Network: Comput. Neural Syst. **9** (1998) 449–466. Printed in the UK PII: S0954-898X(98)91993-5

Analysis of a correlation-based model for the development of orientation-selective receptive fields in the visual cortex

S Wimbauer, W Gerstner† and J L van Hemmen

Physik Department, Technische Universität München, D-85747 Garching bei München, Germany

Received 24 February 1998

Abstract. We analyse a model for the development of orientation-selective receptive fields of simple cells in a locally connected network of cortical neurons. The Hebbian learning rule that underlies the development is described by a linear differential equation. The structure of the emerging cortical map can be predicted by deriving the eigenfunctions corresponding to the leading eigenvalues of the associated matrix. We show that the receptive fields have the typical form of a wavelet. Mathematically, receptive fields are given by a Hermitian polynomial with Gaussian cut-off and a phase factor. Both the phase of the wavelet and the orientation are changing periodically along the surface of the cortical map as suggested by previous simulation studies and as also found in experiments. In order to get orientation-selective receptive fields, the spatial correlation function of the inputs that drive the development must have a zero crossing.

1. Introduction

Processing of visual information proceeds in several stages along the visual pathway. The optic nerve originating from the retina carries the information to the lateral geniculate nucleus (LGN) which is a part of the thalamus, a larger brain nucleus. From there nerve fibres project to the visual cortex, where most of the higher processing takes place.

Cells in the LGN and the visual cortex are characterized by their receptive field defined as the area on the retina that is to be stimulated by an appropriate light pattern so as to evoke a response of the neuron under consideration. In the LGN, two main types of cell can be distinguished, the ON cells and the OFF cells. An ON cell responds best if a bright spot is projected into the centre of its receptive field. Similarly, an OFF cell shows maximal response if a dark spot (surrounded by a weakly illuminated background) covers the centre of its receptive field. The LGN receptive fields are concentric.

Unlike neurons in the LGN [HW61], cortical cells respond selectively to more specific spatial and temporal stimulus parameters, in particular, to the orientation of bars (or edges) located within the cells' receptive field [HW59]. Cortical simple cells receive direct input from the LGN and represent the first processing stage within the visual cortex. According to a classical model of Hubel and Wiesel [HW62], orientation selectivity of cortical cells is based on the convergence of ON and OFF type LGN neurons onto a simple cell. These inputs give rise to ON and OFF subregions within the receptive field of a simple cell that are aligned along its preferred orientation. A considerable amount of experimental evidence is in agreement with this picture [Fer87, CZS91, RAW94, RA95].

† Present address: Center for Neuro-Mimetic Systems MANTRA, EPFL, CH-1015 Lausanne, Switzerland.

0954-898X/98/040449+18$19.50 © 1998 IOP Publishing Ltd

Neurons in the cortex form a thin sheet that covers most of the surface of the brain. Receptive field properties such as the preferred orientation vary systematically along the surface of the cortical sheet. The two-dimensional organization of the neuronal response properties is called a cortical map.

Because of the huge number of neurons (about 10^9 in primate visual cortex) and 10^4 synapses per neuron, it is hard to imagine how the complete wiring that underlies the receptive field properties of cortical cells could be specified genetically. Rather, it seems likely that the development of receptive field properties is driven by an activity-dependent learning process which tunes the connectivity pattern during a critical period, a developmental phase which usually starts shortly before and lasts up to a few weeks after birth; for a review see [SS82].

Beginning with von der Malsburg [vdM73], numerous models have been formulated that describe activity-dependent learning as a self-organization process according to a Hebbian rule [Lin86c, Lin86a, Lin86b, YKC89]. In this article we study analytically the properties of the developmental model of Miller [Mil94]. The Hebbian learning rule driving the development leads to a linear differential equation. We derive the eigenfunctions of the matrix associated with the learning equation. The eigenfunction to the leading eigenvalue determines the shape of the neuronal receptive field and also the map structure of cortical cells. The central quantity driving the developmental process is the correlation function of the activity of LGN cells. It has been proposed [Mil94] that this correlation function must change sign for orientation selectivity to emerge. It will become clear from our analysis why this is indeed the case.

The paper is organized as follows. Section 2 introduces the main elements of the developmental model of [Mil94]. In section 3 the eigenfunction and eigenvalues of the developmental equation are derived. The results of this analysis are then compared with the outcome of a simulation run of the model in section 4. We conclude our considerations with a discussion in section 5.

2. The developmental model

In this section we recapitulate the main elements of the developmental model introduced in [Mil94]. As was mentioned above, the model is based on an activity-dependent learning rule and aims at a description of the emergence of orientation-selective receptive fields of simple cells.

Before turning to the learning equation itself, we have to clarify the structure of the model network. In agreement with Hubel and Wiesel [HW62], the spatial receptive field of a simple cell is determined by the convergence of two types of input, i.e. inputs from ON and OFF type cells in the LGN. In the model under consideration, cortical cells and cells in the LGN are arranged on a two-dimensional grid. The vector $\boldsymbol{x} = (x, y)$ denotes the position of a cell on the cortical grid and $\boldsymbol{\alpha} = (\alpha, \beta)$ labels a position on the LGN grid; see figure 1. Both grids serve to support so-called retinotopic maps that can be observed in the LGN and the visual cortex. By a retinotopic map we mean that neighbouring cells in the cortex and the LGN correspond to neighbouring positions on the retina. That is, a retinotopic map conserves the neighbourhood relations of the visual input. The positions $\boldsymbol{\alpha}$ and $\boldsymbol{x}$ are given in the same retinotopic coordinates and, hence, can be added and subtracted arbitrarily.

To obtain the local potential $h(\boldsymbol{x}, t)$ of a cortical simple cell, the mean firing rates of cells in the LGN $a^c(\boldsymbol{\alpha}, t)$ (with $c = \mathrm{ON}, \mathrm{OFF}$) are summed and weighted with the efficacy $W^c(\boldsymbol{x}, \boldsymbol{\alpha})$ of the feedforward synapses. Furthermore, simple cells receive input

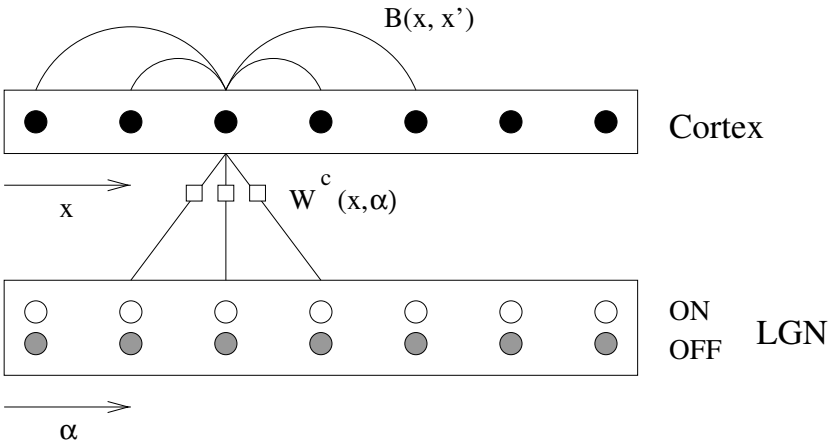

Figure 1. A cortical simple cell receives input from ON and OFF type neurons in the LGN. These inputs are weighted by the feedforward synapses $W^c(\boldsymbol{x}, \boldsymbol{\alpha})$. In addition, other cortical neurons contribute to the local potential of simple cells, where the function $B(\boldsymbol{x}, \boldsymbol{x}')$ denotes the strength of intracortical synapses. It is assumed that only weights between the LGN and the cortex are modified during development.

from other cortical cells with intracortical connection strength $B(\boldsymbol{x}, \boldsymbol{x}')$. In good agreement with experimental results [Fer94], we assume linearity and obtain the following functional form for the local potential of a simple cell at cortex location $\boldsymbol{x}$ and time t:

$$h(\boldsymbol{x}, t) = \sum_{c=\mathrm{ON,OFF}} \sum_{\boldsymbol{\alpha}} W^c(\boldsymbol{x}, \boldsymbol{\alpha})\, a^c(\boldsymbol{\alpha}, t) + \sum_{\boldsymbol{x}'} B(\boldsymbol{x}, \boldsymbol{x}')\, h(\boldsymbol{x}', t). \tag{1}$$

Since equation (1) is purely linear it can be rewritten as

$$h(\boldsymbol{x}, t) = \sum_{\boldsymbol{x}'} I(\boldsymbol{x}, \boldsymbol{x}') \sum_{c=\mathrm{ON,OFF}} \sum_{\boldsymbol{\alpha}} W^c(\boldsymbol{x}', \boldsymbol{\alpha})\, a^c(\boldsymbol{\alpha}, t) \tag{2}$$

with

$$I^{-1}(\boldsymbol{x}, \boldsymbol{x}') = \mathbf{1} - B(\boldsymbol{x}, \boldsymbol{x}') \tag{3}$$

where I^{-1} denotes the inverse of the matrix I. For the sake of simplicity we assume that the long-time average of the rates $a^c(\boldsymbol{\alpha}, t)$ vanishes. In other words, the quantity $a^c(\boldsymbol{\alpha}, t)$ denotes the deviation of the momentary firing rate from the average mean firing rate and can take both positive and negative values.

