of the area V1 [6]. The number of ON- center neurons is denoted by S; the number of postsynaptic neurons is denoted by I; for simplicity $I = S$ is studied here, while the genral case is investigated in [6].

Network dynamics. *In general, the neuronal dynamics* is characterized by a *transition probability* from presynaptic strates $\{s_j(t)\}$ to a postsynaptic state $s_i(t+1)$ according to the corresponding couplings K_{ij}

$$P[s_i(t+1)] = \frac{\exp[h_i(t+1)s_i(t+1)/T]}{2\cosh[h_i(t+1)s_i(t+1)/T]} \text{ with } h_i(t+1) = \sum_j K_{ij}(t)s_j(t). \quad (1)$$

Here T is a formal temperature that models *statistical fluctuations*, h_i is a formal local field and K_{ij} are the *couplings*.

The sensor *stimulation* is uniformly distributed

$$\hat{P}(\{\hat{s}_k\}): \; p(\hat{s}_k = 1) = p(\hat{s}_k = -1) = 1/2. \quad (2)$$

In the spirit of the above general neuronal dynamics, this gives rise to the following stimulation of the ON- center neurons.

$$P(\{s_j\}): \; p(s_j) = \frac{s_j h(\hat{s}_j) + 1}{2} \text{ with } h(\hat{s}_j) = \hat{s}_j/2 - x \sum_{\gamma\gamma}^{4} \hat{s}_{j_{\gamma\gamma}}/2 \text{ with}$$

$$x \leq 1/4, \text{ and with a positive model parameter } x. \quad (3)$$

Thereby the next nearest neighbours are marked by repeated indices $\gamma\gamma$ and the value $x_{max} = 1/4$ provides the dominance of the center neuron $\hat{s}_j$. The *indices of neighbouring neurons* are introduced according to the directions on the clock (see Fig. 3.).

The *particularly modeled neuronal dynamics* is characterized as follows.

$$\tilde{h}_i(t+1) = \frac{1}{2}\sum_j^S [K_{ij} + L^2_{i,i_0}K_{i_0j} + L^2_{i,i_3}K_{i_3j} + L^2_{i,i_6}K_{i_6j} + L^2_{i,i_9}K_{i_9j}]$$
$$[s_j + \alpha s_{j_0} + \alpha s_{j_3} + \alpha s_{j_6} + \alpha s_{j_9}]. \quad (4)$$

For short, the time arguments (t) for sensor neurons and $(t+1)$ for inner neurons $\tilde{s}_i$ and their formal fields $\tilde{h}_i$ are omitted; the time arguments for couplings result from the relations $K_{ij}(t+1) = K_{ij}(t) + \Delta K_{ij}(t)$ and $L_{ij}(t+1) = L_{ij}(t) + \Delta L_{ij}(t)$ and are omitted as well, analogously for W_{ij} below.

The Hebb type [9] *coupling dynamics* is explicated with the following difference equations

$$\Delta K_{ij} = 2aK_{ij}\,[s_j + \alpha s_{j_0} + \alpha s_{j_3} + \alpha s_{j_6} + \alpha s_{j_9}]$$
$$[\tilde{s}_i + L^2_{i_0,i}\tilde{s}_{i_0} + L^2_{i_3,i}\tilde{s}_{i_3} + L^2_{i_6,i}\tilde{s}_{i_6} + L^2_{i_9,i}\tilde{s}_{i_9}]. \quad (5)$$

I. e., the change of couplings corresponds to the contributing electrical potentials, for instance from s_{j_3} via $\tilde{s}_i$ to $\tilde{s}_{i_6}$ (see Fig. 2.)

The couplings are transformed with $K_{ij} = W^2_{ij}$. So one gets

$$\Delta W_{ij} = aW_{ij}[s_j + \alpha(s_{j_0} + s_{j_3} + s_{j_6} + s_{j_9})]$$
$$\times[\tilde{s}_i + L^2_{i_0,i}\tilde{s}_{i_0} + L^2_{i_3,i}\tilde{s}_{i_3} + L^2_{i_6,i}\tilde{s}_{i_6} + L^2_{i_9,i}\tilde{s}_{i_9}]. \quad (6)$$

Analogously one obtains

$$\Delta L_{i,i_\gamma} = aL_{i,i_\gamma}\sum_j^S \tilde{s}_i W^2_{i_\gamma j}\;[s_j + \alpha s_{j_0} + \alpha s_{j_3} + \alpha s_{j_6} + \alpha s_{j_9}]. \quad (7)$$

Coupling norms. According to empirical findings [17], the couplings at one neuron are roughly constant; this is modeled as follows.

