# Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps

K. Obermayer
*Beckman Institute and Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 and Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*

G. G. Blasdel
*Harvard Medical School, Harvard University, Boston, Massachusetts 02115*

K. Schulten
*Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801 and Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801*
(Received 26 December 1991)

We report a detailed analytical and numerical model study of pattern formation during the development of visual maps, namely, the formation of topographic maps and orientation and ocular dominance columns in the striate cortex. Pattern formation is described by a stimulus-driven Markovian process, the self-organizing feature map. This algorithm generates topologically correct maps between a space of (visual) input signals and an array of formal "neurons," which in our model represents the cortex. We define order parameters that are a function of the set of visual stimuli an animal perceives, and we demonstrate that the formation of orientation and ocular dominance columns is the result of a global instability of the retinoptic projection above a critical value of these order parameters. We characterize the spatial structure of the emerging patterns by power spectra, correlation functions, and Gabor transforms, and we compare model predictions with experimental data obtained from the striate cortex of the macaque monkey with optical imaging. Above the critical value of the order parameters the model predicts a lateral segregation of the striate cortex into (i) binocular regions with linear changes in orientation preference, where iso-orientation slabs run perpendicular to the ocular dominance bands, and (ii) monocular regions with low orientation specificity, which contain the singularities of the orientation map. Some of these predictions have already been verified by experiments.

PACS number(s): 42.66.−p, 87.10.+e, 05.40.+j, 89.70.+c

## I. INTRODUCTION

Pattern formation by self-organization is a common phenomenon in brain development. The enormous number of neurons and connections makes it impossible for an organism to completely specify neural connectivity patterns within its genome. Instead, an organism seems to store procedural information—sets of rules that define developmental processes organizing the brain. In contrast to self-organization in "physical" systems, structure in "biological" systems does not necessarily emerge spontaneously from initially homogeneous states. Information is often extracted from the outside world and imprinted into an appropriate spatial pattern. Self-organizing processes, which are characterized by an information flow from the outside world into the biological system are generally called "unsupervised learning." These processes will be the focus of this paper.

The development of the mammalian visual system is one of the best studied areas of self-organization in the brain. After maturation, nerve fibers originating from the ganglion cells in the retina project via the thalamus to an area of the brain called the primary visual cortex. (In primates there exists only one projection area per brain hemisphere, area 17 or the striate cortex.) As the only gateway to higher cortical areas, the primary visual cortex has to process all possible elements natural scenes may contain, including contours, texture, color, shading, disparities, and motion. This leads to a lateral segregation (and specialization) of the striate cortex into regions which are concerned with the detection and processing of certain image elements or combinations thereof, and gives rise to characteristic spatial patterns called *cortical maps* [1–7].

It is now generally accepted that cortical maps are a product of self-organization and are to some extent "learned" by visual experience [8,9]. Although a crude structure is set up in the absence of external input, appropriate visual stimulation is necessary for the visual cortex to obtain its full information-processing capabilities. Experiments have shown that maturation continues after birth and that the final organization depends on the statistical properties of the perceived patterns. However, the degree to which cortical maps are the result of a learning process as well as the origin of the observed spatial structure are still a matter of controversy.

Several activity-dependent mechanisms have been proposed to shape the structure of cortical maps [10–24]. These mechanisms can be characterized by the basic process underlying pattern formation and fall into three

 

classes.

(i) Pattern formation from templates [16,20]. The individual cells gain their properties by a process which amplifies weak preferences introduced by the geometrical arrangements of receptor positions in the retina or by initial synaptic weights between neurons. There is no coupling necessary between neighboring cells in the cortex. Neighboring cells have correlated response properties, because they receive their input from overlapping populations of cells in the preceding layer, e.g., the retina. (These models assume a crude topographic projection between retina and cortex.)

(ii) Pattern formation by spontaneous symmetry breaking [15,17–19]. The individual cells gain their properties by spontaneous symmetry breaking in an energy function that governs the maturation of cortical cells. The symmetry breaking occurs for each cell separately. Correlated spatial patterns emerge due to weak coupling between neighboring cortical cells.

(iii) Pattern formation by unsupervised learning [10–14,21,22]. These models are based on competitive learning networks. The cells within cortex are strongly coupled and the coupling is essential for the individual cells to gain their specific properties as well as for the spatial pattern to emerge.

Unfortunately, most of the numerical simulations of these and other models have not been elaborated enough to be experimentally testable. This led to the unsatisfactory situation in which it has not yet been possible to evaluate the basic mechanisms that have been suggested to shape the spatial structure of cortical maps.

In the following we want to report on an analytical and numerical study of a simple pattern formation process, the self-organizing-feature-map (SOFM) algorithm of Kohonen [13,14]. The model provides an intermediate-level description of neural pattern formation which is based on a small set of abstract but biologically plausible assumptions [25]. The model falls into the third category and explains the development of cortical maps as unsupervised learning and the spatial structure of these patterns as a dimension-reducing map between a high-dimensional feature space and a two-dimensional cortical surface [25,21,22]. In a previous numerical study [22], we investigated the joint formation of a retinotopic projection and an orientation-column system, and we have demonstrated that the resulting maps closely resemble the maps that are observed in the primary visual cortex of a macaque monkey and a cat. In previous analytical studies [26,27] it has been demonstrated in a simplified case that ocular-dominance-column-like patterns can emerge as the result of a global instability of a retinotopic projection under certain stimulus conditions.

In the present contribution we will provide a theoretical framework to study the combined formation of retinotopic maps and orientation and ocular dominance columns and extend our previous studies to analyze a dimension-reducing map between a five-dimensional feature space (two dimensions for the position of a receptive field in retinotopic space, one each for orientation preference, orientation specificity, and ocular dominance) and a two-dimensional network that represents cortex. We will present a detailed analysis of the spatial patterns and compare model predictions quantitatively with experimental data.

The paper is organized into five parts. Section II will provide the biological background and Sec. III gives a short description of the algorithm and its mathematical properties. Following the ideas of [26] we will provide a mathematical analysis of map formation in Sec. IV and we will demonstrate that the orientation and ocular-dominance-column systems emerge as the result of a global instability of the topographic projection. The last section is concerned with a detailed analysis of model predictions with regard to the spatial structure of cortical maps. It contains a comparison with experimental data obtained from the macaque striate cortex using optical imaging [3,5,6].

## II. CORTICAL MAPS IN THE PRIMARY VISUAL CORTEX OF THE MACAQUE

In this section we will provide the biological background and briefly describe the spatial structure of the retinotopic map, the orientation and ocular-dominance-column systems in the striate cortex of the macaque *(Macaca nemestrina).* Since the focus of this paper is on a mathematical analysis of map formation and map structure we will provide only the information necessary to motivate the theoretical approach and to allow a comparison of model predictions with experimental data. If one is interested in details of data acquisition, data postprocessing, and a comparison of the displayed data with results obtained with other techniques, one may wish to refer to [3,5,6].

The cerebral cortex of primates is a thin layer of cells, which covers most of the surface of the brain. Although the cortex can be histologically divided into six layers, the response properties of the neurons seems to vary mainly laterally only and to be roughly identical in all layers. Accordingly, columnar models of cortical organization have been suggested [1] which describe the functional architecture of the primary visual cortex by a two-dimensional spatial pattern of neural response properties across the cortical surface. This description, although a simplification, seems to capture the essential aspects of cortical organization. Therefore we will adopt it for the following investigations.

The nerve fibers originating from the ganglion cells in the retina project via the thalamus to the primary visual cortex in a topographic manner, such that nearby locations in the retina project onto neighboring locations in the cortex. This type of connectivity pattern is called a *topographic map.* On a spatial scale above several millimeters (in the cortex) the magnification factor of the map, the *cortical magnification factor,* varies with cortical location. On a spatial scale below a few millimeters, however, these variations of the cortical magnification factor can be neglected in first approximation except for regions which represent the extreme periphery of the visual field.

The projection from the thalamus to the cortex is slightly divergent, so that each neuron in the cortex receives its input from a small area of the retina. However,

only stimuli that are presented to the center of this region, the *receptive field,* elicit a measurable response. The sharpening of the tuning to the position of localized stimuli in the visual field is, at least in part, the result of lateral interactions between neurons. Receptive fields are the result of the collective dynamics of a large number of cells, and we will, therefore, adopt a mesoscopic description and assign receptive fields not to single cells but to small volumes of the cortex ("columns"), which we will label by their location $\mathbf{r}$ on the cortical surface. In the following we will denote the center of the receptive field of a column $\mathbf{r}$ in visual space by the coordinates $w_{\mathbf{r}1}$ and $w_{\mathbf{r}2}$.

Other prominent features of cells in striate cortex are *orientation selectivity* and *ocular dominance.* Orientation selectivity refers to the pronounced tendency of most cortical cells to respond maximally to oriented visual stimuli, light bars, or edges, within a finite range of orientations. Orientation selectivity of a neuron is described by its response as a function of stimulus orientation, the *tuning curve.* Since tuning curves typically exhibit one peak, they can be reasonably well described by two parameters: preferred orientation, which denotes the stimulus orientation yielding the strongest response, and orientation tuning strength, the difference in response between a stimulus presented in preferred orientation and the corresponding "orthogonal" stimulus. Following Swindale [28] we describe orientation selectivity by a two-dimensional vector $\mathbf{w}_{\mathbf{r}}=(w_{\mathbf{r}3},w_{\mathbf{r}4})$ at each cortical location $\mathbf{r}$. Length and half the angle (the orientation of an edge is $\pi$ periodic) of these vectors denote orientation tuning strength and preferred orientation, respectively. Again, we assign receptive field properties like orientation selectivity not to particular neurons but to cortical locations.

The ganglion cells, which are located in corresponding regions of both eyes, project to the same area of cortex. Within the cortex there are many neurons which respond to stimuli presented to the ipsilateral as well as to the contralateral eye. The degree of response, however, might be different for stimuli presented to different eyes. This effect is called ocular dominance. We will call a region monocular if it responds to stimuli presented to one particular eye only, and binocular if it responds equally well to stimuli presented to either eye. In our model ocular dominance is described by a real-valued function $w_{\mathbf{r}5}$, which denotes the difference in response to stimuli presented to the left and the right eye as a function of cortical location.

The spatial patterns of orientation preference, orientation tuning strength, and ocular dominance in area 17 of the macaque have been studied extensively with optical imaging techniques [3,5,6]. Figure 1(a) shows the spatial pattern of orientation preference and orientation tuning strength recorded from an adult macaque. The image shows a $4\times6$ mm$^2$ patch of striate cortex which is located near the border to area 18. The orientation preference is coded by color, where the sequence red $\rightarrow$ yellow $\rightarrow$ green $\rightarrow$ blue $\rightarrow$ purple $\rightarrow$ red describes a complete cycle of preferred orientation from 0° over 90° to 180°. Brightness indicates the degree of orientation tuning. Dark regions indicate that the response is independent of stimulus orientation, bright regions indicate sharp tuning to the preferred orientation. On a spatial scale below $\approx1$ mm the orientation preferences are correlated and the orientation map is organized into elongated "iso-orientation" patches [slabs, arrow 2 of Fig. 1(a)], which start and end at singularities [arrow 1 of Fig. 1(a)] around which orientation changes by $\pm180°$. On a spatial scale above $\approx1$ mm the pattern is repetitive along every direction in the cortex with approximately the same spatial frequency.

Figure 1(b) shows the spatial distribution of ocular dominance for the same region. Ocular dominance is indicated by brightness, where dark and bright areas denote regions driven preferably by the ipsilateral and contralateral eyes, respectively. The pattern of ocular dominance is more regular than the pattern of orientation preference. Regions of similar eye dominance form parallel bands, which run perpendicular to the border between area 17 and 18 [to the right in Fig. 1(b)], sometimes branching and terminating. The fairly regular pattern is typical for the border region to area 18; it becomes more irregular further inside area 17. Table I summarizes the typical features of the spatial pattern of ocular dominance and orientation columns in the macaque striate cortex. These features should be reproduced by theoretical models.

## III. DESCRIPTION OF THE NEURAL NETWORK MODEL

### A. The dimension-reducing map problem

We consider a five-dimensional feature space $V$ which is spanned by quantities describing the most prominent receptive field properties of cortical cells (Fig. 2). The five coordinate axes correspond to the position of a receptive field in retinotopic space ($v_1$ and $v_2$), to orientation preferences and specificity ($v_3$ and $v_4$), and to ocular dominance ($v_5$), and are represented by the unit vectors $\mathbf{v}_1-\mathbf{v}_5$. (The dimensions $v_3$ and $v_4$ corresponding to orientation selectivity are suppressed in Fig. 2 to allow for visualization.) If all combinations of these properties are represented in striate cortex, each point in this five-dimensional feature space is mapped onto one point in the two-dimensional *cortical surface A.*

### B. The self-organizing feature map algorithm

In order to solve the mapping problem we employ the feature map (SOFM) algorithm of Kohonen [13,14,29] an abstract but biologically plausible pattern-formation process. The cortical surface is divided into small patches which are considered as "units" of a two-dimensional square lattice (Fig. 3) and which are thought to contain cells with common response properties. Each unit is labeled by its discrete position $\mathbf{r}$ in the lattice.

The averaged receptive field properties at the location $\mathbf{r}$ are characterized by a feature vector $\mathbf{w}_{\mathbf{r}}$, whose components $w_{\mathbf{r}k}$ denote the receptive field properties: position of the receptive field centers in visual space $(x_{\mathbf{r}},y_{\mathbf{r}})$, preferred orientation $(\phi_{\mathbf{r}})$, orientation tuning strength

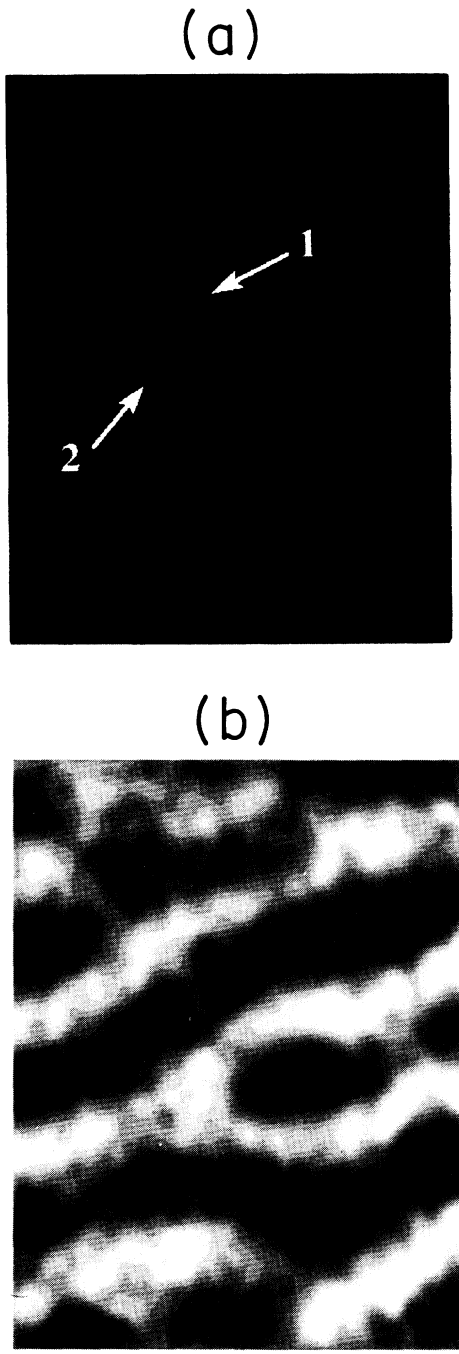


FIG. 1. Spatial pattern of (a) orientation selectivity and (b) ocular dominance in the striate cortex of an adult monkey *(Macaca nemestrina)*. The images were obtained by optical imaging [5,6] and show a $4\times6$ mm$^2$ patch of cortex which is located near the border to area 18. Orientation preference is coded by color, where the sequence red $\rightarrow$ yellow $\rightarrow$ green $\rightarrow$ blue $\rightarrow$ purple $\rightarrow$ red describes a complete cycle of preferred orientation for 0° over 90° to 180°. Orientation tuning strength is normalized and coded by brightness, where bright areas indicate regions with a specific response. Ocular dominance is also coded by brightness, where black and white indicate preference for the contralateral and the ipsilateral eye, respectively. The arrows in (a) mark the two prominent elements of the orientation column system, the *singularities* (arrow 1) and regions where orientation preferences are organized in parallel slabs (arrow 2).

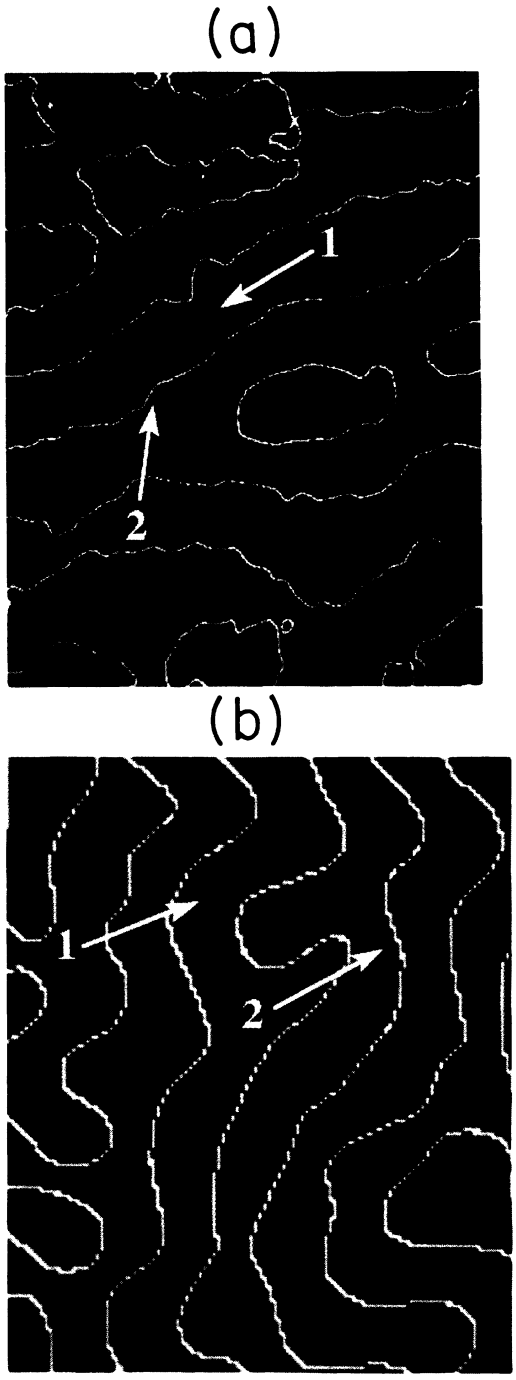


FIG. 11. Correlation between the spatial pattern of orientation selectivity and ocular dominance. Color and brightness code for orientation preference and orientation tuning strength, respectively, the white lines indicate the borders of the ocular dominance bands. (a) Experimental data (taken from [5,6]) obtained from an adult macaque. The image shows a $4\times6$ mm$^2$ patch of cortex, which is located near the border to area 18. (b) Overlay of the orientation and ocular dominance patterns shown in Fig. 10.