It is commonly assumed that the activity-driven modification of synapses takes place according to some type of Hebbian learning rule [Heb49]. A Hebbian synapse is strengthened, if the pre- and postsynaptic activity or depolarization are positively correlated, and is kept fixed or weakened in the presence of negative correlations. Long-term potentiation (LTP) has been proposed as a cellular mechanism that might underlie Hebbian learning [CM95, CPCD90]. In the model under consideration only feedforward synapses $W^c(\boldsymbol{x}, \boldsymbol{\alpha})$ between the LGN and the cortex are modified by Hebbian learning, whereas intracortical synapses are kept fixed. One assumes that the change in synaptic strength at time t is determined by the correlation between the postsynaptic local potential $h(\boldsymbol{x}, t)$ and the deviation $a^c(\boldsymbol{\alpha}, t)$ of the presynaptic cell activity from the average firing rate. We consider the mean change of synaptic efficacy within a learning window of

time Λ and find

$$\frac{\mathrm{d}W^c(\boldsymbol{x},\boldsymbol{\alpha},t)}{\mathrm{d}t} = \eta A(\boldsymbol{x}-\boldsymbol{\alpha})\,\frac{1}{\Lambda}\int_0^\Lambda \mathrm{d}s\, h(\boldsymbol{x},t-s)\,a^c(\boldsymbol{\alpha},t-s), \tag{4}$$

where $A(\boldsymbol{\alpha}-\boldsymbol{\alpha})$ denotes an arbor function that restricts the possible synaptic wiring that might emerge from the learning process to a spatially convergent structure as shown in figure 1. In particular, the arbor function $A(\boldsymbol{\alpha}-\boldsymbol{\alpha})$ favours connections to a presynaptic cell with the same retinotopic coordinate as the cortical cell and suppresses connections to LGN cells further away; cf. equation (7) below.

By inserting (2) into (4) we obtain

$$\frac{\mathrm{d}W^c(\boldsymbol{x},\boldsymbol{\alpha},t)}{\mathrm{d}t} = A(\boldsymbol{x}-\boldsymbol{\alpha})\sum_{\boldsymbol{x}'} I(\boldsymbol{x},\boldsymbol{x}')\sum_{c'=\mathrm{ON,OFF}}\sum_{\boldsymbol{\alpha}'} C^{c,c'}(\boldsymbol{\alpha},\boldsymbol{\alpha}')W^{c'}(\boldsymbol{x}',\boldsymbol{\alpha}',t) \tag{5}$$

where the correlation function between the activity of two cells in the LGN is given by

$$C^{c,c'}(\boldsymbol{\alpha},\boldsymbol{\alpha}') = \frac{\eta}{\Lambda}\int_0^\Lambda \mathrm{d}s\, a^c(\boldsymbol{\alpha},s)\,a^{c'}(\boldsymbol{\alpha}',s). \tag{6}$$

In deriving (5) we have made two additional assumptions. First, $W^c(\boldsymbol{x},\boldsymbol{\alpha},t)$ changes slowly on the time scale of the learning window Λ and can be kept fixed during its duration. Mathematically, this means that '$0 < \eta \ll 1$' so that we can perform a local averaging [SV85] in (5). Second, $a^c(\boldsymbol{\alpha},s)$ fluctuates fast as compared to Λ. Due to the averaging over time and the random character of the input patterns $a^c(\boldsymbol{\alpha},t)$ which are drawn from a given stationary distribution, the correlation function $C^{c,c'}(\boldsymbol{\alpha},\boldsymbol{\alpha}')$ does not depend on the time t. In short, we have performed a separation of time scales.

According to (5), the development of the synaptic weights is now determined by three spatial functions, namely, the arbor function $A(\boldsymbol{x}-\boldsymbol{\alpha})$, the intracortical interaction function $I(\boldsymbol{x},\boldsymbol{x}')$ and the correlation function $C^{c,c'}(\boldsymbol{\alpha},\boldsymbol{\alpha}')$. In the following the arbor function is chosen to be the Gaussian

$$A(\boldsymbol{x}-\boldsymbol{\alpha}) = \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{\alpha}|^2}{2A}\right) \tag{7}$$

so that synapses will develop mainly for short distances $|\boldsymbol{x}-\boldsymbol{\alpha}|$. Similarly, excitatory coupling between neighbouring cortical neurons is modelled by a Gaussian intracortical interaction function

$$I(\boldsymbol{x},\boldsymbol{x}') = I(\boldsymbol{x}-\boldsymbol{x}') = \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{x}'|^2}{2I}\right). \tag{8}$$

It has been suggested previously [Mil94] that the correlation function for the activity of two cells at positions $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}'$ on the LGN grid should change sign at least once as a function of the distance $|\boldsymbol{\alpha}-\boldsymbol{\alpha}'|$ for orientation selectivity to develop. For this reason, we cannot take a simple Gaussian but we have to somehow induce a change of sign. In this article we use the following form for the correlation function,

$$\begin{aligned} C^{\mathrm{ON,ON}}(\boldsymbol{\alpha},\boldsymbol{\alpha}') &= C^{\mathrm{OFF,OFF}}(\boldsymbol{\alpha},\boldsymbol{\alpha}') = \exp\left(-\frac{|\boldsymbol{\alpha}-\boldsymbol{\alpha}'|^2}{2C}\right) - k \\ C^{\mathrm{ON,OFF}}(\boldsymbol{\alpha},\boldsymbol{\alpha}') &= C^{\mathrm{OFF,ON}}(\boldsymbol{\alpha},\boldsymbol{\alpha}') = -\epsilon\, C^{\mathrm{ON,ON}}(\boldsymbol{\alpha},\boldsymbol{\alpha}') \end{aligned} \tag{9}$$

with $0 \leqslant k \leqslant 1$ and $0 \leqslant \epsilon \leqslant 1$. For $k = 0$, the correlation function (9) $C^{\mathrm{ON,ON}}$ is strictly positive whereas for $k > 0$ there is a zero crossing at some finite distance.

The correlation function (9) can be motivated as follows. Before eye-opening, the activity of the photo receptors fluctuates in a spatially and temporally uncorrelated way.

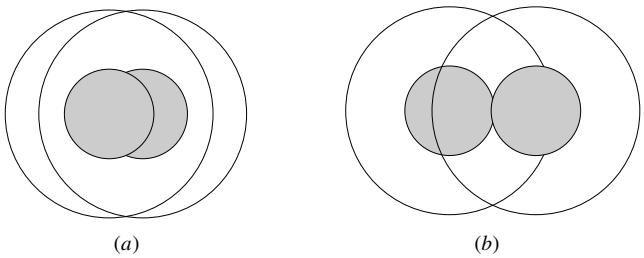

Figure 2. The receptive fields of neurons in the LGN are of the centre-surround type. A centre ON cell responds positively, if the centre of the receptive field is stimulated by a bright dot, and negatively if the bright dot samples its surroundings (bright). We concentrate on the correlation between two centre ON cells. (*a*) For small $|\boldsymbol{\alpha} - \boldsymbol{\alpha}'|$ the receptive field centres of both neurons overlap. Therefore, the cells will respond similarly to a stimulation, and their activities are correlated. (*b*) For large $|\boldsymbol{\alpha} - \boldsymbol{\alpha}'|$ the centre of one receptive field falls onto the surroundings of the other one. Thus, a stimulation results in a response with opposite signs in both cells and their activities are anti-correlated. The correlation function therefore has a zero crossing at some intermediate distance. The behaviour of the correlation function between two OFF or an ON and an OFF cell can be explained in a similar way.

Due to the wiring pattern between the retinal photo receptors and the cells in the LGN, uncorrelated noise in the retina is transformed into a correlated activity of cells in the LGN with a correlation function like the one given in (9); cf. figure 2. The above correlation function is also in agreement with recent experimental findings [MLB95]; for a more detailed discussion, see [Mil94]. The function (9) does not go to zero for $|\boldsymbol{\alpha} - \boldsymbol{\alpha}'| \to \infty$, as one might expect from a biological point of view. It should be noted, however, that the effect of $C^{c,c'}(\boldsymbol{\alpha}, \boldsymbol{\alpha}') \neq 0$ for large $|\boldsymbol{\alpha} - \boldsymbol{\alpha}'|$ is suppressed in equation (5) by the arbor function A, which decreases rapidly for large distances. Thus, the error we make if we take (9) instead of a more realistic correlation function will be negligible.

Since the Hebbian learning equation (5) is an ordinary linear differential equation, it can be solved by expanding $W^c(\boldsymbol{x}, \boldsymbol{\alpha}, t)$ into the eigenfunctions $W_a^c(\boldsymbol{x}, \boldsymbol{\alpha})$ of the operator on the right-hand side of (5). One obtains

$$W^c(\boldsymbol{x}, \boldsymbol{\alpha}, t) = \sum_a j_a(0) \exp(\lambda_a t) W_a^c(\boldsymbol{x}, \boldsymbol{\alpha}) \tag{10}$$

where λ_a is the eigenvalue corresponding to the a-th eigenfunction and $j_a(0)$ the projection of the random initial conditions onto the respective eigenfunction.