$$\sum_j^S W^2_{ij} = q^2; \quad \sum_i^I W^2_{ij} = q^2; \quad L^2_{i_0,i} + L^2_{i_3,i} + L^2_{i_6,i} + L^2_{i_9,i} = \beta^2;$$
$$L^2_{i,i_0} + L^2_{i,i_3} + L^2_{i,i_6} + L^2_{i,i_9} = \beta^2. \quad (8)$$

Moreover, each lateral coupling has a maximal value L_{max}.

$$L_{i_\gamma,i} \leq L_{max}. \tag{9}$$

3. Analysis of the Network Model

Efficient notations. It is adequate to introduce the average of a term y with respect to presynaptic configurations as follows.

$$< y >_\mu := \sum_{\{s_j\}} P(\{s_j\}) y. \tag{10}$$

In the stable states that occur here as a result of the dynamics introduced above [6], each postsynaptic neuron $\tilde{s}_i$ is coupled afferently to exactly one presynaptic ON- center neuron; this ON- center neuron is denoted by $s_{u(i)}$. Analogously, in the stable states that occur here as a result of the dynamics introduced above [6], each presynaptic neuron s_i sends signals to exactly one postsynaptic neuron; this neuron is denoted by $\tilde{s}_{b(i)}$. Next one may introduce the functions

$$Q_i := < \Big([s_{u(i)} + \alpha s_{u(i)_0} + \alpha s_{u(i)_3} + \alpha s_{u(i)_6} + \alpha s_{u(i)_9}]$$
$$+ \beta^2 [s_{u(b(i))} + \alpha s_{u(b(i))_0} + \alpha s_{u(b(i))_3} + \alpha s_{u(b(i))_6} + \alpha s_{u(b(i))_9}] \Big)^2 >_\mu . \tag{11}$$

Cited results. In the human nerveous system, the neuronal changes take place on the time scale of milliseconds while the synaptic changes take place on the time scale of hours to years [10, 17]; so a corresponding *adiabatic limit* (that is the leading order of a systematic adiabatic approximation [8]) for the elimination of neuronal degrees of freedom is adequate. Moreover, the purpose of the present study is to investigate the self- organization of couplings; this is especially difficult for large fluctuations; so the leading order (called high temperature limit here) of the high temperature expansion (1/T- expansion) underestimates the organized structures and establishes a 'conservative approach'; in this sense a *high temperature limit* is particularly adequate. With these two limits one obtains the following

Potential theorem. *In the adiabatic limit of fast neuronal changes and slow coupling changes and in the limit of large fluctuations T holds: The averaged change of couplings generated by the combined neuronal and synaptic dynamics in the coupling space is proportional to the gradient of the potential V as follows:*

$$\langle \Delta W_{ij} \rangle = -\frac{\partial V}{\partial W_{ij}}; \quad \text{and } \langle \Delta L_{i_\gamma,i} \rangle = -\frac{\partial V}{\partial L_{i_\gamma,i}} \tag{12}$$

$$\text{with } V = -\frac{aq^4}{8T} \sum_{i=1}^{I} Q_i. \tag{13}$$

A proof is presented in [6].

A special case that is prototypical for orientation preference. The possible model parameters have been studied systematically in [6], while a prototypical case is presented here. One may consider the case $L_{max}^2 = \beta^2/2$. Then a neuron $\tilde{s}_i$ of area V1 is coupled laterally with *two* other neurons of area V1, say with $\tilde{s}_j$ and $\tilde{s}_k$. Moreover, each neuron $\tilde{s}_i$ in area V1 is coupled afferently with *one* ON- center neuron $s_{u(i)}$. In addition, the three ON- center neurons $s_{u(i)}$, $s_{u(j)}$ and $s_{u(k)}$ are nearest neighbours in the retina, so they are either in a row (∘ − • − ∘) or they form a right angle (⁰|•−∘). The arrangement ∘ − • − ∘ occurs, if the model parameter x is smaller than $8\alpha^2$ otherwise; the arrangement ⁰|•−∘ is stable, see [6]. It is obvious and shown explicitely in [6] that the neurons of area V1 exhibit orientation preference with the arrangement ∘ − • − ∘. This case is studied in the following.

Effective orientation interaction. One may consider two neighbouring neurons $\tilde{s}_i$ and $\tilde{s}_m$ in area V1, each of which has an arrangement ∘ − • − ∘ of presynaptic ON- center neurons in the retina. These two arrangements may be parallel or orthogonal. What potential V results from a parallel or from an orthogonal configuration of these two arrangements ∘ − • − ∘? That is, what potential difference ΔV occurs for parallel versus orthogonal arrangements ∘ − • − ∘? In order to answer this question precisely, one should consider all possible parallel and all possible orthogonal pairs of such arrangements ∘ − • − ∘. For simplicity, the parameter dependence of that potential difference ΔV is determined only roughly in the following.