TABLE I. Properties of orientation and ocular dominance columns in the monkey striate cortex. Experimental data were taken from [3,5,6] and unpublished results.

| | |
|---|---|
| 1 | The maps of orientation selectivity and ocular dominance are highly repetitive across the retinotopic map (feature hierarchy). |
| 2 | Orientation preferences change continuously as a function of cortical location except at isolated points (singularities). |
| 3 | Orientation preferences change by 180° around singularities in a clockwise or counterclockwise fashion. |
| 4 | Both types of singularities appear in equal numbers. |
| 5 | There exist linelike regions (fractures), across which orientation preferences change rapidly with distance. |
| 6 | Below the *coherence length*, 820±140 $\mu$m in the monkey striate cortex, iso-orientation regions are organized as parallel slabs which start and end at singularities. |
| 7 | Slab orientations and orientation preferences are at most weakly correlated. |
| 8 | On a larger scale (above 820 $\mu$m) the pattern of orientation preferences is characterized by an incoherent superposition of waves with similar spatial frequency but arbitrary direction of their wave vector. |
| 9 | Orientation preferences are correlated over small distances (below 200 $\mu$m in the monkey striate cortex), tend to be anticorrelated at an intermediate distance (300 $\mu$m), and are uncorrelated on a global scale (above 820 $\mu$m). |
| 10 | Ocular dominance changes continuously as a function of cortical location. |
| 11 | The ocular dominance pattern is locally organized into parallel strips, which sometimes branch and terminate. |
| 12 | Iso-orientation slabs often cross the borders of ocular dominance bands at approximately right angles. |
| 13 | The singularities tend to align with the center of the ocular dominance bands. |

($q_r$), and ocular dominance ($z_r$). If $q_r$ is zero, then the units are unspecific for orientation, and the larger $q_r$ becomes, the sharper the units are tuned. "Binocular" units are characterized by $z_r=0$, units dominated by one eye by $z_r\neq 0$. These properties are encoded by the five-dimensional feature vector

$$\mathbf{w}_r=(x_r,y_r,q_r\cos(2\phi_r),q_r\sin(2\phi_r),z_r)\ , \tag{1}$$

where the orientation coordinates are given in their Cartesian forms (see [28]). The feature vectors $\mathbf{w}_r$, as a function of unit locations $\mathbf{r}$, describe the spatial pattern of feature preference and selectivity, i.e., the cortical map.

The stimuli which drive the formation of the cortical map are described by a *pattern vector* which is of the same dimensionality as $\mathbf{w}_r$. Its components

$$\mathbf{v}=(x,y,q\cos(2\phi),q\sin(2\phi),z) \tag{2}$$

correspond to the stimulus properties "position in the visual field" $(x,y)$, "orientation" $\phi$, pattern ellipticity $q$, and the distribution of activity between both eyes $z$. $q=0$ corresponds to circular stimuli and $q>0$ to elongated patterns. $z=0$ is to be interpreted as a "binocular" stimulus, while $z\neq 0$ corresponds to more "monocular" stimuli.

Since only little is known about the probability with which certain feature combinations $\mathbf{v}$ are present in the afferent patterns that drive the developmental process, we decided to use an "unbiased" probability distribution

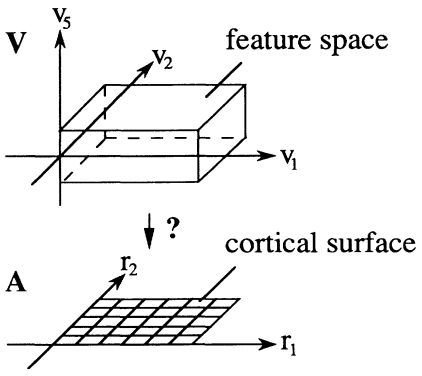


FIG. 2. The problem of dimension reduction. The coordinates $v_1$ and $v_2$ describing the position of a stimulus in visual space are plotted along the $x$ and $y$ axes, the coordinate $v_5$ describing ocular dominance by the $z$ axis. The coordinates $v_3$ and $v_4$ describing orientation selectivity are suppressed to allow for visualization.

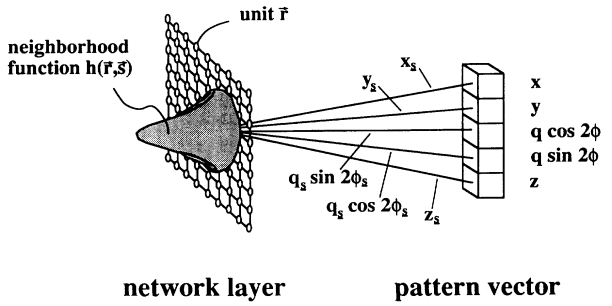


FIG. 3. The neural network model.

$P(\mathbf{v})$ and drew pattern vectors independently and with equal probability from the manifold

$$V=\{\mathbf{v}|x,y\in[0,d],\phi\in[0,\pi],q\langle q_{\text{pat}},|z|\langle z_{\text{pat}}\} \ , \tag{3}$$

i.e., all stimuli characterized by $q$ and $|z|$ smaller than a given value were chosen equally often.

Pattern formation is described by the SOFM algorithm which was introduced by Kohonen [13,14] to describe aspects of the input driven self-organization in the cerebral cortex. (For a biological plausible derivation as well as for a more detailed interpretation of this algorithm see [25].) The algorithm is discrete in space and time. At each time step a pattern vector $\mathbf{v}$ is chosen at random from a given set $\{\mathbf{v}\}$ with probability $P(\mathbf{v})$. Then the unit $\mathbf{s}$, whose feature vector $\mathbf{w}_s$ is closest to the input pattern $\mathbf{v}$, is selected:

$$\mathbf{s}=\min_{\mathbf{r}}|\mathbf{v}-\mathbf{w}_{\mathbf{r}}| \tag{4}$$

and the feature vectors in the network layer are changed according to

$$\mathbf{w}_{\mathbf{r}}(t+1)=\mathbf{w}_{\mathbf{r}}(t)+\varepsilon(t)h(\mathbf{r},\mathbf{s},t)[\mathbf{v}-\mathbf{w}_{\mathbf{r}}(t)] \ , \tag{5}$$

where $h(\mathbf{r},\mathbf{s},t)$, the neighborhood function, is usually given by

$$h(\mathbf{r},\mathbf{s},t)=\exp\left[-\frac{(r_1-s_1)^2}{\sigma_{h1}^2(t)}-\frac{(r_2-s_2)^2}{\sigma_{h2}^2(t)}\right] . \tag{6}$$

The SOFM algorithm is a simple mathematical formulation of a competitive learning network with the additional feature that spatially neighboring units cooperate in the learning process. A pattern vector $\mathbf{v}$ induces a localized "excitation" $h(\mathbf{r},\mathbf{s},t)$, which is assumed to be a continuous region in the network layer. Since the SOFM algorithm was designed to describe processes on a time scale that is a factor of $10^6$ longer than the time scale of the excitation dynamics, the excitation dynamics is not explicitly modeled. (For a neural implementation see [30].) The receptive field properties of the units are then changed by (5) proportional to their "excitation" (6) according to a Hebb-type learning rule with multiplicative constraints.

Thus, the developmental process is modeled as a particular realization of a Markov process whose states are the possible sets $\{\mathbf{w}_{\mathbf{r}}\}$ of feature vectors and whose transitions are triggered by the pattern vectors. The Markov assumption should be a good approximation for the biological system, because the time scale of cortical map formation (several days to several months) is much larger than the time scale on which correlations between subsequent activity patterns may exist (milliseconds to seconds).

### C. Mathematical properties

It is known that the SOFM algorithm generates a "structured representation," a mapping of the multidimensional input space onto the two-dimensional discrete lattice. Kohonen and others [25–27] showed that the vectors asymptotically settle to equilibrium values generating a map which (i) represents most faithfully those dimensions of feature space along which the standard deviation

$$T_i=\left[\frac{\int(v_i-\langle v_i\rangle)^2P(\mathbf{v})d\mathbf{v}}{\int P(\mathbf{v})d\mathbf{v}}\right]^{1/2} \tag{7}$$

is largest, (ii) tries to preserve continuity, such that (metrically) similar patterns are mapped onto neighboring points in the network layer, and (iii) reflects inhomogeneities in the probability density $P(\mathbf{v})$ such that regions with high $P(\mathbf{v})$ are mapped onto larger domains of the network layer. The quantities $T_i$ provide us with a set of order parameters as we will see below.

We consider a network $A$ consisting of $N\times N$ units with periodic boundary conditions to avoid edge effects and we assume periodic boundary conditions for the pattern manifold along $x,y$ coordinates in the input space. For any given state $\{\mathbf{w}_s\}$, $s\in A$ of the network the sets

$$\Gamma(\mathbf{s})=\{\mathbf{v}|\ |\mathbf{v}-\mathbf{w}_s|\le|\mathbf{v}-\mathbf{w}_{\mathbf{r}}|\forall\mathbf{r}\in A\} \tag{8}$$

render a tesselation of the pattern manifold. Each set $\Gamma(\mathbf{s})$ is called the *tesselation cell* of unit $\mathbf{s}$, and it corresponds to the set of all $\mathbf{v}$ for which the unit $\mathbf{s}$ is the winner unit according to (4).

Each selection of an input pattern changes the current state of the map and the corresponding tesselation according to (5). The average change in the feature vectors during one adaptation step is given by the "forces"

$$E(\Delta\mathbf{w}_{\mathbf{r}}|\{\mathbf{w}_{\mathbf{r}}\})=\varepsilon\int h(\mathbf{r},\mathbf{s}(\mathbf{v}))(\mathbf{v}-\mathbf{w}_{\mathbf{s}(\mathbf{v})})P(\mathbf{v})d\mathbf{v} \ . \tag{9}$$

Note, that in general these forces are not conservative and cannot be derived from a potential function (for details see [31,32]). A stationary state of the SOFM algorithm is given by the set $\{\mathbf{w}_{\mathbf{r}}^0\}$ of feature vectors for which all forces vanish:

$$E(\Delta\mathbf{w}_{\mathbf{r}}|\{\mathbf{w}_{\mathbf{r}}^0\})=0, \quad \forall\mathbf{r} \ . \tag{10}$$

If one stationary state has been found other stationary states can be constructed using symmetry properties of $P(\mathbf{v})$. Given a linear transformation in feature space $\underline{A}$ with $\underline{A}^{-1}=\lambda\underline{A}^{T}$, $\lambda\neq0$ real, and a "translation" vector $\mathbf{t}$, then a transformed state $\{\tilde{\mathbf{w}}_{\mathbf{r}}^0\}$

$$\tilde{\mathbf{w}}_{\mathbf{r}}^0=\underline{A}\mathbf{w}_{\mathbf{r}}^0+\mathbf{t} \tag{11}$$

of a stationary state $\{\mathbf{w}_{\mathbf{r}}^0\}$ is again stationary if the condition

$$\|\underline{A}\|P(\underline{A}\mathbf{v}+\mathbf{t})=P(\mathbf{v}) \tag{12}$$

is satisfied (see Appendix A). In the following we will assume that mature cortical maps correspond to the stable stationary states of the SOFM algorithm.

In order to describe the dynamics of the map it is convenient to introduce an ensemble of networks, each element of which is characterized by its own set $\{\mathbf{w}_{\mathbf{r}}\}$, $r\in A$,

of feature vectors, and describe this ensemble by a time-dependent distribution function $S(\{\mathbf{w}_r\},t)$. Since the pattern vectors are chosen randomly the time evolution of the network is a particular realization of a Markov process and the distribution function is transformed according to the Chapman-Kolmogoroff equation,

$$S(\{\mathbf{w}_r\},t+1)=\int Q(\{\mathbf{w}_r\}|\{\mathbf{w}'_r\})S(\{\mathbf{w}'_r\},t)\prod_r d^n\mathbf{w}'_r\ , \quad (13)$$

where the transition probability $Q(\{\mathbf{w}_r\}|\{\mathbf{w}'_r\})$ is given by

$$Q(\{\mathbf{w}_r\}|\{\mathbf{w}'_r\})=\sum_r \int_{\Gamma(s)} \prod_{r'} \delta(\mathbf{w}_r-\{\mathbf{w}'_{r'}+\varepsilon h(\mathbf{r},\mathbf{s},t)[\mathbf{v}-\mathbf{w}'_{r'}(t)]\})P(\mathbf{v})d\mathbf{v} \quad (14)$$

and where $n$ denotes the dimensionality of the input space.

If $\varepsilon$ is sufficiently small, and if the network is already in the vicinity of a stationary state, which is characterized by a distribution function $S(\{\mathbf{w}_r\},t)$ sufficiently peaked around the expectation values of $\{\langle \mathbf{w}_r\rangle\}$, then Eq. (13) can be expanded around the stationary state. Then the learning dynamics can be described by a Fokker-Planck equation

$$\frac{1}{\epsilon}\partial_t S(\{\mathbf{u}_r\},t)=\sum_{p,m,q,n}\frac{\partial}{\partial u_{pm}}B_{pmqn}u_{qn}S(\{\mathbf{u}_r\},t)+\frac{\epsilon}{2}\sum_{p,m,q,n}D_{pmqn}\frac{\partial^2 S(\{\mathbf{u}_r\},t)}{\partial u_{pm}\partial u_{qn}}\ , \quad (15)$$

where $\mathbf{u}_r=\mathbf{w}_r-\mathbf{w}_r^0$ and where the origin of $S(.,t)$ has been shifted to the stationary state $\{\mathbf{w}_r^0\}$. $B_{pmqn}$ are the components of a matrix given by

$$B_{pmqn}\equiv\left[\frac{\partial V_{pm}(\{\mathbf{w}_r\})}{\partial w_{qn}}\right]_{\{\mathbf{w}_r\}=\{\mathbf{w}_r^0\}}\ , \quad (16)$$

and the quantities $V_{rm}$ and $D_{pmqn}$ are the expectation values $\langle \Delta w_{rm}\rangle$ [cf. (9)] and $\langle \Delta w_{rm}\Delta w_{r'n}\rangle$ for one adaptation step, respectively.

This formulation of the learning dynamics near stationary states is derived in detail in [26]. Equations (15) and (16) will serve as a starting point for all subsequent calculations.

The quantities $B_{pmqn}$ and $D_{pmqn}$ are functionals of the probability distribution $P(\mathbf{v})$ and of the tesselation of the pattern manifold induced by $\{\mathbf{w}_r^0\}$. If $\underline{B}+\underline{B}^T$ is a positive-definite matrix, the stationary state under consideration is stable and the eigenvalues of $B_{pmqn}$ determine the time constants of the relaxation of fluctuations. The size of the statistical fluctuations that are present for nonvanishing step size $\epsilon$ are determined by $\underline{B}$ and $\underline{D}$ and are given by

$$\langle u_{rm}u_{sn}\rangle=\epsilon[(\underline{B}+\underline{B}^T)^{-1}\underline{D}]_{rm,sn}\ .$$

## IV. MATHEMATICAL ANALYSIS OF MAP FORMATION

For the pattern manifold given by (3) and any neighborhood function $h(\mathbf{r},\mathbf{s},t)$ monotonously decreasing with $|\mathbf{r}-\mathbf{s}|$, a set of stationary states is given by

$$x_r^0=\frac{d}{N}r_1,\quad y_r^0=\frac{d}{N}r_2,\quad q_r^0=0,\quad z_r^0=0 \quad (18)$$

and by all states, which result from (18) by translations $\mathbf{r}\to\mathbf{r}+\boldsymbol{\alpha}$, reflections $r_i\to -r_i$, and rotations by multiples of $\pi$ in the $x,y$ plane. The magnification factor of the mapping is $d/N$. These states correspond to a topographic representation of visual space, i.e., a retinotopic projection, where the "excess dimensions" $w_{r3}$, $w_{r4}$, and $w_{r5}$ are suppressed by the map.

Due to the invariance under translation of the probability distribution (3) and the stationary state (18), one can simplify Eq. (15) if one represents the distribution function in terms of the Fourier modes

$$\hat{\mathbf{u}}_\mathbf{k}=\frac{1}{N}\sum_r \exp(i\mathbf{k}\mathbf{r})\mathbf{u}_r\ . \quad (19)$$

of $\mathbf{u}_r$, where $\mathbf{k}$ is a two-dimensional wave vector in the network layer. For the biological application the range of the neighborhood function (6) must be large compared to the spacing of the lattice, i.e., $1\ll\sigma_{h1},\sigma_{h2}\ll N$. Inserting (19) into (15) one then obtains

$$\frac{1}{\varepsilon}\frac{\partial}{\partial t}\hat{S}(\hat{\mathbf{u}}_\mathbf{k},t)=\sum_{m,n}\hat{B}_{mn}(\mathbf{k})\frac{\partial}{\partial\hat{u}_{km}}[\hat{S}(\hat{\mathbf{u}}_\mathbf{k},t)\hat{u}_{kn}]+\tfrac{1}{2}\varepsilon\sum_{m,n}\hat{D}_{mn}(\mathbf{k})\frac{\partial^2\hat{S}(\hat{\mathbf{u}}_\mathbf{k},t)}{\partial\hat{u}_{km}\partial\hat{u}_{kn}}\ , \quad (20)$$

where the matrices $\underline{\hat{B}}$ and $\underline{\hat{D}}$ are given by (see Appendix B for their derivation)

$$\underline{\hat{B}}=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[\underline{1}-\underline{K}_B\exp\left[-\frac{\sigma_{h1}^1}{4}k_x^2\right]\times\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right] \quad (21)$$

and

$$\underline{\hat{D}}=\frac{d^2}{N^5}\pi^2\sigma_{h1}^2\sigma_{h2}^2\underline{K}_D\exp\left[-\frac{\sigma_{h1}^2}{2}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{2}k_y^2\right]\ . \quad (22)$$

The matrices $\underline{K}_B$ and $\underline{K}_D$ depend on the probability distribution $P(\mathbf{v})$. For a homogeneous probability distribution (3) these matrices are given by

$$\underline{K}_B = \begin{pmatrix} 1-\frac{1}{2}\sigma_{h1}^2 k_x^2 & \frac{1}{2}\sigma_{h1}^2 k_x k_y & 0 & 0 & 0 \\ -\frac{1}{2}\sigma_{h2}^2 k_x k_y & 1-\frac{1}{2}\sigma_{h2}^2 k_y^2 & 0 & 0 & 0 \\ 0 & 0 & \frac{N^2 T_3^2}{d^2} k^2 & 0 & 0 \\ 0 & 0 & 0 & \frac{N^2 T_4^2}{d^2} k^2 & 0 \\ 0 & 0 & 0 & 0 & \frac{N^2 T_5^2}{d^2} k^2 \end{pmatrix} \tag{23}$$

and

$$\underline{K}_D = \begin{pmatrix} \frac{1}{4}\sigma_{h1}^4 k_x^2 + \frac{1}{12} & \frac{1}{4}\sigma_{h1}^2 \sigma_{h2}^2 k_x k_y & 0 & 0 & 0 \\ \frac{1}{4}\sigma_{h1}^2 \sigma_{h2}^2 k_x k_y & \frac{1}{4}\sigma_{h2}^4 k_y^2 + \frac{1}{12} & 0 & 0 & 0 \\ 0 & 0 & \frac{T_3^2 N^2}{d^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{T_4^2 N^2}{d^2} & 0 \\ 0 & 0 & 0 & 0 & \frac{T_5^2 N^2}{d^2} \end{pmatrix}, \tag{24}$$

where $T_3 = T_4 = q_{\text{pat}}/2$ and $T_5 = z_{\text{pat}}/\sqrt{3}$.