As will become obvious in the next section, positive eigenvalues λ_a always exist and, hence, synaptic weights would increase without bounds. To avoid such a biologically implausible behaviour, an upper and a lower bound for the couplings are introduced explicitly,

$$0 \leqslant W^c(\boldsymbol{x}, \boldsymbol{\alpha}, t) \leqslant W_{\max} A(\boldsymbol{x} - \boldsymbol{\alpha}). \tag{11}$$

The lower bound is zero since only positive (excitatory) synapses from the LGN to a cortical simple cell can be observed experimentally [FL83, Fer88].

If the couplings are distributed randomly at the beginning, all eigenvectors start with a similar amplitude. The components of W belonging to the largest eigenvalues, however, will grow fastest and reach the lower and upper bounds first, where the weights are frozen as long as W wants to move outwards. Thus the eigenfunctions corresponding to the largest eigenvalues determine the emerging receptive fields [MM90, MM94]. To predict

the receptive fields that emerge from the Hebbian learning process, we therefore turn to an analysis of the eigenvalue problem associated with equation (5).

3. Derivation of the eigenfunctions and eigenvalues of the learning equation

In this section we solve the eigenvalue problem corresponding to the learning equation (5),

$$\lambda W^c(\boldsymbol{x},\boldsymbol{\alpha}) = A(\boldsymbol{x}-\boldsymbol{\alpha})\sum_{\boldsymbol{x}'} I(\boldsymbol{x},\boldsymbol{x}') \sum_{c'=\mathrm{ON,OFF}} \sum_{\boldsymbol{\alpha}'} C^{c,c'}(\boldsymbol{\alpha},\boldsymbol{\alpha}')W^{c'}(\boldsymbol{x}',\boldsymbol{\alpha}'). \tag{12}$$

In the following analysis we will proceed in several steps. In the first step we single out the part of the eigenvalue problem that depends on the variable c only. It will become clear that as a result of the development only one type of synapse, i.e. either ON or OFF synapses, will link a cortical position $\boldsymbol{x}$ and an LGN position α. In the second step we consider the dependence of the eigenfunction on the cortical variable $\boldsymbol{x}$. We will show that the dominating eigenfunction is a combination of Fourier modes corresponding to wave vectors of the same length but with different directions. Hence a periodic change of receptive field properties is observed along the cortical sheet. In the third step we investigate the shape of the receptive field for one cortical cell. An exact solution of this part of the eigenvalue problem can be found for the special case of a Gaussian correlation function. This solution that is related to the well known problem of the quantum-mechanical harmonic oscillator can then be used to approximately solve the problem for more general correlation functions. As has been predicted by Miller [Mil94], it turns out that a sign change of the correlation function is a *necessary* prerequisite for the development of orientation-selective receptive fields. We now turn to the details of the mathematical analysis.

Because of the symmetry of the correlation function (9) in the variable c, the eigenvalue problem can be split into a part that depends on c only and another part that depends on $\boldsymbol{\alpha}$ and $\boldsymbol{x}$ only,

$$\begin{aligned}\left\{\lambda_\gamma W_\gamma^c\right\}\left[\lambda_\chi W_\chi(\boldsymbol{x},\boldsymbol{\alpha})\right] = &\left\{\sum_{c'=\mathrm{ON,OFF}} C^{c,c'}W_\gamma^{c'}\right\} \\ &\times\left[A(\boldsymbol{x}-\boldsymbol{\alpha})\sum_{\boldsymbol{x}'} I(\boldsymbol{x},\boldsymbol{x}')\sum_{\boldsymbol{\alpha}'} C(\boldsymbol{\alpha},\boldsymbol{\alpha}')W_\chi(\boldsymbol{x}',\boldsymbol{\alpha}')\right].\end{aligned} \tag{13}$$

Both parts can now be solved separately [MS90].

For the γ part of the eigenvalue problem

$$\lambda_\gamma \boldsymbol{W}_\gamma = \begin{pmatrix} 1 & -\epsilon \\ -\epsilon & 1 \end{pmatrix}\boldsymbol{W}_\gamma \tag{14}$$

one obtains the solution

$$\begin{aligned} \boldsymbol{W}_\gamma^1 &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \lambda_\gamma^1 = 1+\epsilon \\ \boldsymbol{W}_\gamma^2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \lambda_\gamma^2 = 1-\epsilon. \end{aligned} \tag{15}$$

For $\epsilon > 0$ the eigenfunction with opposite sign for the ON and OFF components dominates the dynamics of the learning equation. That is, for fixed coordinates $\boldsymbol{\alpha}$ and $\boldsymbol{x}$, the ON weight will grow and finally saturate at the upper bound, whereas the OFF weight will fall to the lower bound $W_{\min} = 0$, or vice versa. At the end of the developmental process *either* ON *or* OFF weights will therefore exist between a cortical position $\boldsymbol{x}$ and an LGN position $\boldsymbol{\alpha}$, but not both. The question of whether ON *or* OFF synapses develop for a given

coordinate, is determined by the sign of the dominating spatial eigenfunction $W_\chi(\boldsymbol{x}, \boldsymbol{\alpha})$. In particular, ON weights will develop in regions with $W_\chi(\boldsymbol{x}, \boldsymbol{\alpha}) > 0$ and OFF weights in regions with $W_\chi(\boldsymbol{x}, \boldsymbol{\alpha}) < 0$. This is the first important result of our analysis. We will make use of the above considerations throughout the rest of the paper.

We now turn to the solution of the χ part of the eigenvalue problem. We will drop the index χ in the following. To simplify our analysis, the discrete variables $\boldsymbol{\alpha}$ and $\boldsymbol{x}$ are replaced by continuous ones. Using the expressions (7), (8) and (9) for $A(\boldsymbol{x}-\boldsymbol{\alpha})$, $I(\boldsymbol{x}, \boldsymbol{x}')$ and $C^{\mathrm{ON,ON}}(\boldsymbol{\alpha}, \boldsymbol{\alpha}')$, we obtain

$$\lambda W(\boldsymbol{x}, \boldsymbol{\alpha}) = \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{\alpha}|^2}{2A}\right) \int \mathrm{d}\boldsymbol{x}' \int \mathrm{d}\boldsymbol{\alpha}' \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{x}'|^2}{2I}\right) \times \left[\exp\left(-\frac{|\boldsymbol{\alpha}-\boldsymbol{\alpha}'|^2}{2C}\right) - k\right] W(\boldsymbol{x}', \boldsymbol{\alpha}'). \tag{16}$$

We introduce new variables in order to symmetrize the eigenvalue problem

$$\overline{W}(\boldsymbol{x}, \boldsymbol{\alpha}-\boldsymbol{x}) = \exp\left(\frac{|\boldsymbol{x}-\boldsymbol{\alpha}|^2}{4A}\right) W(\boldsymbol{x}, \boldsymbol{\alpha}) \tag{17}$$

and replace $\boldsymbol{\alpha}$ by a new coordinate $\Delta\boldsymbol{\alpha} = \boldsymbol{\alpha} - \boldsymbol{x}$ that denotes the position of an LGN cell relative to the position of a cortical cell. This yields

$$\lambda \overline{W}(\boldsymbol{x}, \Delta\boldsymbol{\alpha}) = \exp\left(-\frac{|\Delta\boldsymbol{\alpha}|^2}{4A}\right) \int \mathrm{d}\boldsymbol{x}' \int \mathrm{d}\Delta\,\boldsymbol{\alpha}' \exp\left(-\frac{|\boldsymbol{x}-\boldsymbol{x}'|^2}{2I}\right) \times \left[\exp\left(-\frac{\left|(\boldsymbol{x}-\boldsymbol{x}') + (\Delta\boldsymbol{\alpha} - \Delta\boldsymbol{\alpha}')\right|^2}{2C}\right) - k\right] \exp\left(-\frac{|\Delta\boldsymbol{\alpha}'|^2}{4A}\right) \overline{W}(\boldsymbol{x}', \Delta\boldsymbol{\alpha}'). \tag{18}$$

Because (18) is a symmetric eigenvalue problem, all eigenvalues are real.

The above problem of determining the eigenvalues can be simplified considerably if a Fourier transformation of (18) is performed with respect to the cortical coordinate $\boldsymbol{x}$. The corresponding variable in Fourier space is the vector $\boldsymbol{l}$ with components l and m. Then (18) reads

$$\lambda \overline{W}^{l,m}(\Delta\boldsymbol{\alpha}) = 2\pi \exp\Big(-\frac{|\Delta\boldsymbol{\alpha}|^2}{4A}\Big) \int \mathrm{d}\Delta\,\boldsymbol{\alpha}' \times \Big[D\,C \exp\left(-\frac{1}{2} D\,C|\boldsymbol{l}|^2\right) \exp\left(-\frac{|\Delta\boldsymbol{\alpha} - \Delta\boldsymbol{\alpha}'|^2}{2F}\right) \exp\left[\mathrm{i} D\boldsymbol{l}\cdot(\Delta\boldsymbol{\alpha} - \Delta\boldsymbol{\alpha}')\right] - k\,I \exp\left(-\frac{1}{2} I|\boldsymbol{l}|^2\right)\Big] \exp\left(-\frac{|\Delta\boldsymbol{\alpha}'|^2}{4A}\right) \overline{W}^{l,m}(\Delta\boldsymbol{\alpha}'). \tag{19}$$

The purpose of a Fourier transformation is to convert a convolution in the variable $\boldsymbol{x}$ in (18) into a product in Fourier space. It is clear from (19) that the Fourier modes with respect to $\boldsymbol{x}$ are eigenfunctions of the developmental equation. This is expressed in (19) by the fact there is no mixing between components $\overline{W}^{l,m}$ and $\overline{W}^{l',m'}$; cf. [MS90] for a similar result.