In effect, two neighbouring neurons of area V1 receive common signals via presynaptic lateral signal transfer proportional to α and via postsynaptic lateral signal transfer proportional to $L^2 \approx \beta^2$ (see Fig. 2.) This gives rise to an attractive orientation interaction that is roughly proportional to $\alpha\beta^2$ and proportional to the length λ of contact between the two presynaptic arrangements ∘ − • − ∘. This length λ is at most *one neuron for an othogonal* configuration of arrangements ∘ − • − ∘ and at most *three neurons for a parallel* configuration of arrangements ∘ − • − ∘. So the potential difference is roughly proportional to the length difference times $\alpha\beta^2$, i. e.,

$$\Delta V \approx \Delta\lambda\alpha\beta^2 \approx 2\alpha\beta^2. \tag{14}$$

Next one may use the considered lattice model as an approximation of a continuum model. In the lattice model, there occur just two orientations of an arrangement ∘ − • − ∘, vertical or horizontal. But in a continuum model, there occur arbitrary orientations. How could one estimate the effective orientation interaction of a continuum model from the above ΔV of the lattice model? For this purpose, a multipole expansion [13] is adequate, because it takes care of the symmetry properties of orientation preferences and because it typically converges quite rapidly. The leading order orientation interaction between two orientations with polar angles ϕ_i and ϕ_m is (in the most simple case [4]) the quadrupole interaction as follows

$$\Delta V_{cont} = -A \cos^2(\phi_i - \phi_m), \tag{15}$$

with a parameter A. This parameter A can be fitted to the results for the lattice model, that is,

$$\Delta V_{cont} = -A = -2\alpha\beta^2 \text{ so } A \approx 2\alpha\beta^2. \tag{16}$$

Accordingly, it is adequate to model continuous orientation preferences in leading order of the multipole expansion with the following effective nearest neighbour orientation interaction

$$\Delta V_{cont} = -2\alpha\beta^2 \cos^2(\phi_i - \phi_m). \tag{17}$$

This model is studied in the rest of the paper.

Mapping to the xy- model. Due to the relation $\cos^2\phi = [\cos(2\phi) + 1]/2$, the statistical phsyics of the quadrupolar model of Eq. (17) and at temperature T is the same [3] as that of the planar vector model (alias xy- model) at temperature T and with the following nearest neighbour interaction

$$\Delta V_{cont} = -\alpha\beta^2 \cos(2[\phi_i - \phi_m]). \tag{18}$$

So far, the positions of the retinal ON- center neurons have been chosen continuously in order to obtain continuous orientations; whereas the positions of the neurons in area V1 are still on a square lattice. So the above xy- model is described by the orientational energy function

$$H = -J \sum_{<i,j>}^{\text{nearest neighbours on square lattice}} \cos(\phi_i - \phi_m) \text{ with } J = \alpha\beta^2. \tag{19}$$

Thereby the sum $\sum_{<i,j>}$ denotes summation of the nearest neighbour pairs on the square lattice. This model is equivalent to the Kosterlitz- Thouless model [12, 1] and its properties can thus be recalled from the literature as follows.

Properties of the ordering of orientation preferences. Each orientation configuration at a local minimum of the energy H is specified by z pairs of singularities with some non- negative number z. The number z is not zero at finite T, because at zero z, the system would exhibit long range orientational order, in contrast to the Mermin Wagner theorem [14]. In the vicinity of a singularity, the orientations are as in Fig. 1.

Each singularity is characterized by a position $\vec{r}_i$ and by an effective charge $q_i = \pm1, \pm2, \pm3, \ldots$. Thereby the sum of effective charges vanishes, $\sum_i q_i = 0$.