We now can perform a stability analysis of the stationary state (18). The eigenvalues of the matrix $\underline{\hat{B}}$, which correspond to eigenmodes within the $x_1$-$x_2$ plane are given by

$$\lambda_{B1} = \frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-(1-\tfrac{1}{2}\sigma_{h1}^2 k_x^2 - \tfrac{1}{2}\sigma_{h2}^2 k_y^2)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right],$$
$$\lambda_{B2} = \frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right]. \tag{25}$$

The eigenvalues are positive except for $\mathbf{k}=\mathbf{0}$, where the eigenvalues vanish. Hence, the stationary state is stable under fluctuations of feature vectors from their stationary values parallel to the $x_1$-$x_2$ plane except for translations, which transform one stationary state into the other. The instability at $\mathbf{k}=\mathbf{0}$ is a result of the translation invariance of the stationary state.

The eigenvalues of the matrix $\underline{\hat{B}}$, which correspond to eigenmodes perpendicular to the $x_1$-$x_2$ plane are given by

$$\lambda_{Bi} = \frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-\frac{N^2 T_i^2}{d^2}k^2\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right], \quad i\in\{3,4,5\}. \tag{26}$$

If the quantities $T_i$ are sufficiently small, these eigenvalues are positive for all $\mathbf{k}$ and the topographic map (18) is stable. If, however, the $T_i$'s increase there exists a threshold

$$T_{\text{thres}} = \tfrac{1}{2}\sqrt{e}\,\frac{d}{N}\min(\sigma_{h1},\sigma_{h2}) \tag{27}$$

at which the eigenvalues for the set of modes characterized by

$$|\mathbf{k}^0| = 2/\sigma_h \quad \text{if } \sigma_{h1}=\sigma_{h2},$$
$$k_x^0 = \pm 2/\sigma_{h1}, \quad k_y^0 = 0 \quad \text{if } \sigma_{h1} < \sigma_{h2} \tag{28}$$

become negative. Accordingly, the stationary state, the topographic map, becomes unstable, the system undergoes a phase transition, and the network folds into feature space as shown in Fig. 4.

The quantities $T_i$, $i\in\{3,4,5\}$, which characterize the set of input patterns, assume the role of order parameters. Small values of the order parameters correspond to a set of input patterns, where the patterns are similar with respect to the features orientation and ocularity. In this regime a stable column system does not form. Only if the order parameters exceed a certain critical value, i.e., if the presented input patterns are sufficiently dissimilar with regard to the features $v_3$, $v_4$, or $v_5$, are ocularity and orientation represented in the network and a stable column system formed. The critical value is proportional to the range of the neighborhood function in the network layer along its minor axis, projected back to feature space.

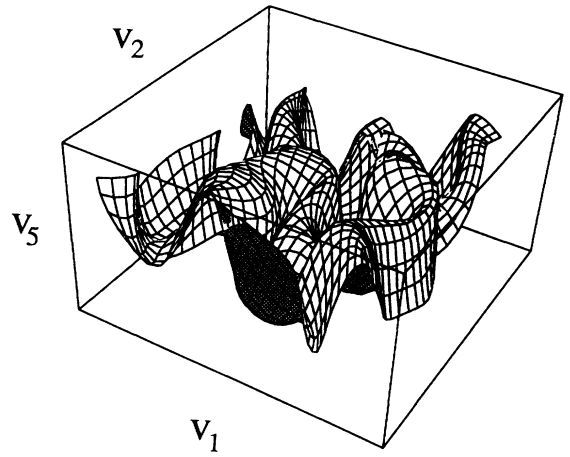


FIG. 4. A typical solution of the map problem (Fig. 2) found by the SOFM algorithm. The locations of the pattern vectors which are mapped to the units in the network are indicated by the intersections of a grid in feature space. Only every fourth vector is presented. The data was taken from a small region of the map shown in Fig. 10.

## V. SPATIAL STRUCTURE OF MAPS

### A. Maps below threshold

Below threshold, the stationary state corresponds to a homogeneous and topographic representation of visual space, where all the units respond equally to stimuli from either eye ($z_r=0$) and to elongated stimuli of all orientations ($q_r=0$). Fluctuations around the stationary state, however, are present for finite learning step size $\varepsilon$. These fluctuations lead to regions in the map that temporarily acquire selectivity for certain oriented stimuli and a biased response to patterns presented to the different eyes and they lead to the temporary formation of column systems.

The size of the statistical fluctuations around the stationary state can be determined by inserting the correlation matrix

$$\hat{C}_{mn}(\mathbf{k})=\langle \hat{u}_{\mathbf{k}m}\hat{u}_{\mathbf{k}n}\rangle \tag{29}$$

into Eq. (20). [Since modes of different wave vector $\mathbf{k}$ do not couple, their amplitudes and phases are uncorrelated. Therefore $\langle \hat{u}_{\mathbf{k}m}\hat{u}_{\mathbf{k}'n}\rangle=0$ for $\mathbf{k}\neq\mathbf{k}'$.] We obtain

$$\underline{\hat{B}}\,\underline{\hat{C}}+(\underline{\hat{B}}\,\underline{\hat{C}})^T=\varepsilon\underline{\hat{D}}\ , \tag{30}$$

where $\partial S/\partial t$ has been set to zero.

For $\sigma_{h1}\neq\sigma_{h2}$ the matrices $\underline{\hat{D}}$ and $\underline{\hat{B}}$ do not commute. Diagonalization of $\hat{C}$ given by (30) leads to rather complicated expressions for the diagonal elements $\hat{C}_{11}$ and $\hat{C}_{22}$ as a function of $\mathbf{k}$, which we omit for brevity. The nonvanishing coefficients $\hat{C}_{mm}, m>2$ are given by

$$\hat{C}'_{mm}=\frac{\varepsilon}{2}\pi T_m^2\sigma_{h1}\sigma_{h2}\times\frac{\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]}{\exp\left[\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[\frac{\sigma_{h2}^2}{4}k_y^2\right]-\frac{N^2T_m^2}{d^2}(k_x^2+k_y^2)}. \tag{31}$$

where $T_3=T_4=q_{\text{pat}}/2$ and $T_5=z_{\text{pat}}/\sqrt{3}$ and $q_{\text{pat}}$ and $z_{\text{pat}}$ are defined in (3).

For isotropic neighborhood functions ($\sigma_{h1}=\sigma_{h2}=\sigma_h$) the matrices $\underline{\hat{D}}$ and $\underline{\hat{B}}$ commute and it is convenient to transform Eq. (30) to the eigenbasis of $\underline{\hat{D}}$ and $\underline{\hat{B}}$ given by

$$\begin{aligned}&\mathbf{e}'_1=\mathcal{N}\{k_x\mathbf{e}_1+k_y\mathbf{e}_2\}\ ,\\ &\mathbf{e}'_2=\mathcal{N}\{k_y\mathbf{e}_1-k_x\mathbf{e}_2\}\ ,\\ &\mathbf{e}'_3=\mathbf{e}_3,\quad \mathbf{e}'_4=\mathbf{e}_4,\quad \mathbf{e}'_5=\mathbf{e}_5\ ,\end{aligned} \tag{32}$$

where $\mathcal{N}$ is a suitable normalization constant. We denote the transformed correlation matrix by $\hat{C}'$. The mean-square amplitudes of fluctuations are then given by

$$\begin{aligned}\hat{C}'_{11}&=\tfrac{1}{24}\varepsilon\frac{d^2}{N^2}\pi\sigma_h^2\frac{(3+\sigma_h^4k^2)\exp\left[-\frac{\sigma_h^2}{4}k^2\right]}{\exp\left[\frac{\sigma_h^2}{4}k^2\right]-1+\frac{1}{2}k^2\sigma_h^2}\ ,\\ \hat{C}'_{22}&=\tfrac{1}{24}\varepsilon\frac{d^2}{N^2}\pi\sigma_h^2\frac{\exp\left[-\frac{\sigma_h^2}{4}k^2\right]}{\exp\left[\frac{\sigma_h^2}{4}k^2\right]-1}\ ,\\ \hat{C}'_{mm}&=\frac{\varepsilon}{2}\pi T_m^2\sigma_h^2\frac{\exp\left[-\frac{\sigma_h^2}{4}k^2\right]}{\exp\left[\frac{\sigma_h^2}{4}k^2\right]-\frac{N^2T_m^2}{d^2}k^2},\quad m>2\ ,\end{aligned} \tag{33}$$

where $k=|\mathbf{k}|$. The common eigenvectors of $\underline{\hat{B}}$ and $\underline{\hat{D}}$ [see Eq. (32)] within the $x_1$-$x_2$ plane are parallel and perpendicular to the wave vector $\mathbf{k}$, respectively, and correspond to "compression" and "shearing" modes of the network layer.

The mean-square amplitudes of the modes, which describe fluctuations parallel to the $x_1$-$x_2$ plane, are proportional to the learning step $\varepsilon$ and independent of the order parameters $T_m, m>2$. The modes exhibit a $1/k$ instability at $\mathbf{k}=\mathbf{0}$ which results from the translational invariance of the system. For large values of $k$ these fluctuations decay exponentially with $|\mathbf{k}|^2$. The compression modes are "softer" than the shearing modes for all but the smallest values of $k$.

Perpendicular to the $x_1$-$x_2$ plane $P(\mathbf{v})$ and $\{\mathbf{w}_r^0\}$ lack translation invariance and fluctuation amplitudes are finite for $\mathbf{k}\to\mathbf{0}$. The mean-square amplitudes are proportional to the learning step $\varepsilon$. For small values of the order parameters, (31) reduces to a Gaussian function of ellipsoidal cross section, which is centered at the origin of the $\mathbf{k}$ plane. The width along its minor and major axes is given by the inverse of the width of the neighborhood function along its major and minor axes, respectively. In this regime fluctuations with long wavelength dominate, which lead to positive correlations between the feature values $w_{r3}$, $w_{r4}$, and $w_{r5}$ of neighboring regions in the network layer.

With increasing value of the order parameters the amplitudes of modes from two regions characterized by $\mathbf{k}$

vectors of opposite sign (anisotropic neighborhood function) or from a ring-shaped region in $\mathbf{k}$ space (isotropic neighborhood function) become larger. Slightly below the critical value the modes corresponding to the peak dominate the spatial structure of the maps and lead to negative spatial correlations between feature values at a distance of $\pi/|\mathbf{k}^0|$, where $|\mathbf{k}^0|$ is the wave number of the peak. The time constants $\tau_i$ given by

$$\tau_i(\mathbf{k})=\lambda_{Bi}^{-1}(\mathbf{k}), \quad i>2 \tag{34}$$

for the relaxation of critical fluctuations diverge proportional to $1/(T_i-T_{\text{thres}})$. Autocorrelation functions have a sombrero shape. For large $k$ the power spectrum decays exponentially with $k^2$.

To test the analytical solutions described Monte Carlo simulations have been carried out using a network of $256\times256$ units and isotropic neighborhood functions. The solution Eq. (33) and the numerically obtained power spectra are shown in Fig. 5. (Note that the mean-square amplitudes of the modes have been averaged over all directions of $\mathbf{k}$ to obtain a better approximation of the power spectrum.) Figures 5(a) and 5(b) show the power spectrum of fluctuations as a function of spatial frequency for the "compression" and "shearing" modes. There is good agreement between the analytical solution and the numerical results except in the vicinity of the $1/k$ instability, where the relaxation time diverges, and for very large values of $k$, where the numerical accuracy was not sufficient.

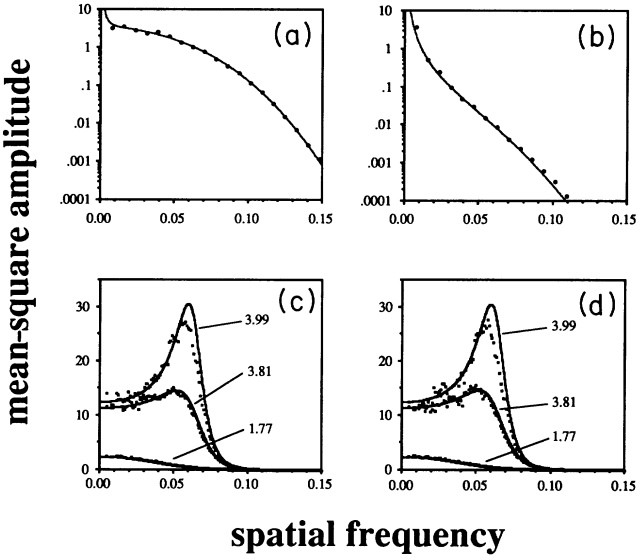


FIG. 5. Power spectra of fluctuations obtained with an isotropic neighborhood function ($\sigma_h=5$). Solid lines indicate the analytical solution (33); dots the results of Monte Carlo simulations. Monte Carlo simulations were carried out with a network of size $N=d=256$. In order to obtain a good approximation of the power spectrum, the squared Fourier spectrum was averaged over 40 maps obtained after 50 000, 300 000, and 600 000 iterations for $T_{3,4,5}=1.77$, 3.81, and 3.99, respectively, from the "topographic" initial state (18). Additionally, the squared Fourier spectrum was averaged over all directions of the $\mathbf{k}$ vector. The step size $\varepsilon$ was 0.02. For this set of parameters the critical values of the order parameters are given by $T_{\text{thres}}=4.12$. (a) Power spectrum of the "compression" modes. (b) Power spectrum of the "shearing" modes. Since the power spectra of "compression" and "shearing" modes are independent of $T_{3,4,5}$ only values for $T_{3,4,5}=1.77$ are shown. (c) Power spectrum of the $w_{r3}$ and $w_{r4}$ modes. Since these modes are degenerate, only one of them ($w_{r3}$) is shown. (d) Power spectrum of the $w_{r5}$ mode.

Figures 5(c) and 5(d) present the power spectra of fluctuations of the modes parallel to $w_{r3}$, $w_{r4}$, and $w_{r5}$ for different values of $T_i$. There is again good agreement between the analytical result and the result of the numerical simulations, except for the parameter regime in the vicinity $[(T_i-T_{\text{thres}})/T_{\text{thres}}<5\%]$ of the phase transition, where the "discreteness" of the network layer comes into play and where the approximations leading from Eq. (50) to Eq. (21) are no longer valid. However, the results of numerical simulations approach the analytical solution in the limit $\sigma\to\infty$, $\sigma/N$ finite, $d/N$ finite, where the higher-order terms in the expansion of (50) become less important.

The corresponding spatial autocorrelation functions (see Appendix C)

$$S_{ii}(\mathbf{p})=\langle u_{ri}u_{(\mathbf{r}-\mathbf{p})i}\rangle, \quad i>2 \tag{35}$$

are presented in Fig. 6 for the case $i=3,4$. (The autocorrelation function of the other coordinate $i=5$ has a similar shape.) The maps were generated with an isotropic neighborhood function. The correlation functions were again averaged over all directions of $\mathbf{p}$. For very small values of $T_i$ neighboring units in the network have positively correlated feature values and the length $l_c$ of correlations is given by $l_c\approx\sqrt{2}\sigma$. If the order parameters approach their critical value the correlation length increases and the autocorrelation functions become sombrero shaped. At the critical value feature coordinates are anticorrelated for units at a distance of $\lambda=1.9\sigma$ and correlations decay as $1/\sqrt{|\mathbf{p}|}$. Since the modes are independent, all cross-correlation functions vanish.

In the following we will visualize the results of this section in a form similar to the experimental data shown in

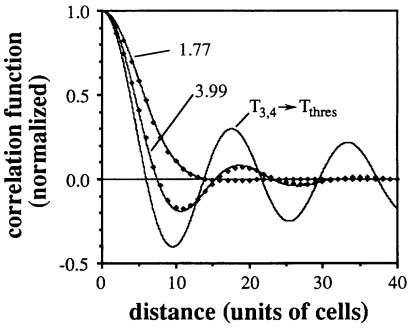


FIG. 6. Autocorrelation functions $S_{33}(\mathbf{p})$ of the orientation coordinate $w_{r3}$ as a function of distance $|\mathbf{p}|$ obtained with an isotropic neighborhood function ($\sigma_h=5$). All correlation functions were normalized to $S_{33}(0)=1$. Solid lines indicate the analytical solution (56), dots indicate the results of the Monte Carlo simulations in Fig. 5. The simulation results were averaged over all directions of $\mathbf{p}$. The dotted line shows the (normalized) autocorrelation function (58) in the limit $T_3\to T_{\text{thres}}$.

Fig. 1. Figure 7(a) shows a snapshot of fluctuations in the orientation coordinates for a small value of the order parameters $T_{3,4}$ and for an isotropic neighborhood function. Orientation preference $\phi_{\mathbf{r}}$ of the units $\mathbf{r}$ is again coded by color, orientation tuning strength by brightness. The map consists of bright patches containing orientation selective units [arrow in Fig. 7(a)], which are embedded into larger unselective areas. If $T_{3,4}$ is small, neighboring units have similar feature values and neighboring patches have similar colors. If $T_{3,4}$ approaches $T_{\text{thres}}$ the correla-

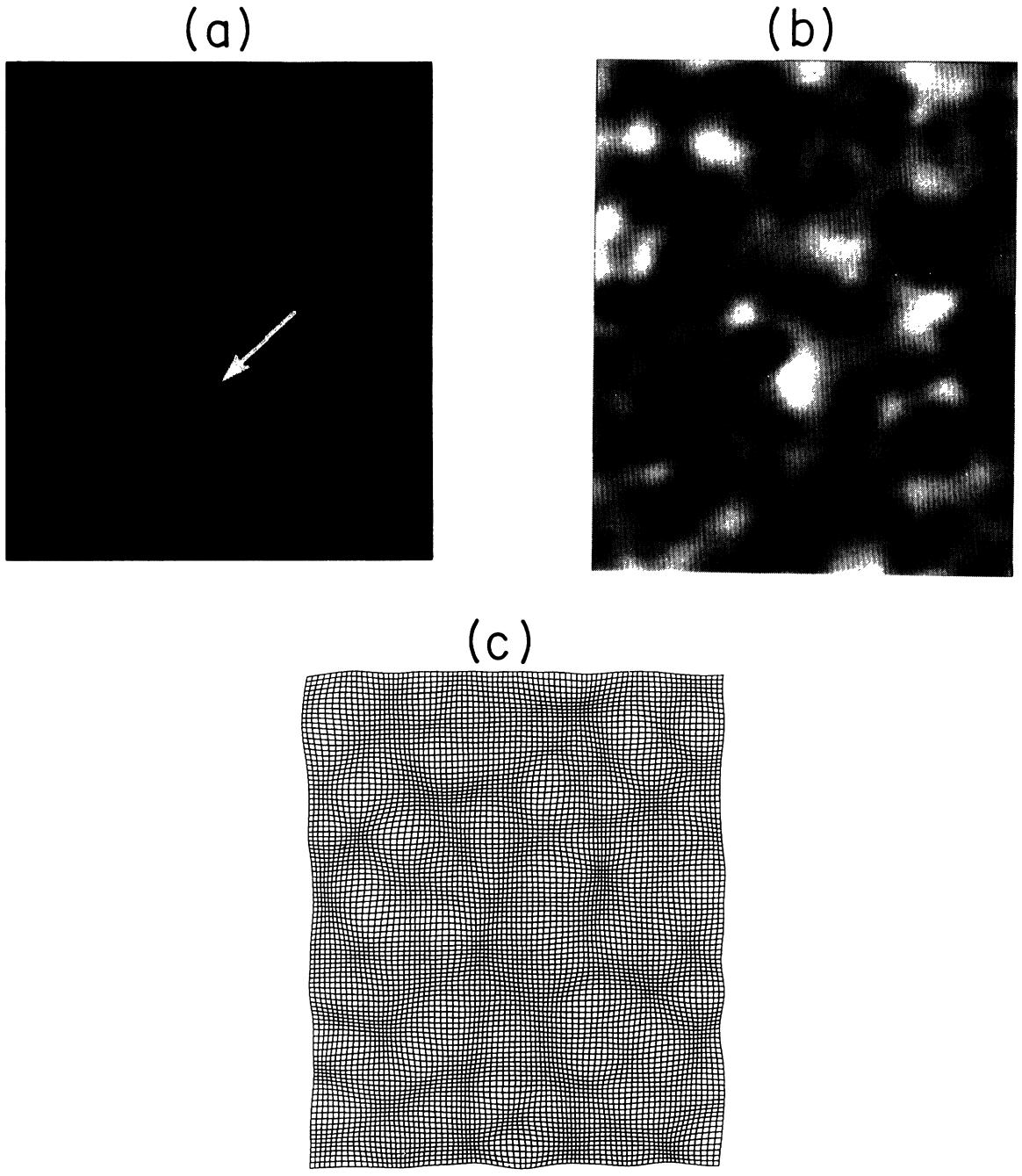


FIG. 7. Snapshot of fluctuations around the stationary state (18). The figure displays an $80\times100$ section of a network of size $N=d=512$. The parameters of the simulation were $\varepsilon=0.1$, $\sigma_h=5$, $T_{3,4,5}=1.24$. For this set of parameters the critical value of the order parameters is $T_{\text{thres}}=4.12$. The snapshot was obtained after 60 000 iterations from the "topographic" initial state (18). (a) Spatial pattern of orientation preference $\phi_{\mathbf{r}}$ (color) and orientation tuning strength $q_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (b) Spatial pattern of ocular dominance values $z_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. Orientation preference, orientation tuning strength, and ocular dominance values were color coded as described in Fig. 1. (c) Spatial structure of the topographic projection. The figure displays the locations $(x_{\mathbf{r}1},x_{\mathbf{r}2})$ of receptive field centers in visual space for all units in the network layer. Receptive field centers of neighboring cells were connected by lines.

tion length increases, but units which are a certain distance apart tend to have anticorrelated feature values. This gives the map a more "colorful" appearance.