In (19) we have introduced new variables

$$D = I/(C+I) \quad \text{and} \quad F = C + I. \tag{20}$$

According to (19) the quantity F can be interpreted as the effective width of the correlation function whose range of influence C has been increased by the intracortical interaction function.

To obtain a solution of equation (19) we will proceed in two steps. First, we consider the special case $k = 0$ which can be solved exactly. As a second step, we approximate eigenfunctions and eigenvalues for the general case $k \neq 0$ by expanding $\overline{W}$ in terms of the eigenfunctions for the case $k = 0$. Eigenfunctions and eigenvalues for $k = 0$ will be marked by a tilde above W and λ.

The eigenvalue problem for the special case $k = 0$ takes the form

$$\begin{aligned}\widetilde{\lambda}\widetilde{W}^{l,m}(\Delta\boldsymbol{\alpha})\exp(-\mathrm{i}D\boldsymbol{l}\cdot\Delta\boldsymbol{\alpha}) = 2\pi D\,C\exp\left(-\frac{1}{2}D\,C|\boldsymbol{l}|^2\right)\exp\left(-\frac{|\Delta\boldsymbol{\alpha}|^2}{4A}\right)\\ \times\int \mathrm{d}\Delta\boldsymbol{\alpha}'\exp\left(-\frac{|\Delta\boldsymbol{\alpha}-\Delta\boldsymbol{\alpha}'|^2}{2F}\right)\exp\left(-\frac{|\Delta\boldsymbol{\alpha}'|^2}{4A}\right)\\ \times\exp\left(-\mathrm{i}D\boldsymbol{l}\cdot\Delta\boldsymbol{\alpha}'\right)\widetilde{W}^{l,m}(\Delta\boldsymbol{\alpha}'). \qquad (21)\end{aligned}$$

Equation (21) factorizes into two parts that depend on Δx or Δy only. To simplify the notation we now introduce the one-dimensional integral operator $H_{\mathrm{int}}[.]$ defined by

$$H_{\mathrm{int}}[f](\xi) = \exp\left(-\frac{\xi^2}{4A}\right)\int \mathrm{d}\xi'\exp\left[-\frac{(\xi-\xi')^2}{2F}\right]\exp\left(-\frac{\xi'^2}{4A}\right)f(\xi'). \qquad (22)$$

Instead of (21) we consider the related one-dimensional problem

$$\lambda^* f(\xi) = H_{\mathrm{int}}[f](\xi). \qquad (23)$$

The eigenvalues and eigenfunctions of (23) are obtained easily since H_{int} *commutes* with the Hamiltonian of the quantum-mechanical harmonic oscillator

$$H_{\mathrm{osc}} = \frac{\mathrm{d}^2}{\mathrm{d}\xi^2} - \frac{\xi^2}{L^2}. \qquad (24)$$

More precisely, it can be demonstrated by partial integration that

$$[H_{\mathrm{int}}, H_{\mathrm{osc}}] = 0 \qquad (25)$$

holds for

$$L = 2A/\sqrt{1+4A/F}. \qquad (26)$$

The eigenfunctions of the harmonic oscillator are given by [Sch68]

$$f_n(\xi) = \frac{1}{\sqrt{2^n n!}}\left(\frac{1}{\pi L}\right)^{1/4}\mathrm{H}_n\left(\frac{\xi}{\sqrt{L}}\right)\exp\left(-\frac{\xi^2}{2L}\right) \qquad (27)$$

where H_n denotes a Hermite polynomial of degree n. The functions (27) form a complete orthonormal set [CH68]. Since the eigenvalues of the harmonic oscillator are non-degenerate, we conclude from the commutator (25) that the complete set of eigenfunctions of (23) is also given by (27).

Given the eigenfunctions, the eigenvalues can be determined in a straightforward manner. For the eigenvalues of (23) we make the ansatz

$$\lambda_n^* = \Lambda_0\, q^n \qquad \text{for} \quad n = 0, 1, 2, \ldots \qquad (28)$$

with parameters Λ_0 and q which are now derived. We use the generating function of the Hermite polynomials [CH68]

$$\exp\left(-t^2 + 2t\frac{\xi}{\sqrt{L}}\right) = \sum_{n=0}^{\infty}\frac{1}{n!}t^n\mathrm{H}_n\left(\frac{\xi}{\sqrt{L}}\right). \qquad (29)$$

If the ansatz (28) is true, the following equation must hold:

$$H_{\rm int}\left[\exp\left(-t^2+2t\frac{\xi'}{\sqrt{L}}\right)\exp\left(-\frac{\xi'^2}{2L}\right)\right](\xi)=\sum_{n=0}^{\infty}\frac{1}{n!}t^n\lambda_n^*\mathrm{H}_n\left(\frac{\xi}{\sqrt{L}}\right)\exp\left(-\frac{\xi^2}{2L}\right)$$
$$\stackrel{!}{=}\sum_{n=0}^{\infty}\frac{1}{n!}\Lambda_0(qt)^n\mathrm{H}_n\left(\frac{\xi}{\sqrt{L}}\right)\exp\left(-\frac{\xi^2}{2L}\right)$$
$$=\Lambda_0\exp\left(-(qt)^2+2tq\frac{\xi}{\sqrt{L}}\right)\exp\left(-\frac{\xi^2}{2L}\right) \tag{30}$$

where the generating function (29) has been used twice. On the other hand, we can directly perform an integration in the first part of (30) which yields an expression of the same form as in the final part of (30). Comparison of coefficients gives, after some algebra,

$$\Lambda_0=\sqrt{\pi\frac{F^2}{A}\left(1+2\frac{A}{F}-\sqrt{1+4\frac{A}{F}}\right)} \tag{31}$$

$$q=\frac{F}{2A}\left(1+2\frac{A}{F}-\sqrt{1+4\frac{A}{F}}\right). \tag{32}$$

Using (27) and (28), we obtain the eigenfunctions and eigenvalues of the two-dimensional eigenvalue problem (21), namely,

$$\widetilde{W}_{c,d}^{l,m}(\Delta\boldsymbol{\alpha})=\frac{1}{\sqrt{\pi\,2^{c+d}\,c!\,d!\,L}}\mathrm{H}_c\left(\frac{\Delta\alpha}{\sqrt{L}}\right)\mathrm{H}_d\left(\frac{\Delta\beta}{\sqrt{L}}\right)\exp\left(-\frac{|\Delta\boldsymbol{\alpha}|^2}{2L}\right)\exp(\mathrm{i}D\,\boldsymbol{l}\cdot\Delta\boldsymbol{\alpha}) \tag{33}$$

and

$$\widetilde{\lambda}_{c,d}^{l,m}=2\pi\,D\,C\,\Lambda_0^2\,q^{c+d}\,\exp\left(-\frac{1}{2}D\,C|\boldsymbol{l}|^2\right). \tag{34}$$

Thus we have found an analytic solution† in the case $k=0$. We now turn to the general problem.

In order to derive a solution for the problem with $k\neq 0$, we expand the eigenfunctions $\overline{W}^{l,m}$ in terms of the eigenfunctions $\widetilde{W}_{c.d}^{l,m}$ of the simplified problem,

$$\overline{W}^{l,m}(\Delta\boldsymbol{\alpha})=\sum_{c\geqslant 0,\,d\geqslant 0}a_{c,d}^{l,m}\widetilde{W}_{c,d}^{l,m}(\Delta\boldsymbol{\alpha}). \tag{35}$$

After this change of the basis, an eigenvalue problem in the coefficients $a_{c,d}^{l,m}$ has to be solved

$$\lambda\,a_{c,d}^{l,m}=\sum_{f\geqslant 0,g\geqslant 0}C_{c,d;\,f,g}^{l,m}\,a_{f,g}^{l,m} \tag{36}$$

with a correlation matrix of the form

$$C_{c,d;\,f,g}^{l,m}=\widetilde{\lambda}_{c,d}^{l,m}\,\delta_{c.f}\delta_{d,g}-k\,2\pi\,I\exp\left(-\frac{1}{2}I|\boldsymbol{l}|^2\right)\langle\widetilde{W}_{c,d}^{l,m}|H_1|\widetilde{W}_{f,g}^{l,m}\rangle. \tag{37}$$