One may denote the lattice spacing by r_0 and separate the spin wave energy from the energy $H_{singularities}$ due to the so called pinwheel singularities. The latter is

$$H_{singularities} \approx -2\pi J \sum_{i \neq j} q_i q_j \ln \left| \frac{\vec{r}_i - \vec{r}_j}{r_0} \right|. \tag{20}$$

The orientation correlation function of two orientations $(\cos\phi_i, \sin\phi_i)$ and $(\cos\phi_j, \sin\phi_j)$ is

$$\langle \cos\phi_i \cos\phi_j + \sin\phi_i \sin\phi_j \rangle \approx \left| \frac{\vec{r}_i - \vec{r}_j}{r_0} \right|^{-\frac{T}{4\pi J}}. \tag{21}$$

The singularities form pairs at temperatures below the Kosterlitz- Thouless transition temperature T_c with

$$T_c \approx \frac{8}{9}\pi J. \tag{22}$$

4. Summary

If one regards the Hebb- rule as the first principles dynamics of self- organization in the cortex, then it has been shown from first principles, how pinwheel structures (see Fig. 1.) form in area V1 of the visual cortex at any nonzero fluctuation rate. This formation dynamics is equivalent to that of the Kosterlitz- Thouless model, i. e. to that of electrical charges in the plane. Thereby the charges correspond to topological singularities of orientation configurations.

It is remarkable that this orientation ordering in area V1 is formally equivalent to the orientation ordering of the xy- model, though the first is due to a nonequilibrium process of the open nerveous system, whereas the second is due to an equilibrium state of a closed statistical mechanical system: The formal equivalence is due to the effective potential dynamics in coupling space (see Eq. (12)).

This study explains the following findings: Orientation preferences can be quantitatively described as equipotential lines of an electrostatic model system with charges at the singularities [19]. The number of singularities with clockwise rotation is equal to that with counter clockwise rotation [16].

So far it was an open question, whether the pinwheel singularities observed in area V1 are in some possibly dynamical equilibrium state, or whether they or some frozen relicts of an initial state. The present study provides a clear answer: they are in a dynamical equilibrium, because a state without singularities contradicts the Mermin Wagner theorem, as a result of the present analysis. A state without singularities would be pathological, because only one orientation preference would occur; in this sense the present result shows why the orientation system is robust against a uniform orientation preference state.

This study predicts a Kosterlitz- Thouless phase transition of the orientation singularites of the visual area V1, at some critical fluctuation rate T_c. It will be intersting to identify the conditions of T_c in physiological terms, and to possibly find the phase transition in the future.

Acknowledgements: I am grateful for stimulating discussions with H. Haken, J. L. van Hemmen, K. Pawelzik, G. Roth, O. Scherf, H. Schwegler and F. Wolf.

References

[1] P. W. Anderson: *Basic Notions of Condensed Matter Physics* (Addison- Wesley, Redwood City 1984).
[2] T. Bonhoeffer, A. Grinvald: *Nature* **343** (1991) 429;
[3] H.-O. Carmesin: *Phys. Lett.* **A 125** (1987) 294;
[4] H.-O. Carmesin: *Z. Phys.* **B 73** (1988) 381;
[5] H.-O. Carmesin: *Acta Physica Slovaca* **44** (1994) 311;
[6] H.-O. Carmesin: *Theorie neuronaler Adaption* (Köster, Berlin 1994, ISBN 3-89574-020-9).
[7] R. Durbin, G. Mitchison: *Nature* **343** (1990) 644;
[8] H. Haken: *Advanced Synergetics* (Springer, Berlin 1983).
[9] D. O. Hebb: *The Organization of Behaviour* (Wiley, New York 1949) p. 62.

[10] E. R. Kandel, J. H. Schwarz, T. M. Jessell: *Principles of Neural Science* (Elsevier, New York 1991).

[11] T. Kohonen: *Self - Organization and Associative Memory* (Springer, Berlin Heidelberg 1989).

[12] J. M. Kosterlitz, D. J. Thouless: *J. Phys.* **C 6** (1973) 1181;

[13] L. D. Landau, E. M. Lifschitz: *Lehrbuch der Theoretischen Physik*, Band II, (Akademie-Verlag, Berlin 1981).

[14] N. D. Mermin, H. Wagner: *Phys. Rev. Lett* **29** (1966) 1133;

[15] K. D. Miller: *J. Neuroscience* **14** (1994) in press;

[16] K. Obermayer, G. G. Blasdel, K. Schulten: *Phys. Rev.* **A45** (1992) 7568;

[17] H. Reichert: *Neurobiologie* (Thieme, Stuttgart - New York 1990).

[18] J.Sirosh, R.Miikkulainen: *Biol. Cybern.* **71** (1994) 1994; 65

[19] F. Wolf, K. Pawelzik, T. Geisel, D.- S. Kim, T. Bonhoeffer: *Verhandlungen DPG (IV)* **29** (1994) 955;

[20] F. Wolf, M. Riesenhuber, T. Geisel: *Proceedings of the 22nd Göttingen Neurobiology Conference*, ed. von N. Elsner & H. Breer (Thieme, Stuttgart 1994), p. 540.