Figure 7(b) shows a snapshot of the corresponding fluctuations in the ocular dominance values. The values $z_{\mathbf{r}}$ are coded by brightness, where white and black denote left eye and right eye preference, respectively. Again the map seems patchy and "monocular" islands (black or white) are embedded into large "binocular" areas (gray).

A snapshot of the structure of the retinotopic projection is depicted in Fig. 7(c). The diagram presents the locations $(x_{\mathbf{r}1},x_{\mathbf{r}2})$ of receptive field centers in visual space for all cells in the network layer. Each location is represented by a dot; receptive field centers of neighboring cells are connected by lines. An ideal topographic projection of visual space to the network layer would give rise to a square lattice with homogeneous mesh size, since the receptive field centroids of neighboring units are equally spaced. The nearly perfect grid demonstrates that the representation of visual space is topographic as predicted by Eq. (18); the wavelike patterns are the result of fluctuations in the "compression" modes. (The amplitude of the "shearing" modes is much smaller.) Since all modes are independent, there exist no correlations between the spatial patterns of retinotopy, orientation selectivity, and ocular dominance.

### B. Maps above threshold

Above the critical value of the order parameters a stable pattern of orientation and ocular dominance is formed. Since there exists no general theory of stationary states of the self-organizing-feature-map algorithm in this regime, the investigation of the spatial pattern of topography, orientation selectivity, and ocular dominance has to rely on numerical simulations. In order to describe these patterns and to compare them with experimental data we will consider the following set of functions and transformations:

Fourier transforms:

$$\hat{w}_{\mathbf{k}j}=\sum_{\mathbf{r}}\exp(i\mathbf{k}\mathbf{r})w_{\mathbf{r}j}\ ;$$

correlation functions:

$$S_{ij}(\mathbf{p})=\langle w_{\mathbf{r}i}w_{(\mathbf{r}+\mathbf{p})j}\rangle_{\mathbf{r}}\ ;$$

feature functions:

$$\{w_{\mathbf{r}j}\}=\{x_{\mathbf{r}},y_{\mathbf{r}},q_{\mathbf{r}},\phi_{\mathbf{r}},z_{\mathbf{r}},q_{\mathbf{r}}\cos\phi_{\mathbf{r}},q_{\mathbf{r}}\sin\phi_{\mathbf{r}},q_{\mathbf{r}}(\cos\phi_{\mathbf{r}}+i\sin\phi_{\mathbf{r}})\}\ ;$$

feature gradients:

$$|\nabla_{\mathbf{r}}w_{\mathbf{r}j}|=\{[w_{(r_1+1,r_2)j}-w_{(r_1,r_2)j}]^2+[w_{(r_1,r_2+1)j}-w_{(r_1,r_2)j}]^2\}^{1/2}\ ;$$

Gabor transforms:

$$g_j(\mathbf{k},\mathbf{r})=(2\pi\sigma_g^2)^{-1/4}\int d^2r'w_{\mathbf{r}'j}\exp\left[-\frac{(\mathbf{r}-\mathbf{r}')^2}{4\sigma_g^2}+i\mathbf{k}(\mathbf{r}'-\frac{1}{2}\mathbf{r})\right]\ . \tag{36}$$

Fourier transforms and correlation functions are the natural quantities from a theoretical point of view. Feature functions and feature gradients are traditionally used to display the experimental data. Gabor transforms have been introduced to objectively find, classify, and count (local) modules of cortical organization (e.g., regions containing singularities or parallel iso-orientation slabs). Since the regime $T_i > T_{\text{thres}}$ generates the typical spatial patterns of mature cortical maps, we will compare model predictions with experimental data as we go along.

#### 1. *Fourier transforms and correlation functions*

The modes which become unstable at $T_{\text{thres}}$ are the modes which have the strongest growth rate for $T_i > T_{\text{thres}}$. Slightly above threshold, these modes dominate the spatial pattern of orientation selectivity and ocular dominance. Figure 8 shows the Fourier transform of the complex orientation coordinate $\hat{u}_{\mathbf{r}3}+i\hat{u}_{\mathbf{r}4}$ for patterns generated with an isotropic [Fig. 8(a)] and an anisotropic [Fig. 8(b)] neighborhood function. The origin of the $\mathbf{k}$ plane is marked by a dot. Each pixel corresponds to a single mode $\mathbf{k}$ and its brightness indicates the mean-square amplitude $|\hat{\mathbf{u}}_{\mathbf{k}}|^2$ of the mode $\mathbf{k}$. For an isotropic neighborhood function the orientation map is characterized by wave vectors from a ring-shaped region in the Fourier domain [Fig. 8(a)], which becomes eccentric with increasing $\sigma_{h1}/\sigma_{h2}$ [Fig. 8(b)] until the ring dissolves into two separate groups of modes. Phases (not shown) seem to be random. The wavelength $\lambda_0$ associated with the modes with high energy determines the average period of the orientation columns which is similar along every direction in the network. The width of the peak indicates the typical length (820 $\mu$m in the monkey striate cortex) over which orientation preferences are correlated. Figure 8(c) shows a typical Fourier spectrum of the spatial pattern of orientation selectivity in the macaque. The orientation map of the macaque has a slightly eccentric Fourier spectrum, which indicates a slight tendency of the iso-orientation slabs to align with the area 17-18 border. Since the spectrum is nearly round, a power spectrum could be approximated by averaging over all directions of the $\mathbf{k}$ vector. Theoretical predictions and experimental data are overlaid in Fig. 8(d) after suitable normalization. The comparison shows that the model is able to reproduce the line shape in the experimentally deter-

mined power spectrum. Thus, the pattern of orientation columns is characterized by local correlation and global disorder.

As a result of the line-shaped power spectrum, the autocorrelation functions of the orientation and ocular dominance coordinates have a sombrero shape. If $\sigma_{h1}=\sigma_{h2}$ the autocorrelation functions have rotational symmetry, and at a separation of about $\lambda_0/2$ the preferred orientations $\phi_{\mathbf{r}}$ are more likely to be orthogonal. For $\sigma_{h1}\neq\sigma_{h2}$ these functions essentially consist of a central "bar" of positive values accompanied by two "lobes" of negative values. The cross-correlation functions are zero in both cases. Figure 9 shows the correlation functions $S_{ij}, i,j\in\{3,4\}$ of the spatial pattern of orientation selectivity in adult monkeys in comparison with simulation results obtained for $\sigma_{h1}=\sigma_{h2}$. The model predicts the shape of the correlation functions well for $S_{34}$ and $S_{44}$ but deviates from $S_{33}$. The difference in the amplitude of $S_{33}$ and $S_{44}$ in the experimental data, however, is likely to be an artifact of the finite map size. Differences in amplitude of the correlation functions obtained from seven animals did not show any consistent trend.

The fact that the patterns of orientation preference and ocular dominance have a simple description in the Fourier domain might also explain the success of most models of cortical map development [10–24]. Isotropic spectra are the result of the invariance of Eq. (20) ($\sigma_{h1}=\sigma_{h2}$) under rotation with respect to cortical coordinates $\mathbf{r}$; global disorder, the finite coherence length, and singularities are a consequence of their invariance under translation. Any model which assumes invariance under translation and rotation with respect to cortical coordinates and, additionally, is characterized by a typical spatial frequency (or a typical length) is very likely to gen-

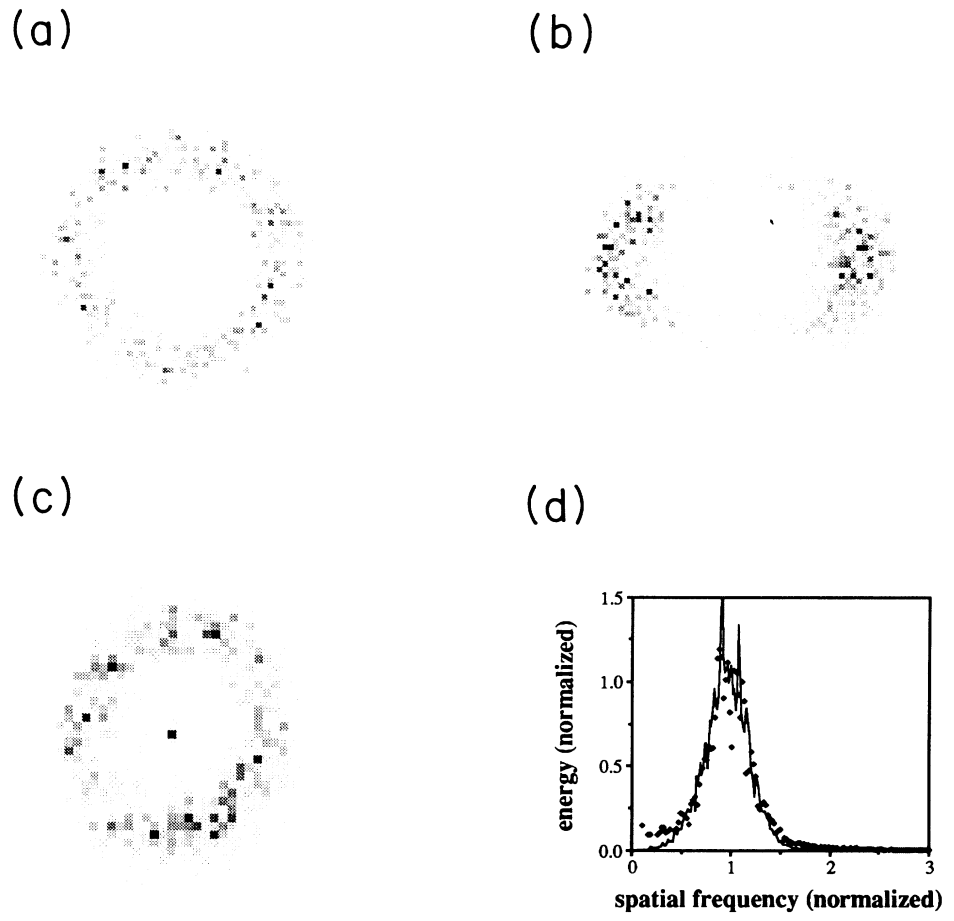


FIG. 8. Complex Fourier transforms $\sum_{\mathbf{r}}\exp(i\mathbf{kr})q_{\mathbf{r}}(\cos\phi_{\mathbf{r}}+i\sin\phi_{\mathbf{r}})$ of the spatial pattern of orientation preference and corresponding power spectra. (a) Fourier transform of the pattern of orientation of a typical stationary state of the SOFM algorithm generated for $T_{3,4,5}>T_{\text{thres}}$ using an isotropic neighborhood function. Parameters of the simulation were $N=d=512$, $\varepsilon=0.02$, $\sigma_h=5$, $T_{3,4}=10.24$, and $T_5=8.87$. For this set of parameters the critical value of the order parameters is $T_{\text{thres}}=4.12$. The figure was obtained after $9\times10^7$ iterations from the "topographic" initial state (18). (b) Fourier transform of the pattern of orientation selectivity for a typical stationary state of the SOFM algorithm generated for $T_{3,4,5}>T_{\text{thres}}$ using an anisotropic neighborhood function. The parameters of the simulation were $N=d=512$, $\varepsilon=0.02$, $\sigma_{h1}=5$, $\sigma_{h2}=7.5$, $T_{3,4}=10.24$, and $T_5=8.87$; the critical value of the order parameters is $T_{\text{thres}}=4.12$. The figure was obtained after $9\times10^7$ iterations from the "topographic" initial state (18). (c) Fourier transform of the pattern of orientation selectivity obtained from a $6\times8$ mm$^2$ patch of macaque striate cortex. (d) Power spectrum of (a) and (c) obtained by averaging over all directions of the $\mathbf{k}$ vector. Mean peak amplitude and peak frequency were normalized to 1. Unit frequency corresponds to 1.38/mm for the experimental data.

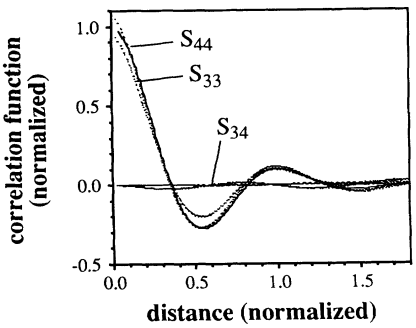


FIG. 9. Correlation functions $S_{33}(\mathbf{p})$, $S_{44}(\mathbf{p})$, and $S_{34}(\mathbf{p})$ as a function of cortical location $\mathbf{p}$ for the experimental data shown in Fig. 8(c) (dotted lines) and simulation results described in Fig. 8(a) (solid lines). The correlation functions were averaged over all directions of $\mathbf{p}$. The labels $S_{33}$, $S_{44}$, and $S_{34}$ refer to the dotted lines. The sombrero-shaped solid line displays the simulation result for $S_{33}$ and $S_{44}$, which have identical graphs; the horizontal line displays $S_{34}$.

erate reasonable patterns. These principles—isotropy, homogeneity, and periodicity—alone, however, cannot account for the recently found correlations between orientation preference and ocular dominance, properties 11–13 in Table I of Sec. II.

### 2. *Structure of the patterns in the spatial domain*

In the following we will again visualize the simulation results in a form similar to the experimental data of Fig. 1. Figure 10(a) shows the spatial pattern of orientation preference and orientation tuning strength generated above threshold using an isotropic neighborhood function. Color and brightness again denote orientation preference and tuning strength, respectively. The pattern shows the following qualitative features of the experimental data [compare, e.g., Fig. 10(a) with Fig. 1(a) or Fig. 11(b) with Fig. 11(a)]. (Figure 11 is located next to Fig. 1.)

(1) Orientation preference changes continuously as a function of cortical location except for isolated points (singularities).

(2) Iso-orientation regions are organized as elongated, parallel patches which start and end at the singularities [see arrow 1 in Fig. 10(a)].

(3) The singularities have a vorticity of $\pm\frac{1}{2}$, and both types of singularities appear in equal numbers. For the numerical simulation presented in Fig. 10(a) the ratio of $+\frac{1}{2}$ to $-\frac{1}{2}$ singularities was 0.98 (855 vortices total); six singularities had a vorticity of $\pm 1$.

(4) All preferred orientations are represented equally in the first approximation. This is a consequence of the invariance of the pattern manifold (3) under rotations within the $v_3$-$v_4$ plane.

The existence of singularities (and of global disorder) can be understood from an entropy argument. Since ~~dimension-reducing maps which exhibit these features~~ have increased entropy, they are generated with higher probability.

Patterns generated with anisotropic neighborhood functions exhibit the same features, except that the iso-orientation slabs are now aligned with the major axis of the neighborhood function. Anisotropic orientation-column systems have not been observed in the macaque but have been observed in areas 17 and 18 of the cat [33,4]. In the cat the iso-orientation slabs align with the direction of extended horizontal connections. This is consistent with the assumptions underlying the self-organizing-feature-map algorithm [25], because an anisotropic neighborhood function is thought to reflect an underlying anisotropic lateral connectivity scheme, where the directions of the extended lateral connections coincide with the major axis of $h(\mathbf{r},\mathbf{s})$.

The spatial pattern of ocular dominance exhibits alternating stripelike regions, which are each dominated by one particular eye. For $\sigma_{h1}=\sigma_{h2}$ the ocular dominance pattern is organized into parallel stripes only locally [Fig. 10(b)] and the emerging pattern is more irregular than the ocular dominance pattern of the macaque. A better fit can be obtained for anisotropic neighborhood functions, where the stripes align and the pattern closely resembles the experimental data. Unfortunately, it is not possible to fit the degree of "anisotropy" of both the macaque's orientation and ocular-dominance column system without introducing different neighborhood functions for different coordinates, i.e., replacing $h(\mathbf{r},\mathbf{s},t)$ by $\{h_i(\mathbf{r},\mathbf{s},t)\}$. In the case of sufficiently "anisotropic" neighborhood functions the pattern shows the following qualitative features of the experimental data [compare Fig. 11(b) with Fig. 11(a)].

(i) Ocular dominance changes continuously as a function of cortical location.

(ii) The ocular dominance pattern is locally organized into parallel stripes, which sometimes branch and run into dead ends.

(iii) The spatial pattern of ocular dominance is characterized by a power spectrum which consists of two groups of modes with opposite wave vectors.

(iv) Monocular regions cover a larger total area than binocular regions.

Properties 1 and 2 are satisfied by patterns generated with isotropic as well as with anisotropic neighborhood functions.

Note that patterns of orientation selectivity and ocular dominance are rather complicated in the spatial domain. This indicates that Fourier space [20] should be preferred over the spatial domain (see [34] for an overview of models) in order to parametrize these patterns. In fact the properties 1–10 (Table I of Sec. II) of the individual column systems emerge if amplitudes and phases are randomly assigned to Fourier modes located in a ring-shaped or eccentric region in Fourier space [20]. Pattern models which are based on colored noise, however, cannot account for correlations between topography, orientation selectivity, and ocular dominance, properties 11–13 of Table I.

### 3. *Correlations between topography, orientation selectivity, and ocular dominance*

In contrast to maps generated below the critical value of the order parameter, the spatial patterns of orientation

selectivity, ocular dominance, and topography are strongly correlated. Figure 11 shows the spatial pattern of orientation selectivity (color and brightness) overlayed by the borders of the ocular dominance bands (white lines) for the result of numerical simulations as well as for experimental data. In both cases the singularities of the orientation map lie in the centers of the ocular dominance bands and, consequently, the iso-orientation slabs intersect ocular dominance bands at steep angles (see arrows).