† K D Miller (private communication) independently noted the following closely related, but less general, result for the case $k=0$. Let $W_C(\boldsymbol{\alpha})$ be an eigenfunction for an isolated postsynaptic cell (i.e. for the case where $I(\boldsymbol{x})$ is a delta function) in the presence of a Gaussian correlation function (i.e. (9) with $k=0$) and for some given arbor function. Then, for a network with the same correlation and arbor functions, but with the Gaussian intracortical interaction function (8), the eigenfunctions for the full network, parameterized by a real variable $\boldsymbol{m}$, are given by $\exp(\mathrm{i}\boldsymbol{m}\cdot\boldsymbol{x})\exp(\mathrm{i}g\boldsymbol{m}\cdot\Delta\boldsymbol{\alpha})W_{C+I}(\Delta\boldsymbol{\alpha})$, with corresponding eigenvalues $\lambda_m=\lambda_{C+I}\widetilde{I}(|\boldsymbol{m}|\sqrt{1-g})$, where the tilde means Fourier transform, $g\equiv I/(C+I)$, and $\widetilde{I}(|\boldsymbol{m}|\sqrt{1-g})=\widetilde{C}(|\boldsymbol{m}|\sqrt{g})$. This expression, along with the solutions for the case of an isolated postsynaptic cell found in [MM90], gives a subset of the solutions given by our general expression. See also Wimbauer *et al* [WGVH94]

The 'perturbation term' is

$$\langle \widetilde{W}^{l,m}_{c,d} | H_1 | \widetilde{W}^{l,m}_{f,g} \rangle = \frac{1}{\pi L} \frac{1}{\sqrt{2^{c+d+f+g}\, c!\, d!\, f!\, g!}} K^{l\,*}_c K^{m\,*}_d K^l_f K^m_g \tag{38}$$

where K^n_a with $a = c, d, f, g$ and $n = l, m$ stands for the one-dimensional integral

$$K^n_a = \int \mathrm{d}\Delta\alpha \, \exp\left(-\frac{\Delta\alpha^2}{4A}\right) \exp(\mathrm{i}D\, n\, \Delta\alpha) \exp\left(-\frac{\Delta\alpha^2}{2L}\right) \mathrm{H}_a\left(\frac{\Delta\alpha}{\sqrt{L}}\right). \tag{39}$$

The integral (39) can be computed and yields [Bat54]

$$K^n_a = p \sqrt{\frac{4\pi AL}{L+2A}} \left(\frac{2A-L}{2A+L}\right)^{(a/2)} \exp\left(-\frac{AL}{L+2A} D^2 n^2\right) \mathrm{H}_a\left(\sqrt{\frac{8A^2L}{4A^2-L^2}}\, D\, n\right) \tag{40}$$

with

$$p = \begin{cases} (-1)^{a/2} & \text{for } a \text{ even} \\ \mathrm{i}(-1)^{(a-1)/2} & \text{for } a \text{ odd.} \end{cases} \tag{41}$$

It is clear from (37) and (38) that the correlation matrix $C^{l,m}_{c,d;\,f,g}$ is Hermitian and, hence, that all eigenvalues are real.

Equation (36) has to be solved numerically. To this end we assume that the eigenfunctions corresponding to the leading eigenvalues can be approximated well by basis functions with a few zero crossings, that is, with indices c and d whose values are small. This seems to be a reasonable approximation since we expect that the leading eigenfunctions of (36) show only a few zero crossings too. Specifically, we restrict the expansion (35) to basis functions with $0 \leqslant c + d \leqslant 6$. In this case, for a fixed wave vector $\boldsymbol{l}$, a complex 28×28 matrix has to be diagonalized; in our case by the routine cheevx of the LAPACK package [ABB+95].

Since the matrix (37) is complex for $\boldsymbol{l} \neq 0$, the coefficients of the expansion (35) may assume complex values too. To simplify the notation, we introduce the following expressions

$$\begin{aligned} \mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha}) &= \sum_{c=0,\,d=0}^{\infty} \Re\left(a^{l,m}_{c,d}\right) \frac{1}{\sqrt{\pi\, 2^{c+d}\, c!\, d!\, L}} \mathrm{H}_c\left(\frac{\Delta\alpha}{\sqrt{L}}\right) \mathrm{H}_d\left(\frac{\Delta\beta}{\sqrt{L}}\right) \exp\left(-\frac{|\Delta\boldsymbol{\alpha}|^2}{2L}\right) \\ \mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha}) &= \sum_{c=0,\,d=0}^{\infty} \Im\left(a^{l,m}_{c,d}\right) \frac{1}{\sqrt{\pi\, 2^{c+d}\, c!\, d!\, L}} \mathrm{H}_c\left(\frac{\Delta\alpha}{\sqrt{L}}\right) \mathrm{H}_d\left(\frac{\Delta\beta}{\sqrt{L}}\right) \exp\left(-\frac{|\Delta\boldsymbol{\alpha}|^2}{2L}\right) \end{aligned} \tag{42}$$

where $\Re\left(a^{l,m}_{c,d}\right)$ and $\Im\left(a^{l,m}_{c,d}\right)$ denote the real and imaginary part of the expansion coefficients.

In figure 3 the quantities defined by (42) and corresponding to the five leading eigenvalues are displayed for a given set of parameters and for a fixed wave vector $\boldsymbol{l}$. As we mentioned before, each Fourier mode in the cortical coordinate $\boldsymbol{x}$ represents an eigenfunction of the developmental equation (18) in the form (19). Using (42) one obtains expressions for the full eigenfunctions in dependence of $\boldsymbol{x}$ and $\Delta\boldsymbol{\alpha}$, namely,

$$\begin{aligned} \overline{W}^{l,m}(\boldsymbol{x}, \Delta\boldsymbol{\alpha}) &= \Re\left\{\left[\mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha}) + \mathrm{i}\,\mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha})\right] \exp\left[\mathrm{i}\,\boldsymbol{l}(\boldsymbol{x} + D\,\Delta\boldsymbol{\alpha})\right]\right\} \\ &= \mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha}) \cos\left[\boldsymbol{l}(\boldsymbol{x} + D\,\Delta\boldsymbol{\alpha})\right] - \mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha}) \sin\left[\boldsymbol{l}(\boldsymbol{x} + D\,\Delta\boldsymbol{\alpha})\right] \end{aligned} \tag{43}$$

and similarly, with $\boldsymbol{l}\boldsymbol{x} = lx + my$ and $\boldsymbol{l}\Delta\boldsymbol{\alpha} = l\Delta\alpha + m\Delta\beta$,

$$\begin{aligned} \overline{W}^{l,m}(\boldsymbol{x}, \Delta\boldsymbol{\alpha}) &= \Im\left\{\left[\mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha}) + \mathrm{i}\,\mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha})\right] \exp\left[\mathrm{i}\,\boldsymbol{l}(\boldsymbol{x} + D\,\Delta\boldsymbol{\alpha})\right]\right\} \\ &= \mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha}) \sin\left[\boldsymbol{l}(\boldsymbol{x} + D\,\Delta\boldsymbol{\alpha})\right] + \mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha}) \cos\left[\boldsymbol{l}(\boldsymbol{x} + D\,\Delta\boldsymbol{\alpha})\right]. \end{aligned} \tag{44}$$

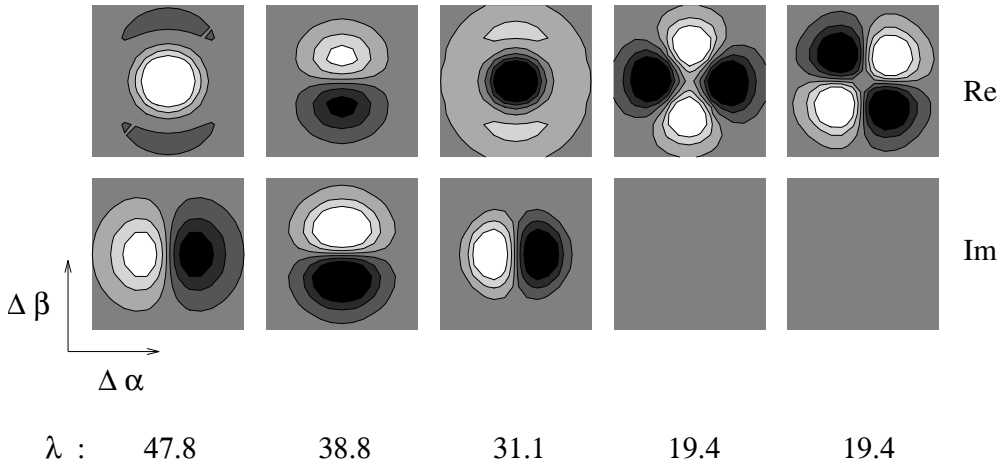

Figure 3. Eigenfunctions for a fixed wave vector $\boldsymbol{l} = (l, 0)$ with $l/(2\pi) = 0.078$; cf. figure 5. The quantities $\mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha})$ and $\mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha})$ as defined by (42) are displayed for the the five leading eigenvalues λ with parameter values $A = 10.25$, $r_c = \sqrt{C/A} = 0.65$, $r_i = \sqrt{I/A} = 0.30$, $k = 0.3$ and for a wave vector with the components $l/(2\pi) = 0.078$ and $m/(2\pi) = 0$. The expansion coefficients (35) have been derived numerically in terms of basis functions with $0 \leqslant c+d \leqslant 6$. The imaginary part of the two degenerate eigenfunctions in the lower right-hand corner vanishes.