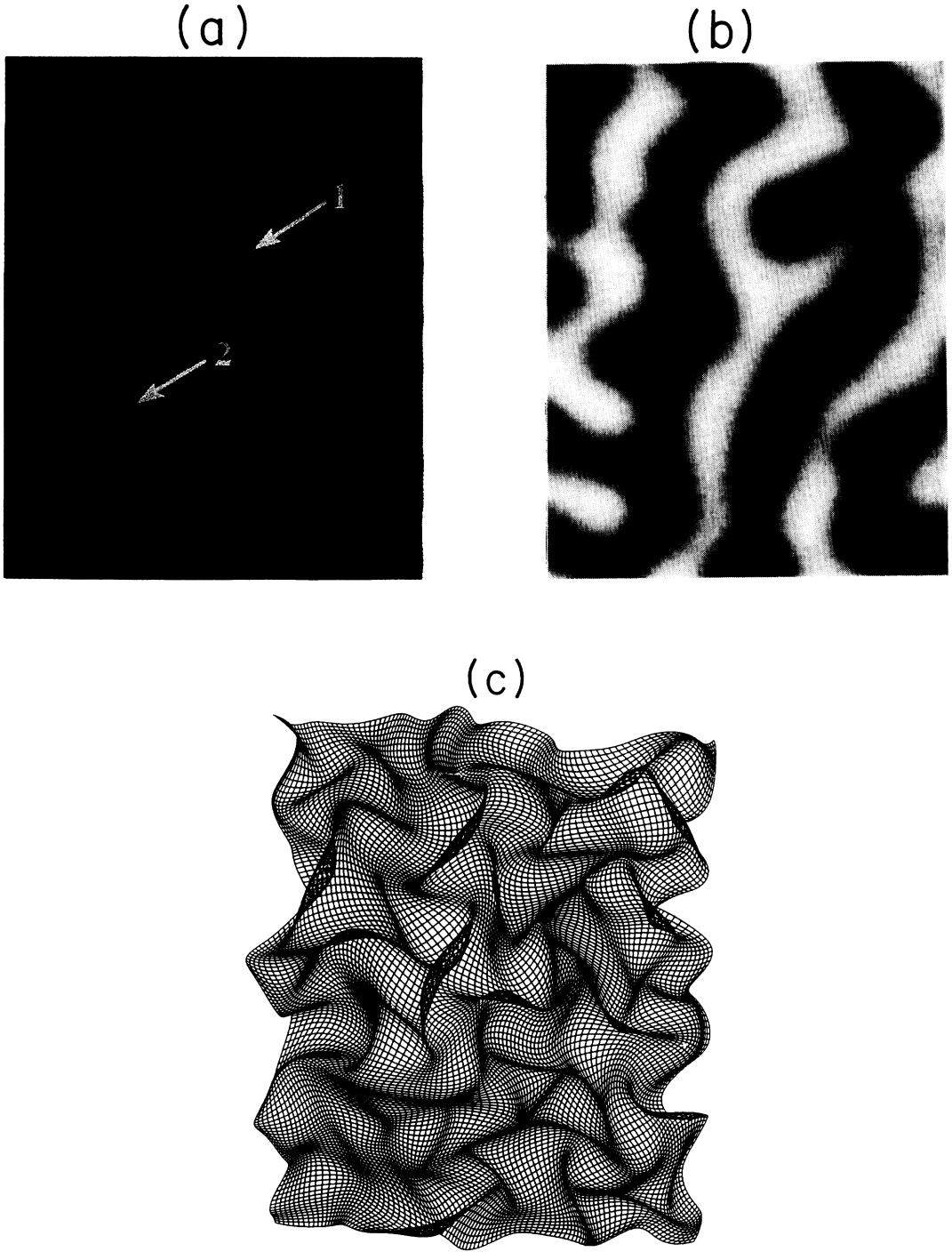


FIG. 10. Typical stationary state of the self-organizing feature maps generated for $T_{3,4,5} > T_{\text{thres}}$ using an isotropic neighborhood function. The figure displays a $110 \times 150$ section of a network of size $N = d = 512$. The parameters of the simulation are given in Fig. 8(a). (a) Spatial pattern of orientation preference $\phi_{\mathbf{r}}$ (color) and orientation tuning strength $q_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. Arrows indicate singularities (arrow 1) and linear regions (arrow 2). (b) Spatial pattern of ocular dominance values $z_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (c) Spatial structure of the topographic projection. The components of the feature vectors were coded as described in Fig. 7.

These correlations can be addressed quantitatively by a Gabor analysis. If the radius $\sigma_g$ of the Gaussian function in the Gabor filter (36) is smaller than the "coherence length," which is proportional to the inverse width of the peak in the power-spectrum, the Gabor transform of any of the feature components $w_{r3}$, $w_{r4}$, and $w_{r5}$ typically consist of two localized regions of high energy in the $\mathbf{k}$ plane (Fig. 12). The length $|\mathbf{k}|$ of the $\mathbf{k}$ vector, which corresponds to the centroids of these regions, fluctuates around the characteristic wave number $2\pi/\lambda_0$ of this pattern; its direction gives the normal to the ocular dominance bands and iso-orientation slabs at the location $\mathbf{r}$, where the Gabor transform was performed.

We define the quantities

$$s_1 = |\psi_3 - \psi_4| \ , \qquad s_2 = |\psi_5 - \tfrac{1}{2}(\psi_3 + \psi_4)| \ , \tag{37}$$

where $\psi_3$, $\psi_4$ and $\psi_5$ denote the angle between the vectors $\mathbf{k}_3$-$\mathbf{k}_5$ and a reference axis for the orientation and ocular dominance coordinates $w_{r3} - w_{r5}$, respectively. If $s_1 \approx 0°$ and $s_2 \approx 90°$, then iso-orientation slabs are parallel and run perpendicular to the ocular dominance stripes and the pattern locally resembles the "ice-cube" architecture proposed by Hubel and Wiesel [1] ("linear regions"). If $s_1 \approx 90°$, then the particular regions contain singularities.

Figure 13 shows the number of map locations as a function of $s_1$ and $s_2$ for a large number of randomly selected locations. Regions where orientation preferences are organized as parallel slabs (back row) are the most abundant in the experimental data as well as in the simulation results. Most of the regions appear on the right ($s_1 \approx 0°$ and $s_2 \approx 90°$) where iso-orientation slabs intersect ocular dominance bands at nearly right angles. Regions containing singularities ($s_1 \approx 90°$, $s_2$ arbitrary) are less abundant.

There is also a strong correlation between regions with low orientation tuning strength [dark areas in Fig. 10(a)] and the vortices and other areas of high magnitude of the orientation gradient. Since vortices are correlated with the centers of ocular dominance stripes and hypercolumnlike regions with their borders, the monocular regions are characterized by lower orientation tuning strength on the average.

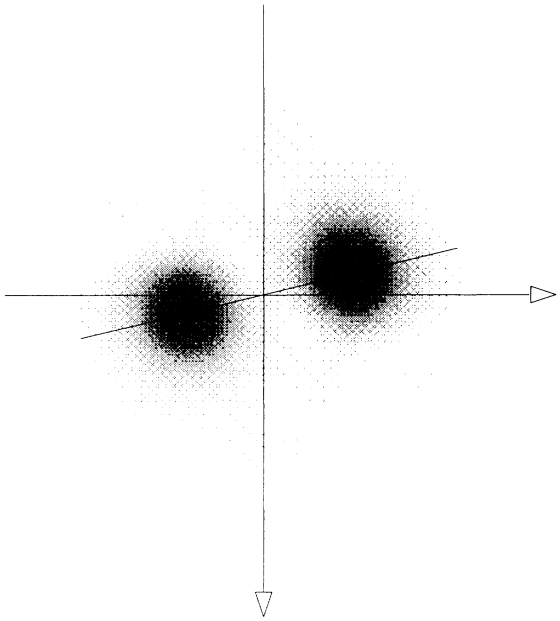

FIG. 12. Gabor transform of $w_{ri}$, $i > 2$ at a typical location in the pattern described in Fig. 10 for $\sigma_g = 6$. Black regions indicate modes with high energy.

In conclusion, patterns generated above the critical value of the order parameters show a lateral segregation into two different regions: (i) binocular regions with linear changes in orientation preferences, where iso-orientation slabs run perpendicular to the ocular dominance bands, and (ii) monocular regions with low orientation tuning strength, which contain the singularities of the orientation map.

The stable representation of orientation selectivity and ocular dominance above the critical value of the order parameter also introduces periodic distortions into the topographic projection [Fig. 10(c)] and gives rise to regions where the receptive field centers of neighboring cells coincide or reverse order. The reason for these distortions as well as for the correlations between orientation selectivity and ocular dominance becomes evident from Fig. 4. In order to continuously represent the five-dimensional feature space on the two-dimensional sur-

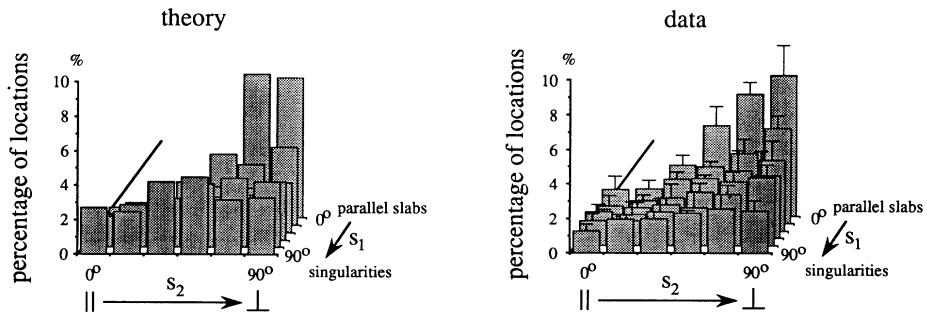


FIG. 13. Number of map locations as a function of $s_1$ and $s_2$ (see text). (a) Analysis of the stationary state of the SOFM algorithm displayed in Fig. 10. The total number of map locations was 1500, $\sigma_g/\lambda_0 = 0.2$. (b) Analysis of experimental data (taken from [3,5,6]) obtained from four adult monkeys. The total number of map locations was 3500, $\sigma_g/\lambda_0 = 0.2$. $\sigma_g$ was 150 $\mu$m, the typical radius of an interneuron dendritic tree located in the superficial layers of the monkey striate cortex [35].

face, regions of high rate of change along the "position" axis have to coincide with regions of zero or low rate of change along the orthogonal "ocular dominance" (and "orientation") axes (crests and troughs). Likewise, regions of high rate of change in the "ocular dominance" (and "orientation") axes (steep slopes) coincide with regions of zero or low rate of change along the "position" axis. Other correlations arise at points where the map exhibits maximal changes in two features. For example, for retinotopic location ($v_1$) and ocular dominance ($v_5$) to vary at a maximal rate the surface in Fig. 4 must be parallel to the $(v_1, v_5)$ plane. Obviously, at such points the directions of maximal change of retinotopic location and ocular dominance are orthogonal on the surface.

In order to compare model predictions with experimental data the surface in the five-dimensional feature space has to be projected into the three-dimensional subspace spanned by orientation preferences ($v_3$ and $v_4$) and ocular dominance ($v_5$). This projection cannot be visualized easily (the surface completely fills the space intersecting itself multiple times) but the same line of reasoning applies: (i) regions where orientation preferences change quickly correlate with regions where ocular dominance changes slowly, and (ii) in regions where orientation preferences change most rapidly along one direction, ocular dominance has to change most rapidly along the orthogonal direction. Consequently we expect discontinuities of the orientation map to be located in the centers of the ocular dominance bands and iso-orientation slabs to intersect ocular dominance bands at steep angles.

There is another way to look at these correlations. These correlations emerge because for a constant probability distribution $P(\mathbf{v})$, the SOFM algorithm distributes the units of the network in feature space such that the distance between neighboring units is approximately the same. Since this distance corresponds to the magnitude of the rate of change $|\nabla_{\mathbf{r}}\mathbf{w}_{\mathbf{r}}|$ of the feature vector $\mathbf{w}_{\mathbf{r}}$, it follows that a high rate of change in one component must be compensated by a small rate of change in the other components.

### 4. *The regime of large $T_i$*

Our simulations show that for increasing values of the order parameters the standard deviation of features values from their stationary state values $\{\mathbf{w}_{\mathbf{r}}^0\}$ increases proportional to $T_i$, $i>2$, and that the characteristic wave number $k^0$ of the column system decreases. Since Monte Carlo simulations of a large enough 2D network require prohibitively much computation time, we have restricted our analysis of the regime of large order parameters to a simpler case, the mapping of a rectangle, onto a one-dimensional "chain" of units. The only order parameter $T_2$ is given by $q/\sqrt{3}$, where $q$ is half the width of the rectangle. Figure 14 shows the standard deviation $\Delta \equiv \{[\sum_i (w_{i2}-w_{i2}^0)^2]/N(N-1)\}^{1/2}$ of feature values and the characteristic wave number as a function of the order parameter. Above the critical value, $T_{\text{thres}}=\sqrt{e}\,(d/N)\sigma_h$, the standard deviation $\Delta$ increases linearly with the order parameter, i.e., $\Delta = T_2 + \text{const}$ except for a small region near the instability. Hence $\Delta$ increases linearly with the width of the rectangle in order to faithfully represent the additional patterns presented to the network. The characteristic wavelength at the critical value is given by $2/\sigma_h$. Above $T_{\text{thres}}$ it also increases proportional to $T_2$, but with different shape.

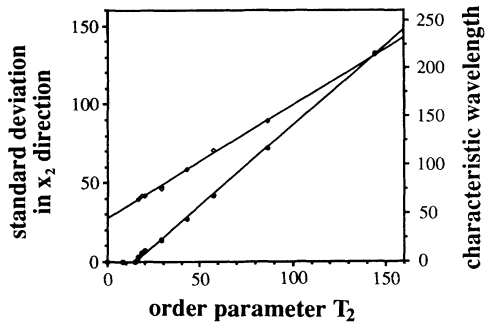


FIG. 14. Standard deviation of features values and characteristic wavelength as a function of the order parameter. Numerical simulations were restricted to the simplest case of a dimension-reducing mapping, the representation of a rectangle of length $d$ and width $2q$ by a one-dimensional network of $N$ units with periodic boundary conditions. Parameters of the simulation were $N=d=8192$, $\sigma_h=20$, $\varepsilon=0.02$. For this set of parameters $T_{\text{thres}}=33.0$. The coordinates of the two-dimensional feature vector are denoted by $w_1$ and $w_2$. Patterns were chosen with equal probability from the manifold $V=\{\mathbf{v}|v_1\in[0,d],|v_2|<q\}$. The filled symbols denote the standard deviation $\Delta$ as a function of the order parameter after the system has settled into a stationary state. The open symbols denote the wavelength corresponding to the peak in the power spectrum of $x_2$. The solid lines indicate the result of a linear regression to fit the standard deviation $\Delta$ and the characteristic wavelength as a function of $T_2$ above the critical value.

We consider again the case of the 5D→2D map. If the $T_i$, $i>2$ are large enough, the feature hierarchy apparent in Fig. 10, namely, "global" order in $x$ and $y$ and "local" order in $q$, $\phi$, and $z$, breaks down. "Orientation preference" or "ocular dominance" now play the role of the primary stimulus variable. Figures 15(a) and 15(b) display orientation selectivity and ocular dominance as a function of unit location $\mathbf{r}$ in this parameter regime, respectively. There is only one continuous region for each interval of preferred orientation and one for each eye, but each of these regions now contains a representation of a large part of visual space. Consequently the position map [Fig. 15(c)] shows multiple representations of visual space. This demonstrates that hierarchical patterns are generated by the feature map algorithm only when there exists a hierarchy in the variances of the set of patterns along the various feature dimensions (in our example a hierarchy in the magnitudes of $T_{1,2}$, $T_{3,4}$, and $T_5$). The features with the largest variance become the primary features (in the case shown in Fig. 10 the positions of receptive field centroids); the other features become secondary features, which are represented multiple times in the network layer (in the case shown in Fig. 10 orientation selectivity and ocular dominance). This property of the SOFM algorithm is known as the "automatic selection of feature dimension [25]."

## VI. SUMMARY AND CONCLUSIONS

In this contribution we have investigated an unsupervised learning algorithm, the self-organizing feature map, which is able to explain the formation of cortical maps in the primary visual cortex of the macaque. We have shown that the spatial pattern of orientation and ocular dominance columns emerges as the result of a global instability of the retinotopic map, when one or several order parameters, which describe properties of the set of input patterns, exceed a critical value $T_{\text{thres}}$. The critical value is proportional to the range of the neighborhood function along its minor axis, multiplied by the inverse magnification factor of the retinotopic mapping. Since

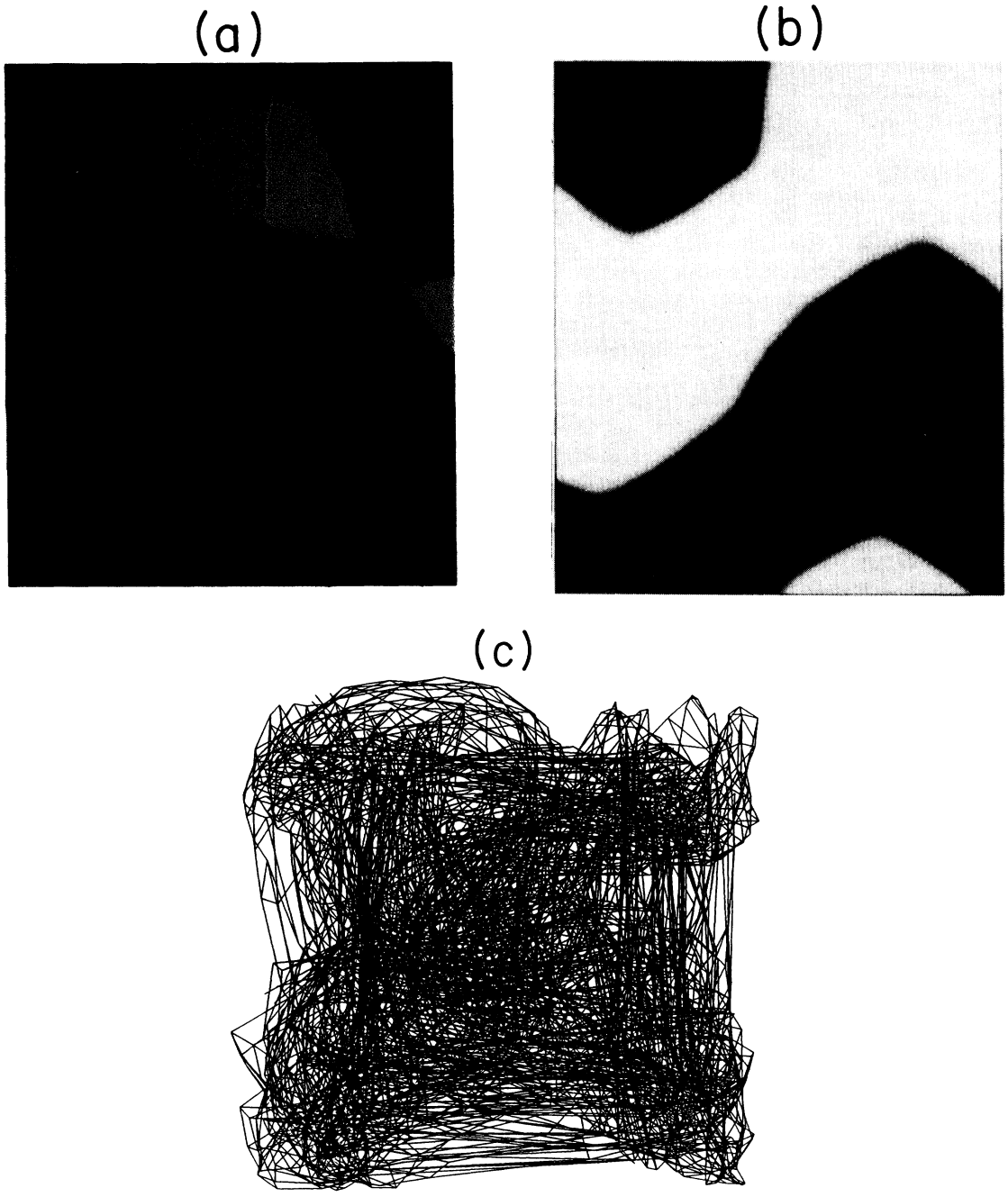


FIG. 15. Typical stationary state of the self-organizing feature maps generated for $T_{3,4,5} \gg T_{\text{thres}}$ using an isotropic neighborhood function. The figure displays a $192 \times 256$ section of a network of size $N = d = 256$. The parameters of the simulation were $\varepsilon = 0.1$, $T_{3,4} = 1280$, and $T_5 = 1478$. The patterns were obtained after 60 000 iterations starting from random initial conditions. $\sigma_h$ was linearly decreased during the simulation from $\sigma_h(0) = 200$ to $\sigma_h(30\,000) = 20$ and then to $\sigma_h(60\,000) = 4$. For this set of parameters the critical value of the order parameters, which depends on $\sigma_h$, lies between 663.0 and 3.3. (a) Spatial pattern of orientation preference $\phi_{\mathbf{r}}$ (color) and orientation tuning strength $q_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (b) Spatial pattern of ocular dominance values $z_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (c) Spatial structure of the topographic projection. The components of the feature vectors were coded as described in Fig. 7.

the range of the neighborhood function is thought to be related to the range of lateral excitatory connections [30], one might speculate that the range of excitatory connections must not exceed a certain value in order to allow column systems to form.

Above, but still near, this critical value the spatial patterns strongly resemble the ocular dominance and orientation columns found in the primary visual cortex of the macaque. The spatial pattern slightly above threshold is dominated by the modes which become unstable at $T_{\text{thres}}$. The spatial patterns of orientation preference, ocular dominance, and topography are correlated giving rise to a lateral segregation of the pattern into binocular regions with linear changes in orientation preferences, where isoorientation slabs run perpendicular to the ocular dominance bands, and monocular regions with low orientation tuning strength which contain the singularities of the orientation map. Also a stable representation of more than two features introduces periodic distortions into the retinotopic projection. The characteristic wave numbers of these distortions are predicted to be twice the characteristic wave number of the orientation and ocular dominance columns, respectively. While experimental results from the primary visual cortex of the macaque seem to confirm correlations between orientation selectivity and ocular dominance and between orientation preference and selectivity, the predicted distortions of the retinotopic projection have not yet been addressed experimentally and it is still an open question if they are present in the visual cortex of the macaque.

Does the regime of small values of the order parameters also have biological significance? Before birth and in very young animals, where the visual system may not be fully functional, afferent activity patterns corresponding to lines and contours may not be present. Hence the afferent patterns would exhibit a weak orientation bias only and the corresponding order parameters $T_{3,4}$ should be below their critical values. In this case the model would predict that (i) the orientation preferences in young monkeys are not correlated with ocular dominance and (ii) the pattern is not stable but changes in time. The regime of small $T_i$ could also be created artificially by depriving young animals of input patterns characterized by large $q$ and $z$ [see Eq. (2)]. In this case the model would lead to the following prediction. If a young animal is deprived of either monocular stimuli or patterns with a strong orientation bias, then stable column systems should *not* form. Both column systems should be uncorrelated; they should fluctuate and randomly change the spatial structure in time.

## ACKNOWLEDGMENTS

We wish to thank H. Ritter for numerous and fruitful discussions as well as for his help in the initial phase of this work. Financial support by the Boehringer-Ingelheim Fonds to K.O. is gratefully acknowledged. This research has been supported by the National Science Foundation (Grant No. DIR 91-22522) and by the National Institute of Health (Grant No. P41RRO5969). Computer time on the Connection Machine CM-2 has been made available by the National Center for Supercomputer Applications at Urbana-Champaign funded by the National Science Foundation.

## APPENDIX A: CONSTRUCTION OF STATIONARY STATES

Let $\underline{A}$ be a linear transformation and $\mathbf{t}$ a translation in feature space, respectively, and define the quantities $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{w}}_\mathbf{r}$ by

$$\begin{aligned}\tilde{\mathbf{v}}&=\underline{A}\mathbf{v}+\mathbf{t}\ ,\\ \tilde{\mathbf{w}}_\mathbf{r}&=\underline{A}\mathbf{w}_\mathbf{r}+\mathbf{t}\ .\end{aligned}\tag{A1}$$

If $\underline{A}^{-1}=\lambda\underline{A}^{T}$, $\lambda\neq 0$ real, then

$$\begin{aligned}\mathbf{s}&=\min_\mathbf{r}|\tilde{\mathbf{v}}-\tilde{\mathbf{w}}_\mathbf{r}|=\min_\mathbf{r}|\underline{A}(\mathbf{v}-\mathbf{w}_\mathbf{r})|\\ &=\min_\mathbf{r}|\mathbf{v}-\mathbf{w}_\mathbf{r}|\ .\end{aligned}\tag{A2}$$

Hence we obtain for the forces $E(\Delta\tilde{\mathbf{w}}_\mathbf{r}|\{\tilde{\mathbf{w}}_\mathbf{r}\})$

$$\begin{aligned}E(\Delta\tilde{\mathbf{w}}_\mathbf{r}|\{\tilde{\mathbf{w}}_\mathbf{r}\})&=\varepsilon\int h(\mathbf{r},\mathbf{s}(\mathbf{v}))(\tilde{\mathbf{v}}-\tilde{\mathbf{w}}_{\mathbf{s}(\mathbf{v})})P(\tilde{\mathbf{v}})d\tilde{\mathbf{v}}\\ &=\underline{A}\cdot\varepsilon\int h(\mathbf{r},\mathbf{s}(\mathbf{v}))(\mathbf{v}-\mathbf{w}_{\mathbf{s}(\mathbf{v})})\\ &\quad\times P(\underline{A}\mathbf{v}+\mathbf{t})\|\underline{A}\|d\mathbf{v}\ .\end{aligned}\tag{A3}$$

If $P(\mathbf{v})=\|\underline{A}\|P(\underline{A}\mathbf{v}+\mathbf{t})$, then

$$E(\Delta\tilde{\mathbf{w}}_\mathbf{r}|\{\tilde{\mathbf{w}}_\mathbf{r}\})=\underline{A}\cdot E(\Delta\mathbf{w}_\mathbf{r}|\{\mathbf{w}_\mathbf{r}\})\ .\tag{A4}$$

Hence, if Eq. (10) is fulfilled, then

$$E(\Delta\tilde{\mathbf{w}}_\mathbf{r}|\{\tilde{\mathbf{w}}_\mathbf{r}\})=0\ .\tag{A5}$$

## APPENDIX B: DERIVATION OF EXPRESSIONS FOR $\hat{\underline{B}}$ AND $\hat{\underline{D}}$

### 1. Derivation of an expression for $\underline{B}$

The "volume" $\bar{P}_\mathbf{r}$ and the coordinates $\bar{v}_{\mathbf{r}m}$ of the "centroid" of the tesselation cell of unit $\mathbf{r}$ are given by

$$\begin{aligned}\bar{P}_\mathbf{r}&=\int_{\Gamma(\mathbf{s})}P(\mathbf{v})d\mathbf{v}\ ,\\ \bar{v}_{\mathbf{r}m}&=\bar{P}_\mathbf{r}^{-1}\cdot\int_{\Gamma(\mathbf{s})}v_{\mathbf{r}m}d\mathbf{v}\ ,\end{aligned}\tag{B1}$$

respectively. Inserting (9) and (B1) into (16) and performing the differentiation we obtain

$$\begin{aligned}B_{\mathbf{p}m\mathbf{q}n}&=\frac{\partial V_{\mathbf{p}m}}{\partial w_{\mathbf{q}n}}\\ &=\sum_\mathbf{s}h(\mathbf{p},\mathbf{s})\frac{\partial\bar{P}_\mathbf{r}}{\partial w_{\mathbf{q}n}}(\bar{v}_{\mathbf{s}m}-w_{\mathbf{p}m})\\ &\quad-\sum_\mathbf{s}h(\mathbf{p},\mathbf{s})\bar{P}_\mathbf{s}\left[\frac{\partial\bar{v}_{\mathbf{s}m}}{\partial w_{\mathbf{q}n}}-\delta(\mathbf{p}m,\mathbf{q}n)\right]\ ,\end{aligned}\tag{B2}$$

which has to be evaluated for the stationary state $\{\mathbf{w}_\mathbf{r}^0\}$ given by Eq. (18).

For the stationary state (18) the relations $\bar{P}_\mathbf{r}=1/N^2$ and $\bar{v}_{\mathbf{r}m}=w_{\mathbf{r}m}$ hold. Inserting these into (B2), and performing the Fourier transform with respect to differences between unit labels, we obtain

$$\begin{aligned}\hat{B}_{mn}(\mathbf{k})&=\frac{1}{N^2}\hat{h}(0)\left[\delta(m,n)-N\frac{\hat{h}(\mathbf{k})}{\hat{h}(0)}\frac{\partial\hat{\bar{v}}_m}{\partial w_n}\right]\\ &\quad+id[\nabla_\mathbf{k}\hat{h}(\mathbf{k})]_m\frac{\partial\bar{P}}{\partial w_n}\ ,\end{aligned}\tag{B3}$$

where $\hat{h}(\mathbf{k})$ denotes the Fourier transform

$$\hat{h}(\mathbf{k})=\frac{1}{N}\sum_{(\mathbf{p}-\mathbf{s})}\exp[i\mathbf{k}(\mathbf{p}-\mathbf{s})]h(\mathbf{p},\mathbf{s}) \tag{B4}$$

of the neighborhood function (6) and where we made use of the fact that the tesselation of the pattern manifold induced by $\{\mathbf{w}_r^0\}$ is translation invariant. In the limit $1\ll\sigma_{h1},\sigma_{h2}\ll N$ relation (B4) can be replaced by the continuous Fourier transform and we get

$$\hat{h}(\mathbf{k})=\frac{1}{N}\pi\sigma_{h1}\sigma_{h2}\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]. \tag{B5}$$

Next we have to evaluate the derivatives $\partial\bar{P}_r/\partial w_{qn}$ and $\partial\bar{v}_{sm}/\partial w_{qn}$. The "volume" and the "centroid" of a tesselation cell $\Gamma_p$ change only under shifts $\partial\mathbf{w}_q$ of feature vectors corresponding to this cell and its neighbors. Figure 16 illustrates the calculation of these changes. If we neglect contributions of second or higher order in $\partial\mathbf{w}_q$ (parts of the hatched areas in Fig. 16) we obtain

$$\frac{\partial\bar{P}_p}{\partial\mathbf{w}_q}=\frac{1}{2dN}\begin{bmatrix}\delta[(p_1+1)p_2,q_1q_2]-\delta[(p_1-1)p_2,q_1q_2]\\ \delta[p_1(p_2+1),q_1q_2]-\delta[p_1(p_2-1),q_1q_2]\\ 0\\ 0\\ 0\end{bmatrix} \tag{B6}$$

and

$$\frac{\partial\bar{v}_{pm}}{\partial w_{qm}}=\begin{bmatrix}
\frac{2}{3}\delta(p_1p_2,q_1q_2)+\frac{1}{4}[\delta((p_1-1)p_2,q_1q_2)+\delta((p_1+1)p_2,q_1q_2)]\\
-\frac{1}{12}[\delta(p_1(p_2-1),q_1q_2)+\delta(p_1(p_2+1),q_1q_2)]\\
\frac{2}{3}\delta(p_1p_2,q_1q_2)-\frac{1}{12}[\delta((p_1-1)p_2,q_1q_2)+\delta((p_1+1)p_2,q_1q_2)]\\
+\frac{1}{4}[\delta(p_1(p_2-1),q_1q_2)+\delta(p_1(p_2+1),q_1q_2)]\\
\frac{q_{\text{pat}}^2N^2}{d^2}\delta(p_1p_2,q_1q_2)-\frac{q_{\text{pat}}^2N^2}{4d^2}[\delta((p_1-1)p_2,q_1q_2)+\delta((p_1+1)p_2,q_1q_2)]\\
-\frac{q_{\text{pat}}^2N^2}{4d^2}[\delta(p_1(p_2-1),q_1q_2)+\delta(p_1(p_2+1),q_1q_2)]\\
\frac{q_{\text{pat}}^2N^2}{d^2}\delta(p_1p_2,q_1q_2)-\frac{q_{\text{pat}}^2N^2}{4d^2}[\delta((p_1-1)p_2,q_1q_2)+\delta((p_1+1)p_2,q_1q_2)]\\
-\frac{q_{\text{pat}}^2N^2}{4d^2}[\delta(p_1(p_2-1),q_1q_2)+\delta(p_1(p_2+1),q_1q_2)]\\
\frac{4z_{\text{pat}}^2N^2}{3d^2}\delta(p_1p_2,q_1q_2)-\frac{z_{\text{pat}}^2N^2}{3d^2}[\delta((p_1-1)p_2,q_1q_2)+\delta((p_1+1)p_2,q_1q_2)]\\
-\frac{z_{\text{pat}}^2N^2}{3d^2}[\delta(p_1(p_2-1),q_1q_2)+\delta(p_1(p_2+1),q_1q_2)]
\end{bmatrix} \tag{B7}$$

All "centroid" changes vanish for $m\neq n$. The fourier transformation of Eqs. (B6) and (B7), and insertion of (B6), (B7), and (B5) into Eq. (B3) lead to

$$\begin{aligned}
\hat{B}_{11}(\mathbf{k})&=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-(\tfrac{2}{3}+\tfrac{1}{2}\cos k_x+\tfrac{1}{2}\sigma_{h1}^2k_x\sin k_x-\tfrac{1}{6}\cos k_y)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right],\\
\hat{B}_{12}(\mathbf{k})&=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}(\tfrac{1}{2}\sigma_{h1}^2k_x\sin k_y)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right],\\
\hat{B}_{21}(\mathbf{k})&=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}(\tfrac{1}{2}\sigma_{h2}^2k_y\sin k_x)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right],\\
\hat{B}_{22}(\mathbf{k})&=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-(\tfrac{2}{3}+\tfrac{1}{6}\cos k_x+\tfrac{1}{2}\sigma_{h2}^2k_y\sin k_y-\tfrac{1}{2}\cos k_y)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right],\\
\hat{B}_{33}(\mathbf{k})&=B_{44}(\mathbf{k})=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-\frac{N^2q_{\text{pat}}^2}{d^2}(1-\tfrac{1}{2}\cos k_x-\tfrac{1}{2}\cos k_y)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right],\\
\hat{B}_{55}(\mathbf{k})&=\frac{1}{N^3}\pi\sigma_{h1}\sigma_{h2}\left[1-\frac{4N^2z_{\text{pat}}^2}{3d^2}(1-\tfrac{1}{2}\cos k_x-\tfrac{1}{2}\cos k_y)\exp\left[-\frac{\sigma_{h1}^2}{4}k_x^2\right]\exp\left[-\frac{\sigma_{h2}^2}{4}k_y^2\right]\right].
\end{aligned} \tag{B8}$$

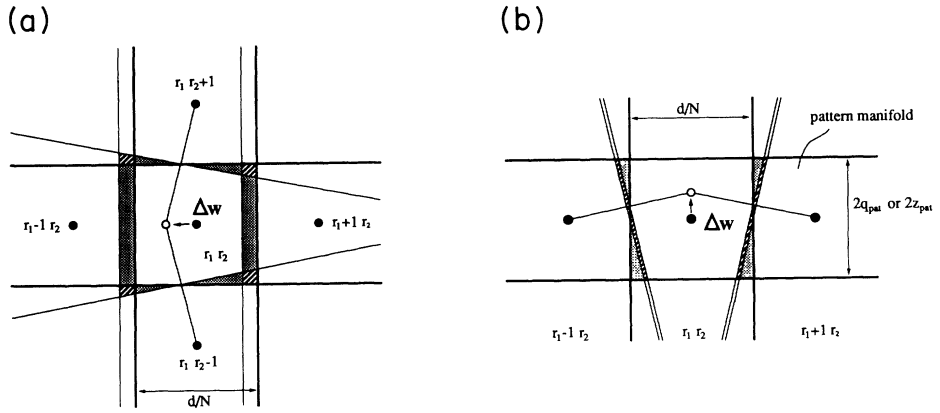


FIG. 16. Change of shape of the tesselation cells corresponding to the stationary state (18) under an infinitesimal shift $\partial\mathbf{w}_r$ of the weight vector $\mathbf{w}_r$. (a) Shift parallel to the coordinates $v_1$ and $v_2$. The drawing shows a projection of the tesselation onto the $v_1,v_2$ plane, where the positions of the weight vectors before and after the shift are indicated by black and white dots, respectively. The dotted areas indicate volume changes linear in $\partial\mathbf{w}_r$; they contribute to Eqs. (B6) and (B7). The hatched areas give an upper bound of corrections of $O(|\partial\mathbf{w}_r|^2)$ to these volume changes; they are added to or subtracted from the dotted areas to keep expressions (B6) and (B7) simple. (b) Shift parallel to the coordinates $v_3$ to $v_5$. The drawing shows a projection of the tesselation onto the $v_1,v_i$ plane where $i\in\{3,4,5\}$. The two horizontal lines indicate the borders of the pattern manifold. Hatched and dotted areas again indicate volume changes of first and higher order in $\partial\mathbf{w}_r$, respectively. Note, that a calculation of volume and centroid changes under $\partial w_{3r}$ and $\partial w_{4r}$ have to take into account the cylindrical shape of the pattern manifold, while changes under $\partial w_{5r}$ have to take into account an essentially boxlike geometry. This leads to the different prefactors in Eq. (B7).

For $\sigma_{h1},\sigma_{h2}\gg 1$ we can either neglect the expressions containing the exponentials or expand the sin and cos terms in the prefactors to leading order. If we further neglect $O(k^2)$ terms relative to $O(\sigma_{h1,h2}^2 k^2)$ we end up with the simplified expressions given in Eqs. (21) and (23).