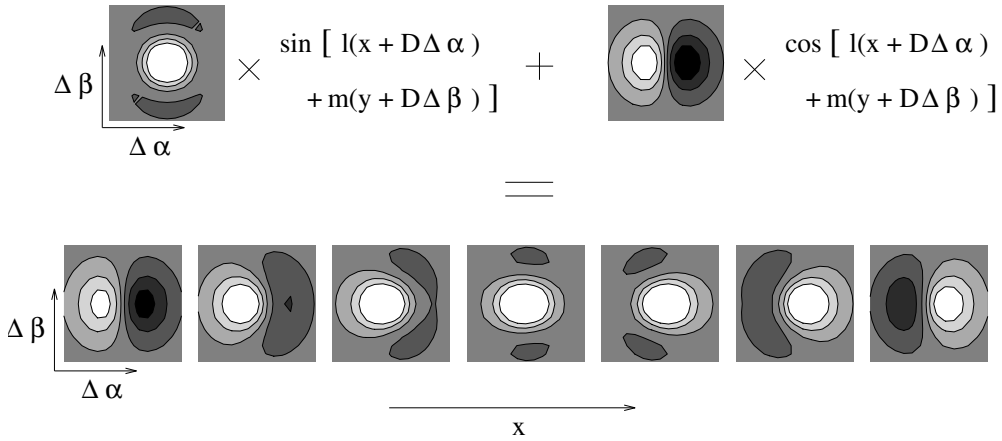

Figure 4. Full eigenfunction corresponding to the leading eigenvalue according to (44). Each square in the second row shows the distribution of the synapses for a fixed cortical cell at position $\boldsymbol{x}$. Two cortical cells are $|\Delta\boldsymbol{x}| = 1$ apart from each other, where the x axis shown in the figure is parallel to the wave vector $\boldsymbol{l}$. Along the direction of the wave vector $\boldsymbol{l}$, an oscillation of the receptive field properties evolves. Half an oscillation period is displayed approximately. The wave vector $\boldsymbol{l}$ and the parameters are the same as in figures 3 and 5.

The solutions (43) and (44) are identical except for a phase shift by $\pi/2$ and we can pick either one in order to discuss the nature of the solution. In figure 4, we show a two-dimensional plot of the function (44) for the leading eigenvalue. Due to the phase factor in the cortical coordinate $\boldsymbol{x}$, an oscillation of the receptive field properties with a wave vector $\boldsymbol{l}$

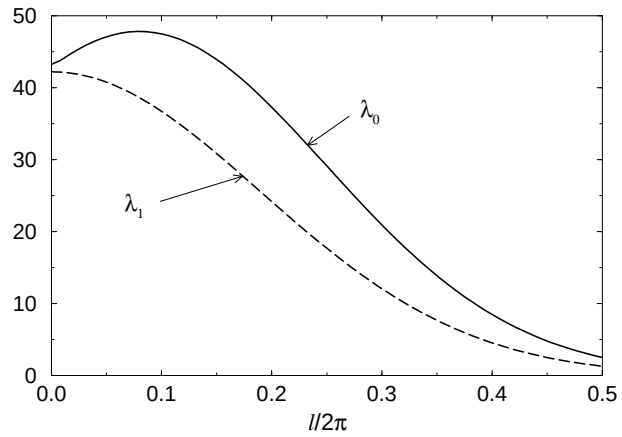

Figure 5. The first and second eigenvalues as a function of the wave vector $\boldsymbol{l}$. Due to the rotational symmetry of the eigenvalue problem we consider the dependence of the eigenvalues on one component of the wave vector (here l) only. The second component is $m = 0$. The parameter values are chosen as in figure 3. A maximum of the leading eigenvalue λ_0 is obtained for $l/(2\pi) = 0.078$. This wave vector has also been chosen in figures 3 and 4 so that for a common set of parameter values the dominating and, hence, relevant eigenfunctions are displayed in these figures.

can be observed. In particular, a transition between an even and an odd weight profile occurs that can be interpreted as a variation of the phase of the receptive field. The notion of a phase stems from the modelling of receptive fields by Gabor functions [JP87, Mar80], that is, by functions of the form

$$W(\boldsymbol{\alpha}) = \exp\left(-\frac{|\boldsymbol{\alpha}|^2}{2A}\right)\cos(\boldsymbol{k}\boldsymbol{\alpha} + \theta). \tag{45}$$

For a phase $\theta = 0$ a positive or a negative subregion can be found in the centre of the receptive field, whereas positive and negative subregions of the same size occupy the left and right half of the receptive field for a phase $\theta = \pi/2$. A variation of the phase of the receptive field from one cortical neuron to the next has also been observed experimentally [HW62]. We would like to stress again that positive and negative regions of the leading eigenfunction correspond to ON and OFF subfields of the receptive fields that emerge from the developmental process, as was explained in more detail at the beginning of this section.

So far we have considered the eigenfunction corresponding to the leading eigenvalue for a *fixed* wave vector only. We are interested, however, in the leading eigenfunction for an *arbitrary* wave vector. We therefore have to diagonalize the matrix (37) for different wave vectors and look for the leading eigenvalue. Since the eigenvalue problem is rotationally symmetric, we can restrict our search to one dimension. In figure 5 the first and the second eigenvalues that have been evaluated numerically have been plotted as a function of l. The second component of the wave vector has been taken to be $m = 0$. The parameter values for the numerical evaluation in figure 5 correspond to those used in figure 3.

The maximum eigenvalue is obtained for $l/(2\pi) = 0.078$, a wave vector that has also been used in figures 3 and 4. The basic structure of the eigenfunction corresponding to λ_0 is not sensitive to the wave vector index l. Whatever l, the real part $\mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha})$ as given by (42) is a sum of Hermite polynomials multiplied by a Gaussian whose index

sum $c + d$ is even. Similarly, the index sum $c + d$ of the non-vanishing coefficients of the imaginary part $\mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha})$ is always odd. The coefficients with indices c, d whose values are small ($a_{0,0}$ for $\mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha})$ and $a_{1,0}$, $a_{0,1}$ for $\mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha})$) dominate in the expansion of the leading eigenfunction. The result is that the function $\mathrm{Re}^{l,m}(\Delta\boldsymbol{\alpha})$ has a large central region whereas $\mathrm{Im}^{l,m}(\Delta\boldsymbol{\alpha})$ changes sign along the axis of the wave vector $\boldsymbol{l}$. Furthermore, since the coefficient for $c + d = 6$ of the leading eigenfunction is always less than 0.03 for all wave vectors $\boldsymbol{l}$, it seems justified that only basis functions up to $c + d = 6$ have been taken into account in the expansion (35).

It becomes obvious from figure 5 that within a range $0 \leqslant l/(2\pi) \leqslant 0.166$ the eigenvalue λ_0 exceeds the maximum of λ_1 at $l = 0$. The eigenfunctions that correspond to λ_0 within this band of wave vectors are all of the same form; they grow fastest during the development and, thus, determine the form of the emerging receptive fields.

It has been proposed [Mil94] that a sign change of the correlation function is crucial to the emergence of orientation-selective receptive fields. We can test this hypothesis by varying the constant k that determines the balance between positive and negative parts of the correlation function (9). It turns out that a critical value of k exists above which the dominating eigenfunctions take a form as that shown in figure 4 ($k \geqslant 0.1$ with the remaining parameters as in figure 3). If k is less than the critical value, the real part of the dominating eigenfunction is given by a Gaussian, whereas the imaginary part is approximately zero. The maximum eigenvalue is then at $\boldsymbol{l} = 0$. Hence the developmental process will not result in orientation-selective receptive fields, and no oscillation of the receptive field properties can be observed in this case.

If the wave vector corresponding to the leading eigenvalue is non-zero, the dominating eigenfunctions are always highly degenerate. The main reason for this degeneracy lies in the rotational symmetry of the eigenvalue problem. Because of this symmetry the eigenvalues for wave vectors with the same length but with different directions are all identical and the corresponding eigenfunctions grow at the same rate during development. The degeneracy has important consequences for the cortical map that emerges from the Hebbian learning process. In particular, the receptive field properties will not only oscillate in one direction from one cortical cell to the next; cf. figure 4. Rather, Fourier modes with wave vectors of different directions but the same fixed length are superimposed within the cortical map. This length corresponds to the maximum of the eigenvalue λ_0 as a function of l, as plotted in figure 5. Furthermore, as a consequence of the additional degeneracy of the eigenfunctions (43) and (44) these Fourier modes may be equipped with an additional and arbitrary phase factor.

4. Simulation of the developmental model

In this section the results of our mathematical analysis will be compared with simulations of a slightly more involved model. The arbor, the intracortical interaction, and the correlation function are chosen as introduced in section 2. Extensive simulations with slightly different arbor and correlation functions can be found in [Mil94].

Cortical simple cells and ON and OFF inputs from the LGN are modelled by three 32×32 grids where retinotopic positions on all grids correspond to each other. Each cortical cell receives inputs from both types of LGN cells. These input neurons lie within a circle centred at the retinotopic position of the cortical cell. The diameter of the circle is set to 13 neurons. In other words, the Gaussian that is chosen for the arbor function is cut at $|\boldsymbol{x} - \boldsymbol{\alpha}| = 6$. This is mainly done for computational reasons, in particular, to save computer time, and does not affect the simulation results.