### 2. Derivation of an expression for $\hat{\underline{D}}$

The matrix elements of $\underline{D}$ are given by

$$D_{pmqn}=\langle \Delta w_{pm}\Delta w_{qn}\rangle = \sum_s h(\mathbf{p},\mathbf{s})h(\mathbf{q},\mathbf{s}) \times[(w^0_{pm}-\bar{v}_{sm})(w^0_{qm}-\bar{v}_{sm})\bar{P}_s+M_{mn}]\ , \quad \text{(B9)}$$

where the matrix elements $M_{mn}$ are given by

$$M_{mn}=\int_{\Gamma_s}(v_m v_n-\bar{v}_{sm}\bar{v}_{sn})P(\mathbf{v})d\mathbf{v}\ . \quad \text{(B10)}$$

Due to the translational invariance of the stationary state the quantities $M_{mn}$ are independent of the location $\mathbf{s}$ in the network layer. Fourier transformation of (B9) gives

$$\hat{D}_{mn}(\mathbf{k})=\frac{d^2}{N^3}\{[\nabla_{\mathbf{k}}\hat{h}(\mathbf{k})]_m[\nabla_{\mathbf{k}}\hat{h}(\mathbf{k})]_n^T+[\hat{h}(\mathbf{k})]^2 M_{mn}\}\ , \quad \text{(B11)}$$

where $\hat{h}(\mathbf{k})$ again denotes the Fourier transform (B5) of the neighborhood function (6). Straightforward integration of (B10) yields $M_{mn}=0$, $\forall m\neq n$ and

$$M_{mm}=\left[\frac{1}{12},\frac{1}{12},\frac{q_{pat}^2N^2}{4d^2},\frac{q_{pat}^2N^2}{4d^2},\frac{z_{pat}^2N^2}{3d^2}\right] \quad \text{(B12)}$$

for the diagonal elements. Inserting (B12) into (B11) leads to the simplified expressions given in Eqs. (22) and (24).

## APPENDIX C: CALCULATION OF CORRELATION FUNCTIONS

For isotropic neighborhood functions ($\sigma_{h1}=\sigma_{h2}$) the appropriately normalized power spectrum of $u_{3r}$, $u_{4r}$, and $u_{5r}$ has the form

$$\langle \hat{u}^2(\mathbf{k})\rangle = a\frac{\exp\left[-\frac{\sigma^2}{4}k^2\right]}{\exp\left[\frac{\sigma^2}{4}k^2\right]-q^2k^2}\ , \quad \text{(C1)}$$

where the constants $q$ and $a$ are proportional to the order parameter and where we have suppressed the component index $i$ for simplicity. The spatial autocorrelation functions are given by the Fourier transform of (C1):

$$S(\mathbf{p})=\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(i\mathbf{kp})\langle \hat{u}^2(\mathbf{k})\rangle d^2k = a\int_0^{\infty}J_0(k|\mathbf{p}|)\frac{k\exp\left[-\frac{\sigma^2}{4}k^2\right]}{\exp\left[\frac{\sigma^2}{4}k^2\right]-q^2k^2}dk\ , \quad \text{(C2)}$$

where $J_0$ is the zeroth-order Bessel function. In the general case (C2) has to be integrated numerically. Analytic expressions can be obtained for the limit cases $q\to 0$,

$$S(\mathbf{p})\to \text{const}\times\exp\left[-\frac{|\mathbf{p}|^2}{2\sigma^2}\right]\ , \quad \text{(C3)}$$

and for $q \to q_{\text{thres}} = \sqrt{e}\,\sigma$,

$$S(\mathbf{p}) \to \text{const} \times \frac{J_0\left[\frac{2|\mathbf{p}|}{\sigma}\right]}{\sqrt{q_{\text{thres}} - q}} . \tag{C4}$$

For large $p$ the Bessel function is approximated by

$$J_0\left[\frac{2|\mathbf{p}|}{\sigma}\right] \approx \left[\frac{\sigma}{\pi|\mathbf{p}|}\right]^{1/2} \cos\left[\frac{2|\mathbf{p}|}{\sigma} - \pi/4\right] . \tag{C5}$$

Thus, at the critical value of $q$, spatial correlations decay with $1/\sqrt{p}$.


[1] D. H. Hubel and T. N. Wiesel, J. Comp. Neurol. **158**, 267 (1974).
[2] D. H. Hubel and M. S. Livingstone, J. Neurosci. **4**, 309 (1984).
[3] G. G. Blasdel and G. Salama, Nature **321**, 579 (1986).
[4] T. Bonhoeffer and A. Grinvald, Nature **353**, 429 (1991).
[5] G. G. Blasdel, J. Neurosci. (to be published).
[6] G. G. Blasdel, J. Neurosci. (to be published).
[7] A. Grinvald, E. Lieke, R. D. Frostig, C. D. Gilbert, and T. N. Wiesel, Nature **324**, 361 (1986).
[8] K. D. Miller, in *Neuroscience and Connectionist Theory,* edited by M. A. Gluck and D. E. Rumelhart (Lawrence Erlbaum Associates, Hillsdale, NJ, 1990), pp. 267–353.
[9] C. J. Shatz, Neuron. **5**, 745 (1990).
[10] C. von der Malsburg, Kybernetik **14**, 85 (1973).
[11] D. J. Willshaw and C. von de Malsburg, Proc. R. Soc. London Ser. B **194**, 431 (1976).
[12] A. Takeuchi and S. Amari, Biol. Cybern. **35**, 63 (1979).
[13] T. Kohonen, Biol. Cybern. **43**, 59 (1982).
[14] T. Kohonen, Biol. Cybern. **44**, 135 (1982).
[15] R. Linsker, Proc. Natl. Acad. Sci. U.S.A. **83**, 8779 (1986).
[16] R. Soodak, Proc. Natl. Acad. Sci. U.S.A. **84**, 3936 (1987).
[17] D. M. Kammen and A. L. Yuille, Biol. Cybern. **59**, 23 (1988).
[18] K. D. Miller, J. B. Keller, and M. P. Stryker, Science **245**, 605 (1989).
[19] K. D. Miller, Neuroreport **3**, 73 (1992).
[20] A. S. Rojer and E. L. Schwarz, Biol. Cybern. **62**, 381 (1990).
[21] R. Durbin and G. Mitchison, Nature **343**, 644 (1990).
[22] K. Obermayer, H. Ritter, and K. Schulten, Proc. Natl. Acad. Sci. U.S.A. **87**, 8345 (1990).
[23] S. Tanaka, Neur. Network **3**, 625 (1990).
[24] S. Tanaka, Biol. Cybern. **64**, 263 (1991).
[25] T. Kohonen, *Self-Organization and Associative Memory* (Springer- Verlag, New York, 1987).
[26] H. Ritter and K. Schulten, Biol. Cybern. **60**, 59 (1988).
[27] H. Ritter, K. Obermayer, K. Schulten, and J. Rubner, in *Physics of Neural Networks,* edited by E. Domani, J. L. van Hemmen, and K. Schulten (Springer-Verlag, New York, 1991), pp. 281–306.
[28] N. V. Swindale, Proc. R. Soc. London Ser. B **215**, 211 (1982).
[29] T. Kohonen, Proc. IEEE **78**, 1464 (1990).
[30] R. Miikkulainen, in *Artificial Neural Networks I,* edited by T. Kohonen, K. Mäkisara, O. Simula, and J. Kangas (North-Holland, Amsterdam, 1991), pp. 415–420.
[31] E. Erwin, K. Obermayer, and K. Schulten, Biol. Cybern. (to be published).
[32] E. Erwin, K. Obermayer, and K. Schulten, Biol. Cybern. (to be published).
[33] N. V. Swindale, J. A. Matsubara, and M. S. Cynader, J. Neurosci. **7**, 1414 (1987).
[34] W. T. Baxter and B. M. Dow, Biol. Cybern. **61**, 171 (1989).
[35] J. S. Lund and S. Yoshioka, J. Comp. Neurol. (to be published).

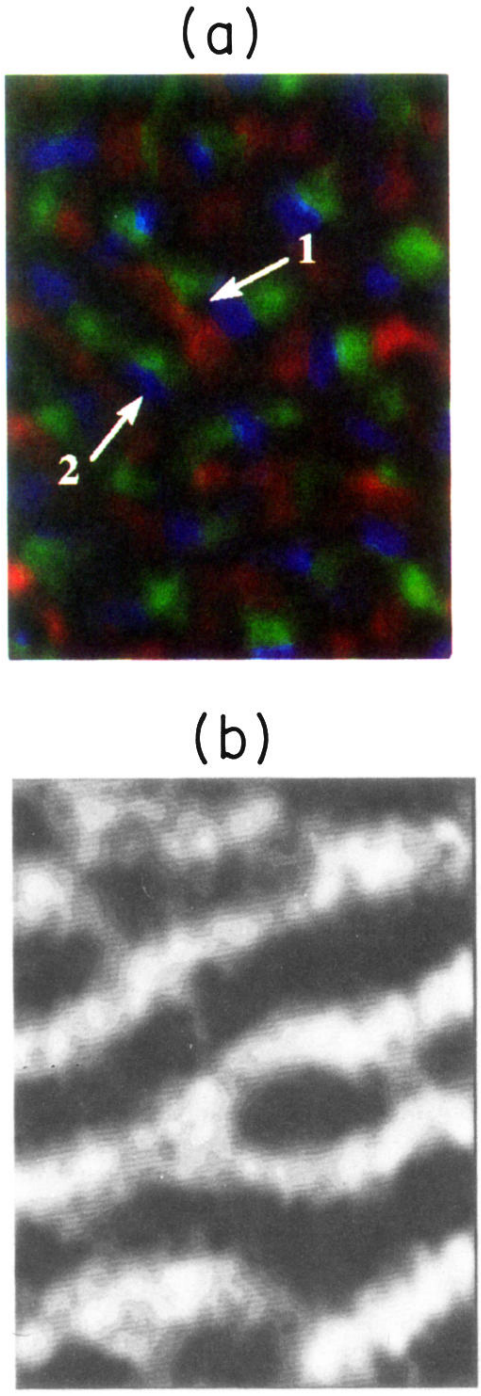


FIG. 1. Spatial pattern of (a) orientation selectivity and (b) ocular dominance in the striate cortex of an adult monkey *(Macaca nemestrina)*. The images were obtained by optical imaging [5,6] and show a $4 \times 6$ mm$^2$ patch of cortex which is located near the border to area 18. Orientation preference is coded by color, where the sequence red $\rightarrow$ yellow $\rightarrow$ green $\rightarrow$ blue $\rightarrow$ purple $\rightarrow$ red describes a complete cycle of preferred orientation for 0° over 90° to 180°. Orientation tuning strength is normalized and coded by brightness, where bright areas indicate regions with a specific response. Ocular dominance is also coded by brightness, where black and white indicate preference for the contralateral and the ipsilateral eye, respectively. The arrows in (a) mark the two prominent elements of the orientation column system, the *singularities* (arrow 1) and regions where orientation preferences are organized in parallel slabs (arrow 2).

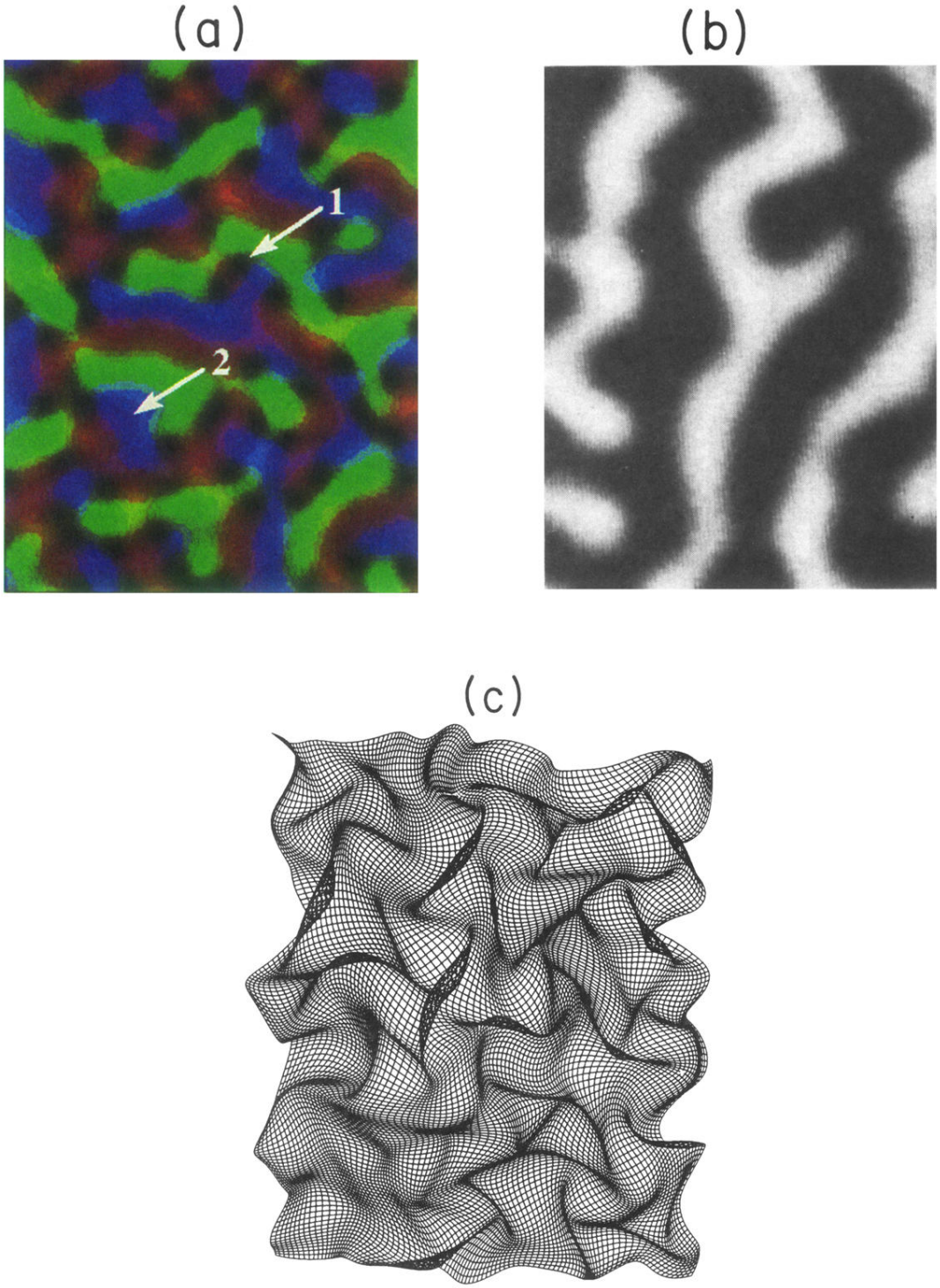


FIG. 10. Typical stationary state of the self-organizing feature maps generated for $T_{3,4,5} > T_{\text{thres}}$ using an isotropic neighborhood function. The figure displays a $110 \times 150$ section of a network of size $N = d = 512$. The parameters of the simulation are given in Fig. 8(a). (a) Spatial pattern of orientation preference $\phi_{\mathbf{r}}$ (color) and orientation tuning strength $q_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. Arrows indicate singularities (arrow 1) and linear regions (arrow 2). (b) Spatial pattern of ocular dominance values $z_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (c) Spatial structure of the topographic projection. The components of the feature vectors were coded as described in Fig. 7.

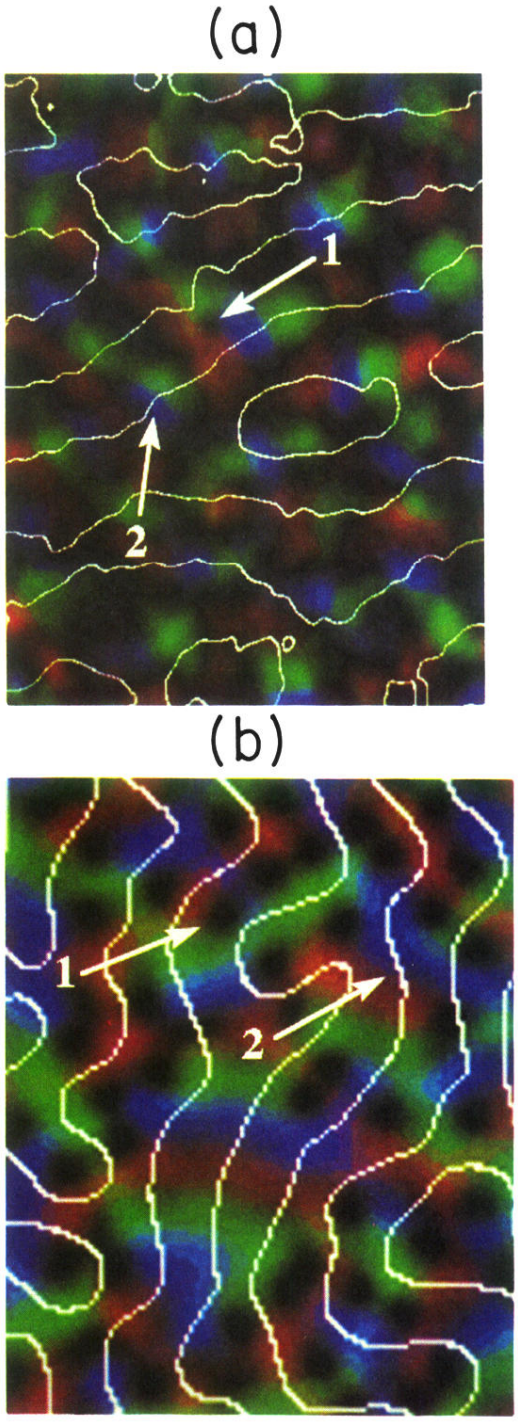


FIG. 11. Correlation between the spatial pattern of orientation selectivity and ocular dominance. Color and brightness code for orientation preference and orientation tuning strength, respectively, the white lines indicate the borders of the ocular dominance bands. (a) Experimental data (taken from [5,6]) obtained from an adult macaque. The image shows a $4\times6$ mm$^2$ patch of cortex, which is located near the border to area 18. (b) Overlay of the orientation and ocular dominance patterns shown in Fig. 10.

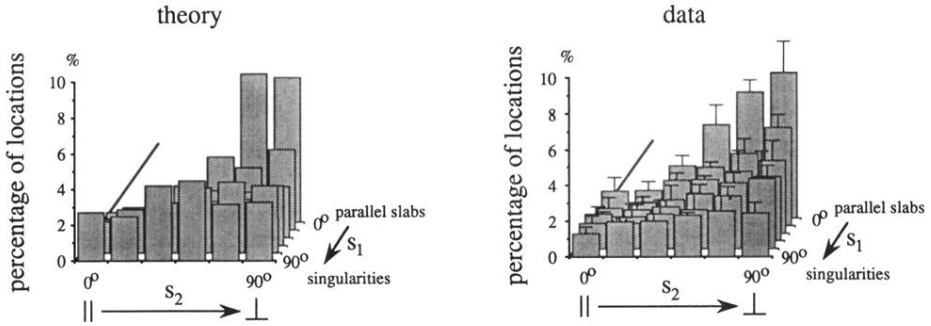


FIG. 13. Number of map locations as a function of $s_1$ and $s_2$ (see text). (a) Analysis of the stationary state of the SOFM algorithm displayed in Fig. 10. The total number of map locations was 1500, $\sigma_g/\lambda_0=0.2$. (b) Analysis of experimental data (taken from [3,5,6]) obtained from four adult monkeys. The total number of map locations was 3500, $\sigma_g/\lambda_0=0.2$. $\sigma_g$ was 150 $\mu$m, the typical radius of an interneuron dendritic tree located in the superficial layers of the monkey striate cortex [35].

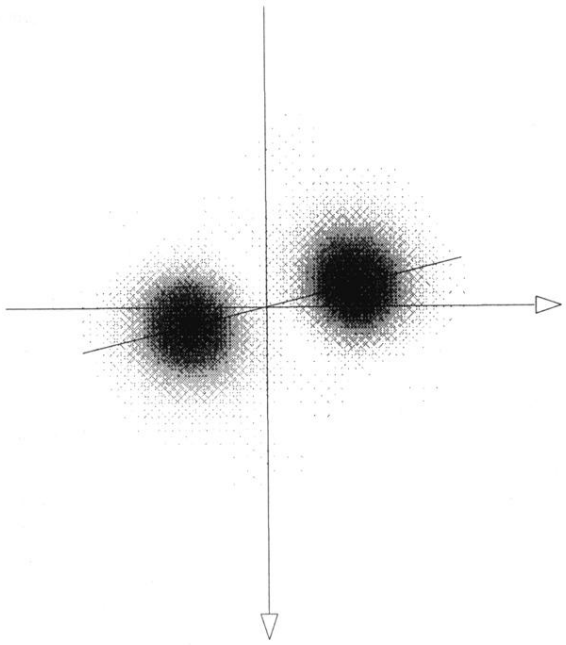

FIG. 12. Gabor transform of $w_{ri}$, $i > 2$ at a typical location in the pattern described in Fig. 10 for $\sigma_g = 6$. Black regions indicate modes with high energy.

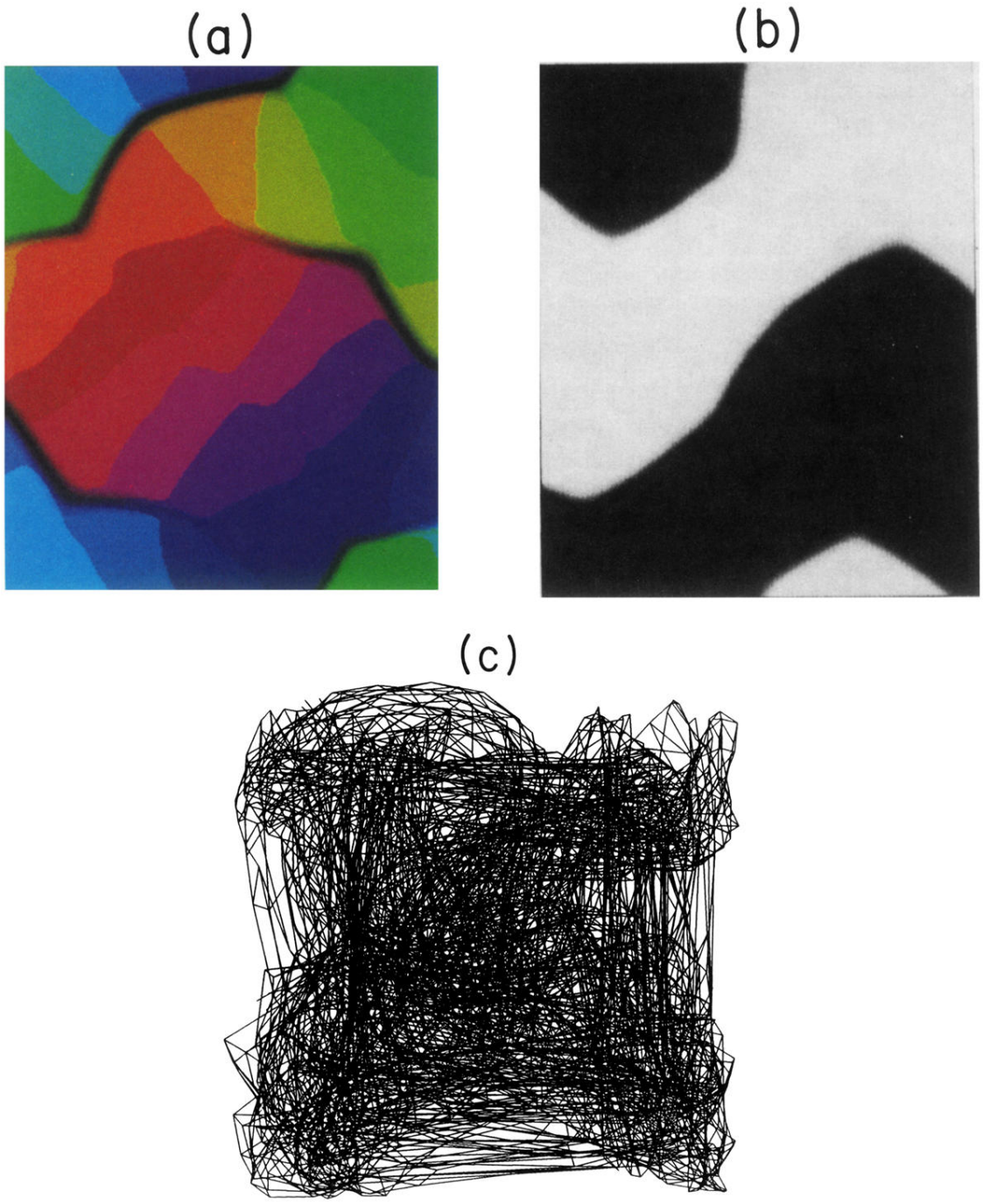


FIG. 15. Typical stationary state of the self-organizing feature maps generated for $T_{3,4,5} \gg T_{\text{thres}}$ using an isotropic neighborhood function. The figure displays a $192 \times 256$ section of a network of size $N = d = 256$. The parameters of the simulation were $\varepsilon = 0.1$, $T_{3,4} = 1280$, and $T_5 = 1478$. The patterns were obtained after 60 000 iterations starting from random initial conditions. $\sigma_h$ was linearly decreased during the simulation from $\sigma_h(0) = 200$ to $\sigma_h(30\,000) = 20$ and then to $\sigma_h(60\,000) = 4$. For this set of parameters the critical value of the order parameters, which depends on $\sigma_h$, lies between 663.0 and 3.3. (a) Spatial pattern of orientation preference $\phi_{\mathbf{r}}$ (color) and orientation tuning strength $q_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (b) Spatial pattern of ocular dominance values $z_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (c) Spatial structure of the topographic projection. The components of the feature vectors were coded as described in Fig. 7.

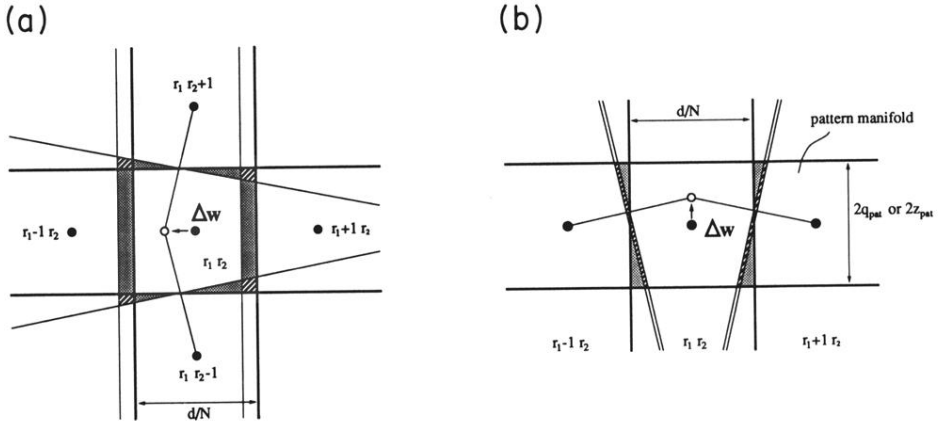


FIG. 16. Change of shape of the tesselation cells corresponding to the stationary state (18) under an infinitesimal shift $\partial \mathbf{w}_r$ of the weight vector $\mathbf{w}_r$. (a) Shift parallel to the coordinates $v_1$ and $v_2$. The drawing shows a projection of the tesselation onto the $v_1, v_2$ plane, where the positions of the weight vectors before and after the shift are indicated by black and white dots, respectively. The dotted areas indicate volume changes linear in $\partial \mathbf{w}_r$; they contribute to Eqs. (B6) and (B7). The hatched areas give an upper bound of corrections of $O(|\partial \mathbf{w}_r|^2)$ to these volume changes; they are added to or subtracted from the dotted areas to keep expressions (B6) and (B7) simple. (b) Shift parallel to the coordinates $v_3$ to $v_5$. The drawing shows a projection of the tesselation onto the $v_1, v_i$ plane where $i \in \{3,4,5\}$. The two horizontal lines indicate the borders of the pattern manifold. Hatched and dotted areas again indicate volume changes of first and higher order in $\partial \mathbf{w}_r$, respectively. Note, that a calculation of volume and centroid changes under $\partial w_{3r}$ and $\partial w_{4r}$ have to take into account the cylindrical shape of the pattern manifold, while changes under $\partial w_{5r}$ have to take into account an essentially boxlike geometry. This leads to the different prefactors in Eq. (B7).

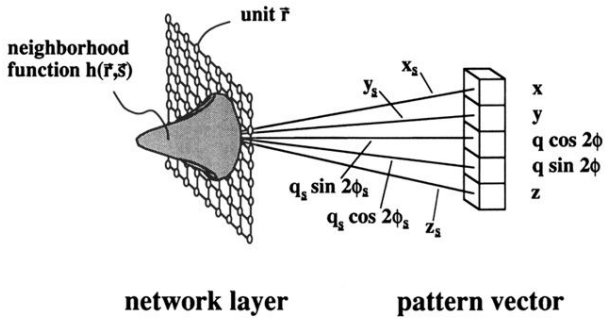


FIG. 3. The neural network model.

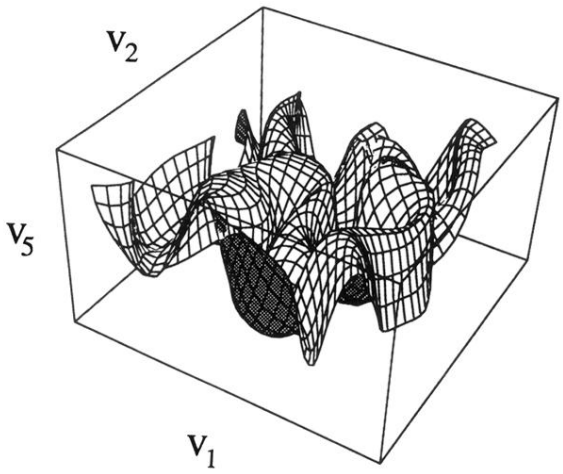


FIG. 4. A typical solution of the map problem (Fig. 2) found by the SOFM algorithm. The locations of the pattern vectors which are mapped to the units in the network are indicated by the intersections of a grid in feature space. Only every fourth vector is presented. The data was taken from a small region of the map shown in Fig. 10.

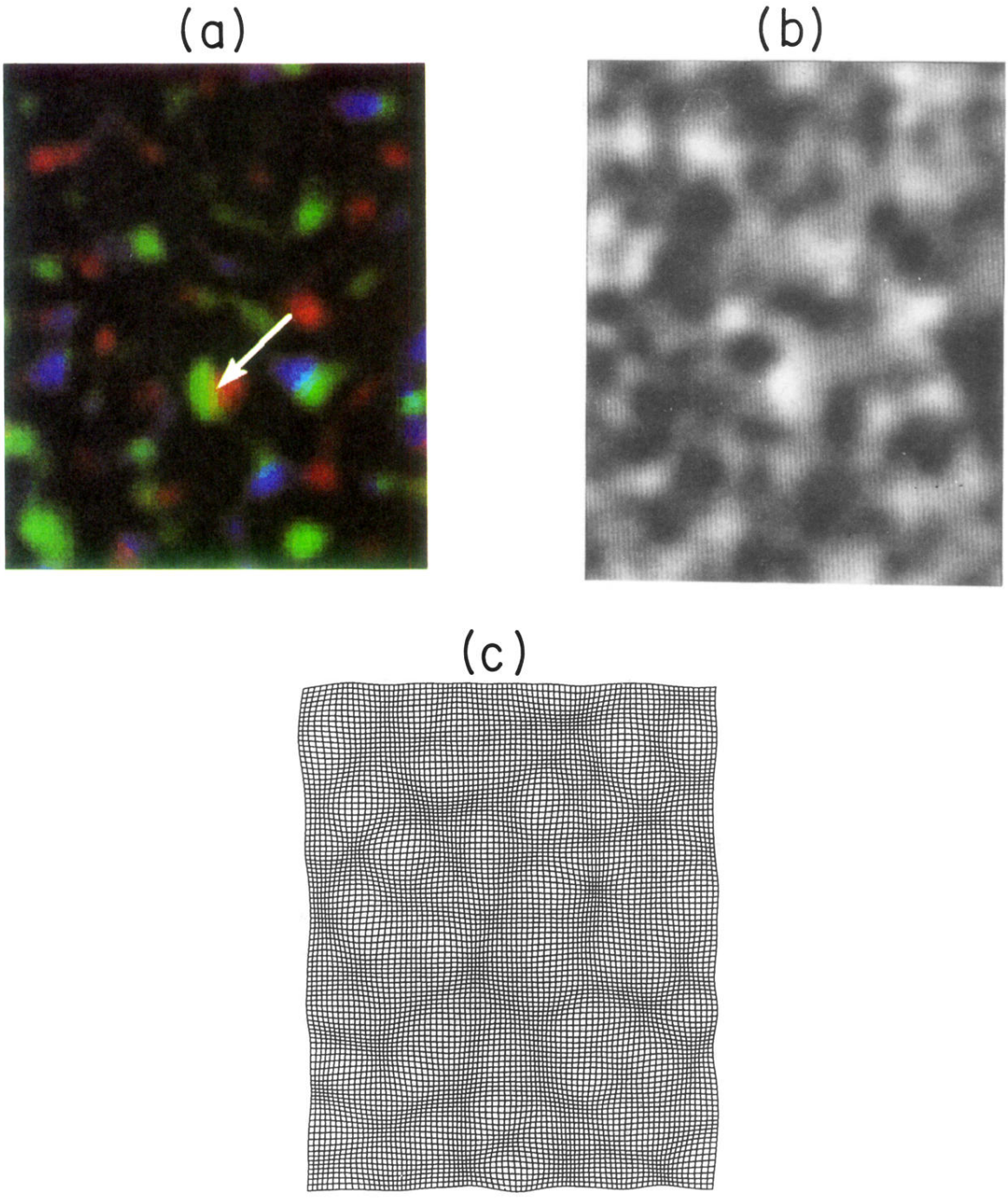


FIG. 7. Snapshot of fluctuations around the stationary state (18). The figure displays an $80\times100$ section of a network of size $N=d=512$. The parameters of the simulation were $\varepsilon=0.1$, $\sigma_h=5$, $T_{3,4,5}=1.24$. For this set of parameters the critical value of the order parameters is $T_{\text{thres}}=4.12$. The snapshot was obtained after 60 000 iterations from the "topographic" initial state (18). (a) Spatial pattern of orientation preference $\phi_{\mathbf{r}}$ (color) and orientation tuning strength $q_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. (b) Spatial pattern of ocular dominance values $z_{\mathbf{r}}$ (brightness) as a function of location $\mathbf{r}$. Orientation preference, orientation tuning strength, and ocular dominance values were color coded as described in Fig. 1. (c) Spatial structure of the topographic projection. The figure displays the locations $(x_{\mathbf{r}1},x_{\mathbf{r}2})$ of receptive field centers in visual space for all units in the network layer. Receptive field centers of neighboring cells were connected by lines.

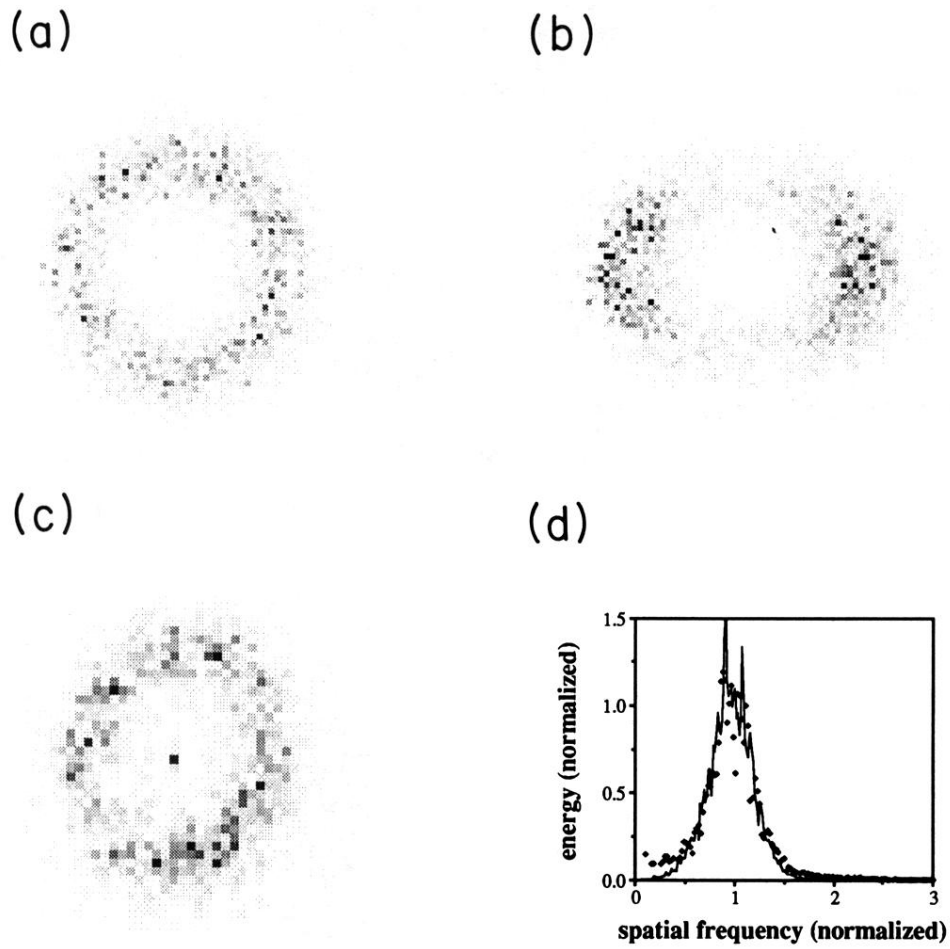


FIG. 8. Complex Fourier transforms $\sum_{\mathbf{r}} \exp(i\mathbf{kr}) q_{\mathbf{r}} (\cos\phi_{\mathbf{r}} + i \sin\phi_{\mathbf{r}})$ of the spatial pattern of orientation preference and corresponding power spectra. (a) Fourier transform of the pattern of orientation of a typical stationary state of the SOFM algorithm generated for $T_{3,4,5} > T_{\text{thres}}$ using an isotropic neighborhood function. Parameters of the simulation were $N = d = 512$, $\varepsilon = 0.02$, $\sigma_h = 5$, $T_{3,4} = 10.24$, and $T_5 = 8.87$. For this set of parameters the critical value of the order parameters is $T_{\text{thres}} = 4.12$. The figure was obtained after $9 \times 10^7$ iterations from the "topographic" initial state (18). (b) Fourier transform of the pattern of orientation selectivity for a typical stationary state of the SOFM algorithm generated for $T_{3,4,5} > T_{\text{thres}}$ using an anisotropic neighborhood function. The parameters of the simulation were $N = d = 512$, $\varepsilon = 0.02$, $\sigma_{h1} = 5$, $\sigma_{h2} = 7.5$, $T_{3,4} = 10.24$, and $T_5 = 8.87$; the critical value of the order parameters is $T_{\text{thres}} = 4.12$. The figure was obtained after $9 \times 10^7$ iterations from the "topographic" initial state (18). (c) Fourier transform of the pattern of orientation selectivity obtained from a $6 \times 8$ mm$^2$ patch of macaque striate cortex. (d) Power spectrum of (a) and (c) obtained by averaging over all directions of the $\mathbf{k}$ vector. Mean peak amplitude and peak frequency were normalized to 1. Unit frequency corresponds to 1.38/mm for the experimental data.