In order to be consistent with the simulations of [Mil94], the developmental equation (5) is complemented by a constraint term that ensures that the sum of the synaptic weights received by one cortical cell is kept fixed. In this way a competitive process between afferent axons is modelled.

The full learning equation then takes the form

$$\frac{\mathrm{d}W^c(\boldsymbol{x},\boldsymbol{\alpha},t)}{\mathrm{d}t} = \eta A(\boldsymbol{x}-\boldsymbol{\alpha})\sum_{\boldsymbol{x}'} I(\boldsymbol{x},\boldsymbol{x}') \sum_{c'=\mathrm{ON,OFF}} \sum_{\boldsymbol{\alpha}'} C^{c,c'}(\boldsymbol{\alpha},\boldsymbol{\alpha}')W^{c'}(\boldsymbol{x}',\boldsymbol{\alpha}',t) - \sigma(\boldsymbol{x},t)A(\boldsymbol{x}-\boldsymbol{\alpha}) \tag{46}$$

where

$$\sigma(\boldsymbol{x},t) = \frac{1}{4\sum_{\boldsymbol{\alpha}'} A(\boldsymbol{x}-\boldsymbol{\alpha}')}\left[\sum_{c=\mathrm{ON,OFF}}\sum_{\boldsymbol{\alpha}'} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{\mathrm{un}} W^c(\boldsymbol{x},\boldsymbol{\alpha}',t)\right] \tag{47}$$

and

$$0 \leqslant W^c(\boldsymbol{x},\boldsymbol{\alpha},t) \leqslant W_{\max} A(\boldsymbol{x}-\boldsymbol{\alpha}). \tag{48}$$

The expression $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{\mathrm{un}} W^c(\boldsymbol{x},\boldsymbol{\alpha}',t)$ denotes the unconstrained developmental equation as given by the first line of (46). The effects of the constraint on the dynamics of the learning equation have been discussed in great detail in [MM94]. It turns out that the basic structure of the emerging receptive fields is still determined by the leading eigenfunction of the unconstrained learning equation, but that the receptive fields are sharpened as compared to the unconstrained case. Furthermore, as in the unconstrained case, all weights saturate either at the upper or at the lower bound given by (48).

The simulation algorithm proceeds along the same steps as the one used in [Mil94]. Both types of synapse have been assigned random initial values uniformly distributed over $[1 \pm s_{\mathrm{noise}} A(\boldsymbol{x}-\boldsymbol{\alpha}))]$ with $s_{\mathrm{noise}} = 0.2$. During each time step of the simulation the change in synaptic strength is calculated according to (46), (47) and (48). The growth constant η is adjusted in such a way that the standard deviation for the change in synaptic strength becomes 0.01 for the first time step. The simulation is stopped if more than 90% of the synapses have reached their upper or lower bounds.

Figure 6 shows the result of a simulation run with the parameter values $A = 10.25$, $r_c = \sqrt{C/A} = 0.45$, $r_i = \sqrt{I/A} = 0.2$, $k = 0.3$ and $\epsilon = 1$; cf. (7), (8) and (9). This is the same set of parameters as used in figures 3, 4 and 5. Thus the simulation results can be compared directly with the analytical predictions. In figure 6 an 8×8 subregion of the cortical grid is displayed. Each small square corresponds to the receptive field of one cortical cell.

As was mentioned above, the fact that *either* ON *or* OFF synapses exist between a cortical cell at a position $\boldsymbol{x}$ and a cell at position $\boldsymbol{\alpha}$ in the LGN is due to the form of the dominating eigenfunction for the γ-part of the eigenvalue problem (cf. (15)). This is also true for the simulation run considered here. In figure 6 white and black dots denote those ON and OFF synapses that have reached the upper bound.

We now turn to the question of how far the structure of the cortical map that has been derived in our analysis of the eigenvalue problem in the last section coincides with the map that emerges from a simulation of the developmental equation. To answer this question we first consider the column on the left-hand side of figure 6. As can be seen clearly, the phase of the receptive field varies from cell to cell. A comparison with figure 4 shows that a similar periodic phase change occurs for an eigenfunction with a wave vector $\boldsymbol{l} \neq 0$. In both cases a transition between a receptive field with an ON and an OFF subregion of equal

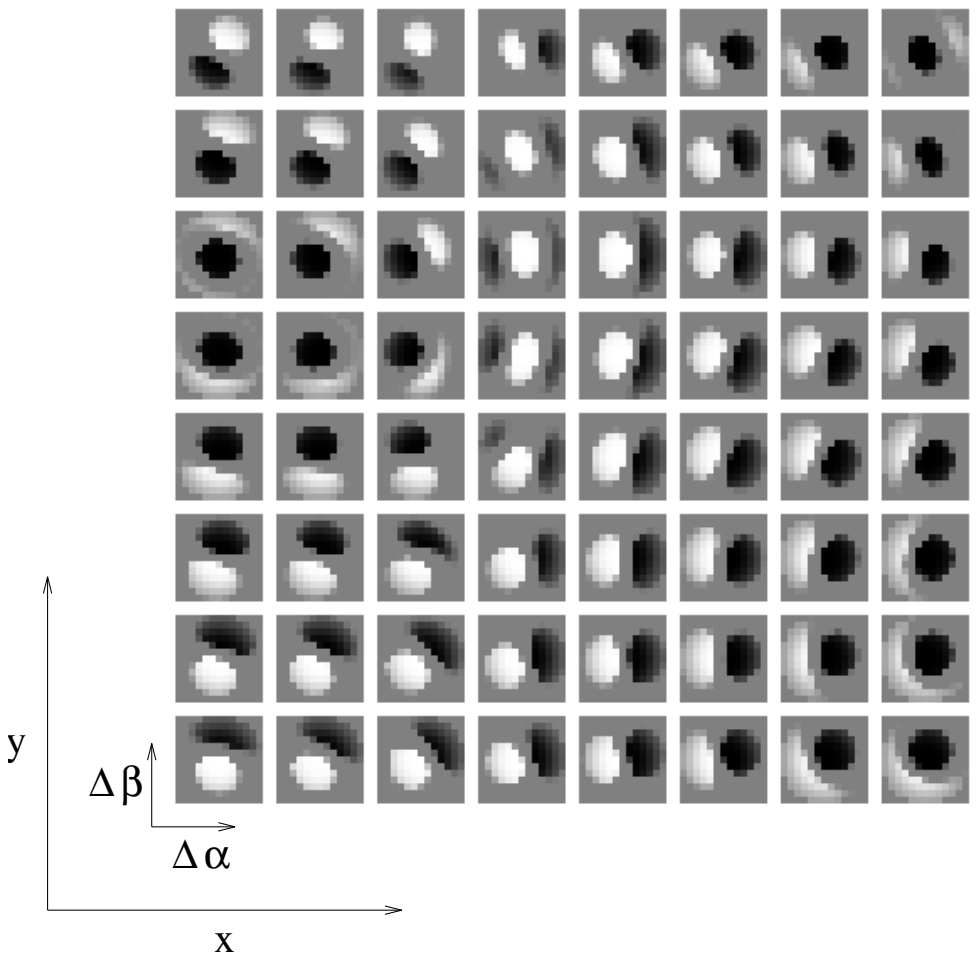

Figure 6. Receptive fields. An 8×8 detail from a grid of 32×32 cortical neurons is displayed for a simulation run with the parameter values $A = 10.25$, $r_c = 0.45$, $r_i = 0.2$, $k = 0.3$ and $\epsilon = 1$. Each small square corresponds to the $W^{\mathrm{ON}}(\boldsymbol{\alpha}, \boldsymbol{x}) - W^{\mathrm{OFF}}(\boldsymbol{\alpha}, \boldsymbol{x})$ distribution for a fixed cortical cell at position $\boldsymbol{x}$ plotted as a function of $\Delta\alpha$, $\Delta\beta$. The cortical cells (and, hence, the set of small squares) are arranged in an x, y coordinate system. White points correspond to ON weights (positive values) and black points to OFF weights (negative values). The question of how the weight distribution displayed here relates to the structure of the dominating eigenfunction is discussed in the text.

size (phase $\pi/2$) and a receptive field with a central ON or OFF region (phase 0) takes place.

As predicted from our analysis of the eigenvalue problem, oscillations of the receptive field properties with different orientations are superimposed in figure 6. In this way one gains the impression that the preferred orientation rotates from one cortical cell to the next and that different preferred orientations are grouped around certain points, so-called pinwheels. The above properties of the cortical map in our simulation model have also been observed experimentally by the technique of optical imaging [BS86, BG91, BG93].

5. Discussion

We have performed an analysis of a developmental model proposed in [Mil94] that describes the emergence of orientation-selective receptive fields of simple cells. To this end we have calculated the eigenfunctions and eigenvalues associated with the linear differential equation that models the Hebbian mechanism of the developmental process. The eigenfunctions that correspond to the leading eigenvalues grow fastest during the development and, hence, determine the receptive fields and the map structure of cortical cells.

There are three central results that can be derived from the above eigenvalue analysis and that can also be observed in simulations of the model. First, as hypothesized in [Mil94], zero crossings of the correlation function are crucial for the emergence of orientation-selective receptive fields. We have investigated the effect of zero crossings in section 3 by a variation of the parameter k of the correlation function. Second, since all Fourier modes in the cortical coordinate $\boldsymbol{u}$ are eigenfunctions of the learning equation, the receptive field properties, in particular, the phase of the receptive field, varies periodically from one cortical cell to the next. Since there is a broad maximum of nearly optimal wave vectors we expect local but no long-range periodicity. Third, because of the rotational symmetry of the problem, Fourier modes with wave vectors of a certain length but an arbitrary direction are superimposed in the cortical map.

There is a large variety of models describing cortical orientation maps; for an overview see [EOS95]. Central to most of these models is a spatial wavelength that characterizes the periodic variation of the receptive field properties within the cortical map [WPG94]. This wavelength is either due to an instability of a developmental learning equation with respect to certain Fourier modes [Swi82, OBS92] or may be introduced by explicit wavelength selection [RS90, NW93]. In the latter case one tries to reproduce the cortical map of a certain receptive field property by simply filtering white noise with a spatially isotropic bandpass.

Our analysis of the eigenvalue problem provides a link between formal models *describing* effective, low-dimensional feature maps such as the orientation map and the more microscopic picture of the development of synaptic wiring patterns. In particular, the processes that take place during Hebbian development lead to orientation-selective receptive fields with phase and orientation that vary along the cortical coordinates. In our developmental model, the random initial conditions are projected onto the eigenfunctions of the learning equation that grow according to the size of the respective eigenvalues. As we have discussed above, the distribution of the eigenvalues in dependence upon the wave vector $\boldsymbol{l}$ has a fairly broad maximum around an optimal length $|\boldsymbol{l}|$ of the wave vector and is rotationally symmetric. The developmental process has therefore a similar effect as filtering with a spatially isotropic bandpass and the function $\lambda_0(\boldsymbol{l})$ as displayed in figure 5 can be interpreted as the corresponding filter profile. The location of the maximum of the wavelength filter depends on the spatial correlations in the input and the typical length of cortical connectivity.

We emphasize that, in the present correlation-based learning model, we did not include inhibitory interactions between cortical neurons of, for example, different preferred orientation. This shows that competitive processes are not necessary to achieve a map-like organization of orientation tuning. As discussed by Wolf *et al* [WPG94], differences between various models of cortical pattern formation may well show up only during the saturation phase where nonlinearities become important. Our analysis is based on a linear learning equation and therefore does not give an adequate picture of saturation. To study nonlinear effects in cortical map formation, more elaborate mathematical techniques have to be used.

Acknowledgments

The authors thank K D Miller for providing the program that was used for the simulation in section 4 and for many useful discussions and suggestions. SW was supported by a grant from the Studienstiftung des deutschen Volkes.

References

[ABB+95] Anderson E, Bai Z, Bischof C, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Ostrouchov S and Sorensen D 1995 *LAPACK User's Guide* (Philadelphia, PA: SIAM)

[Bat54] Bateman H 1954 *Tables of Integral Transforms* vol II (New York: McGraw-Hill)

[BG91] Bonhoeffer T and Grinvald A 1991 Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns *Nature* **353** 429–31

[BG93] Bonhoeffer T and Grinvald A 1993 The layout of iso-orientation domains in area 18 of cat visual cortex: optical imaging reveals a pinwheel-like organization *J. Neurosci.* **13** 4157–80

[BS86] Blasdel G G and Salama G 1986 Voltage sensitive dyes reveal a modular organization in monkey striate cortex *Nature* **321** 579–85

[CH68] Courant R and Hilbert D 1968 *Methoden der mathematischen Physik I* 3rd edn (Berlin: Springer)

[CM95] Crair M C and Malenka R C 1995 A critical period for long-term potentiation at thalamocortical synapses *Nature* **375** 325–8

[CPCD90] Constantine-Paton M, Cline H T and Debski E 1990 Patterned activity, synaptic convergence, and the NMDA receptor in developing visual pathways *Ann. Rev. Neurosci.* **13** 129–54

[CZS91] Chapman B, Zahs K R and Stryker M P 1991 Relation of cortical cell orientation selectivity to alignment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex *J. Neurosci.* **11** 1347–58

[EOS95] Erwin E, Obermayer K and Schulten K 1995 Models of orientation and ocular dominance columns in the visual cortex: a critical comparison *Neural Comput.* **7** 425–68

[Fer87] Ferster D 1987 Origin of orientation-selective EPSPs in simple cells of cat visual cortex *J. Neurosci.* **7** 1780–91

[Fer88] Ferster D 1988 Spatially opponent excitation and inhibition in simple cells of the cat visual cortex *J. Neurosci.* **8** 1172–80

[Fer94] Ferster D 1994 Linearity of synaptic interactions in the assembly of receptive fields in cat visual cortex *Curr. Opin. Neurobiol.* **4** 563–8

[FL83] Ferster D and Lindström S 1983 An intracellular analysis of geniculocortical connectivity in area 17 of the cat *J. Physiol.* **342** 181–215

[Heb49] Hebb D O 1949 *The Organization of Behavior* (New York: Wiley)

[HW59] Hubel D H and Wiesel T N 1959 Receptive fields of single neurons in the cat's striate cortex *J. Physiol.* **148** 574–91

[HW61] Hubel D H and Wiesel T N 1961 Integrative action in the cat's lateral geniculate body *J. Physiol.* **155** 385–98

[HW62] Hubel D H and Wiesel T N 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex *J. Physiol.* **160** 106–54

[JP87] Jones J P and Palmer L A 1987 An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex *J. Neurophysiol.* **58** 1233–58

[Lin86a] Linsker R 1986 From basic network principles to neural architecture: Emergence of orientation-selective cells *Proc. Natl Acad. Sci. USA* **83** 8390–4

[Lin86b] Linsker R 1986 From basic network principles to neural architecture: Emergence of orientation columns *Proc. Natl Acad. Sci. USA* **83** 8779–83

[Lin86c] Linsker R 1986 From basic network principles to neural architecture: Emergence of spatial-opponent cells *Proc. Natl Acad. Sci. USA* **83** 7508–12

[Mar80] Marcelja S 1980 Mathematical description of the responses of simple cortical cells *J. Opt. Soc. Am.* **70** 1297–300

[Mil94] Miller K D 1994 A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs *J. Neurosci.* **14** 409–41

[MLB95] Meister M, Lagnado L and Baylor D A 1995 Concerted signaling by retinal ganglion cells *Science* **270** 1207–10

[MM90] MacKay D J C and Miller K D 1990 Analysis of Linsker's application of Hebbian rules to linear networks *Network: Comput. Neural Syst.* **1** 257–97

[MM94] Miller K D and MacKay D J C 1994 The role of constraints in Hebbian learning *Neural Comput.* **6** 100–26

[MS90] Miller K D and Stryker M P 1990 Ocular dominance column formation: mechanism and models *Connectionist Modeling and Brain Function: the Developing Interface* ed S J Hanson and C R Olson (Cambridge, MA: MIT Press/Bradford) pp 255–350

[NW93] Niebur E and Wörgötter F 1993 Orientation column structure from first principles *Computation and Neural Systems 92* (Dordrecht: Kluwer) sect 62

[OBS92] Obermayer K, Blasdel G G and Schulten K 1992 A statistical mechanical analysis of self-organization and pattern formation during the development of visual maps *Phys. Rev.* A **45** 7568–89

[RA95] Reid R C and Alonso J 1995 Specificity of monosynaptic connections from thalamus to visual cortex *Nature* **378** 281–4

[RAW94] Reid R C, Alonso J and Wiesel T N 1994 Rules determining the monosynaptic connections between LGN cells and simple cells in cat visual cortex *Soc. Neurosci. Abstr.* **20** 1476

[RS90] Rojer A S and Schwartz E L 1990 Cat and monkey cortical columnar patterns modeled by bandpass-filtered white noise *Biol. Cybern.* **62** 381–91

[SV85] Sanders J A and Verhulst F 1985 *Averaging Methods in Nonlinear Dynamical Systems* (Berlin: Springer)

[Sch68] Schiff L I 1968 *Quantum Mechanics* (New York: McGraw-Hill) 3rd edn

[SS82] Sherman S M and Spear P D 1982 Organization of visual pathways in normal and visually deprived cats. *Physiol. Rev.* **62** 738–55

[Swi82] Swindale N V 1982 A model for the formation of orientation columns *Proc. R. Soc. Lond.* B **215** 211–30

[vdM73] von der Malsburg C 1973 Self-organization of orientation-selective cells in the striate cortex *Kybernetik* **14** 85–100

[WGVH94] Wimbauer S, Gerstner W and van Hemmen J L 1994 Emergence of spatiotemporal receptive fields and its application to motion detection. *Biol. Cybern.* **72** 81–92

[WPG94] Wolf F, Pawelzik K and Geisel T 1994 Emergence of long-range order in maps of orientation preference *Proc. ICANN 94* vol 1, ed M Marinaro and P G Morasso (Berlin: Springer)

[YKC89] Yuille A L, Kammen D M and Cohen D S 1989 Quadrature and the development of orientation-selective cortical cells by Hebb rules *Biol. Cybern.* **61** 183–94